Geometry of sub-Riemannian diffusion processes

Karen Habermann

Newnham College

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

August 2017

Abstract

Sub-Riemannian geometry is the natural setting for studying dynamical systems, as noise often has a lower dimension than the dynamics it enters. This makes sub-Riemannian geometry an important field of study. In this thesis, we analysis some of the aspects of sub-Riemannian diffusion processes on manifolds.

We first focus on studying the small-time asymptotics of sub-Riemannian diffusion bridges. After giving an overview of recent work by Bailleul, Mesnager and Norris on small-time fluctuations for the bridge of a sub-Riemannian diffusion, we show, by providing a specific example, that, unlike in the Riemannian case, small-time fluctuations for sub-Riemannian diffusion bridges can exhibit exotic behaviours, that is, qualitatively different behaviours compared to Brownian bridges.

We further extend the analysis by Bailleul, Mesnager and Norris of small-time fluctuations for sub-Riemannian diffusion bridges, which assumes the initial and final positions to lie outside the sub-Riemannian cut locus, to the diagonal and describe the asymptotics of sub-Riemannian diffusion loops. We show that, in a suitable chart and after a suitable rescaling, the small-time diffusion loop measures have a non-degenerate limit, which we identify explicitly in terms of a certain local limit operator. Our analysis also allows us to determine the loop asymptotics under the scaling used to obtain a small-time Gaussian limit for the sub-Riemannian diffusion bridge measures by Bailleul, Mesnager and Norris. In general, these asymptotics are now degenerate and need no longer be Gaussian.

We close by reporting on work in progress which aims to understand the behaviour of Brownian motion conditioned to have vanishing Nth truncated signature in the limit as N tends to infinity. So far, it has led to an analytic proof of the stand-alone result that a Brownian bridge in $\mathbb{R}^d$ from 0 to 0 in time 1 is more likely to stay inside a box centred at the origin than any other Brownian bridge in time 1.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

Chapter 1 consists of a literature review. Chapters 2, 3 and 4 contain original research. Chapter 2 is based on unpublished work, Chapter 3 is published as [Hab17], and the idea for Chapter 4 arose from dialogue with Terry Lyons. The proof of Lemma 4.2.8 is due to Katarzyna Wyczesany.

Acknowledgements

First and foremost, I would like to thank James Norris for his support and guidance over the years, for his words of encouragement, especially during the last year, and for having given me the chance to join the Cambridge Centre for Analysis in the first place.

The company of fellow CCA students made the journey of the last four years less painful and much more fun. Special thanks are due to Henry Jackson, who always made time to distract me when needed, to James Mathews, for organising board game nights and a great hiking trip, to Kim Moore, for bringing the legendary cat calendars into the CMS, to Rob Hocking, who provided a lot of entertainment, and to Tom Begley, for many hours spent working (and not working) in the Core. I am also grateful to Ben Jennings and Ellen Powell for a very enjoyable (and totally non-distracting) office environment, and to Adam Kashlak for treating us with maple syrup biscuits.

I would like to thank Newnham College for having provided me with a home for the last eight years and for having financially supported conference attendances within the UK and around the world.

A huge thank you is due to the porters of past and present at Newnham College, most of all to Dick Keep, Mervyn Russell and Paul McElroy, who looked after me for so many years and who always made me smile. I could not have asked for a better bastion of calm in this manic place that Cambridge can be.

Last but not least to my parents, thank you ever so much for your unconditional support and your unbelievable faith in me.

Contents

Chapter 1

Preliminaries

Sub-Riemannian geometry is the study of geometric structures that arise on spaces where motion is only possible along a given set of trajectories. The subject, which goes all the way back to Carathéodory [Car09] and Cartan [Car31], has found motivation in various areas of mathematics and physics, and has been studied from a number of different viewpoints such as control theory, see Brockett [Bro82, Bro84] and Hermann [Her62, Her73], symplectic and contact geometry, e.g. Chern and Hamilton [CH85], Cauchy-Riemann geometry, e.g. Webster [Web78], or classical mechanics. As these investigations have been carried out more or less independently each area provided its own technical terminology, which has led to the same concepts being known under different names. Even sub-Riemannian geometry is sometimes referred to as non-holonomic Riemannian geometry, Carnot geometry, or singular Riemannian geometry. A notion underlying all different viewpoints is the concept of a bracket generating distribution, also known as completely non-integrable distribution.

Definition (Bracket generating distribution)**.** *Let M be a connected smooth manifold and let H be a distribution, i.e. a subbundle of the tangent bundle TM. We call H a bracket generating distribution if, for all $x \in M$, the sections of H near x together with their commutator brackets of all orders span the tangent space T_xM.*

A connected manifold equipped with a distribution can be considered as a space where motion is only possible along directions given by tangent vectors in the distribution. If, additionally, the distribution is bracket generating then, by the Chow-Rashevskii theorem, any two points on the manifold can be connected by an admissible path.
Let us now compare the notion of a distribution being bracket generating to the similar concept of a collection of vector fields satisfying the Hörmander condition.

Definition (Hörmander condition)**.** *Let M be a connected smooth manifold. A collection of smooth vector fields $X_1, \dots, X_m$ on M is said to satisfy the Hörmander condition if, for all $x \in M$,*

$$\operatorname{span} \bigcup_{k\in\mathbb{N}} \left\{ [X_{i_1}, [X_{i_2}, \dots, [X_{i_{k-1}}, X_{i_k}] \dots]](x) \colon 1 \le i_1, \dots, i_k \le m \right\} = T_xM \ .$$

Both a bracket generating distribution and a collection of smooth vector fields satisfying the Hörmander condition give us a structure in the tangent bundle of a manifold from which we can recover the entire tangent bundle by taking commutator brackets. The main difference is that the structure induced by the collection of vector fields might be rank-varying.

By the Hörmander hypoellipticity theorem, if $X_0, X_1, \ldots, X_m$ are smooth vector fields on a manifold with $X_1, \ldots, X_m$ satisfying the Hörmander condition then the operator

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{m} X_i^2 + X_0$$

is hypoelliptic. Thus, for a diffusion process on a connected manifold whose generator is an operator of the above form, heat flows between any two points on the manifold. This is another manifestation of the Chow-Rashevskii theorem in sub-Riemannian geometry. It shows that the analysis and study of hypoelliptic operators and their associated diffusion processes is yet another facet of sub-Riemannian geometry.

The purpose of this chapter is to provide an introduction to sub-Riemannian geometry, cf. Section 1.1, and to the ideas from Malliavin calculus which are used to study hypoelliptic diffusion processes, cf. Section 1.2. This sets up the relevant background and framework for the research work we report on in Chapters 2, 3 and 4.

1.1 Sub-Riemannian geometry

We give a brief survey of sub-Riemannian geometry, where we shed light on a very limited number of its features. We leave aside a lot of interesting aspects and phenomena, such as the shape of spheres in a sub-Riemannian geometry, the Pansu derivative, or the existence of singular geodesics, and only mention one of many challenging open problems. For more elaborate introductions, consult Agrachev, Barilari and Boscain [ABB16], the collection of lecture notes [BBS16a, BBS16b], Calin and Chang [CC09], Montgomery [Mon02], and Strichartz [Str86, Str89]. Besides, see Hamenstädt [Ham90] for a different approach to the theory of geodesics, Montgomery [Mon95] for a survey of singular curves, as well as Pansu [Pan89] for the Pansu-Rademacher differentiation theorem. For open problems, see Agrachev [Agr14] and Montgomery [Mon02, Chapter 10].

Let M be a connected smooth manifold and recall that a distribution on M is a subbundle of the tangent bundle TM. A space where motion is restricted along a set of admissible paths is understood as a sub-Riemannian manifold.

Definition 1.1.1. *A sub-Riemannian structure on the manifold M consists of a bracket generating distribution H and a fiber inner product $\langle \cdot, \cdot \rangle$ on H. The triple $(M, H, \langle \cdot, \cdot \rangle)$ is called a sub-Riemannian manifold.*

Sub-Riemannian manifolds naturally appear in the study of constrained physical systems. For instance, the motion of robot arms, the orbital dynamics of satellites, the Heisenberg group which plays an important role in quantum mechanics, and the fall of a cat all have an underlying sub-Riemannian structure.

In subsequent chapters, we are mainly concerned with the situation where the bracket generating distribution of a sub-Riemannian structure on M is given by a collection of smooth vector fields $X_1, \ldots, X_m$ on M satisfying the Hörmander condition. In this case, we associate a fiber inner product as follows, cf. Bellaïche [Bel96, Definition 1.1]. For all $x \in M$, we endow the subspace

$$H_x = \operatorname{span}\{X_1(x), \ldots, X_m(x)\} \subset T_x M$$

with the inner product $\langle \cdot, \cdot \rangle_x$ obtained by polarising the quadratic form g_x on H_x satisfying

$$g_x(v) = \inf \left\{ \sum_{i=1}^{m} \left(u^i\right)^2 : u^1, \ldots, u^m \in \mathbb{R} \text{ with } \sum_{i=1}^{m} u^i X_i(x) = v \right\} . \tag{1.1.1}$$

To see that (1.1.1) indeed gives a positive definite quadratic form on H_x, consider the linear map $\sigma_x \colon \mathbb{R}^m \to H_x$ defined by

$$\sigma_x \left(u^1, \ldots, u^m\right) = \sum_{i=1}^{m} u^i X_i(x) .$$

The restriction of the map σ_x to the orthogonal complement $(\ker \sigma_x)^\perp$ of $\ker \sigma_x \subset \mathbb{R}^m$ with respect to the Euclidean inner product is a linear isomorphism. Let $\rho_x \colon H_x \to (\ker \sigma_x)^\perp$ be the inverse of σ_x restricted to $(\ker \sigma_x)^\perp$ and let $\|\cdot\|_2$ denote the Euclidean norm. Then for $v \in H_x$ and any $u = (u^1, \ldots, u^m) \in \mathbb{R}^m$ with $\sigma_x(u) = v$, we have

$$\|\rho_x(v)\|_2^2 \le \|\rho_x(v)\|_2^2 + \|u - \rho_x(v)\|_2^2 = \|\rho_x(v) + u - \rho_x(v)\|_2^2 = \|u\|_2^2$$

due to $u - \rho_x(v) \in \ker \sigma_x$ and $\rho_x(v) \in (\ker \sigma_x)^\perp$. It follows that

$$g_x(v) = \|\rho_x(v)\|_2^2$$

and, as ρ_x is a linear isomorphism, g_x is a positive definite quadratic form on H_x. Moreover, if $X_1, \ldots, X_m$ are linearly independent at $x \in M$ then $\ker \sigma_x = \{0\}$, which implies that

$$\langle X_i, X_i \rangle_x = g_x\left(X_i(x)\right) = 1 \quad \text{for all} \quad i \in \{1, \ldots, m\} .$$

In particular, if the vector fields $X_1, \ldots, X_m$ are linearly independent at every point then $(X_1, \ldots, X_m)$ is a global orthonormal frame with respect to the fiber inner product $\langle \cdot, \cdot \rangle$ of the distribution spanned by $X_1, \ldots, X_m$.

We note that in general a collection of smooth vector fields $X_1, \ldots, X_m$ on M satisfying the Hörmander condition induces a structure in the tangent bundle TM which might be rank-varying. However, if we endow the resulting structure with the fiber inner product obtained by polarising the positive definite quadratic form (1.1.1) the results discussed below still apply. For simplicity of presentation, we therefore choose to not go into the more general setting of rank-varying sub-Riemannian structures. For a complete presentation, see Agrachev, Barilari and Boscain [ABB16, Section 3].

1.1.1 The sub-Riemannian distance

Let $(M, H, \langle\cdot,\cdot\rangle)$ be a sub-Riemannian manifold. We call H the horizontal distribution of the sub-Riemannian manifold. An absolutely continuous path $\omega\colon [0,1] \to M$ is said to be horizontal, or admissible, if $\dot\omega_t \in H(\omega_t)$ for almost all $t \in [0,1]$. The first important result in sub-Riemannian geometry is about the connectability of two points by a horizontal path. It was independently proven by Chow [Cho39] and Rashevskii [Ras38].

Theorem (Chow-Rashevskii theorem)**.** *Any two points on a sub-Riemannian manifold can be connected by a horizontal path in the manifold.* □

The length $l(\omega)$ of a horizontal path $\omega\colon [0,1] \to M$ is defined by

$$l(\omega) = \int_0^1 \sqrt{\langle\dot\omega_t, \dot\omega_t\rangle_{\omega_t}}\, \mathrm{d}t \; . \tag{1.1.2}$$

For $x, y \in M$, let

$$H^{x,y} = \{\omega \in C([0,1], M) : \omega \text{ horizontal path with } \omega_0 = x \text{ and } \omega_1 = y\}$$

be the subset of $C([0,1], M)$ consisting of the horizontal paths connecting x to y, and set

$$d(x,y) = \inf_{\omega \in H^{x,y}} l(\omega) \; . \tag{1.1.3}$$

By the Chow-Rashevskii theorem, the set $H^{x,y}$ is non-empty and (1.1.3) defines a distance function on M which is compatible with the topology of M. This distance function induced by the sub-Riemannian structure on M is called the sub-Riemannian distance, or also the Carnot-Carathéodory distance. Alternatively, one defines the sub-Riemannian distance by considering the energy $I(\omega)$ of a horizontal path ω given as

$$I(\omega) = \int_0^1 \langle\dot\omega_t, \dot\omega_t\rangle_{\omega_t}\, \mathrm{d}t \; ,$$

and then setting

$$d(x,y) = \inf_{\omega \in H^{x,y}} \sqrt{I(\omega)} \; .$$

As a result of the Cauchy-Schwarz inequality, both approaches give rise to the same distance function d on M. If (M, d) is a complete metric space then the corresponding sub-Riemannian manifold is said to be complete. For instance, a sub-Riemannian manifold $(M, H, \langle\cdot,\cdot\rangle)$ where the fiber inner product $\langle\cdot,\cdot\rangle$ arises as the restriction of a complete Riemannian metric on M to the horizontal distribution H is complete.

In the next section, we see that it is possible to define the sub-Riemannian distance using ideas from control theory.

1.1.2 Connections to control theory

In control theory, one is interested in studying smooth control systems on $\mathbb{R}^d$ given, for a smooth function $f\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$, as

$$\dot{q}_t = f(q_t, u_t) \quad \text{for} \quad t \in [0,1] \ , \tag{1.1.4}$$

where $u\colon [0,1] \to \mathbb{R}^m$ is called the control. A solution $q\colon [0,1] \to \mathbb{R}^d$ of the ordinary differential equation (1.1.4) is called a controlled path. The first question which arises in control theory is the question of controllability, i.e. if for any two points $x, y \in \mathbb{R}^d$ there exists a control u such that the associated controlled path $(q_t)_{t\in[0,1]}$ starting from $q_0 = x$ satisfies $q_1 = y$. Note the similarity with the question in sub-Riemannian geometry about the connectability of two points by a horizontal path.

To analyse the controllability of a smooth control system, it is common to consider the first-order approximation of the system. The reason for this is if the linearised system is controllable then so is the original control system near the point of linearisation. Thus, by also extending our consideration to the manifold setting, we become interested in studying linear control systems of the form

$$\dot{q}_t = \sum_{i=1}^{m} u_t^i X_i(q_t) \quad \text{for} \quad t \in [0,1] \ , \tag{1.1.5}$$

where $X_1, \dots, X_m$ are smooth vector fields on a connected smooth manifold M and the path $u = (u^1, \dots, u^m)\colon [0,1] \to \mathbb{R}^m$ is assumed to be measurable. Let us suppose that the vector fields $X_1, \dots, X_m$ satisfy the Hörmander condition. Then $X_1, \dots, X_m$ together with the fiber inner product $\langle\cdot,\cdot\rangle$ on $\operatorname{span}\{X_1, \dots, X_m\}$ defined by polarising the quadratic form (1.1.1) induce a, potentially rank-varying, sub-Riemannian structure on M. From the Chow-Rashevskii theorem, it follows that for any two points $x, y \in M$ there exists a horizontal path $\omega\colon [0,1] \to M$ with $\omega_0 = x$ and $\omega_1 = y$. Since ω is horizontal there exists a measurable path $u\colon [0,1] \to \mathbb{R}^m$ such that

$$\dot{\omega}_t = \sum_{i=1}^{m} u_t^i X_i(\omega_t) \quad \text{for almost all} \quad t \in [0,1] \ .$$

In the language of control theory, this says that $(u_t)_{t\in[0,1]}$ is a control whose associated controlled path $(\omega_t)_{t\in[0,1]}$ starting from x ends at y. Hence, as yet another manifestation of the Chow-Rashevskii theorem, if the vector fields $X_1,\dots,X_m$ satisfy the Hörmander condition then the linear control system (1.1.5) is controllable.
To see how to obtain the sub-Riemannian distance from the viewpoint of control theory, we observe that the expression (1.1.1) of the quadratic form giving the fiber inner product associated with the vector fields $X_1,\dots,X_m$ and the definition (1.1.2) of the length of a horizontal path imply that

$$l(\omega) = \inf\left\{\int_0^1 \sqrt{\sum_{i=1}^m \left(u_t^i\right)^2}\,\mathrm{d}t\colon (u_t)_{t\in[0,1]} \text{ measurable with } \dot{\omega}_t = \sum_{i=1}^m u_t^i X_i(\omega_t)\right\} .$$

Therefore, the problem of finding the sub-Riemannian distance between $x, y \in M$ can be formulated as the optimal control problem

$$\begin{aligned} &\text{minimise} \quad \int_0^1 \sqrt{\sum_{i=1}^m \left(u_t^i\right)^2}\,\mathrm{d}t \\ &\text{subject to} \quad q_0 = x,\ q_1 = y \text{ for } q\colon [0,1]\to M \text{ satisfying } \dot{q}_t = \sum_{i=1}^m u_t^i X_i(q_t)\ . \end{aligned} \tag{1.1.6}$$

We are additionally interested in not only the sub-Riemannian distance between points but in the horizontal paths which achieve this minimal length.

1.1.3 Geodesic curves and the sub-Riemannian cut locus

As in Riemannian geometry, we can use the energy functional to define the notion of a geodesic curve in a sub-Riemannian manifold.

Definition 1.1.2. *A geodesic in a sub-Riemannian manifold is a horizontal path which locally minimises the energy functional.*

Using the Cauchy-Schwarz inequality, we can show that, as in Riemannian geometry, the geodesics in a sub-Riemannian manifold are those horizontal paths $(\omega_t)_{t\in[0,1]}$ which locally minimise the length functional and are parametrised to have constant speed $\sqrt{\langle\dot{\omega}_t,\dot{\omega}_t\rangle_{\omega_t}}$. However, unlike the Riemannian case, it is an open question in sub-Riemannian geometry if geodesics are always smooth, cf. Montgomery [Mon02, Problem 10.1].
There are further complications which make the study of geodesics in sub-Riemannian geometry harder than in Riemannian geometry. Recall that in the Riemannian setting, any maximal geodesic is uniquely determined by its initial point and its initial velocity. This cannot be the case in a general sub-Riemannian geometry. By the Chow-Rashevskii theorem, the geodesics starting from a point x in a sub-Riemannian manifold cover a full neighbourhood of x, whereas the dimension of their admissible initial velocities equals the

dimension of the rank of the sub-Riemannian structure at x, which in general is strictly smaller than the dimension of the manifold. It turns out that the right approach to take in sub-Riemannian geometry is to parametrise a geodesic by its initial point x and an initial covector $\lambda_0 \in T_x^*M$.
In the following, let us suppose that the sub-Riemannian structure on a connected smooth manifold M is induced by smooth vector fields $X_1, \ldots, X_m$ on M satisfying the Hörmander condition. The theorem below, a weak version of the Pontryagin maximum principle, provides a necessary condition satisfied by geodesics in a sub-Riemannian manifold. For a proof, see Agrachev, Barilari and Boscain [ABB16, Section 3].

Theorem 1.1.3. *Suppose that $q\colon [0,1] \to M$ is a solution with constant speed of the optimal control problem (1.1.6), and denote the corresponding control by u. Let $(\phi_{0,t})$ be the flow of the nonautonomous vector field $\sum_{i=1}^m u_t^i X_i$. Then there exists $\lambda_0 \in T_x^*M$ such that the path $(\lambda_t)_{t\in[0,\varepsilon(\lambda_0)]}$ in the cotangent bundle T^*M defined by*

$$\lambda_t = \left(\phi_{0,t}^{-1}\right)^* \lambda_0 \in T_{q_t}^*M$$

satisfies

(N) $\lambda_t(X_i) = u_t^i$ for all $i \in \{1, \ldots, m\}$, or

(A) $\lambda_t(X_i) = 0$ for all $i \in \{1, \ldots, m\}$.

In case (A), we have $\lambda_0 \neq 0$. □

The path λ in the cotangent bundle is called a normal extremal if condition (N) is satisfied, and an abnormal extremal if condition (A) is satisfied. We note that, unless $(q_t)_{t\in[0,1]}$ is a constant path, an associated path λ cannot satisfy both (N) and (A). However, it is possible that for a given solution $(q_t)_{t\in[0,1]}$ there exist two different covectors $\lambda_0^1 \in T_xM$ and $\lambda_0^2 \in T_xM$ such that $\lambda_t^1 = \left(\phi_{0,t}^{-1}\right)^* \lambda_0^1$ defines a normal extremal while $\lambda_t^2 = \left(\phi_{0,t}^{-1}\right)^* \lambda_0^2$ gives an abnormal extremal. Whereas it is known that normal extremals are smooth, it is still an open question if abnormal extremals are always smooth. Moreover, if $(q_t)_{t\in[0,1]}$ admits a normal extremal $(\lambda_t)_{t\in[0,1]}$ then $(q_t)_{t\in[0,1]}$ is shown to be a geodesic, which need not be the case if it admits an abnormal extremal defined up to time 1.
The notion of normal extremals is used in the definition of the sub-Riemannian cut locus. Let $(\psi_t(\lambda)\colon \lambda \in T^*M, t \in (\zeta^-(\lambda), \zeta^+(\lambda)))$ be the maximal flow of the Hamiltonian vector field V on T^*M associated with the Hamiltonian $\mathcal{H}\colon T^*M \to \mathbb{R}$ given by

$$\mathcal{H}(\lambda) = \frac{1}{2}\sum_{i=1}^m \lambda(X_i)^2 \quad \text{for} \quad \lambda \in T^*M \ ,$$

i.e. V is the smooth vector field on T^*M satisfying $\beta(V, \cdot) = \mathrm{d}\mathcal{H}$ with β the canonical symplectic two-form on T^*M. We note that normal extremals are integral curves of the vector field V. Write $\pi\colon T^*M \to M$ for the projection of the bundle. A path $\gamma \in H^{x,y}$ is said to be strongly minimal if there exist $\delta > 0$ and a relatively compact open set $U \subset M$ such that

$$I(\gamma) \le I(\omega) \text{ for all } \omega \in H^{x,y} \quad \text{and} \quad I(\gamma) + \delta \le I(\omega) \text{ for all } \omega \in H^{x,y} \text{ which leave } U \, .$$

Extending Bismut [Bis84] and Ben Arous [BA88], as in [BMN15], to manifolds which are not assumed to be complete, we obtain the following definition of the sub-Riemannian cut locus.

Definition 1.1.4. *The pair $(x, y) \in M \times M$ is said to lie outside the sub-Riemannian cut locus if the following three conditions are satisfied.*

(i) *There is a unique strongly minimal path $\gamma \in H^{x,y}$.*

(ii) *There exists a normal extremal $(\lambda_t)_{t\in[0,1]}$ such that $\gamma_t = \pi\lambda_t$ for all $t \in [0, 1]$.*

(iii) *The linear map $J_1 \colon T_x^*M \to T_yM$ defined by*

$$J_1\xi_0 = \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \pi\psi_1\left(\lambda_0 + \varepsilon\xi_0\right)$$

is invertible.

The original definition of the sub-Riemannian cut locus by Bismut [Bis84] assumes the sub-Riemannian manifold to be complete and does not require the unique minimal path $\gamma \in H^{x,y}$ to be strongly minimal. By the Hopf-Rinow theorem any minimal path on a complete sub-Riemannian manifold is indeed strongly minimal. Therefore, on a complete sub-Riemannian manifold, Definition 1.1.4 reduces to the definition of the sub-Riemannian cut locus by Bismut [Bis84].

The sub-Riemannian cut locus is less well understood than the Riemannian one, and shows some peculiar behaviours which do not occur in Riemannian geometry. For instance, if x is a point in a sub-Riemannian manifold where the rank of the horizontal distribution is less than the dimension of the manifold then any neighbourhood of x contains a point y such that the pair (x, y) lies inside the sub-Riemannian cut locus. What is known is that the sub-Riemannian cut locus is a closed and symmetric subset of $M \times M$, and that the squared sub-Riemannian distance function is a smooth function on the complement of the cut locus, see Bismut [Bis84, Theorem 1.26 and Remark 11].

1.1.4 Sub-Riemannian Laplacians

The sub-Riemannian Laplacian, or short sub-Laplacian, on a sub-Riemannian manifold is defined as the divergence of the horizontal gradient. It is the natural generalisation of the Laplace-Beltrami operator in Riemannian geometry.

Let $(M, H, \langle\cdot,\cdot\rangle)$ be a sub-Riemannian manifold and let ν be a positive smooth measure on the manifold M. We define the horizontal gradient of a smooth function f on M as the unique section $\nabla_H f$ of the distribution H such that, for all sections X of H,

$$\langle \nabla_H f, X\rangle = X(f) \ .$$

Note that this depends on the horizontal distribution H and the fiber inner product $\langle\cdot,\cdot\rangle$ only. In a local orthonormal frame $(Y_1, \ldots, Y_k)$ of H, the horizontal gradient $\nabla_H f$ can be written as

$$\nabla_H f = \sum_{i=1}^{k} Y_i(f) Y_i \ ,$$

because, for all $j \in \{1, \ldots, k\}$, we have

$$\left\langle \sum_{i=1}^{k} Y_i(f) Y_i, Y_j \right\rangle = \sum_{i=1}^{k} Y_i(f) \langle Y_i, Y_j\rangle = Y_j(f) \ .$$

Furthermore, the divergence of a smooth vector field X on M with respect to the positive smooth measure ν is defined to be the smooth function $\operatorname{div} X$ on M which satisfies, for all smooth functions f on M of compact support, that

$$\int_M f \operatorname{div} X \,\mathrm{d}\nu = -\int_M X(f) \,\mathrm{d}\nu \ .$$

This depends on the sub-Riemannian structure $(H, \langle\cdot,\cdot\rangle)$ and on our choice of measure ν. The sub-Riemannian Laplacian Δ_H on the sub-Riemannian manifold $(M, H, \langle\cdot,\cdot\rangle)$ acting on smooth functions f on M is then given by

$$\Delta_H f = \operatorname{div}(\nabla_H f) \ .$$

Due to the dependence of the divergence on the choice of measure ν, the sub-Riemannian Laplacian also depends on this additional structure on M. In a local orthonormal frame $(Y_1, \ldots, Y_k)$ of the horizontal distribution H, we obtain

$$\Delta_H = \sum_{i=1}^{k} \left(Y_i^2 + (\operatorname{div} Y_i) \, Y_i \right) \ .$$

On so-called equiregular sub-Riemannian manifolds, it is possible to define an intrinsic positive smooth measure, the Popp measure, and to define an intrinsic sub-Riemannian Laplacian by taking the divergence with respect to this intrinsic positive smooth measure. See Montgomery [Mon02, Section 10.6] for a construction of the Popp measure.

1.1.5 Examples of sub-Riemannian manifolds

To illustrate the notions introduced above, we provide a few examples of sub-Riemannian manifolds. Some of these examples are revisited in subsequent chapters. We start off with one of the simplest and most important sub-Riemannian geometries.

Example 1.1.5 (Heisenberg group)**.** Let $M = \mathbb{R}^3$ and consider the distribution H on $\mathbb{R}^3$ which is given as

$$H_{(x^1,x^2,x^3)} = \left\{ \left(x^1, x^2, x^3, v_1, v_2, v_3\right) \in T\mathbb{R}^3 \colon v_3 - \frac{1}{2}\left(x^1 v_2 - x^2 v_1\right) = 0 \right\} .$$

Let the fiber inner product $\langle \cdot, \cdot \rangle$ on H be

$$\mathrm{d}x^1 \otimes \mathrm{d}x^1 + \mathrm{d}x^2 \otimes \mathrm{d}x^2 .$$

We note that H is a field of two-planes in $\mathbb{R}^3$ which is generated by the vector fields

$$X_1 = \frac{\partial}{\partial x^1} - \frac{1}{2}x^2 \frac{\partial}{\partial x^3} \qquad \text{and} \qquad X_2 = \frac{\partial}{\partial x^2} + \frac{1}{2}x^1 \frac{\partial}{\partial x^3} .$$

These vector fields are orthonormal with respect to the fiber inner product $\langle \cdot, \cdot \rangle$ on H and are left-invariant on the Lie group obtained by endowing $\mathbb{R}^3$ with the group law

$$\left(x^1, x^2, x^3\right) \star \left(y^1, y^2, y^3\right) = \left(x^1 + y^1, x^2 + y^2, x^3 + y^3 + \frac{1}{2}\left(x^1 y^2 - y^1 x^2\right)\right) .$$

We compute

$$[X_1, X_2] = \frac{\partial}{\partial x^3} ,$$

which implies that the vector fields X_1, X_2 on $\mathbb{R}^3$ satisfy the Hörmander condition. In particular, H is a bracket generating distribution and $(H, \langle \cdot, \cdot \rangle)$ defines a sub-Riemannian structure on $\mathbb{R}^3$. The sub-Riemannian manifold $(\mathbb{R}^3, H, \langle \cdot, \cdot \rangle)$ is called Heisenberg group, indicating its connection with the Lie group $(\mathbb{R}^3, \star)$. The horizontal paths in this geometry are the absolutely continuous curves $(x_t^1, x_t^2, x_t^3)_{t\in[0,1]}$ which satisfy, for almost all $t \in [0, 1]$,

$$\dot{x}_t^3 = \frac{1}{2}\left(x_t^1 \dot{x}_t^2 - x_t^2 \dot{x}_t^1\right) . \tag{1.1.7}$$

We observe that, according to Stokes' theorem, the integral

$$\frac{1}{2}\int_0^t \left(x_s^1 \,\mathrm{d}x_s^2 - x_s^2 \,\mathrm{d}x_s^1\right)$$

gives the signed area of the closed curve in $\mathbb{R}^2$ obtained by first connecting the origin with the point (x_0^1, x_0^2) by a line segment, then traversing the path $(x_s^1, x_s^2)_{s\in[0,t]}$ and finally returning to the origin along a straight line segment. Hence, in the Heisenberg group particles are allowed to move freely in the (x^1, x^2)-plane with the third component being related to the signed area of the curve traced out by this motion. The problem of finding geodesics in the Heisenberg group then reduces to the Dido isoperimetric problem. We find that the geodesics are helices which are arcs of circles lifted to $\mathbb{R}^3$ by relation (1.1.7), with line segments included as a degenerate case, cf. Montgomery [Mon02, Chapter 1]. We finally remark that the Heisenberg group has an intrinsic sub-Riemannian Laplacian, which is the sum of squares operator

$$\Delta_H = X_1^2 + X_2^2 \ .$$

It is the sub-Riemannian Laplacian obtained by taking the divergence with respect to the left Haar measure $\nu = \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3$ on $(\mathbb{R}^3, \star)$.

The Heisenberg group generalises to any odd dimension $2n + 1$ as follows.

Example 1.1.6 (Heisenberg group H^n)**.** Let $M = \mathbb{R}^{2n+1}$. Take H to be the distribution on $\mathbb{R}^{2n+1}$ defined by

$$H_x = \left\{ (x, v) \in T\mathbb{R}^{2n+1} \colon v_{2n+1} - \frac{1}{2} \sum_{i=1}^{n} \left(x^i v_{n+i} - x^{n+i} v_i \right) = 0 \right\}$$

and set the fiber inner product $\langle \cdot, \cdot \rangle$ on H to be

$$\sum_{i=1}^{n} \left(\mathrm{d}x^i \otimes \mathrm{d}x^i + \mathrm{d}x^{n+i} \otimes \mathrm{d}x^{n+i} \right) \ .$$

Let $X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}$ be the vector fields on $\mathbb{R}^{2n+1}$ given, for $i \in \{1, \ldots, n\}$, by

$$X_i = \frac{\partial}{\partial x^i} - \frac{1}{2} x^{n+i} \frac{\partial}{\partial x^{2n+1}} \qquad \text{and} \qquad X_{n+i} = \frac{\partial}{\partial x^{n+i}} + \frac{1}{2} x^i \frac{\partial}{\partial x^{2n+1}} \ .$$

Then $(X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n})$ is a global orthonormal frame of the distribution H. Moreover, the vector fields are left-invariant on the Lie group $(\mathbb{R}^{2n+1}, \star)$ with group law

$$x \star y = \left(x^1 + y^1, \ldots, x^{2n} + y^{2n}, x^{2n+1} + y^{2n+1} + \frac{1}{2} \sum_{i=1}^{n} \left(x^i y^{n+i} - y^i x^{n+i} \right) \right) \ .$$

We have, for all $i \in \{1, \ldots, n\}$,

$$[X_i, X_{n+i}] = \frac{\partial}{\partial x^{2n+1}} \ .$$

Therefore, the vector fields $X_1, \ldots, X_n, X_{n+1}, \ldots, X_{2n}$ satisfy the Hörmander condition and the triple $(\mathbb{R}^{2n+1}, H, \langle \cdot, \cdot \rangle)$ is a sub-Riemannian manifold. It is called the Heisenberg group H^n and it admits the intrinsic sub-Riemannian Laplacian $\Delta_H = \sum_{i=1}^{n} (X_i^2 + X_{n+i}^2)$. We note that H^1 is the Heisenberg group described in Example 1.1.5.

The next example is simpler than the Heisenberg group, but it is in fact a rank-varying sub-Riemannian structure.

Example 1.1.7 (Grushin plane)**.** Let $M = \mathbb{R}^2$ and let X_1, X_2 be the vector fields on $\mathbb{R}^2$ given by

$$X_1 = \frac{\partial}{\partial x^1} \qquad \text{and} \qquad X_2 = x^1 \frac{\partial}{\partial x^2} \ .$$

We observe that $[X_1, X_2] = \frac{\partial}{\partial x^2}$. Thus, the vector fields X_1 and X_2 satisfy the Hörmander condition and therefore, induce a sub-Riemannian structure on $\mathbb{R}^2$. The induced fiber inner product, defined by polarising (1.1.1), is equal to

$$\mathrm{d}x^1 \otimes \mathrm{d}x^1 + \frac{1}{(x^1)^2}\,\mathrm{d}x^2 \otimes \mathrm{d}x^2 \ ,$$

which is in fact Riemannian outside the line $\{x^1 = 0\}$. Since the vector field X_2 vanishes along the line $\{x^1 = 0\}$, all geodesics which cross this line do so parallel to $\{x^2 = 0\}$. For instance, the points $(0,0)$ and $(0,1)$ are connected by the two families of geodesics $(\gamma^{k+} \colon k \in \mathbb{N})$ and $(\gamma^{k-} \colon k \in \mathbb{N})$ given by

$$\gamma_t^{k\pm} = \left(\frac{\pm \sin(k\pi t)}{\sqrt{k\pi/2}}, \frac{2k\pi t - \sin(2k\pi t)}{2k\pi} \right) \quad \text{for} \quad t \in [0,1] \ ,$$

depending on which direction we leave the origin in. For further details about geodesics in the Grushin plane, see Boscain and Laurent [BL13, Section 3.1].

A large class of examples of sub-Riemannian manifolds arises from contact geometry.

Example 1.1.8 (Contact manifold)**.** Let M be a manifold of dimension $2n+1$ and let H be a field of hyperplanes on M, that is, a subbundle of codimension 1. The distribution H can locally be written as the kernel of a one-form α, i.e.

$$H = \ker \alpha = \{X \in TM \colon \alpha(X) = 0\} \ .$$

We call H a contact structure on M if its locally defining one-form α satisfies

$$\alpha \wedge (\mathrm{d}\alpha)^n \neq 0 \tag{1.1.8}$$

at every point. This is referred to as the complete non-integrability condition in contact geometry. It is independent of the local choice of α because, for any smooth function $f \colon M \to \mathbb{R} \setminus \{0\}$, we have

$$(f\alpha) \wedge (\mathrm{d}\,(f\alpha))^n = f^{n+1} \alpha \wedge (\mathrm{d}\alpha)^n \ .$$

If α is a globally defined one-form satisfying (1.1.8), it is called a contact form on M. For a contact structure H on M, the pair (M, H) is called a contact manifold. Note that

from the condition (1.1.8) it follows that H is a bracket generating distribution on M. In particular, if we choose a fiber inner product $\langle\cdot,\cdot\rangle$ on the distribution H then $(M, H, \langle\cdot,\cdot\rangle)$ defines a sub-Riemannian manifold, also called a contact sub-Riemannian manifold.
We see that the Heisenberg group H^n from Example 1.1.6 can be described as a contact manifold by taking the manifold $\mathbb{R}^{2n+1}$ and endowing it with the contact form

$$\alpha = \mathrm{d}x^{2n+1} - \frac{1}{2}\sum_{i=1}^{n}\left(x^i\,\mathrm{d}x^{n+i} - x^{n+i}\,\mathrm{d}x^i\right) .$$

Another important class of contact manifolds is the class of Sasakian manifolds, introduced by Sasaki [Sas60], which now features prominently in theoretical physics and is thought to be important in studying the anti-de Sitter/conformal field theory correspondence in string theory.

The final example we present plays an important role in Chapter 4.

Example 1.1.9 (Carnot group)**.** Let $\mathbb{G}$ be a simply connected Lie group whose associated Lie algebra $\mathfrak{g}$ can be written, for some $N \in \mathbb{N}$, as

$$\mathfrak{g} = \mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_N$$

such that, for all $i, j \in \{1, \ldots, N\}$,

$$[\mathcal{V}_i, \mathcal{V}_j] = \begin{cases} \mathcal{V}_{i+j} & \text{if } i+j \leq N \\ 0 & \text{if } i+j > N \end{cases} . \tag{1.1.9}$$

We call $\mathbb{G}$ a Carnot group of step N. To see that a Carnot group can be considered as a sub-Riemannian manifold, observe that $\mathcal{V}_1$ extends to a left-invariant subbundle H on $\mathbb{G}$ which is bracket generating by (1.1.9). Hence, if we further fix an inner product on $\mathcal{V}_1$ and extend it to a left-invariant fiber inner product $\langle\cdot,\cdot\rangle$ on H, then the triple $(\mathbb{G}, H, \langle\cdot,\cdot\rangle)$ defines a sub-Riemannian manifold.
A Carnot group of step N is said to be free if its associated Lie algebra is isomorphic to the free nilpotent Lie algebra of step N on d generators for some $d \in \mathbb{N}$. Up to isomorphism, there exists a unique free Carnot group with given step and given number of generators. By [Bau04, Proposition 2.8], every free Carnot group is isomorphic to some $\mathbb{R}^m$ endowed with a polynomial group law. In that representation the exponential map reduces to the identity map. Alternatively, a free Carnot group can be represented as follows. Identify the free nilpotent Lie algebra of step N on d generators with the Lie algebra $\mathfrak{g}_N(\mathbb{R}^d)$ generated by d indeterminates inside the set of formal series truncated at order N, i.e.

$$\mathfrak{g}_N\left(\mathbb{R}^d\right) = \mathbb{R}^d \oplus \left[\mathbb{R}^d, \mathbb{R}^d\right] \oplus \cdots \oplus \underbrace{\left[\mathbb{R}^d, \left[\mathbb{R}^d, \ldots, \left[\mathbb{R}^d, \mathbb{R}^d\right] \ldots\right]\right]}_{(N-1)\text{ brackets}} .$$

Here the commutators are taken with respect to the tensor multiplication $\otimes$. The free Carnot group of step N is then given as $\mathbb{G}_N(\mathbb{R}^d) = \exp(\mathfrak{g}_N(\mathbb{R}^d))$, where we use the usual exponential of formal series.
For instance, the Heisenberg group introduced in Example 1.1.5 is isomorphic to the free Carnot group $\mathbb{G}_2(\mathbb{R}^2)$ of step 2 over $\mathbb{R}^2$. On the other hand, by a dimensional argument, we see that the Heisenberg group H^n introduced in Example 1.1.6, which generalises the Heisenberg group from Example 1.1.5, is not free for $n \geq 2$. We also remark that the additive groups $(\mathbb{R}^d, +)$ are the only commutative Carnot groups.

1.2 Malliavin calculus

Malliavin calculus, or the stochastic calculus of variations, has been developed from the program laid out by Paul Malliavin [Mal78a, Mal78b] to give a probabilistic proof of the Hörmander hypoellipticity theorem. Hörmander [Hör67] studied hypoelliptic second order differential operators with smooth coefficients using the theory of partial differential equations, and established the criterion that an operator which can be written as the sum of squares of smooth vector fields satisfying the Hörmander condition plus lower-order terms is hypoelliptic. From this analytic result, it follows that the solution of a stochastic differential equation with a generator of the above form has a smooth density at any fixed positive time. Malliavin outlined a method for directly proving the existence and smoothness of the density for the solution of such a stochastic differential equation, which initiated the theory of an infinite-dimensional differential calculus on the Wiener space. The theory was later expanded in different directions by Bismut [Bis81, Bis84], Kusuoka and Stroock [KS84, KS85, KS87], Shigekawa [Shi80], Stroock [Str81a, Str81b, Str83], Watanabe [Wat84, Wat87], and others.
Whereas Wiener functionals are not in a class of functionals to which the classical calculus of variations can be applied, Malliavin calculus provides the tools to define a derivative operator acting on Wiener functionals. With this in hand, one can investigate regularity properties of the law of Wiener functionals, and in particular analyse when the density for solutions of stochastic differential equations is smooth. A crucial tool in this analysis is an integration by parts formula on Gaussian spaces.
Over the years, Malliavin calculus has also become a powerful mechanism beyond the study of the regularity of probability laws, such as in developing a stochastic calculus for non-adapted processes, cf. Nualart [Nua98], and in mathematical finance, see Karatzas, Ocone and Li [KOL91], as well as Malliavin and Thalmaier [MT06].
In the following, we give an overview of the results and tools from Malliavin calculus which we use for our analysis of the small-time fluctuations for sub-Riemannian diffusion loops in Chapter 3. This exposition is by no means intended to be exhaustive. For the proofs of the results stated and for further reading, we refer to Bell [Bel87], Norris [Nor86] and Nualart [Nua06, Nua09]. We should also remark that, as in [Nor86], we are staying

close to Bismut's approach to Malliavin calculus, whereas [Nua06, Nua09] follow Stroock's formulation of Malliavin calculus. In the two approaches some terms differ by a Jacobian factor, and for instance, what we define to be the Malliavin covariance matrix is called the reduced Malliavin covariance matrix in Stroock's formulation.

1.2.1 The Bismut integration by parts formula

We first introduce the notion of a derived process of a stochastic process, which we then use to present Bismut's integration by parts formula.

Let $Z_0, Z_1, \ldots, Z_m$ be smooth vector fields on $\mathbb{R}^N$ and assume that they have a graded Lipschitz structure in the sense of Norris [Nor86]. This means that the vector fields and their derivatives of all orders satisfy polynomial growth bounds, and that there exist $k \in \mathbb{N}$ and $N_1, \ldots, N_k \in \mathbb{N}$ with $N_1 + \cdots + N_K = N$ such that under the identification of $\mathbb{R}^N$ with $\mathbb{R}^{N_1} \oplus \cdots \oplus \mathbb{R}^{N_k}$, giving the decompositions

$$z = \left(z^1, \ldots, z^k\right) \qquad \text{and} \qquad Z_i(z) = \left(Z_i^1(z), \ldots, Z_i^k(z)\right) \text{ for } i \in \{0, 1, \ldots, m\} ,$$

where $z^j \in \mathbb{R}^{N_j}$ and $Z_i^j(z) \in \mathbb{R}^{N_j}$ for $j \in \{1, \ldots, k\}$, the component $Z_i^j(z)$ depends only on $(z^1, \ldots, z^j)$ and the partial differential $\frac{\partial Z_i^j}{\partial z^j}$ is uniformly bounded. We impose this cascade structure as it ensures the existence and uniqueness of strong solutions to the stochastic differential equations we look at below.

Let $(B_t)_{t\in[0,1]}$ be a Brownian motion in $\mathbb{R}^m$, which is realised as the coordinate process on the path space $\{w \in C([0,1], \mathbb{R}^m) \colon w_0 = 0\}$ under Wiener measure $\mathbb{P}$. Consider the Itô stochastic differential equation in $\mathbb{R}^N$

$$\mathrm{d}z_t = \sum_{i=1}^{m} Z_i(z_t)\, \mathrm{d}B_t^i + Z_0(z_t)\, \mathrm{d}t , \quad z_0 = z , \tag{1.2.1}$$

for $z \in \mathbb{R}^N$. By [Nor86, Proposition 1.3], there exists a unique strong solution $(z_t)_{t\in[0,1]}$ to this stochastic differential equation, and $\sup_{t\in[0,1]} |z_t|$ has moments of all orders. Choose a smooth and bounded function $u \colon \mathbb{R}^N \to \mathbb{R}^m \otimes \mathbb{R}^d$ whose derivatives are of polynomial growth. For $\eta \in \mathbb{R}^d$, define a perturbed process $(B_t^\eta)_{t\in[0,1]}$ in $\mathbb{R}^m$ by

$$\mathrm{d}B_t^\eta = \mathrm{d}B_t + u(z_t)\eta\, \mathrm{d}t , \quad B_0^\eta = 0 .$$

Let $(z_t^\eta)_{t\in[0,1]}$ in $\mathbb{R}^N$ be the strong solution of the stochastic differential equation

$$\mathrm{d}z_t^\eta = \sum_{i=1}^{m} Z_i(z_t^\eta)\, \mathrm{d}B_t^{\eta,i} + Z_0(z_t^\eta)\, \mathrm{d}t , \quad z_0^\eta = z ,$$

which is the equation (1.2.1) with Brownian motion $(B_t)_{t\in[0,1]}$ replaced by the perturbed process $(B_t^\eta)_{t\in[0,1]}$. From [Nor86, Proposition 2.2], it follows that we can choose a version of the family $((z_t^\eta)_{t\in[0,1]} \colon \eta \in \mathbb{R}^d)$ of processes which is almost surely smooth in η, and that the process $(z_t')_{t\in[0,1]}$ in $\mathbb{R}^N \otimes \mathbb{R}^d$ given by

$$z_t' = \left.\frac{\partial}{\partial \eta}\right|_{\eta=0} z_t^\eta$$

satisfies the stochastic differential equation

$$\mathrm{d}z_t' = \sum_{i=1}^m \nabla Z_i(z_t) z_t' \, \mathrm{d}B_t^i + \nabla Z_0(z_t) z_t' \, \mathrm{d}t + \sum_{i=1}^m Z_i(z_t) \otimes u(z_t)^i \, \mathrm{d}t \;, \quad z_0' = 0 \;. \tag{1.2.2}$$

Here $u(z_t)^i$ denotes the ith row of $u(z_t)$. We call $(z_t')_{t\in[0,1]}$ a derived process associated with the stochastic process $(z_t)_{t\in[0,1]}$. Using the notion of derived processes, Bismut's integration by parts formula, cf. [Bis81, Theorem 2.1] and [Nor86, Theorem 2.3], can be stated as follows.

Theorem 1.2.1 (Bismut's integration by parts formula). *Let $(z_t')_{t\in[0,1]}$ be the derived process associated to the process $(z_t)_{t\in[0,1]}$ in $\mathbb{R}^N$ for some choice of $u \colon \mathbb{R}^N \to \mathbb{R}^m \otimes \mathbb{R}^d$ smooth and bounded, with all its derivatives of polynomial growth. Then for any bounded differentiable function $\phi \colon \mathbb{R}^N \to \mathbb{R}$ with bounded first derivatives, and for all $t \in [0,1]$, we have*

$$\mathbb{E}\left[\nabla\phi(z_t) z_t'\right] = \mathbb{E}\left[\phi(z_t) \sum_{i=1}^m \int_0^t u(z_s)^i \, \mathrm{d}B_s^i\right] \;. \qquad \square$$

We note that the stochastic process $(z_t, z_t')_{t\in[0,1]}$ is itself the strong solution of a stochastic differential equation in $\mathbb{R}^N \oplus \mathbb{R}^{Nd}$ with smooth coefficients which have a graded Lipschitz structure. In particular, we can iterate Bismut's integration by parts formula. This is a crucial observation in studying the regularity of probability laws, and is used in proving the criterion for a stochastic process to have a smooth density which is presented next.

1.2.2 Smooth density and the Malliavin covariance matrix

We first define the Malliavin covariance matrix of a stochastic process and then give a criterion for a stochastic process to have a smooth density which is expressed in terms of the associated Malliavin covariance matrix.

Let $X_0, X_1, \ldots, X_m$ be smooth vector fields on $\mathbb{R}^d$ and define a vector field $\underline{X}_0$ on $\mathbb{R}^d$ by

$$\underline{X}_0 = X_0 + \frac{1}{2} \sum_{i=1}^m \nabla_{X_i} X_i \;,$$

where ∇ is understood as the Levi-Civita connection with respect to the Euclidean metric. Assume that the vector fields $\underline{X}_0, X_1, \ldots, X_m$ have bounded first derivatives and higher derivatives of polynomial growth.

Fix $x \in \mathbb{R}^d$, and define processes $(x_t)_{t\in[0,1]}$ in $\mathbb{R}^d$ and $(v_t)_{t\in[0,1]}$ in $\mathbb{R}^d \otimes (\mathbb{R}^d)^*$ as the strong solutions of the system of stochastic differential equations

$$\mathrm{d}x_t = \sum_{i=1}^m X_i(x_t)\,\mathrm{d}B_t^i + \underline{X}_0(x_t)\,\mathrm{d}t\ , \qquad x_0 = x\ , \tag{1.2.3}$$
$$\mathrm{d}v_t = -\sum_{i=1}^m v_t \nabla X_i(x_t)\,\mathrm{d}B_t^i - v_t\left(\nabla\underline{X}_0 - \sum_{i=1}^m (\nabla X_i)^2\right)(x_t)\,\mathrm{d}t\ , \qquad v_0 = I\ .$$

The process $(v_t)_{t\in[0,1]}$ is in fact the inverse of the derivative of the flow associated with the stochastic differential equation defining $(x_t)_{t\in[0,1]}$. It features in the expression for the Malliavin covariance matrix.

Definition 1.2.2. *For $t \in [0,1]$, we call*

$$c_t = \sum_{i=1}^m \int_0^t (v_s X_i(x_s)) \otimes (v_s X_i(x_s))\ \mathrm{d}s$$

the Malliavin covariance matrix of the random variable x_t.

Let $(x_t')_{t\in[0,1]}$ be the derived process associated with the process $(x_t)_{t\in[0,1]}$ for the choice of u having $u(x_t)^i = v_t X_i(x_t)$. The general form (1.2.2) implies that this derived process satisfies the stochastic differential equation in $\mathbb{R}^d \otimes \mathbb{R}^d$

$$\mathrm{d}x_t' = \sum_{i=1}^m \nabla X_i(x_t)x_t'\,\mathrm{d}B_t^i + \nabla\underline{X}_0(x_t)x_t'\,\mathrm{d}t + \sum_{i=1}^m X_i(x_t)\otimes(v_t X_i(x_t))\ \mathrm{d}t\ , \quad x_0' = 0\ . \tag{1.2.4}$$

Indeed, the stochastic process $(x_t, x_t')_{t\in[0,1]}$ is the unique strong solution of the system of stochastic differential equations given by (1.2.3) and (1.2.4). Similarly, $(x_t, v_t^{-1})_{t\in[0,1]}$ is the unique strong solution of the system consisting of (1.2.3) and the stochastic differential equation in $\mathbb{R}^d \otimes (\mathbb{R}^d)^*$

$$\mathrm{d}\left(v_t^{-1}\right) = \sum_{i=1}^m \nabla X_i(x_t)v_t^{-1}\,\mathrm{d}B_t^i + \nabla\underline{X}_0(x_t)v_t^{-1}\,\mathrm{d}t\ , \quad v_0^{-1} = I\ . \tag{1.2.5}$$

Using Definition 1.2.2 and (1.2.5), we further compute that

$$\mathrm{d}\left(v_t^{-1}c_t\right) = \sum_{i=1}^m \nabla X_i(x_t)\left(v_t^{-1}c_t\right)\ \mathrm{d}B_t^i + \nabla\underline{X}_0(x_t)\left(v_t^{-1}c_t\right)\ \mathrm{d}t$$
$$+ v_t^{-1}\sum_{i=1}^m (v_t X_i(x_t))\otimes(v_t X_i(x_t))\ \mathrm{d}t\ .$$

Thus, the process $(v_t^{-1}c_t)_{t\in[0,1]}$ satisfies the stochastic differential equation (1.2.4), and by uniqueness, it follows that almost surely, for all $t \in [0,1]$,

$$x'_t = v_t^{-1} c_t \ .$$

This observation and iterating Bismut's integration by parts formula are the main tools in proving the following criterion for a smooth density. For the details, see [Nor86, Section 3], or [Nua09, Section 5] in Stroock's formulation.

Theorem 1.2.3. *Let $t \in (0,1]$. Suppose that, for all $p < \infty$, we have*

$$\mathbb{E}\left[\left|\det c_t^{-1}\right|^p\right] < \infty \ .$$

Then the law of x_t has a smooth density with respect to Lebesgue measure on $\mathbb{R}^d$. □

Equipped with this criterion, we can study the regularity of probability laws by analysing associated Malliavin covariance matrices. The theorem becomes powerful in conjunction with the next result, cf. [Nor86, Theorem 4.2].

Theorem 1.2.4. *Fix $x \in \mathbb{R}^d$. Suppose that the vectors $X_1(x), \dots, X_m(x)$ together with the collection of vectors*

$$[X_{i_1}, [X_{i_2}, \dots, [X_{i_{k-1}}, X_{i_k}] \dots]](x) \quad \textit{for} \quad k \geq 2 \textit{ and } 0 \leq i_1, \dots, i_k \leq m \tag{1.2.6}$$

span $\mathbb{R}^d$. Then the Malliavin covariance matrix of the process $(x_t)_{t \in [0,1]}$ defined by the stochastic differential equation (1.2.3) satisfies, for all $t \in (0,1]$ and all $p < \infty$,

$$\mathbb{E}\left[\left|\det c_t^{-1}\right|^p\right] < \infty \ . \qquad \square$$

We observe that the collection of vectors (1.2.6) contains commutator brackets, evaluated at x, which use the vector field X_0. Hence, the condition that the vectors $X_1(x), \dots, X_m(x)$ together with the collection of vectors (1.2.6) span $\mathbb{R}^d$ is weaker than requiring the vector fields $X_1, \dots, X_m$ to satisfy the Hörmander condition at $x \in \mathbb{R}^d$. If this weaker condition holds, we say that the vector fields $X_0, X_1, \dots, X_m$ satisfy the weak Hörmander condition at the point x.

By the Kusuoka-Stroock estimate, cf. [AKS93] or see Watanabe [Wat87, Theorem 3.2], we know that under the weak Hörmander condition the quantity $\mathbb{E}[|\det c_t^{-1}|^p]$, for $p < \infty$ fixed, blows up at most polynomially as $t \to 0$. The control provided by this estimate is a crucial ingredient in the proof of Theorem 3.1.3 in Chapter 3.

Theorem 1.2.5 (Kusuoka-Stroock estimate)**.** *Suppose the vector fields $X_0, X_1, \dots, X_m$ satisfy the weak Hörmander condition at x. Then there exist a positive integer n and, for all $p < \infty$, constants $C(p) < \infty$ such that, for all $t \in (0,1]$,*

$$\left(\mathbb{E}\left[\left|\det c_t^{-1}\right|^p\right]\right)^{1/p} \leq C(p) t^{-n} \ . \qquad \square$$

Combining Theorem 1.2.3 with Theorem 1.2.4, or the stronger version Theorem 1.2.5, we obtain the following result on the regularity of probability laws.

Theorem 1.2.6. *Fix $x \in \mathbb{R}^d$. Let $X_0, X_1, \ldots, X_m$ be smooth vector fields on $\mathbb{R}^d$ which satisfy the weak Hörmander condition at x. Suppose that $\underline{X}_0, X_1, \ldots, X_m$ have bounded first derivatives and higher derivatives of polynomial growth. Let $(x_t)_{t\in[0,1]}$ be the unique strong solution of the Itô stochastic differential equation*

$$\mathrm{d}x_t = \sum_{i=1}^m X_i(x_t)\,\mathrm{d}B_t^i + \underline{X}_0(x_t)\,\mathrm{d}t\ , \quad x_0 = x\ .$$

Then, for all $t \in (0,1]$, the law of the random variable x_t has a smooth density with respect to Lebesgue measure on $\mathbb{R}^d$. □

Note that the second order partial differential operator $\mathcal{L}$ on $\mathbb{R}^d$ given by

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^m X_i^2 + X_0$$

is the generator of the process $(x_t)_{t\in[0,1]}$ defined as the unique strong solution of (1.2.3). Thus, Theorem 1.2.6 says that the operator $\frac{\partial}{\partial t} - \mathcal{L}_y^*$ has a smooth fundamental solution. By transferring the consideration from the fundamental solution of $\frac{\partial}{\partial t} - \mathcal{L}_y^*$ to the resolvent kernel of the operator $\mathcal{L}$, Kusuoka and Stroock [KS85] give a complete probabilistic proof of the Hörmander hypoellipticity theorem, which circumvents the use of intermediate subelliptic estimates. It is a consequence of [KS85, Corollary 8.18].

Theorem 1.2.7 (Hörmander's hypoellipticity theorem)**.** *Let M be a connected smooth manifold. Let $X_0, X_1, \ldots, X_m$ be smooth vector fields on M and let f be a smooth function on M. Suppose that the vector fields $X_1, \ldots, X_m$ satisfy the Hörmander condition. Then the operator $\mathcal{L}$ on M given as*

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^m X_i^2 + X_0 + f$$

is hypoelliptic. □

Chapter 2

Example illustrating fluctuations results for sub-Riemannian bridges

We provide an example to illustrate the work by Bailleul, Mesnager and Norris [BMN15] on the small-time fluctuations for the bridge of a sub-Riemannian diffusion process. From the result [BMN15, Theorem 1.3] it follows, as asserted by Molchanov [Mol75], that the law of the small-time fluctuations of a Brownian bridge on a Riemannian manifold between two points which are connected by a unique strongly minimal path is absolutely continuous with respect to the law of the parallel translation of a Brownian bridge from 0 to 0 in the tangent space at the initial position along the unique minimal path. The example we construct demonstrates that in the more general setting of sub-Riemannian geometry, the small-time fluctuations for diffusion bridges can exhibit exotic behaviours, i.e. qualitatively different behaviours compared to Brownian bridges.

2.1 Fluctuations results for sub-Riemannian bridges

We recall the results from [BMN15] on the small-time fluctuations for sub-Riemannian diffusion bridges. To simplify the presentation, we avoid the full generality of [BMN15], and instead restrict our attention to generators in divergence form. As our example in the subsequent section falls into that class, this is sufficient for our considerations.
Let M be a connected smooth manifold of dimension d and let $X_1, \dots, X_m$ be smooth vector fields on M which satisfy the Hörmander condition, i.e. the vector fields together with their commutator brackets of all orders span the tangent space at every point in the manifold. The energy function I on the set of continuous paths $\Omega = C([0,1], M)$ associated with these vector fields can be defined as follows. Suppose that $\omega \in \Omega$ is an absolutely continuous path and that there exists a measurable path $\xi \colon [0,1] \to T^*M$ with $\xi_t \in T^*_{\omega_t}M$ and $\dot{\omega}_t = \sum_{i=1}^m \xi_t(X_i) X_i$ for almost all $t \in [0,1]$. Then ω has energy

$$I(\omega) = \sum_{i=1}^m \int_0^1 \xi_t(X_i)^2 \, \mathrm{d}t \; .$$

Otherwise, we set $I(\omega) = \infty$. For $x, y \in M$, the subset of Ω consisting of the horizontal paths from x to y, which are the paths of finite energy connecting x to y, is given as

$$H^{x,y} = \{\omega \in \Omega \colon I(\omega) < \infty \text{ and } \omega_0 = x,\ \omega_1 = y\} \ .$$

Since $X_1, \ldots, X_m$ satisfy the Hörmander condition, the set $H^{x,y}$ is non-empty by the Chow-Rashevskii theorem, and the topology induced by the sub-Riemannian distance

$$d(x,y) = \inf_{\omega \in H^{x,y}} \sqrt{I(\omega)}$$

is equivalent to the topology of M. Recall that a path $\gamma \in H^{x,y}$ is called strongly minimal if there exist $\delta > 0$ and a relatively compact open set $U \subset M$ such that

$$I(\gamma) \leq I(\omega) \text{ for all } \omega \in H^{x,y} \quad \text{and} \quad I(\gamma) + \delta \leq I(\omega) \text{ for all } \omega \in H^{x,y} \text{ which leave } U \ .$$

We are interested in the small-time fluctuations of the diffusion bridge measures associated with the vector fields $X_1, \ldots, X_m$. Choose a positive smooth measure ν on M and define a second order partial differential operator $\mathcal{L}$ on M by

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^{m} \left(X_i^2 + (\operatorname{div} X_i)\, X_i \right) \ , \tag{2.1.1}$$

where the divergence is understood with respect to ν. If the vector fields $X_1, \ldots, X_m$ are linearly independent at every point then the operator $\mathcal{L}$ is exactly the sub-Riemannian Laplacian with respect to the measure ν associated with the sub-Riemannian structure on M induced by $X_1, \ldots, X_m$. We also remark that $\mathcal{L}$ is an operator in divergence form because, for all smooth functions f of compact support in M, we have

$$\mathcal{L} f = \frac{1}{2} \operatorname{div} \left(\sum_{i=1}^{m} X_i(f) X_i \right) \ .$$

Let p be the Dirichlet heat kernel for $\mathcal{L}$ with respect to ν. Fix $x \in M$ and let $\varepsilon > 0$. Consider the diffusion process $(x_t^\varepsilon)_{t<\zeta}$ defined up to explosion time ζ which starts from x and has generator $\varepsilon \mathcal{L}$. This process may explode with positive probability before time 1, but on the event $\{\zeta > 1\}$, the process $(x_t^\varepsilon)_{t \in [0,1]}$ has a sub-probability law μ_ε^x on Ω. We can disintegrate μ_ε^x uniquely as

$$\mu_\varepsilon^x(\mathrm{d}\omega) = \int_M \mu_\varepsilon^{x,y}(\mathrm{d}\omega) p(\varepsilon, x, y) \nu(\mathrm{d}y) \ ,$$

where $(\mu_\varepsilon^{x,y} \colon y \in M)$ is a family of probability measures on Ω, which is weakly continuous in y, with the diffusion bridge measure $\mu_\varepsilon^{x,y}$ supported on $\Omega^{x,y} = \{\omega \in \Omega \colon \omega_0 = x, \omega_1 = y\}$ for all $y \in M$.

Let the endpoints $x, y \in M$ be such that the pair (x, y) lies outside the sub-Riemannian cut locus. In particular, there exists a unique strongly minimal path $\gamma \in H^{x,y}$. Write $T_\gamma\Omega^{x,y}$ for the set of continuous paths $v\colon [0,1] \to TM$ over γ, i.e. $v_t \in T_{\gamma_t}M$ for all $t \in [0,1]$, with $v_0 = v_1 = 0$. Choose a smooth map $\theta\colon M \to \mathbb{R}^d$ which restricts to a diffeomorphism on a neighbourhood of $\{\gamma_t\colon 0 \le t \le 1\}$, and define a rescaling map $\sigma_\varepsilon\colon \Omega^{x,y} \to T_\gamma\Omega^{x,y}$ by

$$\sigma_\varepsilon(\omega)_t = \frac{(\mathrm{d}\theta_{\gamma_t})^{-1}\left(\theta(\omega_t) - \theta(\gamma_t)\right)}{\sqrt{\varepsilon}} \ .$$

Let $\tilde{\mu}_\varepsilon^{x,y}$ be the pushforward measure of $\mu_\varepsilon^{x,y}$ by σ_ε, i.e. the probability measure on $T_\gamma\Omega^{x,y}$ given by

$$\tilde{\mu}_\varepsilon^{x,y} = \mu_\varepsilon^{x,y} \circ \sigma_\varepsilon^{-1} \ .$$

According to [BMN15, Theorem 1.3], the rescaled diffusion bridge measures $\tilde{\mu}_\varepsilon^{x,y}$ converge weakly to a zero-mean Gaussian measure μ_γ on $T_\gamma\Omega^{x,y}$ as $\varepsilon \to 0$. One way of characterising the resulting limit measure is in terms of the bicharacteristic flow of $\mathcal{L}$. Set $\mathcal{H}\colon T^*M \to \mathbb{R}$ to be the Hamiltonian

$$\mathcal{H}(\lambda) = \frac{1}{2}\sum_{i=1}^m \lambda(X_i)^2 \quad \text{for} \quad \lambda \in T^*M \ .$$

Let β be the canonical symplectic two-form on T^*M and let V denote the smooth vector field on T^*M given by $\beta(V, \cdot) = \mathrm{d}\mathcal{H}$. The bicharacteristic flow of $\mathcal{L}$ is the maximal flow $(\psi_t(\lambda)\colon \lambda \in T^*M, t \in (\zeta^-(\lambda), \zeta^+(\lambda)))$ of the vector field V. This means, for all $\lambda \in T^*M$, we have $\psi_0(\lambda) = \lambda$ as well as $\zeta^-(\lambda) < 0 < \zeta^+(\lambda)$, and

$$\dot{\psi}_t(\lambda) = V\left(\psi_t\left(\lambda\right)\right) \quad \text{for} \quad t \in \left(\zeta^-(\lambda), \zeta^+(\lambda)\right) ,$$

and $\psi_t(\lambda)$ leaves all compact sets in T^*M as $t \to \zeta^+(\lambda)$ if $\zeta^+(\lambda) < \infty$ and as $t \to \zeta^-(\lambda)$ if $\zeta^-(\lambda) > -\infty$. The integral curves of V are called bicharacteristics. Write $\pi\colon T^*M \to M$ for the projection of the bundle. Since (x, y) is assumed to lie outside the sub-Riemannian cut locus, there exists, as detailed in [BA88], a unique bicharacteristic $(\lambda_t)_{t\in[0,1]}$ such that $\gamma_t = \pi\lambda_t$ for all $t \in [0,1]$. The covariance structure of the zero-mean Gaussian limit measure μ_γ on $T_\gamma\Omega^{x,y}$ is given in terms of the following linear maps. For $t \in [0,1]$, define $J_t\colon T_x^*M \to T_{\gamma_t}M$ and $K_t\colon T_y^*M \to T_{\gamma_t}M$ by

$$J_t\xi_0 = \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \pi\psi_t\left(\lambda_0 + \varepsilon\xi_0\right) \quad \text{and} \quad K_t\xi_1 = \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \pi\psi_{-(1-t)}\left(\lambda_1 - \varepsilon\xi_1\right) \ . \tag{2.1.2}$$

Due to Definition 1.1.4 of the sub-Riemannian cut locus, we are guaranteed that the linear map J_1 is invertible. Thus, the linear map, for $0 \le s \le t \le 1$,

$$J_s J_1^{-1} K_t^*\colon T_{\gamma_t}^*M \to T_{\gamma_s}M$$

is well-defined. Combining the result [BMN15, Theorem 1.3] on the weak convergence of the rescaled diffusion bridge measures $\tilde{\mu}_\varepsilon^{x,y}$ and the characterisation [BMN15, Theorem 2.1] of the zero-mean Gaussian limit measure μ_γ on $T_\gamma\Omega^{x,y}$, and by restricting our attention to the class of operators which are of the divergence form (2.1.1), we obtain the following theorem on the small-time fluctuations for the bridge of a sub-Riemannian diffusion.

Theorem 2.1.1. *Let M be a connected smooth manifold and fix $x, y \in M$. Let $\mathcal{L}$ be a second order partial differential operator on M of the form*

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{m}\left(X_i^2 + (\operatorname{div} X_i)\, X_i\right) ,$$

where the divergence is taken with respect to a positive smooth measure ν on M, and where $X_1, \ldots, X_m$ are smooth vector fields on M satisfying the Hörmander condition. Suppose there exists a unique strongly minimal path $\gamma \in H^{x,y}$ and that the pair (x, y) lies outside the sub-Riemannian cut locus. Then, as $\varepsilon \to 0$, the rescaled diffusion bridge measures $\tilde{\mu}_\varepsilon^{x,y}$ converge weakly to the unique zero-mean Gaussian measure μ_γ on $T_\gamma\Omega^{x,y}$ whose covariance is given, for $0 \le s \le t \le 1$, by

$$\int_{T_\gamma\Omega^{x,y}} v_s \otimes v_t\, \mu_\gamma(\mathrm{d}v) = J_s J_1^{-1} K_t^* . \qquad \square$$

In the following section, we determine the two families of linear maps $(J_t\colon t \in [0,1])$ and $(K_t\colon t \in [0,1])$ for a particular choice of sub-Riemannian geometry, and thereby show that the small-time fluctuations for the bridge of a sub-Riemannian diffusion process can exhibit qualitatively different behaviours compared to Brownian bridges.

2.2 Bridge with exotic small-time fluctuations

By means of a specific example, we show that the small-time fluctuations for the bridge of a sub-Riemannian diffusion can exhibit exotic behaviours. Fix $M = \mathbb{R}^3$. Choose a smooth and bounded function $\varphi\colon \mathbb{R} \to \mathbb{R}$ and let X_1, X_2, X_3 be the vector fields on $\mathbb{R}^3$ defined by

$$X_1 = \frac{\partial}{\partial x^1} , \quad X_2 = \left(\varphi(x^1) + x^3\right)\frac{\partial}{\partial x^2} \quad \text{and} \quad X_3 = \frac{\partial}{\partial x^3} .$$

Note that $[X_3, X_2] = \frac{\partial}{\partial x^2}$ and in particular, that $X_1, X_3, [X_3, X_2]$ span the tangent space at every point in $\mathbb{R}^3$. Hence, the vector fields X_1, X_2, X_3 satisfy the Hörmander condition. Let ν be Lebesgue measure on $\mathbb{R}^3$. Since $\sum_{i=1}^3 (\operatorname{div} X_i)\, X_i = 0$ with respect to ν, the operator $\mathcal{L}$ on $\mathbb{R}^3$ given by (2.1.1) is the sum of squares operator

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{3} X_i^2 = \frac{1}{2}\left(\frac{\partial^2}{\partial (x^1)^2} + \left(\varphi(x^1) + x^3\right)^2\frac{\partial^2}{\partial (x^2)^2} + \frac{\partial^2}{\partial (x^3)^2}\right) . \qquad (2.2.1)$$

Fix the initial position to be $x = (0,0,0)$ and the final position to be $y = (1,0,0)$. Let us consider an absolutely continuous path $\omega\colon [0,1] \to \mathbb{R}^3$ with $\omega_0 = x$ and $\omega_1 = y$. Since $X_1 = \frac{\partial}{\partial x^1}$ is orthogonal to both X_2 and X_3 at every point, we obtain

$$I(\omega) \geq \int_0^1 \left(\dot\omega_t^1\right)^2 \, \mathrm{d}t \ ,$$

with equality if and only if $\dot\omega_t^2 = \dot\omega_t^3 = 0$ for almost all $t \in [0,1]$. Using the Cauchy-Schwarz inequality, we further deduce

$$I(\omega) \geq \left(\int_0^1 \dot\omega_t^1 \, \mathrm{d}t\right)^2 = 1 \ ,$$

with equality if and only if $\omega_t^1 = t$ and $\omega_t^2 = \omega_t^3 = 0$ for all $t \in [0,1]$. This shows that the path $\gamma \in H^{x,y}$ given by

$$\gamma_t = (t,0,0) \quad \text{for} \quad t \in [0,1]$$

is the unique minimal path in $H^{x,y}$. Moreover, by the Hopf-Rinow theorem, the path γ is strongly minimal because $\mathbb{R}^3$ endowed with the sub-Riemannian distance function induced by the vector fields X_1, X_2, X_3 is a complete metric space.

Applying the bicharacteristic flow approach from [BMN15, Section 2], which we recalled in Section 2.1, we determine the small-time fluctuations for the bridge from x to y of the sub-Riemannian diffusion process with generator $\mathcal{L}$. Changing to a Hamiltonian point of view, we denote the coordinates on $T^*\mathbb{R}^3$ by $(q,p) = (q^1, q^2, q^3, p_1, p_2, p_3)$. The Hamiltonian $\mathcal{H}\colon T^*\mathbb{R}^3 \to \mathbb{R}$ associated with the operator $\mathcal{L}$ in (2.2.1) is

$$\mathcal{H}(q,p) = \frac{1}{2}\left(p_1^2 + \left(\varphi\left(q^1\right) + q^3\right)^2 p_2^2 + p_3^2\right) \quad \text{for} \quad (q,p) \in T^*\mathbb{R}^3 \ .$$

The bicharacteristics, i.e. the integral curves of the corresponding Hamiltonian vector field, are the solutions to the Hamiltonian equations

$$\dot q^k = \frac{\partial \mathcal{H}}{\partial p_k} \ , \qquad \dot p_k = -\frac{\partial \mathcal{H}}{\partial q^k} \ .$$

In our example, these equations read as follows.

$$\begin{aligned} \dot q_t^1 &= p_{t,1} & \dot p_{t,1} &= -\left(\varphi\left(q_t^1\right) + q_t^3\right)\varphi'\left(q_t^1\right) p_{t,2}^2 \\ \dot q_t^2 &= \left(\varphi\left(q_t^1\right) + q_t^3\right)^2 p_{t,2} & \dot p_{t,2} &= 0 \\ \dot q_t^3 &= p_{t,3} & \dot p_{t,3} &= -\left(\varphi\left(q_t^1\right) + q_t^3\right) p_{t,2}^2 \end{aligned} \tag{2.2.2}$$

In particular, the curve $(\lambda_t)_{t\in[0,1]}$ given by $\lambda_t = (t,0,0,1,0,0)$ is a bicharacteristic which projects onto the unique minimal path $\gamma \in H^{x,y}$. We now aim to determine the linear maps $J_t\colon T_x^*\mathbb{R}^3 \to T_{\gamma_t}\mathbb{R}^3$ and $K_t\colon T_y^*\mathbb{R}^3 \to T_{\gamma_t}\mathbb{R}^3$ which are defined by (2.1.2). It is in fact enough to find the maps $J_{ts}\colon T^*_{\gamma_s}\mathbb{R}^3 \to T_{\gamma_t}\mathbb{R}^3$, for $0 \le s \le t \le 1$, given as

$$J_{ts}\xi_s = \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \pi\psi_{t-s}\left(\lambda_s + \varepsilon\xi_s\right) \ ,$$

since $J_t = J_{t0}$ by definition, and $K_t^* = J_{1t}$ by a generalisation of a calculation performed in [BMN15, Section 2]. The underlying idea for computing the linear maps J_{ts} is to use approximate solutions of the Hamiltonian equations (2.2.2) which are close enough to the actual solutions so that they have the same limit behaviour as $\varepsilon \to 0$.

Before we proceed, let us recall the following theorem on the dependence of solutions of ordinary differential equations on initial conditions, cf. Dieudonné [Die69, Section 10.8]. It ensures the existence of bicharacteristics through $\lambda_s + \varepsilon\xi_s$ up to sufficiently large times, for small enough ε.

Theorem 2.2.1. *Let $U \subset \mathbb{R}^n$ be open and let $V\colon U \to \mathbb{R}^n$ be a locally Lipschitz vector field. For $z \in U$, we denote the lifetime of the unique solution of the ordinary differential equation*

$$\dot{z}_t = V(z_t) \quad \textit{subject to} \quad z_0 = z$$

by $\zeta(z)$. Then, for all $T < \zeta(z)$, there exists some $\varepsilon_0 = \varepsilon_0(T) > 0$ such that $B_{\varepsilon_0}(z) \subset U$ and $\zeta(\tilde{z}) > T$ for all $\tilde{z} \in B_{\varepsilon_0}(z)$. □

Fix $s \in [0,1]$. Let $a, b, c \in \mathbb{R}$ be arbitrary and set $\xi_s = (s, 0, 0, a, b, c)$. Since $(\lambda_t)_{t\in[0,1]}$ extends to an integral curve for all times, Theorem 2.2.1 implies that there exists some $\varepsilon_0 > 0$ such that, for all $\varepsilon \in (0, \varepsilon_0)$, the bicharacteristic

$$(q_t^\varepsilon, p_t^\varepsilon) = \left(q_t^{\varepsilon,1}, q_t^{\varepsilon,2}, q_t^{\varepsilon,3}, p_{t,1}^\varepsilon, p_{t,2}^\varepsilon, p_{t,3}^\varepsilon\right)$$

through $(q_s^\varepsilon, p_s^\varepsilon) = \lambda_s + \varepsilon\xi_s = (s, 0, 0, 1+\varepsilon a, \varepsilon b, \varepsilon c)$ exists for all $t \in [0,1]$. Note that here we fix the initial condition at time $t = s$. Besides, for $t \in [0,1]$, let

$$\begin{aligned} Q_t^{\varepsilon,1} &= t + \varepsilon a(t-s) & \qquad P_{t,1}^\varepsilon &= 1 + \varepsilon a \\ Q_t^{\varepsilon,2} &= \varepsilon b \int_s^t \varphi^2(r)\,\mathrm{d}r & P_{t,2}^\varepsilon &= \varepsilon b \\ Q_t^{\varepsilon,3} &= \varepsilon c(t-s) & P_{t,3}^\varepsilon &= \varepsilon c \end{aligned} \tag{2.2.3}$$

and set

$$(Q_t^\varepsilon, P_t^\varepsilon) = \left(Q_t^{\varepsilon,1}, Q_t^{\varepsilon,2}, Q_t^{\varepsilon,3}, P_{t,1}^\varepsilon, P_{t,2}^\varepsilon, P_{t,3}^\varepsilon\right) \ .$$

We show that (2.2.3) is an approximate solution of the Hamiltonian equations (2.2.2), which is close enough to the actual solution so that the following proposition holds. The result is used in determining the linear map J_{ts}.

Proposition 2.2.2. *For all $t \in [s,1]$, we have*

$$\left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} q_t^\varepsilon = \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} Q_t^\varepsilon \ .$$

In proving this proposition, we need control over how far Q_t^ε deviates from q_t^ε. Observe

$$(Q_s^\varepsilon, P_s^\varepsilon) = (s,0,0,1+\varepsilon a, \varepsilon b, \varepsilon c) = (q_s^\varepsilon, p_s^\varepsilon)\ ,$$

and that for the functions $F, G\colon [0,1]\times T^*\mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$\begin{aligned} F\left(t,q,p\right) &= \left(p_1, \varphi^2(t)p_2, p_3\right)\ , \\ G\left(t,q,p\right) &= 0\ , \end{aligned} \tag{2.2.4}$$

it holds true that

$$\begin{aligned} \dot{Q}_t^\varepsilon &= F\left(t, Q_t^\varepsilon, P_t^\varepsilon\right)\ , \\ \dot{P}_t^\varepsilon &= G\left(t, Q_t^\varepsilon, P_t^\varepsilon\right)\ . \end{aligned}$$

Similarly, let $f, g\colon T^*\mathbb{R}^3 \to \mathbb{R}^3$ be such that the Hamiltonian equations (2.2.2) write as

$$\begin{aligned} \dot{q}_t &= f\left(q_t, p_t\right)\ , \\ \dot{p}_t &= g\left(q_t, p_t\right)\ . \end{aligned}$$

The proof of Proposition 2.2.2 relies on the lemma below, which is used to gain control over the quantity $\|q_t^\varepsilon - Q_t^\varepsilon\|_1$ for $t\in[s,1]$ and $\varepsilon>0$ small enough. Here $\|\cdot\|_1$ denotes the ℓ^1-norm of a vector.

Lemma 2.2.3. *Suppose that $\varepsilon \in (0,1)$. Then there exist constants D_1 and D_2, which depend on a, b and c but are independent of ε, such that, for all $t\in[0,1]$,*

$$\left\|f\left(Q_t^\varepsilon, P_t^\varepsilon\right) - F\left(t, Q_t^\varepsilon, P_t^\varepsilon\right)\right\|_1 \le D_1\varepsilon^2 \qquad \textit{and} \tag{2.2.5}$$

$$\left\|g\left(Q_t^\varepsilon, P_t^\varepsilon\right) - G\left(t, Q_t^\varepsilon, P_t^\varepsilon\right)\right\|_1 \le D_2\varepsilon^2\ . \tag{2.2.6}$$

Proof. From (2.2.3) and the Hamiltonian equations (2.2.2), it follows that

$$\begin{aligned} f\left(Q_t^\varepsilon, P_t^\varepsilon\right) &= \left(P_{t,1}^\varepsilon, \left(\varphi\left(Q_t^{\varepsilon,1}\right) + Q_t^{\varepsilon,3}\right)^2 P_{t,2}^\varepsilon, P_{t,3}^\varepsilon\right) \\ &= \left(1+\varepsilon a, \left(\varphi(t+\varepsilon a(t-s)) + \varepsilon c(t-s)\right)^2 \varepsilon b, \varepsilon c\right)\ . \end{aligned}$$

Using (2.2.4) yields

$$F\left(t, Q_t^\varepsilon, P_t^\varepsilon\right) = \left(P_{t,1}^\varepsilon, \varphi^2(t) P_{t,2}^\varepsilon, P_{t,3}^\varepsilon\right) = \left(1+\varepsilon a, \varphi^2(t)\varepsilon b, \varepsilon c\right)\ ,$$

and therefore, by subtracting the two equations, we obtain

$$f\left(Q_t^\varepsilon, P_t^\varepsilon\right) - F\left(t, Q_t^\varepsilon, P_t^\varepsilon\right) = \left(0, \left((\varphi(t+\varepsilon a(t-s)) + \varepsilon c(t-s))^2 - \varphi^2(t)\right)\varepsilon b, 0\right) . \quad (2.2.7)$$

Applying Taylor's theorem with the Lagrange form of the remainder, we deduce that

$$\varphi^2(t+\varepsilon a(t-s)) = \varphi^2(t) + 2\varepsilon a(t-s)\varphi(t)\varphi'(t) + \varepsilon^2 a^2 (t-s)^2 \left((\varphi'(\eta))^2 + \varphi(\eta)\varphi''(\eta)\right)$$

for some $\eta \in (t - \varepsilon a|t-s|, t + \varepsilon a|t-s|)$. Since $\varphi\colon \mathbb{R} \to \mathbb{R}$ is smooth and as continuous functions on bounded intervals are bounded, there exist constants $A, B, C > 0$ such that, for all $t \in [0,1]$ and all $\varepsilon \in (0,1)$,

$$\left|2\varepsilon a(t-s)\varphi(t)\varphi'(t) + \varepsilon^2 a^2 (t-s)^2 \left((\varphi'(\eta))^2 + \varphi(\eta)\varphi''(\eta)\right)\right| \leq \left(2aAB + a^2\left(B^2 + AC\right)\right)\varepsilon$$

as well as

$$\left|2\varepsilon c(t-s)\varphi(t+\varepsilon a(t-s)) + \varepsilon^2 c^2 (t-s)^2\right| \leq \left(2cA + c^2\right)\varepsilon .$$

In total, we have

$$\left|(\varphi(t+\varepsilon a(t-s)) + \varepsilon c(t-s))^2 - \varphi^2(t)\right| \leq \left(2aAB + a^2\left(B^2 + AC\right) + 2cA + c^2\right)\varepsilon ,$$

which by (2.2.7) implies that, for all $t \in [0,1]$,

$$\left\|f\left(Q_t^\varepsilon, P_t^\varepsilon\right) - F\left(t, Q_t^\varepsilon, P_t^\varepsilon\right)\right\|_1 \leq D_1 \varepsilon^2$$

for some constant D_1 depending on a, b and c but which is independent of $\varepsilon \in (0,1)$. In a similar way, we compute that $g(Q_t^\varepsilon, P_t^\varepsilon) = (g^1(Q_t^\varepsilon, P_t^\varepsilon), g^2(Q_t^\varepsilon, P_t^\varepsilon), g^3(Q_t^\varepsilon, P_t^\varepsilon))$ has

$$\begin{aligned}
g^1\left(Q_t^\varepsilon, P_t^\varepsilon\right) &= -\left(\varphi\left(Q_t^{\varepsilon,1}\right) + Q_t^{\varepsilon,3}\right)\varphi'\left(Q_t^{\varepsilon,1}\right)\left(P_{t,2}^\varepsilon\right)^2 \\
&= -\left(\varphi\left(t+\varepsilon a(t-s)\right) + \varepsilon c(t-s)\right)\varphi'\left(t+\varepsilon a(t-s)\right)\varepsilon^2 b^2 , \\
g^2\left(Q_t^\varepsilon, P_t^\varepsilon\right) &= 0 , \\
g^3\left(Q_t^\varepsilon, P_t^\varepsilon\right) &= -\left(\varphi\left(Q_t^{\varepsilon,1}\right) + Q_t^{\varepsilon,3}\right)\left(P_{t,2}^\varepsilon\right)^2 = -\left(\varphi\left(t+\varepsilon a(t-s)\right) + \varepsilon c(t-s)\right)\varepsilon^2 b^2 .
\end{aligned}$$

Under the assumption that $\varepsilon \in (0,1)$, we have, for all $t \in [0,1]$,

$$\left|\varphi\left(t+\varepsilon a(t-s)\right) + \varepsilon c(t-s)\right| \leq A + c$$

and

$$\left|\left(\varphi\left(t+\varepsilon a(t-s)\right) + \varepsilon c(t-s)\right)\varphi'\left(t+\varepsilon a(t-s)\right)\right| \leq (A+c)\, B .$$

Since $G \equiv 0$, it follows that, for all $t \in [0,1]$,

$$\|g\,(Q_t^\varepsilon, P_t^\varepsilon) - G\,(t, Q_t^\varepsilon, P_t^\varepsilon)\|_1 \le D_2\varepsilon^2$$

for some constant D_2 which depends on a, b and c but is independent of $\varepsilon \in (0,1)$. □

Equipped with this lemma, we can prove our proposition.

Proof of Proposition 2.2.2. Due to the continuous dependence of the solutions to systems of ordinary differential equations on initial conditions, cf. Theorem 2.2.1, the set

$$N_1 = \{(q_t^\varepsilon, p_t^\varepsilon) : 0 \le \varepsilon \le \varepsilon_0, 0 \le t \le 1\} \subset T^*\mathbb{R}^3$$

is compact, where $(q_t^0, p_t^0) = \lambda_t$. Likewise, as an immediate consequence of (2.2.3), the set

$$N_2 = \{(Q_t^\varepsilon, P_t^\varepsilon) : 0 \le \varepsilon \le \varepsilon_0, 0 \le t \le 1\} \subset T^*\mathbb{R}^3$$

is also compact. Since $\varphi\colon \mathbb{R} \to \mathbb{R}$ is smooth, we see from the Hamiltonian equations (2.2.2) that the functions f and g are differentiable. Hence, they are locally Lipschitz on $T^*\mathbb{R}^3$, which implies that f and g are Lipschitz on the compact subset $N = N_1 \cup N_2 \subset T^*\mathbb{R}^3$. Let L_1 and L_2 denote the Lipschitz constants of the functions f and g on the compact set N with respect to the ℓ^1-norm. Using the fact that $(q_s^\varepsilon, p_s^\varepsilon) = (Q_s^\varepsilon, P_s^\varepsilon)$ as well as the estimates (2.2.5) and (2.2.6), we conclude that, for $t \in [s,1]$ and $\varepsilon < \min(1, \varepsilon_0)$,

$$\begin{aligned}
&\|(q_t^\varepsilon, p_t^\varepsilon) - (Q_t^\varepsilon, P_t^\varepsilon)\|_1 \\
&\quad = \left\|\int_s^t (f(q_r^\varepsilon, p_r^\varepsilon) - F(r, Q_r^\varepsilon, P_r^\varepsilon))\,\mathrm{d}r\right\|_1 + \left\|\int_s^t (g(q_r^\varepsilon, p_r^\varepsilon) - G(r, Q_r^\varepsilon, P_r^\varepsilon))\,\mathrm{d}r\right\|_1 \\
&\quad \le \int_s^t \|f(q_r^\varepsilon, p_r^\varepsilon) - F(r, Q_r^\varepsilon, P_r^\varepsilon)\|_1\,\mathrm{d}r + \int_s^t \|g(q_r^\varepsilon, p_r^\varepsilon) - G(r, Q_r^\varepsilon, P_r^\varepsilon)\|_1\,\mathrm{d}r \\
&\quad \le \int_s^t \|f(q_r^\varepsilon, p_r^\varepsilon) - f(Q_r^\varepsilon, P_r^\varepsilon)\|_1\,\mathrm{d}r + \int_s^t \|f(Q_r^\varepsilon, P_r^\varepsilon) - F(r, Q_r^\varepsilon, P_r^\varepsilon)\|_1\,\mathrm{d}r \\
&\qquad\quad + \int_s^t \|g(q_r^\varepsilon, p_r^\varepsilon) - g(Q_r^\varepsilon, P_r^\varepsilon)\|_1\,\mathrm{d}r + \int_s^t \|g(Q_r^\varepsilon, P_r^\varepsilon) - G(r, Q_r^\varepsilon, P_r^\varepsilon)\|_1\,\mathrm{d}r \\
&\quad \le \int_s^t (L_1 + L_2)\|(q_r^\varepsilon, p_r^\varepsilon) - (Q_r^\varepsilon, P_r^\varepsilon)\|_1\,\mathrm{d}r + (D_1 + D_2)\varepsilon^2(t-s)\,.
\end{aligned}$$

By the Gronwall inequality, it follows that, for $t \in [s,1]$,

$$\|(q_t^\varepsilon, p_t^\varepsilon) - (Q_t^\varepsilon, P_t^\varepsilon)\|_1 \le D\varepsilon^2(t-s)\,\mathrm{e}^{L(t-s)}\,,$$

where $D = D_1 + D_2$ and $L = L_1 + L_2$. Thus, there exists some constant $E > 0$, which depends on a, b and c but is independent of $\varepsilon \in (0, \min(1, \varepsilon_0))$, such that, for $t \in [s,1]$,

$$\|q_t^\varepsilon - Q_t^\varepsilon\|_1 \le E\varepsilon^2\,.$$

We deduce that, for all $k \in \{1,2,3\}$,

$$Q_t^{\varepsilon,k} - E\varepsilon^2 \le Q_t^{\varepsilon,k} - \|q_t^\varepsilon - Q_t^\varepsilon\|_1 \le q_t^{\varepsilon,k} \le Q_t^{\varepsilon,k} + \|q_t^\varepsilon - Q_t^\varepsilon\|_1 \le Q_t^{\varepsilon,k} + E\varepsilon^2 \ .$$

Subtracting $Q_t^{0,k} = \gamma_t^k = q_t^{0,k}$ from this chain of inequalities and dividing through by $\varepsilon > 0$ yields

$$\frac{Q_t^{\varepsilon,k} - Q_t^{0,k}}{\varepsilon} - E\varepsilon \le \frac{q_t^{\varepsilon,k} - q_t^{0,k}}{\varepsilon} \le \frac{Q_t^{\varepsilon,k} - Q_t^{0,k}}{\varepsilon} + E\varepsilon \ .$$

Letting ε decrease to 0 gives the desired result. □

Using Proposition 2.2.2, we compute the maps $J_{ts} \colon T^*_{\gamma_s}\mathbb{R}^3 \to T_{\gamma_t}\mathbb{R}^3$, for $0 \le s \le t \le 1$, as follows.

$$\begin{aligned} J_{ts}\xi_s &= \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \pi\psi_{t-s}\left(\lambda_s + \varepsilon\xi_s\right) = \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} q_t^\varepsilon = \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} Q_t^\varepsilon \\ &= \left(t, 0, 0, a(t-s), b\int_s^t \varphi^2(r)\,\mathrm{d}r, c(t-s)\right) \end{aligned}$$

In particular, the linear map $J_1 \colon T^*_x M \to T_y M$ is given by

$$J_1\xi_0 = J_{10}\xi_0 = \left(1, 0, 0, a, b\int_0^1 \varphi^2(r)\,\mathrm{d}r, c\right) \ . \tag{2.2.8}$$

Assume the restriction $\varphi|_{[0,1]} \colon [0,1] \to \mathbb{R}$ is non-zero. Then the expression (2.2.8) implies that J_1 is invertible. As the endpoints $x = (0,0,0)$ and $y = (1,0,0)$ are connected by the unique strongly minimal path $\gamma \in H^{x,y}$, which is the projection of a bicharacteristic, the pair (x, y) lies outside the sub-Riemannian cut locus. Hence, Theorem 2.1.1 applies and the small-time fluctuations for the bridge from x to y are characterised by the zero-mean Gaussian measure μ_γ whose covariance structure is given, for $0 \le s \le t \le 1$, by

$$\begin{aligned} J_s J_1^{-1} K_t^*(t,0,0,a,b,c) &= J_{s0} J_{10}^{-1} J_{1t}(t,0,0,a,b,c) \\ &= \left(s, 0, 0, as(1-t), \frac{b\int_0^s \varphi^2(r)\,\mathrm{d}r \int_t^1 \varphi^2(r)\,\mathrm{d}r}{\int_0^1 \varphi^2(r)\,\mathrm{d}r}, cs(1-t)\right) \ . \end{aligned}$$

Let $(B_t)_{t\in[0,1]}$ be a Brownian motion in $\mathbb{R}^3$. We observe that the measure μ_γ is the law of the Gaussian bridge

$$\left(B_t^1 - tB_1^1, \int_0^t \varphi(r)\,\mathrm{d}B_r^2 - \frac{\int_0^t \varphi^2(r)\,\mathrm{d}r}{\int_0^1 \varphi^2(r)\,\mathrm{d}r}\int_0^1 \varphi(r)\,\mathrm{d}B_r^2, B_t^3 - tB_1^3\right)_{t\in[0,1]} , \tag{2.2.9}$$

which is the Gaussian process

$$\left(B_t^1, \int_0^t \varphi(r)\,\mathrm{d}B_r^2, B_t^3\right)_{t\in[0,1]}$$

conditioned to go from 0 to 0 in time 1. Moreover, if we choose $\varphi\colon \mathbb{R} \to \mathbb{R}$ to be a bump function of the form

$$\varphi(t) = \begin{cases} \exp\left(-\frac{1}{(t-t_1)(t_2-t)}\right) & \text{if } t_1 < t < t_2 \\ 0 & \text{otherwise} \end{cases},$$

for $t_1, t_2 \in (0,1)$ with $t_1 < t_2$ fixed, then the second component

$$\int_0^t \varphi(r)\,\mathrm{d}B_r^2 - \frac{\int_0^t \varphi^2(r)\,\mathrm{d}r}{\int_0^1 \varphi^2(r)\,\mathrm{d}r} \int_0^1 \varphi(r)\,\mathrm{d}B_r^2$$

of the Gaussian bridge (2.2.9) is constant on the intervals $[0, t_1]$ and $[t_2, 1]$. It follows that the corresponding zero-mean Gaussian limit measure μ_γ is not absolutely continuous with respect to the law of a Brownian bridge in $\mathbb{R}^3$ from 0 to 0 in time 1. This shows that the small-time fluctuations for the bridge of a sub-Riemannian diffusion can indeed exhibit exotic behaviours.

Chapter 3

Small-time fluctuations for sub-Riemannian diffusion loops

We study the small-time fluctuations for diffusion processes which are conditioned by their initial and final positions, under the assumptions that the diffusivity has a sub-Riemannian structure and that the drift vector field lies in the span of the sub-Riemannian structure. In the case the endpoints agree and the generator of the diffusion process is non-elliptic at that point, the deterministic Malliavin covariance matrix is always degenerate. We identify, after a suitable rescaling, another limiting Malliavin covariance matrix which is non-degenerate, and we show that, under the same scaling, the diffusion Malliavin covariance matrices are uniformly non-degenerate. We further show that the suitably rescaled fluctuations of the diffusion loop converge to a limiting diffusion loop, which is equal in law to the loop we obtain by taking the limiting process of the unconditioned rescaled diffusion processes and condition it to return to its starting point. The generator of the unconditioned limiting rescaled diffusion process can be described in terms of the original generator.

3.1 Introduction

The small-time asymptotics of heat kernels have been extensively studied over the years, from an analytic, a geometric as well as a probabilistic point of view. Bismut [Bis84] used Malliavin calculus to perform the analysis of the heat kernel in the elliptic case and he developed a deterministic Malliavin calculus to study hypoelliptic heat kernels of Hörmander type. Following this approach, Ben Arous [BA88] found the corresponding small-time asymptotics outside the sub-Riemannian cut locus and Ben Arous [BA89] and Léandre [Léa92] studied the behaviour on the diagonal. In joint work [BAL91a, BAL91b], they also discussed the exponential decay of hypoelliptic heat kernels on the diagonal.

Recently, there has been further progress in the study of heat kernels on sub-Riemannian manifolds. Barilari, Boscain and Neel [BBN12] found estimates of the heat kernel on the cut locus using an analytic approach, and Inahama and Taniguchi [IT17] combined rough paths theory and Malliavin calculus to determine small-time full asymptotic expansions on the off-diagonal cut locus. Moreover, Bailleul, Mesnager and Norris [BMN15] studied the asymptotics of sub-Riemannian diffusion bridges outside the cut locus. We extend their analysis to the diagonal and describe the asymptotics of sub-Riemannian diffusion loops. In a suitable chart, and after a suitable rescaling, we show that the small-time diffusion loop measures have a non-degenerate limit, which we identify explicitly in terms of a certain local limit operator. Our analysis also allows us to determine the loop asymptotics under the scaling used to obtain a small-time Gaussian limit of the sub-Riemannian diffusion bridge measures in [BMN15]. In general, these asymptotics are now degenerate and need no longer be Gaussian.

Let M be a connected smooth manifold of dimension d and let a be a smooth non-negative quadratic form on the cotangent bundle T^*M. Let $\mathcal{L}$ be a second order differential operator on M with smooth coefficients, such that $\mathcal{L}1 = 0$ and such that $\mathcal{L}$ has principal symbol $a/2$. One refers to a as the diffusivity of the operator $\mathcal{L}$. We say that a has a sub-Riemannian structure if there exist $m \in \mathbb{N}$ and smooth vector fields $X_1, \dots, X_m$ on M satisfying the Hörmander condition, i.e. the vector fields together with their commutator brackets of all orders span T_yM for all $y \in M$, such that

$$a(\xi,\xi) = \sum_{i=1}^{m} \langle \xi, X_i(y) \rangle^2 \quad \text{for} \quad \xi \in T_y^*M \ .$$

Thus, we can write

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{m} X_i^2 + X_0$$

for a vector field X_0 on M, which we also assume to be smooth. Note that the vector fields $X_0, X_1, \dots, X_m$ are allowed to vanish and hence, the sub-Riemannian structure $(X_1, \dots, X_m)$ need not be of constant rank. To begin with, we impose the global condition

$$M = \mathbb{R}^d \quad \text{and} \quad X_0, X_1, \dots, X_m \in C_b^\infty\left(\mathbb{R}^d, \mathbb{R}^d\right) \ ,$$

subject to the additional constraint that $X_0(y) \in \operatorname{span}\{X_1(y), \dots, X_m(y)\}$ for all $y \in \mathbb{R}^d$. Subsequently, we follow Bailleul, Mesnager and Norris [BMN15] and insist that there exist a smooth one-form β on M with $\|a(\beta,\beta)\|_\infty < \infty$, and a locally invariant positive smooth measure $\tilde{\nu}$ on M such that, for all $f \in C^\infty(M)$,

$$\mathcal{L}f = \frac{1}{2}\operatorname{div}(a \, \mathrm{d}f) + a(\beta, \mathrm{d}f) \ . \tag{3.1.1}$$

Here the divergence is understood with respect to $\tilde{\nu}$, and the measure $\tilde{\nu}$ is said to be locally invariant for $\mathcal{L}$ if, for all smooth functions f of compact support in M, we have $\int_M a(\beta, \mathrm{d}f)\,\mathrm{d}\tilde{\nu} = 0$. If the operator $\mathcal{L}$ is of the form (3.1.1) then $X_0 = \sum_{i=1}^m \alpha_i X_i$ with

$\alpha_i = \frac{1}{2} \operatorname{div} X_i + \beta(X_i)$ and in particular, $X_0(y) \in \operatorname{span}\{X_1(y), \ldots, X_m(y)\}$ for all $y \in M$. We are interested in the associated diffusion bridge measures. Fix $x \in M$ and let $\varepsilon > 0$. If we do not assume the global condition then the diffusion process $(x_t^\varepsilon)_{t<\zeta}$ defined up to the explosion time ζ starting from x and having generator $\varepsilon\mathcal{L}$ may explode with positive probability before time 1. Though, on the event $\{\zeta > 1\}$, the process $(x_t^\varepsilon)_{t\in[0,1]}$ has a unique sub-probability law μ_ε^x on the set of continuous paths $\Omega = C([0,1], M)$. Choose a positive smooth measure ν on M, which can differ from the locally invariant positive measure $\tilde{\nu}$ on M, and let p denote the Dirichlet heat kernel for $\mathcal{L}$ with respect to ν. We can disintegrate μ_ε^x to give a unique family of probability measures $(\mu_\varepsilon^{x,y} \colon y \in M)$ on Ω such that

$$\mu_\varepsilon^x(\mathrm{d}\omega) = \int_M \mu_\varepsilon^{x,y}(\mathrm{d}\omega) p(\varepsilon, x, y) \nu(\mathrm{d}y) ,$$

with $\mu_\varepsilon^{x,y}$ supported on $\Omega^{x,y} = \{\omega \in \Omega \colon \omega_0 = x, \omega_1 = y\}$ for all $y \in M$ and where the map $y \mapsto \mu_\varepsilon^{x,y}$ is weakly continuous. Bailleul, Mesnager and Norris [BMN15] studied the small-time fluctuations of the diffusion bridge measures $\mu_\varepsilon^{x,y}$ in the limit $\varepsilon \to 0$ under the assumption that (x, y) lies outside the sub-Riemannian cut locus. Due to the latter condition, their results do not cover the diagonal case unless $\mathcal{L}$ is elliptic at x. We show how to extend their analysis in order to understand the small-time fluctuations of the diffusion loop measures $\mu_\varepsilon^{x,x}$.

As a by-product, we recover the small-time heat kernel asymptotics on the diagonal shown by Ben Arous [BA89] and Léandre [Léa92]. Even though our approach for obtaining the small-time asymptotics on the diagonal is similar to [BA89], it does not rely on the Rothschild and Stein lifting theorem, cf. [RS76]. Instead, we use the notion of an adapted chart at x, introduced by Bianchini and Stefani [BS90], which provides suitable coordinates around x. We discuss adapted charts in detail in Section 3.2. The chart Ben Arous [BA89] performed his analysis in is in fact one specific example of an adapted chart, whereas we allow for any adapted chart. In the case where the diffusivity a has a sub-Riemannian structure which is one-step bracket-generating at x, any chart around x is adapted. However, in general these charts are more complex and for instance, even if $M = \mathbb{R}^d$ there is no reason to assume that the identity map is adapted. Paoli [Pao17] successfully used adapted charts to describe the small-time asymptotics of hypoelliptic operators of Hörmander type with non-vanishing drift at a stationary point of the drift field.

To a sub-Riemannian structure $(X_1, \ldots, X_m)$ on M, we associate a linear scaling map $\delta_\varepsilon \colon \mathbb{R}^d \to \mathbb{R}^d$ in a suitable set of coordinates, which depends on the number of brackets needed to achieve each direction, and the so-called nilpotent approximations $\tilde{X}_1, \ldots, \tilde{X}_m$, which are homogeneous vector fields on $\mathbb{R}^d$. For the details see Section 3.2. The map δ_ε allows us to rescale the fluctuations of the diffusion loop to high enough orders so as to obtain a non-degenerate limit measure, and the nilpotent approximations are used to describe this limit measure. Let (U, θ) be an adapted chart around $x \in M$. Smoothly extending this chart to all of M yields a smooth map $\theta\colon M \to \mathbb{R}^d$ whose derivative $\mathrm{d}\theta_x\colon T_xM \to \mathbb{R}^d$ at x is invertible. Write $T\Omega^{0,0}$ for the set of continuous paths $v = (v_t)_{t\in[0,1]}$ in T_xM with $v_0 = v_1 = 0$. Define a rescaling map $\sigma_\varepsilon\colon \Omega^{x,x} \to T\Omega^{0,0}$ by

$$\sigma_\varepsilon(\omega)_t = (\mathrm{d}\theta_x)^{-1}\left(\delta_\varepsilon^{-1}\left(\theta(\omega_t) - \theta(x)\right)\right)$$

and let $\tilde{\mu}_\varepsilon^{x,x}$ be the pushforward measure of $\mu_\varepsilon^{x,x}$ by σ_ε, i.e. $\tilde{\mu}_\varepsilon^{x,x}$ is the unique probability measure on $T\Omega^{0,0}$ given by

$$\tilde{\mu}_\varepsilon^{x,x} = \mu_\varepsilon^{x,x} \circ \sigma_\varepsilon^{-1} .$$

Our main result concerns the weak convergence of these rescaled diffusion loop measures $\tilde{\mu}_\varepsilon^{x,x}$. To describe the limit, assuming that $\theta(x) = 0$, we consider the diffusion process $(\tilde{x}_t)_{t\geq 0}$ in $\mathbb{R}^d$ starting from 0 and having generator

$$\tilde{\mathcal{L}} = \frac{1}{2}\sum_{i=1}^m \tilde{X}_i^2 .$$

A nice cascade structure of the nilpotent approximations $\tilde{X}_1, \dots, \tilde{X}_m$ ensures that this process exists for all time. Let $\tilde{\mu}^{0,\mathbb{R}^d}$ denote the law of the diffusion process $(\tilde{x}_t)_{t\in[0,1]}$ on the set of continuous paths $\Omega(\mathbb{R}^d) = C([0,1], \mathbb{R}^d)$. By disintegrating $\tilde{\mu}^{0,\mathbb{R}^d}$, we obtain the loop measure $\tilde{\mu}^{0,0,\mathbb{R}^d}$ supported on the set $\Omega(\mathbb{R}^d)^{0,0} = \{\omega \in \Omega(\mathbb{R}^d)\colon \omega_0 = \omega_1 = 0\}$. Define a map $\rho\colon \Omega(\mathbb{R}^d)^{0,0} \to T\Omega^{0,0}$ by

$$\rho(\omega)_t = (\mathrm{d}\theta_x)^{-1}\omega_t$$

and set $\tilde{\mu}^{x,x} = \tilde{\mu}^{0,0,\mathbb{R}^d} \circ \rho^{-1}$. This is the desired limit probability measure on $T\Omega^{0,0}$.

Theorem 3.1.1 (Convergence of the rescaled diffusion bridge measures)**.** *Let M be a connected smooth manifold and fix $x \in M$. Let $\mathcal{L}$ be a second order partial differential operator on M such that, for all $f \in C^\infty(M)$,*

$$\mathcal{L}f = \frac{1}{2}\operatorname{div}(a\,\mathrm{d}f) + a(\beta, \mathrm{d}f) ,$$

for the divergence taken with respect to a locally invariant positive smooth measure, and where the smooth non-negative quadratic form a on T^*M has a sub-Riemannian structure and the smooth one-form β on M satisfies $\|a(\beta,\beta)\|_\infty < \infty$. Then the rescaled diffusion loop measures $\tilde{\mu}_\varepsilon^{x,x}$ converge weakly to the probability measure $\tilde{\mu}^{x,x}$ on $T\Omega^{0,0}$ as $\varepsilon \to 0$.

We prove this result by localising Theorem 3.1.2. As a consequence of the localisation argument, Theorem 3.1.1 remains true under the weaker assumption that the smooth vector fields giving the sub-Riemannian structure are only locally defined. The theorem below imposes an additional constraint on the map θ which ensures that we can rely on the tools of Malliavin calculus to prove it. As we see later, the existence of such a diffeomorphism θ is always guaranteed.

Theorem 3.1.2. *Fix $x \in \mathbb{R}^d$. Let $X_0, X_1, \ldots, X_m$ be smooth bounded vector fields on $\mathbb{R}^d$, with bounded derivatives of all orders, which satisfy the Hörmander condition and suppose that $X_0(y) \in \operatorname{span}\{X_1(y), \ldots, X_m(y)\}$ for all $y \in \mathbb{R}^d$. Set*

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{m} X_i^2 + X_0 \ .$$

Assume the smooth map $\theta\colon \mathbb{R}^d \to \mathbb{R}^d$ is a global diffeomorphism with bounded derivatives of all positive orders and an adapted chart at x. Then the rescaled diffusion loop measures $\tilde{\mu}_\varepsilon^{x,x}$ converge weakly to the probability measure $\tilde{\mu}^{x,x}$ on $T\Omega^{0,0}$ as $\varepsilon \to 0$.

Note the limit measures with respect to two different choices of admissible diffeomorphisms θ_1 and θ_2 are related by the Jacobian matrix of the transition map $\theta_2 \circ \theta_1^{-1}$.
The proof of Theorem 3.1.2 follows [BMN15]. The additional technical result needed in our analysis is the uniform non-degeneracy of the δ_ε-rescaled Malliavin covariance matrices. Recall that we consider Malliavin covariance matrices in the sense of Bismut and refer to what is also called the reduced Malliavin covariance matrix simply as the Malliavin covariance matrix. Under the global assumption, there exists a unique diffusion process $(x_t^\varepsilon)_{t\in[0,1]}$ starting at x and having generator $\varepsilon\mathcal{L}$. Choose $\theta\colon \mathbb{R}^d \to \mathbb{R}^d$ as in Theorem 3.1.2 and define $(\tilde{x}_t^\varepsilon)_{t\in[0,1]}$ to be the rescaled diffusion process given by

$$\tilde{x}_t^\varepsilon = \delta_\varepsilon^{-1}\left(\theta(x_t^\varepsilon) - \theta(x)\right) \ .$$

Denote the Malliavin covariance matrix of $\tilde{x}_1^\varepsilon$ by $\tilde{c}_1^\varepsilon$. We know that, for $\varepsilon > 0$, the matrix $\tilde{c}_1^\varepsilon$ is non-degenerate because the vector fields $X_1, \ldots, X_m$ satisfy the Hörmander condition. We prove that these Malliavin covariance matrices are in fact uniformly non-degenerate.

Theorem 3.1.3 (Uniform non-degeneracy of the rescaled Malliavin covariance matrices)**.** *Let $X_0, X_1, \ldots, X_m$ be smooth bounded vector fields on $\mathbb{R}^d$, with bounded derivatives of all orders, which satisfy the Hörmander condition. Suppose $X_0(y) \in \operatorname{span}\{X_1(y), \ldots, X_m(y)\}$ for all $y \in \mathbb{R}^d$. Fix $x \in \mathbb{R}^d$ and consider the diffusion operator*

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{m} X_i^2 + X_0 \ .$$

Then the rescaled Malliavin covariance matrices $\tilde{c}_1^\varepsilon$ are uniformly non-degenerate, i.e. for all $p < \infty$, we have

$$\sup_{\varepsilon\in(0,1]} \mathbb{E}\left[\left|\det\left(\tilde{c}_1^\varepsilon\right)^{-1}\right|^p\right] < \infty \ .$$

We see that the uniform non-degeneracy of the rescaled Malliavin covariance matrices $\tilde{c}_1^\varepsilon$ is a consequence of the non-degeneracy of the limiting diffusion process $(\tilde{x}_t)_{t\in[0,1]}$ with generator $\tilde{\mathcal{L}}$. The latter is implied by the nilpotent approximations $\tilde{X}_1, \ldots, \tilde{X}_m$ satisfying the Hörmander condition on $\mathbb{R}^d$, as proven in Section 3.2.

This chapter is organised as follows. In Section 3.2, we introduce the notion of an adapted chart and define the scaling operator δ_ε with which we rescale the fluctuations of the diffusion loop to obtain a non-degenerate limit. It also sets up notations for subsequent sections and proves preliminary results from which we deduce properties of the limit measure. In Section 3.3, we characterise the leading-order terms of the rescaled Malliavin covariance matrices $\tilde{c}_1^\varepsilon$ as $\varepsilon \to 0$ and use this to prove Theorem 3.1.3. Equipped with the uniform non-degeneracy result, in Section 3.4, we adapt the analysis from [BMN15] to prove Theorem 3.1.2. The approach presented is based on ideas from Azencott, Bismut and Ben Arous and relies on tools from Malliavin calculus. Finally, in Section 3.5, we employ a localisation argument to prove Theorem 3.1.1 and provide an example to illustrate the result. Moreover, we discuss the occurrence of non-Gaussian behaviour in the $\sqrt{\varepsilon}$-rescaled fluctuations of diffusion loops.

3.2 Graded structure and nilpotent approximation

We introduce the notion of an adapted chart and of an associated dilation $\delta_\varepsilon \colon \mathbb{R}^d \to \mathbb{R}^d$ which allows us to rescale the fluctuations of a diffusion loop in a way which gives rise to a non-degenerate limit as $\varepsilon \to 0$. To be able to characterise this limiting measure later, we define the nilpotent approximation of a vector field on M and show that the nilpotent approximations of a sub-Riemannian structure form a sub-Riemannian structure themselves. This section is based on Bianchini and Stefani [BS90] and Paoli [Pao17], but we made some adjustments because the drift term X_0 plays a different role in our setting. At the end, we present an example to illustrate the various constructions.

3.2.1 Graded structure induced by a sub-Riemannian structure

Let $(X_1, \dots, X_m)$ be a sub-Riemannian structure on M and fix $x \in M$. For $k \geq 1$, set

$$\mathcal{A}_k = \left\{ [X_{i_1}, [X_{i_2}, \dots, [X_{i_{k-1}}, X_{i_k}] \dots]] \colon 1 \leq i_1, \dots, i_k \leq m \right\}$$

and, for $n \geq 0$, define a subspace of the space of smooth vector fields on M by

$$C_n = \operatorname{span} \bigcup_{k=1}^{n} \mathcal{A}_k \ ,$$

with the linear combinations taken over $\mathbb{R}$. Note that $C_0 = \{0\}$. Let $C = \operatorname{Lie}\{X_1, \dots, X_m\}$ be the Lie algebra over $\mathbb{R}$ generated by the vector fields $X_1, \dots, X_m$. We observe that $C_n \subset C_{n+1}$ as well as $[C_{n_1}, C_{n_2}] \subset C_{n_1+n_2}$ for $n_1, n_2 \geq 0$ and that $\bigcup_{n \geq 0} C_n = C$. Hence, $\mathcal{C} = \{C_n\}_{n \geq 0}$ is an increasing filtration of the subalgebra C of the Lie algebra of smooth vector fields on M. Consider the subspace $C_n(x)$ of the tangent space T_xM given by

$$C_n(x) = \{X(x) \colon X \in C_n\} \ .$$

Define $d_n = \dim C_n(x)$. Since $X_1, \dots, X_m$ are assumed to satisfy the Hörmander condition, we have $\bigcup_{n \geq 0} C_n(x) = T_xM$, and it follows that

$$N = \min\{n \geq 1 \colon d_n = d\}$$

is well-defined. We call N the step of the filtration $\mathcal{C}$ at x.

Definition 3.2.1. *A chart (U, θ) around $x \in M$ is called an adapted chart to the filtration $\mathcal{C}$ at x if $\theta(x) = 0$ and, for all $n \in \{1, \dots, N\}$,*

(i) $C_n(x) = \operatorname{span}\left\{\frac{\partial}{\partial \theta^1}(x), \dots, \frac{\partial}{\partial \theta^{d_n}}(x)\right\}$, *and*

(ii) $\left(\mathrm{D}\,\theta^k\right)(x) = 0$ *for every differential operator* D *of the form*

$$\mathrm{D} = Y_1 \dots Y_n \quad \text{with} \quad Y_1, \dots, Y_n \in \{X_1, \dots, X_m\}$$

and all $k > d_n$.

Note that condition (ii) is equivalent to requiring that $(\mathrm{D}\,\theta^k)(x) = 0$ for every differential operator $\mathrm{D} \in \operatorname{span}\{Y_1 \cdots Y_j \colon Y_l \in C_{i_l} \text{ and } i_1 + \dots + i_j \leq n\}$ and all $k > d_n$. The existence of an adapted chart to the filtration $\mathcal{C}$ at x is ensured by [BS90, Corollary 3.1], which explicitly constructs such a chart by considering the integral curves of the vector fields $X_1, \dots, X_m$. However, we should keep in mind that even though being adapted at x is a local property, the germs of adapted charts at x need not coincide.

Unlike Bianchini and Stefani [BS90], we choose to construct our graded structure on $\mathbb{R}^d$ instead of on the domain U of an adapted chart, as this works better with our analysis. Define weights $w_1, \dots, w_d$ by setting $w_k = \min\{l \geq 1 \colon d_l \geq k\}$. The definition immediately implies $1 \leq w_1 \leq \dots \leq w_d = N$. Let $\delta_\varepsilon \colon \mathbb{R}^d \to \mathbb{R}^d$ be the anisotropic dilation given by

$$\delta_\varepsilon(y) = \delta_\varepsilon\left(y^1, \dots, y^k, \dots, y^d\right) = \left(\varepsilon^{w_1/2} y^1, \dots, \varepsilon^{w_k/2} y^k, \dots, \varepsilon^{w_d/2} y^d\right) \ ,$$

with $(y^1, \dots, y^d)$ Cartesian coordinates on $\mathbb{R}^d$. For a non-negative integer w, a polynomial g on $\mathbb{R}^d$ is called homogeneous of weight w if it satisfies $g \circ \delta_\varepsilon = \varepsilon^{w/2} g$. For instance, the monomial $y_1^{\alpha_1} \dots y_d^{\alpha_d}$ is homogeneous of weight $\sum_{k=1}^d \alpha_k w_k$. We denote the set of polynomials which are homogeneous of weight w by $\mathcal{P}(w)$. Note that the zero polynomial is contained in $\mathcal{P}(w)$ for all non-negative integers w. Following [BS90], the graded order $\mathcal{O}(g)$ of a polynomial g is defined by the property

$$\mathcal{O}(g) \geq i \quad \text{if and only if} \quad g \in \bigoplus_{w \geq i} \mathcal{P}(w) \ .$$

Thus, the graded order of a non-zero polynomial g is the maximal non-negative integer i such that $g \in \oplus_{w \geq i} \mathcal{P}(w)$ whereas the graded order of the zero polynomial is set to be ∞. Similarly, for a smooth function $f \in C^\infty(V)$, where $V \subset \mathbb{R}^d$ is an open neighbourhood of 0, we define its graded order $\mathcal{O}(f)$ by requiring that $\mathcal{O}(f) \geq i$ if and only if each Taylor approximation of f at 0 has graded order at least i. We see that the graded order of a smooth function is either a non-negative integer or ∞. Furthermore, for an integer a, a polynomial vector field Y on $\mathbb{R}^d$ is called homogeneous of weight a if, for all $g \in \mathcal{P}(w)$, we have $Yg \in \mathcal{P}(w-a)$. Here we set $\mathcal{P}(b) = \{0\}$ for negative integers b. The weight of a general polynomial vector field is defined to be the smallest weight of its homogeneous components. Moreover, the graded order $\mathcal{O}(\mathrm{D})$ of a differential operator D on V is given by saying that

$$\mathcal{O}(\mathrm{D}) \leq i \quad \text{if and only if} \quad \mathcal{O}(\mathrm{D}\, g) \geq \mathcal{O}(g) - i \text{ for all polynomials } g \; .$$

For example, the polynomial vector field $y^1 \frac{\partial}{\partial y^1} + (y^1)^2 \frac{\partial}{\partial y^1}$ on $\mathbb{R}^d$ has weight $-w_1$ but considered as a differential operator it has graded order 0. It also follows that the graded order of a differential operator takes values in $\mathbb{Z} \cup \{\pm\infty\}$ and that the zero differential operator has graded order $-\infty$. In the remainder, we need the notions of the weight of a polynomial vector field and the graded order of a vector field understood as a differential operator. For smooth vector fields X_1 and X_2 on V, it holds true that

$$\mathcal{O}([X_1, X_2]) \leq \mathcal{O}(X_1) + \mathcal{O}(X_2) \; . \tag{3.2.1}$$

Further observe that for any smooth vector field X on V and every integer n, there exists a unique polynomial vector field $X^{(n)}$ of weight at least n such that $\mathcal{O}(X - X^{(n)}) \leq n-1$, namely the sum of the homogeneous vector fields of weight greater than or equal to n in the formal Taylor series of X at 0.

Definition 3.2.2. *Let X be a smooth vector field on an open neighbourhood V of 0. We call $X^{(n)}$ the graded approximation of weight n of X.*

Note that $X^{(n)}$ is a polynomial vector field and hence, it can be considered as a vector field defined on all of $\mathbb{R}^d$.

3.2.2 Nilpotent approximation

Let (U, θ) be an adapted chart to the filtration induced by a sub-Riemannian structure $(X_1, \ldots, X_m)$ on M at x and set $V = \theta(U)$. Note that, for $i \in \{1, \ldots, m\}$, the pushforward vector field $\theta_* X_i$ is a vector field on V and write $\tilde{X}_i$ for the graded approximation $(\theta_* X_i)^{(1)}$ of weight 1 of $\theta_* X_i$.

Definition 3.2.3. *The polynomial vector fields $\tilde{X}_1, \ldots, \tilde{X}_m$ on $\mathbb{R}^d$ are called the nilpotent approximations of the vector fields $X_1, \ldots, X_m$ on M.*

By [BS90, Theorem 3.1], we know that $\mathcal{O}(\theta_* X_i) \leq 1$. Thus, the formal Taylor series of $\theta_* X_i$ at 0 cannot contain any homogeneous components of weight greater than or equal to two. This implies that $\tilde{X}_i$ is a homogeneous vector field of weight 1 and therefore,

$$\left(\delta_\varepsilon^{-1}\right)_* \tilde{X}_i = \varepsilon^{-1/2} \tilde{X}_i \quad \text{for all } i \in \{1, \dots, m\} .$$

Moreover, from $\mathcal{O}(\theta_* X_i - \tilde{X}_i) \leq 0$, we deduce that

$$\sqrt{\varepsilon} \left(\delta_\varepsilon^{-1}\right)_* (\theta_* X_i) \to \tilde{X}_i \quad \text{as} \quad \varepsilon \to 0 \quad \text{for all } i \in \{1, \dots, m\} .$$

This convergence holds on all of $\mathbb{R}^d$ because for $y \in \mathbb{R}^d$ fixed, we have $\delta_\varepsilon(y) \in V$ for $\varepsilon > 0$ sufficiently small.

Remark 3.2.4. The vector fields $\tilde{X}_1, \dots, \tilde{X}_m$ on $\mathbb{R}^d$ have a nice cascade structure. Since $\tilde{X}_i$, for $i \in \{1, \dots, m\}$, contains the terms of weight 1 the component $\tilde{X}_i^k$, for $k \in \{1, \dots, d\}$, does not depend on the coordinates with weight greater than or equal to w_k and depends only linearly on the coordinates with weight $w_k - 1$. □

We show below that the nilpotent approximations $\tilde{X}_1, \dots, \tilde{X}_m$ inherit the Hörmander property from the sub-Riemannian structure $(X_1, \dots, X_m)$. This result plays a crucial role in the subsequent sections as it allows us to describe the limiting measure of the rescaled fluctuations by a stochastic process whose associated Malliavin covariance matrix is non-degenerate.

Lemma 3.2.5. *Let* $\tilde{\mathcal{A}}_k(0) = \left\{ [\tilde{X}_{i_1}, [\tilde{X}_{i_2}, \dots, [\tilde{X}_{i_{k-1}}, \tilde{X}_{i_k}] \dots]](0) \colon 1 \leq i_1, \dots, i_k \leq m \right\}$. *Then*

$$\operatorname{span} \bigcup_{k=1}^{n} \tilde{\mathcal{A}}_k(0) = \operatorname{span} \left\{ \frac{\partial}{\partial y^1}(0), \dots, \frac{\partial}{\partial y^{d_n}}(0) \right\} . \tag{3.2.2}$$

Proof. We prove this lemma by induction. For the base case, note that $\mathcal{O}(\theta_* X_i - \tilde{X}_i) \leq 0$ implies $\tilde{X}_i(0) = (\theta_* X_i)(0)$. Hence, by property (i) of an adapted chart θ, we obtain

$$\operatorname{span} \tilde{\mathcal{A}}_1(0) = \operatorname{span} \left\{ \tilde{X}_1(0), \dots, \tilde{X}_m(0) \right\} = (\theta_* C_1)(0) = \operatorname{span} \left\{ \frac{\partial}{\partial y^1}(0), \dots, \frac{\partial}{\partial y^{d_1}}(0) \right\} ,$$

which proves (3.2.2) for $n = 1$. Let us now assume the result to be true for $n - 1$. Due to $\mathcal{O}(\theta_* X_i - \tilde{X}_i) \leq 0$ and using (3.2.1) as well as the bilinearity of the Lie bracket, it follows that

$$\mathcal{O}\left(\theta_* [X_{i_1}, [X_{i_2}, \dots, [X_{i_{n-1}}, X_{i_n}] \dots]] - [\tilde{X}_{i_1}, [\tilde{X}_{i_2}, \dots, [\tilde{X}_{i_{n-1}}, \tilde{X}_{i_n}] \dots]] \right) \leq n - 1 .$$

Applying the induction hypothesis, we deduce that

$$\left(\theta_*[X_{i_1},[X_{i_2},\ldots,[X_{i_{n-1}},X_{i_n}]\ldots]] - [\tilde{X}_{i_1},[\tilde{X}_{i_2},\ldots,[\tilde{X}_{i_{n-1}},\tilde{X}_{i_n}]\ldots]]\right)(0)$$
$$\in \operatorname{span}\left\{\frac{\partial}{\partial y^1}(0),\ldots,\frac{\partial}{\partial y^{d_{n-1}}}(0)\right\} = \operatorname{span}\bigcup_{k=1}^{n-1}\tilde{\mathcal{A}}_k(0) .$$

This gives

$$\operatorname{span}\left\{\frac{\partial}{\partial y^1}(0),\ldots,\frac{\partial}{\partial y^{d_n}}(0)\right\} = (\theta_* C_n)(0) \subset \operatorname{span}\bigcup_{k=1}^{n}\tilde{\mathcal{A}}_k(0)$$

and since $\mathcal{O}\left([\tilde{X}_{i_1},[\tilde{X}_{i_2},\ldots,[\tilde{X}_{i_{n-1}},\tilde{X}_{i_n}]\ldots]]\right) \le n$, the other inclusion holds as well. Thus, we have established equality, which concludes the induction step. □

The lemma allows us to prove the following proposition.

Proposition 3.2.6. *The nilpotent approximations $\tilde{X}_1,\ldots,\tilde{X}_m$ on $\mathbb{R}^d$ of the vector fields $X_1,\ldots,X_m$ on M satisfy the Hörmander condition everywhere on $\mathbb{R}^d$.*

Proof. By definition, we have $d_N = d$, and Lemma 3.2.5 implies that

$$\operatorname{span}\bigcup_{k=1}^{N}\tilde{\mathcal{A}}_k(0) = \operatorname{span}\left\{\frac{\partial}{\partial y^1}(0),\ldots,\frac{\partial}{\partial y^{d}}(0)\right\} = \mathbb{R}^d ,$$

i.e. $\tilde{X}_1,\ldots,\tilde{X}_m$ satisfy the Hörmander condition at 0. In particular, there are vector fields

$$Y_1,\ldots,Y_d \in \bigcup_{k=1}^{N}\left\{[\tilde{X}_{i_1},[\tilde{X}_{i_2},\ldots,[\tilde{X}_{i_{k-1}},\tilde{X}_{i_k}]\ldots]] \colon 1 \le i_1,\ldots,i_k \le m\right\}$$

such that $Y_1(0),\ldots,Y_d(0)$ are linearly independent, i.e. $\det(Y_1(0),\ldots,Y_d(0)) \neq 0$. By continuity of the map $y \mapsto \det(Y_1(y),\ldots,Y_d(y))$, there exists a neighbourhood V_0 of 0 on which the vector fields $\tilde{X}_1,\ldots,\tilde{X}_m$ satisfy the Hörmander condition. Since the Lie bracket commutes with the pushforward, the homogeneity property $\left(\delta_\varepsilon^{-1}\right)_* \tilde{X}_i = \varepsilon^{-1/2}\tilde{X}_i$ of the nilpotent approximations shows that the Hörmander condition is in fact satisfied on all of $\mathbb{R}^d$. □

We conclude with an example.

Example 3.2.7. Let $M = \mathbb{R}^2$ and fix $x = 0$. Let X_1 and X_2 be the vector fields on $\mathbb{R}^2$ defined by

$$X_1 = \frac{\partial}{\partial x^1} + x^1\frac{\partial}{\partial x^2} \qquad \text{and} \qquad X_2 = x^1\frac{\partial}{\partial x^1} ,$$

with respect to Cartesian coordinates (x^1, x^2) on $\mathbb{R}^2$. We compute

$$[X_1, X_2] = \frac{\partial}{\partial x^1} - x^1\frac{\partial}{\partial x^2} \qquad \text{and} \qquad [X_1,[X_1,X_2]] = -2\frac{\partial}{\partial x^2} .$$

It follows that

$$C_1(0) = C_2(0) = \operatorname{span}\left\{\frac{\partial}{\partial x^1}(0)\right\}\ ,\quad C_3(0) = \mathbb{R}^2 \quad \text{and} \quad d_1 = d_2 = 1\ ,\ \ d_3 = 2\ .$$

We note that the Cartesian coordinates are not adapted to the filtration induced by (X_1, X_2) at 0 because, for instance, $((X_1)^2\, x_2)\,(0) = 1$. Following the constructive proof of [BS90, Corollary 3.1], we find a global adapted chart $\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ at 0 given by

$$\theta^1 = x^1 \quad \text{and} \quad \theta^2 = -\frac{1}{2}(x^1)^2 + x^2\ .$$

The corresponding weights are $w_1 = 1$, $w_2 = 3$ and the associated anisotropic dilation is

$$\delta_\varepsilon(y^1, y^2) = \left(\varepsilon^{1/2} y^1, \varepsilon^{3/2} y^2\right)\ ,$$

where (y^1, y^2) are Cartesian coordinates on our new copy of $\mathbb{R}^2$. For the pushforward vector fields of X_1 and X_2 by θ, we obtain

$$\theta_* X_1 = \frac{\partial}{\partial y^1} \qquad \text{and} \qquad \theta_* X_2 = y^1 \left(\frac{\partial}{\partial y^1} - y^1 \frac{\partial}{\partial y^2}\right)\ .$$

From this we can read off that

$$\tilde{X}_1 = \frac{\partial}{\partial y^1} \qquad \text{and} \qquad \tilde{X}_2 = -\left(y^1\right)^2 \frac{\partial}{\partial y^2}$$

because $y^1 \frac{\partial}{\partial y^1}$ is a vector field of weight 0. We observe that the nilpotent approximations $\tilde{X}_1$ and $\tilde{X}_2$ are indeed homogeneous vector fields of weight 1 on $\mathbb{R}^2$ which satisfy the Hörmander condition everywhere.

3.3 Rescaled diffusion Malliavin covariance matrices

We prove the uniform non-degeneracy of suitably rescaled Malliavin covariance matrices under the global condition

$$M = \mathbb{R}^d \quad \text{and} \quad X_0, X_1, \ldots, X_m \in C_b^\infty\left(\mathbb{R}^d, \mathbb{R}^d\right)\ ,$$

and the additional assumption that $X_0(y) \in \operatorname{span}\{X_1(y), \ldots, X_m(y)\}$ for all $y \in \mathbb{R}^d$. We further suppose that $\theta\colon \mathbb{R}^d \to \mathbb{R}^d$ is a global diffeomorphism with bounded derivatives of all positive orders and an adapted chart to the filtration induced by the sub-Riemannian structure $(X_1, \ldots, X_m)$ at a point $x \in \mathbb{R}^d$ fixed. Such a diffeomorphism always exists as [BS90, Corollary 3.1] guarantees the existence of an adapted chart $\tilde{\theta}\colon U \to \mathbb{R}^d$ and due to [Pal59, Lemma 5.2], we can construct a global diffeomorphism $\theta\colon \mathbb{R}^d \to \mathbb{R}^d$ with bounded derivatives of all positive orders which agrees with $\tilde{\theta}$ on a small enough neighbourhood of x in U. We note that $\theta_* X_0, \theta_* X_1, \ldots, \theta_* X_m$ are also smooth bounded vector fields on $\mathbb{R}^d$ with bounded derivatives of all orders. In particular, to simplify the notation in the subsequent analysis, we may assume $x = 0$ and that θ is the identity map. By Section 3.2, this means that, for Cartesian coordinates $(y_1, \ldots, y_d)$ on $\mathbb{R}^d$ and for all $n \in \{1, \ldots, N\}$, we have

(i) $C_n(0) = \operatorname{span}\left\{\frac{\partial}{\partial y^1}(0), \ldots, \frac{\partial}{\partial y^{d_n}}(0)\right\}$, and

(ii) $\left(\mathrm{D}\, y^k\right)(x) = 0$ for every differential operator

$$\mathrm{D} \in \{Y_1 \cdots Y_j \colon Y_l \in C_{i_l} \text{ and } i_1 + \cdots + i_j \leq n\}$$

and all $k > d_n$.

Write $\langle \cdot, \cdot \rangle$ for the standard inner product on $\mathbb{R}^d$ and, for $n \in \{0, 1, \ldots, N\}$, denote the orthogonal complement of $C_n(0)$ with respect to this inner product by $C_n(0)^\perp$. As defined in the previous section, we further let $\delta_\varepsilon \colon \mathbb{R}^d \to \mathbb{R}^d$ be the anisotropic dilation induced by the filtration at 0 and we consider the nilpotent approximations $\tilde{X}_1, \ldots, \tilde{X}_m$ of the vector fields $X_1, \ldots, X_m$.

Let $(B_t)_{t \in [0,1]}$ be a Brownian motion in $\mathbb{R}^m$, which we assume is realised as the coordinate process on the path space $\{w \in C([0,1], \mathbb{R}^m) \colon w_0 = 0\}$ under Wiener measure $\mathbb{P}$. Define $\underline{X}_0$ to be the vector field on $\mathbb{R}^d$ given by

$$\underline{X}_0 = X_0 + \frac{1}{2} \sum_{i=1}^m \nabla_{X_i} X_i \ ,$$

where ∇ is the Levi-Civita connection with respect to the Euclidean metric. Under our global assumption, the Itô stochastic differential equation in $\mathbb{R}^d$

$$\mathrm{d}x_t^\varepsilon = \sum_{i=1}^m \sqrt{\varepsilon} X_i(x_t^\varepsilon)\, \mathrm{d}B_t^i + \varepsilon \underline{X}_0(x_t^\varepsilon)\, \mathrm{d}t \ , \quad x_0^\varepsilon = 0$$

has a unique strong solution $(x_t^\varepsilon)_{t \in [0,1]}$. Its law on $\Omega = C([0,1], \mathbb{R}^d)$ is μ_ε^0. We consider the rescaled diffusion process $(\tilde{x}_t^\varepsilon)_{t \in [0,1]}$ which is defined by $\tilde{x}_t^\varepsilon = \delta_\varepsilon^{-1}(x_t^\varepsilon)$. It is the unique strong solution of the Itô stochastic differential equation

$$\mathrm{d}\tilde{x}_t^\varepsilon = \sum_{i=1}^m \sqrt{\varepsilon} \left(\left(\delta_\varepsilon^{-1}\right)_* X_i\right)(\tilde{x}_t^\varepsilon)\, \mathrm{d}B_t^i + \varepsilon \left(\left(\delta_\varepsilon^{-1}\right)_* \underline{X}_0\right)(\tilde{x}_t^\varepsilon)\, \mathrm{d}t \ , \quad \tilde{x}_0^\varepsilon = 0 \ .$$

Let us further look at

$$\mathrm{d}\tilde{x}_t = \sum_{i=1}^m \tilde{X}_i(\tilde{x}_t)\, \mathrm{d}B_t^i + \underline{\tilde{X}}_0(\tilde{x}_t)\, \mathrm{d}t \ , \quad \tilde{x}_0 = 0 \ ,$$

where $\underline{\tilde{X}}_0$ is the vector field on $\mathbb{R}^d$ defined by

$$\underline{\tilde{X}}_0 = \frac{1}{2}\sum_{i=1}^{m} \nabla_{\tilde{X}_i}\tilde{X}_i \ .$$

Due to the nice cascade structure noted in Remark 3.2.4 and by [Nor86, Proposition 1.3], there exists a unique strong solution $(\tilde{x}_t)_{t\in[0,1]}$ to this Itô stochastic differential equation in $\mathbb{R}^d$. We recall that $\sqrt{\varepsilon}\,(\delta_\varepsilon^{-1})_* X_i \to \tilde{X}_i$ as $\varepsilon \to 0$ for all $i \in \{1,\dots,m\}$ and because $X_0(y) \in \operatorname{span}\{X_1(y),\dots,X_m(y)\}$ for all $y \in \mathbb{R}^d$, we further have $\varepsilon\,(\delta_\varepsilon^{-1})_* X_0 \to 0$ as $\varepsilon \to 0$. It follows that, for all $t \in [0,1]$,

$$\tilde{x}_t^\varepsilon \to \tilde{x}_t \text{ as } \varepsilon \to 0 \text{ almost surely and in } L^p \text{ for all } p < \infty \ . \tag{3.3.1}$$

For the Malliavin covariance matrices $\tilde{c}_1^\varepsilon$ of $\tilde{x}_1^\varepsilon$ and $\tilde{c}_1$ of $\tilde{x}_1$, we also obtain that

$$\tilde{c}_1^\varepsilon \to \tilde{c}_1 \text{ as } \varepsilon \to 0 \text{ almost surely and in } L^p \text{ for all } p < \infty \ . \tag{3.3.2}$$

Proposition 3.2.6 shows that the nilpotent approximations $\tilde{X}_1,\dots,\tilde{X}_m$ of the vector fields $X_1,\dots,X_m$ satisfy the Hörmander condition, which implies the following non-degeneracy result.

Corollary 3.3.1. *The Malliavin covariance matrix $\tilde{c}_1$ is non-degenerate, i.e. we have, for all $p < \infty$,*

$$\mathbb{E}\left[\left|\det(\tilde{c}_1)^{-1}\right|^p\right] < \infty \ .$$

In particular, the rescaled diffusion processes $(\tilde{x}_t^\varepsilon)_{t\in[0,1]}$ have a non-degenerate limiting diffusion process as $\varepsilon \to 0$. This is an important observation in establishing the uniform non-degeneracy of the rescaled Malliavin covariance matrices $\tilde{c}_1^\varepsilon$. In the following, we first gain control over the leading-order terms of $\tilde{c}_1^\varepsilon$ as $\varepsilon \to 0$, which then allows us to show that the minimal eigenvalue of $\tilde{c}_1^\varepsilon$ can be uniformly bounded below on a set of high probability. Using this property, we prove Theorem 3.1.3 at the end of the section.

3.3.1 Properties of the rescaled Malliavin covariance matrices

Let $(\tilde{v}_t^\varepsilon)_{t\in[0,1]}$ be the unique stochastic process in $\mathbb{R}^d \otimes (\mathbb{R}^d)^*$ such that $(\tilde{x}_t^\varepsilon, \tilde{v}_t^\varepsilon)_{t\in[0,1]}$ is the strong solution of the following system of Itô stochastic differential equations.

$$\begin{aligned}
\mathrm{d}\tilde{x}_t^\varepsilon &= \sum_{i=1}^{m} \sqrt{\varepsilon}\left(\left(\delta_\varepsilon^{-1}\right)_* X_i\right)(\tilde{x}_t^\varepsilon)\,\mathrm{d}B_t^i + \varepsilon\left(\left(\delta_\varepsilon^{-1}\right)_* \underline{X}_0\right)(\tilde{x}_t^\varepsilon)\,\mathrm{d}t\ , & \tilde{x}_0^\varepsilon &= 0\\
\mathrm{d}\tilde{v}_t^\varepsilon &= -\sum_{i=1}^{m} \sqrt{\varepsilon}\tilde{v}_t^\varepsilon \nabla\left(\left(\delta_\varepsilon^{-1}\right)_* X_i\right)(\tilde{x}_t^\varepsilon)\,\mathrm{d}B_t^i \\
&\quad - \varepsilon\tilde{v}_t^\varepsilon\left(\nabla\left(\left(\delta_\varepsilon^{-1}\right)_* \underline{X}_0\right) - \sum_{i=1}^{m}\left(\nabla\left(\left(\delta_\varepsilon^{-1}\right)_* X_i\right)\right)^2\right)(\tilde{x}_t^\varepsilon)\,\mathrm{d}t\ , & \tilde{v}_0^\varepsilon &= I
\end{aligned}$$

The Malliavin covariance matrix $\tilde{c}_t^\varepsilon$ of the rescaled random variable $\tilde{x}_t^\varepsilon$ is then given by

$$\tilde{c}_t^\varepsilon = \sum_{i=1}^m \int_0^t \left(\tilde{v}_s^\varepsilon \left(\sqrt{\varepsilon}\left(\delta_\varepsilon^{-1}\right)_* X_i\right)(\tilde{x}_s^\varepsilon)\right) \otimes \left(\tilde{v}_s^\varepsilon \left(\sqrt{\varepsilon}\left(\delta_\varepsilon^{-1}\right)_* X_i\right)(\tilde{x}_s^\varepsilon)\right) \mathrm{d}s \; .$$

It turns out that we obtain a more tractable expression for $\tilde{c}_t^\varepsilon$ if we write it in terms of $(x_t^\varepsilon, v_t^\varepsilon)_{t\in[0,1]}$, which is the unique strong solution of the following system of Itô stochastic differential equations.

$$\begin{aligned}
\mathrm{d}x_t^\varepsilon &= \sum_{i=1}^m \sqrt{\varepsilon} X_i(x_t^\varepsilon)\,\mathrm{d}B_t^i + \varepsilon \underline{X}_0(x_t^\varepsilon)\,\mathrm{d}t \; , & x_0^\varepsilon &= 0\\
\mathrm{d}v_t^\varepsilon &= -\sum_{i=1}^m \sqrt{\varepsilon} v_t^\varepsilon \nabla X_i(x_t^\varepsilon)\,\mathrm{d}B_t^i - \varepsilon v_t^\varepsilon \left(\nabla \underline{X}_0 - \sum_{i=1}^m (\nabla X_i)^2\right)(x_t^\varepsilon)\,\mathrm{d}t \; , & v_0^\varepsilon &= I
\end{aligned}$$

One can check that the processes $(v_t^\varepsilon)_{t\in[0,1]}$ and $(\tilde{v}_t^\varepsilon)_{t\in[0,1]}$ are related by $\tilde{v}_t^\varepsilon = \delta_\varepsilon^{-1}\, v_t^\varepsilon\, \delta_\varepsilon$, where the map δ_ε is understood as an element in $\mathbb{R}^d \otimes (\mathbb{R}^d)^*$. This implies that

$$\tilde{c}_t^\varepsilon = \sum_{i=1}^m \int_0^t \left(\sqrt{\varepsilon}\, \delta_\varepsilon^{-1}\left(v_s^\varepsilon X_i(x_s^\varepsilon)\right)\right) \otimes \left(\sqrt{\varepsilon}\, \delta_\varepsilon^{-1}\left(v_s^\varepsilon X_i(x_s^\varepsilon)\right)\right) \mathrm{d}s \; . \tag{3.3.3}$$

We are interested in gaining control over the leading-order terms of $\tilde{c}_1^\varepsilon$ as $\varepsilon \to 0$. In the corresponding analysis, we frequently use the lemma stated below.

Lemma 3.3.2. *Let Y be a smooth vector field on $\mathbb{R}^d$. Then*

$$\mathrm{d}(v_t^\varepsilon Y(x_t^\varepsilon)) = \sum_{i=1}^m \sqrt{\varepsilon} v_t^\varepsilon [X_i, Y](x_t^\varepsilon)\,\mathrm{d}B_t^i + \varepsilon v_t^\varepsilon \left([X_0, Y] + \frac{1}{2}\sum_{i=1}^m [X_i, [X_i, Y]]\right)(x_t^\varepsilon)\,\mathrm{d}t \; .$$

Proof. To prove this identity, we switch to the Stratonovich setting. The system of Stratonovich stochastic differential equations satisfied by the processes $(x_t^\varepsilon)_{t\in[0,1]}$ and $(v_t^\varepsilon)_{t\in[0,1]}$ is

$$\begin{aligned}
\partial x_t^\varepsilon &= \sum_{i=1}^m \sqrt{\varepsilon} X_i(x_t^\varepsilon)\,\partial B_t^i + \varepsilon X_0(x_t^\varepsilon)\,\mathrm{d}t \; , & x_0^\varepsilon &= 0\\
\partial v_t^\varepsilon &= -\sum_{i=1}^m \sqrt{\varepsilon} v_t^\varepsilon \nabla X_i(x_t^\varepsilon)\,\partial B_t^i - \varepsilon v_t^\varepsilon \nabla X_0(x_t^\varepsilon)\,\mathrm{d}t \; , & v_0^\varepsilon &= I.
\end{aligned}$$

By the product rule, we have

$$\partial(v_t^\varepsilon Y(x_t^\varepsilon)) = (\partial v_t^\varepsilon) Y(x_t^\varepsilon) + v_t^\varepsilon \nabla Y(x_t^\varepsilon)\,\partial x_t^\varepsilon \; .$$

Using

$$(\partial v_t^\varepsilon) Y(x_t^\varepsilon) = -\sum_{i=1}^m \sqrt{\varepsilon} v_t^\varepsilon \nabla X_i(x_t^\varepsilon) Y(x_t^\varepsilon)\,\partial B_t^i - \varepsilon v_t^\varepsilon \nabla X_0(x_t^\varepsilon) Y(x_t^\varepsilon)\,\mathrm{d}t$$

as well as

$$v_t^\varepsilon \nabla Y(x_t^\varepsilon)\,\partial x_t^\varepsilon = \sum_{i=1}^m \sqrt{\varepsilon} v_t^\varepsilon \nabla Y(x_t^\varepsilon) X_i(x_t^\varepsilon)\,\partial B_t^i + \varepsilon v_t^\varepsilon \nabla Y(x_t^\varepsilon) X_0(x_t^\varepsilon)\,\mathrm{d}t$$

yields the identity

$$\partial(v_t^\varepsilon Y(x_t^\varepsilon)) = \sum_{i=1}^m \sqrt{\varepsilon} v_t^\varepsilon [X_i, Y](x_t^\varepsilon)\,\partial B_t^i + \varepsilon v_t^\varepsilon [X_0, Y](x_t^\varepsilon)\,\mathrm{d}t\ .$$

It remains to change back to the Itô setting. We compute that, for $i \in \{1, \dots, m\}$,

$$\begin{aligned}
&\mathrm{d}\left[\sqrt{\varepsilon} v^\varepsilon [X_i, Y](x^\varepsilon), B^i\right]_t \\
&= \sum_{j=1}^m \varepsilon v_t^\varepsilon \nabla [X_i, Y](x_t^\varepsilon) X_j(x_t^\varepsilon)\,\mathrm{d}[B^j, B^i]_t - \sum_{j=1}^m \varepsilon v_t^\varepsilon \nabla X_j(x_t^\varepsilon)[X_i, Y](x_t^\varepsilon)\,\mathrm{d}[B^j, B^i]_t \\
&= \varepsilon v_t^\varepsilon \nabla [X_i, Y](x_t^\varepsilon) X_i(x_t^\varepsilon)\,\mathrm{d}t - \varepsilon v_t^\varepsilon \nabla X_i(x_t^\varepsilon)[X_i, Y](x_t^\varepsilon)\,\mathrm{d}t \\
&= \varepsilon v_t^\varepsilon [X_i, [X_i, Y]](x_t^\varepsilon)\,\mathrm{d}t
\end{aligned}$$

and the claimed result follows. □

The next lemma, which is enough for our purposes, does not provide an explicit expression for the leading-order terms of $\tilde{c}_1^\varepsilon$. However, its proof shows how one could recursively obtain these expressions if one wishes to do so. To simplify notations, we introduce $(B_t^0)_{t\in[0,1]}$ with $B_t^0 = t$.

Lemma 3.3.3. *For every $n \in \{1, \dots, N\}$, there are finite collections of vector fields*

$$\begin{aligned}
\mathcal{B}_n &= \left\{ Y^{(n,i)}_{j_1,\dots,j_k} : 1 \le k \le n, 0 \le j_1, \dots, j_k \le m, 1 \le i \le m \right\} \subset C_{n+1} \quad \textit{and} \\
\tilde{\mathcal{B}}_n &= \left\{ \tilde{Y}^{(n,i)}_{j_1,\dots,j_k} : 1 \le k \le n, 0 \le j_1, \dots, j_k \le m, 1 \le i \le m \right\} \subset C_{n+2}
\end{aligned}$$

such that, for all $u \in C_n(0)^\perp$ and all $i \in \{1, \dots, m\}$, we have that, for all $\varepsilon > 0$,

$$\begin{aligned}
&\left\langle u, \varepsilon^{-n/2} v_t^\varepsilon X_i(x_t^\varepsilon) \right\rangle \\
&= \left\langle u, \sum_{k=1}^n \sum_{j_1,\dots,j_k=0}^m \int_0^t \int_0^{t_2} \cdots \int_0^{t_k} v_s^\varepsilon \left(Y^{(n,i)}_{j_1,\dots,j_k} + \sqrt{\varepsilon}\, \tilde{Y}^{(n,i)}_{j_1,\dots,j_k} \right)(x_s^\varepsilon)\,\mathrm{d}B_s^{j_k}\,\mathrm{d}B_{t_k}^{j_{k-1}} \dots \mathrm{d}B_{t_2}^{j_1} \right\rangle .
\end{aligned}$$

Proof. We prove this result iteratively over n. For all $u \in C_1(0)^\perp$, we have $\langle u, X_i(0) \rangle = 0$ because $C_1(0) = \operatorname{span}\{X_1(0), \dots, X_m(0)\}$. From Lemma 3.3.2, it then follows that

$$\begin{aligned}
&\left\langle u, \varepsilon^{-1/2} v_t^\varepsilon X_i(x_t^\varepsilon) \right\rangle \\
&= \left\langle u, \sum_{j=1}^m \int_0^t v_s^\varepsilon [X_j, X_i](x_s^\varepsilon)\,\mathrm{d}B_s^j + \int_0^t \sqrt{\varepsilon} v_s^\varepsilon \left([X_0, X_i] + \frac{1}{2} \sum_{j=1}^m [X_j, [X_j, X_i]] \right)(x_s^\varepsilon)\,\mathrm{d}s \right\rangle .
\end{aligned}$$

This gives us the claimed result for $n = 1$ with

$$Y_j^{(1,i)} = \begin{cases} 0 & \text{if } j = 0 \\ [X_j, X_i] & \text{if } 1 \le j \le m \end{cases} \qquad \text{and}$$
$$\tilde{Y}_j^{(1,i)} = \begin{cases} [X_0, X_i] + \frac{1}{2}\sum_{l=1}^m [X_l, [X_l, X_i]] & \text{if } j = 0 \\ 0 & \text{otherwise} \end{cases} .$$

Let us now assume the result to be true for $n - 1$. Due to $C_n(0)^\perp \subset C_{n-1}(0)^\perp$, the corresponding identity also holds for all $u \in C_n(0)^\perp$. Using Lemma 3.3.2, we obtain that

$$\begin{aligned} v_s^\varepsilon Y_{j_1,\dots,j_k}^{(n-1,i)}(x_s^\varepsilon) &= Y_{j_1,\dots,j_k}^{(n-1,i)}(0) + \sum_{j=1}^m \int_0^s \sqrt{\varepsilon} v_r^\varepsilon \left[X_j, Y_{j_1,\dots,j_k}^{(n-1,i)}\right](x_r^\varepsilon)\,\mathrm{d}B_r^j \\ &\quad + \int_0^s \varepsilon v_r^\varepsilon \left(\left[X_0, Y_{j_1,\dots,j_k}^{(n-1,i)}\right] + \frac{1}{2}\sum_{j=1}^m \left[X_j, \left[X_j, Y_{j_1,\dots,j_k}^{(n-1,i)}\right]\right]\right)(x_r^\varepsilon)\,\mathrm{d}r \,. \end{aligned}$$

Note that $Y_{j_1,\dots,j_k}^{(n-1,i)} \in C_n$ implies $\langle u, Y_{j_1,\dots,j_k}^{(n-1,i)}(0)\rangle = 0$ for all $u \in C_n(0)^\perp$. We further observe that

$$\left[X_j, Y_{j_1,\dots,j_k}^{(n-1,i)}\right], \tilde{Y}_{j_1,\dots,j_k}^{(n-1,i)} \in C_{n+1} \quad \text{as well as}$$
$$\left[X_0, Y_{j_1,\dots,j_k}^{(n-1,i)}\right] + \frac{1}{2}\sum_{j=1}^m \left[X_j, \left[X_j, Y_{j_1,\dots,j_k}^{(n-1,i)}\right]\right] \in C_{n+2}$$

and collecting terms shows that the claimed result is also true for n. □

These expressions allow us to characterise the rescaled Malliavin covariance matrix $\tilde{c}_1^\varepsilon$ because, for all $n \in \{0, 1, \dots, N-1\}$ and all $u \in C_{n+1}(0) \cap C_n(0)^\perp$, we have

$$\langle u, \tilde{c}_1^\varepsilon u\rangle = \sum_{i=1}^m \int_0^1 \left\langle u, \varepsilon^{-n/2} v_t^\varepsilon X_i(x_t^\varepsilon)\right\rangle^2 \,\mathrm{d}t \,. \tag{3.3.4}$$

By the convergence result (3.3.2), it follows that, for $u \in C_1(0)$,

$$\langle u, \tilde{c}_1 u\rangle = \lim_{\varepsilon\to 0} \langle u, \tilde{c}_1^\varepsilon u\rangle = \sum_{i=1}^m \int_0^1 \langle u, X_i(0)\rangle^2 \,\mathrm{d}t$$

and Lemma 3.3.3 implies that, for all $n \in \{1, \dots, N-1\}$ and all $u \in C_{n+1}(0) \cap C_n(0)^\perp$,

$$\langle u, \tilde{c}_1 u\rangle = \sum_{i=1}^m \int_0^1 \left\langle u, \sum_{k=1}^n \sum_{j_1,\dots,j_k=0}^m \int_0^t \int_0^{t_2} \dots \int_0^{t_k} Y_{j_1,\dots,j_k}^{(n,i)}(0)\,\mathrm{d}B_s^{j_k}\,\mathrm{d}B_{t_k}^{j_{k-1}} \dots \mathrm{d}B_{t_2}^{j_1}\right\rangle^2 \mathrm{d}t \,, \tag{3.3.5}$$

which describes the limiting Malliavin covariance matrix $\tilde{c}_1$ uniquely.

3.3.2 Proving uniform non-degeneracy

By definition, the Malliavin covariance matrices $\tilde{c}_1^\varepsilon$ and $\tilde{c}_1$ are symmetric tensors. Thus, their matrix representations are symmetric in any basis and we can think of them as symmetric matrices. Let $\lambda_{\min}^\varepsilon$ and $\lambda_{\min}$ denote the minimal eigenvalues of $\tilde{c}_1^\varepsilon$ and $\tilde{c}_1$, respectively. As we frequently use the integrals from Lemma 3.3.3, it is convenient to consider the stochastic processes $(I_t^{(n,i),+})_{t\in[0,1]}$, $(I_t^{(n,i),-})_{t\in[0,1]}$ and $(\tilde{I}_t^{(n,i)})_{t\in[0,1]}$ given by

$$
\begin{aligned}
I_t^{(n,i),+} &= \sum_{k=1}^{n} \sum_{j_1,\dots,j_k=0}^{m} \int_0^t \int_0^{t_2} \cdots \int_0^{t_k} \left(v_s^\varepsilon Y_{j_1,\dots,j_k}^{(n,i)}(x_s^\varepsilon) + Y_{j_1,\dots,j_k}^{(n,i)}(0) \right) \mathrm{d}B_s^{j_k} \, \mathrm{d}B_{t_k}^{j_{k-1}} \dots \mathrm{d}B_{t_2}^{j_1} \,, \\
I_t^{(n,i),-} &= \sum_{k=1}^{n} \sum_{j_1,\dots,j_k=0}^{m} \int_0^t \int_0^{t_2} \cdots \int_0^{t_k} \left(v_s^\varepsilon Y_{j_1,\dots,j_k}^{(n,i)}(x_s^\varepsilon) - Y_{j_1,\dots,j_k}^{(n,i)}(0) \right) \mathrm{d}B_s^{j_k} \, \mathrm{d}B_{t_k}^{j_{k-1}} \dots \mathrm{d}B_{t_2}^{j_1} \,, \\
\tilde{I}_t^{(n,i)} &= \sum_{k=1}^{n} \sum_{j_1,\dots,j_k=0}^{m} \int_0^t \int_0^{t_2} \cdots \int_0^{t_k} v_s^\varepsilon \tilde{Y}_{j_1,\dots,j_k}^{(n,i)}(x_s^\varepsilon) \, \mathrm{d}B_s^{j_k} \, \mathrm{d}B_{t_k}^{j_{k-1}} \dots \mathrm{d}B_{t_2}^{j_1} \,.
\end{aligned}
$$

For $\alpha, \beta, \gamma, \delta > 0$, define subspaces of the path space $\{w \in C([0,1], \mathbb{R}^m) \colon w_0 = 0\}$ by

$$
\begin{aligned}
\Omega^1(\alpha) &= \{\lambda_{\min} \ge 2\alpha\} \,, \\
\Omega_\varepsilon^2(\beta,\gamma) &= \left\{ \sup_{0\le t\le 1} \left| I_t^{(n,i),+} \right| \le \beta^{-1} \,, \ \sup_{0\le t\le 1} \left| \tilde{I}_t^{(n,i)} \right| \le \gamma^{-1} \colon 1 \le i \le m, \ 1 \le n \le N \right\} \,, \quad \text{and} \\
\Omega_\varepsilon^3(\delta) &= \left\{ \sup_{0\le t\le 1} |x_t^\varepsilon| \le \delta \,, \ \sup_{0\le t\le 1} |v_t^\varepsilon - I| \le \delta \right\} \\
&\qquad \cup \left\{ \sup_{0\le t\le 1} \left| I_t^{(n,i),-} \right| \le \delta \colon 1 \le i \le m, \ 1 \le n \le N \right\} \,.
\end{aligned}
$$

Note that the events $\Omega_\varepsilon^2(\beta,\gamma)$ and $\Omega_\varepsilon^3(\delta)$ depend on ε as the processes $(I_t^{(n,i),+})_{t\in[0,1]}$, $(I_t^{(n,i),-})_{t\in[0,1]}$ and $(\tilde{I}_t^{(n,i)})_{t\in[0,1]}$ depend on ε. We show that, for suitable choices of α, β, γ and δ, the rescaled Malliavin covariance matrices $\tilde{c}_1^\varepsilon$ behave nicely on the set

$$
\Omega(\alpha,\beta,\gamma,\delta,\varepsilon) = \Omega^1(\alpha) \cap \Omega_\varepsilon^2(\beta,\gamma) \cap \Omega_\varepsilon^3(\delta)
$$

and that its complement is a set of small probability in the limit $\varepsilon \to 0$. As we are only interested in small values of $\alpha, \beta, \gamma, \delta$ and ε, we may make the non-restrictive assumption that $\alpha, \beta, \gamma, \delta, \varepsilon < 1$.

Lemma 3.3.4. *There exist positive constants χ and κ, which do not depend on ε, such that if*

$$
\chi \varepsilon^{1/6} \le \alpha \,, \qquad \beta = \gamma = \alpha \qquad \text{and} \qquad \delta = \kappa \alpha^2
$$

then, on $\Omega(\alpha,\beta,\gamma,\delta,\varepsilon)$, it holds true that

$$
\lambda_{\min}^\varepsilon \ge \frac{1}{2} \lambda_{\min} \,.
$$

Proof. Throughout, we shall assume that we are on the event $\Omega(\alpha,\beta,\gamma,\delta,\varepsilon)$. Let

$$R^\varepsilon(u)=\frac{\langle u,\tilde c_1^\varepsilon u\rangle}{\langle u,u\rangle}\qquad\text{and}\qquad R(u)=\frac{\langle u,\tilde c_1 u\rangle}{\langle u,u\rangle}$$

be the Rayleigh-Ritz quotients of the rescaled Malliavin covariance matrix $\tilde c_1^\varepsilon$ and of the limiting Malliavin covariance matrix $\tilde c_1$, respectively. As a consequence of the Min-Max Theorem, we have

$$\lambda_{\min}^\varepsilon=\min\{R^\varepsilon(u)\colon u\neq 0\}\qquad\text{as well as}\qquad \lambda_{\min}=\min\{R(u)\colon u\neq 0\}\ .$$

Since $\lambda_{\min}\geq 2\alpha$, it suffices to establish that $|R^\varepsilon(u)-R(u)|\leq\alpha$ for all $u\neq 0$. Set

$$K=\max_{1\leq i\leq m}\sup_{y\in\mathbb{R}^d}|X_i(y)|\ ,\qquad L=\max_{1\leq i\leq m}\sup_{y\in\mathbb{R}^d}|\nabla X_i(y)|$$

and note the global condition ensures $K,L<\infty$. Using the Cauchy-Schwarz inequality, we deduce that, for $u\in C_1(0)\setminus\{0\}$,

$$\begin{aligned}|R^\varepsilon(u)-R(u)|&\leq\frac{\displaystyle\sum_{i=1}^m\int_0^1\left|\langle u,v_t^\varepsilon X_i(x_t^\varepsilon)\rangle^2-\langle u,X_i(0)\rangle^2\right|\,\mathrm{d}t}{\langle u,u\rangle}\\&\leq\sum_{i=1}^m\int_0^1|v_t^\varepsilon X_i(x_t^\varepsilon)+X_i(0)||v_t^\varepsilon X_i(x_t^\varepsilon)-X_i(0)|\,\mathrm{d}t\\&\leq m((1+\delta)K+K)(\delta K+\delta L)\ .\end{aligned}$$

Applying Lemma 3.3.3 as well as the expressions (3.3.4) and (3.3.5), we obtain in a similar way that, for all $n\in\{1,\dots,N-1\}$ and all non-zero $u\in C_{n+1}(0)\cap C_n(0)^\perp$,

$$\begin{aligned}|R^\varepsilon(u)-R(u)|&\leq\sum_{i=1}^m\int_0^1\left|I_t^{(n,i),+}+\sqrt{\varepsilon}\tilde I_t^{(n,i)}\right|\left|I_t^{(n,i),-}+\sqrt{\varepsilon}\tilde I_t^{(n,i)}\right|\,\mathrm{d}t\\&\leq m\left(\beta^{-1}+\sqrt{\varepsilon}\gamma^{-1}\right)\left(\delta+\sqrt{\varepsilon}\gamma^{-1}\right)\ .\end{aligned}$$

It remains to perform the analysis for the cross-terms. For $n_1,n_2\in\{1,\dots,N-1\}$ as well as $u^1\in C_{n_1+1}(0)\cap C_{n_1}(0)^\perp$ and $u^2\in C_{n_2+1}(0)\cap C_{n_2}(0)^\perp$, we polarise (3.3.4) to conclude that

$$\begin{aligned}\frac{\langle u^1,\tilde c_1^\varepsilon u^2\rangle-\langle u^1,\tilde c_1 u^2\rangle}{|u^1||u^2|}&\leq\sum_{i=1}^m\int_0^1\left|\frac{I_t^{(n_1,i),+}+I_t^{(n_1,i),-}}{2}+\sqrt{\varepsilon}\tilde I_t^{(n_1,i)}\right|\left|I_t^{(n_2,i),-}+\sqrt{\varepsilon}\tilde I_t^{(n_2,i)}\right|\,\mathrm{d}t\\&\quad+\sum_{i=1}^m\int_0^1\left|I_t^{(n_1,i),-}+\sqrt{\varepsilon}\tilde I_t^{(n_1,i)}\right|\left|\frac{I_t^{(n_2,i),+}-I_t^{(n_2,i),-}}{2}\right|\,\mathrm{d}t\\&\leq m\left(\beta^{-1}+\delta+\sqrt{\varepsilon}\gamma^{-1}\right)\left(\delta+\sqrt{\varepsilon}\gamma^{-1}\right)\ .\end{aligned}$$

Similarly, if $n_1 = 0$ and $n_2 \in \{1, \dots, N-1\}$, we see that

$$\frac{\langle u^1, \tilde{c}_1^\varepsilon u^2\rangle - \langle u^1, \tilde{c}_1 u^2\rangle}{|u^1||u^2|} \le m\left((1+\delta)K\left(\delta + \sqrt{\varepsilon}\gamma^{-1}\right) + (\delta K + \delta L)\left(\frac{\beta^{-1}+\delta}{2}\right)\right) .$$

Writing a general non-zero $u \in \mathbb{R}^d$ in its orthogonal sum decomposition and combining all the above estimates gives

$$|R^\varepsilon(u) - R(u)| \le \kappa_1\delta + \kappa_2\beta^{-1}\delta + \kappa_3\sqrt{\varepsilon}\beta^{-1}\gamma^{-1} + \kappa_4\varepsilon\gamma^{-2}$$

for constants $\kappa_1, \kappa_2, \kappa_3$ and κ_4, which depend on K, L and m but are independent of $\alpha, \beta, \gamma, \delta$ and ε. If we now choose κ and χ in such a way that both $\kappa \le 1/(4\max\{\kappa_1, \kappa_2\})$ and $\chi^3 \ge 4\max\{\kappa_3, \kappa_4^{1/2}\}$, and provided that $\chi\varepsilon^{1/6} \le \alpha$, $\beta = \gamma = \alpha$ as well as $\delta = \kappa\alpha^2$, then

$$\kappa_1\delta + \kappa_2\beta^{-1}\delta + \kappa_3\sqrt{\varepsilon}\beta^{-1}\gamma^{-1} + \kappa_4\varepsilon\gamma^{-2} \le \kappa_1\kappa\alpha^2 + \kappa_2\kappa\alpha + \kappa_3\chi^{-3}\alpha + \kappa_4\chi^{-6}\alpha^4 \le \alpha .$$

Since κ and χ can always be chosen to be positive, the desired result follows. □

As a consequence of this lemma, we are able to control $\det(\tilde{c}_1^\varepsilon)^{-1}$ on the good set $\Omega(\alpha, \beta, \gamma, \delta, \varepsilon)$. This allows us to prove Theorem 3.1.3.

Proof of Theorem 3.1.3. Recall that by Proposition 3.2.6, the nilpotent approximations $\tilde{X}_1, \dots, \tilde{X}_m$ do satisfy the Hörmander condition everywhere on $\mathbb{R}^d$. Then the proof of [Nor86, Theorem 4.2] shows that

$$\lambda_{\min}^{-1} \in L^p(\mathbb{P}) , \quad \text{for all} \quad p < \infty . \tag{3.3.6}$$

By the Markov inequality, this integrability result implies that, for all $p < \infty$, there exist constants $D(p) < \infty$ such that

$$\mathbb{P}\left(\Omega^1(\alpha)^c\right) \le D(p)\alpha^p. \tag{3.3.7}$$

Using the Burkholder-Davis-Gundy inequality and Jensen's inequality, we further show that, for all $p < \infty$, there are constants $E_1(p), E_2(p) < \infty$ such that

$$\mathbb{E}\left[\sup_{0\le t\le 1} |x_t^\varepsilon|^p\right] \le E_1(p)\varepsilon^{p/2} \qquad \text{and} \qquad \mathbb{E}\left[\sup_{0\le t\le 1} |v_t^\varepsilon - I|^p\right] \le E_2(p)\varepsilon^{p/2} .$$

Similarly, by repeatedly applying the Burkholder-Davis-Gundy inequality and Jensen's inequality, we also see that, for all $p < \infty$ and for all $n \in \{1, \dots, N\}$ and $i \in \{1, \dots, m\}$, there exist constants $E^{(n,i)}(p) < \infty$ and $D^{(n,i)}(p), \tilde{D}^{(n,i)}(p) < \infty$ such that

$$\mathbb{E}\left[\sup_{0\le t\le 1} \left|I_t^{(n,i),-}\right|^p\right] \le E^{(n,i)}(p)\varepsilon^{p/2}$$

as well as

$$\mathbb{E}\left[\sup_{0\le t\le 1}\left|I_t^{(n,i),+}\right|^p\right]\le D^{(n,i)}(p)\qquad\text{and}\qquad\mathbb{E}\left[\sup_{0\le t\le 1}\left|\tilde{I}_t^{(n,i)}\right|^p\right]\le \tilde{D}^{(n,i)}(p)\ .$$

As the sets $\Omega^2_\varepsilon(\beta,\gamma)$ and $\Omega^3_\varepsilon(\delta)$ are defined by only finitely many constraints, the bounds established above and the Markov inequality imply that, for all $p<\infty$, there are constants $D(p)<\infty$ and $E(p)<\infty$ such that

$$\mathbb{P}\left(\Omega^2_\varepsilon(\beta,\gamma)^c\right)\le D(p)\left(\beta^p+\gamma^p\right)\qquad\text{and}\tag{3.3.8}$$

$$\mathbb{P}\left(\Omega^3_\varepsilon(\delta)^c\right)\le E(p)\delta^{-p}\varepsilon^{p/2}\ .\tag{3.3.9}$$

Moreover, from the Kusuoka-Stroock estimate, cf. [AKS93], as stated by Watanabe [Wat87, Theorem 3.2], we know that there exist a positive integer S and, for all $p<\infty$, constants $C(p)<\infty$ such that, for all $\varepsilon\in(0,1]$,

$$\|\det(\tilde{c}_1^\varepsilon)^{-1}\|_p=\left(\mathbb{E}\left[\left|\det\left(\tilde{c}_1^\varepsilon\right)^{-1}\right|^p\right]\right)^{1/p}\le C(p)\varepsilon^{-S/2}\ .$$

Let us now choose $\alpha=\chi^{3/4}\varepsilon^{1/8}$, $\beta=\gamma=\alpha$ and $\delta=\kappa\alpha^2$. We note that $\chi\varepsilon^{1/6}=\alpha^{4/3}\le\alpha$ and hence, from Lemma 3.3.4 it follows that

$$\lambda^\varepsilon_{\min}\ge\frac{1}{2}\lambda_{\min}$$

on $\Omega(\alpha,\beta,\gamma,\delta,\varepsilon)$. Thus, we have

$$\det(\tilde{c}_1^\varepsilon)^{-1}\mathbb{1}_{\Omega(\alpha,\beta,\gamma,\delta,\varepsilon)}\le\left(\lambda^\varepsilon_{\min}\right)^{-d}\mathbb{1}_{\Omega(\alpha,\beta,\gamma,\delta,\varepsilon)}\le 2^d\lambda_{\min}^{-d}\mathbb{1}_{\Omega(\alpha,\beta,\gamma,\delta,\varepsilon)}$$

and therefore,

$$\det(\tilde{c}_1^\varepsilon)^{-1}\le 2^d\lambda_{\min}^{-d}+\det(\tilde{c}_1^\varepsilon)^{-1}\left(\mathbb{1}_{\Omega^1(\alpha)^c}+\mathbb{1}_{\Omega^2_\varepsilon(\beta,\gamma)^c}+\mathbb{1}_{\Omega^3_\varepsilon(\delta)^c}\right)\ .$$

Using the Hölder inequality, the Kusuoka-Stroock estimate as well as the estimates (3.3.7), (3.3.8) and (3.3.9), we further deduce that, for all $q,r<\infty$,

$$\begin{aligned}&\|\det(\tilde{c}_1^\varepsilon)^{-1}\|_p\\&\quad\le 2^d\|\lambda_{\min}^{-1}\|_p^d+C(2p)\varepsilon^{-S/2}\left(\mathbb{P}\left(\Omega^1(\alpha)^c\right)^{1/2p}+\mathbb{P}\left(\Omega^2_\varepsilon(\beta,\gamma)^c\right)^{1/2p}+\mathbb{P}\left(\Omega^3_\varepsilon(\delta)^c\right)^{1/2p}\right)\\&\quad\le 2^d\|\lambda_{\min}^{-1}\|_p^d+C(2p)\varepsilon^{-S/2}\left(\left(D(q)\alpha^q\right)^{1/2p}+\left(E(r)\delta^{-r}\varepsilon^{r/2}\right)^{1/2p}\right).\end{aligned}$$

Hence, we would like to choose q and r in such a way that we can control both $\varepsilon^{-S/2}\alpha^{q/2p}$ and $\varepsilon^{-S/2}\delta^{-r/2p}\varepsilon^{r/4p}$. Since $\delta=\kappa\alpha^2$ and $\alpha=\chi^{3/4}\varepsilon^{1/8}$, we have

$$\varepsilon^{-S/2}\alpha^{q/2p}=\chi^{3q/8p}\varepsilon^{-S/2+q/16p}\quad\text{as well as}\quad\varepsilon^{-S/2}\delta^{-r/2p}\varepsilon^{r/4p}=\left(\kappa\chi^{3/2}\right)^{-r/2p}\varepsilon^{-S/2+r/8p}\ .$$

Thus, picking $q = 8pS$ and $r = 4pS$ ensures both terms remain bounded as $\varepsilon \to 0$ and we obtain

$$\|\det(\tilde{c}_1^\varepsilon)^{-1}\|_p \leq 2^d \|\lambda_{\min}^{-1}\|_p^d + C(2p)\left(D(8pS,\chi)^{1/2p} + E(4pS,\kappa,\chi)^{1/2p}\right) .$$

This together with the integrability (3.3.6) of $\lambda_{\min}^{-1}$ implies the uniform non-degeneracy of the rescaled Malliavin covariance matrices $\tilde{c}_1^\varepsilon$. □

3.4 Convergence of the diffusion bridge measures

We give the proof of Theorem 3.1.2 in this section with the extension to Theorem 3.1.1 left to Section 3.5. For our analysis, we adapt the Fourier transform argument presented in [BMN15] to allow for the higher-order scaling δ_ε. As in Section 3.3, we may assume that the sub-Riemannian structure $(X_1, \dots, X_m)$ has already been pushed forward by the global diffeomorphism $\theta \colon \mathbb{R}^d \to \mathbb{R}^d$ which is an adapted chart at $x = 0$ and which has bounded derivatives of all positive orders.

Define $T\Omega^0$ to be the set of continuous paths $v = (v_t)_{t\in[0,1]}$ in $T_0\mathbb{R}^d \cong \mathbb{R}^d$ with $v_0 = 0$ and set

$$T\Omega^{0,y} = \{v \in T\Omega^0 \colon v_1 = y\} .$$

Let $\tilde{\mu}_\varepsilon^0$ denote the law of the rescaled process $(\tilde{x}_t^\varepsilon)_{t\in[0,1]}$ on $T\Omega^0$ and write $q(\varepsilon, 0, \cdot)$ for the law of v_1 under the measure $\tilde{\mu}_\varepsilon^0$. To obtain the rescaled diffusion bridge measures, we disintegrate $\tilde{\mu}_\varepsilon^0$ uniquely, with respect to the Lebesgue measure on $\mathbb{R}^d$, as

$$\tilde{\mu}_\varepsilon^0(\mathrm{d}v) = \int_{\mathbb{R}^d} \tilde{\mu}_\varepsilon^{0,y}(\mathrm{d}v) q(\varepsilon, 0, y)\,\mathrm{d}y , \tag{3.4.1}$$

where $\tilde{\mu}_\varepsilon^{0,y}$ is a probability measure on $T\Omega^0$ which is supported on $T\Omega^{0,y}$, and the map $y \mapsto \tilde{\mu}_\varepsilon^{0,y}$ is weakly continuous. We can think of $\tilde{\mu}_\varepsilon^{0,y}$ as the law of the process $(\tilde{x}_t^\varepsilon)_{t\in[0,1]}$ conditioned by $\tilde{x}_1^\varepsilon = y$. In particular, this construction is consistent with our previous definition of $\tilde{\mu}_\varepsilon^{0,0}$. Similarly, write $\tilde{\mu}^0$ for the law of the limiting rescaled diffusion process $(\tilde{x}_t)_{t\in[0,1]}$ on $T\Omega^0$, denote the law of v_1 under $\tilde{\mu}^0$ by $\bar{q}(\cdot)$ and let $(\tilde{\mu}^{0,y} \colon y \in \mathbb{R}^d)$ be the unique family of probability measures we obtain by disintegrating the measure $\tilde{\mu}^0$ as

$$\tilde{\mu}^0(\mathrm{d}v) = \int_{\mathbb{R}^d} \tilde{\mu}^{0,y}(\mathrm{d}v) \bar{q}(y)\,\mathrm{d}y . \tag{3.4.2}$$

In order to keep track of the paths of the diffusion bridges, we fix $t_1, \dots, t_k \in (0,1)$ with $t_1 < \cdots < t_k$ as well as a smooth function g on $(\mathbb{R}^d)^k$ of polynomial growth, and consider the smooth cylindrical function G on $T\Omega^0$ defined by $G(v) = g(v_{t_1}, \dots, v_{t_k})$. For $y \in \mathbb{R}^d$

and $\varepsilon > 0$, set

$$G_\varepsilon(y) = q(\varepsilon, 0, y) \int_{T\Omega^{0,y}} G(v)\tilde{\mu}_\varepsilon^{0,y}(\mathrm{d}v) \qquad \text{and}$$
$$G_0(y) = \bar{q}(y) \int_{T\Omega^{0,y}} G(v)\tilde{\mu}^{0,y}(\mathrm{d}v) \ .$$

Both functions are continuous integrable functions on $\mathbb{R}^d$ and in particular, we can consider their Fourier transforms $\hat{G}_\varepsilon(\xi)$ and $\hat{G}_0(\xi)$ given by

$$\hat{G}_\varepsilon(\xi) = \int_{\mathbb{R}^d} G_\varepsilon(y)\, \mathrm{e}^{\mathrm{i}\langle \xi, y\rangle}\ \mathrm{d}y \qquad \text{and} \qquad \hat{G}_0(\xi) = \int_{\mathbb{R}^d} G_0(y)\, \mathrm{e}^{\mathrm{i}\langle \xi, y\rangle}\ \mathrm{d}y \ .$$

Using the disintegration of measure property (3.4.1), we deduce that

$$\begin{aligned}
\hat{G}_\varepsilon(\xi) &= \int_{\mathbb{R}^d} \int_{T\Omega^{0,y}} q(\varepsilon, 0, y) G(v) \tilde{\mu}_\varepsilon^{0,y}(\mathrm{d}v)\, \mathrm{e}^{\mathrm{i}\langle \xi, y\rangle}\ \mathrm{d}y \\
&= \int_{T\Omega^0} G(v)\, \mathrm{e}^{\mathrm{i}\langle \xi, v_1\rangle}\ \tilde{\mu}_\varepsilon^0(\mathrm{d}v) \\
&= \mathbb{E}\left[G(\tilde{x}^\varepsilon) \exp\left\{\mathrm{i}\langle \xi, \tilde{x}_1^\varepsilon\rangle\right\}\right] \ .
\end{aligned}$$

Similarly, by using (3.4.2), we show that

$$\hat{G}_0(\xi) = \mathbb{E}\left[G(\tilde{x}) \exp\left\{\mathrm{i}\langle \xi, \tilde{x}_1\rangle\right\}\right] \ .$$

We recall that $\tilde{x}_t^\varepsilon \to \tilde{x}_t$ as $\varepsilon \to 0$ almost surely and in L^p for all $p < \infty$, which implies that $\hat{G}_\varepsilon(\xi) \to \hat{G}_0(\xi)$ as $\varepsilon \to 0$ for all $\xi \in \mathbb{R}^d$. To be able to use this convergence result to make deductions about the behaviour of the functions G_ε and G_0 we need $\hat{G}_\varepsilon$ to be integrable uniformly in $\varepsilon \in (0, 1]$. This is provided by the following lemma, which is proven at the end of the section.

Lemma 3.4.1. *For all smooth cylindrical functions G on the path space $T\Omega^0$ there are constants $C(G) < \infty$ such that, for all $\varepsilon \in (0, 1]$ and all $\xi \in \mathbb{R}^d$, we have*

$$|\hat{G}_\varepsilon(\xi)| \leq \frac{C(G)}{1 + |\xi|^{d+1}} \ . \tag{3.4.3}$$

Moreover, in the case where $G(v) = |v_{t_1} - v_{t_2}|^4$, there exists a constant $C < \infty$ such that, uniformly in $t_1, t_2 \in (0, 1)$, we can choose $C(G) = C|t_1 - t_2|^2$, i.e. for all $\varepsilon \in (0, 1]$ and all $\xi \in \mathbb{R}^d$,

$$|\hat{G}_\varepsilon(\xi)| \leq \frac{C|t_1 - t_2|^2}{1 + |\xi|^{d+1}} \ . \tag{3.4.4}$$

With this setup, we can prove Theorem 3.1.2.

Proof of Theorem 3.1.2. Applying the Fourier inversion formula and using (3.4.3) from

Lemma 3.4.1 as well as the dominated convergence theorem, we deduce that

$$G_\varepsilon(0) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \hat{G}_\varepsilon(\xi)\,\mathrm{d}\xi \to \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \hat{G}_0(\xi)\,\mathrm{d}\xi = G_0(0) \quad \text{as} \quad \varepsilon \to 0\,. \tag{3.4.5}$$

Let $Q = \sum_{n=1}^N n d_n$ be the homogeneous dimension of the sub-Riemannian structure $(X_1, \dots, X_m)$. Due to the change of variables formula, we have

$$q(\varepsilon, 0, y) = \varepsilon^{Q/2} p(\varepsilon, 0, \delta_\varepsilon(y))\,,$$

where p and q are the Dirichlet heat kernels, with respect to the Lebesgue measure on $\mathbb{R}^d$, associated to the processes $(x_t^1)_{t\in[0,1]}$ and $(\tilde{x}_t^1)_{t\in[0,1]}$, respectively. From (3.4.5), it follows that

$$\varepsilon^{Q/2} p(\varepsilon, 0, 0)\int_{T\Omega^{0,0}} G(v)\tilde{\mu}_\varepsilon^{0,0}(\mathrm{d}v) \to \bar{q}(0)\int_{T\Omega^{0,0}} G(v)\tilde{\mu}^{0,0}(\mathrm{d}v) \quad \text{as} \quad \varepsilon \to 0\,. \tag{3.4.6}$$

Choosing $g \equiv 1$ shows that

$$\varepsilon^{Q/2} p(\varepsilon, 0, 0) \to \bar{q}(0) \quad \text{as} \quad \varepsilon \to 0\,, \tag{3.4.7}$$

which agrees with the small-time heat kernel asymptotics established in [BA89] and [Léa92]. We recall that $\bar{q}\colon \mathbb{R}^d \to [0,\infty)$ is the density of the random variable $\tilde{x}_1$, where $(\tilde{x}_t)_{t\in[0,1]}$ is the limiting rescaled process with generator

$$\tilde{\mathcal{L}} = \frac{1}{2}\sum_{i=1}^m \tilde{X}_i^2\,.$$

By Proposition 3.2.6, the nilpotent approximations $\tilde{X}_1, \dots, \tilde{X}_m$ satisfy the Hörmander condition everywhere on $\mathbb{R}^d$ and since $\tilde{\mathcal{L}}$ has vanishing drift, the discussions in [BAL91b] imply that $\bar{q}(0) > 0$. Hence, we can divide (3.4.6) by (3.4.7) to obtain

$$\int_{T\Omega^{0,0}} G(v)\tilde{\mu}_\varepsilon^{0,0}(\mathrm{d}v) \to \int_{T\Omega^{0,0}} G(v)\tilde{\mu}^{0,0}(\mathrm{d}v) \quad \text{as} \quad \varepsilon \to 0\,.$$

Thus, the finite-dimensional distributions of $\tilde{\mu}_\varepsilon^{0,0}$ converge weakly to those of $\tilde{\mu}^{0,0}$ and it remains to establish tightness in order to deduce the desired convergence result. Taking $G(v) = |v_{t_1} - v_{t_2}|^4$ as well as using the Fourier inversion formula and the estimate (3.4.4) from Lemma 3.4.1, we conclude that

$$\varepsilon^{Q/2} p(\varepsilon, 0, 0)\int_{T\Omega^{0,0}} |v_{t_1} - v_{t_2}|^4\,\tilde{\mu}_\varepsilon^{0,0}(\mathrm{d}v) = G_\varepsilon(0) \le C|t_1 - t_2|^2\,.$$

From (3.4.7) and due to $\bar{q}(0) > 0$, it further follows that there exists a constant $D < \infty$

such that, for all $t_1, t_2 \in (0,1)$,

$$\sup_{\varepsilon\in(0,1]} \int_{T\Omega^{0,0}} |v_{t_1} - v_{t_2}|^4 \, \tilde{\mu}_\varepsilon^{0,0}(\mathrm{d}v) \leq D|t_1 - t_2|^2 \, .$$

Standard arguments finally imply that the family of laws $(\tilde{\mu}_\varepsilon^{0,0} \colon \varepsilon \in (0,1])$ is tight on $T\Omega^{0,0}$ and hence, $\tilde{\mu}_\varepsilon^{0,0} \to \tilde{\mu}^{0,0}$ weakly on $T\Omega^{0,0}$ as $\varepsilon \to 0$. □

It remains to prove Lemma 3.4.1. We closely follow [BMN15, Proof of Lemma 4.1], where the main adjustments needed arise due to the higher-order scaling map δ_ε. In addition to the uniform non-degeneracy of the rescaled Malliavin covariance matrices $\tilde{c}_1^\varepsilon$, which is provided by Theorem 3.1.3, we need the rescaled processes $(\tilde{x}_t^\varepsilon)_{t\in[0,1]}$ and $(\tilde{v}_t^\varepsilon)_{t\in[0,1]}$ defined in Section 3.3.1 to have moments of all orders bounded uniformly in $\varepsilon \in (0,1]$. The latter is ensured by the following lemma.

Lemma 3.4.2. *There are moment estimates of all orders for the stochastic processes $(\tilde{x}_t^\varepsilon)_{t\in[0,1]}$ and $(\tilde{v}_t^\varepsilon)_{t\in[0,1]}$ which are uniform in $\varepsilon \in (0,1]$, i.e. for all $p < \infty$, we have*

$$\sup_{\varepsilon\in(0,1]} \mathbb{E}\left[\sup_{0\leq t\leq 1} |\tilde{x}_t^\varepsilon|^p\right] < \infty \qquad \textit{and} \qquad \sup_{\varepsilon\in(0,1]} \mathbb{E}\left[\sup_{0\leq t\leq 1} |\tilde{v}_t^\varepsilon|^p\right] < \infty \, .$$

Proof. We exploit the graded structure induced by the sub-Riemannian structure $(X_1, \dots, X_m)$ and we make use of the properties of an adapted chart. For $\tau \in [0,1]$, consider the Itô stochastic differential equation in $\mathbb{R}^d$

$$\mathrm{d}x_t^\varepsilon(\tau) = \sum_{i=1}^m \tau\sqrt{\varepsilon} X_i(x_t^\varepsilon(\tau)) \, \mathrm{d}B_t^i + \tau^2 \varepsilon \underline{X}_0(x_t^\varepsilon(\tau)) \, \mathrm{d}t \, , \quad x_0^\varepsilon(\tau) = 0$$

and let $\{(x_t^\varepsilon(\tau))_{t\in[0,1]} \colon \tau \in [0,1]\}$ be the unique family of strong solutions which is almost surely jointly continuous in τ and t. Observe that $x_t^\varepsilon(0) = 0$ and $x_t^\varepsilon(1) = x_t^\varepsilon$ for all $t \in [0,1]$, almost surely. Moreover, for $n \geq 1$, the rescaled nth derivative in τ

$$x_t^{\varepsilon,(n)}(\tau) = \varepsilon^{-n/2} \left(\frac{\partial}{\partial \tau}\right)^n x_t^\varepsilon(\tau)$$

exists for all τ and t, almost surely. For instance, $(x_t^{\varepsilon,(1)}(\tau))_{t\in[0,1]}$ is the unique strong solution of the following Itô stochastic differential equation subject to $x_0^{\varepsilon,(1)}(\tau) = 0$.

$$\begin{aligned}
\mathrm{d}x_t^{\varepsilon,(1)}(\tau) = &\sum_{i=1}^m X_i(x_t^\varepsilon(\tau)) \, \mathrm{d}B_t^i + 2\tau\sqrt{\varepsilon} \underline{X}_0(x_t^\varepsilon(\tau)) \, \mathrm{d}t \\
&+ \sum_{i=1}^m \tau\sqrt{\varepsilon} \nabla X_i(x_t^\varepsilon(\tau)) x_t^{\varepsilon,(1)}(\tau) \, \mathrm{d}B_t^i + \tau^2 \varepsilon \nabla \underline{X}_0(x_t^\varepsilon(\tau)) x_t^{\varepsilon,(1)}(\tau) \, \mathrm{d}t
\end{aligned}$$

In particular, we compute that $x_t^{\varepsilon,(1)}(0) = \sum_{i=1}^m X_i(0) B_t^i$. As $\langle u, X_i(0)\rangle = 0$ holds true for all $i \in \{1, \dots, m\}$ and all $u \in C_1(0)^\perp$, we deduce

$$\left\langle u, x_t^{\varepsilon,(1)}(0) \right\rangle = 0 \quad \text{for all} \quad u \in C_1(0)^\perp . \tag{3.4.8}$$

By looking at the corresponding stochastic differential equation for $(x_t^{\varepsilon,(2)}(\tau))_{t\in[0,1]}$, we further obtain that

$$x_t^{\varepsilon,(2)}(0) = \sum_{i=1}^m \int_0^t 2\nabla X_i(0) x_s^{\varepsilon,(1)}(0) \, \mathrm{d}B_s^i + 2\underline{X}_0(0)t .$$

Due to (3.4.8), the only non-zero terms in $\nabla X_i(0) x_s^{\varepsilon,(1)}(0)$ are scalar multiples of the first d_1 columns of $\nabla X_i(0)$, i.e. where the derivative is taken along a direction lying in $C_1(0)$. Thus, by property (ii) of an adapted chart and since $X_0(0) \in \operatorname{span}\{X_1(0), \dots, X_m(0)\}$, it follows that

$$\left\langle u, x_t^{\varepsilon,(2)}(0) \right\rangle = 0 \quad \text{for all} \quad u \in C_2(0)^\perp .$$

In general, continuing in the same way and by appealing to the Faà di Bruno formula, we prove iteratively that, for all $n \in \{1, \dots, N-1\}$,

$$\left\langle u, x_t^{\varepsilon,(n)}(0) \right\rangle = 0 \quad \text{for all} \quad u \in C_n(0)^\perp . \tag{3.4.9}$$

Besides, let us consider the stochastic process $(x_t^\varepsilon(\tau), x_t^{\varepsilon,(1)}(\tau), \dots, x_t^{\varepsilon,(N)}(\tau))_{t\in[0,1]}$. It is the solution of a stochastic differential equation with graded Lipschitz coefficients in the sense of Norris [Nor86]. As the coefficient bounds of the graded structure are uniform in $\tau \in [0,1]$ and $\varepsilon \in (0,1]$, we obtain, uniformly in τ and ε, moment bounds of all orders for $(x_t^\varepsilon(\tau), x_t^{\varepsilon,(1)}(\tau), \dots, x_t^{\varepsilon,(N)}(\tau))_{t\in[0,1]}$. Finally, due to (3.4.9) we have, for all $n \in \{1, \dots, N\}$ and all $u \in C_n(0) \cap C_{n-1}(0)^\perp$,

$$\langle u, \tilde{x}_t^\varepsilon \rangle = \left\langle u, \varepsilon^{-n/2} x_t^\varepsilon \right\rangle = \left\langle u, \int_0^1 \int_0^{\tau_1} \dots \int_0^{\tau_{n-1}} x_t^{\varepsilon,(n)}(\tau_n) \, \mathrm{d}\tau_n \, \mathrm{d}\tau_{n-1} \dots \mathrm{d}\tau_1 \right\rangle .$$

This together with the uniform moment bounds implies the claimed result that, for all $p < \infty$,

$$\sup_{\varepsilon\in(0,1]} \mathbb{E}\left[\sup_{0\le t\le 1} |\tilde{x}_t^\varepsilon|^p \right] < \infty .$$

We proceed similarly to establish the second estimate. Let $\{(v_t^\varepsilon(\tau))_{t\in[0,1]} \colon \tau \in [0,1]\}$ be the unique family of strong solutions to the Itô stochastic differential equation in $\mathbb{R}^d$

$$\begin{aligned} \mathrm{d}v_t^\varepsilon(\tau) = &- \sum_{i=1}^m \tau\sqrt{\varepsilon} v_t^\varepsilon(\tau) \nabla X_i(x_t^\varepsilon(\tau)) \, \mathrm{d}B_t^i \\ &- \tau^2 \varepsilon v_t^\varepsilon(\tau) \left(\nabla \underline{X}_0 - \sum_{i=1}^m (\nabla X_i)^2 \right) (x_t^\varepsilon(\tau)) \, \mathrm{d}t , \quad v_0^\varepsilon(\tau) = I \end{aligned}$$

which is almost surely jointly continuous in τ and t. We note that $v_t^\varepsilon(0) = I$ and $v_t^\varepsilon(1) = v_t^\varepsilon$ for all $t \in [0,1]$, almost surely. For $n \geq 1$, the derivative

$$v_t^{\varepsilon,(n)}(\tau) = \varepsilon^{-n/2} \left(\frac{\partial}{\partial \tau}\right)^n v_t^\varepsilon(\tau)$$

exists for all τ and t, almost surely. For $n_1, n_2 \in \{1, \dots, N\}$ and $u^1 \in C_{n_1}(0) \cap C_{n_1-1}(0)^\perp$ as well as $u^2 \in C_{n_2}(0) \cap C_{n_2-1}(0)^\perp$, we have

$$\left\langle u^1, \tilde{v}_t^\varepsilon u^2 \right\rangle = \varepsilon^{-(n_1-n_2)/2} \left\langle u^1, v_t^\varepsilon u^2 \right\rangle \ .$$

Therefore, if $n_1 \leq n_2$, we obtain the bound $|\langle u^1, \tilde{v}_t^\varepsilon u^2\rangle| \leq |\langle u^1, v_t^\varepsilon u^2\rangle|$. On the other hand, if $n_1 > n_2$ then $\langle u^1, u^2 \rangle = 0$ and in a similar way to proving (3.4.9), we show that

$$\left\langle u^1, v_t^{\varepsilon,(k)}(0) u^2 \right\rangle = 0 \quad \text{for all} \quad k \in \{1, \dots, n_1 - n_2 - 1\}$$

by repeatedly using property (ii) of an adapted chart. This allows us to write, for $n_1 > n_2$,

$$\left\langle u^1, \tilde{v}_t^\varepsilon u^2 \right\rangle = \left\langle u^1, \left(\int_0^1 \int_0^{\tau_1} \dots \int_0^{\tau_{n_1-n_2-1}} v_t^{\varepsilon,(n_1-n_2)}(\tau_{n_1-n_2}) \, \mathrm{d}\tau_{n_1-n_2} \, \mathrm{d}\tau_{n_1-n_2-1} \dots \mathrm{d}\tau_1 \right) u^2 \right\rangle .$$

As the stochastic process $(x_t^\varepsilon(\tau), v_t^\varepsilon(\tau), x_t^{\varepsilon,(1)}(\tau), v_t^{\varepsilon,(1)}(\tau), \dots, x_t^{\varepsilon,(N)}(\tau), v_t^{\varepsilon,(N)}(\tau))_{t\in[0,1]}$ is the solution of a stochastic differential equation with graded Lipschitz coefficients in the sense of Norris [Nor86], with the coefficient bounds of the graded structure being uniform in $\tau \in [0,1]$ and $\varepsilon \in (0,1]$, the second result claimed follows. □

We finally present the proof of Lemma 3.4.1. For some of the technical arguments which carry over unchanged, we simply refer the reader to [BMN15].

Proof of Lemma 3.4.1. Let $(x_t^\varepsilon)_{t\in[0,1]}$ be the process in $\mathbb{R}^d$ and $(u_t^\varepsilon)_{t\in[0,1]}$ as well as $(v_t^\varepsilon)_{t\in[0,1]}$ be the processes in $\mathbb{R}^d \otimes (\mathbb{R}^d)^*$ which are defined as the unique strong solutions of the following system of Itô stochastic differential equations.

$$\begin{aligned}
\mathrm{d}x_t^\varepsilon &= \sum_{i=1}^m \sqrt{\varepsilon} X_i(x_t^\varepsilon) \, \mathrm{d}B_t^i + \varepsilon \underline{X}_0(x_t^\varepsilon) \, \mathrm{d}t \ , & x_0^\varepsilon &= 0 \qquad (3.4.10)\\
\mathrm{d}u_t^\varepsilon &= \sum_{i=1}^m \sqrt{\varepsilon} \nabla X_i(x_t^\varepsilon) u_t^\varepsilon \, \mathrm{d}B_t^i + \varepsilon \nabla \underline{X}_0(x_t^\varepsilon) u_t^\varepsilon \, \mathrm{d}t \ , & u_0^\varepsilon &= I\\
\mathrm{d}v_t^\varepsilon &= -\sum_{i=1}^m \sqrt{\varepsilon} v_t^\varepsilon \nabla X_i(x_t^\varepsilon) \, \mathrm{d}B_t^i - \varepsilon v_t^\varepsilon \left(\nabla \underline{X}_0 - \sum_{i=1}^m (\nabla X_i)^2 \right)(x_t^\varepsilon) \, \mathrm{d}t \ , & v_0^\varepsilon &= I
\end{aligned}$$

Fix $k \in \{1, \dots, d\}$. For $\eta \in \mathbb{R}^d$, consider the perturbed process $(B_t^\eta)_{t\in[0,1]}$ in $\mathbb{R}^m$ given by

$$\mathrm{d}B_t^{\eta,i} = \mathrm{d}B_t^i + \eta \left(\sqrt{\varepsilon} \, \delta_\varepsilon^{-1} \left(v_t^\varepsilon X_i(x_t^\varepsilon) \right) \right)^k \, \mathrm{d}t \ , \quad B_0^\eta = 0 \ ,$$

where $(\sqrt{\varepsilon}\,\delta_\varepsilon^{-1}(v_t^\varepsilon X_i(x_t^\varepsilon)))^k$ denotes the kth component of the vector $\sqrt{\varepsilon}\,\delta_\varepsilon^{-1}(v_t^\varepsilon X_i(x_t^\varepsilon))$ in $\mathbb{R}^d$. Write $(x_t^{\varepsilon,\eta})_{t\in[0,1]}$ for the strong solution of the stochastic differential equation (3.4.10) with the driving Brownian motion $(B_t)_{t\in[0,1]}$ replaced by $(B_t^\eta)_{t\in[0,1]}$. We choose a version of the family of processes $(x_t^{\varepsilon,\eta})_{t\in[0,1]}$ which is almost surely smooth in η and set

$$((x^\varepsilon)'_t)^k = \left.\frac{\partial}{\partial\eta}\right|_{\eta=0} x_t^{\varepsilon,\eta}\ .$$

The derived process $((x^\varepsilon)'_t)_{t\in[0,1]} = (((x^\varepsilon)'_t)^1,\dots,((x^\varepsilon)'_t)^d)_{t\in[0,1]}$ in $\mathbb{R}^d\otimes\mathbb{R}^d$ associated with the process $(x_t^\varepsilon)_{t\in[0,1]}$ then satisfies the Itô stochastic differential equation

$$\begin{aligned}\mathrm{d}(x^\varepsilon)'_t = \sum_{i=1}^m \sqrt{\varepsilon}\nabla X_i(x_t^\varepsilon)(x^\varepsilon)'_t\,\mathrm{d}B_t^i + \varepsilon\nabla\underline{X}_0(x_t^\varepsilon)(x^\varepsilon)'_t\,\mathrm{d}t&\\ + \sum_{i=1}^m \sqrt{\varepsilon}X_i(x_t^\varepsilon)\otimes\left(\sqrt{\varepsilon}\,\delta_\varepsilon^{-1}\left(v_t^\varepsilon X_i(x_t^\varepsilon)\right)\right)\mathrm{d}t&\end{aligned}$$

subject to $(x^\varepsilon)'_0 = 0$. Using the expression (3.3.3) for the rescaled Malliavin covariance matrix $\tilde{c}_t^\varepsilon$, we show that $(x^\varepsilon)'_t = u_t^\varepsilon\,\delta_\varepsilon\,\tilde{c}_t^\varepsilon$. It follows that for the derived process $((\tilde{x}^\varepsilon)'_t)_{t\in[0,1]}$ associated with the rescaled process $(\tilde{x}_t^\varepsilon)_{t\in[0,1]}$ and the stochastic process $(\tilde{u}_t^\varepsilon)_{t\in[0,1]}$ given by $\tilde{u}_t^\varepsilon = \delta_\varepsilon^{-1}\,u_t^\varepsilon\,\delta_\varepsilon$, we have

$$(\tilde{x}^\varepsilon)'_t = \tilde{u}_t^\varepsilon\tilde{c}_t^\varepsilon\ .$$

Note that both $\tilde{u}_1^\varepsilon$ and $\tilde{c}_1^\varepsilon$ are invertible for all $\varepsilon > 0$ with $(\tilde{u}_1^\varepsilon)^{-1} = \tilde{v}_1^\varepsilon$. Let $(r_t^\varepsilon)_{t\in[0,1]}$ be the process defined by

$$\mathrm{d}r_t^\varepsilon = \sum_{i=1}^m \sqrt{\varepsilon}\,\delta_\varepsilon^{-1}\left(v_t^\varepsilon X_i(x_t^\varepsilon)\right)\mathrm{d}B_t^i\ ,\quad r_0^\varepsilon = 0$$

and set

$$y_t^{\varepsilon,(0)} = \left(x_{t\wedge t_1}^\varepsilon,\dots,x_{t\wedge t_k}^\varepsilon,x_t^\varepsilon,v_t^\varepsilon,r_t^\varepsilon,(x^\varepsilon)'_t\right)\ .$$

The underlying graded Lipschitz structure, in the sense of Norris [Nor86], allows us, for $n\geq 0$, to recursively define

$$z_t^{\varepsilon,(n)} = \left(y_t^{\varepsilon,(0)},\dots,y_t^{\varepsilon,(n)}\right)$$

by first solving for the derived process $((z^{\varepsilon,(n)})'_t)_{t\in[0,1]}$, then writing

$$\left(z^{\varepsilon,(n)}\right)'_t = \left(\left(y^{\varepsilon,(0)}\right)'_t,\dots,\left(y^{\varepsilon,(n)}\right)'_t\right)$$

and finally setting $y_t^{\varepsilon,(n+1)} = (y^{\varepsilon,(n)})'_t$.

Consider the random variable $y^\varepsilon = ((\tilde{x}^\varepsilon)'_1)^{-1}$ in $(\mathbb{R}^d)^*\otimes(\mathbb{R}^d)^*$ and let $\phi = \phi(y^\varepsilon, z_1^{\varepsilon,(n)})$ be a polynomial in y^ε, where the coefficients are continuously differentiable in $z_1^{\varepsilon,(n)}$ and of polynomial growth, along with their derivatives. Going through the deductions made from Bismut's integration by parts formula in [BMN15, Proof of Lemma 4.1] with $R \equiv 0$ and $F \equiv 0$ shows that for any continuously differentiable, bounded function $f\colon \mathbb{R}^d \to \mathbb{R}$ with bounded first derivatives and any $k \in \{1, \dots, d\}$, we have

$$\mathbb{E}\left[\nabla_k f(\tilde{x}_1^\varepsilon)\phi\left(y^\varepsilon, z_1^{\varepsilon,(n)}\right)\right] = \mathbb{E}\left[f(\tilde{x}_1^\varepsilon)\nabla_k^*\phi\left(y^\varepsilon, z_1^{\varepsilon,(n+1)}\right)\right] ,$$

where

$$\begin{aligned}\nabla_k^*\phi&\left(y^\varepsilon, z_1^{\varepsilon,(n+1)}\right)\\ &= \tau_k\left(y^\varepsilon \otimes r_1^\varepsilon + y^\varepsilon(\tilde{x}^\varepsilon)_1'' y^\varepsilon\right)\phi\left(y^\varepsilon, z_1^{\varepsilon,(n)}\right) + \tau_k\left(y^\varepsilon \otimes \left(\nabla_y\phi\left(y^\varepsilon, z_1^{\varepsilon,(n)}\right) y^\varepsilon(\tilde{x}^\varepsilon)_1'' y^\varepsilon\right)\right)\\ &\qquad - \tau_k\left(y^\varepsilon \otimes \left(\nabla_z\phi\left(y^\varepsilon, z_1^{\varepsilon,(n)}\right)\left(z^{\varepsilon,(n)}\right)_1'\right)\right) ,\end{aligned}$$

and $\tau_k\colon (\mathbb{R}^d)^* \otimes (\mathbb{R}^d)^* \otimes \mathbb{R}^d \to \mathbb{R}$ is the linear map given by $\tau_k(e_l^* \otimes e_{k'}^* \otimes e_{l'}) = \delta_{kk'}\delta_{ll'}$. Starting from

$$\phi\left(y^\varepsilon, z_1^{\varepsilon,(0)}\right) = G(\tilde{x}^\varepsilon) = g\left(\tilde{x}_{t_1}^\varepsilon, \dots, \tilde{x}_{t_k}^\varepsilon\right)$$

we see inductively that, for any multi-index $\alpha = (k_1, \dots, k_n)$,

$$\mathbb{E}\left[\nabla^\alpha f(\tilde{x}_1^\varepsilon)G(\tilde{x}^\varepsilon)\right] = \mathbb{E}\left[f(\tilde{x}_1^\varepsilon)(\nabla^*)^\alpha G\left(y^\varepsilon, z_1^{\varepsilon,(n)}\right)\right] .$$

Fixing $\xi \in \mathbb{R}^d$ and choosing $f(\cdot) = \mathrm{e}^{\mathrm{i}\langle\xi,\cdot\rangle}$ in this integration by parts formula yields

$$|\xi^\alpha||\hat{G}_\varepsilon(\xi)| \leq \mathbb{E}\left[\left|(\nabla^*)^\alpha G\left(y^\varepsilon, z_1^{\varepsilon,(n)}\right)\right|\right] .$$

To deduce the bound (3.4.3), it remains to establish that $C_\varepsilon(\alpha, G) = \mathbb{E}[|(\nabla^*)^\alpha G(y^\varepsilon, z_1^{\varepsilon,(n)})|]$ can be controlled uniformly in ε. Due to $y^\varepsilon = (\tilde{c}_1^\varepsilon)^{-1}\tilde{v}_1^\varepsilon$, Theorem 3.1.3 and the second estimate from Lemma 3.4.2 immediately imply that, for all $p < \infty$,

$$\sup_{\varepsilon\in(0,1]} \mathbb{E}\left[|y^\varepsilon|^p\right] < \infty . \tag{3.4.11}$$

Moreover, from the first moment estimate in Lemma 3.4.2, it follows that all processes derived from the rescaled process $(\tilde{x}_t^\varepsilon)_{t\in[0,1]}$ have moments of all orders bounded uniformly in $\varepsilon \in (0, 1]$. Similarly, for $n = d + 1$ and all $p < \infty$, we obtain

$$\sup_{\varepsilon\in(0,1]} \mathbb{E}\left[\left|z_1^{\varepsilon,(n)}\right|^p\right] < \infty , \tag{3.4.12}$$

where we observe that, for all $n \in \{0, 1, \dots, N-1\}$ and all $u \in C_{n+1}(0) \cap C_n(0)^\perp$,

$$\langle u, r_t^\varepsilon\rangle = \sum_{i=1}^m \int_0^t \left\langle u, \varepsilon^{-n/2} v_s^\varepsilon X_i(x_s^\varepsilon)\right\rangle \mathrm{d}B_s^i ,$$

and use Lemma 3.3.3 to show that there is no singularity in the process $(r_t^\varepsilon)_{t\in[0,1]}$ as $\varepsilon \to 0$. Since $(\nabla^*)^\alpha G$ is of polynomial growth in the argument $(y^\varepsilon, z_1^{\varepsilon,(n)})$, the moment estimates (3.4.11) and (3.4.12) show that $C_\varepsilon(\alpha, G)$ is bounded uniformly in $\varepsilon \in (0,1]$. This establishes (3.4.3).

Finally, the same proof as presented in [BMN15, Proof of Lemma 4.1] shows that we have (3.4.4) in the special case where $G(v) = |v_{t_1} - v_{t_2}|^4$ for some $t_1, t_2 \in (0,1)$. Let the process $(\tilde{x}_t^{\varepsilon,(0)})_{t\in[0,1]}$ be given by $\tilde{x}_t^{\varepsilon,(0)} = \tilde{x}_t^\varepsilon$ and, recursively for $n \geq 0$, define $(\tilde{x}_t^{\varepsilon,(n+1)})_{t\in[0,1]}$ by $\tilde{x}_t^{\varepsilon,(n+1)} = (\tilde{x}_t^\varepsilon, (\tilde{x}^{\varepsilon,(n)})_t')$. Then, for all $p \in [1,\infty)$, there exists a constant $D(p) < \infty$ such that, uniformly in $t_1, t_2 \in (0,1)$ and in $\varepsilon \in (0,1]$,

$$\mathbb{E}\left[\left|\tilde{x}_{t_1}^{\varepsilon,(n)} - \tilde{x}_{t_2}^{\varepsilon,(n)}\right|^{4p}\right] \leq D(p)|t_1 - t_2|^{2p} .$$

Furthermore, from the expression for the adjoint operator ∇_k^* we deduce that, for all $n \geq 1$ and any multi-index $\alpha = (k_1, \ldots, k_n)$, there exists a random variable M_α, with moments of all orders which are bounded uniformly in $\varepsilon \in (0,1]$, such that

$$(\nabla^*)^\alpha G\left(y^\varepsilon, z_1^{\varepsilon,(n)}\right) = M_\alpha \left|\tilde{x}_{t_1}^{\varepsilon,(n)} - \tilde{x}_{t_2}^{\varepsilon,(n)}\right|^4 .$$

By using Hölder's inequality, we conclude that there exists a constant $C(\alpha) < \infty$ such that, uniformly in $t_1, t_2 \in (0,1)$ and $\varepsilon \in (0,1]$, we obtain

$$C_\varepsilon(\alpha, G) \leq C(\alpha)|t_1 - t_2|^2 ,$$

which implies (3.4.4). □

3.5 Localisation argument

In proving Theorem 3.1.1 by localising Theorem 3.1.2, we employ the same localisation argument as used in [BMN15, Section 5]. This is possible due to [BMN15, Theorem 6.1], which provides a control over the amount of heat diffusing between two fixed points on a manifold without leaving a fixed closed subset, also covering the diagonal case. After the proof, we give an example to illustrate Theorem 3.1.1 and we remark on deductions made for the $\sqrt{\varepsilon}$-rescaled fluctuations of diffusion loops.

Let $\mathcal{L}$ be a differential operator on M satisfying the conditions of Theorem 3.1.1 and let $(X_1, \ldots, X_m)$ be a sub-Riemannian structure for the diffusivity of $\mathcal{L}$. Define X_0 to be the smooth vector field on M given by requiring

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^m X_i^2 + X_0$$

and recall that $X_0(y) \in \operatorname{span}\{X_1(y), \ldots, X_m(y)\}$ for all $y \in M$. Let (U_0, θ) be an adapted chart to the filtration induced by $(X_1, \ldots, X_m)$ at $x \in M$ and extend it to a smooth map $\theta \colon M \to \mathbb{R}^d$. By passing to a smaller set if necessary, we may assume that the closure of U_0 is compact. Let U be a domain in M containing x and compactly contained in U_0. We start by constructing a differential operator $\bar{\mathcal{L}}$ on $\mathbb{R}^d$ which satisfies the assumptions of Theorem 3.1.2 with the identity map being an adapted chart at 0 and such that $\mathcal{L}(f) = \bar{\mathcal{L}}(f \circ \theta^{-1}) \circ \theta$ for all $f \in C^\infty(U)$.

Set $V = \theta(U)$ and $V_0 = \theta(U_0)$. Let χ be a smooth function on $\mathbb{R}^d$ which satisfies the condition $\mathbb{1}_V \leq \chi \leq \mathbb{1}$ and where $\{\chi > 0\}$ is compactly contained in V_0. The existence of such a function is always guaranteed. Besides, we pick another smooth function ρ on $\mathbb{R}^d$ with $\mathbb{1}_V \leq \mathbb{1} - \rho \leq \mathbb{1}_{V_0}$ and such that $\chi + \rho$ is everywhere positive. Define vector fields $\bar{X}_0, \bar{X}_1, \ldots, \bar{X}_m, \bar{X}_{m+1}, \ldots, \bar{X}_{m+d}$ on $\mathbb{R}^d$ by

$$\begin{aligned} \bar{X}_i(z) &= \begin{cases} \chi(z)\, \mathrm{d}\theta_{\theta^{-1}(z)}\left(X_i\left(\theta^{-1}(z)\right)\right) & \text{if } z \in V_0 \\ 0 & \text{if } z \in \mathbb{R}^d \setminus V_0 \end{cases} && \text{for } i \in \{0, 1, \ldots, m\}\,, \\ \bar{X}_{m+k}(z) &= \rho(z) e_k && \text{for } k \in \{1, \ldots, d\}\,, \end{aligned}$$

where $e_1, \ldots, e_d$ is the standard basis in $\mathbb{R}^d$. Note that $X_0(y) \in \operatorname{span}\{X_1(y), \ldots, X_m(y)\}$ for all $y \in M$ implies that $\bar{X}_0(z) \in \operatorname{span}\{\bar{X}_1(z), \ldots, \bar{X}_m(z)\}$ for all $z \in \mathbb{R}^d$. Moreover, the vector fields $\bar{X}_1, \ldots, \bar{X}_m$ satisfy the Hörmander condition on the set $\{\chi > 0\}$, while $\bar{X}_{m+1}, \ldots, \bar{X}_{m+d}$ themselves span $\mathbb{R}^d$ on $\{\rho > 0\}$. As U_0 is assumed to have compact closure, the vector fields constructed are all bounded with bounded derivatives of all orders. Hence, the differential operator $\bar{\mathcal{L}}$ on $\mathbb{R}^d$ given by

$$\bar{\mathcal{L}} = \frac{1}{2} \sum_{i=1}^{m+d} \bar{X}_i^2 + \bar{X}_0$$

satisfies the assumptions of Theorem 3.1.2. We further observe that, on V,

$$\bar{X}_i = \theta_*(X_i) \quad \text{for all } i \in \{0, 1, \ldots, m\}\,,$$

which yields the the desired property that $\bar{\mathcal{L}} = \theta_* \mathcal{L}$ on V. Additionally, we see that the nilpotent approximations of $(\bar{X}_1, \ldots, \bar{X}_m, \bar{X}_{m+1}, \ldots, \bar{X}_{m+d})$ are $(\tilde{X}_1, \ldots, \tilde{X}_m, 0, \ldots, 0)$ which shows that the limiting rescaled processes on $\mathbb{R}^d$ associated to the processes with generator $\varepsilon\bar{\mathcal{L}}$ and $\varepsilon\mathcal{L}$, respectively, have the same generator $\tilde{\mathcal{L}}$. Since (U_0, θ), and so in particular the restriction (U, θ) is an adapted chart at x, it also follows that the identity map on $\mathbb{R}^d$ is an adapted chart to the filtration induced by the sub-Riemannian structure $(\bar{X}_1, \ldots, \bar{X}_m, \bar{X}_{m+1}, \ldots, \bar{X}_{m+d})$ on $\mathbb{R}^d$ at 0. Thus, Theorem 3.1.2 holds with the identity map as choice for the global diffeomorphism and we associate the same anisotropic dilation $\delta_\varepsilon \colon \mathbb{R}^d \to \mathbb{R}^d$ with the adapted charts (U, θ) at x and (V, I) at 0. We use this to finally prove our main result.

Proof of Theorem 3.1.1. Let $\bar{p}$ be the Dirichlet heat kernel for $\bar{\mathcal{L}}$ with respect to the Lebesgue measure λ on $\mathbb{R}^d$. Choose a positive smooth measure ν on M which satisfies $\nu = (\theta^{-1})_*\lambda$ on U and let p denote the Dirichlet heat kernel for $\mathcal{L}$ with respect to ν. Write $\mu_\varepsilon^{0,0,\mathbb{R}^d}$ for the diffusion loop measure on $\Omega^{0,0}(\mathbb{R}^d)$ associated with the operator $\varepsilon\bar{\mathcal{L}}$ and write $\tilde{\mu}_\varepsilon^{0,0,\mathbb{R}^d}$ for the rescaled loop measure on $T\Omega^{0,0}(\mathbb{R}^d)$, which is the image measure of $\mu_\varepsilon^{0,0,\mathbb{R}^d}$ under the scaling map $\bar{\sigma}_\varepsilon \colon \Omega^{0,0}(\mathbb{R}^d) \to T\Omega^{0,0}(\mathbb{R}^d)$ given by

$$\bar{\sigma}_\varepsilon(\omega)_t = \delta_\varepsilon^{-1}(\omega_t) \ .$$

Moreover, let $\tilde{\mu}^{0,0,\mathbb{R}^d}$ be the loop measure on $T\Omega^{0,0}(\mathbb{R}^d)$ associated with the stochastic process $(\tilde{x}_t)_{t\in[0,1]}$ on $\mathbb{R}^d$ starting from 0 and having generator $\tilde{\mathcal{L}}$ and let $\bar{q}$ denote the probability density function of $\tilde{x}_1$. From Theorem 3.1.2, we know that $\tilde{\mu}_\varepsilon^{0,0,\mathbb{R}^d}$ converges weakly to $\tilde{\mu}^{0,0,\mathbb{R}^d}$ on $T\Omega^{0,0}(\mathbb{R}^d)$ as $\varepsilon \to 0$, and its proof also shows that

$$\bar{p}(\varepsilon, 0, 0) = \varepsilon^{-Q/2}\bar{q}(0)(1 + o(1)) \quad \text{as} \quad \varepsilon \to 0 \ . \tag{3.5.1}$$

Let p_U denote the Dirichlet heat kernel in U of the restriction of $\mathcal{L}$ to U and write $\mu_\varepsilon^{x,x,U}$ for the diffusion bridge measure on $\Omega^{x,x}(U)$ associated with the restriction of the operator $\varepsilon\mathcal{L}$ to U. For any measurable set $A \subset \Omega^{x,x}(M)$, we have

$$p(\varepsilon, x, x)\mu_\varepsilon^{x,x}(A) = p_U(\varepsilon, x, x)\mu_\varepsilon^{x,x,U}(A \cap \Omega^{x,x}(U)) + p(\varepsilon, x, x)\mu_\varepsilon^{x,x}(A \setminus \Omega^{x,x}(U)) \ . \tag{3.5.2}$$

Additionally, by counting paths and since $\nu = (\theta^{-1})_*\lambda$ on U, we obtain

$$\bar{p}(\varepsilon, 0, 0)\mu_\varepsilon^{0,0,\mathbb{R}^d}\left(\theta(A \cap \Omega^{x,x}(U))\right) = p_U(\varepsilon, x, x)\mu_\varepsilon^{x,x,U}\left(A \cap \Omega^{x,x}(U)\right) \ , \tag{3.5.3}$$

where $\theta(A\cap\Omega^{x,x}(U))$ denotes the subset $\{(\theta(\omega_t))_{t\in[0,1]} \colon \omega \in A\cap\Omega^{x,x}(U)\}$ of $\Omega^{0,0}(\mathbb{R}^d)$. Let B be a bounded measurable subset of the set $T\Omega^{x,x}(M)$ of continuous paths $v = (v_t)_{t\in[0,1]}$ in T_xM with $v_0 = 0$ and $v_1 = 0$. For $\varepsilon > 0$ sufficiently small, we have $\sigma_\varepsilon^{-1}(B) \subset \Omega^{x,x}(U)$ and so (3.5.2) and (3.5.3) imply that

$$p(\varepsilon, x, x)\mu_\varepsilon^{x,x}\left(\sigma_\varepsilon^{-1}(B)\right) = \bar{p}(\varepsilon, 0, 0)\mu_\varepsilon^{0,0,\mathbb{R}^d}\left(\theta\left(\sigma_\varepsilon^{-1}(B)\right)\right) \ .$$

Since $\mu_\varepsilon^{x,x}(\sigma_\varepsilon^{-1}(B)) = \tilde{\mu}_\varepsilon^{x,x}(B)$ and

$$\mu_\varepsilon^{0,0,\mathbb{R}^d}\left(\theta\left(\sigma_\varepsilon^{-1}(B)\right)\right) = \mu_\varepsilon^{0,0,\mathbb{R}^d}\left(\bar{\sigma}_\varepsilon^{-1}(\mathrm{d}\theta_x(B))\right) = \tilde{\mu}_\varepsilon^{0,0,\mathbb{R}^d}(\mathrm{d}\theta_x(B)) \ ,$$

we established that

$$p(\varepsilon, x, x)\tilde{\mu}_\varepsilon^{x,x}(B) = \bar{p}(\varepsilon, 0, 0)\tilde{\mu}_\varepsilon^{0,0,\mathbb{R}^d}(\mathrm{d}\theta_x(B)) \ . \tag{3.5.4}$$

Moreover, it holds that $\mu_\varepsilon^{0,0,\mathbb{R}^d}(\theta(\Omega^{x,x}(U)) \to 1$ as $\varepsilon \to 0$. Therefore, taking $A = \Omega^{x,x}(M)$

in (3.5.3) and using (3.5.1) gives

$$p_U(\varepsilon, x, x) = \varepsilon^{-Q/2}\bar{q}(0)(1 + o(1)) \quad \text{as} \quad \varepsilon \to 0 .$$

By [BMN15, Theorem 6.1], we know that

$$\limsup_{\varepsilon\to 0} \varepsilon \log p(\varepsilon, x, M \setminus U, x) \le -\frac{d(x, M \setminus U, x)^2}{2} ,$$

where $p(\varepsilon, x, M \setminus U, x) = p(\varepsilon, x, x) - p_U(\varepsilon, x, x)$ and $d(x, M \setminus U, x)$ is the sub-Riemannian distance from x to x through $M \setminus U$. Since $d(x, M \setminus U, x)$ is strictly positive, it follows that

$$p(\varepsilon, x, x) = p_U(\varepsilon, x, x) + p(\varepsilon, x, M \setminus U, x) = \varepsilon^{-Q/2}\bar{q}(0)(1 + o(1)) \quad \text{as} \quad \varepsilon \to 0 .$$

Hence, due to (3.5.4), we have that $\tilde{\mu}_\varepsilon^{x,x}(B) = \tilde{\mu}_\varepsilon^{0,0,\mathbb{R}^d}(\mathrm{d}\theta_x(B))(1 + o(1))$ for any bounded measurable set $B \subset T\Omega^{x,x}(M)$. From the weak convergence of $\tilde{\mu}_\varepsilon^{0,0,\mathbb{R}^d}$ to $\tilde{\mu}^{0,0,\mathbb{R}^d}$ on the space $T\Omega^{0,0}(\mathbb{R}^d)$ as $\varepsilon \to 0$ and since $\tilde{\mu}^{0,0,\mathbb{R}^d}(\mathrm{d}\theta_x(B)) = \tilde{\mu}^{x,x}(B)$, we conclude that the diffusion loop measures $\tilde{\mu}_\varepsilon^{x,x}$ converge weakly to the loop measure $\tilde{\mu}^{x,x}$ on $T\Omega^{0,0}(M)$ as $\varepsilon \to 0$. □

We close with an example and a remark.

Example 3.5.1. Consider the same setup as in Example 3.2.7, i.e. $M = \mathbb{R}^2$ with $x = 0$ fixed and the vector fields X_1, X_2 on $\mathbb{R}^2$ defined by

$$X_1 = \frac{\partial}{\partial x^1} + x^1 \frac{\partial}{\partial x^2} \qquad \text{and} \qquad X_2 = x^1 \frac{\partial}{\partial x^1}$$

in Cartesian coordinates (x^1, x^2). We recall that these coordinates are not adapted to the filtration induced by (X_1, X_2) at 0 and we start off by illustrating why this chart is not suitable for our analysis. The unique strong solution $(x_t^\varepsilon)_{t\in[0,1]} = (x_t^{\varepsilon,1}, x_t^{\varepsilon,2})_{t\in[0,1]}$ of the Stratonovich stochastic differential equation in $\mathbb{R}^2$

$$\begin{aligned} \partial x_t^{\varepsilon,1} &= \sqrt{\varepsilon}\, \partial B_t^1 + \sqrt{\varepsilon} x_t^{\varepsilon,1}\, \partial B_t^2 \\ \partial x_t^{\varepsilon,2} &= \sqrt{\varepsilon} x_t^{\varepsilon,1}\, \partial B_t^1 \end{aligned}$$

subject to $x_0^\varepsilon = 0$ is given by

$$x_t^\varepsilon = \left(\sqrt{\varepsilon} \int_0^t \mathrm{e}^{\sqrt{\varepsilon}\left(B_t^2 - B_s^2\right)}\, \partial B_s^1, \varepsilon \int_0^t \left(\int_0^s \mathrm{e}^{\sqrt{\varepsilon}\left(B_s^2 - B_r^2\right)}\, \partial B_r^1 \right) \partial B_s^1 \right) .$$

Though the step of the filtration induced by (X_1, X_2) at 0 is $N = 3$, rescaling the stochastic process $(x_t^\varepsilon)_{t\in[0,1]}$ by $\varepsilon^{-3/2}$ in any direction leads to a blow-up in the limit $\varepsilon \to 0$. Instead, the highest-order rescaled process we can consider is $(\varepsilon^{-1/2} x_t^{\varepsilon,1}, \varepsilon^{-1} x_t^{\varepsilon,2})_{t\in[0,1]}$ whose limit process, as $\varepsilon \to 0$, is characterised by

$$\lim_{\varepsilon \to 0} \left(\varepsilon^{-1/2} x_t^{\varepsilon,1}, \varepsilon^{-1} x_t^{\varepsilon,2}\right) \to \left(B_t^1, \frac{1}{2}\left(B_t^1\right)^2\right) .$$

We see that these rescaled processes localise around a parabola in $\mathbb{R}^2$. As the Malliavin covariance matrix of $(B_1^1, \frac{1}{2}(B_1^1)^2)$ is degenerate, the Fourier transform argument from Section 3.4 cannot be used. Rather, we first need to apply an additional rescaling along the parabola to recover a non-degenerate limit process. This is the reason why we choose to work in an adapted chart because it allows us to express the overall rescaling needed as an anisotropic dilation.

Let the map $\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the same global adapted chart as used in Example 3.2.7 and let $\delta_\varepsilon\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the associated anisotropic dilation. We showed that the nilpotent approximations $\tilde{X}_1, \tilde{X}_2$ of the vector fields X_1, X_2 are

$$\tilde{X}_1 = \frac{\partial}{\partial y^1} \qquad \text{and} \qquad \tilde{X}_2 = -\left(y^1\right)^2 \frac{\partial}{\partial y^2} ,$$

with respect to Cartesian coordinates (y^1, y^2) on the second copy of $\mathbb{R}^2$. The convergence result (3.3.1) implies that, for all $t \in [0,1]$,

$$\delta_\varepsilon^{-1}\left(\theta(x_t^\varepsilon)\right) \to \left(B_t^1, -\int_0^t \left(B_s^1\right)^2 \partial B_s^2\right) \quad \text{as} \quad \varepsilon \to 0 .$$

Since $\mathrm{d}\theta_0\colon \mathbb{R}^2 \to \mathbb{R}^2$ is the identity map, Theorem 3.1.1 says that the suitably rescaled fluctuations of the diffusion loop at 0 associated to the stochastic process with generator

$$\mathcal{L} = \frac{1}{2}(X_1^2 + X_2^2)$$

converge weakly to the loop obtained by conditioning the process $(B_t^1, -\int_0^t (B_s^1)^2\, \partial B_s^2)_{t\in[0,1]}$ to return to 0 at time 1.

Remark 3.5.2. We demonstrate that Theorem 3.1.1 and Theorem 3.1.2 allow us to make deductions about the $\sqrt{\varepsilon}$-rescaled fluctuations of diffusion loops. For the rescaling map $\tau_\varepsilon\colon \Omega^{x,x} \to T\Omega^{0,0}$ given by

$$\tau_\varepsilon(\omega)_t = (\mathrm{d}\theta_x)^{-1}\left(\varepsilon^{-1/2}\theta(\omega_t)\right) ,$$

we are interested in the behaviour of the measures $\mu_\varepsilon^{x,x} \circ \tau_\varepsilon^{-1}$ in the limit as $\varepsilon \to 0$. Let $e_1, \dots, e_d$ be the standard basis in $\mathbb{R}^d$ and define $\psi\colon T\Omega^{0,0} \to T\Omega^{0,0}$ by

$$\psi(v)_t = \sum_{i=1}^{d_1} \langle \mathrm{d}\theta_x(v_t), e_i\rangle \left(\mathrm{d}\theta_x\right)^{-1} e_i .$$

The map ψ takes a path in $T\Omega^{0,0}$ and projects it onto the component living in the subspace $C_1(x)$ of T_xM. Since the maps τ_ε and σ_ε are related by

$$\tau_\varepsilon(\omega)_t = (\mathrm{d}\theta_x)^{-1}\left(\varepsilon^{-1/2}\delta_\varepsilon\left(\mathrm{d}\theta_x\left(\sigma_\varepsilon(\omega)_t\right)\right)\right)$$

and because $\varepsilon^{-1/2}\delta_\varepsilon(y)$ tends to $(y^1,\dots,y^{d_1},0,\dots,0)$ in the limit as $\varepsilon\to 0$, it follows that the $\sqrt{\varepsilon}$-rescaled diffusion loop measures $\mu_\varepsilon^{x,x}\circ\tau_\varepsilon^{-1}$ converge weakly to $\tilde\mu^{x,x}\circ\psi^{-1}$ on $T\Omega^{0,0}$ as $\varepsilon\to 0$. Provided $\mathcal{L}$ is non-elliptic at x, the latter is a degenerate measure which is supported on the set of paths $(v_t)_{t\in[0,1]}$ in $T\Omega^{0,0}$ which satisfy $v_t\in C_1(x)$, for all $t\in[0,1]$. Hence, the rescaled diffusion process $(\varepsilon^{-1/2}\theta(x_t^\varepsilon))_{t\in[0,1]}$ conditioned by $\theta(x_1^\varepsilon)=0$ localises around the subspace $(\theta_*C_1)(0)$.

Finally, by considering the limiting diffusion loop from Example 3.5.1, we show that the degenerate limit measure $\tilde\mu^{x,x}\circ\psi^{-1}$ need not be Gaussian. Going back to Example 3.5.1, we first observe that the map ψ is simply projection onto the first component, i.e.

$$\psi(v)_t = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} v_t\;.$$

Thus, to show that the measure $\tilde\mu^{x,x}\circ\psi^{-1}$ is not Gaussian, we have to analyse the process

$$\left(B_t^1, -\int_0^t (B_s^1)^2\,\partial B_s^2\right)_{t\in[0,1]}$$

conditioned to return to 0 at time 1 and show that its first component is not Gaussian. Using the tower property, we first condition on $B_1^1=0$ to see that this component is equal in law to the process $(B_t^1-tB_1^1)_{t\in[0,1]}$ conditioned by $\int_0^1(B_s^1-sB_1^1)^2\,\partial B_s^2=0$, where the latter is in fact equivalent to conditioning on $\int_0^1(B_s^1-sB_1^1)^2\,\mathrm{d}B_s^2=0$. Let μ_B denote the Brownian bridge measure on $\Omega(\mathbb{R})^{0,0}=\{\omega\in C([0,1],\mathbb{R})\colon \omega_0=\omega_1=0\}$ and let ν be the law of $-\int_0^1(B_s^1-sB_1^1)^2\,\mathrm{d}B_s^2$ on $\mathbb{R}$. Furthermore, denote the joint law of

$$\left(B_t^1-tB_1^1\right)_{t\in[0,1]}\qquad\text{and}\qquad -\int_0^1\left(B_s^1-sB_1^1\right)^2\,\mathrm{d}B_s^2$$

on $\Omega(\mathbb{R})^{0,0}\times\mathbb{R}$ by μ. Since $-\int_0^1\omega_s^2\,\mathrm{d}B_s^2$, for $\omega\in\Omega(\mathbb{R})^{0,0}$ fixed, is a normal random variable with mean zero and variance $\int_0^1\omega_s^4\,\mathrm{d}s$, we obtain that

$$\mu(\mathrm{d}\omega,\mathrm{d}y)=\frac{1}{\sqrt{2\pi}\sigma(\omega)}\,\mathrm{e}^{-\frac{y^2}{2\sigma^2(\omega)}}\,\mu_B(\mathrm{d}\omega)\,\mathrm{d}y\quad\text{with}\quad \sigma(\omega)=\left(\int_0^1\omega_s^4\,\mathrm{d}s\right)^{1/2}\;. \tag{3.5.5}$$

On the other hand, we can disintegrate μ as

$$\mu(\mathrm{d}\omega,\mathrm{d}y)=\mu_B^y(\mathrm{d}\omega)\nu(\mathrm{d}y)\;,$$

where μ_B^y is the law of $(B_t^1 - tB_1^1)_{t\in[0,1]}$ conditioned by $-\int_0^1 (B_s^1 - sB_1^1)^2 \, \mathrm{d}B_s^2 = y$, i.e. we are interested in the measure μ_B^0. From (3.5.5), it follows that

$$\mu_B^0(\mathrm{d}\omega) \propto \sigma^{-1}(\omega)\mu_B(\mathrm{d}\omega) = \left(\int_0^1 \omega_s^4 \, \mathrm{d}s\right)^{-1/2} \mu_B(\mathrm{d}\omega) \; .$$

This shows that μ_B^0 is not Gaussian, which implies that the $\sqrt{\varepsilon}$-rescaled fluctuations indeed admit a non-Gaussian limiting diffusion loop. $\square$

Chapter 4

Brownian motion conditioned to have trivial signature

We report on work in progress studying Brownian motion which is conditioned to have vanishing iterated integrals of all orders. The idea for this project resulted from dialogue with Terry Lyons. Chen [Che58] studied the formal series of iterated integrals of a path, called the signature, and proved uniqueness, up to translation and reparametrisation, in a class of piecewise regular paths. Hambly and Lyons [HL10] extended Chen's theorem and showed that two paths of bounded variation have the same signature if and only if they differ by a tree-like path. They left the question open if the same is true for weak geometric p-rough paths with $p > 1$. Le Jan and Qian [LJQ13] first proved that almost all Brownian motion sample paths are determined by their signature, and Boedihardjo, Geng, Lyons and Yang [BGLY16] then positively answered the question in the general case subject to the appropriate definition of a path to be tree-like. This shows that the signature of a path encodes enough information to completely determine it up to tree-like paths. In particular, the law of Brownian motion conditioned to have vanishing iterated integrals up to order N concentrates for large N around the set of tree-like paths. We conjecture that the laws in fact converge weakly to the unit mass at the zero path. Our work in relation with this conjecture has led to an analytic proof of the stand-alone result that a Brownian bridge in $\mathbb{R}^d$ from 0 to 0 in time 1 is more likely to stay inside a box centred at the origin than any other Brownian bridge in time 1.

4.1 Signature of Brownian motion

We set up conventions used and recall results needed in our analysis of the behaviour of Brownian motion which is conditioned to have vanishing iterated integrals of all orders. Following Lyons, Caruana and Lévy [LCL07], the signature and the truncated signature of a continuous path of bounded variation, also called of finite 1-variation, are defined as follows.

Definition 4.1.1. *The signature of a continuous path $z\colon [0,1] \to \mathbb{R}^d$ of bounded variation is the element $S(z)$ in the space of formal series of tensors of $\mathbb{R}^d$ defined by*

$$S(z) = \left(1, \int_0^1 \mathrm{d}z_t, \dots, \int_0^1 \int_0^{t_k} \dots \int_0^{t_2} \mathrm{d}z_{t_1} \otimes \dots \otimes \mathrm{d}z_{t_k}, \dots \right) .$$

The truncated signature of order N of the continuous path z of bounded variation is the element $S_N(z)$ in the truncated tensor algebra of order N of $\mathbb{R}^d$ given by

$$S_N(z) = \left(1, \int_0^1 \mathrm{d}z_t, \dots, \int_0^1 \int_0^{t_N} \dots \int_0^{t_2} \mathrm{d}z_{t_1} \otimes \dots \otimes \mathrm{d}z_{t_N} \right) .$$

Let $(B_t)_{t\in[0,1]}$ be a Brownian motion in $\mathbb{R}^d$, which we assume is realised as the coordinate process on the set $\Omega^0(\mathbb{R}^d) = \{w \in C([0,1],\mathbb{R}^d) \colon w_0 = 0\}$ under Wiener measure $\mathbb{P}$. Note that Definition 4.1.1 does not yet allow us to speak of the signature of Brownian motion as almost all its sample paths are of unbounded variation. By using rough paths theory, it is possible to extend the notion of signature to sets of continuous paths which are not of bounded variation. For details on rough paths theory, see Friz and Victoir [FV10]. Almost all Brownian sample paths are of finite p-variation for $p > 2$, and we obtain the canonical Brownian rough path, using ∂ to denote the Stratonovich differential,

$$\left(1, B_t, \int_0^t \int_0^s \partial B_r \otimes \partial B_s \right)_{t\in[0,1]} ,$$

which is indeed a geometric p-rough path for $2 < p < 3$. As a consequence of the extension theorem [LCL07, Theorem 3.7], every p-rough path has a full signature. Thus, we can use the canonical Brownian rough path to define the signature of Brownian motion. To present the definition obtained by this construction, we change tack and start following notations in Baudoin [Bau04] as our problem can be fully understood in terms of stochastic integrals.

Definition 4.1.2. *The signature of Brownian motion $(B_t)_{t\in[0,1]}$ in $\mathbb{R}^d$ is the element of the non-commutative algebra $\mathbb{R}[[X_1,\dots,X_d]]$ of formal series with d indeterminates defined by*

$$S(B) = 1 + \sum_{k=1}^{\infty} \sum_{i_1,\dots,i_k=1}^{d} X_{i_1} \dots X_{i_k} \int_0^1 \int_0^{t_k} \dots \int_0^{t_2} \partial B_{t_1}^{i_1} \dots \partial B_{t_k}^{i_k} .$$

The truncated signature of order N of Brownian motion $(B_t)_{t\in[0,1]}$ is the element of the set $\mathbb{R}_N[X_1,\cdots,X_d]$ of formal series with d indeterminates truncated at order N given by

$$S_N(B) = 1 + \sum_{k=1}^{N} \sum_{i_1,\dots,i_k=1}^{d} X_{i_1} \dots X_{i_k} \int_0^1 \int_0^{t_k} \dots \int_0^{t_2} \partial B_{t_1}^{i_1} \dots \partial B_{t_k}^{i_k} .$$

This is also called the (truncated) Stratonovich signature of Brownian motion to stress the fact that the integrals are understood in the Stratonovich sense.

The first observation we make is that the truncated signature $S_N(B)$ of Brownian motion in $\mathbb{R}^d$ takes values in the free Carnot group $\mathbb{G}_N(\mathbb{R}^d)$ of step N. This follows from

$$S_N(B) = \exp\left(\sum_{k=1}^{N}\sum_{i_1,\dots,i_k=1}^{d}\Lambda_{(i_1,\dots,i_k)}(B)[X_{i_1},[X_{i_2},\dots,[X_{i_{k-1}},X_{i_k}]\dots]]\right)\;, \tag{4.1.1}$$

where

$$\Lambda_{(i_1,\dots,i_k)}(B) = \sum_{\sigma\in\mathcal{S}_k}\frac{(-1)^{e(\sigma)}}{k^2\binom{k-1}{e(\sigma)}}\int_0^1\int_0^{t_k}\cdots\int_0^{t_2}\partial B_{t_1}^{i_{\sigma^{-1}(1)}}\dots\partial B_{t_k}^{i_{\sigma^{-1}(k)}}$$

with $\mathcal{S}_k$ the group of permutations of the set $\{1,\dots,k\}$ and $e(\sigma)$ denoting the cardinality of the set $\{i\in\{1,\dots,k-1\}\colon \sigma(i)>\sigma(i+1)\}$. The expansion (4.1.1) is a consequence of the Chen-Strichartz development formula [Bau04, Theorem 1.1], which is a restatement of a result by Chen [Che57] and Strichartz [Str87], and whose proof uses the generalised Baker-Campbell-Hausdorff formula.

Let us further understand the $X_1,\dots,X_d$ as left-invariant vector fields on the free Carnot group $\mathbb{G}_N(\mathbb{R}^d)$ and define a process $(\mathbf{B}_t)_{t\in[0,1]}$ in $\mathbb{G}_N(\mathbb{R}^d)$ as the unique strong solution of the Stratonovich stochastic differential equation

$$\partial\mathbf{B}_t = \sum_{i=1}^{d}X_i\left(\mathbf{B}_t\right)\partial B_t^i\;,\quad \mathbf{B}_0 = 1\;.$$

We call $(\mathbf{B}_t)_{t\in[0,1]}$ the lift of Brownian motion to $\mathbb{G}_N(\mathbb{R}^d)$. Note that the generator of this process is $\frac{1}{2}\sum_{i=1}^{d}X_i^2$, which is hypoelliptic by construction. In particular, the law of $\mathbf{B}_1$ on $\mathbb{G}_N(\mathbb{R}^d)$ is absolutely continuous with respect to Lebesgue measure. Hence, using the theory of disintegration of measures, we can make sense of the loop in $\mathbb{G}_N(\mathbb{R}^d)$ obtained by conditioning the lift $(\mathbf{B}_t)_{t\in[0,1]}$ on $\mathbf{B}_1=1$. For details, see Section 4.2. The projection of the loop in $\mathbb{G}_N(\mathbb{R}^d)$ onto the base space $\mathbb{R}^d$ is called the Brownian loop of step N. For an alternative construction using Doob h-transforms, see Baudoin [Bau04, Section 3.6]. We observe that [Bau04, Proposition 2.3] implies that $\mathbf{B}_1=S_N(B)$. Therefore, the study of Brownian motion conditioned to have trivial signature can be seen as analysing Brownian loops of step N in the limit $N\to\infty$. Using this terminology, our conjecture is as follows.

Conjecture 4.1.3. *Let $d\geq 2$. Then the laws of Brownian loops of step N converge to the unit mass δ_0 at the zero path weakly on $\Omega^0(\mathbb{R}^d)$ as $N\to\infty$.*

Our work towards validating this conjecture is included in the next section.

4.2 Outline of ideas

We show how to define a Brownian loop of step N using disintegration of measures and then present ideas which could be useful in proving Conjecture 4.1.3. As a by-product, we obtain an analytic proof of the stand-alone result that a Brownian bridge in $\mathbb{R}^d$ from 0 to 0 in time 1 is more likely to stay inside a box centred at the origin than any other Brownian bridge in time 1.

Throughout, we use the characterisation of the free Carnot group $\mathbb{G}_N(\mathbb{R}^d)$ of step N given by [Bau04, Proposition 2.8]. It allows us to consider $\mathbb{G}_N(\mathbb{R}^d)$ as some $\mathbb{R}^{d_N}$ endowed with a polynomial group law which is unimodular, i.e. the left Haar measure and the right Haar measure are the same on the corresponding Lie group. They agree with the Lebesgue measure and it follows that translations leave the Lebesgue measure invariant. In this setting, the identity element of $\mathbb{G}_N(\mathbb{R}^d)$ is 0, and there exist left-invariant polynomial vector fields $Y_1, \ldots, Y_d$ on $\mathbb{R}^{d_N}$ such that the lift $(\mathbf{B}_t)_{t\in[0,1]}$ of Brownian motion to $\mathbb{G}_N(\mathbb{R}^d)$ becomes the unique strong solution of the Stratonovich stochastic differential equation

$$\partial \mathbf{B}_t = \sum_{i=1}^{d} Y_i\left(\mathbf{B}_t\right) \partial B_t^i \;, \quad \mathbf{B}_0 = 0 \;.$$

Let p_N denote the law of $\mathbf{B}_1 = S_N(B)$ on the free Carnot group $\mathbb{G}_N(\mathbb{R}^d)$ considered as the appropriate $\mathbb{R}^{d_N}$ endowed with a unimodular group law. Since the generator of the lift process $(\mathbf{B}_t)_{t\in[0,1]}$ in $\mathbb{R}^{d_N}$ is hypoelliptic, the law p_N is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^{d_N}$. Disintegrating Wiener measure $\mathbb{P}$ with respect to p_N gives a unique family of probability measures $(\mathbb{P}_N^x \colon x \in \mathbb{R}^{d_N})$ on $\Omega^0(\mathbb{R}^d)$, with $S_N(B) = x$ almost surely under $\mathbb{P}_N^x$ for all $x \in \mathbb{R}^{d_N}$, which is weakly continuous in x, and such that

$$\mathbb{P}(\mathrm{d}w) = \int_{\mathbb{R}^{d_N}} \mathbb{P}_N^x(\mathrm{d}w) p_N(x)\, \mathrm{d}x \;. \tag{4.2.1}$$

Here $S_N(B)$ is indeed well-defined, because $(B_t)_{t\in[0,1]}$ is still a semimartingale under the measure $\mathbb{P}_N^x$, cf. [Bau04, Proposition 3.5]. By construction, $\mathbb{P}_N^0$ is the law of the Brownian loop of step N. Therefore, Conjecture 4.1.3 states that, provided $d \geq 2$, the measures $\mathbb{P}_N^0$ converge to δ_0 weakly on $\Omega^0(\mathbb{R}^d)$ as $N \to \infty$. Our current idea for proving this conjecture is as follows.

Let q_N denote the law of $\int_0^1 (B_t^1)^{N-1}\, \partial B_t^2$ on $\mathbb{R}$, which is absolutely continuous with respect to Lebesgue measure. We disintegrate Wiener measure $\mathbb{P}$ with respect to q_N to obtain a unique family of probability measures $(\mathbb{Q}_N^y \colon y \in \mathbb{R})$ on $\Omega^0(\mathbb{R}^d)$, which is weakly continuous in y, where $\int_0^1 (B_t^1)^{N-1}\, \partial B_t^2 = y$ almost surely under $\mathbb{Q}_N^y$ for all y, and such that

$$\mathbb{P}(\mathrm{d}w) = \int_{\mathbb{R}} \mathbb{Q}_N^y(\mathrm{d}w) q_N(y)\, \mathrm{d}y \;. \tag{4.2.2}$$

It is possible to recover the probability measures $(\mathbb{P}_N^x \colon x \in \mathbb{R}^{d_N})$ from the probability measures $(\mathbb{Q}_N^y \colon y \in \mathbb{R})$ by disintegrating with respect to an appropriate measure. The idea is that first conditioning $(B_t)_{t\in[0,1]}$ on $\int_0^1 (B_t^1)^{N-1}\, \partial B_t^2 = 0$ and then conditioning the resulting process to have trivial truncated signature of order N yields the same process as just conditioning $(B_t)_{t\in[0,1]}$ to have trivial truncated signature of order N. Define $\mathbb{R}^{d_N}_y$ to be the subset of $\mathbb{R}^{d_N}$ of all values taken by the truncated signature of order N of Brownian motion subject to the condition that $\int_0^1 (B_t^1)^{N-1}\,\partial B_t^2 = y$. Let p_N^y be the law of $S_N(B)$ under the measure $\mathbb{Q}_N^y$.

Lemma 4.2.1. *For all $y \in \mathbb{R}$, it holds true that*

$$\mathbb{Q}_N^y(\mathrm{d}w) = \int_{\mathbb{R}^{d_N}_y} \mathbb{P}_N^x(\mathrm{d}w) p_N^y(x)\,\mathrm{d}x\ .$$

Proof. We can disintegrate the measure $\mathbb{Q}_N^y$ uniquely as

$$\mathbb{Q}_N^y(\mathrm{d}w) = \int_{\mathbb{R}^{d_N}_y} \mathbb{P}_N^{x,y}(\mathrm{d}w) p_N^y(x)\,\mathrm{d}x \tag{4.2.3}$$

to obtain a family of probability measures $(\mathbb{P}_N^{x,y} \colon x \in \mathbb{R}^{d_N}_y)$ on $\Omega^0(\mathbb{R}^d)$, which is weakly continuous in x and where $S_N(B) = x$ almost surely under $\mathbb{P}_N^{x,y}$ for all $x \in \mathbb{R}^{d_N}_y$. Since p_N^y is the law of $S_N(B)$ under $\mathbb{Q}_N^y$, we have

$$p_N(x) = \int_{\mathbb{R}} p_N^y(x) q_N(y)\,\mathrm{d}y\ . \tag{4.2.4}$$

Observe that

$$\int_0^1 (B_t^1)^{N-1}\,\partial B_t^2 = (N-1)! \int_0^1 \int_0^{t_N} \cdots \int_0^{t_2} \partial B_{t_1}^1 \ldots \partial B_{t_{N-1}}^1\,\partial B_{t_N}^2$$

is a multiple of the coefficient of $(X_1)^{N-1}X_2$ in $S_N(B)$. In the free Carnot group of step N several coefficients in the truncated signature of order N are combined into one component of $\mathbb{R}^{d_N}$, so that the lift process $(\mathbf{B}_t)_{t\in[0,1]}$ is hypoelliptic. However, it is possible to recover all the coefficients in $S_N(B)$ and we choose $\phi\colon \mathbb{R}^{d_N} \to \mathbb{R}$ such that $\phi(x) = y$ if and only if $x \in \mathbb{R}^{d_N}_y$. Using (4.2.2), (4.2.3) and (4.2.4), we deduce that

$$\begin{aligned}\mathbb{P}(\mathrm{d}w) &= \int_{\mathbb{R}} \mathbb{Q}_N^y(\mathrm{d}w) q_N(y)\,\mathrm{d}y = \int_{\mathbb{R}}\int_{\mathbb{R}^{d_N}_y} \mathbb{P}_N^{x,y}(\mathrm{d}w) p_N^y(x) q_N(y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_{\mathbb{R}^{d_N}} \mathbb{P}_N^{x,\phi(x)}(\mathrm{d}w) p_N(x)\,\mathrm{d}x\ .\end{aligned}$$

Since $S_N(B) = x$ holds almost surely both under $\mathbb{P}_N^{x,\phi(x)}$ and under $\mathbb{P}_N^x$, uniqueness of the disintegration (4.2.1) implies that $\mathbb{P}_N^x = \mathbb{P}_N^{x,\phi(x)}$. The lemma follows due to $\mathbb{P}_N^{x,\phi(x)} = \mathbb{P}_N^{x,y}$ for $x \in \mathbb{R}^{d_N}_y$. □

Our reason for considering the measures $\mathbb{Q}_N^y$ is that, whereas it appears to be challenging to analyse the Brownian loop laws $\mathbb{P}_N^0$ directly, we have an explicit expression for the first marginal of $\mathbb{Q}_N^y$, see Lemma 4.2.2 below. We hope to later use this explicit expression to prove weak convergence of the first marginals of $\mathbb{Q}_N^0$ to the Dirac delta mass δ_0 on $\Omega^0(\mathbb{R})$

as $N \to \infty$. Provided we further obtained sufficient control over the measures $\mathbb{P}_N^x$ in terms of $\mathbb{P}_N^0$, e.g. as in Conjecture 4.2.4 or Conjecture 4.2.5, we could then use Lemma 4.2.1 to bootstrap information about the loop measures $\mathbb{P}_N^0$ to prove weak convergence of their first marginals to δ_0 on $\Omega^0(\mathbb{R})$ as $N \to \infty$. By conditioning on the value of the iterated integral $\int_0^1 (B_t^i)^{N-1}\, \partial B_t^j$ instead of $\int_0^1 (B_t^1)^{N-1}\, \partial B_t^2$, we would similarly deduce that any marginal of $\mathbb{P}_N^0$ converges weakly on $\Omega^0(\mathbb{R})$ to the unit mass δ_0 as $N \to \infty$, which would imply Conjecture 4.1.3.
Let $\pi\colon \Omega^0(\mathbb{R}^d) \to \Omega^0(\mathbb{R})$ denote the projection onto the first component of a path. Our expression for the first marginal of $\mathbb{Q}_N^y$ is given by the following lemma.

Lemma 4.2.2. *For $v \in \Omega^0(\mathbb{R})$, set $\sigma_N(v) = \left(\int_0^1 v_t^{2(N-1)}\, \mathrm{d}t\right)^{1/2}$. Then, for $y \in \mathbb{R}$, we have*

$$(\pi_* \mathbb{Q}_N^y)\,(\mathrm{d}v) = \frac{1}{\sqrt{2\pi}\sigma_N(v) q_N(y)} \exp\left(-\frac{y^2}{2\sigma_N^2\,(v)}\right) (\pi_* \mathbb{P})\,(\mathrm{d}v) \ .$$

Proof. Note that $\int_0^1 (B_t^1)^{N-1}\, \partial B_t^2 = \int_0^1 (B_t^1)^{N-1}\, \mathrm{d}B_t^2$ by independence of the Brownian motions $(B_t^1)_{t\in[0,1]}$ and $(B_t^2)_{t\in[0,1]}$ in $\mathbb{R}$. In particular, the first marginal $\pi_* \mathbb{Q}_N^y$ is the law of the process $(B_t^1)_{t\in[0,1]}$ conditioned on $\int_0^1 (B_t^1)^{N-1}\, \mathrm{d}B_t^2 = y$. Let μ_N denote the joint law of $(B_t^1)_{t\in[0,1]}$ and $\int_0^1 (B_t^1)^{N-1}\, \mathrm{d}B_t^2$ on $\Omega^0(\mathbb{R}) \times \mathbb{R}$. Since, for a path $v \in \Omega^0(\mathbb{R})$, the random variable $\int_0^1 v_t^{N-1}\, \mathrm{d}B_t^2$ is normal with mean zero and variance $\sigma_N^2(v)$, we obtain

$$\mu_N(\mathrm{d}v, \mathrm{d}y) = \frac{1}{\sqrt{2\pi}\sigma_N(v)} \exp\left(-\frac{y^2}{2\sigma_N^2\,(v)}\right) (\pi_* \mathbb{P})\,(\mathrm{d}v)\, \mathrm{d}y \ . \tag{4.2.5}$$

On the other hand, by the disintegration (4.2.2), we also have

$$\mu_N(\mathrm{d}v, \mathrm{d}y) = (\pi_* \mathbb{Q}_N^y)\,(\mathrm{d}v) q_N(y)\, \mathrm{d}y \ . \tag{4.2.6}$$

The result follows by comparing expressions (4.2.5) and (4.2.6). ☐

Taking $y = 0$ in Lemma 4.2.2 yields, for $v \in \Omega^0(\mathbb{R})$,

$$\left(\pi_* \mathbb{Q}_N^0\right)(\mathrm{d}v) = \frac{1}{\sqrt{2\pi}\sigma_N(v) q_N(0)} (\pi_* \mathbb{P})\,(\mathrm{d}v) \ .$$

Due to the reweighting factor $\sigma_N(v) = \left(\int_0^1 v_t^{2(N-1)}\, \mathrm{d}t\right)^{1/2}$, the first marginals $\pi_* \mathbb{Q}_N^0$ appear to localise for large N around the zero path in $\Omega^0(\mathbb{R})$. We make the following conjecture.

Conjecture 4.2.3. *The first marginals $\pi_* \mathbb{Q}_N^0$ converge to δ_0 weakly on $\Omega^0(\mathbb{R})$ as $N \to \infty$.*

We further conjecture that we can control the mass the first marginals $\pi_* \mathbb{P}_N^x$ put on balls around the zero path in terms of the Brownian loop marginals $\pi_* \mathbb{P}_N^0$. For $r > 0$, set

$$D_r = \left\{ v \in \Omega^0(\mathbb{R})\colon \sup_{0 \le t \le 1} |v_t| < r \right\} \ .$$

Conjecture 4.2.4. *For all $x \in \mathbb{R}^{d_N}$ and all $r > 0$, it holds true that*

$$\left(\pi_* \mathbb{P}_N^x\right)(D_r) \le \left(\pi_* \mathbb{P}_N^0\right)(D_r) \,.$$

We show how Conjecture 4.2.3 and Conjecture 4.2.4 imply Conjecture 4.1.3. As argued previously, it is enough to show that they imply the weak convergence of the first marginals $\pi_* \mathbb{P}_N^0$ to δ_0 on $\Omega^0(\mathbb{R})$ as $N \to \infty$. By the Portmanteau theorem, the latter follows if, for all open sets $U \subset \Omega^0(\mathbb{R})$,

$$\delta_0(U) \le \liminf_{N\to\infty} \left(\pi_* \mathbb{P}_N^0\right)(U) \,. \tag{4.2.7}$$

This inequality is indeed implied by Conjecture 4.2.3 and Conjecture 4.2.4. If $U \subset \Omega^0(\mathbb{R})$ is an open subset not containing the zero path then $\delta_0(U) = 0$ and (4.2.7) follows from the non-negativity of measures. Let us now suppose that $U \subset \Omega^0(\mathbb{R})$ is an open subset which does contain the zero path. In particular, there exists some $r > 0$ such that $D_r \subset U$. Assuming Conjecture 4.2.4 and using Lemma 4.2.1, we can deduce that

$$\begin{aligned}\left(\pi_* \mathbb{Q}_N^0\right)(D_r) &= \int_{\mathbb{R}_0^{d_N}} \left(\pi_* \mathbb{P}_N^x\right)(D_r)\, p_N^0(x)\, \mathrm{d}x \\ &\le \int_{\mathbb{R}_0^{d_N}} \left(\pi_* \mathbb{P}_N^0\right)(D_r)\, p_N^0(x)\, \mathrm{d}x = \left(\pi_* \mathbb{P}_N^0\right)(D_r) \le \left(\pi_* \mathbb{P}_N^0\right)(U) \,.\end{aligned}$$

Conjecture 4.2.3 says that the first marginals $\pi_* \mathbb{Q}_N^0$ converge to δ_0 weakly as $N \to \infty$ and therefore,

$$1 = \delta_0\left(D_r\right) \le \liminf_{N\to\infty} \left(\pi_* \mathbb{Q}_N^0\right)(D_r) \,. \tag{4.2.8}$$

Combining the last two inequalities gives

$$\delta_0(U) = 1 \le \liminf_{N\to\infty} \left(\pi_* \mathbb{Q}_N^0\right)(D_r) \le \liminf_{N\to\infty} \left(\pi_* \mathbb{P}_N^0\right)(U) \,.$$

Thus, assuming Conjecture 4.2.3 and Conjecture 4.2.4, we obtain (4.2.7), as desired. As argued below, the inequality (4.2.7) still holds if, instead of Conjecture 4.2.4, we assume the following weaker conjecture.

Conjecture 4.2.5. *For all $x \in \mathbb{R}^{d_N}$ and all $r > 0$, it holds true that*

$$\left(\pi_* \mathbb{P}_N^x\right)(D_r) \le \sqrt{\left(\pi_* \mathbb{P}_N^0\right)\left(D_{\sqrt{2}r}\right)} \,.$$

Assuming Conjecture 4.2.5 instead of Conjecture 4.2.4, we get (4.2.7) as follows. By using Lemma 4.2.1, we this time deduce

$$\begin{aligned}\left(\pi_* \mathbb{Q}_N^0\right)\left(D_{r/\sqrt{2}}\right) &= \int_{\mathbb{R}_0^{d_N}} \left(\pi_* \mathbb{P}_N^x\right)\left(D_{r/\sqrt{2}}\right) p_N^0(x)\, \mathrm{d}x \\ &\le \int_{\mathbb{R}_0^{d_N}} \sqrt{\left(\pi_* \mathbb{P}_N^0\right)(D_r)}\, p_N^0(x)\, \mathrm{d}x = \sqrt{\left(\pi_* \mathbb{P}_N^0\right)(D_r)} \le \sqrt{\left(\pi_* \mathbb{P}_N^0\right)(U)} \,,\end{aligned}$$

which together with (4.2.8) yields

$$\delta_0(U) = 1 \leq \liminf_{N\to\infty} \left(\left(\pi_* \mathbb{Q}_N^0\right) \left(D_{r/\sqrt{2}}\right)\right)^2 \leq \liminf_{N\to\infty} \left(\pi_* \mathbb{P}_N^0\right)(U) \ ,$$

as required.

Our proofs of Conjecture 4.2.4 and Conjecture 4.2.5 for $N = 1$, i.e. in the case of Brownian bridges, make it apparent why we include both conjectures. The proof of Conjecture 4.2.5 for $N = 1$, which uses a coupling argument, is much shorter than the analytic proof we have of Conjecture 4.2.4 for $N = 1$. Both conjectures remain open for $N \geq 2$.

Lemma 4.2.6 (Conjecture 4.2.5 for $N = 1$)**.** *For all $x \in \mathbb{R}^d$ and all $r > 0$, we have*

$$\left(\pi_* \mathbb{P}_1^x\right)(D_r) \leq \sqrt{\left(\pi_* \mathbb{P}_1^0\right)\left(D_{\sqrt{2}r}\right)} \ .$$

Proof. We note that the marginal $\pi_* \mathbb{P}_1^x$ is the law of a Brownian bridge in $\mathbb{R}$ from 0 to x^1 in time 1, and that $(\pi_* \mathbb{P}_1^x)(D_r) = (\pi_* \mathbb{P}_1^{-x})(D_r)$ by symmetry. Let $(W_t^1)_{t\in[0,1]}$ and $(W_t^2)_{t\in[0,1]}$ be independent standard Brownian motions in $\mathbb{R}$. Consider the Brownian bridge $(\mathbf{W}_t^1)_{t\in[0,1]}$ in $\mathbb{R}$ from 0 to x^1 given by

$$\mathbf{W}_t^1 = W_t^1 - tW_1^1 + tx^1$$

and the Brownian bridge $(\mathbf{W}_t^2)_{t\in[0,1]}$ in $\mathbb{R}$ from 0 to $-x^1$ obtained as

$$\mathbf{W}_t^2 = W_t^2 - tW_1^2 - tx^1 \ .$$

By independence of the Brownian motions $(W_t^1)_{t\in[0,1]}$ and $(W_t^2)_{t\in[0,1]}$, the Brownian bridges $(\mathbf{W}_t^1)_{t\in[0,1]}$ and $(\mathbf{W}_t^2)_{t\in[0,1]}$ are also independent, and the process $\frac{1}{\sqrt{2}}(W_t^1 + W_t^2)_{t\in[0,1]}$ is a standard Brownian motion in $\mathbb{R}$. We further observe

$$\frac{\mathbf{W}_t^1 + \mathbf{W}_t^2}{\sqrt{2}} = \frac{W_t^1 + W_t^2}{\sqrt{2}} - t\frac{W_1^1 + W_1^2}{\sqrt{2}} \ ,$$

and it follows that $\frac{1}{\sqrt{2}}(\mathbf{W}_t^1 + \mathbf{W}_t^2)_{t\in[0,1]}$ is a Brownian bridge in $\mathbb{R}$ from 0 to 0 in time 1. Phrased differently, the law of $(\mathbf{W}_t^1)_{t\in[0,1]}$ is $\pi_* \mathbb{P}_1^x$, the law of $(\mathbf{W}_t^2)_{t\in[0,1]}$ is $\pi_* \mathbb{P}_1^{-x}$ and the law of $\frac{1}{\sqrt{2}}(\mathbf{W}_t^1 + \mathbf{W}_t^2)_{t\in[0,1]}$ is $\pi_* \mathbb{P}_1^0$. By using

$$\left\{ \sup_{0\leq t\leq 1} |\mathbf{W}_t^1| < r, \sup_{0\leq t\leq 1} |\mathbf{W}_t^2| < r \right\} \subset \left\{ \sup_{0\leq t\leq 1} \frac{|\mathbf{W}_t^1 + \mathbf{W}_t^2|}{\sqrt{2}} < \sqrt{2}r \right\}$$

as well as the property $(\pi_* \mathbb{P}_1^x)(D_r) = (\pi_* \mathbb{P}_1^{-x})(D_r)$, and the independence of the Brownian bridges $(\mathbf{W}_t^1)_{t\in[0,1]}$ and $(\mathbf{W}_t^2)_{t\in[0,1]}$, we deduce that

$$\left(\left(\pi_* \mathbb{P}_1^x\right)(D_r)\right)^2 = \left(\left(\pi_* \mathbb{P}_1^x\right)(D_r)\right)\left(\left(\pi_* \mathbb{P}_1^{-x}\right)(D_r)\right) \leq \left(\pi_* \mathbb{P}_1^0\right)\left(D_{\sqrt{2}r}\right) \ .$$

The claimed result follows upon taking square roots. $\square$

We turn to the proof of Conjecture 4.2.4 for $N = 1$.

Lemma 4.2.7 (Conjecture 4.2.4 for $N = 1$)**.** *For all $x \in \mathbb{R}^d$ and all $r > 0$, we have*

$$\left(\pi_* \mathbb{P}_1^x\right)(D_r) \leq \left(\pi_* \mathbb{P}_1^0\right)(D_r) \ .$$

To prove Lemma 4.2.7, we need the following result from convex geometry. Its proof is due to Katarzyna Wyczesany.

Lemma 4.2.8. *Let $U, V \subset \mathbb{R}^k$ be open convex subsets which are point-symmetric about the origin. Let μ denote Lebesgue measure on $\mathbb{R}^k$. Then, for all $x \in \mathbb{R}^k$, we have*

$$\mu\left(U \cap (V + x)\right) \leq \mu\left(U \cap V\right) \ ,$$

where $V + x = \{y + x \colon y \in V\} \subset \mathbb{R}^k$.

Proof. Since U and V are point-symmetric about the origin, it follows that

$$U \cap (V + x) = -\left(U \cap (V - x)\right) \ ,$$

and therefore

$$\mu\left(U \cap (V + x)\right) = \mu\left(U \cap (V - x)\right) \ . \tag{4.2.9}$$

The convexity of U implies that $\frac{y+z}{2} \in U$ for $y, z \in U$. Similarly, if $y \in V + x$ and $z \in V - x$ then the convexity of V and

$$\frac{y + z}{2} = \frac{y - x + z + x}{2}$$

give $\frac{y+z}{2} \in V$. We deduce that

$$\begin{aligned}
\frac{1}{2}\left(U \cap (V + x)\right) + \frac{1}{2}\left(U \cap (V - x)\right) &= \left\{\frac{y + z}{2} \colon y \in U \cap (V + x), z \in U \cap (V - x)\right\} \\
&= \left\{\frac{y + z}{2} \colon y, z \in U \text{ and } y \in V + x, z \in V - x\right\} \\
&\subset U \cap V \ ,
\end{aligned}$$

which yields

$$\mu\left(\frac{1}{2}\left(U \cap (V + x)\right) + \frac{1}{2}\left(U \cap (V - x)\right)\right) \leq \mu\left(U \cap V\right) \ . \tag{4.2.10}$$

By a multiplicative version of the Brunn-Minkowski inequality, cf. [Gru07, Theorem 8.15], we have

$$\mu\left(U \cap (V + x)\right)^{1/2} \mu\left(U \cap (V - x)\right)^{1/2} \leq \mu\left(\frac{1}{2}\left(U \cap (V + x)\right) + \frac{1}{2}\left(U \cap (V - x)\right)\right) \ .$$

From (4.2.9) we further obtain $\mu(U \cap (V + x))^{1/2} \mu(U \cap (V - x))^{1/2} = \mu(U \cap (V + x))$, and the claimed result follows by (4.2.10). $\square$

This lemma allows us to generalise standard rearrangement inequalities as needed in the following proof.

Proof of Lemma 4.2.7. For $t_1, \ldots, t_k \in (0,1)$, we deduce that

$$\begin{aligned}
&(\pi_* \mathbb{P}_1^x)\left(\{v \colon v_{t_1} \in (-r,r), \ldots, v_{t_k} \in (-r,r)\}\right) \\
&\quad = (\pi_* \mathbb{P})\left(\left\{v \colon v_{t_1} - t_1 v_1 + t_1 x^1 \in (-r,r), \ldots, v_{t_k} - t_k v_1 + t_k x^1 \in (-r,r)\right\}\right) \\
&\quad = (\pi_* \mathbb{P})\left(\left\{v \colon v_{t_1} - t_1 v_1 \in (-r - t_1 x^1, r - t_1 x^1), \ldots, v_{t_k} - t_k v_1 \in (-r - t_k x^1, r - t_k x^1)\right\}\right) \\
&\quad = \left(\pi_* \mathbb{P}_1^0\right)\left(\left\{v \colon v_{t_1} \in (-r - t_1 x^1, r - t_1 x^1), \ldots, v_{t_k} \in (-r - t_k x^1, r - t_k x^1)\right\}\right) . \qquad (4.2.11)
\end{aligned}$$

Let $R^x \subset \mathbb{R}^k$ be the open subset which is enclosed by the hyperrectangle whose vertices are

$$(-r - t_1 x^1, 0, \ldots, 0), (r - t_1 x^1, 0, \ldots, 0), \ldots, (0, \ldots, 0, -r - t_k x^1), (0, \ldots, 0, r - t_k x^1) .$$

Recall that the random vector $(B^1_{t_1} - t_1 B^1_1, \ldots, B^1_{t_k} - t_k B^1_1)$ in $\mathbb{R}^k$ has a multivariate normal distribution with mean zero and $k \times k$ covariance matrix Σ given by

$$\Sigma_{ij} = \min\{t_i, t_j\} - t_i t_j \quad \text{for} \quad 1 \le i, j \le k .$$

In particular, the level sets of the corresponding density function $\rho \colon \mathbb{R}^k \to \mathbb{R}$ are ellipsoids, all of which are similar to each other and oriented along the same axes. By the layer cake representation of a non-negative measurable function, we have

$$\rho(z) = \int_0^\infty \mathbb{1}_{\{y \colon \rho(y) > s\}}(z) \, \mathrm{d}s .$$

Using this expression, we obtain

$$\begin{aligned}
&\left(\pi_* \mathbb{P}_1^0\right)\left(\left\{v \colon v_{t_1} \in (-r - t_1 x^1, r - t_1 x^1), \ldots, v_{t_k} \in (-r - t_k x^1, r - t_k x^1)\right\}\right) \\
&\qquad = \int_{\mathbb{R}^k} \rho(z) \mathbb{1}_{R^x}(z) \, \mathrm{d}z \\
&\qquad = \int_{\mathbb{R}^k} \int_0^\infty \mathbb{1}_{\{y \colon \rho(y) > s\}}(z) \mathbb{1}_{R^x}(z) \, \mathrm{d}s \, \mathrm{d}z \\
&\qquad = \int_0^\infty \mu\left(\{\rho > s\} \cap R^x\right) \, \mathrm{d}s , \qquad (4.2.12)
\end{aligned}$$

where μ denotes Lebesgue measure on $\mathbb{R}^k$. We observe that, for all $s \in (0, \infty)$, the set $\{\rho > s\} \subset \mathbb{R}^k$ is an open subset enclosed by an ellipsoid centred at the origin. Thus, both $\{\rho > s\}$ and R^0 are open convex subsets of $\mathbb{R}^k$ which are point-symmetric about the origin. We further have $R^x = R^0 - (t_1 x^1, \ldots, t_k x^1)$. Therefore, Lemma 4.2.8 applies to give

$$\mu\left(\{\rho > s\} \cap R^x\right) \le \mu\left(\{\rho > s\} \cap R^0\right) .$$

It follows, upon reversing the steps in (4.2.12), that

$$\begin{aligned}
\left(\pi_*\mathbb{P}_1^0\right)&\left(\left\{v\colon v_{t_1}\in(-r-t_1x^1, r-t_1x^1),\ldots,v_{t_k}\in(-r-t_kx^1,r-t_kx^1)\right\}\right)\\
&=\int_0^\infty \mu\left(\{\rho>s\}\cap R^x\right)\,\mathrm{d}s\\
&\le\int_0^\infty \mu\left(\{\rho>s\}\cap R^0\right)\,\mathrm{d}s\\
&=\left(\pi_*\mathbb{P}_1^0\right)\left(\{v\colon v_{t_1}\in(-r,r),\ldots,v_{t_k}\in(-r,r)\}\right)\ .
\end{aligned}$$

Together with the identity (4.2.11) and using the notation

$$\left(\pi_*\mathbb{P}_1^x\right)_{t_1,\ldots,t_k}(D_r)=\left(\pi_*\mathbb{P}_1^x\right)\left(\{v\colon v_{t_1}\in(-r,r),\ldots,v_{t_k}\in(-r,r)\}\right)\ ,$$

this implies that, for all $k\in\mathbb{N}$ and $t_1,\ldots,t_k\in(0,1)$,

$$\left(\pi_*\mathbb{P}_1^x\right)_{t_1,\ldots,t_k}(D_r)\le\left(\pi_*\mathbb{P}_1^0\right)_{t_1,\ldots,t_k}(D_r)\ . \tag{4.2.13}$$

By continuity of almost all Brownian motion sample paths, we have

$$\left(\pi_*\mathbb{P}_1^x\right)(D_r)=\left(\pi_*\mathbb{P}\right)\left(\bigcap_{t\in\mathbb{Q}\cap(0,1)}\left\{v\colon v_t-tv_1+tx^1\in(-r,r)\right\}\right)\ .$$

Let $(t_k)_{k\in\mathbb{N}}$ be an enumeration of the rationals inside the interval $(0,1)$. Due to the reverse monotone convergence theorem, it follows that

$$\left(\pi_*\mathbb{P}\right)\left(\bigcap_{t\in\mathbb{Q}\cap(0,1)}\left\{v\colon v_t-tv_1+tx^1\in(-r,r)\right\}\right)=\lim_{k\to\infty}\left(\pi_*\mathbb{P}_1^x\right)_{t_1,\ldots,t_k}(D_r)\ .$$

Using (4.2.13), we conclude

$$\left(\pi_*\mathbb{P}_1^x\right)(D_r)=\lim_{k\to\infty}\left(\pi_*\mathbb{P}_1^x\right)_{t_1,\ldots,t_k}(D_r)\le\lim_{k\to\infty}\left(\pi_*\mathbb{P}_1^0\right)_{t_1,\ldots,t_k}(D_r)=\left(\pi_*\mathbb{P}_1^0\right)(D_r)\ ,$$

as claimed. □

We finally remark that Lemma 4.2.7 together with independence of the components of a Brownian bridge in $\mathbb{R}^d$ give the following result.

Theorem 4.2.9. *Set*

$$E_r=\left\{v\in\Omega^0(\mathbb{R}^d)\colon\ \max_{1\le i\le d}\sup_{0\le t\le 1}\left|v_t^i\right|<r\right\}\ .$$

Then, for all $x\in\mathbb{R}^d$ and all $r>0$, we have $\mathbb{P}_1^x(E_r)\le\mathbb{P}_1^0(E_r)$.

Bibliography

[ABB16] Andrei A. Agrachev, Davide Barilari, and Ugo Boscain. *Introduction to Riemannian and Sub-Riemannian geometry*, Version 20 November 2016. `https://webusers.imj-prg.fr/~davide.barilari/Notes.php`.

[Agr14] Andrei A. Agrachev. Some open problems. In *Geometric Control Theory and sub-Riemannian Geometry*, volume 5 of *Springer INdAM Series*, pages 1–13. Springer, Cham, 2014.

[AKS93] Shigeki Aida, Shigeo Kusuoka, and Daniel W. Stroock. On the support of Wiener functionals. In *Asymptotic problems in probability theory: Wiener functionals and asymptotics. Proceedings of the 26th Taniguchi international symposium, Sanda and Kyoto, 1990*, pages 3–34. Longman Scientific & Technical, Harlow, 1993.

[BA88] Gérard Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Annales scientifiques de l'É.N.S.*, 21(3):307–331, 1988.

[BA89] Gérard Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale. *Annales de l'institut Fourier*, 39(1):73–99, 1989.

[BAL91a] Gérard Ben Arous and Rémi Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale (I). *Probability Theory and Related Fields*, 90(2):175–202, 1991.

[BAL91b] Gérard Ben Arous and Rémi Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale (II). *Probability Theory and Related Fields*, 90(3):377–402, 1991.

[Bau04] Fabrice Baudoin. *An Introduction to the Geometry of Stochastic Flows*. Imperial College Press, 2004.

[BBN12] Davide Barilari, Ugo Boscain, and Robert W. Neel. Small-time heat kernel asymptotics at the sub-Riemannian cut locus. *Journal of Differential Geometry*, 92(3):373–416, 2012.

[BBS16a] Davide Barilari, Ugo Boscain, and Mario Sigalotti, editors. *Geometry, Analysis and Dynamics on sub-Riemannian Manifolds. Volume I. Lecture notes from the IHP Trimester, Institut Henri Poincaré, Paris, France and from the CIRM summer school "Sub-Riemannian manifolds: from geodesics to hypoelliptic diffusion", Luminy, 2014.* European Mathematical Society, Zürich, 2016.

[BBS16b] Davide Barilari, Ugo Boscain, and Mario Sigalotti, editors. *Geometry, Analysis and Dynamics on sub-Riemannian Manifolds. Volume II.* European Mathematical Society, Zürich, 2016.

[Bel87] Denis R. Bell. *The Malliavin Calculus*, volume 34 of *Pitman Monographs and Surveys in Pure and Applied Mathematics.* Longman Scientific & Technical, Harlow, 1987.

[Bel96] André Bellaïche. The tangent space in sub-Riemannian geometry. In *Sub-Riemannian Geometry*, volume 144 of *Progress in Mathematics*, pages 1–78. Birkhäuser, 1996.

[BGLY16] Horatio Boedihardjo, Xi Geng, Terry Lyons, and Danyu Yang. The signature of a rough path: Uniqueness. *Advances in Mathematics*, 293:720–737, 2016.

[Bis81] Jean-Michel Bismut. Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 56:469–505, 1981.

[Bis84] Jean-Michel Bismut. *Large Deviations and the Malliavin Calculus*, volume 45 of *Progress in Mathematics.* Birkhäuser, 1984.

[BL13] Ugo Boscain and Camille Laurent. The Laplace-Beltrami operator in almost-Riemannian Geometry. *Université de Grenoble. Annales de l'Institut Fourier*, 63(5):1739–1770, 2013.

[BMN15] Ismael Bailleul, Laurent Mesnager, and James R. Norris. *Small-time fluctuations for the bridge of a sub-Riemannian diffusion.* `arXiv:1505.03464`, 13 May 2015.

[Bro82] Roger W. Brockett. Control theory and singular Riemannian geometry. In *New directions in Applied Mathematics, Cleveland, Ohio, 1980*, pages 11–27. Springer, Berlin, 1982.

[Bro84] Roger W. Brockett. Nonlinear control theory and differential geometry. In *Proceedings of the International Congress of Mathematicians, Warsaw, 1983*, volume 2, pages 1357–1368. Polish Scientific Publishers, Warsaw, 1984.

[BS90] Rosa Maria Bianchini and Gianna Stefani. Graded approximations and controllability along a trajectory. *SIAM Journal on Control and Optimization*, 28(4):903–924, 1990.

[Car09] Constantin Carathéodory. Untersuchungen über die Grundlagen der Thermodynamik. *Mathematische Annalen*, 67(3):355–386, 1909.

[Car31] Élie Cartan. Sur la représentation géométrique des systèmes matériels non holonomes. In *Proceedings of the International Congress of Mathematicians, Bologna, 1928*, volume 4, pages 253–261. Bologna, 1931.

[CC09] Ovidiu Calin and Der-Chen Chang. *Sub-Riemannian Geometry*, volume 126 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2009.

[CH85] Shiing-Shen Chern and Richard S. Hamilton. On Riemannian metrics adapted to three-dimensional contact manifolds. In *Workshop Bonn 1984, Bonn, 1984*, volume 1111 of *Lecture Notes in Mathematics*, pages 279–308. Springer, Berlin, 1985.

[Che57] Kuo-Tsai Chen. Integration of Paths, Geometric Invariants and a Generalized Baker-Hausdorff Formula. *Annals of Mathematics. Second Series*, 65(1):163–178, 1957.

[Che58] Kuo-Tsai Chen. Integration of paths – a faithful representation of paths by non-commutative formal power series. *Transactions of the American Mathematical Society*, 89(2):395–407, 1958.

[Cho39] Wei-Liang Chow. Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. *Mathematische Annalen*, 117:98–105, 1939.

[Die69] Jean A. Dieudonné. *Foundations of Modern Analysis*. Academic Press, New York and London, 1969. Enlarged and corrected printing, Pure and Applied Mathematics, Vol. 10-I.

[FV10] Peter K. Friz and Nicolas B. Victoir. *Multidimensional Stochastic Processes as Rough Paths. Theory and Applications*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[Gru07] Peter M. Gruber. *Convex and Discrete Geometry*, volume 336 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Berlin, 2007.

[Hab17] Karen Habermann. Small-time fluctuations for sub-Riemannian diffusion loops. *Probability Theory and Related Fields*, Published online 19 June 2017. DOI 10.1007/s00440-017-0788-7.

[Ham90] Ursula Hamenstädt. Some regularity theorems for Carnot-Carathéodory metrics. *Journal of Differential Geometry*, 32(3):819–850, 1990.

[Her62] Robert Hermann. Some differential-geometric aspects of the Lagrange variational problem. *Illinois Journal of Mathematics*, 6:634–673, 1962.

[Her73] Robert Hermann. Geodesics of singular Riemannian metrics. *Bulletin of the American Mathematical Society*, 79:780–782, 1973.

[HL10] Ben Hambly and Terry J. Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Annals of Mathematics. Second Series*, 171(1):109–167, 2010.

[Hör67] Lars Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, 119:147–171, 1967.

[IT17] Yuzuru Inahama and Setsuo Taniguchi. Short time full asymptotic expansion of hypoelliptic heat kernel at the cut locus. *Forum of Mathematics, Sigma*, 5:e16, 74 pages, 2017.

[KOL91] Ioannis Karatzas, Daniel L. Ocone, and Jinlu Li. An extension of Clark's formula. *Stochastics and Stochastics Reports*, 37(3):127–131, 1991.

[KS84] Shigeo Kusuoka and Daniel W. Stroock. Applications of the Malliavin calculus. I. In *Stochastic Analysis, Katata and Kyoto, 1982*, volume 32 of *North-Holland Mathematical Library*, pages 271–306. North-Holland, Amsterdam, 1984.

[KS85] Shigeo Kusuoka and Daniel W. Stroock. Applications of the Malliavin calculus. II. *Journal of the Faculty of Science. University of Tokyo. Section IA. Mathematics*, 32(1):1–76, 1985.

[KS87] Shigeo Kusuoka and Daniel W. Stroock. Applications of the Malliavin calculus. III. *Journal of the Faculty of Science. University of Tokyo. Section IA. Mathematics*, 34(2):391–442, 1987.

[LCL07] Terry J. Lyons, Michael J. Caruana, and Thierry Lévy. *Differential equations driven by rough paths*, volume 1908 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.

[Léa92] Rémi Léandre. Développement asymptotique de la densité d'une diffusion dégénérée. *Forum Mathematicum*, 4(1):45–75, 1992.

[LJQ13] Yves Le Jan and Zhongmin Qian. Stratonovich's signatures of Brownian motion determine Brownian sample paths. *Probability Theory and Related Fields*, 157(1-2):209–223, 2013.

[Mal78a] Paul Malliavin. C^k-hypoellipticity with degeneracy. In *Stochastic Analysis*, pages 199–214. Academic Press, New York and London, 1978.

[Mal78b] Paul Malliavin. Stochastic Calculus of Variation and Hypoelliptic Operators. In *Proceedings of the International Symposium on Stochastic Differential Equations, Kyoto, 1976*, pages 195–263. Wiley, New York, 1978.

[Mol75] Stanislav A. Molchanov. Diffusion processes and Riemannian geometry. *Russian Mathematical Surveys*, 30(1):1–63, 1975.

[Mon95] Richard Montgomery. A survey of singular curves in sub-Riemannian geometry. *Journal of Dynamical and Control Systems*, 1(1):49–90, 1995.

[Mon02] Richard Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, volume 91 of *Mathematical Surveys and Monographs.* American Mathematical Society, 2002.

[MT06] Paul Malliavin and Anton Thalmaier. *Stochastic Calculus of Variations in Mathematical Finance.* Springer Finance. Springer, Berlin, 2006.

[Nor86] James R. Norris. Simplified Malliavin calculus. *Séminaire de probabilités (Strasbourg)*, 20:101–130, 1986.

[Nua98] David Nualart. Analysis on Wiener space and anticipating stochastic calculus. In *Lectures on Probability Theory and Statistics, Saint-Flour, 1995*, volume 1690 of *Lecture Notes in Mathematics*, pages 123–227. Springer, Berlin, 1998.

[Nua06] David Nualart. *The Malliavin Calculus and Related Topics.* Probability and its Applications. Springer, Berlin, 2nd edition, 2006.

[Nua09] David Nualart. *Malliavin Calculus and Its Applications*, volume 110 of *CBMS Regional Conference Series in Mathematics.* Published for the Conference Board of the Mathematical Sciences by the American Mathematical Society, Providence, RI, 2009.

[Pal59] Richard S. Palais. Natural Operations on Differential Forms. *Transactions of the American Mathematical Society*, 92(1):125–141, 1959.

[Pan89] Pierre Pansu. Métriques de Carnot-Carathéodory et Quasiisométries des Espaces Symétriques de rang un. *Annals of Mathematics. Second Series*, 129(1):1–60, 1989.

[Pao17] Elisa Paoli. Small time asymptotics on the diagonal for Hörmander's type hypoelliptic operators. *Journal of Dynamical and Control Systems*, 23(1):111–143, 2017.

[Ras38] Pjotr K. Rashevskii. About connecting two arbitrary points of a complete non-holonomic space by an admissible curve (in Russian). *Uchenye Zapiski – Moskovskii pedagogicheskii institut imeni Karla Libknekhta*, 2:83–94, 1938.

[RS76] Linda P. Rothschild and Elias M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Mathematica*, 137(1):247–320, 1976.

[Sas60] Shigeo Sasaki. On differentiable manifolds with certain structures which are closely related to almost contact structure. I. *The Tohoku Mathematical Journal. Second Series*, 12:459–476, 1960.

[Shi80] Ichiro Shigekawa. Derivatives of Wiener functionals and absolute continuity of induced measures. *Journal of Mathematics of Kyoto University*, 20(2):263–289, 1980.

[Str81a] Daniel W. Stroock. The Malliavin calculus: a functional analytic approach. *Journal of Functional Analysis*, 44(2):212–257, 1981.

[Str81b] Daniel W. Stroock. The Malliavin calculus and its application to second order parabolic differential equations. II. *Mathematical Systems Theory*, 14(2):141–171, 1981.

[Str83] Daniel W. Stroock. Some applications of stochastic calculus to partial differential equations. In *Eleventh Saint Flour Probability Summer School, Saint Flour, 1981*, volume 976 of *Lecture Notes in Mathematics*, pages 267–382. Springer, Berlin, 1983.

[Str86] Robert S. Strichartz. Sub-Riemannian geometry. *Journal of Differential Geometry*, 24(2):221–263, 1986.

[Str87] Robert S. Strichartz. The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations. *Journal of Functional Analysis*, 72(2):320–345, 1987.

[Str89] Robert S. Strichartz. Corrections to: "Sub-Riemannian geometry". *Journal of Differential Geometry*, 30(2):595–596, 1989.

[Wat84] Shinzo Watanabe. *Lectures on Stochastic Differential Equations and Malliavin Calculus*, volume 73 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics*. Springer, Berlin, 1984.

[Wat87] Shinzo Watanabe. Analysis of Wiener Functionals (Malliavin Calculus) and its Applications to Heat Kernels. *The Annals of Probability*, 15(1):1–39, 1987.

[Web78] Sidney M. Webster. Pseudo-Hermitian structures on a real hypersurface. *Journal of Differential Geometry*, 13(1):25–41, 1978.