INTRINSIC RANDOM WALKS IN RIEMANNIAN AND SUB-RIEMANNIAN GEOMETRY VIA VOLUME SAMPLING

ANDREI AGRACHEV[1], UGO BOSCAIN[2], ROBERT NEEL[3], AND LUCA RIZZI[2]

Abstract. We relate some basic constructions of stochastic analysis to differential geometry, via random walk approximations. We consider walks on both Riemannian and sub-Riemannian manifolds in which the steps consist of travel along either geodesics or integral curves associated to orthonormal frames, and we give particular attention to walks where the choice of step is influenced by a volume on the manifold. A primary motivation is to explore how one can pass, in the parabolic scaling limit, from geodesics, orthonormal frames, and/or volumes to diffusions, and hence their infinitesimal generators, on sub-Riemannian manifolds, which is interesting in light of the fact that there is no completely canonical notion of sub-Laplacian on a general sub-Riemannian manifold. However, even in the Riemannian case, this random walk approach illuminates the geometric significance of Ito and Stratonovich stochastic differential equations as well as the role played by the volume.

Contents

1. Introduction	2
2. Convergence of random walks	4
3. Geodesic random walks in the Riemannian setting	6
3.1. Ito SDEs via geodesic random walks	6
3.2. Volume sampling through the exponential map	7
4. Geodesic random walks in the sub-Riemannian setting	10
4.1. Geodesics and exponential map	10
4.2. Sub-Laplacians	11
4.3. The sub-Riemannian geodesic random walk with volume sampling	11
4.4. Contact Carnot groups	12
4.5. Alternative constructions for the sub-Riemannian random walk	15
4.6. The 3D Heisenberg group	17
5. Flow random walks	17
5.1. Stratonovich SDEs via flow random walks	18
5.2. Volume sampling through the flow	18
5.3. Flow random walks with volume sampling in the Riemannian setting	19
5.4. Flow random walks with volume sampling in the sub-Riemannian setting	20
6. Proof of the results	21
6.1. Proof of Theorem 8	21
6.2. Proof of Theorem 15	22
6.3. Proof of Theorem 20	25
6.4. Proof of Theorem 25	27
Appendix A. Volume sampling as Girsanov-type change-of-measure	27
Acknowledgments	29
References	29

[1]SISSA, Trieste, Italy
[2]CNRS, CMAP École Polytechnique, and Équipe INRIA GECO Saclay Île-de-France, Paris
[3]Lehigh University

Date: January 6, 2016.

1. Introduction

Consider a Riemannian or a sub-Riemannian manifold M and assume that $\{X_1, \ldots, X_k\}$ is a global orthonormal frame. It is well known that, under mild hypotheses, the solution q_t to the stochastic differential equation in Stratonovich sense

$$dq_t = \sum_{i=1}^{k} X_i(q_t) \circ \left(\sqrt{2}\, dw_t^i\right) \tag{1}$$

produces a solution to the heat-like equation

$$\partial_t \varphi = \sum_{i=1}^{k} X_i^2 \varphi \tag{2}$$

by taking $\varphi_t(q) = \mathbb{E}\left(\varphi_0(q_t) | q_0 = q\right)$, where φ_0 gives the initial condition. (Here the driving processes w_t^i are independent real Brownian motions, and $\sqrt{2}$ factor is there so that the resulting sum-of-squares operator doesn't need a $1/2$, consistent with the convention favored by analysts.) One can interpret (2) as the equation satisfied by a random walk with parabolic scaling following the integral curves of the vector fields $X_1, \ldots X_m$, when the step of the walk tends to zero. This construction is very general (works in Riemannian and in the sub-Riemannian case) and does not use any notion of volume on the manifold.[1]

However the operator $\sum_{i=1}^{k} X_i^2$ is not completely satisfactory to describe a diffusion process for the following reasons:

- the construction works only if a global orthonormal frame $X_1, \ldots, X_k$ exists;
- it is not intrinsic in the sense that it depends on the choice of the orthonormal frame;
- it is not essentially self-adjoint w.r.t. a natural volume and one cannot guarantee a priori a "good" evolution in L^2 (existence and uniqueness of a contraction semigroup etc...).

In the Riemannian context a heat operator that is globally well defined, frame independent and essentially self-adjoint w.r.t. the Riemannian volume (at least under the hypotheses of completeness) is the Laplace-Beltrami operator $\Delta = \mathrm{div} \circ \mathrm{grad}$. A heat-like equation

$$\partial_t \varphi = \Delta \varphi \tag{3}$$

has an associated diffusion given by the solution of the stochastic differential equation

$$dq_t = \sum_{i=1}^{k} X_i(q_t) \left(\sqrt{2}\, dw_t^i\right) \qquad \text{(in this case } k = n \text{ is equal to the dimension of } M) \tag{4}$$

in the Ito sense (for instance using the Bismut contruction on the bundle of orthonormal frame or the Emery approach [9, 12]). Also, this equation can be interpreted as the equation satisfied by the limit of a random walk that, instead of integral curves of the vector fields of the orthonormal frame, follows geodesics. The geodesics starting from a given point are weighted with a uniform probability given by the Riemannian metric on the tangent space at the point.

The purpose of this paper is to extend this more invariant construction of random walks to the sub-Riemannian context, to obtain a definition of an intrinsic Laplacian in sub-Riemannian geometry and to compare it with the divergence of the horizontal gradient.

In Section 2, we introduce a general scheme of convergence of random walks of sufficiently general class to include all our constructions.

[1] In the Riemannian case avoiding the use of a volume is not crucial since an intrinsic volume (the Riemannian one) can always be defined. But in the sub-Riemannian case, how to define an intrinsic volume is a subtle question, as discussed below.

The task of determining the appropriate random walk is not obvious for several reasons. First it is not obvious how to give a uniform measure on the space of geodesics starting from a given point. Indeed in sub-Riemannian geometry geodesics starting from a given point are always parameterized by a non-compact subset of the cotangent space at the point, on which there is no canonical "uniform" probability measure. Second, in sub-Riemannian geometry for every ε there exist geodesics of length ε that have already lost their optimality, and one has to choose between a construction involving all geodesics (including the non-optimal ones) or only those that are optimal. Third, one should decide what to do with abnormal extremals. Finally, there is the problem of defining an intrinsic volume in sub-Riemannian geometry, to compute the divergence.

It is not the first time that this problem has been attacked. In [15, 16, 10, 17, 18], the authors compare the divergence of the gradient with the Laplacian corresponding to a random walk induced by a splitting of the cotagent bundle (see [10, Section 1.4] for a more detailed summary of this literature).

In this paper we use another approach, trying to induce a measure on the space of geodesics, from the ambient space, using the idea of "sampling" the volume.

This idea works very well in the Riemannian case, permitting a random walk interpretation of the divergence of the gradient also when the divergence is computed w.r.t. an arbitrary volume. From these results one also recognizes a particular role played by the Riemannian volume (see Section 3 and Corollary 9).

In the sub-Riemannian case the picture appears richer and more complicated. Even for contact Carnot groups (see Section 4) the volume sampling procedure is non-trivial, as one requires an explicit knowledge of the exponential map, and we get some surprising results. In the 3D Heisenberg group one gets that the limit process is generated by the divergence of the horizontal gradient if and only if at least one of the two conditions are satisfied: (i) one is using the Popp volume; (ii) a suitable parameter used to realize the "volume sampling" is equal to 1/2, evoking reminiscences of the Stratonovich integral.

In higher dimension one gets that in general the generator of the limit process is not the expected divergence of the horizontal gradient (even the second-order terms are not the expected ones); however, the generator will be the divergence of the horizontal gradient with respect to a different metric on the same distribution, as shown in Section 4.4.1.

Motivated by these unexpected results and difficulties in manipulating the exponential map in more general sub-Riemannian cases, in Section 5, we try an alternative construction in the general contact case (that we call the flow random walk with volume sampling), inspired by the classical Stratonovich integration and including a volume sampling procedure. This construction, a priori not-intrinsic (as it depends on the choice of some vector fields), gives rise in the limit to an intrinsic operator showing the particular role played by the Popp volume. This construction gives some interesting hints also in the Riemannian case; unfortunately this construction cannot be easily generalized to situations of higher step or corank.

In the process of developing the material just described, we naturally obtain an intuitively appealing description of the solution to a Stratonovich SDE on a manifold as a randomized flow along the vector fields $X_1, \ldots, X_k$ (as already outlined above) while the solution to an Ito SDE is a randomized geodesic tangent to the vector fields $X_1, \ldots, X_k$. This difference corresponds to the infinitesimal generator being based on second Lie derivatives along the X_i versus second covariant derivatives. Of course, such an approximation procedure by random walks yields nothing about the diffusions that is not contained in standard stochastic calculus, but the the explicit connection to important geometric objects seems compelling and also something that has not been succinctly described before, to the best of our knowledge. Further, it is then natural to round out this perspective on the basic objects of stochastic calculus on manifolds by highlighting the way in which the volume sampling procedure can be viewed as a random walk approximation of the Girsanov change of measure, at least in the Riemannian case (see Appendix A).

For the benefit of exposition, all the proofs are collected in Section 6.

2. Convergence of random walks

We recall some preliminaries in sub-Riemannian geometry (see [1], but also [24, 26, 20]).

Definition 1. A *(sub-)Riemannian manifold* is a triple $(M, \blacktriangle, \mathbf{g})$ where M is smooth, connected manifold, $\blacktriangle \subset TM$ is a vector distribution of constant rank $k \leq n$ and $\mathbf{g}$ is a smooth scalar product on $\blacktriangle$. We assume that $\blacktriangle$ satisfies the *Hörmander's condition*

$$\operatorname{span}\{[X_{i_1}, [X_{i_2}, [\dots, [X_{i_{m-1}}, X_{i_m}]]] \mid m \geq 0, \quad X_{i_\ell} \in \Gamma(\blacktriangle)\}_q = T_qM, \qquad \forall q \in M. \tag{5}$$

By the Chow-Rashevskii theorem, any two points in M can be joined by a Lipschitz continuous curve whose velocity is a.e. in $\blacktriangle$. We call such curves *horizontal.* Horizontal curves $\gamma : I \to M$ have a well-defined length, given by

$$\ell(\gamma) = \int_I \|\gamma(t)\| dt, \tag{6}$$

where $\|\cdot\|$ is the norm induced by $\mathbf{g}$. The *sub-Riemannian distance* between $p, q \in M$ is

$$d(p, q) = \inf\{\ell(\gamma) \mid \gamma \text{ horizontal curve connecting } q \text{ with } p\}. \tag{7}$$

This distance turns $(M, \blacktriangle, \mathbf{g})$ into a metric space that has the same topology of M. A sub-Riemannian manifold is *complete* if (M, d) is complete as a metric space.

Our definition of (sub-)Riemannian structures includes Riemannian ones, when $k = n$. We use the term "sub-Riemannian" (without parenthesis) to denote structures that are not Riemannian, i.e. $k < n$.

Definition 2. If M is a (sub-)Riemannian manifold, (following the basic construction of Stroock and Varadhan [32]) let $\Omega(M)$ be the space of continuous paths from $[0, \infty)$ to M. If $\gamma \in \Omega$ is such a curve on M (with $\gamma(t)$ giving the position of the path at time t), then the metric on M induces a metric on Ω by

$$d_{\Omega_M}\left(\gamma^1, \gamma^2\right) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{\sup_{0\leq t\leq i} d_M\left(\gamma^1(t), \gamma^2(t)\right)}{1 + \sup_{0\leq t\leq i} d_M\left(\gamma^1(t), \gamma^2(t)\right)}$$

making Ω_M into a Polish space. We give Ω_M its Borel σ-algebra. We are primarily interested in the weak convergence of probability measures on Ω_M.

A choice of probability measure P on Ω_M determines a continuous, random process on M, and we will generally denote the random position of the path at time t by q_t. Moreover, we will use the measure P and the process q_t interchangeably.

One can consider families of random walks in varying degrees of generality. For our purposes, we are interested in what one might call bounded-step-size, parabolically-scaled families of random walks, which for simplicity in what follows, we will just call a family of random walks. We will index our families by a "spatial parameter" $\varepsilon > 0$ (this will be clearer below), and we let $\delta = \varepsilon^2/(2k)$ be the corresponding time step (here k is again the dimension of the distribution on M).

Definition 3. A family of random walks on a (sub-)Riemannian manifold M, indexed by $\varepsilon > 0$ and starting from $q \in M$, is a family of probability measures P_q^ε on $\Omega(M)$ with $P_q^\varepsilon(q_0^\varepsilon = q) = 1$ and having the following property. For every ε, and every $\tilde{q} \in M$, there exists a probability measure $\Pi_{\tilde{q}}^\varepsilon$ on continuous paths $\gamma : [0, \delta] \to M$ with $\gamma(0) = \tilde{q}$ such that for every $m = 0, 1, 2, \dots$, the distribution of $q_{[m\delta,(m+1)\delta]}^\varepsilon$ under P_q^ε is given by $\Pi_{q_{m\delta}^\varepsilon}^\varepsilon$, independently of the position of the path q_t^ε prior to time $m\delta$. Further, there exists some constant κ, independent of $\tilde{q}$ and ε, such that the length of $\gamma_{[0,\delta]}$ is almost surely less than or equal to $\kappa\varepsilon$ under $\Pi_{\tilde{q}}^\varepsilon$. (So the position of the path at times $m\delta$ for $m = 0, 1, 2, \dots$ is a

Markov chain, starting from q, with transition probabilities $P^\varepsilon_q\left(q^\varepsilon_{(m+1)\delta}\in A \mid q^\varepsilon_{m\delta}=\tilde q\right)=\Pi^\varepsilon_{\tilde q}\left(\gamma_\delta\in A\right)$ for any Borel $A\subset M$. That $\Pi^\varepsilon_{\tilde q}$ is a measure on the entire path $\gamma_{[0,\delta]}$ means that is also encodes the continuous interpolation between $q^\varepsilon_{m\delta}$ and $q^\varepsilon_{(m+1)\delta}$.)

Remark 1. In what follows $\Pi^\varepsilon_{\tilde q}$ will, in most cases, be supported on paths of length exactly ε (allowing us to take $\kappa=1$). For example, on a Riemannian manifold, one might choose a direction at $q^\varepsilon_{m\delta}$ at random and then follow a geodesic in this direction for length ε (and in time δ). Alternatively, on a Riemannian manifold with a global orthonormal frame, one might choose a random linear combination of the vectors in the frame, still having length 1, and then flow along this vector field for length ε. In both of these cases, $\Pi^\varepsilon_{\tilde q}$ is itself built on a probability measure on the unit sphere in $T_{\tilde q}M$ according to a kind of scaling by ε. These walks, and variations and sub-Riemannian versions thereof, form the bulk of what we consider, and should be sufficient to illuminate the definition.

Further, while the introduction of the "next step" measure $\Pi^\varepsilon_{\tilde q}$ is suitable for the general definition and accompanying convergence result, it is overkill for the geometrically natural steps that we specifically consider. Instead, we will generally describe the steps of our random walks in simpler geometric terms (as in the case of a choosing a random geodesic segment of length ε just mentioned), and leave the specification of $\Pi^\varepsilon_{\tilde q}$ implicit, though in a straightforward way.

We note that, for some constructions like that of solutions to a Stratonovich SDE, there need not be a metric on M, but instead a smooth structure is sufficient. Unfortunately, the machinery of convergence of random walks just discussed is formulated in terms of metrics, and thus we will generally proceed by choosing some (Riemannian or sub-Riemannian) metric on M when desired. However, note that if M is compact, any two Riemannian metrics induce Lipschitz-equivalent distances on M, and thus the induced distances on Ω_M are comparable. This means that the resulting topologies on Ω_M are the same, just metrized differently, and thus statements about the convergence of probability measures on Ω_M (which is how we formalize the convergence of random walks) don't depend on what metric on M is chosen. This suggests that a more general framework could be developed, avoiding the need to introduce a metric on M when the smooth structure should suffice, but such an approach will not be pursued here.

Definition 4. Let $\varepsilon>0$. To the family of random walks q^ε_t (in our terminology, and with the above notation), we associated the family of smooth operators

$$(L^\varepsilon\phi)(q):=\frac{1}{\delta}\mathbb{E}[\phi(q^\varepsilon_\delta)-\phi(q)\mid q^\varepsilon_0=q],\qquad \forall q\in M \tag{8}$$

Definition 5. Let L be a differential operator on M. We say that a family L^ε of differential operators converge to L if for any $\phi\in C^\infty(M)$ we have $L^\varepsilon\phi\to L\phi$ uniformly on compact sets. In this case, we write $L^\varepsilon\to L$.

Let L be a smooth second-order differential operator with no zeroth-order term. If the principal symbol of L is also non-negative definite, then there is a unique diffusion associated to L starting from any point $q\in M$, at least up until a possible explosion time. However, since our analysis in fundamentally local, we will assume that the diffusion does not explode. In that case, this diffusion is given by the measure P_q on Ω that solves the martingale problem for L, so that

$$\phi(q_t)-\int_0^t L\phi(q_s)\,ds$$

is a martingale under P_q for any smooth, compactly supported ϕ, and $P_q\left(q_0=q\right)=1$.

Theorem 6. *Let M be a (sub-) Riemannian manifold, let P^ε_q be the probability measures on $\Omega(M)$ corresponding to a sequence of random walks q^ε_t (in our terminology), with*

$q_0^\varepsilon = q$, and let L^ε be the associated family of operators. Suppose that $L^\varepsilon \to L^0$, where L^0 is a smooth second-order operator with non-negative definite principal symbol and without zeroth-order term. Further, suppose that the diffusion generated by L, which we call q_t^0, does not explode, and let P_q^0 be the corresponding probability measure on $\Omega(M)$ starting from q. Then $P_q^\varepsilon \to P_q^0$ as $\varepsilon \to 0$, in the sense of weak convergence of probability measures.

Proof. The theorem is a special case of Theorem 70 and Remark 26 of [10]. First note that a random walk q_t^ε as described here corresponds to a random walk X_t^h in the notation of that paper under the change of variables $h = \varepsilon^2/2k$, with each step being given either by a continuous curve (which may or may not be a geodesic), as addressed in Remark 26. Then all of the assumptions of that theorem are included in the hypotheses above and our definition of the convergence $L^\varepsilon \to L^0$, except the assumption that the random walks q_t^ε satisfy Eq. (19) of [10]. However, every random walk in our class has the property that, during any step, the path never goes more than distance $\kappa\varepsilon$ from the starting point of the step for some fixed $\kappa > 0$, by construction. This immediately shows that every random walk in our class satisfies this Eq. (19). But then the conclusion of Theorem 70 of [10], namely $P_q^\varepsilon \to P_q^0$ as $\varepsilon \to 0$ follows. $\square$

3. Geodesic random walks in the Riemannian setting

3.1. **Ito SDEs via geodesic random walks.** Let $(M, \mathbf{g})$ be a Riemannian manifold. We consider a set of smooth vector fields, and since we are interested in local phenomena, we are free to assume that the V_i have bounded lengths and that $(M, \mathbf{g})$ is geodesically complete. We now consider the Ito SDE

$$dq_t = \sum_{i=1}^{k} V_i\,(q_t)\,dw_t^i, \quad q_0 = q \tag{9}$$

for some $q \in M$, where $w_t^1, \dots, w_t^k$ are independent, one-dimensional Brownian motions[2]. To construct a corresponding sequence of random walks, we choose a random vector $V = \beta_1 V_1 + \beta_2 V_2 + \cdots + \beta_k V_k$ by choosing $(\beta_1, \dots, \beta_k)$ uniformly from the unit sphere. Then, we follow the geodesic $\gamma(s)$ determined by $\gamma(0) = q$ and $\gamma'(0) = \frac{2k}{\varepsilon} V$ for time $\delta = \varepsilon^2/(2k)$. Equivalently, we travel a distance of $\varepsilon|V|$ in the direction of V (along a geodesic). This determines the first step, q_t^ε with $t \in [0, \delta]$, of a random walk (and thus, implicitly, the measure Π_q^ε). Determining each additional step in the same way produces a family of piecewise geodesic random walks q_t^ε, $t \in [0, \infty)$, which we call the *geodesic random walk* at scale ε associated with the SDE (9) (note that, in terms of definition 3, we can take $\kappa = \sup_{q,(\beta_1,\dots,\beta_k)} V$).

We now study the convergence of this family of walks as $\varepsilon \to 0$. Let $x_1, \dots, x_n$ be Riemannian normal coordinates around $q_0^\varepsilon = q$, and write the random vector V as

$$V(x) = \sum_{m=1}^{k} \beta_m V_m(x) = \sum_{m=1}^{k} \beta_i \sum_{i=1}^{n} V_m^i \partial_i + O(r) = \sum_{i=1}^{n} A^i \partial_{x_i} + O(r), \tag{10}$$

where $r = \sqrt{x_1^2 + \ldots + x_n^2}$. In normal coordinates, Riemannian geodesics correspond to Euclidean geodesics up to second order, and thus $\gamma_V(t)$ has i-th coordinate $A^i t + O(t^3)$. In particular, for any smooth function ϕ we have

$$\phi(\gamma_V(\varepsilon)) - \phi(q) = \sum_{i=1}^{n} A^i (\partial_i \phi)(q)\varepsilon + \frac{1}{2} \sum_{i,j=1}^{n} A^i A^j (\partial_i \partial_j \phi)(q)\varepsilon^2 + O(\varepsilon^3). \tag{11}$$

[2]One approach to interpreting and solving (9), as well as verifying that q_t will be a martingale, is via lifting it to the bundle of orthonormal frames; see the first two chapters of [19] for background on stochastic differential geometry, connection-martingales, and the bundle of orthonormal frames. Alternatively, [12, Chapter 7] gives a treatment of Ito integration on manifolds.

Taking the average w.r.t. the uniform probability measure on the sphere $\beta_1^2 + \ldots + \beta_k^2 = 1$ we find that

$$\begin{aligned}\frac{1}{\delta}\mathbb{E}\left[\phi\left(q_\delta^\varepsilon\right) - \phi(q)\,|q_0^\varepsilon = q\right] &\to \sum_{m=1}^{k}\sum_{i,j=1}^{n} V_m^i V_m^j (\partial_i\partial_j\phi)(q) \\ &= \sum_{m=1}^{k} \nabla^2_{V_m,V_m}(q) \qquad \text{as } \varepsilon \to 0,\end{aligned} \tag{12}$$

where ∇^2 denotes the Hessian, with respect to the Levi-Civita connection, and where we recall that $\sum_{j=1}^n V_m^j \partial_j$ is $V_m(q)$ and the x_i are a system of normal coordinates at q. Note that right-hand side of (12) determines a second-order operator at q which is independent of the choice of normal coordinates (and thus depends only on the V_i). Moreover, this same construction works at any point, and thus we have a second-order operator L on all of M. Because the V_i are smooth, so is L.

We see that the martingale problem associated to L has a unique solution (at least until explosion, but since we are interested in local questions, we can assume that there is no explosion). Further, this solution is the law of the process q^0 that solves (9). If we again let P^ε and P^0 be the probability measures on $\Omega(M)$ corresponding to q_t^ε and q_t^0, respectively, Theorem 6 implies that $P^\varepsilon \to P^0$ (weakly) as $\varepsilon \to 0$.

Of course, we see that our geodesic random walks, as well as the diffusion q^0 and thus the interpretation of the SDE (9), depend on the Riemannian structure. This is closely related to the fact that neither Ito SDEs, normal coordinates, covariant derivatives, nor geodesics are preserved under diffeomorphisms, in general, and to the non-standard calculus of Ito's rule for Ito integrals, in contrast to Stratonovich integrals.

The most important special case of a geodesic random walk is when $k = n$ and the vector fields $V_1, \ldots, V_n$ are an orthonormal frame. In that case, q_t^ε is an isotropic random walk, as described in [10] (see also [25] for a related family of processes) and

$$L^\varepsilon \to \Delta, \tag{13}$$

where $\Delta = \operatorname{div} \circ \operatorname{grad}$ is the Laplace-Beltrami operator (here the divergence is computed with respect to the Riemannian volume). In particular q_t^0 is the Brownian motion on M.

If we further specialize to Euclidean space, we see that the convergence of the random walk to Eucldiean Brownian motion is just a special case of Donsker's invariance principle. The development of Brownian motion on a Riemannian manifold via approximations is also not new; one approach can be found in [31].

3.2. **Volume sampling through the exponential map.** Let $(M, \mathbf{g})$ be a n-dimensional Riemannian manifold equipped with general volume form ω, that might be different from the Riemannian one $\mathcal{R}$. This freedom is motivated by the forthcoming applications to sub-Riemannian geometry, where there are several choices of intrinsic volumes and in principle there is not a preferred one [2, 7]. Besides, also in the Riemannian case, one might desire to study operators which are symmetric w.r.t. a generic *weighted measure* $\omega = e^h\mathcal{R}$ where $h \in C^\infty(M)$.

We recall that the gradient $\operatorname{grad}(\phi)$ of a smooth function depends only on the Riemannian structure, while the divergence $\operatorname{div}_\omega(X)$ of a smooth vector field depends on the choice of the volume. In this setting we introduce an intrinsic diffusion operator, symmetric in $L^2(M,\omega)$, with domain $C_c^\infty(M)$ as the divergence of the gradient:

$$\Delta_\omega := \operatorname{div}_\omega \circ \operatorname{grad} = \sum_{i=1}^n X_i^2 + \operatorname{div}_\omega(X_i)X_i, \tag{14}$$

where in the last equality, that holds locally, $X_1, \ldots, X_n$ is a orthonormal frame.

If one would like to define a random walk converging to (the law of) Δ_ω, one should make a construction in such a way that the choice of the volume enters in the definition of the walk. One way to do this is to "sample the volume along the walk". For all $s \geq 0$, consider the Riemannian exponential map

$$\exp_q(s;\cdot) : S_qM \to M, \qquad q \in M, \tag{15}$$

where $S_qM \subset T_qM$ is the Riemannian tangent sphere. In particular, for $v \in S_qM$, then $\gamma_v(s) = \exp_q(s;v)$ is the unit-speed geodesic starting at q with initial vector v. Then $|\iota_{\dot\gamma_v(s)}\omega|$ is a density[3] on the Riemannian sphere of radius s. By pulling this back through the exponential map, we obtain a probability measure that "gives more weight to geodesics arriving where there is more volume".

Definition 7. For any $q \in M$, and $\varepsilon > 0$, we define the family of densities μ_q^ε on S_qM

$$\mu_q^\varepsilon(v) := \frac{1}{N(q,\varepsilon)} \left|(\exp_q(\varepsilon;\cdot)^* \iota_{\dot\gamma_v(\varepsilon)}\omega)(v)\right|, \qquad \forall v \in S_qM, \tag{16}$$

where $N(q,\varepsilon)$ is such that $\int_{S_qM} \mu_q^\varepsilon = 1$. For $\varepsilon = 0$, we set μ_q^0 to be the standard normalized Riemannian density form on S_qM.

Remark 2. Observe that for small $\varepsilon > 0$ the Jacobian determinant of $\exp_q(\varepsilon;\cdot)$ does not change sign, hence the absolute value in the definition above is not strictly necessary to obtain a well defined probability measure on S_qM. By assuming that the sectional curvature $\mathrm{Sec} \leq K$ is bounded from above, we can get rid globally of the need for the absolute value. Since our concerns are local, this would not be much of a loss of generality.

Thus, we define a random walk b_t^ε as follows:

$$b^\varepsilon_{(i+1)\delta} := \exp_{b^\varepsilon_{i\delta}}(\varepsilon; v), \qquad v \in S_qM \text{ with probability } \mu_q^\varepsilon. \tag{17}$$

(see Definition 3 and Remark 1). Let P_ω^ε (we drop a q from the notation as the starting point is fixed) be the probability measure on the space of continuous paths on M starting at q, associated with b_t^ε and consider the associated family of operators

$$(L_\omega^\varepsilon \phi)(q) := \frac{1}{\delta}\mathbb{E}[\phi(b_\delta^\varepsilon) - \phi(q) \mid b_0^\varepsilon = q] \tag{18}$$

$$:= \frac{1}{\delta}\int_{S_qM} [\phi(\exp_q(\varepsilon;v)) - \phi(q)]\mu_q^\varepsilon(v), \qquad \forall q \in M, \tag{19}$$

(see Definition 5), for any $\phi \in C^\infty(M)$. A special case of Theorem 8 gives

$$\lim_{\varepsilon \to 0} L_\omega^\varepsilon = \underbrace{\Delta_{\mathcal{R}} + \mathrm{grad}(h)}_{\Delta_\omega} + \mathrm{grad}(h), \tag{20}$$

where $\mathrm{grad}(h)$ is understood as a derivation. By Theorem 6, P_ω^ε converges to a well-defined diffusion generated by r.h.s. of (20). This result is not satisfactory, as one would expect $L_\omega^\varepsilon \to \Delta_\omega$. Indeed, in (20), we observe that the correction $2\,\mathrm{grad}(h)$ provided by the volume sampling construction is twice the expected one.

To address this problem we introduce a parameter $c \in [0,1]$ and consider, instead, the family $\mu_q^{c\varepsilon}$. This corresponds to sampling the volume not at the final point of the geodesic segment, but at an intermediate point. We define a discrete process as follows:

$$b^\varepsilon_{(i+1)\delta,c} := \exp_{q^\varepsilon_{i\delta,c}}(\varepsilon, v), \qquad v \in S_qM \text{ with probability } \mu_q^{c\varepsilon}, \tag{21}$$

that we call *Ito random walk with volume sampling* (with volume ω and sampling ratio c).

Remark 3. The case $c = 0$ does not depend on the choice of ω and correspond to the construction of Section 3.1, while the case $c = 1$ corresponds to the process of Equation 17.

[3] If η is an m-form on an m-dimensional manifold, the symbol $|\eta|$ denotes the associated density, in the sense of tensors.

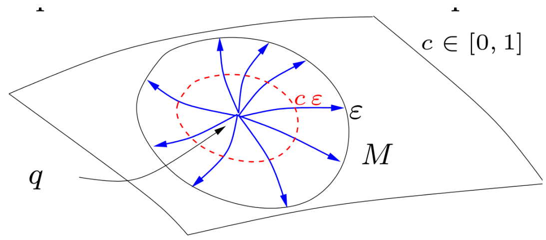

FIGURE 1. Geodesic random walk with sampling of the volume ω and ratio c. The case $c = 0$ corresponds to a geodesic random walk with no volume sampling. For each ε, the paths of the walk are piece-wise smooth geodesics.

For $\varepsilon > 0$, let $P^{\varepsilon}_{\omega,c}$ be the probability measure on the space of continuous paths on M associated with the process $b^{\varepsilon}_{t,c}$, and consider the family of operators

$$(L^{\varepsilon}_{c,\omega}\phi)(q) := \frac{1}{\delta}\mathbb{E}[\phi(b^{\varepsilon}_{\delta}) - \phi(q) \mid b^{\varepsilon}_0 = q] \tag{22}$$

$$:= \frac{1}{\delta}\int_{S_qM}[\phi(\exp_q(\varepsilon; v)) - \phi(q)]\mu^{c\varepsilon}_q(v), \qquad \forall q \in M, \tag{23}$$

for any $\phi \in C^\infty(M)$. Our first result is that the family of Riemannian geodesic random walk with volume sampling converges to a well-defined diffusion.

Theorem 8. *Let $(M, \mathbf{g})$ be a Riemannian manifold with a volume $\omega = e^h\mathcal{R}$, where $\mathcal{R}$ is the Riemannian volume and $h \in C^\infty(M)$. Let $c \in [0, 1]$. Then $L^{\varepsilon}_{c,\omega} \to L_{\omega,c}$, where*

$$\begin{aligned} L_{\omega,c} &= \Delta_{\mathcal{R}} + 2c\,\mathrm{grad}(h) \\ &= \Delta_\omega + (2c-1)\,\mathrm{grad}(h). \end{aligned} \tag{24}$$

Moreover, if $(M, \mathbf{g})$ is complete, then $P^{\varepsilon}_{\omega,c} \to P_{\omega,c}$ weakly, where $P_{\omega,c}$ is the law of the process associated with $L_{\omega,c}$.

Remark 4. We have these alternative forms of (24), obtained by unraveling the definitions:

$$L_{\omega,c} = \Delta_{e^{(2c-1)h}\omega} \tag{25}$$

$$= \Delta_{e^{2ch}\mathcal{R}} \tag{26}$$

$$= \sum_{i=1}^n X_i^2 + (2c\,\mathrm{div}_\omega(X_i) + (1-2c)\,\mathrm{div}_{\mathcal{R}}(X_i))\,X_i, \tag{27}$$

where, in the last line, $X_1, \dots, X_n$ is a local orthonormal frame.

As a simple consequence of (24) or its alternative formulations, we have the following statement, which appears to be new even in the Riemannian case.

Corollary 9. *Let $(M, \mathbf{g})$ be a complete Riemannian manifold. The operator $L_{\omega,c}$ with domain $C^\infty_c(M)$ is essentially self-adjoint on $L^2(M, \omega)$ if and only at least one of the following two conditions hold:*

(i) *$c = 1/2$;*
(ii) *ω is proportional to the Riemannian volume (i.e. h is constant).*

The previous discussion stresses the particular role played by the Riemannian volume. Not only does it coincide with the Hausdorff measure, but according to the above construction, it is the only volume (up to constant rescaling) that gives the correct self-adjoint operator for any choice of the parameter c.

Remark 5. If we want the volume-sampling scheme to produce the Laplacian with respect to the general volume being sampled ω, we should take $c = 1/2$. With hindsight, this might not be surprising. By linearity, we see that sampling with $c = 1/2$ is equivalent to sampling the volume along the entire step, uniformly with respect to time (recall that the geodesics of the steps are traversed with constant speed), rather than privileging any particular point along the path.

Remark 6. One can prove that the limit operator corresponding to the geodesic random walk with volume sampling ratio $c = 1$ is equal, up to a constant (given by the ratio of the area of the euclidean unit sphere and the volume of the unit ball in dimension n), to the limit operator corresponding to a more general class of random walk where we jump to points of the metric ball $B_q(\varepsilon)$ of radius ε, uniformly with respect to normalized measure $\omega/\omega(B_q(\varepsilon))$. This kind of random walk for the Riemannian measure has also been considered in [23], in relation with the study of its spectral properties.

4. Geodesic random walks in the sub-Riemannian setting

We want to define a sub-Riemannian version of the geodesic random walk with volume sampling, extending the Riemannian construction of the previous section. Recall the definition of (sub-)Riemannian manifold in Section 2.

4.1. Geodesics and exponential map.

As in Riemannian geometry, *geodesics* are horizontal curves that have constant speed and locally minimize the length between their endpoints. Define the *sub-Riemannian Hamiltonian* $H : T^*M \to \mathbb{R}$ as

$$H(\lambda) := \frac{1}{2}\sum_{i=1}^{k}\langle \lambda, X_i\rangle^2, \tag{28}$$

for any local orthonormal frame $X_1, \ldots, X_k \in \Gamma(\blacktriangle)$. Let σ be the natural symplectic structure on T^*M, and $\pi : T^*M \to M$. The *Hamiltonian vector field* $\vec{H}$ is the unique vector field on T^*M such that $dH = \sigma(\cdot, \vec{H})$. Then the Hamilton equations are

$$\dot{\lambda}(t) = \vec{H}(\lambda(t)). \tag{29}$$

Solutions of (29) are smooth curves on T^*M called *extremals*, and their projections $\gamma(t) := \pi(\lambda(t))$ on M are geodesics. In the Riemannian setting, all geodesics can be recovered uniquely in this way. In the sub-Riemannian one, this is no longer true, as *abnormal geodesics* can appear. These are minimizing trajectories that might not follow the Hamiltonian dynamics of (29).

For any $\lambda \in T^*M$ we consider the geodesic $\gamma_\lambda(t)$, obtained as the projection of the solution of (29) with initial condition $\lambda(0) = \lambda$. Observe that the Hamiltonian function, which is constant on $\lambda(t)$, measures the speed of the associated geodesic:

$$2H(\lambda) = \|\dot{\gamma}_\lambda(t)\|^2, \qquad \lambda \in T^*M. \tag{30}$$

Since H is fiber-wise homogeneous of degree 2, we have the following rescaling property:

$$\gamma_{\alpha\lambda}(t) = \gamma_\lambda(\alpha t), \qquad \alpha > 0. \tag{31}$$

This justifies the restriction to the subset of initial covectors lying in the level set $2H = 1$.

Definition 10. The *unit cotangent bundle* is the set of initial covectors such that the associated geodesic has unit speed, namely

$$\text{⌭} := \{\lambda \in T^*M \mid 2H(\lambda) = 1\} \subset T^*M. \tag{32}$$

For any $\lambda \in \text{⌭}$, the geodesic $\gamma_\lambda(t)$ is *parametrized by arc-length*, namely $\ell(\gamma|_{[0,T]}) = T$.

Remark 7. We stress that, in the genuinely sub-Riemannian case, $H|_{T_q^*M}$ is a degenerate quadratic form. It follows that the fibers ⌭_q are non-compact cylinders, in sharp contrast with the Riemannian case (where the fibers ⌭_q are spheres).

For any $\lambda \in \mathcal{C}$, the *cut time* $t_c(\lambda)$ is defined as the time where $\gamma_\lambda(t)$ loses optimality

$$t_c(\lambda) := \sup\{t > 0 \mid d(\gamma_\lambda(0), \gamma_\lambda(t)) = t\}. \tag{33}$$

In particular, for a fixed $\varepsilon > 0$ we define

$$\mathcal{C}_q^\varepsilon := \{\lambda \in \mathcal{C}_q \mid t_c(\lambda) \geq \varepsilon\} \subset \mathcal{C}_q, \tag{34}$$

as the set of unit covector such that the associated geodesic is optimal up to time ε.

Definition 11. Let $D_q \subseteq [0, \infty) \times T_q^*M$ the set of the pairs (t, λ) such that γ_λ is well defined up to time t. The *exponential map* at $q \in M$ is the map $\exp_q : D_q \to M$ that associates with (t, λ) the point $\gamma_\lambda(t)$.

Even if $\mathcal{C}$ is not compact, all arc-length parametrized geodesics are well defined for a sufficiently small time. The next lemma is a consequence of the form of Hamilton's equations and the compactness of small balls (see [1]).

Lemma 12. *For any compact $K \subset M$, there exists $\bar\varepsilon(K) > 0$ such that, for any $\varepsilon < \bar\varepsilon(K)$ we have $[0, \varepsilon) \times \mathcal{C}_q \subseteq D_q$. In other words all arc-length parametrized normal geodesics $\gamma_\lambda(t)$ are well defined on the interval $[0, \varepsilon)$.*

4.2. **Sub-Laplacians.** For any function $\phi \in C^\infty(M)$, the *horizontal gradient* $\mathrm{grad}(\phi) \in \Gamma(\blacktriangle)$ is, at each point, the horizontal direction of steepest slope of ϕ, that is

$$\mathbf{g}(\mathrm{grad}(\phi), X) = \langle d\phi, X\rangle, \qquad \forall X \in \Gamma(\blacktriangle). \tag{35}$$

Since in the Riemannian case this coincides with the usual gradient, this notation will cause no confusion. If $X_1, \dots, X_k$ is a local orthonormal frame, we have

$$\mathrm{grad}(\phi) = \sum_{i=1}^k X_i(\phi) X_i. \tag{36}$$

For any fixed volume form $\omega \in \Lambda^n M$ (or density if M is not orientable), the *divergence* of a smooth vector field X is defined by the relation $\mathcal{L}_X \omega = \mathrm{div}_\omega(X)$, where $\mathcal{L}$ denotes the Lie derivative. Notice that the sub-Riemannian structure does not play any role in the definition of div_ω. Following [24, 3], the *sub-Laplacian* on $(M, \blacktriangle, \mathbf{g})$ associated with ω is

$$\Delta_\omega \phi := \mathrm{div}_\omega(\mathrm{grad}(\phi)), \qquad \forall \phi \in C^\infty(M). \tag{37}$$

The sub-Laplacian is symmetric on the space $C_c^\infty(M)$ of smooth functions with compact support with respect to the $L^2(M, \omega)$ product. If (M, d) is complete and there are no non-trivial abnormal minimizers, then Δ_ω is essentially self-adjoint on $C_c^\infty(M)$ and has a smooth positive heat kernel [29, 30].

The sub-Laplacian will be intrinsic if we choose an intrinsic volume. See [10, Sec. 3] for a discussion of intrinsic volumes in sub-Riemannian geometry. A natural choice, at least in the equiregular setting, is Popp volume [7, 24], that is smooth. Other choices are possible, for example the Hausdorff or the spherical Hausdorff volume that, however, are not always smooth [2]. For the moment we let ω be a general smooth volume.

4.3. **The sub-Riemannian geodesic random walk with volume sampling.** In contrast with the Riemannian case, where S_qM has a well defined probability measure induced by the Riemannian structure, we have no such a construction on $\mathcal{C}_q$. Thus, it is not clear how to define a simple geodesic random walk in the sub-Riemannian setting.

For $\varepsilon > 0$, consider the sub-Riemannian exponential map

$$\exp_q(\varepsilon; \cdot) : \mathcal{C}_q \to M, \qquad q \in M. \tag{38}$$

For $\lambda \in \mathcal{C}_q$, then $\gamma_\lambda(\varepsilon) = \exp_q(\varepsilon; \lambda)$ is the associated unit speed geodesic starting at q.

One might be tempted to repeat Definition 7, using the exponential map to induce a density on $\mathcal{C}_q$, through the formula $\mu_q^\varepsilon(\lambda) \propto |(\exp_q(\varepsilon; \cdot)^* i_{\dot\gamma_\lambda(\varepsilon)} \omega)(\lambda)|$. However, there are non-trivial difficulties arising in the genuine sub-Riemannian setting.

- The exponential map is not a local diffeomorphism at $\varepsilon = 0$, and Riemannian normal coordinates are not available. This tool is used for proving the convergence of walks in the Riemannian setting;
- Due to the presence of zeroes in the Jacobian determinant of $\exp_q(\varepsilon;\cdot)$ for arbitrarily small ε, the absolute value in the definition of μ_q^ε is strictly necessary (in contrast with the Riemannian case, see Remark 2.
- Since $\mathcal{C}_q$ is not compact, there is no guarantee that $\int_{\mathcal{C}_q} \mu_q^\varepsilon < +\infty$;

Assuming that $\int_{\mathcal{C}_q} \mu_q^\varepsilon < +\infty$, we generalize Definition 7 as follows.

Definition 13. For any $q \in M$, and $\varepsilon > 0$, we define the family of densities μ_q^ε on $\mathcal{C}_q$

$$\mu_q^\varepsilon(\lambda) := \frac{1}{N(q,\varepsilon)} \left|(\exp_q(\varepsilon;\cdot)^* i_{\dot\gamma_\lambda(\varepsilon)}\omega)(\lambda)\right|, \qquad \forall \lambda \in \mathcal{C}_q, \tag{39}$$

where $N(q,\varepsilon)$ is fixed by the condition $\int_{\mathcal{C}_q} \mu_q^\varepsilon = 1$.

As we did in Section 3.2, and for $c \in (0,1]$, we build a random walk

$$b^\varepsilon_{(i+1)\delta,c} := \exp_{b^\varepsilon_{i\delta,c}}(\varepsilon;\lambda), \qquad \lambda \in \mathcal{C}_q \text{ with probability } \mu_q^{c\varepsilon}. \tag{40}$$

Let $P^\varepsilon_{\omega,c}$ be the associated probability measure on the space of continuous paths on M starting from q, and consider the corresponding family of operators, which in this case is

$$\begin{aligned}(L^\varepsilon_{c,\omega}\phi)(q) &= \frac{1}{\delta}\mathbb{E}[\phi(b^\varepsilon_{\delta,c}) - \phi(q) \mid b^\varepsilon_{0,c} = q] \\ &= \frac{1}{\delta}\int_{\mathcal{C}_q} [\phi(\exp_q(\varepsilon;\lambda)) - \phi(q)]\mu_q^{c\varepsilon}(\lambda), \qquad \forall q \in M,\end{aligned} \tag{41}$$

for any $\phi \in C^\infty(M)$. Clearly when $k = n$, (41) is the same family of operators associated with a Riemannian geodesic random walk with volume sampling discussed in Section 3.2, and this is why - without risk of confusion - we used the same symbol.

The problem of convergence is extremely complicated in general sub-Riemannian setting ($k < n$), due to the difficulties outlined above. We treat in detail the case of contact Carnot groups, where we find some surprising results. This class of structures is particularly important as they arise as Gromov-Hausdorff tangent cones of contact sub-Riemannian structures [8], and play the same role that Euclidean space plays in Riemannian geometry.

Remark 8. As we anticipated, in sub-Riemannian geometry abnormal geodesics may appear. More precisely, *strictly abnormal geodesics* do not arise as projections of solutions of (29). The class of processes that we defined never walk along these trajectories, but can walk along abnormal segments that are not strictly abnormal.

The (minimizing) Sard conjecture states that the set of endpoints of strictly abnormal (minimizing) geodesics starting from a given point has measure zero in M. However, this remains an hard open problem in sub-Riemannian geometry [5]. See also [22, 27, 4] for recent progresses on the subject.

4.4. **Contact Carnot groups.** Let $M = \mathbb{R}^{2d+1}$, with coordinates $(x,z) \in \mathbb{R}^{2d} \times \mathbb{R}$. Consider the following global vector fields

$$X_i = \partial_{x_i} - \frac{1}{2}(Ax)_i\partial_z, \qquad i = 1,\dots,2d, \tag{42}$$

where

$$A = \begin{pmatrix} \alpha_1 J & & \\ & \ddots & \\ & & \alpha_d J \end{pmatrix}, \qquad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{43}$$

is a skew-symmetric, non-degenerate matrix with singular values $0 < \alpha_1 \leq \ldots \leq \alpha_d$. A *contact Carnot group* is the sub-Riemannian structure on $M = \mathbb{R}^{2d+1}$ such that $\blacktriangle = \mathrm{span}\{X_1, \ldots, X_{2d}\}$ and $\mathbf{g}(X_i, X_j) = \delta_{ij}$. Notice that

$$[X_i, X_j] = A_{ij}\partial_z. \tag{44}$$

Set $\mathfrak{g}_1 := \mathrm{span}\{X_1, \ldots, X_{2d}\}$ and $\mathfrak{g}_2 := \mathrm{span}\{\partial_z\}$. The algebra $\mathfrak{g}$ generated by the X_i's and ∂_z admits a nilpotent stratification of step 2, that is

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2, \qquad \mathfrak{g}_1, \mathfrak{g}_2 \neq \{0\}, \tag{45}$$

with

$$[\mathfrak{g}_1, \mathfrak{g}_1] = \mathfrak{g}_2, \qquad \text{and} \qquad [\mathfrak{g}_1, \mathfrak{g}_2] = [\mathfrak{g}_2, \mathfrak{g}_2] = \{0\}. \tag{46}$$

There is a unique connected, simply connected Lie group G such that $\mathfrak{g}$ is its Lie algebra of left-invariant vector fields. The group exponential map,

$$\exp_G : \mathfrak{g} \to G, \tag{47}$$

associates with $v \in \mathfrak{g}$ the element $\gamma(1)$, where $\gamma : [0,1] \to G$ is the unique integral line of the vector field v such that $\gamma(0) = 0$. Since G is simply connected and $\mathfrak{g}$ is nilpotent, $\exp_G$ is a smooth diffeomorphism. Thus we can identify $G \simeq \mathbb{R}^{2d+1}$ with a polynomial product law $\star$ given by

$$(x, z) \star (x', z') = \left(x + x', z + z' + \tfrac{1}{2}x^* A x'\right). \tag{48}$$

Denote by L_q the left-translation $L_q(p) := q \star p$. The field X_i are left-invariant, and as a consequence also the sub-Riemannian distance is left-invariant, in the sense that $d(L_q(p_1), L_q(p_2)) = d(p_1, p_2)$.

Remark 9. As consequence of left-invariance, contact Carnot groups are complete as metric spaces. Moreover all abnormal minimizers are trivial. Hence, for each volume ω, the symmetric operator Δ_ω with domain $C_c^\infty(M)$ is also essentially self-adjoint in $L^2(M, \omega)$.

Example 1. The $2d+1$ dimensional *Heisenberg group* $\mathbb{H}_{2d+1}$, for $d \geq 1$, is the contact Carnot group where $\alpha_1 = \ldots = \alpha_d = 1$. The term Heisenberg group usually denotes the simplest 3D case, that is $\mathbb{H}_3$.

Example 2. The *bi-Heisenberg group* is the 5-dimensional contact Carnot group with $0 < \alpha_1 < \alpha_2$. That is, A has two distinct singular values.

A natural volume is the Popp volume $\mathcal{P}$. It turns out that that, up to constant scaling (see [7]), it coincides with the Lebesgue volume of $\mathbb{R}^{2d+1}$, that is

$$\mathcal{P} = \frac{1}{2\sum_{i=1}^d \alpha_i^2} dx_1 \wedge \ldots \wedge dx_{2d} \wedge dz, \tag{49}$$

and coincides also with the left-invariant Haar volume. One can check that $\mathrm{div}_{\mathcal{P}}(X_i) = 0$ hence the sub-Laplacian w.r.t. $\mathcal{P}$ is the sum of squares[4]:

$$\Delta_{\mathcal{P}} = \sum_{i=1}^{2d} X_i^2, \tag{50}$$

In this setting, we are able to prove the convergence of the sub-Riemannian random walk with volume sampling, with fixed volume $\mathcal{P}$.

[4]This is the case for any sub-Riemannian left-invariant structure on unimodular Lie group [3, 6].

Theorem 14. *Let $\mathbb{H}_{2d+1}$ be the Heisenberg group, with a general volume $\omega = e^h\mathcal{P}$. Then $L^\varepsilon_{\omega,c} \to L_{\omega,c}$, where*

$$L_{\omega,c} = \sigma(c)\left(\sum_{i=1}^{2d} X_i^2 + 2cX_i(h)\right) = \sigma(c)\left(\operatorname{div}_\omega \circ \operatorname{grad} + (2c-1)\operatorname{grad}(h)\right), \tag{51}$$

and $\sigma(c)$ is a constant (see Remark 10).

In particular $L_{\omega,c}$ it is essentially self-adjoint in $L^2(M,\omega)$ if and only if $c = 1/2$ or $\omega = \mathcal{P}$ (i.e. h is constant). The proof of the above theorem is omitted, as it is a consequence of the next, more general, result. In the general case, the picture is different, and quite surprising, since not even the principal symbol is the expected one.

Theorem 15. *Let $(\mathbb{R}^{2d+1}, \blacktriangle, \mathbf{g})$ be a contact Carnot group, equipped with a volume $\omega = e^h\mathcal{P}$ and let $c \in (0,1]$. Then $L^\varepsilon_{\omega,c} \to L_{\omega,c}$, where*

$$L_{\omega,c} = \sum_{i=1}^{d} \sigma_i(c)\left(X_{2i-1}^2 + X_{2i}^2\right) + 2c\sum_{i=1}^{d} \sigma_i(c)\left(X_{2i-1}(h)X_{2i-1} + X_{2i}(h)X_{2i}\right), \tag{52}$$

where $\sigma_1(c), \ldots, \sigma_d(c) \in \mathbb{R}$ are

$$\sigma_i(c) = \frac{cd}{(d+1)\sum_{i=1}^d \int_{-\infty}^{+\infty} |g_i(y)|dy} \sum_{\ell=1}^{d} (1+\delta_{\ell i}) \int_{-\infty}^{+\infty} |g_\ell(cp_z)| \frac{\sin(\frac{\alpha_i p_z}{2})^2}{(\alpha_i p_z/2)^2} dp_z, \tag{53}$$

and, for $i = 1, \ldots, d$

$$g_i(y) = \left(\prod_{j\neq i} \sin\left(\tfrac{\alpha_j y}{2}\right)\right)^2 \frac{\sin\left(\frac{\alpha_i y}{2}\right)\left(\frac{\alpha_i y}{2}\cos\left(\frac{\alpha_i y}{2}\right) - \sin\left(\frac{\alpha_i y}{2}\right)\right)}{(y/2)^{2d+2}}. \tag{54}$$

Moreover, $P^\varepsilon_{\omega,c} \to P_{\omega,c}$ weakly, where $P_{\omega,c}$ is the law of the process associated with $L_{\omega,c}$.

Remark 10. If $\alpha_1 = \ldots = \alpha_d = 1$ (Heisenberg), the functions $g_i = g$ are equal and

$$\sigma(c) := \sigma_i(c) = \frac{c}{\int_{\mathbb{R}} |g(y)|dy} \int_{\mathbb{R}} |g(cy)| \frac{\sin(y/2)^2}{(y/2)^2}, \tag{55}$$

does not depend on i. In general, however, $\sigma_i \neq \sigma_j$ (see Figure 2).

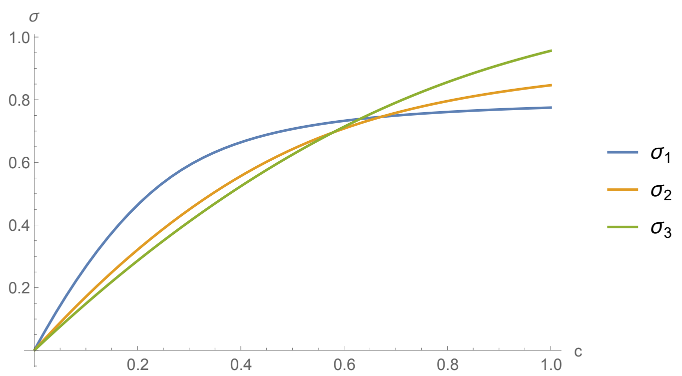

FIGURE 2. Values of $\sigma_i(c)$ for a contact Carnot group with $d = 3$ and $\alpha_1 = 1$, $\alpha_2 = 2$ and $\alpha_3 = 3$.

4.4.1. *An intrinsic formula.* We rewrite the operator of Theorem 15 in a more intrinsic form. Define a new contact Carnot group structure $(\mathbb{R}^{2d+1}, \blacktriangle, \mathbf{g}')$, by defining

$$X'_{2i-1} := \sqrt{\sigma_i(c)} X_{2i-1}, \qquad X'_{2i} := \sqrt{\sigma_i(c)} X_{2i}, \qquad i = 1, \dots, d, \tag{56}$$

to be a new orthonormal frame. Observe that this construction does not depend on the choice of ω. Let grad and grad$'$ denote the horizontal gradients w.r.t. the sub-Riemannian metrics $\mathbf{g}$ and $\mathbf{g}'$, respectively.

Lemma 16. *The limit operator $L_{\omega,c}$ of Theorem 15 is*

$$L_{\omega,c} = \mathrm{div}_\omega \circ \mathrm{grad}' + (2c-1)\,\mathrm{grad}'(h), \tag{57}$$

where $\mathrm{grad}'(h) = \sum_{i=1}^{2d} X'_i(h) X'_i$ *is understood as a derivation.*

Again $L_{\omega,c}$ it is essentially self-adjoint in $L^2(M,\omega)$ if and only if $c = 1/2$ or $\omega = \mathcal{P}$ (i.e. h is constant). In both cases it is a "divergence of the gradient", i.e. a well defined, intrinsic and symmetric operator but, surprisingly, not the expected one. In particular, the behavior of associated heat kernel (e.g. its asymptotics) depend not on the original sub-Riemannian metric $\mathbf{g}$, but the new one $\mathbf{g}'$.

4.4.2. *On the symbol.* We recall that the (principal) symbol of a smooth differential operator D on a smooth manifold M can be seen as a function $\Sigma(D) : T^*M \to \mathbb{R}$. The symbol associated with the sub-Riemannian geodesic random walk with volume sampling is

$$\Sigma(L_{\omega,c})(\lambda) = \sum_{i=1}^{d} \sigma_i(c)(\langle \lambda, X_{2i-1}\rangle^2 + \langle \lambda, X_{2i}\rangle^2), \qquad \lambda \in T^*M, \tag{58}$$

and does not depend on ω. On the other hand, the principal symbol of Δ_ω is

$$\Sigma(\Delta_\omega)(\lambda) = \sum_{i=1}^{2d} \langle \lambda, X_i \rangle^2 = 2H(\lambda), \qquad \lambda \in T^*M. \tag{59}$$

The two symbols are different, for any value of the sampling ratio $c > 0$. The reason behind this discrepancy is that the family of operators $L^\varepsilon_{\omega,c}$ keeps track of the different eigenspace associated with the generically different singular values $\alpha_i \neq \alpha_j$, through the Jacobian determinant of the exponential map.

4.5. **Alternative constructions for the sub-Riemannian random walk.** An alternative construction of the sub-Riemannian random walk of Section 4 is the following. For any fixed step length $\varepsilon > 0$, one follows only minimizing geodesics segments, that is $\lambda \in \text{⌭}_q^\varepsilon$ (the latter is defined in 34). In other words, for $\varepsilon > 0$, and $c \in (0,1]$, we consider the restriction of $\mu_q^{c\varepsilon}$ to ⌭_q^ε (which we renormalize in such a way that $\int_{\text{⌭}_q^\varepsilon} \mu_q^{c\varepsilon} = 1$).

Remark 11. In the the original construction the endpoints of the first step of the walk lie on the *front of radius ε centered at q*, that is the set $F_q(\varepsilon) = \exp_q(\varepsilon; \text{⌭}_q)$. With this alternative construction, the endpoints lie on the *metric sphere of radius ε, centered at q*, that is the set $S_q(\varepsilon) = \exp_q(\varepsilon; \text{⌭}_q^\varepsilon)$.

Remark 12. In the Riemannian setting, locally, for $\varepsilon > 0$ sufficiently small, all geodesics starting from q are optimal at least up to length ε, and the two constructions coincide.

This construction requires the explicit knowledge of ⌭_q^ε. For contact Carnot groups this is known [2]. We obtain the following convergence theorem, whose proof is similar to the one of Theorem 15, and thus omitted.

Theorem 17. *Consider the geodesic sub-Riemannian random walk with volume sampling, with volume ω and ratio c, defined according to the alternative constructions. Then the statement of Theorem 15 holds, replacing the constants $\sigma_1(c),\dots,\sigma_d(c)\in\mathbb{R}$ with*

$$\sigma_i^{alt}(c) = \frac{cd}{(d+1)\sum_{i=1}^d \int_{-2\pi/\alpha_d c}^{2\pi/\alpha_d c} |g_i(y)|dy} \sum_{\ell=1}^d (1+\delta_{\ell i}) \int_{-2\pi/\alpha_d}^{2\pi/\alpha_d} |g_\ell(cp_z)| \frac{\sin(\frac{\alpha_i p_z}{2})^2}{(\alpha_i p_z/2)^2} dp_z. \tag{60}$$

We call $L^{alt}_{\omega,c}$ the corresponding operator.

Remark 13 (The case $c=0$). In the Riemannian setting the case $c=0$ represents the geodesic random walk with no volume sampling of Section 3.1. In fact, by Theorem 8

$$L_{\omega,0} = \lim_{c\to 0^+} L_{\omega,c} = \mathrm{div}_{\mathcal{R}} \circ \mathrm{grad}, \qquad \text{(Riemannian geodesic RW)}, \tag{61}$$

is the Laplace-Beltrami operator, for any choice of the initial volume ω. In the sub-Riemannian setting the case $c=0$ is not defined, but we can still consider the limit for $c\to 0^+$ of the operator. In the original construction, $\lim_{c\to 0^+}\sigma_i(c)=0$ and by Theorem 15 we have:

$$\lim_{c\to 0^+} L_{\omega,c} = 0, \qquad \text{(sub-Riemannian geodesic RW)}. \tag{62}$$

For the alternative sub-Riemannian geodesic RW discussed above, we have:

$$\lim_{c\to 0^+} \sigma_i^{alt}(c) = \frac{d}{4\pi(d+1)}\left(1+\frac{\alpha_i^2}{\sum_{\ell=1}^d \alpha_\ell^2}\right)\int_{-2\pi}^{2\pi} \mathrm{sinc}\left(\frac{\alpha_i x}{2\alpha_d}\right)^2 dx, \qquad \forall i=1,\dots,d. \tag{63}$$

In particular from Lemma 16 and the change of volume formula, by Theorem 17 we have:

$$L^{alt}_{\omega,0} := \lim_{c\to 0^+} L^{alt}_{\omega,c} = \mathrm{div}_{\mathcal{P}} \circ \mathrm{grad}', \qquad \text{(alternative sub-Riemannian geodesic RW)}. \tag{64}$$

This is a non-zero operator, symmetric w.r.t. Popp volume, independently on the choice of the initial volume ω. Here grad' is the horizontal gradient defined w.r.t. the metric $\mathbf{g}'$ discussed in Section 4.4.1, with the coefficients $\sigma_i^{alt}(0) := \lim_{c\to 0^+}\sigma_i^{alt}(c) > 0$. Unless all α_i are equal, in general $\sigma_i^{alt}(0)\neq\sigma_j^{alt}(0)$ and grad' is not proportional to grad.

Notice that $L^{alt}_{\omega,0}\neq 0$ and it does not depend on the choice of the volume ω. This makes $L^{alt}_{\omega,0}$ (and the corresponding diffusion) an intriguing candidate for an intrinsic sub-Laplacian (and an intrinsic Brownian motion) for contact Carnot groups. For the Heisenberg group $\mathbb{H}_{2d+1}$, where $\alpha_i=1$ for all i, by Theorem 14 we have:

$$L^{alt}_{\omega,0} = \sigma^{alt}(0)\,\mathrm{div}_{\mathcal{P}} \circ \mathrm{grad}, \quad \text{where} \quad \sigma^{alt}(0) = \frac{1}{4\pi}\int_{-2\pi}^{2\pi} \mathrm{sinc}(x)^2 dx. \tag{65}$$

Remark 14 (Signed measures). A further alternative construction is the one in which we remove the absolute value in the definition 13 of μ_q^ε on $\boxminus_q$. In this case we lose the probabilistic interpretation, and we deal with signed measure. We still have an analogue of Theorem 15 for the operators, replacing the constants $\sigma_1(c),\dots,\sigma_d(c)\in\mathbb{R}$ with

$$\widetilde{\sigma}_i(c) = \frac{cd}{(d+1)\sum_{i=1}^d \int_{-\infty}^{+\infty} g_i(y)dy} \sum_{\ell=1}^d (1+\delta_{\ell i}) \int_{-\infty}^{+\infty} g_\ell(cp_z) \frac{\sin(\frac{\alpha_i p_z}{2})^2}{(\alpha_i p_z/2)^2} dp_z. \tag{66}$$

We observe the same qualitative behavior of the initial construction highligthed in Section 4.4.1 and 4.4.2.

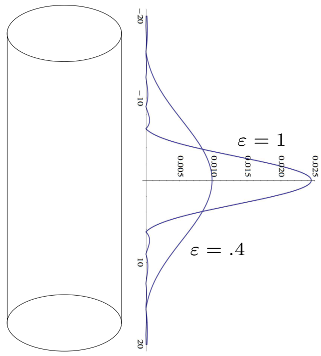

FIGURE 3. Measures on the cylinder for $c = 1$ in the Heisenberg group $\mathbb{H}_3$ for the original construction. Each zero corresponds to a conjugate point.

4.6. **The 3D Heisenberg group.** We give more details for the sub-Riemannian geodesic random walk in the 3D Heisenberg group. This is a contact Carnot group with $d = 1$ and $\alpha_1 = 1$. The origin of the group is $(x, z) = 0$. In coordinates $(p_x, p_z) \in T_0^*M$ we have

$$⌭_0 = \{(p_x, p_z) \in \mathbb{R}^2 \times \mathbb{R} \mid \|p_x\|^2 = 1\}, \tag{67}$$

$$⌭_0^\varepsilon = \{(p_x, p_z) \in \mathbb{R}^2 \times \mathbb{R} \mid \|p_x\|^2 = 1, \quad |p_z| \leq 2\pi/\varepsilon\}. \tag{68}$$

see [2]. For instance, we set ω equal to the Lebesgue volume. From the proof of Theorem 15, we obtain, in cylindrical coordinates $(\theta, p_z) \in \mathbb{S}^1 \times \mathbb{R} \simeq T_0^*M$

$$\mu_0^{c\varepsilon} = \begin{cases} \dfrac{c\varepsilon |g(c\varepsilon p_z)|}{2\pi \int_{-\infty}^{\infty} |g(y)| dy} d\theta \wedge dp_z & \text{original construction,} \\ \dfrac{c\varepsilon |g(c\varepsilon p_z)|}{2\pi \int_{2\pi c}^{2\pi c} |g(y)| dy} d\theta \wedge dp_z & \text{alternative construction,} \end{cases} \tag{69}$$

where

$$g(y) = \frac{\sin\left(\frac{y}{2}\right)\left(\frac{y}{2}\cos\left(\frac{y}{2}\right) - \sin\left(\frac{y}{2}\right)\right)}{(y/2)^4}. \tag{70}$$

The normalizations are determined by the conditions

$$\begin{cases} \int_{⌭_0} |\mu_0^{c\varepsilon}| = 1 & \text{original construction,} \\ \int_{⌭_0^\varepsilon} |\mu_0^{c\varepsilon}| = 1 & \text{alternative construction.} \end{cases} \tag{71}$$

The density corresponding to $\mu_0^{c\varepsilon}$, in coordinates (p_x, p_z) depends only on p_z. For any fixed $c > 0$, the density has larger and larger tails for $\varepsilon \to 0$, thus the probability to follow a geodesic with large p_z increases (see Fig. 3).

5. Flow random walks

The main difficulties to deal with convergence in the sub-Riemannian geodesic random walk with volume sampling scheme were related to the non-compactness of $⌭_q$, and the lack of a general asymptotics for μ_q^ε. To overcome these difficulties, we discuss a different class of walks. This approach is inspired by the classical integration of a Stratonovich SDE, and can be implemented on Riemannian and sub-Riemannian structures alike (the only requirement being a set of vector fields $X_1, \ldots, X_k$ on a smooth manifold M, and a volume ω for volume sampling).

5.1. **Stratonovich SDEs via flow random walks.** Let M be a smooth n-dimensional manifold, and let $V_1, \dots, V_k$ be smooth vector fields on M. Since SDEs are fundamentally local objects (at least in the case of smooth vector fields, where the SDE has a unique, and thus Markov, solution), we don't worry about the global behavior of the V_i, and thus we can assume, without loss of generality, that the flow along any vector field $V = \beta_1 V_1 + \beta_2 V_2 + \cdots + \beta_k V_k$ for any constants β_i exists for all time. Further, we can assume that there exists a Riemannian metric $\mathbf{g}$ on M such that the V_i all have bounded length.

We consider the Stratonovich SDE

$$dq_t = \sum_{i=1}^{k} V_i(q_t) \circ dw_t^i, \qquad q_0 = q, \tag{72}$$

for some $q \in M$, where $w_t^1, \dots, w_t^k$ are independent, one-dimensional Brownian motions. We recall that solving this SDE is essentially equivalent to solving the martingale problem for the operator $\sum_{i=1}^k V_i^2$. (See [21, Chapter 5] for the precise relationship between solutions to SDEs and solutions to martingale problems, although in this case, because of strong uniqueness of the solution to the (72), the situation is relatively simple.) We also assume that the solution to (72), which we denote q_t^0, does not explode.

The sequence of random walks which we associate to (72) is as follows. We take $\varepsilon > 0$. Consider the k-dimensional vector space of all linear combinations $\beta_1 V_1 + \beta_2 V_2 + \cdots + \beta_k V_k$. Then we can naturally identify $\mathbb{S}^{k-1}$ with the set $\sum_{i=1}^k \beta_i^2 = 1$, and thus choose a k-tuple $(\beta_1, \dots, \beta_k)$ from the sphere according to the uniform probability measure. This gives a random linear combination $V = \beta_1 V_1 + \beta_2 V_2 + \cdots + \beta_k V_k$. Now, starting from q, we flow along the vector field $\frac{2k}{\varepsilon} V$ for time $\delta = \varepsilon^2/(2k)$, traveling a curve of length $\varepsilon \|V\|_{\mathbf{g}}$. This determines the first step, q_t^ε with $t \in [0, \delta]$, of a random walk (and the measure Π_q^ε). Determining each additional step in the same way produces a family of random walks q_t^ε, that we call *flow random walk* at scale ε associated with the SDE (72).

We associate to each process q_t^ε and q_t^0 the corresponding probability measures P^ε and P^0 on $\Omega(M)$. Then Theorem 6 shows that P^ε converges to P^0 weakly as $\varepsilon \to 0$. Note that since this holds for any metric $\mathbf{g}$ as described above, this is really a statement about processes on M as a smooth manifold, and the occurrence of $\mathbf{g}$ is just an artifact of the formalism of Theorem 6.

The relationship of Stratonovich integration to ODEs, and thus flows of vector fields, is not new. Approximating the solution to a Stratonovich SDE by an ODE driven by an approximation to Brownian motion is considered in [11] and [33]. Here, we have tried to give a simple, random walk approach emphasizing the geometry of the situation. Nonetheless, because M is locally diffeomorphic to $\mathbb{R}^n$ (or a ball around the origin in $\mathbb{R}^n$, depending on one's preferences) and the entire construction is preserved by diffeomorphism, there is nothing particularly geometric about the above, except perhaps the observation that the construction is coordinate independent.

5.2. **Volume sampling through the flow.** The random walk defined in the previous section, which depends only on the choice of k smooth, complete, vector field $X_1, \dots, X_k$ fits in the general class of walks of Section 2. Moreover, the construction can be generalized to include a volume sampling technique, as we now describe.

Here $X_1, \dots, X_k$ are a fixed set of global orthonormal fields of a (sub-)Riemannian structure, and all our definitions will depend on this choice. We will discuss in which cases the limit diffusion does not depend on such a choice.

Definition 18. For any $q \in M$, and $\varepsilon > 0$, the *end-point map* $E_{q,\varepsilon} : \mathbb{R}^k \to M$ gives the point $E_{q,\varepsilon}(u)$ at time ε of integral curve of the vector field $X_u := \sum_{i=1}^k u_i X_i$ starting from $q \in M$. Moreover, let $S_{q,\varepsilon} := E_{q,\varepsilon}(\mathbb{S}^{k-1})$.

Remark 15. For small $\varepsilon \geq 0$, $E_{q,\varepsilon} : \mathbb{S}^{k-1} \to S_{q,\varepsilon}$ is a diffeomorphism, and for any unit $u \in \mathbb{S}^{k-1}$, then $\gamma_u(\varepsilon + \tau) := E_{q,\varepsilon+\tau}(u)$ is a segment of flow line transverse to $S_{q,\varepsilon}$.

The next step is to induce a probability measure μ_q^ε on $\mathbb{S}^{k-1}$ via volume sampling through the endpoint map. We start with the Riemannian case.

5.3. **Flow random walks with volume sampling in the Riemannian setting.** In this case $k = n$ and the choice for the volume sampling scheme is quite natural.

Definition 19. Let $(M, \mathbf{g})$ be a Riemannian manifold. For any $q \in M$ and $\varepsilon > 0$, we defined the family of densities on μ_q^ε on $\mathbb{S}^{n-1}$

$$\mu_q^\varepsilon(u) := \frac{1}{N(q,\varepsilon)} \left| E^*_{q,\varepsilon} \iota_{\dot\gamma_u(\varepsilon)} \omega)(u) \right|, \qquad \forall u \in \mathbb{S}^{n-1}. \tag{73}$$

where $N(q,\varepsilon)$ is fixed by the condition $\int_{\mathbb{S}^{n-1}} \mu_q^\varepsilon = 1$. For $\varepsilon = 0$, we set μ_q^0 the standard normalized density on $\mathbb{S}^{n-1}$.

Then, we define a random walk by choosing $u \in \mathbb{S}^{k-1}$ according to μ_q^ε, and following the corresponding integral line. That is, for $\varepsilon > 0$

$$r^\varepsilon_{(i+1)\delta,c} := E_{r^\varepsilon_{i\delta,c},\varepsilon}(u), \qquad u \in \mathbb{S}^{n-1} \text{ with probability } \mu_q^{c\varepsilon}, \tag{74}$$

where we also introduced the parameter $c \in [0,1]$ in the volume sampling. This class of walks include the one described in the above section (by setting $c = 0$).

Let $P^\varepsilon_{\omega,c}$ be the probability measure on the space of continous paths on M associated with $r^\varepsilon_{t,c}$ and consider the associated family of operators that, in this case are

$$(L^\varepsilon_{\omega,c}\phi)(q) := \frac{1}{\delta} \int_{\mathbb{S}^{n-1}} [\phi(E_{q,\varepsilon}(u)) - \phi(q)] \mu_q^{c\varepsilon}(u), \qquad \forall q \in M, \tag{75}$$

for any $\phi \in C^\infty(M)$.

Theorem 20. *Let $(M, \mathbf{g})$ be a Riemannian manifold and $X_1, \dots, X_n$ be a global set of orthonormal vector fields. Let $c \in [0,1]$ and $\omega = e^h \mathcal{R}$ be a fixed volume on M, for some $h \in C^\infty(M)$. Then $L^\varepsilon_{c,\omega} \to L_{c,\omega}$, where*

$$L_{\omega,c} = \Delta_\omega + c \operatorname{grad}(h) + (c-1) \sum_{i=1}^n \operatorname{div}_\omega(X_i) X_i. \tag{76}$$

Moreover, if $(M, \mathbf{g})$ is complete, $P^\varepsilon_{\omega,c} \to P_{\omega,c}$ weakly, where $P_{\omega,c}$ is the law of the process associated with $L_{\omega,c}$.

The limit operator is not intrinsic in general, as it clearly depends on the choice of the orthonormal frame. However, thanks to the explicit formula, we have the following.

Corollary 21. *The operator $L_{\omega,c}$ does not depend on the choice of the orthonormal frame if and only if $c = 1$. In this case*

$$L_{\omega,1} = \Delta_\omega + \operatorname{grad}(h) = \Delta_{e^h\omega} = \Delta_{e^{2h}\mathcal{R}}. \tag{77}$$

Once again, we get a surprising result. In fact, even if $L_{\omega,1}$ is intrinsic and depends only on the Riemannian structure and on the volume ω, it is not symmetric on $L^2(M,\omega)$ unless we choose h to be constant.

Corollary 22. *The operator $L_{\omega,c}$ with domain $C_c^\infty(M)$ is essentially self-adjoint on $L^2(M,\omega)$ if and only if $c = 1$ and ω is proportional to the Riemannian volume.*

The natural request of independence on the choice of the orthonormal frame and symmetry on $L^2(M,\omega)$ leaves no choice and, in particular, selects a preferred volume $\omega = \mathcal{R}$, up to a proportionality constant.

On the other hand, by setting $c = 0$, we recover the "sum of squares" generator of the solution of the Stratonovich SDE (72).

Corollary 23. *The operator $L_{\omega,0}$ depends on the choice of the vector fields $X_1, \dots, X_n$, but not on the choice of the volume ω, in particular*

$$L_{\omega,0} = \sum_{i=1}^{n} X_i^2. \tag{78}$$

5.4. **Flow random walks with volume sampling in the sub-Riemannian setting.** To extend the flow random walk construction to the sub-Riemannian setting we need $Z_1, \dots, Z_{n-k}$ vector fields on M, transverse to $\blacktriangle$, in such a way that $\iota_{Z_1,\dots,Z_{n-k}}\omega$ is a well defined k-form that we can use to induce a measure on $\mathbb{S}^{k-1}$ as in Definition 19.

In general there is no natural choice of these $Z_1, \dots, Z_{n-k}$. We explain the construction in detail for contact sub-Riemannian structures, where such a natural choice exists. Indeed this class contains contact Carnot groups.

5.4.1. *Contact sub-Riemannian structures.* A sub-Riemannian structure $(M, \blacktriangle, \mathbf{g})$ is *contact* if there exists a global one-form η such that $\blacktriangle = \ker \eta$. This forces $\dim(M) = 2d + 1$ and $\operatorname{rank} \blacktriangle = 2d$, for some $d \geq 1$. Consider the skew-symmetric *contact endomorphism* $J : \Gamma(\blacktriangle) \to \Gamma(\blacktriangle)$, defined by the relation

$$g(X, JY) = d\eta(X, Y), \qquad \forall X, Y \in \Gamma(\blacktriangle). \tag{79}$$

We assume that J is non-degenerate. Multiplying η by a non-zero smooth function f gives the same contact structure, with contact endomorphism fJ. We fix η up to sign by taking

$$\operatorname{Tr}(JJ^*) = 1. \tag{80}$$

The Reeb vector field is defined as the unique vector X_0 such that

$$\eta(X_0) = 1, \qquad \iota_{X_0} d\eta = 0. \tag{81}$$

In this case (see [7]), the Popp density is the unique density such that $\mathcal{P}(X_0, X_1, \dots, X_{2d}) = 1$ for any orthonormal frame $X_1, \dots, X_{2d}$ of $\blacktriangle$.

The flow random walk with volume sampling construction, with volume ω and sampling ratio c, can be implemented as follows.

Definition 24. Let $(M, \blacktriangle, \mathbf{g})$ be a contact sub-Riemannian structure with Reeb vector field X_0. For any $q \in M$ and $\varepsilon > 0$ we define the family of densities μ_q^ε on $\mathbb{S}^{k-1}$

$$\mu_q^\varepsilon(u) := \frac{1}{N(q,\varepsilon)} \left| (E_{q,\varepsilon}^* \circ i_{X_0, \dot\gamma_u(\varepsilon)} \omega)(u) \right|, \qquad \forall u \in \mathbb{S}^{k-1}, \tag{82}$$

where $N(q, \varepsilon)$ is fixed by the condition $\int_{\mathbb{S}^{k-1}} \mu_q^\varepsilon = 1$. For $\varepsilon = 0$, we set μ_q^0 the standard normalized density on $\mathbb{S}^{k-1}$.

We define a random walk $r_{t,c}^\varepsilon$ as in (74), with sampling ratio $c \in [0, 1]$, and we call the associated family of operators $L_{\omega,c}^\varepsilon$ as in (75), with no risk of confusion.

Theorem 25. *Let $(M, \blacktriangle, \mathbf{g})$ be a contact sub-Riemannian manifold and $X_1, \dots, X_{2d}$ be a global set of orthonormal vector fields. Let $c \in [0, 1]$ and $\omega = e^h \mathcal{P}$ be a fixed volume on M, for some $h \in C^\infty(M)$. Then $L_{c,\omega}^\varepsilon \to L_{c,\omega}$, where*

$$L_{\omega,c} = \Delta_\omega + c \operatorname{grad}(h) + (c-1) \sum_{i=1}^{k} \operatorname{div}_\omega(X_i) X_i. \tag{83}$$

Moreover, if $(M, \mathbf{g})$ is complete, $P_{\omega,c}^\varepsilon \to P_{\omega,c}$ weakly, where $P_{\omega,c}$ is the law of the process associated with $L_{\omega,c}$.

This construction, in the contact sub-Riemannian case, has the properties of the Riemannian one, where the Riemannian volume is replaced by Popp one. In particular we have the following analogues of Corollaries 21 22 and 23.

Corollary 26. *The operator $L_{\omega,c}$ does not depend on the choice of the orthonormal frame if and only if $c = 1$. In this case*

$$L_{\omega,1} = \Delta_\omega + \mathrm{grad}_H(h) = \Delta_{e^h\omega} = \Delta_{e^{2h}\mathcal{P}}. \tag{84}$$

Corollary 27. *The operator $L_{\omega,c}$ with domain $C_c^\infty(M)$ is essentially self-adjoint on $L^2(M,\omega)$ if and only if $c = 1$ and ω is proportional to the Popp volume.*

Corollary 28. *The operator $L_{\omega,0}$ depends on the choice of the vector fields $X_1, \dots, X_k$, but not on the choice of the volume ω, in particular*

$$L_{\omega,0} = \sum_{i=1}^k X_i^2. \tag{85}$$

6. Proof of the results

6.1. **Proof of Theorem 8.** Let $q \in M$, and fix normal coordinates $(x_1, \dots, x_n)$ on a compact neighborhood K of q. In these coordinates, length parametrized geodesics are straight lines εv, with $v \in S_qM \simeq \mathbb{S}^{n-1}$. In particular

$$\phi(\exp_q(\varepsilon, v)) - \phi(q) = \varepsilon \sum_{i=1}^n v_i \partial_i \phi + \frac{1}{2}\varepsilon^2 \sum_{i,j=1}^n v_i v_j \partial_{ij}^2 \phi + \varepsilon^3 O_q, \tag{86}$$

where all the derivatives are computed at q, and the remainder term O_q is uniformly bounded on K by a constant M_K. When $\omega = \mathcal{R}$ is the Riemannian volume, a well known asymptotics (see, for instance, [14]) gives

$$\mu_q^{c\varepsilon}(v) = (1 + \varepsilon^2 O_q) d\Omega, \tag{87}$$

where $d\Omega$ is the normalized euclidean measure on $\mathbb{S}^{n-1}$. When $\omega = e^h\mathcal{R}$, the above formula is multiplied by a factor $e^{h(\exp_q(c\varepsilon, v))}$, and taking in account the normalization we obtain

$$\mu_q^{c\varepsilon}(v) = (1 + \varepsilon c v_i \partial_i h + \varepsilon^2 O_q) d\Omega. \tag{88}$$

Then, for the operator $L_{\omega,c}^\varepsilon \phi$, evaluated at q, we obtain

$$(L_{\omega,c}^\varepsilon \phi)|_q = \frac{2n}{\varepsilon^2} \int_{S_qM} [\phi(\exp_q(\varepsilon, v)) - \phi(q)] \mu_q^{c\varepsilon}(v) \tag{89}$$

$$= \frac{2n}{\varepsilon} \sum_{i=1}^n \partial_i \phi \int_{\mathbb{S}^{n-1}} v_i d\Omega + 2n \sum_{i,j=1}^n \left(c \partial_i h \partial_j \phi + \frac{1}{2} \partial_{ij}^2 \phi \right) \int_{\mathbb{S}^{n-1}} v_j v_i d\Omega + \varepsilon O_q. \tag{90}$$

The first integral is zero, while $\int_{\mathbb{S}^{n-1}} v_i v_j d\Omega = \delta_{ij}/n$. Thus we obtain

$$(L_{\omega,c}^\varepsilon \phi)|_q = \sum_{i=1}^n \partial_{ii}^2 \phi + 2c(\partial_i h)(\partial_i \phi) + \varepsilon O_q. \tag{91}$$

The first term is the Laplace-Beltrami operator applied to ϕ, written in normal coordinates, while the second term coincides with the action of the derivation $2c\,\mathrm{grad}(h)$ on ϕ, evaluated at q. Since the l.h.s. is invariant by changes of coordinates, we have $L_{\omega,c}^\varepsilon \to L_{\omega,c}$, where

$$L_{\omega,c} = \Delta_{\mathcal{R}} + 2c\,\mathrm{grad}(h). \tag{92}$$

The alternative forms of the statement follow from the change of volume formula $\Delta_{e^h\omega} = \Delta_\omega + \mathrm{grad}(h)$. The weak convergence $P_{\omega,c}^\varepsilon \to P_{\omega,c}$ follows from Theorem 6. □

6.2. **Proof of Theorem 15.** We start with the case $h = 0$ and $q = 0$. Hamilton equations for a contact Carnot groups are readily solved, and the geodesic with initial covector $(p_x, p_z) \in T_0^*M \simeq \mathbb{R}^{2d} \times \mathbb{R}$ is

$$x(t) = \int_0^t e^{sp_z A} p_x ds, \qquad z(t) = -\frac{1}{2}\int_0^t \dot{x}^*(s)Ax(s)ds. \tag{93}$$

It is convenient to split $p_x = (p_x^1, \dots, p_x^d)$, where $p_x^i = (p_{x_{2i-1}}, p_{x_{2i}}) \in \mathbb{R}^2$ is the projection of p_x in the real eigenspace of A corresponding to the singular value α_i. We get

$$\exp_0(t; p_x, p_z) = \begin{pmatrix} B(t;\alpha_1 p_z)p_x^1 \\ \vdots \\ B(t;\alpha_d p_z)p_x^d \\ \sum_{i=1}^d b(t;\alpha_i p_z)\alpha_i \|p_x^i\|^2 \end{pmatrix}, \tag{94}$$

where

$$B(t;y) := \frac{\sin(ty)}{y}\mathbb{I} + \frac{\cos(ty)-1}{y}J, \qquad b(t;y) := \frac{ty - \sin(ty)}{2y^2}. \tag{95}$$

If $p_z = 0$, the equations above must be understood in the limit, thus $\exp_0(t; p_x, 0) = (tp_x, 0)$. The Jacobian determinant is computed in [2] (see also [28] for the more general case of a corank 1 Carnot group with a notation closer to the one of this paper):

$$\det(d_{p_x,p_z}\exp_0(t;\cdot)) = \frac{t^{2d+3}}{4\alpha^2}\sum_{i=1}^d g_i(tp_z)\|p_x^i\|^2, \tag{96}$$

where $\alpha = \prod_{i=1}^d \alpha_i$ and

$$g_i(y) := \left(\prod_{j\neq i}\sin\left(\tfrac{\alpha_j y}{2}\right)\right)^2 \frac{\sin\left(\frac{\alpha_i y}{2}\right)\left(\frac{\alpha_i y}{2}\cos\left(\frac{\alpha_i y}{2}\right) - \sin\left(\frac{\alpha_i y}{2}\right)\right)}{(y/2)^{2d+2}}. \tag{97}$$

Lemma 29. *For any $\lambda \in T_q^*M$ and $t > 0$, we have (up to the normalization)*

$$(\exp_q(t;\cdot)^* i_{\dot{\gamma}_\lambda(t)}\omega)(\lambda) = \frac{1}{t} i_\lambda(\exp_q(t;\cdot)^*\omega)(\lambda). \tag{98}$$

Proof. It follows from the homogeneity property $\exp_q(t;\alpha\lambda) = \exp_q(\alpha t;\lambda)$, for all $\alpha \in \mathbb{R}$:

$$\dot{\gamma}_\lambda(t) = \left.\frac{d}{d\tau}\right|_{\tau=0}\exp_q(t+\tau;\lambda) = \frac{1}{t}\left.\frac{d}{d\tau}\right|_{\tau=0}\exp_q(t;(1+\tau)\lambda) = \frac{1}{t}d_\lambda\exp_q(t;\cdot)\lambda, \tag{99}$$

where we used the standard identification $T_\lambda(T_q^*M) = T_q^*M$. □

The cylinder is $⌭_0 = \{(p_x, p_z) \mid \|p_x\|^2 = 1\} \subset T_0^*M$ and $\lambda \simeq p_z\partial_{p_z} + p_x\partial_{p_x}$. The Lebesgue volume is $\mathscr{L} = dx \wedge dz$. By Lemma 29 and reintroducing the normalization factor, we obtain that the restriction to $⌭_0$ of μ_0^t is

$$\mu_0^t = \frac{1}{N(t)}\sum_{i=1}^d |g_i(p_z t)|\|p_x^i\|^2 |d\Omega \wedge dp_z|, \tag{100}$$

where $d\Omega$ is the normalized volume of $\mathbb{S}^{2d-1}$. Observe that each $|g_i| \in L^1(\mathbb{R})$. Thus

$$N(t) = \sum_{i=1}^d \int_{\mathbb{S}^{2d-1}} \|p_x^i\|^2 d\Omega \int_{\mathbb{R}} dp_z |g_i(p_z t)| = \frac{1}{dt}\sum_{i=1}^d \int_{\mathbb{R}} dy|g_i(y)|. \tag{101}$$

To compute $\mathbb{E}[\phi(\exp_q(\varepsilon;\lambda)) - \phi(q)]$, we can assume $\phi(q) = 0$. Hence

(102)
$$\int_{\mathbb{C}_0} \phi(\exp_0(\varepsilon;\lambda))\mu_0^{c\varepsilon}(\lambda) = \frac{1}{N(c\varepsilon)}\sum_{i=1}^d \int_{\mathbb{S}^{2d-1}} d\Omega \int_{\mathbb{R}} dp_z |g_i(p_z c\varepsilon)| \|p_x^i\|^2 \phi(\exp_0(\varepsilon; p_x, p_z))$$

$$(103)\qquad = \frac{c}{\varepsilon N(\varepsilon)}\sum_{i=1}^d \int_{\mathbb{S}^{2d-1}} \|p_x^i\|^2 d\Omega \int_{\mathbb{R}} dy |g_i(cy)| \phi(\exp_0(\varepsilon; p_x, y/\varepsilon))$$

$$(104)\qquad = \frac{cd}{\sum_{i=1}^d \int_{\mathbb{R}} |g_i(y)| dy}\sum_{i=1}^d \int_{\mathbb{S}^{2d-1}} \|p_x^i\|^2 d\Omega \int_{\mathbb{R}} dp_z |g_i(cp_z)| \phi(\exp_0(1; \varepsilon p_x, p_z)),$$

where we used the rescaling property of the exponential map. From (94) we get

$$(105)\qquad \exp_0(1;\varepsilon p_x, p_z) = \left(\begin{pmatrix} B(\alpha_1 p_z) & & \\ & \ddots & \\ & & B(\alpha_d p_z) \end{pmatrix} \varepsilon p_x, \sum_{i=1}^d b(\alpha_i p_z)\alpha_i \|p_x^i\|^2 \varepsilon^2 \right),$$

where, with a slight abuse of notation

$$(106)\qquad B(y) = \frac{\sin(y)}{y}\mathbb{I} + \frac{\cos(y)-1}{y}J, \qquad b(y) = \frac{y - \sin(y)}{2y^2}.$$

We observe here that

$$(107)\qquad B(y)B(y)^* = \frac{\sin(y/2)^2}{(y/2)^2}\mathbb{I}.$$

It is convenient to rewrite

$$(108)\qquad \exp_0(1;\varepsilon p_x, p_z) = (\mathbf{B}(p_z)\varepsilon p_x, \varepsilon^2 p_x^* \mathbf{b}(p_z) p_x),$$

where $\mathbf{B}(p_z)$ is a block-diagonal $2d \times 2d$ matrix, whose blocks are $B(\alpha_i p_z)$, and $\mathbf{b}$ is a $2d \times 2d$ diagonal matrix. Notice that $\exp_0(1;\varepsilon p_x, p_z)$ is contained in the compact metric ball of radius ε. Hence, we have

$$(109)\qquad \begin{aligned} \phi(\exp_0(\varepsilon; tp_x, p_z)) &= (\partial_x\phi)(\mathbf{B}(p_z)\varepsilon p_x) + (\partial_z\phi)p_x^*\mathbf{b}(p_z)p_x\varepsilon^2 \\ &\quad + \frac{1}{2}\varepsilon^2(\mathbf{B}(p_z)p_x)^*(\partial_x^2\phi)(\mathbf{B}(p_z)p_x) + \varepsilon^3 R_{(p_x,p_z)}(\varepsilon). \end{aligned}$$

The derivatives of ϕ are computed at 0. If $\varepsilon \leq \varepsilon_0$, the remainder is uniformly bounded by the derivatives, up to order three, of ϕ on the compact metric ball of radius ε_0, that is $|R_{(p_x,p_z)}(\phi)| \leq M_0$. When plugging (109) back in (104), we observe that the term proportional to

$$(110)\qquad \int_{\mathbb{S}^{2d-1}} \|p_x^i\|^2 d\Omega \int_{\mathbb{R}} dp_z |g_i(cp_z)| (\partial_x\phi)\mathbf{B}(p_z)\varepsilon p_x$$

vanishes, as the integral of any odd-degree monomial in p_x on the sphere is zero. Furthermore, the term proportional to

$$(111)\qquad \int_{\mathbb{S}^{2d-1}} \|p_x^i\|^2 d\Omega \int_{\mathbb{R}} dp_z |g_i(cp_z)| (\partial_z\phi)p_x^*\mathbf{b}(p_z)p_x t^2$$

vanishes, as the integrand is an odd function of p_z. The last second order (in ε) term is

$$(112)\qquad \frac{cd}{\sum_{i=1}^d \int_{\mathbb{R}} |g_i(y)| dy}\sum_{i=1}^d \int_{\mathbb{S}^{2d-1}} \|p_x^i\|^2 d\Omega \int_{\mathbb{R}} dp_z |g_i(cp_z)| \frac{1}{2}\varepsilon^2(\mathbf{B}(p_z)p_x)^*(\partial_x^2\phi)(\mathbf{B}(p_z)p_x).$$

If all the α_i are equal, then all $g_i = g$, and (112) the sum $\sum_{i=1}^d \|p_x^i\|^2 = \|p_x\|^2 = 1$ simplifies. In this case we have a simple average of a quadratic form on $\mathbb{S}^{2d-1}$. When the α_i are distinct, we need the following results.

Lemma 30 (see [13])**.** *Let $P(x) = x_1^{a_1} \dots x_n^{a_n}$ a monomial in $\mathbb{R}^n$, with $a_i, \dots, a_n \in \{0, 1, 2, \dots\}$. Set $b_i := \frac{1}{2}(a_i + 1)$ Then*

$$\int_{\mathbb{S}^{n-1}} P(x) d\Omega = \frac{\Gamma(n/2)}{2\pi^{n/2}} \begin{cases} 0 & \text{if some } a_j \text{ is odd,} \\ \frac{2\Gamma(b_1)\Gamma(b_2)\cdots\Gamma(b_n)}{\Gamma(b_1+b_2+\cdots+b_n)} & \text{if all } a_j \text{ are even,} \end{cases} \tag{113}$$

where $d\Omega$ is the normalized measure on the sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$.

Lemma 31. *Let $Q(x) = x^* Q x$ and $R(x) = x^* R x$ be two quadratic forms on $\mathbb{R}^n$, such that $QR = RQ$. Then*

$$\int_{\mathbb{S}^{n-1}} Q(x) R(x) d\Omega = \frac{2 \operatorname{Tr}(QR) + \operatorname{Tr}(Q) \operatorname{Tr}(R)}{n(n+2)}. \tag{114}$$

If $R = \mathbb{I}$, we recover the usual formula $\int_{\mathbb{S}^{n-1}} Q d\Omega = \frac{1}{n} \operatorname{Tr}(Q)$.

Proof. Up to an orthogonal transformation, we can assume that Q and R are diagonal. Hence (for brevity we omit the domain of integration and the measure)

$$\int Q(x) R(x) = \sum_{i,j=1}^n Q_{ii} R_{jj} \int x_i^2 x_j^2. \tag{115}$$

By Lemma 30

$$\int x_i^2 x_j^2 = \begin{cases} \frac{3}{n(n+2)} & i = j, \\ \frac{1}{n(n+2)} & i \neq j. \end{cases} \tag{116}$$

Thus

$$\int Q(x) R(x) = \sum_{i,j=1}^n Q_{ii} R_{jj} \int x_i^2 x_j^2 \left(\delta_{ij} + (1 - \delta_{ij})\right) \tag{117}$$

$$= \frac{1}{n(n+2)} \sum_{i,j}^n Q_{ii} R_{jj} (3\delta_{ij} + (1 - \delta_{ij})) = \frac{2 \operatorname{Tr}(QR) + \operatorname{Tr}(Q) \operatorname{Tr}(R)}{n(n+2)}. \tag{118}$$

If Q and R do not commute we cannot expect such a simple expression, that in general depends on the whole set of invariants of the pair of quadratic forms. □

We can write (112), as the sum of integrals of products of quadratic forms over $\mathbb{S}^{2d-1}$

$$\frac{1}{2} \varepsilon^2 \frac{cd}{\sum_{i=1}^d \int_{\mathbb{R}} |g_i(y)| dy} \sum_{i=1}^d \int_{\mathbb{R}} dp_z |g_i(cp_z)| \int_{\mathbb{S}^{2d-1}} Q_i(p_x) R(p_x) d\Omega, \tag{119}$$

where the quadratic forms are (we omit the explicit dependence on p_z)

$$Q_i(p_x) := \|p_x^i\|^2, \qquad R(p_x) := (\mathbf{B}(p_z) p_x)^* (\partial_x^2 \phi) (\mathbf{B}(p_z) p_x). \tag{120}$$

A direct check shows that Q and R are commuting, block diagonal matrices. Thus, applying Lemma 31 to (119), we obtain

$$\frac{1}{2} \varepsilon^2 \frac{cd}{\sum_{i=1}^d \int_{\mathbb{R}} |g_i(y)| dy} \sum_{i=1}^d \int_{\mathbb{R}} dp_z |g_i(cp_z)| \int_{\mathbb{S}^{2d-1}} Q_i(p_x) R(p_x) d\Omega \tag{121}$$

$$= \frac{1}{2} \varepsilon^2 \frac{cd}{\sum_{i=1}^d \int_{\mathbb{R}} |g_i(y)| dy} \sum_{i=1}^d \int_{\mathbb{R}} dp_z |g_i(cp_z)| \frac{2 \operatorname{Tr}(Q_i R) + \operatorname{Tr}(Q_i) \operatorname{Tr}(R)}{2d(2d+2)}. \tag{122}$$

Observe that $\operatorname{Tr}(Q_i) = 2$, and $\sum_{\ell=1}^d Q_\ell = \mathbb{I}$. Therefore we rewrite (122) as

$$\varepsilon^2 \frac{c}{\sum_{i=1}^d \int_{\mathbb{R}} |g_i(y)| dy} \sum_{i,\ell=1}^d \int_{\mathbb{R}} dp_z |g_i(cp_z)| \frac{(1 + \delta_{i\ell}) \operatorname{Tr}(Q_\ell R)}{4(d+1)}. \tag{123}$$

To compute $\mathrm{Tr}(Q_\ell R)$ denote, for $\ell = 1, \dots, d$

$$D_\ell^2 \phi := \begin{pmatrix} \partial^2_{x_{2\ell-1}} \phi & \partial_{x_{2\ell-1}} \partial_{x_{2\ell}} \phi \\ \partial_{x_{2\ell}} \partial_{x_{2\ell-1}} \phi & \partial^2_{x_{2\ell}} \phi \end{pmatrix}, \qquad B_\ell := B(\alpha_\ell p_z). \tag{124}$$

We obtain thus

$$\mathrm{Tr}(Q_\ell R) = \mathrm{Tr}(B_\ell^*(D_\ell^2\phi)B_\ell) = \mathrm{Tr}(B_\ell B_\ell^*(D_\ell^2\phi)) = \frac{\sin(\frac{\alpha_\ell p_z}{2})^2}{(\alpha_\ell p_z/2)^2}(\partial^2_{x_{2\ell-1}}\phi + \partial^2_{x_{2\ell}}\phi), \tag{125}$$

where we used (107). Thus (123) becomes

$$\frac{\varepsilon^2}{4d} \sum_{i=1}^d \sigma_i(c)(\partial_{x_{2i-1}}\phi + \partial_{x_{2i}}\phi), \tag{126}$$

where the constants $\sigma_i(c)$ are

$$\sigma_i(c) = \frac{dc}{(d+1)\sum_{i=1}^d \int_{\mathbb{R}} |g_i(y)|dy} \sum_{\ell=1}^d (1+\delta_{\ell i}) \int_{\mathbb{R}} |g_\ell(cp_z)| \frac{\sin(\frac{\alpha_i p_z}{2})^2}{(\alpha_i p_z/2)^2}, \tag{127}$$

Taking in account also the remainder term, we obtain

$$\frac{4d}{t^2} \int_{\square_0} \phi(\exp_0(\varepsilon; p_x, p_z))\mu_0^{c\varepsilon}(p_x, p_z) = \sum_{i=1}^d \sigma_i(c)(\partial^2_{x_{2i-1}}\phi + \partial^2_{x_{2i}}\phi)|_0 + 4d\varepsilon O_0(1), \tag{128}$$

where $O_0(1) \leq M_0$ is a remainder term that, when $\varepsilon \leq \varepsilon_0$, is bounded by a constant that depends only on the derivatives of ϕ in a compact metric ball of radius ε_0 centered in 0. A straightforward left-invariance argument shows that, for any other $q \in M$

$$\frac{4d}{\varepsilon^2} \int_{\square_q} [f(\exp_q(\varepsilon; \lambda)) - f(q)]\mu_q^{c\varepsilon}(\lambda) = \sum_{i=1}^d \sigma_i(c)(X^2_{2i-1}\phi + X^2_{2i}\phi)|_q + 4d\varepsilon O_q(1), \tag{129}$$

where $O_q(1) \leq M_q$ is a remainder term bounded by a constant that depends only on the derivatives of ϕ in a compact metric ball of radius ε_0 centered in q. Thus

$$(L_{c,\mathscr{L}}\phi)|_q = \lim_{\varepsilon \to 0} \frac{4d}{\varepsilon^2} \int_{\square_q} [\phi(\exp_q(\varepsilon; \lambda)) - \phi(q)]\mu_q^{c\varepsilon}(\lambda) = \sum_{i=1}^d \sigma_i(c)(X_{2i-1}\phi + X_{2i}\phi)|_q, \tag{130}$$

and the convergence is uniform on compact sets. This completes the proof for $\omega = \mathscr{L}$.

Let, instead, $\omega = e^h \mathscr{L}$ for some $h \in C^\infty(M)$. This leads to an extra factor $e^{h(\exp_q(c\varepsilon;\lambda))}$ in front of $\mu_q^{c\varepsilon}(\lambda)$ (up to re-normalization). After a moment of reflection one realizes that

$$(L^\varepsilon_{\omega,c}\phi)|_q = (L^\varepsilon_{\mathscr{L},c}\tilde{\phi})|_q + \varepsilon O_q(1), \qquad \text{with} \qquad \tilde{\phi} = e^{c(h-h(q))}(\phi - \phi(q)). \tag{131}$$

This observation yields the general statement, after noticing that

$$X_i^2(\tilde{\phi}) = X_i^2(\phi) + 2cX_i(h)X_i(\phi), \qquad \forall i = 1, \dots, 2d, \tag{132}$$

where everything is evaluated at the fixed point q. □

6.3. **Proof of Theorem 20.** We expand the function ϕ along the path $\gamma_u(\varepsilon) = E_{q,\varepsilon}(u)$:

$$\phi(E_{q,\varepsilon}(u)) - \phi(q) = \varepsilon X_u(\phi) + \frac{1}{2}\varepsilon^2 X_u(X_u(\phi)) + O(\varepsilon^3), \tag{133}$$

where everything on the r.h.s. is computed at q (as a convention, in the following when the evaluation point is not explicitly displayed, we understand evaluation at q).

Lemma 32. *For any one-form $\nu \in T_q^*M$ and any vector $v \in T_u\mathbb{S}^{n-1}$*

$$(E^*_{q,\varepsilon}\nu)|_u(v) = \varepsilon\nu(X_v) + \frac{1}{2}\varepsilon^2\nu([X_v, X_u]) + O(\varepsilon^3). \tag{134}$$

Proof of Lemma 32. The differential of the endpoint map (with constant controls) is

$$d_u E_{q,\varepsilon}(v) = e_*^{\varepsilon X_u} \int_0^\varepsilon e_*^{-\tau X_u} X_v d\tau, \qquad v \in \mathbb{R}^n, \tag{135}$$

where $e^{\varepsilon Y}$ is the flow of the field Y (see [1]). By definition of Lie derivative $\mathcal{L}$ we get

$$\frac{d}{d\varepsilon}(E_{q,\varepsilon}^* \nu)|_u(v) = \frac{d}{d\varepsilon}(e^{\varepsilon X_u *}\nu)|_q \left(\int_0^\varepsilon e_*^{-\tau X_u} X_v d\tau\right) \tag{136}$$

$$= (e^{\varepsilon X_u *}\mathcal{L}_{X_u}\nu)|_q \left(\int_0^\varepsilon e_*^{-\tau X_u} X_v d\tau\right) + (e^{\varepsilon X_u *}\nu)|_q \left(e_*^{-\varepsilon X_u} X_v\right). \tag{137}$$

Taking another derivative, and evaluating at $t = 0$, we get

$$\left.\frac{d^2}{d\varepsilon^2}\right|_{\varepsilon=0} (E_{q,\varepsilon}^* \nu)|_u(v) = 2(\mathcal{L}_{X_u}\nu)|_q(X_v) + \nu|_q(\mathcal{L}_{X_u}(X_v)) = \nu([X_v, X_u]), \tag{138}$$

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} (E_{q,\varepsilon}^* \nu)|_u(v) = \nu|_q(X_v).$$

□

Lemma 33. *We have the following Taylor expansion for the measure*

$$\mu_q^\varepsilon(u) = \left(1 + \frac{\varepsilon}{2}\operatorname{div}_{\mathcal{R}}(X_u) + \varepsilon X_u(h) + O(\varepsilon^2)\right)\Omega(u), \tag{139}$$

where Ω is the normalized Euclidean measure on $\mathbb{S}^{n-1}$.

Proof of Lemma 33. Let $\nu_1, \dots, \nu_n$ be the dual frame to $X_1, \dots, X_n$, that is $\nu_i(X_j) = \delta_{ij}$. Since $\omega = e^h \mathcal{R} = e^h \nu_1 \wedge \ldots \wedge \nu_n$, we obtain (ignoring normalization factors)

$$\mu_q^\varepsilon(u) \propto D_q(\varepsilon) e^{h(\gamma_u(\varepsilon))}\Omega(u), \qquad u \in \mathbb{S}^{n-1}, \tag{140}$$

where $D_q(\varepsilon)$ is the determinant of the matrix $(E_{q,\varepsilon}^* \nu_i)(e_j)$, for $i, j = 1, \dots, n$. Using Lemma 32, since $X_{e_j} = X_j$, we obtain

$$(E_{q,\varepsilon}^* \nu_i)(e_j) = \varepsilon \nu_i(X_j) + \frac{\varepsilon^2}{2}\nu_i([X_j, X_u]) + O(\varepsilon^3) \tag{141}$$

where everything is computed at q. Since $\det(\mathbb{I} + \varepsilon M) = 1 + \varepsilon \operatorname{Tr}(M) + O(\varepsilon^2)$ for any matrix M, we get

$$D_q(\varepsilon) = \varepsilon^n \left(1 + \frac{\varepsilon}{2}\sum_{i=1}^n \nu_i([X_i, X_u]) + O(\varepsilon^2)\right) \tag{142}$$

$$= \varepsilon^n \left(1 + \frac{\varepsilon}{2}\sum_{i,j=1}^n u_j c_{ij}^i + O(\varepsilon^2)\right) = \varepsilon^n \left(1 + \frac{\varepsilon}{2}\operatorname{div}_{\mathcal{R}}(X_u) + O(\varepsilon^2)\right). \tag{143}$$

Plugging this in (140), and expanding the function $e^{h(\gamma_u(\varepsilon))}$, we get

$$\mu_q^\varepsilon \propto \varepsilon^n \left(1 + \frac{\varepsilon}{2}\operatorname{div}_{\mathcal{R}}(X_u) + O(\varepsilon^2)\right) e^{h(q)} \left(1 + \varepsilon X_u(h) + O(\varepsilon^2)\right)\Omega \tag{144}$$

$$\propto \varepsilon^n e^{h(q)} \left(1 + \frac{\varepsilon}{2}\operatorname{div}_{\mathcal{R}}(X_u) + t X_u(h) + O(\varepsilon^2)\right)\Omega. \tag{145}$$

Taking in account the normalization (recall that $\int_{\mathbb{S}^{n-1}} X_u = 0$), we obtain the result. □

We are ready to compute the expectation value

$$\int_{\mathbb{S}^{n-1}} [\phi(E_{q,\varepsilon}(u)) - \phi(q)]\mu_q^{c\varepsilon} = \int_{\mathbb{S}^{n-1}} \left[\varepsilon X_u(\phi) + \frac{1}{2}\varepsilon^2 X_u(X_u(\phi)) + O(\varepsilon^3)\right] \times \tag{146}$$

$$\times \left[1 + \frac{c\varepsilon}{2}\operatorname{div}_{\mathcal{R}}(X_u) + c\varepsilon X_u(h) + O(\varepsilon^2)\right]\Omega.$$

Since $\int_{\mathbb{S}^{n-1}} X_u = 0$ and $\int_{\mathbb{S}^{n-1}} Q_{ij} u_i u_j = \mathrm{Tr}(Q)/n$, we get

$$
\begin{aligned}
(L_{\omega,c}\phi)(q) &= \lim_{\varepsilon\to 0^+} \frac{2n}{\varepsilon^2}\left(\frac{c\varepsilon^2}{2n}\operatorname{div}_{\mathcal{R}}(X_i)X_i(\phi) + \frac{c\varepsilon^2}{n}X_i(\phi)X_i(h) + \frac{\varepsilon^2}{2n}X_i^2(\phi) + O(\varepsilon^3)\right)\\
&= \sum_{i=1}^n X_i^2(\phi) + c\operatorname{div}_{\mathcal{R}}(X_i)X_i(\phi) + 2cX_i(\phi)X_i(h).
\end{aligned}
$$

From the last equation we obtain the different forms of the statements using the change of volume formula $\operatorname{div}_\omega(X_i) = \operatorname{div}_{e^h\mathcal{R}}(X_i) = \operatorname{div}_{\mathcal{R}}(X_i) + X_i(h)$. The covergence is uniform on compact set since the domain of integration $\mathbb{S}^{n-1}$ is compact and smoothness of all objects. □

6.4. **Proof of Theorem 25.** The proof follows the same lines of the one of Theorem 20. The expansion of the function ϕ along the path $\gamma_u(\varepsilon) = E_{q,\varepsilon}(u)$ remains unchanged:

$$
\phi(E_{q,\varepsilon}(u)) - \phi(q) = \varepsilon X_u(\phi) + \frac{1}{2}\varepsilon^2 X_u(X_u(\phi)) + O(\varepsilon^3). \tag{147}
$$

where, this time $X_u = \sum_{i=1}^k u_i X_i$. Also Lemma 32 remains unchanged, replacing n with k. The following contact version of Lemma 33 also holds.

Lemma 34. *We have the following Taylor expansion for the measure*

$$
\mu_q^\varepsilon(u) = \left(1 + \frac{\varepsilon}{2}\operatorname{div}_{\mathcal{P}}(X_u) + \varepsilon X_u(h) + O(\varepsilon^2)\right)\Omega(u), \tag{148}
$$

where Ω is the normalized Euclidean measure on $\mathbb{S}^{k-1}$.

Proof of the Lemma. Since $\omega = e^h\mathcal{P} = e^h\nu_0\wedge\nu_1\wedge\dots\nu_k$, we have $i_{X_0}\omega = e^h\nu_1\wedge\dots\wedge\nu_k$. Hence the proof is similar to proof of Lemma 33, with n replaced by k. In fact, up to normalization

$$
\mu_q^\varepsilon(u) \propto (E^*_{\varepsilon,q} i_{\dot\gamma_u(\varepsilon),X_0}\omega) = D_q(\varepsilon)e^{h(\gamma_u(\varepsilon))}\Omega, \qquad u\in\mathbb{S}^{k-1}, \tag{149}
$$

where $D_q(\varepsilon)$ is the determinant of the matrix $(E^*_{q,\varepsilon}\nu_i)(X_j)$, for $i,j = 1,\dots,k$. This is a $k\times k$ matrix. With a computation analogue to the one of the proof of Lemma 33, we obtain $D_q(\varepsilon) = \varepsilon^k(1+\varepsilon\,\mathrm{Tr}(M) + O(\varepsilon^2))$, with

$$
\mathrm{Tr}(M) = \frac{1}{2}\sum_{i=1}^k \nu_i([X_i,X_u]) = \frac{1}{2}\sum_{i,j=1}^k u_j c_{ij}^i = \frac{1}{2}\sum_{j=1}^k u_j\sum_{i=0}^k c_{ij}^i = \frac{1}{2}\operatorname{div}_{\mathcal{P}}(X_u), \tag{150}
$$

where we have been able to complete the sum, including the index 0 since, in the contact case, $c_{0j}^0 = \eta([X_0,X_j]) = -d\eta(X_0,X_j) = 0$ for all $j = 1,\dots,k$. From here, we conclude proof as in the one of Lemma 33. □

The computation of the limit operator is analogue to the one of the proof of Theorem 20, replacing the Riemannian volume $\mathcal{R}$ with the Popp volume $\mathcal{P}$. □

Appendix A. Volume sampling as Girsanov-type change-of-measure

In both the geodesic and flow random walks defined in Sections 3.1 and 5.1, the probability measure used to select the vector $V = \sum \beta_i V_i$ was the uniform probability measure on the unit sphere with respect to the covariance structure of the w_t^i (which gives an inner product on the vector space of such V). In the volume sampling scheme just discussed for the geodesic random walk with respect to an orthonormal frame on a Riemannian manifold, that is, for the isotropic random walk that approximates Brownian motion, the probability measure on the sphere is replaced by a different probability measure, absolutely continuous with respect to the uniform one. In terms of the random walk, the volume-sampled walk is supported on the same set of paths as the original walk, with a different probability measure, absolutely continuous with respect to the original. In the scaling limit as $\varepsilon \to 0$, this change in measure produces a drift in the limiting diffusion, and we recognize this as a Girsanov-type phenomenon. We now take a moment to explore this interpretation in a bit more detail.

The standard finite-dimensional model for Girsanov's theorem, as given at the beginning of [21, Section 3.5], is as follows. With slightly loose notation, we let $N(0, I_n)$ denote the centered (multivariate) normal distribution on $\mathbb{R}^n$ with covariance structure given by the identity matrix (that is, the n Euclidean coordinates are i.i.d. normals with expectation 0 and variance 1). Let Z be a random variable (on some probability space with probability denoted P) with distribution $N(0, I_n)$, and let $\mu \in \mathbb{R}^n$. We have a new probability measure $\tilde{P}$, absolutely continuous with respect to P, given by

$$\tilde{P}(d\lambda) = e^{\langle \mu, Z(\lambda) \rangle - \frac{1}{2}\langle \mu, \mu \rangle} P(d\lambda),$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^n$. Then the random variable $Z - \mu$ has distribution $N(0, I_n)$ under $\tilde{P}$. So adjusting the measure in this way compensates for the translation, which equivalently means that one can create a translation by adjusting the measure. The infinite-dimensional version of this (for Brownian motion on Euclidean space) is Girsanov's theorem.

Next, we rephrase this. Another way of determining $\tilde{P}$ is to say that it comes from adjusting the "likelihood ratios" for P by

$$\frac{\tilde{P}(d\lambda_2)}{\tilde{P}(d\lambda_1)} = e^{\langle \mu, Z(\lambda_2) \rangle - \langle \mu, Z(\lambda_1) \rangle} \frac{P(d\lambda_2)}{P(d\lambda_1)}. \tag{151}$$

The accounts for the $e^{\langle \mu, Z(\lambda) \rangle}$ factor in the Radon-Nikodym derivative above, which is the important part; the $e^{-\frac{1}{2}\langle \mu, \mu \rangle}$ is just the normalizing constant that turns $\tilde{P}$ into a probability measure.

For the isotropic random walk, we have that P is μ_q^0, the uniform probability measure on the sphere of radius $\sqrt{n}$ in T_qM, with respect to the Riemannian inner product. (Here we choose to normalize the sphere to include the $\sqrt{n}$ factor in order to make the connection to Girsanov's theorem clearer.) Of course, μ_q^0 is not a multivariate normal on $T_qM \simeq \mathbb{R}^n$. However, μ_q^0 has expectation 0 and covariance matrix I_n, so that μ_q^0 has the same first two moments as $N(0, I_n)$. In light of Donsker's invariance principle, it is not surprising that "getting the first two moments right" is enough. Now $\mu_q^{c\varepsilon}$ is absolutely continuous with respect to μ_q^0, and, as we will see in the proof of Theorem 8, the relationship is given by

$$\begin{aligned}
\frac{\mu_q^{c\varepsilon}(d\lambda_2)}{\mu_q^{c\varepsilon}(d\lambda_1)} &= \frac{\frac{1}{\mathrm{vol}(\mathbb{S}^{n-1})}\left(1 + c\varepsilon \langle \mathrm{grad}(h), \lambda_2 \rangle + O\left(\varepsilon^2\right)\right)}{\frac{1}{\mathrm{vol}(\mathbb{S}^{n-1})}\left(1 + c\varepsilon \langle \mathrm{grad}(h), \lambda_1 \rangle + O\left(\varepsilon^2\right)\right)} \cdot \frac{\mu_q^0(d\lambda_2)}{\mu_q^0(d\lambda_1)} \\
&= e^{c\varepsilon(\langle \mathrm{grad}(h), \lambda_2 \rangle - \langle \mathrm{grad}(h), \lambda_1 \rangle) + O\left(\varepsilon^2\right)} \cdot \frac{\mu_q^0(d\lambda_2)}{\mu_q^0(d\lambda_1)}.
\end{aligned}$$

Note that, as we have developed things, the random variable that has distribution μ_q^0, which is analogous to Z above, is implicitly just the identity on the sphere. (Also, $\mu_q^{c\varepsilon}$ is a probability measure by construction, so there's no need for a normalizing factor, partially explaining our focus on the likelihood ratio.)

Comparing this to (151), we see that the role of μ is played by the quantity $c\varepsilon\, \mathrm{grad}(h) + O(\varepsilon^2)$. To take into account the parabolic scaling limit (taking in account the analysts normalization), note that this non-centered measure on the sphere of radius $\sqrt{n}$ (namely $\mu_q^{c\varepsilon}$) is mapped onto geodesics of length ε, and that this step takes place in time $\delta = 2n/\varepsilon^2$, so that the difference quotient (expected spatial displacement over elapsed time) is $2c\, \mathrm{grad}(h) + O(\varepsilon)$. Thus, in the limit as $\varepsilon \to 0$, we expect an infinitesimal translation given by the tangent vector $2c\, \mathrm{grad}(h)$, which is exactly what we see in Theorem 8 (appearing as a first-order differential operator). Namely, the random walk corresponding to μ_q^0 has infinitesimal generator $\Delta_{\mathcal{R}}$ in the limit, while the random walk corresponding to $\mu_q^{c\varepsilon}$ has infinitesimal generator $\Delta_{\mathcal{R}} + 2c\,\mathrm{grad}(h)$ in the limit. Thus, this volume sampling gives a natural random walk version of the Girsanov change of measure.

Acknowledgments. This research has been partially supported by the European Research Council, ERC StG 2009 "GeCoMethods", contract n. 239748, by the iCODE institute (research project of the Idex Paris-Saclay), by the SMAI project "BOUM", the Grant ANR-15-CE40-0018 of the ANR, the National Security Agency under Grant Number H98230-15-1-0171. This research benefited from the support of the "FMJH Program Gaspard Monge in optimization and operation research" and from the support to this program from EDF.

References

[1] A. Agrachev, D. Barilari, and U. Boscain. Introduction to Riemannian and sub-Riemannian geometry (Lecture Notes), http://people.sissa.it/agrachev/agrachev_files/notes.html. 2012.

[2] A. Agrachev, D. Barilari, and U. Boscain. On the Hausdorff volume in sub-Riemannian geometry. *Calc. Var. Partial Differential Equations*, 43(3-4):355–388, 2012.

[3] 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.

[4] A. A. Agrachëv. Any sub-Riemannian metric has points of smoothness. *Dokl. Akad. Nauk*, 424(3):295–298, 2009.

[5] A. A. Agrachev. Some open problems. In *Geometric control theory and sub-Riemannian geometry*, volume 5 of *Springer INdAM Ser.*, pages 1–13. Springer, Cham, 2014.

[6] M. Asaad and M. Gordina. Hypoelliptic heat kernel on nilpotent Lie groups. *ArXiv e-prints*, May 2015.

[7] D. Barilari and L. Rizzi. A formula for Popp's volume in sub-Riemannian geometry. *Anal. Geom. Metr. Spaces*, 1:42–57, 2013.

[8] A. Bellaïche. The tangent space in sub-Riemannian geometry. In *Sub-Riemannian geometry*, volume 144 of *Progr. Math.*, pages 1–78. Birkhäuser, Basel, 1996.

[9] J.-M. Bismut. *Large deviations and the Malliavin calculus*, volume 45 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1984.

[10] U. Boscain, R. Neel, and L. Rizzi. Intrinsic random walks and sub-Laplacians in sub-Riemannian geometry. *ArXiv e-prints*, Mar. 2015.

[11] H. Doss. Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 13(2):99–125, 1977.

[12] M. Émery. *Stochastic calculus in manifolds*. Universitext. Springer-Verlag, Berlin, 1989. With an appendix by P.-A. Meyer.

[13] G. B. Folland. How to integrate a polynomial over a sphere. *Amer. Math. Monthly*, 108(5):446–448, 2001.

[14] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian geometry*. Universitext. Springer-Verlag, Berlin, third edition, 2004.

[15] M. Gordina and T. Laetsch. Sub-Laplacians on sub-Riemannian manifolds. *ArXiv e-prints*, Nov. 2014.

[16] M. Gordina and T. Laetsch. Weak Convergence to Brownian Motion on Sub-Riemannian Manifolds. *Transactions of the AMS (to appear)*, Mar. 2014.

[17] E. Grong and A. Thalmaier. Curvature-dimension inequalities on sub-riemannian manifolds obtained from riemannian foliations, part 1. *Mathematische Zeitschrift (to appear)*, 2014.

[18] E. Grong and A. Thalmaier. Curvature-dimension inequalities on sub-riemannian manifolds obtained from riemannian foliations, part 2. *Mathematische Zeitschrift (to appear)*, 2014.

[19] E. P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[20] F. Jean. *Control of Nonholonomic Systems: from Sub-Riemannian Geometry to Motion Planning*. Springer International Publishing, SpringerBriefs in Mathematics, 2014.

[21] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[22] E. Le Donne, R. Montgomery, A. Ottazzi, P. Pansu, and D. Vittone. Sard Property for the endpoint map on some Carnot groups. *Annales de l'Institut Henri Poincare, Analyse non lineaire (to appear)*, Mar. 2015.

[23] G. Lebeau and L. Michel. Semi-classical analysis of a random walk on a manifold. *Ann. Probab.*, 38(1):277–315, 2010.

[24] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

[25] M. A. Pinsky. Isotropic transport process on a Riemannian manifold. *Trans. Amer. Math. Soc.*, 218:353–360, 1976.

[26] L. Rifford. *Sub-Riemannian geometry and optimal transport.* Springer, 2014.

[27] L. Rifford and E. Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.

[28] L. Rizzi. Measure contraction properties of Carnot groups. *ArXiv e-prints*, Oct. 2015.

[29] R. S. Strichartz. Sub-Riemannian geometry. *J. Differential Geom.*, 24(2):221–263, 1986.

[30] R. S. Strichartz. Corrections to: "Sub-Riemannian geometry" [J. Differential Geom. **24** (1986), no. 2, 221–263; MR0862049 (88b:53055)]. *J. Differential Geom.*, 30(2):595–596, 1989.

[31] D. W. Stroock. *An introduction to the analysis of paths on a Riemannian manifold*, volume 74 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.

[32] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

[33] H. J. Sussmann. On the gap between deterministic and stochastic ordinary differential equations. *Ann. Probability*, 6(1):19–41, 1978.