HAL
open science

Kolmogorov-Fokker-Planck operators in dimension two: heat kernel and curvature

Davide Barilari, Francesco Boarotto

To cite this version:

Davide Barilari, Francesco Boarotto. Kolmogorov-Fokker-Planck operators in dimension two: heat kernel and curvature. Journal of Evolution Equations, 2018, 18 (3), pp.1115-1146. hal-01593179

HAL Id: hal-01593179

https://hal.science/hal-01593179

Submitted on 25 Sep 2017

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

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

KOLMOGOROV-FOKKER-PLANCK OPERATORS IN DIMENSION TWO: HEAT KERNEL AND CURVATURE

DAVIDE BARILARI AND FRANCESCO BOAROTTO

Abstract. We consider the heat equation associated with a class of hypoelliptic operators of Kolmogorov-Fokker-Planck type in dimension two. We explicitly compute the first meaningful coefficient of the small time asymptotic expansion of the heat kernel on the diagonal, and we interpret it in terms of curvature-like invariants of the optimal control problem associated with the diffusion. This gives a first example of geometric interpretation of the small-time heat kernel asymptotics of non-homogeneous Hörmander operators which are not associated with a sub-Riemannian structure, i.e., whose second-order part does not satisfy the Hörmander condition.

Contents

1. Introduction 1
2. The fundamental solution and its perturbation 7
3. Perturbing the Kolgomorov operator 11
4. Curvature-like invariants 17
5. Geometric interpretation of the small-time asymptotics 21
References 23

1. Introduction

During the last years, many results revealing the interaction between analysis and geometry have been obtained in the study of the heat equation. In particular, many of them relate the small-time asymptotics of the fundamental solution of the heat equation, the so-called *heat kernel*, to geometric quantities underlying the partial differential operator, in several smooth and non-smooth contexts.

These investigations are inspired by the nowadays classical results in Riemannian geometry. In the Riemannian context there is a precise understanding of the small-time asymptotics of the heat kernel and its relation to distance, its singularities (such as cut and conjugate loci) and curvature invariants. Some fundamental results in these directions can be found in [54, 45, 46, 22]. For a more comprehensive account of the literature we refer to the books [21, 30, 49] and references therein.

We state here the following result, that can be found for instance in [24, Chapter IV], and which is the closest to the spirit of the investigation presented in this paper.

Date: September 25, 2017.
Key words and phrases. Kolmogorov operator, Fokker-Planck equation, heat kernel asymptotics, curvature, optimal control.

Theorem 1. *Let (M, g) be a complete smooth n-dimensional Riemannian manifold, and let Δ_g be the Laplace-Beltrami operator of (M, g). Let $p(t, x, y)$ be the heat kernel associated with $L = X_0 + \frac{1}{2}\Delta_g$, where X_0 is a smooth vector field. Fix $x_0 \in M$. Then for $t \to 0$ one has*

$$p(t, x_0, x_0) = \frac{1}{(2\pi t)^{n/2}} \left(1 + t\left(-\frac{1}{2}\mathrm{div}_\mu(X_0)(x_0) - \frac{1}{2}|X_0(x_0)|^2 + \frac{1}{12}S(x_0)\right) + O(t^2)\right), \tag{1.1}$$

where μ is the Riemannian volume, and S is the scalar curvature of the Riemannian metric g.

For later purposes, let us mention that, in terms of an orthonormal basis $X_1, \dots, X_n$ for the Riemannian metric, L rewrites as follows

$$L = X_0 + \frac{1}{2}\left(\sum_{i=1}^{n} X_i^2 + (\mathrm{div}_\mu X_i)X_i\right), \tag{1.2}$$

where μ is the Riemannian volume form. Recall that the divergence $\mathrm{div}_\mu X$ of a vector field X with respect to the measure defined by the volume form μ is the function characterized by the identity

$$\int_M Xf\, d\mu = -\int_M f\mathrm{div}_\mu X\, d\mu, \qquad f \in C_c^\infty(M)$$

for every smooth function f having compact support in M.

The extension of results in the spirit of Theorem 1 to non-Riemannian situations, such as sub-Riemannian geometry or more in general hypoelliptic operators, when possible, is non-trivial: some results have been obtained relating the hypoelliptic heat kernel of a sub-Riemannian Laplacian with its associated Carnot-Carathéodory distance [42, 43, 4] and its cut locus [8, 9], but much less is known concerning the relation with curvature. A result for 3D contact structures is contained in [7], where an invariant κ playing the role of the curvature for sub-Riemannian structures appears. Only partial results are known even in the higher-dimensional contact case: for structures with symmetries in [17, 50] the first coefficient has been related to the scalar Tanaka-Webster curvature. See also [15, 16, 55] for a recent account on higher-dimensional model spaces.

This is related on the one hand to the difficulty of defining a general notion of (scalar) curvature associated with a given Carnot-Carathéodory metric [3, 12, 2, 13], and on the other hand (at least using the perturbation technique exploited in the present paper) to the fact that the precise knowledge of the heat kernel of the principal part of the operator is non-trivial as soon as the operator is not elliptic. The main idea of the paper is to exploit the knowledge of the principal part of the operator when it is an Hörmander operator of Kolmogorov type. A characterization of the full small-time heat kernel asymptotics in terms of curvature-like invariants for this kind of operators have been considered in [11].

The goal of this paper is to provide an analogue of Theorem 1 for the heat equation associated with a class of hypoelliptic operators of Kolmogorov-Fokker-Planck type in dimension two. To our best knowledge this is the first example of this kind of results for non-homogeneous Hörmander operators which are not associated with a sub-Riemannian structure, i.e., whose second-order part does not satisfy the Hörmander condition.

We explicitly compute the first meaningful coefficient of the small-time asymptotic expansion of the heat kernel on the diagonal, and we study its relation with curvature-like invariants of the optimal control problem associated with the diffusion operator.

1.1. **The Kolmogorov operator.** Consider on $[0,+\infty)\times\mathbb{R}^2$ the second-order partial differential operator $\mathcal{K}$ defined as follows:

$$\mathcal{K}=\partial_t-x_1\partial_{x_2}-\frac{1}{2}\partial_{x_1}^2,\qquad (t,x_1,x_2)\in[0,+\infty)\times\mathbb{R}^2 \tag{1.3}$$

In the literature, $\mathcal{K}$ is often called Kolmogorov operator, since it was first introduced in Kolmogorov's paper [39] in the study of Brownian motion and kinetic theory of gases.

In fact, Kolmogorov wrote in [40] an explicit smooth (which always means C^∞ in what follows) fundamental solution for the operator $\mathcal{K}$ with respect to the Lebesgue measure, thus proving its hypoellipticity. Introducing the matrices

$$A=\begin{pmatrix}0&0\\1&0\end{pmatrix},\quad\text{and}\quad D_t=\begin{pmatrix}t&t^2/2\\t^2/2&t^3/3\end{pmatrix},$$

the fundamental solution $p_0:\mathbb{R}^+\times(\mathbb{R}^2)^2\to\mathbb{R}$ associated with (1.3) can be explicitly rewritten as follows

$$p_0(t,x,y)=\frac{e^{-\frac{1}{2}(y-e^{tA}x)^*D_t^{-1}(y-e^{tA}x)}}{2\pi\sqrt{\det D_t}},\qquad (t,x,y)\in[0,+\infty)\times(\mathbb{R}^2)^2,$$

where v^* denotes the transpose of a vector v. Notice in particular that choosing $x=y=0$ one has the on-the-diagonal identity

$$p_0(t,0,0)=\frac{\sqrt{12}}{2\pi t^2}. \tag{1.4}$$

The Kolmogorov operator $\mathcal{K}$ is one of the simplest examples of a homogeneous, left-invariant and hypoelliptic operator that we can define on an homogeneous group structure $\mathbb{G}$ on $\mathbb{R}^3$. To see this, let us endow $\mathbb{G}=\mathbb{R}^3$ with the following translations and dilations:

$$(x_1,x_2,t)\circ(y_1,y_2,s):=(x_1+y_1,x_2+y_2-x_1s,t+s),$$
$$\delta_\lambda(x_1,x_2,t):=(\lambda x_1,\lambda^3x_2,\lambda^2t),\quad \lambda\in\mathbb{R}^+.$$

Then one can see $\mathbb{G}$ as a homogeneous and stratified Lie group (in the sense of [25]) where $X_0=x_1\partial_{x_2}$, $X_1=\partial_{x_1}$ and ∂_t are left-invariant vector fields on $\mathbb{G}$. Thus

$$\mathcal{K}=\partial_t-\left(X_0+\frac{1}{2}X_1^2\right) \tag{1.5}$$

is indeed an hypoelliptic partial differential operator on $\mathbb{G}$, that is left-invariant with respect to the product and homogeneous of degree two with respect to the family of dilations.

Notice that $[X_0,X_1]=-\partial_{x_2}$. Then the vector fields X_0,X_1 satisfy the Hörmander condition

$$\operatorname{span}_x\{X_1,[X_0,X_1]\}=T_xM,\qquad\forall\,x\in M, \tag{1.6}$$

which guarantees the hypoellipticity of $\mathcal{K}$ thanks to the celebrated result of [32].

The study of such operators has recently known a quick development. For the interested reader we refer for example to [41, 48, 27, 29, 14, 10] and references therein.

1.2. **Perturbing the Kolmogorov operator.** We consider in this paper a class of hypoelliptic Hörmander operators that can be modeled on $\mathcal{K}$. In what follows M denotes $\mathbb{R}^2$ or a smooth, connected and compact two-dimensional manifold, endowed with a volume form μ (or a smooth density), we study the class of second-order differential operators of Fokker-Planck type on $\mathbb{R}\times M$ written in the form ∂_t-L where

$$L=X_0+\frac{1}{2}\left(X_1^2+\operatorname{div}_\mu(X_1)X_1\right), \tag{1.7}$$

and the vector fields X_0, X_1 satisfy condition (1.6). This class of operators does indeed contain (1.3), as explained above, cf. (1.5). Notice that the vector field $X_1 = \partial_{x_1}$ has divergence zero with respect μ, choosing μ as the Lebesgue measure in $M = \mathbb{R}^2$.

Remark. The presence of the divergence term in (1.7) does not really affect the class of operators we are considering since the first-order term can be always included in the drift by defining $Y_0 := X_0 + \frac{1}{2}\mathrm{div}_\mu(X_1)X_1$ and writing

$$L = Y_0 + \frac{1}{2}X_1^2.$$

Moreover the pair X_0, X_1 satisfies (1.6) if and only if Y_0, X_1 does.

Writing L in the form (1.7) is mainly useful for the final geometric interpretation of the small-time heat kernel expansion. Even in the case of a second-order elliptic operator, for the geometric interpretation of the heat kernel it is useful to write it as a sum of a self-adjoint part (i.e., a Laplace-Beltrami operator for some suitable Riemannian metric) plus a drift, as Theorem 1 and formula (1.2) suggest.

The Hörmander condition guarantees the existence of a (local) smooth fundamental solution $p \in C^\infty(\mathbb{R}^+ \times M \times M)$ for $\partial_t - L$ with respect to the measure μ, and our objective is to retrieve geometrical informations on M from the small-time asymptotics of p on the diagonal $\Delta = \{(x, x) \mid x \in M\} \subset M \times M$.

More precisely, given any reference point $x_0 \in M$, we want to understand the expansion $p(t, x_0, x_0)$ for $t \to 0$, and see how it reflects the local geometry of the manifold around x_0. It is well-known from the works of Ben Arous and Léandre [19, 20] that in order to expect a polynomial decay for the heat kernel on the diagonal (as in the elliptic case, cf. (1.1)), we have to make the additional assumption that $X_0(x_0)$ is parallel to $X_1(x_0)$, with $X_0(x_0)$ possibly zero.

When the two vectors $X_1(x_0)$ and $X_0(x_0)$ are independent, one may expect an exponential decay for the heat on Δ. An explicit example of this behavior is showed in [11] for the class of higher-dimensional Kolmogorov operators.

The relation between the anisotropic diffusion generated by $\partial_t - L$ with its underlying geometry is primarily given by the probabilistic characterization of the fundamental solution $p(t, x, y)$ as the probability density of the stochastic process $t \mapsto \xi_t$ that solves the stochastic differential equation

$$d\xi_t = X_0(\xi_t) + X_1(\xi_t) \circ dw, \quad \xi_0 = x \in M.$$

Here w denotes the standard Brownian motion on M and $\circ$ indicates the integration in the Stratonovich sense. Moreover, it is known from the pioneering ideas of [44, 51, 6] that when $t \to 0$, the diffusion tends to concentrate along the optimal trajectories of the following optimal control problem

$$\dot{x} = X_0(x) + uX_1(x), \quad x \in M, \quad u \in \mathbb{R}, \tag{1.8}$$

where one minimizes the energy

$$J_t(u) = \frac{1}{2}\int_0^t |u(s)|^2 ds \tag{1.9}$$

among all controls $u \in L^2([0, t], \mathbb{R})$ such that the corresponding solution to (1.8) joins x and y in time t. In other words, for every fixed $x_0 \in M$ and $t > 0$, we consider the *attainable set* $A^t_{x_0} := \{x_u(t) \mid u \in \mathcal{U},\, x_u(0) = x_0\}$, where $\mathcal{U} \subset L^2([0, t], \mathbb{R})$ is the open subset of $L^2([0, t], \mathbb{R})$ such that the corresponding solution $x_u : [0, t] \to M$ to (1.8) is defined on the interval $[0, t]$. We

introduce also the *value function* $S^t_{x_0} : M \to \mathbb{R} \cup \{+\infty\}$, by

$$S^t_{x_0}(y) := \inf\{J_t(u) \mid u \in \mathcal{U},\, x_u(0) = x_0,\, x_u(t) = y\},$$

with the convention that $\inf \emptyset = +\infty$. For $y \in A^t_{x_0}$, solving the optimal control problem means then finding the controls u realizing the infimum in $S^t_{x_0}(y)$.

Under the assumption (1.6), the attainable set has non-empty interior, and the value function is always smooth along optimal trajectories for small time, see e.g. [3, Appendix A]. If moreover $X_0(x_0)$ is parallel to $X_1(x_0)$, it follows from [23, 47] that for small times t we can relate the attainable set $A^t_{x_0}$ to the attainable set $A^{t,\mathcal{K}}_{x_0}$, reached by solutions to (1.8) where $X_0 = x_1\partial_{x_2}$ and $X_1 = \partial_{x_1}$ are the vector fields defining the Kolmogorov operator $\mathcal{K}$. In particular, since $A^{t,\mathcal{K}}_{x_0}$ contains a full neighborhood of x_0, the same holds for $A^t_{x_0}$ and the value function $S^t_{x_0}$ is well-defined in a neighborhood of x_0.

Curvature-like invariants of the control system (1.8) are obtained by higher-order derivatives of the value function along optimal trajectories, and as such really describe the local geometry of M around x_0. For more details to curvature-like invariants in this setting we refer the reader to Section 4, and for a more comprehensive presentation to [3].

We stress that these curvature-like invariants are indeed a generalization of the Riemannian curvature. More precisely, when one applies this approach to a control systems associated with a Riemannian geodesic problem, the value function is related to the squared Riemannian distance and the curvature-like invariants recover the full Riemann curvature tensor [3, Chapter 4].

1.3. **Main result and comments.** To have a clear geometric interpretation of the small-time asymptotics, it is natural to choose for the operator L defined as in (1.7), a volume form μ that represents some canonical volume form associated with the problem, as in the Riemannian case where one chooses the canonical Riemannian volume.

In view of the assumption (1.6), it is natural to consider on the two-dimensional manifold M the volume form μ defined by the requirement that

$$\mu(X_1, [X_0, X_1]) = 1. \tag{1.10}$$

Inspired by Theorem 1, we expect the small-time asymptotics of $p(t, x_0, x_0)$ to contain informations on the curvature-like invariants associated with the system (1.8), and to depend on the pointwise values both of $X_0(x_0)$ and $\mathrm{div}_\mu(X_0)(x_0)$. The main result of this paper is the following.

Theorem 2. *Let M be a two-dimensional smooth, connected and compact manifold, or $M = \mathbb{R}^2$. Fix $x_0 \in M$ and let X_0 and X_1 be two smooth vector fields on M, with bounded derivatives of all orders, such that*

(a) *$X_0(x_0)$ is parallel to $X_1(x_0)$,*
(b) *$\mathrm{span}_x\{X_1, [X_0, X_1]\} = T_xM$, for every $x \in M$.*

Let μ be the volume form defined in (1.10) *and $p(t, x, y)$ be the heat kernel associated with*

$$\partial_t - X_0 - \frac{1}{2}\left(X_1^2 + (\mathrm{div}_\mu X_1)X_1\right). \tag{1.11}$$

Then the following asymptotic expansion holds for $t \to 0$

$$p(t, x_0, x_0) = \frac{\sqrt{12}}{2\pi t^2}\left(1 + t\left(-\frac{1}{2}\mathrm{div}_\mu(X_0)(x_0) - \frac{1}{2}|X_0(x_0)|^2 + \frac{1}{14}K_1(x_0) - \frac{1}{70}K_2(x_0)^2\right) + o(t)\right). \tag{1.12}$$

The main-order coefficient in (1.12) is the one of Kolmogorov operator (1.4), while the contribution of the drift is analogous to what happens in (1.1). The two remaining coefficients

are curvature-like invariants associated with the control problem and play the role of the scalar curvature. All coefficients appearing in (1.12) have invariant meaning.

Few comments on the statement of Theorem 2 are in order.

(i). The manifold M is not endowed by a canonical Riemannian metric, the symbol $|X_0(x_0)|$ here denotes the norm of the vector $X_0(x_0)$ in the vector space $D_{x_0} := \mathbb{R}X_1(x_0)$ endowed with the scalar product defined by declaring $X_1(x_0)$ as an orthonormal basis.

(ii). The assumption for the vector fields to have bounded derivatives of all orders is only technical (and automatically satisfied when M is compact). Indeed one can weaken the condition above to a "polynomial growth condition" on the vector fields to prove Hörmander Theorem, cf. [31, Assumption 4.2 and Theorem 4.5]. This plays the role of a completeness assumption (cf. Theorem 1) and, as soon as the operator (1.11) admits a fundamental solution, thanks to the localization arguments contained in Section 3 one reduces for the small-time asymptotics on the diagonal to the assumptions of Theorem 2.

(iii). The quantities K_1 and K_2 are curvature-like invariants associated with the control problem (1.8)-(1.9), defined in Section 4. In particular K_1 can be interpreted as the contraction along the controlled vector field X_1 of the second-order part of the curvature operator associated with the control problem, while K_2 plays the role of a "higher order" curvature. The presence of both terms is related with the fact that the drift X_0 is necessary to fulfill the Hörmander condition.

For a detailed and general presentation of curvature-like invariants associated with an affine optimal control problem, we refer to the monograph [3]. In Section 4 we give an adapted presentation to the specific context of the paper. We define and compute the constants K_1 and K_2 in terms of the structure constants of the Lie algebra defined by X_0 and X_1.

1.4. **On the role of the volume.** Let us denote by L^μ the operator defined in (1.7) with respect to the volume μ. If one is interested in the study of the operator L^ω associated with a different smooth volume ω, which is then absolutely continuous with respect to μ, it is not difficult to understand that this is actually equivalent to the study of the same operator with a new drift vector field. More precisely, writing $\omega = e^\psi \mu$ for some smooth function $\psi \in C^\infty(M)$, the formula

$$\mathrm{div}_\omega(X_1) = \mathrm{div}_\mu(X_1) + (X_1\psi)$$

shows that the study of the operator L^ω reduces to the study of the original one (i.e., where the second-order term is symmetric with respect to μ) but with modified drift $X_0' = X_0 + \frac{1}{2}(X_1\psi)X_1$. See also [1] for a discussion about the role of the curvature in the variation of the volume along a geodesic flow.

1.5. **Structure of the paper.** In Section 2 we introduce some preliminaries, we present the perturbation formula of Duhamel and we explain how the blow-up procedure applied to a hypoelliptic operator $\partial_t - L$ permits to recover its nilpotent part. Section 3 constitutes the computational core of this work. We specialize the theory previously exposed to the specific case of a two-dimensional manifold and a Kolmogorov-like operator. We show how a fundamental solution of $\partial_t - L$ can be reconstructed from a "weighted" parametrix method applied to the fundamental solution of $\mathcal{K}$. Despite the low dimensionality of the problem the calculations involved are heavy, nonetheless everything can be done explicitly.

Section 4 is devoted to a brief account on curvature-like invariants associated with an affine control system associated to our operator, and we compute them explicitly. Finally, in Section 5, we collect together the results to interpret the first coefficient appearing in the small-time asymptotics of the fundamental solution, proving Theorem 2.

2. The fundamental solution and its perturbation

In this section M denotes either $\mathbb{R}^n$, or a smooth, connected, compact manifold of dimension n. Let also μ be a given smooth volume form on M.

Given any collection $X_0, X_1, \dots, X_k$, of smooth (i.e., C^∞) vector fields on M with bounded derivatives of all orders and that satisfy the Hörmander condition

$$\mathrm{Lie}_x \left\{ (\mathrm{ad}\, X_0)^j X_i \mid j \geq 0,\, i = 1, \dots, k \right\} = T_x M \quad \text{for every } x \in M, \tag{2.1}$$

we consider the differential operator $\partial_t - L$, defined on the space of distributions $\mathcal{D}'(\mathbb{R}^+ \times M)$ as

$$(\partial_t - L)(\varphi) = \partial_t(\varphi) - X_0(\varphi) - \frac{1}{2}\sum_{i=1}^k X_i^2(\varphi), \quad \varphi \in \mathcal{D}'(\mathbb{R}^+ \times M). \tag{2.2}$$

Definition 3 (Fundamental solution)**.** A *fundamental solution* to the operator $\partial_t - L$ with respect to μ is a distribution $p(t,x,y) \in \mathcal{D}'(\mathbb{R}^+ \times M \times M)$ such that

(a) $(\partial_t - L)p(t,x,y) = 0$ (here L acts on the x-variable) for every $x, y \in M$,
(b) $\lim_{t\to 0} p(t,x,y) = \delta_y(x)$, where $\delta_y(x)$ denotes the Dirac delta function concentrated at y, for every $x, y \in M$.

An immediate consequence of (b) is that for every $\varphi \in L^2(M,\mu)$ there holds

$$\lim_{t\to 0} \int_M p(t,x,y)\varphi(y)d\mu(y) = \varphi(x), \quad \forall\, x \in M, \quad \text{uniformly in } t.$$

A crucial result of Hörmander [32] states that, under condition (2.1), the operator $\partial_t - L$ in (2.2) is hypoelliptic, that is for every distribution v defined on a domain $\Omega \subset \mathbb{R} \times M$, the condition $(\partial_t - L)(v) \in C^\infty(\Omega)$ implies $v \in C^\infty(\Omega)$. In particular, it admits a smooth fundamental solution p with respect to the given volume μ, given explicitly as the probability density of the process $t \mapsto \xi_t$ that solves the following stochastic differential equation (in the Stratonovich sense)

$$\begin{cases} d\xi_t = X_0(\xi_t) + \displaystyle\sum_{i=1}^k X_i(\xi_t) \circ db_i, \\ \xi_0 = x, \end{cases}$$

where $b = (b_1, \dots, b_k)$ is a k-dimensional Brownian motion, see for instance [31].

Definition 4 (Weighted coordinates)**.** Let (U, x) be a coordinated neighborhood of x_0, and let $w = (w_1, \dots, w_n)$ be an n-tuple of positive integers. For any $0 < \varepsilon \leq 1$ we define the ε-dilation δ_ε as follows:

$$\delta_\varepsilon(x_1, \dots, x_n) = (\varepsilon^{w_1} x_1, \dots, \varepsilon^{w_n} x_n), \quad \forall\, x = (x_1, \dots, x_n) \in U.$$

The action of δ_ε extends to the coordinated fields ∂_{x_i} and to the volume form μ by

$$(\delta_{1/\varepsilon *})\partial_{x_i} = \varepsilon^{-w_i}\partial_{x_i}, \quad (\delta_\varepsilon)^*\mu = \varepsilon^{\sum_{i=1}^k w_i}\mu.$$

If we perform a rescaling of time and space around x_0, the fundamental solution p changes according to the following proposition [47, Proposition 2.4].

Proposition 5. *Fix weights $(w_1, \dots, w_n)$, $0 < \varepsilon \leq 1$ and $\gamma \in \mathbb{N}$. Then a fundamental solution q_ε to the rescaled operator*

$$\partial_t - \varepsilon^\gamma \left(\delta_{1/\varepsilon *} X_0 + \frac{1}{2}\sum_{i=1}^k (\delta_{1/\varepsilon *} X_i)^2 \right) \tag{2.3}$$

is given by

$$q_\varepsilon(t,x,y) = \varepsilon^{\sum_{i=1}^k w_i} p(\varepsilon^\gamma t, \delta_\varepsilon x, \delta_\varepsilon y), \tag{2.4}$$

where $p(t,x,y)$ is a fundamental solution associated with the operator (2.2).

When we apply the push-forward of a dilation δ_ε to the vector field X_i, depending on the values of the weights $w = (w_1, \dots, w_n)$ and on its Taylor expansion, there exist an integer α_i and a vector field $\widehat{X}_i$, homogeneous of degree $-\alpha_i$ with respect to the dilation, such that

$$\delta_{1/\varepsilon *} X_i = \frac{1}{\varepsilon^{\alpha_i}} \widehat{X}_i + o(\varepsilon^{-\alpha_i}).$$

In particular, we can rewrite (2.3) as

$$\partial_t - \varepsilon^\gamma \left(\frac{1}{\varepsilon^{\alpha_0}} \widehat{X}_0 + \frac{1}{2} \sum_{i=1}^k \frac{1}{\varepsilon^{2\alpha_i}} (\widehat{X}_i)^2 \right) + o(\varepsilon^{\gamma-\alpha}), \quad \alpha = \max\{\alpha_0, 2\alpha_1, \dots, 2\alpha_k\}.$$

If we can find suitable coordinates and a good choice of the weights, so that all the vector fields X_0 and X_i rescale with the same degree γ, then for every $0 < \varepsilon \le 1$ the principal part of

$$L_\varepsilon := \varepsilon^\gamma \left(\delta_{1/\varepsilon *} X_0 + \frac{1}{2} \sum_{i=1}^k (\delta_{1/\varepsilon *} X_i)^2 \right) \tag{2.5}$$

reduces to

$$L_0 = \widehat{X}_0 + \frac{1}{2} \sum_{i=1}^k \widehat{X}_i^2, \tag{2.6}$$

and we are able to write a decomposition of L_ε into a principal part L_0 plus a remainder term $\mathcal{X}_\varepsilon$ going to zero for $\varepsilon \to 0$.

2.1. **Duhamel's perturbation formula.** In this section we give only a brief overview of the method to keep the paper self-contained. The interested reader is referred to [35, 49, 7] for a more comprehensive exposition.

Let $\mathcal{L}$ be an operator that acts on $L^2(M,\mu)$ and admits a fundamental solution p. Then we may consider the evolution operator $e^{\mathcal{L}} : \mathbb{R}^+ \times L^2(M,\mu) \to L^2(M,\mu)$, defined by

$$e^{t\mathcal{L}}\varphi(x) := e^{\mathcal{L}}(t,\varphi)(x) = \int_M p(t,x,y)\varphi(y) d\mu(y).$$

Moreover, in this context, we refer to p as an *heat kernel* (or a *Schwartz kernel*) for the operator $e^{\mathcal{L}}$. It is then easy to show that, for every $\varphi \in L^2(M,\mu)$, there hold both

$$\partial_t e^{t\mathcal{L}}\varphi = \mathcal{L} e^{t\mathcal{L}}\varphi \quad \text{and} \quad \lim_{t\to 0} e^{t\mathcal{L}}\varphi = \varphi.$$

In particular $e^{\mathcal{L}}$ is a *heat operator*. Consider now the case where $\mathcal{L}$ admits a decomposition

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{X},$$

into a principal part plus a perturbation. If also $\mathcal{L}_0$ admits a fundamental solution p_0, we may reconstruct $e^{t\mathcal{L}}$ from $e^{t\mathcal{L}_0}$ as follows.

Theorem 6 (Duhamel)**.** *For every $t \ge 0$ the following identity holds:*

$$e^{t\mathcal{L}} = e^{t\mathcal{L}_0} + \int_0^t e^{(t-s)\mathcal{L}} \mathcal{X} e^{s\mathcal{L}_0} ds = e^{t\mathcal{L}_0} + e^{t\mathcal{L}} * \mathcal{X} e^{t\mathcal{L}_0}.$$

Here $*$ denotes the convolution of two operators $A(t,\cdot)$ and $B(t,\cdot)$, acting on the Hilbert space $L^2(M,\mu)$, and reads as follows:

$$(A*B)(t,\varphi)=\left(\int_0^t A(t-s)B(s)ds\right)(\varphi):L^2(M,\mu)\to L^2(M,\mu).$$

Moreover, assume that both $A(t,\cdot)$ and $B(t,\cdot)$ possess an heat kernel, denoted by $a(t,x,y)$ and $b(t,x,y)$ respectively. Then a heat kernel of $(A*\mathcal{X}B)(t)$ can be computed as:

$$(a*\mathcal{X}b)(t,x,y):=\int_0^t\int_M a(s,x,z)\mathcal{X}_z b(t-s,z,y)d\mu(z)ds,$$

where the operator $\mathcal{X}$ acts on the z-variable within the integral.

Choosing $A(t)=e^{t\mathcal{L}}$ and $B(t)=e^{t\mathcal{L}_0}$, Theorem 6 permits to write

$$p(t,x,y)=p_0(t,x,y)+(p*\mathcal{X}p_0)(t,x,y)$$

or, more precisely

$$p(t,x,y)=p_0(t,x,y)+\sum_{i=1}^r(p_0*^i\mathcal{X}p_0)(t,x,y)+(p*^{r+1}\mathcal{X}p_0)(t,x,y),$$

where $*^i\mathcal{X}p_0$ indicates the i-th repeated convolution.

2.2. **Perturbation of the fundamental solution.** We have seen how, by means of an appropriate weighted dilation δ_ε, for every $0<\varepsilon\le 1$ we can express L_ε in (2.5) as the sum $L_\varepsilon=L_0+(L_\varepsilon-L_0)=L_0+\mathcal{X}_\varepsilon$. If additionally L_0 admits a fundamental solution $q_0(t,x,y)$, then Duhamel's formula implies that

$$q_\varepsilon(t,x_0,x_0)=q_0(1,x_0,x_0)+\sum_{i=1}^r(q_0*^i\mathcal{X}_\varepsilon q_0)(1,x_0,x_0)+(q_\varepsilon*^{r+1}\mathcal{X}_\varepsilon q_0)(1,x_0,x_0). \tag{2.7}$$

Since by (2.4) one has $p(t,x,y)=\varepsilon^{-\sum_{i=1}^k w_i}q_\varepsilon(\varepsilon^{-\gamma}t,\delta_{1/\varepsilon}x,\delta_{1/\varepsilon}y)$, and the dilation δ_ε is centered at x_0, we eventually find that

$$\begin{aligned}p(t,x_0,x_0)&=t^{-\frac{\sum_{i=1}^k w_i}{\gamma}}q_{\sqrt[\gamma]{t}}(1,x_0,x_0)\\&=t^{-\frac{\sum_{i=1}^k w_i}{\gamma}}\left(q_0+\sum_{j=1}^r(q_0*^j\mathcal{X}_{\sqrt[\gamma]{t}}q_0)+(q_{\sqrt[\gamma]{t}}*^{r+1}\mathcal{X}_{\sqrt[\gamma]{t}}q_0)\right)(1,x_0,x_0).\end{aligned} \tag{2.8}$$

2.3. **Graded coordinate chart.** Evidently, the role of the drift field X_0 is different from that of the X_i. Since the latter are applied twice as many times as the former, it is natural to assign X_0 the weight 2, and to all the other vector fields the weight 1. The length of any bracket $\Lambda\in\mathrm{Lie}\{X_0,\dots,X_k\}$ is then computed by:

$$|\Lambda|=2|\Lambda|_0+\sum_{i=1}^k|\Lambda|_i,\quad |0|=0,$$

where $|\Lambda|_j$ counts the number of occurrences of X_j within Λ.

By means of these weights, we construct a filtration $\{\mathcal{G}_i\}_{i\in\mathbb{N}}\subset TM$ as follows:

$$\mathcal{G}_0=\{0\},\quad \mathcal{G}_i=\mathrm{span}\{\Lambda\in\mathrm{Lie}\{X_0,\dots,X_k\}:|\Lambda|\le i\}.$$

Notice that $\mathcal{G}_i\subset\mathcal{G}_{i+1}$, $[\mathcal{G}_i,\mathcal{G}_l]\subset\mathcal{G}_{i+l}$ and that eventually, by Hörmander's condition (2.1), $\mathcal{G}_m$ becomes equal to TM, for some minimal integer m called the *step* of the filtration.

For a fixed point $x_0 \in M$, the collection $\{\mathcal{G}_i(x_0)\}_{i=1}^m$ stratifies $T_{x_0}M$. If we set $d_0 = 0$ and $d_i = \dim \mathcal{G}_i(x_0)$, then this stratification induces a particular choice of adapted coordinates around x_0. Concerning this construction, we report here only the results which are relevant for this paper. For an exhaustive description of the graded structure associated with such coordinates, and how it defines a grading also on the algebra of differential operators, we refer the interested reader to the original construction in [23] (see also [7, 47] for its relation with heat kernel).

Proposition 7 (Bianchini, Stefani)**.** *There exists a chart (U, x), centered at x_0, such that for every $1 \le i \le m$:*

a) $\mathcal{G}_i(x_0) = \mathrm{span}\{\partial_{x_1}, \dots, \partial_{x_{d_i}}\}$,
b) $D_{x_h}(x_0) = 0$, *for every differential operator* $D \in \mathcal{A}_i = \{Z_1 \dots Z_l, \mid Z_s \in \mathcal{G}_{i_s}(x_0)$ *and* $i_1 + \dots + i_l \le i\}$ *and every* $h > d_i$.

With coordinates chosen as in the previous proposition, we define a splitting

$$\mathbb{R}^n = \mathbb{R}^{k_1} \oplus \dots \oplus \mathbb{R}^{k_m}, \quad k_i = d_i - d_{i-1}.$$

Then, the action of the dilation $\delta_\varepsilon : U \to U$ can be described as follows:

$$\delta_\varepsilon(x) = (\varepsilon x^1, \dots, \varepsilon^m x^m), \quad x^i = (x_{d_{i-1}+1}, \dots, x_{d_i}).$$

In particular, if we further assume that

$$X_0(x_0) \in \mathrm{span}\left\{X_i(x_0), [X_i, X_j](x_0) \mid i, j \in \{1, \dots, k\}\right\}, \tag{2.9}$$

the choice of the weighted dilation δ_ε induced by Proposition 7 implies that

$$\mathrm{Lie}_{x_0}\left\{\left(\mathrm{ad}\,\widehat{X}_0\right)^j \widehat{X}_i \mid j \ge 0,\, i = 1, \dots, k\right\} = T_{x_0}M,$$

i.e., the Lie algebra generated by the principal parts of X_0 and X_i, for $1 \le i \le k$, still satisfies the Hörmander condition (2.1) in a neighborhood of x_0. As a consequence, the operator L_0 in (2.6) is hypoelliptic, and possesses a smooth fundamental solution p_0 around x_0.

2.4. **Convergence of the fundamental solution.** Let

$$\mathcal{N}(x_0) := \sum_{i=1}^m i k_i = \sum_{i=1}^m i\left(\dim(\mathcal{G}_i(x_0)) - \dim(\mathcal{G}_{i-1}(x_0))\right).$$

Within the adapted chart (U, x) discussed in Proposition 7, and with the associated choice of the weights, we have the following convergences (in the C^∞ topology on U)[1]

$$\varepsilon \delta_{1/\varepsilon *} X_i \to \widehat{X}_i, \quad \varepsilon^2 \delta_{1/\varepsilon *} X_0 \to \widehat{X}_0, \quad \frac{1}{\varepsilon^{\mathcal{N}(x_0)}} \delta_\varepsilon^*(\mu) \equiv \mu,$$

where $\widehat{X}$ denotes the principal part of the vector field X with respect to the weighted dilation δ_ε. Assume that, by an iterated use of Duhamel's formula (2.7), for any $0 < \varepsilon \le 1$ we are able to write

$$q_\varepsilon = q_0 + \varepsilon \mathcal{X}_1 + \varepsilon^2 \mathcal{X}_2 + \dots \quad \text{on } U.$$

Then by the Trotter-Kato theorem [38, 53] we derive the pointwise convergence for the heat evolution operators generated by L_ε, $0 < \varepsilon \le 1$, and L_0, that is

$$e^{tL_\varepsilon} \xrightarrow{\varepsilon \to 0} e^{tL_0}, \quad \forall\, t \in [0, +\infty).$$

[1]More generally, for every smooth measure μ, $\lim_{\varepsilon \to 0} \frac{1}{\varepsilon^{\mathcal{N}(x_0)}} \delta_\varepsilon^*(\mu) = c d^n \mathcal{L}$ where $d^n \mathcal{L}$ denotes the n-dimensional Lebesgue measure, and c is a real constant.

This, in turn, permits to conclude that $q_\varepsilon \to q_0$ in the (weak) topology of distributions $\mathcal{D}'(\mathbb{R}^+ \times U \times U)$. If we can show that q_ε remains uniformly bounded in $C^\infty(\mathbb{R}^+ \times U \times U)$ topology as ε goes to zero, we can conclude that the convergence $q_\varepsilon \to q_0$ takes place also in the (strong) C^∞ topology. This is a consequence of the fact that $C^\infty(\mathbb{R}^+ \times U \times U)$ is a Montel space, and that in a Montel space every closed and bounded subset is necessarily compact.[2]

However, we observe that, in general, uniform estimates for q_ε in the C^∞ topology are usually delicate and rather hard to obtain.

3. Perturbing the Kolgomorov operator

In this section M denotes either $\mathbb{R}^2$, or a smooth, connected and orientable manifold of dimension 2. Let also μ be a given smooth volume form (or smooth density) on M.

The goal of this section is the study of the small-time asymptotics on the diagonal at a point $x_0 \in M$ of a fundamental solution p of the perturbed Kolmogorov operator

$$\partial_t - L = \partial_t - X_0 - \frac{1}{2}\left(X_1^2 + \mathrm{div}_\mu(X_1)X_1\right) \tag{3.1}$$

defined on $\mathbb{R}\times M$, with the assumptions that X_0, X_1 are smooth vector fields on M with bounded derivatives of all orders, and that for every $x \in M$

$$\dim\left(\mathrm{span}_x\{X_1, [X_0, X_1]\}\right) = 2. \tag{3.2}$$

The volume form μ is defined by the equality $\mu(X_1, [X_0, X_1]) = 1$. To apply the technical machinery developed in the previous section, we interpret in what follows $Y_0 := X_0 + \frac{1}{2}\mathrm{div}_\mu(X_1)X_1$ as a new drift.

In order to avoid exponential behaviors [20, 18] for $p(t, x_0, x_0)$, $t > 0$, and to recover meaningful geometrical informations from this expansion, we assume $X_0(x_0) \in \mathrm{span}\{X_1(x_0)\}$.

Fix a neighborhood V of x_0, on which X_1 and $[X_0, X_1]$ are linearly independent, and let $U \subset V$ be a coordinate chart centered at x_0, constructed as in Proposition 7, compactly contained in V. With the aid of a localization argument on the vector fields X_0 and X_1 within U, we may assume that

$$X_1\big|_U = \partial_{x_1}, \quad X_0\big|_U = \alpha_1\partial_{x_1} + \alpha_2\partial_{x_2}, \quad \alpha_1, \alpha_2 \in C^\infty(U), \tag{3.3}$$

with $\alpha_2(0) = 0$ and $\partial_{x_1}\alpha_2(0) \neq 0$. Accordingly, the condition $\mu(X_1, [X_0, X_1]) = 1$ translates into

$$\mu = -\frac{1}{(\partial_{x_1}\alpha_2)} dx_1 \wedge dx_2. \tag{3.4}$$

Consider the localized version $\partial_t - L\big|_U$ of the operator in (3.1). This means to consider $\partial_t - L\big|_U$ as an operator on $L^2(U, \mu\big|_U)$. Its domain is a subset of

$$\{u \in L^2(U, \mu\big|_U) \mid (\partial_t - L\big|_U)(u) \in L^2(U, \mu\big|_U)\},$$

possibly encoding informations on $u\big|_{\partial U}$ if U has a non-trivial boundary (e.g. Dirichlet boundary conditions).

This operator retains all the properties of the former, and as such it possesses a well-defined fundamental solution $p^U(t, x, y) \in C^\infty(\mathbb{R}^+ \times U \times U)$ on its own.

The following result asserts that $p(t, x, y) = p^U(t, x, y) + O(t^\infty)$ on U. This is an instance of the so-called *principle of not feeling the boundary* [37, 33]. The proof presented here was suggested to us from Emmanuel Trélat and is inspired by previous arguments in [34], which use only the Hörmander condition to establish the locality of the heat kernel on the diagonal.

[2]This argument is borrowed from the ongoing work [28], with the agreement of the authors.

This says in particular, if one is interested in a finite Taylor expansion of $p(t, x_0, x_0)$, it is not restrictive to work on a neighborhood of x_0.

Proposition 8. *Let $x_0 \in M$ and $U \subset M$ be a neighborhood of x_0 such that*

$$\dim\left(\operatorname{span}_x\{X_1, [X_0, X_1]\}\right) = 2, \quad \textit{for every } x \in U.$$

Then, for all $(l, \alpha, \beta) \in \mathbb{N} \times \mathbb{N}^2 \times \mathbb{N}^2$, every compact set $K \subset U$ and every $N \in \mathbb{N} \setminus \{0\}$, there exist $t_0 > 0$ and $c_N > 0$ such that

$$\left| \left(\partial_t^l \partial_x^\alpha \partial_y^\beta (p - p^U)\right)(t, x, y) \right| \leq c_N t^N,$$

for every $0 < t < t_0$, $(x, y) \in K \times K$.

Proof. [3] Extend at first both p and p^U by 0 if $t < 0$. Define $g(t, x, y) := p(t, x, y) - p^U(t, x, y)$, for $(t, x, y) \in \mathbb{R} \times U \times U$. Notice that g is smooth on $\mathbb{R}^+ \times U \times U$. We employ the notation L^x or L^y, meaning that we regard L as a differential operator acting either on the x or on the y variable respectively. By the properties defining a fundamental solution (compare with Definition 3), for any fixed $y \in U$ we deduce that $(\partial_t - L^x)(g(t, x, y)) = \delta_{(0,y)}(t, x)$ on $\mathbb{R} \times U$ in the sense of distributions. Since as soon as $p(t, x, y)$ is a fundamental solution to L^x, then so is $p(t, y, x)$ to the operator L^y, the same argument proves that, for any fixed $x \in U$, $(\partial_t - L^y)(g(t, y, x)) = \delta_{(0,x)}(t, y)$ on $\mathbb{R} \times U$. Defining $Y := L^x + L^y - 2\partial_t$, we infer that $Y(g) = 0$. Due to the hypoellipticity of Y, g is smooth. Since it vanishes, together with its derivatives, for $t < 0$, the same happens at $t = 0$ thus completing the proof. □

As a consequence of Proposition 8, it is then not restrictive for what concerns the finite-order, small-time asymptotics of p, to work within U with the localized version of (3.1). For the sake of clarity, with a slight abuse of notation we will omit the restriction sign in the sequel. The weighted dilation δ_ε, ensuring for X_0 twice the weight of X_1 as explained in Section 2, is given by

$$\delta_\varepsilon(x_1, x_2) = (\varepsilon x_1, \varepsilon^3 x_2).$$

The volume μ rescales then by a factor ε^4, and, by choosing $\gamma = 2$ in Proposition 5, a fundamental solution $q_\varepsilon(t, x, y)$ of the operator

$$\partial_t - L_\varepsilon := \partial_t - \varepsilon^2 \left(\delta_{1/\varepsilon *} X_0 + \frac{1}{2} (\delta_{1/\varepsilon *} X_1)^2 \right), \quad 0 < \varepsilon \leq 1,$$

is equal to $q_\varepsilon(t, x, y) = \varepsilon^4 p(\varepsilon^2 t, \delta_\varepsilon x, \delta_\varepsilon y)$. In particular, this construction also ensures that the operator L_0, appearing as the principal part of the limit $\lim_{\varepsilon \to 0^+} L_\varepsilon$ is hypoelliptic, with q_0 as a fundamental solution (cf. also with (2.9)).

Using (3.3), we can expand L_ε as a polynomial series with respect to ε as follows

$$L_\varepsilon = L_0 + \varepsilon \mathcal{X} + \varepsilon^2 \mathcal{Y} + \mathcal{Z}_{\varepsilon^3},$$

[3] This argument is borrowed from the ongoing work [28], with the agreement of the authors.

where

$$
\begin{aligned}
L_0 &= \left(x_1 \partial_{x_1} \alpha_2(0) \partial_{x_2} + \frac{1}{2} \partial_{x_1}^2 \right), \\
\mathcal{X} &= \left(\alpha_1(0) \partial_{x_1} + \frac{1}{2} x_1^2 \partial_{x_1}^2 \alpha_2(0) \partial_{x_2} - \frac{1}{2} \frac{\partial_{x_1}^2 \alpha_2(0)}{\partial_{x_1} \alpha_2(0)} \partial_{x_1} \right), \\
\mathcal{Y} &= \left(x_1 \partial_{x_1} \alpha_1(0) \partial_{x_1} + x_2 \partial_{x_2} \alpha_2(0) \partial_{x_2} + \frac{1}{6} x_1^3 \partial_{x_1}^3 \alpha_2(0) \partial_{x_2} \right. \\
&\qquad \left. - \frac{x_1}{2} \left(\frac{\partial_{x_1}^3 \alpha_2(0)}{\partial_{x_1} \alpha_2(0)} - \left(\frac{\partial_{x_1}^2 \alpha_2(0)}{\partial_{x_1} \alpha_2(0)} \right)^2 \right) \partial_{x_1} \right).
\end{aligned}
$$

Moreover, the term $\mathcal{Z}_{\varepsilon^3}$ can be presented as a smooth vector field on U in the form

$$
\mathcal{Z}_{\varepsilon^3} = \varepsilon^3 \left(\beta_1^\varepsilon(x_1, x_2) \partial_{x_1} + \beta_2^\varepsilon(x_1, x_2) \partial_{x_2} \right),
$$

where β_1^ε and β_2^ε are smooth functions on U that remain uniformly bounded for all values $0 < \varepsilon \leq 1$. The remaining point, as anticipated in Section 2.4, will be to show that the sequence of fundamental solutions q_ε, $0 < \varepsilon \leq 1$, associated with the (localized) rescaled operators $\partial_t - L_\varepsilon$, properly converges to the fundamental solution q_0 of $\partial_t - L_0$. We show in Theorem 12 that, in our case, this follows explicitly from the computations and the Gaussian nature of q_0, which permits in particular to control all its derivatives.

3.1. **Linear-quadratic operators.** The principal part $\partial_t - L_0$ of the operator $\partial_t - L_\varepsilon$ is an hypoelliptic operator of the form

$$
\partial_t \varphi - \sum_{i=1}^n (Ax)_i \partial_{x_i} \varphi - \frac{1}{2} \sum_{j,k=1}^n (BB^*)_{jk} \partial_{x_j, x_k}^2 \varphi, \quad \varphi \in \mathcal{D}'(\mathbb{R} \times \mathbb{R}^n), \tag{3.5}
$$

where A is a real $n \times n$ real matrix, B is a $n \times k$ matrix, with $k \leq n$, and whose hypoellipticity is guaranteeed by *Kalman's controllability condition*

$$
\operatorname{rank} \left[B, AB, \ldots, A^{n-1} B \right] = n. \tag{3.6}
$$

Its fundamental solution is classical in the literature and has the form of a Gaussian density as presented below. The explicit expression of the fundamental solution is classical, while the coefficients of the full small-time heat kernel asymptotic and their relations with curvature-like invariants have been considered in [11].

Proposition 9. *A fundamental solution with respect to the Lebesgue measure on $\mathbb{R}^n$ of the differential operator* (3.5) *is given by*

$$
q_0(t, x, y) = \frac{e^{-\frac{1}{2}(y - e^{tA}x)^* D_t^{-1} (y - e^{tA}x)}}{(2\pi)^{\frac{n}{2}} \sqrt{\det(D_t)}}, \quad \textit{with} \quad D_t = e^{tA} \int_0^t e^{-\tau A} BB^* e^{-\tau A^*} d\tau \, e^{tA^*}.
$$

Remark. Observe that Kalman's condition (3.6) ensures indeed that D_t is invertible, as soon as $t > 0$. For future purposes we also define the matrix

$$
\Gamma_t = \left(\int_0^t e^{-A\tau} BB^* e^{-A^* \tau} d\tau \right).
$$

Finally, notice that for our problem we have

$$
A = \begin{pmatrix} 0 & 0 \\ \partial_{x_1} \alpha_2(0) & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix},
$$

and that $\det(\Gamma_t) = \det(D_t)$ for all $t \in [0, +\infty)$.

3.2. **Main terms of the expansion.** Throughout this section we denote $S = \partial_{x_1}\alpha_2(0)$, which is non-zero by the assumption (3.2).

Iterating Duhamel's formula (2.7) three times, we obtain the following expression for $q_\varepsilon(1,0,0)$, namely:

$$q_\varepsilon(1,0,0) = q_0(1,0,0) + \varepsilon(q_0 * \mathcal{X}q_0)(1,0,0) + \varepsilon^2(q_0 * \mathcal{X}q_0 * \mathcal{X}q_0 + q_0 * \mathcal{Y}q_0)(1,0,0) + \mathcal{R}_{\varepsilon^3}. \tag{3.7}$$

The leading term $q_0(1,0,0)$ is simply evaluated, with the notations of Proposition 9 by the formula

$$q_0(1,0,0) = \frac{1}{2\pi\sqrt{\det(D_1)}}, \quad \text{where} \quad \det(D_1) = \det\begin{pmatrix} 1 & \frac{S}{2} \\ \frac{S}{2} & \frac{S^2}{3} \end{pmatrix} = \frac{S^2}{12}.$$

The computation of the convolutions

$$(q_0 * \mathcal{X}q_0)(1,0,0), \quad (q_0 * \mathcal{X}q_0 * \mathcal{X}q_0)(1,0,0) \quad \text{and} \quad (q_0 * \mathcal{Y}q_0)(1,0,0),$$

needs a technical lemma which is crucial in what follows.

Lemma 10. *For every pair $(z,w) \in \mathbb{R}^2 \times \mathbb{R}^2$, $0 \le s \le 1$ and $0 \le r \le 1-s$, the following identity holds*

$$\frac{e^{-\frac{1}{2}(w-e^{rA}z)^* D_r^{-1}(w-e^{rA}z)}}{2\pi\sqrt{\det(D_r)}} \cdot \frac{e^{-\frac{1}{2}w^*\Gamma_{1-s-r}^{-1}w}}{2\pi\sqrt{\det(D_{1-s-r})}} = \frac{e^{-\frac{1}{2}z^*\Gamma_{1-s}^{-1}z}}{2\pi\sqrt{\det(\Gamma_{1-s})}} \cdot \frac{e^{-\frac{1}{2}(w-\nu)^*\Sigma^{-1}(w-\nu)}}{2\pi\sqrt{\det(\Sigma)}},$$

with $\Sigma = \Gamma_{1-s-r}e^{-rA^*}\Gamma_{1-s}^{-1}\Gamma_r e^{rA^*}$ and $\nu = \Gamma_{1-s-r}e^{-rA^*}\Gamma_{1-s}^{-1}z$.

Proof. The product of two Gaussian densities is again a Gaussian density, according to the following classical identity

$$(v-a)^*\mathcal{A}(v-a) + (v-b)^*\mathcal{B}(v-b) = (v-c)^*(\mathcal{A}+\mathcal{B})(v-c) + (a-b)^*\mathcal{C}(a-b),$$

which holds for every $v, a, b \in \mathbb{R}^2$, $\mathcal{A}, \mathcal{B}$ square 2×2 real matrices, and for $\mathcal{C} := (\mathcal{A}^{-1}+\mathcal{B}^{-1})^{-1} \in \mathrm{M}_2(\mathbb{R})$ and $c := (\mathcal{A}+\mathcal{B})^{-1}(\mathcal{A}a+\mathcal{B}b) \in \mathbb{R}^2$.

In our case, we apply this identity with (compare with Remark 3.1)

$$\mathcal{A} := \Gamma_{1-s-r}^{-1}, \quad \mathcal{B} := D_r^{-1}, \quad a := 0, \quad b := e^{rA}z,$$

and the only non-trivial identity to prove is the following:

$$\begin{aligned} \mathcal{A}^{-1} + \mathcal{B}^{-1} &= D_r + \Gamma_{1-s-r} \\ &= e^{rA}\left(\Gamma_r + e^{-rA}\Gamma_{1-s-r}e^{-rA^*}\right)e^{rA^*} \\ &= e^{rA}\left(\Gamma_r + e^{-rA}\int_r^{1-s} e^{(-\tau+r)A}BB^*e^{(-\tau+r)A^*}d\tau\, e^{-rA^*}\right)e^{rA^*} \\ &= e^{rA}\Gamma_{1-s}e^{rA^*}. \end{aligned} \tag{3.8}$$

Indeed, once (3.8) is established, one computes $\Gamma_{1-s-r}^{-1} + D_r^{-1} = \mathcal{A}+\mathcal{B}$ using the identity $\mathcal{A}+\mathcal{B} = \mathcal{B}\left(\mathcal{A}^{-1}+\mathcal{B}^{-1}\right)\mathcal{A}$. The claim then follows from a routine computation. □

Theorem 11. *The following list of equalities holds true:*

$$
\begin{aligned}
(q_0 * \partial_{x_1} q_0)(1,0,0) &= 0,\\
(q_0 * x_1^2 \partial_{x_2} q_0)(1,0,0) &= 0,\\
(q_0 * \partial_{x_1} q_0 * \partial_{x_1} q_0)(1,0,0) &= \frac{1}{2\pi\sqrt{\det(D_1)}}\left(-\frac{1}{2}\right),\\
(q_0 * x_1^2 \partial_{x_2} q_0 * \partial_{x_1} q_0)(1,0,0) &= \frac{1}{2\pi\sqrt{\det(D_1)}}\left(\frac{3}{14S}\right),\\
(q_0 * \partial_{x_1} q_0 * x_1^2 \partial_{x_2} q_0)(1,0,0) &= \frac{1}{2\pi\sqrt{\det(D_1)}}\left(-\frac{3}{14S}\right),\\
(q_0 * x_1 \partial_{x_1} q_0)(1,0,0) &= \frac{1}{2\pi\sqrt{\det(D_1)}}\left(-\frac{1}{2}\right),\\
(q_0 * x_2 \partial_{x_2} q_0)(1,0,0) &= \frac{1}{2\pi\sqrt{\det(D_1)}}\left(-\frac{1}{2}\right),\\
(q_0 * x_1^3 \partial_{x_2} q_0)(1,0,0) &= \frac{1}{2\pi\sqrt{\det(D_1)}}\left(-\frac{3}{14S}\right),\\
(q_0 * x_1^2 \partial_{x_2} q_0 * x_1^2 \partial_{x_2} q_0)(1,0,0) &= \frac{1}{2\pi\sqrt{\det(D_1)}}\left(\frac{9}{70S^2}\right).
\end{aligned}
$$

In particular, we have that

$$
(q_0 * \mathcal{X} q_0)(1,0,0) = 0,
$$

$$
\begin{aligned}
(q_0 * \mathcal{X} q_0 * \mathcal{X} q_0 + q_0 * \mathcal{Y} q_0)(1,0,0) = \frac{1}{2\pi\sqrt{\det(D_1)}}\Bigg(&-\frac{1}{2}\alpha_1^2(0) - \frac{12}{35}\left(\frac{\partial_{x_1}^2 \alpha_2(0)}{\partial_{x_1}\alpha_2(0)}\right)^2\\
&-\frac{1}{2}\partial_{x_1}\alpha_1(0) - \frac{1}{2}\partial_{x_2}\alpha_2(0) + \frac{3}{14}\frac{\partial_{x_1}^3\alpha_2(0)}{\partial_{x_1}\alpha_2(0)} - \frac{1}{2}\alpha_1(0)\frac{\partial_{x_1}^2\alpha_2(0)}{\partial_{x_1}\alpha_2(0)}\Bigg).
\end{aligned}
$$

Proof. We show as an example how to find the value of $(q_0 * \partial_{x_1} q_0)(1,0,0)$ and $(q_0 * \partial_{x_1} q_0 * \partial_{x_1} q_0)(1,0,0)$, the other cases being all similar. In the following we tacitly assume that the values of all the moments up to the sixth order of a Gaussian distribution are known, and we constantly use the equality

$$
\partial_{x_i} q_0(t,x,y) = -\left[\Gamma_t^{-1}\left(x - e^{-At}y\right)\right]_i q_0(t,x,y), \quad \text{for} \quad i = 1,2.
$$

Here $[v]_i$ for $i = 1,2$ denotes the i-th component of a vector $v \in \mathbb{R}^2$. Working in a local chart, it is not restrictive to evaluate all the integrals appearing throughout the proof as if they were on $\mathbb{R}^2$. For the first one, using Lemma 10, we have

$$
\begin{aligned}
(q_0 * \partial_{x_1} q_0)(1,0,0) &= \int_0^1 \int_{\mathbb{R}^2} q_0(s,0,z)\partial_{x_1} q_0(1-s,z,0)\,dz\,ds\\
&= -\int_0^1 \int_{\mathbb{R}^2} \left[\Gamma_{1-s}^{-1}(z)\right]_1 \frac{e^{-\frac{1}{2}z^* D_s^{-1} z}}{2\pi\sqrt{\det(D_s)}} \frac{e^{-\frac{1}{2}z^* \Gamma_{1-s}^{-1} z}}{2\pi\sqrt{\det(D_{1-s})}}\,dz\,ds = 0,
\end{aligned}
$$

since the product of two Gaussian densities of zero mean is again a Gaussian density of zero mean, $\left[\Gamma_{1-s}^{-1}(z)\right]_1$ is a polynomial of degree 1 in z, and all the centered moments of odd order of a zero-mean Gaussian are zero.

For the second one we proceed as follows: by Lemma 10 we have the identity

$$q_0(r,z,w)q_0(1-s-r,w,0) = \frac{e^{-\frac{1}{2}z^*\Gamma_{1-s}^{-1}z}}{2\pi\sqrt{\det(\Gamma_{1-s})}} \cdot \frac{e^{-\frac{1}{2}(w-\nu)^*\Sigma^{-1}(w-\nu)}}{2\pi\sqrt{\det(\Sigma)}},$$

with $\Sigma := \Gamma_{1-s-r}e^{-rA^*}\Gamma_{1-s}^{-1}\Gamma_r e^{rA^*}$ and $\nu := \Gamma_{1-s-r}e^{-rA^*}\Gamma_{1-s}^{-1}z$.

On a second step, using similar arguments to those exploited above, one computes the product

$$\frac{e^{-\frac{1}{2}z^*D_s^{-1}z}}{2\pi\sqrt{\det(D_s)}} \cdot \frac{e^{-\frac{1}{2}z^*\Gamma_{1-s}^{-1}z}}{2\pi\sqrt{\det(D_{1-s})}} = \frac{1}{2\pi\sqrt{\det(\Gamma_1)}} \cdot \frac{e^{-\frac{1}{2}z^*\Sigma'^{-1}z}}{2\pi\sqrt{\det(\Sigma')}},$$

where $\Sigma' := \Gamma_{1-s}e^{-sA^*}\Gamma_1^{-1}\Gamma_s e^{sA}$.

From here we deduce (for Σ, ν and Σ' defined as above):

$$\begin{aligned}
&(q_0 * \partial_{x_1}q_0 * \partial_{x_1}q_0)(1,0,0) = \\
&\int_0^1\int_{\mathbb{R}^2} q_0(s,0,z)\int_0^{1-s}\int_{\mathbb{R}^2}\partial_{x_1}q_0(r,z,w)\partial_{x_1}q_0(1-s-r,w,0)dwdrdzds \\
&= \int_0^1\int_{\mathbb{R}^2} q_0(s,0,z)\frac{e^{-\frac{1}{2}z^*\Gamma_{1-s}^{-1}z}}{2\pi\sqrt{\det(D_{1-s})}}\times \\
&\qquad\qquad \times\int_0^{1-s}\int_{\mathbb{R}^2}\left[\Gamma_r^{-1}\left(z-e^{-rA}w\right)\right]_1\left[\Gamma_{1-s-r}^{-1}w\right]_1\frac{e^{-\frac{1}{2}(w-\nu)^*\Sigma^{-1}(w-\nu)}}{2\pi\sqrt{\det(\Sigma)}}dwdrdzds \\
&= \int_0^1\int_{\mathbb{R}^2}\frac{e^{-\frac{1}{2}z^*D_s^{-1}z}}{2\pi\sqrt{\det(D_s)}}\cdot\frac{e^{-\frac{1}{2}z^*\Gamma_{1-s}^{-1}z}}{2\pi\sqrt{\det(D_{1-s})}}Q(z)dzds \\
&= \frac{1}{2\pi\sqrt{\det(\Gamma_1)}}\int_0^1\int_{\mathbb{R}^2}\frac{e^{-\frac{1}{2}z^*\Sigma'^{-1}z}}{2\pi\sqrt{\det(\Sigma')}}Q(z)dzds,
\end{aligned}$$

where $Q(z)$ is a polynomial in z that can be determined explicitly and whose coefficients are rational functions in $1-s$. Their expression depends on the moments of even order of a Gaussian density. To arrive to the results of the theorem is now just a routine computation. □

3.3. **Bounds on the remainder term.** Iterating three times Duhamel's formula, and collecting all the terms containing at least the factor ε^3, wee see that the remainder term $\mathcal{R}_{\varepsilon^3}$ in (3.7) is a finite sum of items of these three kinds:

a) $(q_0 * \varepsilon^{\alpha_1}\mathcal{D}_1 q_0)(1,0,0)$, with $\alpha_1 \geq 3$, or

b) $(q_0 * \varepsilon^{\beta_1}\mathcal{D}_1 q_0 * \varepsilon^{\beta_2}\mathcal{D}_2 q_0)(1,0,0)$, with $\beta_1+\beta_2 \geq 3$, or

c) $(q_\varepsilon * \varepsilon^{\gamma_1}\mathcal{D}_1 q_0 * \varepsilon^{\gamma_2}\mathcal{D}_2 q_0 * \varepsilon^{\gamma_3}\mathcal{D}_3 q_0)(1,0,0)$, with $\gamma_1+\gamma_2+\gamma_3 \geq 3$.

Here the $\mathcal{D}_i$ are chosen among a finite set of polynomial vector fields which are all supported and bounded on U, uniformly with respect to the value of $\varepsilon \in (0,1]$.

The objective of this section is to show that the ratio $\mathcal{R}_\varepsilon^3/\varepsilon^3$ remains uniformly bounded in ε, for such values of the parameter. To this end, it suffices to control, uniformly with respect to ε, the absolute values of all the terms appearing in each of the three families.

For what concerns summands belonging either to a) or b), the result follows easily applying the same kind of computations exploited in Theorem 11. For items appearing in c), their uniform estimate is a consequence of the following result.

Theorem 12. *Let f_1, f_2, f_3 be smooth, bounded functions supported on U, and $i, j, k \in \{1, 2\}$. Then there exists $C > 0$, depending on all the previous data, such that*

$$\left|\left(q_\varepsilon * f_1(x_1,x_2)\partial_{x_i} q_0 * f_2(x_1,x_2)\partial_{x_j} q_0 * f_3(x_1,x_2)\partial_{x_k} q_0\right)(1,0,0)\right| \leq C.$$

Proof. Again all calculations will be carried out on $\mathbb{R}^2$, our statement being local. Begin with the obvious inequality (recall that $q_\varepsilon \geq 0$, being a probability density)

(3.9)
$$\begin{aligned}&\left|\left(q_\varepsilon * f_1(x_1,x_2)\partial_{x_i} q_0 * f_2(x_1,x_2)\partial_{x_j} q_0 * f_3(x_1,x_2)\partial_{x_k} q_0\right)(1,0,0)\right| \leq \\ &C_1 \int_0^1 \int_{\mathbb{R}^2} q_\varepsilon(s,0,z)\Bigg(\int_0^{1-s}\int_{\mathbb{R}^2} |\partial_{x_i} q_0(r,z,w)| \Bigg(\int_0^{1-s-r}\int_{\mathbb{R}^2} \left|\partial_{x_j} q_0(l,w,t)\right| \times \\ &\qquad \times |\partial_{x_k} q_0(1-s-r-l,t,0)|\, dtdl\Bigg) dwdr\Bigg) dzds,\end{aligned}$$

with $C_1 := \sup_U |f_1| \cdot \sup_U |f_2| \cdot \sup_U |f_3|$. Using Lemma 10 iteratively together with an argument close to the one contained in the proof of Theorem 11, one can treat all the nested integrals in (3.9), and majorize them by

$$C_1 \int_0^1 \int_{\mathbb{R}^2} q_\varepsilon(s,0,z) P(|z_1|,|z_2|) \frac{e^{-\frac{1}{2} z^* \Gamma_{1-s}^{-1} z}}{2\pi\sqrt{\det(\Gamma_{1-s})}} dzds,$$

where P is a polynomial in $|z_1|, |z_2|$ whose coefficients are rational functions of $(1-s)$. Notice that P can be explicitly computed in terms of the absolute moments of a Gaussian distribution. We now claim that there exists a positive constant C_2 such that

$$P(|z_1|,|z_2|) \frac{e^{-\frac{1}{2} z^* \Gamma_{1-s}^{-1} z}}{2\pi\sqrt{\det(\Gamma_{1-s})}} \leq C_2 \quad \text{on} \quad [0,1] \times \mathbb{R}^2. \tag{3.10}$$

In fact, recalling the notation $S = \partial_{x_1}\alpha_2(0)$, we have

$$e^{-\frac{1}{2} z^* \Gamma_{1-s}^{-1} z} = e^{\frac{2}{S^2(s-1)^3}(S^2(s-1)^2 z_1^2 - 3S(s-1) z_1 z_2 + 3 z_2^2)}.$$

In particular, as $s \nearrow 1$, $e^{-\frac{1}{2} z^* \Gamma_{1-s}^{-1} z} \sim e^{\frac{6 z_2^2}{S^2(s-1)^3}}$ on the set $\{z_2 \neq 0\}$, while $e^{-\frac{1}{2} z^* \Gamma_{1-s}^{-1} z} \sim e^{\frac{2 z_1^2}{(s-1)}}$ on $\{z_2 = 0\}$. Then (3.10) remains uniformly bounded on $[0,1] \times \mathbb{R}^2$ as claimed.

The integral of $q_\varepsilon(s,0,\cdot)$ is bounded independently on ε, since q_ε is a probability density. Hence, we conclude that

$$C_1 \int_0^1 \int_{\mathbb{R}^2} q_\varepsilon(s,0,z) P(|z_1|,|z_2|) \frac{e^{-\frac{1}{2} z^* \Gamma_{1-s}^{-1} z}}{2\pi\sqrt{\det(\Gamma_{1-s})}} dzds \leq C_1 \cdot C_2,$$

and the theorem follows with $C = C_1 \cdot C_2$. □

4. Curvature-like invariants

Let M be a two-dimensional smooth and connected manifold, and X_0, X_1 be a pair of smooth vector fields on M, such that for every $x \in M$

$$\operatorname{span}\{X_1(x), [X_0, X_1](x)\} = T_x M. \tag{4.1}$$

Consider the differential system

$$\dot{x} = X_0(x) + uX_1(x), \quad x(0) = x_0 \in M, \quad u \in \mathbb{R}, \tag{4.2}$$

and let $T > 0$. An *admissible control* $t \mapsto u(t)$ is an element of $L^2([0,T],\mathbb{R})$ such that the corresponding trajectory $x_u(\cdot)$, solution to (4.2), is defined on $[0,T]$. The set of admissible controls $\Omega^T_{x_0} \subset L^2([0,T],\mathbb{R})$ is an open set. If $u \in \Omega^T_{x_0}$, we say that $x_u(\cdot)$ is an *admissible trajectory*, and we define the *attainable set* from x_0 at time T as $A^T_{x_0} = \{x_u(T) \mid u \in \Omega_{x_0}\}$. By Sussmann-Jurdjevic Theorem [52, 36], $A^T_{x_0}$ has non-empty interior thanks to our assumptions.

Introduce now the energy functional $J_T : L^2([0,T],\mathbb{R}) \to \mathbb{R}$, defined by

$$J_T(u) = \frac{1}{2}\int_0^T |u(t)|^2 dt.$$

For a fixed final point $x \in M$, we are interested into solving the optimal control problem

$$S^T_{x_0}(x) = \inf\left\{J_T(u) \mid u \in \Omega^T_{x_0},\, x_u(T) = x\right\}, \tag{4.3}$$

with the convention that $S^T_{x_0}(x) = +\infty$ if $x \notin A^T_{x_0}$. The geodesic flow of this problem can be seen as a flow on T^*M, associated with the Hamiltonian

$$H(p,x) = \langle p, X_0(x)\rangle + \frac{1}{2}\langle p, X_1(x)\rangle^2, \quad (p,x) \in T^*M.$$

Then Hamilton's equations[4] are written in the form $\dot{\lambda} = \vec{H}(\lambda)$, where $\lambda = (p,x) \in T^*M$ and $\vec{H}$ is the Hamiltonian vector field associated with H. Integral curves of $\vec{H}$ are called *extremals*, and their projections on the manifold permit to find the so-called *normal* geodesics, whose short enough pieces solve the optimal problem (4.3) between their end-points. It is well-known that, under the assumption (4.1), the optimal control problem has no singular minimizers, i.e., every solution to the optimal problem satisfy this necessary condition. Moreover, the control u associated with a normal geodesic, called also an *optimal control*, is smooth on $[0,T]$. A detailed presentation on the subject may be found, for example, in [5].

4.1. **The canonical frame.** Given a geodesic $t \mapsto x(t)$, consider the *geodesic flag* associated with $x(t)$

$$\mathcal{F}^1_{x(t)} \subset \mathcal{F}^2_{x(t)} \subset \ldots \subset T_{x(t)}M, \quad t \in [0,T],$$

defined by $\mathcal{F}^i_{x(t)} = \mathrm{span}\{X_1, \ldots, X_i\}\big|_{x(t)}$ for every $i \in \mathbb{N}$. Here $X_j = \mathrm{ad}(X_0 + \alpha X_1)^{j-1}X_1$, and $\alpha \in C^\infty(M)$ is such that $\alpha(x(t)) = u(t)$ for every $t \in [0,T]$. The vector fields X_i are smooth along $x(t)$. We say that $x(\cdot)$ is *equiregular* at t if $\dim\mathcal{F}^i_{x(t)}$ is locally constant for every $i > 0$, and *ample* at t if there exists $m > 0$ such that $\mathcal{F}^m_{x(t)} = T_{x(t)}M$.

Under the assumption (4.1), it follows immediately that, in our setting, all geodesics are indeed ample and equiregular. For a comprehensive presentation of these notions in the general setting the reader is referred to [3, Chapter 3].

If $x(t)$ is ample and equiregular, then the geodesic flag terminates with $\mathcal{F}^2_{x(t)} = T_{x(t)}M$ for every $t \in [0,T]$, and is described with the basis X_1 and $X_2 = [X_1, X_0]$ along $x(t)$. Then the following result, which is a specification of a more general statement adapted to our context, holds true (cf. also [3, Chapter 7]).

[4]By means of the canonical symplectic form σ on T^*M we define, for every $f : T^*M \to \mathbb{R}$ and for every $\lambda \in T^*M$, the Hamiltonian lift $\vec{f}$ of f through the identity $\sigma_\lambda(\cdot, \vec{f}) = d_\lambda f(\cdot)$.

Proposition 13 (Zelenko-Li, [56]). *Assume that $\lambda(t)$ is the lift of an ample and equiregular geodesic $x(t)$. Then, there exists a smooth moving frame $\{E_1, F_1, E_2, F_2\}$ along $\lambda(t)$ such that, for every $t \in [0, T]$,*

i) $\pi_*(E_i)\big|_{\lambda(t)} = 0$, *where $\pi : T^*M \to M$ is the canonical projection.*
ii) $\sigma(E_i, E_j)\big|_{\lambda(t)} = \sigma(F_i, F_j)\big|_{\lambda(t)} = 0$, $\sigma(E_i, F_j)\big|_{\lambda(t)} = \delta_{ij}$, *where σ is the canonical symplectic form defined on the cotangent space T^*M. In particular the canonical frame is a Darboux basis for $T_{\lambda(t)}T^*M$.*
iii) *The structural equations satisfied by this frame are*

$$\dot{E}_2(t) = E_1(t), \qquad \dot{E}_1(t) = -F_1(t),$$
$$\dot{F}_2(t) = R_{22}(t)E_2(t), \qquad \dot{F}_1(t) = R_{11}(t)E_1(t) - F_2(t).$$

The remaining part of this section will be thus devoted to the computation of the term $R_{22}(t)$.

4.2. **Structural equations for the canonical frame.** Let us consider on M the smooth vector fields X_0, X_1 and denote $X_2 := [X_0, X_1]$. Define the associated fiber-wise linear functions $h_i : T^*M \to \mathbb{R}$ for $i = 0, 1, 2$, defined by $h_i(p, x) := \langle p, X_i(x)\rangle$.

Owing to the identity $\{h_i, h_j\}(p, x) = \langle p, [X_i, X_j](x)\rangle$, the Lie bracket relations between X_0, X_1 and X_2 translate into Poisson bracket relations as follows:

$$\begin{aligned} \{h_0, h_1\} &= h_2, \\ \{h_1, h_2\} &= c_{12}^1 h_1 + c_{12}^2 h_2, \\ \{h_0, h_2\} &= c_{02}^1 h_1 + c_{02}^2 h_2, \end{aligned} \tag{4.4}$$

for some structural functions $c_{ij}^k \in C^\infty(M)$.

The functions $h_1, h_2 : T^*M \to \mathbb{R}$ define a coordinate system on the fibers of T^*M. Let us denote by $\partial_{h_1}, \partial_{h_2}$ the corresponding vector fields. These vector fields are vertical, namely $\ker \pi_* = \mathrm{span}\{\partial_{h_1}, \partial_{h_2}\}$, where $\pi_* : T(T^*M) \to TM$ is the differential of the canonical projection $\pi : T^*M \to M$. We complete $\{\partial_{h_1}, \partial_{h_2}\}$ to a basis of $T(T^*M)$ by means of the Hamiltonian lifts $\vec{h}_1$ and $\vec{h}_2$.

In this formalism, the geodesic Hamiltonian H reads $H = h_0 + \frac{1}{2}h_1^2$, and its Hamiltonian lift becomes $\vec{H} = \vec{h}_0 + h_1\vec{h}_1$. For every integral curve $t \mapsto \lambda(t)$, $t \in [0, T]$ of the vector field $\vec{H}$, and every smooth function $f : T^*M \to \mathbb{R}$, we have

$$\dot{f}(\lambda(t)) = \vec{H}(f(\lambda(t))) = \{H, f\}(\lambda(t)),$$

the dot meaning the derivative with respect to t. On the other hand, for the vector fields $\vec{h}_i$, the differentiation along the geodesic lift becomes

$$\dot{\vec{h}}_i(\lambda(t)) = [\vec{H}, \vec{h}_i](\lambda(t)) = \overrightarrow{\{h_0, h_i\}} + h_1\overrightarrow{\{h_1, h_i\}} - \delta_{i1}\vec{h}_1,$$

the dot meaning here the Lie derivative with respect to $\vec{H}$. To compute the quantity $\overrightarrow{\{h_i, h_j\}}$ one uses the structural equations in (4.4). Moreover we recall that for $g \in C^\infty(T^*M)$ fiber-wise constant, we have

$$\dot{g}(\lambda(t)) = \{H, g\}(\lambda(t)) = (X_0 g)(\lambda(t)) + (X_1 g)(\lambda(t))h_1(\lambda(t)),$$

where in the right hand side g is treated as a function on the manifold. To compute $\dot{\partial}_{h_i}$, for $i = 1, 2$, one has

$$\dot{\partial}_{h_i}(\lambda(t)) = [\vec{h}_0 + h_1\vec{h}_1, \partial_{h_i}](\lambda(t)) = [\vec{h}_0, \partial_{h_i}](\lambda(t)) + [\vec{h}_1, \partial_{h_i}](\lambda(t)) - \delta_{i1}\vec{h}_1(\lambda(t)).$$

For the last term $[\vec{h}_j, \partial_{h_i}](\lambda(t))$ it is useful to observe that $[\vec{h}_1, \partial_{h_i}](\lambda(t))$ is vertical, that is $\pi_*[\vec{h}_1, \partial_{h_i}] = 0$, hence one can write

$$[\vec{h}_j, \partial_{h_i}](\lambda(t)) = [\vec{h}_j, \partial_{h_i}](h_1)\partial_{h_1}(\lambda(t)) + [\vec{h}_j, \partial_{h_i}](h_2)\partial_{h_2}(\lambda(t)).$$

4.3. **Computations.** We provide in this sections all the main computations. For clarity's sake, we systematically drop the explicit dependence on t along the extremal. The next two lemmas follows by direct computations as explained in Section 4.2.

Lemma 14 (Fundamental computations)**.** *Setting*

$$\begin{aligned} A &:= c^1_{12}h_1 + c^1_{02}, \\ B &:= c^2_{12}h_1 + c^2_{02}, \\ C &:= c^1_{12}h_1 + c^2_{12}h_2, \\ M_i &:= X_i(c^1_{02})h_1 + X_i(c^2_{02})h_2 + X_i(c^1_{12})h_1^2 + X_i(c^2_{12})h_1h_2, \quad i = 1, 2, \end{aligned}$$

along any extremal $\lambda(t)$, the following relations hold true:

$$\begin{aligned} \dot{\partial}_{h_1} &= -\vec{h}_1 - A\partial_{h_2}, \\ \dot{\partial}_{h_2} &= -\partial_{h_1} - B\partial_{h_2}, \\ \dot{\vec{h}}_1 &= \vec{h}_2, \\ \dot{\vec{h}}_2 &= (A + C)\vec{h}_1 + B\vec{h}_2 - M_1\partial_{h_1} - M_2\partial_{h_2}, \\ \ddot{\partial}_{h_2} &= \vec{h}_1 + B\partial_{h_1} + (A - \dot{B} + B^2)\partial_{h_2}, \\ \dddot{\partial}_{h_2} &= \vec{h}_2 - B\vec{h}_1 - (A - 2\dot{B} + B^2)\partial_{h_1} + (\dot{A} - \ddot{B} - 2AB + 3B\dot{B} - B^3)\partial_{h_2}. \end{aligned}$$

Lemma 15 (Computation of symplectic products)**.** *We have:*

$$\begin{aligned} \sigma(\partial_{h_1}, \vec{h}_1) &= \sigma(\partial_{h_2}, \vec{h}_2) = 1, \\ \sigma(\vec{h}_1, \vec{h}_2) &= \{h_1, h_2\} = C, \\ \sigma(\partial_{h_2}, \partial_{h_2}) &= 0, \\ \sigma(\partial_{h_2}, \dot{\partial}_{h_2}) &= 0, \\ \sigma(\partial_{h_2}, \ddot{\partial}_{h_2}) &= 0 \text{ implies also } \sigma(\dot{\partial}_{h_2}, \dot{\partial}_{h_2}) = 0, \\ \sigma(\partial_{h_2}, \dddot{\partial}_{h_2}) &= 1 \text{ implies also } \sigma(\dot{\partial}_{h_2}, \ddot{\partial}_{h_2}) = -1, \\ \sigma(\dot{\partial}_{h_2}, \dddot{\partial}_{h_2}) &= 0 \text{ implies also } \sigma(\ddot{\partial}_{h_2}, \ddot{\partial}_{h_2}) = 0. \end{aligned}$$

We are now able to compute the coefficient R_{22} in Proposition 13.

Proposition 16. *Along any extremal $\lambda(t)$, the following relation holds true:*

$$\begin{aligned} R_{22} = {} & h_1^2[-(c^2_{12})^2 + 3X_1(c^2_{12})] \\ & + h_1[-3c^1_{12} - 2c^2_{12}c^2_{02} + 3X_0(c^2_{12}) + 3X_1(c^2_{02})] \\ & + h_2[2c^2_{12}] \\ & + (-2c^1_{02} - (c^2_{02})^2 + 3X_0(c^2_{02})). \end{aligned} \tag{4.5}$$

Proof. The vector E_1 is uniquely determined by the conditions

(i) $\pi_* E_1 = 0$,
(ii) $\pi_* \dot{E}_1 = 0$,

(iii) $\sigma(\ddot{E}_1, \dot{E}_1) = 1$.

Conditions (i) and (ii) imply that $E_1 = f\partial_{h_2}$, so that

$$\dot{E}_1 = \dot{f}\partial_{h_2} + f\dot{\partial}_{h_2}, \qquad \ddot{E}_1 = \ddot{f}\partial_{h_2} + 2\dot{f}\dot{\partial}_{h_2} + f\ddot{\partial}_{h_2}.$$

From the previous formulas we find $\sigma(\ddot{E}_1, \dot{E}_1) = f^2$, and thus condition (iii) finally implies that $f = 1$. Notice that this means $E_1 = \partial_{h_2}$, $E_2 = \dot{\partial}_{h_2}$, $F_2 = -\ddot{\partial}_{h_2}$, and we eventually arrive to the expression

$$R_{22} = \sigma(\dot{F}_2, F_2) = \sigma(\dddot{\partial}_{h_2}, \ddot{\partial}_{h_2}).$$

By Lemmas 14 and 15, the latter expression reduces to

$$R_{22} = -C - 2A - B^2 + 3\dot{B}.$$

Recalling the fact that

$$\dot{B} = X_0(c_{02}^2) + (X_1(c_{02}^2) + X_0(c_{12}^2))h_1 + c_{12}^2 h_2 + X_1(c_{12}^2)h_1^2,$$

we conclude. □

Definition 17. We define K_1 and K_2 as the coefficients of h_1^2 and h_2, respectively, in the expression of R_{22}, that is

$$K_1 := -(c_{12}^2)^2 + 3X_1(c_{12}^2),$$
$$K_2 := 2c_{12}^2.$$

The coefficient K_1 can be interpreted as the contraction along the controlled vector field X_1 of the second-order part of the curvature operator associated with the control problem, that is a sort of scalar curvature. Indeed, in the case of a control problem associated to Riemannian geodesic problem, this quantity coincides, up to a scalar factor, to the scalar curvature of the Riemannian metric.

The coefficient K_2 plays the role of a "higher order" curvature. Its presence here is necessary to give a correct interpretation since the drift X_0 is necessary to fulfill the Hörmander condition.

5. Geometric interpretation of the small-time asymptotics

The purpose of this section is to interpret the first coefficient appearing in the small-time asymptotics of the heat kernel $p(t, x_0, x_0)$, by means of the curvature-like invariants that can be deduced from (4.5), and more in general in terms of the geometry associated with the control system (4.2).

5.1. **Canonical volume form.** Owing to Proposition 13, we complete the canonical frame along the extremal $t \mapsto \lambda(t)$, $t \in [0, T]$. Referring to the same kind of calculations as in the previous section, we find

$$F_1(t) = R_{22}(t)E_2(t) - \dot{F}_2(t) = R_{22}(t)\dot{\partial}_{h_2}(t) + \dddot{\partial}_{h_2}(t) \in T^*_{\lambda(t)}TM.$$

Then we declare $\theta_1(t)$, $\theta_2(t)$ to be the coframe dual to $Z_i(t) = \pi_*(F_i)(t) \in T_{x(t)}M$, $i = 1, 2$, and set $\xi_{x(t)} := \theta_1(t) \wedge \theta_2(t)$. Intuitively, $\xi_{x(\cdot)}$ is the volume form canonically associated with (4.2), along the geodesic $t \mapsto x(t)$. By explicit computations, we find that $Z_1 = X_2 - BX_1$, and $Z_2 = -X_1$. In local coordinates, it is immediate to establish that

$$\xi(x_1, x_2) = -\frac{1}{\partial_{x_1}\alpha_2(x_1, x_2)} dx_1 \wedge dx_2, \tag{5.1}$$

so that ξ coincides indeed with the volume form μ chosen in (1.10) and described in (3.4). Alternatively, we could have directly argued that by definition, $\theta_1(X_2)(t) = -\theta_2(X_1)(t) = 1$, implying $\xi_{x(t)}(X_1, [X_0, X_1]) = 1$ and $\xi_{x(t)} = \mu_{x(t)}$.

Since by assumption $S = \partial_{x_1}\alpha_2(0) \neq 0$, it is not restrictive to assume that μ does not change sign on the whole coordinate chart.

5.2. **Proof of Theorem 2.** We begin by observing that

$$\partial_t - X_0 - \frac{1}{2}\left(X_1^2 + \mathrm{div}_\mu(X_1)X_1\right), \tag{5.2}$$

can be interpreted in the form $\partial_t - X_0 - \Delta_H^\mu$, where Δ_H^μ denotes the horizontal Laplacian [4, 26] (also called sub-Laplacian) naturally associated with (4.2) and the canonical volume form μ characterized in (5.1).

On the one hand, by the general theory developed for the expansion in (2.8) and the results of Theorems 11 and 12, with $\sqrt{t}$ in place of ε, we know that the asymptotic development of $p(t, 0, 0)$ takes the form

(5.3)

$$p(t,0,0) = \frac{1}{2\pi t^2\sqrt{\det(D_1)}}\Bigg(1 + t\Bigg(-\frac{1}{2}\alpha_1^2(0) - \frac{1}{2}\Bigg(\partial_{x_1}\alpha_1(0) + \partial_{x_2}\alpha_2(0) - \alpha_1(0)\frac{\partial^2_{x_1}\alpha_2(0)}{\partial_{x_1}\alpha_2(0)}\Bigg) - \frac{12}{35}\left(\frac{\partial^2_{x_1}\alpha_2(0)}{\partial_{x_1}\alpha_2(0)}\right)^2 + \frac{3}{14}\frac{\partial^3_{x_1}\alpha_2(0)}{\partial_{x_1}\alpha_2(0)}\Bigg) + O(t^2)\Bigg)$$

within a coordinate chart U centered at x_0.

On the other hand, reading in coordinates (x_1, x_2) the structural equations (compare also with (3.3))

$$X_2 = [X_0, X_1], \quad [X_1, X_2] = c_{12}^1 X_1 + c_{12}^2 X_2, \quad [X_0, X_2] = c_{02}^1 X_1 + c_{02}^2 X_2,$$

we see that, in particular,

$$c_{12}^2(x_1,x_2) = \frac{\partial^2_{x_1}\alpha_2(x_1,x_2)}{\partial_{x_1}\alpha_2(x_1,x_2)} \quad \text{and} \quad X_1(c_{12}^2)(x_1,x_2) = \frac{\partial^3_{x_1}\alpha_2(x_1,x_2)}{\partial_{x_1}\alpha_2(x_1,x_2)} - \left(\frac{\partial^2_{x_1}\alpha_2(x_1,x_2)}{\partial_{x_1}\alpha_2(x_1,x_2)}\right)^2.$$

It is now easy to reinterpret (5.3) in an intrinsic way. More precisely, we have shown that, under the hypotheses of Theorem 2, the fundamental solution p to (5.2) admits the following small-time asymptotic expansion on the diagonal around x_0, namely:

$$p(t,x_0,x_0) = \frac{\sqrt{12}}{2\pi t^2}\left(1 + t\left(-\frac{1}{2}|X_0(x_0)|^2 - \frac{1}{2}\mathrm{div}_\mu(X_0)(x_0) + \frac{1}{14}K_1(x_0) - \frac{1}{70}K_2(x_0)^2\right) + O(t^2)\right),$$

where K_1 and K_2 are defined as in Definition 17. This ends the proof of Theorem 2.

Acknowledgements. We are grateful to Emmanuel Trélat and Yves Colin de Verdière for useful advices concerning some technical points presented in Section 2.4 and in Proposition 8. We also warmly thank Elisa Paoli who gave a first input to this project.

This research has been supported by the ANR project SRGI Sub-Riemannian Geometry and Interactions", contract number ANR-15-CE40-0018.

References

[1] A. Agrachev, D. Barilari, and E. Paoli. Volume geodesic distortion and Ricci curvature for Hamiltonian dynamics. *ArXiv e-prints*, Feb. 2016. `https://arxiv.org/pdf/1602.08745.pdf`.

[2] A. Agrachev, D. Barilari, and L. Rizzi. Sub-Riemannian curvature in contact geometry. *J. Geom. Anal.*, 27(1):366–408, 2017.

[3] A. A. Agrachev, D. Barilari, and L. Rizzi. Curvature: a variational approach. *Memoirs of the AMS, in press*, 2013. `https://arxiv.org/pdf/1306.5318.pdf`.

[4] A. A. Agrachev, U. Boscain, J.-P. Gauthier, and F. Rossi. The intrinsic hypoelliptic Laplacian and its heat kernel on unimodular Lie groups. *J. Funct. Anal.*, 256(8):2621–2655, 2009.

[5] A. A. Agrachev and Y. L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.

[6] R. Azencott. Grandes déviations et applications. In *Eighth Saint Flour Probability Summer School—1978 (Saint Flour, 1978)*, volume 774 of *Lecture Notes in Math.*, pages 1–176. Springer, Berlin, 1980.

[7] D. Barilari. Trace heat kernel asymptotics in 3d contact sub-riemannian geometry. *Journal of Mathematical Sciences*, 195(3):391–411, Dec 2013.

[8] D. Barilari, U. Boscain, G. Charlot, and R. W. Neel. On the Heat Diffusion for Generic Riemannian and Sub-Riemannian Structures. *International Mathematics Research Notices*, 2017(15):4639–4672, 2017.

[9] D. Barilari, U. Boscain, and R. W. Neel. Small-time heat kernel asymptotics at the sub-Riemannian cut locus. *J. Differential Geom.*, 92(3):373–416, 2012.

[10] D. Barilari, U. Boscain, and M. Sigalotti, editors. *Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. 1.* EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich, 2016.

[11] D. Barilari and E. Paoli. Curvature terms in small time heat kernel expansion for a model class of hypoelliptic Hörmander operators. *Nonlinear Analysis, in press.*

[12] D. Barilari and L. Rizzi. Comparison theorems for conjugate points in sub-Riemannian geometry. *ESAIM Control Optim. Calc. Var.*, 22(2):439–472, 2016.

[13] D. Barilari and L. Rizzi. On Jacobi fields and a canonical connection in sub-Riemannian geometry. *Archivum Mathematicum*, (53):77–92, 2017.

[14] F. Baudoin. Sub-Laplacians and hypoelliptic operators on totally geodesic Riemannian foliations. In *Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. 1*, EMS Ser. Lect. Math., pages 259–321. Eur. Math. Soc., Zürich, 2016.

[15] F. Baudoin and J. Wang. The subelliptic heat kernel on the CR sphere. *Math. Z.*, 275(1-2):135–150, 2013.

[16] F. Baudoin and J. Wang. The subelliptic heat kernels of the quaternionic Hopf fibration. *Potential Anal.*, 41(3):959–982, 2014.

[17] R. Beals, P. C. Greiner, and N. K. Stanton. The heat equation on a CR manifold. *J. Differential Geom.*, 20(2):343–387, 1984.

[18] G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Ann. Sci. École Norm. Sup. (4)*, 21(3):307–331, 1988.

[19] G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale. *Ann. Inst. Fourier (Grenoble)*, 39(1):73–99, 1989.

[20] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale. II. *Probab. Theory Related Fields*, 90(3):377–402, 1991.

[21] M. Berger. *A panoramic view of Riemannian geometry.* Springer-Verlag, Berlin, 2003.
[22] M. Berger, P. Gauduchon, and E. Mazet. *Le spectre d'une variété riemannienne.* Lecture Notes in Mathematics, Vol. 194. Springer-Verlag, Berlin, 1971.
[23] R. M. Bianchini and G. Stefani. Graded approximations and controllability along a trajectory. *SIAM J. Control Optim.*, 28(4):903–924, 1990.
[24] J.-M. Bismut. *Large deviations and the Malliavin calculus*, volume 45 of *Progress in Mathematics.* Birkhäuser Boston Inc., Boston, MA, 1984.
[25] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni. *Stratified Lie groups and potential theory for their sub-Laplacians.* Springer Monographs in Mathematics. Springer, Berlin, 2007.
[26] U. Boscain and C. Laurent. The Laplace-Beltrami operator in almost-Riemannian geometry. *Ann. Inst. Fourier (Grenoble)*, 63(5):1739–1770, 2013.
[27] M. Bramanti. *An invitation to hypoelliptic operators and Hörmander's vector fields.* SpringerBriefs in Mathematics. Springer, Cham, 2014.
[28] Y. C. de Verdière, L. Hillairet, and E. Trélat. Spectral asymptotics for sub-Riemannian Laplacians, II. *Ongoing work.*
[29] N. Garofalo. Hypoelliptic operators and some aspects of analysis and geometry of sub-Riemannian spaces. In *Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. 1*, EMS Ser. Lect. Math., pages 123–257. Eur. Math. Soc., Zürich, 2016.
[30] A. Grigor'yan. *Heat kernel and analysis on manifolds*, volume 47 of *AMS/IP Studies in Advanced Mathematics.* American Mathematical Society, Providence, RI; International Press, Boston, MA, 2009.
[31] M. Hairer. On Malliavin's proof of Hörmander's theorem. *Bull. Sci. Math.*, 135(6-7):650–666, 2011.
[32] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.
[33] E. P. Hsu. On the principle of not feeling the boundary for diffusion processes. *J. London Math. Soc. (2)*, 51(2):373–382, 1995.
[34] D. S. Jerison and A. Sánchez-Calle. Estimates for the heat kernel for a sum of squares of vector fields. *Indiana Univ. Math. J.*, 35(4):835–854, 1986.
[35] F. John. *Partial differential equations*, volume 1 of *Applied Mathematical Sciences.* Springer-Verlag, New York, fourth edition, 1991.
[36] V. Jurdjevic. *Geometric control theory*, volume 52 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1997.
[37] M. Kac. On some connections between probability theory and differential and integral equations. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, pages 189–215. University of California Press, Berkeley and Los Angeles, 1951.
[38] T. Kato. *Perturbation theory for linear operators.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.
[39] A. Kolmogoroff. Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.*, 104(1):415–458, 1931.
[40] A. Kolmogoroff. Zur Grössenordnung des Restgliedes Fourierscher Reihen differenzierbarer Funktionen. *Ann. of Math. (2)*, 36(2):521–526, 1935.
[41] E. Lanconelli and S. Polidoro. On a class of hypoelliptic evolution operators. *Rend. Sem. Mat. Univ. Politec. Torino*, 52(1):29–63, 1994. Partial differential equations, II (Turin, 1993).
[42] R. Léandre. Majoration en temps petit de la densité d'une diffusion dégénérée. *Probab. Theory Related Fields*, 74(2):289–294, 1987.

[43] R. Léandre. Minoration en temps petit de la densité d'une diffusion dégénérée. *J. Funct. Anal.*, 74(2):399–414, 1987.

[44] P. Malliavin. Stochastic calculus of variation and hypoelliptic operators. In *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*, pages 195–263. Wiley, New York-Chichester-Brisbane, 1978.

[45] S. A. Molčanov. Diffusion processes, and Riemannian geometry. *Uspehi Mat. Nauk*, 30(1(181)):3–59, 1975.

[46] R. Neel and D. Stroock. Analysis of the cut locus via the heat kernel. In *Surveys in differential geometry. Vol. IX*, Surv. Differ. Geom., IX, pages 337–349. Int. Press, Somerville, MA, 2004.

[47] E. Paoli. Small Time Asymptotics on the Diagonal for Hörmander's Type Hypoelliptic Operators. *J. Dyn. Control Syst.*, 23(1):113–143, 2017.

[48] S. Polidoro. On a class of ultraparabolic operators of Kolmogorov-Fokker-Planck type. *Matematiche (Catania)*, 49(1):53–105 (1995), 1994.

[49] S. Rosenberg. *The Laplacian on a Riemannian manifold*, volume 31 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997. An introduction to analysis on manifolds.

[50] N. K. Stanton and D. S. Tartakoff. The heat equation for the $\bar{\partial}_b$-Laplacian. *Comm. Partial Differential Equations*, 9(7):597–686, 1984.

[51] D. W. Stroock. The Malliavin calculus and its application to second order parabolic differential equations. I. *Math. Systems Theory*, 14(1):25–65, 1981.

[52] H. J. Sussmann and V. Jurdjevic. Controllability of nonlinear systems. *J. Differential Equations*, 12:95–116, 1972.

[53] H. F. Trotter. Approximation of semi-groups of operators. *Pacific J. Math.*, 8:887–919, 1958.

[54] S. R. S. Varadhan. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Comm. Pure Appl. Math.*, 20:431–455, 1967.

[55] J. Wang. The subelliptic heat kernel on the anti-de Sitter space. *Potential Anal.*, 45(4):635–653, 2016.

[56] I. Zelenko and C. Li. Differential geometry of curves in Lagrange Grassmannians with given Young diagram. *Differential Geom. Appl.*, 27(6):723–742, 2009.

Université Paris Diderot - Paris 7, Institut de Mathematique de Jussieu, UMR CNRS 7586 - UFR de Mathématiques.

E-mail address: `davide.barilari@imj-prg.fr`

Sorbonne Universités, UPMC Univ Paris 06, CNRS UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France and INRIA - Team CaGe

E-mail address: `boarotto@ljll.math.upmc.fr`