

# Hitting Probabilities for Hypoelliptic Differential Equations Driven by Fractional Brownian Motion

Xi Geng\* and Sheng Wang†

## Abstract

The main goal of this article is to derive a two-sided estimate for hitting probabilities of a hypoelliptic stochastic differential equation (SDE) driven by fractional Brownian motion (fBM) with Hurst parameter  $H \in (1/4, 1)$  in terms of Newtonian-type capacities that are defined with respect to the (sub-Riemannian) control distance associated with the vector fields. As a starting point, we first establish the existence and smoothness of joint densities for the finite-dimensional distributions of the solution in the general context of hypoelliptic SDEs driven by Gaussian rough paths. We then turn to the fBM setting and derive a local upper bound for the joint density in terms of the control distance. As an application of these results, we establish our main estimate on hitting probabilities which generalises a well-known elliptic result of [6] to the hypoelliptic case.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Existence and smoothness of joint density . . . . .	3
1.2	Local upper bound for joint density . . . . .	5
1.3	Hitting probabilities and hypoelliptic capacities . . . . .	6
1.4	Summary of contributions . . . . .	8
<b>2</b>	<b>Preliminaries</b>	<b>9</b>
2.1	Gaussian rough paths . . . . .	9
2.2	The Mandelbrot-van Ness representation . . . . .	11
2.3	Conditional decomposition of fractional Brownian motion . . . . .	14
2.4	Conditional small ball estimates . . . . .	15
<b>3</b>	<b>Existence and smoothness of joint density</b>	<b>18</b>

---

\*School of Mathematics and Statistics, University of Melbourne, Parkville VIC 3010, Australia.  
Email: xi.geng@unimelb.edu.au.

†School of Mathematics and Statistics, University of Melbourne, Parkville VIC 3010, Australia.  
Email: shewang4@student.unimelb.edu.au.

<b>4</b>	<b>Local upper estimate for the joint density of <math>(Y_s, Y_t)</math></b>	<b>22</b>
4.1	Free nilpotent Lie groups and Lie algebras . . . . .	22
4.2	Truncated logarithmic signatures of fractional Brownian motion . . . . .	24
4.2.1	Estimate of $\mathbb{P}(\mathbf{U}_{s,t} \geq u \mid \mathcal{F}_s^W)$ . . . . .	25
4.2.2	Estimate of $\ \mathbf{D}_W \mathbf{U}_{s,t}\ _{k,p,s}$ . . . . .	27
4.2.3	Estimate of $\gamma(\mathbf{U}_{s,t})^{-1}$ . . . . .	30
4.2.4	Proof of Proposition 4.3 . . . . .	30
4.3	Proof of Theorem 1.5 . . . . .	31
<b>5</b>	<b>Hitting probabilities and Newtonian-type capacities</b>	<b>36</b>
5.1	Lower bound for hitting probabilities . . . . .	36
5.2	Upper bound for hitting probabilities . . . . .	45
<b>A</b>	<b>Proof of Lemma 4.8</b>	<b>53</b>

# 1 Introduction

Consider a stochastic differential equation (SDE)

$$dY_t = V(Y_t) dX_t + V_0(Y_t) dt, \quad Y_0 = y_0 \in \mathbb{R}^N, \quad (1.1)$$

where  $X$  is a certain stochastic process in  $\mathbb{R}^d$  and  $V_0, \dots, V_d$  are suitably regular vector fields on  $\mathbb{R}^N$ . When  $X$  is a semimartingale (e.g. Brownian motion) with the stochastic integral being interpreted in the sense of Itô, the SDE (1.1) has been well studied within the framework of Itô's calculus. In a non-semimartingale setting, a typical situation which has gained much interest is when  $X$  is a fractional Brownian Motion (fBM). If the Hurst parameter  $H > 1/2$ , the stochastic integral in (1.1) can be defined in the sense of Young's integration, and the well-posedness of the SDE was established in [20, 23]. If  $1/4 < H < 1/2$ , the well-posedness of the SDE (1.1) requires techniques from rough path theory, which was originally introduced in the seminal work of Lyons [24] in 1998. Controlled differential equations of the type (1.1) with driven process  $X$  being a general rough path are more commonly known as rough differential equations (RDEs). Over the past decades, applications of rough path techniques to the study of quantitative properties of RDEs have been extensively developed by various authors in the literature.

On the other hand, as a counterpart to this analytic development there has been growing interest in the study of probabilistic properties of the SDE (1.1). An important aspect of this is the nondegeneracy and regularity of solutions. In the 1960s, Hörmander [19] proved the existence and smoothness of fundamental solution to the heat equation associated with a hypoelliptic second order differential operator. Later on, in the 1970s Malliavin [25] obtained a probabilistic proof of Hörmander's result by showing that the solution to the SDE (1.1) driven by Brownian motion admits a smooth density under Hörmander's hypoellipticity condition on the vector fields. Along

this pathway, Malliavin developed a powerful technique known as the Malliavin calculus, which has far-reaching applications to various regularity problems in stochastic analysis. A quantitative version of the Doob-Meyer decomposition was obtained by Norris [27] which simplified Malliavin’s proof and served as a key ingredient in this methodology. In the context of fBm with  $H > 1/2$  and elliptic vector fields, Nualart-Hu [20] established the existence of a smooth density for the solution. This result was extended to the hypoelliptic case by Baudoin-Hairer [2]. More recently, by using rough path techniques Cass-Friz [8] obtained the existence of density under hypoelliptic vector fields where the driving process  $X$  can be realised as a wide range of Gaussian processes, including fBm with  $H > 1/4$ . The smoothness of density was established by Cass-Hairer-Litterer-Tindel [10] utilising a pathwise Norris’ lemma in [18] together with a novel interpolation inequality.

A natural stream of questions along this direction are related to the study of quantitative properties of solutions. For instance, in the diffusion context, the seminal paper of Kusuoka-Strook [22] established global upper and lower estimates for the transition density under Hörmander’s condition. Cass-Litterer-Lyons [9] derived the integrability of the Jacobian flow of the solution to RDEs driven by Gaussian rough paths, resulting in a well-known Weibull-type upper bound on the tail distribution of the solution. In a complementary direction, Boedihardjo-Geng [7] showed that the Cass-Litterer-Lyons estimate is optimal by proving a matching lower bound for a large class of rough integrals. For systems driven by fBm with  $H > 1/4$  under elliptic vector fields, Baudoin-Nualart-Ouyang-Tindel [6] obtained a global Gaussian upper bound on density of the solution, while a local lower bound was later obtained by Geng-Ouyang-Tindel [17] in the hypoelliptic case. In [6], the authors also obtained existence, smoothness and quantitative estimates of the joint density (over a pair of times) in the elliptic case, which enabled them to establish a two-sided estimate for hitting probabilities of the solution in terms of Newtonian capacities.

In this article, we consider a hypoelliptic differential equation driven by fBm in the rough path regime  $H \in (1/4, 1)$ . Our main objective is to establish a quantitative relationship between hitting probabilities of the solution and Newtonian-type capacities associated with the sub-Riemannian metric induced by the vector fields. This is a natural generalisation of the elliptic result obtained by Baudoin-Nualart-Ouyang-Tindel [6]. In what follows, we provide an overview of our main results towards this goal. The main theorem is stated in Theorem 1.9.

## 1.1 Existence and smoothness of joint density

Consider an RDE (1.1) driven by fBm with Hurst parameter  $H \in (1/4, 1)$ . While the existence and smoothness of density at a fixed time  $t$  is well-known under Hörmander’s condition, a natural question arises: does the joint distribution  $(Y_{t_1}, \dots, Y_{t_m})$  admit a (smooth) density? If so, what quantitative estimates can one obtain for the joint density? As seen in [6, 11] for instance, joint density estimates play an important role in the analysis of hitting probabilities, which is often studied using potential-theoretic

techniques in the Markovian setting.

Our first main result provides an affirmative answer to this question. We should point out that the existence of joint density is not an immediate consequence of the existence of marginal density, due to the lack of Markovian properties in the current setting. To formulate the result precisely, we first present the basic assumptions.

**Hypothesis 1.1.** The vector fields  $V_0, \dots, V_d$  belong to  $C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$  (i.e. smooth with bounded derivatives of all orders).

To state the key non-degeneracy condition, we need to introduce some notation. Given a word  $I = (i_1, \dots, i_k)$  over the letters  $\{0, 1, 2, \dots, d\}$ , we set

$$V_I = [V_{i_1}, [V_{i_2}, \dots [V_{i_{k-1}}, V_{i_k}] \dots]],$$

where  $[V_i, V_j]$  denotes the Lie bracket between two vector fields; as a differential operator it is defined by  $V_i V_j - V_j V_i$  and one identifies a vector field with a differential operator (taking directional derivatives). The length of a word  $I$  is denoted as  $|I|$ . For each  $l \geq 1$  and  $x \in \mathbb{R}^N$ , we introduce the following family of vectors at  $x$ :

$$\mathcal{W}_l(x) = \{V_I(x) : |I| = k \leq l, 1 \leq i_k \leq d \text{ and } 0 \leq i_r \leq d \text{ for } 1 \leq r \leq k-1\}.$$

**Hypothesis 1.2.** There exist an integer  $\bar{l} \geq 1$  and a positive real number  $\lambda$  such that

$$\sum_{W \in \mathcal{W}_{\bar{l}}(x)} \langle W, u \rangle_{\mathbb{R}^N}^2 \geq \lambda \|u\|_{\mathbb{R}^N}^2$$

holds uniformly for all  $x, u \in \mathbb{R}^N$ . The number  $\bar{l}$  is referred to as the *hypoellipticity constant*.

This is often known as the *uniform Hörmander condition* which plays a basic role in studying regularity properties of the solution as well as quantitative estimates for the density. Our first result is state as follows (see Theorem 3.1 for a more precise formulation).

**Theorem 1.3.** *Consider an RDE*

$$d\mathbf{Y}_t = V(\mathbf{Y}_t) d\mathbf{X}_t + V_0(\mathbf{Y}_t) dt, \quad Y_0 = y_0 \in \mathbb{R}^N. \quad (1.2)$$

Here  $\mathbf{X}$  is a Gaussian rough path including fBM with Hurst parameter  $H \in (1/4, 1)$ , Ornstein-Uhlenbeck process, Brownian bridge as examples (the precise conditions on  $\mathbf{X}$  are provided in Section 2.1 below). The vector fields  $\mathcal{V} = \{V_i\}_{i=1}^d$  are assumed to satisfy Hypotheses 1.1 and 1.2. Let  $Y$  be the solution to the the RDE. Then for any  $m \geq 1$  and  $0 < t_1 < \dots < t_m \leq T$ , the distribution of  $(Y_{t_1}, \dots, Y_{t_m})$  admits a smooth joint density with respect to the Lebesgue measure on  $\mathbb{R}^{mN}$ .

## 1.2 Local upper bound for joint density

A crucial ingredient in the analysis of hitting probabilities is a precise local estimate for the joint density of solution over a pair of times  $(s, t)$ . In what follows, we assume that  $V_0 = 0$  and the driving process is a  $d$ -dimensional fBM with Hurst parameter  $H \in (1/4, 1)$ . In other words, we consider the following RDE

$$d\mathbf{Y}_t = V(Y_t) d\mathbf{B}_t, \quad Y_0 = y_0 \in \mathbb{R}^N, \quad (1.3)$$

where  $\mathbf{B}$  is the canonical rough path lifting of  $B$ . It is known from Theorem 1.3 that  $(Y_s, Y_t)$  admits a smooth joint density. To state our second main result, we need to introduce the following essential definition.

**Definition 1.4.** The *control distance function* associated with the vector fields  $\{V_i\}_{i=1}^d$  is defined by

$$d(x, y) := \inf \left\{ \|\dot{h}\|_{L^2([0,1])} : \Pi_1(x, h) = y \right\}. \quad (1.4)$$

Here  $\Pi_1(x, h)$  denotes the solution to the equation (1.3) at time  $t = 1$  with driving path  $h$  and initial location  $x$ . The infimum ranges over all those  $h : [0, 1] \rightarrow \mathbb{R}^d$  which are absolutely continuous with  $L^2$ -derivatives and satisfy the condition in (1.4).

We use  $B_d(x, r) = \{z \in \mathbb{R}^N : d(x, z) < r\}$  to denote the ball centered at  $x$  with radius  $r$  under the control distance  $d$  and  $|B_d(x, r)|_{\text{Vol}}$  to denote its Lebesgue measure. Our second main result provides a (sharp) local upper estimate of the joint density of  $(Y_s, Y_t)$ .

**Theorem 1.5.** *Suppose that the vector fields  $\mathcal{V} = \{V_i\}_{i=1}^d$  satisfy Hypotheses 1.1 and 1.2. Let  $Y$  be the solution to the RDE (1.3) on  $[0, T]$  and let  $p(s, t, x, y)$  denote the joint density of  $(Y_s, Y_t)$ . For any  $\delta > 0$ , there exist constants  $C, \varepsilon > 0$  depending on  $H, \mathcal{V}, T, \delta$ , such that*

$$p(s, t, x, y) \leq \frac{C}{|B_d(x, (t-s)^H)|_{\text{Vol}}} \quad (1.5)$$

holds uniformly for all  $s < t \in (\delta, T]$ ,  $x, y \in \mathbb{R}^N$  with  $d(x, y) \leq (t-s)^H$  and  $t-s < \varepsilon$ .

**Remark 1.6.** If the vector fields are elliptic (e.g.  $\bar{l} = 1$  in Hypothesis 1.2), the control distance is locally equivalent to the Euclidean distance: there exist  $C, \delta > 0$  such that

$$\frac{1}{C} \|x - y\|_{\mathbb{R}^N} \leq d(x, y) \leq C \|x - y\|_{\mathbb{R}^N}, \quad (1.6)$$

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

Some related quantitative density estimates in the literature are discussed as follows.

- In the diffusion context ( $B$  is the Brownian motion), Kusuoka-Strook [22] obtained a global upper and lower bound for the transition density  $p(t, x, y)$  under the uniform Hörmander condition:

$$\frac{1}{M |B_d(x, t^{1/2})|_{\text{Vol}}} e^{-Md(x,y)^2/t} \leq p(t, x, y) \leq \frac{M}{|B_d(x, t^{1/2})|_{\text{Vol}}} e^{-d(x,y)^2/Mt}.$$

- For the fBM driven case with elliptic vector fields, Baudoin-Nualart-Ouyang-Tindel [6] established the following upper estimate for the joint density  $p(s, t, x, y)$  of  $(Y_s, Y_t)$ :

$$p(s, t, x, y) \leq \frac{C_M}{(t-s)^{nH}} \left( \frac{|t-s|^H}{|x-y|} \wedge 1 \right)^p$$

for all  $x, y \in [-M, M]^N$ .

- For the fBM driven case with hypoelliptic vector fields (i.e. under the uniform Hörmander condition), Geng-Ouyang-Tindel [17] established the following local lower bound for the density:

$$p(t, x, y) \geq \frac{C}{|B_d(x, t^H)|_{\text{Vol}}}$$

for all  $(t, x, y) \in (0, 1] \times \mathbb{R}^N \times \mathbb{R}^N$  satisfying  $d(x, y) \leq t^H$  and  $t$  small.

If one attempts to replicate the argument in [6] for the hypoelliptic case, one would only obtain a non-sharp upper bound for the joint density which does not yield optimal capacity estimates for hitting probabilities. We will use Kusuoka-Stroock's method in [22], which respects the sub-Riemannian metric associated with the vector fields and hence yields sharp estimates.

### 1.3 Hitting probabilities and hypoelliptic capacities

Finally, as an application of Theorem 1.5 we establish a quantitative relationship between hitting probabilities of sets in  $\mathbb{R}^N$  by the solution to the RDE (1.3) and the associated Newtonian-type capacities. A classical problem in potential theory is to derive conditions under which a process hits a given set  $A$  with positive probability. Typically, such conditions are described in terms of certain capacities. In our setting, we shall look for the following type of relations:

$$\mathbb{P}(Y([0, T]) \cap A \neq \emptyset) > 0 \quad \text{iff} \quad \text{Cap}_\alpha(A) > 0.$$

Here  $\text{Cap}_\alpha(A)$  is a suitable notion of  $\alpha$ -dimensional capacity, which is defined in terms of the control distance function  $d$  (see (1.4)).

Let us now define the capacity precisely. Given any real number  $\alpha$ , the  $\alpha$ -dimensional Newtonian kernel is defined by

$$K_\alpha(r) := \begin{cases} r^{-\alpha}, & \text{if } \alpha > 0; \\ \log(N_0/r), & \text{if } \alpha = 0; \\ 1, & \text{if } \alpha < 0, \end{cases} \quad (1.7)$$

where  $N_0 > 0$  is a suitable constant whose value is not particularly important. Let  $m(\cdot, \cdot)$  be a given metric on  $\mathbb{R}^N$ . For any probability measure  $\mu$  on  $\mathbb{R}^N$  with compact

support, its  $\alpha$ -dimensional Newtonian energy with respect to the metric  $m$  is defined by

$$\mathcal{E}_\alpha(\mu) := \int_{\mathbb{R}^N \times \mathbb{R}^N} K_\alpha(m(x, y)) \mu(dx) \mu(dy). \quad (1.8)$$

**Definition 1.7.** Let  $A$  be a Borel subset of  $\mathbb{R}^N$ . The  $\alpha$ -dimensional capacity of  $A$  with respect to the metric  $m$  is the reciprocal of the minimal energy:

$$\text{Cap}_\alpha(A) := \left[ \inf_{\mu \in \mathcal{P}(A)} \mathcal{E}_\alpha(\mu) \right]^{-1},$$

where  $\mathcal{P}(A)$  denotes the set of probability measures that are compactly supported within  $A$ .

To relate the capacity to dynamical properties of the process, the choice of the metric  $m(\cdot, \cdot)$  is connected with both the dimensionality of the state space and the way the process propagates. For elliptic systems, the process diffuses uniformly in all directions, and therefore the Euclidean distance accurately measures its propagation. Taking  $m$  to be the Euclidean distance, i.e.  $m(x, y) = \|x - y\|_{\mathbb{R}^N}$ , a considerable amount of work has been done to characterise hitting probabilities for a wide range of stochastic processes (see [16] and references therein).

In the Markovian setting, this is effectively studied by means of techniques from classical potential theory. In a non-Markovian setting, Baudoin-Nualart-Ouyang-Tindel [6] proved the following quantitative estimate. Let  $Y$  be the solution to the RDE (1.3) driven by a  $d$ -dimensional fBM with Hurst parameter  $H \in (1/4, 1)$ . Suppose that the vector fields are uniformly elliptic (i.e. satisfying Hypothesis 1.2 with  $\bar{l} = 1$ ). Fix  $0 < a < b \leq 1$ ,  $M > 0$  and  $\eta > 0$ . Then there exist positive constants  $C_1, C_2$  depending on  $a, b, M, \eta$ , such that the estimate

$$C_1 \text{Cap}_{N-1/H}(A) \leq \mathbb{P}(Y([a, b]) \cap A \neq \emptyset) \leq C_2 \text{Cap}_{N-1/H-\eta}(A) \quad (1.9)$$

holds uniformly for all compact sets  $A \subseteq [-M, M]^N$ . Here the capacity is defined with respect to the Euclidean distance. This generalises the well-known result that “a  $d$ -dimensional fBM hits points iff  $d < 1/H$ ” (see [30] and references therein).

On the other hand, for hypoelliptic systems, the geometry induced by the dynamics is not uniform in all directions since those directions along Lie brackets of the vector fields are harder to explore. The control distance  $d$  defined by (1.4) captures the underlying sub-Riemannian geometry through which the process propagates. We now define the intrinsic dimension associated with this geometry (see [1, Section 20.1]).

Let  $\mathcal{V} = \{V_1, \dots, V_d\}$  be a family of smooth vector fields which satisfies Hypothesis 1.2. Let  $\mathcal{D}$  denote the  $C^\infty(\mathbb{R}^N; \mathbb{R})$ -module generated by  $\mathcal{V}$ . For each  $k \geq 1$ , we define the submodule  $\mathcal{D}^k$  of  $C^\infty(\mathbb{R}^N; \mathbb{R}^N)$  recursively by

$$\mathcal{D}^1 := \mathcal{D}, \quad \mathcal{D}^{k+1} := \mathcal{D}^k + [\mathcal{D}, \mathcal{D}^k].$$

We also set  $\mathcal{D}_x^k := \{W_x : W \in \mathcal{D}^k\}$ . The uniform Hörmander condition ensures that for each  $x \in \mathbb{R}^N$  there is a smallest integer  $r(x) \leq \bar{l}$  such that  $\mathcal{D}_x^{r(x)} = \mathbb{R}^N$ . This gives rise to a *canonical flag of  $\mathcal{D}$*  at  $x$ :

$$\{0\} =: \mathcal{D}_x^0 \subseteq \mathcal{D}_x^1 \subseteq \dots \subseteq \mathcal{D}_x^{r(x)} = \mathbb{R}^N.$$

The list  $\{\dim \mathcal{D}_x^k\}_{1 \leq k \leq r(x)}$  of integers is called the *growth vector* of  $\mathcal{V}$  at  $x$ . The vector fields in  $\mathcal{V}$  are said to be *equiregular* if the growth vector is constant when  $x$  varies over  $\mathbb{R}^N$ .

**Definition 1.8.** Suppose that  $\mathcal{V}$  is equiregular. The *homogeneous dimension* of  $\mathcal{V}$  is defined by

$$Q := \sum_{k=1}^{r(x)} k (\dim \mathcal{D}_x^k - \dim \mathcal{D}_x^{k-1}),$$

which is independent of  $x$  by assumption.

We now state our final result which is also the main theorem of the present article. In what follows, we choose  $m(\cdot, \cdot)$  to be the control distance  $d$  in Definition 1.7 of the capacity.

**Theorem 1.9.** *Suppose that the family  $\mathcal{V}$  of vector fields satisfies Hypotheses 1.1, 1.2 and is equiregular with homogeneous dimension  $Q$ . Let  $Y$  be the solution to the RDE (1.3) driven by a  $d$ -dimensional fractional Brownian motion  $B$  with Hurst parameter  $H \in (1/4, 1)$ . For any fixed real numbers  $0 < a < b$ ,  $M > 0$ ,  $\eta_1 > 0$  and  $0 < \eta_2 < 1/H$ , there exist positive constants*

$$C_1 = C_1(a, b, \mathcal{V}, H, M, \eta_1, Q), \quad C_2 = C_2(a, b, \mathcal{V}, H, M, \eta_2, Q)$$

such that the estimate

$$C_2 \text{Cap}_{Q-1/H+\eta_2}(A) \leq \mathbb{P}(Y([a, b]) \cap A \neq \emptyset) \leq C_1 \text{Cap}_{Q-1/H-\eta_1}(A) \quad (1.10)$$

holds uniformly for all compact sets  $A \subseteq B_a(o, M)$ . In addition, if  $Q = N$  one can take  $\eta_2 = 0$ .

**Remark 1.10.** Theorem 1.9 is a natural extension of the elliptic result (1.9) of [6]. In view of (1.9), the estimate (1.10) is nearly optimal since  $\eta_1, \eta_2$  are arbitrarily small numbers.

## 1.4 Summary of contributions

In what follows, we summarise the novelty and contributions of the current work.

(A) To prove the existence and smoothness of joint density for  $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_m})$ , instead of successively conditioning on  $\mathcal{F}_{t_m}, \dots, \mathcal{F}_{t_1}$  with step-by-step applications of

Malliavin calculus, we analyse the joint Malliavin covariance matrix directly. Then, by a suitable change of variable, the required nondegeneracy of the joint Malliavin matrix is seen as a consequence of the nondegeneracy of the vector fields over finitely many points, which is ensured by Hypothesis 1.2.

(B) By invoking a disintegration formula from Riemannian manifold, we are able to relate the joint density of  $(Y_s, Y_t)$  with the distribution of the truncated signature process  $\mathbf{B}_{s,t}$  on free nilpotent Lie groups. The problem thus reduces to establishing a smooth upper bound for the conditional density of the signature of fBM which is more manageable due to its explicit structure.

(C) We employ the control distance (1.4) (rather than Euclidean) in Definition 1.7 of the capacity. It accurately captures the hypoelliptic behaviour of the dynamics and therefore effectively controls the hitting probabilities through the main estimate (1.10).

(D) In the hypoelliptic case, the Newtonian kernel  $K(d(x, y))$  lacks translation invariance which prevents the use of Fourier analysis and invalidates the elliptic argument. To address this, we slightly increase the energy dimension to  $\alpha + \eta$  under which one can effectively estimate the energy of smoothed measures.

**Organisation.** In Section 2, we discuss some basic notions and preliminary tools that are needed for our later analysis. In Section 3, we prove the existence of smooth joint densities for the finite-dimensional distributions of the solution. In Section 4, we derive a local upper estimate for the joint density, which is a crucial ingredient for establishing our main hitting probability estimate. In Section 5, we develop the proof of the main Theorem 1.9.

## 2 Preliminaries

In this section, we first review some basic notions on Gaussian rough paths and then restrict to the fractional Brownian motion setting where more tools are discussed.

### 2.1 Gaussian rough paths

We begin by recalling the following definition. Given a mean-zero, real-valued Gaussian process  $\{Z_t : t \in [0, T]\}$  in  $\mathbb{R}$ , let  $R(s, t) := \mathbb{E}[Z_s Z_t]$  denote its covariance function.

**Definition 2.1.** Let  $1 \leq \rho < 2$ . The covariance function  $\phi$  is said to have *finite (2D)  $\rho$ -variation* if

$$V_\rho(R; [0, T] \times [0, T])^\rho := \sup_{\mathcal{D}, \mathcal{D}'} \sum_{\substack{[s,t] \in \mathcal{D} \\ [u,v] \in \mathcal{D}'}} |R(s, t; v, u)|^\rho < \infty,$$

where  $R(s, t; v, u) := \mathbb{E}[(Z_t - Z_s)(Z_u - Z_v)]$  and  $\mathcal{D}, \mathcal{D}'$  are any finite partitions of  $[0, T]$ . The function  $\phi$  is said to have *finite (2D) Hölder-controlled  $\rho$ -variation*, if there exists

$C > 0$  such that

$$V_\rho(R; [s, t] \times [s, t]) \leq C(t - s)^{1/\rho}$$

for all  $0 \leq s \leq t \leq T$ .

**Theorem 2.2** (see [14]). *Let  $(X_t)_{0 \leq t \leq T} = (X_t^1, \dots, X_t^d)$  be a mean-zero Gaussian process with i.i.d. components. Let  $R$  denote the covariance function of any of its components. Suppose that  $R$  is of finite 2D  $\rho$ -variation for some  $\rho \in [1, 2)$ . Then the process  $X$  admits a canonical lifting to a geometric  $p$ -rough path  $\mathbf{X}$  for all  $p > 2\rho$ . In addition, its (intrinsic) Cameron-Martin space  $\bar{\mathcal{H}}$  can be embedded into the space  $C^{\rho\text{-var}}([0, T], \mathbb{R}^d)$ , more precisely, one has*

$$\|h\|_{\bar{\mathcal{H}}} \geq \frac{|h|_{\rho\text{-var}; [0, T]}}{\sqrt{V_\rho(R; [0, T] \times [0, T])}}.$$

**Remark 2.3.** Writing  $\bar{\mathcal{H}}_B$  for the Cameron-Martin space of fBm for  $H \in (1/4, 1/2)$ , the variation embedding in [12] gives the stronger result that

$$\bar{\mathcal{H}}_B \hookrightarrow C^{q\text{-var}}([0, T], \mathbb{R}^d) \quad \forall q > (H + 1/2)^{-1}.$$

In what follows, we always assume that  $(X_t)_{0 \leq t \leq T} = (X_t^1, \dots, X_t^d)$  is a mean-zero Gaussian process with i.i.d. components. Let  $R$  be the covariance function of any of its components. To study regularity properties of rough differential equations driven by a Gaussian process, one needs the following standard assumption.

**Condition 2.4.** We assume that  $R$  has finite Hölder controlled  $\rho$ -variation for some  $\rho \in [1, 2)$ . In particular,  $X$  admits a canonical lifting to a geometric  $p$ -rough path  $\mathbf{X}$ . We also assume that  $\bar{\mathcal{H}}$ , the Cameron-Martin space associated with  $X$ , has *Young complementary regularity* with respect to  $X$  in the following sense: one has the continuous embedding

$$\bar{\mathcal{H}} \hookrightarrow C^{q\text{-var}}([0, T], \mathbb{R}^d)$$

for some  $q \geq 1$  satisfying  $1/p + 1/q > 1$ .

**Remark 2.5.** Complementary Young regularity for fBM holds when  $H > 1/4$  by taking  $\rho \in [1, 3/2)$ .

The following two conditions are fundamental in the study of non-degeneracy (existence and smoothness of density) for RDEs driven by a Gaussian process (see [10]). Let  $Z_t$  denote any component of  $X_t$  (recall that they are i.i.d. Gaussian).

**Condition 2.6.** We assume that there exists  $\alpha > 0$  such that

$$\inf_{0 \leq s < t \leq T} \frac{1}{(t - s)^\alpha} \text{Var}(Z_{s,t} | \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}) > 0. \quad (2.1)$$

The *index of nondeterminism* is the smallest  $\alpha$  for which (2.1) is true.

**Condition 2.7.** For every  $[u, v] \subseteq [s, t] \subseteq [0, S] \subseteq [0, T]$ , one has

$$\text{Cov}(Z_{s,t}, Z_{u,v} | \mathcal{F}_{0,s} \vee \mathcal{F}_{t,S}) \geq 0.$$

**Remark 2.8.** It was shown in [10, Lemma 4.1] that for fBm with Hurst parameter  $H \in (0, 1/2)$ , the relation (2.1) holds for any  $\alpha \in (1, 2]$ . The case when  $H \in (1/2, 1)$  was handled in [29] where  $\alpha$  is shown to be  $H + 1$ . For the Ornstein-Uhlenbeck process and the Brownian bridge, one has  $\alpha = 1$ .

**Definition 2.9.** Let  $\theta \in (0, 1)$ . A path  $X : [0, T] \rightarrow \mathbb{R}^d$  is called  $\theta$ -Hölder rough if there exists a constant  $c > 0$  such that for every  $s \in [0, T]$ , every  $\varepsilon$  in  $(0, T/2]$ , and every unit vector  $\zeta$  in  $\mathbb{R}^d$ , there exists  $t$  in  $[0, T]$  such that  $\varepsilon/2 < |t - s| < \varepsilon$  and

$$|\langle \zeta, X_{s,t} \rangle| > c\varepsilon^\theta.$$

The largest such constant is called the *modulus of  $\theta$ -Hölder roughness* and is denoted as  $L_\theta(X)$ .

**Remark 2.10.** It was proven in [10, Corollary 5.10] that under Condition 2.6, one has  $L_\theta(X) \in L^p$  for all  $p > 1$ .

## 2.2 The Mandelbrot-van Ness representation

Let  $B = (B_t)_{0 \leq t \leq T}$  be a one-dimensional fBM with Hurst parameter  $H \in (0, 1)$ . Its *canonical Hilbert space* (non-intrinsic Cameron-Martin space)  $\mathcal{H}_B$  is defined to be the completion of the linear span of  $\{\mathbf{1}_{[0,t]} : t \in [0, T]\}$  with respect to the inner product

$$\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \rangle_{\mathcal{H}_B} = R_H(t, s), \quad (2.2)$$

where

$$R_H(t, s) := \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H})$$

denotes the covariance function of  $B$ . Note that the canonical Hilbert space of the standard Brownian motion  $W = (W_t)_{t \in \mathbb{R}}$  is  $\mathcal{H}_W = L^2(\mathbb{R})$  with

$$\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \rangle_{\mathcal{H}_W} = \min(t, s).$$

The fBM admits the so-called *Mandelbrot-van Ness representation* which will be useful to us. More precisely, the covariance  $R_H$  admits the integral form

$$R_H(t, s) = \int_{-\infty}^{t \wedge s} k_H(t, u) k_H(s, u) du, \quad (2.3)$$

where  $k_H$  is the kernel defined by

$$k_H(t, u) := c_H \left( (t - u)_+^{H-1/2} - (-u)_+^{H-1/2} \right)$$

and  $c_H$  is an explicit constant ensuring (2.3) to hold. This leads to the representation

$$B_t = \int_{-\infty}^t k_H(t, s) dW_s, \quad 0 \leq t \leq T \quad (2.4)$$

which expresses  $B_t$  as a Wiener integral with respect to  $W$ . This representation defines an isometric embedding of  $\mathcal{H}_B$  into  $\mathcal{H}_W$  via

$$\mathcal{K}(\mathbf{1}_{[0,t]}) := k_H(t, \cdot) \mathbf{1}_{(-\infty, t]}(\cdot), \quad 0 \leq t \leq T.$$

Indeed, according to (2.3) one has

$$\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \rangle_{\mathcal{H}_B} = \langle \mathcal{K}(\mathbf{1}_{[0,t]}), \mathcal{K}(\mathbf{1}_{[0,s]}) \rangle_{\mathcal{H}_W}.$$

For any step function  $g = \sum_{i=1}^n a_i \mathbf{1}_{[t_{i-1}, t_i]}$  on  $[0, T]$ , we define

$$\mathcal{I}(g) = \sum_{i=1}^n a_i (B_{t_i} - B_{t_{i-1}}).$$

By the definition of the inner product (2.2), one has

$$\mathbb{E}[\mathcal{I}(f)\mathcal{I}(g)] = \langle f, g \rangle_{\mathcal{H}_B} \quad (2.5)$$

for any step functions  $f, g$ . As a result, the map  $\mathcal{I}$  can be extended to an isometry  $\mathcal{I} : \mathcal{H}_B \rightarrow L^2(\Omega)$ . The (intrinsic) *Cameron-Martin* space  $\bar{\mathcal{H}}_B$  is defined to be

$$\bar{\mathcal{H}}_B := \{ \bar{h} \mid \bar{h}_t = \mathbb{E}[B_t \mathcal{I}(h)], \quad 0 \leq t \leq T, \quad h \in \mathcal{H}_B \},$$

where the inner product in  $\bar{\mathcal{H}}_B$  is induced by  $\langle \bar{f}, \bar{g} \rangle_{\bar{\mathcal{H}}_B} = \langle f, g \rangle_{\mathcal{H}_B}$ . For any  $h \in \mathcal{H}_B$ , we use  $\bar{h}$  to denote its corresponding element in  $\bar{\mathcal{H}}_B$ .

To avoid ambiguity, we adopt the notation  $\mathbf{D}_W$  (resp.  $\mathbf{D}_B$ ) for the Malliavin derivative with respect to  $W$  (resp.  $B$ ), and denote the corresponding Sobolev spaces by  $\mathbb{D}_W^{k,p}$  (resp.  $\mathbb{D}_B^{k,p}$ ), where  $k$  is the differentiability index and  $p$  is the integrability index. Here  $F \in \mathbb{D}_W^{k,p}$  means

$$\sum_{j=0}^k \left( \mathbb{E} \left[ \|\mathbf{D}_W^j F\|_{(\mathcal{H}_W)^{\otimes j}}^p \right] \right)^{1/p} < \infty,$$

and analogously for  $\mathbb{D}_B^{k,p}$ . The connection between  $\mathbb{D}_W^{k,p}$  and  $\mathbb{D}_B^{k,p}$  is described by the following result.

**Proposition 2.11.** *Let  $(W_u)_{-\infty < u \leq T}$  be a Brownian motion on the canonical Wiener space  $(\Omega, \mathcal{F}^W, \mathbb{P})$ . Given fixed  $H \in (0, 1)$ , we define the process  $(B_u)_{0 \leq u \leq T}$  by*

$$B_u = \int_{-\infty}^u k_H(u, s) dW_s, \quad 0 \leq u \leq T,$$

where  $k_H$  is the Volterra kernel associated with the Hurst parameter  $H$ . Then the following relations hold true:

(i)  $\mathbb{D}_B^{1,2} = (\mathcal{K})^{-1}(\mathbb{D}_W^{1,2})$ ;

(ii) For all  $F \in \mathbb{D}_B^{1,2}$ , one has  $\mathbf{D}_W F = \mathcal{K} \mathbf{D}_B F$ .

We shall consider the Malliavin calculus with respect to the filtration

$$\mathcal{F}_s^W := \sigma(W_u, -\infty < u \leq s)$$

and define the following conditional Malliavin-Sobolev norm:

$$\|F\|_{k,p,s} := \left( \sum_{j=0}^k \mathbb{E} \left[ \|\mathbf{D}_W^j F\|_{(\mathcal{H}_{W,s})^{\otimes j}}^p \mid \mathcal{F}_s^W \right] \right)^{1/p}, \quad (2.6)$$

where  $\mathcal{H}_{W,s} := L^2([s, T])$  for any  $0 \leq s \leq T$ .

To ensure compatibility with the Hilbert space  $\mathcal{H}_{W,0}$ , we introduce the kernel

$$L_H(t, s) := c_H(t-s)^{H-1/2}, \quad \text{for all } 0 \leq s \leq t \leq T.$$

This defines an operator

$$\mathcal{L}(\mathbf{1}_{[0,t]}) := L_H(t, \cdot) \mathbf{1}_{[0,t]}(\cdot), \quad 0 \leq t \leq T,$$

and the associated Volterra-type process

$$L_t := c_H \int_0^t (t-r)^{H-1/2} dW_r, \quad 0 \leq t \leq T.$$

Note that for any  $0 \leq s \leq t \leq T$ , one has the identities

$$\begin{aligned} \mathcal{K}(\mathbf{1}_{[0,t]}) &= \mathcal{L}(\mathbf{1}_{[0,t]}) \quad \text{on } [0, T] \\ \langle \mathcal{K}(\mathbf{1}_{[0,t]}), \mathcal{K}(\mathbf{1}_{[0,s]}) \rangle_{\mathcal{H}_{W,0}} &= \langle \mathcal{L}(\mathbf{1}_{[0,t]}), \mathcal{L}(\mathbf{1}_{[0,s]}) \rangle_{\mathcal{H}_{W,0}}. \end{aligned} \quad (2.7)$$

This immediately leads to the following result.

**Proposition 2.12.** For all  $F \in \mathbb{D}_B^{1,2}$ , one has

$$\mathbf{D}_W F = \mathcal{L}(\mathbf{D}_B F)$$

as a  $L^2([0, T])$ -valued random variable.

*Proof.* This follows directly from Proposition 2.11 and (2.7).  $\square$

The following two lemmas are taken from [3, Proposition 9 and Lemma 2 respectively]. Two processes are said to be *equivalent* if their laws are mutually absolutely continuous.

**Lemma 2.13.** For any  $H \in (0, 1)$ , the process  $(B_t)_{0 \leq t \leq T}$  is equivalent to  $(L_t)_{0 \leq t \leq T}$ .

**Lemma 2.14.** Let  $X_t^1, X_t^2$  be Gaussian processes and assume that they are equivalent on  $[0, T]$ . Then there exists  $C_T > 0$  such that

$$E \left[ \left( \int_0^T f(s) dX_s^1 \right)^2 \right] \leq C_T E \left[ \left( \int_0^T f(s) dX_s^2 \right)^2 \right]$$

for any step function  $f$ .

### 2.3 Conditional decomposition of fractional Brownian motion

Let  $W_t = (W_t^1, \dots, W_t^d)$  be a  $d$ -dimensional Brownian motion defined on  $(-\infty, T]$ . The Mandelbrot-van Ness representation defines an fBM with Hurst parameter  $H \in (0, 1)$  as a Volterra-type process given by

$$B_t = \int_{-\infty}^t k_H(t, u) dW_u, \quad 0 \leq t \leq T.$$

For any  $0 \leq s \leq u \leq T$ , we decompose the increment  $B_{s,u} := B_u - B_s$  into two parts according to their  $\mathcal{F}_s^W$ -measurability:

$$B_{s,u} = L_{s,u} + Q_{s,u},$$

where

$$\begin{aligned} L_{s,u} &:= c_H \int_s^u (u-r)^{H-1/2} dW_r, \\ Q_{s,u} &:= c_H \int_{-\infty}^s [(u-r)^{H-1/2} - (s-r)^{H-1/2}] dW_r. \end{aligned} \tag{2.8}$$

Clearly,  $(L_{s,u})_{s \leq u \leq T}$  is independent of  $\mathcal{F}_s^W$  while  $(Q_{s,u})_{s \leq u \leq T}$  is  $\mathcal{F}_s^W$ -measurable.

The following simple observation will be used for several times in the sequel.

**Lemma 2.15.** *For any  $\varepsilon > 0$  and  $s \geq 0$ , the joint distribution of the process*

$$(L_{s,s+\varepsilon u}, Q_{s,s+\varepsilon u})_{0 \leq u \leq 1}$$

*coincides with that of*

$$(\varepsilon^H L_{0,u}, \varepsilon^H Q_{0,u})_{0 \leq u \leq 1}.$$

*Proof.* This is a simple consequence of scaling invariance. □

To compare the Malliavin derivative of functionals of Gaussian processes, we introduce the following notion of equivalence.

**Definition 2.16.** Let  $X$  and  $Y$  be random variables defined on the Wiener space  $(\Omega, \mathcal{F}^W, \mathbb{P})$ . We say that  $X$  and  $Y$  are *equivalent in the sense of Malliavin* and simply write  $X \stackrel{M}{=} Y$ , if one has

$$\|\mathbf{D}_W^k X\|_{\mathcal{H}_W^{\otimes k}} \stackrel{d}{=} \|\mathbf{D}_W^k Y\|_{\mathcal{H}_W^{\otimes k}}$$

for every integer  $k \geq 0$  provided that both sides are well-defined. Here  $\mathbf{D}_W^k$  denotes the  $k$ -th Malliavin derivative with respect to the Brownian motion  $W$ .

The following lemma provides a sufficient condition under which two random variables are equivalent in the above sense.

**Lemma 2.17.** *Let  $\Phi^1, \Phi^2: C([0, 1]) \rightarrow \mathbb{R}$  be measurable functionals such that  $\Phi^1(L_{0,\cdot})$  and  $\Phi^2(Q_{0,\cdot})$  are smooth in the sense of Malliavin, i.e.,*

$$\mathbb{E} \left[ \left\| \mathbf{D}_W^k \Phi^1(L_{0,\cdot}) \right\|_{\mathcal{H}_W^{\otimes k}}^p \right] < \infty, \quad \mathbb{E} \left[ \left\| \mathbf{D}_W^k \Phi^2(Q_{0,\cdot}) \right\|_{\mathcal{H}_W^{\otimes k}}^p \right] < \infty$$

for all  $k \geq 0$  and  $p \geq 1$ . Then for any  $s \geq 0$  and  $\varepsilon > 0$ , one has

$$\Phi^1(L_{0,\cdot}) \stackrel{M}{=} \Phi^1(\varepsilon^{-H} L_{s,s+\varepsilon,\cdot}), \quad \Phi^2(Q_{0,\cdot}) \stackrel{M}{=} \Phi^2(\varepsilon^{-H} Q_{s,s+\varepsilon,\cdot}).$$

*Proof.* To ease notation, we write  $L_u = L_{0,u}$  and  $K_u = \varepsilon^{-H} L_{s,s+\varepsilon,u}$  for  $u \geq 0$ . It suffices to show that, for any integers  $l \geq 1$ ,  $k \geq 0$ , any  $f \in C_c^\infty(\mathbb{R}^l; \mathbb{R})$ , and any finite sequence  $0 < t_1 < t_2 < \dots < t_l \leq 1$ , the following two random variables

$$\left\| \sum_{1 \leq i_1, \dots, i_k \leq l} \partial_{x_{i_1} \dots x_{i_k}} f(L_{t_1}, \dots, L_{t_l}) \mathbf{D}_W L_{t_{i_1}} \otimes \dots \otimes \mathbf{D}_W L_{t_{i_k}} \right\|_{\mathcal{H}_W^{\otimes k}}$$

and

$$\left\| \sum_{1 \leq i_1, \dots, i_k \leq l} \partial_{x_{i_1} \dots x_{i_k}} f(K_{t_1}, \dots, K_{t_l}) \mathbf{D}_W K_{t_{i_1}} \otimes \dots \otimes \mathbf{D}_W K_{t_{i_k}} \right\|_{\mathcal{H}_W^{\otimes k}}$$

are equal in distribution. But this follows from the facts that both processes  $L_t$  and  $K_t$  are Wiener integrals with respect to the same Brownian motion  $(W_u)_{-\infty < u \leq T}$ , and that

$$(L_{t_1}, \dots, L_{t_l}) \stackrel{d}{=} (K_{t_1}, \dots, K_{t_l}).$$

The same argument applies to the processes  $(Q_{0,u})_{0 \leq u \leq 1}$  and  $(\varepsilon^{-H} Q_{s,s+\varepsilon,u})_{0 \leq u \leq 1}$ , which completes the proof of the lemma.  $\square$

## 2.4 Conditional small ball estimates

Integrability of the modulus of  $\theta$ -Hölder roughness plays an important role in the study of the existence of a smooth density. In our study, we need a conditional version of this property and in this subsection, we first derive some related conditional small ball estimates.

**Lemma 2.18.** *Let  $(B_t)_{t \geq 0}$  be a  $d$ -dimensional fBM with Hurst parameter  $H \in (0, 1)$ . Then there exists a positive constant  $C$  such that for any  $\delta > 0$ , any unit vector  $\phi \in \mathbb{R}^d$ , and any  $x \in (0, 1)$ , one has*

$$\mathbb{P} \left( \sup_{u,v \in [0, \delta]} |\langle \phi, B_{u,v} \rangle| \leq x \mid \mathcal{F}_0^W \right) \leq \exp(-C\delta x^{-1/H}),$$

where  $B_{u,v} := B_v - B_u$  denotes the increment of the fBM.

*Proof.* Firstly, by the scaling property of fBM, one has

$$\mathbb{P} \left( \sup_{u,v \in [0,\delta]} |\langle \phi, B_{u,v} \rangle| \leq x \mid \mathcal{F}_0^W \right) = \mathbb{P} \left( \sup_{u,v \in [0,1]} |\langle \phi, B_{u,v} \rangle| \leq \frac{x}{\delta^H} \mid \mathcal{F}_0^W \right).$$

Secondly, for any  $\phi \in \mathbb{R}^d$  with  $\|\phi\|_{\mathbb{R}^d} = 1$ , one notes that

$$(\langle \phi, B_t \rangle)_{t \geq 0} \stackrel{d}{=} (B_t^1)_{t \geq 0},$$

where  $(B_t^1)_{t \geq 0}$  denotes the first component  $B$ . Therefore, it suffices to establish the desired bound in the case  $\delta = 1$  and  $\phi = e_1$ , i.e.

$$\mathbb{P} \left( \sup_{u,v \in [0,1]} |B_{u,v}^1| \leq x \mid \mathcal{F}_0^W \right) \leq \exp(-Cx^{-1/H})$$

for some constant  $C > 0$ .

Fix  $x \in (0, 1)$ . We define a partition of the interval  $I = [0, 1]$  in the follow way. Let  $n := \lfloor x^{-1/H} \rfloor \geq 1$  and define the partition sequence  $0 = t_0 < t_1 < \dots < t_{n+1} = 1$  by

$$t_i := ix^{1/H}, \quad \text{for } i = 0, 1, \dots, n.$$

Since  $B_0 = 0$ , it is obvious that

$$\begin{aligned} \mathbb{P} \left( \sup_{u,v \in [0,1]} |B_{u,v}^1| \leq x \mid \mathcal{F}_0^W \right) &\leq \mathbb{P} \left( \sup_{u \in [0,1]} |B_u^1| \leq x \mid \mathcal{F}_0^W \right) \\ &\leq \mathbb{P} \left( \max_{i=1, \dots, n} |B_{t_i}^1| \leq x \mid \mathcal{F}_0^W \right) \end{aligned} \quad (2.9)$$

To estimate (2.9), we successively condition on the following filtrations in order:

$$\mathcal{F}_{t_{n-1}}^W, \mathcal{F}_{t_{n-2}}^W, \dots, \mathcal{F}_0^W.$$

According to Lemma 2.15, the distribution of  $B_{t_{i-1}, t_i}^1$  conditional on  $\mathcal{F}_{t_{i-1}}^W$  is equal to the distribution of  $(t_i - t_{i-1})^H B_{0,1}^1$  conditional on  $\mathcal{F}_0^W$ . Furthermore, one has  $B_{0,1}^1 = L_{0,1}^1 + Q_{0,1}^1$ . Here  $L_{0,1}^1, Q_{0,1}^1$  are Gaussian variables, where  $L_{0,1}^1$  is independent of  $\mathcal{F}_0^W$  and  $Q_{0,1}^1$  is measurable with respect to  $\mathcal{F}_0^W$ . If we denote  $\sigma^2$  as the variance of  $L_{0,1}^1$ , then one has

$$\begin{aligned} \mathbb{P} \left( |B_{t_{i-1}, t_i}^1| \leq x \mid \mathcal{F}_{t_{i-1}}^W \right) &\stackrel{M}{=} \mathbb{P} \left( |(t_i - t_{i-1})^H B_{0,1}^1| \leq x \mid \mathcal{F}_0^W \right) \\ &\leq \int_{-\frac{x}{\sigma(t_i - t_{i-1})^H}}^{\frac{x}{\sigma(t_i - t_{i-1})^H}} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy. \end{aligned}$$

Here we used the fact that for a Gaussian random variable with fixed variance, the probability  $\mathbb{P}(|Z| \leq x)$  is maximised when the mean is zero. Note that the right hand

side also serves as an upper bound for  $\mathbb{P}(|B_{t_i}^1| \leq x \mid \mathcal{F}_{t_{i-1}}^W)$ , since  $B_{t_i}^1$  can be viewed as the increment  $B_{t_{i-1}, t_i}^1$  shifted by  $B_{t_{i-1}}^1$ , and the bound depends only on the variance. By iterating the above estimate, one obtains that

$$\begin{aligned} & \mathbb{P}\left(\max_{i=1, \dots, n} |B_{t_i}^1| \leq x \mid \mathcal{F}_0^W\right) \\ & \leq \mathbb{P}\left(\max_{i=1, \dots, n-1} |B_{t_i}^1| \leq x \mid \mathcal{F}_0^W\right) \int_{-\frac{x}{\sigma(t_n - t_{n-1})^H}}^{\frac{x}{\sigma(t_n - t_{n-1})^H}} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\ & \dots \\ & \leq \prod_{i=1}^n \int_{-\frac{x}{\sigma(t_n - t_{n-1})^H}}^{\frac{x}{\sigma(t_n - t_{n-1})^H}} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy. \end{aligned}$$

Since  $\frac{x}{\sigma(t_n - t_{n-1})^H} \leq \frac{1}{\sigma}$  by the construction of the partition, one has

$$\mathbb{P}\left(\max_{i=1, \dots, n} |B_{t_{i-1}, t_i}^1| \leq x \mid \mathcal{F}_0^W\right) \leq e^{-Cn} \leq e^{-Cx^{-1/H}/2}, \quad (2.10)$$

where  $C := -\log \int_{-1/\sigma}^{1/\sigma} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$ . The desired estimate thus follows.  $\square$

**Proposition 2.19.** *Let  $(B_t)_{t \geq 0}$  be a  $d$ -dimensional fBM with Hurst parameter  $H \in (0, 1)$ . Then there exist positive constants  $C_1, C_2$ , such that one has*

$$\mathbb{P}\left(\inf_{\|\phi\|_{\mathbb{R}^d}=1} \sup_{u, v \in [s, s+\delta]} |\langle \phi, B_{u,v} \rangle| \leq x \mid \mathcal{F}_s^W\right) \leq C_1 \exp(-C_2 \delta x^{-1/H})$$

for all  $\delta > 0$ ,  $s \geq 0$  and  $x \in (0, 1)$ .

*Proof.* As in the proof of Lemma 2.18, it suffices to prove the desired bound in the case  $\delta = 1$  and  $s = 0$ , i.e.

$$\mathbb{P}\left(\inf_{\|\phi\|_{\mathbb{R}^d}=1} \sup_{u, v \in [0, 1]} |\langle \phi, B_{u,v} \rangle| \leq x \mid \mathcal{F}_0^W\right) \leq C_1 \exp(-C_2 x^{-1/H}).$$

Since  $B_0 = 0$ , one first has

$$\mathbb{P}\left(\inf_{\|\phi\|_{\mathbb{R}^d}=1} \sup_{u, v \in [0, 1]} |\langle \phi, B_{u,v} \rangle| \leq x \mid \mathcal{F}_0^W\right) \leq \mathbb{P}\left(\inf_{\|\phi\|_{\mathbb{R}^d}=1} \sup_{u \in [0, 1]} |\langle \phi, B_u \rangle| \leq x \mid \mathcal{F}_0^W\right).$$

The right hand side is what we aim to control in the following argument.

According to (2.8), the process  $(B_t)_{t \geq 0}$  can be decomposed as

$$B_t = L_{0,t} + Q_{0,t},$$

where  $(L_{0,t})_{t \geq 0}$  and  $(Q_{0,t})_{t \geq 0}$  are Gaussian processes, with  $(L_{0,t})_{t \geq 0}$  being independent of  $\mathcal{F}_0^W$  and  $(Q_{0,t})_{t \geq 0}$  being measurable with respect to  $\mathcal{F}_0^W$ . For any unit vector  $\phi \in \mathbb{R}^d$ , we define

$$\langle \phi, B_t \rangle = \langle \phi, L_{0,t} \rangle + \langle \phi, Q_{0,t} \rangle := L_t^\phi + Q_t^\phi.$$

It follows from [18, Equation (3.12)] that

$$\mathbb{P} \left( \sup_{u \in [0,1]} |\langle \phi, B_u \rangle| \leq x \mid \mathcal{F}_0^W \right) \leq C_1 \exp \left( -C_2 \sup_{u \in [0,1]} |Q_u^\phi| \right) \mathbb{P} \left( \sup_{u \in [0,1]} |L_u^\phi| \leq x \right),$$

where  $C_1, C_2 > 0$  do not depend on  $B$  and  $\phi$ . Due to the scaling property of  $L^\phi$ , the same argument leading to (2.10) shows that

$$\sup_{\|\phi\|_{\mathbb{R}^d}=1} \mathbb{P} \left( \sup_{u \in [0,1]} |L_u^\phi| \leq x \right) \leq \exp(-C_3 x^{-1/H})$$

for some  $C_3 > 0$ .

Combining the above estimates, it follows that

$$\mathbb{P} \left( \sup_{u \in [0,1]} |\langle \phi, B_u \rangle| \leq x \mid \mathcal{F}_0^W \right) \leq C_1 \exp \left( -C_2 \sup_{u \in [0,1]} |Q_u^\phi| - C_3 x^{-1/H} \right). \quad (2.11)$$

One can now apply the compactness argument as [18, Lemma 2], specifically the steps following Equation (3.13) there, to bound the right hand side of (2.11) uniformly over all unit vectors  $\phi$ . This yields the estimate

$$\mathbb{P} \left( \inf_{\|\phi\|_{\mathbb{R}^d}=1} \sup_{u \in [0,1]} |\langle \phi, B_u \rangle| \leq x \mid \mathcal{F}_0^W \right) \leq C_4 \exp(-C_5 x^{-1/H}),$$

which completes the proof of the result.  $\square$

### 3 Existence and smoothness of joint density

Our goal of this section is to prove the existence of a smooth joint density for the solution to the RDE (1.2) over multiple times under the hypoellipticity assumption. The main result is stated as follows.

**Theorem 3.1.** *Let  $(X_t)_{t \in [0,T]} = (X_t^1, \dots, X_t^d)_{t \in [0,T]}$  be a continuous Gaussian process with i.i.d. components. Assume that every component of  $X$  satisfies Condition 2.4 for some  $\rho \in [1, 2)$ , Condition 2.6 with index  $\alpha < 2/\rho$  and Condition 2.7. Consider the RDE*

$$d\mathbf{Y}_t = V(Y_t) d\mathbf{X}_t + V_0(Y_t) dt, \quad Y_0 = y_0 \in \mathbb{R}^N,$$

where  $\mathbf{X}$  is viewed as geometric rough path with roughness  $p > 2\rho$ . Suppose that the vector fields  $\{V_i\}_{i=1}^d$  satisfy Hypothesis 1.1 and 1.2. Then for any  $m \geq 1$  and  $0 < t_1 < \dots < t_m \leq T$ , the joint distribution of  $(Y_{t_1}, \dots, Y_{t_m})$  admits a smooth joint density with respect to the Lebesgue measure.

The rest of this section is devoted to the proof of Theorem 3.1.

As a standard route from the perspective of Malliavin's calculus, the crucial point is to investigate the non-degeneracy of the Malliavin covariance matrix which we first define.

**Definition 3.2.** Let  $F \in \mathbb{D}_X^{1,2}(\mathbb{R}^N)$  be a random vector in  $\mathbb{R}^N$ . The *Malliavin covariance matrix* (or simply *Malliavin matrix*) of  $F$  is the  $N \times N$  random matrix  $\gamma(F)$  whose  $(i, j)$ -entry is defined by

$$\gamma_{ij}(F) := \langle \mathbf{D}_X F^i, \mathbf{D}_X F^j \rangle_{\mathcal{H}_X}.$$

The Malliavin matrix of the solution  $Y_t$  is closely related to the Jacobian flow. Let us first recall that  $\mathbf{J}_t := \frac{\partial Y_t^x}{\partial x}$  ( $Y_t^x$  denotes the solution to (1.2) with initial value  $x$ ) is the unique solution to the linear equation

$$\mathbf{J}_t = \text{Id}_n + \int_0^t DV_0(Y_s^x) \mathbf{J}_s ds + \sum_{j=1}^d \int_0^t DV_j(Y_s^x) \mathbf{J}_s dX_s^j,$$

and that the following result holds true (see [10]).

**Proposition 3.3.** Let  $Y_t$  be the solution to equation (1.2) and suppose that the vector fields  $\{V_i\}_{i=1}^d$  satisfy Hypothesis 1.1. Then one has  $Y_t \in \mathbb{D}_X^\infty(\mathbb{R}^N)$  and

$$(\mathbf{D}_X^{(k)} Y_t)(s) = \mathbf{J}_{t,s} V_k(Y_s), \quad \text{for all } k = 1, \dots, d \text{ and } 0 \leq s \leq t,$$

where  $\mathbf{D}_X^{(k)}$  means the  $k$ -th component of  $\mathbf{D}_X$  and  $\mathbf{J}_{t,s} = \mathbf{J}_t \mathbf{J}_s^{-1}$ .

We are now in a position to introduce the Malliavin covariance matrix.

Given  $0 < t_1 < \dots < t_m \leq T$ , let  $F = (Y_{t_1}, \dots, Y_{t_m})$  being viewed as a  $N \times m$ -dimensional column vector. We write its Malliavin matrix as a block matrix

$$\gamma(F) = \begin{bmatrix} \langle \mathbf{D}_X Y_{t_1}, \mathbf{D}_X Y_{t_1} \rangle_{\mathcal{H}_X} & \dots & \langle \mathbf{D}_X Y_{t_1}, \mathbf{D}_X Y_{t_m} \rangle_{\mathcal{H}_X} \\ \vdots & \ddots & \vdots \\ \langle \mathbf{D}_X Y_{t_m}, \mathbf{D}_X Y_{t_1} \rangle_{\mathcal{H}_X} & \dots & \langle \mathbf{D}_X Y_{t_m}, \mathbf{D}_X Y_{t_m} \rangle_{\mathcal{H}_X} \end{bmatrix} \quad (3.1)$$

with blocks

$$\langle \mathbf{D}_X Y_{t_i}, \mathbf{D}_X Y_{t_j} \rangle_{\mathcal{H}_X} = \sum_{k=1}^d \int_{[0, t_i] \times [0, t_j]} \mathbf{J}_{t_i, u} V_k(Y_u) \otimes \mathbf{J}_{t_j, v} V_k(Y_v) d\phi_{u,v}.$$

According to Proposition 3.3,  $(\mathbf{D}_X^{(k)} Y_{t_i})(s) = \mathbf{J}_{t_i, s} V_k(Y_s) 1_{[0, t_i]}(s)$ , and the inner product within  $\mathcal{H}_X$  is understood as a  $2D$ -Young integral with respect to the covariance function  $R(u, v) := \mathbb{E}[X_u^i X_v^i]$ . The following result establishes the non-degeneracy of the Malliavin matrix  $\gamma(F)$ .

**Proposition 3.4.** *For any  $q > 1$ , there exists  $C_q > 0$  such that*

$$\mathbb{P} \left( \inf_{\|a\|=1} \langle a, \gamma(F)a \rangle < \varepsilon \right) \leq C_q \varepsilon^q$$

for all  $\varepsilon > 0$ .

*Proof.* We only consider the case  $m = 2$  (the general case only involves more notation). Let the random variable  $R_X$  be the corresponding quantity  $\mathcal{R}$  in the pathwise Norris' lemma (see [10, Theorem 5.6, Equation (5.8)]). According to Remark 2.10 and Cass–Litterer–Lyons [9],  $R_X$  possesses finite moments of all orders. Our proof is divided into the following four steps.

*Step 1: Lower bound for  $\langle a, \gamma(F)a \rangle$ .* We first derive a lower bound for  $\langle a, \gamma(F)a \rangle$  in terms of the supremum norm. Let  $a^* = (a_1^*, a_2^*)$ , where each  $a_i$  ( $i = 1, 2$ ) is a column vector in  $\mathbb{R}^N$  and the superscript  $*$  means transposition. By the definition of  $\gamma(F)$  in (3.1), it is readily checked that

$$\langle a, \gamma(F)a \rangle = \sum_{k=1}^d \Lambda^k, \quad \text{where} \quad \Lambda^k := \int_{[0,T]^2} f_u^k f_v^k dR(u, v),$$

with

$$f_s^k := \begin{cases} (a_1^* \mathbf{J}_{t_1,0} + a_2^* \mathbf{J}_{t_2,0}) \mathbf{J}_{0,s} V_k(Y_s), & 0 \leq s < t_1, \\ a_2^* \mathbf{J}_{t_2,t_1} \mathbf{J}_{t_1,s} V_k(Y_s), & t_1 \leq s \leq t_2. \end{cases}$$

To simplify notation, we write  $(\hat{a}_1^*, \hat{a}_2^*) := (a_1^* \mathbf{J}_{t_1,0} + a_2^* \mathbf{J}_{t_2,0}, a_2^* \mathbf{J}_{t_2,t_1})$  and one has

$$f_s^k = \begin{cases} \hat{a}_1^* \mathbf{J}_{0,s} V_k(Y_s), & 0 \leq s < t_1, \\ \hat{a}_2^* \mathbf{J}_{t_1,s} V_k(Y_s), & t_1 \leq s \leq t_2. \end{cases} \quad (3.2)$$

By applying [10, Corollary 6.10], one sees that  $\|f^k\|_{\infty, [t_{i-1}, t_i]}$  is upper bounded by the maximum of

$$2\text{Var}(X_{t_{i-1}, t_i} | \mathcal{F}_{0, t_{i-1}} \vee \mathcal{F}_{t_i, T})^{-1/2} \left( \int_{[0, T]^2} f_u^k f_v^k dR(u, v) \right)^{1/2}$$

and

$$C \left( \int_{[0, T]^2} f_u^k f_v^k dR(u, v) \right)^{\frac{\gamma}{2\gamma + \alpha}} \|f^k\|_{\gamma, [t_{i-1}, t_i]}^{\frac{\alpha}{2\gamma + \alpha}}$$

where  $C > 0$  is some constant and  $\gamma + 1/\rho > 1$ . In particular, to obtain an upper bound on  $\|f^k\|_{\infty, [t_{i-1}, t_i]}$  one needs to control the following three terms separately:

$$\|f^k\|_{\gamma, [t_{i-1}, t_i]}, \quad \Lambda^k, \quad \text{Var}(X_{t_{i-1}, t_i} | \mathcal{F}_{0, t_{i-1}} \vee \mathcal{F}_{t_i, T})^{-1}.$$

1. It is easy to see that  $\|f^k\|_{\gamma, [t_{i-1}, t_i]}$  is bounded by  $R_X$  according to the definition of  $R_X$ .

2.  $\Lambda^k$  is bounded above by  $\langle a, \gamma(F)a \rangle$ .
3.  $\text{Var}(X_{t_{i-1}, t_i} | \mathcal{F}_{0, t_{i-1}} \vee \mathcal{F}_{t_i, T})^{-1}$  is bounded above by a positive constant due to Condition 2.6.

As a result, there exist constants  $C, q_1, q_2 > 0$  such that

$$\|f^k\|_{\infty, [t_{i-1}, t_i]} \leq C \langle a, \gamma(F)a \rangle^{q_1} R_X^{q_2}.$$

Combining this with (3.2), it follows that

$$\|\hat{a}_i^* \mathbf{J}_{t_{i-1}, \cdot} V_k(Y)\|_{\infty, [t_{i-1}, t_i]} \leq C \langle a, \gamma(F)a \rangle^{q_1} R_X^{q_2} \quad (3.3)$$

for any  $i = 1, 2$  and  $k = 1, \dots, d$ .

*Step 2: Iterating the bound (3.3) to Lie brackets.* The next observation is that

$$\hat{a}_i^* \mathbf{J}_{t_{i-1}, t} V_I(Y_t) = \hat{a}_i^* V_I(Y_{t_{i-1}}) + \sum_{k=1}^d \int_{t_{i-1}}^t \hat{a}_i^* \mathbf{J}_{t_{i-1}, u} V_{(k, I)}(Y_u) dX_u^k + \int_{t_{i-1}}^t \hat{a}_i^* \mathbf{J}_{t_{i-1}, u} V_{(0, I)}(Y_u) du$$

for all  $I \in \mathcal{A}_1(\bar{l} - 1)$ ,  $i = 1, 2$  and  $t_{i-1} \leq t \leq t_i$ . This is easily obtained through the integration by parts formula. Together with Norris' lemma (see [10, Theorem 5.6]), the above formula implies that there exist constants  $C, q_1, q_2 > 0$ , such that

$$\|\hat{a}_i^* \mathbf{J}_{t_{i-1}, \cdot} V_{(k, I)}(Y)\|_{\infty, [t_{i-1}, t_i]} \leq C \|\hat{a}_i^* \mathbf{J}_{t_{i-1}, \cdot} V_I(Y)\|_{\infty, [t_{i-1}, t_i]}^{q_1} R_X^{q_2}$$

for any  $i = 1, 2$ ,  $k = 0, 1, \dots, d$  and  $I \in \mathcal{A}_1(\bar{l} - 1)$ . Starting from the base estimate (3.3), an induction argument shows that

$$\|\hat{a}_i^* \mathbf{J}_{t_{i-1}, \cdot} V_I(Y)\|_{\infty, [t_{i-1}, t_i]} \leq C \langle a, \gamma(F)a \rangle^{q_1} R_X^{q_2} \quad (3.4)$$

for any  $i = 1, 2$  and  $I \in \mathcal{A}_1(\bar{l})$ .

*Step 3: Nondegeneracy of  $\langle a, \gamma(F)a \rangle R_X$ .* We are going to apply Hypothesis 1.2 to find a constant lower bound for (3.4). Indeed, if one evaluates  $\hat{a}_i^* \mathbf{J}_{t_{i-1}, \cdot} V_I(Y)$  at the starting points of every interval  $[t_{i-1}, t_i]$ , the estimate (3.4) implies that there are constants  $C, q_1, q_2 > 0$  such that

$$\sup_{I \in \mathcal{A}_1(\bar{l})} |\hat{a}_1^* V_I(y_0)|^2 + \sup_{I \in \mathcal{A}_1(\bar{l})} |\hat{a}_2^* V_I(Y_{t_1})|^2 \leq C \langle a, \gamma(F)a \rangle^{q_1} R_X^{q_2}.$$

According to Hypothesis 1.2, one has

$$\sup_{I \in \mathcal{A}_1(\bar{l})} |\hat{a}_1^* V_I(y_0)|^2 + \sup_{I \in \mathcal{A}_1(\bar{l})} |\hat{a}_2^* V_I(Y_{t_1})|^2 \geq \lambda (\|\hat{a}_1\|_{\mathbb{R}^N}^2 + \|\hat{a}_2\|_{\mathbb{R}^N}^2).$$

Recalling  $(\hat{a}_1^*, \hat{a}_2^*) = (a_1^* \mathbf{J}_{t_1, 0} + a_2^* \mathbf{J}_{t_2, 0}, a_2^* \mathbf{J}_{t_2, t_1})$ , one can write

$$\begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{J}_{0, t_1} & 0 \\ -I & \mathbf{J}_{t_1, t_2} \end{bmatrix}^* \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} := \mathbf{Q}_{t_1, t_2} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix},$$

and one has

$$1 = \|a_1\|_{\mathbb{R}^N}^2 + \|a_2\|_{\mathbb{R}^N}^2 \leq \|\mathbf{Q}_{t_1, t_2}\|_{\mathbb{R}^{N \times N}}^2 (\|\hat{a}_1\|_{\mathbb{R}^N}^2 + \|\hat{a}_2\|_{\mathbb{R}^N}^2)$$

Assembling together the above estimates and taking infimum on both sides, one concludes that

$$1 \leq \lambda^{-1} C \|\mathbf{Q}_{t_1, t_2}\|_{\mathbb{R}^{N \times N}}^2 R_X^{q_1} \inf_{\|a\|=1} \langle a, \gamma(F)a \rangle^{q_2} \leq C_1 R_X^{q_3} \inf_{\|a\|=1} \langle a, \gamma(F)a \rangle^{q_4} \quad (3.5)$$

for some  $C_1, q_1, q_2, q_3, q_4 > 0$ , where we also used the fact that  $\|\mathbf{Q}_{t_1, t_2}\|_{\mathbb{R}^{N \times N}} \leq C' R_X^p$  for some  $p \geq 1$ .

*Step 4: Integrability of  $\gamma(F)^{-1}$ .* Let  $q > 1$  be given fixed. As a consequence of (3.5), one obtains that

$$\mathbb{P} \left( \inf_{\|a\|=1} \langle a, \gamma(F)a \rangle < \varepsilon \right) \leq \mathbb{P}(1 \leq C_1 R_X^{q_3} \varepsilon^{q_4}) \leq C_2 \varepsilon^q \mathbb{E}[R_X^{q_5}]$$

for suitable constants  $C_2, q_5 > 0$ . The desired estimate thus follows.  $\square$

*Proof of Theorem 3.1.* Proposition 3.4 implies that the Malliavin covariance matrix of  $(Y_{t_1}, \dots, Y_{t_m})$  has inverse moments of all orders. The desired result thus follows from standard regularity theorems from Malliavin's calculus (see [28]).  $\square$

## 4 Local upper estimate for the joint density of $(Y_s, Y_t)$

In this section, we prove Theorem 1.5, which provides a precise local upper bound for the joint density of  $(Y_s, Y_t)$  and will be essential for establishing capacity estimates for hitting probabilities in the next section.

Our approach is inspired by Kusuoka-Stroock's technique in [22]. The main idea is to approximate the actual solution by its stochastic Taylor expansion, whose nondegeneracy can be obtained from the nondegeneracy of the truncated signature process of  $X$  together with the hypoellipticity of the vector fields  $\mathcal{V}$ .

### 4.1 Free nilpotent Lie groups and Lie algebras

To study the truncated signature process, we shall first recall some basic notions on free nilpotent Lie groups.

The truncated tensor algebra of order  $l$  over  $\mathbb{R}^d$  is denoted by  $T^{(l)}$ . Under addition and tensor product,  $(T^{(l)}, +, \otimes)$  is an associative algebra. The set of homogeneous Lie polynomials of degree  $k$  is denoted as  $\mathcal{L}_k$ , and the free nilpotent Lie algebra of order  $l$  is denoted as  $\mathfrak{g}^{(l)}$ . The exponential map on  $T^{(l)}$  is defined by

$$\exp(a) := \sum_{k=0}^{\infty} \frac{1}{k!} a^{\otimes k} \in T^{(l)}, \quad (4.1)$$

where the sum is indeed finite and hence well-defined.  $T^{(l)}$  is equipped with a natural inner product induced by the Euclidean structure on  $\mathbb{R}^d$  and the Hilbert-Schmidt tensor norm on each of the tensor products. To be more precise, the inner product on  $(\mathbb{R}^d)^{\otimes k}$  is induced by

$$\langle v_1 \otimes \cdots \otimes v_k, w_1 \otimes \cdots \otimes w_k \rangle_{(\mathbb{R}^d)^{\otimes k}} := \langle v_1, w_1 \rangle_{\mathbb{R}^d} \cdots \langle v_k, w_k \rangle_{\mathbb{R}^d}.$$

The components  $(\mathbb{R}^d)^{\otimes k}$  and  $(\mathbb{R}^d)^{\otimes m}$  ( $k \neq m$ ) are assumed to be orthogonal. The Hilbert-Schmidt tensor norm is denoted as  $\|\cdot\|_{\text{HS}}$ .

The following algebraic structure plays a fundamental role in rough path theory.

**Definition 4.1.** The free nilpotent Lie group  $G^{(l)}$  of order  $l$  is defined by

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

Note that the exponential map is a diffeomorphism under which  $\mathfrak{g}^{(l)}$  is the Lie algebra of  $G^{(l)}$ . We define the operation  $\star$  to be the multiplication on  $\mathfrak{g}^{(l)}$  induced from  $G^{(l)}$  through the exponential map, namely

$$v \star u := \log(\exp(v) \otimes \exp(u)), \quad v, u \in \mathfrak{g}^{(l)}.$$

It will be useful in the sequel to have some basis elements available for the algebras introduced above. Let  $\mathcal{A}(l)$  (respectively,  $\mathcal{A}_1(l)$ ) denote the set of words (respectively, non-empty words) over  $\{1, \dots, d\}$  of length at most  $l$ . We set  $e_\emptyset := 1$ , and for each word  $I = (i_1, \dots, i_r) \in \mathcal{A}_1(l)$ , we define

$$e_{(I)} := e_{i_1} \otimes \cdots \otimes e_{i_r}, \quad \text{and} \quad e_I := [e_{i_1}, \cdots [e_{i_{r-2}}, [e_{i_{r-1}}, e_{i_r}]] \cdots],$$

where  $\{e_1, \dots, e_d\}$  denotes the canonical basis of  $\mathbb{R}^d$ . Then  $\{e_{(I)} : I \in \mathcal{A}(l)\}$  forms an orthonormal basis of  $T^{(l)}$  under the Hilbert-Schmidt tensor norm. In addition, one also has

$$\mathcal{L}_k = \text{Span}\{e_I : |I| = k\}, \quad \mathfrak{g}^{(l)} = \text{Span}\{e_I : I \in \mathcal{A}_1(l)\}.$$

As a closed subspace,  $\mathfrak{g}^{(l)}$  inherits 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 previously defined product  $\star$ . In addition, for each  $\lambda > 0$ , there is a dilation operation  $\delta_\lambda : T^{(l)} \rightarrow T^{(l)}$  induced by  $\delta_\lambda(a) \triangleq \lambda^k a$  if  $a \in (\mathbb{R}^d)^{\otimes k}$ , which satisfies the relation  $\delta_\lambda \circ \exp = \exp \circ \delta_\lambda$ . One can easily show that

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

The number  $\nu$  is known as the *homogeneous dimension* of  $G^{(l)}$ .

We always fix the Euclidean norm on  $\mathbb{R}^d$  in the remainder of the article. As far as the free nilpotent group  $G^{(l)}$  is concerned, there are several useful metric structures.

Among them we will use the extrinsic Hilbert-Schmidt metric  $\rho_{\text{HS}}$  induced from  $T^{(l)}$ , namely we set

$$\rho_{\text{HS}}(g_1, g_2) := \|g_2 - g_1\|_{\text{HS}}, \quad g_1, g_2 \in G^{(l)}.$$

Another important metric which we will use is the intrinsic *Carnot-Carathéodory metric* (or just *CC metric*). For a bounded variation (BV) path  $w: [0, 1] \rightarrow \mathbb{R}^d$ , its *truncated signature* up to level  $l$  is defined as

$$\mathbf{w}_{0,1}^{(l)} := \sum_{k=0}^l \int_{0 < t_1 < \dots < t_k < 1} dw_{t_1} \otimes \dots \otimes dw_{t_k} \in G^{(l)}.$$

It is known that every element in  $G^{(l)}$  is the truncated signature of some BV path. The CC-norm of  $g \in G^{(l)}$ , denoted as  $\|g\|_{\text{CC}}$ , is the minimal length of all those BV paths whose truncated signature is  $g$ . Given  $g_1, g_2 \in G^{(l)}$ , their CC-distance is defined by

$$\rho_{\text{CC}}(g_1, g_2) := \|g_1^{-1} \otimes g_2\|_{\text{CC}}.$$

For  $u \in \mathfrak{g}^{(l)}$ , we set  $\|u\|_{\text{CC}} := \|\exp(u)\|_{\text{CC}}$ .

## 4.2 Truncated logarithmic signatures of fractional Brownian motion

To define the truncated signature path of fBM, we need some definitions from rough path theory. Recall that the order- $l$  *truncated signature path* of a BV path  $w: [0, T] \rightarrow \mathbb{R}^d$  is the functional defined by

$$(s, t) \mapsto \mathbf{w}_{s,t} = \sum_{k=0}^l \int_{s < t_1 < \dots < t_k < t} dw_{t_1} \otimes \dots \otimes dw_{t_k} \in G^{(l)}.$$

It is a *multiplicative functional* in the sense that  $\mathbf{w}_{s,u} \otimes \mathbf{w}_{u,t} = \mathbf{w}_{s,t}$  for any  $s < u < t$ .

**Definition 4.2.** Let  $\alpha \in (0, 1)$ . A multiplicative functional  $\mathbf{z}: \{(s, t) \in [0, T]^2 : s \leq t\} \rightarrow T^{([1/\alpha])}$  is called a *weakly geometric  $\alpha$ -Hölder rough path* if it takes values in  $G^{([1/\alpha])}$  and satisfies

$$\|\mathbf{z}\|_{\alpha; \text{HS}} := \sum_{i=1}^{[1/\alpha]} \sup_{0 \leq s < t \leq 1} \frac{\|z_{s,t}^i\|_{\text{HS}}}{|t - s|^{i\alpha}} < \infty.$$

It is a basic result in rough path theory (Lyons' extension theorem) that given a weakly geometric rough path  $\mathbf{z}$ , its truncated signature paths of all orders are well-defined (see [24]).

Now let  $(X_t)_{0 \leq t \leq T}$  be a  $d$ -dimensional fBM with Hurst parameter  $H \in (1/4, 1)$  (viewed as a  $\gamma$ -Hölder weakly geometric rough path for  $\gamma < H$ ; see Theorem 2.2). Let  $l \geq 1$  be given fixed. We use  $\mathbf{X}_{s,t}^{(l)}$  to denote its truncated signature path of order  $l$  and also set

$$\mathbf{U}_{s,t} := \exp^{-1}(\mathbf{X}_{s,t}^{(l)}) \in \mathfrak{g}^{(l)}.$$

The process  $\mathbf{U}_{s,t}$  is known as the *truncated logarithmic signature path* of order  $l$ . It is well-known that  $\mathbf{U}_{s,t}$  satisfies a hypoelliptic RDE on the Lie algebra (see [15, Remark 7.43])  $\mathfrak{g}^{(l)}$  and thus admits a smooth density with respect to the Lebesgue measure  $du$  on  $\mathfrak{g}^{(l)}$  (equivalently, the Haar measure on  $G^{(l)}$ ).

The aim of this section is to establish the following upper bound for the conditional density of the logarithmic signature  $\mathbf{U}_{s,t}$ .

**Proposition 4.3.** *Let*

$$\rho_{s,t}(u) := \mathbb{P}(\mathbf{U}_{s,t} \in du \mid \mathcal{F}_s^W), \quad u \in \mathfrak{g}^{(l)}(\mathbb{R}^d)$$

be the conditional density of  $\mathbf{U}_{s,t}$  given  $\mathcal{F}_s^W$ . There exist a constant  $C > 0$ , depending only on  $d$  and  $H$ , and a family of positive random variables  $\{\Psi_{s,t}\}_{0 \leq s < t}$  such that

$$\rho_{s,t}(u) \leq \frac{C}{(t-s)^{H\nu}} \exp\left(-\frac{\|u\|_{\text{CC}}^2}{C(t-s)^{2H}}\right) \Psi_{s,t},$$

for all  $0 \leq s < t \leq T$ , where the Malliavin-Sobolev norm of  $\Psi_{s,t}$  does not depend on  $s$  or  $t$ , and  $\nu = \sum_{k=1}^l k \dim(\mathcal{L}_k)$  is the homogeneous dimension of  $\mathfrak{g}^{(l)}(\mathbb{R}^d)$ .

The starting point for proving Proposition 4.3 is the following inequality:

$$\mathbb{P}(\mathbf{U}_{s,t} \in du \mid \mathcal{F}_s^W) \leq C_n \mathbb{P}(\mathbf{U}_{s,t} \geq u \mid \mathcal{F}_s^W)^{1/2} \|\gamma(\mathbf{U}_{s,t})^{-1}\|_{n,2^{n+2},s}^n \|\mathbf{D}_W \mathbf{U}_{s,t}\|_{n,2^{n+2},s}^n. \quad (4.2)$$

This is obtained in the same way as [6, Inequality (24)]. Here  $n$  is arbitrary and  $C_n$  is a positive constant depending on  $n$ . The event  $\{\mathbf{U}_{s,t} \geq u\}$  is understood coordinate-wisely with respect to a given basis of  $\mathfrak{g}^{(l)}$ . Our task is now reduced to estimating each individual term on the right hand side of (4.2).

#### 4.2.1 Estimate of $\mathbb{P}(\mathbf{U}_{s,t} \geq u \mid \mathcal{F}_s^W)$

Let  $\mathbf{X}$  denote the canonical lifting of  $X$  as a (random) geometric rough path. We first define its Besov norm which is served as a smooth upper bound for the Hölder norm.

**Definition 4.4.** Let  $\gamma < H$  and  $q \geq 1$  be given fixed. The  $(\gamma, q)$ -Besov norm of  $\mathbf{X}$  over  $[s, t]$  is defined to be

$$\mathcal{N}_{\gamma,q}^{s,t}(\mathbf{X}) := \int_s^t \int_s^t \frac{D(\mathbf{X}_{u,v}, \mathbf{X}_v)^q}{|v-u|^{\gamma q}} du dv, \quad 0 \leq s < t \leq T,$$

where

$$D(\mathbf{X}_{u,v}) := \sum_{k=1}^{\lfloor 1/H \rfloor} \left\| \int_{u \leq s_1 \leq \dots \leq s_k \leq v} dX_{s_1} \otimes \dots \otimes dX_{s_k} \right\|_{\text{HS}}^{1/k}.$$

A benefit of working with these Besov norms is that they are smooth in the sense of Malliavin. In addition, standard Fernique-type estimate shows that there exists a constant  $\eta_0 > 0$  such that

$$\mathbb{E} \left[ \exp \left( \eta_0 \left( \mathcal{N}_{\gamma, 2q}^{0,1}(\mathbf{X}) \right)^{1/q} \right) \right] < \infty. \quad (4.3)$$

By using the Garsia-Rodemich-Rumsey inequality in Carnot groups, for any  $\varepsilon \in (0, H - \gamma)$  one has the following estimate

$$\|\mathbf{X}\|_{\gamma\text{-Hö}; [s,t]} \leq C \left( \mathcal{N}_{\gamma+\varepsilon, 2q}^{s,t}(\mathbf{X}) \right)^{1/2q}, \quad 0 \leq s \leq t \leq T. \quad (4.4)$$

See [15] for a discussion of these facts.

The goal of this part is to prove the following lemma.

**Lemma 4.5.** *There exist a constant  $C > 0$  and a family of positive random variables  $(\Psi_{s,t}^1)_{0 \leq s < t}$ , where the Malliavin-Sobolev norm of each  $\Psi_{s,t}^1$  does not depend on  $s$  or  $t$ , such that*

$$\mathbb{P}(\mathbf{U}_{s,t} \geq u \mid \mathcal{F}_s^W) \leq \exp \left( -\frac{\|u\|_{\text{CC}}^2}{C(t-s)^{2H}} \right) \Psi_{s,t}^1 \quad (4.5)$$

for all  $u \in \mathfrak{g}^{(l)}$  and  $0 \leq s < t \leq T$ .

*Proof.* For any given  $\varepsilon > 0, s \geq 0$ , by using the decomposition (2.8) the conditional probability  $\mathbb{P}(\mathbf{U}_{s,s+\varepsilon} \geq u \mid \mathcal{F}_s^W)$  can be viewed as a functional of  $(Q_{s,s+\varepsilon u})_{0 \leq u \leq 1}$ . According to Lemma 2.15, there exists a functional  $F : C([0, 1]) \rightarrow \mathbb{R}$  such that

$$\mathbb{P}(\mathbf{U}_{s,s+\varepsilon} \geq u \mid \mathcal{F}_s^W) = F(Q_{s,s+\varepsilon}) \quad \text{and} \quad \mathbb{P}(\delta_{\varepsilon^H} \mathbf{U}_{0,1} \geq u \mid \mathcal{F}_0^W) = F(\varepsilon^H Q_{0,\cdot})$$

By applying Lemma 2.17, one has

$$\mathbb{P}(\mathbf{U}_{s,t} \geq u \mid \mathcal{F}_s^W) \stackrel{M}{=} \mathbb{P}(\delta_{(t-s)^H} \mathbf{U}_{0,1} \geq u \mid \mathcal{F}_0^W)$$

for any  $0 \leq s < t$ , where the equivalence notion  $\stackrel{M}{=}$  is introduced in Definition 2.16.

Now it suffices to show

$$\mathbb{P}(\delta_{(t-s)^H} \mathbf{U}_{0,1} \geq u \mid \mathcal{F}_0^W) \leq \exp \left( -\frac{\|u\|_{\text{CC}}^2}{C(t-s)^{2H}} \right) \Psi^1 \quad (4.6)$$

for certain random variable  $\Psi^1$  with bounded Malliavin-Sobolev norms. We assume without loss of generality that each coordinate  $u^\alpha \geq 0$ . An argument justifying this reduction can be found in the proof of [4, Theorem 3.13]. We also assume that  $\|u\|_{\text{CC}} > 0$ , for otherwise there is nothing to prove.

Let us define the homogeneous norm on  $\mathfrak{g}^{(l)}$  by  $\|u\| \triangleq \max_\alpha |u^\alpha|^{1/q_\alpha}$ , where  $q_\alpha = k$  if  $u^\alpha$  is the coordinate of  $u$  along a basis element in  $\mathcal{L}_k$ . For any random variable  $\Xi : \Omega \rightarrow \mathfrak{g}^{(l)}$ , one has

$$\{\Xi \geq u\} = \left\{ \delta_{\|u\|_{\text{CC}}^{-1}} \Xi \geq \delta_{\|u\|_{\text{CC}}^{-1}} u \right\} \subseteq \left\{ \|\delta_{\|u\|_{\text{CC}}^{-1}} \Xi\| \geq \|\delta_{\|u\|_{\text{CC}}^{-1}} u\| \right\}.$$

The above inclusion uses the assumption  $u^\alpha \geq 0$  for all  $\alpha$ . Since all homogeneous norms on  $\mathfrak{g}^{(l)}$  are equivalent, there are positive constants  $C_1, C_2$  such that

$$C_1 \|\cdot\|_{\text{CC}} \leq \|\cdot\| \leq C_2 \|\cdot\|_{\text{CC}} \quad \text{on } \mathfrak{g}^{(l)}.$$

In particular,  $\|\delta_{\|u\|_{\text{CC}}^{-1}} u\| \geq C_1$ . It follows that

$$\{\Xi \geq u\} \subseteq \left\{ \|\delta_{\|u\|_{\text{CC}}^{-1}} \Xi\| \geq C_1 \right\} = \left\{ \|u\|_{\text{CC}}^{-1} \|\Xi\| \geq C_1 \right\} \subseteq \left\{ C_2 \|\Xi\|_{\text{CC}} \geq C_1 \|u\|_{\text{CC}} \right\}. \quad (4.7)$$

Now let us apply (4.7) to  $\Xi := \delta_{(t-s)^H} \mathbf{U}_{0,1}$ . One finds that

$$\begin{aligned} \mathbb{P}(\delta_{(t-s)^H} \mathbf{U}_{0,1} \geq u \mid \mathcal{F}_0^W) &\leq \mathbb{P}\left(\|\delta_{(t-s)^H} \mathbf{U}_{0,1}\|_{\text{CC}} \geq \frac{C_1}{C_2} \|u\|_{\text{CC}} \mid \mathcal{F}_0^W\right) \\ &= \mathbb{P}\left(\|\mathbf{U}_{0,1}\|_{\text{CC}} \geq \frac{C_1 \|u\|_{\text{CC}}}{C_2 (t-s)^H} \mid \mathcal{F}_0^W\right). \end{aligned}$$

Recall from our convention that  $\|\mathbf{U}_{0,1}\|_{\text{CC}} = \|\mathbf{X}_{0,1}\|_{\text{CC}}$ . By the Chebyshev inequality and the Besov-type bound (4.4), one concludes that

$$\begin{aligned} &\mathbb{P}\left(\|\mathbf{U}_{0,1}\|_{\text{CC}} \geq \frac{C_1 \|u\|_{\text{CC}}}{C_2 (t-s)^H} \mid \mathcal{F}_0^W\right) \\ &\leq C \exp\left(-\frac{\|u\|_{\text{CC}}^2}{C (t-s)^{2H}}\right) \mathbb{E}\left[\exp\left(\eta (\mathcal{N}_{\gamma, 2q}^{0,1}(\mathbf{B}))^{1/q}\right) \mid \mathcal{F}_0^W\right], \end{aligned}$$

for some  $C > 0$ ,  $q \geq 1$ ,  $\eta \in (0, \eta_0]$  and  $\gamma < H$ . The estimate (4.6) thus follows by taking

$$\Psi^1 = \mathbb{E}\left[\exp\left(\eta (\mathcal{N}_{\gamma, 2q}^{0,1}(\mathbf{B}))^{1/q}\right) \mid \mathcal{F}_0^W\right]. \quad (4.8)$$

To complete the proof of the lemma, observe that  $\Psi^1$  is a functional of the process  $(Q_{0,u})_{0 \leq u \leq 1}$ , i.e.  $\Psi^1 = F(Q_{0,\cdot})$  with some path functional  $F : C([0, 1]) \rightarrow \mathbb{R}$ . We then set

$$\Psi_{s,t}^1 := F\left((t-s)^{-H} Q_{s, s+(t-s)}\right). \quad (4.9)$$

According to Lemma 2.17,  $\Psi^1$  and  $\Psi_{s,t}^1$  have the same Malliavin-Sobolev norms for all  $s, t \in [0, T]$ . The desired estimate (4.5) thus follows.  $\square$

#### 4.2.2 Estimate of $\|\mathbf{D}_W \mathbf{U}_{s,t}\|_{k,p,s}$

Our next step is to estimate  $\|\mathbf{D}_W \mathbf{U}_{s,t}\|_{k,p,s}$  and the main result is stated as follows.

**Lemma 4.6.** *For any integer  $k \geq 1$  and any real number  $p > 0$ , there exist a family of random variables  $(\Psi_{s,t}^2)_{0 \leq s < t}$  (depending on  $k$  and  $p$ ) whose Malliavin-Sobolev norms do not depend on  $s$  or  $t$ , such that*

$$\|\mathbf{D}_W \mathbf{U}_{s,t}\|_{k,p,s} \leq (t-s)^H \Psi_{s,t}^2$$

for all  $0 \leq s < t \leq T$ .

Recall that  $\mathbf{U}_{s,t} = \exp^{-1}(\mathbf{X}_{s,t}^{(l)})$ . In particular, each component of  $\mathbf{U}_{s,t}$  is a polynomial of components of  $\mathbf{B}_{s,t}$ . As a result, it suffices to estimate the components of  $\mathbf{B}_{s,t}$ . Note that the  $m$ -th level component of  $\mathbf{B}_{s,t}$  is

$$I_{s,t}^m := \int_{s < t_1 < \dots < t_m < t} dB_{t_1} \otimes \dots \otimes dB_{t_m}$$

We first need the following lemma.

**Lemma 4.7.** *For any given  $0 \leq s < t \leq T$ , one has*

$$\|I_{s,t}^m\|_{k,p,s} \stackrel{M}{=} (t-s)^{mH} \|I_{0,1}^m\|_{k,p,0}.$$

*Proof.* Clearly, it is equivalent to proving that

$$\left( \mathbb{E} \left[ \|\mathbf{D}_W^k I_{s,s+\varepsilon}^m\|_{(\mathcal{H}_{W,s})^{\otimes k}}^p \middle| \mathcal{F}_s^W \right] \right)^{1/p} \stackrel{M}{=} \varepsilon^{Hm} \left( \mathbb{E} \left[ \|\mathbf{D}_W^k I_{0,1}^m\|_{(\mathcal{H}_{W,0})^{\otimes k}}^p \middle| \mathcal{F}_0^W \right] \right)^{1/p}. \quad (4.10)$$

for any  $s \geq 0$  and  $\varepsilon > 0$ . To this end, let us write

$$\begin{aligned} \Phi^1(\varepsilon^{-H} Q_{s,s+\varepsilon}) &= \varepsilon^{-Hmp} \mathbb{E} \left[ \|\mathbf{D}_W^k I_{s,s+\varepsilon}^m\|_{(\mathcal{H}_{W,s})^{\otimes k}}^p \middle| \mathcal{F}_s^W \right] \\ \Phi^2(Q_{0,\cdot}) &= \mathbb{E} \left[ \|\mathbf{D}_W^k I_{0,1}^m\|_{(\mathcal{H}_{W,0})^{\otimes k}}^p \middle| \mathcal{F}_0^W \right] \end{aligned}$$

with two path functionals  $\Phi^1, \Phi^2 : C([0,1]) \rightarrow \mathbb{R}$ , where  $Q_t, L_t$  are defined in (2.8). We prove (4.10) by showing

$$\Phi^1(\varepsilon^{-H} Q_{s,s+\varepsilon}) \stackrel{M}{=} \Phi^2(Q_{0,\cdot}). \quad (4.11)$$

Indeed, according to Lemma 2.17, it suffices to show that  $\Phi^1 = \Phi^2$ . This is straightforward, since for any  $\eta \in C([0,1])$ ,

$$\begin{aligned} \Phi^1(\eta) &= \mathbb{E} \left[ \|\mathbf{D}_W^k (\varepsilon^{-Hm} I_{s,s+\varepsilon}^m)\|_{(\mathcal{H}_{W,s})^{\otimes k}}^p \middle| \varepsilon^{-H} Q_{s,s+\varepsilon} = \eta \right] \\ &= \mathbb{E} \left[ \|\mathbf{D}_W^k I_{0,1}^m\|_{(\mathcal{H}_{W,0})^{\otimes k}}^p \middle| Q_{0,\cdot} = \eta \right] = \Phi^2(\eta). \end{aligned} \quad (4.12)$$

The second equality holds because conditional on  $\{Q_{0,\cdot} = \eta\}$  and  $\{\varepsilon^{-H} Q_{s,s+\varepsilon} = \eta\}$ , there exists a functional  $F_\eta : C([0,1]) \rightarrow \mathbb{R}$  such that

$$F_\eta(\varepsilon^{-H} L_{s,s+\varepsilon}) = \varepsilon^{-Hm} I_{s,s+\varepsilon}^m, \quad \text{and} \quad F_\eta(L_{0,\cdot}) = I_{0,1}^m.$$

Since  $(Q_u)_{u \geq 0}$  and  $(L_u)_{u \geq 0}$  are independent, Lemma 2.17 implies that

$$\varepsilon^{-Hm} I_{s,s+\varepsilon}^m = F_\eta(\varepsilon^{-H} L_{s,s+\varepsilon}) \stackrel{M}{=} F_\eta(L_{0,\cdot}) = I_{0,1}^m$$

on the event  $\{Q_{0,\cdot} = \eta\} \cap \{\varepsilon^{-H} Q_{s,s+\varepsilon} = \eta\}$ . This completes the proof of (4.12).  $\square$

Now we give the proof of Lemma 4.6.

*Proof of Lemma 4.6.* It suffices to prove that for any fixed  $k, p$ , there exist a family of random variables  $(\Psi_{s,t}^2)_{0 \leq s < t}$  whose Malliavin-Sobolev norms do not depend on  $s$  or  $t$ , such that

$$\|I_{s,t}^m\|_{k,p,s} \leq (t-s)^{mH} \Psi_{s,t}^2.$$

To this end, we adopt a similar idea as in [21]. Let  $\tilde{W}$  be an independent copy of  $W$ , and we use  $\tilde{\mathbb{E}}$  and  $\mathcal{F}_s^{\tilde{W}}$  to represent the expectation and filtration with respect to  $\tilde{W}$ . Define  $\tilde{B}$  by

$$\tilde{B}_t^i = \int_{-\infty}^t k_H(t, u) d\tilde{W}_u^i, \quad 0 \leq t \leq T, \quad 1 \leq i \leq d.$$

We also set

$$\Theta_1 = \sum_{i=1}^m \int_{0 < t_1 < \dots < t_m < 1} \dots dB_{t_{i-1}} \otimes d\tilde{B}_{t_i}^i \otimes dB_{t_{i+1}} \dots$$

To simplify notation, we denote  $I^m := I_{0,1}^m$ . For any  $h \in \mathcal{H}_B$ , one has

$$(\mathbf{D}_B^1 I^m)(h) = \sum_{i=1}^m \int_{0 < t_1 < \dots < t_m < 1} \dots dB_{t_{i-1}} \otimes d\bar{h}_{t_i} \otimes dB_{t_{i+1}} \dots = (\mathbf{D}_{\tilde{B}}^1 \Theta_1)(h),$$

where  $\bar{h}$  is the corresponding element of  $h$  in the intrinsic Cameron-Martin space  $\tilde{\mathcal{H}}_B$ . Therefore,  $\mathbf{D}_B^1 I^m = \mathbf{D}_{\tilde{B}}^1 \Theta_1$ . It follows from Proposition 2.12 that

$$\mathbf{D}_W^1 I^m = \mathcal{L}(\mathbf{D}_B^1 I^m) = \mathcal{L}(\mathbf{D}_{\tilde{B}}^1 \Theta_1) = \mathbf{D}_{\tilde{W}}^1 \Theta_1$$

holds almost surely as a functional from  $\Omega \rightarrow L^2([0, 1])$ .

Since all  $\mathbb{D}_W^{k,2}$ -norms are equivalent on a fixed inhomogeneous Wiener chaos, one has

$$\|\mathbf{D}_W^1 I^m\|_{\mathcal{H}_{W,0}} = \|\mathbf{D}_{\tilde{W}}^1 \Theta_1\|_{\mathcal{H}_{W,0}} \lesssim \left( \tilde{\mathbb{E}} \left[ \Theta_1^2 \middle| \mathcal{F}_0^{\tilde{W}} \right] \right)^{1/2}.$$

In addition, by Hölder's inequality, one has

$$\left( \mathbb{E} \left[ \|\mathbf{D}_W^1 I^m\|_{\mathcal{H}_{W,0}}^p \middle| \mathcal{F}_0^W \right] \right)^{1/p} \leq \left( \mathbb{E} \tilde{\mathbb{E}} \left[ \Theta_1^p \middle| \mathcal{F}_0^W \vee \mathcal{F}_0^{\tilde{W}} \right] \right)^{1/p}$$

for all  $p \geq 2$ . As in Inahama [21], one can similarly define  $\Theta_k$  for  $1 \leq k \leq m$  and show that

$$\left( \mathbb{E} \left[ \|\mathbf{D}_W^k I^m\|_{\mathcal{H}_{W,0}}^p \middle| \mathcal{F}_0^W \right] \right)^{1/p} \leq \left( \mathbb{E} \tilde{\mathbb{E}} \left[ \Theta_k^p \middle| \mathcal{F}_0^W \vee \mathcal{F}_0^{\tilde{W}} \right] \right)^{1/p}$$

for all  $p \geq 2$ . It follows that for any  $k \geq 1, p \geq 2$ , there exists a constant  $C_{k,p}$  such that

$$\|I^m\|_{k,p,0} \leq C_{k,p} \sum_{r=1}^k \left( \mathbb{E} \tilde{\mathbb{E}} \left[ \Theta_r^p \middle| \mathcal{F}_0^W \vee \mathcal{F}_0^{\tilde{W}} \right] \right)^{1/p}, \quad (4.13)$$

where

$$\Theta_r = \sum_{1 \leq i_1 < \dots < i_r \leq m} \int_{0 < t_1 < \dots < t_m < 1} \dots \otimes d\tilde{B}_{t_{i_1}} \dots \otimes d\tilde{B}_{t_{i_2}} \otimes \dots \otimes d\tilde{B}_{t_{i_r}} \dots \quad (4.14)$$

and the omitted terms are filled with  $\otimes dB_{t_i}$ .

To eliminate the dependence on  $\tilde{W}$ , we take  $L^p$ -norm with respect to the conditional probability  $\mathbb{P}(\cdot | \mathcal{F}_0^W)$  on both sides of (4.13). This gives

$$\|I_{s,t}^m\|_{k,p,s} \stackrel{M}{=} (t-s)^{mH} \|I_{0,1}^m\|_{k,p,0} \leq C_{k,p} (t-s)^{mH} \sum_{r=1}^k (\mathbb{E} [\Theta_r^p | \mathcal{F}_0^W])^{1/p}. \quad (4.15)$$

We thus define  $\Psi^2 := C_{k,p} (\mathbb{E} [\Theta_r^p | \mathcal{F}_0^W])^{1/p}$ . According to the definition of  $\Theta_i$  in (4.14), each  $\Theta_i$  is in fact a component of the signature lifted from the  $2d$ -dimensional fractional Brownian motion  $(B_u, \tilde{B}_u)_{u \geq 0}$ . It is easily seen that  $\Psi^2$  is smooth in the Malliavin sense. The random variable  $\Psi_{s,t}^2$  is now defined in a similar way as in (4.9) from  $\Psi^2$ . This completes the proof of the lemma.  $\square$

### 4.2.3 Estimate of $\gamma(\mathbf{U}_{s,t})^{-1}$

The last ingredient for estimating the right hand side of (4.2) is the inverse Malliavin matrix of  $\mathbf{U}_{s,t}$ . The main result for this part is summarised in the lemma below.

**Lemma 4.8.** *For any integer  $k \geq 1$  and any real number  $p > 0$ , there exist a constant  $\alpha > 0$  and a family of random variables  $(\Psi_{s,t}^3)_{0 \leq s < t}$  (depending on  $k$  and  $p$ ) such that*

$$\|\gamma(\mathbf{U}_{s,t})^{-1}\|_{k,p,s} \leq \frac{1}{(t-s)^\alpha} \Psi_{s,t}^3, \quad 0 \leq s < t \leq T,$$

where the Malliavin-Sobolev norm of each  $\Psi_{s,t}^3$  does not depend on  $s$  or  $t$ .

The proof of Lemma 4.8 is quite technical and therefore postponed to the appendix.

### 4.2.4 Proof of Proposition 4.3

We now combine Lemmas 4.5, 4.6, and 4.8 to derive an upper bound for the conditional density of the logarithmic signature process  $\mathbf{U}_{s,t}$ .

*Proof of Proposition 4.3.* Based on Lemma 2.17, one can apply a similar argument as in the derivation of (4.10) to show that

$$\mathbb{P}(\mathbf{U}_{s,s+\varepsilon} \in du \mid \mathcal{F}_s^W) \stackrel{M}{=} \mathbb{P}(\delta_{\varepsilon H} \mathbf{U}_{0,1} \in du \mid \mathcal{F}_0^W).$$

for any  $\varepsilon > 0$  and  $s \geq 0$ . By the change of variable  $u = \delta_{\varepsilon H} \hat{u}$ , one obtains

$$\mathbb{P}(\delta_{\varepsilon H} \mathbf{U}_{0,1} \in du \mid \mathcal{F}_0^W) = \varepsilon^{-H\nu} \mathbb{P}(\mathbf{U}_{0,1} \in d\hat{u} \mid \mathcal{F}_0^W).$$

It follows that

$$\rho_{s,s+\varepsilon}(u) = \mathbb{P}(\mathbf{U}_{s,s+\varepsilon} \in du \mid \mathcal{F}_s^W) \stackrel{M}{=} \varepsilon^{-H\nu} \rho_{0,1}(\hat{u})$$

for all  $u \in \mathfrak{g}^{(l)}(\mathbb{R}^d)$  with  $\hat{u} = \delta_{\varepsilon^{-H}} u$ .

By applying Lemmas 4.5, 4.6 and 4.8 to the inequality (4.2), one finds that

$$\rho_{0,1}(\hat{u}) \leq C \exp\left(-\frac{\|\hat{u}\|_{\text{CC}}^2}{C}\right) \prod_{i=1}^3 (1 + \Psi^i)^k$$

for some constants  $C, k > 0$ . As a result, for any  $0 \leq s < t \leq T$ , one has

$$\rho_{s,t}(u) \leq \frac{C}{(t-s)^{H\nu}} \exp\left(-\frac{\|u\|_{\text{CC}}^2}{C(t-s)^{2H}}\right) \prod_{i=1}^3 (1 + \Psi_{s,t}^i)^k,$$

where we define

$$\Psi := \prod_{i=1}^3 (1 + \Psi^i)^k,$$

and  $\Psi_{s,t}$  is defined in a similar way as in (4.9) from  $\Psi$ . It is clear that the Malliavin-Sobolev norm of  $\Psi$  is finite, which completes the proof of the proposition.  $\square$

### 4.3 Proof of Theorem 1.5

We now proceed to develop the proof of the local upper bound for the joint density of  $(Y_s, Y_t)$  (Theorem 1.5). More specifically, we consider the RDE (1.3). Given smooth vector fields  $\{V_i\}_{i=1}^d$ , we set  $V_\emptyset = 1$  and for each word  $I = (i_1, \dots, i_r)$ ,

$$V_{(I)} := V_{i_1}(\dots V_{i_{r-1}}(V_{i_r}) \dots), \quad \text{and} \quad V_I := [V_{i_1}, \dots [V_{i_{r-2}}, [V_{i_{r-1}}, V_{i_r}]] \dots].$$

Here we identify a vector field with a first order differential operator (directional derivative) in the obvious way. The following definition plays an essential role in our analysis.

**Definition 4.9.** For each  $l \geq 1$ , we define the *Taylor approximation function*  $F_l : \mathfrak{g}^{(l)} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$  of order  $l$  associated with the RDE (1.3) by

$$F_l(u, x) := \sum_{\alpha \in \mathcal{A}_1(l)} V_{(\alpha)}(x) \cdot (\exp u)^\alpha, \quad (u, x) \in \mathfrak{g}^{(l)} \times \mathbb{R}^N, \quad (4.16)$$

where the exponential function is defined on  $T^{(l)}$  by (4.1) and  $(\exp u)^\alpha$  is the coefficient of  $\exp u$  with respect to the tensor basis element  $e_{(\alpha)}$ . We say that  $u \in \mathfrak{g}^{(l)}$  *joins*  $x$  to  $y$  in the sense of Taylor approximation if  $y = x + F_l(u, x)$ .

Let  $l \geq \bar{l}$  be given fixed, where we recall that  $\bar{l}$  is the hypoelliptic constant appearing in Hypothesis 1.2. According to [17, Corollary 4.9], there exists a constant  $r > 0$  such that for each  $x \in \mathbb{R}^N$  and  $y \in \{x + F_l(u, x) : \|u\|_{\text{HS}} < r\}$ , the "bridge space"

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

is a submanifold of  $\{u \in \mathfrak{g}^{(l)} : \|u\|_{\text{HS}} < r\}$  with dimension  $\dim \mathfrak{g}^{(l)} - N$ . In addition, since both  $\mathfrak{g}^{(l)}$  and  $\mathbb{R}^N$  are oriented Riemannian manifolds, the manifold  $M_{x,y}$  carries a natural orientation and hence a volume form which will be denoted as  $m_{x,y}$ . The following result is a standard disintegration formula in Riemannian geometry.

**Proposition 4.10.** *For any  $\varphi \in C_c^\infty(\{u \in \mathfrak{g}^{(l)} : \|u\|_{\text{HS}} < r\})$ , one has*

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

where the kernel  $K$  is given by

$$K(v, x) = (\det(JF_l(v, x) \cdot JF_l(v, x)^*))^{-1/2}, \quad (4.17)$$

$JF_l(v, x) : \mathfrak{g}^{(l)} \rightarrow \mathbb{R}^N$  is the Jacobian of  $F_l(\cdot, x)$  at  $v \in \mathfrak{g}^{(l)}$  and we define  $m_{x,y} = 0$  if  $M_{x,y} = \emptyset$ .

The above disintegration formula immediately leads to the following formula for the localised conditional density of the Taylor approximation process  $Y_l(s, t, x) := x + F_l(\mathbf{U}_{s,t}, x)$ .

**Lemma 4.11.** *Let  $r > 0$  be given as before. Let  $\eta : \mathfrak{g}^{(l)} \rightarrow [0, 1]$  be a smooth bump function such that  $\eta(u) = 1$  when  $\|u\|_{\text{HS}} < r/2$  and  $\eta(u) = 0$  when  $\|u\|_{\text{HS}} > r$ . Let  $\mathbb{P}_l^\eta(s, t, \cdot)$  be the conditional measure defined by*

$$\mathbb{P}_l^\eta(s, t, A) := \mathbb{E} [\eta(\mathbf{U}_{s,t}) \mathbf{1}_{\{Y_l(s,t,x) \in A\}} | \mathcal{F}_s^W] \Big|_{x=Y_s}, \quad A \in \mathcal{B}(\mathbb{R}^N).$$

Then  $\mathbb{P}_l^\eta(s, t, \cdot)$  is absolutely continuous with respect to the Lebesgue measure and its density is given by

$$p_l^\eta(s, t, y) = \int_{M_{x,y}} \eta(u) K(u, x) \rho_{s,t}(u) m_{x,y}(du) \Big|_{x=Y_s}, \quad (4.18)$$

where  $\rho_{s,t}$  is the density of  $\mathbf{U}_{s,t}$  conditional on  $\mathcal{F}_s^W$  and  $K$  is given by (4.17).

The following result is taken from [22, Lemma 3.23].

**Lemma 4.12.** *There are constants  $0 < r_1 < r_2$  and  $\zeta \in (0, 1]$  such that for any  $x \in \mathbb{R}^N$  and  $y \in \{x + F_l(u, x) : u \in \mathfrak{g}^{(l)}, \|u\|_{\text{HS}} < r_1\}$ , there is a diffeomorphism  $S_{x,y}$  between  $M_{x,y} \cap \{u \in \mathfrak{g}^{(l)} : \|u\|_{\text{HS}} < r_1\}$  and  $M_{x,x} \cap \{u \in \mathfrak{g}^{(l)} : \|u\|_{\text{HS}} < r_2\}$ , which satisfies*

$$\zeta \cdot m_{x,x} \leq (K(\cdot, x) m_{x,y}) \circ S_{x,y}^{-1} \leq \frac{1}{\zeta} \cdot m_{x,x} \text{ on } \{S_{x,y}(u); u \in M_{x,y} \text{ and } \|u\|_{\text{HS}} < r_1\} \quad (4.19)$$

and

$$\zeta \cdot (\|S_{x,y}(u)\|_{\text{CC}} + d(x, y)) \leq \|u\|_{\text{CC}} \leq \frac{1}{\zeta} \cdot (\|S_{x,y}(u)\|_{\text{CC}} + d(x, y)) \quad (4.20)$$

for any  $u \in M_{x,y}$  with  $\|u\|_{\text{HS}} < r_1$ .

We are now in a position to prove Theorem 1.5. In fact, we will prove the following more general result which will be used in the next section.

**Theorem 4.13.** *Suppose that the vector fields  $\{V_\alpha\}_{\alpha=1}^d$  satisfy Hypothesis 1.1 and Hypothesis 1.2. Let  $p_{s,t}(x, y)$  be the joint density of the random vector  $(Y_s, Y_t)$  defined by the RDE (1.3). For any  $\delta > 0$ , there exist constants  $C_1, C_2, \tau > 0$  depending only on  $\delta, H, N$  and the vector fields  $\{V_\alpha\}_{\alpha=1}^d$  such that*

$$p_{s,t}(x, y) \leq \frac{C_1}{|B_d(x, (t-s)^H)|_{\text{Vol}}} \exp\left(-\frac{d(x, y)^2}{C_1(t-s)^{2H}}\right) + C_2|t-s| \quad (4.21)$$

for all  $(s, t, x, y) \in (\delta, T] \times \mathbb{R}^N \times \mathbb{R}^N$ . Moreover, if one restricts to the cone region

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

then  $C_2$  can be taken to be 0.

*Proof.* According to Lemma 4.11 and Lemma 4.12, one has

$$\begin{aligned} p_l^\eta(s, t, y) &= \int_{M_{x,y} \cap \{u \in \mathfrak{g}^{(l)}: \|u\|_{\text{HS}} < r_1\}} \eta(u) \rho_{s,t}(u) K(u, x) m_{x,y}(du) \Big|_{x=Y_s} \\ &\leq \int_{M_{x,y} \cap \{u \in \mathfrak{g}^{(l)}: \|u\|_{\text{HS}} < r_1\}} \rho_{s,t}(u) K(u, x) m_{x,y}(du) \Big|_{x=Y_s} \\ &= \int_{M_{x,x} \cap \{v \in \mathfrak{g}^{(l)}: \|v\|_{\text{HS}} < r_2\}} \cap \text{Im}(S_{x,y}) \rho_{s,t}(S_{x,y}^{-1}v) m_{x,x}(dv) \Big|_{x=Y_s}, \end{aligned} \quad (4.22)$$

Thanks to Proposition 4.3, one has

$$\rho_{s,t}(u) \leq \frac{C}{(t-s)^{H\nu}} \exp\left(-\frac{\|u\|_{\text{CC}}^2}{C(t-s)^{2H}}\right) \Psi_{s,t}$$

with some  $\mathcal{F}_s^W$ -measurable  $\Psi_{s,t}$  which is smooth in the sense of Malliavin and whose Malliavin-Sobolev norms do not depend on  $s, t$ . By applying (4.20), one obtains that

$$\rho_{s,t}(S_{x,y}^{-1}v) \leq \frac{C}{(t-s)^{H\nu}} e^{-\frac{\zeta^2 d(x,y)^2}{C(t-s)^{2H}}} e^{-\frac{\zeta^2 \|v\|_{\text{CC}}^2}{C(t-s)^{2H}}} \Psi_{s,t} \quad (4.23)$$

for all  $v \in \text{Im}(S_{x,y})$ . Combining (4.22) and (4.23), it follows that

$$p_l^\eta(s, t, y) \leq \frac{C}{(t-s)^{H\nu}} e^{-\frac{\zeta^2 d(x,y)^2}{C(t-s)^{2H}}} \Psi_{s,t} \int_{M_{x,x} \cap \{v \in \mathfrak{g}^{(l)}: \|v\|_{\text{HS}} < r_2\}} e^{-\frac{\zeta^2 \|v\|_{\text{CC}}^2}{C(t-s)^{2H}}} m_{x,x}(dv) \Big|_{x=Y_s}. \quad (4.24)$$

The integral on the right hand side can be estimated as

$$\begin{aligned}
& \int_{M_{x,x} \cap \{v \in \mathfrak{g}^{(l)} : \|v\|_{\text{HS}} < r_2\}} e^{-\frac{\zeta^2 \|v\|_{\text{CC}}^2}{C(t-s)^{2H}}} m_{x,x}(dv) \Big|_{x=Y_s} \\
& \leq \int_{M_{x,x} \cap \{\|v\|_{\text{CC}} \leq 1\}} e^{-\frac{\zeta^2 \|v\|_{\text{CC}}^2}{C(t-s)^{2H}}} m_{x,x}(dv) \Big|_{x=Y_s} \\
& \quad + \int_{M_{x,x} \cap \{\|v\|_{\text{CC}} > 1, \|v\|_{\text{HS}} < r_2\}} e^{-\frac{\zeta^2 \|v\|_{\text{CC}}^2}{C(t-s)^{2H}}} m_{x,x}(dv) \Big|_{x=Y_s} \\
& := I_1(Y_s) + I_2(Y_s).
\end{aligned}$$

By Fubini's theorem, one has

$$\begin{aligned}
I_1(x) &= \frac{2\zeta^2}{C(t-s)^{2H}} \int_0^\infty u e^{-\frac{\zeta^2 u^2}{C(t-s)^{2H}}} m_{x,x}(\{v \in M_{x,x} : \|v\|_{\text{CC}} \leq u \wedge 1\}) du \\
&= \frac{2\zeta^2}{C} \int_0^\infty r e^{-\frac{\zeta^2 r^2}{C}} m_{x,x}(\{v \in M_{x,x} : \|v\|_{\text{CC}} \leq (t-s)^H r \wedge 1\}) dr,
\end{aligned}$$

where we used the change of variable  $u = (t-s)^H r$  to obtain the second equality. According to [22, Equation (3.33)], there exists  $K \in (0, \infty)$  such that the estimate

$$\begin{aligned}
& m_{x,x}(\{v \in M_{x,x} : \|v\|_{\text{CC}} \leq (t-s)^H r \wedge 1\}) \\
& \leq K^{r \wedge (t-s)^{-H} + 1} m_{x,x}(\{v \in M_{x,x} : \|v\|_{\text{CC}} \leq (t-s)^H\})
\end{aligned}$$

holds for all  $r \in (0, \infty)$  and  $x \in \mathbb{R}^N$ . In addition, by [22, Equation (3.32)], there exists  $\delta \in (0, 1)$  such that for all  $(\eta, x) \in (0, 1] \times \mathbb{R}^N$ , one has

$$\eta^{-\nu} m_{x,x}(\{v \in M_{x,x} : \|v\|_{\text{CC}} < \eta\}) |B_d(x, \eta)|_{\text{Vol}} \in [\delta, 1/\delta].$$

It follows that

$$\begin{aligned}
I_1(x) &\leq C_{\zeta, K} m_{x,x}(\{v \in M_{x,x} : \|v\|_{\text{CC}} \leq (t-s)^H\}) \\
&\leq C_{\zeta, K, \delta} (t-s)^{H\nu} |B_d(x, (t-s)^H)|_{\text{Vol}}^{-1}.
\end{aligned} \tag{4.25}$$

On the other hand, note that  $e^{-\frac{\zeta^2 \|v\|_{\text{CC}}^2}{C(t-s)^{2H}}}$  is upper bounded by  $e^{-\frac{\zeta^2}{C(t-s)^{2H}}}$  on  $\{\|v\|_{\text{CC}} > 1\}$ .

As a result,  $I_2(x) \leq e^{-\frac{\zeta^2}{C(t-s)^{2H}}}$  holds for some strictly positive constant  $\zeta$ , and in particular,  $I_2(x)$  is upper bounded by  $C_\alpha (t-s)^\alpha$  for all  $\alpha > 0$  and  $s, t \in [0, T]$ . By choosing a small  $\alpha$ , it follows that

$$I_2(x) \lesssim (t-s)^{H\nu} |B_d(x, (t-s)^H)|_{\text{Vol}}^{-1},$$

where we used the simple relation  $r^\alpha \lesssim |B_d(x, r)|_{\text{Vol}}$  for some small  $\alpha$  when  $r$  is small. By substituting (4.25) into (4.24), one concludes that

$$p_l^\eta(s, t, y) \leq \frac{C}{|B_d(x, (t-s)^H)|_{\text{Vol}}} e^{-\frac{d(x,y)^2}{C(t-s)^{2H}}} \Psi_{s,t} \Big|_{x=Y_s}. \tag{4.26}$$

On the other hand, by applying a similar argument as in [17, Proposition 4.23] but with  $\mathbb{P}(\cdot)$  replaced by  $\mathbb{P}(\cdot | \mathcal{F}_s^W)$ , one can show that

$$\left| \mathbb{P}(Y_t \in dy | \mathcal{F}_s^W) - p_t^\eta(s, t, y) \right| \leq |t - s| \times \Phi_{s,t}, \quad (4.27)$$

where  $\Phi_{s,t}$  is smooth in the sense of Malliavin and its Malliavin-Sobolev norms do not depend on  $s, t$ . By combining the estimates (4.26) and (4.27), one obtains that

$$\mathbb{E}[\delta_y(Y_t) | \mathcal{F}_s^W] = \mathbb{P}(Y_t \in dy | \mathcal{F}_s^W) \leq \frac{C}{|B_d(x, (t-s)^H)|_{\text{Vol}}} e^{-\frac{d(x,y)^2}{C(t-s)^{2H}}} \Psi_{s,t} \Big|_{x=Y_s} + |t-s| \times \Phi_{s,t}.$$

It follows that

$$\begin{aligned} p_{s,t}(x, y) &= \mathbb{E}[\delta_x(Y_s) \mathbb{E}[\delta_y(Y_t) | \mathcal{F}_s^W]] \\ &\leq \frac{C}{|B_d(x, (t-s)^H)|_{\text{Vol}}} e^{-C \frac{d(x,y)^2}{(t-s)^{2H}}} \mathbb{E}[\delta_x(Y_s) \Psi_{s,t}] + (t-s) \mathbb{E}[\delta_x(Y_s) \Phi_{s,t}], \end{aligned} \quad (4.28)$$

where  $\Psi_{s,t}$  and  $\Phi_{s,t}$  are two random variables whose Malliavin-Sobolev norms do not depend on  $s, t$ .

Following the notation and the integration by parts formula in [28, Propositions 2.1.4, 2.1.5], one has

$$\mathbb{E}[\delta_x(Y_s) \Psi_{s,t}] = \mathbb{E}[\mathbf{1}_{\{Y_s > x\}} H_{(1, \dots, N)}(Y_s, \Psi_{s,t})].$$

For  $1 \leq p < q < \infty$ , there exist constants  $\beta, \gamma > 1$  and integers  $n, m$  such that

$$\|H_\alpha(F, G)\|_p \leq c_{p,q} \|\det \gamma_F^{-1}\|_\beta^m \|DF\|_{k,\gamma}^n \|G\|_{k,q}.$$

Because the Malliavin-Sobolev norms of  $\Psi_{s,t}$  are independent of  $s, t$ ,  $\|\Psi_{s,t}\|_{k,q}$  only depends on  $k, p$ . In addition, for any  $s \in [\delta, T]$ , it is well-known that (see [10]) the following three quantities

$$\mathbb{P}(Y_s > x), \quad \|\det \gamma_{Y_s}^{-1}\|_\beta^m, \quad \|\mathbf{D}_W Y_s\|_{k,\gamma}^n,$$

are uniformly bounded above by a constant depending on  $\delta$ . The same argument applies to the other term  $\mathbb{E}[\delta_x(Y_s) \Phi_{s,t}]$ . Now one can integrate (4.28) by parts to obtain that

$$p_{s,t}(x, y) \leq \frac{C}{|B_d(x, (t-s)^H)|_{\text{Vol}}} e^{-\frac{d(x,y)^2}{C(t-s)^{2H}}} + C(t-s).$$

If  $d(x, y) \leq (t-s)^H$ , the first term grows faster than the second one, hence yielding the desired upper estimate (1.5).

Now the proof of the theorem is complete.  $\square$

## 5 Hitting probabilities and Newtonian-type capacities

In this section, we develop the proof of Theorem 1.9, which establishes a two-sided estimate for hitting probabilities of a hypoelliptic SDE driven by fBM with Hurst parameter  $H \in (1/4, 1)$  in terms of Newtonian-type capacities. A main novel point here is that the capacities used to control the hitting probabilities effectively are the ones induced by the *sub-Riemannian control distance*  $d$  (rather than the Euclidean metric as in the elliptic case). We recall from Definition 1.7 that they are defined by

$$\text{Cap}_\alpha(A) := \left[ \inf_{\mu \in \mathcal{P}(A)} \mathcal{E}_\alpha(\mu) \right]^{-1} \quad (5.1)$$

for  $\alpha \in \mathbb{R}$ , where

$$\mathcal{E}_\alpha(\mu) := \iint K_\alpha(d(x, y)) \mu(dx) \mu(dy) \quad (5.2)$$

and  $K_\alpha(d(x, y))$  is the kernel defined by (1.7).

The strategies for proving the upper and lower bounds in (1.10) are quite different, and we present them in separate subsections. The two main theorems are stated in Theorem 5.4 and Theorem 5.8 respectively.

As a preliminary tool, here we recall a standard ball-box volume estimate from sub-Riemannian geometry which will be used in both parts of the argument. A detailed proof can be found in [1, Lemma 20.17].

**Lemma 5.1.** *Let  $\mathcal{V} = \{V_\alpha\}_{\alpha=1}^d \subseteq C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$  be a family of equiregular  $C_b^\infty$ -vector fields on  $\mathbb{R}^N$ . For every compact set  $K \subseteq \mathbb{R}^N$ , there exist positive constants  $\varepsilon_0$  and  $C_{1,\mathcal{V},K} < C_{2,\mathcal{V},K}$ , such that*

$$C_{1,\mathcal{V},K} \varepsilon^Q \leq |B_d(q, \varepsilon)|_{\text{vol}} \leq C_{2,\mathcal{V},K} \varepsilon^Q \quad (5.3)$$

for all  $q \in K$  and  $\varepsilon < \varepsilon_0$ , where  $Q$  denotes the homogeneous dimension (see Definition 1.8).

### 5.1 Lower bound for hitting probabilities

We first demonstrate in a general way how certain estimates on joint densities will lead to a lower bound of hitting probabilities in terms of capacities. After that, we show that our hypoelliptic SDE satisfy the presumed density estimates.

Let  $(u_t)_{0 \leq t \leq T}$  be a stochastic process in  $\mathbb{R}^N$  with continuous sample paths. Suppose that  $(u_s, u_t)$  has a joint density function  $p_{s,t}(\cdot, \cdot)$  with respect to the Lebesgue measure for all  $s \neq t \in (0, T]$ . We make the following assumptions on the process  $u_t$ .

**(A1)** For any  $0 < a < b \leq T$  and  $M > 0$ , there exists a positive constant  $C_1 = C(a, b, T, M)$  such that

$$\int_a^b p_t(y) dt \geq C_1$$

holds for all  $y \in B_d(o, M)$ .

(A2) Let  $\alpha < Q$  be given fixed ( $Q$  is the homogeneous dimension as before). For any  $0 < \delta < T$  and  $M > 0$ , there exists  $C_2 = C(\delta, T, M, \alpha) > 0$  such that

$$\int_a^b \int_a^b p_{s,t}(x, y) ds dt \leq C_2 K_\alpha(d(x, y))$$

holds for all  $\delta \leq a < b \leq T$  and  $x, y \in B_d(o, M)$ .

**Lemma 5.2.** *Let  $(u_t)_{0 \leq t \leq T}$  be a stochastic process satisfying assumptions (A1) and (A2). Let  $0 < a < b \leq T$  and  $M > 0$  be given fixed. Then for any  $\eta \in (0, Q - \alpha)$  (if  $\alpha < 0$ , we further require  $\alpha + \eta < 0$ ), there exists a positive constant  $C = C(a, b, T, M, \alpha, \eta, V)$  such that*

$$\mathbb{P}(u([a, b]) \cap A \neq \emptyset) \geq C \text{Cap}_{\alpha+\eta}(A) \quad (5.4)$$

holds for all compact sets  $A \subseteq B_d(o, M)$

*Proof.* Let  $A \subseteq B_d(o, M)$  be a given compact subset such that  $\text{Cap}_{\alpha+\eta}(A) > 0$ .

Case 1:  $\alpha < 0$ .

In this case,  $\alpha + \eta < 0$  and by the definition (1.7) of the kernel  $K_\alpha$ , one has  $\text{Cap}_{\alpha+\eta}(A) = 1$ . Therefore, one only needs to prove that

$$\mathbb{P}(u([a, b]) \cap A \neq \emptyset) \geq C \quad (5.5)$$

with some positive constant  $C$  depending only on  $a, b, M, \alpha, \eta$ . In what follows, we always use  $C$  to denote such a constant which may differ from line to line.

Let us consider the following random variable

$$J_\varepsilon(A) := \frac{1}{|A_\varepsilon|_{\text{Vol}}} \int_a^b \mathbf{1}_{A_\varepsilon}(u_t) dt, \quad \varepsilon \in (0, 1),$$

where  $A_\varepsilon := \{x \in \mathbb{R}^N : d(x, A) \leq \varepsilon\}$ . We claim that the estimate

$$\mathbb{P}(J_\varepsilon(A) > 0) \geq C$$

holds uniformly for all  $\varepsilon \in (0, 1)$  and compact sets  $A \subseteq B_d(o, M)$ , where the constant  $C > 0$  does not depend on  $\varepsilon$  or the set  $A$ . Indeed, by Assumption (A1) one has

$$\mathbb{E}[J_\varepsilon(A)] = \frac{1}{|A_\varepsilon|_{\text{Vol}}} \int_a^b \int_{A_\varepsilon} p_t(y) dy dt \geq C.$$

On the other hand, Assumption (A2) (also noting that  $\alpha < 0$ ) shows that

$$\begin{aligned} \mathbb{E}[J_\varepsilon(A)^2] &= \frac{1}{|A_\varepsilon|_{\text{Vol}}^2} \int_a^b \int_a^b \int_{A_\varepsilon} \int_{A_\varepsilon} p_{s,t}(x, y) dx dy ds dt \\ &\leq \frac{C_2}{|A_\varepsilon|_{\text{Vol}}^2} \int_{A_\varepsilon} \int_{A_\varepsilon} K_\alpha(d(x, y)) dx dy \leq C. \end{aligned}$$

By the Cauchy-Schwarz inequality, one obtains that

$$\mathbb{P}(J_\varepsilon(A) > 0) = \mathbb{E}[\mathbf{1}_{\{J_\varepsilon(A) > 0\}}^2] \geq \frac{\mathbb{E}[J_\varepsilon(A)]^2}{\mathbb{E}[J_\varepsilon(A)^2]} \geq C.$$

Observe that

$$\{\omega : J_\varepsilon(A) > 0\} \subseteq \{\omega : u([a, b]) \cap A_\varepsilon \neq \emptyset\}.$$

Therefore, one has

$$\mathbb{P}(u([a, b]) \cap A_\varepsilon \neq \emptyset) \geq \mathbb{P}(J_\varepsilon(A) > 0) \geq C.$$

The inequality (5.5) now follows by taking  $\varepsilon \downarrow 0$ .

Case 2:  $0 < \alpha < Q$ .

For any  $\varepsilon \in (0, 1)$  and  $\mu \in \mathcal{P}(A)$ , we define the random variable

$$J_\varepsilon(\mu) := \int_{\mathbb{R}^N} \mu(dz) \int_a^b \frac{\mathbf{1}_{B_d(z, \varepsilon)}(u_t)}{|B_d(z, \varepsilon)|_{\text{Vol}}} dt = \int_a^b \mu_\varepsilon(u_t) dt,$$

where the mollified density  $\mu_\varepsilon: \mathbb{R}^N \rightarrow \mathbb{R}$  is given by

$$\mu_\varepsilon(x) := \int_{\mathbb{R}^N} \frac{\mathbf{1}_{B_d(z, \varepsilon)}(x)}{|B_d(z, \varepsilon)|_{\text{Vol}}} \mu(dz).$$

It follows from (5.3) and assumption **(A1)** that there exists a positive constant  $C = C(a, b, M, \mathcal{V})$  such that for all  $\varepsilon \in (0, 1)$ ,

$$\mathbb{E}[J_\varepsilon(\mu)] = \int_{\mathbb{R}^N} \frac{\mu(dz)}{|B_d(z, \varepsilon)|_{\text{Vol}}} \int_{B_d(z, \varepsilon)} \left( \int_a^b p_t(y) dt \right) dy \geq C. \quad (5.6)$$

Moreover, one has

$$\begin{aligned} \mathbb{E}[J_\varepsilon(\mu)^2] &= \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \mu_\varepsilon(x) \mu_\varepsilon(y) \left( \int_a^b \int_a^b p_{s,t}(x, y) ds dt \right) dx dy \\ &\leq C_2 \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} K_\alpha(d(x, y)) \mu_\varepsilon(dx) \mu_\varepsilon(dy) = C_2 \mathcal{E}_\alpha(\mu_\varepsilon), \end{aligned} \quad (5.7)$$

where the inequality follows from Assumption **(A2)** and the last equality is due to the definition of the energy functional (by abuse of notation we write  $\mu_\varepsilon$  for both the measure and the density).

Recall that  $\text{Cap}_{\alpha+\eta}(A) > 0$ . The main observation is that there exists a probability measure  $\bar{\mu} \in \mathcal{P}(A)$  such that

$$\mathcal{E}_\alpha(\bar{\mu}_\varepsilon) \leq \frac{C_{M, \alpha, \mathcal{V}, \eta}}{\text{Cap}_{\alpha+\eta}(A)} \quad (5.8)$$

holds for all  $\varepsilon \in (0, 1)$ , where  $C_{M,\alpha,\nu,\eta}$  is a positive constant independent of  $\varepsilon$  and the set  $A$ . We postpone the proof of this technical fact in Lemma 5.3 below. Presuming the correctness of (5.8), one can deduce by using the Cauchy-Schwarz inequality along with the estimates (5.6, 5.7, 5.8) that

$$\mathbb{P}(u([a, b]) \cap A_\varepsilon \neq \emptyset) \geq \mathbb{P}(J_\varepsilon(\bar{\mu}) > 0) = \mathbb{E}[\mathbf{1}_{\{J_\varepsilon(\bar{\mu}) > 0\}}^2] \geq \frac{\mathbb{E}[J_\varepsilon(\bar{\mu})]^2}{\mathbb{E}[J_\varepsilon(\bar{\mu})^2]} \geq C \text{Cap}_{\alpha+\eta}(A) \quad (5.9)$$

for all  $\varepsilon \in (0, 1)$  with some positive constant  $C = C(a, b, M, n, \alpha, \nu, \eta)$ . The desired estimate (5.4) follows by taking  $\varepsilon \downarrow 0$ .

Case 3:  $\alpha = 0$ .

By definition, one has  $K_0(d(x, y)) = \log(N_0/d(x, y))$ . It is straightforward to see that there exists a positive constant  $C = C_{M,\eta,N_0}$  such that

$$K_0(d(x, y)) \leq C d(x, y)^{-\eta/2} = C K_{\eta/2}(d(x, y))$$

for all  $x, y \in B_d(o, M)$ . The desired estimate follows from the same argument as in Cast II.  $\square$

Now we come back to prove the existence of  $\bar{\mu}$  used in the above proof.

**Lemma 5.3.** *Fix a real number  $M > 0$  and a compact set  $A \subseteq B_d(o, M)$ . Suppose that  $\alpha \geq 0, \eta > 0, \alpha + \eta < Q$  and  $\text{Cap}_{\alpha+\eta}(A) > 0$ . Then there exists a probability measure  $\bar{\mu} \in \mathcal{P}(A)$  such that*

$$\mathcal{E}_\alpha(\bar{\mu}_\varepsilon) \leq \frac{C_{M,\alpha,\nu,\eta}}{\text{Cap}_{\alpha+\eta}(A)} \quad (5.10)$$

holds for all  $\varepsilon \in (0, 1)$ , where  $C_{M,\alpha,\nu,\eta}$  is a positive constant independent of  $\varepsilon$  and the set  $A$ .

*Proof.* Without loss of generality, we only consider the case when  $M \geq 2$  and  $A \subseteq B_d(o, M - 1)$ . According to the definition of the capacity  $\text{Cap}_{\alpha+\eta}(A)$ , there exists a probability measure  $\nu \in \mathcal{P}(A)$  such that

$$0 < \mathcal{E}_{\alpha+\eta}(\nu) \leq 2 \inf_{\mu \in \mathcal{P}(A)} \mathcal{E}_{\alpha+\eta}(\mu) = \frac{2}{\text{Cap}_{\alpha+\eta}(A)}. \quad (5.11)$$

Let us write

$$F := \mathcal{E}_{\alpha+\eta}(\nu) = \iint d(x, y)^{-(\alpha+\eta)} \nu(dy) \nu(dx).$$

*Step 1.* In this step, we provide the following sufficient condition for (5.10) to hold. More precisely, suppose that there exists a probability measure  $\bar{\mu} \in \mathcal{P}(A)$  such that

$$\bar{\mu}_\varepsilon(B_d(x, r)) \leq C_{1,\nu}^{-1} \cdot C_{2,\nu} \cdot 6F \cdot 4^{\alpha+\eta} r^{\alpha+\eta} \quad \text{for all } x \in \mathbb{R}^N, r > 0, \varepsilon > 0 \quad (5.12)$$

and

$$\bar{\mu}_\varepsilon(\mathbb{R}^N) \leq C_{1,\nu}^{-1} \cdot C_{2,\nu} \quad \text{for all } \varepsilon \in (0, 1), \quad (5.13)$$

where  $C_{1,\nu}, C_{2,\nu}$  are given by the volume comparison (5.3). Then we claim that (5.10) follows.

Indeed, one first notes that

$$\begin{aligned} \int_{\mathbb{R}^N} d(x, y)^{-\alpha} \bar{\mu}_\varepsilon(dy) &= \int_0^\infty \bar{\mu}_\varepsilon(\{y : d(x, y)^{-\alpha} \geq u\}) du \\ &= \int_0^\infty \bar{\mu}_\varepsilon(B_d(x, u^{-1/\alpha})) du \\ &= \alpha \int_0^\infty r^{-\alpha-1} \bar{\mu}_\varepsilon(B_d(x, r)) dr \\ &= \alpha \int_0^{2M} r^{-\alpha-1} \bar{\mu}_\varepsilon(B_d(x, r)) dr + \alpha \int_{2M}^\infty r^{-\alpha-1} \bar{\mu}_\varepsilon(B_d(x, r)) dr \\ &=: I_1(x) + I_2(x), \end{aligned}$$

for all  $x \in \mathbb{R}^n$ , where the first equality follows from Fubini's theorem and the third equality is due to the change of variables  $r = u^{-1/\alpha}$ . By the assumption (5.12), one has

$$\frac{I_1(x)}{C_{1,V}^{-1} \cdot C_{2,V} \cdot 6F \cdot 4^{\alpha+\eta} \cdot \alpha} \leq \int_0^{2M} r^{\eta-1} dr = \frac{2^\eta M^\eta}{\eta}$$

for all  $x \in \mathbb{R}^N$ . If one further assumes  $x \in A_\varepsilon$ , then

$$\frac{I_2(x)}{C_{1,V}^{-1} \cdot C_{2,V} \cdot 6F \cdot 4^{\alpha+\eta} \cdot \alpha} \leq \int_{2M}^\infty r^{-\alpha-1} (2M)^{\alpha+\eta} dr = \frac{2^\eta M^\eta}{\alpha}.$$

To reach the above inequality, we used the fact that  $A_\varepsilon \subseteq B_d(x, 2M)$  and thus

$$\bar{\mu}_\varepsilon(B_d(x, r)) = \bar{\mu}_\varepsilon(B_d(x, 2M))$$

for all  $r \geq 2M$ . It follows from (5.13) and (5.11) that

$$\begin{aligned} \mathcal{E}_\alpha(\bar{\mu}_\varepsilon) &= \int_{A_\varepsilon} (I_1(x) + I_2(x)) \bar{\mu}_\varepsilon(dx) \\ &\leq 6 \cdot C_{1,V}^{-2} \cdot C_{2,V}^2 \cdot 2^{2\alpha+3\eta} \cdot (\alpha\eta^{-1} + 1) \cdot M^\eta \cdot F \\ &\leq \frac{12 \cdot C_{1,V}^{-2} \cdot C_{2,V}^2 \cdot 2^{2\alpha+3\eta} \cdot (\alpha\eta^{-1} + 1) \cdot M^\eta}{\text{Cap}_{\alpha+\eta}(A)} \end{aligned}$$

for all  $\varepsilon \in (0, 1)$ . This proves the desired estimate (5.10).

*Step 2.* We will verify (5.13) and (5.12) in this step. Define the set

$$B := \left\{ x \in A : \int_{\mathbb{R}^N} d(x, y)^{-(\alpha+\eta)} \nu(dy) \leq 3F \right\} \subseteq A.$$

We claim that  $\nu(B) \geq 1/2$ . Assuming the contrary, one has

$$\begin{aligned} F &= \iint d(x, y)^{-(\alpha+\eta)} \nu(dy) \nu(dx) \\ &\geq \int_{B^c} \nu(dx) \int_{\mathbb{R}^N} d(x, y)^{-(\alpha+\eta)} \nu(dy) \\ &\geq \int_{B^c} 3F \nu(dx) \geq 3F \cdot \frac{1}{2} = \frac{3}{2}F, \end{aligned}$$

which is a contradiction. Therefore, one must have  $\nu(B) \geq 1/2$ .

Let  $\mu := \nu(\cdot \cap B)$  be the restriction of  $\nu$  to the set  $B$ . For all  $x \in B$  and  $r > 0$ , one has

$$r^{-(\alpha+\eta)} \mu(B_d(x, r)) \leq \int_{B_d(x, r)} d(x, y)^{-(\alpha+\eta)} \nu(dy) \leq 3F.$$

Now we define

$$\bar{\mu}(\cdot) := \frac{\nu(\cdot \cap B)}{\nu(B)} = \frac{\mu(\cdot)}{\nu(B)}.$$

Since  $\nu(B) \geq 1/2$ , it follows that  $\bar{\mu} \in \mathcal{P}(B)$  and

$$\bar{\mu}(B_d(x, r)) = \nu(B)^{-1} \mu(B_d(x, r)) \leq 2 \cdot 3F \cdot r^{\alpha+\eta} = 6Fr^{\alpha+\eta} \quad (5.14)$$

for all  $x \in B$ ,  $r > 0$ . In fact, the estimate (5.14) holds for all  $x \in \mathbb{R}^N$  and  $r > 0$  which is seen as follows.

- If  $B_d(x, r) \cap B = \emptyset$ , then  $\bar{\mu}(B_d(x, r)) = 0$  trivially.
- Otherwise, there exists some  $z \in B_d(x, r) \cap B$ . By applying (5.14) at  $z$  with radius  $2r$  and noting that  $B_d(x, r) \subseteq B_d(z, 2r)$ , one has

$$\bar{\mu}(B_d(x, r)) \leq \bar{\mu}(B_d(z, 2r)) \leq 6F \cdot 2^{\alpha+\eta} r^{\alpha+\eta}. \quad (5.15)$$

Thus, the estimate (5.15) holds uniformly for all  $x \in \mathbb{R}^N$  and  $r > 0$ . Now we proceed to prove the estimates (5.12) and (5.13) for the mollified measure  $\bar{\mu}_\varepsilon$ .

- Case 1:  $0 < r < \varepsilon$ . In this case, one has

$$\begin{aligned} \bar{\mu}_\varepsilon(B_d(x, r)) &= \int_{B_d(x, r)} dy \int_{\mathbb{R}^N} \frac{\mathbf{1}_{B_d(z, \varepsilon)}(y)}{|B_d(z, \varepsilon)|_{\text{Vol}}} \bar{\mu}(dz) \\ &\leq C_{2, V} r^Q \cdot C_{1, V}^{-1} \varepsilon^{-Q} \cdot 6F \cdot 2^{\alpha+\eta} \varepsilon^{\alpha+\eta} \\ &= C_{2, V} \cdot C_{1, V}^{-1} \cdot 6F \cdot 2^{\alpha+\eta} \cdot r^{\alpha+\eta}, \end{aligned} \quad (5.16)$$

where we used the volume comparison (5.3) and the fact that  $\alpha + \eta < Q$ .

- Case 2:  $r \geq \varepsilon$ . By using Fubini's theorem and the triangle inequality, one has

$$\begin{aligned}
\bar{\mu}_\varepsilon(B_d(x, r)) &= \int_{B_d(x, r)} dy \int_{\mathbb{R}^N} \frac{\mathbf{1}_{B_d(z, \varepsilon)}(y)}{|B_d(z, \varepsilon)|_{\text{Vol}}} \bar{\mu}(dz) \\
&= \int_{B_d(x, 2r)} \frac{\bar{\mu}(dz)}{|B_d(z, \varepsilon)|_{\text{Vol}}} \int_{B_d(x, r)} \mathbf{1}_{B_d(z, \varepsilon)}(y) dy \\
&\leq \bar{\mu}(B_d(x, 2r)) \cdot C_{1, \mathcal{V}}^{-1} \varepsilon^{-Q} \cdot C_{2, \mathcal{V}} \varepsilon^Q \\
&\leq C_{1, \mathcal{V}}^{-1} \cdot C_{2, \mathcal{V}} \cdot 6F \cdot 4^{\alpha+\eta} r^{\alpha+\eta}, \tag{5.17}
\end{aligned}$$

where we used (5.15) and the volume comparison in (5.3) to reach the third line.

Since  $\bar{\mu}$  is a probability measure, one also has from the third line of (5.17) that

$$\bar{\mu}_\varepsilon(\mathbb{R}^N) \leq C_{1, \mathcal{V}}^{-1} \cdot C_{2, \mathcal{V}}.$$

Combining (5.16) and (5.17), one obtains the estimate

$$\bar{\mu}_\varepsilon(B_d(x, r)) \leq C_{1, \mathcal{V}}^{-1} \cdot C_{2, \mathcal{V}} \cdot 6F \cdot 4^{\alpha+\eta} r^{\alpha+\eta} \tag{5.18}$$

for all  $x \in \mathbb{R}^N$  and  $r > 0$ , which completes the proof.  $\square$

We now specialise in our SDE context and prove the following main result of this section.

**Theorem 5.4.** *Assume that the family of vector fields  $\mathcal{V} = \{V_\alpha\}_{\alpha=1}^d$  is equiregular on  $\mathbb{R}^N$  and satisfies Hypotheses 1.1 and 1.2. Let  $Y_t$  be the solution to the RDE (1.3) where  $\mathbf{B}$  is the canonical lift of a fractional Brownian motion  $(B_t)_{0 \leq t \leq T}$  with Hurst parameter  $H \in (1/4, 1)$ . Fix  $0 < a < b$  and  $M > 0$ . Then for any  $\eta \in (0, 1/H)$ , there exists a constant  $C = C(a, b, M, Q, H, \mathcal{V}, \eta) > 0$  such that the estimate*

$$\mathbb{P}(Y([a, b]) \cap A \neq \emptyset) \geq C \text{Cap}_{Q-1/H+\eta}(A)$$

holds for all compact sets  $A \subseteq B_d(o, M)$ . Here  $Q$  denotes the homogeneous dimension.

*Proof.* According to Lemma 5.2, it remains to verify conditions (A1) and (A2). The density function  $p_t(\cdot)$  of  $Y_t$  is known to be continuous (see [10, Theorem 3.5]) and everywhere strictly positive (see [17, Theorem 1.5]). In particular, one has

$$\int_a^b p_t(y) dt \geq \inf_{y \in B_d(o, M)} \int_a^b p_t(y) dt =: C(a, b, M) > 0.$$

This verifies condition (A1).

To establish the upper bound in condition (A2), recall from the joint density upper estimate (4.21) that there exist constants  $c_1, c_2, c_3 > 0$  such that

$$p_{s,t}(x, y) \leq \frac{c_1}{|B_d(x, (t-s)^H)|_{\text{Vol}}} \exp\left(-c_2 \frac{d(x, y)^2}{(t-s)^{2H}}\right) + c_3 |t-s|$$

holds for all  $s, t \in [a, b]$  and  $x, y \in \mathbb{R}^N$ . If one further restricts to  $x \in B_d(o, M)$ , by Lemma 5.1 the volume of the ball  $B_d(x, (t-s)^H)$  is comparable to  $|t-s|^{HQ}$ . Namely, there exist constants  $c_4, c_5 > 0$  depending only on  $M$  and the vector fields  $\mathcal{V}$ , such that

$$c_4 (t-s)^{QH} \leq |B_d(x, (t-s)^H)| \leq c_5 (t-s)^{QH}. \quad (5.19)$$

It follows that

$$p_{s,t}(x, y) \leq \frac{C_{1,M,\mathcal{V}}}{(t-s)^{QH}} \exp\left(-C_{2,M,\mathcal{V}} \frac{d(x,y)^2}{(t-s)^{2H}}\right) + C_{3,M,\mathcal{V}} |t-s|$$

holds uniformly for all  $s, t \in [a, b]$  and  $x, y \in B_d(o, M)$ , where  $C_{i,M,\mathcal{V}}$  ( $i = 1, 2, 3$ ) are positive constants depending on  $a, b, M, \mathcal{V}$ .

Our aim is to show that

$$\int_a^b \int_a^b p_{s,t}(x, y) ds dt \leq C K_{Q-1/H}(d(x, y)),$$

with some positive constant  $C = C(a, b, M, Q, H, \mathcal{V})$ .

First of all, one observes that

$$\int_a^b \int_a^b |t-s| ds dt \leq C K_{Q-1/H}(d(x, y)). \quad (5.20)$$

This is a trivial consequence of the fact that  $K_{Q-1/H}(d(x, y))$  is bounded away from zero for all  $x, y \in B_d(o, M)$  (in the case  $Q = 1/H$ , one needs to choose a suitably large  $N_0$  depending on  $M$ ; recall (1.7) for the definition of the kernel  $K$ ). Now it suffices to establish the following estimate:

$$\int_a^b \int_a^b \frac{1}{(t-s)^{QH}} \exp\left(-C_{2,M,\mathcal{V}} \frac{d(x,y)^2}{(t-s)^{2H}}\right) ds dt \leq C K_{Q-1/H}(d(x, y)), \quad (5.21)$$

where  $C = C(a, b, M, Q, H, \mathcal{V}) > 0$ . To this end, one first obtains by some simple changes of variables that

$$\begin{aligned} & \int_a^b \int_a^b \frac{1}{(t-s)^{QH}} \exp\left(-C_{2,M,\mathcal{V}} \frac{d(x,y)^2}{(t-s)^{2H}}\right) ds dt \\ & \leq 2(b-a) \int_0^{b-a} \frac{1}{u^{QH}} \exp\left(-C_{2,M,\mathcal{V}} \frac{r^2}{u^{2H}}\right) du \\ & = \frac{2(b-a)}{Hr^\alpha} \int_0^{\frac{(b-a)^H}{r}} q^{-Q-1+1/H} \exp\left(-C_{2,M,\mathcal{V}} \frac{1}{q^2}\right) dq \\ & = \frac{2(b-a)}{Hr^\alpha} \int_0^{\frac{(b-a)^H}{r}} q^{-\alpha-1} \exp\left(-C_{2,M,\mathcal{V}} \frac{1}{q^2}\right) dq, \end{aligned} \quad (5.22)$$

where we set  $r := d(x, y) \leq 2M$  and  $\alpha := Q - 1/H$ . We now verify the inequality (5.21) by considering three separate cases.

- Case 1:  $Q < 1/H$  (i.e.  $\alpha < 0$ ). We split the integral (5.22) as follows:

$$\begin{aligned}
& \frac{2(b-a)}{Hr^\alpha} \int_0^{\frac{(b-a)^H}{r}} q^{-\alpha-1} \exp\left(-C_{2,M,\nu} \frac{1}{q^2}\right) dq \\
& \leq \frac{2(b-a)}{Hr^\alpha} \left[ \int_0^{\frac{(b-a)^H}{2M}} q^{-\alpha-1} \exp\left(-C_{2,M,\nu} \frac{1}{q^2}\right) dq + \int_{\frac{(b-a)^H}{2M}}^{\frac{(b-a)^H}{r}} q^{-\alpha-1} dq \right] \\
& = \frac{2(b-a)}{Hr^\alpha} \left[ \xi_{a,b,M,\alpha,H,\nu} + \frac{1}{-\alpha} \left( \frac{r^\alpha}{(b-a)^{\alpha H}} - \frac{(2M)^\alpha}{(b-a)^{\alpha H}} \right) \right] \\
& \leq \frac{2(b-a)}{H} \left[ \xi_{a,b,M,\alpha,H,\nu} (2M)^{-\alpha} + \frac{1}{-\alpha \cdot (b-a)^{\alpha H}} \right],
\end{aligned}$$

where

$$\xi_{a,b,M,\alpha,H,\nu} := \int_0^{\frac{(b-a)^H}{2M}} q^{-\alpha-1} \exp\left(-C_{2,M,\nu} \frac{1}{q^2}\right) dq.$$

Since  $K_\alpha(d(x,y)) \equiv 1$  in this case, the estimate (5.21) holds with

$$C = \frac{2(b-a)}{H} \left[ \xi_{a,b,M,\alpha,H,\nu} (2M)^{-\alpha} + \frac{1}{(-\alpha)(b-a)^{\alpha H}} \right].$$

- Case 2:  $Q = 1/H$  (i.e.  $\alpha = 0$ ). In this case, one has

$$\begin{aligned}
& \frac{2(b-a)}{H} \int_0^{\frac{(b-a)^H}{r}} q^{-1} \exp\left(-C_{2,M,\nu} \frac{1}{q^2}\right) dq \\
& \leq \frac{2(b-a)}{H} \left[ \int_0^{\frac{(b-a)^H}{2M}} q^{-1} \exp\left(-C_{2,M,\nu} \frac{1}{q^2}\right) dq + \int_{\frac{(b-a)^H}{2M}}^{\frac{(b-a)^H}{r}} q^{-1} dq \right] \\
& = \frac{2(b-a)}{H} \left[ \zeta_{a,b,M,H,\nu} + \log\left(\frac{2M}{r}\right) \right],
\end{aligned}$$

where

$$\zeta_{a,b,M,H,\nu} := \int_0^{\frac{(b-a)^H}{2M}} q^{-1} \exp\left(-C_{2,M,\nu} \frac{1}{q^2}\right) dq.$$

By setting  $N_0 := e^{\zeta_{a,b,M,H,\nu}} \cdot 2M$ , one obtains that

$$\int_a^b \int_a^b \frac{1}{(t-s)^{QH}} \exp\left(-C_{2,M,\nu} \frac{d(x,y)^2}{(t-s)^{2H}}\right) ds dt \leq \frac{2(b-a)}{H} \log\left(\frac{N_0}{d(x,y)}\right),$$

which matches the definition of  $K_0(d(x,y))$ . Therefore, the estimate (5.21) holds with  $C = \frac{2(b-a)}{H}$ .

- Case 3:  $Q > 1/H$  (i.e.  $\alpha > 0$ ). In this case, one has

$$\begin{aligned} & \frac{2(b-a)}{Hr^\alpha} \int_0^{\frac{(b-a)H}{r}} q^{-\alpha-1} \exp\left(-C_{2,M,\nu} \frac{1}{q^2}\right) dq \\ & \leq \frac{2(b-a)}{Hr^\alpha} \int_0^\infty q^{-\alpha-1} \exp\left(-C_{2,M,\nu} \frac{1}{q^2}\right) dq = \frac{2(b-a)}{H} \eta_{\alpha,M,\nu} \cdot r^{-\alpha}, \end{aligned}$$

where

$$\eta_{\alpha,M,\nu} := \int_0^\infty q^{-\alpha-1} \exp\left(-C_{2,M,\nu} \frac{1}{q^2}\right) dq.$$

Since  $K_\alpha(d(x,y)) = d(x,y)^{-\alpha}$ , the estimate (5.21) holds with  $C = \frac{2(b-a)}{H} \eta_{\alpha,M,\nu}$ .

Now the verification of the condition (A2) and thus the proof of the theorem is complete.  $\square$

## 5.2 Upper bound for hitting probabilities

We now proceed to establish the upper bound in Theorem 1.9. As in the lower bound argument, we first establish a general result and restrict to the RDE context.

According to [6, Theorem 5.12], the upper bound stated in (1.10) can be obtained under a general criterion involving Hausdorff measures. We first recall some basic definitions. We continue to assume that the collection of vector fields  $\{V_\alpha\}_{\alpha=1}^d$ , each belonging to  $C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$ , is equiregular throughout  $\mathbb{R}^N$ , with the corresponding homogeneous dimension denoted by  $Q$ . The metric  $d(\cdot, \cdot)$  refers to the control distance defined by (1.4). Given any non-negative real number  $\alpha$ , the *Hausdorff measure of dimension  $\alpha$*  for a subset  $A$  of  $\mathbb{R}^N$  is defined by

$$\begin{aligned} \mathcal{H}_\alpha(A) & := \lim_{\delta \downarrow 0} \mathcal{H}_\alpha^\delta(A); \\ \mathcal{H}_\alpha^\delta(A) & := \inf \left\{ \sum_i (2r_i)^\alpha : A \subseteq \bigcup_i B_d(x_i, r_i), \sup_i r_i \leq \delta \right\}. \end{aligned} \quad (5.23)$$

For any Borel subset  $A \subseteq \mathbb{R}^N$ , we define the *order- $\alpha$  weighted Hausdorff measure* as

$$\begin{aligned} \lambda_\alpha(A) & = \lim_{\delta \downarrow 0} \lambda_\alpha^\delta(A), \\ \lambda_\alpha^\delta(A) & = \inf \left\{ \sum_i c_i d(E_i)^\alpha : 1_A \leq \sum_i c_i 1_{E_i}, c_i > 0, \sup_i d(E_i) \leq \delta \right\}. \end{aligned} \quad (5.24)$$

where for each set  $E$ , the diameter  $d(E)$  refers to the greatest distance between any two points in  $E$  with respect to the metric  $d$ .

In order to apply Hausdorff-type criteria to control hitting probabilities, we will need a generalised version of Frostman's lemma adapted to compact metric spaces. We refer the reader to [26, Theorem 8.17] for its proof.

**Lemma 5.5** (Frostman's Lemma). *Let  $U$  be a compact metric space and fix  $0 \leq \alpha < \infty$ ,  $0 < \delta \leq \infty$ . Then there exists a Radon measure  $\mu$  on  $U$  such that  $\mu(U) = \lambda_\alpha^\delta(U)$  and*

$$\mu(E) \leq d(E)^\alpha \quad \text{for all } E \subseteq U \text{ with } d(E) < \delta. \quad (5.25)$$

*In particular, if  $\mathcal{H}_\alpha(U) > 0$ , then there exist  $\delta > 0$  and a Radon measure  $\mu$  satisfying (5.25) with  $\mu(U) > 0$ .*

Let  $(u_t)_{0 \leq t \leq T}$  be a continuous stochastic process taking values in  $\mathbb{R}^N$ . We consider the following assumption.

**(A3)** The estimate

$$\mathbb{P}(u([s, t]) \cap B_d(z, \varepsilon) \neq \emptyset) \leq C\varepsilon^\beta \quad (5.26)$$

holds uniformly for all  $z \in B_d(o, M)$ ,  $\varepsilon > 0$ ,  $s, t \in [a, b]$  with  $|t - s| = \varepsilon^{1/H}$ . Here  $0 < a < b \leq T$ ,  $\beta > 0$ ,  $M > 0$ ,  $H \in (0, 1)$  are suitable constants and  $C > 0$  possibly depends on them but not on  $s, t, z$ .

Assuming condition **(A3)**, we will first derive the desired upper bound for the hitting probabilities. The justification of **(A3)** in the RDE context will be addressed in a later part.

**Lemma 5.6.** *Let  $(u_t)_{0 \leq t \leq T}$  be a continuous stochastic process satisfying assumption **(A3)**. Fix  $\beta > 0$ ,  $0 < a < b \leq T$  and  $M > 0$ . Then the following statements hold true.*

- *There exists a constant  $C_1 > 0$  such that*

$$\mathbb{P}(u([a, b]) \cap A \neq \emptyset) \leq C_1 \mathcal{H}_{\beta-1/H}(A) \quad (5.27)$$

*for all Borel sets  $A \subseteq B_d(o, M)$ .*

- *For every  $\eta > 0$ , there exists a constant  $C_2 > 0$  depending additionally on  $\eta$ , such that*

$$\mathbb{P}(u([a, b]) \cap A \neq \emptyset) \leq C_2 \text{Cap}_{\beta-1/H-\eta}(A) \quad (5.28)$$

*for all compact sets  $A \subseteq B_d(o, M)$ .*

*Proof.* We assume that  $A \neq \emptyset$  for otherwise there is nothing to prove. We proceed by considering two cases.

*Case 1:*  $\beta \leq 1/H$ . From the definitions of the Hausdorff measure and capacities in (5.23) and (1.7), one has

$$\text{Cap}_{\beta-1/H-\eta}(A) = 1 \quad \text{and} \quad \mathcal{H}_{\beta-1/H}(A) = \infty$$

for any  $A \subseteq B_d(o, M)$  and any  $\eta > 0$ . Hence, both upper bounds in (5.27) and (5.28) are satisfied in this case, with constants  $C_1 = C_2 = 1$ .

*Case 2:  $\beta > 1/H$ .* We first establish (5.27), which is an easier consequence of (A3). Fix  $\varepsilon \in (0, 1)$  and consider a partition  $\{I_k\}_{k \geq 1}$  of the interval  $[a, b]$ , where each subinterval  $I_k$  has length  $\varepsilon^{1/H}$  and the intervals have pairwise disjoint interiors. Then

$$\begin{aligned} \mathbb{P}(u([a, b]) \cap B_d(z, \varepsilon) \neq \emptyset) &\leq \sum_{k: I_k \cap [a, b] \neq \emptyset} \mathbb{P}(u(I_k) \cap B_d(z, \varepsilon) \neq \emptyset) \\ &\leq \left( \left\lfloor \frac{b-a}{\varepsilon^{1/H}} \right\rfloor + 1 \right) C_H \varepsilon^\beta \leq C_{a,b,\beta,H} \varepsilon^{\beta-1/H}. \end{aligned}$$

It follows that for any countable collection of open balls  $\{B_d(x_i, r_i)\}_{i=1}^\infty$  such that  $A \subseteq \bigcup_{i=1}^\infty B_d(x_i, r_i)$  and  $\sup_i r_i \leq \delta$ , one has

$$\begin{aligned} \mathbb{P}(u([a, b]) \cap A \neq \emptyset) &\leq \sum_{i=1}^\infty \mathbb{P}(u([a, b]) \cap B_d(x_i, r_i) \neq \emptyset) \\ &\leq C_{a,b,\beta,H} \cdot 2^{1/H-\beta} \sum_{i=1}^\infty (2r_i)^{\beta-1/H}. \end{aligned} \tag{5.29}$$

Taking the infimum of both sides of (5.29) over all such coverings  $\{B_d(x_i, r_i)\}_{i=1}^\infty$  with  $\sup_i r_i \leq \delta$ , one obtains that

$$\mathbb{P}(u([a, b]) \cap A \neq \emptyset) \leq C_1 \mathcal{H}_{\beta-1/H}^\delta(A), \tag{5.30}$$

where the constant is given by  $C_1 = 2^{1/H-\beta} C_{a,b,\beta,H}$ . Since the left hand side of (5.30) is independent of  $\delta$ , sending  $\delta \downarrow 0$  yields

$$\mathbb{P}(u([a, b]) \cap A \neq \emptyset) \leq C_1 \mathcal{H}_{\beta-1/H}(A).$$

This establishes (5.27).

Next, we proceed to establish (5.28). We assume that

$$\gamma := \mathbb{P}(u([a, b]) \cap A \neq \emptyset) > 0.$$

To prove (5.28), we propose the following sufficient condition. Suppose that for some  $\delta \geq 2M$ , there exists a probability measure  $\bar{\mu} \in \mathcal{P}(A)$  such that

$$\bar{\mu}(E) \leq \gamma^{-1} C_1 \cdot 30^{\beta-1/H} \cdot d(E)^{\beta-1/H} \quad \text{for all } E \subseteq \mathbb{R}^N \text{ with } d(E) < \delta. \tag{5.31}$$

We claim that the estimate (5.28) follows from this assumption.

- **Sufficiency of (5.31).** Fix  $\eta > 0$  such that  $\alpha := \beta - 1/H - \eta > 0$ . Recall that the  $\alpha$ -energy of the measure  $\bar{\mu}$  is defined by

$$\mathcal{E}_\alpha(\bar{\mu}) = \int \int d(x, y)^{-\alpha} \bar{\mu}(dx) \bar{\mu}(dy).$$

One has

$$\begin{aligned}
\int d(x, y)^{-\alpha} \bar{\mu}(dy) &= \int_0^\infty \bar{\mu}(\{y : d(x, y)^{-\alpha} \geq u\}) du \\
&= \int_0^\infty \bar{\mu}(B_d(x, u^{-1/\alpha})) du \stackrel{u^{-1/\alpha}=r}{=} \alpha \int_0^\infty r^{-\alpha-1} \bar{\mu}(B_d(x, r)) dr \\
&= \alpha \int_0^{2M} r^{-\alpha-1} \bar{\mu}(B_d(x, r)) dr + \alpha \int_{2M}^\infty r^{-\alpha-1} \bar{\mu}(B_d(x, r)) dr \\
&:= U_1(x) + U_2(x).
\end{aligned}$$

For the  $U_1$ -term, (5.31) implies that

$$\frac{U_1(x)}{\gamma^{-1}C_1 \cdot 30^{\alpha+\eta} \cdot \alpha} \leq \int_0^{2M} r^{-\alpha-1} \cdot (2r)^{\alpha+\eta} dr = \frac{2^{\alpha+2\eta} M^\eta}{\eta},$$

for all  $x \in \mathbb{R}^N$ , where we used the fact that  $d(B_d(x, r)) \leq 2r$ . For the  $U_2$ -term, (5.31) implies that

$$\frac{U_2(x)}{\gamma^{-1}C_1 \cdot 30^{\alpha+\eta} \cdot \alpha} \leq \int_{2M}^\infty r^{-\alpha-1} \cdot (4M)^{\alpha+\eta} dr = \frac{2^{\alpha+2\eta} M^\eta}{\alpha}$$

for all  $x \in A$ , where we used the fact that  $A \subseteq B_d(x, 2M)$  for any  $x \in A$ , so that

$$\bar{\mu}(B_d(x, r)) = \bar{\mu}(B_d(x, 2M)) \quad \text{for all } r \geq 2M.$$

As a result, one has

$$\begin{aligned}
\mathcal{E}_\alpha(\bar{\mu}) &= \int_A (U_1(x) + U_2(x)) \bar{\mu}(dx) \\
&\leq \gamma^{-1}C_1 \cdot 30^{\alpha+\eta} \cdot (\alpha\eta^{-1} + 1) \cdot 2^{\alpha+2\eta} M^\eta \\
&:= \gamma^{-1}C_2,
\end{aligned}$$

where  $C_2 := C_1 \cdot 30^{\alpha+\eta} \cdot (\alpha\eta^{-1} + 1) \cdot 2^{\alpha+2\eta} M^\eta$ . It follows that

$$\mathbb{P}(u([a, b]) \cap A \neq \emptyset) = \gamma \leq C_2 \left[ \inf_{\mu \in \mathcal{P}(A)} \mathcal{E}_{\beta-1/H-\eta}(\mu) \right]^{-1} = C_2 \text{Cap}_{\beta-1/H-\eta}(A).$$

This completes the proof of (5.28).

Now we come back to construct a measure  $\bar{\mu}$  that satisfies (5.31). According to (5.30), one has

$$0 < \gamma = \mathbb{P}(u([a, b]) \cap A \neq \emptyset) \leq C_1 \mathcal{H}_{\beta-1/H}^\delta(A), \quad \text{for all } \delta > 0. \quad (5.32)$$

A key tool is [26, Lemma 8.16], which states that for any compact metric space  $U$  and real numbers  $0 < \delta \leq \infty$ ,  $0 \leq \alpha < \infty$ , one has

$$\mathcal{H}_\alpha^{30\delta}(U) \leq 30^\alpha \lambda_\alpha^\delta(U). \quad (5.33)$$

Combining (5.32) and (5.33), one finds that

$$0 < \gamma \leq C_1 \mathcal{H}_{\beta-1/H}^{30\delta}(A) \leq 30^{\beta-1/H} C_1 \lambda_{\beta-1/H}^\delta(A), \quad \text{for all } \delta > 0. \quad (5.34)$$

Fix a  $\delta$  sufficiently large (say,  $\delta \geq 2M$ ). Applying Lemma 5.5 to the set  $A$  and the measure  $\lambda_{\beta-1/H}^\delta$ , one obtains a Radon measure  $\mu$  such that

$$\begin{aligned} \mu(A) &= \lambda_{\beta-1/H}^\delta(A), \quad \text{and} \quad \text{supp}(\mu) \subseteq A, \\ \mu(E) &\leq d(E)^{\beta-1/H} \quad \text{for all } E \subseteq \mathbb{R}^N \text{ with } d(E) < \delta. \end{aligned}$$

Since  $A \subseteq B_d(o, M)$  is a non-empty bounded set, the definition of  $\lambda_\alpha^\delta$  ensures

$$0 < \mu(A) = \lambda_{\beta-1/H}^\delta(A) < \infty.$$

One then define a probability measure  $\bar{\mu} \in \mathcal{P}(A)$  by

$$\bar{\mu}(\cdot) := \frac{\mu(\cdot \cap A)}{\mu(A)}.$$

Duo to (5.34) and the relation  $\mu(A) = \lambda_{\beta-1/H}^\delta(A)$ , one has

$$\bar{\mu}(E) \leq \gamma^{-1} C_1 \cdot 30^{\beta-1/H} \cdot d(E)^{\beta-1/H}$$

for all  $E \subseteq \mathbb{R}^N$  with  $d(E) < \delta$ . This justifies the condition (5.31) and therefore completes the proof of the lemma.  $\square$

Now we restrict to the RDE context and establish the main upper bound in Theorem 1.9. We first present a moment continuity estimate with respect to the control distance  $d$  which may also be of independent interest.

**Lemma 5.7.** *Under the assumptions in Theorem 5.4, there exists a constant  $C_1 = C(H, \mathcal{V}, p, T)$  such that*

$$\mathbb{E} [d(Y_s, Y_t)^p] \leq C_1 |t - s|^{pH} \quad \forall s, t \in [0, T].$$

Moreover, for any  $\eta \in (0, H)$  and  $p \geq 1$ , there exists a constant  $C_2 = C(H, \mathcal{V}, p, T, \eta)$ , such that

$$\mathbb{E} \left[ \sup_{s, t \in [0, T]} \frac{d(Y_s, Y_t)^p}{|t - s|^{p\eta}} \right] \leq C_2.$$

*Proof.* We assume without loss of generality that  $s = 0$  and aim at proving the following estimate:

$$\mathbb{E} [d(y_0, Y_t)^p] \leq C t^{pH}, \quad \text{for all } t \in [0, T], \quad (5.35)$$

where the constant  $C$  depends only on  $H$ ,  $\mathcal{V}$ ,  $p$ , and  $T$ . To this end, consider the rescaled rough differential equation

$$dZ_r = \varepsilon^H \sum_{i=1}^d V_i(Z_r) dB_r, \quad Z_0 = z_0 \in \mathbb{R}^N, \quad (5.36)$$

and let  $(Y_r^\varepsilon)_{r \geq 0}$  denote its solution. By the scaling property of fBM, it is obvious that

$$d(y_0, Y_t) \stackrel{d}{=} d(y_0, Y_1^t)$$

for all  $t \in [0, T]$  where  $Y_1^t$  is the solution at time 1 to the equation (5.36) with  $\varepsilon = t$ . Therefore, the estimate (5.35) is equivalent to the following bound:

$$\mathbb{E} [d(y_0, Y_1^t)^p] \leq C t^{pH} \quad \forall t \in [0, T]. \quad (5.37)$$

To establish this estimate, we localise the expectation on the events  $\{\omega : \|Y_1^t - y_0\|_{\mathbb{R}^N} \geq 1\}$  and  $\{\omega : \|Y_1^t - y_0\|_{\mathbb{R}^N} < 1\}$  separately.

*Case 1 : Localisation on  $\{\omega : \|Y_1^t - y_0\|_{\mathbb{R}^N} \geq 1\}$ .* To proceed, we need a comparison result between the control distance  $d(x, y)$  and the Euclidean distance  $\|x - y\|_{\mathbb{R}^N}$  in  $\mathbb{R}^N$ . According to [22, Equation (3.22)], there exists a constant  $K \geq 1$ , depending only on the vector fields, such that

$$\frac{1}{K} \|x - y\|_{\mathbb{R}^N} \leq d(x, y) \leq K \|x - y\|_{\mathbb{R}^N}^{1/\bar{l}} \quad (5.38)$$

for all  $x, y \in \mathbb{R}^N$  with  $\|x - y\|_{\mathbb{R}^N} \wedge d(x, y) \leq 1$ , where  $\bar{l}$  is the hypoellipticity constant appearing in Hypothesis 1.2.

Using (5.38), one can derive an upper bound for  $d(x, y)$  when  $\|x - y\|_{\mathbb{R}^N} \geq 1$ . Fix such  $x, y$  and let  $m \geq 1$  satisfy  $m \leq \|x - y\|_{\mathbb{R}^N} < m + 1$ . Consider a partition  $\{x_i\}_{i=0}^m \subseteq \overline{xy}$ , the line segment between  $x$  and  $y$ , such that  $x_0 = x$ ,  $\|x_k - x_{k-1}\|_{\mathbb{R}^N} = 1$  for  $1 \leq k \leq m$ , and  $\|x_m - y\|_{\mathbb{R}^N} < 1$ . It follows from the triangle inequality and the upper bound in (5.38) that

$$\begin{aligned} d(x, y) &\leq \sqrt{m+1} \left( \sum_{i=0}^{m-1} d(x_i, x_{i+1}) + d(x_m, y) \right) \\ &\leq \sqrt{m+1} (Km + K) = K(m+1)^{3/2}. \end{aligned}$$

Since  $m+1 \leq 2\|x - y\|_{\mathbb{R}^N}$ , the right hand side can be further bounded by  $2^{3/2} K \|x - y\|_{\mathbb{R}^N}^{3/2}$ . It follows that

$$d(x, y) \leq 2^{3/2} K \|x - y\|_{\mathbb{R}^N}^{3/2}$$

for all  $x, y$  with  $\|x - y\|_{\mathbb{R}^N} \geq 1$ . By applying this bound to the random variable  $Y_1^t$ , one obtains that

$$\begin{aligned} d(y_0, Y_1^t)^p \mathbf{1}_{\{\|Y_1^t - y_0\|_{\mathbb{R}^N} \geq 1\}} &\leq (2^{3/2} K)^p \|Y_1^t - y_0\|_{\mathbb{R}^N}^{\frac{3p}{2}} \mathbf{1}_{\{\|Y_1^t - y_0\|_{\mathbb{R}^N} \geq 1\}} \\ &\leq (2^{3/2} K)^p t^{\frac{3pH}{2}} \left| \int_0^1 V(Y_r^t) dB_r \right|^{\frac{3p}{2}}, \end{aligned} \quad (5.39)$$

where the last inequality is due to the definition of  $(Y_r^t)_{r \geq 1}$  in (5.36).

Case 2: Localisation on the event  $\{\omega : \|Y_1^t - y_0\|_{\mathbb{R}^N} < 1\}$ . By (5.38), one has

$$d(y_0, Y_1^t) \leq K \|Y_1^t - y_0\|_{\mathbb{R}^N}^{1/\bar{l}} < K.$$

Hence,

$$d(y_0, Y_1^t)^p \mathbf{1}_{\{\|Y_1^t - y_0\|_{\mathbb{R}^N} < 1\}} \leq d(y_0, Y_1^t)^p \mathbf{1}_{\{d(y_0, Y_1^t) < K\}}. \quad (5.40)$$

We now insert the Euler scheme approximation  $\mathcal{E}_{(\nu)}(y_0, \delta_{t^H} \mathbf{B})$  for  $Y_1^t$ , which starts from the initial point  $y_0$  and runs up to time 1 (see [13, Definition 10.1]) to obtain that

$$d(y_0, Y_1^t) \leq d(y_0, y_0 + \mathcal{E}_{(\nu)}(y_0, \delta_{t^H} \mathbf{B})) + d(y_0 + \mathcal{E}_{(\nu)}(y_0, \delta_{t^H} \mathbf{B}), Y_1^t), \quad (5.41)$$

where  $\mathbf{B}$  denotes the step- $\bar{l}$  signature of the fBM  $(B_r)_{0 \leq r \leq 1}$  and  $\delta_{t^H}$  is the dilation operator. According to [13, Corollary 10.15], one has

$$|y_0 + \mathcal{E}_{(\nu)}(y_0, \delta_{t^H} \mathbf{B}) - Y_1^t| \leq C_{\nu, \bar{l}} \|\delta_{t^H} \mathbf{B}\|_{p\text{-var}; [0, 1]}^{\bar{l}+1} = C_{\nu, \bar{l}} \|\mathbf{B}\|_{p\text{-var}; [0, 1]}^{\bar{l}+1} t^{H(\bar{l}+1)}.$$

To obtain an appropriate upper bound in (5.41), we localise on the event

$$L := \{\|\mathbf{B}_{0,1}\|_{\text{CC}} \cdot t^H \leq 1\} \cap \{C_{\nu, \bar{l}} \|\mathbf{B}\|_{p\text{-var}; [0, 1]}^{\bar{l}+1} \cdot t^{H(\bar{l}+1)} \leq 1\}.$$

On this event, by using [22, Equation 3.21] and (5.38) one finds that

$$d(y_0, y_0 + \mathcal{E}_{(\nu)}(y_0, \delta_{t^H} \mathbf{B})) \leq C_{\nu} \|\delta_{t^H} \mathbf{B}_{0,1}\|_{\text{CC}} = C_{\nu} \|\mathbf{B}_{0,1}\|_{\text{CC}} \cdot t^H, \quad (5.42)$$

and

$$\begin{aligned} d(y_0 + \mathcal{E}_{(\nu)}(y_0, \delta_{t^H} \mathbf{B}), Y_1^t) &\leq K |y_0 + \mathcal{E}_{(\nu)}(y_0, \delta_{t^H} \mathbf{B}) - Y_1^t|^{1/\bar{l}} \\ &\leq K C_{\nu, \bar{l}}^{1/\bar{l}} \|\mathbf{B}\|_{p\text{-var}; [0, 1]}^{(\bar{l}+1)/\bar{l}} t^{H(\bar{l}+1)/\bar{l}}. \end{aligned} \quad (5.43)$$

Combining (5.40), (5.41), (5.42) and (5.43), one obtains that

$$\begin{aligned} d(y_0, Y_1^t)^p \mathbf{1}_{\{\|Y_1^t - y_0\|_{\mathbb{R}^N} < 1\}} &\leq d(y_0, Y_1^t)^p \mathbf{1}_{\{d(y_0, Y_1^t) < K\}} \\ &\leq d(y_0, Y_1^t)^p \mathbf{1}_{\{d(y_0, Y_1^t) < K\} \cap L} + K^p \mathbf{1}_{L^c} \\ &\leq 2^{p/2} \left( C_{\nu} \|\mathbf{B}_{0,1}\|_{\text{CC}} t^H + K C_{\nu, \bar{l}}^{1/\bar{l}} \|\mathbf{B}\|_{p\text{-var}; [0, 1]}^{(\bar{l}+1)/\bar{l}} t^{H(\bar{l}+1)/\bar{l}} \right)^p \\ &\quad + K^p \left( \|\mathbf{B}_{0,1}\|_{\text{CC}}^p t^{Hp} + C_{\nu, \bar{l}}^p \|\mathbf{B}\|_{p\text{-var}; [0, 1]}^{(\bar{l}+1)p} t^{H(\bar{l}+1)p} \right). \end{aligned}$$

Since  $0 \leq t \leq T$  and  $\|\mathbf{B}_{0,1}\|_{\text{CC}} \leq \|\mathbf{B}\|_{p\text{-var}; [0, 1]}$ , there exists a constant  $C_3 := C(H, \nu, p, T)$  such that

$$d(y_0, Y_1^t)^p \mathbf{1}_{\{\|Y_1^t - y_0\|_{\mathbb{R}^N} < 1\}} \leq C_3 \left( 1 + \|\mathbf{B}\|_{p\text{-var}; [0, 1]}^{(\bar{l}+1)p} \right) t^{Hp}. \quad (5.44)$$

Now one concludes from (5.39) and (5.44) that

$$\begin{aligned} d(y_0, Y_1^t)^p &= d(y_0, Y_1^t)^p \mathbf{1}_{\{\|Y_1^t - y_0\|_{\mathbb{R}^N} < 1\}} + d(y_0, Y_1^t)^p \mathbf{1}_{\{\|Y_1^t - y_0\|_{\mathbb{R}^N} \geq 1\}} \\ &\leq (2^{3/2} K)^p t^{\frac{3pH}{2}} \left| \int_0^1 V(Y_r^t) dB_r \right|^{\frac{3p}{2}} + C_3 \left( 1 + \|\mathbf{B}\|_{p\text{-var}; [0, 1]}^{(\bar{l}+1)p} \right) t^{Hp}. \end{aligned}$$

Clearly,  $\|\mathbf{B}\|_{p\text{-var};[0,1]}$  has finite moments of all orders. The rough integral  $\int_0^1 V(Y_r^t) dB_r$  also has finite moments of all orders (uniformly in  $t \in [0, T]$ ) due to [13, Theorem 10.36, Theorem 10.47]. One therefore arrives at the following estimate

$$\mathbb{E}[d(y_0, Y_t)^p] = \mathbb{E}[d(y_0, Y_1^t)^p] \leq C_{H, \mathcal{V}, p, T} t^{Hp}.$$

This proves the first part of the lemma. The second part follows from Kolmogorov's continuity theorem.  $\square$

We are now in a position to state and prove the main result of this section.

**Theorem 5.8.** *Under the assumptions in Theorem 5.4, let  $0 < a < b$  and  $M > 0$  be given fixed. For any  $\eta > 0$  (if  $Q - 1/H > 0$ , we further require  $\eta < Q - 1/H$ ), there exists a positive constant  $C = C(a, b, M, Q, H, \mathcal{V}, \eta)$  such that*

$$\mathbb{P}\left(Y([a, b]) \cap A \neq \emptyset\right) \leq C \text{Cap}_{Q-1/H-\eta}(A) \quad (5.45)$$

for all compact sets  $A \subseteq B_d(o, M)$ .

*Proof.* According to Lemma 5.6, it suffices to verify that  $(Y_t)_{0 \leq t \leq T}$  satisfies Assumption (A3) for any fixed  $\beta \in (0, Q)$ . Fix  $z \in B_d(o, M)$ . For any  $\varepsilon \in (0, 1)$  and  $s, t \in [a, b]$  satisfying  $|t - s| = \varepsilon^{1/H}$ , we define

$$U := d(Y_s, z), \quad \text{and} \quad R := \sup_{r, v \in [s, t]} d(Y_r, Y_v).$$

One easily checks by using the triangle inequality that

$$\mathbb{P}\left(Y([s, t]) \cap B_d(z, \varepsilon) \neq \emptyset\right) \leq \mathbb{P}(U \leq \varepsilon + R). \quad (5.46)$$

The right hand side is estimated as

$$\begin{aligned} \mathbb{P}(U \leq \varepsilon + R) &\leq \mathbb{P}(U \leq \varepsilon + R, U \geq 2\varepsilon) + \mathbb{P}(U < 2\varepsilon) \\ &\leq \mathbb{P}(U/2 \leq R) + \mathbb{P}(U < 2\varepsilon), \end{aligned} \quad (5.47)$$

where the second inequality uses the observation that if  $U \leq \varepsilon + R$  and  $U \geq 2\varepsilon$ , then  $\varepsilon \leq R$ , which further implies  $U \leq 2R$  and hence  $U/2 \leq R$ . Now for fix  $\alpha \in (0, 1)$ , one has (5.47):

$$\begin{aligned} \mathbb{P}(U/2 \leq R) &\leq \mathbb{P}(U/2 \leq R, R \leq \varepsilon^\alpha) + \mathbb{P}(R > \varepsilon^\alpha) \\ &\leq \mathbb{P}(U/2 \leq \varepsilon^\alpha) + \mathbb{P}(R > \varepsilon^\alpha). \end{aligned} \quad (5.48)$$

Combining (5.47) and (5.48), one obtains the following estimate:

$$\mathbb{P}(U \leq \varepsilon + R) \leq \mathbb{P}(U/2 \leq \varepsilon^\alpha) + \mathbb{P}(R > \varepsilon^\alpha) + \mathbb{P}(U < 2\varepsilon).$$

We will estimate the above three terms separately.

First of all, recall that  $Y_s$  admits a smooth density for all  $0 < s \leq T$ . Then there exists a constant  $C_1 = C_{a,b,M} > 0$  such that

$$\sup_{x \in B_d(o,M)} |p_s(x)| \leq C_1$$

holds for any  $s \in [a, b]$ . By using Lemma 5.1, one therefore obtains that

$$\begin{aligned} \mathbb{P}(U/2 \leq \varepsilon^\alpha) &\leq C_1 |B_d(z, 2\varepsilon^\alpha)| \leq C_1 C_{2,\mathcal{V},M} \cdot 2^Q \varepsilon^{\alpha Q}, \\ \mathbb{P}(U < 2\varepsilon) &\leq C_1 |B_d(z, 2\varepsilon)| \leq C_1 C_{2,\mathcal{V},M} \cdot 2^Q \varepsilon^Q, \end{aligned}$$

where  $C_{2,\mathcal{V},M}$  is the constant appearing in the volume comparison (5.3).

In addition, let  $q \in (0, 1 - \alpha)$  and  $p \geq 1$ . By using Lemma 5.7 and Chebyshev's inequality, one finds that

$$\mathbb{P}(R > \varepsilon^\alpha) \leq \mathbb{P}\left(\sup_{r,v \in [s,t]} \frac{d(Y_r, Y_v)}{|v - r|^{H(\alpha+q)}} > \varepsilon^{-q}\right) \leq C_2 \varepsilon^{qp},$$

where  $|s - t| = \varepsilon^{1/H}$  and  $C_2$  depends only on  $H, \mathcal{V}, p, b, \eta$ . Now fix any  $\beta \in (0, Q)$ . By choosing  $\alpha$  and  $p$  sufficiently large such that  $\alpha Q \geq \beta$  and  $qp \geq \beta$ , one concludes that

$$\begin{aligned} \mathbb{P}(Y([s, t]) \cap B_d(z, \varepsilon) \neq \emptyset) &\leq \mathbb{P}(U \leq \varepsilon + R) \\ &\leq \mathbb{P}(U/2 \leq \varepsilon^\alpha) + \mathbb{P}(R > \varepsilon^\alpha) + \mathbb{P}(U < 2\varepsilon) \\ &\leq 2^Q C_1 C_{2,\mathcal{V},M} (\varepsilon^{\alpha Q} + \varepsilon^Q) + C_2 \varepsilon^{qp} \\ &\leq C \varepsilon^\beta, \end{aligned}$$

where  $C$  depends on  $C_1, C_2$ , and  $C_{2,\mathcal{V},M}$ . This yields the desired result.  $\square$

## A Proof of Lemma 4.8

In this appendix, we prove Lemma 4.8. As a starting point, we make the following observations.

- Using the stationarity of increments proved by Lemma 2.15, we may fix  $s = 0$  without loss of generality. More precisely, Lemma 2.15, combined with Lemma 2.17, gives that

$$\|\gamma(\mathbf{U}_{s,s+\varepsilon})^{-1}\|_{k,p,s} \stackrel{M}{=} \|\gamma(\mathbf{U}_{0,\varepsilon})^{-1}\|_{k,p,0}.$$

- By the scaling property of fBm, it suffices to consider  $t = 1$  by introducing the rescaled dynamics over  $r \in [0, 1]$ :

$$d\mathbf{U}_r^{\varepsilon,x} = \sum_{i=1}^d A_i^\varepsilon(\mathbf{U}_r^{\varepsilon,x}) dB_r^i = \sum_{i=1}^d \varepsilon^H A_i(\mathbf{U}_r^{\varepsilon,x}) dB_r^i, \quad \mathbf{U}_0^{\varepsilon,x} = x \in \mathfrak{g}^{(l)}(\mathbb{R}^d). \quad (\text{A.1})$$

Here  $A_i$  is the left invariant vector field on  $\mathfrak{g}^{(l)}(\mathbb{R}^d)$  induced by the canonical basis vector  $e_i$  and note that the vector fields  $\{A_i\}_{i=1}^d$  satisfy Hörmander's condition on  $\mathfrak{g}^{(l)}(\mathbb{R}^d)$ . The vector fields  $A_i^\varepsilon$  are defined by  $A_i^\varepsilon(u) = \varepsilon^H A_i(u)$ . We use  $o \in \mathfrak{g}^{(l)}(\mathbb{R}^d)$  to denote the origin and to ease notation, we will use  $\mathbf{U}_t^\varepsilon$  instead of  $\mathbf{U}_t^{\varepsilon,o}$  when the initial value is  $o$ . Note that  $\mathbf{U}_1^\varepsilon = \delta_{\varepsilon^H} \mathbf{U}_{0,1} \stackrel{d}{=} \mathbf{U}_{0,\varepsilon}$ . This formulation allows one to reduce the analysis to the unit time interval  $[0, 1]$ , where the dependence on  $\varepsilon$  encodes small-time behaviour in the original problem.

Based on Lemma 2.17, one can apply a similar argument as in the derivation of (4.10) to show that

$$\|\gamma(\mathbf{U}_{s,s+\varepsilon})^{-1}\|_{k,p,s} \stackrel{M}{=} \|\gamma(\mathbf{U}_1^\varepsilon)^{-1}\|_{k,p,0}$$

for any  $s \geq 0, \varepsilon > 0, k \geq 1$  and  $p > 0$ . Now the task is reduced to showing that

$$\|\gamma(\mathbf{U}_1^\varepsilon)^{-1}\|_{k,p,0} \leq C\varepsilon^{-\alpha}\Psi \quad (\text{A.2})$$

for all  $0 < \varepsilon \leq T$  with some smooth random variable  $\Psi$ .

First of all, it is well-known that for any  $\varepsilon \in (0, 1]$  and  $t > 0$ , the map  $x \rightarrow \mathbf{U}_t^{\varepsilon,x}$  is a flow of  $C^\infty$  diffeomorphism. We denote the Jacobian of  $\mathbf{U}_t^{\varepsilon,x}$  by  $\mathbf{J}_t^{\varepsilon,x} = \partial_x \mathbf{U}_t^{\varepsilon,x}$  and let  $\beta_I^{J,\varepsilon}(t, x)$  ( $I, J \in \mathcal{A}_1(l)$ ) satisfy

$$(\mathbf{J}_t^{\varepsilon,x})^{-1} A_I^\varepsilon(\mathbf{U}_t^{\varepsilon,x}) = \sum_{J \in \mathcal{A}_1(l)} \beta_I^{J,\varepsilon}(t, x) A_J^\varepsilon(x), \quad 0 \leq t \leq 1. \quad (\text{A.3})$$

The precise definition of  $\{\beta_I^{J,\varepsilon}(\cdot, \cdot)\}_{I, J \in \mathcal{A}_1(l)}$  can be found in [5, Equation (3.5)]. We define the Malliavin matrix  $\mathbf{M}^\varepsilon(x)$  by setting

$$M_{I,J}^\varepsilon = (\mathbf{M}^\varepsilon(x))_{I,J} = \langle \mathcal{L}(\beta^{\varepsilon,I}(\cdot, x)), \mathcal{L}(\beta^{\varepsilon,J}(\cdot, x)) \rangle_{\mathcal{H}_{W,0}} \quad (\text{A.4})$$

for any  $1 \leq I, J \leq \mathcal{A}_1(l)$ , where  $\mathcal{L}$  is defined in (2.7) and  $\beta^{\varepsilon,I}(\cdot, x)$  means the column vector  $(\beta_1^{\varepsilon,I}(\cdot, x), \dots, \beta_d^{\varepsilon,I}(\cdot, x))$ . The following result is taken from [5, Lemma 3.9].

**Lemma A.1.** *Let  $\mathbf{U}^{\varepsilon,x}$ ,  $\beta_I^{J,\varepsilon}$  and  $M_{I,J}^\varepsilon$  defined by (A.1), (A.3) and (A.4) respectively. Then one has*

$$\lambda_{\max}(\gamma(\mathbf{U}_1^{\varepsilon,x})^{-1}) \leq C\varepsilon^{-2Hl} \lambda_{\max}((\mathbf{M}^\varepsilon(x))^{-1}) \lambda_{\max}((\mathbf{J}_1^{\varepsilon,x}(\mathbf{J}_1^{\varepsilon,x})^*)^{-1}),$$

where  $C$  is a positive constant independent of  $\varepsilon$  and  $\lambda_{\max}(\cdot)$  means taking the maximal eigenvalue of a matrix.

In view of the above Lemma, it remains to control the eigenvalues  $\lambda_{\max}((\mathbf{M}^\varepsilon(o))^{-1})$  and  $\lambda_{\max}((\mathbf{J}_1^{\varepsilon,o}(\mathbf{J}_1^{\varepsilon,o})^*)^{-1})$  individually. The latter is easier which we shall first address.

**Proposition A.2.** *There exists a smooth random variable  $\Psi$  such that*

$$\lambda_{\max}((\mathbf{J}_1^{\varepsilon,o}(\mathbf{J}_1^{\varepsilon,o})^*)^{-1}) \leq \Psi,$$

for any  $\varepsilon \in (0, 1]$ .

*Proof.* It is enough to prove that each entry of  $(\mathbf{J}_1^{\varepsilon,o})^{-1}$  is bounded by certain smooth random variable  $\Psi$ . Note that  $\mathbf{U}_t^{\varepsilon,x} = x \star \mathbf{U}_t^{\varepsilon,o}$  where  $\star$  is the group law induced from  $G^{(l)}(\mathbb{R}^d)$ . Since the inverse of  $x \mapsto x \star \mathbf{U}_1^{\varepsilon,o}$  is  $y \mapsto y \star (\mathbf{U}_1^{\varepsilon,o})^{-1} =: H(y)$ , it follows that

$$(\mathbf{J}_1^{\varepsilon,o})^{-1} = \frac{H(y)}{\partial y} \Big|_{y=\mathbf{U}_1^{\varepsilon,o}}$$

Observe that every component of  $\mathbf{U}_1^{\varepsilon,o}$  and  $(\mathbf{U}_1^{\varepsilon,o})^{-1}$  is a polynomial of

$$\varepsilon^{Hm} \int_{0 < t_1 < \dots < t_m < 1} dB_{t_1}^{i_1} \dots dB_{t_m}^{i_d}.$$

As in the argument of Lemma 4.5, this can be bounded above by the random variable  $\exp(\eta (\mathcal{N}_{\gamma,2q}^{0,1}(\mathbf{B}))^{1/q})$  for some  $q \geq 1$ ,  $\eta \in (0, \eta_0]$  and  $\gamma < H$ .  $\square$

The analysis of  $\lambda_{\max}((\mathbf{M}^\varepsilon(o))^{-1})$  is more involved and we need several technical lemmas.

**Lemma A.3.** *Let  $I, J \in \mathcal{A}_1(l)$  with  $|I| \leq |J|$ . Then there is no  $(\varepsilon, x)$ -dependence in  $\beta_I^{J,\varepsilon}$  and more specifically, one has*

$$\beta_I^{J,\varepsilon}(t, x) = \begin{cases} (-1)^{|K|} B_t^K, & \text{if } J = (I, K) \text{ for some } K \in \mathcal{A}_1(l), \\ 0, & \text{otherwise.} \end{cases}$$

Here  $B_t^K$  denotes the projection of  $\mathbf{B}_{0,t}$  onto the tensor basis  $e_{k_1} \otimes \dots \otimes e_{k_m}$  for any word  $K = (k_1, \dots, k_m)$ .

*Proof.* This is just Lemma 4.2 and Remark 4.3 of [5] applied to the special context of the truncated signature process.  $\square$

We define the supreme norm

$$\|f\|_{\infty,[s,t]} := \sup_{u \in [s,t]} \|f(u)\|_{\text{HS}}$$

for any continuous  $f$  taking values in  $T^{(l)}(\mathbb{R}^d)$ .

**Lemma A.4.** *For any  $m \in \mathbb{N}$  and  $p \geq 1$ , there exists a smooth function  $\Psi_{m,p}$  such that*

$$\sup_{\sum_{I \in \mathcal{A}(m)} a_I^2 = 1} \mathbb{P} \left( \left\| \sum_{I \in \mathcal{A}(m)} a_I B_t^I \right\|_{\infty,[0,1]} < \varepsilon \left| \mathcal{F}_0^W \right. \right) \leq \varepsilon^p \Psi_{m,p}$$

for all  $\varepsilon > 0$ .

*Proof.* When  $m = 0, 1$ , one can use the same argument as in [5, Lemma 4.4], with the small ball estimate replaced by the corresponding conditional version given by Lemma 2.18, to show that

$$\sup_{\sum_{I \in \mathcal{A}(m)} a_I^2 = 1} \mathbb{P} \left( \left\| \sum_{I \in \mathcal{A}(m)} a_I B_t^I \right\|_{\infty, [0,1]} < \varepsilon \mid \mathcal{F}_0^W \right) \leq C_{p,H} \varepsilon^p$$

with some constant  $C_{p,H}$  only depending on  $p, H$ .

We prove the general case by induction. Assume that the result holds for  $k = 0, 1, \dots, m$ . We only consider the case  $a_\emptyset = 0$ ; the case  $a_\emptyset \neq 0$  can be handled similarly as in [5, Lemma 4.4]. Let  $f(t) = \sum_{I \in \mathcal{A}_1(m+1)} a_I B_t^I$  with  $\sum_{I \in \mathcal{A}_1(m+1)} a_I^2 = 1$ . One can write

$$f(t) = \int_0^t A_s dB_s,$$

where  $B_t = (B_t^1, \dots, B_t^d)$  and  $A_t = (\sum_{J \in \mathcal{A}(m)} a_{J*1} B_t^J, \dots, \sum_{J \in \mathcal{A}(m)} a_{J*d} B_t^J)$ . We choose a  $k$ ,  $1 \leq k \leq d$ , such that  $\sum_{J \in \mathcal{A}(m)} a_{J*k}^2 \geq 1/d$ . According to the pathwise Norris' lemma ([10, Theorem 5.6]),

$$\left\| \sum_{J \in \mathcal{A}(m)} a_{J*k} B_t^J \right\|_{\infty, [0,1]} \leq CR^\eta \|f(t)\|_{\infty, [0,1]}^r$$

where  $C, \eta > 0$ ,

$$R^\eta \lesssim 1 + L_\theta^{-\eta} + \exp \left( \zeta (\mathcal{N}_{\gamma, 2q}^{0,1}(\mathbf{B}))^{1/q} \right)$$

for some  $\zeta \in (0, \eta_0]$  (see (4.3) for  $\eta_0$ ) and  $L_\theta$  is the  $\theta$ -Hölder roughness of  $B$  with  $2\gamma = \theta \in (H, 2H)$ . It follows that

$$\begin{aligned} & \mathbb{P} \left( \|f(t)\|_{\infty, [0,1]} < \varepsilon \mid \mathcal{F}_0^W \right) \\ & \leq \mathbb{P} \left( \frac{\left\| \sum_{J \in \mathcal{A}(m)} a_{J*k} B_t^J \right\|_{\infty, [0,1]}}{CR^\eta} < \varepsilon^r \mid \mathcal{F}_0^W \right) \\ & \leq \mathbb{P} \left( \left\| \sum_{J \in \mathcal{A}(m)} a_{J*k} B_t^J \right\|_{\infty, [0,1]} < \varepsilon^{r/2} \mid \mathcal{F}_0^W \right) + \mathbb{P} (CR^\eta \geq \varepsilon^{-r/2} \mid \mathcal{F}_0^W) \\ & \leq \mathbb{P} \left( \left\| \sum_{J \in \mathcal{A}(m)} \frac{a_{J*k}}{\sqrt{\sum a_{J*k}^2}} B_t^J \right\|_{\infty, [0,1]} < \sqrt{d} \varepsilon^{r/2} \mid \mathcal{F}_0^W \right) + \mathbb{P} (CR^\eta \geq \varepsilon^{-r/2} \mid \mathcal{F}_0^W). \end{aligned}$$

The first term is bounded above by  $\varepsilon^p \Psi_p$  due to the induction hypothesis.

For the second term, one has

$$\begin{aligned} & \mathbb{P} (CR^\eta \geq \varepsilon^{-r/2} \mid \mathcal{F}_0^W) \lesssim \varepsilon^p \mathbb{E} \left[ R^{\frac{2p\eta}{r}} \mid \mathcal{F}_0^W \right] \\ & \lesssim \varepsilon^p \left( 1 + \mathbb{E} \left[ L_\theta^{-\frac{2p\eta}{r}} \mid \mathcal{F}_0^W \right] + \mathbb{E} \left[ \exp \left( \frac{2p\zeta}{r} (\mathcal{N}_{\gamma, 2q}^{0,1}(\mathbf{B}))^{1/q} \right) \mid \mathcal{F}_0^W \right] \right) \end{aligned}$$

One can use Proposition 2.19 (in the same way as Lemma 5.8 and Corollary 5.10 in [10]) to conclude that

$$\mathbb{P}(L_\theta < x \mid \mathcal{F}_0^W) \leq C_{1,p,H} \exp(-C_{2,p,H} x^{-1/H}), \quad 0 < x < 1$$

for some constants  $C_{1,p,H}, C_{2,p,H} > 0$ . This further implies that

$$\mathbb{E} \left[ L_\theta^{-\frac{2p\eta}{r}} \mid \mathcal{F}_0^W \right] \leq C_{p,H}.$$

Note that  $\mathbb{E} \left[ \exp \left( \frac{2p\zeta}{r} (\mathcal{N}_{\gamma,2q}^{0,1}(\mathbf{B}))^{1/q} \right) \mid \mathcal{F}_0^W \right]$  is similar to  $\Psi^1$  in Lemma 4.5, so one can choose  $0 < \zeta \leq \frac{\eta_0 r}{2p}$  and use the same argument as the one leading to (4.8) to show that it is a smooth random variable. Therefore, the conclusion holds for  $k = m + 1$ , which completes the induction step.  $\square$

**Corollary A.5.** *For any  $m \geq 0$  and  $p > 1$ , one has*

$$\mathbb{E} \left[ \inf \left\{ \int_0^1 \left( \sum_{I \in \mathcal{A}(m)} a_I B_t^I \right)^2 dt ; \sum_{I \in \mathcal{A}(m)} a_I^2 = 1 \right\}^{-p} \mid \mathcal{F}_0^W \right] \leq \Psi_{m,p}, \quad (\text{A.5})$$

where  $\Psi_{m,p}$  is a smooth random variable depending only on  $m, p$ .

*Proof.* An unconditional version of this result was obtained in [5, Corollary 4.5]. Here we use a similar argument but with all the estimates replaced by the conditional ones. We first show that

$$\sup_{\sum_{I \in \mathcal{A}(m)} a_I^2 = 1} \mathbb{P} \left( \int_0^1 \left( \sum_{I \in \mathcal{A}(m)} a_I B_t^I \right)^2 dt < \varepsilon \mid \mathcal{F}_0^W \right) \leq \varepsilon^p \Psi_{m,p} \quad (\text{A.6})$$

for all  $\varepsilon > 0$  with some smooth random variable  $\Psi_{m,p}$ . Setting  $f(t) = \sum_{I \in \mathcal{A}(m)} a_I B_t^I$ , one has

$$\mathbb{P} \left( \int_0^1 \left( \sum_{I \in \mathcal{A}(m)} a_I B_t^I \right)^2 dt < \varepsilon \mid \mathcal{F}_0^W \right) = \mathbb{P}(\|f\|_{L^2}^2 < \varepsilon \mid \mathcal{F}_0^W) = \mathbb{P}(\|f\|_{L^2} < \sqrt{\varepsilon} \mid \mathcal{F}_0^W).$$

By using the interpolation inequality in [10, Equation (6.5)], i.e.

$$\|f\|_\infty \leq 2 \max \left\{ \|f\|_{L^2}, \|f\|_{L^2}^{\frac{2r}{2r+1}} \|f\|_r^{\frac{1}{2r+1}} \right\},$$

one finds that

$$\begin{aligned} \{ \|f\|_{L^2} < \sqrt{\varepsilon} \} &\subseteq \left\{ \frac{\|f\|_\infty}{2} < \sqrt{\varepsilon}, \|f\|_{L^2} > \|f\|_r \right\} \\ &\cup \left\{ \left( \frac{\|f\|_\infty}{2\|f\|_r^{\frac{1}{2r+1}}} \right)^{\frac{2r+1}{2r}} < \sqrt{\varepsilon}, \|f\|_{L^2} < \|f\|_r \right\}. \end{aligned}$$

As a consequence,

$$\begin{aligned}
& \mathbb{P}(\|f\|_{L^2} < \sqrt{\varepsilon} \mid \mathcal{F}_0^W) \\
& \leq \mathbb{P}(\|f\|_\infty < 2\sqrt{\varepsilon} \mid \mathcal{F}_0^W) + \mathbb{P}\left(\|f\|_\infty^{\frac{2r+1}{2r}} < \varepsilon^{1/4} \mid \mathcal{F}_0^W\right) \\
& \quad + \mathbb{P}\left(\left(2\|f\|_r^{\frac{1}{2r+1}}\right)^{\frac{2r+1}{2r}} > \varepsilon^{-1/4} \mid \mathcal{F}_0^W\right) \\
& \leq \mathbb{P}(\|f\|_\infty < 2\sqrt{\varepsilon} \mid \mathcal{F}_0^W) + \mathbb{P}\left(\|f\|_\infty < \varepsilon^{\frac{r}{4r+2}} \mid \mathcal{F}_0^W\right) + \mathbb{P}(\|f\|_r > 2^{-2r-1}\varepsilon^{-r/2} \mid \mathcal{F}_0^W)
\end{aligned}$$

The first and second terms on the right hand side are estimated by Lemma A.4 and the third term is upper bounded by smooth random variable due to (4.4).

To obtain the desired estimate (A.5) from (A.6), one only needs to apply a standard compactness argument as in [28, Lemma 2.3.1].  $\square$

Now we can derive the key nondegeneracy estimate for the Malliavin matrix  $\mathbf{M}^\varepsilon(x)$ .

**Proposition A.6.** *For any  $p \in (1, \infty)$ , there exists a smooth function  $\Psi_{\bar{l}, p}$  depending only on  $\bar{l}, p$  such that*

$$\mathbb{E}[\lambda_{\max}((\mathbf{M}^\varepsilon(x))^{-1})^p \mid \mathcal{F}_0^W] \leq \Psi_{\bar{l}, p}$$

holds for all  $\varepsilon \in (0, 1)$ .

*Proof.* Recalling the definition of  $(\mathbf{M}^\varepsilon(x))_{I, J}$  in (A.4) and the definition of  $\beta_I^J(t, x)$  in Lemma A.3, one has

$$\begin{aligned}
(\mathbf{M}^\varepsilon(x))_{I, J} &= \langle \mathcal{L}(\beta^{\varepsilon, I}(\cdot, x)), \mathcal{L}(\beta^{\varepsilon, J}(\cdot, x)) \rangle_{\mathcal{H}_{W, 0}} \\
&= (-1)^{|I|+|J|} \delta_{i_1}^{j_1} \langle \mathcal{L}(B^I), \mathcal{L}(B^J) \rangle_{\mathcal{H}_{W, 0}} := M_{I, J}.
\end{aligned}$$

We first assume  $1/4 < H \leq 1/2$ . For any multi-index  $a$  such that  $\sum_{I \in \mathcal{A}_1(\bar{l})} a_I^2 = 1$ , one has

$$\begin{aligned}
& \sum_{I, J \in \mathcal{A}_1(\bar{l})} (-1)^{|I|+|J|} a_I a_J M_{I, J} \\
&= \sum_{i=1}^d \left\langle \mathcal{L}\left(\sum_{I \in \mathcal{A}(\bar{l}-1)} a_{(i, I)} B^{(i, I)}\right), \mathcal{L}\left(\sum_{J \in \mathcal{A}(\bar{l}-1)} a_{(i, J)} B^{(i, J)}\right) \right\rangle_{\mathcal{H}_{W, 0}} \\
&\geq C_H \sum_{i=1}^d \left\| \sum_{J \in \mathcal{A}(\bar{l}-1)} a_{(i, J)} B^{(i, J)} \right\|_{\mathcal{H}_B} \geq C_H \sum_{i=1}^d \int_0^1 \left( \sum_{J \in \mathcal{A}(\bar{l}-1)} a_{(i, J)} B_t^{(i, J)} \right)^2 dt,
\end{aligned}$$

where the second line follows from Lemma 2.13, Lemma 2.14 and the third line is implied by the inequality  $\|f\|_{\mathcal{H}_B} \geq C\|f\|_{L^2}$  for every  $f \in \mathcal{H}_B$  (see [5, P. 414]). Let  $k$  be such that  $\sum_{J \in \mathcal{A}(\bar{l}-1)} a_{(k, J)}^2 \geq 1/d$ . Then one has

$$\sqrt{d} \sum_{I, J \in \mathcal{A}_1(\bar{l})} (-1)^{|I|+|J|} a_I a_J M_{I, J} \geq C_H \int_0^1 \left( \sum_{J \in \mathcal{A}(\bar{l}-1)} \frac{a_{(k, J)}}{\sqrt{\sum_{J \in \mathcal{A}(\bar{l}-1)} a_{(k, J)}^2}} B_t^{(k, J)} \right)^2 dt.$$

It follows that

$$\begin{aligned} & \mathbb{E} \left[ \lambda_{\max} \left( (M^\varepsilon(x))^{-1} \right)^p \mid \mathcal{F}_0^W \right] \\ & \lesssim \mathbb{E} \left[ \inf \left\{ \int_0^1 \left( \sum_{I \in \mathcal{A}(\bar{l})} a_I B_t^I \right)^2 dt ; \sum_{I \in \mathcal{A}(\bar{l})} a_I^2 = 1 \right\}^{-p} \mid \mathcal{F}_0^W \right] \leq \Psi_{\bar{l}, p}. \end{aligned}$$

If  $1/2 < H < 1$ , one uses the interpolation inequality  $\|f\|_{\mathcal{H}_B} \geq C \frac{\|f\|_\infty^{3+1/\alpha}}{\|f\|_\alpha^{2+1/\gamma}}$  instead (See [2, Lemma 4.4]). The similar analysis will be omitted for conciseness.  $\square$

*Proof of Lemma 4.8.* The conclusion follows from Lemma A.1, Proposition A.2 and Proposition A.6.  $\square$

## Acknowledgement

XG gratefully acknowledges the support from ARC grant DE210101352. SW gratefully acknowledges the support from Melbourne Research Scholarship.

## References

- [1] A. Agrachev, D. Barilari, and U. Boscain. *A comprehensive introduction to sub-Riemannian geometry*. Cambridge University Press, 2019.
- [2] F. Baudoin and M. Hairer. “A version of Hörmander’s theorem for the fractional Brownian motion”. In: *Probab. Theory Relat. Fields*. 139.3 (2007), pp. 373–395.
- [3] F. Baudoin and D. Nualart. “Equivalence of Volterra processes”. In: *Stoch Proc Appl*. 107.2 (2003), pp. 327–350.
- [4] F. Baudoin, C. Ouyang, and S. Tindel. “Upper bounds for the density of solutions to stochastic differential equations driven by fractional Brownian motions”. In: *Ann. Inst. Henri Poincaré Probab. Stat.* 50.1 (2014), pp. 111–135.
- [5] F. Baudoin, C. Ouyang, and X. Zhang. “Smoothing effect of rough differential equations driven by fractional Brownian motions”. In: *Ann. Inst. Henri Poincaré Probab. Stat.* (2016), pp. 412–428.
- [6] F. Baudoin et al. “On probability laws of solutions to differential systems driven by a fractional Brownian motion”. In: *Ann. Probab.* (2016), pp. 2554–2590.
- [7] H. Boedihardjo and X. Geng. “On the lack of Gaussian tail for rough line integrals along fractional Brownian paths”. In: *Probab. Theory Relat. Fields*. 188.3 (2024), pp. 1287–1313.
- [8] T. Cass and P.K. Friz. “Densities for rough differential equations under Hörmander’s condition”. In: *Ann. of Math.* (2010), pp. 2115–2141.

- [9] T. Cass, C. Litterer, and T. Lyons. “Integrability and tail estimates for Gaussian rough differential equations”. In: *Ann. Probab.* (2013), pp. 3026–3050.
- [10] T. Cass et al. “Smoothness of the density for solutions to Gaussian rough differential equations”. In: *Ann. Probab.* (2015), pp. 188–239.
- [11] R.C. Dalang, D. Khoshnevisan, and E. Nualart. “Hitting probabilities for systems of non-linear stochastic heat equations with additive noise”. In: *ALEA Lat. Am. J. Probab. Math. Stat.* (2007), pp. 231–271.
- [12] P. Friz and N. Victoir. “A variation embedding theorem and applications”. In: *J. Funct. Anal.* 239.2 (2006), pp. 631–637.
- [13] P.K Friz and N.B. Victoir. *Multidimensional stochastic processes as rough paths: theory and applications*. Cambridge University Press, 2010.
- [14] P.K. Friz and N.B. Victoir. “Differential equations driven by Gaussian signals”. In: *Ann. Inst. Henri Poincaré Probab. Stat.* 46 (2010), pp. 369–413.
- [15] P.K. Friz and N.B. Victoir. *Multidimensional stochastic processes as rough paths: theory and applications*. Cambridge University Press, 2010.
- [16] M. Fukushima. “Basic properties of Brownian motion and a capacity on the Wiener space”. In: *J. Math. Soc. Japan* 36.1 (1984), pp. 161–176.
- [17] X. Geng, C. Ouyang, and S. Tindel. “Precise local estimates for differential equations driven by fractional Brownian motion: hypoelliptic case”. In: *Ann. Probab.* 50.2 (2022), pp. 649–687.
- [18] M. Hairer and N.S. Pillai. “Regularity of laws and ergodicity of hypoelliptic SDEs driven by rough paths”. In: *Ann. Probab.* (2013).
- [19] L. Hörmander. “Hypoelliptic second order differential equations”. In: *Acta Math.* 119 (1967), pp. 147–171.
- [20] Y. Hu and D. Nualart. “Differential equations driven by Hölder continuous functions of order greater than  $1/2$ ”. In: *Stochastic Analysis and Applications: The Abel Symposium 2005*. Springer. 2007, pp. 399–413.
- [21] Y. Inahama. “Malliavin differentiability of solutions of rough differential equations”. In: *J. Funct. Anal.* 267.5 (2014), pp. 1566–1584.
- [22] S. Kusuoka and D. Stroock. “Applications of the Malliavin calculus, Part III”. In: *J. Fac. Sci. Univ. Tokyo* 34 (1987), pp. 391–442.
- [23] T.J. Lyons. “Differential equations driven by rough signals (I): An extension of an inequality of LC Young”. In: *Math. Res. Lett.* 1.4 (1994).
- [24] T.J. Lyons. “Differential equations driven by rough signals”. In: *Rev. Mat. Iberoamericana* 14.2 (1998), pp. 215–310.
- [25] P. Malliavin. “Stochastic calculus of variation and hypoelliptic operators”. In: *Proc. Intern. Symp. SDE Kyoto 1976*. Kinokuniya. 1978, pp. 195–263.

- [26] P. Mattila. *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*. 44. Cambridge university press, 1999.
- [27] J. Norris. “Simplified Malliavin calculus”. In: *Séminaire de Probabilités XX 1984/85: Proceedings*. Springer, 2006, pp. 101–130.
- [28] D. Nualart. *The Malliavin calculus and related topics*. Springer-Verlag, 2006.
- [29] P. Paulovics. *Hormander’s theorem for Gaussian rough differential equations, Master thesis*. Master Thesis, Warwick University, 2023.
- [30] Y. Xiao. “Sample path properties of anisotropic Gaussian random fields”. In: *A mini-course on stochastic partial differential equations. Lecture Notes in Mathematics*. Vol. 1962. 2009, pp. 145–212.