Non-degeneracy of Stochastic Line Integrals

Xi Geng* and Sheng Wang†

Abstract

We derive quantitative criteria for the existence of density for stochastic line integrals and iterated line integrals along solutions of hypoelliptic differential equations driven by fractional Brownian motion. As an application, we also study the signature uniqueness problem for these rough differential equations.

1 Introduction and summary of main results

It is classical that there is a natural pairing between a $\mathcal{C}^1$-path $\gamma : [0,T] \to M$ in a differentiable manifold M and a differential one-form ϕ on M, which is defined by integration:

$$\int_0^T \phi(d\gamma_t) \triangleq \int_0^T \langle \phi, \dot{\gamma}_t \rangle dt.$$

Here $\langle \cdot, \cdot \rangle$ denotes the pairing between cotangent and tangent vectors. This notion of integration, sometimes known as *line integrals*, has an intrinsic geometric meaning in the sense that it does not rely on local coordinates or embeddings of M into ambient Euclidean spaces. More generally, given a finite sequence of one-forms $(\phi_1, \cdots, \phi_m)$, one can consider an associated iterated line integral

$$\int_{0<t_1<\cdots<t_m<T} \phi_1(d\gamma_{t_1}) \cdots \phi_m(d\gamma_{t_m}) \triangleq \int_0^T \int_0^{t_m} \cdots \int_0^{t_2} \langle \phi_1, \dot{\gamma}_{t_1} \rangle \cdots \langle \phi_m, \dot{\gamma}_{t_m} \rangle dt_1 \cdots dt_m.$$

*School of Mathematics and Statistics, University of Melbourne, Parkville VIC 3010, Australia. Email: xi.geng@unimelb.edu.au. XG acknowledges the support from ARC Grant DE210101352.

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

The definition of such integrals can be naturally extended to the rough path context under suitable regularity conditions on the path γ and the one-forms (cf. [LQ02, CDL15]). In the rough path literature, these iterated line integrals are often referred to as *extended signatures* of γ (cf. [LQ12] for their use in the context of Brownian motion).

A natural reason of considering line integrals is that they encode rich geometric/topological information about the original path γ. For instance, if $\gamma = (x_t, y_t)$ is a simple closed curve in $\mathbb{R}^2$, the line integral of γ against the one-form

$$\phi \triangleq \frac{1}{2}(xdy - ydx) \tag{1.1}$$

gives the (signed) area enclosed by the path γ. The integral against the one-form

$$d\theta \triangleq \frac{1}{r^2}(xdy - ydx)$$

on the punctured plane gives the winding number of γ around the origin. Other topological properties associated with paths, e.g. turning number and linking number, can also be defined in a similar way in terms of line integrals. In the probabilistic context, one can considder distributional properties of stochastic line integrals along stochastic processes such as diffusion paths. A well-known example is Lévy's formula for the characteristic function of the area process associated with a planar Brownian motion, i.e. the stochastic line integral of Brownian motion against the area one-form defined by (1.1) (cf. [Lev40]). Another famous example is Spitzer's asymptotic Cauchy law for the Brownian winding number (cf. [Spi58]). Stochastic line integrals are also essential in the study of diffusions/martingales on manifolds (cf. [Hsu02]).

A more fundamental reason of considering (iterated) line integrals is that the original path γ is uniquely determined by these integrals when one varies the degree n and the one-forms $\phi_1, \cdots, \phi_n$ in a suitably rich class. Indeed, when $M = \mathbb{R}^d$, the collection of numbers (known as the *signature* of γ)

$$\Big\{ \int_{0<t_1<\cdots<t_m<T} d\gamma_{t_1}^{i_1} \cdots d\gamma_{t_m}^{i_m} : m \in \mathbb{N},\ i_1, \cdots, i_m = 1, \cdots, d \Big\}$$

uniquely determines the path $\gamma : [0, T] \to \mathbb{R}^d$ up to tree-like pieces (cf. [Che58, HL10, BGLY16]). In [Che73], the author used iterated line integrals against differential forms to construct a de Rham cohomology theory on loop spaces over manifolds and proved that such a theory is canonically isomorphic to the singular cohomology theory in classical algebraic topology.

In the probabilistic context, in the pioneering work of Le Jan and Qian [LQ12], the authors developed an explicit method of recovering a generic Brownian trajectory from the knowledge of its extended signatures. Their underlying idea can be summarised as follows. Given an arbitrary bounded domain D in $\mathbb{R}^d$, by constructing a suitable one-form ϕ supported on D one can detect whether the Brownian motion B has visited D from the knowledge of the line integral against ϕ. More generally, given a discretisation of $\mathbb{R}^d$ into disjoint cubes with suitably constructed one-forms supported inside each of them, one can detect the discrete route of the motion from the knowledge of iterated line integrals against these on-forms. By refining the space discretisation, one recovers the original trajectory in the limit under this mechanism (cf. Section 4 below for more discussion).

In the method of [LQ12], an essential property of the required one-form ϕ is that

$$\int_0^T \phi(dB_t) \neq 0 \Longleftrightarrow B \text{ visits the } D \quad \text{a.s.}$$

where B is a Brownian motion in $\mathbb{R}^d$. Such a property can be trivially implied by a much stronger non-degeneracy property that the conditional law of $\int_0^T \phi(dB_t)$ given that B visits the domain D is absolutely continuous with respect to the Lebesgue measure. This motivates the following general question which is the main object of study in the present work.

We consider the following SDE on M ($M = \mathbb{R}^n$ or a compact differentiable manifold):

$$\begin{cases} dX_t = \sum_{\alpha=1}^d V_\alpha(X_t) dB_t^\alpha, \quad 0 \leqslant t \leqslant T; \\ X_0 = x_0 \in M. \end{cases} \tag{1.2}$$

Here $B = (B^1, \cdots, B^d)$ is assumed to be a d-dimensional fractional Brownian motion with Hurst parameter $H > 1/4$. This falls into the rough path framework under which the SDE is well-posed in the sense of rough paths. The vector fields $V_1, \cdots, V_d$ on M are assumed to be of class C_b^∞ and satisfy the so-called *Hörmander's condition* (cf. Definition 3.5). This is a natural non-degeneracy condition under which the solution X_t is known to have a smooth density function with respect to the Lebesgue measure (cf. [CHLT15]). Throughout the rest, we use C_p^∞ to mean the class of functions/one-forms whose derivatives (of all orders) have at most polynomial growth. This property ensures the L^p-integrability (for all $p > 1$) of all relevant random variables under consideration.

Question. Let ϕ be a C_p^∞ one-form on M. Can we identify an explicit quantitative condition on ϕ, such that the conditional distribution of the stochastic line integral

$\int_0^T \phi(dX_t)$, given that X visits the interior of the support of ϕ, admits a density function with respect to the Lebesgue measure?

We first make a few comments. It is necessary to restrict on the event that X visits $(\mathrm{supp}\phi)^\circ$, for otherwise the line integral is trivially zero. In addition, suppose that $M = \mathbb{R}^n$, $\mathrm{supp}\phi \neq M$ and $x_0 \in (\mathrm{supp}\phi)^\circ$. For the stochastic line integral $\int_0^T \phi(dX_t)$ to have a density function, it is necessary that ϕ is not closed. Indeed, if $d\phi = 0$, then $\phi = df$ for some smooth function f (every closed one-form on $\mathbb{R}^n$ is exact). In this case, we have

$$\int_0^T \phi(dX_t) = f(X_T) - f(x_0).$$

This integral will have constant value on the non-trivial event $\{X_T \notin \mathrm{supp}\phi\}$. As a result, the line integral cannot have a density function in this case.

As we will see, in the elliptic case, the non-closedness of ϕ is essentially sufficient for the line integral to have a density.

Theorem 1.1. *Suppose that the vector fields $V_1, \cdots, V_d$ are elliptic. Let ϕ be a C_p^∞ one-form such that*

$$d\phi \neq 0 \quad \textit{a.e. on } \mathrm{supp}\phi.$$

Then the conditional distribution of $\int_0^T \phi(dX_t)$, given that X visits $(\mathrm{supp}\phi)^\circ$, admits a density with respect to the Lebesgue measure.

The hypoelliptic case requires a stronger condition and more delicate analysis. The general result is given by Theorem 3.9 below. Here we state the special version in the step-two hypoelliptic case.

Theorem 1.2. *Consider the case when $M = \mathbb{R}^3$ and $d = 2$. Suppose that the vector fields $\mathcal{V} = \{V_1, V_2, [V_1, V_2]\}$ linear span T_xM at every point $x \in M$. Let ϕ be a C_p^∞ one-form on M. Suppose that*

$$d\big(\phi + d\phi(V_1, V_2)\omega^3\big) \neq 0 \quad \textit{a.e. on } \mathrm{supp}\phi,$$

where $\{\omega^1, \omega^2, \omega^3\}$ is the cotangent frame field dual to $\mathcal{V}$. Then the conditional distribution of $\int_0^T \phi(dX_t)$, given that X visits $(\mathrm{supp}\phi)^\circ$, admits a density with respect to the Lebesgue measure.

In Theorem 3.13 below, we also derive an explicit method of constructing one-forms that satisfy the general non-degeneracy criterion given by Theorem 3.9. In the above step-two hypoelliptic case, the method is summarised in the following result. In this case, the class of one-forms that satisfy such a condition is as generic as pairs of C_p^∞-functions.

Proposition 1.3. *Under the setting of Theorem 1.2, consider a one-form ϕ given by*

$$\phi \triangleq c_1\omega^1 + c_2\omega^2 + (V_1c_2 - V_2c_1)\omega^3$$

where $c_1, c_2 \in C_p^\infty(M)$. Suppose that

$$d\phi \neq 0 \quad \text{a.e. on } \mathrm{supp}\phi. \tag{1.3}$$

Then the conditional distribution of $\int_0^T \phi(dX_t)$, given that X visits $(\mathrm{supp}\phi)^\circ$, admits a density with respect to the Lebesgue measure.

Our analysis can be extended to the case of iterated line integrals

$$F \triangleq \int_{0<t_1<\cdots<t_m<T} \phi_1(dX_{t_1}) \cdots \phi_m(dX_{t_m}).$$

As we will see, if $\phi_1, \cdots, \phi_m$ have disjoint supports, our general condition given by (3.10) imposed on each ϕ_i continues to guarantee the conditional non-degeneracy of F. On the other hand, if these one-forms have a common compact support, when $m \geqslant 2$ it is indeed possible to have all the ϕ_i's being exact while F is non-degenerate. Recall what we explained earlier that this is not possible when $m = 1$. Our results for iterated line integrals are discussed in Section 3.2 below.

Our study is motivated by the signature uniqueness problem in the spirit of [LQ12]. As an application, in Section 4 we prove a signature uniqueness theorem for the SDE (1.2) in the elliptic or step-two hypoelliptic case, which asserts that with probability one the solution path $t \mapsto X_t$ is uniquely determined by its signature transform up to reparametrisation. Under existing methodology, the key ingredient is the explicit construction of compactly supported one-forms that satisfy our non-degeneracy conditions. The main result for this part is stated in Theorem 4.3 below.

Finally, we remark that our results hold for more general Gaussian driving processes essentially without changing any part of the proofs. The required Gaussian setting is precisely the one formulated in the work of [CHLT15] concerning the smoothness of density for Gaussian rough differential equations. We formulate our results in the context of fractional Brownian motion simply to avoid the non-rewarding effort of restating all the assumptions proposed in [CDL15].

Organisation. The present article is organised in the following way. In Section 2, we recall basic notions from rough path theory and some terminology from differential geometry. In Section 3.1, we derive our quantitative criteria for the non-degeneracy of single stochastic line integrals as well as an explicit method of construction. We begin with the elliptic case and then proceed to the hypoelliptic case. The analysis is made more transparent in the step-two hypoelliptic case after the general discussion. In Section 3.2, we extend our analysis to the case of iterated line integrals. In Section 4, we discuss the application of our results to the signature uniqueness problem for rough differential equations.

2 Preliminary notions from rough path theory and differential geometry

In this section, we recall some basic tools and discuss the basic kind of pathwise analysis that will be performed frequently in the sequel. We first give a notational comment which will be applied throughout the rest of the article.

Notation. Above all, we will adopt Einstein's convention of summation, i.e. doubly repeated indices are summed automatically. We will also use matrix notation exclusively. For instance, a vector field $V = V^i \frac{\partial}{\partial x^i}$ on $\mathbb{R}^n$ is identified as an $n \times 1$ column vector function. DV is the $n \times n$ matrix whose (i, j)-entry is $\frac{\partial V^i}{\partial x^j}$. A one-form $\phi = \phi_i dx^i$ on $\mathbb{R}^n$ is identified as a $1 \times n$ row vector function. If $f \in C^\infty(\mathbb{R}^n)$, df is the one-form defined by $df \triangleq \frac{\partial f}{\partial x^i} dx^i$. Given a smooth function f and vector field V, we write $Vf \triangleq df \cdot V = V^i \frac{\partial f}{\partial x^i}$. The pairing between a one-form ϕ and a vector field V is obviously $\phi \cdot V$, while on the other hand we write $V\phi$ as the $1 \times n$ row vector defined by $V\phi \triangleq (V\phi_1, \cdots, V\phi_n)$. Note that DV and $V\phi$ are local quantities that do not have intrinsic geometric meaning.

2.1 Pathwise differential calculus

Let $\{X_t : t \geqslant 0\}$ be the solution to the SDE (1.2) where $M = \mathbb{R}^n$ for now, B_t is a d-dimensional fBM with Hurst parameter $H > 1/4$ and the vector fields $V_1, \cdots, V_d \in C_b^\infty$. Here B_t is regarded as a geometric rough path and the SDE is solved under the framework of rough path theory (cf. [LQ02]). Throughout the rest, we will assume that B_t is realised on the canonical path space. More specifically, the underlying probability space is $(\mathcal{W}, \mathcal{B}(\mathcal{W}), \mathbb{P})$ where $\mathcal{W}$ is the Banach space of $\mathbb{R}^d$-valued continuous paths starting at the origin, $\mathcal{B}(\mathcal{W})$ is the Borel σ-algebra over $\mathcal{W}$ and $\mathbb{P}$ is the law of the fBM. The process B is taken to be the coordinate process on $\mathcal{W}$. Under this set-up, solutions of differential equations driven by B and stochastic line integrals along B are regarded as functionals over $\mathcal{W}$.

The rough path nature of (1.2) allows us to justify and make use of pathwise differential calculus in the ordinary manner. To illustrate this, we first recall the following definition of the Cameron-Martin subspace (cf. [FH14]).

Definition 2.1. The *Cameron-Martin subspace* associated with fBM is the subspace $\mathcal{H}$ of paths $h \in \mathcal{W}$ that can be represented in the form

$$h_t = \mathbb{E}[ZB_t], \quad 0 \leqslant t \leqslant T,$$

where Z is an element in the first Wiener chaos (i.e. the L^2-closure of linear functions on $\mathcal{W}$ under $\mathbb{P}$). $\mathcal{H}$ is a Hilbert space with respect to the inner product

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

where Z_i is the chaos element associated with h_i in its definition ($i = 1, 2$).

We use the following lemma to illustrate an example of the type of pathwise calculation that will appear frequently later on. If $F : \mathcal{W} \to \mathbb{R}$ is a functional of B and $h \in \mathcal{H}$ is a Cameron-Martin path, we write

$$D_h F(w) \triangleq \frac{d}{d\varepsilon}\,|_{\varepsilon=0}\, F(w + \varepsilon h)$$

as the derivative of F along direction h at the location $w \in \mathcal{W}$. The *Malliavin derivative* of F is the $\mathcal{H}$-valued random variable defined by

$$DF \triangleq [h \mapsto D_h F] \in \mathcal{H}^* \cong \mathcal{H}.$$

Lemma 2.2. *Let $\phi = \phi_i dx^i$ be a C_p^∞ one-form on $\mathbb{R}^n$. Consider the stochastic line integral*

$$F \triangleq \int_0^T \phi(dX_t).$$

Then

$$D_h F(w) = \int_0^T \big((\zeta_T(w) - \zeta_t(w)) \cdot \Phi_t^{-1}(w) + \phi(X_t(w)) \big) \cdot V_\alpha(X_t(w)) dh_t^\alpha. \quad (2.1)$$

Here $\Phi_t(w) \triangleq \frac{\partial X_t(w)}{\partial x_0}$ denotes the Jacobian of the RDE (1.2) and

$$\zeta_t(w) \triangleq \int_0^t d(\phi_i V_\alpha^i)(X_s(w)) \cdot \Phi_s(w) dw_s^\alpha. \quad (2.2)$$

Proof. To simplify notation we will omit the dependence on w. By the definition of $D_hF(w)$, we have

$$\begin{aligned} D_hF(w) &= \frac{d}{d\varepsilon}|_{\varepsilon=0}\int_0^T \phi_i(X_t(w+\varepsilon h))dX_t^i(w+\varepsilon h) \\ &= \int_0^T \frac{\partial \phi_i}{\partial x^j}(X_t)D_hX_t^j dX_t^i + \int_0^1 \phi_i(X_t)dD_hX_t^i. \end{aligned}$$

By differentiating the SDE (1.2) along the direction h, it is seen that D_hX_t satisfies the differential equation

$$dD_hX_t^i = \frac{\partial V_\alpha^i}{\partial x^j}(X_t)D_hX_t^j dw_t^\alpha + V_\alpha^i(X_t)dh_t^\alpha. \tag{2.3}$$

As a result, we have

$$\begin{aligned} D_hF(w) &= \int_0^T \frac{\partial \phi_i}{\partial x^j}(X_t)D_hX_t^j V_\alpha^i(X_t)dw_t^\alpha \\ &\quad + \int_0^T \phi_i(X_t)\big(\frac{\partial V_\alpha^i}{\partial x^j}(X_t)D_hX_t^j dw_t^\alpha + V_\alpha^i(X_t)dh_t^\alpha\big) \\ &= \int_0^T \frac{\partial}{\partial x^j}(\phi_i V_\alpha^i)(X_t)D_hX_t^j dw_t^\alpha + \int_0^1 \phi_i(X_t)V_\alpha^i(X_t)dh_t^\alpha. \end{aligned} \tag{2.4}$$

On the other hand, the Jabocian Φ_t satisfies the homogeneous linear equation

$$d\Phi_t = DV_\alpha(X_t)\Phi_t dw_t^\alpha, \ \Phi_0 = \mathrm{Id}.$$

By the variational principle, it is standard that

$$D_hX_t = \Phi_t \int_0^t \Phi_s^{-1}V_\alpha(X_s)dh_s^\alpha. \tag{2.5}$$

By using the formula (2.5) and integration by parts, the first integral in (2.4) can be written as

$$\begin{aligned} &\int_0^T \big(d\zeta_t \cdot \int_0^t \Phi_s^{-1}V_\alpha(X_s)dh_s^\alpha\big) \\ &= \zeta_T \cdot \int_0^T \Phi_s^{-1}V_\alpha(X_s)dh_s^\alpha - \int_0^T \zeta_t\Phi_t^{-1}V_\alpha(X_t)dh_t^\alpha \\ &= \int_0^T (\zeta_T - \zeta_t)\Phi_t^{-1}V_\alpha(X_t)dh_t^\alpha, \end{aligned}$$

where ζ_t is the integral path defined by (2.2). The equation (2.1) thus follows. $\square$

A technical remark. In the above proof, we have performed pathwise integration essentially using principles of ordinary calculus. This type of calculations can all be made rigorous under the framework of rough path theory (cf. [CHLT15, Ina14, FH14] in which such type of calculation was used frequently and justified carefully). For instance, the integral on the right hand side of (2.1) is understood in the sense of Young, due to a variational embedding theorem for the Cameron Martin space $\mathcal{H}$ proved by Friz-Victoir (cf. [FH14]). Another example is that the path ζ_t can be understood in the sense of RDE, namely the last component of the triple $\Xi_t \triangleq (X_t, \Phi_t, \zeta_t)$ which is defined through the RDE

$$\begin{cases} dX_t = V_\alpha(X_t)dB_t^\alpha, \\ d\Phi_t = DV_\alpha(X_t)\Phi_t dB_t^\alpha, \\ d\zeta_t = d(\phi \cdot V_\alpha)(X_t) \cdot \Phi_t dB_t^\alpha. \end{cases}$$

The assumption of $\phi \in C_p^\infty$ ensures that the stochastic line integral F has moments of all orders (indeed smooth in the sense of Malliavin), due to standard estimates of rough integrals and the exponential integrability of the p-variation of B ($p > 1/H$). This is seen in exactly the same way as in [Ina14]. Throughout the rest, we will perform pathwise differential calculus of similar kind without further justification.

The following theorem is a standard regularity result in the Malliavin calculus which will be used in Section 3. Its proof can be found in [Nua06].

Theorem 2.3. *Let F be a twice differentiable random variable on $\mathcal{W}$ (in the sense of Malliavin) and $F, DF, D^2F \in L^p$ for some $p > 1$. Then conditional on the event $\{DF \neq 0\}$, the distribution of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.*

2.2 Some terminology from differential geometry

The notion of RDEs, stochastic line integrals as well as our non-degeneracy criteria in Section 3 are intrinsic properties, in the sense that they are defined in terms of the underlying vector fields and one-forms. In particular, they are independent of the choice of local coordinates or embedding of the state manifold into an ambient Euclidean space. It is thus beneficial to perform some of the analysis in geometric terms. The benefit is particularly clear in the hypoelliptic analysis developed in Section 3.1.2. In this section, we recall some notation from differential geometry that will be used later on (cf. Chern-Chen-Lam [CCL00]).

Let M be a differentiable manifold. We denote $\Omega^k(M)$ ($0 \leqslant k \leqslant n$) as the space of (smooth) k-forms on M. Given a (smooth) vector field X, the *interior* *product* $i(X) : \Omega^k(M) \to \Omega^{k-1}(M)$ is defined by

$$(i(X)\omega)(Y_1, \cdots, Y_{k-1}) \triangleq \omega(X, Y_1, \cdots, Y_{k-1}). \tag{2.6}$$

The *Lie derivative* $L_X : \Omega^k(M) \to \Omega^k(M)$ is defined by

$$(L_X\omega)(Y_1, \cdots, Y_k) \triangleq X(\omega(Y_1, \cdots, Y_k)) - \sum_{i=1}^{k} \omega(Y_1, \cdots, Y_{i-1}, [X, Y], Y_{i+1}, \cdots, Y_k). \tag{2.7}$$

Here $Y_1, \cdots, Y_{k-1}$ are arbitrary vector fields, a k-form is viewed as an antisymmetric k-linear functional on vector fields and Xf is the directional derivative of f along X. These two operators are related through the so-called *Cartan's identity*:

$$d \circ i(X) + i(X) \circ d = L_X, \tag{2.8}$$

where $d : \Omega^k(M) \to \Omega^{k+1}(M)$ is the exterior derivative operator. We also recall that the *exterior product* of two one-forms α, β is defined by

$$\alpha \wedge \beta(X, Y) = \alpha(X)\beta(Y) - \beta(X)\alpha(Y)$$

and the exterior derivative of one-form α has the following characterisation:

$$d\alpha(X, Y) = X(\alpha(Y)) - Y(\alpha(X)) - \alpha([X, Y]), \tag{2.9}$$

where X, Y are arbitrary vector fields. A one-form α is *closed* if $d\alpha = 0$. It is *exact* if $\alpha = df$ for some smooth function f. Every exact form is closed, and on $\mathbb{R}^n$ the converse is also true. Throughout the rest, we will simply write $\alpha \cdot X$ for the pairing $\alpha(X)$ which is also consistent with matrix notation in the Euclidean case.

It is convenient to re-interpret Lemma 2.2 in geometric terms. For instance, the Jacobian $\Phi_t : T_{x_0}M \to T_{X_t}M$ is the linear isomorphism that pushes tangent vectors at x_0 forward along the solution path by the flow of diffeomorphisms associated with the RDE (1.2). The Malliavin derivative $t \mapsto D_h X_t \in T_{X_t}M$ is a path on the tangent bundle (cf. (2.5)). The formula (2.1) can be expressed as

$$D_h F(w) = \int_0^T \Big(\int_t^T \langle \Phi_s^* d(\phi \cdot V_\beta)(X_s), \Phi_t^{-1} V_\alpha(X_t) \rangle dw_s^\beta + (\phi \cdot V_\alpha)(X_t) \Big) dh_t^\alpha, \tag{2.10}$$

where $\langle \cdot, \cdot \rangle$ denotes the pairing between cotangent and tangent vectors at the starting point x_0. Equation (2.10) clearly has an intrinsic meaning.

Finally, we recall an equation for the pull-back of vector fields by the Jacobian. This equation, which played an essential role in the proof of Hörmander's theorem for RDEs (cf. [CF10, CHLT15]), will also be crucial for our analysis. Its proof is a straight forward application of the chain rule.

Lemma 2.4. *Let W be a smooth vector field on M. Then the path $t \mapsto \Phi_t^{-1}W(X_t) \in T_{x_0}M$ satisfies the equation*

$$\Phi_t^{-1}W(X_t) = W(x_0) + \int_0^t \Phi_s^{-1}[V_\alpha, W](X_s)dw_s^\alpha.$$

3 Non-degeneracy criteria for stochastic line integrals

In this section, we establish quantitative criteria for the non-degeneracy (i.e. existence of density) of stochastic/rough line integrals (extended signatures) of the form

$$\int_{0<t_1<\cdots<t_m<T} \phi_1(dX_{t_1})\cdots\phi_m(dX_{t_m}),$$

where $X_t \in M$ is the solution to the RDE (1.2) and $\phi_1, \cdots, \phi_m$ are C_p^∞ one-forms. Our results hold when M is either $\mathbb{R}^n$ or a (compact) manifold, but we will only work with the case when $M = \mathbb{R}^n$ and the vector fields $V_\alpha \in C_b^\infty$. Extension of the argument to the manifold case is routine by either working in local charts or embedding M into an ambient Euclidean space. To illustrate the idea better, we first consider the case when $m = 1$ and then extend the analysis to the case of iterated integrals.

3.1 Single line integrals

We begin by considering a single line integral $F \triangleq \int_0^T \phi(dX_t)$. We divide the discussion into two cases: elliptic and hypoelliptic. In the elliptic case, the result is particularly simple and neat, while the hypoelliptic case requires a strong condition as well as more delicate analysis.

3.1.1 The elliptic case

By *ellipticity*, we assume that $n = d$ and the vector fields $V_1, \cdots, V_d$ in the RDE (1.2) linearly span $\mathbb{R}^d$ at every point. In the introduction, we saw that the line integral F may fail to have a density if the one-form ϕ is closed. In the elliptic case, it turns out that non-closedness is essentially sufficient for the non-degeneracy of F. The main result is stated as follows.

Theorem 3.1. *Let ϕ be a C_p^∞ one-form on $\mathbb{R}^d$. Suppose that $d\phi \neq 0$ for almost everywhere inside the support of ϕ. Let E denote the event that "X_t enters the interior of* $\mathrm{supp}\phi$ *at some time t". Then conditional on E, the line integral F has a density with respect to the Lebesgue measure.*

Our proof of Theorem 3.1, as well as its hypoelliptic counterpart, relies crucially on the following two properties of fBM. Its proof is contained in [CHLT15].

Lemma 3.2. *(i) Let $f = (f_1, \cdots, f_d) : [0,T] \to \mathbb{R}^d$ be a deterministic path such that $\int_0^T f_t dh_t$ is well-defined in the sense of Young for all $h \in \mathcal{H}$. If $\int_0^T f_\alpha(t) dh_t^\alpha = 0$ for all $h \in \mathcal{H}$, then $f \equiv 0$.*
(ii) The fBM is a.s. truly rough in the sense of [FH14]. As a result, with probability one we have

$$\int_0^t y_s dB_s = 0 \ \forall t \in [0,T] \implies y \equiv 0,$$

whenever y is a rough path controlled by B so that the rough integral is well-defined.

Remark 3.3. In [CHLT15], these two properties are implied by a nondeterminism-type condition which was used by the authors to establish the smoothness of density for the RDE solution. It was proved in the same paper that fBM satisfies their nondeterminism condition.

Proof of Theorem 3.1. According to Theorem 2.3, the point is to show that $E \subseteq \{DF \neq 0\}$ modulo some $\mathbb{P}$-null set N. First of all, let $N_1 \subseteq \mathcal{W}$ be a null set such that w admits a canonical rough path lifting and is truly rough (so that Lemma 3.2 (ii) holds) for all $w \in N_1^c$.

Now suppose that $w \in E \cap N_1^c$ is a path such that $DF(w) = 0$. According to Lemma 2.2 and Lemma 3.2 (i), we have

$$\big((\zeta_T - \zeta_t) \cdot \Phi_t^{-1} + \phi(X_t)\big) \cdot V_\alpha(X_t) = 0 \quad \forall t \in [0,T], \alpha = 1, \cdots d \tag{3.1}$$

at the driving path w, where ζ_t is defined by (2.2) and Φ_t is the Jacobian of the RDE. Since the vector fields are assumed to be elliptic, the matrix $V \triangleq (V_1, \cdots, V_d)$ is invertible everywhere. After multiplying (3.1) by $V(X_t)^{-1}\Phi_t$, we obtain that

$$\zeta_T - \zeta_t + \phi(X_t) \cdot \Phi_t = 0.$$

Recall from the equations for X_t and Φ_t that

$$d\big(\phi(X_t) \cdot \Phi_t\big) = \big(\frac{\partial \phi}{\partial x^i} V_\alpha^i\big)(X_t)\Phi_t + (\phi \cdot DV_\alpha)(X_t) \cdot \Phi_t\big) dw_t^\alpha. \tag{3.2}$$

In view of the definition of ζ_t and (3.2), Lemma 3.2 (ii) implies that

$$\big(-d(\phi_i V_\alpha^i)+\frac{\partial\phi}{\partial x^i}V_\alpha^i+\phi\cdot DV_\alpha\big)(X_t)=0\quad\forall t\in[0,T],\ \alpha=1,\cdots,d.$$

By taking the j-th component of this equation, it is seen that

$$\big(\frac{\partial\phi_j}{\partial x^i}-\frac{\partial\phi_i}{\partial x^j}\big)(X_t)V_\alpha^i(X_t)=0\quad\forall t,\alpha,j. \tag{3.3}$$

By ellipticity, the equation (3.3) is equivalent to the property that $(d\phi)(X_t)=0$ for all t. Note that this property holds at the particular path w. To summarise, we have shown that

$$w\in E\cap N_1^c,\ DF(w)=0\implies(d\phi)(X_t(w))=0\ \forall t\in[0,T]. \tag{3.4}$$

By continuity, this is particularly true for $t\in\mathbb{Q}\cap[0,T]$.

On the other hand, by the definition of E and continuity, there is a rational time r such that $X_r(w)\in(\mathrm{supp}\phi)^\circ$. In addition, we know from [CF10] that the law of X_r is absolutely continuous with respect to the Lebesgue measure. Since $\Lambda\triangleq\{x\in(\mathrm{supp}\phi)^\circ:(d\phi)(x)=0\}$ is a Lebesgue null set by assumption, we have

$$\mathbb{P}\big(X_r\in\Lambda\big)=0,\quad\forall r\in\mathbb{Q}\cap[0,T].$$

In view of (3.4), by further excluding the $\mathbb{P}$-null set

$$N_2\triangleq\bigcup_{r\in\mathbb{Q}\cap[0,T]}\{X_r\in\Lambda\},$$

we conclude that

$$w\in E\backslash(N_1\cup N_2)\implies DF(w)\neq0.$$

The result thus follows from Theorem 2.3.

□

Examples that satisfy the assumptions of Theorem 3.1 are generic and easy to construct.

Example 3.4. Let $h(t)\in C_c^\infty(\mathbb{R})$ be a function such that

$$h(t)>0,\ t\in(-1,1);\ h(t)=0,\ t\notin(-1,1),$$

and $h'(t)$ is everywhere nonzero in $(-1,1)$ except at $t=0$. Define the following one-form on $\mathbb{R}^2$:

$$\phi=h(x)h(y)e^{h(y)^2}dx.$$

Then ϕ is supported on $[-1,1]^2$ and

$$d\phi = -h(x)h'(y)(1+2h(y)^2)e^{h(y)^2}dx \wedge dy.$$

Inside its support, $d\phi = 0$ precisely on the slice $y = 0$ which has zero Lebesgue measure.

3.1.2 The hypoelliptic case

We now extend the previous analysis to the hypoelliptic case. We first give the following key definition. Let $V_1, \cdots, V_d$ be a family of smooth vector fields on a differentiable manifold M.

Definition 3.5. We say that $V_1, \cdots, V_d$ satisfy *Hörmander's condition* if the following family of vector fields

$$V_i, [V_i, V_j], [V_i, [V_j, V_k]], [V_i, [V_j, [V_k, V_l]]], \cdots \quad (i, j, k, l \text{ etc.} = 1, \cdots, d)$$

linearly span T_xM at every $x \in M$.

It is a well-known fact that under Hörmander's condition, the solution X_t to the RDE (1.2) admits a smooth density function with respect to the Lebesgue measure (cf. [CF10, CHLT15]). In the diffusion case, this result was first established in the seminal work of fHörmander [Hor67].

We now consider a stochastic line integral $F = \int_0^T \phi(dX_t)$, where X_t is the solution to the RDE (1.2) and $V_1, \cdots, V_d$ are C_b^∞-vector fields on $M = \mathbb{R}^n$ that satisfy Hörmander's condition. We shall obtain a quantitative criterion for the non-degeneracy of F and derive an explicit method of constructing one-forms that satisfy such criterion.

A general criterion

In order to derive a non-degeneracy criterion for F, as in the elliptic case we start by assuming that $DF(w) = 0$ at a given fBM path w. We aim at obtaining a geometric constraint on ϕ which holds at paths w satisfying $DF(w) = 0$ (in the elliptic case, the constraint is closedness: $d\phi = 0$). Our non-degeneracy criterion will simply be that "ϕ does not satisfy such a geometric constraint" (in the elliptic case, $d\phi \neq 0$ a.e. inside suppϕ).

We begin by fixing the following notation. Given a word $I = (i_1, \cdots, i_k)$ over the letters $\{1, \cdots, d\}$, we set

$$V_I \triangleq [V_{i_1}, [V_{i_2}, \cdots, [V_{i_{k-1}}, V_{i_k}]]].$$

The set of finite words (respectively, of length k) is denoted as $\mathcal{W}$ (respectively, $\mathcal{W}_k$). The following lemma is crucial for us.

Lemma 3.6. *We define $\{\psi_I : I \in \mathcal{W}\}$ inductively in the following way:*

$$\psi_I \triangleq 0 \quad \textit{for } I \in \mathcal{W}_1,$$

and

$$\psi_I \triangleq d\phi(V_i, V_J) + V_i\psi_J \quad \textit{for } I = (i, J). \tag{3.5}$$

Suppose that $DF(w) = 0$. Then at the path w, we have

$$(\eta \cdot \Phi^{-1} + \phi) \cdot V_I + \psi_I = 0 \quad \forall I \in \mathcal{W}, \tag{3.6}$$

where

$$\eta_t(w) \triangleq \int_t^T d(\phi \cdot V_\alpha)(X_s(w)) \cdot \Phi_s(w) dw_s^\alpha.$$

Proof. We prove the claim by induction on the length of I. When $I \in \mathcal{W}_1$, this is a restatement of (3.1), whose proof clearly does not rely on ellipticity. Suppose that (3.6) is true for all words of length $\leqslant k$. By taking differential with $I \in \mathcal{W}_k$, we find that

$$d\eta \cdot (\Phi^{-1} V_I) + \eta \cdot d(\Phi^{-1} \cdot V_I) + V_i(\phi \cdot V_I + \psi_I) dw_t^i = 0.$$

By the definition of η and Lemma 2.4, we have

$$\big(\eta \cdot \Phi^{-1}[V_i, V_I] - V_I(\phi \cdot V_i) + V_i(\phi \cdot V_I) + V_i\psi_I\big) dw_t^i = 0.$$

It follows from Lemma 3.2 (ii) that

$$\eta \cdot \Phi^{-1}[V_i, V_I] - V_I(\phi \cdot V_i) + V_i(\phi \cdot V_I) + V_i\psi_I = 0 \quad \forall i = 1, \cdots, d.$$

In addition, note that

$$\begin{aligned} V_i(\phi \cdot V_I) - V_I(\phi \cdot V_i) &= (V_i\phi)V_I - (V_I\phi)V_i + \phi \cdot (DV_I \cdot V_i - DV_i \cdot V_I) \\ &= d\phi(V_i, V_I) + \phi \cdot [V_i, V_I]. \end{aligned}$$

Therefore, we obtain that

$$(\eta \cdot \Phi^{-1} + \phi)[V_i, V_I] + d\phi(V_i, V_I) + V_i\psi_I = 0 \quad \forall i.$$

According to the definition of $\psi_{I'}$, the above relation is precisely the desired property for the word $I' = (i, I)$ $(i = 1, \cdots, d)$. This completes the induction step. $\square$

We shall make use of properly chosen local frame fields (i.e. family of vector fields that form a basis of $T_x\mathbb{R}^n$ at every point x) associated with Hörmander's condition. As a result, our condition on ϕ will be expressed locally in terms of these frame fields. For each $x \in \mathbb{R}^n$, according to Hörmander's condition and continuity, there exists a neighbourhood U of x together with subsets $\mathcal{I}_1, \cdots, \mathcal{I}_r$ of words ($\mathcal{I}_k \subseteq \mathcal{W}_k$), such that

$$\{V_I : I \in \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r\}$$

form a local frame field of $\mathbb{R}^n$ on U. We may assume that suppϕ is covered by such local "charts".

Now suppose that $X(w)$ passes through a local chart U of suppϕ on which a local frame field

$$\mathcal{V} = \{V_I : I \in \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r\}$$

is chosen and fixed. Note that

$$|\mathcal{I}_1| + \cdots + |\mathcal{I}_r| = n.$$

Let W be the Mat(n, n)-valued function on U defined by

$$W \triangleq (V_I)_{I \in \mathcal{I}_k, 1 \leqslant k \leqslant r}$$

and set

$$\Theta \triangleq (\psi_I)_{I \in \mathcal{I}_k, 1 \leqslant k \leqslant r},$$

where ψ_I is defined by (3.5). Under the assumption $DF(w) = 0$, the relation (3.6) can be written in matrix form as

$$(\eta \cdot \Phi^{-1} + \phi) \cdot W + \Theta = 0$$

Since W is invertible, we have

$$\eta + \phi \cdot \Phi = \Xi \cdot \Phi \tag{3.7}$$

where $\Xi \triangleq -\Theta \cdot W^{-1}$. Note that (3.7) holds at w for all times in

$$L_w \triangleq \{t \in [0, 1] : X_t(w) \in U\}.$$

Let $\{\omega^I\}$ be the coframe field dual to $\mathcal{V}$. As a one-form on U, we have

$$\Xi = -\sum_{k=2}^{r} \sum_{I \in \mathcal{I}_k} \psi_I \omega^I \quad \text{on } U, \tag{3.8}$$

Lemma 3.7. *Let $\omega = \omega_i dx^i$ be a one-form and $V = V^i \partial_i$ be a vector field. Then the following two identities hold true:*

(i) $-d(\omega \cdot V) + V\omega + \omega \cdot DV = i(V)d\omega$;
(ii) $V\omega + \omega \cdot DV = L_V \omega$.

Proof. By the definition (2.7) of the Lie derivative, we have

$$\begin{aligned}(L_V\omega)(\partial_i) &= V\omega_i - \omega \cdot [V, \partial_i] = V^j \frac{\partial \omega_i}{\partial x^j} + \omega_j \frac{\partial V^j}{\partial x^i} \\ &= (V\omega + \omega \cdot DV) \cdot \partial_i.\end{aligned}$$

This justifies the relation in (ii). The relation of (i) is a simple consequence of (ii) and Cartan's identity (2.8). $\square$

Lemma 3.8. *Suppose that $DF(w) = 0$ and $X(w)$ passes through the local chart U. Then at the path w, we have*

$$i(V_\alpha)d\phi = L_{V_\alpha}\Xi \tag{3.9}$$

for all $\alpha = 1, \cdots, d$ and $t \in L_w$.

Proof. By taking differential of the relation (3.7), we obtain that

$$\big(-d(\phi \cdot V_\alpha) \cdot \Phi + (V_\alpha \cdot \phi) \cdot \Phi + \phi \cdot DV_\alpha \cdot \Phi\big)dw^\alpha = \big((V_\alpha \Xi) \cdot \Phi + \Xi \cdot DV_\alpha \cdot \Phi\big)dw^\alpha.$$

After cancelling Φ on both sides, Lemma 3.2 (ii) implies that

$$-d(\phi \cdot V_\alpha) + V_\alpha \phi + \phi \cdot DV_\alpha = V_\alpha \Xi + \Xi \cdot DV_\alpha \quad \forall \alpha = 1, \cdots, d \text{ and } t \in L_w.$$

The result follows immediately from Lemma 3.7. $\square$

Equivalently, Lemma 3.8 suggests that if the relation (3.9) does not hold on U and if $X(w)$ passes through U, then $DF(w) \neq 0$. As a consequence, along the same lines of argument as in the elliptic case, we have proved the following result, which is the main theorem in this section giving a quantitative non-degeneracy criterion for the line integral F.

Theorem 3.9. *Let ϕ be a C_p^∞ one-form on $\mathbb{R}^n$. Suppose that the support of ϕ is covered by local charts U on which suitable local frame fields $\mathcal{V}$ are chosen and fixed. For each U, define the local one-form Ξ_U on U by (3.8) with respect to the coframe dual to $\mathcal{V}$. Suppose that on each chart U, we have*

$$i(V_\alpha)d\phi - L_{V_\alpha}\Xi_U \neq 0 \quad a.e. \text{ on } U \cap \mathrm{supp}\phi \tag{3.10}$$

for some $\alpha = 1, \cdots, d$. Then conditional on the event that "X enters the support of ϕ", the stochastic line integral $\int_0^T \phi(dX_t)$ has a density with respect to the Lebesgue measure.

Remark 3.10. The condition (3.10) is stronger than the non-closedness condition $d\phi \neq 0$ a.e. obtained in the elliptic case. Indeed, it is obvious that

$$d\phi = 0 \implies \Xi_U = 0 \implies i(V_\alpha)d\phi - L_{V_\alpha}\Xi_U = 0.$$

An explicit method of construction

The next basic question is whether there are rich examples of one-forms that satisfy the non-degeneracy criteria derived in the previous sections. In the elliptic case, the non-closedness condition is fairly easy to achieve. In the hypoelliptic case, there is also a rich class of one-forms (at least as generic as pairs of smooth functions) that satisfy the condition (3.10). In what follows, we discuss a general and explicit method of constructing them.

We first recall some basic notation from sub-Riemannian geometry. Suppose that $\{V_1, \cdots, V_d\}$ are given smooth vector fields on a differentiable manifold M which satisfy Hörmander's condition. Define $\mathcal{D}_1$ to be the $C^\infty(M)$-module generated by $\{V_1, \cdots, V_d\}$. Equivalently, for each $x \in M$, $\mathcal{D}_1(x)$ is the subspace of $T_x(M)$ defined by

$$\mathcal{D}_1(x) = \mathrm{Span}\{V_1(x), \cdots, V_d(x)\}, \quad x \in M.$$

Inductively, define

$$\mathcal{D}_k \triangleq \mathcal{D}_{k-1} + [\mathcal{D}_1, \mathcal{D}_{k-1}], \quad k \geqslant 2,$$

where $[\mathcal{D}_1, \mathcal{D}_{k-1}]$ denote the $C^\infty(M)$-module generated by $\{[X, Y] : X \in \mathcal{D}_1, Y \in \mathcal{D}_{k-1}\}$. Elements in $\mathcal{D}_k$ are linear combinations of $\{V_I : |I| \leqslant k\}$ with smooth coefficients. According to Hörmander's condition, at every $x \in M$ there is a smallest integer $r(x)$ such that $\mathcal{D}_{r(x)}(x) = T_x M$. Observe that

$$\{0\} =: \mathcal{D}_0(x) \subseteq \mathcal{D}_1(x) \subseteq \mathcal{D}_2(x) \subseteq \cdots \subseteq \mathcal{D}_{r(x)}(x).$$

The list of integers

$$\dim \mathcal{D}_1(x) < \dim \mathcal{D}_2(x) < \cdots < \dim \mathcal{D}_{r(x)}(x)$$

is known as the *growth vector* of $\{V_1, \cdots, V_d\}$ at x. A point x is a *regular point* if the growth vector is constant near x. The set of regular points is open and dense in M.

The following simple algebraic lemma allows us to choose preferable local frame fields to work with.

Lemma 3.11. *Let $x_0 \in M$ be a regular point. There exists a neighbourhood U of x_0 and a collection of words $(\mathcal{I}_1, \cdots, \mathcal{I}_r)$ ($r \triangleq r(x_0)$, $\mathcal{I}_k \subseteq \mathcal{W}_k$), such that the following two properties hold true for each $k = 1, \cdots, r$:*

(i) $\mathcal{I}_k \subseteq \mathcal{I}_1 \times \mathcal{I}_{k-1}$;
(ii) $\{V_I : I \in \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_k\}$ *is a local frame field of* $\mathcal{D}_k$ *on* U.

Proof. We construct $\mathcal{I}_k$ by induction. First of all, let U_0 be a neighbourhood of x_0 on which the growth vector is constant. Choose $\mathcal{I}_1$ so that $\{V_i(x_0) : i \in \mathcal{I}_1\}$ form a basis of $\mathcal{D}_1(x_0)$. By continuity, there exists $U_1 \subseteq U_0$ such that $\{V_i(x) : i \in \mathcal{I}_1\}$ are linearly independent for each $x \in U_1$. Since $\dim \mathcal{D}_1$ is constant on U_1, we see that $\{V_i : i \in \mathcal{I}_1\}$ is a local frame field of $\mathcal{D}_1$ on U_1.

Now suppose that a neighbourhood U_k and $\mathcal{I}^{(k)} = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_k$ have been obtained to satisfy the required properties. We claim that

$$\mathcal{D}_{k+1}(x_0) = \mathrm{Span}\big\{V_I(x_0), [V_i, V_J](x_0) : i \in \mathcal{I}_1, I \in \mathcal{I}^{(k)}, J \in \mathcal{I}_k\big\}. \tag{3.11}$$

Indeed, let $W \in \mathcal{D}_1(U_k)$ and $Z \in \mathcal{D}_k(U_k)$. By the induction hypothesis, we can write

$$W = \sum_{i \in \mathcal{I}_1} f_i V_i, \ Z = \sum_{J \in \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_k} g_J V_J$$

where $f_i, g_I \in C^\infty(U_k)$. It follows that

$$[W, Z] = \sum_{i \in \mathcal{I}_1} \sum_{J \in \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_k} \big((f_i V_i g_J) V_J - (g_J V_J f_i) V_i + f_i g_J [V_i, V_J]\big).$$

For $J \in \mathcal{I}^{(k-1)}$, since $[V_i, V_J] \in \mathcal{D}_k$ is a $C^\infty(M)$-linear combination of V_I ($I \in \mathcal{I}^{(k)}$) on U_k, the claim (3.11) follows immediately. Note that $\{V_I(x_0)\} : I \in \mathcal{I}^{(k)}\}$ are already linearly independent. As a result, we can choose a collection $\mathcal{I}_{k+1}$ of (i, J) with $i \in \mathcal{I}_1, J \in \mathcal{I}_k$ such that

$$\{V_I(x_0), [V_i, V_J](x_0) : I \in \mathcal{I}^{(k)}, (i, J) \in \mathcal{I}_{k+1}\}$$

form a basis of $\mathcal{D}_{k+1}(x_0)$. By continuity and the constant dimensionality of $\mathcal{D}_{k+1}$ on U_k, we see that $\{V_I : I \in \mathcal{I}^{(k+1)}\}$ is a local frame field of $\mathcal{D}_{k+1}$ on some $U_{k+1} \subseteq U_k$. From the construction, it is also clear that $\mathcal{I}_{k+1} \subseteq \mathcal{I}_1 \times \mathcal{I}_k$. □

Remark 3.12. We know from Property (i) that $\mathcal{I}_k \neq \emptyset$ for all k.

We now derive a general method of constructing one-forms that satisfy Theorem 3.9. Let $(U; \mathcal{V} = \{V_I : I \in \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r\})$ be a chosen local frame field that satisfies the properties in Lemma 3.11. For $I = (i, J)$, we denote $\langle\!\langle d\phi, V_I \rangle\!\rangle \triangleq d\phi(V_i, V_J)$.

From the definition of ψ_I (cf. (3.5)) and Property (i) of Lemma 3.11, it is not hard to see that

$$\langle\!\langle d\phi, V_I \rangle\!\rangle = 0 \ \forall I \in \mathcal{I}_2 \cup \cdots \cup \mathcal{I}_r \implies \Xi = 0,$$

where Ξ is the one-form defined by (3.8) with respect to the local frame field $\mathcal{V}$. In addition, since $\Xi \cdot V_\alpha = 0$ for all $\alpha \in \mathcal{I}_1$, we have

$$i(V_\alpha)d\phi - L_{V_\alpha}\Xi = i(V_\alpha)d(\phi - \Xi). \tag{3.12}$$

As a result, a sufficient condition for (3.10) to hold on U is that:

(A) $\langle\!\langle d\phi, V_I \rangle\!\rangle = 0$ for all $I \in \mathcal{I}_2 \cup \cdots \cup \mathcal{I}_r$;
(B) $i(V_\alpha)d\phi \neq 0$ a.e. on U for some $\alpha \in \mathcal{I}_1$.

We shall reduce the above two conditions to a more explicit set of relations in terms of coefficients. To this end, let $\{\omega^I : I \in \mathcal{I}^{(r)} \triangleq \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r\}$ be the coframe dual to $\mathcal{V}$ and express ϕ on U as

$$\phi = \sum_{I \in \mathcal{I}^{(r)}} c_I \omega^I,$$

where $c_I \in C^\infty(U)$. Let us fix a total ordering $\prec$ on $\mathcal{I}^{(r)}$ such that $I \prec J$ if $|I| < |J|$. For $I, J, K \in \mathcal{I}^{(r)}$, we set

$$\Lambda^I_{JK} \triangleq d\omega^I(V_J, V_K) = V_J(\omega^I(V_K)) - V_K(\omega^I(V_J)) - \omega^I([V_J, V_K]) = -\omega^I([V_J, V_K]).$$

It follows that

$$\begin{aligned} d\phi &= \sum_I \big(dc_I \wedge \omega^I + \sum_{J \prec K} c_I \Lambda^I_{JK} d\omega^J \wedge d\omega^K \big) \\ &= \sum_I \big(\sum_J V_J c_I \omega^J \wedge \omega^I - \sum_{J \prec K} c_I \omega^I([V_J, V_K]) d\omega^J \wedge d\omega^K \big) \\ &= \sum_{I \prec J} (V_I c_J - V_J c_I - \sum_K c_K \omega^K([V_I, V_J])) \omega^I \wedge \omega^J. \end{aligned}$$

Let c_i ($i \in \mathcal{I}_1$) be an arbitrary family of smooth functions on U. Given $I = (i, j) \in \mathcal{I}_2$, since $i, j \in \mathcal{I}_1$, we have

$$\langle\!\langle d\phi, V_I \rangle\!\rangle = \pm(V_i c_j - V_j c_i - \sum_K c_K \omega^K(V_I))) = \pm(V_i c_j - V_j c_i - c_I\big).$$

As a result, by setting

$$c_I \triangleq V_i c_j - V_j c_i, \quad I = (i,j) \in \mathcal{I}_2,$$

we conclude that $\langle\!\langle d\phi, V_I \rangle\!\rangle = 0$ for all $I \in \mathcal{I}_2$. Inductively on k, for $I = (i, J) \in \mathcal{I}_k$ we set

$$c_I \triangleq V_i c_J - V_J c_i$$

where c_J has already been defined since $J \in \mathcal{I}_{k-1}$. It then follows that

$$\langle\!\langle d\phi, V_I \rangle\!\rangle = 0 \quad \forall I \in \mathcal{I}_2 \cup \cdots \cup \mathcal{I}_r$$

on U. In particular, the aforementioned Condition (A) holds. For Condition (B), note that

$$i(V_\alpha)d\phi = \sum_{J:J\neq\alpha} \Big(V_\alpha c_J - V_J c_\alpha - \sum_K c_K \omega^K([V_\alpha, V_J])\Big)\omega^J \tag{3.13}$$

for each $\alpha \in \mathcal{I}_1$. As a result, Condition (B) boils down to requiring that at least one of the ω^J-coefficients in (3.13) is a.e. nonzero on U.

To summarise, we have obtained the following result which provides an explicit method of constructing one-forms that satisfy the criterion (3.10).

Theorem 3.13. *Let $c_i \in C_c^\infty(U)$ $(i \in \mathcal{I}_1)$ be given and define c_I $(I \in \mathcal{I}_k)$ inductively by*

$$c_I \triangleq V_i c_J - V_J c_i, \quad I = (i, J) \in \mathcal{I}_k.$$

Suppose that for some $\alpha \in \mathcal{I}_1$ and $J \in \mathcal{I}_r$, we have

$$V_\alpha c_J - V_J c_\alpha - \sum_K c_K \omega^K([V_\alpha, V_J]) \neq 0 \quad \text{a.e. on } U. \tag{3.14}$$

Then conditional on the event that "X enters the support of ϕ", the stochastic line integral $\int_0^T \phi(dX_t)$ has a density with respect to the Lebesgue measure.

Remark 3.14. The left hand side of (3.14) is an expression involving up to the r-th derivatives of c_i $(i \in \mathcal{I}_1)$. The property (3.14) is essentially generic for functions $c_i \in C_c^\infty(U)$ $(i \in \mathcal{I}_1)$.

The step-two case and the Heisenberg group

Let us consider the simplest hypoelliptic situation, i.e. when $d = 2$, $\dim M = 3$ and the vector fields

$$\mathcal{V} = \{V_1, V_2, V_3 \triangleq [V_1, V_2]\}$$

form a basis of $T_x M$ at every point $x \in M$(i.e. a global frame field over M). In this case, Theorems 3.9 and 3.13 are simplified substantially. Let $\{\omega^1, \omega^2, \omega^3\}$ be the coframe of $\mathcal{V}$. The definition (3.8) of the one-form Ξ reads

$$\Xi = -d\phi(V_1, V_2)\omega^3.$$

According to the identity (3.12) and the anti-symmetry of $d(\phi - \Xi)$ as a bilinear form on vector fields, the condition (3.10) in Theorem 3.9 is equivalent to that

$$d(\phi + d\phi(V_1, V_2)\omega^3) \neq 0 \quad \text{a.e. on supp}\phi.$$

In addition, Conditions (A) and (B) in the last section simply reads

$$d\phi(V_1, V_2) = 0 \text{ and } d\phi \neq 0 \text{ a.e. on supp}\phi.$$

In terms of coefficients of ϕ with respect to $\{\omega^1, \omega^2, \omega^3\}$, we have the following direct corollary of Theorem 3.13.

Corollary 3.15. *Consider a one-form*

$$\phi = c_1\omega^1 + c_2\omega^2 + (V_1 c_2 - V_2 c_1)\omega^3, \tag{3.15}$$

where $c_1, c_2 \in C_p^\infty(M)$. Suppose that $d\phi \neq 0$ a.e. inside the support of ϕ. Then conditional on the event that "X enters suppϕ"*, the stochastic line integral $\int_0^T \phi(dX_t)$ has a density with respect to the Lebesgue measure.*

We conclude with an explicit example: the *Heisenberg group*. More precisely, we consider $M = \mathbb{R}^3$, where the vector fields V_1, V_2 are given by

$$V_1 = \partial_x - y\partial_z, \;\; V_2 = \partial_y + x\partial_z$$

respectively. In this case, the solution to the RDE (1.2) is explicitly given by the original fBM B coupled with the associated *Lévy area process*

$$X_t = \big(B_t^x, B_t^y, \int_0^t B_s^x dB_s^y - \int_0^t B_s^y dB_s^x\big)_{0\leqslant t\leqslant T}.$$

By explicit calculation, it is easily seen that $[V_1, V_2] = 2\partial_z$. In particular, $\mathcal{V} \triangleq \{V_1, V_2, [V_1, V_2]\}$ is a global frame field. Its coframe is found to be

$$\omega^1 = dx,\ \omega^2 = dy,\ \omega^3 = \frac{y}{2}dx - \frac{x}{2}dy + \frac{1}{2}dz.$$

Let $\phi = c_i\omega^i$ where $c_i \in C_p^\infty(\mathbb{R}^3)$. Under Cartesian coordinates, we have

$$\phi = \big(c_1 + \frac{yc_3}{2}\big)dx + \big(c_2 - \frac{xc_3}{2}\big)dy + \frac{1}{2}c_3 dz. \tag{3.16}$$

Let us further assume that c_1, c_2 depend only on the x, y coordinates. Define

$$c_3 \triangleq -V_2c_1 + V_1c_1 = -\partial_y c_1 + \partial_x c_2,$$

so that $d\phi(V_1, V_2) = 0$ as seen before. Note that c_3 also depends only on x, y. We obtain from (3.16) that

$$\begin{cases} d\phi(\partial_x, \partial_z) = \frac{1}{2}\partial_x c_3 = \frac{1}{2}\big(-\partial^2_{xy}c_1 + \partial^2_{xx}c_2\big), \\ d\phi(\partial_y, \partial_z) = \frac{1}{2}\partial_y c_3 = \frac{1}{2}\big(-\partial^2_{yy}c_1 + \partial^2_{xy}c_2\big). \end{cases} \tag{3.17}$$

As a consequence, as long as the functions (c_1, c_2) are chosen such that

$$\big(-\partial^2_{xy}c_1 + \partial^2_{xx}c_2\big)\cdot\big(-\partial^2_{yy}c_1 + \partial^2_{xy}c_2\big) \neq 0 \quad \text{a.e. in supp}\phi, \tag{3.18}$$

the non-degeneracy of the line integral $\int_0^T \phi(dX_t)$ holds. Since there are no a priori constraints on c_1, c_2, the property (3.18) is apparently generic.

3.2 Iterated line integrals

We now turn to the stuy of an extended signature

$$F = \int_{0<t_1<\cdots<t_m<T} \phi_1(dX_{t_1})\cdots\phi_m(dX_{t_m}) \quad (m \geqslant 2).$$

We consider two typical situations: (i) the supports of the one-forms $\phi_1, \cdots, \phi_m$ are mutually disjoint, or (ii) they all have common support. As we will see, in the first case the conditions provided by Theorem 3.9 (imposed on each ϕ_i) continue to ensure the non-degeneracy of F. In the second case, we demonstrate that it is possible to have all ϕ_i's being exact while F is non-degenerate, which is surprising in contrast to the case of $m = 1$.

We first prepare a lemma that will be used in both cases. It is a natural extension of (3.1).

Lemma 3.16. *For $k=1,\cdots,m$, we set*

$$G_t^k \triangleq \int_{0<t_1<\cdots<t_{k-1}<t} \phi_1(dX_{t_1})\cdots\phi_{k-1}(dX_{t_{k-1}}), \tag{3.19}$$

$$H_t^k \triangleq \int_{t<t_{k+1}<\cdots<t_m<T} \phi_{k+1}(dX_{t_{k+1}})\cdots\phi_m(dX_{t_m}), \tag{3.20}$$

where $G_t^1 = H_t^m \triangleq 1$. Suppose that $DF(w)=0$. Then at the path w we have

$$\sum_{k=1}^m \big(\int_t^T G_s^k H_s^k d\zeta_s^k \cdot \Phi_t^{-1} + G_t^k H_t^k \phi_k(X_t)\big)\cdot V_\alpha(X_t) = 0 \tag{3.21}$$

for all $\alpha = 1,\cdots,d$ and $t\in[0,T]$, where

$$\zeta_t^k \triangleq \int_0^t d(\phi_k\cdot V_\alpha)(X_s)\Phi_s dw_s^\alpha. \tag{3.22}$$

Proof. As in the proof of Lemma 2.2, given any $h\in\mathcal{H}$ we have

$$\begin{aligned}
D_hF(w) &= \sum_{k=1}^m \int_{0<t_1<\cdots<t_k<\cdots<t_m<T} \phi_1(dX_{t_1})\cdots D_h\phi_k(dX_{t_k})\cdots\phi_m(dX_{t_m}) \\
&= \sum_{k=1}^m \big(\int_{0<\cdots<t_k<\cdots<T} \cdots(d\zeta_{t_k}^k\cdot\eta_{t_k})\cdots \\
&\quad + \int_{0<\cdots<t_k<\cdots<T} \cdots(\phi_k\cdot V_\alpha)(X_{t_k})dh_{t_k}^\alpha\cdots\big) \\
&= \sum_{k=1}^m \big(\int_0^T G_t^k H_t^k (d\zeta_t^k\cdot\eta_t) + \int_0^T G_t^k H_t^k(\phi_k\cdot V_\alpha)(X_t)dh_t^\alpha\big) \\
&=: A_1 + A_2,
\end{aligned}$$

where $\eta_t \triangleq \int_0^t \Phi_s^{-1}V_\alpha(X_s)dh_s^\alpha$. The same integration by parts argument as in the proof of Lemma 2.2 yields that

$$A_1 = \sum_{k=1}^m \int_0^T \int_t^T G_s^k H_s^k d\zeta_s^k\cdot\Phi_t^{-1}V_\alpha(X_t)dh_t^\alpha.$$

As a consequence, we have

$$D_hF(w) = \sum_{k=1}^m \int_0^T \big(\int_t^T G_s^k H_s^k d\zeta_s^k\cdot\Phi_t^{-1}V_\alpha(X_t) + G_t^k H_t^k(\phi_k\cdot V_\alpha)(X_t)\big)dh_t^\alpha.$$

Since $D_hF(w)=0$ for all $h\in\mathcal{H}$, the result thus follows from Lemma 3.2 (i). □

3.2.1 The case of disjoint supports

Let $\phi_1, \cdots, \phi_m$ be smooth one-forms such that $\mathrm{supp}\phi_i \cap \mathrm{supp}\phi_j = \emptyset$ for all $i \neq j$. Define E to be the event that "there exist times $t_1 < \cdots < t_m$ such that $X_{t_i} \in (\mathrm{supp}\phi_i)^\circ$ for all i". From the definition of F, it is not hard to see in a deterministic way that the line integral F is identically zero on E^c. Our main result in this case is stated as follows.

Theorem 3.17. *Suppose that each ϕ_i satisfies the conditions in Theorem 3.9. Then conditional on the event E, the extended signature F has a density with respect to the Lebesgue measure.*

The following lemma, which is an extension of Lemma 3.6, is needed for our proof of Theorem 3.17.

Lemma 3.18. *For each $k = 1, \cdots, m$, we define $\{\psi_{k,I} : I \in \mathcal{W}\}$ by $\psi_{k,I} \triangleq 0$ if $I \in \mathcal{W}_1$ and*

$$\psi_{k,I} \triangleq d\phi_k(V_i, V_J) + V_i \psi_{k,J}$$

for $I = (i, J)$. Suppose that $DF(w) = 0$. Then at the path w, we have

$$(\rho_t \cdot \Phi_t^{-1} + \sum_{k=1}^m G_t^k H_t^k \phi_k) \cdot V_I + \sum_{k=1}^m G_t^k H_t^k \psi_{k,I} = 0 \quad \forall I \in \mathcal{W}, t \in [0, T], \tag{3.23}$$

where $\rho_t \triangleq \sum_{k=1}^m \int_t^T G_s^k H_s^k d\zeta_s^k$ and G_t^k, H_t^k, ζ_t^k are defined by (3.19, 3.20, 3.22) respectively.

Proof. We prove the claim by induction on the length of the word I. When $I \in \mathcal{W}_1$, the claim reduces to the equation (3.21). Suppose that (3.23) is true for all words of length $\leqslant k$. By differentiating (3.23) with $I \in \mathcal{W}_k$, we have

$$\begin{aligned} d\rho \cdot (\Phi^{-1} V_I) + \rho \cdot d(\Phi^{-1} V_I) + \sum_k d(G_t^k H_t^k) \cdot (\phi_k \cdot V_I + \psi_{k,I}) \\ + \sum_k G_t^k H_t^k V_i (\phi_k \cdot V_I + \psi_{k,I}) dw^i = 0. \end{aligned}$$

Recall that

$$d\rho_t = -\sum_k G_t^k H_t^k d(\phi_k \cdot V_i) \cdot \Phi_t dw_t^i, \; d(\Phi_t^{-1} V_I) = \Phi_t^{-1} \cdot [V_i, V_I] dw_t^i.$$

As a result, we have

$$\rho \cdot \Phi^{-1}[V_i, V_I] + \sum_k G_t^k H_t^k\big(- V_I(\phi_k \cdot V_i) + V_i(\phi_k \cdot V_I) + V_i \psi_{k,I}\big)$$
$$+ \sum_k G_t^k H_t^{k+1} \phi_k \wedge \phi_{k+1}(V_i, V_I) + \sum_k G_t^k H_t^{k+1}(\psi_{k+1,I}\phi_k - \psi_{k,I}\phi_{k+1}) \cdot V_i = 0 \tag{3.24}$$

for all i. Since $\mathrm{supp}\phi_k \cap \mathrm{supp}\phi_{k+1} = \emptyset$, it is readily seen that

$$\phi_k \wedge \phi_{k+1} = 0, \;\; \psi_{k+1,I}\phi_k - \psi_{k,I}\phi_{k+1} = 0.$$

In addition, note that

$$V_i(\phi_k \cdot V_I) - V_I(\phi_k \cdot V_i) = d\phi_k(V_i, V_I) + \phi_k \cdot [V_i, V_I].$$

The equation (3.24) thus reduces to

$$\big(\rho \cdot \Phi^{-1} + \sum_k G_t^k H_t^k \phi_k\big) \cdot [V_i, V_I] + \sum_k G_t^k H_t^k \big(d\phi_k(V_i, V_I) + V_i(\psi_{k,I})\big) = 0.$$

By the definition of $\{\psi_{k,I} : I \in \mathcal{W}\}$, the last expression is equivalent to that

$$\big(\rho \cdot \Phi^{-1} + \sum_k G_t^k H_t^k \phi_k\big) \cdot V_{I'} + \sum_k G_t^k H_t^k \psi_{k,I'} = 0$$

where $I' = (i, I)$. Since $I \in \mathcal{W}_k$ and $i \in \{1, \cdots, d\}$ are arbitrary, we conclude that (3.23) is true for words of length $k+1$. $\square$

We now prove Theorem 3.17 by induction on the degree of F.

Proof of Theorem 3.17. Consider the following slightly more general claim:

($\mathbf{P}_m$) Let $\phi_1, \cdots, \phi_m$ be smooth one-forms with disjoint support and each of them satisfies the conditions in Theorem 3.9. For each pair of $s < t \in [0, T]$, let $E_{s,t}$ be the event that "X visits $(\mathrm{supp}\phi_1)^\circ, \cdots, (\mathrm{supp}\phi_m)^\circ$ in order over $[s, t]$". Then

$$\int_{s<t_1<\cdots<t_m<t} \phi_1(dX_{t_1}) \cdots \phi_m(dX_{t_m})\Big|_{E_{s,t}}$$

admits a density with respect the Lebesgue measure.

We are going to prove $(\mathbf{P}_m)$ by induction on m. The case when $m=1$ is just Theorem 3.9. Suppose that the claim is true for iterated integrals of degree less than m and consider an m-th order integral

$$F=\int_{s<t_1<\cdots<t_m<t}\phi_1(dX_{t_1})\cdots\phi_m(dX_{t_m}).$$

We wish to show that

$$w\in E_{s,t}\cap N^c \implies DF(w)\neq 0, \tag{3.25}$$

where N is a suitable $\mathbb{P}$-null set to be excluded.

Suppose that $w\in E_{s,t}$ and $DF(w)=0$. Let k be fixed and consider a time u such that $X_u\in(\mathrm{supp}\phi_k)^\circ$ and $X|_{[s,u]}$ (respectively, $X|_{[u,t]}$) visits the supports of $\phi_1,\cdots,\phi_{k-1}$ (respectively, of $\phi_{k+1},\cdots,\phi_m$). Such a time u exists as $w\in E_{s,t}$. By the assumption of the theorem, we may take a chart U near X_u on which a local frame field $\{V_I: I\in\mathcal{I}_1\cup\cdots\cup\mathcal{I}_r\}$ is defined and

$$i(V_\alpha)d\phi_k-L_{V_\alpha}\Xi_k\neq 0\quad \text{a.e. on } U \tag{3.26}$$

for some α, where under the notation of Section 3.1.2 we set

$$\Xi_k\triangleq\Theta_kW^{-1},\ \Theta_k\triangleq(\psi_{k,I})_{I\in\mathcal{I}_l,1\leqslant l\leqslant r}, W\triangleq(V_I)_{I\in\mathcal{I}_l,1\leqslant l\leqslant r}\quad \text{on } U.$$

In a small time neighbourhood $v\in(u-\varepsilon,u+\varepsilon)$, the equation (3.23) yields

$$(\rho\cdot\Phi^{-1}+G_v^kH_v^k\phi_k)\cdot W+G_v^kH_v^k\cdot\Theta_k=0 \iff \rho+G_v^kH_v^k\phi_k\cdot\Phi=G_v^kH_v^k\Xi_k\cdot\Phi.$$

Note that the above relation holds at k (not summing over k!) near X_u. By differentiating both sides with respect to w_t^α, we obtain that

$$\begin{aligned}&G_v^kH_v^k\big(-d(\phi_k\cdot V_\alpha)+V_\alpha\phi_k+\phi_k\cdot DV_\alpha\big)+d(G_v^kH_v^k)\phi_k\\ &\quad=G_v^kH_v^k\cdot\big(V_\alpha\Xi_k+\Xi_k\cdot DV_\alpha\big)+d(G_v^kH_v^k)\Xi_k\end{aligned}\tag{3.27}$$

for all α and $v\in(u-\varepsilon,u+\varepsilon)$.

Next, we observe that

$$d(G_v^kH_v^k)=G_v^{k-1}H_v^k\phi_{k-1}(dX_v)-G_v^kH_v^{k+1}\phi_{k+1}(dX_v)=0,$$

since $X_v\in\mathrm{supp}\phi_k$ for v close to u. As a result, the equation (3.27) reduces to

$$G_v^kH_v^k\big(-d(\phi_k\cdot V_\alpha)+V_\alpha\phi_k+\phi_k\cdot DV_\alpha\big)=G_v^kH_v^k\cdot\big(V_\alpha\Xi_k+\Xi_k\cdot DV_\alpha\big).$$

By using the relations

$$-d(\phi_k \cdot V_\alpha) + V_\alpha \phi_k + \phi_k \cdot DV_\alpha = i(V_\alpha)d\phi_k,$$
$$V_\alpha \Xi_k + \Xi_k \cdot DV_\alpha = L_{V_\alpha}\Xi_k,$$

we obtain that

$$G_v^k H_v^k \big(i(V_\alpha)d\phi_k - L_{V_\alpha}\Xi_k\big) = 0 \tag{3.28}$$

for all α and $v \in (u-\varepsilon, u+\varepsilon)$. According to the assumption (3.26), we conclude that either X_v lives on some Lebesgue null set $C \subseteq U$, or $G_v^k H_v^k = 0$.

For each v, we set

$$E'_v \triangleq \{\exists t_1 < \cdots < t_{k-1} \in (0,v) : X_{t_i} \in (\mathrm{supp}\phi_i)^\circ\},$$
$$E''_v \triangleq \{\exists t_{k+1} < \cdots < t_m \in (0,v) : X_{t_i} \in (\mathrm{supp}\phi_i)^\circ\}$$

respectively. To summarise, by continuity we have obtained from (3.28) that

$$\begin{aligned} & w \in E_{s,t} \cap \{DF = 0\} \\ & \implies w \in N \triangleq \bigcup_{r \in \mathbb{Q}\cap(s,t)} \big(\{X_r \in C\} \cup (\{G_r^k = 0\} \cap E'_r) \cup (\{H_r^k = 0\} \cap E''_r)\big). \end{aligned}$$

Since X_r has a density, we know that $\{X_r \in C\}$ is a $\mathbb{P}$-null set. In addition, since G_r^k and H_r^k are iterated line integrals with degree less than m, by the induction hypothesis both of $G_r^k|_{E'_r}$ and $H_r^k|_{E''_r}$ have densities. In particular,

$$\{G_r^k = 0\} \cap E'_r) \cup (\{H_r^k = 0\} \cap E''_r)$$

is also a $\mathbb{P}$-null set. As a result, $\mathbb{P}(N) = 0$ and the desired relation (3.25) follows. In other words, we conclude that $DF \neq 0$ a.s. on $E_{s,t}$, which implies the existence of conditional density by Theorem 2.3. This completes the induction step for the claim $(\mathbf{P}_m)$.

□

3.2.2 The case of common support

Next, we assume that the supports of $\phi_1, \cdots, \phi_m$ have a common intersection S. Our aim here is to demonstrate a surprising fact that the extended signature F can still be non-degenerate even when all the ϕ_i's are exact and compactly supported (i.e. $\phi_i = df_i$ where $f_i \in C_c^\infty(S)$). As we mentioned in the introduction, this is not possible when $m = 1$ (cf. Remark 3.20 as well). Our result in this case is stated as follows. We only consider the elliptic situation.

Proposition 3.19. *Consider an elliptic RDE (1.2) where $X_0 = x_0 \in \mathbb{R}^d$. Let $f_1, \cdots, f_m$ be compactly supported smooth functions. Suppose that the two-forms*

$$df_1 \wedge df_2, \cdots, df_{m-1} \wedge df_m$$

are linearly independent at x_0. Then the extended signature

$$F \triangleq \int_{0<t_1<\cdots<t_m<T} (df_1)(dX_{t_1}) \cdots (df_m)(dX_{t_m})$$

has a density with respect to the Lebesgue measure.

Proof. Write $\phi_k \triangleq df_k$. Let w be an fBM path such that $DF(w) = 0$. According to the equation (3.21) and ellipticity, we have

$$\sum_k \int_t^1 G_s^k H_s^k d\zeta_s^k + \sum_k G_t^k H_t^k \phi_k \cdot \Phi_t = 0 \quad \forall t \in [0, T].$$

By taking differentiation with respect to w_t^α, we find that

$$\sum_k G_t^k H_t^k \big(- d(\phi_k \cdot V_\alpha) + V_\alpha(\phi_k) + \phi_k \cdot DV_\alpha \big)$$
$$+ \sum_k G_t^k H_t^{k+1} \big((\phi_k \cdot V_\alpha)\phi_{k+1} - (\phi_{k+1} \cdot V_\alpha)\phi_k = 0,$$

which is equivalent to that

$$i(V_\alpha) \sum_k \left(G_t^k H_t^k d\phi_k + G_t^k H_t^{k+1} \phi_k \wedge \phi_{k+1} \right) = 0$$

for all $\alpha = 1, \cdots, d$ and $t \in [0, T]$. Again by ellipticity and the fact that $d\phi_k = d^2 f_k = 0$, we have

$$\sum_k G_t^k H_t^{k+1} \phi_k \wedge \phi_{k+1} = 0 \tag{3.29}$$

for all $t \in [0, T]$ at the path w.

We first consider the case when $m = 2$. In this case, the relation (3.29) simply reads

$$(\phi_1 \wedge \phi_2)(X_t(w)) = 0 \quad \forall t \in [0, T].$$

By taking $t = 0$, we reach a contradiction as $\phi_1 \wedge \phi_2(x_0) \neq 0$ by the assumption. Next, we consider the case when $m = 3$. In this case, the relation (3.29) becomes

$$H_t^2 \phi_1 \wedge \phi_2 + G_t^2 \phi_2 \wedge \phi_3 = 0.$$

By the linear independence assumption and continuity, when t is small we have $H_t^2 = G_t^2 = 0$. In particular,

$$G_t^2 = \int_0^t \phi_1(dX_s) = f_1(X_t) - f_1(x_0) = 0 \quad \forall t \text{ small.} \tag{3.30}$$

On the other hand, since $df_1(x_0) \neq 0$ (otherwise the linear independence assumption cannot hold), there exists a neighbourhood U of x_0 such that

$$P \triangleq \{x \in U : f_1(x) = f_1(x_0)\}$$

is an $(n-1)$-dimensional submanifold in U. In particular, the event

$$N \triangleq \bigcup_{r \in \mathbb{Q}_+} \{X_r \in P\}$$

is a $\mathbb{P}$-null set. Note that the property (3.30) implies that N happens. Consequently, in both cases $m = 2, 3$, we see that $DF(w) \neq 0$ a.s. The existence of density thus follows.

Now suppose that the claim is true for iterated integrals of degree $m-2$ where $m \geqslant 4$. . For the degree m case, by taking $k = m-1$ in (3.29) we have

$$DF(w) = 0 \implies G_t^{m-1} = \int_{0<t_1<\cdots<t_{m-2}<t} \phi_1(dX_s) \cdots \phi_{m-2}(dX_s) = 0$$

when t is small. In particular,

$$\{DF = 0\} \subseteq \bigcup_{r \in \mathbb{Q}_+} \{G_r^{m-1} = 0\},$$

which is a $\mathbb{P}$-null set since G_r^{m-1} has a density by the induction hypothesis. Therefore, $DF \neq 0$ a.s. and the claim holds for the degree-m case. The result thus follows by induction. □

Remark 3.20. In contrast, when $m = 1$, the stochastic line integral of a compactly supported exact form will never have a density. Indeed, let f be a compactly supported smooth function. Then

$$F \triangleq \int_0^T (df)(dX_t) = f(X_T) - f(x_0).$$

According to [GOT21, Theorem 1.5], the density of X_T is everywhere strictly positive. It follows that

$$\mathbb{P}(X_T \in (\mathrm{supp} f)^c) > 0.$$

In particular, there is a positive probability that $F = -f(x_0)$. As a result, F cannot have a density. Nonetheless, if we allow $\mathrm{supp} f = \mathbb{R}^n$ it is clearly possible that F has a density. For instance, take $f(x) = |x|^2$ with X_t being a Brownian motion.

4 An application: signature uniqueness for RDEs

In this section, we discuss an application of Theorem 3.13 to the probabilistic signature uniqueness problem. We first give the definition of the signature transform of a rough path (cf. [LCL07]). Let $T((\mathbb{R}^n)) \triangleq \prod_{m=0}^{\infty} (\mathbb{R}^n)^{\otimes m}$ denote the algebra of formal tensor series over $\mathbb{R}^n$ where $(\mathbb{R}^n)^{\otimes 0} \triangleq \mathbb{R}$.

Definition 4.1. Let $\mathbf{X} = (\mathbf{X}_t)_{0 \leqslant t \leqslant T}$ be a rough path over $\mathbb{R}^n$. The *signature* of $\mathbf{X}$ is the formal tensor series defined by

$$S(\mathbf{X}) \triangleq \big(1, \int_0^T d\mathbf{X}_t, \cdots, \int_{0<t_1<\cdots<t_m<T} d\mathbf{X}_{t_1} \otimes \cdots \otimes d\mathbf{X}_{t_m}, \cdots\big) \in T((\mathbb{R}^n)). \quad (4.1)$$

Remark 4.2. If $\mathbf{X}$ is a continuous path in $\mathbb{R}^n$ with bounded variation, the iterated integrals in (4.1) are all defined in the classical sense of Lebesgue-Stieltjes. In the rough path case, the well-definedness of $S(\mathbf{X})$ follows from a basic extension theorem of Lyons (cf. [LQ02]).

After extracting coordinates, the signature $S(\mathbf{X})$ consists of a countable family of numbers associated with the path $\mathbf{X}$. It can be viewed as the pathwise / deterministic analogue of moments of a random variable. There are two basic reasons of considering the signature transform:

(i) [*The signature uniqueness theorem*] Every (geometric) rough path is uniquely determined by its signature up to tree-like pieces (cf. [HL10, BGLY16]). Here a tree-like piece is a portion along which the path travels out and reverses back to cancel itself.
(ii) The signature $S(\mathbf{X})$ has nice algebraic and analytic properties that are concealed at the level of paths (cf. [LCL07, Reu93]).

In the probabilistic setting, the signature uniqueness theorem may take a stronger form as we do not expect tree-like pieces to appear for a suitably non-degenerate stochastic process. Below is the main result in this section which extends earlier probabilistic works [LQ12, GQ16, BG15]. To reduce technicalities, we only consider the elliptic or step-two hypoelliptic case.

Theorem 4.3. *Consider an n-dimensional RDE (1.2) driven by a d-dimensional fractional Brownian motion. Suppose that the vector fields $\{V_1, \cdots, V_d\}$ are C_b^∞ and we are in one of the following two situations:*

(i) $n = d$ and the vector fields are elliptic;
(ii) $n = 3$, $d = 2$ and the vector fields satisfy Hörmander's condition.

Then with probability one, every sample path of the solution process $X = \{X_t : 0 \leqslant t \leqslant T\}$ is uniquely determined by its signature up to reparametrisation.

Remark 4.4. We expect the result to be true for the general hypoelliptic case of arbitrary order, although the construction of relevant one-forms (cf. Condition (ND) below) may be technically more involved in the general case.

Such a probabilistic uniqueness theorem was first established by Le Jan and Qian [LQ12] for the Brownian motion case. The result was later extended to the cases of hypoelliptic diffusions in [GQ16] and Gaussian processes in [BG15]. These works were largely based on a technique developed in [LQ12], which was formalised in [BG15] down to the verification of three key conditions in the context of a general stochastic process X. The first two conditions are: (i) X can be lifted as a rough path in a canonical way and (ii) X_t has a density for each $t > 0$. These two conditions are naturally satisfied for hypoelliptic RDEs. The last condition is stated as follows.

Non-degeneracy Condition (ND). For any cube H in $\mathbb{R}^n$, there exists a smooth one-form ϕ supported in H, such that conditional on the event that "X enters H at some time", the stochastic line integral $\int_0^T \phi(dX_t)$ is a.s. non-zero.

It was proved in [BG15] that the above three conditions imply the signature uniqueness theorem for a general stochastic process X. As a result, in order to prove the aforementioned Theorem 4.3, it remains to verify Condition (ND) under the given assumptions. Before doing so, for the sake of completeness, we briefly recapture the main strategy of [LQ12] and explain at a conceptual level how Condition (ND) leads to the signature uniqueness property.

Step one. Decompose the state space $\mathbb{R}^n$ into disjoint cubes of order ε with narrow gaps δ ($\delta << \varepsilon$). Label the cubes by a set L ($L = \mathbb{Z}^n$ in [LQ12]).

Step two. For each cube H_z ($z \in L$), construct a one-form ϕ_z supported in H_z according to Condition (ND). For each word $w = (z_1, \cdots, z_m)$ over L, one can define the associated extended signature

$$[\phi_{z_1}, \cdots, \phi_{z_m}]_{0,T} \triangleq \int_{0<t_1<\cdots<t_m<T} \phi_{z_1}(dX_{t_1}) \cdots \phi_{z_m}(dX_{t_m})$$

along the path X. As a consequence of an algebraic property of the signature, these extended signatures are all uniquely determined by the signature of X.
Step three. Due to Condition (ND), there exists a unique word w of maximal length, with respect to which the extended signature is non-zero. This word precisely corresponds to the discrete route of the path X in the given space discretisation. As a result, the signature of X uniquely determines its discrete route.
Step four. As we refine the discretisation (i.e. sending $\varepsilon, \delta \to 0$), the discrete route converges to the original sample path X in a suitable sense. Therefore, the signature uniquely determines the trajectory of X.

The rest of this section is devoted to the proof of Theorem 4.3.

Proof of Theorem 4.3: Verification of Condition (ND)

In the elliptic case, we can use Example 3.4 to explicitly construct one-forms satisfying Condition (ND). According to Theorem 3.1, conditional on X entering the cube H, the associated line integral $\int_0^T \phi(dX_t)$ (for ϕ given by Example 3.4) has a density. This clearly implies that its value is a.s. non-zero.

We now consider the step-two hypoelliptic case. Suppose that $n = 3$, $d = 2$ and $\mathcal{V} = \{V_1, V_2, [V_1, V_2]\}$ form a global frame field of $\mathbb{R}^3$. We use the method of Corollary 3.15 to construct suitable one-forms. Recall from (3.15) that such one-forms are given by

$$\phi = c_1\omega^1 + c_2\omega^2 + (V_1c_2 - V_2c_1)\omega^3,$$

where $\{\omega^i\}$ is the coframe of $\mathcal{V}$ and c_1, c_2 are arbitrary smooth functions supported in the cube H. We want to choose ϕ with $\mathrm{supp}\phi = H$ and $d\phi \neq 0$ a.e. in H. Note that $d\phi(V_1, V_2) = 0$. Hence we have to look at $d\phi(V_i, [V_1, V_2])$. Straightforward calculation yields

$$d\phi(V_i, [V_1, V_2]) = V_i(V_1c_2 - V_2c_1) - [V_1, V_2]c_i - \langle \phi, [V_i, [V_1, V_2]] \rangle, \quad i = 1, 2.$$

We will set $c_2 = 0$, so that

$$d\phi(V_2, [V_1, V_2]) = -V_2(V_2c_1) - \langle \omega^1, [V_2, [V_1, V_2]] \rangle \cdot c_1 + \langle \omega^3, [V_2, [V_1, V_2]] \rangle \cdot V_2c_1. \quad (4.2)$$

In other words, we want to construct c_1 with $\text{supp}c_1 = H$, such that the above expression is a.e. non-zero in H.

According to [CCL00, Chap. 1, Theorem 4.3], a non-degenerate vector field locally generates coordinate curves. Since we will eventually refine the space discretisation, we may assume without loss of generality that H is contained in a coordinate chart $[U; x, y, z]$ of $\mathbb{R}^3$ where $V_2 = \partial_x$. To simplify notation, we further assume that H is the unit cube

$$H = \{(x, y, z) : \max\{|x|, |y|, |z|\} < 1\}$$

under the above coordinate system. We define

$$c_1(x, y, z; \lambda) \triangleq h_\lambda(x)\eta(y, z),$$

where $\lambda > 0$ is a parameter to be chosen later on,

$$h_\lambda(x) \triangleq \begin{cases} e^{-\frac{\lambda}{1-x^2}}, & |x| < 1; \\ 0, & |x| \geqslant 1, \end{cases}$$

and $\eta(y, z)$ is a given smooth function supported on $\bar{H}_{y,z} \triangleq \{(y, z) : \max\{|y|, |z|\} \leqslant 1\}$ which is strictly positive in the interior. Under such choice of c_1, the equation (4.2) can be concisely written as

$$-d\phi(V_2, [V_1, V_2]) = \big(h''_\lambda(x) + f(x, y, z)h'_\lambda(x) + g(x, y, z)h_\lambda(x)\big)\eta(y, z),$$

where f, g are known C^∞-functions. Our proof will be concluded from the following lemma.

Lemma 4.5. *There exists $\lambda > 0$, such that*

$$N_\lambda \triangleq \{(x, y, z) \in H : h''_\lambda(x) + f(x, y, z)h'_\lambda(x) + g(x, y, z)h_\lambda(x) = 0\}$$

is a Lebesgue null set.

Remark 4.6. It will be clear from the proof below that Lemma 4.5 holds for all λ outside a suitable null set of $(0, \infty)$. For our purpose, we only need one such λ.

Proof. Explicit calculation shows that

$$h''_\lambda(x) + f(x, y, z)h'_\lambda(x) + g(x, y, z)h_\lambda(x) = \frac{h_\lambda(x)}{(1 - x^2)^4} \cdot \Phi_\lambda(x, y, z),$$

where

$$\Phi_\lambda(x,y,z) = 4x^2\lambda^2 - 2(1-x^2)(1+3x^2+x(1-x^2)f)\lambda + (1-x^2)^4 g. \tag{4.3}$$

Observe that $\Phi_\lambda(x,y,z)$ is a quadratic polynomial in λ. It is easy to see that $(x,y,z) \in N_\lambda \cap \{x \neq 0\}$ if and only if

$$\lambda = \frac{-p \pm \sqrt{\Delta}}{8x^2} \text{ and } \Delta \geqslant 0,$$

where p, Δ are known C^∞-functions on H that can be expressed explicitly in terms of f, g (Δ is the discriminant of (4.3)).

We now consider the following three smooth functions:

$$\psi_\pm \triangleq \frac{-p \pm \sqrt{\Delta}}{8x^2}, \ q \triangleq \frac{-p}{8x^2},$$

where $\psi_\pm$ are defined on $E \triangleq \{\Delta > 0\} \cap \{x \neq 0\}$ (could possibly be empty) and q is defined on $H \cap \{x \neq 0\}$. Recall that the critical set of a smooth function $F : U \to \mathbb{R}$ consists of those points in U at which $\nabla F = 0$. The classical Sard's theorem (cf. [Mil97, Chap. 2]) asserts that the image of the critical set of a smooth function is a Lebesgue null set. Let $Y_\pm, Z$ be the critical sets of $\psi_\pm, q$ respectively. It follows that

$$C \triangleq \psi_+(Y_+) \cup \psi_-(Y_-) \cup q(Z)$$

is a Lebesgue null set in $\mathbb{R}$. As a result, there exists at least one $\lambda \in (0,\infty) \cap C^c$. We fix one such λ. Then each of $\psi_+^{-1}(\lambda)$, $\psi_-^{-1}(\lambda)$, $q^{-1}(\lambda)$ is either empty or a two-dimensional sub-manifold in H. The result thus follows from the observation that

$$N_\lambda \cap \{x \neq 0\} \subseteq \psi_+^{-1}(\lambda) \cup \psi_-^{-1}(\lambda) \cup q^{-1}(\lambda).$$

Note that the slice $\{x = 0\}$ is a Lebesgue null set and has no effect on our discussion. □

If we choose λ as in Lemma 4.5, for the resulting one-form ϕ we have

$$d\phi(V_2, [V_1, V_2]) \neq 0$$

except on a low dimensional manifold which has zero Lebesgue measure. Therefore, $d\phi \neq 0$ a.e. inside the support of ϕ. The Condition (ND) is then a consequence of Corollary 3.15.

References

[BG15] H. Boedihardjo and X. Geng. The uniqueness of signature problem in the non-Markov setting. *Stochastic Process. Appl.* 125 (12) (2015): 4674-4701.

[BGLY16] H. Boedihardjo, X. Geng, T. Lyons and D. Yang. The signature of a rough path: uniqueness. *Adv. Math.* 293 (2016): 720-737.

[CDL15] T. Cass, B.K. Driver and C. Litterer. Constrained rough paths. *Proc. Lond. Math. Soc.* 111 (6) (2015): 1471–1518.

[CF10] T. Cass and P.K. Friz. Densities for rough differential equations under Hörmander's condition. *Ann. of Math.* 171 (2010): 2115–2141.

[CHLT15] T. Cass, M. Hairer, C. Litterer and S. Tindel. Smoothness of the density for solutions to Gaussian rough differential equations. *Ann. Probab.* 43 (1) (2015): 188–239.

[Che58] K. Chen. Integration of paths-a faithful representation of paths by noncommutative formal power series. *Trans. Amer. Math. Soc.* 89 (1958): 395–407.

[Che73] K.T. Chen. Iterated integrals of differential forms and loop space homology. *Ann. of Math.* 97 (2) (1973): 217–246.

[CCL00] S.S. Chern, W.H. Chen and K.S. Lam. *Lectures on differential geometry*. World Scientific, 2000.

[FH14] P.K. Friz and M. Hairer. *A course on rough paths*. Universitext, 2014.

[GOT21] X. Geng, C. Ouyang and S. Tindel. Precise local estimates for differential equations driven by fractional Brownian motion: hypoelliptic case, preprint. To appear in *Ann. Probab.*, 2021.

[GQ16] X. Geng and Z. Qian. On an inversion theorem for Stratonovich's signatures of multidimensional diffusion paths. *Ann. Inst. Henri Poincaré Probab. Stat.* 52 (1) (2016): 429-447.

[HL10] B. Hambly and T. Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Ann. of Math.* 171 (1) (2010): 109–167.

[Hor67] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.* 119 (1967): 147–171.

[Hsu02] E. Hsu. *Stochastic analysis on manifolds*. Americal Mathematical Society, 2002.

[Ina14] Y. Inahama. Malliavin Differentiability of Solutions of Rough Differential Equations. *J. Funct. Anal.* 267 (2014): 1566–1584.

[Lev40] P. Lévy. Le mouvement brownien plan. *Amer. J. Math.* 62 (1940): 487–550.

[LQ12] Y. Le Jan and Z. Qian. Stratonovich's signatures of Brownian motion determine Brownian sample paths, *Probab. Theory Relat. Fields* 157 (2012): 440–454.

[LCL07] T.J. Lyons, M. Caruana and T. Lévy. *Differential equations driven by rough paths.* Lecture Notes in Mathematics, Vol. 1908. Springer, Berlin, 2007.

[LQ02] T. Lyons and Z. Qian. *System control and rough path*s. Oxford University Press, 2002.

[Mil97] J.W. Milnor. *Topology from the differentiable viewpoint.* Princeton University Press, Princeton, 1997.

[Nua06] D. Nualart. *The Malliavin calculus and related topics*. Second Edition. Springer-Verlag, 2006.

[Reu93] C. Reutenauer. *Free Lie algebras.* Clarendon Press, Oxford, 1993.

[Spi58] F. Spitzer. Some theorems concerning 2-dimensional Brownian motion. *Trans. Amer. Math. Soc.* 87:187–197, 1958.