The Annals of Probability
2016, Vol. 44, No. 6, 4049–4082
DOI: 10.1214/15-AOP1068

CHARACTERISTIC FUNCTIONS OF MEASURES ON GEOMETRIC ROUGH PATHS

BY ILYA CHEVYREV[1] AND TERRY LYONS[2]

University of Oxford

We define a characteristic function for probability measures on the signatures of geometric rough paths. We determine sufficient conditions under which a random variable is uniquely determined by its expected signature, thus partially solving the analogue of the moment problem. We furthermore study analyticity properties of the characteristic function and prove a method of moments for weak convergence of random variables. We apply our results to signature arising from Lévy, Gaussian and Markovian rough paths.

1. Introduction. Paths serve as a natural description of an ordered progression of events and are abundant throughout mathematics. Furthermore, measures on paths are almost as common in nature as paths. Considering the flow of infinitesimal elements, one sees that any system involving rigid motions can be represented as a measure on paths; the same can be said of a gas or fluid flow. For this reason, the ability to characterize paths, and measures on them, becomes of value.

It was first shown by Chen [7] that an irreducible piecewise regular continuous path in Euclidean space (which includes all paths that are smooth when parameterized at unit speed) may be faithfully represented, up to reparametrization, by the collection of its iterated integrals known as the signature. The representation of a path through its signature has been recently explored in much greater detail due to its connection with rough paths theory [31]. The exact geometric equivalence of paths of bounded variation possessing the same signature was first described by Hambly and Lyons [19], and recently extended to all geometric rough paths [2]. Methods to recover information encoded by the signature have also been explored and, in general, pose a difficult problem [30].

The signature may be viewed concretely as the universal solution to the exponential differential equation $dS(X)_t = S(X)_t \otimes dX_t$, and serves as the fully non-commutative analogue of the classical exponential function for points in $\mathbb{R}$. Its importance is further emphasized when one considers a general differential equation

$$dY_t = M(Y_t)\, dX_t, \tag{1.1}$$

Received November 2014; revised October 2015.

[1]Supported by the University of Oxford Clarendon Fund Scholarship.

[2]Supported by ERC (Grant Agreement No. 291244 Esig) and by the Oxford-Man Institute of Quantitative Finance.

MSC2010 subject classifications. Primary 60B11; secondary 43A05.

Key words and phrases. Rough paths, expected signature, accumulated local variation.

since the solution Y_t is invariant under reparametrizations of the driving signal X_t. This relationship is most evident in the case of linear differential equations, where $X_t \in V$ and $Y_t \in W$ lie in Banach spaces, and $M : V \mapsto \mathbf{L}(W)$ is a continuous linear map. In this case, the extension of M to an algebra homomorphism $M : T(V) \mapsto \mathbf{L}(W)$, when applied to the signature of X_t, provides a series converging rapidly to the flow of (1.1) [31]. In particular, when M takes values in a Lie algebra, the flow of (1.1) corresponds to the Cartan development of X_t in the corresponding Lie group, thus naturally inducing a representation of the group of signatures.

In the case of a one-dimensional path X_t in $\mathbb{R}$, the signature takes the simple form $(1, X_t - X_0, (X_t - X_0)^2/2!, \ldots)$. When X_t is a random variable, the sequence of expectations $(1, \mathbb{E}[X_t - X_0], \mathbb{E}[(X_t - X_0)^2/2!], \ldots)$, whenever it exists, describes precisely the moments of $X_t - X_0$. Thus, for a stochastic process X_t, the expectations of its iterated integrals, termed the expected signature, naturally form the generalization of the moments of the process.

The expected signature has been exploited in high order approximation schemes [24] and is explicitly known for certain stochastic processes [14, 27]. Moreover, the fundamental property that every polynomial function on signatures may be realized as a linear functional implies that the expected signature distinguishes any two random variables of compact support [10] and implicitly demonstrates the potential of the path signature in applications to numerical analysis and machine learning [26, 29].

The moments of a random variable are of course closely related to the characteristic function $\phi_X(\lambda) = \mathbb{E}[e^{i\lambda X}]$. For a topological group G, a classical extension of the characteristic function to a G-valued random variable X is $\phi_X(M) = \mathbb{E}[M(X)]$ where M is a unitary representation of G [21]. Under suitable conditions, particularly the existence of sufficiently many unitary representations, ϕ_X uniquely determines the law of X.

This paper aims to study the characteristic function $\phi_X(M) = \mathbb{E}[M(X)]$ where X is a random signature and M is a unitary representation arising from a linear map $M : V \mapsto \mathfrak{u}$ into a unitary Lie algebra. Our main result asserts that ϕ_X uniquely determines every random variable X and greatly extends the analogous result for the expected signature beyond the case of compact support.

We now briefly outline the structure of the paper. Section 2 studies a universal topological algebra $E(V)$ in which we embed the group of signatures. Roughly speaking, the induced topology is such that a sequence of signatures converges if and only if the solution to (1.1) converges for every continuous linear map $M : V \mapsto \mathbf{L}(W)$. In Section 3, we derive important properties of probability measures on the set $G(V)$ of group-like elements of $E(V)$. In Section 4, we study representations of $E(V)$. Our first main result is Theorem 4.8, which describes explicitly a family of representations of $E(\mathbb{R}^d)$ which preserves unitary elements and separates the points. An immediate consequence is that one is able to define a meaningful characteristic function for $G(\mathbb{R}^d)$-valued random variables (Corollary 4.12). Though our results for uniqueness of random variables are restricted

to the case $V = \mathbb{R}^d$, we mostly work in the general setting of Banach spaces and make precise whenever finite dimensionality is required.

In Section 5, we recall elements of rough paths theory and show that the signatures of geometric rough paths form a topological subgroup of $G(\mathbb{R}^d)$. In Section 6, we describe applications of our results to stochastic rough paths, particularly in connection with the expected signature. We split Section 6 into three parts.

In Section 6.1, we study the analogue of the moment problem. Proposition 6.1 provides a general criterion under which a $G(\mathbb{R}^d)$-valued random variable is uniquely determined by its expected signature. In turn, Theorem 6.3 provides a method to verify this criterion without explicit knowledge of the expected signature itself. We demonstrate applications of these results to the Lévy–Khintchine formula derived in [14] and to families of Gaussian and Markovian rough paths studied in [4] and [5].

In Section 6.2, we study analyticity properties of the characteristic function. The main result is Theorem 6.13 (and its Corollaries 6.17 and 6.18), which provides a criterion to establish analyticity of the characteristic function and solve the moment problem within a restricted family of random variables. We demonstrate an application to Markovian rough paths stopped upon exiting a domain.

In Section 6.3, we conclude with Theorem 6.31, which demonstrates a method of moments for weak convergence of $G(\mathbb{R}^d)$-valued random variables.

2. Universal locally *m*-convex algebra. Throughout the paper, all vector spaces are assumed real and all algebras are assumed unital. For topological vector spaces V, W, let $\mathbf{L}(V, W)$ be the space of continuous linear maps from V to W, and denote $\mathbf{L}(V) = \mathbf{L}(V, V)$ and $V' = \mathbf{L}(V, \mathbb{R})$. For terminology and basic properties of topological algebras we refer to [32].

For a topological vector space V, a topological algebra A, and a topology on $T(V) = \bigoplus_{k\geq 0} V^{\otimes k}$, consider the statement:

(2.1) For all $M \in \mathbf{L}(V, A)$, the extension $M : T(V) \mapsto A$ is continuous.

One may then topologize $T(V)$ by requiring that (2.1) holds for all topological algebras A of a given category. In this paper, we consider the category of locally m-convex algebras.

DEFINITION 2.1. Let V be a locally convex space. Let $E_a(V) = T(V)$ equipped with the coarsest topology such that (2.1) holds for all locally m-convex algebras A (or equivalently, all normed algebras A). Denote by $E(V)$ the completion of $E_a(V)$.

Thus for any normed algebra A, the set of continuous algebra homomorphisms $\mathrm{Hom}(E_a, A)$ is in bijection with $\mathbf{L}(V, A)$. For any $M \in \mathbf{L}(V, A)$, we shall usually denote by the same letter M the corresponding element in $\mathrm{Hom}(E_a, A)$, but shall write $M_E \in \mathrm{Hom}(E_a, A)$ whenever a clear distinction is needed.

Though in most parts of the paper we shall assume that V is normed, most results in this section are more easily understood for locally convex spaces and so unless stated otherwise, we only assume V is locally convex.

In most of our notation, we shall drop the reference to V when it is clear from the context. It holds that E_a and E are locally m-convex algebra ([32], pages 14 and 22). While most results in this section are stated for E, it is easy to verify which remain valid for E_a.

This method to obtain a universal topological algebra of a specific category is very natural, and we note that this construction is not new; the same construction (and essentially Proposition 2.3 below) appeared in [8] in relation to cyclic cohomology, while analogous constructions were investigated for locally convex algebras with continuous multiplication in [37] and for commutative locally m-convex algebras (particularly in relation to nuclear spaces) in [9] Section 6.4.

REMARK 2.2. If we start with V as a general topological vector space, an easy verification shows that we arrive at the same space E_a as when we equip V with the finest locally convex topology coarser than its original.

A family of semi-norms Ψ on V is called *fundamental* if for every semi-norm ξ on V, there exist $\gamma \in \Psi$ and $\varepsilon > 0$ such that $\varepsilon\xi \leq \gamma$ (note that by a semi-norm we always mean a continuous semi-norm). For any collection of semi-norms Ψ on V, define $\Psi^* = \{n\gamma \mid n \geq 1, \gamma \in \Psi\}$.

For semi-norms γ, ξ on locally convex spaces V, W, respectively, let $\gamma \otimes \xi$ denote the projective semi-norm on $V \otimes W$. Denote by $V \otimes_\pi W$ the projective tensor product and $V \widehat{\otimes} W$ its completion. For a normed space F, and $M \in \mathbf{L}(V, F)$ denote $\gamma(M) = \sup_{\gamma(v)=1} \|Mv\|$ (possibly infinite).

Define the *projective extension* of a semi-norm γ on V as the semi-norm $\exp(\gamma) = \sum_{k\geq 0} \gamma^{\otimes k}$ on E_a. Remark that $\exp(\gamma)$ is a sub-multiplicative semi-norm on E_a. Moreover for any normed algebra A, $M \in \mathbf{L}(V, A)$, and a semi-norm γ on V such that $\gamma(M) \leq 1$, it holds that $\exp(\gamma)(M_E) \leq 1$. We thus readily obtain the following.

PROPOSITION 2.3. *Let Ψ be a family of semi-norms on V. Then Ψ is a fundamental family of semi-norms on V if and only if* $\exp(\Psi^*)$ *is a fundamental family of semi-norms on E.*

COROLLARY 2.4. *The space E is Hausdorff* (*resp.*, *metrizable*, *separable*) *if and only if V is Hausdorff* (*resp.*, *metrizable*, *separable*).

Whenever we speak of a topological space, we shall henceforth always assume it is Hausdorff. The following result identifies E with a subspace of $P(V) := \prod_{k\geq 0} V^{\widehat{\otimes} k}$. For $x \in P$, we write x^k for the projection of x onto $V^{\widehat{\otimes} k}$, so that $x = (x^0, x^1, x^2, \ldots)$.

COROLLARY 2.5. *Let Ψ be a fundamental family of semi-norms on V. Then* $E = \{x \in P \mid \forall \gamma \in \Psi^*, \sum_{k\geq 0} \gamma^{\otimes k}(x^k) < \infty\}$.

By noting the identification $P^{\widehat{\otimes} 2} = \prod_{i,j\geq 0} V^{i,j}$, where $V^{i,j} \cong V^{\widehat{\otimes}(i+j)}$, the same considerations show that

$$E^{\widehat{\otimes} 2} = \left\{x \in P^{\widehat{\otimes} 2} \mid \forall \gamma \in \Psi^*, \sum_{i,j\geq 0} \gamma^{\otimes(i+j)}(x^{i,j}) < \infty\right\}.$$

Let $\rho^k : E \mapsto V^{\widehat{\otimes} k}$ denote the projection $\rho^k(x) = x^k$. The following result shall also be useful later and is another consequence of Proposition 2.3.

COROLLARY 2.6. *The operators $T^{(n)} := \sum_{k=0}^{n} \rho^k : E \mapsto E$ converge uniformly on bounded sets to the identity operator on E.*

When V is a normed space, we always equip $V^{\otimes k}$ with the projective norm unless stated otherwise. For an element $x \in P$ define its radius of convergence $R(x)$ as the radius of convergence of the series $\sum_{k\geq 0} \|x^k\| \lambda^k$. Corollary 2.5 then implies that $x \in E$ if and only if $R(x) = \infty$.

We now come to a more interesting permanence property. For a semi-normed space (W, γ) denote the quotient normed space $W_\gamma = (W/\operatorname{Ker}(\gamma), \gamma)$ and $\widehat{W}_\gamma$ its completion. For a locally convex space W and a Banach space A, a map $M \in \mathbf{L}(W, A)$ is called compact (resp., nuclear) if there exists a semi-norm γ on W such that the $\gamma(M) < \infty$ and the induced map $M_\gamma : \widehat{W}_\gamma \mapsto A$ is compact (resp., nuclear). Recall that W is called Schwartz (resp., nuclear) if every $M \in \mathbf{L}(W, A)$ is compact (resp., nuclear) for every Banach space A.

PROPOSITION 2.7. *The space E is Schwartz (resp., nuclear) if and only if V is Schwartz (resp., nuclear).*

We note that the case when V is simply Schwartz shall not be used later in the paper and is recorded simply for completeness. Moreover, nuclearity of E shall only be applied in Section 6.3 to the case $V = \mathbb{R}^d$. However the equivalent statement for $V = \mathbb{R}^d$ uses essentially the same proof and thus we record the result in full generality.

Let Ψ be a fundamental family of sub-multiplicative semi-norms of a locally m-convex algebra F. Equipping Ψ^* with its natural partial order, $(\widehat{F}_\gamma)_{\gamma\in\Psi^*}$ is a projective system of Banach algebras and one obtains a dense topological algebra embedding $F \hookrightarrow \varprojlim_{\gamma\in\Psi^*} \widehat{F}_\gamma$ known as the Arens–Michael decomposition (see [32], Chapter III). As compact (resp., nuclear) operators form an operator ideal, we obtain the following.

LEMMA 2.8. *Let F be a locally m-convex algebra. Then F is Schwartz (resp., nuclear) if and only if every continuous algebra homomorphism $M : F \mapsto A$ is compact (resp., nuclear) for every Banach algebra A.*

For a normed space V and Banach space W, denote by $\mathcal{N}(V, W)$ the Banach space of nuclear operators from V to W with the nuclear norm $\|\cdot\|_N$.

LEMMA 2.9. *Let (V, γ) be a normed space and A a Banach algebra. Let $M \in \mathcal{N}(V, A)$ with $\|M\|_N < 1$. Equip $T(V)$ with the norm* $\exp(\gamma)$. *Then the extension $M_E : T(V) \mapsto A$ is nuclear and $\|M_E\|_N \le (1 - \|M\|_N)^{-1}$.*

PROOF. It holds that product map $M^{\otimes k} : V^{\otimes_\pi k} \mapsto A^{\otimes_\pi k}$ is nuclear with nuclear norm bounded by $\|M\|_N^k$ ([22] Theorem 3.7—the bound is clear from the proof therein), and the multiplication map $A^{\otimes_\pi k} \mapsto A$ has unit operator norm. It follows that $M^{\otimes k} : V^{\otimes_\pi k} \mapsto A$ is nuclear with nuclear norm at most $\|M\|_N^k$ ([18], page 84). The conclusion follows since $M_E = \sum_{k \ge 0} M^{\otimes k}$ is an absolutely convergent series in $\mathcal{N}(T(V), A)$. $\square$

For a semi-norm γ on V, let $B_\gamma = \{v \in V \mid \gamma(v) < 1\}$, and for a subset $B \subseteq V$, let $\Gamma(B)$ be the absolutely convex hull of B.

PROOF OF PROPOSITION 2.7. The "only if" direction is clear. Let A be a Banach algebra, $M \in \mathbf{L}(V, A)$, and let Ψ be a fundamental family of semi-norms on V. For a semi-norm γ on V, recall that $B_{\gamma^{\otimes k}} = \Gamma(B_\gamma^{\otimes k}) \subset V^{\otimes k}$.

Suppose V is Schwartz. Take $\gamma \in \Psi^*$ such that $M(B_\gamma) \subset A$ is relatively compact and $\gamma(M) < 1$. It follows that $M(B_{\gamma^{\otimes k}})$ is relatively compact in A ([36] Proposition 7.11). Since the unit ball $B_{\exp(\gamma)}$ is given by $\Gamma(\bigcup_{k \ge 0} B_{\gamma^{\otimes k}})$, we obtain that $M(B_{\exp(\gamma)})$ is totally bounded in A. Thus E is Schwartz by Lemma 2.8.

Suppose V is moreover nuclear. Take $\gamma \in \Psi^*$ such that the induced map $M_\gamma : V_\gamma \mapsto A$ is nuclear with $\|M_\gamma\|_N < 1$. As $(V^{\otimes_\pi k})_{\gamma^{\otimes k}}$ and $(V_\gamma)^{\otimes_\pi k}$ are isometrically isomorphic ([18], page 38), we have the natural identification $T(V)_{\exp(\gamma)} \cong (T(V_\gamma), \exp(\gamma))$. It follows that $M_E : T(V)_{\exp(\gamma)} \mapsto A$ is nuclear by Lemma 2.9. Thus, E is nuclear again by Lemma 2.8. $\square$

One may also ask when the extension map $\cdot_E : \mathbf{L}(V, A) \mapsto \mathrm{Hom}(E, A)$ is continuous under certain topologies. In the case of the strong topology when V is normed, we obtain a homeomorphism by the following proposition. First, remark that if $\|x_j\| \le c$ and $\|x_j - y_j\| \le \varepsilon$ for $x_1, \dots, x_n, y_1, \dots, y_n \in A$, where A is a normed algebra, then

$$\|x_1 \cdots x_n - y_1 \cdots y_n\| \le \sum_{j=1}^{n} \binom{n}{j} \varepsilon^j c^{n-j} = (c + \varepsilon)^n - c^n. \tag{2.2}$$

PROPOSITION 2.10. *Let V be a normed space and A a Banach algebra. The extension map $\cdot_E : M \mapsto M_E$ from $\mathbf{L}(V, A)$ to* $\mathrm{Hom}(E, A)$ *is continuous* (*and thus a homeomorphism*) *when one equips both sides with the strong topology.*

PROOF. Let $(M_j)_{j\geq 1} \to M$ in $\mathbf{L}(V, A)$. Let γ be a norm on V such that $\gamma(M) \leq 1$ and $\gamma(M_j) \leq 1$ for all $j \geq 1$.

Remark that for any bounded set $B \subset E$ and $\varepsilon > 0$, there exists $k_\varepsilon \geq 1$ such that $\sup_{x\in B} \gamma^{\otimes k}(x^k) \leq \varepsilon^k$ for all $k \geq k_\varepsilon$ [if not, then take a sequence $x_n \in B$ such that $\gamma^{\otimes n}(x_n^n) > \varepsilon^n$. Then $\exp(c\gamma)(x_n) > c^n\varepsilon^n$ for any $c > 1$ and $n \geq 1$, which is implies that $\exp(c\gamma)$ is not bounded on B for some $c > 1$ which is a contradiction].

Remark that every bounded set in $V^{\otimes k}$ is contained in $\overline{\Gamma}(B_1 \otimes \cdots \otimes B_k)$ for bounded sets $B_1, \dots, B_k \subset V$, and that the supremum of a convex function on a set is equal to its supremum on the set's convex hull. Together with (2.2), this implies that for any fixed n,

$$\sup_{x\in B} \sum_{0\leq k\leq n} \|M_i^{\otimes k}(x^k) - M^{\otimes k}(x^k)\| \to 0.$$

Hence,

$$\sup_{x\in B}\|M_j(x) - M(x)\| \leq \sup_{x\in B} \sum_{0\leq k\leq n} \|M_i^{\otimes k}(x^k) - M^{\otimes k}(x^k)\| + 2\sup_{x\in B}\sum_{k>n} \gamma^{\otimes k}x^k$$

can be made arbitrarily small with sufficiently large n and j. $\square$

REMARK 2.11. If we assume only that V is locally convex and $M \in \mathbf{L}(V, A)$, where A is a real (resp., complex) Banach algebra, applying the above proposition to the semi-norm $\gamma(x) := \|M(x)\|$ on V implies in particular that the map $\lambda \mapsto (\lambda M)_E$ is continuous from $\mathbb{R}$ (resp., $\mathbb{C}$) to $\mathrm{Hom}(E, A)$, where the latter is equipped with the strong topology.

3. Group-like elements. We recall that $T(V)$ is a Hopf algebra with coproduct $\Delta v = 1 \otimes v + v \otimes 1$ for all $v \in V$ and antipode $\alpha(v_1 \cdots v_k) = (-1)^k v_k \cdots v_1$ for all $v_1 \cdots v_k \in V^{\otimes k}$ ([33] Proposition 1.10).

Consider now V a locally convex space. Since $E^{\widehat{\otimes}2}$ is itself a locally m-convex algebra ([32], page 378), and since $\Delta \in \mathbf{L}(V, E^{\otimes_\pi 2})$, the extension $\Delta : E \mapsto E^{\widehat{\otimes}2}$ is continuous by the universal property of E. Moreover, the antipode α extends to a continuous linear map $\alpha : E \mapsto E$. This endows E with an "almost" Hopf algebra structure ("almost" since E is not mapped to $E^{\otimes 2}$ under the coproduct Δ as for Hopf algebras, but to its completion $E^{\widehat{\otimes}2}$).

Denote by $U(V) = \{g \in E \mid \alpha(g) = g^{-1}\}$ and $G(V) = \{g \in E \mid \Delta(g) = g \otimes g, g \neq 0\}$ the groups of *unitary* elements and *group-like* elements of E, respectively. Note that since multiplication and inversion in E are continuous (and indeed in every locally m-convex algebra; [32], pages 5 and 52), U and G are topological groups when endowed with the subspace topology. Moreover, U is closed in E since the map $\phi : x \mapsto (\alpha(x)x, x\alpha(x))$ from E into $E \times E$ is continuous and $U = \phi^{-1}\{(1, 1)\}$. Likewise, G is closed in E since $g^0 = 1$ for all $g \in G$ and $G = \psi^{-1}\{0\} \setminus \{0\}$ for the continuous map $\psi : x \mapsto x \otimes x - \Delta(x)$ from E into $E^{\widehat{\otimes}2}$. Finally, note the inclusion $G \subset U$.

In this section, we collect several results concerning measures on G. While these results shall later be applied mostly to the case $V = \mathbb{R}^d$, we find making this assumption does not simplify the proofs, and thus make most statements in full generality.

All measures (resp., random variables) are assumed to be Borel. Denote by $\mathcal{P}(S)$ the space of probability measures on a topological space S endowed with the topology of weak convergence on $C_b(S, \mathbb{C})$.

Recall that for a locally convex space F, an F-valued random variable X is weakly (Gelfand–Pettis) integrable, or that $\mathbb{E}[X]$ exists, if $f(X)$ is integrable for all $f \in F'$ and if there exists $\mathbb{E}[X] := x \in F$ such that $\mathbb{E}[f(X)] = f(x)$. Letting μ be the probability measure associated with X, we denote by $\mu^* = \mathbb{E}[X]$ its barycenter. Unless stated otherwise, we shall always assume that μ is the measure associated to X and that integrals are taken in the weak sense.

DEFINITION 3.1. For an E-valued random variable X, we call the sequence

$$\mathrm{ESig}(X) := \left(\mathbb{E}[X^0], \mathbb{E}[X^1], \dots\right) \in P = \prod_{k \geq 0} V^{\widehat{\otimes} k}$$

the *expected signature* of X whenever X^k is integrable for all $k \geq 0$.

When V is normed, define $r_1(X)$ as the radius of convergence of the series

$$\sum_{k \geq 0} \mathbb{E}[\|X^k\|]\lambda^k$$

[setting $r_1(X) = 0$ whenever X^k is not norm-integrable for some $k \geq 0$], and $r_2(X)$ as the radius of convergence of the series

$$\sum_{k \geq 0} \|\mathbb{E}[X^k]\|\lambda^k,$$

[setting $r_2(X) = 0$ whenever X^k is not integrable for some $k \geq 0$].

Note that $r_2(X) = R(\mathrm{ESig}(X))$. Remark also that $r_1(X) \leq r_2(X)$ and that Proposition 3.4 provides a partial converse when $V = \mathbb{R}^d$ and X is $G(\mathbb{R}^d)$-valued.

Note that $\mathrm{ESig}(X)$ exists whenever X is integrable as an E-valued random variable. The following proposition now provides a converse when X is G-valued. Recall that we identify E as a subspace of P (Corollary 2.5).

PROPOSITION 3.2. *Let X be a G-valued random variable. Then X is weakly integrable if and only if* $\mathrm{ESig}(X)$ *exists and lies in E. In this case* $\mathbb{E}[X] = \mathrm{ESig}(X)$.

In the case that V is normed, note that in order to conclude that a G-valued random variable X is (weakly) integrable (as an E-valued random variable), Proposition 3.2 implies that one only needs to check that each projection X^k is (weakly)

integrable and that $\|\mathbb{E}[X^k]\|$ decays sufficiently fast as $k \to \infty$. Remark that this is certainly not true for an arbitrary E-valued random variable.

We observe that for any $f \in E'$, it holds that $f^{\otimes 2} \circ \Delta \in E'$ and $f(g)^2 = f^{\otimes 2}(\Delta g)$ for all $g \in G$. In particular, for all $\mu \in \mathcal{P}(G)$, we have

$$\mu(|f|) \leq \sqrt{\mu(f^2)} = \sqrt{\mu(f^{\otimes 2} \circ \Delta)}. \tag{3.1}$$

This simple observation allows for very easy control of a measure through its barycenter. For example, whenever $\mu \in \mathcal{P}(G)$ and $\mathbb{E}[X]$ exists, it follows immediately that for all $f \in E'$, the real random variable $f(X)$ has finite moments of all orders.

The main idea behind the proof of Proposition 3.2 is that given the existence of $\mathbb{E}[X^k]$ for all $k \geq 0$, we wish to approximate $\mathbb{E}[f(X)]$ by $\sum_{k=0}^n \mathbb{E}[f(X^k)]$. Using the estimate (3.1) and the grading of the coproduct Δ, we apply dominated convergence to obtain $\mathbb{E}[f(X)] = \sum_{k \geq 0} \mathbb{E}[f(X^k)]$.

PROOF OF PROPOSITION 3.2. The "only if" direction is clear. Assume that $\mathrm{ESig}(X)$ exists and $\mathrm{ESig}(X) \in E$. As usual, let μ be the measure on G associated to X. We are required to show that f is μ-integrable and that $\mu(f) = \langle f, \mathrm{ESig}(X) \rangle$ for all $f \in E'$.

Recall the projection $\rho^k : E \mapsto V^{\widehat{\otimes} k}$. Treating $V^{\widehat{\otimes} k}$ as a subspace of E, for all $f \in E'$ denote $f^k := f \circ \rho^k \in E'$. Furthermore, we canonically embed $(V^{\widehat{\otimes} k})'$ into E' by $f \mapsto f \circ \rho^k$ for all $f \in (V^{\widehat{\otimes} k})'$. Observe that for all $f \in E'$, by Corollary 2.6, $\sum_{k=0}^n f^k$ converges uniformly on bounded sets (and a fortiori pointwise) to f.

Remark that for any $f \in E'$, $f \in (V^{\widehat{\otimes} k})'$ if and only if $f = f^k$. Recall that Δ is a graded linear map from $T(V)$ to $T(V)^{\otimes 2}$. In particular, for all $f_1 \in (V^{\widehat{\otimes} k})'$, $f_2 \in (V^{\widehat{\otimes} m})'$, and $x \in T(V)$, it holds that

$$(f_1 \otimes f_2)\Delta(x) = (f_1 \otimes f_2)\Delta(x^{k+m}). \tag{3.2}$$

As $T(V)$ is dense in E, (3.2) holds for all $x \in E$, from which it follows that $(f_1 \otimes f_2) \circ \Delta \in (V^{\widehat{\otimes}(k+m)})'$.

Let $f \in E'$ and note that $\mu(f^k) = \langle f^k, \mathbb{E}[X^k] \rangle$ for all $k \geq 0$. Since μ has support on G, it follows from (3.1) and (3.2) that

$$\mu\left(\sum_{k \geq 0} |f^k|\right) \leq \sum_{k \geq 0} \sqrt{\mu((f^k)^{\otimes 2} \circ \Delta)} = \sum_{k \geq 0} \sqrt{(f^k)^{\otimes 2} \Delta \mathbb{E}[X^{2k}]}. \tag{3.3}$$

Without loss of generality, we can assume that $|f(1)| \leq 1$. Let γ be a semi-norm on V such that $\exp(\gamma) \geq |f|$ and ξ a semi-norm on E such that $\xi \geq \exp(\gamma)^{\otimes 2} \circ \Delta$. It follows that $\exp(\gamma) \geq |f^k|$ for all $k \geq 0$, and thus $\xi \geq |(f^k)^{\otimes 2} \circ \Delta|$ for all $k \geq 0$.

Since $\mathrm{ESig}(X) \in E$, it follows from Corollary 2.5 that $\sum_{k \geq 0} \sqrt{\xi(\mathbb{E}[X^k])}$ is finite, and hence (3.3) is finite. By dominated convergence, we obtain

$$\mu(f) = \lim_{n \to \infty} \mu\left(\sum_{k=0}^n f^k\right).$$

It then follows that $\mu(f) = \langle f, \mathrm{ESig}(X)\rangle$ as desired since

$$\lim_{n\to\infty} \mu\left(\sum_{k=0}^{n} f^k\right) = \lim_{n\to\infty} \sum_{k=0}^{n} \langle f^k, \mathbb{E}[X^k]\rangle = \langle f, \mathrm{ESig}(X)\rangle. \qquad \square$$

COROLLARY 3.3. *Let V be a normed space and X a G-valued random variable. Then $\mathbb{E}[X] \in E$ exists if and only if $r_2(X) = \infty$, that is,* $\mathrm{ESig}(X)$ *exists and has an infinite radius of convergence. In this case* $\mathbb{E}[X] = \mathrm{ESig}(X)$.

We are moreover able to show explicit bounds between $r_1(X)$ and $r_2(X)$ when $V = \mathbb{R}^d$. Suppose first that V is a normed space. Remark that $\|\Delta v\| = 2\|v\|$ for all $v \in V$, from which it follows that $\|\Delta|_{V^{\hat{\otimes} k}}\| = 2^k$ and thus

$$\|\Delta x^k\| \le 2^k \|x^k\| \qquad \text{for all } x \in E. \tag{3.4}$$

Let $V = \mathbb{R}^d$ equipped with the ℓ^1 norm from its standard basis $e_1, \dots, e_d$, and denote $e_I = e_{i_1} \cdots e_{i_k} \in V^{\otimes k}$ for a word $I = i_1 \cdots i_k$ in the alphabet $\{1, \dots, d\}$. Then the grading of Δ gives

$$\begin{aligned}
\mathbb{E}[\|X^k\|^2] &= \mathbb{E}\left[\left(\sum_{|I|=k} |\langle e_I, X^k\rangle|\right)^2\right] \le d^k \mathbb{E}\left[\sum_{|I|=k} \langle e_I, X^k\rangle^2\right] \\
&= d^k \sum_{|I|=k} e_I^{\otimes 2} \Delta \mathbb{E}[X^{2k}] \\
&\le d^k \|\Delta \mathbb{E}[X^{2k}]\|,
\end{aligned}$$

where the last inequality follows since $(e_I \otimes e_J)_{|I|=|J|=k}$ is an ℓ^1 basis for $V^{\otimes 2k}$. Using (3.4), we now obtain the following.

PROPOSITION 3.4. *Let X be a $G(\mathbb{R}^d)$-valued random variable. It follows that $\mathbb{E}[\|X^k\|^2] \le d^k 2^{2k} \|\mathbb{E}[X^{2k}]\|$. In particular, $r_1(X) \le r_2(X) \le 2d^{1/2} r_1(X)$.*

4. Representations. Recall that for any Hopf algebra, one may define the tensor product and dual of representations via the coproduct and antipode by $M_1 \otimes M_2(x) := (M_1 \otimes M_2)\Delta(x)$ and $M^*(x) := M(\alpha(x))^*$. By virtue of continuity of Δ and α, we observe that the family of continuous representations of E over finite dimensional Hilbert spaces is closed under tensor products and duals.

DEFINITION 4.1. Denote by $\mathcal{A}(V)$ the family of finite dimensional representations of E which arise from extensions of all linear maps $M \in \mathbf{L}(V, \mathfrak{u}(H_M))$, where H_M ranges over all finite dimensional complex Hilbert spaces and $\mathfrak{u}(H_M)$ denotes the Lie algebra of the anti-Hermitian operators on H_M. Denote by $\mathcal{C}(V)$ the set of corresponding matrix coefficients, that is, the set of linear functionals $M_{u,v} \in \mathbf{L}(E, \mathbb{C})$, $M_{u,v}(x) = \langle M(x)u, v\rangle$ for all $M \in \mathcal{A}$ and $u, v \in H_M$.

The family $\mathcal{A}$ possesses the desirable property that it is closed under taking tensor products and duals of representations. Moreover, we see that $\mathcal{A}$ contains exactly those finite dimensional representations of E which preserve involution, that is, $M(\alpha x) = M(x)^*$ for all $x \in E$. It follows that every $M \in \mathcal{A}$ is a unitary representation of the group U, and thus of G.

Observe that the tensor product $M_1 \otimes M_2$ (of any representations M_1, M_2 of E) coincides on G with the usual group-theoretic tensor product of representations. Moreover, the dual representation M^* of $M \in \mathcal{A}$ can be identified on U with the conjugate representation of M on U. It follows that $\mathcal{C}\,|_G$ forms a $*$-subalgebra of $C_b(G, \mathbb{C})$.

Let S be a topological space and F a separating $*$-subalgebra of $C_b(S, \mathbb{C})$. Recall that for Radon measures μ and ν on S (see [3], Definition 7.1.1), it follows from the Stone–Weierstrass theorem that $\mu = \nu$ if and only if $\mu(f) = \nu(f)$ for all $f \in F$ ([3] Exercise 7.14.79). We now obtain the following from the above discussion.

LEMMA 4.2. *Assume that $\mathcal{A}$ separates the points of G. Then for tight Borel measures μ, ν on G, $\mu = \nu$ if and only if $\mu(f) = \nu(f)$ for all $f \in \mathcal{C}$, or equivalently, $\mu(M) = \nu(M)$ for all $M \in \mathcal{A}$.*

We show in Theorem 4.8 that in fact $\mathcal{A}(\mathbb{R}^d)$ separates the points of $E(\mathbb{R}^d)$.

4.1. *Separation of points.* We investigate conditions under which algebra homomorphisms of E separate points. Though ultimately we apply the theory to the case $V = \mathbb{R}^d$, the arguments used in the general case are exactly the same and we provide them here.

For a Banach algebra A and $M \in \mathbf{L}(V, A)$, let (λM) denote the algebra homomorphism on E induced by $\lambda M \in \mathbf{L}(V, A)$ (λ possibly complex if A is over $\mathbb{C}$). For $\lambda \in \mathbb{R}$, let $\delta_\lambda : E \mapsto E$ denote the dilation operator $\delta_\lambda(x^0, x^1, \ldots) = (\lambda^0 x^0, \lambda^1 x^1, \ldots)$ [note that $(\lambda M) = M\delta_\lambda$ for $\lambda \in \mathbb{R}$].

LEMMA 4.3. *Let V be locally convex, A a Banach algebra and $M \in \mathbf{L}(V, A)$. Let $x \in E$ such that $M(x^k) \neq 0$ for some $k \geq 0$. Then there exists $\varepsilon > 0$ sufficiently small such that $(\varepsilon M)(x) \neq 0$.*

PROOF. Since $\|M(x)\|$ is a semi-norm on E, $\sum_{k\geq 0} \|M(x^k)\|$ converges by Corollary 2.5, from which the conclusion follows. $\square$

Let $\mathbb{F}$ be a field and A an $\mathbb{F}$-algebra. A polynomial identity over $\mathbb{F}$ on a subset $Q \subseteq A$ is a polynomial in non-commuting indeterminates $x_1, \ldots, x_k$, with coefficients in $\mathbb{F}$, which is non-zero (i.e., not every coefficient is zero) and which vanishes under all substitutions of variables $x_1, \ldots, x_k \in Q$. We refer to Giambruno and Zaicev [17] for further details.

Let V be a vector space with Hamel basis Θ. Then the set of pure tensors $\Theta^{\otimes k} = \{v_1 \cdots v_k \mid v_j \in \Theta, 1 \le j \le k\}$ is a Hamel basis for $V^{\otimes k}$. Thus for every $x \in V^{\otimes k}$ define Θ_x as the finite set of vectors in Θ which appear in the representation of x in the basis $\Theta^{\otimes k}$. Define f_x^{Θ} the canonical formal non-commuting polynomial in indeterminates Θ_x associated with x. As Θ_x is a finite set, the following is a consequence of the Hahn–Banach theorem.

LEMMA 4.4. *Let V be a locally convex space with Hamel basis Θ, A an algebra which is a topological vector space, and $Q \subseteq A$ a subset. Let $k \ge 0$ and $x \in V^{\otimes k}$. The following two assertions are equivalent.*

(i) *f_x^{Θ} is not a polynomial identity over $\mathbb{R}$ on Q.*

(ii) *There exists a continuous linear map $M : V \mapsto \operatorname{span}(Q)$ such that $M(x)$ is non-zero and $M(v)$ is in Q for all $v \in \Theta_x$.*

REMARK 4.5. If one is not interested in the topological aspects, the same statement holds if one replaces $\mathbb{R}$ by a field $\mathbb{F}$, V by a vector space over $\mathbb{F}$, A by an $\mathbb{F}$-algebra, and drops the continuity assumption in (ii).

4.2. *Polynomial identities over Lie algebras.* From Lemmas 4.3 and 4.4, it is clear that to study how representations in $\mathcal{A}(\mathbb{R}^d)$ separate the points of $E(\mathbb{R}^d)$, we must look at polynomial identities in unitary Lie algebras. Let $m \ge 1$ be an integer and denote by $\cdot^s$ the symplectic involution on $M_{2m}(\mathbb{C})$, which we recall is an involution of the first kind (see [16]).

Recall the real Lie algebra $\mathfrak{sp}(m) = \{u \in \mathfrak{u}(\mathbb{C}^{2m}) \mid u^s + u = 0\}$ [$\mathfrak{sp}(m)$ is the Lie algebra of the compact symplectic group $Sp(m)$]. A closely related complex Lie subalgebra of $\mathfrak{gl}(\mathbb{C}^{2m})$ is $\mathfrak{sp}(m, \mathbb{C}) = \{u \in M_{2m}(\mathbb{C}) \mid u^s + u = 0\}$. It holds that $\mathfrak{sp}(m, \mathbb{C})$ is the complexification of $\mathfrak{sp}(m)$.

We now illustrate our interest in the Lie algebras $\mathfrak{sp}(m)$ and $\mathfrak{sp}(m, \mathbb{C})$. From the remark that $\mathfrak{sp}(m, \mathbb{C}) = \{u - u^s \mid u \in M_{2m}(\mathbb{C})\}$, we may reformulate a result due to Giambruno and Valenti as follows.

THEOREM 4.6 ([16] Theorem 6). *Let $m \ge 2$ and $f(x_1, \dots, x_k)$ a polynomial identity over $\mathbb{C}$ on $\mathfrak{sp}(m, \mathbb{C}) \subset M_{2m}(\mathbb{C})$. Then $\deg(f) > 3m$.*

The following is a slight generalization of [17] Theorem 1.3.2 and follows from exactly the same inductive proof.

LEMMA 4.7. *Let $\mathbb{F}$ be an infinite field, A an $\mathbb{F}$-algebra and Q a linear subspace of A. If f is a polynomial identity over $\mathbb{F}$ on Q, then every multi-homogeneous component of f is a polynomial identity over $\mathbb{F}$ on Q.*

We remark that every multi-homogeneous polynomial identity over $\mathbb{C}$ (and a fortiori over $\mathbb{R}$) on $\mathfrak{sp}(m) \subset M_{2m}(\mathbb{C})$ is also a polynomial identity over $\mathbb{C}$ on its complexification $\mathfrak{sp}(m, \mathbb{C})$. Thus if f is a polynomial identity over $\mathbb{R}$ on $\mathfrak{sp}(m)$ for $m \geq 2$, then by Theorem 4.6 and Lemma 4.7, every multi-homogeneous component of f has degree greater than $3m$. Together with Lemmas 4.3 and 4.4, we have the following result.

THEOREM 4.8. *Let $x \in E(\mathbb{R}^d)$ such that $x^k \neq 0$ for some $k \geq 0$. Then for any integer $m \geq \max\{2, k/3\}$ there exists $M \in \mathbf{L}(\mathbb{R}^d, \mathfrak{sp}(m))$ such that $M(x) \neq 0$. In particular, $\mathcal{A}(\mathbb{R}^d)$ separates the points of $E(\mathbb{R}^d)$.*

REMARK 4.9. The necessity that $V = \mathbb{R}^d$ only came into the above argument to ensure that $V^{\otimes k} = V^{\widehat{\otimes} k}$. If one was able to find an analogue of Lemma 4.4 for elements $x \in V^{\widehat{\otimes} k}$, or an analogue of Theorem 4.6 for appropriate series of polynomials of bounded degree but an unbounded number of indeterminates, then one could readily extend Theorem 4.8 to the case when V is infinite dimensional.

COROLLARY 4.10. *The group $U(\mathbb{R}^d)$ is maximally almost periodic.*

REMARK 4.11. For $d \geq 2$, the topological group $G(\mathbb{R}^d)$ [and thus $U(\mathbb{R}^d)$] is not locally compact. To observe this, let $V = \mathbb{R}^d$ and $L(V)$ be the smallest Lie algebra in $T(V)$ containing V. Since every $\ell \in L(V)$ satisfies $\Delta(\ell) = 1 \otimes \ell + \ell \otimes 1$ ([33] Theorem 1.4), a direct calculation shows that $\exp(\ell) \in G$.

Let $u, v \in V$ be linearly independent elements and $W = \operatorname{span}(u, v)$. Observe that $L(W)$ contains a non-zero element in $W^{\otimes k}$ for every $k \geq 1$. In light of Proposition 2.3, for any neighborhood of zero B of $L(W)$ one can construct a sequence $(\ell_n)_{n \geq 1} \in B$ such that $\gamma(\exp(\ell_i) - \exp(\ell_j)) \geq 1$ for all $i \neq j$ and some semi-norm γ on E. Since $\exp : L(W) \mapsto G$ is continuous ([1] Theorem 3), it follows that no neighborhood of the identity in G is contained in a sequentially compact set (the same argument more generally applies whenever V is metrizable).

It follows from Corollary 2.4 that E is Polish whenever V is metrizable and separable, and thus G, as a closed subset of E, is also Polish. By Lemma 4.2 and Theorem 4.8, we have the following.

COROLLARY 4.12. *For Borel probability measures μ and ν on $G(\mathbb{R}^d)$, it holds that $\mu = \nu$ if and only if $\mu(f) = \nu(f)$ for all $f \in \mathcal{C}(\mathbb{R}^d)$, or equivalently, $\mu(M) = \nu(M)$ for all $M \in \mathcal{A}(\mathbb{R}^d)$.*

For a Borel probability measure μ on $G(\mathbb{R}^d)$, with associated random variable X, we are thus able to define its characteristic function (or Fourier transform) by $\phi_X = \widehat{\mu} := \mu\,|_{\mathcal{A}}$, which uniquely characterizes μ.

5. Signatures of paths. We now discuss the space E in the setting of rough paths theory. The main connection is that the signature of any geometric rough path in $\mathbb{R}^d$ lies in $G(\mathbb{R}^d)$. We treat rough paths in the sense of Lyons and refer to [15] and [28] for details and terminology.

Let V be a Banach space, $p \geq 1$, $T > 0$, and $\Delta_{[0,T]} = \{(s,t) \mid 0 \leq s \leq t \leq T\}$. Let ω denote a control function and $T^n(V) = \bigoplus_{0 \leq k \leq n} V^{\widehat{\otimes} k}$ the truncated tensor algebra. We recall that the space of p-rough paths $\Omega_p(V)$ is the collection of all continuous multiplicative maps $\mathbf{x} : \Delta_{[0,T]} \mapsto T^{\lfloor p \rfloor}$ with p-variation controlled by some control ω, that is:

(a) $\mathbf{x}^0_{s,t} = 1$ and $\mathbf{x}_{s,t}\mathbf{x}_{t,u} = \mathbf{x}_{s,u}$ for all $0 \leq s \leq t \leq u \leq T$, and
(b) for some control ω one has

$$\sup_{1 \leq k \leq \lfloor p \rfloor} \left((k/p)!\beta_p \|\mathbf{x}^k_{s,t}\|\right)^{p/k} \leq \omega(s,t), \qquad \forall (s,t) \in \Delta_{[0,T]}, \tag{5.1}$$

where β_p is a constant that only depends on p.

The map $\mathbf{x}$ may alternatively be viewed as a path $\mathbf{x}_{0,\cdot} : [0,T] \mapsto T^{\lfloor p \rfloor}$, $t \mapsto \mathbf{x}_{0,t}$ of finite p-variation, that is,

$$\|\mathbf{x}\|_{p\text{-var};[0,T]} := \sum_{k=1}^{\lfloor p \rfloor} \sup_{\mathcal{D} \subset [0,T]} \left(\sum_{t_j \in \mathcal{D}} \left((k/p)!\beta_p \|\mathbf{x}^k_{t_j,t_{j+1}}\|\right)^{p/k} \right)^{1/p} \tag{5.2}$$

is finite, which completely characterizes $\mathbf{x}$ due to the multiplicative property (a) (noting that $\mathbf{x}_{s,t} = \mathbf{x}^{-1}_{0,s}\mathbf{x}_{0,t}$).

Let $\mathbf{x} \in \Omega_p$ satisfy (5.1) for some control ω. A fundamental result of rough paths theory is that for all $n \geq \lfloor p \rfloor$ there exists a unique lift $S_n(\mathbf{x}) : \Delta_{[0,T]} \mapsto T^n$ such that (a) and (b) remain true for the same ω and with $\sup_{1 \leq k \leq \lfloor p \rfloor}$ replaced by $\sup_{1 \leq k \leq n}$ in (5.1) ([28] Theorem 3.1.2). Equivalently, there exists a unique lift to the entire product space $S(\mathbf{x}) : \Delta_{[0,T]} \mapsto P = \prod_{k \geq 0} V^{\widehat{\otimes} k}$ such that (a) and (b) remain true for the same ω and with $\sup_{1 \leq k \leq \lfloor p \rfloor}$ replaced by $\sup_{1 \leq k}$ in (5.1).

An immediate consequence of the factorial decay in (5.1) is that the lift $S(\mathbf{x})$ takes values in the space E for any $p \geq 1$ (see Corollary 2.5).

REMARK 5.1. While the value of β_p does not affect the definition of the space Ω_p, its existence is crucial to ensure the factorial decay arising from the lift. On this point, we mention the work of Hara and Hino [20] who have resolved a conjecture on the optimal possible value of β_p.

We thus make a canonical extension of the space Ω_p.

DEFINITION 5.2. Define the space ΩE_p as the set of maps $\mathbf{x} : \Delta_{[0,T]} \mapsto E$ which satisfy (a) and (b) with $\sup_{1 \leq k \leq \lfloor p \rfloor}$ replaced by $\sup_{1 \leq k}$ in (5.1).

It follows that the lift S is a bijective map from Ω_p to ΩE_p, with inverse provided naturally by the $\lfloor p \rfloor$th level truncation $(\mathbf{x}^0_{s,t}, \mathbf{x}^1_{s,t}, \dots) \mapsto (\mathbf{x}^0_{s,t}, \mathbf{x}^1_{s,t}, \dots, \mathbf{x}^{\lfloor p \rfloor}_{s,t})$.

The element $S(\mathbf{x})_{0,T} \in E$ is called the *signature* of a rough path $\mathbf{x} \in \Omega_p$. For $1 \le p < 2$, $S(\mathbf{x})_{0,T}$ is precisely the sequence of iterated integrals of the path $\mathbf{x}_{0,\cdot} : [0, T] \mapsto V$ taken in the sense of Young.

REMARK 5.3. The only property of the projective tensor norm used above is that the projective extension provides a sub-multiplicative system of norms. Completely analogous results hold true if one equips $T(V)$ with any system of sub-multiplicative norms and defines E as the completion of $T(V)$ under scalar dilations of these norms. Note that in the case $V = \mathbb{R}^d$, all these systems lead to identical definitions and topologies on the space E.

The lift S moreover exhibits a natural continuity property with respect to the p-variation topology on Ω_p, which we briefly recall. For $\mathbf{x}, (\mathbf{x}(n))_{n\ge1} \in \Omega_p$, a control ω and a sequence of reals $(a_n)_{n\ge1}$ with $a_n \ge 0$, consider the statement

$$\begin{aligned} &\omega \text{ controls the } p\text{-variation of } \mathbf{x} \text{ and } \mathbf{x}(n) \text{ for all } n \ge 1, \text{ and} \\ &\sup_{1\le k\le \lfloor p \rfloor} \big((k/p)!\beta_p a_n \|\mathbf{x}(n)^k_{s,t} - \mathbf{x}^k_{s,t}\|\big)^{p/k} \le \omega(s,t), \qquad \forall (s,t) \in \Delta_{[0,T]}. \end{aligned} \tag{5.3}$$

When (5.3) is satisfied for some control ω and a sequence $(a_n)_{n\ge1}$ such that $a_n \ge 0$ and $\lim_{n\to\infty} a_n = \infty$, we say that $\mathbf{x}(n)$ converges to $\mathbf{x}$ in the p-variation topology of Ω_p. One makes the same definition for $\mathbf{x}, (\mathbf{x}(n))_{n\ge1} \in \Omega E_p$ with $\sup_{1\le k\le\lfloor p \rfloor}$ replaced by $\sup_{1\le k}$ in (5.3).

The following is an immediate consequence of the continuity of the individual lifts S_n for $n \ge \lfloor p \rfloor$ ([28] Theorem 3.1.3).

PROPOSITION 5.4. *If $\mathbf{x}, (\mathbf{x}(n))_{n\ge1} \in \Omega_p$ satisfy (5.3) for some ω and $(a_n)_{n\ge1}$ with $a_n \ge 0$, then $S(\mathbf{x}), (S(\mathbf{x}(n)))_{n\ge1} \in \Omega E_p$ satisfy (5.3) for the same control ω and sequence $(a_n)_{n\ge1}$.*

In particular, S is continuous (and thus a homeomorphism) when Ω_p and ΩE_p are equipped with their respective p-variation topologies.

We equip Ω_p with the p-variation topology and define the evaluation map $\mathcal{I}_p : \Omega_p \mapsto E, \mathbf{x} \mapsto S(\mathbf{x})_{0,T}$.

COROLLARY 5.5. *The map $\mathcal{I}_p$ is continuous.*

Define the space of geometric p-rough paths $G\Omega_p$ as the closure of $S_{\lfloor p \rfloor}(\Omega_1)$ in Ω_p.

DEFINITION 5.6. For $p \ge 1$, define $R_p(V) = \{S(\mathbf{x})_{0,T} \mid \mathbf{x} \in G\Omega_p\} \subset E$ as the set of signatures of all geometric p-rough paths.

We equip R_p with the subspace topology from E. Observe that R_1 is dense in R_p as a consequence of Corollary 5.5.

Remark that R_p is closed under multiplication in E and that for all $\mathbf{x} \in G\Omega_p$, the inverse of $S(\mathbf{x})_{0,T}$ is $S(\mathbf{y})_{0,T} = \alpha(S(\mathbf{x})_{0,T})$, where $\mathbf{y} \in G\Omega_p$ is the reversal of $\mathbf{x}$ and α is the antipode of E defined in Section 3 ([28] Theorem 3.3.3). Thus R_p is a subgroup of $U = \{g \in E \mid \alpha(g) = g^{-1}\}$.

5.1. *Finite dimensional case.* In this section, we consider $V = \mathbb{R}^d$. It follows that $P(\mathbb{R}^d)$ [resp., $E(\mathbb{R}^d)$] can be identified with the algebra of non-commuting formal power series in d indeterminates (resp., with an infinite radius of convergence).

We remark that the coproduct Δ of $E(\mathbb{R}^d)$ is given by a locally finite formula involving the shuffle product ([33] Proposition 1.8) and an element $g \in E(\mathbb{R}^d)$ is in $G(\mathbb{R}^d)$ precisely when $(g^0, g^1, \dots, g^n)$ is in the free n-step nilpotent Lie group $G^n(\mathbb{R}^d)$ for all $n \geq 1$ ([31] Lemma 2.24). Unless otherwise stated, we always equip $G^n(\mathbb{R}^d)$ with the metric induced by the Carnot–Carathéodory norm.

A fundamental result of Chen [6] is that the signature of a bounded variation path in $\mathbb{R}^d$ is a group-like element of $E(\mathbb{R}^d)$ ([31] Section 2.2.5), and thus $R_1(\mathbb{R}^d) \subset G(\mathbb{R}^d)$. Since G is closed in E, we immediately obtain the inclusions $R_p(\mathbb{R}^d) \subset \overline{R_1(\mathbb{R}^d)} \subseteq G(\mathbb{R}^d)$ for all $p \geq 1$.

A closely related set to $G\Omega_p(\mathbb{R}^d)$ is the space $WG\Omega_p(\mathbb{R}^d) \subset \Omega_p(\mathbb{R}^d)$ of *weakly* geometric p-rough paths, that is, those p-rough paths $\mathbf{x} \in \Omega_p(\mathbb{R}^d)$ which take values in the free $\lfloor p \rfloor$-step nilpotent Lie group, that is,

$$\left(\mathbf{x}^0_{s,t}, \mathbf{x}^1_{s,t}, \dots, \mathbf{x}^{\lfloor p \rfloor}_{s,t}\right) \in G^{\lfloor p \rfloor}\left(\mathbb{R}^d\right), \qquad \forall (s,t) \in \Delta_{[0,T]}.$$

We note the strict inclusions

$$G\Omega_p\left(\mathbb{R}^d\right) \subset WG\Omega_p\left(\mathbb{R}^d\right) \subset G\Omega_{p'}\left(\mathbb{R}^d\right)$$

for all $p' > p \geq 1$ ([15] Section 8.5), and thus $WR_p(\mathbb{R}^d) \subset R_{p'}(\mathbb{R}^d)$, where

$$WR_p\left(\mathbb{R}^d\right) = \left\{S(\mathbf{x})_{0,T} \mid \mathbf{x} \in WG\Omega_p\left(\mathbb{R}^d\right)\right\}.$$

PROPOSITION 5.7. *Let $p \geq 1$. Then $WR_p(\mathbb{R}^d)$ is σ-compact in $G(\mathbb{R}^d)$. In particular, $WR_p(\mathbb{R}^d)$ is a Borel set of $G(\mathbb{R}^d)$.*

For the proof, we recall the (homogeneous) p-variation metric $d_{p\text{-var}}$ on $\Omega_p(\mathbb{R}^d)$ under which $(\Omega_p(\mathbb{R}^d), d_{p\text{-var}})$ is a complete metric space with a coarser topology than the p-variation topology, but for which convergence of a sequence in $d_{p\text{-var}}$ implies the existence of a subsequence which converges in the p-variation topology (see [15], Section 8, and [28], Proposition 3.3.3, but note the differing notations for homogeneous and inhomogeneous metrics in the two texts; we use the notation of [15]).

PROOF OF PROPOSITION 5.7. For $r > 0$, consider the set

$$B_p^r = \{\mathbf{x} \in WG\Omega_p(\mathbb{R}^d) \mid \|\mathbf{x}\|_{p\text{-var};[0,T]} \le r\}.$$

For every $\mathbf{x} \in B_p^r$ there exists $\mathbf{y} \in B_p^r$ such that $t \mapsto \mathbf{y}_{0,t}$ is $(1/p)$-Hölder continuous, with Hölder coefficient depending only on r and p, and such that $\mathbf{x}$ is a reparametrization of $\mathbf{y}$ (cf. [15] Proposition 5.14). In particular, $\|\mathbf{x}\|_{p\text{-var};[0,T]} = \|\mathbf{y}\|_{p\text{-var};[0,T]}$ and $S(\mathbf{x})_{0,T} = S(\mathbf{y})_{0,T}$. Let $C_p^r \subset B_p^r$ denote the set of all such $\mathbf{y}$ associated with all $\mathbf{x} \in B_p^r$.

Let $p' > p$ be such that $\lfloor p' \rfloor = \lfloor p \rfloor$. It follows from an interpolation estimate and the Arzelà–Ascoli theorem ([15] Lemma 5.12, Proposition 8.17) that C_p^r is compact in $(G\Omega_{p'}(\mathbb{R}^d), d_{p'\text{-var}})$ and thus sequentially compact in $G\Omega_{p'}(\mathbb{R}^d)$ under the p'-variation topology.

Since $\mathcal{I}_{p'} : G\Omega_{p'}(\mathbb{R}^d) \mapsto WR_{p'}(\mathbb{R}^d)$ is continuous by Corollary 5.5, and since $\mathcal{I}_{p'}(C_p^r) = \mathcal{I}_{p'}(B_p^r)$, it follows that $\mathcal{I}_{p'}(B_p^r)$ is sequentially compact in $WR_{p'}(\mathbb{R}^d)$, and thus compact. Since $WR_p(\mathbb{R}^d) = \bigcup_{r \ge 1} \mathcal{I}_{p'}(B_p^r)$, it follows that $WR_p(\mathbb{R}^d)$ is σ-compact in $G(\mathbb{R}^d)$. $\square$

We lastly record here a consequence of Theorem 4.8 and Theorem 4 of [19], which strengthens Corollary 1.7 therein, and which was originally observed by Prof. Thierry Lévy.

COROLLARY 5.8. *A path of bounded variation in $\mathbb{R}^d$ is tree-like if and only if its Cartan development into every finite-dimensional compact Lie group is trivial.*

6. Expected signature. Our main focus in this section is the expected signature of G-valued random variables and its connection with the characteristic function defined at the end of Section 4.

6.1. *Moment problem.* In this section, we study the moment problem for $G(\mathbb{R}^d)$-valued random variables, that is, conditions under which a $G(\mathbb{R}^d)$-valued random variable is uniquely determined by its expected signature. We mention here that a large part of the results in this section arose from discussions with Dr. Ni Hao, and we hope to soon jointly expand on this material in a future paper.

When V is a normed space, recall from Corollary 3.3 that if X is a G-valued random variable such that $\mathrm{ESig}(X)$ exists and has an infinite radius of convergence, then $\mathbb{E}[X]$ exists as an element of E and is equal to $\mathrm{ESig}(X)$. Thus $\mathbb{E}[f(X)]$ is completely determined by $\mathrm{ESig}(X)$ for all $f \in E'$, and in particular for all $M \in \mathcal{A}$. The following is now a consequence of the uniqueness of probability measures from Corollary 4.12.

PROPOSITION 6.1. *Let X and Y be $G(\mathbb{R}^d)$-valued random variables such that* $\mathrm{ESig}(X) = \mathrm{ESig}(Y)$ *and* $\mathrm{ESig}(X) \in E$, *that is,* $\mathrm{ESig}(X)$ *has an infinite radius of convergence. Then $X \overset{\mathcal{D}}{=} Y$.*

Recall from Corollary 5.5 that the evaluation map $\mathcal{I}_p : \Omega_p \mapsto E$ is continuous. It follows that the signature $S(\mathbf{X})_{0,T}$ of any Ω_p-valued (resp., $G\Omega_p$-valued) random variable $\mathbf{X}$ is a well-defined (Borel) E-valued [resp., U-valued, or $G(\mathbb{R}^d)$-valued in case $V = \mathbb{R}^d$] random variable.

EXAMPLE 6.2. We apply Proposition 6.1 to the Lévy–Khintchine formula established in [14]. Recall that [14] Theorem 52 describes a class of Lévy processes in $G^2(\mathbb{R}^d)$ which, by adding appropriate adjustments for jumps, may be treated as $G\Omega_p(\mathbb{R}^d)$-valued random variables for some $2 < p < 3$ [which in particular includes all Lévy processes in $\mathbb{R}^d$ lifted to $G^2(\mathbb{R}^d)$ by solving the associated Marcus stochastic differential equation].

Let $(\mathbf{a}, \mathbf{b}, \mathbf{K})$ be the triplet of a Lévy process in $G^2(\mathbb{R}^d)$ satisfying the conditions of [14] Theorem 52. Let $\mathbf{X}$ be the associated $G\Omega_p(\mathbb{R}^d)$-valued random variable for some $2 < p < 3$, and $X = S(\mathbf{X})_{0,T}$ its signature. It follows from [14] Theorem 55 that whenever the measure $\mathbf{1}_{|\log x|>1}\mathbf{K}(dx)$ integrates all powers of $|\log x|$, $\mathrm{ESig}(X)$ exists as an element of $P(\mathbb{R}^d) = \prod_{k\geq 0}(\mathbb{R}^d)^{\otimes k}$ and is given by

$$\mathrm{ESig}(X) = \exp\Bigg[T\Bigg(\frac{1}{2}\sum_{i,j=1}^{d}\mathbf{a}^{i,j}e_ie_j + \sum_{i=1}^{d}\mathbf{b}^i e_i + \sum_{j<k}\mathbf{b}^{j,k}[e_j, e_k] + \int_{G^2(\mathbb{R}^d)}\big[\exp(\log x) - x\mathbf{1}_{|\log x|<1}\big]\mathbf{K}(dx)\Bigg)\Bigg],$$

where both appearances of exp are understood as exponentiation in $P(\mathbb{R}^d)$. In particular, defining the norm $\|x\|_\lambda = \sum_{k\geq 0}\lambda^k\|x^k\|$ on E for $\lambda > 0$, one can readily see that $\mathrm{ESig}(X) \in E$ whenever

$$\int_{G^2(\mathbb{R}^d)}\big\|\exp(\log x) - x\mathbf{1}_{|\log x|<1}\big\|_\lambda\mathbf{K}(dx) < \infty \qquad \text{for all } \lambda > 0. \tag{6.1}$$

Note that if $\mathbf{X}$ is the Marcus lift of a Lévy process in $\mathbb{R}^d$ with triplet (a, b, K), then (6.1) holds precisely when

$$\int_{\mathbb{R}^d}\big(e^{\lambda|x|} - 1 - \mathbf{1}_{|x|<1}\lambda|x|\big)K(dx) < \infty \qquad \text{for all } \lambda > 0.$$

It follows by Proposition 6.1 that whenever (6.1) is satisfied, $S(\mathbf{X})_{0,T}$ is uniquely determined as a $G(\mathbb{R}^d)$-valued random variable by its expected signature.

Recall the radius of convergence $r_1(X)$ from Definition 3.1. Theorem 6.3 below provides sufficient conditions to ensure that $r_1(X) > 0$ or $r_1(X) = \infty$ without explicit knowledge of $\mathrm{ESig}(X)$.

For a subset $B \subseteq A$ of an algebra A and $n \geq 1$, define $B^n = \{x_1, \dots, x_n \mid x_1, \dots, x_n \in B\}$. For an element $x \in A$, define $B(x) = \inf\{n \geq 1 \mid x \in B^n\}$ [taking $B(x) = \infty$ if $x \notin B^n$ for all $n \geq 1$].

Note that for a topological algebra A with (jointly) continuous multiplication, an A-valued random variable X, and a (Borel) measurable set $B \subset A$, $B(X)$ is a well-defined random variable in $\{1, 2, \dots\} \cup \{\infty\}$.

THEOREM 6.3. *Let V be a normed space and X an E-valued random variable. Suppose there exists a bounded, measurable set $B \subset E$ such that $B(X)$ has an exponential tail, that is, $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for some $\lambda > 0$. Then $r_1(X) > 0$. If moreover $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for all $\lambda > 0$, then $r_1(X) = \infty$.*

PROOF. Equip E with the projective extension of the norm on V. For any $r > 0$ and $\lambda > 0$ such that $\sup_{x \in B} \|\delta_r(x)\| < e^{\lambda}$, it holds that

$$\sum_{k \geq 0} r^k \mathbb{E}[\|X^k\|] = \mathbb{E}[\|\delta_r(X)\|] \leq \mathbb{E}[e^{\lambda B(X)}], \tag{6.2}$$

where the inequality follows from the fact that $\delta_r(X) = \delta_r(X_1) \cdots \delta_r(X_{B(X)})$ for some $X_1, \dots, X_{B(X)} \in B$.

Suppose first that $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for all $\lambda > 0$. For any $r > 0$ let $\lambda > 0$ be sufficiently large such that $\sup_{x \in B} \|\delta_r(x)\| < e^{\lambda}$. Then (6.2) implies that $r_1(X) \geq r$, and thus $r_1(X) = \infty$.

Suppose now that $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for some $\lambda > 0$. By Proposition 2.10, the functions δ_r converge strongly to δ_0 as $r \to 0$ and, in particular, uniformly on B. Thus, there exists $r > 0$ such that $\sup_{x \in B} \|\delta_r(x)\| < e^{\lambda}$. Then (6.2) implies that $r_1(X) \geq r > 0$ as desired. $\square$

We demonstrate how to apply Theorem 6.3 to random variables arising from signatures of geometric rough paths.

Let V be a Banach space and $p \geq 1$. We note that for any $\mathbf{x} \in \Omega_p$, $\omega_{\mathbf{x}}(s,t) := \|\mathbf{x}\|^p_{p\text{-var};[s,t]}$ defines a control for which (5.1) is satisfied. Thus for all $k \geq 1$, the lift $S(\mathbf{x}) : \Delta_{[0,T]} \mapsto E$ satisfies

$$\|S(\mathbf{x})^k_{0,T}\| \leq \frac{\omega(0,T)^{k/p}}{\beta_p (k/p)!}.$$

We hence define

$$K_p = \Big\{ x \in E \mid \sup_{k \geq 0} \beta_p (k/p)! \|x^k\| \leq 1 \Big\}$$

and observe that $S(\mathbf{x})_{0,T} \in K_p$ for every $\mathbf{x} \in \Omega_p$ with $\|\mathbf{x}\|_{p\text{-var};[0,T]} \leq 1$. Observe furthermore that K_p is bounded and measurable in E.

For $\mathbf{x} \in \Omega_p$, define $k_p(\mathbf{x}) = K_p(S(\mathbf{x})_{0,T})$, that is, the minimum positive integer k for which there exist $x_1, \dots, x_k \in K_p$ such that $S(\mathbf{x})_{0,T} = x_1, \dots, x_k$.

We briefly recall the construction of the greedy sequence and function $N_{\kappa,[0,T],p}(\mathbf{x})$ introduced in [4]. For $\kappa > 0$, define the sequence of times $\tau_0 = 0$,

$$\tau_{j+1} = \inf\{t > \tau_j \mid \omega_{\mathbf{x}}(\tau_j, t) \geq \kappa\} \wedge T,$$

so that $\omega_{\mathbf{x}}(\tau_j, \tau_{j+1}) = \kappa$ for all $0 \le j < N = N_{\kappa,[0,T],p}(\mathbf{x}) := \sup\{j \ge 0 \mid \tau_j < T\}$ and $\omega_{\mathbf{x}}(\tau_N, \tau_{N+1}) \le \kappa$ (see [4], Definition 4.7, or [13], page 158). Note that $k_p(\mathbf{x}) \le N_{1,[0,T],p}(\mathbf{x}) + 1$.

REMARK 6.4. For any $p, q \ge 1$ and $\mathbf{x} \in \Omega_q$, note that the signature $S(\mathbf{x})_{0,T}$ exists and so $k_p(\mathbf{x})$ is meaningfully defined. Moreover, in case $q \le p$, $\mathbf{x}$ can canonically be viewed as an element of Ω_p by its lift $S_{\lfloor p \rfloor}\mathbf{x} \in \Omega_p$, and we have $S(\mathbf{x})_{0,T} = S(S_{\lfloor p \rfloor}\mathbf{x})_{0,T}$.

However, if $q < \lfloor p \rfloor$, and $N_{1,[0,T],p}(\mathbf{x})$ and the greedy sequence $(\tau_j)_{j=1}^{\infty}$ are defined in terms of $\mathbf{x}$ (not its lift $S_{\lfloor p \rfloor}\mathbf{x}$), then $N_{1,[0,T],p}(\mathbf{x})$ does not yield a deterministic bound on $k_p(\mathbf{x})$ since the individual signatures $S(\mathbf{x})_{\tau_j,\tau_{j+1}}$ will in general fail to be elements of K_p.

To obtain a bound on $k_p(\mathbf{x})$, one needs to consider $N_{1,[0,T],p}(S_{\lfloor p \rfloor}\mathbf{x})$ and $(\tau_j)_{j=1}^{\infty}$ defined in terms of $S_{\lfloor p \rfloor}\mathbf{x} \in \Omega_p$. Then $S(\mathbf{x})_{\tau_j,\tau_{j+1}} = S(S_{\lfloor p \rfloor}\mathbf{x})_{\tau_j,\tau_{j+1}} \in K_p$ for all $j = 0, 1, \dots,$ and so $k_p(\mathbf{x}) \le N_{1,[0,T],p}(S_{\lfloor p \rfloor}\mathbf{x}) + 1$.

Let $\mathcal{K}_p(V)$ be the family of Ω_p-valued random variables $\mathbf{X}$ such that $\mathbb{E}[e^{\lambda k_p(\mathbf{X})}] < \infty$ for all $\lambda > 0$.

COROLLARY 6.5. *Let V be a Banach space, $p \ge 1$ and $\mathbf{X} \in \mathcal{K}_p(V)$. Then* $\mathrm{ESig}[S(\mathbf{X})_{0,T}]$ *has an infinite radius of convergence.*

COROLLARY 6.6. *Let $p \ge 1$ and $\mathbf{X} \in \mathcal{K}_p(\mathbb{R}^d)$ such that $\mathbf{X}$ is $G\Omega_p(\mathbb{R}^d)$-valued. Then $S(\mathbf{X})_{0,T}$ is the unique $G(\mathbb{R}^d)$-valued random variable whose expected signature is* $\mathrm{ESig}[S(\mathbf{X})_{0,T}]$.

We now demonstrate two important examples of $G\Omega_p(\mathbb{R}^d)$-valued random variables in $\mathcal{K}_p(\mathbb{R}^d)$. Remark that a non-negative random variable Z satisfies $\mathbb{E}[e^{\lambda Z}] < \infty$ for all $\lambda > 0$ whenever Z^θ has a Gaussian tail for some $\theta > 1/2$, that is, $\mathbb{P}[Z^\theta > z] \le C^{-1}e^{-Cz^2}$ for all $z > 0$ and a constant $C > 0$. In both of the following examples, $[0, T]$ is a fixed time interval.

EXAMPLE 6.7 (Gaussian rough paths). Recall that every centred continuous Gaussian process in $\mathbb{R}^d$ with independent components and covariance matrix of finite 2D ρ-variation, $\rho \in [1, 2)$, admits a natural lift to a $G\Omega_p(\mathbb{R}^d)$-valued random variable $\mathbf{X}$ for any $p > 2\rho$ ([15] Theorem 15.33).

Recall that [4] Theorem 6.3 provides conditions which lead to bounds on the tail of $N_{1,[0,T],p}(\mathbf{X})$. In particular, due to [12] Theorem 1, the conditions of [4] Theorem 6.3 are verified whenever the covariance matrix of $\mathbf{X}$ possesses finite mixed $(1, \rho)$-variation for some $\rho < 2$, from which it follows that $N_{1,[0,T],p}(\mathbf{X})^\theta$ has a Gaussian tail, where $\theta = (2\rho)^{-1} + 1/2$. We refer to [12] Section 2.3 for a collection of examples satisfying the finite mixed $(1, \rho)$-variation condition, which

in particular includes the natural lift of fractional Brownian motion with Hurst parameter $H > 1/4$ and $\rho = (2H)^{-1}$ ([12] Example 2.8).

Since $k_p(\mathbf{X}) \leq N_{1,[0,T],p}(\mathbf{X}) + 1$ as remarked before, it follows that whenever the finite mixed $(1, \rho)$-variation condition is satisfied, $\mathbf{X}$ is a $G\Omega_p(\mathbb{R}^d)$-valued random variable in $\mathcal{K}_p(\mathbb{R}^d)$ for any $p > 2\rho$.

EXAMPLE 6.8 (Markovian rough paths). Consider $V = \mathbb{R}^d$, $n \geq 1$, and $\mathfrak{g} = \mathfrak{g}^n(\mathbb{R}^d)$, which for convenience we identify with the Lie group $G^n(\mathbb{R}^d)$ via the exponential map. Let $\mathbf{X} = \mathbf{X}^{a,x}$ be a Markovian rough path constructed from a Dirichlet form $\mathcal{E}^a$ on $L^2(\mathfrak{g})$ for $a \in \Xi^{n,d}(\Lambda)$, $\Lambda \geq 1$, and starting point $\mathbf{X}_0^{a,x} = x \in \mathfrak{g}$ (taking the natural lift when $n = 1$; see [15] Chapter 16 for definitions). The sample paths of this process are almost surely geometric p-rough paths for any $p > 2$.

A recent result of Cass and Ogrodnik ([5] Theorem 5.3) implies that $N_{1,[0,T],p}(\mathbf{X})^{1-1/p}$ has a Gaussian tail for any $p > 2$ (moreover the constant determining the tail bounds depends only on Λ, p, n, d and T). It follows that $\mathbf{X}$ is a $G\Omega_p(\mathbb{R}^d)$-valued random variable in $\mathcal{K}_p(\mathbb{R}^d)$ for all $p > 2$.

6.2. *Analyticity.* In this section, we investigate conditions under which the characteristic function is analytic. We apply these results to situations where the expected signature does not necessarily have an infinite radius of convergence.

DEFINITION 6.9. Let X be an E-valued random variable, H a finite dimensional complex Hilbert space, and $M \in \mathbf{L}(V, \mathbf{L}(H))$. For $\lambda \in \mathbb{C}$, define $\phi_{X,M}(\lambda) = \mathbb{E}[(\lambda M)(X)]$ whenever $\|(\lambda M)(X)\|$ is integrable.

The above definition of $\phi_{X,M}$ does not introduce any new concept to the previously defined ϕ_X and simply makes the results in this section easier to state.

Recall that for a real random variable X, if $\mathbb{E}[|e^{\lambda X}|] < \infty$ for all $\lambda \in (-\varepsilon, \varepsilon)$, then $\phi_X(\lambda) = \mathbb{E}[e^{i\lambda X}]$ is well defined and analytic on the strip $|\operatorname{Im}(z)| < \varepsilon$. This property is known as the propagation of regularity (and similar results hold for C^{2k} regularity of ϕ_X on $\mathbb{R}$; see, e.g., [25]).

We start by showing that the analogue of this property is not in general true for G-valued random variables whenever $\dim(V) \geq 2$. The propagation of regularity for real [or equivalently $G(\mathbb{R})$-valued] random variables relies crucially on commutativity in $E(\mathbb{R})$, and we show how the lack of commutativity prevents the same phenomenon from occurring when $\dim(V) \geq 2$. Recall the radius of convergence $r_1(X)$ from Definition 3.1.

EXAMPLE 6.10. Let V be a normed space with $\dim(V) \geq 2$. We construct a G-valued random variable X such that:

(1) $r_1(X) > 0$, thus in particular, $\operatorname{ESig}(X)$ exists and $\phi_{X,M}$ is analytic in a neighbourhood of zero for all $M \in \mathcal{A}$,

(2) there exists $M \in \mathbf{L}(V, \mathfrak{u}(\mathbb{C}^2))$ such that the set of $\lambda \in \mathbb{C}$ for which $|\lambda| > 1$ and $\mathbb{E}[\|(\lambda M)(X)\|] = \infty$ forms a dense subset of $\{z \in \mathbb{C} \mid |z| > 1\}$, and

(3) $\phi_{X,M}$ is nowhere differentiable on $(1, \infty)$.

Let e_1, e_2 be fixed linearly independent vectors in V of unit length, s a nonnegative real random variable, and N a random variable in $\mathbb{N} = \{0, 1, 2, \dots\}$ independent of s. Define the G-valued random variable $X = \exp(sf_N)$, where $f_N = [e_1, [\dots, [e_1, e_2], \dots]$ with e_1 appearing N times.

Suppose there exists $r_0 > 0$ such that that $\mathbb{E}[e^{rs}] < \infty$ for all $0 \le r < r_0$. We claim this implies (1). Indeed, remark that $\|f_n\| \le 2^n$, and thus $\|\delta_r X\| \le \exp(2^N r^N s)$. Denote $p_n = \mathbb{P}[N = n]$. It follows that for $r > 0$ sufficiently small

$$\mathbb{E}\big[\|\delta_r X\|\big] \le \sum_{n \ge 0} p_n \mathbb{E}\big[\exp\big(2^n r^n s\big)\big] < \infty,$$

which implies $r_1(X) > 0$ as claimed.

Let $\mathfrak{su}(2)$ be the special unitary Lie algebra of dimension 3 with the standard basis u_1, u_2, u_3 satisfying $[u_1, u_2] = u_3$, $[u_2, u_3] = u_1$, $[u_3, u_1] = u_2$. Let $M : V \mapsto \mathfrak{su}(2)$ defined by $e_i \mapsto u_i$ for $i = 1, 2$ and arbitrary otherwise.

Suppose moreover that $\mathbb{E}[e^{rs}] = \infty$ for all $r > r_0$ and that N has unbounded support. We claim this implies (2). Indeed, let $v_n = M(f_n)$ (thus $v_0 = u_2, v_1 = u_3, v_2 = -u_2, v_3 = -u_3, v_4 = v_0, \dots$). Denote $\lambda = re^{i\theta}$ for $r, \theta \ge 0$, so that $(\lambda M)(X) = \exp(sr^N e^{iN\theta} v_N)$. We obtain

$$\begin{aligned}\mathbb{E}\big[\|(\lambda M)(X)\|\big] &= \sum_{n \ge 1} p_n \mathbb{E}\big[\|\exp\big(sr^n e^{in\theta} v_n\big)\|\big] \\ &= \sum_{n \ge 1} p_n \mathbb{E}\Big[\exp\Big(\frac{1}{2}\big|sr^n \sin(n\theta)\big|\Big)\Big],\end{aligned}$$

where the last equality follows since iv_n is Hermitian with eigenvalues $\pm\frac{1}{2}$.

Let D be any open subset of $\{z \in \mathbb{C} \mid |z| > 1\}$. We observe that there exists $n \ge 1$ sufficiently large and $r > 1, \theta > 0$, such that $p_n > 0, re^{i\theta} \in D$ and $\mathbb{E}[\exp(\frac{1}{2}|sr^n \sin(n\theta)|)] = \infty$. Thus, (2) holds as claimed.

Finally, we make specific choices for s and N to obtain (3). Observe by Fubini's theorem that for all $r \ge 0$

$$\mathbb{E}\big[(rM)(X)\big] = \sum_{n \ge 1} p_n \mathbb{E}\big[\exp\big(sr^n v_n\big)\big].$$

Suppose $p_n > 0$ only if $n = 4m$ for some integer m. Since

$$\exp(tu_3) = \begin{pmatrix} e^{it/2} & 0 \\ 0 & e^{-it/2} \end{pmatrix},$$

it follows that

$$\mathbb{E}\big[(rM)(X)\big] = \sum_{n \ge 0} p_{4n} \begin{pmatrix} \exp\big(ir^{4n} s/2\big) & 0 \\ 0 & \exp\big(-ir^{4n} s/2\big) \end{pmatrix}.$$

It follows that $\phi_{X,M}$ has the same regularity at $r > 0$ as $\sum_{n\geq 0} p_{4n}\phi_s(r^{4n}/2)$, where ϕ_s is the characteristic function of s.

It is now easy to find ϕ_s and p_n such that the above series defines a nowhere differentiable function on $(1, \infty)$. For example, let $\phi_s(\lambda) = (1-q)(1-qe^{i\lambda})^{-1}$ for any $0 < q < 1$, that is, s is geometrically distributed with parameter $1-q$, and let p_n decay faster than any geometric sequence, that is, for any $\alpha \in (0, 1)$ there exists n_α such that $p_n < \alpha^n$ for all $n \geq n_\alpha$. The statement of (3) then follows by Dini's general construction of a nowhere differentiable function ([23], page 24).

REMARK 6.11. The random variable X constructed above is the exponential of a Lie polynomial of degree N. Thus, when $V = \mathbb{R}^d$, X is the signature of a random weakly geometric N-rough path ([15], Exercise 9.17), and thus of a random geometric p-rough path for $p > N$. Moreover, as the decay of $\|X^k\|$ is exactly of the order $(k/N)!^{-1}$, there does not exist a fixed $p \geq 1$ such that X is almost surely the signature of a random geometric p-rough path.

One can however approximate each sample of X by the signature $S(\mathbf{X})_{0,T}$ of a bounded variation path $\mathbf{X}_{0,\cdot} : [0, T] \mapsto \mathbb{R}^d$ in such as way that (1) and (2) in Example 6.10 still hold for the $G(\mathbb{R}^d)$-valued random variable $S(\mathbf{X})_{0,T}$ with the change that the stated λ in (2) will be dense in the annulus $\{z \in \mathbb{C} \mid 1 < |z| < R\}$ for any fixed $R > 1$ (where the random variable $\mathbf{X}$ depends on R).

DEFINITION 6.12. Let V be a normed space. Denote by $\Phi(V)$ the set of G-valued random variables X which satisfy:

(P1) $r_1(X) > 0$, and
(P2) $\phi_{X,M}$ is (weakly) analytic on $\mathbb{R}$ for all $M \in \mathcal{A}$.

The importance of the set Φ is that when $V = \mathbb{R}^d$ and $X, Y \in \Phi(\mathbb{R}^d)$ such that $\mathrm{ESig}(X) = \mathrm{ESig}(Y)$, we have $X \overset{\mathcal{D}}{=} Y$. To observe this, remark that for V normed and X an E-valued random variable with $r_1(X) =: \varepsilon > 0$, it follows by dominated convergence that $\mathbb{E}[M(X)] = \sum_{k\geq 0} M^{\otimes k}\mathbb{E}[X^k]$ whenever $\|M\| < \varepsilon$. Hence for all $M \in \mathcal{A}$, $\phi_{X,M}(\lambda)$ is completely determined by $\mathrm{ESig}(X)$ whenever $|\lambda|$ is sufficiently small, and the claim follows by Corollary 4.12.

Theorem 6.13 is the main result of this section and provides a criterion to ensure that $X \in \Phi$.

THEOREM 6.13. *Let V be a normed space and X a U-valued random variable. Suppose there exists a bounded, measurable set $B \subset U$ such that $B(X)$ has an exponential tail, that is, $\mathbb{E}[e^{\lambda B(X)}] < \infty$ for some $\lambda > 0$. Then* (P1) *and* (P2) *hold for X.*

PROOF OF THEOREM 6.13, (P1). This follows immediately from Theorem 6.3. □

For the proof of (P2), we require the following two lemmas.

LEMMA 6.14. *Let F be a topological algebra, A a normed algebra, $M \in \mathrm{Hom}(F, A)$, $B \subset F$ a bounded measurable set, and let $c = \sup_{x \in B} \|M(x)\|$. Let X be an F-valued random variable such that $\|M(X)\|$ and $(c+\varepsilon)^{B(X)}$ are integrable for some $\varepsilon > 0$.*

Then $\sup_{M' \in \mathcal{U}} \|M'(X)\|$ is an integrable random variable, where

$$\mathcal{U} = \Big\{ M' \in \mathrm{Hom}(F, A) \mid \sup_{x \in B} \big\| M(x) - M'(x) \big\| < \varepsilon \Big\}$$

is an open subset of $\mathrm{Hom}(F, A)$ (*under the strong topology*).

PROOF. By definition of the strong topology, $\sup_{x \in B} \| \cdot (x) \|$ is a semi-norm on $\mathbf{L}(F, A)$ and so $\mathcal{U}$ is indeed an open subset.

Moreover, $x \mapsto \sup_{M' \in \mathcal{U}} \|M'(x)\|$ is the supremum of a family of continuous functions, thus lower semi-continuous, and thus measurable. The claim now follows by a direct application of (2.2). $\square$

For a bounded complex domain $D \subset \mathbb{C}$, denote by $H_b(\overline{D})$ the space of continuous functions on $\overline{D}$ which are analytic on D. Recall that $H_b(\overline{D})$ equipped with the uniform norm is a separable Banach space.

LEMMA 6.15. *Let V be a normed space, A a separable Banach algebra and $M \in \mathbf{L}(V, A)$. Let X be an E-valued random variable. Assume that for a bounded domain $D \subset \mathbb{C}$, $\sup_{\lambda \in \overline{D}} \|(\lambda M)(X)\|$ is an integrable random variable. Then for every $f \in A'$, the map $\lambda \mapsto \mathbb{E}[\langle f, (\lambda M)(X) \rangle]$ is in $H_b(\overline{D})$.*

PROOF. Let $f \in A'$. For $x \in E$ consider the map $\phi_{M,f}(x) : \lambda \mapsto \langle f, (\lambda M)(x) \rangle$, which is an entire function on $\mathbb{C}$.

We claim that the corresponding linear map $\phi_{M,f} : E \mapsto H(\mathbb{C})$, where $H(\mathbb{C})$ is the space of entire functions on $\mathbb{C}$, is bounded when we equip $H(\mathbb{C})$ with the compact-open topology. Indeed, since $\lambda \mapsto (\lambda M)$ is a continuous map from $\mathbb{C}$ into $\mathrm{Hom}(E, A)$ by Proposition 2.10, the collection of maps $(\lambda M)_{\lambda \in K}$ is strongly bounded in $\mathrm{Hom}(E, A)$ for any bounded set $K \subset \mathbb{C}$. Thus for every bounded set $L \subset E$, it holds that

$$\sup_{x \in L} \sup_{\lambda \in K} \big\| (\lambda M)(x) \big\| < \infty.$$

In particular, this implies that $\phi_{M,f}(L)$ is a bounded subset of $H(\mathbb{C})$ for every bounded set $L \subset E$ as claimed.

Since E is a Fréchet space (hence bornological), it follows moreover that $\phi_{M,f} : E \mapsto H(\mathbb{C})$ is continuous. Hence $\phi_{M,f}(X) |_{\overline{D}}$ is a norm-integrable $H_b(\overline{D})$-valued random variable and thus possesses a barycenter $h \in H_b(\overline{D})$.

Let $\lambda \in \overline{D}$. Since the evaluation map $\langle \cdot, \lambda \rangle : x \mapsto x(\lambda)$ is in the continuous dual of $H_b(\overline{D})$, it follows that

$$h(\lambda) = \mathbb{E}\big[\langle \phi_{M,f}(X), \lambda \rangle\big] = \mathbb{E}\big[\langle f, (\lambda M)(X) \rangle\big] = \langle f, \mathbb{E}\big[(\lambda M)(X)\big] \rangle,$$

where the last equality follows since $(\lambda M)(X)$ is a norm-integrable A-valued random variable and is thus weakly integrable by the separability of A. As h is in $H_b(\overline{D})$, the conclusion follows. $\square$

PROOF OF THEOREM 6.13, (P2). Let $M \in \mathcal{A}$. Since $\|M(g)\| = 1$ for all $g \in U$, one obtains from Proposition 2.10 and Lemma 6.14 that there exists a domain D containing $1 \in \mathbb{C}$ such that $\sup_{\lambda \in \overline{D}} \|(\lambda M)(X)\|$ is an integrable random variable. The conclusion now follows by applying Lemma 6.15. $\square$

Following the discussion at the end of Section 6.1, define

$$N_p = K_p \cap U = \Big\{ g \in U \mid \sup_{k \geq 0} \beta_p(k/p)! \|g^k\| \leq 1 \Big\}.$$

We observe that $S(\mathbf{x})_{0,T} \in N_p$ for every $\mathbf{x} \in G\Omega_p$ with $\|\mathbf{x}\|_{p\text{-var};[0,T]} \leq 1$. As with K_p, N_p is bounded and measurable in E.

For $\mathbf{x} \in G\Omega_p$, define $n_p(\mathbf{x}) = N_p(S(\mathbf{x})_{0,T})$, that is, the minimum positive integer n for which there exist $g_1, \dots, g_n \in N_p$ such that $S(\mathbf{x})_{0,T} = g_1 \cdots g_n$. Recall the functions k_p and $N_{1,[0,T],p}$ from Section 6.1 and note that $k_p(\mathbf{x}) \leq n_p(\mathbf{x}) \leq N_{1,[0,T],p}(\mathbf{x}) + 1$.

REMARK 6.16. As in Remark 6.4, we mention again that for $1 \leq q \leq p$, every $\mathbf{x} \in G\Omega_q$ is canonically defined as an element of $G\Omega_p$ via its lift $S_{\lfloor p \rfloor}\mathbf{x} \in G\Omega_p$. However, one cannot bound $n_p(\mathbf{x})$ in terms $N_{1,[0,T],p}(\mathbf{x})$ computed directly in terms of $\mathbf{x}$; instead one has $n_p(\mathbf{x}) \leq N_{1,[0,T],p}(S_{\lfloor p \rfloor}\mathbf{x}) + 1$.

Let $\mathcal{N}_p(V)$ be the family of $G\Omega_p$-valued random variables $\mathbf{X}$ such that $n_p(\mathbf{X})$ has an exponential tail. Note that if $\mathbf{X} \in \mathcal{K}_p$ and is $G\Omega_p$-valued, then $\mathbf{X} \in \mathcal{N}_p$.

COROLLARY 6.17. *Let V be a Banach space. Then for all $p \geq 1$ and $\mathbf{X} \in \mathcal{N}_p$, the signature $S(\mathbf{X})_{0,T}$ is a U-valued random variable satisfying* (P1) *and* (P2).

In the finite dimensional setting, we obtain a result analogous to Corollary 6.6 but with weaker assumptions and a weaker conclusion.

COROLLARY 6.18. *Let $p \geq 1$ and $\mathbf{X} \in \mathcal{N}_p(\mathbb{R}^d)$. Then $S(\mathbf{X})_{0,T} \in \Phi(\mathbb{R}^d)$. In particular, $S(\mathbf{X})_{0,T}$ is the unique $G(\mathbb{R}^d)$-valued random variable in $\Phi(\mathbb{R}^d)$ whose expected signature is* $\mathrm{ESig}[S(\mathbf{X})_{0,T}]$.

REMARK 6.19. For a random variable $X \in \Phi(\mathbb{R}^d)$, we cannot exclude the possibility that there exists a $G(\mathbb{R}^d)$-valued random variable Y (which might arise as the signature of a geometric rough path) such that $Y \notin \Phi(\mathbb{R}^d)$ and $\mathrm{ESig}(X) = \mathrm{ESig}(Y)$. Whether this is possible currently remains unknown.

However, note that Corollaries 6.17 and 6.18 apply for all $p \geq 1$. Thus for any $p, q \geq 1$ and random geometric rough paths $\mathbf{X} \in \mathcal{N}_p(\mathbb{R}^d)$ and $\mathbf{Y} \in G\Omega_q(\mathbb{R}^d)$, if $S(\mathbf{X})_{0,T}$ and $S(\mathbf{Y})_{0,T}$ are not equal in law [as $G(\mathbb{R}^d)$-valued random variables] and $\mathrm{ESig}(S(\mathbf{X})_{0,T}) = \mathrm{ESig}(S(\mathbf{Y})_{0,T})$, then the lift $S_{\lfloor q' \rfloor}\mathbf{Y}$ cannot be in $\mathcal{N}_{q'}(\mathbb{R}^d)$ for any $q' \geq q$.

EXAMPLE 6.20 (Markovian rough paths stopped upon exiting a domain). Recall the notation of Example 6.8 and the result of Cass and Ogrodnik [5] that $N_{1,[0,1],p}(\mathbf{X}^{a,x})^{1-1/p}$ has a Gaussian tail for any $p > 2$.

In this example, we shall replace the interval $[0, 1]$ by $[0, T]$, where T is the first exit time of $\mathbf{X}^{a,x}$ from a suitable set. In particular, we shall show that $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$ has an exponential tail and that this result is asymptotically sharp.

Throughout the example, we fix $\Lambda \geq 1$ and $\mathfrak{g} = \mathfrak{g}^n(\mathbb{R}^d)$. We first give a slight extension of the support theorem [15] Theorem 16.33 in the Hölder topology. Recall the Sobolev path space $W_x^{1,2}([s, t], \mathfrak{g})$ with starting point $x \in \mathfrak{g}$. In particular, recall that for all $\mathbf{h} \in W_x^{1,2}([s, t], \mathfrak{g})$ and $\alpha \in [0, 1/2]$

$$\|\mathbf{h}\|_{\alpha\text{-Höl};[s,t]} \leq (t - s)^{1/2-\alpha} \|\mathbf{h}\|_{W^{1,2};[s,t]}.$$

For $\theta > 0$ consider the ball

$$W_{\theta;x} := \{\mathbf{h} \in W_x^{1,2}([0, 1], \mathfrak{g}) \mid \|\mathbf{h}\|_{W^{1,2};[0,1]} < \theta\}.$$

LEMMA 6.21. *For any $\alpha \in [0, 1/4)$, $\theta > 0$ and $c > 0$, there exists $\delta > 0$ such that*

$$\mathbb{P}^{a,x}\big[d_{\alpha\text{-Höl};[0,1]}(\mathbf{X}, \mathbf{h}) < c\big] > \delta$$

for all $a \in \Xi^{n,d}(\Lambda)$, starting points $x \in \mathfrak{g}$, and $\mathbf{h} \in W_{\theta;x}$.

The proof is essentially the same as that in [15] and we defer it to the end of the example. Recall now the greedy sequence $(\tau_j)_{j=1}^{\infty}$ associated with $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$. For ease of notation, we shall not stop τ_j at T for $j > N := N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$ (i.e., we do not necessarily have $\tau_{N+1} = T$). Note this causes no confusion since $\mathbf{X}_t^{a,x}$ is defined for all times $t \geq 0$ as a diffusion in $\mathfrak{g}$.

Consider first $\mathbf{X}^{a,x} : [0, 1] \mapsto \mathfrak{g}$. Taking $\mathbf{h} \equiv x$ the trivial path, Lemma 6.21 implies that for any $p > 4$ and $\kappa > 0$, there exists $\delta > 0$ such that

$$\inf_{x\in\mathfrak{g}} \mathbb{P}^{a,x}\big[\|\mathbf{X}\|_{1/p\text{-Höl};[0,1]} < \kappa\big] \geq \delta.$$

It follows that

$$\inf_{x\in\mathfrak{g}} \mathbb{P}^{a,x}[\tau_1 > 1] \geq \inf_{x\in\mathfrak{g}} \mathbb{P}^{a,x}\big[\|\mathbf{X}\|_{1/p\text{-Höl};[0,1]} < \kappa\big] \geq \delta,$$

so by the (strong) Markov property of $\mathbf{X}^{a,x}$ and properties of conditional expectation

$$\mathbb{P}^{a,x}\big[N_{\kappa,[0,1],p}(\mathbf{X}) \geq k\big] = \mathbb{P}^{a,x}[\tau_k < 1] \leq (1 - \delta)^k.$$

That is, $N_{\kappa,[0,1],p}(\mathbf{X}^{a,x})$ has an exponential tail [moreover δ does not depend on $x \in \mathfrak{g}$ or $a \in \Xi^{n,d}(\Lambda)$].

While this argument yields a strictly weaker asymptotic bound than that in [5], the advantage is that by choosing appropriate $\mathbf{h}$ in Lemma 6.21, a very similar argument gives upper and lower bounds on the tail of $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$, where T is now the first exit time of $\mathbf{X}^{a,x}$ from a suitable open set. We first show the lower bound.

Recall that $\mathfrak{g}$ is equipped with the (left-invariant) metric d induced by the Carnot–Carathéodory norm (or any other symmetric sub-additive homogeneous norm). For any $r > 0$ and $x \in \mathfrak{g}$, define $B_r(x) = \{y \in \mathfrak{g} \mid d(x, y) \le r\}$.

PROPOSITION 6.22. *Let $p > 4$, $\kappa, r > 0$. Define $T = \inf\{t > 0 \mid \mathbf{X}_t^{a,x} \notin B_r(x)\}$ the first exit time of $\mathbf{X}^{a,x}$ from $B_r(x)$. Then there exists $\delta > 0$ such that*

$$\mathbb{P}^{a,x}\big[N_{\kappa,[0,T],p}(\mathbf{X}) \ge k\big] \ge \delta^k$$

for all $a \in \Xi^{n,d}(\Lambda)$ and $x \in \mathfrak{g}$.

PROOF. Let $\theta > 0$ sufficiently large such for all $x \in \mathfrak{g}$ and $y \in B_{r/2}(x)$ there exists $\mathbf{h}^y \in W_{\theta;y}$ such that $\mathbf{h}_1^y = x$, $\mathbf{h}_t^y \in B_{r/2}(x)$ for all $t \in [0, 1]$, and $\|\mathbf{h}^y\|_{p\text{-var};[0,1]} > \kappa + r/2$ (e.g., take $t \mapsto \mathbf{h}_t^y$ as a geodesic from y to x on $[0, 1/2]$ and then $\mathbf{h}_t^y = x\mathbf{h}_{2t-1}$ for $t \in [1/2, 1]$ for a fixed $\mathbf{h} \in W_0^{1,2}([0, 1], \mathfrak{g})$ with $\mathbf{h}_1 = 0$, $\mathbf{h}_t \in B_{r/2}(0)$ for all $t \in [0, 1]$, and $\|\mathbf{h}\|_{p\text{-var};[0,1]} > \kappa + r/2$).

Since $\|\mathbf{X}\|_{p\text{-var};[0,1]} \ge \|\mathbf{h}^y\|_{p\text{-var};[0,1]} - d_{1/p\text{-Höl};[0,1]}(\mathbf{X}, \mathbf{h}^y)$, we have for all $x \in \mathfrak{g}$, $y \in B_{r/2}(x)$ and $a \in \Xi^{n,d}(\Lambda)$

$$\begin{aligned}\mathbb{P}^{a,y}\big[&\mathbf{X}_t \in B_r(x) \text{ for all } t \in [0, 1], \|\mathbf{X}\|_{p\text{-var};[0,1]} > \kappa, \mathbf{X}_1 \in B_{r/2}(x)\big] \\ &\ge \mathbb{P}^{a,y}\big[d_{1/p\text{-Höl}}(\mathbf{X}, \mathbf{h}^y) < r/2\big].\end{aligned}$$

Applying Lemma 6.21 with $c = r/2$ and $\alpha = 1/p$, along with the (weak) Markov property and conditional expectation, concludes the proof. $\square$

REMARK 6.23. Note that Proposition 6.22 deals only with the quantity $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$ and does not provide a lower bound on the tail of $n_p(\mathbf{X}_{[0,T]}^{a,x})$. In particular, one cannot conclude that $\mathrm{ESig}[S(\mathbf{X}^{a,x})_{0,T}]$ does not have an infinite radius of convergence.

We now show an upper bound on the tail of $N_{\kappa,[0,T],p}(\mathbf{X}^{a,x})$ which will imply that $S(\mathbf{X}^{a,x})_{0,T} \in \Phi(\mathbb{R}^d)$ (see, however, Remark 6.26). For a subset $D \subset \mathfrak{g}$, consider the following property:

(6.3) There exist $r, c > 0$ such that $\sup_{h \in B_r(0)} \inf_{y \in D} d(xh, y) > c$ for all $x \in D$.

REMARK 6.24. For $1 \le k \le n$, let $\pi^k : \mathfrak{g}^n(\mathbb{R}^d) \mapsto \mathfrak{g}^k(\mathbb{R}^d)$ denote the projection. Then whenever the image $\pi^k(D)$ satisfies (6.3) for some $1 \le k \le n$ [for the respective metric on $\mathfrak{g}^k(\mathbb{R}^d)$], then so does D (with a different choice of r, c).

Indeed, on the one hand $d(x, y) \ge d(\pi^k(x), \pi^k(y))$ for all $x, y \in \mathfrak{g}^n(\mathbb{R}^d)$. On the other hand, for every $r > 0$, there exists $R > 0$ such that $B_r(0) \subset \pi^k(B_R(0)) \subset \mathfrak{g}^k(\mathbb{R}^d)$. The conclusion readily follows since π^k is a group homomorphism.

PROPOSITION 6.25. *Let $p > 4$, $\kappa > 0$, and $D \subset \mathfrak{g}$ be an open set satisfying (6.3) for some $r, c > 0$. Define $T = \inf\{t > 0 \mid \mathbf{X}_t^{a,x} \notin D\}$ the first exit time of $\mathbf{X}^{a,x}$ from D. Then there exists $\delta > 0$ such that*

$$\mathbb{P}^{a,x}\big[N_{\kappa,[0,T],p}(\mathbf{X}) \ge k\big] \le (1-\delta)^k$$

for all $a \in \Xi^{n,d}(\Lambda)$ and $x \in D$.

PROOF. Let $\theta > 0$ be sufficiently large such that for every $h \in B_r(0)$ there exists $\mathbf{h} \in W_{\theta;0}$ such that $\mathbf{h}_1 = h$. Note that it suffices to prove the statement for any fixed $\kappa > 0$. In particular, we may assume that $\kappa > \theta + c$.

It follows that to every point $x \in D$, we can assign $h^x \in B_r(x)$ and $\mathbf{h}^x \in W_{\theta;x}$ such that $\inf_{y \in D} d(h^x, y) > c$ and $\mathbf{h}_1^x = h^x$. Then for all $a \in \Xi^{n,d}(\Lambda)$ and $x \in D$

$$\begin{aligned}
\mathbb{P}^{a,x}[\tau_1 > 1 \ge T] &\ge \mathbb{P}^{a,x}\big[\|\mathbf{X}\|_{p\text{-var};[0,1]} < \theta + c, \mathbf{X}_1 \notin D\big] \\
&\ge \mathbb{P}^{a,x}\big[\|\mathbf{X}\|_{1/p\text{-Höl};[0,1]} < \theta + c, d\big(\mathbf{X}_1, h^x\big) < c\big] \\
&\ge \mathbb{P}^{a,x}\big[d_{1/p\text{-Höl};[0,1]}\big(\mathbf{X}, \mathbf{h}^x\big) < c\big].
\end{aligned}$$

Applying Lemma 6.21 with $\alpha = 1/p$, along with the (strong) Markov property and conditional expectation, concludes the proof. $\square$

REMARK 6.26. The diffusion $\mathbf{X}^{a,x}$ is constructed on the space $\mathfrak{g}^n = \mathfrak{g}^n(\mathbb{R}^d)$ [or equivalently on $G^n(\mathbb{R}^d)$], and Proposition 6.25 gives an exponential bound on the tail of $N_{1,[0,T],p}(\mathbf{X}^{a,x})$ computed in terms of $\mathbf{X}^{a,x}$ for any $p > 4$. Fixing $4 < p < 5$, Corollary 6.18 thus implies that $S(\mathbf{X}^{a,x})_{0,T} \in \Phi(\mathbb{R}^d)$ for $n \ge 4$.

One could extend this to the case $n = 2$ or 3 (recall for $n = 1$ we consider the diffusion $\mathbf{X}^{a\circ\pi^1,x}$ on $\mathfrak{g}^2$) if the analogue of Lemma 6.21 were true for all $\alpha \in [0, 1/2)$. However such a support theorem is currently unknown.

Nonetheless, in light of Remarks 6.4 and 6.16, for $n = 2$ or 3 we can still show that $S(\mathbf{X}^{a,x})_{0,T} \in \Phi(\mathbb{R}^d)$ by showing that $N_{1,[0,T],p}(S_4\mathbf{X}^{a,x})$ has an exponential tail.

To show this, note we can apply Proposition 6.25 to the diffusion $\mathbf{X}^{a\circ\pi^n,y}$ on $\mathfrak{g}^4$ and the open set $(\pi^n)^{-1}(D) \subset \mathfrak{g}^4$ [which indeed satisfies (6.3) due to Remark 6.24]. We thus obtain that $N_{1,[0,\widetilde{T}],p}(\mathbf{X}^{a\circ\pi^n,y})$ has an exponential tail, where $\widetilde{T}$ is the first exit time of $\pi^n\mathbf{X}^{a\circ\pi^n,y}$ from D.

To conclude that $N_{1,[0,T],p}(S_4\mathbf{X}^{a,x})$ has an exponential tail, it suffices to show that $\mathbf{Y}_\cdot^y := y * S_4\mathbf{X}_{0,\cdot}^{a,\pi^n y}$ is equal in law to $\mathbf{X}_\cdot^{a\circ\pi^n,y}$ for all $y \in \mathfrak{g}^4$ as processes on

$\mathfrak{g}^4$ ($*$ denoting group multiplication in $\mathfrak{g}^4$). This follows by a similar argument as [11] Section 6: observe that the Markov process $\mathbf{Y}_t^y$ is the solution of an RDE with starting point $y \in \mathfrak{g}^4$ and driven by $\pi^2(\mathbf{X}_t^{a,\pi^n y})$ (which is non-Markov in general) along the (unbounded) canonical left-invariant vector fields $u_1, \dots, u_d$ on $\mathfrak{g}^4$. Denoting by $\mathcal{P}_t$ the semi-group on $C_b(\mathfrak{g}^4)$ of $\mathbf{Y}_t^y$, it suffices to show that

$$\lim_{t\to 0}\langle t^{-1}(f - \mathcal{P}_t f), g\rangle_{L^2(\mathfrak{g}^4)} = \mathcal{E}^{a\circ\pi^n}(f, g)$$

for all $f, g \in C_c^\infty(\mathfrak{g}^4)$.

Consider $f, g \in C_c^\infty(\mathfrak{g}^4)$ with support in a ball $B_R(0) \subset \mathfrak{g}^4$ and fix smooth vector fields u_i^R which agree with u_i on $B_{2R}(0)$ and have compact support. Let $\mathbf{Y}_t^{R,y}$ denote the RDE driven by $\pi_2(\mathbf{X}_t^{a,\pi^n(y)})$ along u_i^R starting at $\mathbf{Y}_0^{R,y} = y$. For all $y \in B_R(0)$ and $t \in [0, 1]$, we have $\mathbf{Y}_t^{R,y} = \mathbf{Y}_t^y$ whenever $\mathbf{Y}_s^y \in B_{2R}(0)$ for all $s \in [0, t]$. The probability that $\mathbf{Y}_s^y$ leaves $B_{2R}(0)$ in $[0, t]$ is bounded above by $C^{-1}\exp(-Ct^{-2/p})$ for any $2 < p < 3$ and some $C = C(R, p)$ (which follows from Fernique estimates on $\|\mathbf{X}^{a,x}\|_{1/p\text{-Höl};[0,1]}$).

Defining $\mathcal{P}_t^R f(y) := \mathbb{E}[f(\mathbf{Y}_t^{R,y})]$, it follows readily that

$$\lim_{t\to 0}\langle t^{-1}(f - \mathcal{P}_t f), g\rangle = \lim_{t\to 0}\langle t^{-1}(f - \mathcal{P}_t^R f), g\rangle.$$

Finally, the latter limit is now seen to equal $\mathcal{E}^{a\circ\pi^n}(f, g)$ following [11] Lemmas 26, 27 and the proof of Proposition 28 (note that one readily extends Lemma 27 to diffusions on $\mathfrak{g}^n$ for $n > 2$, cf. [15] Proposition 16.20).

PROOF OF LEMMA 6.21. We mimic the proofs of [15] Lemma 16.32 and Theorem 16.33 while keeping track of constants.

For $\alpha \in [0, 1/4)$, $\mathbf{h} \in W_x^{1,2}([0, 1], \mathfrak{g})$ and $\varepsilon > 0$ define the set

$$B_{\varepsilon;\alpha}^{\mathbf{h}} = \{\mathbf{x} \in C_x^{\alpha\text{-Höl}}([0, 1], \mathfrak{g}) \mid \|\mathbf{x}\|_{\alpha\text{-Höl}} \le 2\|\mathbf{h}\|_{\alpha\text{-Höl}} + 1, d_\infty(\mathbf{x}, \mathbf{h}) \le \varepsilon\}.$$

We claim that for all $\alpha \in [0, 1/4)$ and $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\mathbb{P}^{a,x}[\mathbf{X} \in B_{\varepsilon;\alpha}^{\mathbf{h}}] > \delta$$

for all $a \in \Xi^{n,d}(\Lambda)$, $x \in \mathfrak{g}$, and $\mathbf{h} \in W_{\theta;x}$.

Indeed, we follow the proof [15] Lemma 16.32 (we also mention here that, directly as stated, [15] Lemma 16.32 contains the minor error that it fails to hold for the trivial path $\|\mathbf{h}\|_{\alpha\text{-Höl}} = 0$; this is readily fixed by modifying the definition of their $B_\varepsilon^{\mathbf{h}}$ to our definition above; moreover the proof of [15] Theorem 16.33 then goes through unchanged).

Using Step 1 of the proof of [15] Lemma 16.32, we obtain that for any $\beta \in (\alpha, 1/2)$, $\mathbb{P}^{a,x}[\mathbf{X} \in B_{\varepsilon;\alpha}^{\mathbf{h}}] \ge \Delta_1 - \Delta_2$, where $\Delta_1 = \mathbb{P}^{a,x}[d_\infty(\mathbf{X}, \mathbf{h}) \le \varepsilon]$ and

$$\Delta_2 = \mathbb{P}^{a,x}\big[\|\mathbf{X}\|_{\beta\text{-Höl}} > \big(\|\mathbf{h}\|_{\alpha\text{-Höl}} + 1\big)^{\beta/\alpha}(2\varepsilon)^{1-\beta/\alpha}\big].$$

The claim will follow once we show that $\Delta_2/\Delta_1 \to 0$ as $\varepsilon \to 0$ uniformly over $a \in \Xi^{n,d}(\Lambda)$, $x \in \mathfrak{g}$, and $\mathbf{h} \in W_{\theta;x}$.

By [15] Theorem E.21, we have $\log(\Delta_1) \geq -c_1\varepsilon^{-2}$ where $c_1 = C(1 + \|\mathbf{h}\|_{W^{1,2}})^2$ and C is a constant depending only on the doubling and Poincaré constants of $\mathcal{E}^a$, which in turn depend only on Λ, n and d ([15] Proposition 16.5 and Theorem E.8).

On the other hand, the Fernique estimate in [15] Corollary 16.12 implies that

$$\log(\Delta_2) \leq -c_2\big(\|\mathbf{h}\|_{\alpha\text{-Höl}} + 1\big)^{2\beta/\alpha}\varepsilon^{2-2\beta/\alpha} \leq -c_2\varepsilon^{2-2\beta/\alpha},$$

where c_2 depends only on β and Λ. So for fixed $\alpha \in [0, 1/4)$, choose any $\beta \in (2\alpha, 1/2)$. Since $2 - 2\beta/\alpha < -2$, we see $\Delta_2/\Delta_1 \to 0$ as $\varepsilon \to 0$ uniformly over the desired variables, which proves the claim.

To conclude, we follow the proof of [15] Theorem 16.33. By the d_0/d_∞ estimate on $\mathfrak{g}$ ([15] Proposition 8.15),

$$d_0(\mathbf{x}, \mathbf{h}) \leq Cd_\infty(\mathbf{x}, \mathbf{h}) + Cd_\infty(\mathbf{x}, \mathbf{h})^{1/n}\big(\|\mathbf{x}\|_\infty + \|\mathbf{h}\|_\infty\big)^{1-1/n},$$

where $C = C(n, d)$, and so by interpolation ([15] Lemma 8.16) we have for all $x \in \mathfrak{g}$, $\mathbf{x}, \mathbf{h} \in C_x^{\alpha\text{-Höl}}([0, 1], \mathfrak{g})$ and $0 \leq \alpha' < \alpha < 1/4$ that

$$\begin{aligned} d_{\alpha'\text{-Höl}}(\mathbf{x}, \mathbf{h}) &\leq \big(\|\mathbf{x}\|_{\alpha\text{-Höl}} + \|\mathbf{h}\|_{\alpha\text{-Höl}}\big)^{\alpha'/\alpha} d_0(\mathbf{x}, \mathbf{h})^{1-\alpha'/\alpha} \\ &\leq C^{1-\alpha'/\alpha}\big(\|\mathbf{x}\|_{\alpha\text{-Höl}} + \|\mathbf{h}\|_{\alpha\text{-Höl}}\big)^{\alpha'/\alpha} \\ &\quad \times \big[d_\infty(\mathbf{x}, \mathbf{h}) + d_\infty(\mathbf{x}, \mathbf{h})^{1/n}\big(\|\mathbf{x}\|_\infty + \|\mathbf{h}\|_\infty\big)^{1-1/n}\big]^{1-\alpha'/\alpha}. \end{aligned}$$

Since $\|\mathbf{h}\|_\infty \leq \|\mathbf{h}\|_{\alpha\text{-Höl}} \leq \|\mathbf{h}\|_{W^{1,2}}$, it follows for all $\mathbf{h} \in W_{\theta;x}$ and $\mathbf{x} \in B^{\mathbf{h}}_{\varepsilon;\alpha}$ that

$$d_{\alpha'\text{-Höl}}(\mathbf{x}, \mathbf{h}) \leq c_3\big(\varepsilon + \varepsilon^{1/n}\big)^{1-\alpha'/\alpha},$$

where c_3 depends only on n, d, α, α' and θ.

Choosing $\varepsilon > 0$ so that $c_3(\varepsilon + \varepsilon^{1/n})^{1-\alpha'/\alpha} < c$, it follows that there exists $\delta > 0$ such that

$$\mathbb{P}^{a,x}\big[d_{\alpha'\text{-Höl}}(\mathbf{X}, \mathbf{h}) < c\big] \geq \mathbb{P}^{a,x}\big[\mathbf{X} \in B^{\mathbf{h}}_{\varepsilon;\alpha} < c\big] > \delta$$

for all $a \in \Xi^{n,d}(\Lambda)$, $x \in \mathfrak{g}$, and $\mathbf{h} \in W_{\theta;x}$, which concludes the proof. $\square$

6.3. *Convergence of measures*. We conclude the paper with a result analogous to the method of moments for weak convergence of $G(\mathbb{R}^d)$-valued random variables. We work first with a slightly general notion of coproduct spaces as this is the only structure of E which we require.

DEFINITION 6.27. A *coproduct space* (F, Δ) is a locally convex space F and a continuous linear map $\Delta : F \mapsto F^{\widehat{\otimes}2}$, with the additional property that $G(F) := \{g \in F \mid \Delta(g) = g \otimes g, g \neq 0\}$ is closed in F. Let $\mathcal{P}_G(F)$ be the set of (weakly) integrable probability measures on $G(F)$.

The extra condition that $G(F)$ is closed in F will only arise in Lemma 6.30 to ensure that $G(F)$ is Polish whenever F is. Remark that (3.1) remains true for any coproduct space (F, Δ), $f \in F'$, and $\mu \in \mathcal{P}(F)$ with support on $G(F)$.

LEMMA 6.28. *Let (F, Δ) be a nuclear coproduct space and γ a semi-norm on F. There exists a semi-norm ξ on F such that $\mu(\gamma) \le \sqrt{\xi(\mu^*)}$ for all $\mu \in \mathcal{P}_G(F)$.*

PROOF. Let ζ be a semi-norm on F such that the canonical map $\widehat{F}_\zeta \mapsto \widehat{F}_\gamma$ is nuclear. Increasing ζ by a scalar multiple if necessary, it follows that there exist $(f_n)_{n\ge1} \in F'$ such that $\sum_{n\ge1} \zeta(f_n) \le 1$ and $\gamma \le \sum_{n\ge1} |f_n|$. The conclusion then follows from (3.1) for any semi-norm ξ on F such that $\xi \ge (\zeta^{\otimes 2}) \circ \Delta$. $\square$

LEMMA 6.29. *Let (F, Δ) be a Fréchet nuclear coproduct space. Let $R \subseteq \mathcal{P}_G(F)$ be a family of probability measures on $G(F)$ such that $(\mu^*)_{\mu\in R}$ is bounded. Then R is uniformly tight.*

PROOF. Let $(\gamma_n)_{n\ge1}$ be a defining non-decreasing sequence of semi-norms on F. By Lemma 6.28 there exists a sequence of semi-norms $(\xi_n)_{n\ge1}$ on F such that $\mu(\gamma_n) \le \sqrt{\xi_n(\mu^*)}$ for all $\mu \in \mathcal{P}_G(F)$. Since $(\mu^*)_{\mu\in R}$ is bounded, $\sup_{\mu\in R} \xi_n(\mu^*) < \infty$ for every $n \ge 1$.

Let $B_n = \{x \in F \mid \gamma_n(x) < 1\}$. For any sequence of positive reals $(\lambda_n)_{n\ge1}$, the set $K := \bigcap_{n\ge1} \lambda_n B_n$ is bounded in H and thus relatively compact ([36], page 520). For all $\mu \in \mathcal{P}_G(F)$ we have that

$$\mu(K^c) \le \sum_{n\ge1} \mu(\{x \mid \gamma_n(x) \ge \lambda_n\}) \le \sum_{n\ge1} \lambda_n^{-1} \mu(\gamma_n) \le \sum_{n\ge1} \lambda_n^{-1} \sqrt{\xi_n(\mu^*)}.$$

Taking λ_n sufficiently large, it follows that $\sup_{\mu\in R} \mu(K^c)$ can be made arbitrarily small. $\square$

LEMMA 6.30. *Let (F, Δ) be a Fréchet nuclear coproduct space and let $(\mu_n)_{n\ge1}$ be a sequence of measures in $\mathcal{P}_G(F)$ such that $\mu_n^* \to x$ weakly for some $x \in F$. Then there exists $\mu \in \mathcal{P}_G(F)$ and a subsequence $(n(k))_{k\ge1}$ such that $\mu_{n(k)} \xrightarrow{\mathcal{D}} \mu$ and $x = \mu^*$.*

PROOF. Recall that a Fréchet Montel space (thus in particular a Fréchet nuclear space) is always separable ([35], page 195), and hence Polish. As a closed subset of F, $G(F)$ is also Polish.

The sequence $(\mu_n^*)_{n\ge1}$ is bounded ([34], Theorem 3.18) thus there exists a convergent subsequence $\mu_{n(k)} \to \mu$ for some probability measure μ on $G(F)$ by Lemma 6.29.

Let $f \in F'$. Since $\sup_{n\geq 1} \mu_n(f^2) = \sup_{n\geq 1}(f^{\otimes 2})(\Delta\mu_n^*) < \infty$, the sequence of image measures $(\mu_n f^{-1})_{n\geq 1}$ on $\mathbb{R}$ is uniformly integrable. It follows that f is μ-integrable and $f(\mu_n^*) = \mu_n(f) \to \mu(f)$ ([3] Lemma 8.4.3). Thus, $x = \mu^*$ and $\mu \in \mathcal{P}_G(F)$. $\square$

Recall that E is Fréchet and nuclear whenever V is. The following is now a consequence of Lemma 6.30 and Proposition 6.1.

THEOREM 6.31. *Let $(X_n)_{n\geq 1}$ be a sequence of $G(\mathbb{R}^d)$-valued random variables such that $\mathbb{E}[X_n] \in E(\mathbb{R}^d)$ exists [i.e., $r_2(X_n) = \infty$] for all $n \geq 1$. Suppose that $\mathbb{E}[X_n]$ converges to some $x \in E(\mathbb{R}^d)$ in the weak topology of $E(\mathbb{R}^d)$. Then there exists a unique integrable $G(\mathbb{R}^d)$-valued random variable X such that $X_n \xrightarrow{\mathcal{D}} X$ and $x = \mathbb{E}[X]$.*

REMARK 6.32. We remark that $F := P(\mathbb{R}) = \prod_{k\geq 0}(\mathbb{R})^{\otimes k}$ is also a Fréchet nuclear coproduct space under the product topology. Moreover the exponential map $\exp : \mathbb{R} \mapsto G(\mathbb{R}) = G(F)$ is a homeomorphism. One may then directly apply Lemma 6.30 to obtain a proof of the classical method of moments for real random variables: if μ_n are probability measures on $\mathbb{R}$ with finite moments $(m_n(j))_{j\geq 1}$ such that $\lim_{n\to\infty} m_n(j) = m(j)$ for every $j \geq 1$, then $(m(j))_{j\geq 1}$ are the moments of a probability measure μ on $\mathbb{R}$ for which $\mu_{n(k)} \xrightarrow{\mathcal{D}} \mu$ along a subsequence $(n(k))_{k\geq 1}$ [if μ is moment-determined then in fact $\mu_n \xrightarrow{\mathcal{D}} \mu$].

Acknowledgments. We would like to thank Dr. Ni Hao and Prof. Peter Friz for numerous constructive discussions and suggestions, particularly during the Berlin–Oxford Meetings on Applied S.A. in 2013/2014. We would also like to thank Prof. Thierry Lévy for several helpful conversations, and Dr. Horatio Boedihardjo for valuable comments on an earlier draft. Finally, we would like to thank the anonymous referee for his detailed comments and for pointing out the work of [12].

REFERENCES

[1] ARENS, R. (1946). The space L^ω and convex topological rings. *Bull. Amer. Math. Soc.* **52** 931–935. MR0017524

[2] BOEDIHARDJO, H., GENG, X., LYONS, T. and YANG, D. (2014). The signature of a rough path: Uniqueness. Preprint. Available at arXiv:1406.7871.

[3] BOGACHEV, V. I. (2007). *Measure Theory. Vol. I, II*. Springer, Berlin. MR2267655

[4] CASS, T., LITTERER, C. and LYONS, T. (2013). Integrability and tail estimates for Gaussian rough differential equations. *Ann. Probab.* **41** 3026–3050. MR3112937

[5] CASS, T. and OGRODNIK, M. (2015). Tail estimates for Markovian rough paths. Preprint. Available at arXiv:1411.5189.

[6] CHEN, K.-T. (1957). Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula. *Ann. of Math.* (2) **65** 163–178. MR0085251

[7] CHEN, K.-T. (1958). Integration of paths—a faithful representation of paths by non-commutative formal power series. *Trans. Amer. Math. Soc.* **89** 395–407. MR0106258

[8] CUNTZ, J. (1997). Excision in periodic cyclic theory for topological algebras. In *Cyclic Cohomology and Noncommutative Geometry* (*Waterloo*, *ON*, 1995). *Fields Inst. Commun.* **17** 43–53. Amer. Math. Soc., Providence, RI. MR1478700

[9] DINEEN, S. (1981). *Complex Analysis in Locally Convex Spaces*. *North-Holland Mathematics Studies* **57**. North-Holland, Amsterdam. MR0640093

[10] FAWCETT, T. (2003). Problems in stochastic analysis: Connections between rough paths and non-commutative harmonic analysis. Ph.D. thesis, Univ. Oxford.

[11] FRIZ, P. and VICTOIR, N. (2008). On uniformly subelliptic operators and stochastic area. *Probab. Theory Related Fields* **142** 475–523. MR2438699

[12] FRIZ, P. K., GESS, B., GULISASHVILI, A. and RIEDEL, S. (2016). The Jain–Monrad criterion for rough paths and applications to random Fourier series and non-Markovian Hörmander theory. *Ann. Probab.* **44** 684–738.

[13] FRIZ, P. K. and HAIRER, M. (2014). *A Course on Rough Paths*: *With an Introduction to Regularity Structures*. Springer, Cham. MR3289027

[14] FRIZ, P. K. and SHEKHAR, A. (2014). General rough integration, Lévy Rough paths and a Lévy–Kintchine type formula. Preprint. Available at arXiv:1212.5888.

[15] FRIZ, P. K. and VICTOIR, N. B. (2010). *Multidimensional Stochastic Processes as Rough Paths*: *Theory and Applications*. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge Univ. Press, Cambridge. MR2604669

[16] GIAMBRUNO, A. and VALENTI, A. (1995). On minimal $*$-identities of matrices. *Linear Multilinear Algebra* **39** 309–323. MR1365451

[17] GIAMBRUNO, A. and ZAICEV, M. (2005). *Polynomial Identities and Asymptotic Methods*. *Mathematical Surveys and Monographs* **122**. Amer. Math. Soc., Providence, RI. MR2176105

[18] GROTHENDIECK, A. (1955). Produits tensoriels topologiques et espaces nucléaires. *Mem. Amer. Math. Soc.* **1955** 140. MR0075539

[19] HAMBLY, B. and LYONS, T. (2010). Uniqueness for the signature of a path of bounded variation and the reduced path group. *Ann. of Math.* (2) **171** 109–167. MR2630037

[20] HARA, K. and HINO, M. (2010). Fractional order Taylor's series and the neo-classical inequality. *Bull. Lond. Math. Soc.* **42** 467–477. MR2651942

[21] HEYER, H. (1977). *Probability Measures on Locally Compact Groups*. Springer, Berlin. MR0501241

[22] HOLUB, J. R. (1970). Tensor product mappings. *Math. Ann.* **188** 1–12. MR0284828

[23] KNOPP, K. (1918). Ein einfaches Verfahren zur Bildüng stetiger nirgends differenzierbarer Funktionen. *Math. Z.* **2** 1–26. MR1544308

[24] LITTERER, C. and LYONS, T. (2012). High order recombination and an application to cubature on Wiener space. *Ann. Appl. Probab.* **22** 1301–1327. MR2985162

[25] LUKACS, E. (1970). *Characteristic Functions*, 2nd ed. Hafner Publishing Co., New York. MR0346874

[26] LYONS, T. (2014). Rough paths, signatures and the modelling of functions on streams. Preprint. Available at arXiv:1405.4537. To appear in the Proceedings of the International Congress of Mathematicians 2014, Korea.

[27] LYONS, T. and NI, H. (2015). Expected signature of Brownian motion up to the first exit time from a bounded domain. *Ann. Probab.* **43** 2729–2762. MR3395473

[28] LYONS, T. and QIAN, Z. (2002). *System Control and Rough Paths*. Oxford Univ. Press, Oxford. MR2036784

[29] LYONS, T. and VICTOIR, N. (2004). Cubature on Wiener space: Stochastic analysis with applications to mathematical finance. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460** 169–198. MR2052260

[30] LYONS, T. and XU, W. (2014). Inverting the signature of a path. Preprint. Available at arXiv:1406.7833.

[31] LYONS, T. J., CARUANA, M. and LÉVY, T. (2007). *Differential Equations Driven by Rough Paths*. *Lecture Notes in Math*. **1908**. Springer, Berlin. MR2314753

[32] MALLIOS, A. (1986). *Topological Algebras*. *Selected Topics*. *North-Holland Mathematics Studies* **124**. North-Holland, Amsterdam. MR0857807

[33] REUTENAUER, C. (1993). *Free Lie Algebras*. *London Mathematical Society Monographs*. *New Series* **7**. Oxford Univ. Press, New York. MR1231799

[34] RUDIN, W. (1973). *Functional Analysis*. McGraw-Hill, New York. MR0365062

[35] SCHAEFER, H. H. (1971). *Topological Vector Spaces*. Springer, New York. MR0342978

[36] TRÈVES, F. (1967). *Topological Vector Spaces*, *Distributions and Kernels*. Academic Press, New York. MR0225131

[37] VALQUI, C. (2001). Universal extension and excision for topological algebras. *K-Theory* **22** 145–160. MR1828613

MATHEMATICAL INSTITUTE
UNIVERSITY OF OXFORD
ANDREW WILES BUILDING
RADCLIFFE OBSERVATORY QUARTER
WOODSTOCK ROAD
OXFORD OX2 6GG
UNITED KINGDOM
E-MAIL: chevyrev@maths.ox.ac.uk

OXFORD-MAN INSTITUTE
OF QUANTITATIVE FINANCE
UNIVERSITY OF OXFORD
WALTON WELL ROAD
OXFORD OX2 6ED
UNITED KINGDOM
E-MAIL: tlyons@maths.ox.ac.uk