Precise Local Estimates for Hypoelliptic Differential Equations driven by Fractional Brownian Motion

Xi Geng* Cheng Ouyang† Samy Tindel‡

Abstract

This article is concerned with stochastic differential equations driven by a d dimensional fractional Brownian motion with Hurst parameter $H > 1/4$, understood in the rough paths sense. Whenever the coefficients of the equation satisfy a uniform hypoellipticity condition, we establish a sharp local estimate on the associated control distance function and a sharp local lower estimate on the density of the solution. Our methodology relies heavily on the rough paths structure of the equation.

Contents

*School of Mathematics and Statistics, University of Melbourne, Melbourne, Australia. Email: xi.geng@unimelb.edu.au.

†Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, United States. Email: couyang@math.uic.edu.

‡Department of Mathematics, Purdue University, West Lafayette, United States. Email: stindel@purdue.edu.

1 Introduction.

We will split our introduction into two parts. In Section 1.1, we recall some background on the stochastic analysis of stochastic differential equations driven by a fractional Brownian motion. In Section 1.2 we describe our main results. Section 1.3 is then devoted to a brief explanation about the methodology we have used in order to obtain our main results.

1.1 Background and motivation.

One way to envision Malliavin calculus is to see it as a geometric and analytic framework on an infinite dimensional space (namely the Wiener space) equipped with a Gaussian measure. This is already apparent in Malliavin's seminal contribution [20] giving a probabilistic proof of Hörmander's theorem. The same point of view has then been pushed forward in the celebrated series of papers by Kusuoka and Stroock, which set up the basis for densities and probabilities expansions for diffusion processes within this framework.

On the other hand, the original perspective of Lyons' rough path theory (cf. [18, 19]) is quite different. Summarizing very briefly, it asserts that a reasonable differential calculus with respect to a noisy process X can be achieved as long as one can define enough iterated integrals of X. One of the first processes to which the theory has been successfully applied is a fractional Brownian motion, and we shall focus on this process in the present paper.

Namely a $\mathbb{R}^d$-valued fractional Brownian motion is a continuous centered Gaussian process $B = \{(B_t^1, \dots, B_t^d);\, t \geq 0\}$ with independent coordinates, such that each B^j satisfies

$$\mathbb{E}\left[(B_t^j - B_s^j)^2\right] = |t-s|^{2H}, \qquad s,t \geq 0,$$

for a given $H \in (0,1)$. The process B can be seen as a natural generalization of Brownian motion allowing any kind of Hölder regularity (that is a Hölder exponent $H - \varepsilon$ for an arbitrary small ε, whenever H is given). We are interested in the following differential equation driven by B:

$$\begin{cases} dX_t = \sum_{\alpha=1}^d V_\alpha(X_t) dB_t^\alpha, & 0 \leq t \leq 1, \\ X_0 = x \in \mathbb{R}^N. \end{cases} \tag{1.1}$$

Here the V_α's are C_b^∞ vector fields, and the Hurst parameter is assumed to satisfy the condition $H > 1/4$. In this setting, putting together the results contained in [8] and [19], the stochastic differential equation (1.1) can be understood in the framework of rough path theory. Although we will give an account on the notion of rough path solution in Section 2.3, the simplest way of looking at equation (1.1) is the following. Let $B_t^{(n)}$ be a dyadic linear interpolation of B_t. Let $X_t^{(n)}$ be the solution to equation (1.1) in which the driving process B_t is replaced by $B_t^{(n)}$. From standard ODE theory, $X_t^{(n)}$ is pathwisely well-defined. The solution to the SDE (1.1) is then proved (cf. [12] for instance) to be the limit of $X_t^{(n)}$ as $n \to \infty$.

With the solution of (1.1) in hand, a natural problem one can think of is the following: can we extend the aforementioned analytic studies on Wiener's space to the process B? In particular can we complete Kusuoka-Stroock's program in the fractional Brownian motion setting? This question has received a lot of attention in the recent years, and previous efforts along this line include Hörmander type theorems for the process X defined by (1.1) (cf. [2, 6, 7]), some upper Gaussian bounds on the density $p(t,x,y)$ of X_t (cf. [3]), as well as Varadhan type estimates for $\log(p(t,x,y))$ in small time [4]. One should stress at this point that the road from the Brownian to the fractional Brownian case is far from being trivial. This is essentially due to the lack of independence of the fBm increments and Markov property, as well as to the technically demanding characterization of the Cameron-Martin space whenever B is not a Brownian motion. We shall go back to those obstacles throughout the article.

Our contribution can be seen as a step in the direction mentioned above. More specifically, we shall obtain a sharp local estimate on the associated control distance function and a sharp local estimates for the density of X_t under hypoelliptic conditions on the vector fields V_α. This will be achieved thanks to a combination of geometric and analytic tools which can also be understood as a mix of stochastic analysis and rough path theory. We describe our main results more precisely in the next subsection.

1.2 Statement of main results.

Let us recall that equation (1.1) is our main object of concern. We are typically interested in the degenerate case where the vector fields $V = \{V_1, \dots, V_d\}$ satisfy the so-called *uniform*

hypoellipticity assumption to be defined shortly. This is a standard degenerate setting where one can expect that the solution of the SDE (1.1) admits a smooth density with respect to the Lebesgue measure. As mentioned in Section 1.1, we wish to obtain quantitative information for the density in this context.

We first formulate the uniform hypoellipticity condition which will be assumed throughout the rest of the paper. For $l \geq 1$, define $\mathcal{A}(l)$ to be the set of words over letters $\{1, \ldots, d\}$ with length at most l (including the empty word), and $\mathcal{A}_1(l) \triangleq \mathcal{A}(l) \backslash \{\emptyset\}$. Denote $\mathcal{A}_1$ as the set of all non-empty words. Given a word $\alpha \in \mathcal{A}_1$, we define the vector field $V_{[\alpha]}$ inductively by $V_{[i]} \triangleq V_i$ and $V_{[\alpha]} \triangleq [V_i, V_{[\beta]}]$ for $\alpha = (i, \beta)$ with i being a letter and $\beta \in \mathcal{A}_1$.

Uniform Hypoellipticity Assumption. *The vector fields $(V_1, \ldots, V_d)$ are C_b^∞, and there exists an integer $l_0 \geq 1$, such that*

$$\inf_{x \in \mathbb{R}^N} \inf_{\eta \in S^{N-1}} \left\{ \sum_{\alpha \in \mathcal{A}_1(l_0)} \langle V_{[\alpha]}(x), \eta \rangle_{\mathbb{R}^N}^2 \right\} > 0. \tag{1.2}$$

The smallest such l_0 is called the hypoellipticity constant for the vector fields.

Remark 1.1. The uniform hypoellipticity assumption is a quantitative description of the standard uniform Hörmander condition that the family of vectors $\{V_{[\alpha]}(x) : \alpha \in \mathcal{A}_1(l_0)\}$ span the tangent space $T_x\mathbb{R}^N$ uniformly with respect to $x \in \mathbb{R}^N$.

Under condition (1.2), it was proved by Cass-Friz [6] and Cass-Hairer-Litterer-Tindel [7] that the solution to the SDE (1.1) admits a smooth density $y \mapsto p(t, x, y)$ with respect to the Lebesgue measure on $\mathbb{R}^N$ for all $(t, x) \in (0, 1] \times \mathbb{R}^N$. Our contribution aims at getting quantitative small time estimates for $p(t, x, y)$.

In order to describe our bounds on the density $p(t, x, y)$, let us recall that the small time behavior of $p(t, x, y)$ is closely related to the so-called control distance function associated with the vector fields. This fact was already revealed in the Varadhan-type asymptotics result proved by Baudoin-Ouyang-Zhang [4]:

$$\lim_{t \to 0} t^{2H} \log p(t, x, y) = -\frac{1}{2} d(x, y)^2. \tag{1.3}$$

The control distance function $d(x, y)$ in (1.3) is defined in the following way. For any continuous path $h : [0, 1] \to \mathbb{R}^d$ with finite q-variation for some $1 \leq q < 2$, denote $\Phi_t(x; h)$ as the solution to the ODE

$$\begin{cases} dx_t = \sum_{\alpha=1}^d V_\alpha(x_t) dh_t^\alpha, & 0 \leq t \leq 1, \\ x_0 = x, \end{cases} \tag{1.4}$$

which is well-posed in the sense of Young [25] and Lyons [18]. Let $\bar{\mathcal{H}}$ be the Cameron-Martin subspace for the fractional Brownian motion with Hurst parameter H, whose definition will be recalled in Section 2.1. According to a variational embedding theorem due to Friz-Victoir [12] (cf. Proposition 2.6 in Section 2.1), every Cameron-Martin path $h \in \bar{\mathcal{H}}$ has finite q-variation for some $1 \leq q < 2$ so that the ODE (1.4) can be solved for all such h. With those preliminary considerations in hand, the distance d in (1.3) is given in the next definition.

Definition 1.2. For $x, y \in \mathbb{R}^N$ and $\Phi_t(x;h)$ defined as in (1.4), set

$$\Pi_{x,y} \triangleq \left\{ h \in \bar{\mathcal{H}} : \Phi_1(x;h) = y \right\} \tag{1.5}$$

to be the space of Cameron-Martin paths which *join x to y in the sense of differential equations*. The *control distance function* $d(x,y) = d_H(x,y)$ is defined by

$$d(x,y) \triangleq \inf \left\{ \|h\|_{\bar{\mathcal{H}}} : h \in \Pi_{x,y} \right\}, \quad x, y \in \mathbb{R}^N.$$

According to (1.3), one can clearly expect that the Cameron-Martin structure and the control distance function will play an important role in understanding the small time behavior of $p(t,x,y)$. However, unlike the diffusion case and due to the complexity of the Cameron-Martin structure, the function $d(x,y)$ is far from being a metric and its shape is not clear. Our first main result is thus concerned with the local behavior of $d(x,y)$. It establishes a comparison between d and the Euclidian distance.

Theorem 1.3. *Under the same assumptions as in Theorem 1.4, let l_0 be the hypoellipticity constant in assumption (1.2) and d be the control distance given in Definition 1.2. There exist constants $C_1, C_2, \delta > 0$, where C_1, C_2 depend only on H, l_0 and the vector fields, and where δ depends only on l_0 and the vector fields, such that*

$$C_1|x-y| \leq d(x,y) \leq C_2|x-y|^{\frac{1}{l_0}}, \tag{1.6}$$

for all $x, y \in \mathbb{R}^N$ with $|x-y| < \delta$.

We are in fact able to establish a stronger result, namely, the local equivalence of d to the sub-Riemannian distance induced by the vector fields $\{V_1, ..., V_d\}$. More specifically, let us write the distance given in Definition 1.2 as $d_H(x,y)$, in order to emphasize the dependence on the Hurst parameter H. Our second main result asserts that all the distances d_H are locally equivalent.

Theorem 1.4. *Assume that the vector fields $(V_1, \ldots, V_d)$ satisfy the uniform hypoellipticity condition (1.2). For $H \in (1/4, 1)$, consider the distance d_H given in Definition 1.2. Then for any $H_1, H_2 \in (1/4, 1)$, there exist constants $C = C(H_1, H_2, V) > 0$ and $\delta > 0$ such that*

$$\frac{1}{C} d_{H_1}(x,y) \leq d_{H_2}(x,y) \leq C d_{H_1}(x,y), \tag{1.7}$$

for all $x, y \in \mathbb{R}^N$ with $|x-y| < \delta$. In particular, all distances d_H are locally equivalent to $d_{\mathrm{BM}} \equiv d_{1/2}$, where d_{BM} stands for the controlling distance of the system (1.4) driven by a Brownian motion, i.e. the sub-Riemannian distance induced by the vector fields $\{V_1, ..., V_d\}$.

Remark 1.5. In the special case when (1.1) reads as $dX_t = X_t \otimes dB_t$, that is, when X_t is the truncated signature of B up to order $l > 0$, it is proved in [1] that all $d_H(x,y)$ are globally equivalent. The proof crucially depends on the fact that the signature of B is homogeneous with respect to the dilation operator on $G^{(l)}(\mathbb{R}^d)$, the free nilpotent Lip group over $\mathbb{R}^d$ of order l. In the current general nonlinear case, the local equivalence is much more technically challenging. In addition, we believe that the global equivalence of the distances d_H does not hold.

Our second main result asserts that the density $p(t,x,y)$ of X_t is strictly positive everywhere whenever $t>0$. It generalizes for the first time the result of [3, Theorem 1.4] to a general hypoelliptic case, by affirming that Hypothesis 1.2 in that theorem is always verified under our assumption (1.2). Recall that a *distribution* over a differentiable manifold is a smooth choice of subspace of the tangent space at every point with constant dimension.

Theorem 1.6. *Let $\{V_1,\dots,V_d\}$ be a family of C_b^∞-vector fields on $\mathbb{R}^N$, which span a distribution over $\mathbb{R}^N$ and satisfy the uniform hypoellipticity assumption (1.2). Let X_t^x be the solution to the stochastic differential equation (1.1), where B_t is a d-dimensional fractional Brownian motion with Hurst parameter $H>1/4$. Then for each $t\in(0,1]$, the density of X_t is everywhere strictly positive.*

As we will see in Section 5.1, the proof of the above result is based on finite dimensional geometric arguments such as the classical Sard theorem, as well as a general positivity criteria for densities on the Wiener space. We believe that the positivity result in Theorem 1.6 is non-trivial and interesting in its own right.

Let us now turn to a description of our third main result. It establishes a sharp local lower estimate for the density function $p(t,x,y)$ of the solution to the SDE (1.1) in small time.

Theorem 1.7. *Under the uniform hypoelliptic assumption (1.2), let $p(t,x,y)$ be the density of the random variable X_t defined by equation (1.1). There exist some constants $C,\tau>0$ depending only on H,l_0 and the vector fields V_α, such that*

$$p(t,x,y)\geq \frac{C}{|B_d(x,t^H)|}, \tag{1.8}$$

for all $(t,x,y)\in(0,1]\times\mathbb{R}^N\times\mathbb{R}^N$ satisfying the following local condition involving the distance d introduced in Definition 1.2:

$$d(x,y)\leq t^H, \quad \text{and} \quad t<\tau.$$

In relation (1.8), $B_d(x,t^H)\triangleq\{z\in\mathbb{R}^N : d(x,z)<t^H\}$ denotes the ball with respect to the distance d and $|\cdot|$ stands for the Lebesgue measure.

The sharpness of Theorem 1.7 can be seen from the fractional Brownian motion case, i.e. when $N=d$ and $V=\mathrm{Id}$. As we will see, the technique we use to prove Theorem 1.3 will be an essential ingredient for establishing Theorem 1.7. Theorem 1.4 and Theorem 1.6 will also be proved as byproducts along our path of proving Theorem 1.7.

1.3 Strategy and outlook.

Let us say a few words about the methodology we have used in order to obtain our main results. Although we will describe our overall strategy with more details in Section 5, let us mention here that it is based on the reduction of the problem to a finite dimensional one, plus some geometric type arguments.

More specifically, the key point in our proofs is that the solution X_t to (1.1) can be approximated by a simple enough function of the so-called truncated signature of order l for the fractional Brownian motion B. This object is formally defined, for a given $l \geq 1$, as the following $\oplus_{k=0}^{l}(\mathbb{R}^d)^{\otimes k}$-valued process:

$$\Gamma_t = 1 + \sum_{k=1}^{l} \int_{0<t_1<\cdots<t_k<t} dB_{t_1} \otimes \cdots \otimes dB_{t_k},$$

and it enjoys some convenient algebraic and analytic properties. The truncated signature is the main building block of the rough path theory (see e.g [19]), and was also used in [17] in a Malliavin calculus context. Part of our challenge in the current contribution is to combine the properties of the process Γ, together with the Cameron-Martin space structure related to the fractional Brownian motion B, in order to achieve efficient bounds for the density of X_t.

As mentioned above, the truncated signature gives rise to a l-th order local approximation of X_t in a neighborhood of its initial condition x. Namely if we set

$$F_l(\Gamma_t, x) \triangleq \sum_{k=1}^{l} \sum_{i_1,\ldots,i_k=1}^{d} V_{(i_1,\ldots,i_k)}(x) \int_{0<t_1<\cdots<t_k<t} dB_{t_1}^{i_1} \cdots dB_{t_k}^{i_k}, \tag{1.9}$$

then classical rough paths considerations assert that $F_l(\Gamma_t, x)$ is an approximation of order t^{Hl} of X_t for small t. In the sequel we will heavily rely on some non degeneracy properties of F_l derived from the uniform hypoelliptic assumption (1.2), in order to get the following information:

(i) One can construct a path h in the Cameron-Martin space of B which joins x and any point y in a small enough neighborhood of x. This task is carried out thanks to a complex iteration procedure, whose building block is the non-degeneracy of the function F_l. It is detailed in Section 4.2. In this context, observe that the computation of the Cameron-Martin norm of h also requires a substantial effort. This will be the key step in order to prove Theorems 1.4 and 1.3 concerning the distance d given in Definition 1.2.

(ii) The proof of the lower bound given in Theorem 1.7 also hinges heavily on the approximation F_l given by (1.9). Indeed the preliminary results about the density of Γ_t (see Remark 1.5 above), combined with the non-degeneracy of F_l, yield good properties for the density of $F_l(\Gamma_t, x)$. One is then left with the task of showing that $F_l(\Gamma_t, x)$ approximates X_t properly at the density level.

In conclusion, although the steps performed in the remainder of the article might look technically and computationally involved, they rely on a natural combination of analytic and geometric bricks as well as a reduction to a finite dimensional problem. Let us also highlight the fact that our next challenge is to iterate the local estimates presented here in order to get Gaussian type lower bounds for the density $p(t, x, y)$ of X_t. This causes some further complications due to the complex (non Markovian) dependence structure for the increments of the fractional Brownian motion B. We defer this project to a future publication.

Organization of the present paper. In Section 2, we present some basic notions from the analysis of fractional Brownian motion and rough path theory. In Section 3, we give an

independent discussion in the elliptic case in which the analysis is considerably simpler. In Section 4 and Section 5, we develop the proofs of Theorem 1.3 and Theorem 1.7 respectively in the hypoelliptic case. Theorem 1.4 and Theorem 1.6 are proved in the steps towards proving Theorem 1.7.

Notation. Throughout the rest of this paper, we use "$\text{Letter}_{\text{subscript}}$" to denote constants whose value depend only on objects specified in the "subscript" and may differ from line to line. For instance, C_{H,V,l_0} denotes a constant depending only on the Hurst parameter H, the vector fields V and the hypoellipticity constant l_0. Unless otherwise stated, a constant will implicitly depend on H, V, l_0. We will always omit the dependence on dimension.

2 Preliminary results.

This section is devoted to some preliminary results on the Cameron-Martin space related to a fractional Brownian motion. We shall also recall some basic facts about rough paths solutions to noisy equations.

2.1 The Cameron-Martin subspace of fractional Brownian motion.

Let us start by recalling the definition of fractional Brownian motion.

Definition 2.1. A d-dimensional *fractional Brownian motion* with Hurst parameter $H \in (0,1)$ is an $\mathbb{R}^d$-valued continuous centered Gaussian process $B_t = (B_t^1, \dots, B_t^d)$ whose covariance structure is given by

$$\mathbb{E}[B_s^i B_t^j] = \frac{1}{2}\left(s^{2H} + t^{2H} - |s-t|^{2H}\right)\delta_{ij} \triangleq R(s,t)\delta_{ij}. \tag{2.1}$$

This process is defined and analyzed in numerous articles (cf. [10, 23, 24] for instance), to which we refer for further details. In this section, we mostly focus on a proper definition of the Cameron-Martin subspace related to B. We also prove two general lemmas about this space which are needed for our analysis of the density $p(t,x,y)$. Notice that we will frequently identify a Hilbert space with its dual in the canonical way without further mentioning.

In order to introduce the Hilbert spaces which will feature in the sequel, consider a one dimensional fractional Brownian motion $\{B_t : 0 \le t \le 1\}$ with Hurst parameter $H \in (0,1)$. The discussion here can be easily adapted to the multidimensional setting with arbitrary time horizon $[0,T]$. Denote W as the space of continuous paths $w : [0,1] \to \mathbb{R}^1$ with $w_0 = 0$. Let $\mathbb{P}$ be the probability measure over W under which the coordinate process $B_t(w) = w_t$ becomes a fractional Brownian motion. Let $\mathcal{C}_1$ be the associated first order Wiener chaos, i.e. $\mathcal{C}_1 \triangleq \overline{\text{Span}\{B_t : 0 \le t \le 1\}}$ in $L^2(W, \mathbb{P})$.

Definition 2.2. Let B be a one dimensional fractional Brownian motion as defined in (2.1). Define $\bar{\mathcal{H}}$ to be the space of elements $h \in W$ which can be written as

$$h_t = \mathbb{E}[B_t Z], \quad 0 \le t \le 1, \tag{2.2}$$

where $Z \in \mathcal{C}_1$. We equip $\bar{\mathcal{H}}$ with an inner product structure given by

$$\langle h_1, h_2 \rangle_{\bar{\mathcal{H}}} \triangleq \mathbb{E}[Z_1 Z_2], \quad h_1, h_2 \in \bar{\mathcal{H}},$$

whenever h^1, h^2 are defined by (2.2) for two random variables $Z_1, Z_2 \in \mathcal{C}_1$. The Hilbert space $(\bar{\mathcal{H}}, \langle \cdot, \cdot \rangle_{\bar{\mathcal{H}}})$ is called the *Cameron-Martin subspace* of the fractional Brownian motion.

One of the advantages of working with fractional Brownian motion is that a convenient analytic description of $\bar{\mathcal{H}}$ in terms of fractional calculus is available (cf. [10]). Namely recall that given a function f defined on $[a, b]$, the right and left *fractional integrals* of f of order $\alpha > 0$ are respectively defined by

$$(I^{\alpha}_{a^+} f)(t) \triangleq \frac{1}{\Gamma(\alpha)} \int_a^t f(s)(t-s)^{\alpha-1} ds, \quad \text{and} \quad (I^{\alpha}_{b^-} f)(t) \triangleq \frac{1}{\Gamma(\alpha)} \int_t^b f(s)(s-t)^{\alpha-1} ds. \tag{2.3}$$

In the same way the right and left *fractional derivatives* of f of order $\alpha > 0$ are respectively defined by

$$(D^{\alpha}_{a^+} f)(t) \triangleq \left(\frac{d}{dt}\right)^{[\alpha]+1} (I^{1-\{\alpha\}}_{a^+} f)(t), \quad \text{and} \quad (D^{\alpha}_{b^-} f)(t) \triangleq \left(-\frac{d}{dt}\right)^{[\alpha]+1} (I^{1-\{\alpha\}}_{b^-} f)(t), \tag{2.4}$$

where $[\alpha]$ is the integer part of α and $\{\alpha\} \triangleq \alpha - [\alpha]$ is the fractional part of α. The following formula for $D^{\alpha}_{a^+}$ will be useful for us:

$$(D^{\alpha}_{a^+} f)(t) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(t)}{(t-a)^{\alpha}} + \alpha \int_a^t \frac{f(t) - f(s)}{(t-s)^{\alpha+1}} ds \right), \quad t \in [a, b]. \tag{2.5}$$

The fractional integral and derivative operators are inverse to each other. For this and other properties of fractional derivatives, the reader is referred to [14].

Let us now go back to the construction of the Cameron-Martin space for B, and proceed as in [10]. Namely define an isomorphism K between $L^2([0,1])$ and $I^{H+1/2}_{0+}(L^2([0,1]))$ in the following way:

$$K\varphi \triangleq \begin{cases} C_H \cdot I^1_{0+} \left(t^{H-\frac{1}{2}} \cdot I^{H-\frac{1}{2}}_{0+} \left(s^{\frac{1}{2}-H} \varphi(s) \right)(t) \right), & H > \frac{1}{2}; \\ C_H \cdot I^{2H}_{0+} \left(t^{\frac{1}{2}-H} \cdot I^{\frac{1}{2}-H}_{0+} \left(s^{H-\frac{1}{2}} \varphi(s) \right)(t) \right), & H \leq \frac{1}{2}, \end{cases} \tag{2.6}$$

where c_H is a universal constant depending only on H. One can easily compute K^{-1} from the definition of K in terms of fractional derivatives. Moreover, the operator K admits a kernel representation, i.e. there exits a function $K(t, s)$ such that

$$(K\varphi)(t) = \int_0^t K(t, s)\varphi(s) ds, \quad \varphi \in L^2([0,1]).$$

The kernel $K(t, s)$ is defined for $s < t$ (taking zero value otherwise). One can write down $K(t, s)$ explicitly thanks to the definitions (2.3) and (2.4), but this expression is not included here since it will not be used later in our analysis. A crucial property for $K(t, s)$ is that

$$R(t, s) = \int_0^{t \wedge s} K(t, r) K(s, r) dr, \tag{2.7}$$

where $R(t,s)$ is the fractional Brownian motion covariance function introduced in (2.1). This essential fact enables the following analytic characterization of the Cameron-Martin space in [10, Theorem 3.1].

Theorem 2.3. *Let $\bar{\mathcal{H}}$ be the space given in Definition 2.2. As a vector space we have $\bar{\mathcal{H}} = I_{0^+}^{H+1/2}(L^2([0,1]))$, and the Cameron-Martin norm is given by*

$$\|h\|_{\bar{\mathcal{H}}} = \|K^{-1}h\|_{L^2([0,1])}. \tag{2.8}$$

In order to define Wiener integrals with respect to B, it is also convenient to look at the Cameron-Martin subspace in terms of the covariance structure. Specifically, we define another space $\mathcal{H}$ as the completion of the space of simple step functions with inner product induced by

$$\langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}} \triangleq R(s,t). \tag{2.9}$$

The space $\mathcal{H}$ is easily related to $\bar{\mathcal{H}}$. Namely define the following operator

$$\mathcal{K}^* : \mathcal{H} \to L^2([0,1]), \quad \text{such that} \quad \mathbf{1}_{[0,t]} \mapsto K(t,\cdot). \tag{2.10}$$

We also set

$$\mathcal{R} \triangleq K \circ \mathcal{K}^* : \mathcal{H} \to \bar{\mathcal{H}}, \tag{2.11}$$

where the operator K is introduced in (2.6). Then it can be proved that $\mathcal{R}$ is an isometric isomorphism (cf. Lemma 2.7 below for the surjectivity of $\mathcal{K}^*$). In addition, under this identification, $\mathcal{K}^*$ is the adjoint of K, i.e. $\mathcal{K}^* = K^* \circ \mathcal{R}$. This can be seen by acting on indicator functions and then passing limit. As mentioned above, one advantage about the space $\mathcal{H}$ is that the fractional Wiener integral operator $I : \mathcal{H} \to \mathcal{C}_1$ induced by $\mathbf{1}_{[0,t]} \mapsto B_t$ is an isometric isomorphism. According to relation (2.7), B_t admits a Wiener integral representation with respect to an underlying Wiener process W:

$$B_t = \int_0^t K(t,s) dW_s. \tag{2.12}$$

Moreover, the process W in (2.12) can be expressed as a Wiener integral with respect to B, that is $W_s = I((\mathcal{K}^*)^{-1}\mathbf{1}_{[0,s]})$ (cf. [23, relation (5.15)]).

Let us also mention the following useful formula for the natural pairing between $\mathcal{H}$ and $\bar{\mathcal{H}}$.

Lemma 2.4. *Let $\mathcal{H}$ be the space defined as the completion of the indicator functions with respect to the inner product (2.9). Also recall that $\bar{\mathcal{H}}$ is introduced in Definition 2.2. Then through the isometric isomorphism $\mathcal{R}$ defined by (2.11), the natural pairing between $\mathcal{H}$ and $\bar{\mathcal{H}}$ is given by*

$${}_{\mathcal{H}}\langle f, h \rangle_{\bar{\mathcal{H}}} = \int_0^1 f_s \, dh_s. \tag{2.13}$$

Proof. First of all, let $h \in \bar{\mathcal{H}}$ and $g \in \mathcal{H}$ be such that $\mathcal{R}(g) = h$. It is easy to see that g can be constructed in the following way. According to Definition 2.2, there exists a random

variable Z in the first chaos $\mathcal{C}_1$ such that $h_t = \mathbb{E}[B_t Z]$. The element $g \in \mathcal{H}$ is then given via the Wiener integral isomorphism between $\mathcal{H}$ and $\mathcal{C}_1$, that is, the element $g \in \mathcal{H}$ such that $Z = I(g)$. Also note that we have $h_t = \mathbb{E}[B_t\, I(g)]$.

Now consider $f \in \mathcal{H}$. The natural pairing between f and h is thus given by

$$ {}_{\mathcal{H}}\langle f, h\rangle_{\bar{\mathcal{H}}} = {}_{\mathcal{H}}\langle f, g\rangle_{\mathcal{H}} = \mathbb{E}[Z \cdot I(f)]. $$

A direct application of Fubini's theorem then yields:

$$ {}_{\mathcal{H}}\langle f, h\rangle_{\bar{\mathcal{H}}} = \mathbb{E}[Z \cdot I(f)] = \mathbb{E}\left[Z \cdot \int_0^1 f_s dB_s\right] = \int_0^1 f_s\, \mathbb{E}[Z dB_s] = \int_0^1 f_s dh_s. $$

□

The space $\mathcal{H}$ can also be described in terms of fractional calculus (cf. [24]), since the operator $\mathcal{K}^*$ defined by (2.10) can be expressed as

$$ (\mathcal{K}^* f)(t) = \begin{cases} C_H \cdot t^{\frac{1}{2}-H} \cdot \left(I_{1^-}^{H-\frac{1}{2}}\left(s^{H-\frac{1}{2}} f(s)\right)\right)(t), & H > \frac{1}{2}; \\ C_H \cdot t^{\frac{1}{2}-H} \cdot \left(D_{1^-}^{\frac{1}{2}-H}\left(s^{H-\frac{1}{2}} f(s)\right)\right)(t), & H \leq \frac{1}{2}. \end{cases} \tag{2.14} $$

Starting from this expression, it is readily checked that when $H > 1/2$ the space $\mathcal{H}$ coincides with the following subspace of the Schwartz distributions $\mathcal{S}'$:

$$ \mathcal{H} = \left\{ f \in \mathcal{S}';\, t^{1/2-H} \cdot (I_{1^-}^{H-1/2}(s^{H-1/2} f(s)))(t) \text{ is an element of } L^2([0,1]) \right\}. \tag{2.15} $$

In the case $H \leq 1/2$, we simply have

$$ \mathcal{H} = I_{1^-}^{1/2-H}(L^2([0,1])). \tag{2.16} $$

Remark 2.5. As the Hurst parameter H increases, $\mathcal{H}$ gets larger (and contains distributions when $H > 1/2$) while $\bar{\mathcal{H}}$ gets smaller. This fact is apparent from Theorem 2.3 and relations (2.15)-(2.16). When $H = 1/2$, the process B_t coincides with the usual Brownian motion. In this case, we have $\mathcal{H} = L^2([0,1])$ and $\bar{\mathcal{H}} = W_0^{1,2}$, the space of absolutely continuous paths starting at the origin with square integrable derivative.

Next we mention a variational embedding theorem for the Cameron-Martin subspace $\bar{\mathcal{H}}$ which will be used in a crucial way. The case when $H > 1/2$ is a simple exercise starting from the definition (2.2) of $\bar{\mathcal{H}}$ and invoking the Cauchy-Schwarz inequality. The case when $H \leq 1/2$ was treated in [12]. From a pathwise point of view, this allows us to integrate a fractional Brownian path against a Cameron-Martin path or vice versa (cf. [25]), and to make sense of ordinary differential equations driven by a Cameron-Martin path (cf. [18]).

Proposition 2.6. *If $H > \frac{1}{2}$, then $\bar{\mathcal{H}} \subseteq C_0^H([0,1];\mathbb{R}^d)$, the space of H-Hölder continuous paths. If $H \leq \frac{1}{2}$, then for any $q > (H+1/2)^{-1}$, we have $\bar{\mathcal{H}} \subseteq C_0^{q\text{-var}}([0,1];\mathbb{R}^d)$, the space of continuous paths with finite q-variation. In addition, the above inclusions are continuous embeddings.*

Finally, we prove two general lemmas on the Cameron-Martin subspace that are needed later on. These properties do not seem to be contained in the literature and they require some care based on fractional calculus. The first one claims the surjectivity of $\mathcal{K}^*$ on properly defined spaces.

Lemma 2.7. *Let $H \in (0,1)$, and consider the operator $\mathcal{K}^* : \mathcal{H} \to L^2([0,1])$ defined by (2.10). Then $\mathcal{K}^*$ is surjective.*

Proof. If $H > 1/2$, we know that the image of $\mathcal{K}^*$ contains all indicator functions (cf. [23, Equation (5.14)]). Therefore, $\mathcal{K}^*$ is surjective.

If $H < 1/2$, we first claim that the image of $\mathcal{K}^*$ contains functions of the form $t^{1/2-H}p(t)$ where $p(t)$ is a polynomial. Indeed, given an arbitrary $\beta \geq 0$, consider the function

$$f_\beta(t) \triangleq t^{\frac{1}{2}-H}(1-t)^{\beta+\frac{1}{2}-H}.$$

It is readily checked that $D_{1-}^{\frac{1}{2}-H} f_\beta \in L^2([0,1])$, and hence $f_\beta \in I_{1-}^{\frac{1}{2}-H}(L^2([0,1])) = \mathcal{H}$. Using the analytic expression (2.14) for $\mathcal{K}^*$, we can compute $\mathcal{K}^* f_\beta$ explicitly (cf. [14, Chapter 2, Equation (2.45)]) as

$$(\mathcal{K}^* f_\beta)(t) = C_H \frac{\Gamma\left(\beta + \frac{3}{2} - H\right)}{\Gamma(\beta+1)} t^{\frac{1}{2}-H}(1-t)^\beta.$$

Since β is arbitrary and $\mathcal{K}^*$ is linear, the claim follows.

Now it remains to show that the space of functions of the form $t^{\frac{1}{2}-H}p(t)$ with $p(t)$ being a polynomial is dense in $L^2([0,1])$. To this end, let $\varphi \in C_c^\infty((0,1))$. Then $\psi(t) \triangleq t^{-(1/2-H)}\varphi(t) \in C_c^\infty((0,1))$. According to Bernstein's approximation theorem, for any $\varepsilon > 0$, there exists a polynomial $p(t)$ such that

$$\|\psi - p\|_\infty < \varepsilon,$$

and thus

$$\sup_{0 \leq t \leq 1} |\varphi(t) - t^{\frac{1}{2}-H}p(t)| < \varepsilon.$$

Therefore, functions in $C_c^\infty((0,1))$ (and thus in $L^2([0,1])$) can be approximated by functions of the desired form. □

Our second lemma gives some continuous embedding properties for $\mathcal{H}$ and $\bar{\mathcal{H}}$ in the irregular case $H < 1/2$.

Lemma 2.8. *For $H < 1/2$, the inclusions $\mathcal{H} \subseteq L^2([0,1])$ and $W_0^{1,2} \subseteq \bar{\mathcal{H}}$ are continuous embeddings.*

Proof. For the first assertion, let $f \in \mathcal{H}$. We wish to prove that

$$\|f\|_{L^2([0,1])} \leq C_H \|f\|_{\mathcal{H}}. \tag{2.17}$$

Towards this aim, define $\varphi \triangleq \mathcal{K}^* f$, where $\mathcal{K}^*$ is defined by (2.10). Observe that $\mathcal{K}^* : \mathcal{H} \to L^2([0,1])$ and thus $f \in L^2([0,1])$. By solving f in terms of φ using the analytic expression (2.14) for $\mathcal{K}^*$, we have

$$f(t) = C_H t^{\frac{1}{2}-H} \left(I_{1-}^{\frac{1}{2}-H} \left(s^{H-\frac{1}{2}} \varphi(s) \right) \right)(t). \tag{2.18}$$

We now bound the right hand side of (2.18). Our first step in this direction is to notice that according to the definition (2.3) of fractional integral we have

$$\begin{aligned} \left| \left(I_{1-}^{\frac{1}{2}-H} (s^{H-\frac{1}{2}} \varphi(s)) \right)(t) \right| &= C_H \left| \int_t^1 (s-t)^{-\frac{1}{2}-H} s^{H-\frac{1}{2}} \varphi(s) ds \right| \\ &\leq C_H \int_t^1 (s-t)^{-\frac{1}{2}-H} s^{H-\frac{1}{2}} |\varphi(s)| ds \\ &= C_H \int_t^1 (s-t)^{-\frac{1}{4}-\frac{H}{2}} \left((s-t)^{-\frac{1}{4}-\frac{H}{2}} s^{H-\frac{1}{2}} |\varphi(s)| \right) ds. \end{aligned}$$

Hence a direct application of Cauchy-Schwarz inequality gives

$$\begin{aligned} \left| \left(I_{1-}^{\frac{1}{2}-H} (s^{H-\frac{1}{2}} \varphi(s)) \right)(t) \right| &\leq C_H \left(\int_t^1 (s-t)^{-\frac{1}{2}-H} ds \right)^{\frac{1}{2}} \left(\int_t^1 (s-t)^{-\frac{1}{2}-H} s^{2H-1} |\varphi(s)|^2 ds \right)^{\frac{1}{2}} \\ &= C_H (1-t)^{\frac{1}{2}\left(\frac{1}{2}-H\right)} \left(\int_t^1 (s-t)^{-\frac{1}{2}-H} s^{2H-1} |\varphi(s)|^2 ds \right)^{\frac{1}{2}}, \end{aligned} \tag{2.19}$$

where we recall that C_H is a positive constant which can change from line to line. Therefore, plugging (2.19) into (2.18) we obtain

$$\|f\|^2_{L^2([0,1])} \leq C_H \int_0^1 t^{1-2H} (1-t)^{\frac{1}{2}-H} \int_t^1 (s-t)^{-\frac{1}{2}-H} s^{2H-1} |\varphi(s)|^2 ds.$$

We now bound all the terms of the form s^β with $\beta > 0$ by 1. This gives

$$\begin{aligned} \|f\|^2_{L^2([0,1])} &\leq C_H \int_0^1 dt \int_t^1 (s-t)^{-\frac{1}{2}-H} |\varphi(s)|^2 ds = C_H \int_0^1 |\varphi(s)|^2 ds \int_0^s (s-t)^{-\frac{1}{2}-H} dt \\ &= C_H \int_0^1 s^{\frac{1}{2}-H} |\varphi(s)|^2 ds \leq C_H \|\varphi\|^2_{L^2([0,1])} = C_H \|f\|^2_{\mathcal{H}}, \end{aligned}$$

which is our claim (2.17).

For the second assertion about the embedding of $W_0^{1,2}$ in $\bar{\mathcal{H}}$, let $h \in W_0^{1,2}$. We thus also have $h \in \bar{\mathcal{H}}$ and we can write $h = K\varphi$ for some $\varphi \in L^2([0,1])$. We first claim that

$$\int_0^1 f(s) dh(s) = \int_0^1 \mathcal{K}^* f(s) \varphi(s) ds \tag{2.20}$$

for all $f \in \mathcal{H}$. This assertion can be reduced in the following way: since $\mathcal{H} \hookrightarrow L^2([0,1])$ continuously and $\mathcal{K}^* : \mathcal{H} \to L^2([0,1])$ is continuous, one can take limits along indicator functions

in (2.20). Thus it is sufficient to consider $f = \mathbf{1}_{[0,t]}$ in (2.20). In addition, relation (2.20) can be checked easily for $f = \mathbf{1}_{[0,t]}$. Namely we have

$$\int_0^1 \mathbf{1}_{[0,t]}(s)dh(s) = h(t) = \int_0^t K(t,s)\varphi(s)ds = \int_0^1 \left(\mathcal{K}^*\mathbf{1}_{[0,t]}\right)(s)\varphi(s)ds.$$

Therefore, our claim (2.20) holds true. Now from Lemma 2.7, if $\varphi \in L^2([0,1])$ there exists $f \in \mathcal{H}$ such that $\varphi = \mathcal{K}^* f$. For this particular f, invoking relation (2.20) we get

$$\int_0^1 f(s)dh(s) = \|\varphi\|^2_{L^2([0,1])}. \tag{2.21}$$

But we also know that

$$\|\varphi\|_{L^2([0,1])} = \|h\|_{\bar{\mathcal{H}}} = \|f\|_{\mathcal{H}}, \quad \text{and thus} \quad \|\varphi\|^2_{L^2([0,1])} = \|h\|_{\bar{\mathcal{H}}}\|f\|_{\mathcal{H}}. \tag{2.22}$$

In addition recall that the $W^{1,2}$ norm can be written as

$$\|h\|_{W^{1,2}} = \sup_{\psi \in L^2([0,1])} \frac{\left|\int_0^1 \psi(s)dh(s)\right|}{\|\psi\|_{L^2([0,1])}}$$

Owing to (2.21) and (2.22) we thus get

$$\|h\|_{W^{1,2}} \geq \frac{\int_0^1 f(s)dh(s)}{\|f\|_{L^2([0,1])}} = \frac{\|h\|_{\bar{\mathcal{H}}}\|f\|_{\mathcal{H}}}{\|f\|_{L^2([0,1])}} \geq C_H\|h\|_{\bar{\mathcal{H}}},$$

where the last step stems from (2.17). The continuous embedding $W_0^{1,2} \subseteq \bar{\mathcal{H}}$ follows. □

2.2 Free nilpotent groups.

In this section we introduce some geometrical objects which are of fundamental importance for the definition of equation (1.1) as well as our study on density estimates.

Definition 2.9. For $l \in \mathbb{N}$, the *truncated tensor algebra* $T^{(l)}$ of order l is defined by

$$T^{(l)} = \bigoplus_{n=0}^{l} (\mathbb{R}^d)^{\otimes n},$$

with the convention $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$. The set $T^{(l)}$ is equipped with a straightforward vector space structure, plus an operation $\otimes$ defined by

$$[g \otimes h]^n = \sum_{k=0}^{n} g^{n-k} \otimes h^k, \qquad g, h \in T^{(l)}, \tag{2.23}$$

where g^n designates the projection onto the n-th degree component of g for $n \leq l$.

Notice that $T^{(l)}$ should be denoted $T^{(l)}(\mathbb{R}^d)$. We have dropped the dependence on $\mathbb{R}^d$ for notational simplicity. Also observe that with Definition 2.9 in hand, $(T^{(l)}, +, \otimes)$ is an associative algebra with unit element $\mathbf{1} \in (\mathbb{R}^d)^{\otimes 0}$. The polynomial terms in the expansions which will be considered later on are contained in a subspace of $T^{(l)}$ that we proceed to define now.

Definition 2.10. The *free nilpotent Lie algebra* $\mathfrak{g}^{(l)}$ of order l is defined to be the graded sum

$$\mathfrak{g}^{(l)} \triangleq \bigoplus_{k=1}^{l} \mathcal{L}_k \subseteq T^{(l)}.$$

Here $\mathcal{L}_k$ is the space of homogeneous Lie polynomials of degree k defined inductively by $\mathcal{L}_1 \triangleq \mathbb{R}^d$ and $\mathcal{L}_k \triangleq [\mathbb{R}^d, \mathcal{L}_{k-1}]$, where the Lie bracket is defined to be the commutator of the tensor product.

We now define some groups related to the algebras given in Definitions 2.9 and 2.10. To this aim, introduce the subspace $T_0^{(l)} \subseteq T^{(l)}$ of tensors whose scalar component is zero and recall that $\mathbf{1} \triangleq (1, 0, \ldots, 0)$. For $u \in T_0^{(l)}$, one can define the inverse $(\mathbf{1}+u)^{-1}$, the exponential $\exp(u)$ and the logarithm $\log(\mathbf{1}+u)$ in $T^{(l)}$ by using the standard Taylor expansion formula with respect to the tensor product. For instance,

$$\exp(a) \triangleq \sum_{k=0}^{\infty} \frac{1}{k!} a^{\otimes k} \in T^{(l)}, \tag{2.24}$$

where the sum is indeed locally finite and hence well-defined. We can now introduce the following group.

Definition 2.11. The *free nilpotent Lie group* $G^{(l)}$ of order l is defined by

$$G^{(l)} \triangleq \exp(\mathfrak{g}^{(l)}) \subseteq T^{(l)}.$$

The exponential function is a diffeomorphism under which $\mathfrak{g}^{(l)}$ in Definition 2.10 is the Lie algebra of $G^{(l)}$.

Remark 2.12. One can also include the case of $l = \infty$ in the definitions of $T^{(l)}$, $\mathfrak{g}^{(l)}$ and $G^{(l)}$. But in this case we need to be careful that the direct sums should all be understood as formal series instead of polynomials. Also all the spaces we have mentioned are defined over the given underlying vector space (which is $\mathbb{R}^d$ in our case), and recall that we have omitted such dependence in the notation for simplicity.

It will be useful in the sequel to have some basis available for the algebras introduced above. We shall resort to the following families: for each word $\alpha = (i_1, \ldots, i_r) \in \mathcal{A}_1(l)$, set

$$\mathrm{e}_{(\alpha)} \triangleq \mathrm{e}_{i_1} \otimes \cdots \otimes \mathrm{e}_{i_r}, \quad \text{and} \quad \mathrm{e}_{[\alpha]} \triangleq [\mathrm{e}_{i_1}, \cdots [\mathrm{e}_{i_{r-2}}, [\mathrm{e}_{i_{r-1}}, \mathrm{e}_{i_r}]]], \tag{2.25}$$

where $\{\mathrm{e}_1, \ldots, \mathrm{e}_d\}$ denotes the canonical basis of $\mathbb{R}^d$. Then it can be shown that $\{\mathrm{e}_{(\alpha)} : \alpha \in \mathcal{A}(l)\}$ is the canonical basis of $T^{(l)}$, and we also have $\mathfrak{g}^{(l)} = \mathrm{Span}\{\mathrm{e}_{[\alpha]} : \alpha \in \mathcal{A}_1(l)\}$.

As a closed subspace, $\mathfrak{g}^{(l)}$ induces a canonical Hilbert structure from $T^{(l)}$ which makes it into a flat Riemannian manifold. The associated volume measure du (the Lebesgue measure) on $\mathfrak{g}^{(l)}$ is left invariant with respect to the product induced from the group structure on $G^{(l)}$ through the exponential diffeomorphism. In addition, for each $\lambda > 0$, there is a dilation operation $\delta_\lambda : T^{(l)} \to T^{(l)}$ induced by $\delta_\lambda(a) \triangleq \lambda^k a$ if $a \in (\mathbb{R}^d)^{\otimes k}$, which satisfies the relation $\delta_\lambda \circ \exp = \exp \circ \delta_\lambda$ when restricted on $\mathfrak{g}^{(l)}$. Thanks to the fact that $\delta_\lambda(a) = \lambda^k a$ for any $a \in (\mathbb{R}^d)^{\otimes k}$, one can easily show that

$$du \circ \delta_\lambda^{-1} = \lambda^{-\nu} du, \quad \text{where} \quad \nu \triangleq \sum_{k=1}^{l} k \dim(\mathcal{L}_k). \tag{2.26}$$

We always fix the Euclidean norm on $\mathbb{R}^d$ in the remainder of the paper. As far as the free nilpotent group $G^{(l)}$ is concerned, there are several useful metric structures. Among them we will use an extrinsic metric ρ_{HS} which can be defined easily due to the fact that $G^{(l)}$ is a subspace of $T^{(l)}$. Namely for $g_1, g_2 \in G^{(l)}$ we set:

$$\rho_{\mathrm{HS}}(g_1, g_2) \triangleq \|g_2 - g_1\|_{\mathrm{HS}}, \quad g_1, g_2 \in G^{(l)}, \tag{2.27}$$

where the right hand side is induced from the Hilbert-Schmidt norm on $T^{(l)}$.

2.3 Path signatures and the fractional Brownian rough path.

The stochastic differential equation (1.1) governed by a fractional Brownian motion B is standardly solved in the rough paths sense. In this section we recall some basic facts about this notion of solution. We will also give some elements of rough paths expansions, which are at the heart of our methodology in order to obtain lower bounds for the density.

The link between free nilpotent groups and noisy equations like (1.1) is made through the notion of signature. Recall that a continuous map $\mathbf{x} : \{(s,t) \in [0,1]^2 : s \leq t\} \to T^{(l)}$ is called a *multiplicative functional* if for $s < u < t$ one has $\mathbf{x}_{s,t} = \mathbf{x}_{s,u} \otimes \mathbf{x}_{u,t}$. A particular occurrence of this kind of map is given when one considers a path w with finite variation and sets for $s \leq t$,

$$\mathbf{w}^n_{s,t} = \int_{s<u_1<\cdots<u_n<t} dw_{u_1} \otimes \cdots \otimes dw_{u_n}. \tag{2.28}$$

Then the so-called *truncated signature path* of order l associated with w is defined by the following object:

$$S_l(w)_{\cdot,\cdot} : \{(s,t) \in [0,1]^2;\, s \leq t\} \to T^{(l)}, \qquad (s,t) \mapsto S_l(w)_{s,t} := 1 + \sum_{n=1}^{l} \mathbf{w}^n_{s,t}. \tag{2.29}$$

It can be shown that the functional $S_l(w)_{\cdot,\cdot}$ is multiplicative and takes values in the free nilpotent group $G^{(l)}$. The *truncated signature* of order l for w is the tensor element $S_l(w)_{0,1} \in G^{(l)}$. It is simply denoted as $S_l(w)$.

A rough path can be seen as a generalization of the truncated signature path (2.29) to the non-smooth situation. Specifically, the definition of Hölder rough paths can be summarized as follows.

Definition 2.13. Let $\gamma \in (0,1)$. The space of *weakly geometric γ-Hölder rough paths* is the set of multiplicative paths $\mathbf{x} : \{(s,t) \in [0,1]^2;\ s \le t\} \to G^{[1/\gamma]}$ such that the following norm is finite:

$$\|\mathbf{x}\|_{\gamma;\mathrm{HS}} = \sup_{0 \le s < t \le 1} \frac{\|\mathbf{x}_{s,t}\|_{\mathrm{HS}}}{|t-s|^{\gamma}}. \tag{2.30}$$

An important subclass of weakly geometric γ-Hölder rough paths is the set of *geometric* γ-Hölder rough paths. These are multiplicative paths $\mathbf{x}$ with values in $G^{\lfloor 1/\gamma \rfloor}$ such that $\|\mathbf{x}\|_{\gamma;\mathrm{HS}}$ is finite and such that there exists a sequence $\{x_\varepsilon;\ \varepsilon > 0\}$ with $x_\varepsilon \in C^\infty([0,T];\mathbb{R}^d)$ satisfying

$$\lim_{\varepsilon \to 0} \|\mathbf{x} - S_{[1/\gamma]}(x_\varepsilon)\|_{\gamma;\mathrm{HS}} = 0. \tag{2.31}$$

The notion of signature allows to define a more intrinsic distance (with respect to the HS-distance given by (2.27)) on the free group $G^{(l)}$. This metric is known as the *Carnot-Caratheodory metric* and given by

$$\rho_{\mathrm{CC}}(g_1, g_2) \triangleq \|g_1^{-1} \otimes g_2\|_{\mathrm{CC}}, \quad g_1, g_2 \in G^{(l)},$$

where the CC-norm $\|\cdot\|_{\mathrm{CC}}$ is defined by

$$\|g\|_{\mathrm{CC}} \triangleq \inf\left\{\|w\|_{1-\mathrm{var}} : w \in C^{1-\mathrm{var}}([0,1];\mathbb{R}^d) \text{ and } S_l(w) = g\right\}. \tag{2.32}$$

It can be shown that the infimum in (2.32) is attainable.

Remark 2.14. It is well-known that for any $g \in G^{(l)}$, one can find a piecewise linear path w such that $S_l(w) = g$ (cf. [13] for instance). Moreover, one can do better, and find a smooth path w whose derivative is compactly supported such that $S_l(w) = g$. Indeed, this could be achieved by (i) reparametrizing the piecewise linear path so that the resulting path is smooth but the trajectory itself is still the same piecewise linear one; and (ii) adding trivial pieces to the beginning and the end of the path. This will not change the truncated signature as it is invariant under reparametrization.

The HS and CC metrics are equivalent as seen from the following so-called *ball-box estimate*.

Proposition 2.15. *Let ρ_{HS} and ρ_{CC} be the distances on $G^{(l)}$ respectively defined by* (2.27) *and* (2.32)*. For each $l \ge 1$, there exists a constant $C = C_l > 0$, such that*

$$\rho_{\mathrm{CC}}(g_1, g_2) \le C \max\left\{\rho_{\mathrm{HS}}(g_1, g_2), \rho_{\mathrm{HS}}(g_1, g_2)^{\frac{1}{l}} \cdot \max\left\{1, \|g_1\|_{\mathrm{CC}}^{1-\frac{1}{l}}\right\}\right\} \tag{2.33}$$

and

$$\rho_{\mathrm{HS}}(g_1, g_2) \le C \max\left\{\rho_{\mathrm{CC}}(g_1, g_2)^l, \rho_{\mathrm{CC}}(g_1, g_2) \cdot \max\left\{1, \|g_1\|_{\mathrm{CC}}^{l-1}\right\}\right\}$$

for all $g_1, g_2 \in G^{(l)}$. In particular,

$$\|g\|_{\mathrm{CC}} \le 1 \implies \|g - \mathbf{1}\|_{\mathrm{HS}} \le C\|g\|_{\mathrm{CC}}$$

and

$$\|g - \mathbf{1}\|_{\mathrm{HS}} \le 1 \implies \|g\|_{\mathrm{CC}} \le C\|g - \mathbf{1}\|_{\mathrm{HS}}^{\frac{1}{l}}.$$

One of the main application of the abstract rough path theory is the ability to extend most stochastic calculus tools to a large class of Gaussian processes. The following result, borrowed from [8, 12], establishes this link for fractional Brownian motion.

Proposition 2.16. *Let B be a fractional Brownian motion with Hurst parameter $H > 1/4$. Then B admits a lift $\mathbf{B}$ as a geometric rough path of order $[1/\gamma]$ for any $\gamma < H$.*

Let us now turn to the definition of rough differential equations. There are several equivalent ways to introduce this notion, among which we will choose to work with Taylor type expansions, since they are more consistent with our later developments. To this aim, let us first consider a bounded variation path w and the following ordinary differential equation driven by w:

$$dx_t = \sum_{\alpha=1}^{d} V_\alpha(x_t)\, dw_t^\alpha, \tag{2.34}$$

where the V_α's are C_b^∞ vector fields. For any given word $\alpha = (i_1, \dots, i_r)$ over the letters $\{1, \dots, d\}$, we define the vector field $V_{(\alpha)} \triangleq (V_{i_1} \cdots (V_{i_{r-2}}(V_{i_{r-1}} V_{i_r})))$, where we have identified a vector field with a differential operator, so that $V_i V_j$ means differentiating V_j along direction V_i. Classically, a *formal Taylor expansion* of the solution x_t to (2.34) is then given by

$$x_{s,t} \sim \sum_{k=1}^{\infty} \sum_{i_1,\dots,i_k=1}^{d} V_{(i_1,\dots,i_k)}(x_s) \int_{s<u_1<\cdots<u_k<t} dw_{u_1}^{i_1} \cdots dw_{u_k}^{i_k}, \tag{2.35}$$

where we have set $x_{s,t} = x_t - x_s$. This expansion can be rephrased in more geometrical terms. Specifically, we define the following Taylor approximation function on $\mathfrak{g}^{(l)}$.

Definition 2.17. Let $\{V_\alpha;\, 1 \le \alpha \le d\}$ be a family of C_b^∞ vector fields on $\mathbb{R}^N$, and recall that the sets of words $\mathcal{A}(l), \mathcal{A}_1(l)$ are introduced at the beginning of Section 1.2. For each $l \ge 1$, we define the *Taylor approximation function* $F_l : \mathfrak{g}^{(l)} \times \mathbb{R}^N \to \mathbb{R}^N$ of order l associated with the ODE (2.34) by

$$F_l(u, x) \triangleq \sum_{\alpha \in \mathcal{A}_1(l)} V_{(\alpha)}(x) \cdot (\exp u)^\alpha, \quad (u, x) \in \mathfrak{g}^{(l)} \times \mathbb{R}^N,$$

where the exponential function is defined on $T^{(l)}$ by (2.24) and $(\exp(u))^\alpha$ is the coefficient of $\exp(u)$ with respect to the tensor basis element $\mathrm{e}_{(\alpha)}$ (we recall that the notation $\mathrm{e}_{(\alpha)}$ is introduced in (2.25)). We also say that $u \in \mathfrak{g}^{(l)}$ *joins x to y in the sense of Taylor approximation* if $y = x + F_l(u, x)$.

With Definition 2.17 in hand, we can recast the formal expansion (2.35) (truncated at an arbitrary degree l) in the following way:

$$x_{s,t} \sim F_l\left(\log\left(S(w)_{s,t}\right), x_s\right), \tag{2.36}$$

where the function log is the inverse of the exponential map for $G^{(l)}$ which can also be defined by a truncated Taylor's formula on $G^{(l)}$ similar to the exponential, and $S(w)_{s,t}$ is the truncated signature path of w defined by (2.29). In order to define rough differential equations, a natural idea is to extend this approximation scheme to rough paths. We get a definition which is stated below in the fractional Brownian motion case.

Definition 2.18. Let B be a fractional Brownian motion with Hurst parameter $H > 1/4$, and consider its rough path lift $\mathbf{B}$ as in Proposition 2.16. Let $\{V_\alpha;\, 1 \leq \alpha \leq d\}$ be a family of C_b^∞ vector fields on $\mathbb{R}^N$. We say that X is a *solution to the rough differential equation* (1.1) if for all $(s,t) \in [0,1]^2$ such that $s < t$ we have

$$X_{s,t} = F_{[1/\gamma]-1}\left(\log\left(S(\mathbf{B})_{s,t}\right), X_s\right) + R_{s,t}\,, \tag{2.37}$$

where $R_{s,t}$ is an $\mathbb{R}^N$-valued remainder such that there exists $\varepsilon > 0$ satisfying

$$\sup_{0 \leq s < t \leq 1} \frac{|R_{s,t}|}{|t-s|^{1+\varepsilon}} < \infty.$$

Roughly speaking, Definition 2.18 says that the expansion of the solution X to a rough differential equation should coincide with (2.35) up to a remainder with Hölder regularity greater than 1. This approach goes back to Davie [9], and it can be shown to coincide with more classical notions of solutions. We close this section by recalling an existence and uniqueness result which is fundamental in rough path theory.

Proposition 2.19. *Under the same conditions as in Definition 2.18, there exists a unique solution to equation* (1.1) *considered in the sense of (2.37).*

2.4 Malliavin calculus for fractional Brownian motion.

In this section we review some basic aspects of Malliavin calculus. The reader is referred to [23] for further details.

We consider the fractional Brownian motion $B = (B^1, \dots, B^d)$ as in Definition (2.1), defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For sake of simplicity, we assume that $\mathcal{F}$ is generated by $\{B_t;\, t \in [0,T]\}$. An $\mathcal{F}$-measurable real valued random variable F is said to be *cylindrical* if it can be written, with some $m \geq 1$, as

$$F = f\left(B_{t_1}, \dots, B_{t_m}\right), \quad \text{for} \quad 0 \leq t_1 < \dots < t_m \leq 1,$$

where $f : \mathbb{R}^m \to \mathbb{R}$ is a C_b^∞ function. The set of cylindrical random variables is denoted by $\mathcal{S}$.

The Malliavin derivative is defined as follows: for $F \in \mathcal{S}$, the derivative of F in the direction $h \in \mathcal{H}$ is given by

$$\mathbf{D}_h F = \sum_{i=1}^m \frac{\partial f}{\partial x_i}\left(B_{t_1}, \dots, B_{t_m}\right) h_{t_i}.$$

More generally, we can introduce iterated derivatives. Namely, if $F \in \mathcal{S}$, we set

$$\mathbf{D}^k_{h_1,\dots,h_k} F = \mathbf{D}_{h_1} \dots \mathbf{D}_{h_k} F.$$

For any $p \geq 1$, it can be checked that the operator $\mathbf{D}^k$ is closable from $\mathcal{S}$ into $\mathbf{L}^p(\Omega; \mathcal{H}^{\otimes k})$. We denote by $\mathbb{D}^{k,p}(\mathcal{H})$ the closure of the class of cylindrical random variables with respect to the norm

$$\|F\|_{k,p} = \left(\mathbb{E}\left[|F|^p\right] + \sum_{j=1}^{k} \mathbb{E}\left[\left\|\mathbf{D}^j F\right\|^p_{\mathcal{H}^{\otimes j}}\right] \right)^{\frac{1}{p}}, \tag{2.38}$$

and we also set $\mathbb{D}^\infty(\mathcal{H}) = \cap_{p\geq 1} \cap_{k\geq 1} \mathbb{D}^{k,p}(\mathcal{H})$.

Estimates of Malliavin derivatives are crucial in order to get information about densities of random variables, and Malliavin covariance matrices as well as non-degenerate random variables will feature importantly in the sequel.

Definition 2.20. Let $F = (F^1, \ldots, F^n)$ be a random vector whose components are in $\mathbb{D}^\infty(\mathcal{H})$. Define the *Malliavin covariance matrix* of F by

$$\gamma_F = (\langle \mathbf{D}F^i, \mathbf{D}F^j \rangle_{\mathcal{H}})_{1\leq i,j\leq n}. \tag{2.39}$$

Then F is called *non-degenerate* if γ_F is invertible *a.s.* and

$$(\det \gamma_F)^{-1} \in \cap_{p\geq 1} L^p(\Omega).$$

It is a classical result that the law of a non-degenerate random vector $F = (F^1, \ldots, F^n)$ admits a smooth density with respect to the Lebesgue measure on $\mathbb{R}^n$.

3 The elliptic case.

In this section, for a better understanding of our overall strategy, we first prove Theorem 1.3 and Theorem 1.7 in the uniformly elliptic case. The analysis in this case is more explicit and straightforward, and our methodology might be more apparent. More precisely we still consider the SDE (1.1), and we assume that the coefficients satisfy the following hypothesis:

Uniform Ellipticity Assumption. The $\mathcal{C}_b^\infty$ vector fields $V = \{V_1, \ldots, V_d\}$ are such that

$$\Lambda_1 |\xi|^2 \leq \xi^* V(x) V(x)^* \xi \leq \Lambda_2 |\xi|^2, \quad \forall x, \xi \in \mathbb{R}^N, \tag{3.1}$$

with some constants $\Lambda_1, \Lambda_2 > 0$, where $(\cdot)^*$ denotes matrix transpose.

Notice that the condition (3.1) can be seen as a special case of the uniform hypoellipticity condition (1.2), where $l_0 = 1$.

One of the major simplifications of the elliptic (vs. hypoelliptic) situation concerns the control distance d. Indeed, recall that for $x, y \in \mathbb{R}^N$, $\Pi_{x,y}$ is the set of Cameron-Martin paths that join x to y in the sense of differential equation (cf. (1.5)). Under our assumption (3.1) it is easy to construct an $h \in \bar{\mathcal{H}} \in \Pi_{x,y}$ explicitly, which will ease our computations later on.

Lemma 3.1. *Let $V = \{V_1, \dots, V_d\}$ be vector fields satisfying the uniform elliptic assumption (3.1). Given $x, y \in \mathbb{R}^N$, define*

$$h_t \triangleq \int_0^t V^*(z_s) \cdot (V(z_s)V^*(z_s))^{-1} \cdot (y - x) ds, \tag{3.2}$$

where $z_t \triangleq (1-t)x + ty$ is the line segment from x to y. Then $h \in \Pi_{x,y}$, where $\Pi_{x,y}$ is defined by relation (1.5).

Proof. Since $\bar{\mathcal{H}} = I_{0^+}^{H+1/2}(L^2([0,1]))$ contains smooth paths, it is obvious that $h \in \bar{\mathcal{H}}$. As far as z_t is concerned, the definition $z_t = (1-t)x + ty$ clearly implies that $z_0 = x, z_1 = y$ and $\dot{z}_t = y - x$. In addition, since $VV^*(\xi)$ is invertible for all $\xi \in \mathbb{R}^N$ under our condition (3.1), we get

$$\dot{z}_t = y - x = \left(VV^*(VV^*)^{-1}\right)(z_t) \cdot (y - x) = V(z_t)\dot{h}_t,$$

where the last identity stems from the definition (3.2) of h. Therefore $h \in \Pi_{x,y}$ according to our definition (1.5). □

Remark 3.2. The intuition behind Lemma 3.1 is very simple. Indeed, given any smooth path x_t with $x_0 = x, x_1 = y$, since the vector fields are elliptic, there exist smooth functions $\lambda^1(t), \dots, \lambda^d(t)$, such that

$$\dot{x}_t = \sum_{\alpha=1}^{d} \lambda^\alpha(t) V_\alpha(x_t), \quad 0 \le t \le 1.$$

In matrix notation, $\dot{x}_t = V(x_t) \cdot \lambda(t)$. A canonical way to construct $\lambda(t)$ is writing it as $\lambda(t) = V^*(x_t)\eta(t)$ so that from ellipticity we can solve for $\eta(t)$ as

$$\eta(t) = (V(x_t)V^*(x_t))^{-1}\dot{x}_t.$$

It follows that the path $h_t \triangleq \int_0^t \lambda(s)ds$ belongs to $\Pi_{x,y}$.

Now we can prove the following result which asserts that the control distance function is locally comparable with the Euclidean metric, that is Theorem 1.7 under elliptic assumptions.

Theorem 3.3. *Let $V = \{V_1, \dots, V_d\}$ be vector fields satisfying the uniform elliptic assumption (3.1). Consider the control distance $d = d_H$ given in Definition 1.2 for a given $H > \frac{1}{4}$. Then there exist constants $C_1, C_2 > 0$ depending only on H and the vector fields, such that*

$$C_1|x - y| \le d(x, y) \le C_2|x - y| \tag{3.3}$$

for all $x, y \in \mathbb{R}^N$ with $|x - y| \le 1$.

Proof. We first consider the case when $H \le 1/2$, which is simpler due to Lemma 2.8. Given $x, y \in \mathbb{R}^N$, define $h \in \Pi_{x,y}$ as in Lemma 3.1. According to Lemma 2.8 and Definition 1.2 we have

$$d(x, y)^2 \le \|h\|_{\bar{\mathcal{H}}}^2 \le C_H \|h\|_{W^{1,2}}^2.$$

Therefore, according to the definition (3.2) of h, we get

$$d(x,y)^2 \leq C_H \int_0^1 |V^*(z_s)(V(z_s)V^*(z_s))^{-1} \cdot (y-x)|^2 ds \leq C_{H,V}|y-x|^2,$$

where the last inequality stems from the uniform ellipticity assumption (3.1) and the fact that V^* is bounded. This proves the upper bound in (3.3).

We now turn to the lower bound in (3.3). To this aim, consider $h \in \Pi_{x,y}$. We assume (without loss of generality) in the sequel that

$$\|h\|_{\bar{\mathcal{H}}} \leq 2d(x,y) \leq 2C_2, \tag{3.4}$$

where the last inequality is due to the second part of inequality (3.3) and the fact that $|x-y| \leq 1$. Then recalling the definition (1.5) of $\Pi_{x,y}$ we have

$$y - x = \int_0^1 V(\Phi_t(x;h))dh_t.$$

According to Proposition 2.6 (specifically the embedding $\bar{\mathcal{H}} \subseteq C_0^{q-\mathrm{var}}([0,1];\mathbb{R}^d)$ for $q > (H+1/2)^{-1}$) and the pathwise variational estimate given by [13, Theorem 10.14], we have

$$|y-x| \leq C_{H,V}\left(\|h\|_{q-\mathrm{var}} \vee \|h\|_{q-\mathrm{var}}^q\right) \leq C_{H,V}\left(\|h\|_{\bar{\mathcal{H}}} \vee \|h\|_{\bar{\mathcal{H}}}^q\right). \tag{3.5}$$

Since $q \geq 1$ and owing to (3.4), we conclude that

$$|y-x| \leq C_{H,V}\|h\|_{\bar{\mathcal{H}}}$$

for all x, y with $|y-x| \leq 1$. Since $h \in \Pi_{x,y}$ is arbitrary provided (3.4) holds true, the lower bound in (3.3) follows again by a direct application of Definition 1.2.

Next we consider the case when $H > 1/2$. The lower bound in (3.3) can be proved with the same argument as in the case $H \leq 1/2$, the only difference being that in (3.5) we replace $\bar{\mathcal{H}} \subseteq C_0^{q-\mathrm{var}}([0,1];\mathbb{R}^d)$ by $\bar{\mathcal{H}} \subseteq C_0^H([0,1];\mathbb{R}^d)$ and the pathwise variational estimate of [13, Theorem 10.14] by a Hölder estimate borrowed from [11, Proposition 8.1].

For the upper bound in (3.3), we again take $h \in \Pi_{x,y}$ as given by Lemma 3.1 and estimate its Cameron-Martin norm. Note that due to our uniform ellipticity assumption (3.1), one can define the function

$$\gamma_t \equiv \int_0^t (V^*(VV^*)^{-1})(z_s)ds = \int_0^t g((1-s)x + sy)ds, \tag{3.6}$$

where g is a matrix-valued C_b^∞ function. We will now prove that γ can be written as $\gamma = K\varphi$ for $\varphi \in L^2([0,1])$. Indeed, one can solve for φ in the analytic expression (2.6) for $H > 1/2$ and get

$$\varphi(t) = C_H t^{H-\frac{1}{2}} \left(D_{0+}^{H-\frac{1}{2}}\left(s^{\frac{1}{2}-H}\dot{\gamma}_s\right)\right)(t).$$

We now use the expression (2.4) for $D_{0^+}^{H-1/2}$, which yield (after an elementary change of variable)

$$\begin{aligned}\varphi(t) &= C_H t^{H-\frac{1}{2}} \frac{d}{dt} \int_0^t s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H} g((1-s)x+sy)ds \\ &= C_H t^{H-\frac{1}{2}} \frac{d}{dt} \left(t^{2-2H} \int_0^1 (u(1-u))^{\frac{1}{2}-H} g((1-tu)x+tuy)du \right) \\ &= C_H t^{\frac{1}{2}-H} \int_0^1 (u(1-u))^{\frac{1}{2}-H} g((1-tu)x+tuy)du \\ &\quad + C_H t^{\frac{3}{2}-H} \int_0^1 (u(1-u))^{\frac{1}{2}-H} u \nabla g((1-tu)x+tuy)\cdot(y-x)du.\end{aligned}$$

Hence, thanks to the fact that g and ∇g are bounded plus the fact that $t \leq 1$, we get

$$|\varphi(t)| \leq C_{H,V}(t^{\frac{1}{2}-H} + |y-x|),$$

from which φ is clearly an element of $L^2([0,1])$. Since $|y-x| \leq 1$, we conclude that

$$\|\gamma\|_{\bar{\mathcal{H}}} = \|\varphi\|_{L^2([0,1])} \leq C_{H,V}.$$

Therefore, recalling that h is given by (3.2) and γ is defined by (3.6), we end up with

$$\begin{aligned}d(x,y) \leq \|h\|_{\bar{\mathcal{H}}} &= \left\| \left(\int_0^{\cdot} (V^*(VV^*)^{-1})(z_s)ds \right) \cdot (y-x) \right\|_{\bar{\mathcal{H}}} \\ &= \|\gamma\|_{\bar{\mathcal{H}}} |y-x| \leq C_{H,V}|y-x|.\end{aligned}$$

This concludes the proof. □

From Theorem 3.3, we know that $|B_d(x,t^H)| \asymp t^{NH}$ when t is small. Therefore, the elliptic version of Theorem 1.7 becomes the following result, which is consistent with the intuition that the density $p(t,x,y)$ of the solution to equation (1.1) should behave like the Gaussian kernel

$$p(t,x,y) \asymp \frac{C_1}{t^{NH}} \exp\left(-\frac{C_2|y-x|^2}{t^{2H}} \right).$$

Theorem 3.4. *Let $p(t,x,y)$ be the density of the solution X_t to equation (1.1). Under the uniform ellipticity assumption (3.1), there exist constants $C_1, C_2, \tau > 0$ depending only on H and the vector fields V, such that*

$$p(t,x,y) \geq \frac{C_1}{t^{NH}} \tag{3.7}$$

for all $(t,x,y) \in (0,1] \times \mathbb{R}^N \times \mathbb{R}^N$ satisfying $|x-y| \leq C_2 t^H$ and $t < \tau$.

The main idea behind the proof of Theorem 3.4 is to translate the small time estimate in (3.7) into a large deviation estimate. To this aim, we will first recall some preliminary

notions taken from [4]. By a slight abuse of notation with respect to (1.4), we will call $w \mapsto \Phi_t(x;w)$ the solution map of the SDE (1.1) (or (2.37)). From the scaling invariance of fractional Brownian motion, it is not hard to see that

$$\Phi_t(x;B) \overset{\text{law}}{=} \Phi_1(x;\varepsilon B), \tag{3.8}$$

where $\varepsilon \triangleq t^H$. Therefore, since the random variable $\Phi_t(x;B)$ is nondegenerate under our standing assumption (3.1), the density $p(t,x,y)$ can be written as

$$p(t,x,y) = \mathbb{E}\left[\delta_y\left(\Phi_1(x;\varepsilon B)\right)\right]. \tag{3.9}$$

Starting from expression (3.9), we now label a proposition which gives a lower bound on $p(t,x,y)$ in terms of some conveniently chosen shifts on the Wiener space.

Proposition 3.5. *In this proposition, Φ_t stands for the solution map of equation (1.1). The vector fields $\{V_1,\dots,V_d\}$ are supposed to satisfy the uniform elliptic assumption (3.1). Then the following holds true.*
(i) Let Φ_t be the solution map of equation (1.1), $h \in \bar{\mathcal{H}}$, and let

$$X^\varepsilon(h) \triangleq \frac{\Phi_1(x;\varepsilon B + h) - \Phi_1(x;h)}{\varepsilon}. \tag{3.10}$$

Then $X^\varepsilon(h)$ converges in $\mathbb{D}^\infty$ to $X(h)$, where $X(h)$ is a $\mathbb{R}^N$-valued centered Gaussian random variable whose covariance matrix will be specified below.
(ii) Let $\varepsilon > 0$ and consider $x, y \in \mathbb{R}^N$ such that $d(x,y) \le \varepsilon$, where $d(\cdot,\cdot)$ is the distance considered in Theorem 3.3. Choose $h \in \Pi_{x,y}$ so that

$$\|h\|_{\bar{\mathcal{H}}} \le d(x,y) + \varepsilon \le 2\varepsilon. \tag{3.11}$$

Then we have

$$\mathbb{E}\left[\delta_y\left(\Phi_1(x;\varepsilon B)\right)\right] \ge C\varepsilon^{-N} \cdot \mathbb{E}\left[\delta_0\left(X^\varepsilon(h)\right) \mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right]. \tag{3.12}$$

Proof. The first statement is proved in [4]. For the second statement, according to the Cameron-Martin theorem, we have

$$\mathbb{E}\left[\delta_y\left(\Phi_1(x;\varepsilon B)\right)\right] = \mathrm{e}^{-\frac{\|h\|_{\bar{\mathcal{H}}}^2}{2\varepsilon^2}} \mathbb{E}\left[\delta_y\left(\Phi_1(x;\varepsilon B + h)\right) \mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right],$$

where we have identified $\bar{\mathcal{H}}$ with $\mathcal{H}$ through $\mathcal{R}$ and recall that $I : \mathcal{H} \to \mathcal{C}_1$ is the Wiener integral operator introduced in Section 2.1. Therefore, thanks to inequality (3.11), we get

$$\mathbb{E}\left[\delta_y\left(\Phi_1(x;\varepsilon B)\right)\right] \ge C \cdot \mathbb{E}\left[\delta_y\left(\Phi_1(x;\varepsilon B + h)\right) \mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right].$$

In addition we have chosen $h \in \Pi_{x,y}$, which means that $\Phi_1(x;h) = y$. Thanks to the scaling property of the Dirac delta function in $\mathbb{R}^N$, we get

$$p(t,x,y) = \mathbb{E}\left[\delta_y\left(\Phi_1(x;\varepsilon B)\right)\right] \ge C\varepsilon^{-N} \cdot \mathbb{E}\left[\delta_0\left(\frac{\Phi_1(x;\varepsilon B + h) - \Phi_1(x;h)}{\varepsilon}\right) \mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right].$$

Our claim (3.12) thus follows from the definition (3.10) of $X^\varepsilon(h)$. □

Let us now describe the covariance matrix of $X(h)$ introduced in Proposition 3.5. For this, we recall again that Φ is the application defined on $\bar{\mathcal{H}}$ by (1.4). The Jacobian of $\Phi_t(\cdot\,;h):\mathbb{R}^N\to\mathbb{R}^N$ is denoted by $J(\cdot\,;h)$. Then it is standard (cf. [4]) that the deterministic Malliavin differential of Φ satisfies

$$\langle D\Phi_t(x;h),l\rangle_{\bar{\mathcal{H}}}=J_t(x;h)\cdot\int_0^t J_s^{-1}(x;h)\cdot V(\Phi_s(x;h))dl_s,\quad \text{for all } l\in\bar{\mathcal{H}}, \tag{3.13}$$

where D is the Malliavin derivative operator. According to the pairing (2.13), when viewed as an $\mathcal{H}$-valued functional, we have

$$(D\Phi_t^i(x;h))_s=\left(J_t(x;h)J_s^{-1}(x;h)V(\Phi_s(x;h))\right)^i\mathbf{1}_{[0,t]}(s),\quad 1\le i\le N. \tag{3.14}$$

Then the $N\times N$ covariance matrix of $X(h)$ admits the following representation taken from the reference [4]:

$$\mathrm{Cov}(X(h))\equiv\Gamma_{\Phi_1(x;h)}=\langle D\Phi_1(x;h),D\Phi_1(x;h)\rangle_{\mathcal{H}}. \tag{3.15}$$

With (3.15) in hand, a crucial point for proving Theorem 3.4 is the fact that $\Gamma_{\Phi_1(x;h)}$ is uniformly non-degenerate with respect to all h. This is the content of the following result which is another special feature of ellipticity that fails in the hypoelliptic case. Its proof is an adaptation of the argument in [4] to the deterministic context.

Lemma 3.6. *Let $M>0$ be a localizing constant. Consider the Malliavin covariance matrix $\Gamma_{\Phi_1(x;h)}$ defined by (3.15). Under the uniform ellipticity assumption (3.1), there exist $C_1,C_2>0$ depending only on H,M and the vector fields, such that*

$$C_1\le\det\Gamma_{\Phi_1(x;h)}\le C_2 \tag{3.16}$$

for all $x\in\mathbb{R}^N$ and $h\in\bar{\mathcal{H}}$ with $\|h\|_{\bar{\mathcal{H}}}\le M$.

Proof. We consider the cases of $H>1/2$ and $H\le 1/2$ separately. We only study the lower bound of $\Gamma_{\Phi_1(x;h)}$ since the upper bound is standard from pathwise estimates by (3.14) and (3.15), plus the fact that $\|h\|_{\mathcal{H}}\le M$.

(i) Proof of the lower bound when $H>1/2$. According to relation (3.15) and the expression for the inner product in $\mathcal{H}$ given by [23, equation (5.6)], we have

$$\Gamma_{\Phi_1(x;h)}=C_H\sum_{\alpha=1}^d\int_{[0,1]^2}J_1J_s^{-1}V_\alpha(\Phi_s)V_\alpha^*(\Phi_t)(J_t^{-1})^*J_1^*|t-s|^{2H-2}dsdt,$$

where we have omitted the dependence on x and h for Φ and J inside the integral for notational simplicity. It follows that for any $z\in\mathbb{R}^N$, we have

$$z^*\Gamma_{\Phi_1(x;h)}z=C_H\int_{[0,1]^2}\langle\xi_s,\xi_t\rangle_{\mathbb{R}^d}|t-s|^{2H-2}dsdt, \tag{3.17}$$

where ξ is the function in $\mathcal{H}$ defined by

$$\xi_t \triangleq V^*(\Phi_t)(J_t^{-1})^* J_1^* z. \tag{3.18}$$

According to an interpolation inequality proved by Baudoin-Hairer (cf. [2, Proof of Lemma 4.4]), given $\gamma > H - 1/2$, we have

$$\int_{[0,1]^2} \langle f_s, f_t \rangle_{\mathbb{R}^d} |t-s|^{2H-2} ds dt \geq C_\gamma \frac{\left(\int_0^1 v^\gamma (1-v)^\gamma |f_v|^2 dv \right)^2}{\|f\|_\gamma^2} \tag{3.19}$$

for all $f \in C^\gamma([0,1];\mathbb{R}^d)$. Observe that, due to our uniform ellipticity assumption (3.1) and the non-degeneracy of J_t, we have

$$\inf_{0 \leq t \leq 1} |\xi_t|^2 \geq C_{H,V,M} |z|^2. \tag{3.20}$$

Furthermore, recall that Φ_t is defined in (1.4) and is driven by $h \in \bar{\mathcal{H}}$. We have also seen that $\bar{\mathcal{H}} \hookrightarrow C_0^H$ whenever $H > 1/2$. Thus for $H - 1/2 < \gamma < H$, we get $\|\Phi_t\|_\gamma \leq C_{H,V} \|h\|_\gamma$; and the same inequality holds true for the Jacobian J in (3.18). Therefore, going back to equation (3.18) again, we have

$$\|\xi\|_\gamma^2 \leq C_{H,V,M} \|h\|_{\bar{\mathcal{H}}} |z|^2 \leq C_{H,V,M} |z|^2, \tag{3.21}$$

where the last inequality stems from our assumption $\|h\|_{\bar{\mathcal{H}}} \leq M$. Therefore, taking $f_t = \xi_t$ in (3.19), plugging inequalities (3.20) and (3.21) and recalling inequality (3.17), we conclude that

$$z^* \Gamma_{\Phi_1(x;h)} z \geq C_{H,V,M} |z|^2$$

uniformly for $\|h\|_{\bar{\mathcal{H}}} \leq M$ and the result follows.

(ii) Proof of the lower bound when $H \leq 1/2$. Recall again that (3.15) yields

$$z^* \Gamma_{\Phi_1(x;h)} z = \|z^* D\Phi_1(x;h)\|_{\mathcal{H}}^2.$$

Then owing to the continuous embedding $\mathcal{H} \subseteq L^2([0,1])$ proved in Lemma 2.8, and expression (3.14) for $D\Phi_t$, we have for any $z \in \mathbb{R}^N$,

$$\begin{aligned} z^* \Gamma_{\Phi_1(x;h)} z &\geq C_H \|z^* D\Phi_1(x;h)\|_{L^2([0,1])}^2 \\ &= C_H \int_0^1 z^* J_1 J_t^{-1} V(\Phi_t) V^*(\Phi_t) (J_t^{-1})^* J_1^* z dt. \end{aligned}$$

We can now invoke the uniform ellipticity assumption (3.1) and the non-degeneracy of J_t in order to obtain

$$z^* \Gamma_{\Phi_1(x;h)} z \geq C_{H,V,M} |z|^2$$

uniformly for $\|h\|_{\bar{\mathcal{H}}} \leq M$. Our claim (3.16) now follows as in the case $H > 1/2$. □

With the preliminary results of Proposition 3.5 and Lemma 3.6 in hand, we are now able to complete the proof of Theorem 3.4.

Proof of Theorem 3.4. Recall that $X^\varepsilon(h)$ is defined by (3.10). According to our preliminary bound (3.12), it remains to show that

$$\mathbb{E}\left[\delta_0\left(X^\varepsilon(h)\right)\mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right] \geq C_{H,V} \tag{3.22}$$

uniformly in h for $\|h\|_{\bar{\mathcal{H}}} \leq 2\varepsilon$ when ε is small enough. The proof of this fact consists of the following two steps:

(i) Prove that $\mathbb{E}[\delta_0(X(h))\mathrm{e}^{-I(h/\varepsilon)}] \geq C_{H,V}$ for all $\varepsilon > 0$ and $h \in \bar{\mathcal{H}}$ with $\|h\|_{\bar{\mathcal{H}}} \leq 1$;
(ii) Upper bound the difference

$$\mathbb{E}\left[\delta_0\left(X^\varepsilon(h)\right)\mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right] - \mathbb{E}\left[\delta_0(X(h))\mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right],$$

and show that it is small uniformly in h for $\|h\|_{\bar{\mathcal{H}}} \leq 2\varepsilon$ when ε is small. We now treat the above two parts separately.

Proof of item (i): Recall that the first chaos $\mathcal{C}_1$ has been defined in Section 2.1. then observe that the random variable $X(h) = (X^1(h), ..., X^N(h))$ introduced in Proposition 3.5 sits in $\mathcal{C}_1$. We decompose the Wiener integral $I(h/\varepsilon)$ as

$$I\left(h/\varepsilon\right) = G_1^\varepsilon + G_2^\varepsilon,$$

where G_1^ε and G_2^ε satisfy

$$G_1^\varepsilon \in \mathrm{Span}\{X^i(h); 1 \leq i \leq N\}, \quad G_2^\varepsilon \in \mathrm{Span}\{X^i(h); 1 \leq i \leq N\}^\perp$$

where the orthogonal complement is considered in $\mathcal{C}_1$. With this decomposition in hand, we get

$$\mathbb{E}\left[\delta_0(X(h))\mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right] = \mathbb{E}\left[\delta_0(X(h))\mathrm{e}^{-G_1^\varepsilon}\right] \cdot \mathbb{E}\left[\mathrm{e}^{-G_2^\varepsilon}\right].$$

Furthermore, $\mathbb{E}[\mathrm{e}^G] \geq 1$ for any centered Gaussian random variable G. Thus

$$\mathbb{E}\left[\delta_0(X(h))\mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right] \geq \mathbb{E}\left[\delta_0(X(h))\mathrm{e}^{-G_1^\varepsilon}\right]. \tag{3.23}$$

Next we approximate δ_0 above by a sequence of function $\{\psi_n; n \geq 1\}$ compactly supported in $B(0, 1/n) \subset \mathbb{R}^N$. Taking limits in the right hand-side of (3.23) and recalling that $G_1^\varepsilon \in \mathrm{Span}\{X^i(h); 1 \leq i \leq N\}$, we get

$$\mathbb{E}\left[\delta_0(X(h))\mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right] \geq \mathbb{E}[\delta_0(X(h))].$$

We now resort to the fact that $X(h)$ is a Gaussian random variable with covariance matrix $\Gamma_{\Phi_1(x;h)}$ by (3.15), which satisfies relation (3.16). This yields

$$\mathbb{E}\left[\delta_0(X(h))\mathrm{e}^{-I\left(\frac{h}{\varepsilon}\right)}\right] \geq \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\det \Gamma_{\Phi_1(x;h)}}} \geq C_{H,V},$$

uniformly for $\|h\|_{\bar{\mathcal{H}}} \leq 1$. This ends the proof of item (i).

Proof of item (ii): By using the integration by parts formula in Malliavin's calculus, the expectation $\mathbb{E}[\delta_0(X^\varepsilon(h))\mathrm{e}^{-I(h/\varepsilon)}]$ can be expressed in terms of the Malliavin derivatives of $I(h/\varepsilon)$, $X^\varepsilon(h)$ and the inverse Malliavin covariance matrix $M_{X^\varepsilon(h)}$ of $X^\varepsilon(h)$, and similarly for $\mathbb{E}[\delta_0(X(h))\mathrm{e}^{-I(h/\varepsilon)}]$. In addition, from standard argument (cf. [4, Lemma 3.4]), one can show that $\det M_{X^\varepsilon(h)}$ has negative moments of all orders uniformly for all $\varepsilon \in (0,1)$ and bounded $h \in \bar{\mathcal{H}}$. Together with the convergence $\mathbb{D}\text{-}\lim_{\varepsilon\to 0} X^\varepsilon(h) = X(h)$ in Proposition 3.5, we conclude that

$$\det M^{-1}_{X^\varepsilon(h)} \xrightarrow{L^p} \det M^{-1}_{\Phi_1(x;h)}, \qquad \text{as } \varepsilon \to 0$$

uniformly for $\|h\|_{\bar{\mathcal{H}}} \leq 1$ for each $p \geq 1$. Therefore, the assertion of item (ii) holds.

Once item (i) and (ii) are proved, it is easy to obtain (3.22) and the details are omitted. This finishes te proof of Theorem 3.4. □

4 Hypoelliptic case: local estimate for the control distance function.

In this section, we prove Theorem 1.3 in the hypoelliptic case. In contrast to the elliptic case, it should be noticed that one cannot explicitly construct a Cameron-Martin path joining two points in the sense of Definition 1.2 in any easy way (i.e. no simple analogue of Lemma 3.1 is possible). The analysis of Cameron-Martin norms also becomes more involved. We detail those steps below, starting with some technical lemmas.

4.1 Preliminary results.

As we mentioned above, it is quite difficult to explicitly construct a Cameron-Martin path joining x to y in the sense of differential equation in the hypoelliptic case. However, it is possible to find some $u \in \mathfrak{g}^{(l)}$ joining x to y *in the sense of Taylor approximation*, i.e. $y = x + F_l(u,x)$ as introduced in Definition 2.17. This is the content of the following fundamental lemma proved in [17], which will be crucial for us in the proofs of both Theorem 1.3 and Theorem 1.7. Recall that l_0 is the hypoellipticity constant in the assumption (1.2).

Lemma 4.1. *For each $l \geq l_0$, there exist constants $r, A > 0$ depending only on l and the vector fields, and a C_b^∞-function*

$$\Psi_l : \left\{u \in \mathfrak{g}^{(l)} : \|u\|_{\mathrm{HS}} < r\right\} \times \mathbb{R}^N \times \left\{\eta \in \mathbb{R}^N : |\eta| < r\right\} \to \mathfrak{g}^{(l)},$$

such that for all u, x, η in the domain of Ψ_l, we have:
(i) $\Psi_l(u, x, 0) = u$;
(ii) $\|\Psi_l(u, x, \eta) - u\|_{\mathrm{HS}} \leq A|\eta|$;
(iii) $F_l(\Psi_l(u, x, \eta), x) = F_l(u, x) + \eta$.

The intuition behind the function Ψ_l is the following. Let $y \triangleq x + F_l(u, x)$ so that u joins x to y in the sense of Taylor approximation. Then $v \triangleq \Psi_l(u, x, \eta)$ joins x to $y + \eta$, i.e. $x + F_l(v, x) = y + \eta$. In particular, $\Psi_l(0, x, y - x)$ gives an element in $\mathfrak{g}^{(l)}$ which joins x to y in the sense of Taylor approximation, provided $|y - x| < r$. The proof of this lemma, for which we refer again to [17], is based on a non-degeneracy property of F_l stated in Lemma 5.7 due to hypoellipticity, as well as a parametrized version of the classical inverse function theorem.

We begin with some easy preliminary steps toward the proof of Theorem 1.3, namely the lower bound on $d(\cdot, \cdot)$ and the upper bound for the case $H < 1/2$.

Lemma 4.2. *Assume that the vector fields in equation (1.1) satisfy the uniform hypoellipticity assumption (1.2) with constant l_0. Let $d = d_H$ be the control distance introduced in Definition 1.2. Then the following bounds hold true.*
(i) For all $H \in (1/4, 1)$ and x, y such that $|x - y| \leq 1$, we have

$$d(x, y) \geq C_1 |x - y|.$$

(ii) Whenever $H \in (1/4, 1/2)$ we have

$$d(x, y) \leq C_2 |x - y|^{\frac{1}{l_0}}.$$

Proof. Claim (i) follows from the exact same argument as in the proof of Theorem 3.3. Claim (ii) stems from the fact that when $H < 1/2$,

$$d(x, y) \leq C_H \, d_{\mathrm{BM}}(x, y) \tag{4.1}$$

where d_{BM} stands for the distance for the Brownian motion case. Note that (4.1) can be easily justified by the fact that, according to Lemma 2.8,

$$d(x, y) \leq \|h\|_{\bar{\mathcal{H}}} \leq C_H \|h\|_{W^{1,2}},$$

for any $h \in \Pi_{x,y}$. Then, with (4.1) in hand, our claim (ii) follows from the Brownian hypoelliptic analysis [17].

□

In the remainder of the section, we focus on the case $H > 1/2$. It is not surprising that this is the hardest case since the Cameron-Martin subspace $\bar{\mathcal{H}}$ gets smaller as H increases. First, we need to make use of the following scaling property of Cameron-Martin norm. Namely denote $\bar{\mathcal{H}}([0, T])$ (respectively, $d_T(x, y)$) as the Cameron-Martin subspace (respectively, the control distance function) associated with fractional Brownian motion over $[0, T]$. Then the following property holds true.

Lemma 4.3. *Let $0 < T_1 < T_2$, and consider $H > 1/2$. Given $h \in \bar{\mathcal{H}}([0,T_1])$, define $\tilde{h}_t \triangleq h_{T_1 t/T_2}$ for $0 \le t \le T_2$. Then $\tilde{h} \in \bar{\mathcal{H}}[0,T_2]$, and*

$$\|\tilde{h}\|_{\bar{\mathcal{H}}([0,T_2])} = \left(\frac{T_1}{T_2}\right)^H \|h\|_{\bar{\mathcal{H}}([0,T_1])}. \tag{4.2}$$

In particular, let d_T be the distance introduced in Definition 1.2 associated with a fractional Brownian motion over $[0,T]$. Then we have

$$d_1(x,y) = T^H d_T(x,y), \quad \forall T > 0,\ x,y \in \mathbb{R}^N. \tag{4.3}$$

Proof. Recall that, thanks to relation (2.8), we have

$$\|\tilde{h}\|_{\bar{\mathcal{H}}([0,T_2])} = \|K^{-1}\tilde{h}\|_{L^2([0,T_2])}. \tag{4.4}$$

Moreover, invoking relation (2.6) for $H > 1/2$, we get

$$(K^{-1}h)_t = C_H \cdot t^{H-\frac{1}{2}} D_{0^+}^{H-\frac{1}{2}} \left(s^{\frac{1}{2}-H}\dot{h}_s\right)(t). \tag{4.5}$$

Plugging (4.5) into (4.4) and performing an elementary change of variables, one ends up with

$$\|\tilde{h}\|_{\bar{\mathcal{H}}([0,T_2])} = \|K^{-1}\tilde{h}\|_{L^2([0,T_2])} = \left(\frac{T_1}{T_2}\right)^H \|K^{-1}h\|_{L^2([0,T_1])} = \left(\frac{T_1}{T_2}\right)^H \|h\|_{\bar{\mathcal{H}}([0,T_1])},$$

and the assertion (4.2) follows. The second claim (4.3) is now easily deduced. □

Remark 4.4. In fact Lemma 4.3 also holds true for $H \le 1/2$. However, it will only be invoked for the case $H > 1/2$.

We also need the following lemma about the free nilpotent group $G^{(l)}$ which allows us to choose a "regular" path γ with $S_l(\gamma) = u$ for all $u \in G^{(l)}$.

Lemma 4.5. *Let $l \ge 1$. For each $M > 0$, there exists a constant $C = C_{l,M} > 0$, such that for every $u \in G^{(l)}$ with $\|u\|_{\mathrm{CC}} \le M$, one can find a smooth path $\gamma : [0,1] \to \mathbb{R}^d$ which satisfies:*
(i) $S_l(\gamma) = u$;
(ii) $\dot{\gamma}$ is supported on $[1/3, 2/3]$;
(iii) $\|\ddot{\gamma}\|_{\infty;[0,1]} \le C$.

Proof. We first prove the claim for a generic element $u \in \exp(\mathcal{L}_k)$, seen as an element of $G^{(k)}$. Let $\{a_1, \dots, a_{d_k}\}$ be a basis of $\mathcal{L}_k$ where $d_k \triangleq \dim \mathcal{L}_k$. Given $u \in \exp(\mathcal{L}_k)$, we can write $u = \exp(a)$ with

$$a = \lambda_1 a_1 + \cdots + \lambda_{d_k} a_{d_k} \in \mathcal{L}_k \tag{4.6}$$

for some $\lambda_1, \dots, \lambda_{d_k} \in \mathbb{R}$. Since we assume that $\|u\|_{\mathrm{CC}} \le M$, according to the ball-box estimate (cf. Proposition 2.15) and the fact that $a \in \mathcal{L}_k$, we have

$$\|a\|_{\mathrm{HS}} = \|u - \mathbf{1}\|_{\mathrm{HS}} \le C_{1,l,M}. \tag{4.7}$$

Moreover, $\mathcal{L}_k$ is a finite dimensional vector space, on which all norms are equivalent. Thus relation (4.7) yields

$$\max_{1\le i\le d_k} |\lambda_i| \le C_{2,l,M}. \tag{4.8}$$

Now recall from Remark 2.14 that for each a_i in (4.6) one can choose a smooth path $\alpha_i : [0,1] \to \mathbb{R}^d$ such that $S_k(\alpha_i) = \exp(a_i)$ and $\dot\alpha_i$ is supported on $[1/3, 2/3]$. Set

$$R_k \triangleq \max\left\{ \|\ddot\alpha_i\|_{\infty;[0,1]} : 1 \le i \le d_k \right\}.$$

Note that R_k is a constant depending only on k. We construct a smooth path $\gamma : [0, d_k] \to \mathbb{R}^d$ by

$$\gamma \triangleq \left(|\lambda_1|^{\frac{1}{k}} \alpha_1^{\mathrm{sgn}(\lambda_1)} \right) \sqcup \cdots \sqcup \left(|\lambda_{d_k}|^{\frac{1}{k}} \alpha_{d_k}^{\mathrm{sgn}(\lambda_{d_k})} \right), \tag{4.9}$$

where α_i^{-1} denotes the reverse of α_i, and $\sqcup$ denotes path concatenation. Then $\dot\gamma$ is obviously compactly supported, and we also claim that $S_k(\gamma) = u$. Indeed, it follows from (4.9) that

$$\begin{aligned} S_k(\gamma) &= S_k\left(|\lambda_1|^{\frac{1}{k}} \alpha_1^{\mathrm{sgn}(\lambda_1)} \right) \otimes \cdots \otimes S_k\left(|\lambda_{d_k}|^{\frac{1}{k}} \alpha_{d_k}^{\mathrm{sgn}(\lambda_{d_k})} \right) \\ &= \delta_{|\lambda_1|^{\frac{1}{k}}}\left(S_k\left(\alpha_1^{\mathrm{sgn}(\lambda_1)} \right) \right) \otimes \cdots \otimes \delta_{|\lambda_{d_k}|^{\frac{1}{k}}}\left(S_k\left(\alpha_{d_k}^{\mathrm{sgn}(\lambda_{d_k})} \right) \right) \\ &= \delta_{|\lambda_1|^{\frac{1}{k}}}\left(\exp(\mathrm{sgn}(\lambda_1) a_1) \right) \otimes \cdots \otimes \delta_{|\lambda_{d_k}|^{\frac{1}{k}}}\left(\exp(\mathrm{sgn}(\lambda_{d_k}) a_{d_k}) \right), \end{aligned} \tag{4.10}$$

where we have used the properties of the dilation, recalled in Section 2.2, and the relation between signatures and $G^{(l)}$ given in (2.28) – (2.29). In addition, since each element $\exp(\lambda_i a_i)$ above sits in $\exp(\mathcal{L}_k)$, the tensor product in $G^{(k)}$ is reduced to

$$S_k(\gamma) = \exp(\lambda_1 a_1) \otimes \cdots \otimes \exp(\lambda_{d_k} a_{d_k}) = \exp(a) = u. \tag{4.11}$$

We have thus found a path γ with compactly supported derivative such that $S_k(\gamma) = u$. In addition, from the definition of R_k and (4.8), we have

$$\|\ddot\gamma\|_{\infty;[0,d_k]} \le R_k \cdot \left(\max_{1\le i\le d_k} |\lambda_i| \right)^{\frac{1}{k}} \le C_{3,l,M}.$$

By suitable rescaling and adding trivial pieces on both ends if necessary, we may assume that γ is defined on $[0,1]$ and $\dot\gamma$ is supported on $[1/3, 2/3]$. In this way, we have

$$\|\ddot\gamma\|_{\infty;[0,1]} \le C_k \cdot C_{3,k,M} \triangleq C_{4,k,M},$$

where C_k is the constant coming from the rescaling. Therefore, our assertion (i)–(iii) holds for u which are elements of $\exp(\mathcal{L}_k)$.

With the help of the previous special case, we now prove the lemma by induction on l. The case when $l = 1$ is obvious, as we can simply choose γ to be a straight line segment. Suppose now that the claim is true on $G^{(l-1)}$. We let $M > 0$ and $u \in G^{(l)}$ with $\|u\|_{\mathrm{CC}} \le M$.

Define $v \triangleq \pi^{(l-1)}(u)$ where $\pi^{(l-1)}: \ G^{(l)} \to G^{(l-1)}$ is the canonical projection. We obviously have

$$\|v\|_{\mathrm{CC}} \leq \|u\|_{\mathrm{CC}} \leq M,$$

where the CC-norm of v is taken on the group $G^{(l-1)}$. According to the induction hypothesis, there exists a constant $C_{l-1,M}$, such that we can find a smooth path $\alpha : [0,1] \to \mathbb{R}^d$ which satisfies (i)–(iii) in the assertion of Lemma 4.5, for $v = S_{l-1}(\alpha)$ and constant $C_{l-1,M}$. Define

$$w \triangleq (S_l(\alpha))^{-1} \otimes u, \tag{4.12}$$

where the tensor product is defined on $G^{(l)}$. Then note that owing to the fact that $\|u\|_{\mathrm{CC}} \leq M$, we have

$$\|w\|_{\mathrm{CC}} \leq \|S_l(\alpha)\|_{\mathrm{CC}} + \|u\|_{\mathrm{CC}} \leq \|\alpha\|_{1-\mathrm{var};[0,1]} + \|u\|_{\mathrm{CC}} \leq \frac{1}{2}\|\ddot{\alpha}\|_{\infty;[0,1]} + M.$$

Therefore, thanks to the induction procedure applied to $v = S_{l-1}(\alpha)$, we get

$$\|w\|_{\mathrm{CC}} \leq \frac{1}{2}C_{l-1,M} + M \triangleq C_{5,l,M}.$$

We claim that $w \in \exp(\mathcal{L}_l)$. This can be proved in the following way.
(i) Write $u = \exp(l_0 + l_h)$, where $l_0 \in \mathfrak{g}^{(l-1)}$ and $l_h \in \mathcal{L}_l$. Recall $v \triangleq \pi^{(l-1)}(u)$. We argue that $v = \exp(l_0) \in G^{(l-1)}$ as follows: since $l_h \in \mathcal{L}_l$, any product of the form $l_h^p \otimes l_0^q = 0$ whenever $p, q > 0$. Taking into account the definition (2.24) of the exponential function, we get that

$$u = \exp(l_0 + l_h) \implies v = \exp(l_0) \in G^{(l-1)}. \tag{4.13}$$

(ii) Recall that our induction hypothesis asserts that $v = S_{l-1}(\alpha)$, thus according to (4.13) we have $S_{l-1}(\alpha) = \exp(l_0)$. Thanks to the same kind of argument as in (i), we get $S_l(\alpha) = \exp(l_0 + l_h') \in G^{(l)}$ for some $l_h' \in \mathcal{L}_l$.
(iii) In order to conclude that $w \in \exp(\mathcal{L}_l)$, we go back to relation (4.12), which can now be read as

$$w = (\exp(l_0 + l_h'))^{-1} \otimes \exp(l_0 + l_h).$$

According to Campbell-Baker-Hausdorff formula and taking into account the fact that

$$[l_0, l_0] = [l_0, l_h] = [l_0, l_h'] = [l_h, l_h'] = 0 \in \mathfrak{g}^{(l)},$$

we conclude that $w = \exp(l_h - l_h')$ and thus $w \in \exp(\mathcal{L}_l)$.

We are now ready to summarize our information and conclude our induction procedure. Namely, for $u \in G^{(l)}$, we can recast relation (4.12) as

$$u = S_l(\alpha) \otimes w, \tag{4.14}$$

and we have just proved that $w \in \exp(\mathcal{L}_l)$. Hence relation (4.11) asserts that w can be written as $w = S_l(\beta)$, where $\beta : [0,1] \to \mathbb{R}^d$ satisfying relation (i)-(iii) in Lemma 4.5 with

$C = C_{6,l,M}$. Now set $\gamma \triangleq \alpha \sqcup \beta$ and rescale it so that it is defined on $[0,1]$ and its derivative path is supported on $[1/3, 2/3]$. Then, recalling our decomposition (4.14), we have

$$S_l(\gamma) = S_l(\alpha) \otimes S_l(\beta) = S_l(\alpha) \otimes w = u,$$

and, moreover, the following upper bound holds true

$$\|\ddot{\gamma}\|_{\infty;[0,1]} \leq 36 \max\left\{\|\ddot{\alpha}\|_{\infty;[0,1]}, \|\ddot{\beta}\|_{\infty;[0,1]}\right\} \leq C_{7,l,M}.$$

Therefore our induction procedure is established, which finishes the proof.

□

We conclude this subsection by stating a convention on the group $G^{(l)}$ which will ease notation in our future computations.

Convention 4.6. Since $\mathfrak{g}^{(l)}$ is a finite dimensional vector space on which differential calculus is easier to manage, we will frequently identify $G^{(l)}$ with $\mathfrak{g}^{(l)}$ through the exponential diffeomorphism without further mention. This is not too beneficial when proving Theorem 1.3 but will be very convenient when proving Theorem 1.7. In this way, for instance, $S_l(w) = u$ means $S_l(w) = \exp(u)$ if $u \in \mathfrak{g}^{(l)}$. The same convention will apply to other similar relations when the meaning is clear from context. For norms on $\mathfrak{g}^{(l)}$, we denote $\|u\|_{\mathrm{CC}} \triangleq \|\exp(u)\|_{\mathrm{CC}}$. As for the HS-norm, note that

$$C_{1,l}\|u\|_{\mathrm{HS}} \leq \|\exp(u) - \mathbf{1}\|_{\mathrm{HS}} \leq C_{2,l}\|u\|_{\mathrm{HS}}$$

for all $u \in \mathfrak{g}^{(l)}$ satisfying $\|\exp(u) - \mathbf{1}\|_{\mathrm{HS}} \wedge \|u\|_{\mathrm{HS}} \leq 1$. Therefore, up to a constant depending only on l, the notation $\|u\|_{\mathrm{HS}}$ can either mean the HS-norm of u or $\exp(u) - \mathbf{1}$. This will not matter because we are only concerned with local estimates. The same convention applies to the distance functions ρ_{CC} and ρ_{HS}.

4.2 Proof of Theorem 1.3.

In this section we give the details in order to complete the proof of Theorem 1.3. Notice that thanks to our preliminary Lemma 4.2, we only focus on the upper bound on the distance d for $H > 1/2$.

Recall that $\Psi_l(u, x, \eta)$ is the function given by Lemma 4.1. This function allows us to construct elements in $\mathfrak{g}^{(l)}$ joining two points in the sense of Taylor approximation locally. In the following, we take $l = l_0$ (where l_0 stands for the hypoellipticity constant) and we will omit the subscript l for simplicity (e.g., $F = F_l$ and $\Psi = \Psi_l$) . We will also identify $G^{(l)}$ with $\mathfrak{g}^{(l)}$ in the way mentioned in Convention 4.6. We now divide our proof in several steps.

Step 1: Construction of an approximating sequence. Let $\delta < r$ be a constant to be chosen later on, where r is the constant appearing in the domain of Ψ in Lemma 4.1. Consider $x, y \in \mathbb{R}^N$ with $|x - y| < \delta$.

We are going to construct three sequences $\{x_m\} \subseteq \mathbb{R}^N$, $\{u_m\} \subseteq \mathfrak{g}^{(l_0)}$, $\{h_m\} \subseteq C^\infty([0,1];\mathbb{R}^d)$ inductively. We start with $x_1 \triangleq x$ and define the rest of them by the following general procedure in the order

$$u_1 \to h_1 \to x_2 \to u_2 \to h_2 \to x_3 \to \cdots .$$

To this aim, suppose we have already defined x_m. Set

$$u_m \triangleq \Psi(0, x_m, y - x_m), \quad \text{and} \quad \bar{u}_m \triangleq \delta_{\|u_m\|_{\mathrm{CC}}^{-1}} u_m. \tag{4.15}$$

By Lemma 4.1, the first condition in (4.15) states that u_m is an element of $\mathfrak{g}^{(l_0)}$ such that

$$x_m + F(u_m, x_m) = y, \tag{4.16}$$

while the second condition in (4.15) ensures that $\|\bar{u}_m\|_{\mathrm{CC}} = 1$. Once u_m is defined, we construct h_m in the following way: let $\bar{h}_m : [0,1] \to \mathbb{R}^d$ be the smooth path given by Lemma 4.5 such that $S_{l_0}(\bar{h}_m) = \bar{u}_m$, $\dot{\bar{h}}_m$ is supported on $[1/3, 2/3]$, and $\|\ddot{\bar{h}}_m\|_{\infty;[0,1]} \leq C_{l_0}$. Define

$$h_m \triangleq \|u_m\|_{\mathrm{CC}} \bar{h}_m, \tag{4.17}$$

so that the truncated signature of h_m is exactly u_m (here recall the Convention 4.6). More specifically, we have:

$$S_{l_0}(h_m) = S_{l_0}(\|u_m\|_{\mathrm{CC}} \cdot \bar{h}_m) = \delta_{\|u_m\|_{\mathrm{CC}}}(S_{l_0}(\bar{h}_m)) = \delta_{\|u_m\|_{\mathrm{CC}}}(\overline{u}_m) = u_m.$$

Taking into account the definition (2.32) of the CC-norm, it is immediate that

$$\|u_m\|_{\mathrm{CC}} \leq \|h_m\|_{1\text{-var};[0,1]} \leq \|u_m\|_{\mathrm{CC}} \|\bar{h}_m\|_{1\text{-var};[0,1]} \leq C_{l_0} \|u_m\|_{\mathrm{CC}}\,, \tag{4.18}$$

where the last inequality stems from the fact that $\bar{h}_m$ has a bounded second derivative. Eventually we define

$$x_{m+1} \triangleq \Phi_1(x_m; h_m), \tag{4.19}$$

where recall that $\Phi_t(x; h)$ is the solution flow of the ODE (2.34) driven by h over $[0,1]$.

Step 2: Checking the condition $|y - x_m| < r$. Recall that in Lemma 4.1 we have to impose $\|u\|_{\mathrm{HS}} < r$ and $|\eta| < r$ in order to apply Ψ. In the context of (4.15) it means that we should make sure that

$$|y - x_m| < r, \quad \text{for all } m. \tag{4.20}$$

We will now choose δ_1 small enough such that if $|y - x| < \delta_1$, then (4.20) is satisfied. This will guarantee that u_m is well-defined by Lemma 4.1 and we will also be able to write down several useful estimates for x_m and u_m. Our first condition on δ_1 is that $\delta_1 \leq r$, so that if $|x - y| < \delta_1$, we can define u_1 by a direct application of Lemma 4.1. We will now prove by induction that if δ_1 is chosen small enough, then condition (4.20) is satisfied. To this aim,

assume that $|x_m - y| < \delta_1$. Then one can apply Lemma 4.1 in order to define u_m, h_m and x_{m+1}. We also get the following estimate:

$$\|u_m\|_{\mathrm{HS}} \leq A|x_m - y| < A\delta_1, \tag{4.21}$$

where A is the constant appearing in Lemma 4.1. In addition, let us require $\delta_1 \leq 1/A$ so that $\|u_m\|_{\mathrm{HS}} \leq 1$. Recalling relations (4.16) and (4.19) we get

$$|x_{m+1} - y| = |\Phi_1(x_m, h_m) - x_m - F(S_{l_0}(h_m), x_m)|.$$

Thus applying successively the Taylor type estimate of [13, Proposition 10.3] and relation (4.18) we end up with

$$|x_{m+1} - y| \leq C_{V,l_0}\|h_m\|_{1-\mathrm{var};[0,1]}^{1+l_0} \leq C_{V,l_0}\|u_m\|_{\mathrm{CC}}^{1+l_0}.$$

The quantity $\|u_m\|_{\mathrm{CC}}$ above can be bounded thanks to the ball-box estimate of Proposition 2.15, for which we observe that the dominating term in (2.33) is $\rho_{\mathrm{HS}}(g_1, g_2)^{1/l_0}$ since our element u_m is bounded by one in HS-norm. We get

$$|x_{m+1} - y| \leq C_{V,l_0}\|u_m\|_{\mathrm{CC}}^{1+l_0} \leq C_{V,l_0}\|u_m\|_{\mathrm{HS}}^{1+\frac{1}{l_0}} \leq C_{V,l_0}A^{1+\frac{1}{l_0}}|x_m - y|^{1+\frac{1}{l_0}}.$$

Summarizing our considerations so far, we have obtained the estimate

$$|x_{m+1} - y| \leq C_{1,V,l_0}\|u_m\|_{\mathrm{CC}}^{1+l_0} \leq C_{2,V,l_0}|x_m - y|^{1+\frac{1}{l_0}}. \tag{4.22}$$

On top of the inequalities $\delta_1 < r$ and $\delta_1 \leq 1/A$ imposed previously, we will also assume that $C_{2,V,l_0}\delta_1^{1/l_0} \leq 1/2$, which easily yields the relation

$$|x_{m+1} - y| \leq \frac{1}{2}|x_m - y| < \frac{1}{2}\delta_1 < \delta_1. \tag{4.23}$$

For our future computations we will thus set

$$\delta_1 \triangleq r \wedge A^{-1} \wedge (2C_{2,V,l_0})^{-l_0}.$$

According to our bound (4.23), we can guarantee that if $|x - y| < \delta_1$, then $|x_m - y| < \delta_1 < r$ for all m. In addition, an easy induction procedure performed on inequality (4.23) leads to the following relation, valid for all $m \geq 1$:

$$|x_m - y| \leq 2^{-(m-1)}|x - y|. \tag{4.24}$$

Together with the second inequality of (4.22), we obtain that

$$\|u_m\|_{\mathrm{CC}} \leq C_{3,V,l_0}2^{-\frac{m}{l_0}}|x - y|^{\frac{1}{l_0}}, \quad \forall m \geq 1. \tag{4.25}$$

We will now choose a constant $\delta_2 \leq \delta_1$ such that the sequence $\{\|u_m\|_{\mathrm{CC}}; m \geq 1\}$ is decreasing with m when $|x - y| < \delta_2$. This property will be useful for our future considerations. Towards this aim, observe that applying successively (2.33), (4.21) and (4.22) we get

$$\|u_{m+1}\|_{\mathrm{CC}} \leq C_{l_0}\|u_{m+1}\|_{\mathrm{HS}}^{\frac{1}{l_0}} \leq C_{4,V,l_0}\|u_m\|_{\mathrm{CC}}^{1+\frac{1}{l_0}}. \tag{4.26}$$

Hence involving the second inequality in (4.22) we have

$$\|u_{m+1}\|_{\mathrm{CC}} \leq C_{5,V,l_0}|x-y|^{\frac{1}{l_0^2}}\|u_m\|_{\mathrm{CC}}. \tag{4.27}$$

Therefore, let us consider a new constant $\delta_2 > 0$ such that

$$C_{5,V,l_0}\delta_2^{\frac{1}{l_0^2}} < 1.$$

If we choose $|x-y| < \delta$ with $\delta \triangleq \delta_1 \wedge \delta_2$, equation (4.27) can be recast as

$$\|u_{m+1}\|_{\mathrm{CC}} \leq \|u_m\|_{\mathrm{CC}}. \tag{4.28}$$

Note that $\delta = \delta_1 \wedge \delta_2$ depends only on l_0 and the vector fields, but not on the Hurst parameter H. We have thus shown that the application of Lemma 4.1 is valid in our context.

Step 3: Construction of a path joining x and y in the sense of differential equation. Our next aim is to obtain a path $\tilde{h}$ joining x and y along the flow of equation (2.34). Our first step in this direction is to rescale h_m in a suitable way. Namely, set $a_1 \triangleq 0$, and for $m \geq 1$, define recursively the following sequence:

$$a_{m+1} \triangleq \sum_{k=1}^{m} \|u_m\|_{\mathrm{CC}}, \quad I_m \triangleq [a_m, a_{m+1}], \quad I \triangleq \overline{\bigcup_{m=1}^{\infty} I_m}.$$

It is clear that $|I_m| = \|u_m\|_{\mathrm{CC}}$, and I is a compact interval since the sequence $\{\|u_m\|_{\mathrm{CC}};\, m \geq 1\}$ is summable according to (4.25). We also define a family of function $\{\tilde{h}_m, m \geq 1\}$ by

$$\tilde{h}_m(t) \triangleq h_m\left(\frac{t-a_m}{a_{m+1}-a_m}\right), \quad t \in I_m, \tag{4.29}$$

and the concatenation of the first $\tilde{h}_m$'s is

$$\tilde{h}^{(m)} \triangleq \tilde{h}_1 \sqcup \cdots \sqcup \tilde{h}_m : \ [0, a_{m+1}] \to \mathbb{R}^d. \tag{4.30}$$

We will now bound the derivative of $\tilde{h}_m$. Specifically, we first use equation (4.29) to get

$$\sup_{m\geq 1} \|\dot{\tilde{h}}^{(m)}\|_{\infty;[0,a_{m+1}]} = \sup_{m\geq 1} \|\dot{\tilde{h}}_m\|_{\infty;I_m} = \sup_{m\geq 1} \frac{1}{|I_m|} \cdot \|\dot{h}_m\|_{\infty;[0,1]}.$$

Then resort to relation (4.17), which yields

$$\sup_{m\geq 1} \|\dot{\tilde{h}}^{(m)}\|_{\infty;[0,a_{m+1}]} = \sup_{m\geq 1} \left\{ \frac{\|u_m\|_{\mathrm{CC}}}{|I_m|} \cdot \|\dot{\tilde{h}}_m\|_{\infty;[0,1]} \right\}.$$

Since $\|u_m\|_{\mathrm{CC}} = |I_m|$ we end up with

$$\sup_{m\geq 1} \|\dot{\tilde{h}}^{(m)}\|_{\infty;[0,a_{m+1}]} = \sup_{m\geq 1} \left\{ \|\dot{\tilde{h}}_m\|_{\infty;[0,1]} \right\} \leq C_{l_0}, \tag{4.31}$$

where the last inequality stems from the fact that $\|\ddot{\bar{h}}_m\|_{\infty;[0,1]} \leq C_{l_0}$.

We can now proceed to the construction of the announced path joining x and y. Namely, set

$$\tilde{h} \triangleq \sqcup_{m=1}^{\infty} \tilde{h}_m : I \to \mathbb{R}^d. \tag{4.32}$$

Then according to (4.31) we have that $\tilde{h}$ is a smooth function from I to $\mathbb{R}^d$. We also claim that $\Phi_1(x;\tilde{h}) = y$, where Φ has to be understood in the sense of equation (1.4). Indeed, set

$$z_t = \Phi_t(x;\tilde{h}), \quad t \in I.$$

From the construction of x_m in (4.19) and the fact that $\tilde{h}|_{[0,a_{m+1}]} = \tilde{h}^{(m)}$ asserted in (4.32), we have

$$x_{m+1} = x + \sum_{\alpha=1}^{d} \int_0^{a_{m+1}} V_\alpha(z_t) d\tilde{h}_t^\alpha. \tag{4.33}$$

Since $x_{m+1} \to y$ as $m \to \infty$, which can be easily seen from (4.24), one can take limits in (4.33) and we conclude that

$$y = x + \sum_{\alpha=1}^{d} \int_0^{|I|} V_\alpha(z_t) d\tilde{h}_t^\alpha.$$

We have thus proved that $\tilde{h}$ is a smooth path joining x and y in the sense of differential equations.

Step 4: Strategy for the upper bound. Let us recall that the family of distances $\{d_T; T > 0\}$ has been introduced in Lemma 4.3, and that they satisfy the scaling property (4.3). Therefore we get

$$\begin{aligned} d(x,y) &= |I|^H d_{|I|}(x,y) \leq |I|^H \|\tilde{h}\|_{\bar{\mathcal{H}}([0,|I|])} \\ &= \lim_{m\to\infty} \left(\left(\sum_{k=1}^{m} |I_k| \right)^H \|\tilde{h}^{(m)}\|_{\bar{\mathcal{H}}([0,a_{m+1}])} \right), \end{aligned} \tag{4.34}$$

where the last relation stems from the definition (4.32) of $\tilde{h}$.

In order to estimate the right hand-side of (4.34), we use the definition (2.8) of the Cameron-Martin norm to get

$$\|\tilde{h}^{(m)}\|^2_{\bar{\mathcal{H}}([0,a_{m+1}])} = \|K^{-1}\tilde{h}^{(m)}\|^2_{L^2([0,a_{m+1}];dt)}.$$

We now invoke the formula (2.6) for K, from which a formula for K^{-1} is easily deduced. We end up with

$$\|\tilde{h}^{(m)}\|^2_{\bar{\mathcal{H}}([0,a_{m+1}])} = C_H \int_0^{a_{m+1}} \left| t^{H-\frac{1}{2}} D_{0+}^{H-\frac{1}{2}} (s^{\frac{1}{2}-H} \dot{\tilde{h}}^{(m)}(s))(t) \right|^2 dt.$$

Taking into account formula (2.5) for the fractional derivative, this yields

$$\|\tilde{h}^{(m)}\|^2_{\bar{\mathcal{H}}([0,a_{m+1}])} = C_H \cdot \int_0^{a_{m+1}} \Big| t^{H-\frac{1}{2}} \Big(t^{1-2H} \dot{\tilde{h}}^{(m)}(t) + \left(H - \frac{1}{2}\right) \int_0^t \frac{t^{\frac{1}{2}-H} \dot{\tilde{h}}^{(m)}(t) - s^{\frac{1}{2}-H} \dot{\tilde{h}}^{(m)}(s)}{(t-s)^{H+\frac{1}{2}}} ds \Big) \Big|^2 dt.$$

We now split the interval $[0, a_{m+1}]$ as $[0, a_{m+1}] = \cup_{k=0}^m I_k$ and use the elementary inequality $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$ in order to get

$$\|\tilde{h}^{(m)}\|^2_{\bar{\mathcal{H}}([0,a_{m+1}])} \leq Q_1 + Q_2 + Q_3, \tag{4.35}$$

with

$$Q_1 \triangleq \sum_{k=1}^m \int_{I_k} \left| t^{H-\frac{1}{2}} \left(t^{1-2H} \dot{\tilde{h}}_k(t) \right) \right|^2 dt \triangleq \sum_{k=1}^m Q_{1,k}, \tag{4.36}$$

$$Q_2 \triangleq \sum_{k=1}^m \int_{I_k} \left| t^{H-\frac{1}{2}} \sum_{l=1}^{k-1} \int_{I_l} \frac{t^{\frac{1}{2}-H} \dot{\tilde{h}}_k(t) - s^{\frac{1}{2}-H} \dot{\tilde{h}}_l(s)}{(t-s)^{H+\frac{1}{2}}} ds \right|^2 dt \triangleq \sum_{k=1}^m Q_{2,k}, \tag{4.37}$$

$$Q_3 \triangleq \sum_{k=1}^m \int_{I_k} \left| t^{H-\frac{1}{2}} \int_{a_k}^t \frac{t^{\frac{1}{2}-H} \dot{\tilde{h}}_k(t) - s^{\frac{1}{2}-H} \dot{\tilde{h}}_k(s)}{(t-s)^{H+\frac{1}{2}}} ds \right|^2 dt \triangleq \sum_{k=1}^m Q_{3,k}. \tag{4.38}$$

We now bound the above three terms separately.

Step 5: Bound for Q_1. In order to bound each $Q_{1,k}$ in the definition (4.36) of Q_1, we just resort to (4.29) which allows to write

$$Q_{1,k} = \int_{I_k} t^{1-2H} \left| \dot{\tilde{h}}_k(t) \right|^2 dt = \int_{I_k} t^{1-2H} \left| \frac{1}{|I_k|} \dot{h}_k \left(\frac{t - a_k}{a_{k+1} - a_k} \right) \right|^2 dt.$$

Then the elementary change of variable

$$v = (t - a_k)/(a_{k+1} - a_k)$$

yields

$$Q_{1,k} = \frac{1}{|I_k|} \int_0^1 (a_k + v|I_k|)^{1-2H} \left| \dot{h}_k(v) \right|^2 dv.$$

We now wish to express Q_1 in terms of $\bar{h}_m$. To this aim, recall from (4.17) that we have

$$Q_{1,k} = \frac{\|u_k\|^2_{\mathrm{CC}}}{|I_k|} \int_0^1 |\varphi_k(v)|^2 dv,$$

where we have set

$$\varphi_k(v) = (a_k + v|I_k|)^{\frac{1}{2}-H} \dot{\bar{h}}_k(v). \tag{4.39}$$

We also recall that each $\bar{h}_k$ has bounded second derivative and is supported in $[1/3, 2/3]$. In addition, we have seen previously that $\|u_k\|_{\mathrm{CC}} = |I_k|$. Thus we have

$$Q_{1,k} = |I_k| \int_{1/3}^{2/3} |\varphi_k(v)|^2 dv$$

We now bound the terms $a_k + v|I_k|$ in the definition of $\varphi_k(v)$ uniformly by $\sum_{j=1}^k |I_j|$. We obtain

$$Q_{1,k} \leq C_H |I_k| \left(\sum_{j=1}^k |I_k| \right)^{1-2H} \|\dot{\bar{h}}_k\|^2_{\infty;[0,1]} \leq C_{H,l_0} \frac{|I_k|}{\left(\sum_{j=1}^m |I_k|\right)^{2H-1}}, \tag{4.40}$$

where in the last step we have used the fact that $\dot{\bar{h}}_k$ is bounded. Therefore, summing relation (4.40) over k and recalling that $Q_1 = \sum_{k=1}^m Q_{1,k}$ we get

$$Q_1 \leq C_{H,l_0} \frac{\sum_{k=1}^m |I_k|}{\left(\sum_{j=1}^m |I_k|\right)^{2H-1}} \leq C_{H,l_0} \sum_{k=1}^m |I_k|^{2(1-H)} \leq C_{H,l_0} \left(\sum_{k=1}^m |I_k| \right)^{2(1-H)}, \tag{4.41}$$

where we use the relation $2(1-H) < 1$ for the last step. This concludes our estimate for the term Q_1.

Step 6: Bound for Q_3. As in the previous step, we first upper bound each term $Q_{3,k}$ separately. To this aim, we perform the same elementary change of variable as for $Q_{1,k}$ above, which allow to express $Q_{3,k}$ in terms of $\bar{h}_k$ instead of $\tilde{h}_k$. We let the patient reader check that we have

$$Q_{3,k} = |I_k|^{2(1-H)} \int_0^1 \left| (a_k + v_1|I_k|)^{H-\frac{1}{2}} \cdot \int_0^{v_1} \frac{\varphi_k(v_1) - \varphi_k(v_2)}{(v_1 - v_2)^{H+\frac{1}{2}}} dv_2 \right|^2 dv_1, \tag{4.42}$$

where we recall that the function φ_k has been introduced in (4.39).

Next we express the derivative of each φ_k in the following way,

$$\frac{d\varphi_k}{du} = \frac{\ddot{\bar{h}}_k(u)}{(a_k + u|I_k|)^{H-\frac{1}{2}}} + \left(\frac{1}{2} - H\right) \cdot \frac{|I_k| \, \dot{\bar{h}}_k(u)}{(a_k + u|I_k|)^{H+\frac{1}{2}}}. \tag{4.43}$$

Hence, since $\bar{h}_k$ is supported on $[1/3, 2/3]$ and $\|\ddot{\bar{h}}_k\|_{\infty;[0,1]} \leq C_{l_0}$, it is readily checked from (4.43) that

$$\left| \frac{d\varphi_k}{du} \right| \leq \frac{C_{H,l_0}}{(\sum_{j=1}^k |I_j|)^{H-\frac{1}{2}}}.$$

Plugging this information into (4.42) and bounding all the terms $a_k + v|I_k|$ uniformly by $\sum_{j=1}^k |I_j|$, we end up with

$$\begin{aligned} Q_{3,k} &\leq C_{H,l_0} |I_k|^{2(1-H)} \int_0^1 \left(\sum_{j=1}^k |I_j| \right)^{2H-1} \cdot \left| \int_0^{v_1} \frac{dv_2}{(\sum_{j=1}^k |I_j|)^{H-\frac{1}{2}} (v_1 - v_2)^{H-\frac{1}{2}}} \right|^2 dv_1 \\ &\leq C_{H,l_0} |I_k|^{2(1-H)}. \end{aligned}$$

As for relation (4.41), we can now sum the previous bounds over k, which yields the following estimate for Q_3 ,

$$Q_3 \leq C_{H,l_0} \sum_{k=1}^{m} |I_k|^{2(1-H)}. \tag{4.44}$$

Step 7: Bound for Q_2. We now turn to the estimation of Q_2, which is more involved than Q_1 and Q_3. We adopt the same strategy as in the previous steps, that is, we handle each $Q_{2,k}$ in (4.37) separately and we resort to the elementary change of variables

$$u \triangleq \frac{t-a_k}{a_{k+1}-a_k}, \quad \text{and} \quad v \triangleq \frac{s-a_l}{a_{l+1}-a_l}.$$

We also express the terms $\dot{\bar{h}}_k$ in (4.37) in terms of $\dot{\bar{h}}_k$. Thanks to some easy algebraic manipulations, we get

$$Q_{2,k} = \int_0^1 \left| \sum_{l=1}^{k-1} \int_0^1 \frac{\frac{\dot{h}_k(u)}{|I_k|} - \left(\frac{a_k+u|I_k|}{a_l+v|I_l|}\right)^{H-\frac{1}{2}} \cdot \frac{\dot{h}_l(v)}{|I_l|}}{(a_k+u|I_k|-a_l-v|I_l|)^{H+\frac{1}{2}}} |I_l| dv \right|^2 |I_k| du. \tag{4.45}$$

In the expression above, notice that for $l \leq k-1$ we have

$$a_k + u|I_k| - a_l - v|I_l| = q_{k,l}(u,v),$$

where

$$q_{k,l}(u,v) = (1-v)|I_l| + |I_{l+1}| + \cdots + |I_{k-1}| + u|I_k|. \tag{4.46}$$

Therefore, invoking the trivial bounds $a_k + u|I_k| \leq \sum_{j_1=1}^{k} |I_{j_1}|$ and $a_l + v|I_l| \geq \sum_{j_2=1}^{l-1} |I_{j_2}|$, and bounding trivially the differences by sums, we obtain

$$Q_{2,k} \leq C_H \int_0^1 \left| \sum_{l=1}^{k-1} \int_0^1 \frac{\left|\dot{\bar{h}}_k(u)\right| + \left(\frac{\sum_{j_1=1}^{k}|I_{j_1}|}{\sum_{j_2=1}^{l-1}|I_{j_2}|}\right)^{H-\frac{1}{2}} \cdot \left|\dot{\bar{h}}_l(v)\right|}{|q_{k,l}(u,v)|^{H+\frac{1}{2}}} |I_l| dv \right|^2 |I_k| du. \tag{4.47}$$

In order to obtain a sharp estimate in (4.47), we want to take advantage of the fact that $\dot{\bar{h}}_l$ is supported on $[1/3, 2/3]$ and therefore avoids the singularities in u, v close to 0 and 1. We thus introduce the intervals

$$J_1 \triangleq [0, 1/3], \;\; J_2 \triangleq [1/3, 2/3], \;\; J_3 \triangleq [2/3, 1]$$

and decompose the expression (4.47) as follows,

$$Q_{2,k} \leq C_H \sum_{p,q=1}^{3} L_{k,p,q},$$

where the quantity $L_{k,p,q}$ is defined by

$$L_{k,p,q} \triangleq \int_{J_p} \left| \sum_{l=1}^{k-1} \int_{J_q} \frac{\left|\dot{\bar{h}}_k(u)\right| + \left(\frac{|I_1|+\cdots+|I_k|}{|I_1|+\cdots+|I_{l-1}|}\right)^{H-\frac{1}{2}} \cdot \left|\dot{\bar{h}}_l(v)\right|}{|q_{k,l}(u,v)|^{H+\frac{1}{2}}} |I_l| dv \right|^2 |I_k| du, \tag{4.48}$$

for all $p, q = 1, 2, 3$. Notice again that since all the $\dot{\bar{h}}_k$ are supported on $[1/3, 2/3]$, the only non-vanishing $L_{k,p,q}$'s are those for which $p = 2$ or $q = 2$. Let us show how to handle the terms $L_{k,p,q}$ given by (4.48), according to $q = 1, 2$ and $q = 3$.

Whenever $q = 1$ or $q = 2$, regardless of the value of p, it is easily seen from (4.46) that we can bound $q_{k,l}(u, v)$ from below uniformly by $C \sum_{j=l}^{k-1} |I_j|$. Thanks again to the fact that $\dot{\bar{h}}_k$ is uniformly bounded for all k, we obtain

$$\begin{aligned} &\frac{\left|\dot{\bar{h}}_k(u)\right| + \left(\frac{|I_1|+\cdots+|I_k|}{|I_1|+\cdots+|I_{l-1}|}\right)^{H-\frac{1}{2}} \cdot \left|\dot{\bar{h}}_l(v)\right|}{q_{k,l}(u,v)^{H+\frac{1}{2}}} \\ &\leq \frac{C_{H,l_0}}{(|I_l| + \cdots + |I_{k-1}|)^{H+\frac{1}{2}}} \cdot \left(\frac{|I_1| + \cdots + |I_k|}{|I_1| + \cdots + |I_{l-1}|}\right)^{H-\frac{1}{2}}. \end{aligned}$$

Summing the above quantity over l and integrating over $[0, 1]$, we end up with

$$L_{k,p,q} \leq C_{H,l_0} |I_k| \cdot \left(\sum_{l=1}^{k-1} \frac{|I_l|}{(|I_l| + \cdots + |I_{k-1}|)^{H+\frac{1}{2}}} \cdot \left(\frac{|I_1| + \cdots + |I_k|}{|I_1| + \cdots + |I_{l-1}|}\right)^{H-\frac{1}{2}} \right)^2.$$

By lower bounding the quantity $|I_1| + \cdots + |I_{l-1}|$ above uniformly by $|I_1|$, we get

$$L_{k,p,q} \leq C_{H,l_0} |I_k| \cdot \left(\frac{\sum_{j=1}^{k} |I_j|}{|I_1|}\right)^{2H-1} \cdot \left(\sum_{l=1}^{k-1} |I_l|^{\frac{1}{2}-H}\right)^2. \tag{4.49}$$

Recall that we have shown in (4.28) that $m \mapsto \|u_m\|_{\mathrm{CC}}$ is a decreasing sequence. Since $\|u_m\|_{\mathrm{CC}} = |I_m|$ we can bound uniformly $\sum_{l=1}^{k-1} |I_l|^{1/2-H}$ by $k|I_1|^{1/2-H}$ and $|I_1|^{-1} \sum_{j=1}^{k} |I_j|$ by k. Plugging this information into (4.49) we obtain,

$$L_{k,p,q} \leq C_{H,l_0} k^{2H+1} |I_k|^{2(1-H)}, \tag{4.50}$$

which is our bound for $L_{k,p,q}$ when $q \in \{1, 2\}$.

Let us now bound $L_{k,p,q}$ for $q = 3$ and $p = 2$. In this case, going back to the definition (4.48) of $L_{k,p,q}$, we have that $\dot{\bar{h}}_l(v) = 0$ for $v \in J_q$. Thus we get

$$\begin{aligned} L_{k,2,3} &= \int_{J_2} \left| \sum_{l=1}^{k-1} \int_{J_3} \frac{\left|\dot{\bar{h}}_k(u)\right| |I_l| dv}{((1-v)|I_l| + |I_{l+1}| + \cdots + |I_{k-1}| + u|I_k|)^{H+\frac{1}{2}}} \right|^2 |I_k| du \\ &\leq C_{H,l_0} \left| \sum_{l=1}^{k-1} \int_{J_3} \frac{|I_l| dv}{((1-v)|I_l| + |I_{l+1}| + \cdots + |I_k|)^{H+\frac{1}{2}}} \right|^2 \cdot |I_k|, \end{aligned} \tag{4.51}$$

where we have used the boundedness of $\dot{\bar{h}}_k$ for the second inequality. We can now evaluate the above v-integral explicitly, which yields

$$\begin{aligned}
&\int_{J_3} \frac{|I_l| dv}{((1-v)|I_l| + |I_{l+1}| + \cdots + |I_k|)^{H+\frac{1}{2}}} \\
=&\frac{1}{\left(H-\frac{1}{2}\right)} \left(\frac{1}{(|I_{l+1}| + \cdots + |I_k|)^{H-\frac{1}{2}}} - \frac{1}{\left(\frac{1}{3}|I_l| + |I_{l+1}| + \cdots + |I_k|\right)^{H-\frac{1}{2}}} \right) \leq \frac{C_H}{|I_k|^{H-\frac{1}{2}}},
\end{aligned}$$

where the second inequality is obtained by lower bounding trivially $|I_{l+1} + \cdots + |I_k|$ by $|I_k|$. Summing this inequality over l and plugging this information into (4.51), we get

$$L_{k,2,3} \leq C_{H,l_0} |I_k| \left(\frac{k}{|I_k|^{H-\frac{1}{2}}} \right)^2 \leq C_{H,l_0} k^{2H+1} |I_k|^{2(1-H)}. \tag{4.52}$$

Summarizing our considerations in this step, we have handled the cases $q = 1, 2$ and $(q, p) = (3, 2)$ in (4.50) and (4.52) respectively. Therefore, we obtain

$$Q_2 \leq \sum_{k=1}^{m} k^{2H+1} |I_k|^{2(1-H)}. \tag{4.53}$$

Step 8: Conclusion. Let us go back to the decomposition (4.35) and plug our bounds (4.41), (4.53) and (4.44) on Q_1, Q_2 and Q_3. We get

$$\|\tilde{h}^{(m)}\|^2_{\tilde{\mathcal{H}}([0,a_{m+1}])} \leq C_H(Q_1 + Q_2 + Q_3) \leq C_{H,l_0} \sum_{k=1}^{m} k^{2H+1} |I_k|^{2(1-H)}.$$

In addition, we have $|I_k| = \|u_k\|_{\mathrm{CC}}$ and relation (4.25) asserts that $k \mapsto \|u_k\|_{\mathrm{CC}}$ decays exponentially. Thus we get

$$\begin{aligned}
&\left(\sum_{k=1}^{m} |I_k| \right)^{2H} \|\tilde{h}^{(m)}\|^2_{\tilde{\mathcal{H}}([0,a_{m+1}])} \leq C_{H,l_0} \left(\sum_{k=1}^{m} |I_k| \right)^{2H} \left(\sum_{k=1}^{m} k^{2H+1} |I_k|^{2(1-H)} \right) \\
&\leq C_{H,V,l_0} \left(\sum_{k=1}^{m} 2^{-\frac{k}{l_0}} \right)^{2H} \cdot \left(\sum_{k=1}^{m} k^{2H+1} 2^{-\frac{2(1-H)}{l_0} k} \right) \cdot |x-y|^{\frac{2}{l_0}} \\
&\leq C_{H,V,l_0} |x-y|^{\frac{2}{l_0}},
\end{aligned} \tag{4.54}$$

where we have trivially bounded the partial geometric series for the last step. Hence the left hand-side of (4.54) converges to a quantity which is lower bounded by $d^2(x, y)$ as $m \to \infty$, thanks to (4.34). Therefore, letting $m \to \infty$ in (4.54) we have obtained

$$d(x, y)^2 \leq C_{H,V,l_0} |x-y|^{\frac{2}{l_0}}, \tag{4.55}$$

which concludes our proof of Theorem 1.3.

5 Hypoelliptic case: local lower estimate for the density of solution.

In this section, we develop the proof of Theorem 1.7 under the uniform hypoellipticity assumption (1.2). As in Section 4, one faces a much more complex situation than in the elliptic case. More specifically, the deterministic Malliavin covariance matrix of X_t^x will not be uniformly non-degenerate (i.e. Lemma 3.6 is no longer true). Without this key ingredient, the whole elliptic argument will break down and one needs new approaches. Our strategy follows the main philosophy of Kusuoka-Stroock [17] in the diffusion case. However, as we will see when we develop the analysis, there are several non-trivial challenges in several key steps for the fractional Brownian setting, which require new ideas and methods. In particular, we shall see how to marry Kusuoka-Stroock's approach and the rough paths formalism.

To increase readability, we first summarize the main strategy of the proof. Our analysis starts from the existence of the truncated signature of order l for the fractional Brownian motion, as asserted in Proposition 2.16. Specifically, with our notation (2.29) in mind, we will write $\Gamma_t \equiv S_l(B)_{0,t} \in G^{(l)}$ as

$$\Gamma_t = S_l(B)_{0,t} = 1 + \sum_{i=1}^{l} \int_{0<t_1<\cdots<t_i<t} dB_{t_1} \otimes \cdots \otimes dB_{t_i}. \tag{5.1}$$

In the sequel we will also use the truncated $\mathfrak{g}^{(l)}$-valued *log-signature* of B, defined by

$$U_t^{(l)} \triangleq \log S_l(B)_{0,t}. \tag{5.2}$$

Notice that $U_t^{(l)}$ features in relation (2.35), and more precisely the process

$$X_l(t,x) \triangleq x + F_l(U_t^{(l)}, x) \tag{5.3}$$

is the Taylor approximation of order l for the solution of the rough equation (1.1) in small time (cf. relation (2.36)).

With those preliminary notation in hand, we decompose the strategy towards the proof of Theorem 1.7 into three major steps.

Step One. According to the scaling property of fractional Brownian motion, a precise local lower estimate on the density of $U_t^{(l)}$ can be easily obtained from a general positivity property.

Step Two. When $l \geq l_0$, the hypoellipticity of the vector fields allows us to obtain a precise local lower estimate on the density of the process $X_l(t,x)$ defined by (5.3) from the estimate on $U_t^{(l)}$ derived in step one.

Step Three. When t is small, the density of $X_l(t,x)$ is close to the density of the actual solution in a reasonable sense, and the latter inherits the lower estimate obtained in step two.

The above philosophy was first proposed by Kusuoka-Stroock [17] in the diffusion case. However, in the fractional Brownian setting, there are several difficulties when implementing

these steps precisely. Conceptually the main challenge arises from the need of respecting the fractional Brownian scaling and the Cameron-Martin structure in each step in order to obtain sharp estimates. More specifically, for Step 1 we need a new idea to prove the positivity for the density of $U_t^{(l)}$ when the Markov property is not available. For Step 2 we rely on the technique we used for proving Theorem 1.3 in Section 4, which yields sharp estimates for the density of $X_l(t,x)$. In Step 3, a new ingredient is needed to prove uniformity for an upper estimate for the density of $X_l(t,x)$ with respect to the degree l of expansion. In the following, we develop the above three steps mathematically.

5.1 Step one: local lower estimate for the signature density of fractional Brownian motion.

We fix $l \geq 1$. Recall that the truncated signature Γ is defined by (5.1). We will now write Γ as the solution of a simple enough rough differential equation. To this aim, let $\{\mathrm{e}_1, \dots, \mathrm{e}_d\}$ be the standard basis of $\mathbb{R}^d$. By viewing this family as vectors in $\mathfrak{g}^{(l)} \cong T_{\mathbf{1}}G^{(l)}$, we denote the associated left invariant vector fields on $G^{(l)}$ by $\{\tilde{W}_1, \dots, \tilde{W}_d\}$. It is standard (cf. [13, Remark 7.43]) that Γ_t satisfies the following intrinsic stochastic differential equation on $G^{(l)}$:

$$\begin{cases} d\Gamma_t = \sum_{\alpha=1}^d \tilde{W}_\alpha(\Gamma_t) dB_t^\alpha, \\ \Gamma_0 = \mathbf{1}. \end{cases} \tag{5.4}$$

Let $U_t \triangleq \log \Gamma_t \in \mathfrak{g}^{(l)}$ be the truncated log-signature path, as defined in (5.2). Since $\{\tilde{W}_1, \dots, \tilde{W}_d\}$ satisfies Hörmander's condition by the definition of $\mathfrak{g}^{(l)}$, we know that U_t admits a smooth density with respect to the Lebesgue measure du on $\mathfrak{g}^{(l)}$. Denote this density by $\rho_t(u)$.

Next we show that the density function ρ_t is everywhere strictly positive. This fact will be important for us. In the Brownian case, this was proved in [17] using support theorem and the semigroup property (or the Markov property). In the fractional Brownian setting, the argument breaks down although general support theorems for Gaussian rough paths are still available. It turns out that there is a simple neat proof based on Sard's theorem and a general positivity criteria of Baudoin-Nualart-Ouyang-Tindel [3]. We mention that Baudoin-Feng-Ouyang [1] also has an independent proof of this fact.

We first recall the classical Sard's theorem, and we refer the reader to [22] for a beautiful presentation. Let $f : M \to N$ be a smooth map between two finite dimensional differentiable manifolds M and N. A point $x \in M$ is said to be a *critical point* of f if the differential $df_x : T_xM \to T_{f(x)}N$ is not surjective. A *critical value* of f in N is the image of a critical point in M. Also recall that a subset $E \subseteq N$ is a *Lebesgue null set* if its intersection with any coordinate chart has zero Lebesgue measure in the corresponding coordinate space.

Theorem 5.1 (Sard's theorem)**.** *Let $f : M \to N$ be a smooth map between two finite dimensional differentiable manifolds. Then the set of critical values of f is a Lebesgue null set in N.*

We now prove the positivity result announced above, which will be important for our future considerations.

Lemma 5.2. *For each $t > 0$, the density ρ_t of the truncated signature path U_t is everywhere strictly positive.*

Proof. We only consider the case when $t = 1$. The general case follows from the scaling property (5.15) below. Our strategy relies on the fact that $\Gamma_t = \exp(U_t)$ solves equation (5.4). In addition, recall our Convention 4.6 about the identification of $\mathfrak{g}^{(l)}$ and $G^{(l)}$. Therefore we can get the desired positivity by applying [3, Theorem 1.4]. To this aim, recall that the standing assumptions in [3, Theorem 1.4] are the following:

(i) The Malliavin covariance matrix of U_t is invertible with inverse in $L^p(\Omega)$ for all $p > 1$;
(ii) The skeleton of equation (5.4), defined similarly to (1.4), generates a submersion. More specifically, we need to show that for any $u \in \mathfrak{g}^{(l)}$, there exists $h \in \bar{\mathcal{H}}$ such that $\log S_l(h) = u$ and

$$(d \log S_l)_h : \bar{\mathcal{H}} \to \mathfrak{g}^{(l)} \quad \text{is surjective,} \tag{5.5}$$

where $S_l(h) \triangleq S_l(h)_{0,1}$ is the truncated map.

Notice that item (i) is proved in [1]. We will thus focus on condition (ii) in the remainder of the proof.

In order to prove relation (5.5) in item (ii) above, let us introduce some additional notation. First we shall write $G \triangleq G^{(l)}$ for the sake of simplicity. Then for all $n \geq 1$ we introduce a linear map $H_n : (\mathbb{R}^d)^n \to \bar{\mathcal{H}}$ in the following way. Given $y = (y_1, \dots, y_n)$, the function $H_n(y)$ is defined to be the piecewise linear path obtained by concatenating the vectors $y_1, \dots, y_n$ successively. We also define a set $\bar{\mathcal{H}}_0$ of piecewise linear paths by

$$\bar{\mathcal{H}}_0 \triangleq \bigcup_{n=1}^{\infty} H_n\left((\mathbb{R}^d)^n\right) \subseteq \bar{\mathcal{H}}.$$

Note that $\bar{\mathcal{H}}_0$ is closed under concatenation, and $S_l(\bar{\mathcal{H}}_0) = G$ by the Chow-Rashevskii theorem (cf. Remark 2.14). Now we claim that:

(**P**) For any $g \in G$, there exists $h \in \bar{\mathcal{H}}_0$ such that $S_l(h) = g$ and the differential $(dS_l)_h|_{\bar{\mathcal{H}}_0} : \bar{\mathcal{H}}_0 \to T_g G$ is surjective.

Note that the property (**P**) is clearly stronger than the original desired claim (5.5). To prove (**P**), let $\mathcal{P}$ be the set of elements in G which satisfy (**P**). We first show that $\mathcal{P}$ is either $\emptyset$ or G. The main idea behind our strategy is that if there exists $g_0 \in \mathcal{P}$, such that $(dS_l)_{h_0}$ is a submersion for some $h_0 \in \bar{\mathcal{H}}_0$ satisfying $S_l(h_0) = g_0$, then one can obtain every point $g \in G$ by a left translation L_a, since dL_a is an isomorphism. To be more precise, suppose that $g_0 \in G$ is an element satisfying (**P**). By definition, there exists a path $h_0 \in \bar{\mathcal{H}}_0$ such that $S_l(h_0) = g_0$ and $(dS_l)_{h_0}|_{\bar{\mathcal{H}}_0}$ is surjective. Now pick a generic element $a \in G$ and choose a path $\alpha \in \bar{\mathcal{H}}_0$ so that $S_l(\alpha) = a$. Then $S_l(\alpha \sqcup h_0) = a \otimes g_0$. We want to show that $(dS_l)_{\alpha \sqcup h_0} : \bar{\mathcal{H}}_0 \to T_{a \otimes g_0} G$ is surjective. For this purpose, let $\xi \in T_{a \otimes g_0} G$ and set

$$\xi_0 \triangleq dL_{a^{-1}}(\xi) \in T_{g_0} G.$$

By the surjectivity of $(dS_l)_{h_0}|_{\bar{\mathcal{H}}_0}$, there exists $\gamma \in \bar{\mathcal{H}}_0$ such that $(dS_l)_{h_0}(\gamma) = \xi_0$. It follows that, for $\varepsilon > 0$ we have

$$S_l(\alpha \sqcup (h_0 + \varepsilon \cdot \gamma)) = a \otimes S_l(h_0 + \varepsilon \cdot \gamma).$$

By differentiation with respect to ε at $\varepsilon = 0$, we obtain that

$$(dS_l)_{\alpha \sqcup h_0}(0 \sqcup \gamma) = (dL_a)_{S_l(h_0)} \circ (dS_l)_{h_0}(\gamma) = (dL_a)_{g_0}(\xi_0) = \xi.$$

Therefore, $(dS_l)_{\alpha \sqcup h_0}|_{\bar{\mathcal{H}}_0}$ is surjective. Since a is arbitrary, we conclude that if $\mathcal{P}$ is non-empty, then $\mathcal{P} = G$.

To complete the proof, it remains to show that $\mathcal{P} \neq \emptyset$. This will be a simple consequence of Sard's theorem. Indeed, for each $n \geq 1$, define

$$f_n \triangleq S_l \circ H_n : (\mathbb{R}^d)^n \to G, \tag{5.6}$$

where we recall that $H_n(y)$ is the piecewise linear path obtained by concatenating $y_1, ..., y_n$. The map f_n is simply given by

$$f_n(y_1, \ldots, y_n) = \exp(y_1) \otimes \cdots \otimes \exp(y_n),$$

where we recall that the exponential maps is defined by (2.24). It is readily checked that f_n is a smooth map. According to Sard's theorem (cf. Theorem 5.1), the set of critical values of f_n, denoted as E_n, is a Lebesgue null set in G. It follows that $E \triangleq \cup_{n=1}^{\infty} E_n$ is also a Lebesgue null set in G. We have thus obtained that,

$$G \backslash E = \left(\bigcup_{n=1}^{\infty} f_n((\mathbb{R}^d)^n) \right) \backslash E \neq \emptyset,$$

where the first equality is due to the fact that $S_l(\bar{\mathcal{H}}_0) = G$ by the Chow-Rashevskii theorem. Pick any element $g \in G \backslash E$. Then for some $n \geq 1$, we have $g \in f_n((\mathbb{R}^d)^n) \backslash E_n$. In particular, there exists $y \in (\mathbb{R}^d)^n$ such that $f_n(y) = g$ and $(df_n)_y$ is surjective. We claim that $g \in \mathcal{P}$ with $h \triangleq H_n(y) \in \bar{\mathcal{H}}_0$ being the associated path. Indeed, it is apparent that $S_l(h) = g$. In addition, let $\xi \in T_g G$ and $w \in (\mathbb{R}^d)^n$ be such that $(df_n)_y(w) = \xi$. The existence of w follows from the surjectivity of $(df_n)_y$. Since H_n is linear, we obtain that

$$\begin{aligned}(dS_l)_h(H_n(w)) &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} S_l(H_n(y) + \varepsilon \cdot H_n(w)) = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} S_l(H_n(y + \varepsilon \cdot w)) \\ &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} f_n(y + \varepsilon \cdot w) = (df_n)_y(w) = \xi.\end{aligned}$$

Therefore, the pair (h, g) satisfies property $(\mathbf{P})$ and thus $\mathcal{P}$ is non-empty. □

Remark 5.3. Some mild technical care is needed in the above proof which we have postponed until now so not to distract the reader from getting the key idea in the proof. One point is that, Theorem 1.4 in [3] was stated for SDEs in which the vector fields are of class C_b^∞. Nevertheless, that theorem relies on properties of the skeleton of a non-degenerate random variable F, which are clearly satisfied for the truncated signature path Γ_t defined by (5.1).

Remark 5.4. Another point is that, when $H > 1/2$ it is not clear whether $\bar{\mathcal{H}}$ contains the space of piecewise linear paths. It is though obvious from $\bar{\mathcal{H}} = I_{0+}^{H+1/2}(L^2([0,1]))$ that it contains all smooth paths. One simple way to fix this issue is to reparametrize the piecewise linear path $y_1 \sqcup \cdots \sqcup y_n$ in a way depending only on n, so that the resulting path is smooth but the trajectory remains unchanged. This will not change the truncated signature as it is invariant under reparametrization. For instance, one can define $H_n(y)$ in a way that on $[(i-1)/n, i/n]$ it is given by

$$H_n(y)_t = y_1 + \cdots y_{i-1} + \left(\int_0^t \eta_i(s)ds\right) y_i,$$

where η_i is a positive smooth function supported on $[\frac{i-1}{n} + \frac{1}{3n}, \frac{i}{n} - \frac{1}{3n}]$ with $\int_{\frac{i-1}{n}}^{\frac{i}{n}} \eta_i(t)dt = 1$. See also [1] for a direct strategy.

Essentially the same amount of effort allows us to adapt the argument in the proof of Lemma 5.2 to establish the general positivity result for hypoelliptic SDEs as stated in Theorem 1.6 which is of independent interest. This complements the result of [3, Theorem 1.4] by affirming that Hypothesis 1.2 in that theorem is always verified under hypoellipticity.

Proof of Theorem 1.6. Without loss of generality we only consider $t = 1$. Continuing to denote by $\Phi_t(x;h)$ the skeleton of equation (1.1), defined by (1.4), let $F : \bar{\mathcal{H}} \to \mathbb{R}^N$ be the end point map defined by $F(h) \triangleq \Phi_1(x;h)$. As in the proof of Lemma 5.2, we wish to check the assumptions of [3, Theorem 1.4]. Recall that this means that we should prove that the Malliavin covariance matrix of X_1 admits an inverse in $L^p(\Omega)$, and that (5.5) holds for the map F. Furthermore, under our standing assumptions, the fact that the Malliavin covariance matrix of X_1 is in $L^p(\Omega)$ is already proved in [7]. We will thus focus on an equivalent of condition (5.5) in the remainder of the proof. Summarizing our considerations so far, we wish to prove that for any $y \in \mathbb{R}^N$ there exists $h \in \bar{\mathcal{H}}$ such that

$$F(h) = y, \quad \text{and} \quad (dF)_h : \bar{\mathcal{H}} \to \mathbb{R}^N \ \text{ is surjective.} \tag{5.7}$$

Along the same lines as in the proof of Lemma 5.2, we define $\mathcal{P}$ to be the set of points $y \in \mathbb{R}^N$ satisfying (5.7) for some $h \in \bar{\mathcal{H}}$. We first show that $\mathcal{P}$ is non-empty which then implies $\mathcal{P} = \mathbb{R}^N$ again by a translation argument.

To show that $\mathcal{P}$ is non-empty, we first define $H_n : (\mathbb{R}^d)^n \to \bar{\mathcal{H}}$ and $\bar{\mathcal{H}}_0 \subseteq \bar{\mathcal{H}}$ in the same way as in the proof of Lemma 5.2. Also define a map F_n by

$$F_n \triangleq F \circ H_n : (\mathbb{R}^d)^n \to \mathbb{R}^N.$$

According to Sard's theorem, the set of critical values of F_n, again denoted as E_n, is a Lebesgue null set in $\mathbb{R}^N$, and so is $E \triangleq \cup_n E_n$.

Next consider a given $q \in \mathbb{R}^N$. Thanks to the hypoellipticity assumption (1.2), we can equip a neighborhood U_q of q with a sub-Riemannian metric, by requiring that a certain subset of $\{V_1, ..., V_d\}$ is an orthonormal frame near q. Then according to the Chow-Rashevskii theorem (cf. [21], Theorem 2.1.2), every point in U_q is reachable from q by a horizontal path.

And if one examines the proof of the theorem in Section 2.4 of [21] carefully, this horizontal path is controlled by a piecewise linear path in $\mathbb{R}^d$, i.e. $U_q \subseteq \cup_n \Phi_1(q; H_n((\mathbb{R}^d)^n))$. Now for given $y \in \mathbb{R}^N$, choose an arbitrary continuous path γ joining x to y. By compactness, we can cover the image of γ by finitely many open sets of the form U_{q_i} such that $U_{q_i} \cap U_{q_{i+1}} \neq \emptyset$ for all i where $q_i \in \mathrm{Im}(\gamma)$. It follows that y can be reached from x by a horizontal path controlled by a piecewise linear path in $\mathbb{R}^d$. In other words, we have $y \in F_n((\mathbb{R}^d)^n)$ for some n. This establishes the property that $\mathbb{R}^N = \cup_n F_n((\mathbb{R}^d)^n)$.

Now the same argument as in the proof of Lemma 5.2 allows us to conclude that

$$\mathbb{R}^N \backslash E = \bigcup_{n=1}^{\infty} F_n((\mathbb{R}^d)^n) \backslash E \subseteq \mathcal{P},$$

showing that $\mathcal{P}$ is non-empty since E is a Lebesgue null set.

Finally, we show that $\mathcal{P} = \mathbb{R}^N$. To this aim, first note that, for any $h_0, \gamma, \alpha \in \bar{\mathcal{H}}$ and $\varepsilon > 0$, we have

$$\Phi_1(x; (h_0 + \varepsilon \cdot \gamma) \sqcup \alpha) = \Phi_1\left(\Phi_1(x; h_0 + \varepsilon \cdot \gamma); \alpha\right),$$

where paths are always assumed to be parametrized on $[0, 1]$. Therefore, by differentiating with respect to ε at $\varepsilon = 0$, we obtain that

$$(dF)_{h_0 \sqcup \alpha}(\gamma \sqcup 0) = J_1(F(h_0); \alpha) \circ (dF)_{h_0}(\gamma),$$

where recall that $J_t(\cdot; \cdot)$ is the Jacobian of the flow Φ_t. This shows that

$$(dF)_{h_0 \sqcup \alpha} = J_1(F(h_0); \alpha) \circ (dF)_{h_0}. \tag{5.8}$$

Now pick any fixed $y_0 \in \mathcal{P}$ with an associated $h_0 \in \bar{\mathcal{H}}$ satisfying (5.7). For any $\eta \in \mathbb{R}^N$, choose $\alpha \in \bar{\mathcal{H}}$ such that $F(\alpha) = \eta$. Then $F(h_0 \sqcup \alpha) = y + \eta$ and the surjectivity of $(dF)_{h_0 \sqcup \alpha}$ follows from (5.8), the surjectivity of $(dF)_{h_0}$ and the invertibility of the Jacobian. In particular, $y + \eta \in \mathcal{P}$. Since η is arbitrary, we conclude that $\mathcal{P} = \mathbb{R}^N$.

□

Remark 5.5. A general support theorem for hypoelliptic SDEs allows one to show that the support of the density $p_t(x, y)$ is dense. In the diffusion case, together with the semigroup property

$$p(s + t, x, y) = \int_{\mathbb{R}^N} p(s, x, z) p(t, z, y) dz$$

one immediately sees that $p(t, x, y)$ is everywhere strictly positive. This argument clearly breaks down in the fractional Brownian setting.

Finally, we present the main result in this part which gives a precise local lower estimate for the density $\rho_t(u)$. In order to get this estimate a first idea wold be to use the stochastic differential equation for U_t, which is obtained by taking logarithm in relation (5.4). Instead of following this strategy, we will resort to some more elementary scaling properties, which stems from the left invariance of the vector fields $\tilde{W}_\alpha$ in (5.4). This is why dealing with U_t is considerably easier than studying the solution to the general SDE (1.1).

Proposition 5.6. *For each $M > 0$, define $\beta_M \triangleq \inf\{\rho_1(u) : \|u\|_{\mathrm{CC}} \leq M\}$. Then β_M is strictly positive and for all $(u,t) \in \mathfrak{g}^{(l)} \times (0,1)$ with $\|u\|_{\mathrm{CC}} \leq Mt^H$, we have*

$$\rho_t(u) \geq \beta_M t^{-H\nu}, \tag{5.9}$$

where the constant ν is given by $\nu \triangleq \sum_{k=1}^{l} k \dim \mathcal{L}_k$, and $\mathcal{L}_k$ is introduced in Definition 2.10.

Proof. First observe that the strict positivity of β_M is an easy consequence of Lemma 5.2 plus the fact that the set $\{u \in \mathfrak{g}^{(l)}; \|u\|_{\mathrm{CC}} \leq M\}$ is compact. Next recall that if $\tilde{W}$ is a left invariant vector field on $G^{(l)}$, the push-forward of $\tilde{W}$ by δ_λ (denoted by $(\delta_\lambda)_*\tilde{W}$) is defined by

$$\left[(\delta_\lambda)_*\tilde{W}\right](\delta_\lambda u) = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \delta_\lambda(u \otimes \exp(\varepsilon \cdot \tilde{W}(\mathbf{1}))).$$

Using this definition, it is readily checked that

$$(\delta_\lambda)_*\tilde{W} = \lambda \cdot \tilde{W}. \tag{5.10}$$

Therefore, applying a change of variable formula to (5.4) with $f(\Gamma_t) = \delta_\lambda \Gamma_t$ and resorting to (5.10) we obtain

$$d(\delta_\lambda \Gamma_t) = \sum_{\alpha=1}^{d} \left((\delta_\lambda)_*\tilde{W}_\alpha\right)(\delta_\lambda \Gamma_t) dB_t^\alpha = \lambda \sum_{\alpha=1}^{d} \tilde{W}_\alpha(\delta_\lambda \Gamma_t) dB_t^\alpha. \tag{5.11}$$

On the other hand, for $\lambda > 0$ set $\Gamma_t^\lambda \triangleq \Gamma_{\lambda^{1/H}t}$. Then we have the following series of identities:

$$d\Gamma_t^\lambda = \lambda^{1/H} d\Gamma_{\lambda^{1/H}t} = \lambda^{1/H} \sum_{\alpha=1}^{d} \tilde{W}_\alpha(\Gamma_t^\lambda) dB_{\lambda^{1/H}t}^\alpha.$$

Therefore, setting $B_t^{\alpha,\lambda} = B_{\lambda^{1/H}t}^\alpha$, we get

$$d\Gamma_t^\lambda = \sum_{\alpha=1}^{d} \tilde{W}_\alpha(\Gamma_t^\lambda) d(B_t^{\alpha,\lambda}) = \lambda \sum_{\alpha=1}^{d} \tilde{W}_\alpha(\Gamma_t^\lambda) d\left(\lambda^{-1} B_t^{\alpha,\lambda}\right)_t. \tag{5.12}$$

Now observe that the usual scaling for the fractional Brownian motion yields

$$\{\lambda^{-1} B_t^\lambda; t \geq 0\} \stackrel{d}{=} \{B_t; t \geq 0\}.$$

We have thus obtained that $\Gamma^\lambda \stackrel{d}{=} \hat{\Gamma}^\lambda$, where $\hat{\Gamma}^\lambda$ solves the system

$$d\hat{\Gamma}_t^\lambda = \lambda \sum_{\alpha=1}^{d} \tilde{W}_\alpha(\Gamma_t^\lambda) dB_t^\alpha. \tag{5.13}$$

Comparing (5.13) and (5.11), we conclude that $\delta_\lambda \Gamma_\cdot \overset{\text{law}}{=} \Gamma_{\lambda^{\frac{1}{H}}\cdot}$. If we define Q_t to be the law of U_t on $\mathfrak{g}^{(l)}$, it follows that $Q_s \circ \delta_\lambda^{-1} = Q_{\lambda^{\frac{1}{H}} s}$ for all $s > 0$. In particular, by setting $s = 1$ and $\lambda = t^H$, we obtain that

$$Q_t = Q_1 \circ \delta_{t^H}^{-1}. \tag{5.14}$$

Now suppose that $\rho_t(u)$ is the density of Q_t with respect to the Lebesgue measure du on $\mathfrak{g}^{(l)}$. Then for any $f \in C_b^\infty(\mathfrak{g}^{(l)})$, we have

$$\begin{aligned}\int_{\mathfrak{g}^{(l)}} f(u)\rho_t(u)du &= \int_{\mathfrak{g}^{(l)}} f(u)Q_t(du) = \int_{\mathfrak{g}^{(l)}} f(\delta_{t^H}u)Q_1(du) \\ &= \int_{\mathfrak{g}^{(l)}} f(\delta_{t^H}u)\rho_1(u)du = \int_{\mathfrak{g}^{(l)}} t^{-H\nu} f(u)\rho_1(\delta_{t^{-H}}u)du,\end{aligned}$$

where the equality follows from the change of variables $u \leftrightarrow \delta_{t^{-H}}u$ and the fact that $du \circ \delta_{t^H}^{-1} = t^{-H\nu}du$ (cf. relation (2.26)). Therefore, we conclude that

$$\rho_t(u) = t^{-H\nu}\rho_1(\delta_{t^H}^{-1}u), \quad \text{for all} \quad (u,t) \in \mathfrak{g}^{(l)} \times (0,1), \tag{5.15}$$

from which our result (5.9) follows. □

5.2 Step two: local lower estimate for the density of the Taylor approximation process.

Recall that according to our definition (2.37), $X_l(t,x) = x + F_l(U_t^{(l)}, x)$ is the Taylor approximation process of order l for the actual solution of the SDE (1.1). Due to hypoellipticity, it is natural to expect that when $l \geq l_0$, F_l is "non-degenerate" in certain sense. In addition, $X_l(t,x)$ should have a density, and a precise local lower estimate for the density should follow from Proposition 5.6 in Step One, combined with the "non-degeneracy" of F_l. Here the main subtlety and challenge lies in finding a way of respecting the fractional Brownian scaling and Cameron-Martin structure so that the estimate we obtain on $X_l(t,x)$ is sharp. In this part, we always fix $l \geq l_0$.

I. Non-degeneracy of F_l and a disintegration formula.

We first review a basic result in [17] on the (local) non-degeneracy of F_l, which then allows us to obtain a formula for the (localized) density of $X_l(t,x)$ by disintegration. This part is purely analytic and does not rely on the structure of the underlying process.

Let $JF_l(u,x) : \mathfrak{g}^{(l)} \to \mathbb{R}^N$ be the Jacobian of F_l with respect to u. Since $\mathfrak{g}^{(l)}$ has a canonical Hilbert structure induced from $T^{(l)}(\mathbb{R}^d)$, we can also consider the adjoint map $JF_l(u,x)^* : \mathbb{R}^N \to \mathfrak{g}^{(l)}$. The non-degeneracy of JF_l is summarized in the following lemma.

Lemma 5.7. *Let F_l be the approximation map given in Definition 2.17 and Let $JF_l(u,x) : \mathfrak{g}^{(l)} \to \mathbb{R}^N$ be its Jacobian. Then there exists a constant $c > 0$ depending only on l_0 and the vector fields, such that*

$$JF_l(0,x) \cdot JF_l(0,x)^* \geq c \cdot \mathrm{Id}_{\mathbb{R}^N}$$

for all $l \geq l_0$ and $x \in \mathbb{R}^N$.

Sketch of proof. This is Lemma 3.13 of [17]. Because of its importance, we outline the idea of the proof so that one may see how the hypoellipticity property comes into play. Recall the definitions of $\mathrm{e}_{(\alpha)}, \mathrm{e}_{[\alpha]}, V_{(\alpha)}$ from equation (2.25) and $V_{[\alpha]}$ from Section 1.2. Define a linear map $\Xi : T_0^{(l)} \to C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$ by setting $\Xi(\mathrm{e}_{(\alpha)}) \triangleq V_{(\alpha)}$ for each $\alpha \in \mathcal{A}_1(l)$. A crucial property is that Ξ respects Lie brackets, i.e.

$$\Xi(\mathrm{e}_{[\alpha]}) = V_{[\alpha]}, \quad \forall \alpha \in \mathcal{A}_1(l). \tag{5.16}$$

Now let $\{u_\mu : 1 \leq \mu \leq m_l\}$ be an orthonormal basis of $\mathfrak{g}^{(l)}$, where $m_l \triangleq \dim \mathfrak{g}^{(l)}$, and set $V_\mu = \Xi(u_\mu)$. Based on (5.16), it is not hard to show that

$$\mathrm{Span}\{V_{[\alpha]}(x) : \alpha \in \mathcal{A}_1(l)\} = \mathrm{Span}\{V_\mu(x) : 1 \leq \mu \leq m_l\},$$

for each $x \in \mathbb{R}^N$. Let us now relate these notions to the non-degeneracy of JF_l. To this aim, taking Definition 2.17 into account, it is easily seen that

$$JF_l(0,x)(u) = \sum_{\alpha \in \mathcal{A}_1(l)} V_{(\alpha)}(x) u^\alpha.$$

In particular we have $JF_l(0,x)(e_{(\alpha)}) = V_{(\alpha)}(x)$. Hence invoking relation (5.16) we end up with

$$JF_l(0,x)(e_{[\alpha]}) = V_{[\alpha]}(x).$$

By definition of our orthonormal basis $\{u_\mu; 1 \leq \mu \leq m_l\}$ we thus get

$$JF_l(0,x) \cdot JF_l(0,x)^* = \sum_{\mu=1}^{m_l} V_\mu(x) \otimes V_\mu(x).$$

Therefore, the non-degeneracy of $JF_l(0,x)$ follows from the hypoellipticity assumption (1.2) of the vector fields V_α. □

An immediate corollary of Lemma 5.7 is the following.

Corollary 5.8. *Given $l \geq l_0$, there exists $r > 0$ depending on l and the vector fields, such that $\det(JF_l(u,x) \cdot JF_l(u,x)^*)$ is uniformly positive on $\{u \in \mathfrak{g}^{(l)} : \|u\|_{\mathrm{HS}} < r\} \times \mathbb{R}^N$. In particular, the map*

$$\{u \in \mathfrak{g}^{(l)} : \|u\|_{\mathrm{HS}} < r\} \to \mathbb{R}^N, \qquad u \mapsto x + F_l(u,x),$$

is a submersion in the sense of differential geometry.

Remark 5.9. Note that the map F_l and the constant r in Corollary 5.8 depend on l. For technical reasons, we will assume that r is chosen (still depending on l) so that for all $l_0 \leq l' \leq l$, the map $JF_{l'}(\pi^{(l')}(u), x)$ has full rank whenever $(u,x) \in \mathfrak{g}^{(l)} \times \mathbb{R}^N$ with $\|u\|_{\mathrm{HS}} < r$, where $\pi^{(l')} : \mathfrak{g}^{(l)} \to \mathfrak{g}^{(l')}$ is the canonical projection. This property will be used in the proof of Lemma 5.21 in Step Three below.

Now let r be the constant given in Remark 5.9 . It is standard from differential geometry that for each $x \in \mathbb{R}^N$ and $y \in \{x + F_l(u,x) : \|u\|_{\mathrm{HS}} < r\}$, the "bridge space"

$$M_{x,y} \triangleq \{u \in \mathfrak{g}^{(l)} : \|u\|_{\mathrm{HS}} < r \text{ and } x + F_l(u,x) = y\} \tag{5.17}$$

is a submanifold of $\{u \in \mathfrak{g}^{(l)} : \|u\|_{\mathrm{HS}} < r\}$ with dimension $\dim \mathfrak{g}^{(l)} - N$. In addition, since both of $\mathfrak{g}^{(l)}$ and $\mathbb{R}^N$ are oriented Riemannian manifolds, we know from differential topology that $M_{x,y}$ carries a natural orientation and hence a volume form which we denote as $m_{x,y}$. The following result is the standard disintegration formula in Riemannian geometry (cf. Appendix for a proof).

Proposition 5.10. *For any $\varphi \in C_c^\infty(\{u \in \mathfrak{g}^{(l)} : \|u\|_{\mathrm{HS}} < r\})$, we have*

$$\int_{\mathfrak{g}^{(l)}} \varphi(u)du = \int_{\mathbb{R}^N} dy \int_{M_{x,y}} K(u,x)\varphi(u)m_{x,y}(du), \tag{5.18}$$

where the kernel K is given by

$$K(u,x) \triangleq (\det(JF_l(u,x) \cdot JF_l(u,x)^*))^{-\frac{1}{2}}, \tag{5.19}$$

and we define $m_{x,y} \triangleq 0$ if $M_{x,y} = \emptyset$.

The disintegration formula (5.18) immediately leads to a formula for the (localized) density of the Taylor approximation process $X_l(t,x)$. We summarize this fact in the folloiwng proposition.

Proposition 5.11. *Let $\eta \in C_c^\infty(\{u \in \mathfrak{g}^{(l)} : \|u\|_{\mathrm{HS}} < r\})$ be a bump function so that $0 \le \eta \le 1$ and $\eta = 1$ when $\|u\|_{\mathrm{HS}} < r/2$, where r is the constant featuring in Proposition 5.10. Define $\mathbb{P}_l^\eta(t,x,\cdot)$ to be the measure*

$$\mathbb{P}_l^\eta(t,x,A) \triangleq \mathbb{E}\left[\eta(U_t)\mathbf{1}_{\{X_l(t,x)\in A\}}\right], \quad A \in \mathcal{B}(\mathbb{R}^N),$$

where $U_t = \log \Gamma_t$, Γ_t is defined by (5.4) and $X_l(t,x) = x + F_l(U_t,x)$ is the approximation given by (2.37). The measure $\mathbb{P}_l^\eta(t,x,\cdot)$ is absolutely continuous with respect to the Lebesgue measure, and its density is given by

$$p_l^\eta(t,x,y) \triangleq \int_{M_{x,y}} \eta(u)K(u,x)\rho_t(u)m_{x,y}(du). \tag{5.20}$$

where ρ_t is the density of U_t alluded to in Lemma 5.2 and K is given by (5.19).

Proof. Since $X_l((t,x) = F_l(U_t,x)$, we have

$$\mathbb{P}_l^\eta(t,x,A) = \mathbb{E}\left[\eta(U_t)\mathbf{1}_{\{F_l(U_t,x)\in A\}}\right],$$

and one can recast this expression in terms of the density of U_t, namely

$$\mathbb{P}_l^\eta(t,x,A) = \int_{\mathfrak{g}^{(l)}} \eta(u)\mathbf{1}_{\{F_l(u,x)\in A\}}\rho_t(u)du.$$

Then our conclusion (5.20) stems from a direct application of (5.18). □

II. Estimating the volume form $m_{x,y}$.

To obtain a sharp lower estimate on $p_l^\eta(t,x,y)$ from formula (5.20) and the lower estimate of $\rho_t(u)$ given by (5.9), one needs to estimate the volume form $m_{x,y}$ precisely. For this purpose, we apply a change of variables $S_{x,y} : M_{x,y} \to M_{x,x}$ introduced in [17]. In contrast to the Brownian motion case, our main challenge lies in respecting the Cameron-Martin structure in order to obtain sharp estimates. In this section, we will use the technique implemented in the proof of Theorem 1.3 and pathwise estimates for Cameron-Martin paths to achieve this.

The construction of the function $S_{x,y}$ alluded to above is based on the simple idea that in order to transform a loop α from x to x into a path from x to y, we just concatenate a generic path from x to y to the loop α. However, we are looking at this construction from the Taylor approximation point of view. More specifically, we define the operation $\times$ to be the multiplication induced from $G^{(l)}$ through the exponential map, namely

$$v \times u \triangleq \log(\exp(v) \otimes \exp(u)), \quad v, u \in \mathfrak{g}^{(l)}. \tag{5.21}$$

As mentioned above, we would ideally like the operation $\times$ to transform elements $v \in M_{x,x}$ into elements $v \times u \in M_{x,y}$ where $u \in M_{x,y}$ is fixed. However, due to the fact that F_l is only an approximation of the flow Φ, this property will in general not be fulfilled. Nevertheless, if $\|u\|_{\mathrm{HS}}$ and $\|v\|_{\mathrm{HS}}$ are small enough the product $v \times u$ is close to an element of $M_{x,y}$, so that the function Ψ_l of Lemma 4.1 can be applied. We summarize those heuristic considerations in the following lemma (cf. [17, Lemma 3.23]), which gives the precise construction of the change of variables $S_{x,y} : M_{x,y} \to M_{x,x}$.

Lemma 5.12. *For $x \in \mathbb{R}^N$ and $h \in \bar{\mathcal{H}}$, we set $y = \Phi_1(x;h)$ where recall that Φ is the flow defined by (1.4). Let r be the constant arising in Lemma 4.1. For $u \in M_{x,y}$ and $v \in M_{x,x}$, recall that $v \times u$ is defined by (5.21). Then the following holds true:*

(i) There exist $\varepsilon, \rho_1 > 0$ and $\rho_2 \in (0,r)$, such that for any given $x \in \mathbb{R}^N$ and $h \in \bar{\mathcal{H}}$ with $\|h\|_{\bar{\mathcal{H}}} < \rho_1$, the map

$$\tilde{\Psi}_{x,h}(v) \triangleq \Psi_l\left(v \times u, x, y - x - F_l(v \times u, x) + F_l(v,x)\right)$$

defines a diffeomorphism from an open neighbourhood $V_{x,h} \subset \mathfrak{g}^{(l)}$ of 0 containing the ball $\{v \in \mathfrak{g}^{(l)} : \|v\|_{\mathrm{HS}} < \varepsilon\}$ onto $W \triangleq \{w \in \mathfrak{g}^{(l)} : \|w\|_{\mathrm{HS}} < \rho_2\}$, where recall that the function Ψ_l is defined in Lemma 4.1. In addition,

$$v \in V_{x,h} \cap M_{x,x} \qquad \text{iff} \qquad w \triangleq \tilde{\Psi}_{x,h}(v) \in W \cap M_{x,y}.$$

(ii) Recall that $d(x,y)$ is the control distance function associated with the SDE (1.1) and that the set $\Pi_{x,y}$ is defined by (1.5). Given $x, y \in \mathbb{R}^N$ with $d(x,y) < \rho_1/2$, choose $h \in \Pi_{x,y}$ satisfying

$$d(x,y) \le \|h\|_{\bar{\mathcal{H}}} \le 2d(x,y) < \rho_1, \tag{5.22}$$

and define

$$S_{x,y} \triangleq \tilde{\Psi}_{x,h}^{-1}\Big|_{W \cap M_{x,y}} : W \cap M_{x,y} \to V_{x,h} \cap M_{x,x}. \tag{5.23}$$

Then there exist a constant $C > 0$, such that

$$\frac{1}{C} \cdot m_{x,x}(\cdot) \leq m_{x,y} \circ S_{x,y}^{-1}(\cdot) \leq C \cdot m_{x,x}(\cdot) \tag{5.24}$$

on $V_{x,h} \cap M_{x,x}$.

The previous lemma sets the stage for a useful change of variable in (5.20). Our next step is to provide some useful bounds for the integral in (5.20). We first need the following crucial lemma.

Lemma 5.13. *There exist constants $C, \kappa > 0$ such that for any $u \in \mathfrak{g}^{(l)}$ with $\|u\|_{\mathrm{HS}} < \kappa$, we have*

$$d(x, x + F_l(u, x)) \leq C\|u\|_{\mathrm{CC}}. \tag{5.25}$$

Proof. We only consider the case when $H > 1/2$, as the other case follows from Lemma 2.8 and the result for the diffusion case proved in [17]. When $H > 1/2$, the argument is an adaptation of the proof of Theorem 1.3. We use the same notation as in that proof, except for the fact that l_0 is replaced by a general $l \geq l_0$ exclusively.

We set up an inductive procedure as in the proof of Theorem 1.3. Namely, denote $u_1 \triangleq u$, $x_1 \triangleq x$, and $y \triangleq x + F_l(u, x)$. Choose $\kappa_1 > 0$ so that

$$\|u\|_{\mathrm{CC}} < \kappa_1 \implies |y - x| < \delta, \tag{5.26}$$

where δ is again the constant arising in the proof of Theorem 1.3. By constructing successively elements $u_m \in \mathfrak{g}^{(l)}$ and intervals I_m, we obtain exactly as in (4.26) that

$$|I_m| = \|u_m\|_{\mathrm{CC}} \leq C_{V,l}\|u_{m-1}\|^{1+\frac{1}{l}} = C_{V,l}|I_{m-1}|^{1+\frac{1}{l}}. \tag{5.27}$$

In addition, along the same lines as (4.54) and (4.55), we have

$$d(x, x + F_l(u, x))^2 \leq C_{H,l} \lim_{m\to\infty} \left(\sum_{k=1}^{m} |I_k|\right)^{2H} \left(\sum_{k=1}^{m} k^{2H+1}|I_k|^{2(1-H)}\right). \tag{5.28}$$

We will now bound the right hand side of (5.28).

Let us set $\alpha \triangleq 1 + 1/l$. By iterating (5.27), we obtain that

$$|I_m| \leq (C_{V,l}|I_1|)^{\alpha^{m-1}}, \quad \forall m \geq 1.$$

Therefore, we can bound the two terms on the right hand-side of (5.28) as follows:

$$\sum_{k=1}^{m} |I_k| \leq C_{V,l}|I_1| \cdot \left(\sum_{k=1}^{m} (C_{V,l}|I_1|)^{\alpha^{k-1}-1}\right) \tag{5.29}$$

and

$$\sum_{k=1}^{m} k^{2H+1}|I_k|^{2(1-H)} \leq (C_{V,l}|I_1|)^{2(1-H)} \cdot \left(\sum_{k=1}^{m} k^{2H+1} (C_{V,l}|I_1|)^{2(1-H)(\alpha^{k-1}-1)} \right). \tag{5.30}$$

Let us now bound the term $|I_1|$ in (5.29) and (5.30). To this aim, recall that we have chosen $u_1 = u$. Therefore one can choose $\kappa_2 > 0$ so that

$$\|u\|_{\mathrm{CC}} < \kappa_2 \implies C_{V,l}|I_1| = C_{V,l}\|u_1\|_{\mathrm{CC}} \leq \frac{1}{2}. \tag{5.31}$$

We will assume that both (5.26) and (5.31) are satisfied in the sequel, under the condition that $\|u\|_{\mathrm{CC}} < \kappa$ with $\kappa = \kappa_1 \wedge \kappa_2$. In order to bound (5.29) and (5.30) above, also observe that the series

$$\sum_{m=1}^{\infty} \left(\frac{1}{2}\right)^{\alpha^{m-1}-1}, \quad \text{and} \quad \sum_{m=1}^{\infty} m^{2H+1} \left(\frac{1}{2}\right)^{2(1-H)(\alpha^{m-1}-1)} \tag{5.32}$$

are convergent. Therefore plugging (5.32) and (5.31), and then (5.29) and (5.30) into (5.28), we get that for $\|u\|_{\mathrm{CC}} < \kappa$ we have

$$d(x, x + F_l(u,x))^2 \leq C_{H,V,l}|I_1|^{2H} \cdot |I_1|^{2(1-H)} = C_{H,V,l}|I_1|^2 = C_{H,V,l}\|u\|_{\mathrm{CC}}^2.$$

Therefore, our result (5.25) follows. □

Lemma 5.13 yields the following two-sided estimate on the transformation $S_{x,y}$.

Lemma 5.14. *Keeping the same notation as in Lemma 5.12, let $x, y \in \mathbb{R}^N$ be such that $d(x,y) < \rho_1/2$ and let $h \in \Pi_{x,y}$ fulfilling condition (5.22). Then the function $S_{x,y}$ defined by (5.23) satisfies the following inequality for all $v \in V_{x,h} \cap M_{x,x}$,*

$$\frac{1}{\Lambda} \cdot (\|v\|_{\mathrm{CC}} + d(x,y)) \leq \|S_{x,y}^{-1}(v)\|_{\mathrm{CC}} \leq \Lambda \cdot (\|v\|_{\mathrm{CC}} + d(x,y)) \tag{5.33}$$

for some constant $\Lambda > 0$.

Proof. We prove the two inequalities in our claim separately.

Step 1: proof of the upper bound in (5.33). To this aim, we set $w \triangleq S_{x,y}^{-1}(v)$ and $u \triangleq \log S_l(h)$. Note that according to the definition of $S_{x,y}$ given in Lemma 5.12 we have $w = \tilde{\Psi}_{x,h}(v)$. We wish to upper bound $\|w\|_{\mathrm{CC}}$ in order to get the second part of (5.33).

According to Lemma 5.12, $\|w\|_{\mathrm{CC}}$ is close to $\|v \times u\|_{\mathrm{CC}}$. Specifically, Proposition 2.15 entails

$$|\|w\|_{\mathrm{CC}} - \|v \times u\|_{\mathrm{CC}}| \leq \rho_{\mathrm{CC}}(w, v \times u) \leq C_1 \|w - v \times u\|_{\mathrm{HS}}^{\frac{1}{l}}.$$

Let $\gamma \in C^{1-\mathrm{var}}([0,1];\mathbb{R}^d)$ be such that $v = \log(S_l(\gamma))$ and $\|v\|_{\mathrm{CC}} = \|\gamma\|_{1-\mathrm{var}}$. By the definition of $\tilde{\Psi}_{x,h}$, the fact that $w = \tilde{\Psi}_{x,h}(v)$ and Lemma 4.1 (ii), we have

$$\begin{aligned} &\|w - v \times u\|_{\mathrm{HS}} \\ &\leq A|y - x - F_l(v \times u, x)| \\ &\leq A\left(|y - \Phi_1(x; \gamma \sqcup h)| + |\Phi_1(x; \gamma \sqcup h) - x - F_l(v \times u, x)|\right) \triangleq A(Q_1 + Q_2). \end{aligned} \tag{5.34}$$

We will now bound the terms Q_1 and Q_2 respectively.

In order to estimate the term Q_1 in (5.34), let us recall that $y = \Phi_1(x;h)$. Therefore using the flow property of Φ we get

$$Q_1 = |\Phi_1(x;h) - \Phi_1(\Phi_1(x,\gamma);h)|.$$

We can now use some standard estimate on the flow of differential equations (cf. [13, Theorem 10.26]) plus the fact that $\|h\|_{\bar{\mathcal{H}}}$ is bounded by ρ_1 (cf. condition (5.22)) in order to get

$$Q_1 \leq C_2|\Phi_1(x;\gamma) - x|.$$

In addition, v is assumed to be an element of $M_{x,x}$. According to (5.17), this means in particular that $F_l(v,x) = 0$. Hence we obtain

$$Q_1 \leq C_2|\Phi_1(x;\gamma) - x - F_l(v,x)|.$$

Thanks to the Euler estimate of [13, Corollary 10.15], and the fact that $v = \log(S_l(\gamma))$, we end up with

$$Q_1 \leq C_2 C_{V,l}\|\gamma\|_{1-\mathrm{var}}^{\bar{l}} = C_2 C_{V,l}\|v\|_{\mathrm{CC}}^{\bar{l}}, \tag{5.35}$$

where $\bar{l}$ is any given number in $(l, l+1)$.

In order to handle the term Q_2 in (5.34), let us observe that whenever $H > 1/4$, Proposition 2.6 entails that $h \in C^{q-\mathrm{var}}$ for $q \in [1,2)$, for all $h \in \bar{\mathcal{H}}$. Moreover, the following inequality holds true,

$$\|h\|_{q-\mathrm{var}} \leq C_H\|h\|_{\bar{\mathcal{H}}}. \tag{5.36}$$

Therefore, since $v \times u = \log(S_l(\gamma \sqcup h))$, the rough path estimates in [13, Corollary 10.15] yield

$$Q_2 \leq C_{V,l}\|\gamma \sqcup h\|_{q-\mathrm{var}}^{\bar{l}} \leq C_{V,l}2^{\bar{l}-\frac{\bar{l}}{q}}\left(\|\gamma\|_{1-\mathrm{var}} + \|h\|_{q-\mathrm{var}}\right)^{\bar{l}},$$

where the last inequality stems from the simple relation

$$\|\gamma \sqcup h\|_{q-\mathrm{var}} \leq 2^{1-\frac{1}{q}}\left(\|\gamma\|_{q-\mathrm{var}} + \|h\|_{q-\mathrm{var}}\right), \quad \text{and} \quad \|\gamma\|_{q-\mathrm{var}} \leq \|\gamma\|_{1-\mathrm{var}}.$$

Furthermore, since we have chosen v such that $v = \log(S_l(\gamma))$ and $\|v\|_{\mathrm{CC}} = \|\gamma\|_{1-\mathrm{var}}$, we get

$$Q_2 \leq C_{H,V,l}(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}})^{\bar{l}}, \tag{5.37}$$

where we have also invoked (5.36) in order to upper bound $\|h\|_{q-\mathrm{var}}$.

Summarizing our considerations so far, if we plug our estimate (5.35) on Q_1 and our bound (5.37) on Q_2 into relation (5.34), we obtain the following inequality

$$|\|w\|_{\mathrm{CC}} - \|v \times u\|_{\mathrm{CC}}| \leq C_3\left(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}}\right)^{1+\frac{\bar{l}-l}{l}}. \tag{5.38}$$

Next, we claim that $\|u\|_{\mathrm{CC}} \leq C_{H,l}\|h\|_{\bar{\mathcal{H}}}$. Indeed, recall that $u = \log(S_l(h))$ and set $S_l(h) = g$. This means that $g = \exp(u)$. Since $h \in C^{q-\mathrm{var}}$ with $q \in [1,2)$, Lyons' extension theorem (cf. [19, Theorem 2.2.1]) implies that for all $i = 1, ..., l$ we have

$$\|g_i\|_{\mathrm{HS}} \leq C_{H,l}\|h\|^i_{q-\mathrm{var}},$$

where g_i is the i-th component of g. If we define the homogeneous norm $|||\cdot|||$ on $G^{(l)}$ by

$$|||\xi||| \triangleq \max_{1\leq i\leq l} \|\xi_i\|_{\mathrm{HS}}^{\frac{1}{i}}, \quad \xi \in G^{(l)},$$

we get the following estimate:

$$|||g||| \leq C_{H,l}\|h\|_{q-\mathrm{var}}, \quad \text{and} \quad \|u\|_{\mathrm{CC}} \leq C_{H,l}\, |||\, g|||, \tag{5.39}$$

where the second inequality stems from the equivalence of homogeneous norms in $G^{(l)}$ (cf. [13, Theorem 7.44]). Now combining the two inequalities in (5.39) and relation (5.36), we end up with

$$\|u\|_{\mathrm{CC}} \leq C_{H,l}\|h\|_{q-\mathrm{var}} \leq C_{H,l}\|h\|_{\bar{\mathcal{H}}}. \tag{5.40}$$

Let us now go back to (5.38), from which we easily deduce

$$\begin{aligned}\|w\|_{\mathrm{CC}} &\leq \|v \times u\|_{\mathrm{CC}} + C_3\left(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}}\right)^{1+\frac{\bar{l}-l}{l}} \\ &\leq \|v\|_{\mathrm{CC}} + \|u\|_{\mathrm{CC}} + C_3\left(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}}\right)^{1+\frac{\bar{l}-l}{l}}\end{aligned}$$

Plugging (5.40) into this inequality and resorting to the fact that $v \in V_{x,h}$ and h satisfies (5.22), we end up with

$$\begin{aligned}\|w\|_{\mathrm{CC}} &\leq \|v\|_{\mathrm{CC}} + C_{H,l}\|h\|_{\bar{\mathcal{H}}} + C_3\left(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}}\right)^{1+\frac{\bar{l}-l}{l}} \\ &\leq C_4\left(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}}\right) \\ &\leq C_4\left(\|v\|_{\mathrm{CC}} + 2d(x,y)\right).\end{aligned}$$

Recalling that we have set $w = S_{xy}^{-1}(v)$, this proves the upper bound in (5.33).

Step 2: proof of the lower bound in (5.33). We start from inequality (5.38), which yields

$$\begin{aligned}\|w\|_{\mathrm{CC}} &\geq \|v \times u\|_{\mathrm{CC}} - C_3\left(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}}\right)^{1+\frac{\bar{l}-l}{l}} \\ &\geq \|v\|_{\mathrm{CC}} - \|u\|_{\mathrm{CC}} - C_3\left(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}}\right)^{1+\frac{\bar{l}-l}{l}}\end{aligned}$$

Furthermore, thanks to (5.40) we obtain

$$\|w\|_{\mathrm{CC}} \geq \|v\|_{\mathrm{CC}} - C_{H,l}\|h\|_{\bar{\mathcal{H}}} - C_3\left(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}}\right)^{1+\frac{\bar{l}-l}{l}}. \tag{5.41}$$

We now invoke the fact that $\|v\|_{\mathrm{HS}} < r$ and $\|h\|_{\bar{\mathcal{H}}} < \rho_1$, thus by possibly shrinking r and ρ_1, we may assume that

$$C_3\left(\|v\|_{\mathrm{CC}} + \|h\|_{\bar{\mathcal{H}}}\right)^{\frac{\bar{l}-l}{l}} < \frac{1}{2}.$$

Putting this information into (5.41), we thus obtain

$$\|w\|_{\mathrm{CC}} \geq \frac{1}{2}\|v\|_{\mathrm{CC}} - \left(\frac{1}{2} + C_{H,l}\right)\|h\|_{\bar{\mathcal{H}}} \geq \frac{1}{2}\|v\|_{\mathrm{CC}} - (1 + 2C_{H,l})d(x,y). \tag{5.42}$$

On the other hand, according to Lemma 5.13 (one may further shrink r so that $\|w\|_{\mathrm{CC}} < \kappa$, where κ is the constant featuring in Lemma 5.13), we have

$$d(x,y) = d(x, x + F_l(w,x)) \leq C\|w\|_{\mathrm{CC}}. \tag{5.43}$$

Putting together inequalities (5.42) and (5.43), we easily get the lower bound in (5.33), which finishes our proof. □

Before proceeding to analyze the density $p_l^\eta(t,x,y)$, with all the preparations above we take a short detour to prove the local equivalence of controlling distances claimed in Theorem 1.4.

Proof of Theorem 1.4. For each fixed $x \in \mathbb{R}^N$, define a function on $\mathbb{R}^N$ by

$$g(x,y) = \inf\{\|u\|_{\mathrm{CC}} : u \in \mathfrak{g}^{(l)} \text{ and } x + F_l(u,x) = y\}.$$

Observe that $g(x,y)$ is an intrinsic quantity that does not depend on H. In order to prove Theorem 1.4, it suffices to show that $d(x,y)$ is equivalent to $g(x,y)$ in a Euclidean neighborhood of x.

To this aim, first note that by Corollary 5.8 the map $G(\cdot) : u \mapsto x + F_l(u,x)$ is a submersion in a neighborhood of $0 \in \mathfrak{g}^{(l)}$. Recall that the set $W \cap M_{x,y}$ is introduced in Lemma 5.12. We can thus choose a small enough $\delta > 0$ such that $\{G^{-1}(y) : |x-y| \leq \delta\} \subset W \cap M_{x,y}$ and that both Lemma 5.13 and Lemma 5.14 can be applied.

Now we fix such a choice of δ. For any $y \in \mathbb{R}^N$ with $|x-y| \leq \delta$, we can first apply Lemma 5.13 to conclude that

$$d(x,y) \leq Cg(x,y).$$

Next, we use the second inequality in Lemma 5.14 for $v = 0$ to conclude that

$$g(x,y) \leq \Lambda d(x,y).$$

The proof is thus completed. □

Now we come back to the main goal of this part. Namely starting from Proposition 5.11, we will apply a change of variables involving $S_{x,y}$ and express the density p_l^η in terms of the measure $m_{x,x}$ not depending on y.

Lemma 5.15. *Let $p_l^\eta(t,x,y)$ be the density defined by (5.20), and recall that the exponent ν is defined by $\nu \triangleq \sum_{k=1}^l k \dim \mathcal{L}_k$. Then there exist constants $C, \tau > 0$, such that for all x, y, t with $d(x,y) \leq t^H$ and $0 < t < \tau$, we have*

$$p_l^\eta(t,x,y) \geq Ct^{-H\nu} m_{xx}\left(\{v \in M_{x,x} : \|v\|_{\mathrm{CC}} \leq t^H\}\right), \tag{5.44}$$

where $m_{x,x,}$ is the volume form on $M_{x,x}$ given by (5.17)

Proof. Lemma 5.12 asserts that there exists $\rho_1 > 0$, such that if $d(x,y) < \rho_1/2$, then $S_{x,y}$ given by (5.23) defines a change of variables (i.e. a diffeomorphism) for (5.20). Specifically we have

$$\begin{aligned} p_l^\eta(t,x,y) &\geq \int_{M_{x,y}\cap W} \eta(u)K(u,x)\rho_t(u)m_{x,y}(du) \\ &= \int_{M_{x,x}\cap V_{x,h}} \eta(S_{x,y}^{-1}v)K(S_{x,y}^{-1}v,x)\rho_t(S_{x,y}^{-1}v)\, m_{x,y}\circ S_{x,y}^{-1}(dv). \end{aligned}$$

In addition, since $V_{x,h}$ contains the ball $\{v \in \mathfrak{g}^{(l)} : \|v\|_{\mathrm{HS}} < \varepsilon\}$, owing to relation (5.24) and thanks to the fact that K defined by (5.19) is bounded below, we obtain

$$p_l^\eta(t,x,y) \geq C_{H,V,l} \int_{M_{x,x}\cap\{v\in\mathfrak{g}^{(l)}:\|v\|_{\mathrm{HS}}<\varepsilon\}} \rho_t(S_{x,y}^{-1}v)\, m_{x,x}(dv).$$

Now choose $\tau < (\rho_1/2)^{\frac{1}{H}}$ to be such that

$$0 < t < \tau \implies \{v \in \mathfrak{g}^{(l)} : \|v\|_{\mathrm{CC}} \leq t^H\} \subseteq \left\{v \in \mathfrak{g}^{(l)} : \|v\|_{\mathrm{HS}} < \varepsilon\right\}.$$

We will thus lower bound $p_l^\eta(t,x,y)$ as follows

$$p_l^\eta(t,x,y) \geq C_{H,V,l} \int_{M_{x,x}\cap\{v\in\mathfrak{g}^{(l)}:\|v\|_{\mathrm{CC}}<t^H\}} \rho_t(S_{x,y}^{-1}v)m_{x,x}(dv). \tag{5.45}$$

Next, according to the second inequality of (5.33), if $d(x,y) \leq t^H$ and $t < \tau$ (so that $d(x,y) < \rho_1/2$), then

$$\|S_{x,y}^{-1}v\|_{\mathrm{CC}} \leq 2Ct^H,$$

provided that $v \in M_{x,x}$ with $\|v\|_{\mathrm{CC}} \leq t^H$. For such x,y,t,v, by Proposition 5.6 we have

$$\rho_t(S_{x,y}^{-1}v) \geq \beta_{2C}t^{-H\nu}.$$

Plugging this inequality into (5.45), we arrive at

$$p_l^\eta(t,x,y) \geq C_{H,V,l}\beta_{2C}t^{-H\nu}m_{x,x}\left(\{v \in M_{x,x} : \|v\|_{\mathrm{CC}} \leq t^H\}\right),$$

which is our claim (5.44).

□

The next result tells us that the right hand side of (5.44) is comparable with the inverse volume of $B_d(x,t^H)$. This seems to be surprising as the first quantity does not capture the Gaussian structure at all while the second quantity relies crucially on the Cameron-Martin structure. The key reason behind this lies in the precise two-sided estimate (5.33) of $S_{x,y}$ in terms of $d(x,y)$, which is also the key point leading to the local equivalence of all the control distance functions as we just proved.

Lemma 5.16. *Let $M_{x,x}$ be the set defined by (5.17) and recall that $m_{x,x}$ is the volume measure on $M_{x,x}$. There exist constants $C, \tau > 0$, such that*

$$\frac{1}{C|B_d(x,t^H)|} \leq t^{-H\nu} m_{x,x}\left(\{v \in M_{x,x} : \|v\|_{\mathrm{CC}} \leq t^H\}\right) \leq \frac{C}{|B_d(x,t^H)|} \tag{5.46}$$

for all $x \in \mathbb{R}^N$ and $0 < t < \tau$.

Proof. The upper and lower bounds in (5.46) follow the same pattern, therefore we focus on the proof of the lower bound. To this aim, let $\psi \in C_c^\infty((-1,1);\mathbb{R}^1)$ be such that $0 \leq \psi \leq 1$ and $\psi(\xi) = 1$ when $|\xi| \leq 1/2$. We further localize the measure P_l^η defined in Proposition 5.11 by considering the following measure

$$P_l^{\eta,\psi}(t,x,\Gamma) \triangleq \mathbb{E}\left[\eta(U_t)\psi\left(\frac{\Lambda\|U_t\|_{\mathrm{CC}}}{t^H}\right)\mathbf{1}_{\{X_l(t,x)\in\Gamma\}}\right], \quad \Gamma \in \mathcal{B}(\mathbb{R}^N),$$

where Λ is the constant appearing in Lemma 5.14. Along the same lines as for the disintegration formula (5.18), the measure $P_l^{\eta,\psi}(t,x,\cdot)$ has a density given by

$$\begin{aligned} p_l^{\eta,\psi}(t,x,y) &= \int_{M_{x,y}} \eta(u)\psi\left(\frac{\Lambda\|u\|_{\mathrm{CC}}}{t^H}\right) K(u,x)\rho_t(u) m_{x,y}(du) \\ &= \int_{M_{x,y}\cap W} \eta(u)\psi\left(\frac{\Lambda\|u\|_{\mathrm{CC}}}{t^H}\right) K(u,x)\rho_t(u) m_{x,y}(du), \end{aligned}$$

where the kernel K is given by (5.19) and provided that τ further satisfies

$$0 < t < \tau \implies \left\{u \in \mathfrak{g}^{(l)} : \Lambda\|u\|_{\mathrm{CC}} \leq t^H\right\} \subseteq W.$$

As in the proof of Lemma 5.15, we now apply the change of variables $S_{x,y}u = v$ and the fact that $\psi(\xi = 0)$ if $|\xi| \geq 1$ in order to get

$$\begin{aligned} p_l^{\eta,\psi}(t,x,y) = \int_{V_{x,h}\cap\left\{v\in M_{x,x}:\Lambda\|S_{x,y}^{-1}v\|_{\mathrm{CC}}\leq t^H\right\}} &\eta(S_{x,y}^{-1}v)\,\psi\left(\frac{\Lambda\|S_{x,y}^{-1}v\|_{\mathrm{CC}}}{t^H}\right) \\ &\times K(S_{x,y}^{-1}v,x)\,\rho_t(S_{x,y}^{-1}v)\,m_{x,y}\circ S_{x,y}^{-1}(dv). \end{aligned}$$

Furthermore, due to the fact that ψ is supported on $(-1,1)$, K is bounded owing to (5.19) and according to the upper bound in (5.24), we get

$$p_l^{\eta,\psi}(t,x,y) \leq C_{H,V,l}\int_{V_{x,h}\cap\left\{v\in M_{x,x}:\Lambda\|S_{x,y}^{-1}v\|_{\mathrm{CC}}\leq t^H\right\}} \rho_t(S_{x,y}^{-1}v)\,m_{x,x}(dv).$$

Therefore identity (5.15) yields

$$p_l^{\eta,\psi}(t,x,y) \leq C_{H,V,l}t^{-H\nu} m_{x,x}\left(\{v \in M_{x,x} : \Lambda\|S_{x,y}^{-1}v\|_{\mathrm{CC}} \leq t^H\}\right).$$

Finally the lower bound on $\|S_{x,y}^{-1}(v)\|_{\mathrm{CC}}$ in (5.33) implies that whenever $d(x,y) \le t^H$ and $0 < t < \tau$ we have

$$p_l^{\eta,\psi}(t,x,y) \le C_{H,V,l} t^{-H\nu} m_{x,x}\left(\{v \in M_{x,x} : \|v\|_{\mathrm{CC}} \le t^H\}\right). \tag{5.47}$$

We now lower bound the density $p_l^{\eta,\psi}$ on the ball $B_d(x,t^H)$. To this aim we first write

$$\begin{aligned}
&\int_{B_d(x,t^H)} p_l^{\eta,\psi}(t,x,y)dy \\
&= \mathbb{E}\left[\eta(U_t)\psi\left(\frac{\Lambda\|U_t\|_{\mathrm{CC}}}{t^H}\right)\mathbf{1}_{\{d(x,x+F(U_t,x))<t^H\}}\right] \\
&\ge \mathbb{P}\left(d(x,x+F(U_t,x)) < t^H,\ \|U_t\|_{\mathrm{CC}} \le \frac{t^H}{2\Lambda},\ \|U_t\|_{\mathrm{HS}} < \kappa \wedge \frac{r}{2}\right),
\end{aligned} \tag{5.48}$$

where the second inequality stems from the fact that in Proposition 5.11 we have assumed that $\eta = 1$ where $\|u\|_{\mathrm{HS}} < r/2$ and we also have $\psi(r) = 1$ if $|r| \le 1/2$. Next we resort to Lemma 5.13, which can be rephrased as follows: there exist $C, \kappa, \tau > 0$ such that if $t < \tau$ and $\|u\|_{\mathrm{CC}} < \gamma t^H$ with $\gamma \triangleq (\max\{C, 2\Lambda\})^{-1}$, then we have

$$\|u\|_{\mathrm{CC}} \le \gamma t^H \implies d(x,x+F_l(u,x)) < t^H,\ \|u\|_{\mathrm{CC}} \le \frac{t^H}{2\Lambda},\ \|u\|_{\mathrm{HS}} < \kappa \wedge \frac{r}{2}. \tag{5.49}$$

Plugging (5.49) into (5.48), we thus get that for $t < \tau$ we have

$$\int_{B_d(x,t^H)} p_l^{\eta,\psi}(t,x,y)dy \ge \mathbb{P}\left(\|U_t\|_{\mathrm{CC}} \le \gamma t^H\right) = \mathbb{P}\left(\|\delta_{t^H}^{-1} U_t\|_{\mathrm{CC}} \le \gamma\right).$$

Eventually, owing to the scaling property of U_t alluded to in (5.14), we end up with

$$\int_{B_d(x,t^H)} p_l^{\eta,\psi}(t,x,y)dy \ge \int_{\{u \in \mathfrak{g}^{(l)} : \|u\|_{\mathrm{CC}} \le \gamma\}} \rho_1(u)du \triangleq C_{\gamma,l}. \tag{5.50}$$

Now the lower bound in (5.46) follows from integrating both sides of (5.47) over $B_d(x,t^H)$ and (5.50). □

Summarizing the content of Lemma 5.15 and Lemma 5.16, we have obtained the following lower bound on $p_l^{\eta}(t,x,y)$, which finishes the second step of the main strategy.

Corollary 5.17. *Let $p_l^{\eta}(t,x,y)$ be the density given by (5.20), and recall the notations of Lemma 5.15. Then there exist constants $C, \tau > 0$ depending only on H, l and the vector fields, such that*

$$p_l^{\eta}(t,x,y) \ge \frac{C}{|B_d(x,t^H)|} \tag{5.51}$$

for all x, y, t satisfying $d(x,y) \le t^H$ and $0 < t < \tau$.

5.3 Step three: comparing approximating and actual densities.

The last step towards the proof of Theorem 1.7 will be to show that the approximating density $p_l^\eta(t,x,y)$ and the actual density $p(t,x,y)$ of X_t^x are close to each other when t is small. For this part, we combine the Fourier transform approach developed in [17] with general estimates for Gaussian rough differential equations. As we will see, there is a quite subtle point related to the uniformity in l (the degree of approximation) when obtaining upper bound of $p_l^\eta(t,x,y)$ which is the main challenge for this part. In our modest opinion, we believe that there is a gap in the argument in [17] for the diffusion case, and we therefore propose an alternative proof in the fractional Brownian setting which also covers the diffusion result. As before, we assume that $l \geq l_0$.

Recall that the Fourier transform of a function $f(y)$ on $\mathbb{R}^N$ is defined by

$$\mathcal{F}f(\xi) = \hat{f}(\xi) \triangleq \int_{\mathbb{R}^N} f(y) \mathrm{e}^{2\pi i \langle \xi, y\rangle}\, dy, \quad \xi \in \mathbb{R}^N,$$

where we highlight the fact that $\mathcal{F}f$ and $\hat{f}$ are used indistinctly to designate our Fourier transform. In the sequel we will consider the Fourier transform $\hat{p}(t,x,\xi)$ (respectively, $\hat{p}_l^\eta(t,x,\xi)$) of the density $p(t,x,y)$ (respectively, $p_l^\eta(t,x,y)$) with respect to the y-variable. We will invoke the following trivial bound on $p - p_l$ in terms of $\hat{p}$ and $\hat{p}_l^\eta$:

$$|p(t,x,y) - p_l^\eta(t,x,y)| \leq \int_{\mathbb{R}^N} |\hat{p}(t,x,\xi) - \hat{p}_l^\eta(t,x,\xi)|\, d\xi. \tag{5.52}$$

Therefore our aim in this section will be to estimate the right hand side of (5.52) by considering two regions $\{|\xi| \leq R\}$ and $\{|\xi| > R\}$ separately in the integral, where R is some large number to be chosen later on.

I. Integrating relation (5.52) in a neighborhood of the origin.

We first integrate our Fourier variable ξ in (5.52) over the region $\{|\xi| \leq R\}$. In this case, we make use of a tail estimate for the error of the Taylor approximation of X_t^x which is provided below.

Lemma 5.18. *Let X_t^x be the solution to the SDE (1.1) and consider its approximation $X_l(t,x)$ of order $l \geq l_0$, as given in (5.3). Fix $\bar{l} \in (l, l+1)$ and assume that the vector fields V_α are C_b^∞. There exist constants C_1, C_2 depending only on H, l and the vector fields, such that for all $t \in (0,1]$ and $x, y \in \mathbb{R}^N$ we have*

$$\mathbb{P}\left(|X_t^x - X_l(t,x)| \geq \lambda\right) \leq C_1 \exp\left(-\frac{C_2 \lambda^{\frac{2}{\bar{l}}}}{t^{2H}}\right), \qquad \textit{for all } \lambda > 0. \tag{5.53}$$

Proof. According to [13, Corollary 10.15], we have the following almost sure pathwise estimate

$$|X(t,x) - X_l(t,x)| \leq C \cdot \|\mathbf{B}\|_{p-\mathrm{var};[0,t]}^{\bar{l}}, \tag{5.54}$$

with $C = C_{H,V,l} > 0$, and where $\mathbf{B}$ is the rough path lifting of B alluded to in Proposition 2.16. In equation (5.54), the parameter p is any number greater that $1/H$ and the p-variation norm is defined with respect to the CC-norm. It follows from (5.54) that for any $\lambda > 0$ and $\eta > 0$, we have

$$\mathbb{P}\left(|X_t^x - X_l(t,x)| \geq \lambda\right) \leq \mathbb{P}\left(\|\mathbf{B}\|_{p-\mathrm{var};[0,t]}^{\bar{l}} \geq \lambda/C\right).$$

In addition, the fBm signature satisfies the identity in law

$$(\mathbf{B}_s)_{0\leq s\leq t} \stackrel{d}{=} \left(\delta_{t^H} \circ \mathbf{B}_{\frac{s}{t}}\right)_{0\leq s\leq t}.$$

Owing to the scaling properties of the CC-norm, we thus get that for an arbitrary $\zeta > 0$ we have

$$\begin{aligned}\mathbb{P}(|X_t^x - X_l(t,x) \geq \lambda) &\leq \mathbb{P}\left(\|\mathbf{B}\|_{p-\mathrm{var};[0,1]} \geq \frac{(\lambda/C)^{\frac{1}{\bar{l}}}}{t^H}\right) \\ &\leq \exp\left(-\frac{\zeta(\lambda/C)^{\frac{2}{\bar{l}}}}{t^{2H}}\right) \cdot \mathbb{E}\left[\mathrm{e}^{\zeta\|\mathbf{B}\|_{p-\mathrm{var};[0,1]}^2}\right], \qquad (5.55)\end{aligned}$$

where we have simply involved Markov's inequality for the last inequality. Now notice that a Fernique type estimate holds for the fractional Brownian rough path (cf. [13, Theorem 15.33]), namely there exists $\zeta = \zeta_H > 0$ such that

$$\mathbb{E}\left[\mathrm{e}^{\zeta\|\mathbf{B}\|_{p-\mathrm{var};[0,1]}^2}\right] < \infty.$$

Plugging this inequality into (5.55), our conclusion (5.53) is easily obtained. □

We are now ready to derive a Fourier transform estimate for small values of ξ.

Lemma 5.19. *Keep the same notation and hypothesis as in Lemma 5.18, and also assume that the uniform hypoellipticity condition (1.2) is fulfilled. Let $p(t,x,y)$ be the density of the random variable X_t^x and denote by p_l^η the approximating density defined by (5.20). Then the Fourier transforms $\hat{p} = \mathcal{F}p$ and $\hat{p}_l^\eta = \mathcal{F}p_l^\eta$ satisfy the following inequality over the region $\{|\xi| \leq R\}$,*

$$|\hat{p}(t,x,\xi) - \hat{p}_l^\eta(t,x,\xi)| \leq C_{H,V,l}(1+|\xi|)t^{H\bar{l}}, \qquad (5.56)$$

provided that $t < \tau_1$ for some constant τ_1 depending on H, l and the vector fields.

Proof. Notice that according to our definition (5.20) of p_l^η, we have

$$\hat{p}(t,x,\xi) = \mathbb{E}\left[e^{2\pi i\langle\xi,X_t^x\rangle}\right], \quad \text{and} \quad \hat{p}_l^\eta(t,x,\xi) = \mathbb{E}\left[\eta(U_t)e^{2\pi i\langle\xi,X_l(t,x)\rangle}\right].$$

Hence it is easily seen that

$$\begin{aligned}|\hat{p}(t,x,\xi) - \hat{p}_l^\eta(t,x,\xi)| &\leq \mathbb{E}\left[\left|\mathrm{e}^{-2\pi i\langle\xi,X_t^x\rangle} - \mathrm{e}^{-2\pi i\langle\xi,X_l(t,x)\rangle}\right|\right] + \mathbb{E}[1-\eta(U_t)] \\ &\leq 2\pi|\xi| \cdot \mathbb{E}\left[|X_t^x - X_l(t,x)|\right] + \mathbb{E}[1-\eta(U_t)]. \qquad (5.57)\end{aligned}$$

Now in order to bound the right hand-side of (5.57), we first invoke Lemma 5.18. This yields

$$\begin{aligned}\mathbb{E}\left[|X_t^x - X_l(t,x)|\right] &= \int_0^\infty \mathbb{P}\left(|X_t^x - X_l(t,x)| \geq \lambda\right) d\lambda \\ &\leq C_1 \int_0^\infty \exp\left(-\frac{C_2 \lambda^{\frac{2}{l}}}{t^{2H}}\right) d\lambda = C_3 t^{H\bar{l}}. \end{aligned} \tag{5.58}$$

On the other hand, using a similar argument to the proof of Lemma 5.18, there exists a strictly positive exponent $\alpha_{H,l}$ such that

$$\mathbb{E}[1-\eta(U_t)] \leq \mathbb{P}\left(\|U_t\|_{\mathrm{HS}} \geq \frac{r}{2}\right) \leq C_4 \cdot \mathrm{e}^{-\frac{C_5}{t^{\alpha_{H,l}}}}. \tag{5.59}$$

Therefore taking t small enough, we can make the right hand-side of (5.59) smaller than $C_6 t^{H\bar{l}}$. Hence there exists $\tau_1 > 0$ such that if $t \leq \tau_1$ we have

$$\mathbb{E}[1-\eta(U_t)] \leq C_6 t^{H\bar{l}}. \tag{5.60}$$

Now combining (5.58) and (5.60), we easily get our conclusion (5.56). □

II. Integrating relation (5.52) for large Fourier modes.

We now integrate the Fourier variable ξ over the region $\{|\xi| > R\}$. In this case, we make use of certain upper estimates for $p(t,x,y)$ and $p_l^\eta(t,x,y)$. We start with a bound on the density of X_t^x which is also of independent interest. The main ingredients of the proof are basically known in the literature, but to our best knowledge the result (for the hypoelliptic case) has not been formulated elsewhere.

Proposition 5.20. *Let $p(t,x,y)$ be the density of the random variable X_t^x and as in Lemma 5.19 we assume that the uniform hypoellipticity condition (1.2) is satisfied. Then for each $n \geq 1$, there exist constants $C_{1,n}, C_{2,n}, \nu_n > 0$ depending on n, H and the vector fields such that*

$$|\partial_y^n p(t,x,y)| \leq C_{1,n} t^{-\nu_n} \exp\left(-\frac{C_{2,n}|y-x|^{2\wedge(2H+1)}}{t^{2H}}\right), \tag{5.61}$$

for all $(t,x,y) \in (0,1] \times \mathbb{R}^N \times \mathbb{R}^N$, where ∂_y^n denotes the n-th order derivative operator with respect to the y variable.

Proof. Elaborating on the integration by parts invoked for example in [3, Relation (24)], there exist exponents $\alpha, \beta, p, q > 1$ such that

$$|\partial_y^n p(t,x,y)| \leq C_{1,n} \mathbb{P}(|X_t^x - x| \geq |y-x|)^{\frac{1}{2}} \cdot \|\gamma_{X_t^x}^{-1}\|_{\alpha,p}^{\alpha} \cdot \|\mathbf{D}X_t^x\|_{\beta,q}^{\beta}, \tag{5.62}$$

where the Malliavin covariance matrix $\gamma_{X_t^x}$ is defined by (2.39) and the Sobolev norm $\|\cdot\|_{k,p}$ is introduced in (2.38). Then with (5.62) in hand, we proceed in the following way:

(i) An exponential tail estimate for X_t^x yield the exponential term in (5.61). This step is achieved as in [3, Relation (25)].
(ii) The Malliavin derivatives of X_t^x are estimated as in [4, Lemma 3.5 (1)]. This produces some positive powers of t in (5.61).
(iii) The inverse of the Malliavin covariance matrix is bounded as in [4, Lemma 3.5 (2)]. It gives some negative powers of t in (5.61).

For the sake of conciseness, we will not detail the steps outlined as above. We refer the reader to [3, 4] for the details. □

We now state a lemma which parallels Proposition 5.20 for the approximation X_l. Its proof is somehow delicate and is thus postponed to a separate paragraph.

Lemma 5.21. *Assume the same hypothesis as in Proposition 5.20. Recall that the approximating density p_l^η is defined by (5.20). Fix $l \geq l_0$. Then for each $n \geq 1$ there exists constants $C_n = C_n(H, l)$ and $\gamma_n = \gamma_n(H, l_0)$ such that for all $(t, x) \in (0, 1] \times \mathbb{R}^N$ the following bound holds true*

$$\|\partial_y^n p_l^\eta(t, x, \cdot)\|_{C_b^n(\mathbb{R}^N)} \leq C_n \cdot t^{-\gamma_n}. \tag{5.63}$$

Moreover, the function $\partial_y^n p_l^\eta(t, x, \cdot)$ is compactly supported in $\mathbb{R}^N$.

Remark 5.22. Let us highlight the fact that γ_n in (5.63) depends on l_0 instead of l. This subtle technical point is crucial and requires a non-trivial amount of analysis, which is carried out in the next paragraph.

We now take Lemma 5.21 for granted and we come back to the estimate (5.52) for the region $|\xi| > R$. We are able to state the following result.

Lemma 5.23. *Using the same notation and hypothesis as in Lemma 5.19, the Fourier transforms $\hat{p}$ and $\hat{p}_l^\eta$ are such that for all $|\xi| > R$ we have*

$$|\xi|^{N+2} \left(|\widehat{p}(t, x, \xi)| + |\widehat{p}_l^\eta(t, x, \xi)|\right) \leq C \cdot t^{-\mu}, \tag{5.64}$$

for some strictly positive constants $C = C_{N,H,V,l}$ and $\mu = \mu_{N,H,V,l_0}$.

Proof. According to standard compatibility rules between Fourier transform and differentiation, we have (recall that $\mathcal{F}f$ and $\hat{f}$ are both used to designate the Fourier transform of a function f):

$$\begin{aligned}
&|\xi|^{N+2}(|\hat{p}(t, x, \xi)| + |\hat{p}_l^\eta(t, x, \xi)|) \\
&\leq C_N \left(|\mathcal{F}(\partial_y^{N+2} p(t, x, y)| + |\mathcal{F}(\partial_y^{N+2} p_l^\eta(t, x, y)|\right).
\end{aligned}$$

Plugging (5.61) and (5.63) into this relation and using the fact that $\partial_y^{N+2} p_l^\eta(t, x, \cdot)$ is compactly supported, our claim (5.64) is easily proved. □

III. Comparison of the densities.

Combining the previous preliminary results on Fourier transforms we get the following uniform bound on the difference $p - p_l^\eta$.

Proposition 5.24. *We still keep the same notation and assumptions of Lemma 5.19. Then there exists $\tau > 0$ such that for all $t \leq \tau$ and $x, y \in \mathbb{R}^N$ we have*

$$|p(t,x,y) - p_l^\eta(t,x,y)| \leq C_{H,V,l}\, t. \tag{5.65}$$

Proof. Thanks to (5.52) we can write

$$|p(t,x,y) - p_l^\eta(t,x,y)| \leq \left(\int_{|\xi| \leq \mathbb{R}} + \int_{|\xi| > \mathbb{R}} \right) |\hat{p}(t,x,\xi) - \hat{p}_l^\eta(t,x,\xi)|\, d\xi.$$

Next we invoke the bounds (5.56) and (5.64), which allows to write

$$|p(t,x,y) - p_l^\eta(t,x,y)| \leq C_1 \left(R^{N+1} t^{H\bar{l}} + t^{-\mu} \int_{|\xi|>R} |\xi|^{-N-2} d\xi \right),$$

whenever $t \in (0,1]$, and where we recall that $\bar{l}$ is a fixed number in $[l, l+1]$ introduced in Lemma 5.18. Now an elementary change of variable yields

$$\begin{aligned} &|p(t,x,y) - p_l^\eta(t,x,y)| \\ &\leq C_1 \left(R^{N+1} t^{H\bar{l}} + t^{-\mu} R^{-1} \int_{|\xi| \geq 1} |\xi|^{-(N+1)} d\xi \right) \\ &\leq C_2 \left(R^{N+1} t^{H\bar{l}} + t^{-\mu} R^{-1} \right). \end{aligned} \tag{5.66}$$

We can easily optimize expression (5.66) with respect to R by choosing $R = t^{-(\mu+1)}$. It follows that

$$|p(t,x,y) - p_l^\eta(t,x,y)| \leq C_2 t^{-(N+1)(\mu+1)+H\bar{l}} + t, \tag{5.67}$$

for all $t \in (0,1]$. In addition, recall that a crucial point in our approach is that the exponent μ in (5.67) does not depend on l. Therefore we can choose $l \geq l_0$ large enough, so that

$$-(N+1)(\mu+1) + H\bar{l} \geq 1.$$

For this value of l, the upper bound (5.65) is easily deduced from (5.67).

□

IV. A new proof of Lemma 5.21.

Before proving Lemma 5.21, we mention that the independence on l for the exponent γ_n

was already observed in [17] for the diffusion case. However, in our modest opinion we believe that there is a gap in the argument. We explain the reason as follows. Recall that for a differentiable random vector $Z = (Z^1, \dots, Z^n)$ in the sense of Malliavin, we use the notation $\gamma_Z \triangleq (\langle DZ^i, DZ^j \rangle_{\mathcal{H}})_{1 \le i,j \le n}$ to denote its Malliavin covariance matrix. By the definition (5.3) of $X_l(t,x)$, it is immediate that

$$\gamma_{X_l(t,x)} = JF_l(U_t^{(l)}, x) \cdot \gamma_{U_t^{(l)}} \cdot JF_l(U_t^{(l)}, x)^*. \tag{5.68}$$

Next, let $\pi_{l,l_0} : \mathfrak{g}^{(l)} \to \mathfrak{g}^{(l_0)} \subseteq \mathfrak{g}^{(l)}$ be the canonical orthogonal projection. The matrix form of π_{l,l_0} as a linear function on $\mathfrak{g}(l)$ is given by

$$\pi_{l,l_0} = \begin{pmatrix} \mathrm{Id}_{\mathfrak{g}^{(l_0)}} & 0 \\ 0 & 0 \end{pmatrix}.$$

In [17, Page 420], it was asserted that

$$JF_l(U_t^{(l)}, x) \cdot \gamma_{U_t^{(l)}} \cdot JF_l(U_t^{(l)}, x)^* \ge JF_l(U_t^{(l)}, x) \cdot \pi_{l,l_0} \cdot \gamma_{U_t^{(l)}} \cdot \pi_{l,l_0} \cdot JF_l(U_t^{(l)}, x)^*, \tag{5.69}$$

which we believe was crucial for proving the l-independence in the argument. Now if we write

$$\gamma_{U_t^{(l)}} = \begin{pmatrix} \gamma_{U_t^{(l_0)}} & P \\ Q & R \end{pmatrix},$$

then it is readily checked that (5.69) is equivalent to

$$JF_l(U_t^{(l)}, x) \cdot \begin{pmatrix} 0 & P \\ Q & R \end{pmatrix} \cdot JF_l(U_t^{(l)}, x)^* \ge 0.$$

However, we do not see a reason why this nonnegative definiteness property should hold even if we know that the Malliavin covariance matrices are always nonnegative definite. Therefore the considerations below are devoted to an alternative proof of Lemma 5.21 in the fractional Brownian setting, which also covers the diffusion case.

In view of the decomposition (5.68), we first need the following lemma from [1], which gives an estimate of the Malliavin covariance matrix of $U_t^{(l)}$.

Lemma 5.25. *Given $l \ge 1$, let $U_t^{(l)}$ be the truncated log-signature defined by (5.2). We consider the Malliavin covariance matrix $\gamma_{U_t^{(l)}}$ of the random variable $U_t^{(l)}$, and denote by $\mu_t^{(l)}$ the smallest eigenvalue. Then for any $q > 1$, we have*

$$\sup_{t \in (0,1]} \left\| \frac{t^{2Hl}}{\mu_t^{(l)}} \right\|_q < \infty. \tag{5.70}$$

Now we are able to give the proof of Lemma 5.21.

Proof of Lemma 5.21. As mentioned earlier, the uniform upper bound for the derivatives of $p_l^\eta(t,x,y)$ follows from the same lines as in the proof of Proposition 5.20 (with the same three main ingredients (i)-(ii)-(iii)), based on the integration by parts formula. In the remainder of the proof, we show that the exponent γ_n can be chosen depending only on l_0 but not on l (note, however, that it also depends on H, n and the vector fields). We now divide the proof in several steps.

Step 1: A decomposition based on lowest eigenvalue. As recalled in (5.62) and the strategy of proof of Proposition 5.20, the exponent γ_n in (5.63) comes from integrability estimates for the inverse of the Malliavin covariance matrix of $X_l(t,x)$. To prove the claim, by the definition of $\mathbb{P}_l^\eta(t,x,\cdot)$, it is sufficient to establish the following property: for each $q>1$, we have

$$\sup_{t\in(0,1]} \mathbb{E}\left[\left|\frac{t^{2Hl_0}}{\lambda_t^{(l)}}\right|^q ; \|U_t^{(l)}\|_{\mathrm{HS}} < r\right] < \infty, \tag{5.71}$$

where the random variable $\lambda_t^{(l)}$ is defined by

$$\lambda_t^{(l)} \triangleq \inf_{\eta\in S^{N-1}} \langle \eta, \gamma_{X_l(t,x)}\eta\rangle_{\mathbb{R}^N},$$

that is, $\lambda_t^{(l)}$ is the smallest eigenvalue of $\gamma_{X_l(t,x)}$. In order to lower bound $\lambda_t^{(l)}$, we write $X_l(t,x) = X_{l_0}(t,x) + R_t$, where by the definition (5.3) of $X_l(t,x)$ we have

$$R_t \triangleq \sum_{l_0<|\alpha|\le l} V_{(\alpha)}(x)(\exp U_t^{(l)})^\alpha. \tag{5.72}$$

Then for every $\eta\in S^{N-1}$, we have

$$\begin{aligned}\langle \eta, \gamma_{X_l(t,x)}\eta\rangle_{\mathbb{R}^N} &= \|D\left(\langle \eta, X_l(t,x)\rangle_{\mathbb{R}^N}\right)\|_{\bar{\mathcal{H}}}^2 \\ &= \|D\left(\langle \eta, X_{l_0}(t,x)\rangle_{\mathbb{R}^N}\right) + D\left(\langle \eta, R_t\rangle_{\mathbb{R}^N}\right)\|_{\bar{\mathcal{H}}}^2 .\end{aligned}$$

By invoking the definition (2.39) of $\gamma_{X_{l_0}}$, we get

$$\begin{aligned}\langle \eta, \gamma_{X_l(t,x)}\eta\rangle_{\mathbb{R}^N} &\ge \frac{1}{2}\|D\left(\langle \eta, X_{l_0}(t,x)\rangle_{\mathbb{R}^N}\right)\|_{\bar{\mathcal{H}}}^2 - \|D\left(\langle \eta, R_t\rangle_{\mathbb{R}^N}\right)\|_{\bar{\mathcal{H}}}^2 \\ &= \frac{1}{2}\langle \eta, \gamma_{X_{l_0}(t,x)}\eta\rangle_{\mathbb{R}^N} - \langle \eta, \gamma_{R_t}\eta\rangle_{\mathbb{R}^N} \\ &\ge \frac{1}{2}\langle \eta, \gamma_{X_{l_0}(t,x)}\eta\rangle_{\mathbb{R}^N} - \|\gamma_{R_t}\|_{\mathrm{F}},\end{aligned}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm of a matrix, and we have used the simple inequality $\|a+b\|_E^2 \ge \frac{1}{2}\|a\|_E^2 - \|b\|_E^2$ which is valid in any Hilbert space E. It follows that

$$\lambda_t^{(l)} \ge \frac{1}{2}\lambda_t^{(l_0)} - \|M_{R_t}\|_{\mathrm{F}}. \tag{5.73}$$

In order to go from (5.73) to our desired estimate (5.71), consider $q>1$ and the following decomposition:

$$\mathbb{E}\left[\left|\frac{t^{2Hl_0}}{\lambda_t^{(l)}}\right|^q ; \|U_t^{(l)}\|_{\mathrm{HS}} < r\right] = I_t + J_t, \tag{5.74}$$

where I_t and J_t are respectively defined by

$$I_t = \mathbb{E}\left[\left|\frac{t^{2Hl_0}}{\lambda_t^{(l)}}\right|^q ; \frac{1}{2}\lambda_t^{(l_0)} - \|M_{R_t}\|_{\mathrm{F}} \geq \frac{1}{4}\lambda_t^{(l_0)},\ \|U_t^{(l)}\|_{\mathrm{HS}} < r\right],$$
$$J_t = \mathbb{E}\left[\left|\frac{t^{2Hl_0}}{\lambda_t^{(l)}}\right|^q ; \frac{1}{2}\lambda_t^{(l_0)} - \|M_{R_t}\|_{\mathrm{F}} < \frac{1}{4}\lambda_t^{(l_0)},\ \|U_t^{(l)}\|_{\mathrm{HS}} < r\right].$$

Now we estimate I_t and J_t separately.

Step 2: Upper bound for I_t. To estimate I_t, observe that according to (5.73) we have

$$I_t \leq \mathbb{E}\left[\left|\frac{4t^{2Hl_0}}{\lambda_t^{(l_0)}}\right|^q ; \|U_t^{(l)}\|_{\mathrm{HS}} < r\right]. \tag{5.75}$$

Furthermore, since $X_{l_0}(t,x) = x + F_{l_0}(U_t^{(l_0)}, x)$, we know that

$$\gamma_{X_{l_0}(t,x)} = JF_{l_0}(U_t^{(l_0)}, x) \cdot \gamma_{U_t^{(l_0)}} \cdot JF_{l_0}(U_t^{(l_0)}, x)^*.$$

Therefore, for each $\eta \in S^{N-1}$,

$$\begin{aligned}
&\langle \eta, \gamma_{X_{l_0}(t,x)}\eta\rangle_{\mathbb{R}^N} \\
&= \eta^* \cdot JF_{l_0}(U_t^{(l_0)}, x) \cdot \gamma_{U_t^{(l_0)}} \cdot JF_{l_0}(U_t^{(l_0)}, x)^*\eta \\
&\geq \mu_t^{(l_0)} \cdot \left(\eta^* \cdot JF_{l_0}(U_t^{(l_0)}, x) \cdot JF_{l_0}(U_t^{(l_0)}, x)^* \cdot \eta\right),
\end{aligned} \tag{5.76}$$

where recall that $\mu_t^{(l_0)}$ denotes the smallest eigenvalue of $\gamma_{U_t^{(l_0)}}$. In addition, we choose the constant r in (5.75) as in Corollary 5.8 and Remark 5.9. We hence know that the matrix

$$JF_{l_0}(\pi^{(l_0)}(u), x) \cdot JF_{l_0}(\pi^{(l_0)}(u), x)$$

is uniformly positive definite on $\{(u,x) \in \mathfrak{g}^{(l)} \times \mathbb{R}^N : \|u\|_{\mathrm{HS}} < r\}$. In particular, there exists a constant $c_{V,l} > 0$, such that on the event $\{\|U_t^{(l)}\|_{\mathrm{HS}} < r\}$, we have

$$\eta^* \cdot JF_{l_0}(U_t^{(l_0)}, x) \cdot JF_{l_0}(U_t^{(l_0)}, x)^* \cdot \eta \geq c_{V,l}|\eta|^2, \quad \forall \eta \in \mathbb{R}^N,$$

Therefore, according to (5.76), we conclude that on the event $\{\|U_t^{(l)}\|_{\mathrm{HS}} < r\}$ we have

$$\lambda_t^{(l_0)} = \inf_{\eta \in S^{N-1}} \langle \eta, M_{X_{l_0}(t,x)}\eta\rangle_{\mathbb{R}^N} \geq c_{V,l}\mu_t^{(l_0)}, \tag{5.77}$$

where we recall that $\mu_t^{(l_0)}$ is the smallest eigenvalue of $\gamma_{U_t^{(l_0)}}$. Putting (5.77) into (5.75) it follows that

$$I_t \leq \mathbb{E}\left[\left|\frac{4t^{2Hl_0}}{c_{V,l}\mu_t^{(l_0)}}\right|^q : \|U_t^{(l)}\|_{\mathrm{HS}} < r\right].$$

Hence a direct application of Lemma 5.25 yields

$$\sup_{t\in(0,1]} I_t < \infty.$$

Step 3: Upper bound for J_t. To estimate J_t, according to Hölder's inequality, we have

$$\begin{aligned} J_t &\leq \left(\mathbb{E}\left[\left|\frac{t^{2Hl_0}}{\lambda_t^{(l)}}\right|^{2q}\right]; \|U_t^{(l)}\|_{\mathrm{HS}} < r\right)^{\frac{1}{2}} \cdot \left(\mathbb{P}\left(\frac{1}{2}\lambda_t^{(l_0)} - \|M_{R_t}\|_{\mathrm{F}} < \frac{1}{4}\lambda_t^{(l_0)}; \|U_t^{(l)}\|_{\mathrm{HS}} < r\right)\right)^{\frac{1}{2}} \\ &= \left(\mathbb{E}\left[\left|\frac{t^{2Hl_0}}{\lambda_t(l)}\right|^{2q}\right]; \|U_t^{(l)}\|_{\mathrm{HS}} < r\right)^{\frac{1}{2}} \cdot \left(\mathbb{P}\left(\|M_{R_t}\|_{\mathrm{F}} > \frac{1}{4}\lambda_t^{(l_0)},\ \|U_t^{(l)}\|_{\mathrm{HS}} < r\right)\right)^{\frac{1}{2}}. \end{aligned} \tag{5.78}$$

On the one hand, according to (5.77) applied to general l and Lemma 5.25, we have

$$C_{1,q,l} \triangleq \sup_{t\in(0,1]} \mathbb{E}\left[\left|\frac{t^{2Hl}}{\lambda_t^{(l)}}\right|^{2q} : \|U_t^{(l)}\|_{\mathrm{HS}} < r\right] < \infty.$$

It follows that

$$\begin{aligned} \mathbb{E}\left[\left|\frac{t^{2Hl_0}}{\lambda_t^{(l)}}\right|^{2q} : \|U_t^{(l)}\|_{\mathrm{HS}} < r\right] &= \frac{1}{t^{4qH(l-l_0)}}\mathbb{E}\left[\left|\frac{t^{2Hl}}{\lambda_t^{(l)}}\right|^{2q} : \|U_t^{(l)}\|_{\mathrm{HS}} < r\right] \\ &\leq \frac{C_{1,q,l}}{t^{4qH(l-l_0)}}. \end{aligned} \tag{5.79}$$

In order to bound the right hand-side of (5.78), we also write

$$\mathbb{P}\left(\|M_{R_t}\|_{\mathrm{F}} > \frac{1}{4}\lambda_t^{(l_0)},\ \|U_t^{(l)}\|_{\mathrm{HS}} < r\right) = \mathbb{P}\left(t^{-2Hl_0}\|M_{R_t}\|_{\mathrm{F}} \geq \frac{1}{4}t^{-2Hl_0}\lambda_t^{(l_0)},\ \|U_t^{(l)}\|_{\mathrm{HS}} < r\right).$$

Now a crucial observation is that R_t is defined in terms of signature components of order at least $l_0 + 1$ for the fractional Brownian motion, as easily seen from (5.72). According to the scaling property of the signature, if we define $\xi_t \triangleq t^{-2H(l_0+1)}\|M_{R_t}\|_{\mathrm{F}}$, then ξ_t has moments

of all orders uniformly in $t \in (0,1]$. It follows that

$$\begin{aligned}
&\mathbb{P}\left(\|M_{R_t}\|_{\mathrm{F}} > \frac{1}{4}\lambda_t^{(l_0)},\ \|U_t^{(l)}\|_{\mathrm{HS}} < r\right) \\
&= \mathbb{P}\left(t^{2H}\xi_t \geq \frac{1}{4}t^{-2Hl_0}\lambda_t^{(l_0)},\ \|U_t^{(l)}\|_{\mathrm{HS}} < r\right) \\
&\leq \mathbb{E}\left[\left|\frac{4t^{2H}\xi_t}{t^{-2Hl_0}\lambda_t^{(l_0)}}\right|^{2q(l-l_0)} ;\ \|U_t^{(l)}\|_{\mathrm{HS}} < r\right] \\
&= t^{4qH(l-l_0)}\mathbb{E}\left[\left|\frac{4\xi_t t^{2Hl_0}}{\lambda_t(l_0)}\right|^{2q(l-l_0)} ;\ \|U_t^{(l)}\|_{\mathrm{HS}} < r\right] \\
&\leq t^{4qH(l-l_0)}\left(\mathbb{E}[|4\xi_t|^{4q(l-l_0)}]\right)^{\frac{1}{2}} \cdot \left(\mathbb{E}\left[\left|\frac{t^{2Hl_0}}{\lambda_t(l_0)}\right|^{4q(l-l_0)}\right]\right)^{\frac{1}{2}} \\
&\leq C_{2,q,l}t^{4qH(l-l_0)}. \qquad (5.80)
\end{aligned}$$

Plugging (5.79) and (5.80) into (5.78), we arrive at

$$J_t \leq \sqrt{C_{1,q,l}\cdot C_{2,q,l}} < \infty, \quad \text{for all } t \in (0,1].$$

Step 4: Conclusion. Putting together our estimates on I_t and J_t and inserting them into (5.74), our claim (5.71) is readily proved.

□

5.4 Completing the proof of Theorem 1.7.

Finally, we are in a position to complete the proof of Theorem 1.7. Indeed, recall that (5.51) and (5.65) assert that for x and y such that $d(x,y) \leq t^H$ and $t < \tau$ we have

$$p^\eta(t,x,y) \geq \frac{C_1}{|B_d(x,t^H)|}, \quad \text{and} \quad |p(t,x,y) - p_l^\eta(t,x,y)| \leq C_{H,V,l}t. \qquad (5.81)$$

In addition, owing to (1.6), for small t we get

$$\frac{1}{|B_d(x,t^H)|} \geq \frac{C}{t^{HN/l_0}}. \qquad (5.82)$$

Putting together (5.81) and (5.82), it is thus easily seen that when t is small enough we have

$$p(t,x,y) \geq \frac{C_2}{|B_d(x,t^H)|}.$$

The proof of Lemma 5.21 is thus complete.

Appendix A A disintegration formula on Riemannian manifolds.

Since we are not aware of a specific reference in the literature, for completeness we include a proof of a general disintegration formula on Riemannian manifolds, which is used for Proposition 5.10.

Recall that, if V is an m-dimensional real inner product space, then for each $0 \leqslant p \leqslant m$, the p-th exterior power $\Lambda^p V$ of V carries an inner product structure defined by

$$\langle v_1 \wedge \cdots \wedge v_p, w_1 \wedge \cdots \wedge w_p \rangle_{\Lambda^p V} \triangleq \det\left(\langle v_i, w_j \rangle_{1 \leqslant i,j \leqslant p}\right).$$

In particular, if M is a Riemannian manifold, then for each p, the space of differential p-forms carries a canonical pointwise inner product structure induced from the Riemannian structure of M. A norm on differential p-forms is thus defined pointwisely on M.

Let N be an oriented n-dimensional Riemannian manifold. Suppose that $F: M \triangleq \mathbb{R}^m \to N$ $(m \geqslant n)$ is a non-degenerate C^∞-map in the sense that $(dF)_p$ is surjective everywhere. Then we know that for each $q \in F(\mathbb{R}^m)$, $F^{-1}(q)$ is a closed submanifold of $\mathbb{R}^m$, which carries a canonical Riemannian structure induced from $\mathbb{R}^m$. From differential topology we also know that $F^{-1}(q)$ carries a natural orientation induced from the ones on $\mathbb{R}^m$ and N. In particular, the volume form on $F^{-1}(q)$ is well-defined for every q.

Now we have the following disintegration formula.

Theorem A.1. *Let* vol_N *be the volume form on* N*. Then for every* $\varphi \in C_c^\infty(\mathbb{R}^m)$*, we have*

$$\int_{\mathbb{R}^m} \varphi(x)dx = \int_{q \in N} \mathrm{vol}_N(dq) \int_{x \in F^{-1}(q)} \frac{\varphi(x)}{\|F^* \mathrm{vol}_N\|} \mathrm{vol}_{F^{-1}(q)}(dx), \tag{A.1}$$

where $\mathrm{vol}_{F^{-1}(q)} \triangleq 0$ *if* $F^{-1}(q) = \emptyset$.

Proof. By a partition of unity argument, it suffices to prove the formula locally under coordinate charts on N. Fix $p \in \mathbb{R}^m$ and $q \triangleq F(p) \in N$. Let $(V; y^i)$ be a chart around q. Then the Jacobian matrix $\frac{\partial y}{\partial x}$ has full rank (i.e. rank n) at p. Without loss of generality, we may assume that

$$\frac{\partial y}{\partial x_1} \triangleq \left(\frac{\partial y^i}{\partial x^j}\right)_{1 \leqslant i,j \leqslant n}$$

is non-degenerate, where we write $x_1 = (x^1, \cdots, x^n)$ and $x_2 = (x^{n+1}, \cdots, x^m)$. Define a map $\overline{F}: \mathbb{R}^m \to N \times \mathbb{R}^{m-n}$ by $\overline{F}(x_1, x_2) \triangleq (F(x_1, x_2), x_2)$. It follows that locally around p, we have

$$\frac{\partial \overline{F}}{\partial x} = \begin{pmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} \\ 0 & \mathrm{I}_{m-n} \end{pmatrix}.$$

In particular, $\frac{\partial \overline{F}}{\partial x}$ is non-denegerate at p. Therefore, $\overline{F}$ defines a local diffeomorphism between $p \in U$ and $W = V' \times (a, b)$ for some U, $V' \subseteq V$ and $(a, b) \subseteq \mathbb{R}^{m-n}$. We use $(y, z) \in V' \times (a, b)$

to denote the new coordinates on $U \subseteq \mathbb{R}^m$. Note that every slice $\{(y,z) \in W : \ y = y_0\}$ ($y_0 \in V'$) defines a parametrization of the fiber $F^{-1}(y_0) \cap U$.

By change of coordinates from $x = (x_1, x_2)$ to (y, z), we have

$$dx = \frac{1}{\det\left(\frac{\partial \overline{F}}{\partial x}\right)} dy \wedge dz = \frac{1}{\det\left(\frac{\partial y}{\partial x_1}\right)} dy \wedge dz. \tag{A.2}$$

Since for each $y \in V'$, $z \in (a,b) \mapsto \overline{F}^{-1}(y,z) \in F^{-1}(y)$ defines of parametrization of the fiber $F^{-1}(y) \cap U$, we know that

$$dz = \frac{\mathrm{vol}_{F^{-1}(y)}}{\sqrt{\det\left((\langle \partial_i z, \partial_j z\rangle)_{n+1 \leqslant i,j \leqslant m}\right)}}$$

for each fixed $y \in V'$, where the inner product is defined by the induced Riemannian structure on $F^{-1}(y)$. But we know that $\{\partial_i x : \ 1 \leqslant i \leqslant m\}$ is an orthonormal basis of $T_x\mathbb{R}^m$ for every x. Therefore,

$$\begin{aligned}\langle \partial_i z, \partial_j z\rangle &= \sum_{\alpha,\beta=1}^{m} \frac{\partial x^\alpha}{\partial z^i}\frac{\partial x^\beta}{\partial z^j}\langle \partial_\alpha x, \partial_\beta x\rangle \\ &= \sum_{\alpha=1}^{m} \frac{\partial x^\alpha}{\partial z^i}\frac{\partial x^\alpha}{\partial z^j}.\end{aligned}$$

It follows that

$$(\langle \partial_i z, \partial_j z\rangle)_{n+1 \leqslant i,j \leqslant m} = \left(\frac{\partial x}{\partial z}\right)^* \cdot \frac{\partial x}{\partial z}.$$

On the other hand, we know that

$$\begin{pmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial z} \\ \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial z} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} \\ 0 & \mathrm{I}_{m-n} \end{pmatrix} = \mathrm{I}_m.$$

By comparing components, we get

$$\begin{cases} \frac{\partial x_1}{\partial y} = \left(\frac{\partial y}{\partial x_1}\right)^{-1}, \\ \frac{\partial x_1}{\partial z} = -\left(\frac{\partial y}{\partial x_1}\right)^{-1} \cdot \frac{\partial y}{\partial x_2}, \\ \frac{\partial x_2}{\partial y} = 0, \\ \frac{\partial x_2}{\partial z} = \mathrm{I}_{m-n}. \end{cases}$$

Therefore,

$$\frac{\partial x}{\partial z} = \begin{pmatrix} -\left(\frac{\partial y}{\partial x_1}\right)^{-1} \cdot \frac{\partial y}{\partial x_2} \\ \mathrm{I}_{m-n} \end{pmatrix},$$

and

$$\begin{aligned}
\det\left(\left(\langle\partial_i z,\partial_j z\rangle\right)_{n+1\leqslant i,j\leqslant m}\right) &= \begin{pmatrix} -\left(\frac{\partial y}{\partial x_2}\right)^*\left(\frac{\partial y}{\partial x_1}\right)^{*-1} & \mathrm{I}_{m-n}\end{pmatrix}\cdot\begin{pmatrix} -\left(\frac{\partial y}{\partial x_1}\right)^{-1}\cdot\frac{\partial y}{\partial x_2} \\ \mathrm{I}_{m-n}\end{pmatrix} \\
&= \det\left(\left(\frac{\partial y}{\partial x_2}\right)^*\left(\frac{\partial y}{\partial x_1}\right)^{*-1}\left(\frac{\partial y}{\partial x_1}\right)^{-1}\frac{\partial y}{\partial x_2}+\mathrm{I}_{m-n}\right) \\
&= \det\left(\left(\frac{\partial y}{\partial x_1}\right)^{*-1}\left(\frac{\partial y}{\partial x_1}\right)^{-1}\frac{\partial y}{\partial x_2}\left(\frac{\partial y}{\partial x_2}\right)^*+\mathrm{I}_n\right),
\end{aligned}$$

where in the last equality we have used Sylvester's determinant identity (i.e. $\det(\mathrm{I}_m+AB)=\det(\mathrm{I}_n+BA)$ if A,B are $m\times n$ and $n\times m$ matrices respectively).

Consequently, according to (A.2), we get

$$\begin{aligned}
dx &= \frac{1}{\det\left(\frac{\partial y}{\partial x_1}\right)\cdot\sqrt{\det\left(\left(\langle\partial_i z,\partial_j z\rangle\right)_{n+1\leqslant i,j\leqslant m}\right)}}dy\wedge\mathrm{vol}_{F^{-1}(y)} \\
&= \frac{1}{\sqrt{\det\left(\frac{\partial y}{\partial x_1}\left(\frac{\partial y}{\partial x_1}\right)^*\right)\cdot\det\left(\left(\frac{\partial y}{\partial x_1}\right)^{*-1}\left(\frac{\partial y}{\partial x_1}\right)^{-1}\frac{\partial y}{\partial x_2}\left(\frac{\partial y}{\partial x_2}\right)^*+\mathrm{I}_n\right)}}dy\wedge\mathrm{vol}_{F^{-1}(y)} \\
&= \frac{1}{\sqrt{\det\left(\frac{\partial y}{\partial x_1}\left(\frac{\partial y}{\partial x_1}\right)^*+\frac{\partial y}{\partial x_2}\left(\frac{\partial y}{\partial x_2}\right)^*\right)}}dy\wedge\mathrm{vol}_{F^{-1}(y)} \\
&= \frac{1}{\sqrt{\det\left(\frac{\partial y}{\partial x}\left(\frac{\partial y}{\partial x}\right)^*\right)}}dy\wedge\mathrm{vol}_{F^{-1}(y)}.
\end{aligned}$$

But we also know that

$$\begin{aligned}
\|F^*dy\| &= \sqrt{\langle dy^1\wedge\cdots\wedge dy^n, dy^1\wedge\cdots\wedge dy^n\rangle} \\
&= \sqrt{\det\left(\left(\langle dy^i,dy^j\rangle\right)_{1\leqslant i,j\leqslant n}\right)} \\
&= \sqrt{\det\left(\frac{dy}{dx}\left(\frac{dy}{dx}\right)^*\right)}.
\end{aligned}$$

Therefore, we arrive at

$$\begin{aligned}
dx &= \frac{1}{\|F^*dy\|}dy\wedge\mathrm{vol}_{F^{-1}(y)} \\
&= \frac{1}{\|F^*\mathrm{vol}_N\|}\mathrm{vol}_N\wedge\mathrm{vol}_{F^{-1}(y)}
\end{aligned}$$

on U, where in the last equality we have used the fact that F^* is a linear map.

Now the proof of the theorem is complete. □

Remark A.2. Note that the disintegration formula (A.1) is intrinsic, i.e. it does not depend on coordinates over N. However, the formula is not true when M is not flat. Indeed, the left hand side of the formula depends on the entire Riemannian structure of M since the volume form is defined in terms of the Riemannian metric on M. However, the right hand side of the formula depends only on the Riemannian structure of N and of those fibers. In general, the volume form on M cannot be recovered intrinsically from the geometry of N and the geometry of those fibers.

A particularly useful case of the disintegration formula is when $N = \mathbb{R}^n$. In this case, the formula reads

$$\int_{\mathbb{R}^m} \varphi(x)dx = \int_{y\in\mathbb{R}^m} dy \int_{F^{-1}(y)} \frac{\varphi(x)}{\sqrt{\det\left(\frac{\partial y}{\partial x}\cdot\left(\frac{\partial y}{\partial x}\right)^*\right)}} \mathrm{vol}_{F^{-1}(y)}(dx).$$

References

[1] F. Baudoin, Q. Feng and C. Ouyang, Density of the signature process of fBM, preprint, *arXiv:1904.09384*, 2019.

[2] F. Baudoin and M. Hairer, A version of Hörmander's theorem for the fractional Brownian motion, *Probab. Theory Related Fields* 139 (3-4) (2007) 373–395.

[3] F. Baudoin, E. Nualart, C. Ouyang and S. Tindel, On probability laws of solutions to differential systems driven by a fractional Brownian motion, *Ann. Probab.* 44 (4) (2016) 2554–2590.

[4] F. Baudoin, C. Ouyang and X. Zhang, Varadhan estimates for rough differential equations driven by fractional Brownian motions, *Stochastic Process. Appl.* 125 (2015) 634–652.

[5] F. Baudoin, C. Ouyang and X. Zhang, Smoothing effect of rough differential equations driven by fractional Brownian motions, *Ann. Inst. H. Poincaré Probab. Statist.* 52 (1) (2016) 412–428.

[6] T. Cass and P. Friz, Densities for rough differential equations under Hörmander condition, *Ann. of Math.* 171 (3) (2010) 2115–2141.

[7] T. Cass, M. Hairer, C. Litterer and S. Tindel, Smoothness of the density for solutions to Gaussian rough differential equations, *Ann. Probab.* 43 (1) (2015) 188–239.

[8] L. Coutin, Z. Qian, Stochastic analysis, rough path analysis and fractional Brownian motions. *Probab. Theory Related Fields* 122 (2002), no. 1, 108–140.

[9] A. M. Davie, Differential Equations Driven by Rough Paths: An Approach via Discrete Approximation. *Applied Mathematics Research Express. AMRX*, 2 (2007).

[10] L. Decreusefond and A. Üstünel, Stochastic analysis of the fractional Brownian motion, *Potential Anal.* 10 (1997) 177–214.

[11] P. Friz and M. Hairer, *A course on rough paths with an introduction to regularity structures*, Springer, 2014.

[12] P. Friz and N. Victoir, A variation embedding theorem and applications, *J. Funct. Anal.* 239 (2006) 631–637.

[13] P. Friz and N. Victoir, *Multidimensional stochastic processes as rough paths: theory and applications*, Cambridge University Press, 2010.

[14] A. Kilbas, O. Marichev and S. Samko, *Fractional integrals and derivatives: theory and applications*, Gordon and Breach, Amsterdam, 1993.

[15] S. Kusuoka and D. Stroock, Applications of the Malliavin calculus. I. Stochastic analysis (Katata/Kyoto, 1982), 271–306, North-Holland Math. Library, 32, North-Holland, Amsterdam, 1984.

[16] S. Kusuoka and D. Stroock, Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo* Sect. IA Math. 32 (1985), no. 1, 1-76.

[17] S. Kusuoka and D. Stroock, Applications of the Malliavin calculus, Part III, *J. Fac. Sci. Univ. Tokyo* 34 (1987) 391–442.

[18] T. Lyons, Differential equations driven by rough signals (I): an extension of an inequality of L.C.Young, *Math. Res. Lett.* 1 (1994) 451–464.

[19] T. Lyons, Differential equations driven by rough signals, *Rev. Mat. Iberoam.* 14 (2) (1998) 215–310.

[20] P. Malliavin, Stochastic calculus of variation and hypoelliptic operators. *Proceedings of the International Symposium on Stochastic Differential Equations* (*Res. Inst. Math. Sci.*, Kyoto Univ., Kyoto, 1976), pp. 195–263, Wiley, 1978.

[21] R. Montgomery, *A tour of subriemannian geometries, their geodesics and applications*, American Mathematical Society, Providence, 2002.

[22] J.W. Milnor, *Topology from the differentiable viewpoint*, Princeton University Press, Princeton, 1997.

[23] D. Nualart, *The Malliavin calculus and related topics*, Springer-Verlag, 2006.

[24] V. Pipiras, M. Taqqu, Integration questions related to fractional Brownian motion. *Probab. Theory Related Fields* 118 (2000), no. 2, 251–291.

[25] L.C. Young, an inequality of Höder type, connected with Stieltjes integration, *Acta Math.* 67 (1936) 251–282.