Path Developments and Tail Asymptotics of Signature for Pure Rough Paths

Horatio Boedihardjo[*], Xi Geng[†] and Nikolaos P. Souris[‡]

Abstract

Solutions to linear controlled differential equations can be expressed in terms of global iterated path integrals along the driving path. This collection of iterated integrals encodes essentially all information about the underlying path. While upper bounds for iterated path integrals are well known, lower bounds are much less understood, and it is known only relatively recently that some types of asymptotics for the n-th order iterated integral can be used to recover some intrinsic quantitative properties of the path, such as the length for C^1 paths.

In the present paper, we investigate the simplest type of rough paths (the rough path analogue of line segments), and establish uniform upper and lower estimates for the tail asymptotics of iterated integrals in terms of the local variation of the underlying path. Our methodology, which we believe is new for this problem, involves developing paths into complex semisimple Lie algebras and using the associated representation theory to study spectral properties of Lie polynomials under the Lie algebraic development.

[*]Department of Mathematics and Statistics, University of Reading, Reading RG6 6AX, United Kingdom. Email: h.s.boedihardjo@reading.ac.uk.

[†]School of Mathematics and Statistics, University of Melbourne, Melbourne VIC 3010, Australia. Email: xi.geng@unimelb.edu.au.

[‡]Department of Mathematics and Statistics, University of Reading, Reading RG6 6AX, United Kingdom. Email: ntgeva@hotmail.com.

Contents

1 Introduction

Controlled differential equations of the form

$$dY_t = \sum_{i=1}^{d} V_i(Y_t)dX_t^i \tag{1.1}$$

where $V_i : \mathbb{R}^N \to \mathbb{R}^N$, $X : [0,T] \to \mathbb{R}^d$, $Y : [0,T] \to \mathbb{R}^N$, frequently appear in many interesting problems in stochastic analysis and applications to stochastic modelling (cf. [4],

[19], [26], [30] and the references therein). The most well known and fundamental example is perhaps when X_t is a Brownian motion. The rough path theory initiated by Lyons [21] and further developed by many authors (cf. [8], [13], [14]), identifies a wide class of "rough" paths including Brownian motion for which the equation (1.1) is well defined. The theory is analytically consistent with the classical viewpoint, in the sense that it is a continuous extension of the Lebesgue-Stieltjes theory with respect to the rough path topology and reduces to the classical setting when the underlying paths have finite lengths. Rough path theory naturally motivates the study of analytic properties of solutions to (1.1) driven by rough paths.

One particularly tractable class of examples is when the vector fields $(V_i)_{1\leqslant i\leqslant d}$ are linear. In this case, the solution at time $t=T$ can be represented explicitly as

$$Y_T=\sum_{n=0}^{\infty}\sum_{i_1,\cdots,i_n=1}^{d}V_{i_1}\cdots V_{i_n}(Y_0)\cdot\int_{0<t_1<\cdots<t_n<T}dX_{t_1}^{i_1}\cdots dX_{t_n}^{i_n}.$$

In particular, Y_T depends on the driving path X through the collection of iterated coordinate integrals

$$S(X)\triangleq\left\{\int_{0<t_1<\cdots<t_n<T}dX_{t_1}^{i_1}\cdots dX_{t_n}^{i_n}:\ n\geqslant 1,\ 1\leqslant i_1,\cdots,i_n\leqslant d\right\}.$$

For algebraic reasons, it is useful to think of this collection as a single element of the infinite tensor algebra $T((\mathbb{R}^d))\triangleq\prod_{n=0}^{\infty}(\mathbb{R}^d)^{\otimes n}$, more intrinsically as

$$S(X)=1+\sum_{n=1}^{\infty}\int_{0<t_1<\cdots<t_n<T}dX_{t_1}\otimes\cdots\otimes dX_{t_n}.$$

This tensor element $S(X)$, known as the *signature* of the path X, plays an essential role in rough path theory. The significance and usefulness of path signature is based on a fundamental theorem which asserts that every (weakly geometric) rough path is uniquely determined by its signature up to tree-like pieces (cf. [15] and [2]). However, the proof of this uniqueness result is non-constructive and does not contain information about how one can reconstruct a rough path from its signature. The general reconstruction problem was studied by many authors (cf. [25], [12], [3]).

On the other hand, combining with algebraic properties of signature, the uniqueness result ensures that essentially all information about the rough path is encoded in the *tail* of its signature, i.e. when looking at the component $\int dX_{t_1}\otimes\cdots\otimes dX_{t_n}$ in the asymptotics as $n\to\infty$. An interesting question arises naturally as follows.

Question: Are there explicit and elegant formulae allowing us to recover intrinsic properties of the path from its signature tail asymptotics?

The study of this question begins by observing the following elementary estimate

$$\left\|\int_{0<t_1<\cdots<t_n<T} dX_{t_1}\otimes\cdots\otimes dX_{t_n}\right\| \leqslant \frac{\|X\|_{1\text{-var}}^n}{n!} \tag{1.2}$$

when the path X has finite length. A surprising and highly non-trivial fact is that this simple estimate becomes asymptotically sharp as $n\to\infty$, at least for the class of C^1 paths. In a precise and elegant way, it was shown by Hambly-Lyons [15], and subsequently by Lyons-Xu [24] that the tail asymptotics of the normalized signature recovers the length of a C^1 path with unit speed parametrization:

$$\lim_{n\to\infty}\left(n!\left\|\int_{0<t_1<\cdots<t_n<T} dX_{t_1}\otimes\cdots\otimes dX_{t_n}\right\|\right)^{\frac{1}{n}} = \|X\|_{1\text{-var}}. \tag{1.3}$$

It was conjectured by Chang-Lyons-Ni [5] that the same formula should hold true for general paths with finite length, which remains an important and challenging open problem.

We are interested in the analogue of the formula (1.3) in the rough path context. First of all, the rough path analogue of the factorial estimate (1.2) becomes much deeper, and the following type of uniform upper estimate for rough paths with finite p-variation ($p\geqslant 1$) was due to Lyons [21] (cf. Theorem 2.5 below):

$$\left\|\int_{0<t_1<\cdots<t_n<T} dX_{t_1}\otimes\cdots\otimes dX_{t_n}\right\| \leqslant \frac{C_p\cdot\|\mathbf{X}\|_{p\text{-var}}^n}{(n/p)!}.$$

If one believes that the above estimate is asymptotically sharp as $n\to\infty$ for paths whose intrinsic roughness is p, we are naturally led to considering the quantity

$$L_p(\mathbf{X}) \triangleq \limsup_{n\to\infty}\left(\left(\frac{n}{p}\right)!\left\|\int_{0<t_1<\cdots<t_n<T} dX_{t_1}\otimes\cdots\otimes dX_{t_n}\right\|\right)^{\frac{p}{n}} \tag{1.4}$$

constructed from the tail of signature, and looking for its connection with intrinsic properties of the path $\mathbf{X}$. The quantity $L_p(\mathbf{X})$ certainly does not recover the usual p-variation, since $L_p(\mathbf{X})=0$ for any bounded variation path if $p>1$ as a simple consequence of (1.2), while the p-variation of a bounded variation path need not be zero . The first hint about the meaning of $L_p(\mathbf{X})$ was provided by Boedihardjo and Geng [1], in which the authors showed that, when $\mathbf{X}$ is a Brownian motion and $p=2$, with probability one $L_p(\mathbf{X})$ is a deterministic constant multiple of the quadratic variation of Brownian motion. To some extent, this is suggesting that, $L_p(\mathbf{X})$ may be intimately related to certain notion of *local* p-variation defined in a similar way to the usual p-variation but along partitions with arbitrarily fine scales, which can also be interpreted as an additive notion of length in the rough path context.

The main goal of the present paper is to investigate this problem at a precise quantitative level for the class of rough paths that are natural extensions of classical line segments. These paths, known as *pure rough paths*, are of the form $\mathbf{X}_t=\exp(tl)$ ($0\leqslant t\leqslant 1$) where l is a Lie polynomial of degree $m\geqslant 1$. If $m=1$, $\mathbf{X}_t$ becomes a classical line segment represented by the vector $l\in V$. In general, $\mathbf{X}_t$ carries an intrinsic roughness of m. This is also a natural class of rough paths in the sense that their signatures are given by the exponential of Lie polynomials (cf. Proposition 2.11 below).

We are going to show that, for any pure rough path $\mathbf{X}_t = \exp(tl)$ over $\mathbb{R}^d$ with roughness m, under the projective tensor norm, the signature tail asymptotics $L_m(\mathbf{X})$ defined by (1.4) with $p = m$ is precisely related to the highest degree component l_m of the Lie polynomial l through the uniform estimate

$$c(m, d) \cdot \|l_m\| \leqslant L_m(\mathbf{X}) \leqslant \|l_m\|, \tag{1.5}$$

where $c(m, d) \in (0, 1]$ is a constant depending only on the roughness m and the dimension d which also admits an explicit lower estimate. The quantity $\|l_m\|$ coincides with the local m-variation of $\mathbf{X}$ interpreted by Proposition 2.10 below. When $d = 2$ and $m = 2, 3$, we have $c(m, d) = 1$ and therefore

$$L_m(\mathbf{X}) = \|l_m\|. \tag{1.6}$$

The same conclusion also holds for some cases in degrees $m = 4, 5$. The precise formulation of our main result is given by Theorem 2.13 below. On the other hand, if one works with the Hilbert-Schmidt tensor norm, there is also a class of pure rough paths for which $c(m, d) = 1$. We conjecture that the formula (1.6) is true for arbitrary pure rough paths. This can be viewed as the analogue of the formula (1.3) in the pure rough path context.

Our proof of the upper estimate in (1.5) has a combinatorial flavour that relies on Stirling's approximation and the multivariate neo-classical inequality proved by Friz-Riedel [10]. The core of our work, which lies in establishing a matching lower estimate, is a novel method based on the representation theory of complex semisimple Lie algebras. To be more precise, our starting point is a general representation of the tensor algebra that allows us to develop paths onto an automorphism group from Cartan's viewpoint. Specific choices of such representations were already used by Hambly-Lyons [15] and Lyons-Xu [24] for proving (1.3) for C^1 paths, and also by Chevyrev-Lyons [7] and Lyons-Sidorova [23] for studying other signature-related properties. Using such an arbitrary representation already allows us to establish a general intermediate lower bound of the signature tail asymptotics quantity in terms of eigenvalues of the Lie polynomial defining the pure rough path. The key ingredient in our approach, is to allow such representation factor through a complex semisimple Lie algebra $\mathfrak{g}$ and develop the highest degree Lie component into a so-called Cartan subalgebra of $\mathfrak{g}$. It turns out that, under this semisimple picture, the associated representation theory for $\mathfrak{g}$ enables us to study spectral properties of the highest degree Lie component in an effective and quantitative way, leading us to the main lower estimate. We explain the strategy and elaborate these points more precisely in Section 4.2 as we develop the mathematical details.

It is also worthwhile to mention that, as an immediate application of our methodology, one can prove a separation of points property for path signatures. More specifically, if g_1 and g_2 are two distinct group-like elements as the signatures of two different rough paths over $\mathbb{R}^d$, then one can find a finite dimensional semisimple Lie algebra $\mathfrak{g}$ and an embedding $F : \mathbb{R}^d \to \mathfrak{g}$, such that $F(g_1) \neq F(g_2)$ where F also denotes the natural extension to the tensor algebra over $\mathbb{R}^d$. The precise formulation and proof of this fact is given in Corollary

4.30 below. Such a separation property was first obtained by Chevyrev-Lyons [7] as an essential ingredient for proving their uniqueness result for the expected signature of stochastic processes.

Organization of the paper. In Section 2, we recall some basic notions from rough path theory and then formulate our main result in Theorem 2.13. In Section 3, we give some heuristics on the underlying problem by discussing some special examples. Another result that complements our main result is stated in Theorem 3.1. In Section 4, we develop the proof of our main result. Section 4.1 is devoted to the upper estimate, and Section 4.2 is devoted to the lower estimate in which we divide the proof into several intermediate steps and results. In Section 5, we give the proof of Theorem 3.1.

2 Notions from rough path theory and statement of the main result

In this section, we recall some basic ideas and notions from the rough path theory developed by Lyons [21]. We refer the reader to the monographs by Lyons-Qian [22] and Friz-Victoir [11] for a systematic introduction. After that, we formulate the main result of the present paper.

2.1 The rough path structure

The fundamental insight of rough path theory is that, beyond certain level of regularity, the structure encoded in a given path living in some Banach space V becomes no longer sufficient for yielding an analytically consistent notion of integration and differential equations, and thus higher order structures (iterated path integrals) need to be specified along with the underlying path as a priori information. Mathematically, a rough path should be viewed as a generic path living inside some tensor group in which the state space V is embedded as the first order structure.

Let $(V, \|\cdot\|)$ be a given fixed Banach space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$.

Definition 2.1. A sequence $\{\|\cdot\|_{V^{\otimes_a m}} : m \geqslant 1\}$ of norms on the algebraic tensor products $\{V^{\otimes_a m} : m \geqslant 1\}$ are called *reasonable tensor algebra norms* if

(i) $\|\cdot\|_V = \|\cdot\|$;
(ii) $\|\xi \otimes \eta\|_{V^{\otimes_a (m+n)}} \leqslant \|\xi\|_{V^{\otimes_a m}} \cdot \|\eta\|_{V^{\otimes_a n}}$ for $\xi \in V^{\otimes_a m}$ and $\eta \in V^{\otimes_a n}$;
(iii) $\|\phi \otimes \psi\| \leqslant \|\phi\| \cdot \|\psi\|$ for $\phi \in (V^{\otimes_a m})^*$ and $\psi \in (V^{\otimes_a n})^*$ where the norms are the induced dual norms;
(iv) $\|P^\sigma(\xi)\|_{V^{\otimes_a m}} = \|\xi\|_{V^{\otimes_a m}}$ for $\xi \in V^{\otimes_a m}$ and σ being a permutation of order m, where P^σ is the permutation operator induced by σ on m-tensors.

It is known that the inequalities in (ii) and (iii) automatically become equalities (cf. [9]). The completion of $V^{\otimes_a m}$ under $\|\cdot\|_{V^{\otimes_a m}}$ is denoted as $(V^{\otimes m}, \|\cdot\|_{V^{\otimes m}})$.

Examples of reasonable tensor algebra norms include the projective tensor norm, the injective tensor norm, and the Hilbert-Schmidt tensor norm if V is a Hilbert space. Since the projective tensor norm is mostly relevant to us, we recall its definition here. Given $\xi \in V^{\otimes_a m}$, the *projective tensor norm* of ξ is defined by

$$\|\xi\|_{\text{proj}} \triangleq \inf\left\{\sum_{i=1}^{r} \|v_1^i\| \cdots \|v_m^i\| : \xi = \sum_{i=1}^{r} v_1^i \otimes \cdots \otimes v_m^i \text{ with } r \geqslant 1, v_j^i \in V\right\}.$$

Given a fixed norm on V, the associated projective tensor norm is the largest among all reasonable tensor algebra norms. It admits the following dual characterization (cf. [28]):

$$\|\xi\|_{\text{proj}} = \sup\left\{|B(\xi)| : B \in \mathcal{L}(V \times \cdots \times V; \mathbb{R}),\ \|B\| \leqslant 1\right\}. \tag{2.1}$$

When $V = \mathbb{R}^d$ is equipped with the l^1-norm with respect to the standard basis, the associated projective tensor norm on $V^{\otimes m}$ coincides with the l^1-norm with respect to the canonical tensor basis.

From now on, we assume that a sequence of reasonable tensor algebra norms are given and fixed. We often omit the subscript when the norms are clear from the context.

Let $T((V))$ be the *infinite tensor algebra* consisting of tensor series $\xi = (\xi_0, \xi_1, \xi_2, \cdots)$ with $\xi_n \in V^{\otimes n}$ for each n ($V^{\otimes 0} \triangleq \mathbb{F}$). Given $n \geqslant 1$, let $T^{(n)}(V) \triangleq \oplus_{k=0}^{n} V^{\otimes k}$ be the *truncated tensor algebra* of degree n. There are natural notions of exponential and logarithm over these tensor algebras defined by using the standard Taylor expansion formula with respect to the tensor product. For instance, the exponential function over $T((V))$ is given by

$$\exp(\xi) \triangleq \sum_{n=0}^{\infty} \frac{1}{n!}\xi^{\otimes n}, \quad \xi \in T((V)),$$

while over $T^{(n)}(V)$ it is defined by the same formula but truncated up to degree n.

Definition 2.2. A *multiplicative functional* of degree n is a continuous functional

$$\mathbf{X} = (1, X^1, \cdots, X^n) : \Delta_T \triangleq \{(s,t) : 0 \leqslant s \leqslant t \leqslant T\} \to T^{(n)}(V)$$

such that $\mathbf{X}_{s,u} = \mathbf{X}_{s,t} \otimes \mathbf{X}_{t,u}$ for all $s \leqslant t \leqslant u$. Given a real number $p \geqslant 1$, $\mathbf{X}$ is said to have *finite total p-variation* if

$$\|\mathbf{X}\|_{p\text{-var}} \triangleq \sum_{k=1}^{n} \sup_{\mathcal{P}} \left(\sum_{t_i \in \mathcal{P}} \|X_{t_{i-1},t_i}^k\|^{\frac{p}{k}}\right)^{\frac{k}{p}} < \infty, \tag{2.2}$$

where the supremum is taken over all finite partitions of $[0,T]$. A *rough path* with roughness p (or simply a p-rough path) is a multiplicative functional of degree $\lfloor p \rfloor$ which has finite total p-variation, where $\lfloor p \rfloor$ denotes the largest integer not exceeding p.

Remark 2.3. Due to multiplicativity, a rough path $\mathbf{X}_{s,t}$ can be equivalently regarded as an actual path $\mathbf{X}_t \triangleq \mathbf{X}_{0,t}$ and vice versa by $\mathbf{X}_{s,t} \triangleq \mathbf{X}_s^{-1} \otimes \mathbf{X}_t$. We do not distinguish these two viewpoints.

The notion of rough paths is mostly useful when a crucial Lie algebraic property is satisfied. Recall that there is a natural Lie structure on the tensor algebra given by $[\xi,\eta] \triangleq \xi\otimes\eta-\eta\otimes\xi$. The space of *homogeneous Lie polynomials* of degree n, denoted as $\mathcal{L}_n(V)$, is the norm completion of the algebraic space $\mathcal{L}_n^a(V)$ defined inductively by $\mathcal{L}_1^a(V)\triangleq V$ and $\mathcal{L}_{n+1}^a(V)\triangleq[V,\mathcal{L}_n^a(V)]$. Define the space of *Lie polynomials* of degree n by

$$\mathcal{L}^{(n)}(V)\triangleq\oplus_{k=1}^n\mathcal{L}_k(V).$$

This is also called the *free nilpotent Lie algebra* of degree n. Correspondingly, the *free nilpotent Lie group* of degree n is defined by

$$G^{(n)}(V)\triangleq\exp(\mathcal{L}^{(n)}(V)).$$

They are both canonically embedded inside $T^{(n)}(V)$.

Definition 2.4. A p-rough path is said to be *weakly geometric* if it takes values in the group $G^{(\lfloor p\rfloor)}(V)$.

Weakly geometric rough paths cover a wide range of interesting examples, for instance bounded variation paths ($p=1$), Brownian motion and continuous semimartingales ($2<p<3$), wide classes of Gaussian processes and Markov processes ($p>3$) etc. This is the appropriate class of paths which the rough path theory of integration and differential equations is based on.

2.2 The signature of a rough path

An important aspect of rough path theory is the characterization of rough paths in terms of the so-called path signature, which is a generalized notion of iterated path integrals. Its definition is based on the following basic property of rough paths proved by Lyons [21].

Theorem 2.5 (Lyons' Extension Theorem)**.** *Let $\mathbf{X}=(\mathbf{X}_{s,t})_{0\leqslant s\leqslant t\leqslant T}$ be a p-rough path. Then there exists a unique extension of $\mathbf{X}$ to a multiplicative functional $\mathbb{X}:\Delta_T\to T((V))$:*

$$(s,t)\mapsto\mathbb{X}_{s,t}=(1,X_{s,t}^1,\cdots,X_{s,t}^{\lfloor p\rfloor},\cdots,X_{s,t}^n,\cdots),$$

whose restriction to $T^{(n)}(V)$ has finite total p-variation for all $n\geqslant\lfloor p\rfloor+1$. Moreover, there exist a universal constant β_p depending only on p and a non-negative function $\omega_{\mathbf{X}}(s,t)$ related to the p-variation of $\mathbf{X}$, such that

$$\|X_{s,t}^n\|\leqslant\frac{\omega_{\mathbf{X}}(s,t)^{n/p}}{\beta_p(n/p)!},\quad\text{for all } n\geqslant1 \text{ and } (s,t)\in\Delta_T,\tag{2.3}$$

where the factorial $(n/p)!$ is defined by using the Gamma function.

Definition 2.6. The tensor series $\mathbb{X}_{0,T}\in T((V))$ is called the *signature* of $\mathbf{X}$. It is usually denoted as $S(\mathbf{X})$.

Example 2.7. If $(X_t)_{0\leqslant t\leqslant T}$ is a bounded variation path, then its signature is precisely the sequence of iterated path integrals

$$\left(1, X_T - X_0, \int_{0<s<t<T} dX_s \otimes dX_t, \cdots\right) \in T((V))$$

defined in the sense of Lebesgue-Stieltjes. In this case, the factorial estimate (2.3) reduces to the elementary estimate (1.2). If $(B_t)_{0\leqslant t\leqslant T}$ is a multidimensional Brownian motion, then its (pathwise) signature coincides with the sequence of iterated stochastic integrals defined in the sense of Stratonovich.

It is a fundamental result (cf. [15] and [2]) that every weakly geometric rough path over a real Banach space is uniquely determined by its signature up to tree-like pieces. In addition, it is a consequence of the weakly geometric property that any given component of signature can be embedded into arbitrary higher degree components by raising tensor powers (cf. [5]). Therefore, the *tail* of signature (in the asymptotics as degree tends to infinity) encodes essentially all information about the underlying path.

In view of the factorial estimate (2.3), a natural quantity one can construct from the tail of signature is the normalized component $((n/p)!\|X_{0,T}^n\|)^{p/n}$ as $n\to\infty$. Since signature components can vanish infinitely often, we are led to considering the functional

$$L_p(\mathbf{X}) \triangleq \limsup_{n\to\infty} \left(\left(\frac{n}{p}\right)! \left\|X_{0,T}^n\right\|\right)^{\frac{p}{n}}. \tag{2.4}$$

Our goal is to investigate at a quantitative level how the tail asymptotics quantity $L_p(\mathbf{X})$ is related to certain notion of local p-variation of $\mathbf{X}$ for a natural class of rough paths known as pure rough paths. These are straight forward analogues of line segments in the rough path context, and it is also the class of rough paths whose signature is given by the exponential of a Lie polynomial. They form the first non-trivial class of rough paths for the underlying problem.

2.3 Pure rough paths and formulation of the main result

Now we give the precise definition of the aforementioned class of rough paths that we will be working wtih. Let $m\geqslant 1$ be a given integer.

Definition 2.8. A *pure m-rough path* is a weakly geometric rough path of the form

$$\mathbf{X}_t = \exp(tl) \in G^{(m)}(V), \quad 0\leqslant t\leqslant 1,$$

where $l\in\mathcal{L}^{(m)}(V)$ is a Lie polynomial of degree m.

Example 2.9. When $m=1$, a pure 1-rough path is simply a line segment in V.

We list a few basic properties of pure rough paths that are relevant to us and leave the proofs in the appendix so as not to distract the reader from the main picture.

Proposition 2.10. *A pure m-rough path $\mathbf{X}_t = \exp(tl)$ is a rough path with roughness m in the sense of Definition 2.2. In addition, the local m-variation of $\mathbf{X}$ coincides with the norm of the highest degree component of l, in the sense that*

$$\lim_{n\to\infty}\sum_{k=1}^{m}\left(\sum_{t_i\in\mathcal{P}_n}\|X^k_{t_{i-1},t_i}\|^{\frac{m}{k}}\right)^{\frac{k}{m}} = \|\pi_m(l)\|$$

for any sequence of finite partitions $\mathcal{P}_n$ over $[0,1]$ whose mesh size tends to zero, where $\pi_m : T^{(m)}(V) \to V^{\otimes m}$ is the canonical projection.

Proposition 2.11. *Let $\mathbf{X}_t = \exp(tl)$ be a pure m-rough path. Then its signature is equal to* $\exp(l)$ *where the exponential is now taken over the infinite tensor algebra $T((V))$. In addition, up to tree-like equivalence this is the only weakly geometric rough path whose signature is* $\exp(l)$.

In the case of pure rough paths, we believe that the relationship between the signature tail asymptotics and the local m-variation is as simple and neat as stated in the following conjectural formula.

Conjecture 2.12. *For every pure m-rough path $\mathbf{X}_t = \exp(tl) \in G^{(m)}(V)$, the tail asymptotics quantity $L_m(\mathbf{X})$ of signature equals the local m-variation of $\mathbf{X}$. In view of Proposition 2.10, that is $L_m(\mathbf{X}) = \|\pi_m(l)\|$.*

As a first major step towards understanding this problem, our main result can be summarized as a uniform upper and lower estimate of $L_m(\mathbf{X})$ in terms of $\|\pi_m(l)\|$ for pure m-rough paths.

Theorem 2.13. *Let V be a finite dimensional Banach space, and let every tensor product $V^{\otimes n}$ be equipped with the associated projective tensor norm. Then for each $m \geqslant 1$, there exists a constant $c(m,d) \in (0,1]$ depending only on m and $d \triangleq \dim V$, such that*

$$c(m,d)\|\pi_m(l)\| \leqslant L_m(\mathbf{X}) \leqslant \|\pi_m(l)\|$$

for all pure m-rough paths $\mathbf{X}_t = \exp(tl) \in G^{(m)}(V)$. The factor $c(m,d)$ admits an explicit lower estimate

$$c(m,d) \geqslant \Lambda_d^{-m}\cdot 2^{-(\nu_{m,d}!)^{\gamma\nu_{m,d}}},$$

where Λ_d is a constant depending only on d, $\nu_{m,d} \triangleq \dim \mathcal{L}_m(V)$, and $\gamma > 1$ is a universal constant.

In addition, if $V = \mathbb{R}^2$ is equipped with the l^1-norm with respect to the canonical basis, then for degrees $m = 2,3$, we further have $c(m,d) = 1$, showing that Conjecture 2.12 holds for these cases. The same conclusion holds for some cases in degrees $m = 4,5$ specified in Section 4.2.4, Part II.

Remark 2.14. When $m = 1$, Conjecture 2.12 boils down to the bounded variation formula (1.3) which holds trivially in this case since the underlying path is now a classical line segment. Apart from the low degree cases stated in Theorem 2.13, the conjecture also holds true for the following two classes of pure rough paths with arbitrary roughness: homogeneous (cf. Section 3 below) and inhomogeneous with two components satisfying certain parity condition (cf. Theorem 3.1 below under the Hilbert-Schmidt tensor norm).

Although the main problem and result are motivated from rough path theory, we also give a parallel algebraic formulation which might raise potential interests from other fields.

Conjecture 2.12'. *Let $(V, \|\cdot\|)$ be a finite dimensional Banach space, and let the tensor products be equipped with some given reasonable tensor algebra norms. Then for any Lie polynomial l, the following asymptotics formula holds true:*

$$\limsup_{n\to\infty}\Big(\Big(\frac{n}{m}\Big)!\|\pi_n(\exp(l))\|\Big)^{\frac{m}{n}} = \|\pi_m(l)\|,$$

where m is the degree of the Lie polynomial l.

Theorem 2.13'. *Let $(V, \|\cdot\|)$ be a d dimensional Banach space, and let the tensor products be equipped with the associated projective tensor norm. Then for each $m \geqslant 1$, there exists a constant $c(m,d) \in (0,1]$ depending only on m and d, such that for any Lie polynomial l of degree m, the following estimate holds true:*

$$c(m,d)\|\pi_m(l)\| \leqslant \limsup_{n\to\infty}\Big(\Big(\frac{n}{m}\Big)!\|\pi_n(\exp(l))\|_{V^{\otimes n}}\Big)^{\frac{m}{n}} \leqslant \|\pi_m(l)\|.$$

The factor $c(m,d)$ admits an explicit lower estimate, and for some low degree cases we further have $c(m,d) = 1$ giving the sharp result, precisely as stated in Theorem 2.13.

3 Some special examples and heuristic calculations

Before developing the proof of Theorem 2.13, we examine a few special examples in order to get a better sense of the problem.

In the first place, the problem is trivial when (and only when) $\mathbf{X}_t$ is defined by a homogeneous polynomial. More precisely, if $\mathbf{X}_t = \exp(tl)$ with $l \in V^{\otimes m}$, it is immediate that

$$X^n = \pi_n(\exp(l)) = \sum_{k=0}^{\infty}\frac{1}{k!}\pi_n(l^{\otimes k}) = \begin{cases}\frac{1}{(n/m)!}l^{\otimes(n/m)}, & m \mid n,\\ 0, & m \nmid n.\end{cases}$$

Therefore,

$$L_m(\mathbf{X}) = \lim_{k\to\infty}\big(k!\|X^{km}\|\big)^{\frac{1}{k}} = \lim_{k\to\infty}\Big(\|l^{\otimes k}\|^{\frac{1}{k}}\Big) = \|l\|, \tag{3.1}$$

and Conjecture 2.12 holds trivially for $\mathbf{X}_t$.

A less trivial example is $l = \mathrm{e}_1 + [\mathrm{e}_1, \mathrm{e}_2]$, in which we have

$$X^{2n} = \pi_{2n}(\exp(l)) = \sum_{k=n}^{2n}\frac{1}{k!}\pi_{2n}\left((\mathrm{e}_1 + [\mathrm{e}_1,\mathrm{e}_2])^{\otimes k}\right). \tag{3.2}$$

A rather special observation in this example is that, the expansion of $\pi_{2n}((\mathrm{e}_1 + [\mathrm{e}_1, \mathrm{e}_2])^{\otimes k})$ is supported on disjoint sets of words for different k's. Suppose we work with the projective tensor norm induced from the standard l^1-norm on $\mathbb{R}^2$. It then follows that

$$\|X^{2n}\| = \sum_{k=n}^{2n} \frac{1}{k!} \left\| \pi_{2n} \left((\mathrm{e}_1 + [\mathrm{e}_1, \mathrm{e}_2])^{\otimes k} \right) \right\| \geqslant \frac{2^n}{n!}.$$

In particular,

$$L_2(\mathbf{X}) \geqslant \limsup_{n \to \infty} \left(n! \|X^{2n}\| \right)^{\frac{1}{n}} = 2 = \|\pi_2(l)\|.$$

Combining with the general upper bound to be established in Theorem 4.1 below, we see that Conjecture 2.12 holds for $\mathbf{X}_t$.

However, it becomes much less clear how similar calculations can be done even for the next simple candidate $l = \mathrm{e}_1 + \mathrm{e}_2 + [\mathrm{e}_1, \mathrm{e}_2]$. Brute force calculation does not give us much insight to proceed further. The main challenge of the problem lies in understanding the complicated interactions (possibly cancellations) among different degree components of l when looking at the signature expansion at arbitrarily high degrees.

On the other hand, some extra mileage can still be achieved if we work with the Hilbert-Schmidt tensor norm. Recall that the *Hilbert-Schmidt tensor norm* over the tensor product of two Hilbert spaces H_1, H_2 is induced by

$$\langle v_1 \otimes w_1, v_2 \otimes w_2 \rangle_{H_1 \otimes H_2} \triangleq \langle v_1, v_2 \rangle_{H_1} \cdot \langle w_1, w_2 \rangle_{H_2}, \quad v_1, v_2 \in H_1, \ w_1, w_2 \in H_2.$$

In this context, we can prove the following result. We postpone the proof to Section 5, in which the strategy, based on orthogonality properties in free Lie algebras, is very different from the main approach of proving Theorem 2.13.

Theorem 3.1. *Let V be a finite dimensional Hilbert space, and let the tensor products be equipped with the induced Hilbert-Schmidt tensor norm. Suppose that $\mathbf{X}_t = \exp(t(l_a + l_b))$, where l_a, l_b are homogeneous Lie polynomials of degrees a, b respectively for $a < b$. If $(b-a)/\gcd(a,b)$ is an odd integer where "gcd" denotes the greatest common divisor, then Conjecture 2.12 holds for $\mathbf{X}_t$.*

As an example, we immediately see that Conjecture 2.12 holds for $l = \mathrm{e}_1 + \mathrm{e}_2 + [\mathrm{e}_1, \mathrm{e}_2]$ under the Hilbert-Schmidt tensor norm. However, the argument breaks down if $(b-a)/\gcd(a,b)$ is an even number, or if l has more than two homogeneous components.

The above special examples seem to suggest that, the key to getting the lower bound is the concentration of the degree km signature expansion around the term $\pi_m(l)^{\otimes k}/k!$ as $k \to \infty$. However, the picture can be much subtler in general. Some heuristic estimates on magnitudes suggest that the signature expansion at degree km is concentrated at a number of terms near $\pi_m(l)^{\otimes k}/k!$, each possibly having comparable magnitudes. As $k \to \infty$, the total number of these terms seem to be of order $o(k)$, and there can be delicate cancellations among them which are hard to analyze.

One contribution of the present paper is to develop a general strategy which on the one hand allows us to overcome the above difficulties to some extent and on the other hand is specific enough to be implemented computationally in order to generate explicit quantitative estimates in many interesting examples.

4 Proof of the main theorem

Throughout the rest of this section, unless otherwise stated, let $(V, \|\cdot\|)$ be a finite dimensional Banach space and let each tensor product $V^{\otimes n}$ $(n \geqslant 1)$ be equipped with the projective tensor norm. We work with a given pure m-rough path $\mathbf{X}_t = \exp(tl)$ defined by some Lie polynomial $l \in \mathcal{L}^{(m)}(V)$.

We aim at studying the relationship between the signature tail asymptotics of $\mathbf{X}$ defined by $L_p(\mathbf{X})$ in (2.4) with $p = m$, and the local m-variation of $\mathbf{X}$ which is also equal to $\|\pi_m(l)\|$ by Proposition 2.10. Our main result, as stated in Theorem 2.13, is a uniform upper and lower estimate of $L_m(\mathbf{X})$ in terms of $\|\pi_m(l)\|$. The techniques we develop for proving the two estimates are drastically different. The upper estimate is based on combinatorial arguments while the lower estimate relies on the representation theory of complex semisimple Lie algebras.

4.1 The upper estimate

We start by establishing the (sharp) upper bound. In this part, more generality can be pursued: V can be infinite dimensional, tensor norms only need to be reasonable and l need not be of Lie type.

Theorem 4.1. *We have the following upper estimate*

$$L_m(\mathbf{X}) \leqslant \|\pi_m(l)\|$$

for all rough paths of the form $\mathbf{X}_t = \exp(tl)$ with l being an arbitrary element in $T^{(m)}(V)$.

Our proof of Theorem 4.1 relies on a multivariate neo-classical inequality proved by Friz-Riedel [10]. The bivariate version was proved by Hara-Hino [16].

Lemma 4.2 (cf. [10], Lemma 1)**.** *Suppose that $a_1, \cdots, a_m > 0$, $p \geqslant 1$ and $n \in \mathbb{N}$. Then we have*

$$\sum_{\substack{0 \leqslant k_1, \cdots, k_m \leqslant n \\ k_1 + \cdots + k_m = n}} \frac{a_1^{k_1/p} \cdots a_m^{k_m/p}}{(k_1/p)! \cdots (k_m/p)!} \leqslant p^{m-1} \cdot \frac{(a_1 + \cdots + a_m)^{n/p}}{(n/p)!}.$$

We also need the following analytic lemma.

Lemma 4.3. *Suppose that $0 < \alpha < \beta \leqslant 1$ and $a, b > 0$. Then we have*

$$\limsup_{n \to \infty} \left((n\alpha)! \sum_{j=0}^{n} \frac{a^{j\alpha} b^{(n-j)\alpha}}{(j\beta)!\,((n-j)\alpha)!} \right)^{\frac{1}{n\alpha}} \leqslant b.$$

Proof. From Stirling's approximation, we know that

$$\frac{(j\alpha)!}{(j\beta)!} \sim \sqrt{\frac{\alpha}{\beta}} \left(\frac{\alpha^\alpha \mathrm{e}^{\beta-\alpha}}{\beta^\beta} \right)^j j^{(\alpha-\beta)j}, \quad j \to \infty$$

In particular, given any arbitrary $\varepsilon > 0$, there exists $J \geqslant 1$ such that

$$\frac{(j\alpha)!}{(j\beta)!} \leqslant \varepsilon^j \quad \forall j \geqslant J.$$

It follows that

$$\sum_{j=0}^{n} \frac{a^{j\alpha} b^{(n-j)\alpha}}{(j\beta)!\,((n-j)\alpha)!} \leqslant \sum_{j=0}^{J-1} \frac{a^{j\alpha} b^{(n-j)\alpha}}{(j\beta)!\,((n-j)\alpha)!} + \sum_{j=J}^{n} \frac{\left(\varepsilon^{1/\alpha} a\right)^{j\alpha} b^{(n-j)\alpha}}{(j\alpha)!\,((n-j)\alpha)!}. \tag{4.1}$$

To estimate the first term on the right hand side of (4.1), using Stirling's approximation again, it is easily seen that

$$\frac{(n\alpha)!}{(j\beta)!\,((n-j)\alpha)!} a^{j\alpha} b^{(n-j)\alpha} \leqslant C n^{J\alpha} b^{n\alpha} \quad \text{for all } 0 \leqslant j < J,$$

where C is a constant depending on a, b, α, β and J. To estimate the second term on the right hand side of (4.1), using Lemma 4.2 with $m = 2$ and $p = 1/\alpha$, we have

$$\sum_{j=J}^{n} \frac{\left(\varepsilon^{1/\alpha} a\right)^{j\alpha} b^{(n-j)\alpha}}{(j\alpha)!\,((n-j)\alpha)!} \leqslant \frac{\left(\varepsilon^{1/\alpha} a + b\right)^{n\alpha}}{\alpha (n\alpha)!}.$$

By substituting the above two estimates into (4.1), we have

$$(n\alpha)! \sum_{j=0}^{n} \frac{a^{j\alpha} b^{(n-j)\alpha}}{(j\beta)!\,((n-j)\alpha)!} \leqslant C n^{J\alpha} b^{n\alpha} + \frac{\left(\varepsilon^{1/\alpha} a + b\right)^{n\alpha}}{\alpha}.$$

Therefore, by taking $n \to \infty$, we arrive at

$$\limsup_{n\to\infty} \left((n\alpha)! \sum_{j=0}^{n} \frac{a^{j\alpha} b^{(n-j)\alpha}}{(j\beta)!\,((n-j)\alpha)!} \right)^{\frac{1}{n\alpha}} \leqslant \varepsilon^{1/\alpha} a + b,$$

which yields the result since ε is arbitrary. $\square$

With the help of the above two lemmas, we can now give the proof of Theorem 4.1.

Proof of Theorem 4.1. Given $\mathbf{X}_t = \exp(tl)$ with $l \in T^{(m)}(V)$, we write $l = l_1 + \cdots + l_m$ where $l_i \in V^{\otimes i}$. For each $n \geqslant 1$, the n-th degree signature of $\mathbf{X}$ can be estimated by

$$\begin{aligned}
\|X^n\| &= \|\pi_n\left(\exp(l)\right)\| \\
&= \left\| \sum_{k=0}^{\infty} \frac{1}{k!} \pi_n \left((l_1 + \cdots + l_m)^{\otimes k} \right) \right\| \\
&\leqslant \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\substack{1 \leqslant i_1, \cdots, i_k \leqslant m \\ i_1 + \cdots + i_k = n}} \|l_{i_1}\| \cdots \|l_{i_k}\| \\
&= \sum_{\substack{j_1, \cdots, j_m \geqslant 0 \\ j_1 + 2j_2 + \cdots + mj_m = n}} \frac{\|l_1\|^{j_1} \cdots \|l_m\|^{j_m}}{j_1! \cdots j_m!}.
\end{aligned}$$

To reach the last equality, we have used a different way to count terms that have a total degree of n in the expansion of $(\|l_1\| + \cdots + \|l_m\|)^k$. By applying change of variables $k_r = rj_r$ $(1 \leqslant r \leqslant m)$, we arrive at

$$\|X^n\| \leqslant \sum_{\substack{k_1, \cdots, k_m \geqslant 0 \\ k_1 + \cdots + k_m = n}} \frac{\|l_1\|^{k_1} \|l_2\|^{k_2/2} \cdots \|l_m\|^{k_m/m}}{k_1!(k_2/2)! \cdots (k_m/m)!}. \tag{4.2}$$

Next, for each fixed k_m, by using Lemma 4.2 with $p = m-1$, we see that

$$\sum_{\substack{k_1, \cdots, k_{m-1} \geqslant 0 \\ k_1 + \cdots + k_{m-1} = n - k_m}} \frac{\|l_1\|^{k_1} \cdots \|l_{m-1}\|^{k_{m-1}/(m-1)}}{k_1! \cdots (k_{m-1}/(m-1))!} \leqslant (m-1)^{m-2} \cdot \frac{a^{(n-k_m)/m}}{((n-k_m)/(m-1))!},$$

where

$$a \triangleq \left(\sum_{r=1}^{m-1} \|l_r\|^{\frac{m-1}{r}} \right)^{\frac{m}{m-1}}.$$

By substituting this into (4.2), we obtain

$$\|X^n\| \leqslant (m-1)^{m-2} \sum_{k_m=0}^{n} \frac{a^{(n-k_m)/m} \|l_m\|^{k_m/m}}{(k_m/m)!((n-k_m)/(m-1))!}.$$

Now the result follows from Lemma 4.3 with $\alpha = 1/m$, $\beta = 1/(m-1)$ and $b = \|l_m\|$.

□

Remark 4.4. The upper bound given by Theorem 4.1 is sharp, which can be easily seen by considering the case when l is homogeneous (i.e. when $l \in V^{\otimes m}$).

4.2 The core of the matter: Lie algebraic developments and the lower estimate

Now we turn our attention to establishing a matching lower bound, which is the core of the present paper. The philosophy of our main strategy can be briefly summarized as follows.

Our starting point is to look at the development of paths into a space of automorphisms associated with a given representation of the tensor algebra. This enables us to obtain an intermediate lower estimate of $L_m(\mathbf{X})$ in terms of eigenvalues of the highest degree Lie component defining $\mathbf{X}$ under the given representation, and thus also allows us to eliminate the subtle contributions arising from the presence of lower degree Lie components.

The problem then reduces to designing suitable type of representations under which one can study spectral properties of Lie polynomials effectively. For this purpose, as the next key point in our approach which leads us to the main lower estimate, we design a representation that factors through a complex semisimple Lie algebra. In this way, the associated representation theory enables us to study eigenvalues of the highest degree Lie component at an explicit and quantitative level. This is largely due to the presence of an abelian subalgebra (a so-called Cartan subalgebra) consisting of semisimple elements, a basic feature of semisimple Lie algebras that is quite different from nilpotent (or more generally, solvable) Lie algebras. A crucial step towards making good use of such feature is to develop the highest degree Lie polynomials into this Cartan subalgebra.

Our plan of proving the main lower estimate is organized in the following way, which also underlines the main ingredients of our strategy.

Organization of this subsection. In Section 4.2.1, we introduce the notion of Lie algebraic developments, which is a main tool we will be using for proving our lower estimate. In Section 4.2.2, we prove an intermediate lower estimate using path developments and finite dimensional perturbation theory. Section 4.2.3 is devoted to reaching our main lower estimate from the intermediate one, and for this purpose it is further divided into four parts. Part I contains a quick review on several notions from the representation theory of complex semisimple Lie algebras that are needed in our approach. In Part II, we develop ways of mapping a space of homogeneous Lie polynomials into a Cartan subalgebra, by using basic root patterns from semisimple Lie theory. Part III is devoted to the proof of a consistency lemma for certain polynomial systems, which is a crucial ingredient in order to obtain a uniform lower estimate. In Part IV, having all necessary ingredients at hand, we give the proof of our main lower estimate by designing appropriate Lie algebraic developments. In Section 4.2.4, we perform explicit calculations in low degree cases to demonstrate how our strategy can be implemented specifically, leading to the sharp lower bound in certain situations.

4.2.1 Lie algebraic developments of rough paths

To describe the necessary structures efficiently, we start with the following definition. Given a Banach space W, we use the notation $\mathrm{End}(W)$ (respectively, $\mathrm{Aut}(W)$) to denote the space of continuous linear endomorphisms (respectively, automorphisms) over W, equipped with the operator norm.

Definition 4.5. Let V be a real or complex Banach space. A *Lie algebraic development* Φ of V consists of a linear map $F : V \to \mathfrak{g}$ into a complex Lie algebra $\mathfrak{g}$ and a representation $\rho : \mathfrak{g} \to \mathrm{End}(W)$ of $\mathfrak{g}$ on a complex Banach space W such that $\Phi = \rho \circ F$ is continuous. The development Φ is said to be *finite dimensional* if $\mathfrak{g}$ and W are both finite dimensional. In situations when the intermediate Lie algebra $\mathfrak{g}$ is not relevant, we simply refer to $\Phi : V \to \mathrm{End}(W)$ as a *development*.

Remark 4.6. When V is real, linearity is understood over $\mathbb{R}$ by regarding a complex vector space as a real vector space in the obvious way.

Let $\Phi : V \to \mathrm{End}(W)$ be a given development. According to the universal property of the projective tensor product (cf. [22], Theorem 5.6.3), for each $n \geqslant 1$, Φ induces a continuous linear map $\Phi^{(n)} : V^{\otimes n} \to \mathrm{End}(W)$ such that

$$\Phi^{(n)}(v_1 \otimes \cdots \otimes v_n) = \Phi(v_1) \cdots \Phi(v_n)$$

and

$$\|\Phi^{(n)}\|_{V^{\otimes n} \to \mathrm{End(W)}} \leqslant \|\Phi\|^n_{V \to \mathrm{End}(W)}. \tag{4.3}$$

It follows that Φ induces a natural algebra homomorphism from a subspace of $T((V))$ to $\mathrm{End}(W)$, which is defined by (still denoted as Φ)

$$\Phi\left((\xi_0, \xi_1, \xi_2, \cdots)\right) \triangleq \xi_0 \cdot \mathrm{Id} + \sum_{n=1}^{\infty} \Phi^{(n)}(\xi_n),$$

provided that the sum on the right hand side is convergent under the operator norm on $\mathrm{End}(W)$. In addition, Φ descends to a natural Lie algebra homomorphism from the free Lie algebra $\mathcal{L}(V) = \oplus_{n=1}^{\infty} \mathcal{L}_n(V)$ over V into $\mathrm{End}(W)$.

Under the given development Φ, every rough path $(\mathbf{X}_t)_{0 \leqslant t \leqslant T}$ over V can be developed onto the automorphism group $\mathrm{Aut}(W)$ by solving the linear ODE

$$\begin{cases} d\Gamma_t = \Gamma_t \cdot \Phi(d\mathbf{X}_t), \ 0 \leqslant t \leqslant T, \\ \Gamma_0 = \mathrm{Id}. \end{cases} \tag{4.4}$$

Using Picard's iteration, it is easily seen that

$$\begin{aligned} \Gamma_t &= \sum_{n=0}^{\infty} \int_{0<t_1<\cdots<t_n<t} \Phi(d\mathbf{X}_{t_1}) \cdots \Phi(d\mathbf{X}_{t_n}) \\ &= \sum_{n=0}^{\infty} \Phi^{(n)} \left(\int_{0<t_1<\cdots<t_n<t} d\mathbf{X}_{t_1} \otimes \cdots \otimes d\mathbf{X}_{t_n} \right) \\ &= \Phi(\mathbb{X}_{0,t}), \end{aligned} \tag{4.5}$$

where $\mathbb{X}_{0,t}$ is the Lyons extension of $\mathbf{X}$ given by Theorem 2.5. Note that by the factorial decay estimate in the same theorem, $\Phi(\mathbb{X}_{0,t})$ is well defined. In particular, we have $\Gamma_T = \Phi(S(\mathbf{X}))$.

Remark 4.7. In the above discussion, the intermediate Lie algebra $\mathfrak{g}$ and the complex structure appearing in Definition 4.5 are not so relevant yet, and the structure used here is simply a representation of the tensor algebra. Their roles will become clear later on when we look for explicit quantitative lower estimates for the signature.

The viewpoint of developing Euclidean paths onto a Lie group was essentially due to Cartan and had been used by many authors for geometric reasons. We give an example which is related to studies on path signatures.

Example 4.8. Hambly-Lyons [15] proved the asymptotics formula (1.3) for C^1-paths (with certain extra regularity condition) by developing the underlying path onto the space of constant negative curvature. Using the notion of Lie algebraic developments, in their case $V = \mathbb{R}^d$,

$$\mathfrak{g} = \left\{ \begin{pmatrix} A & x \\ x^T & 0 \end{pmatrix} \in \mathrm{Mat}(d+1;\mathbb{R}) : A \in \mathfrak{so}(d), x \in \mathbb{R}^d \right\}$$

is the Lie algebra of the isometry group for the standard d-dimensional hyperboloid. The embedding $F : V \to \mathfrak{g}$ is given by

$$F(x) \triangleq \begin{pmatrix} 0 & x \\ x^T & 0 \end{pmatrix}, \quad x \in \mathbb{R}^d,$$

and $\rho : \mathfrak{g} \to \mathrm{End}(W)$ is the canonical matrix representation with $W = \mathbb{R}^{d+1}$. Rather than looking at the developed path Γ_t in the isometry group, the authors worked with the trajectory on the hyperboloid traced out by the action of Γ_t on a base point of the hyperboloid. Their main philosophy, which is rather geometric, is to make use of exotic properties of hyperbolic geodesics which do not have Euclidean counterparts. Related results by Lyons-Xu [24] for studying signature inversion and by Boedihardjo-Geng [1] for studying tail asymptotics of the Brownian signature are also based on similar geometric intuitions. In this hyperbolic picture, there is no need to work with complex structure appearing in Definition 4.5.

In contrast to the hyperbolic geometric ideas, our approach deviates from the aforementioned works by not projecting the path onto a base manifold which the group acts on. Instead of following geometric intuitions, we look at path developments from an algebraic viewpoint, which provides a more suitable framework for the implementation of representation-theoretic techniques.

4.2.2 An intermediate lower estimate

Using the notion of developments, we can first establish a general lower estimate which holds for arbitrary rough paths. A similar estimate already appeared in [1] for the hyperbolic development of Brownian motion. Given any p-rough path $\mathbf{X}_t$ and $\lambda > 0$, we use $\delta_\lambda(\mathbf{X}_t)$ to denote the dilated path $(1, \lambda X_t^1, \cdots, \lambda^{\lfloor p \rfloor} X_t^{\lfloor p \rfloor})$.

Proposition 4.9. *Let $(\mathbf{X}_t)_{0 \leqslant t \leqslant T}$ be a p-rough path over some Banach space V. For any given nonzero development $\Phi : V \to \mathrm{End}(W)$, we have the following lower estimate for the* *signature tail asymptotics of* $\mathbf{X}$*:*

$$L_p(\mathbf{X}) \geqslant \limsup_{\lambda\to\infty} \frac{\log \|\Gamma_T^\lambda\|_{W\to W}}{\left(\lambda \|\Phi\|_{V\to\mathrm{End}(W)}\right)^p}, \tag{4.6}$$

where for $\lambda > 0$, $(\Gamma_t^\lambda)_{0\leqslant t\leqslant T}$ *denotes the development of the dilated path* $\delta_\lambda(\mathbf{X}_t)$ *under* Φ*, defined by the ODE (4.4).*

Proof. According to the formula (4.5) for path developments, we have

$$\Gamma_T^\lambda = \sum_{n=0}^{\infty} \lambda^n \Phi^{(n)}\left(X^n\right),$$

where X^n is the degree n component of the signature of $\mathbf{X}$. For given $N \geqslant 1$, define

$$L_N \triangleq \sup_{n\geqslant N} \left(\left(\frac{n}{p}\right)! \|X^n\|\right)^{\frac{p}{n}},$$

which is finite according to the factorial estimate (2.3). Note that if $L_N = 0$ for some N, then the right hand side of (4.6) is zero since Γ_T^λ becomes polynomial in λ in this case. Therefore, we will assume that $L_N > 0$ for all N. It then follows from (4.3) that

$$\begin{aligned}
&\|\Gamma_T^\lambda\|_{W\to W}\\
&\leqslant \sum_{n=0}^{N-1} (\lambda\|\Phi\|)^n \|X^n\| + \sum_{n=N}^{\infty} (\lambda\|\Phi\|)^n \|X^n\|\\
&\leqslant \sum_{n=0}^{N-1} (\lambda\|\Phi\|)^n \|X^n\| + \sum_{n=N}^{\infty} \frac{(\lambda^p\|\Phi\|^p L_N)^{n/p}}{(n/p)!}\\
&= \sum_{n=0}^{\infty} \frac{(\lambda^p\|\Phi\|^p L_N)^{n/p}}{(n/p)!} + \sum_{n=0}^{N-1} \left((\lambda\|\Phi\|)^n \|X^n\| - \frac{(\lambda^p\|\Phi\|^p L_N)^{n/p}}{(n/p)!}\right),
\end{aligned} \tag{4.7}$$

where for notational simplicity we have omitted the subscript for the operator norm of Φ.

To understand the asymptotic behaviour of the right hand side as $\lambda \to \infty$, we first consider the explicit function defined by

$$f(x) \triangleq \sum_{n=0}^{\infty} \frac{x^{n/p}}{(n/p)!}.$$

We claim that

$$f(x) \leqslant (p+1)x\mathrm{e}^x \quad \text{for all } x \geqslant 1. \tag{4.8}$$

Indeed, for each $m \geqslant 0$, define R_m to be the set of real numbers $r \in [0, p)$ such that $mp + r$ is an integer. Then $R_m \subseteq [0, p)$ consists of no more than $p + 1$ elements. Therefore,

$$f(x) = \sum_{m=0}^{\infty} \sum_{r\in R_m} \frac{x^{m+r/p}}{(m+r/p)!} \leqslant \sum_{m=0}^{\infty} \frac{x^m}{m!} \sum_{r\in R_m} x^{r/p} \leqslant (p+1)x \sum_{m=0}^{\infty} \frac{x^m}{m!} = (p+1)x\mathrm{e}^x.$$

By applying (4.8) to the first term on the right hand side of (4.7) and denoting the second term as $q_N(\lambda)$, we obtain that

$$\|\Gamma_T^\lambda\|_{W\to W} \leqslant (p+1)\lambda^p\|\Phi\|^p L_N \exp\left(\lambda^p\|\Phi\|^p L_N\right) + q_N(\lambda).$$

Note that $q_N(\lambda)$ has polynomial growth in λ. Therefore, by taking $\lambda\to\infty$, we have

$$\limsup_{\lambda\to\infty}\frac{\log\|\Gamma_T^\lambda\|_{W\to W}}{(\lambda\|\Phi\|)^p} \leqslant L_N.$$

Since N is arbitrary, we conclude that

$$\limsup_{\lambda\to\infty}\frac{\log\|\Gamma_T^\lambda\|_{W\to W}}{(\lambda\|\Phi\|)^p} \leqslant \inf_{N\geqslant 1} L_N = L_p(\mathbf{X}).$$

$\square$

At first glance, the estimate (4.6) does not seem to be as useful as it will be. We now unwind the shape of its right hand side in the context of pure rough paths. This leads us to the following intermediate lower estimate. From now on, we confine ourselves in finite dimensional developments, which is the main situation where useful calculations can be done explicitly.

Theorem 4.10. *Let $\mathbf{X}_t = \exp(tl)$ be a pure m-rough path with $l \in \mathcal{L}^{(m)}(V)$. For any given finite dimensional development $\Phi : V \to \mathrm{End}(W)$, we have*

$$L_m(\mathbf{X}) \geqslant \frac{\sup\left\{\mathrm{Re}(\mu) : \mu\in\sigma(\Phi(l_m))\right\}}{\|\Phi\|^m_{V\to\mathrm{End}(W)}}, \tag{4.9}$$

where $l_m \triangleq \pi_m(l)$ is the highest degree component of l, and $\sigma(\Phi(l_m))$ denotes the set of eigenvalues of $\Phi(l_m)\in\mathrm{End}(W)$.

Proof. The proof is an application of perturbation theory in finite dimensions. Let μ be an eigenvalue of $\Phi(l_m)$ and write $l = l_1 + \cdots + l_m$ as the sum of homogeneous components. According to [20], Chapter 2, Theorem 5.1 and Theorem 5.2 applied to the continuous family

$$(0,\infty)\ni\lambda\mapsto T(\lambda)\triangleq\Phi(l_m)+\frac{1}{\lambda}\Phi(l_{m-1})+\cdots+\frac{1}{\lambda^{m-1}}\Phi(l_1)\in\mathrm{End}(W)$$

of bounded linear transformations, we know that there exists a complex valued continuous function $\phi(\lambda)$, such that $\phi(\lambda)$ is an eigenvalue of $T(\lambda)$ for all λ and $\phi(\lambda)\to\mu$ as $\lambda\to\infty$. On the other hand, let $(\Gamma_t^\lambda)_{0\leqslant t\leqslant 1}$ be the development of the dilated path $\delta_\lambda(\mathbf{X}_t)$ under Φ. By (4.5) and the definition of operator norm, we have

$$\begin{aligned}\|\Gamma_1^\lambda\|_{W\to W} &= \|\Phi\left(\delta_\lambda(\exp(l))\right)\|_{W\to W} = \|\exp(\Phi(\delta_\lambda(l)))\|_{W\to W}\\ &= \|\exp(\lambda^m T(\lambda))\|_{W\to W} \geqslant \left|\exp\left(\lambda^m\phi(\lambda)\right)\right| = \exp\left(\lambda^m\mathrm{Re}(\phi(\lambda))\right).\end{aligned}$$

Therefore,

$$\frac{\log\|\Gamma_1^\lambda\|_{W\to W}}{\lambda^m} \geqslant \mathrm{Re}(\phi(\lambda))$$

for all $\lambda>0$. Now the result follows from Proposition 4.9 by taking $\lambda\to\infty$. $\square$

Remark 4.11. Note that the right hand side of (4.9) does not depend on lower degree components of l. In other words, Theorem 4.10 provides a possible way of eliminating the complicated interactions among different degree components of l in the signature tail asymptotics. Nonetheless, it is *not* true that this fact allows us to conclude Conjecture 2.12 directly from the homogeneous case (i.e. when $l = l_m$) for which we know the result holds trivially (cf. (3.1) in Section 3). The subtle point is that, as suggested by (4.9), the elimination of lower degree effects is only achieved through a given development Φ. Therefore, even though we know the result holds for the homogeneous case, one needs to see that the lower bound can be attained at some specific choice of developments in order to conclude the result for the inhomogeneous case. Designing such developments is the main goal of what follows.

4.2.3 The main lower estimate

In view of Theorem 4.10, to obtain useful lower bounds on $L_m(\mathbf{X})$, one needs to design suitable Lie algebraic developments under which we can estimate eigenvalues of l_m effectively. This is where the intermediate Lie algebra $\mathfrak{g}$ and the complex structure in Definition 4.5 come into play. In particular, we will choose $\mathfrak{g}$ to be a finite dimensional complex semisimple Lie algebra and rely on the associated representation theory.

I. Notions from the representation theory of complex semisimple Lie algebras

To explain how the semisimple structure plays a role, it is helpful to first recall some relevant notions from Lie theory. We refer the reader to [18] for more details. Unless otherwise stated, all Lie algebras and representations are finite dimensional over the complex field. The main benefit of this setting is the existence of eigenvalues for linear transformations, which significantly simplifies the associated representation theory.

Definition 4.12. A complex Lie algebra $\mathfrak{g}$ is called *semisimple* if it is isomorphic to a direct sum $\mathfrak{g} \cong \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$ of Lie algebras, where each summand $\mathfrak{g}_i$ is *simple* in the sense that it does not contain non-trivial proper ideals.

It can be shown that semisimpleness is equivalent to the non-degeneracy of the *Killing form*, which is the bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ defined by

$$B(X, Y) \triangleq \mathrm{Tr}\left(\mathrm{ad}(X) \circ \mathrm{ad}(Y)\right),$$

where Tr means taking trace and $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ denotes the *adjoint representation* of $\mathfrak{g}$ given by $\mathrm{ad}(X)(Z) \triangleq [X, Z]$.

A central concept in semisimple Lie theory that is also crucial for us is the following.

Definition 4.13. A *Cartan subalgebra* of $\mathfrak{g}$ is a subspace $\mathfrak{h} \subseteq \mathfrak{g}$ such that:

(i) $\mathfrak{h}$ is a maximal abelian subalgebra of $\mathfrak{g}$;
(ii) for each $H \in \mathfrak{h}$, the linear transformation $\mathrm{ad}(H) \in \mathrm{End}(\mathfrak{g})$ is semisimple (over $\mathbb{C}$ this is equivalent to being diagonalizable).

For a complex semisimple Lie algebra, a Cartan subalgebra always exists and it is unique up to conjugation in $\mathfrak{g}$. Let $\mathfrak{h}$ be a given Cartan subalgebra of $\mathfrak{g}$. By its definition and a standard application of linear algebra, given an arbitrary representation $\rho : \mathfrak{g} \to \mathrm{End}(W)$, all elements of $\mathfrak{h}$ are simultaneously diagonalizable when viewed as linear transformations over W under ρ. More specifically, a complex linear functional $\mu \in \mathfrak{h}^*$ is called a *weight* for the given representation ρ if the subspace

$$W^{\mu} \triangleq \{w \in W : \rho(H)w = \mu(H)w \text{ for all } H \in \mathfrak{h}\} \tag{4.10}$$

is non-trivial. It follows that there are at most finitely many weights for ρ. Denote their collection by $\Pi(\rho)$. The space W then admits the decomposition (simultaneous diagonalization) $W = \oplus_{\mu \in \Pi(\rho)} W^{\mu}$, in which for every $H \in \mathfrak{h}$, W^{μ} is an eigenspace of $\rho(H)$ with eigenvalue $\mu(H)$ $(\mu \in \Pi(\rho))$.

Indeed, much more can be said in the semisimple setting. We first look at the adjoint representation of $\mathfrak{g}$. Given a complex linear functional $\alpha \in \mathfrak{h}^*$, define the subspace

$$\mathfrak{g}^{\alpha} \triangleq \{X \in \mathfrak{g} : \mathrm{ad}(H)(X) = \alpha(H)X \text{ for all } H \in \mathfrak{h}\}$$

in the same way as (4.10). It is easy to verify that $\mathfrak{g}^0 = \mathfrak{h}$, and $[\mathfrak{g}^{\alpha}, \mathfrak{g}^{\beta}] \subseteq \mathfrak{g}^{\alpha+\beta}$ for all $\alpha, \beta \in \mathfrak{h}^*$. A complex linear functional $\alpha \in \mathfrak{h}^*$ is called a *root* of $\mathfrak{g}$ with respect to $\mathfrak{h}$ if it is a weight for the adjoint representation, i.e. if $\mathfrak{g}^{\alpha} \neq \{0\}$. In this case, $\mathfrak{g}^{\alpha}$ is called the *root space* associated with the root α. As before, there are at most finitely many roots. A basic result in semisimple Lie theory is the following so-called *root space decomposition*.

Theorem 4.14. *Let $\Delta \subseteq \mathfrak{h}^*$ be the set of nonzero roots with respect to a given Cartan subalgebra $\mathfrak{h}$. Then $\mathfrak{g}$ can be written as the direct sum*

$$\mathfrak{g} = \mathfrak{h} + \sum_{\alpha \in \Delta} \mathfrak{g}^{\alpha}.$$

In addition, $\dim \mathfrak{g}^{\alpha} = 1$ *for each $\alpha \in \Delta$, and if α, β are two roots with $\alpha + \beta \in \Delta$, then*

$$[\mathfrak{g}^{\alpha}, \mathfrak{g}^{\beta}] = \mathfrak{g}^{\alpha+\beta}. \tag{4.11}$$

It is possible to study general representations using the structure of roots. Before stating relevant results, we need a few more definitions. Let E be the vector space generated by Δ over $\mathbb{R}$. A subset Δ_0 of Δ is called a *base* if:

(i) Δ_0 is a basis of E;
(ii) each root $\beta \in \Delta$ can be expressed as $\beta = \sum_{\alpha \in \Delta_0} k_{\alpha}\alpha$ with integral coefficients k_{α} either being all non-negative or all non-positive.

The roots in Δ_0 are called *simple roots*. The choice of Δ_0 is not unique but its cardinality is. The Lie algebra $\mathfrak{g}$ is said to have *rank* m if Δ_0 has m elements, which is equivalent to saying that $\dim_{\mathbb{C}} \mathfrak{h} = m$. Let $\Delta_0 = \{\alpha_1, \cdots, \alpha_m\}$ be a given set of simple roots. The Killing form B restricted to $\mathfrak{h}$ is also non-degenerate. It follows that, for each $\alpha_i \in \Delta_0$, there exists $T_i \in \mathfrak{h}$ such that $\alpha_i(H) = B(T_i, H)$ for all $H \in \mathfrak{h}$. We define the normalized element $H_i \triangleq 2T_i/B(T_i, T_i)$.

Definition 4.15. A linear functional $\lambda \in \mathfrak{h}^*$ is called a *dominant integral functional* if all $\lambda(H_i)$ $(1 \leqslant i \leqslant m)$ are non-negative integers. The set $\{\lambda_1, \cdots, \lambda_m\}$ of *fundamental dominant integral functionals* are defined by the duality relation $\lambda_i(H_j) = \delta_{ij}$.

The main result in the representation theory of complex semisimple Lie algebras is stated as follows. Recall that a representation $\rho : \mathfrak{g} \to \mathrm{End}(W)$ is *irreducible* if W does not contain non-trivial proper $\mathfrak{g}$-invariant subspaces.

Theorem 4.16. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra with a given Cartan subalgebra $\mathfrak{h}$ and a given base Δ_0 of simple roots. There is a one-to-one correspondence between dominant integral functionals and isomorphism classes of (finite dimensional) irreducible representations.*

We must point out that representation theory provides richer quantitative information than the statement of the above classification theorem itself. A main consequence of the theory which is relevant to us, is that for each dominant integral functional λ, the set of weights for the associated irreducible representation can be described in a rather quantitative way, making the computation of eigenvalues of elements in $\mathfrak{h}$ quite tractable. We use an important example to illustrate this point, in which all the aforementioned notions and results can also be worked out explicitly. The implementation of our main technique is largely based on this example.

Consider $\mathfrak{g} = \mathfrak{sl}(m, \mathbb{C})$ $(m \geqslant 2)$, the set of $m \times m$ complex matrices with zero trace. Then $\mathfrak{g}$ is a complex semisimple (in fact, simple) Lie algebra of rank $m-1$. A Cartan subalgebra $\mathfrak{h}$ can be chosen as the subspace of diagonal matrices in $\mathfrak{g}$. For each $1 \leqslant i \leqslant m$, define $\mu_i \in \mathfrak{h}^*$ to be linear functional of taking the i-th diagonal entry. Then the set of nonzero roots with respect to $\mathfrak{h}$ is given by

$$\Delta = \left\{ \alpha_{i,j} \triangleq \mu_i - \mu_j : 1 \leqslant i \neq j \leqslant m \right\}.$$

In addition, for each $i \neq j$, the root space $\mathfrak{g}^{\alpha_{i,j}} = \mathbb{C} \cdot E_{i,j}$, where $E_{i,j}$ is the matrix whose (i, j)-entry is 1 and all other entries are 0's. To summarize, the root space decomposition takes the form

$$\mathfrak{g} = \mathfrak{h} + \sum_{1 \leqslant i \neq j \leqslant m} \mathbb{C} \cdot E_{i,j}.$$

A base of simple roots can be chosen as

$$\Delta_0 = \left\{ \alpha_i \triangleq \mu_i - \mu_{i+1} : 1 \leqslant i \leqslant m-1 \right\}. \tag{4.12}$$

For each simple root $\alpha_i \in \Delta_0$, the associated $H_i \in \mathfrak{h}$ is given by the diagonal matrix in which the i-th diagonal entry is 1, the $(i+1)$-th diagonal entry is -1, and all other entries are zero. The representation theory of $\mathfrak{g} = \mathfrak{sl}(m, \mathbb{C})$ can be summarized as the following theorem. Note that $\mathfrak{g}$ acts on $W \triangleq \mathbb{C}^m$ in the canonical way by matrix multiplication. We call this canonical matrix representation ρ. For each $k \geqslant 1$, ρ induces a representation $\rho^{\otimes k}$ (respectively, $\rho^{\wedge k}$) on $W^{\otimes k}$ (respectively, on $\Lambda^k(W)$, the k-th exterior power of W) in the natural way. For $1 \leqslant k \leqslant m-1$, denote $W_k \triangleq \Lambda^k(W)$.

Theorem 4.17. *Choose a Cartan subalgebra and a base of simple roots for $\mathfrak{g} = \mathfrak{sl}(m, \mathbb{C})$ as before.*

(1) The set $\{\lambda_1, \cdots, \lambda_{m-1}\}$ of fundamental dominant integral functionals are given by $\lambda_k = \mu_1 + \cdots + \mu_k$ $(1 \leqslant k \leqslant m-1)$. For each k, the irreducible representation associated with λ_k is given by $\rho^{\wedge k} : \mathfrak{g} \to \mathrm{End}(W_k)$, whose set of weights is precisely

$$\Pi(\lambda_k) = \{\mu_{i_1} + \cdots + \mu_{i_k} : 1 \leqslant i_1 < \cdots < i_k \leqslant m\}.$$

(2) For each dominant integral functional $\lambda = a_1\lambda_1 + \cdots + a_{m-1}\lambda_{m-1}$ with a_i's being non-negative integers, the representation $\rho^\lambda : \mathfrak{g} \to \mathrm{End}(W_1^{\otimes a_1} \otimes \cdots \otimes W_{m-1}^{\otimes a_{m-1}})$ contains exactly one copy of the irreducible representation associated with λ, whose set of weights is a subset of

$$\left\{ \sum_{k=1}^{m-1} \sum_{j=1}^{a_k} \nu_{k,j} : \nu_{k,j} \in \Pi(\lambda_k) \right\}.$$

Remark 4.18. In the second part of the above theorem, by using Schur polynomials and Young tableaux, it is possible to identify the precise copy of irreducible representation contained in the tensor product representation as well as the associated set of weights. However, at the moment we do not see the need of pursuing this generality.

To conclude this part, we mention as an example that the adjoint representation of $\mathfrak{sl}(m, \mathbb{C})$ is the irreducible representation associated with the dominant integral functional $\lambda_1 + \lambda_{m-1} = \mu_1 - \mu_m$.

II. An essential step: developing the highest degree Lie component into a Cartan subalgebra

Returning to our signature problem, let $\mathbf{X}_t = \exp(tl)$ be a pure m-rough path, where $l \in \mathcal{L}^{(m)}(V)$ whose the highest degree component is denoted by l_m. An essential step in our approach, is to choose $\mathfrak{g}$ to be a finite dimensional complex semisimple Lie algebra in the Lie algebraic development, together with a linear embedding $F : V \to \mathfrak{g}$ such that the space $\mathcal{L}_m(V)$ of homogeneous Lie polynomials of degree m is mapped into a Cartan subalgebra of $\mathfrak{g}$ under the induced homomorphism on the free Lie algebra $\mathcal{L}(V)$. In this way, according to Theorem 4.10, we are immediately led to the lower estimate

$$L_m(\mathbf{X}) \geqslant \frac{\sup \{\mathrm{Re}(\mu(F(l_m))) : \mu \in \Pi(\rho)\}}{\|\Phi\|^m_{V \to \mathrm{End}(W)}} \tag{4.13}$$

under the given Lie algebraic development $\Phi = \rho \circ F$, where recall that $\Pi(\rho) \subseteq \mathfrak{h}^*$ is the set of weights for the representation ρ. Representation theory then provides tractable methods of computing weights for given representations, hence leading us to more explicit lower bounds on $L_m(\mathbf{X})$.

The simplest way of mapping $\mathcal{L}_m(V)$ into a Cartan subalgebra is through the Lie algebra $\mathfrak{sl}(m, \mathbb{C})$, which can be seen by straight forward matrix calculation. However, the essential reason behind such calculation is hidden in the root pattern as stated in the following lemma. Working with root patterns also allows one to identify other semisimple Lie algebras which are not isomorphic to $\mathfrak{sl}(m;\mathbb{C})$ but achieve the same property (cf. Example 4.21 and Example 4.22 below).

Lemma 4.19. *Suppose that there exist $m-1$ nonzero roots $\alpha_1,\cdots,\alpha_{m-1}$ with respect to $\mathfrak{h}$, such that all nonzero roots one can construct from them as integral linear combinations are precisely of the form $\pm(\alpha_i+\alpha_{i+1}+\cdots+\alpha_j)$ with $1\leqslant i\leqslant j\leqslant m-1$. Define the subspace*

$$E \triangleq \mathfrak{g}^{\alpha_1}\oplus\cdots\oplus\mathfrak{g}^{\alpha_{m-1}}\oplus\mathfrak{g}^{-(\alpha_1+\cdots+\alpha_{m-1})}. \tag{4.14}$$

Then

$$E^{(m-1)} \triangleq \underbrace{[\cdots[[[E,E],E],E]\cdots]}_{m-1 \text{ brackets}} \subseteq \mathfrak{h}.$$

Proof. For each $1\leqslant i\leqslant m-1$ and $1\leqslant j\leqslant m-i$, define

$$\alpha_{i;j} \triangleq \alpha_i+\alpha_{i+1}+\cdots+\alpha_{i+j-1}.$$

According to the assumption, the $\alpha_{i;j}$'s are precisely all the nonzero roots one can build from $\alpha_1,\cdots,\alpha_{m-1}$ as integral linear combinations. It follows from the graded property (4.11) of root spaces that

$$\begin{aligned}
E^{(1)} &= \left(\sum_{i=1}^{m-2}\mathfrak{g}^{\alpha_{i;2}}\right)\oplus\mathfrak{g}^{-\alpha_{1;m-2}}\oplus\mathfrak{g}^{-\alpha_{2;m-2}},\\
&\cdots\\
E^{(k)} &= \left(\sum_{i=1}^{m-1-k}\mathfrak{g}^{\alpha_{i;k+1}}\right)\oplus\left(\sum_{j=1}^{k+1}\mathfrak{g}^{-\alpha_{j;m-1-k}}\right),\\
&\cdots\\
E^{(m-2)} &= \mathfrak{g}^{\alpha_1+\cdots+\alpha_{m-1}}\oplus\left(\sum_{j=1}^{m-1}\mathfrak{g}^{-\alpha_j}\right).
\end{aligned}$$

Finally, by using property (4.11) again as well as the assumption of the lemma, we obtain

$$E^{(m-1)} = [E^{(m-2)},E]\subseteq\mathfrak{g}^0=\mathfrak{h}.$$

□

Lemma 4.19 tells us that, if we design $F:V\to\mathfrak{g}$ so that $F(V)\subseteq E$, then under the induced homomorphism on the free Lie algebra, $\mathcal{L}_m(V)$ is mapped into the Cartan subalgebra $\mathfrak{h}$.

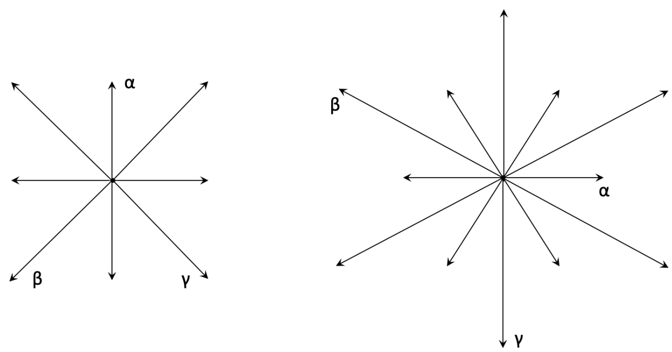

Figure 4.1: Root Systems of $\mathfrak{so}(5,\mathbb{C})$ and $\mathfrak{g}_2$

Example 4.20. Consider $\mathfrak{g} = \mathfrak{sl}(m,\mathbb{C})$, with a Cartan subalgebra $\mathfrak{h}$ given by the subspace of diagonal matrices in $\mathfrak{g}$. In this case, it is easy to see that the simple roots $\alpha_i = \mu_i - \mu_{i+1}$ $(1 \leqslant i \leqslant m-1)$ given by (4.12) satisfy the assumption of Lemma 4.19. In this case, we have

$$E = \left\{ \begin{pmatrix} 0 & z_1 & & 0 \\ & & \ddots & \\ & & & z_{m-1} \\ z_m & & & 0 \end{pmatrix} : z_1, \cdots, z_m \in \mathbb{C} \right\},$$

where omitted entries in the matrix are all 0's. Indeed, the semisimple Lie algebra associated with the root system generated by the roots given in Lemma 4.19 is isomorphic to $\mathfrak{sl}(m,\mathbb{C})$.

Using root patterns, we give two other examples that are not isomorphic to $\mathfrak{sl}(m,\mathbb{C})$ but also allow one to map the highest degree Lie polynomials into a Cartan subalgebra. In each example, the underlying Lie algebra is of rank two. The nonzero roots are drawn as planar Euclidean vectors, in which integral linear combinations follow usual vector operation rules. The corresponding conclusion is immediate by manipulating the root vectors based on the graded property (4.11) and $\mathfrak{g}^0 = \mathfrak{h}$. Although possible, there is no need to work with the actual Lie algebra $\mathfrak{g}$ and the associated root spaces at this level.

Example 4.21. Consider $\mathfrak{g} = \mathfrak{so}(5,\mathbb{C})$, the Lie algebra of 5×5 complex skew-symmetric matrices. The associated root system is given by the left hand side of Figure 4.1. If we require $F : V \to E \triangleq \mathfrak{g}^\alpha \oplus \mathfrak{g}^\beta \oplus \mathfrak{g}^\gamma$, then $\mathcal{L}_4(V)$ is mapped into a Cartan subalgebra (cf. Section 4.2.4 II below for more explicit calculations in degree $m = 4$ based on this structure). The property can be generalized to higher degrees by considering $\mathfrak{so}(n,\mathbb{C})$ with larger n.

Example 4.22. Consider $\mathfrak{g} = \mathfrak{g}_2$, the smallest exceptional simple Lie algebra. It arises from the classification of simple Lie algebras, and can be identified as the Lie algebra of the subgroup of Spin(7) preserving a point on S^7. The associated root system is given by the right hand side of Figure 4.1. If we require $F : V \to E \triangleq \mathfrak{g}^\alpha \oplus \mathfrak{g}^\beta \oplus \mathfrak{g}^\gamma$, then $\mathcal{L}_6(V)$ is mapped into a Cartan subalgebra.

III. A consistency lemma for certain symmetric polynomial systems

Note that the homogeneous Lie polynomial $l_m \in \mathcal{L}_m(V)$ has the general form $l_m = c_1 h_1 + \cdots + c_\nu h_\nu$, where $\{h_1, \cdots, h_\nu\}$ is a given basis of $\mathcal{L}_m(V)$. In order to produce a lower bound on $L_m(\mathbf{X})$ in the form of Theorem 2.13, with a factor *independent of the coefficients* c_i's, a natural idea is to require each h_i to have the right eigenvalue individually. In this way, properties of Cartan subalgebra will guarantee that l_m has a desired eigenvalue $\|l_m\|$ and the operator norm of the Lie algebraic development will depend only on the roughness m but not on the coefficients c_i. This viewpoint leads us to the consideration of certain type of polynomial systems. A consistency lemma for these systems, stated as follows, will be needed for the proof of our main lower estimate. The lemma may also be of independent interest.

Lemma 4.23. *Let $p_1, \cdots, p_\nu$ be homogeneous polynomials over $\mathbb{C}^n$ of the same degree. Suppose that they are linearly independent. Then for any $k \geqslant 4^{\nu-1}\nu!$ and $c_1, \cdots, c_\nu \in \mathbb{C}$, the polynomial system*

$$\begin{cases} p_1(\mathbf{z}_1) + \cdots + p_1(\mathbf{z}_k) = c_1, \\ \cdots \\ p_\nu(\mathbf{z}_1) + \cdots + p_\nu(\mathbf{z}_k) = c_\nu \end{cases}$$

has at least one solution in $\mathbb{C}^{kn}$, where $\mathbf{z}_1, \cdots, \mathbf{z}_k$ are independent variables each having dimension n.

Remark 4.24. Lemma 4.23 is not as obvious as one may expect and the special structure of the system has to play an essential role. In general, a polynomial system in which the number of variables is greater than the number of equations may not always possess a solution, even when assuming that the underlying polynomials are algebraically independent. For instance, the system

$$x^2 y = 0, \ xyz = 1$$

does not have a solution! The precise level of independence is given by the renowned Hilbert's nullstellensatz in algebraic geometry. It is to some extent surprising that linear independence is sufficient for the assertion to hold in our case.

Proof of Lemma 4.23. [1]We are going to prove the claim that there exists $k \geqslant 1$ such that the system is consistent. It will be clear in the argument that $k \geqslant 4^{\nu-1}\nu!$. For this purpose, we treat the claim as a property depending on ν (the number of polynomials involved) and prove it by induction. When $\nu = 1$, since $p_1 \neq 0$ we know that $p_1(\mathbf{z}) \neq 0$ for some $\mathbf{z} \in \mathbb{C}^n$. Since p_1 is homogeneous, it follows from scaling that the image of p_1 must be $\mathbb{C}$. Therefore, the assertion holds with $k = 1$.

Suppose that the assertion is true for ν polynomials, and assume that we are now given $\nu + 1$ linearly independent homogeneous polynomials $p_1, \cdots, p_{\nu+1}$ of the same degree. By

[1]From the algebraic geometric viewpoint, it is not obvious how one can approach by using a general dimension argument, since in the associated projective space one needs to rule out the possibility that the underlying projective variety lies in the hyperplane at infinity. The proof we give here is entirely elementary.

induction hypothesis, there exists $k \geqslant 1$, such that for any $1 \leqslant i \leqslant \nu+1$, the map $\mathbf{p}_{\hat{i}}^{(k)}$: $\mathbb{C}^{kn} \to \mathbb{C}^{\nu}$ defined by

$$\mathbf{p}_{\hat{i}}^{(k)}(\mathbf{z}_1, \cdots, \mathbf{z}_k) \triangleq \begin{pmatrix} p_1(\mathbf{z}_1) + \cdots + p_1(\mathbf{z}_k) \\ \vdots \\ p_{i-1}(\mathbf{z}_1) + \cdots + p_{i-1}(\mathbf{z}_k) \\ p_{i+1}(\mathbf{z}_1) + \cdots + p_{i+1}(\mathbf{z}_k) \\ \vdots \\ p_{\nu+1}(\mathbf{z}_1) + \cdots + p_{\nu+1}(\mathbf{z}_k) \end{pmatrix}$$

is surjective. We claim that, for every $1 \leqslant i \leqslant \nu+1$, the following system

$$\begin{cases} p_i(\mathbf{z}_1) + \cdots + p_i(\mathbf{z}_{4k}) \neq 0, \\ p_j(\mathbf{z}_1) + \cdots + p_j(\mathbf{z}_{4k}) = 0, \quad \text{for all } j \neq i, \end{cases} \tag{4.15}$$

must have a solution. Observe that if this is true, then the induction step finishes with $k' = 4k(\nu+1)$. Indeed, let $c_1, \cdots, c_{\nu+1} \in \mathbb{C}$. For each i, by homogenuity and scaling, the consistency of the system (4.15) implies the consistency of the system

$$\begin{cases} p_1(\mathbf{z}_1) + \cdots + p_1(\mathbf{z}_{4k}) = 0, \\ \cdots \\ p_i(\mathbf{z}_1) + \cdots + p_i(\mathbf{z}_{4k}) = c_i, \\ \cdots \\ p_{\nu+1}(\mathbf{z}_1) + \cdots + p_{\nu+1}(\mathbf{z}_{4k}) = 0. \end{cases}$$

Let $\mathbf{Z}^{(i)} \in \mathbb{C}^{4kn}$ be a solution to the above system. By adding up the $\nu+1$ cases, we know that the system

$$\begin{cases} p_1(\mathbf{z}_1) + \cdots + p_1(\mathbf{z}_{4k(\nu+1)}) = c_1, \\ \cdots \\ p_{\nu+1}(\mathbf{z}_1) + \cdots + p_{\nu+1}(\mathbf{z}_{4k(\nu+1)}) = c_{\nu+1}, \end{cases}$$

has a solution given by $\mathbf{Z} = (\mathbf{Z}^{(1)}, \cdots, \mathbf{Z}^{(\nu+1)}) \in \mathbb{C}^{4kn(\nu+1)}$. In other words, the assertion holds with $k' = 4k(\nu+1)$.

Now it remains to show the consistency of the system (4.15). Suppose on the contrary that the system is inconsistent for some i. Without loss of generality, we may assume that $i = 1$. We first introduce some notation to simplify the presentation. It is convenient to call

$$\mathbf{Z} = (\mathbf{z}_1, \cdots, \mathbf{z}_k), \ \mathbf{Z}' = (\mathbf{z}_{k+1}, \cdots, \mathbf{z}_{2k})$$

and

$$\mathbf{W} = (\mathbf{z}_{2k+1}, \cdots, \mathbf{z}_{3k}), \ \mathbf{W}' = (\mathbf{z}_{3k+1}, \cdots, \mathbf{z}_{4k}).$$

We also define

$$P_i(\mathbf{Z}) = p_i(\mathbf{z}_1) + \cdots + p_i(\mathbf{z}_k)$$

and similarly for the other parts of the variables. In particular, we have

$$p_i(\mathbf{z}_1)+\cdots+p_i(\mathbf{z}_{4k})=P_i(\mathbf{Z})+P_i(\mathbf{Z}')+P_i(\mathbf{W})+P_i(\mathbf{W}').$$

Under the above notation and assumption, we know that $P_1(\mathbf{Z})+P_1(\mathbf{Z}')+P_1(\mathbf{W})+P_1(\mathbf{W}')$ vanishes identically on the algebraic variety

$$\mathcal{V}\triangleq\{(\mathbf{Z},\mathbf{Z}',\mathbf{W},\mathbf{W}'):P_i(\mathbf{Z})+P_i(\mathbf{Z}')+P_i(\mathbf{W})+P_i(\mathbf{W}')=0 \text{ for } 2\leqslant i\leqslant \nu+1\}$$

defined by the remaining polynomials.

We claim that, there exists a function $F:\mathbb{C}^\nu\to\mathbb{C}$, such that

$$P_1(\mathbf{W})+P_1(\mathbf{W}')=F\left(P_2(\mathbf{W})+P_2(\mathbf{W}'),\cdots,P_{\nu+1}(\mathbf{W})+P_{\nu+1}(\mathbf{W}')\right)\tag{4.16}$$

for all $(\mathbf{W},\mathbf{W}')\in\mathbb{C}^{2kn}$. Indeed, define $\Xi:\mathbb{C}^{2kn}\to\mathbb{C}^\nu$ by

$$\Xi(\mathbf{W},\mathbf{W}')\triangleq(P_2(\mathbf{W})+P_2(\mathbf{W}'),\ldots,P_{\nu+1}(\mathbf{W})+P_{\nu+1}(\mathbf{W}')).$$

By the induction hypothesis, we know that Ξ is surjective. We then define $F:\mathbb{C}^\nu\to\mathbb{C}$ by

$$F(\xi)\triangleq P_1(\mathbf{W})+P_1(\mathbf{W}'),$$

where $(\mathbf{W},\mathbf{W}')$ is any element such that $\xi=\Xi(\mathbf{W},\mathbf{W}')$. To verify that F is well defined, suppose that $\xi=\Xi(\mathbf{W},\mathbf{W}')=\Xi(\tilde{\mathbf{W}},\tilde{\mathbf{W}}')$. Then

$$P_j(\mathbf{W})+P_j(\mathbf{W}')=P_j(\tilde{\mathbf{W}})+P_j(\tilde{\mathbf{W}}'),\quad \text{for all } 2\leqslant j\leqslant\nu+1.$$

Let

$$(\mathbf{Z},\mathbf{Z}')\triangleq(-1)^{1/m}\cdot(\mathbf{W},\mathbf{W}'),$$

where m is the degree of the underlying polynomials. It follows that both of $(\mathbf{Z},\mathbf{Z}',\mathbf{W},\mathbf{W}')$ and $(\mathbf{Z},\mathbf{Z}',\tilde{\mathbf{W}},\tilde{\mathbf{W}}')$ are elements in $\mathcal{V}$, and thus they are both zeros of the polynomial at $i=1$. In particular, we have

$$P_1(\mathbf{W})+P_1(\mathbf{W}')=P_1(\tilde{\mathbf{W}})+P_1(\tilde{\mathbf{W}}'),$$

showing that F is well defined.

By taking $\mathbf{W}'=0$ in (4.16), we arrive at

$$P_1(\mathbf{W})=F(P_2(\mathbf{W}),\cdots P_{\nu+1}(\mathbf{W})).$$

Now the key observation is that, F must be linear. Indeed, given $\lambda\in\mathbb{C}$, we have

$$\begin{aligned}&\lambda F\left(P_2(\mathbf{W}),\cdots,P_{\nu+1}(\mathbf{W})\right)\\&=\lambda P_1(\mathbf{W})\\&=P_1(\lambda^{1/m}\mathbf{W})\\&=F\left(P_2(\lambda^{1/m}\mathbf{W}),\cdots,P_{\nu+1}(\lambda^{1/m}\mathbf{W})\right)\\&=F\left(\lambda P_2(\mathbf{W}),\cdots,\lambda P_2(\mathbf{W})\right).\end{aligned}$$

In addition, let $\xi, \eta \in \mathbb{C}^{\nu}$. Again by induction hypothesis, there exist $\mathbf{W}$ and $\mathbf{W}'$ in $\mathbb{C}^{kn}$, such that

$$\xi = (P_2(\mathbf{W}), \cdots, P_{\nu+1}(\mathbf{W}))\,,\ \eta = (P_2(\mathbf{W}'), \cdots, P_{\nu+1}(\mathbf{W}'))\,.$$

It follows that

$$\begin{aligned} F(\xi+\eta) &= F\left(P_2(\mathbf{W}) + P_2(\mathbf{W}'), \cdots, P_{\nu+1}(\mathbf{W}) + P_{\nu+1}(\mathbf{W}')\right) \\ &= P_1(\mathbf{W}) + P_1(\mathbf{W}') \\ &= F(\xi) + F(\eta). \end{aligned}$$

Therefore, F is linear. This leads to a contradiction with the linear independence among $P_1, \cdots, P_{\nu+1}$. Consequently, the system (4.15) is consistent, finishing the proof of the induction step. $\square$

IV. Establishing the main lower estimate

Using the representation theory of $\mathfrak{sl}(n;\mathbb{C})$ and Lemma 4.23, we can now establish our main lower estimate for the signature tail asymptotics of pure rough paths. The result contains two parts, a general lower estimate involving a multiplicative factor $c(m,d)$, and an explicit further lower estimate on this factor. We state and prove them separately.

First of all, our general lower bound is stated as follows. The proof is based on designing appropriate Lie algebraic developments.

Theorem 4.25. *Let V be a d-dimensional Banach space and let every tensor product $V^{\otimes n}$ be equipped with the associated projective tensor norm. For each $m \geqslant 1$, there exists a constant $c(m,d) \in (0,1]$ depending only on m and d, such that*

$$L_m(\mathbf{X}) \geqslant c(m,d)\|\pi_m(l)\|$$

for all pure m-rough paths $\mathbf{X}_t = \exp(tl) \in G^{(m)}(V)$ over V.

Proof. We write the highest degree component of l in the form $l_m = c_1h_1 + \cdots + c_\nu h_\nu$, where $\{h_1, \cdots, h_\nu\}$ is a given basis of $\mathcal{L}_m(V)$. Using the dual characterization (2.1) of the projective tensor norm, let B be a given m-linear functional over V whose norm is bounded by 1. We aim at constructing a Lie algebraic development $\Phi : V \to \mathfrak{g} \to \mathrm{End}(W)$ such that:

(i) $\mathfrak{g}$ is semisimple, and the space $\mathcal{L}_m(V)$ is mapped into a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ under the Lie homomorphism induced by F;
(ii) there exists a weight $\mu \in \mathfrak{h}^*$ for ρ such that $\mu(F(l_m)) = B(l_m)$;
(iii) the operator norm of Φ is bounded above by a constant which is independent of B and the specific values of the coefficients c_i.

If this can be achieved, the general lower bound will follow from (4.13) and (2.1) since B is arbitrary.

One way of constructing such a development is the following. For simplicity we assume that $\dim V = 2$ with a given basis $\{\mathrm{e}_1, \mathrm{e}_2\}$ (there is only notational difference in higher dimensions). We choose $\mathfrak{g} = \mathfrak{sl}(k \cdot m, \mathbb{C})$ where $k \geqslant 1$ is a large number to be specified later on. We choose a Cartan subalgebra $\mathfrak{h}$ and a base of simple roots according to the discussion in Part I of the current section. Define the embedding $F : V \to \mathfrak{g}$ in the following block diagonal form

$$F(\mathrm{e}_1) = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_k \end{pmatrix}_{km \times km}, \quad F(\mathrm{e}_2) = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_k \end{pmatrix}_{km \times km},$$

where each $A_i, B_j \in \mathfrak{sl}(m, \mathbb{C})$ $(1 \leqslant i, j \leqslant k)$ has the form

$$A_i = \begin{pmatrix} 0 & a_{i,1} & & 0 \\ & & \ddots & \\ & & & a_{i,m-1} \\ a_{i,m} & & & 0 \end{pmatrix}_{m \times m}, \quad B_j = \begin{pmatrix} 0 & b_{j,1} & & 0 \\ & & \ddots & \\ & & & b_{j,m-1} \\ b_{j,m} & & & 0 \end{pmatrix}_{m \times m},$$

with all the $a_{i,r}, b_{j,s}$'s being complex parameters to be specified later on. There are totally $2km$ independent variables to determine F. According to Lemma 4.19 and Example 4.20, under the induced homomorphism (still denoted as F) on the free Lie algebra, $\mathcal{L}_m(V)$ is mapped into the given Cartan subalgebra $\mathfrak{h}$.

Finally, we choose $\rho : \mathfrak{g} \to \mathrm{End}(W)$ to be the irreducible representation of $\mathfrak{g}$ associated with the k-th fundamental dominant integral functional λ_k according to Theorem 4.16, and more explicitly by Theorem 4.17 in the $\mathfrak{sl}(n, \mathbb{C})$ case, we have $W = \Lambda^k(\mathbb{C}^{km})$ and ρ being the k-th exterior power of the canonical matrix representation. According to the same theorem, a weight for this representation is given by

$$\mu = \mu_1 + \mu_{m+1} + \mu_{2m+1} + \cdots + \mu_{(k-1)m+1} \in \mathfrak{h}^*,$$

where recall that μ_i is the linear functional of taking the i-th diagonal entry.

To specify the parameters in order to fulfil the eigenvalue condition (ii) while respecting the uniformity condition (iii), we are led to setting up a system of equations:

$$\mu(F(h_i)) = B(h_i), \ 1 \leqslant i \leqslant \nu.$$

This is a polynomial system with ν equations and $2km$ independent complex variables. It has the form

$$\begin{cases} p_1(A_1, B_1) + \cdots + p_1(A_k, B_k) = B(h_1), \\ \cdots \\ p_\nu(A_1, B_1) + \cdots + p_\nu(A_k, B_k) = B(h_\nu), \end{cases} \tag{4.17}$$

where each p_i is a homogeneous polynomial of degree m in $2m$ complex variables. More precisely, $p_i(A, B)$ is the first entry of the diagonal polynomial matrix $G(h_i)$, where G is the homomorphism induced from the linear map $V \to \mathfrak{sl}(m, \mathbb{C}[a_i, b_j])$ given by

$$\mathrm{e}_1 \mapsto A \triangleq \begin{pmatrix} 0 & a_1 & & 0 \\ & & \ddots & \\ & & & a_{m-1} \\ a_m & & & 0 \end{pmatrix}, \ \mathrm{e}_2 \mapsto B \triangleq \begin{pmatrix} 0 & b_1 & & 0 \\ & & \ddots & \\ & & & b_{m-1} \\ b_m & & & 0 \end{pmatrix}.$$

It is important to view G as a homomorphism into the polynomial matrix algebra in $2m$ complex variables.

We claim that, the polynomial system (4.17) has a solution in $\mathbb{C}^{2km}$ for some large $k \geqslant 1$, which according to Lemma 4.23, boils down to showing that the polynomials $p_1, \cdots, p_\nu \in \mathbb{C}[a_i, b_j]$ are linearly independent. To this end, consider the linear map $T : V^{\otimes m} \to \mathbb{C}[a_i, b_j]$ defined by

$$T(\mathrm{e}_{i_1} \otimes \cdots \otimes \mathrm{e}_{i_m}) \triangleq (1,1)\text{-entry of } G(\mathrm{e}_{i_1} \otimes \cdots \otimes \mathrm{e}_{i_m}).$$

Explicit calculation then shows that

$$T(\mathrm{e}_{i_1} \otimes \cdots \otimes \mathrm{e}_{i_m}) = w_{i_1} \cdots w_{i_m}, \tag{4.18}$$

where $w_{i_j} = a_j$ or b_j according to whether $i_j = 1$ or 2. In particular, we see that T is injective. Since $h_1, \cdots, h_\nu$ is a basis of $\mathcal{L}_m(V) \subseteq V^{\otimes m}$, we conclude that the polynomials

$$p_i(A, B) = T(h_i), \ 1 \leqslant i \leqslant \nu$$

are linearly independent. Therefore, by Lemma 4.23, the polynomial system (4.17) has a solution for some large k. Any solution can be used to determine the Lie algebraic development $\Phi = \rho \circ F$ specified in the previously given form. Under such development, the eigenvalue condition (ii) holds, and it follows from Theorem 4.10 that

$$L_m(\mathbf{X}) \geqslant \frac{B(l_m)}{\|\Phi\|^m_{V \to \mathrm{End}(W)}}.$$

Now it remains to estimate the operator norm of Φ, which reduces to estimating a solution to the polynomial system (4.17). For this purpose, according to Lemma 4.23, there exists $k \geqslant 1$, such that for each $1 \leqslant i \leqslant \nu$, the polynomial system

$$\begin{cases} p_i(A_1, B_1) + \cdots + p_i(A_k, B_k) = 1, \\ p_j(A_1, B_1) + \cdots + p_j(A_k, B_k) = 0, \quad j \neq i, \end{cases} \tag{4.19}$$

has a solution $\mathbf{Z}^{(i)} \in \mathbb{C}^{2km}$. It follows that with $\tilde{\mathbf{Z}}^{(i)} \triangleq B(h_i)^{1/m}\mathbf{Z}^{(i)}$, the vector $\tilde{\mathbf{Z}} \triangleq (\tilde{\mathbf{Z}}^{(1)}, \cdots, \tilde{\mathbf{Z}}^{(\nu)}) \in \mathbb{C}^{2k\nu m}$ is a solution to the system (4.17) with k being enlarged to $k\nu$. Since $\|B\| \leqslant 1$, we see that $\tilde{\mathbf{Z}}$, and thus the operator norm of Φ, is bounded above by a constant depending only on the roughness m and the dimension d. Since B is arbitrary, this implies the desired lower bound with a multiplicative factor $c(m, d)$ depending only on m and d. $\square$

It is clear from the last paragraph of the previous proof that, the key to estimating the multiplicative factor $c(m, d)$ is an explicit estimate on a solution to the system (4.17). In general, selecting solutions to a consistent polynomial system with a priori bounds is an important topic in computational algebraic geometry that has been studied by many authors. We state a result of Vorob'ev [29] that is relevant to us. Recall that the *bitsize* of a nonzero integer n is the unique natural number τ such that $2^{\tau-1} \leqslant |n| < 2^{\tau}$. The bitsize of a rational number is the sum of the bitsizes of its numerator and denominator.

Lemma 4.26 (cf. [29], Theorem 3)**.** *Let $\mathcal{V}$ be the set of real solutions to a consistent system of polynomial equations $f_1 = \cdots = f_\nu = 0$ where each $f_i \in \mathbb{Q}[x_1, \cdots, x_n]$. Let L be the maximum of the bitsizes of the coefficients of the system, $D \triangleq \sum_{i=1}^{\nu} \deg f_i$ and $r \triangleq \begin{pmatrix} n+2D \\ n \end{pmatrix}$. Then there exists a point $x = (x_1, \cdots, x_n) \in \mathcal{V}$, such that*

$$|x_i| \leqslant 2^{H(r,L)} \quad \text{for all } 1 \leqslant i \leqslant n,$$

where H is some universal bivariate polynomial independent of the original system.

Remark 4.27. In Vorob'ev's result (and other results of similar type), having rational or sometimes integral coefficients is a crucial assumption. In addition, it presumes the consistency of the system before locating an a priori bounded solution. In particular, it does not provide a proof on whether the system admits a solution.

With the help of Vorob'ev's estimate, we can now establish an explicit estimate on the factor $c(m, d)$ arising from Theorem 4.25.

Theorem 4.28. *Keeping the same notation as in Theorem 4.25, the multiplicative factor $c(m, d)$ satisfies*

$$c(m, d) \geqslant \Lambda_d^{-m} \cdot 2^{-(\nu_{m,d}!)^{\gamma \nu_{m,d}}},$$

where Λ_d is a constant depending only on d, $\nu_{m,d} \triangleq \dim \mathcal{L}_m(V)$, and $\gamma > 1$ is a universal constant.

Proof. Essentially we just need to keep track of the quantities appearing in the proof of Theorem 4.25 in a precise way.

First of all, in that proof we fix a basis $\{\mathrm{e}_1, \cdots, \mathrm{e}_d\}$ of V with norm 1, and assume that $\{h_1, \cdots, h_\nu\}$ is a Hall basis of $\mathcal{L}_m(V)$ built over the letters $\mathrm{e}_1, \cdots, \mathrm{e}_d$. Next, in the representation $\rho : \mathfrak{sl}(k \cdot m, \mathbb{C}) \to \Lambda^k(\mathbb{C}^{km})$, we work with the l^1-norm on $\Lambda^k(\mathbb{C}^{km})$ with respect to the canonical exterior basis. In addition, according to Lemma 4.23 we choose $k = 4^{\nu-1}\nu!$ for the system (4.19). Recall that $\mathbf{Z}^{(i)}$ (respectively, $\tilde{\mathbf{Z}}$) is a solution to the system (4.19) (respectively, (4.17)). Now we presume that for each i, all components of $\mathbf{Z}^{(i)}$ are bounded by a number M. Using the observation that $\|h_i\| \leqslant 2^m$, we know that all components of $\tilde{\mathbf{Z}}$ are bounded by $2M$. It then follows from a simple unwinding of definitions that

$$\|\Phi\|_{V \to \mathrm{End}(W)} \leqslant 2\Lambda_d k M, \tag{4.20}$$

where Λ_d is the constant depending only on d which arises from the comparison between the given norm $\|\cdot\|_V$ on V and the l^1-norm $\|\cdot\|_1$ with respect to the basis $\{\mathrm{e}_1, \cdots, \mathrm{e}_d\}$, i.e. $\|\cdot\|_1 \leqslant \Lambda_d \|\cdot\|_V$.

It remains to work out M explicitly. The first observation is that, the system (4.19) has integral coefficients each being bounded by 2^m. To apply Vorob'ev's estimate, we need to turn the system into an equivalent system over real variables, which can be done by viewing each complex variable as a pair of real variables. In this way, (4.19) becomes a system with 2ν equations and $2kdm$ real variables. The next observation is that, the new system again has integer coefficients, and more importantly when transforming from complex to real, the coefficients are not enlarged. This is due to the fact that the polynomial p_i is linear with respect to every single complex variable when the others are frozen (cf. (4.18) for the shape of relevant monomials). Therefore, using the notation in Lemma 4.26, we find that

$$L \leqslant m, \ D = 2m\nu, \ n = \frac{1}{2}dm4^\nu \nu!, \ r = \begin{pmatrix} n+2D \\ n \end{pmatrix}.$$

It follows from Stirling's approximation and Vorob'ev's estimate that $M \leqslant 2^{(\nu!)^{\kappa\nu}}$ with some universal constant $\kappa > 1$ independent of the system. Now the result follows by substituting this into (4.20) and using Theorem 4.25. $\square$

Remark 4.29. The proof of Theorem 4.25 does not provide the optimal way of constructing the Lie algebraic development Φ in general, and the explicit lower bound given by Theorem 4.28 does not seem to be optimal either. To improve the estimate, among the class of Lie algebraic developments Φ in which $\|\pi_m(l)\|$ is an eigenvalue of $\Phi(\pi_m(l))$, one needs to minimize the operator norm of Φ. As we will see in low degree cases, there are plenty of rooms for reducing the operator norm of Φ and hence improving the factor $c(m,d)$. The sharp lower bound (Conjecture 2.12) will hold if one can achieve $\|\Phi\|_{V\to\mathrm{End}(W)} = 1$.

As an immediate corollary of our methodology, we prove the following separation of points property for signatures. Such a separation property was first obtained by Chevyrev-Lyons [7] as an essential ingredient of proving their uniqueness result for the expected signature of stochastic processes.

Corollary 4.30. *Let V be a finite dimensional vector space.*

(1) Let $l, l' \in \mathcal{L}(V)$ be two distinct Lie polynomials over V. Then there exists a finite dimensional complex semisimple Lie algebra $\mathfrak{g}$ and a linear embedding $F: V \to \mathfrak{g}$, such that $F(l) \neq F(l')$.
(2) Let g_1, g_2 be the signatures of two weakly geometric rough paths over V. Suppose that $g_1 \neq g_2$. Then there exists a finite dimensional complex semisimple Lie algebra $\mathfrak{g}$ and a linear embedding $F: V \to \mathfrak{g}$, such that $F(g_1) \neq F(g_2)$.

Proof. (1) Let $m \geqslant 1$ be the smallest integer such that $\pi_m(l) \neq \pi_m(l')$. According to the proof of Theorem 4.25, there exists a finite dimensional Lie algebraic development

$$\Phi : V \xrightarrow{F} \mathfrak{g} \xrightarrow{\rho} \mathrm{End}(W)$$

such that

$$\Phi(\pi_m(l)) \neq \Phi(\pi_m(l')).$$

More specifically, we have $\mathfrak{g} = \mathfrak{sl}(k \cdot m, \mathbb{C})$ and $W = \Lambda^k(\mathbb{C}^{km})$ with $k = 4^{\nu-1}\nu!$ and $\nu = \dim \mathcal{L}_m(V)$. For given $\varepsilon > 0$, define $\Phi_\varepsilon \triangleq \rho \circ (\varepsilon \cdot F)$. It follows that

$$\begin{aligned}
\Phi_\varepsilon(l - l') &= (\rho \circ (\varepsilon \cdot F))(l - l') \\
&= \rho\left(\varepsilon^m \cdot F(\pi_m(l - l')) + \sum_{n>m} \varepsilon^n \cdot F(\pi_n(l - l'))\right) \\
&= \varepsilon^m \cdot \Phi(\pi_m(l - l')) + \sum_{n>m} \varepsilon^n \cdot \Phi(\pi_n(l - l')).
\end{aligned}$$

Note that the summation is indeed finite since l, l' are Lie polynomials. Therefore, we see that

$$\Phi_\varepsilon(l - l') = \varepsilon^m \cdot \Phi(\pi_m(l - l')) + o(\varepsilon^m),$$

which implies that $\Phi_\varepsilon(l - l') \neq 0$ when ε is small. The embedding $\varepsilon \cdot F$ will satisfy the desired property.

(2) Write $g = \exp(l)$ and $g' = \exp(l')$ where l, l' are Lie series respectively. In the same way as the proof of the first part, let $m \geqslant 1$ be the smallest integer such that $\pi_m(l) \neq \pi_m(l')$, and choose a finite dimensional Lie algebraic development $\Phi = \rho \circ F : V \to \mathfrak{g} \to \mathrm{End}(W)$ separating $\pi_m(l)$ and $\pi_m(l')$. Since g and g' are path signatures, it is known that (cf. [23] and [6]), l and l' both have positive radius of convergence when viewed as formal tensor series. In particular, both of

$$\varepsilon \mapsto \Phi_\varepsilon(l),\ \ \varepsilon \mapsto \Phi_\varepsilon(l')$$

are analytic functions in some neighbourhood of $\varepsilon = 0$ where $\Phi_\varepsilon \triangleq \rho \circ (\varepsilon \cdot F)$. Therefore, we see that

$$\Phi_\varepsilon(l) = \varepsilon^m \cdot \Phi(\pi_m(l)) + o(\varepsilon^m), \quad \Phi_\varepsilon(l') = \varepsilon^m \cdot \Phi(\pi_m(l')) + o(\varepsilon^m),$$

when ε is small. Note that we also have

$$\Phi_\varepsilon(g) = \exp\left(\Phi_\varepsilon(l)\right), \quad \Phi_\varepsilon(g') = \exp\left(\Phi_\varepsilon(l')\right).$$

Since the exponential map for the group $\mathrm{Aut}(W)$ is a local diffeomorphism at the identity, we conclude that $\Phi_\varepsilon(g) \neq \Phi_\varepsilon(g')$ when ε is small. In particular, the embedding $\varepsilon \cdot F$ will satisfy the desired property.

□

Remark 4.31. Apart from the $\mathfrak{sl}(n;\mathbb{C})$ structure, the above argument can be extended to other types of semisimple Lie algebras (possibly of compact type) for instance over $\mathfrak{so}(n;\mathbb{C})$, as long as the development is chosen so that the highest degree Lie component is mapped into a Cartan subalgebra (cf. Example 4.21). In [7], the separation property was proved using the (semi)simple Lie algebra $\mathfrak{sp}(n)$.

Remark 4.32. One advantage of stating the separation property at the level of free Lie algebra (Part (1) of the corollary) is that the property becomes purely algebraic. Even at the level of signature, the dependence on analytic properties is rather mild. Indeed, the proof of the positive radius of convergence for the logarithmic signature given in [6] requires only the faster-than-geometric decay for signature components. This is the only analytic condition needed here.

4.2.4 Explicit calculations in low degrees

We perform some more explicit calculations in low degrees to illustrate the methodology better. We consider $V = \mathbb{R}^2$ equipped with the l^1-norm with respect to the standard basis $\{\mathrm{e}_1, \mathrm{e}_2\}$. The associated projective tensor norm then coincides with the l^1-norm with respect to the canonical tensor basis. In this context, we are going to show that, the sharp lower bound holds in degrees $m = 2, 3$ and some cases in degrees $m = 4, 5$. When $m = 4$, we have $c(4, 2) \geqslant 5/32$ in general.

I. Sharp lower bound in degrees 2 and 3

Let $\mathbf{X}_t = \exp(tl)$ be a pure 2-rough path, and write $\pi_2(l) = c[\mathrm{e}_1, \mathrm{e}_2] \in \mathcal{L}_2(V)$. In order to develop $\mathcal{L}_2(V)$ into a Cartan subalgebra, according to Lemma 4.19 and Example 4.20, we choose $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C})$, and define $F : V \to \mathfrak{g}$ by

$$F(\mathrm{e}_1) = \begin{pmatrix} 0 & a_1 \\ a_2 & 0 \end{pmatrix}, \ F(\mathrm{e}_2) = \begin{pmatrix} 0 & b_1 \\ b_2 & 0 \end{pmatrix},$$

where a_1, a_2, b_1, b_2 are parameters to be specified. In addition, we choose $\rho : \mathfrak{g} \to \mathrm{End}(\mathbb{C}^2)$ to be the canonical matrix representation, where $\mathbb{C}^2$ is equipped with the standard l^1-norm.

Note that

$$F([\mathrm{e}_1, \mathrm{e}_2]) = \begin{pmatrix} a_1 b_2 - a_2 b_1 & 0 \\ 0 & a_2 b_1 - a_1 b_2 \end{pmatrix} \in \mathfrak{h}.$$

Since $\|\pi_2(l)\| = 2|c|$, we set up the equation

$$a_1 b_2 - a_2 b_1 = +2 \text{ or } -2, \tag{4.21}$$

depending on whether c is positive or negative. This will allow us to produce $\|\pi_2(l)\|$ as an eigenvalue of $\Phi(\pi_2(l)) \in \mathrm{End}(\mathbb{C}^2)$. Among all solutions, the minimum $\|\Phi\|_{\mathbb{R}^2 \to \mathrm{End}(\mathbb{C}^2)} = 1$ is obtained at

$$a_1 = a_2 = 1, \ b_1 = \mp 1, \ b_2 = \pm 1,$$

where the signs are chosen depending on whether c is positive or negative. According to Theorem 4.1 and Theorem 4.10, we conclude that $L_2(\mathbf{X}) = \|\pi_2(l)\|$ and thus Conjecture 2.12 holds for roughness $m = 2$.

Next we consider the case when $l \in \mathcal{L}^{(3)}(V)$. In this case, $\pi_3(l) \in \mathcal{L}_3(V)$ takes the form

$$\pi_3(l) = c_1[\mathrm{e}_1, [\mathrm{e}_1, \mathrm{e}_2]] + c_2[[\mathrm{e}_1, \mathrm{e}_2], \mathrm{e}_2].$$

To develop $\mathcal{L}_3(V)$ into a Cartan subalgebra, we choose $\mathfrak{g} = \mathfrak{sl}(3, \mathbb{C})$, define $F : V \to \mathfrak{g}$ by

$$F(\mathrm{e}_1) = \begin{pmatrix} 0 & a_1 & 0 \\ 0 & 0 & a_2 \\ a_3 & 0 & 0 \end{pmatrix}, \; F(\mathrm{e}_2) = \begin{pmatrix} 0 & b_1 & 0 \\ 0 & 0 & b_2 \\ b_3 & 0 & 0 \end{pmatrix}$$

where a_i, b_j's are parameters to be determined, and choose $\rho : \mathfrak{g} \to \mathrm{End}(\mathbb{C}^3)$ to be the canonical matrix representation where $\mathbb{C}^3$ is equipped with the standard l^1-norm.

Suppose that $c_1, c_2 > 0$, under which we have $\|\pi_3(l)\| = 4c_1 + 4c_2$. To match the eigenvalues, we set up a system of equations

$$\mu_1(F([\mathrm{e}_1, [\mathrm{e}_1, \mathrm{e}_2]])) = 4, \;\; \mu_1(F([[\mathrm{e}_1, \mathrm{e}_2], \mathrm{e}_2])) = 4,$$

where recall that μ_1 is a weight for ρ defined by taking the first diagonal entry. By direct calculation, the system reads

$$\begin{cases} a_1 a_2 b_3 + a_2 a_3 b_1 - 2a_1 a_3 b_2 = 4, \\ a_1 b_2 b_3 - 2a_2 b_1 b_3 + a_3 b_1 b_2 = 4. \end{cases}$$

Among all its solutions, the minimum $\|\Phi\|_{\mathbb{R}^2 \to \mathrm{End}(\mathbb{C}^3)} = 1$ is achieved at

$$a_1 = a_2 = 1, \; a_3 = -1, \; b_1 = -1, \; b_2 = b_3 = 1.$$

The cases for other sign conditions on c_1, c_2 are treated similarly. Therefore, Conjecture 2.12 holds for roughness $m = 3$.

II. The degree 4 case

Now consider $l \in \mathcal{L}^{(4)}(V)$ with $\pi_4(l) = c_1 h_1 + c_2 h_2 + c_3 h_3$, where

$$h_1 = [[\mathrm{e}_1, [\mathrm{e}_1, \mathrm{e}_2]], \mathrm{e}_1], \;\; h_2 = [[[\mathrm{e}_1, \mathrm{e}_2], \mathrm{e}_2], \mathrm{e}_2], \;\; h_3 = [\mathrm{e}_1, [[\mathrm{e}_1, \mathrm{e}_2], \mathrm{e}_2]]$$

form a Hall basis of $\mathcal{L}_4(V)$. In this case, we demonstrate the possibility of using other root systems that are not isomorphic to $\mathfrak{sl}(n, \mathbb{C})$, and show that

$$L_4(\mathbf{X}) \geqslant \begin{cases} \frac{5}{32}\|\pi_4(l)\|, & c_1 \cdot c_2 \geqslant 0, \\ \frac{\sqrt{7}}{8}\|\pi_4(l)\|, & c_1 \cdot c_2 < 0. \end{cases} \tag{4.22}$$

To be precise, we choose $\mathfrak{g} = \mathfrak{so}(5, \mathbb{C})$ and develop $\mathcal{L}_4(V)$ into a Cartan subalgebra according to Example 4.21. A Cartan subalgebra $\mathfrak{h}$ is generated by the two elements

$$H_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; H_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The generators of the three root spaces $\mathfrak{g}^\alpha, \mathfrak{g}^\beta, \mathfrak{g}^\gamma$ corresponding to the specified roots α, β, γ in that example can be chosen as

$$X_\alpha = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & i \\ 0 & 0 & -1 & -i & 0 \end{pmatrix},$$

$$X_\beta = \begin{pmatrix} 0 & 0 & -1 & i & 0 \\ 0 & 0 & i & 1 & 0 \\ 1 & -i & 0 & 0 & 0 \\ -i & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; X_\gamma = \begin{pmatrix} 0 & 0 & -1 & i & 0 \\ 0 & 0 & -i & -1 & 0 \\ 1 & i & 0 & 0 & 0 \\ -i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

respectively. We refer the reader to [17], Chapter III, Section 8 for an explicit description of the root space decomposition of $\mathfrak{g}$, from which one will see how the above matrices arise naturally.

Now we define $F : V \to \mathfrak{g}$ by

$$F(\mathrm{e}_1) = a_1 X_\alpha + a_2 X_\beta + a_3 X_\gamma, \;\; F(\mathrm{e}_2) = b_1 X_\alpha + b_2 X_\beta + b_3 X_\gamma,$$

where a_i, b_j's are parameters to be chosen. According to Example 4.21, we have $F(\mathcal{L}_4(V)) \subseteq \mathfrak{h}$. We choose $\rho : \mathfrak{g} \to \mathbb{C}^5$ to be the canonical matrix representation, where $\mathbb{C}^5$ is equipped with the standard Hermitian norm. A common eigenbasis of $\mathbb{C}^5$ for all elements in $\mathfrak{h}$ under ρ is given by

$$w_1 = \varepsilon_5, \;\; w_2 = i\varepsilon_1 + \varepsilon_2, \;\; w_3 = -i\varepsilon_1 + \varepsilon_2, \;\; w_4 = i\varepsilon_3 + \varepsilon_4, \;\; w_5 = -i\varepsilon_3 + \varepsilon_4,$$

where $\{\varepsilon_1, \cdots, \varepsilon_5\}$ is the canonical basis of $\mathbb{C}^5$. For $H = xH_1 + yH_2 \in \mathfrak{h}$, the set of eigenvalues of $\rho(H)$ with respect to the above eigenbasis (listed in the same order) is $\{0, -ix, ix, -iy, iy\}$. Denote μ as the weight defined by $H = xH_1 + yH_2 \mapsto iy$, the eigenvalue with respect to the common eigenvector w_5.

Suppose that $c_1, c_2, c_3 > 0$, under which we have $\|\pi_4(l)\| = 8c_1 + 8c_2 + 6c_3$. We then set up a polynomial system

$$\mu(F(h_1)) = 8, \;\; \mu(F(h_2)) = 8, \;\; \mu(F(h_3)) = 6. \tag{4.23}$$

The left hand side consists of homogeneous polynomials of degree 4 in six variables a_i, b_j. To simplify computation, we restrict ourselves to solutions satisfying $a_2 = a_3, b_2 = b_3$. Under this constraint, by explicit calculation it is seen that $\pm\mu(F(h_i))$ become the only possibly nonzero eigenvalues of $\Phi(h_i)$ $(i = 1, 2, 3)$, and the system (4.23) reads

$$\begin{cases} 4a_1 a_3(a_1 b_3 - a_3 b_1) = 1, \\ 4b_1 b_3(a_1 b_3 - a_3 b_1) = -1, \\ 8(a_1^2 b_3^2 - a_3^2 b_1^2) = 3. \end{cases}$$

Treating a_1 as a free variable, the above system can be solved explicitly to yield precisely four scenarios:

$$\begin{cases} a_3 = \pm\frac{\sqrt{10}}{10a_1}, \\ b_1 = -\frac{a_1}{2}, \\ b_3 = \pm\frac{\sqrt{10}}{5a_1}, \end{cases} \quad , \quad \begin{cases} a_3 = \pm\frac{\sqrt{10}i}{10a_1}, \\ b_1 = 2a_1, \\ b_3 = \mp\frac{\sqrt{10}}{20a_1}. \end{cases}$$

In other words, the solution set $\Sigma \subseteq \mathbb{C}^4$ has complex dimension one and consists of four irreducible components $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4$, each being globally parametrized by $a_1 \in \mathbb{C}\backslash\{0\}$.

Finally, we try to minimize the operator norm of Φ over Σ. To this end, first recall that given an $n \times n$ complex matrix A, when viewed as a linear transformation over $\mathbb{C}^n$, the operator norm of A with respect to the standard Hermitian norm on $\mathbb{C}^n$ coincides with the maximal singular value of A. By direct calculation, on the component Σ_1, the sets of singular values of $\Phi(\mathrm{e}_1), \Phi(\mathrm{e}_2) \in \mathrm{End}(\mathbb{C}^5)$ are

$$\left\{0, \sqrt{2}|a_1|, \frac{2\sqrt{5}}{5|a_1|}\right\}, \ \left\{0, \frac{|a_1|}{\sqrt{2}}, \frac{4\sqrt{5}}{5|a_1|}\right\}$$

respectively. Therefore, we have

$$\|\Phi\|_{\mathbb{R}^2\to\mathrm{End}(\mathbb{C}^5)} = \max\left\{\|\Phi(\mathrm{e}_1)\|_{\mathbb{C}^5\to\mathbb{C}^5}, \|\Phi(\mathrm{e}_2)\|_{\mathbb{C}^5\to\mathbb{C}^5}\right\} = \max\left\{\sqrt{2}|a_1|, \frac{4\sqrt{5}}{5|a_1|}\right\}.$$

It is now elementary to see that the minimum of $\|\Phi\|_{\mathbb{R}^2\to\mathrm{End}(\mathbb{C}^5)}$ over Σ_1 is achieved at $|a_1| = 2 \cdot 10^{-1/4}$, and the minimum equals $2\sqrt{2} \cdot 10^{-1/4}$. Similar calculation over the other three components of Σ yields exactly the same minimum. Therefore, we conclude that

$$\inf_{\Sigma} \|\Phi\|_{\mathbb{R}^2\to\mathrm{End}(\mathbb{C}^5)} = 2\sqrt{2} \cdot 10^{-1/4},$$

and the infimum is achieved at a Lie algebraic development determined by, for instance,

$$a_1 = 2 \cdot 10^{-1/4}, \ a_2 = a_3 = \frac{1}{2} \cdot 10^{-1/4}, \ b_1 = -10^{-1/4}, \ b_2 = b_3 = 10^{-1/4}.$$

Under this development, we have the lower bound

$$L_4(\mathbf{X}) \geqslant \frac{8c_1 + 8c_2 + 6c_3}{\|\Phi\|^4_{\mathbb{R}^2\to\mathrm{End}(\mathbb{C}^5)}} = \frac{5}{32}\|\pi_4(l)\|.$$

The discussion for other sign conditions on the coefficients c_1, c_2, c_3 is entirely analogous by adjusting the signs on the right hand side of the system (4.23) accordingly. This eventually leads us to precisely two scenarios of the desired lower bound (4.22). We omit the lengthy and repeating calculations.

On the other hand, if one of the coefficients c_1, c_2, c_3 is zero, the lower bound can be improved further, since one equation from the system (4.23) is removed which produces a higher dimensional solution set. Indeed, when $c_3 = 0$, one obtains the sharp lower bound and hence Conjecture 2.12 holds for this case. A simple choice of Lie algebraic developments achieving the sharp lower bound is the following. Choose $\mathfrak{g}$ to be $\mathfrak{sl}(4,\mathbb{C})$, the representation ρ to be the canonical matrix representation, and the embedding $F: V \to \mathfrak{g}$ to be given by

$$\mathrm{e}_1 \mapsto A \triangleq \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \ \mathrm{e}_2 \mapsto B \triangleq \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

if $c_1 \cdot c_2 \geqslant 0$, and

$$\mathrm{e}_1 \mapsto \mathrm{e}^{\frac{5\pi i}{8}} \cdot A, \ \mathrm{e}_2 \mapsto \mathrm{e}^{\frac{\pi i}{8}} \cdot B$$

if $c_1 \cdot c_2 < 0$, respectively. The same conclusion is true in degree 5 when $\pi_5(l)$ consists of a single Hall polynomial. We again omit the similar type of calculations.

5 The Hilbert-Schmidt tensor norm: proof of Theorem 3.1

As we mentioned earlier (cf. Theorem 3.1), Conjecture 2.12 can be proved for a special class of pure rough paths if we work with the Hilbert-Schmidt tensor norm instead. Here we give an independent proof of this result.

Let $V = \mathbb{R}^d$ be equipped with the l^2-metric with respect to the standard basis $\{\mathrm{e}_1, \cdots, \mathrm{e}_d\}$. We equip each $V^{\otimes m}$ with the l^2-metric with respect to the standard tensor basis. They extend to an inner product structure $\langle \cdot, \cdot \rangle$ on the subalgebra $T(V)$ of $T((V))$ consisting of finite tensors by requiring that $V^{\otimes m}$ and $V^{\otimes n}$ are orthogonal if $m \neq n$. By considering basis elements and using bilinearity, it is immediate that

$$\langle \xi_m \otimes \xi_n, \eta_m \otimes \eta_n \rangle = \langle \xi_m, \eta_m \rangle \cdot \langle \xi_n, \eta_n \rangle$$

for all $\xi_m, \eta_m \in V^{\otimes m}$ and $\xi_n, \eta_n \in V^{\otimes n}$.

Recall from the assumption that $\mathbf{X}_t = \exp(t(l_a + l_b)) \in G^{(b)}(V)$, where $a < b$ and l_a, l_b are homogeneous Lie polynomials of degrees a, b respectively. Suppose that $(b-a)/\gcd(a,b)$ is an odd integer. We aim at showing that $L_b(\mathbf{X}) = \|l_b\|$.

For each $k \geqslant 1$, we write

$$\pi_{bk}\left(\exp(l_a + l_b)\right) = \frac{l_b^{\otimes k}}{k!} + Q, \tag{5.1}$$

where the exponential is now taken over $T((V))$, and Q is sum of all remaining terms in the expansion. The key step is to show that, if $(b-a)/\gcd(a,b)$ is odd, then $l_b^{\otimes k}$ and Q are orthogonal for all large k. This can be proved by making use of an anti-automorphism on the tensor algebra together with symmetry properties of the signature expansion. The orthogonality property clearly leads to the lower estimate

$$\|\pi_{bk}(\exp(l_a + l_b))\| \geqslant \frac{\|l_b\|^k}{k!}.$$

Combining with the general upper estimate given by Theorem 4.1, the result then follows.

To prove (5.1), first consider the linear map $\alpha : T(V) \to T(V)$ induced by

$$\alpha(\mathrm{e}_{i_1} \otimes \cdots \otimes \mathrm{e}_{i_m}) = (-1)^m \mathrm{e}_{i_m} \otimes \cdots \otimes \mathrm{e}_{i_1}.$$

By definition, α is an anti-involution, i.e. $\alpha(\xi \otimes \eta) = \alpha(\eta) \otimes \alpha(\xi)$ and $\alpha^2 = \mathrm{Id}$. In addition, for any $\xi, \eta \in T(V)$, we have $\langle \alpha(\xi), \alpha(\eta) \rangle = \langle \xi, \eta \rangle$. A crucial property of α is that $\alpha(l) = -l$ for any Lie polynomial l. The notion of α and the above properties can be found in [27], Chapter 1. An immediate consequence of using the anti-involution α is the following lemma. Recall that the *symmetrized product* of $\xi_1, \cdots, \xi_n \in T(V)$ is defined by

$$\mathrm{Sym}(\xi_1, \cdots, \xi_n) \triangleq \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \xi_{\sigma(1)} \otimes \cdots \otimes \xi_{\sigma(n)},$$

where $\mathcal{S}_n$ is the permutation group of order n. For convenience, we also define the *reduced symmetrized product*

$$\mathrm{RSym}(\underbrace{\xi_1, \cdots, \xi_1}_{k_1 \text{ times}}, \cdots, \underbrace{\xi_n, \cdots, \xi_n}_{k_n \text{ times}}) \triangleq \frac{1}{k_1! \cdots k_n!} \mathrm{Sym}(\underbrace{\xi_1, \cdots, \xi_1}_{k_1 \text{ times}}, \cdots, \underbrace{\xi_n, \cdots, \xi_n}_{k_n \text{ times}}).$$

Lemma 5.1. *Let $l_0, l_1, \cdots, l_n$ be Lie polynomials and $k \geqslant 1$. If $k+n$ is an odd integer, then*

$$\langle l_0^{\otimes k}, \mathrm{Sym}(l_1, \cdots, l_n) \rangle = 0.$$

The same result holds for the reduced symmetrized product.

Proof. Observe that

$$\alpha\left(\mathrm{Sym}(l_1, \cdots, l_n)\right) = (-1)^n \mathrm{Sym}(l_1, \cdots, l_n).$$

Therefore, we have

$$\begin{aligned} \langle l_0^{\otimes k}, \mathrm{Sym}(l_1, \cdots, l_n) \rangle &= \langle \alpha(l_0^{\otimes k}), \alpha(\mathrm{Sym}(l_1, \cdots, l_n)) \rangle \\ &= (-1)^{k+n} \langle l_0^{\otimes k}, \mathrm{Sym}(l_1, \cdots, l_n) \rangle. \end{aligned}$$

The first assertion follows since $k+n$ is odd by assumption. The second assertion is obvious. □

Now we are in a position to give the proof of Theorem 3.1.

Proof of Theorem 3.1. We express the remainder Q in the expression (5.1) in a more explicit way:

$$Q = \sum_{x>0,\ ax+by=bk} \mathrm{RSym}(\underbrace{l_a, \cdots, l_a}_{x \text{ times}}, \underbrace{l_b, \cdots, l_b}_{y \text{ times}}). \tag{5.2}$$

An important observation is that, for each summand, since $ax + by = bk$, we have

$$\frac{a}{r} x = \frac{b}{r}(k - y)$$

where $r \triangleq \gcd(a, b)$, showing that $b/r \mid x$ and thus $x \geqslant b/r$. For the equation to make sense, one also needs $k \geqslant a/r$.

Firstly, if $x = b/r$, then $y = k - a/r$. In this case, we have

$$k + x + y = 2k + \frac{b-a}{r},$$

which is an odd integer by assumption. According to Lemma 5.1, we conclude that

$$\langle l_b^{\otimes k}, \mathrm{RSym}(\underbrace{l_a, \cdots, l_a}_{x \text{ times}}, \underbrace{l_b, \cdots, l_b}_{y \text{ times}})\rangle = 0. \tag{5.3}$$

Next, consider a given $x > b/r$ from the sum in (5.2). For each single term ξ in the corresponding reduced symmetrized product, ξ can be uniquely written as $\xi = \xi_1 \otimes \xi_2$, where ξ_1 contains exactly b/r number of l_a's and ξ_2 starts with l_a. Let S be the set of all such ξ_2's arising in this way. Denote $y(\xi_2)$ as the number of l_b's in each given $\xi_2 \in S$. Then the reduced symmetrized product can further be written as

$$\begin{aligned}
&\mathrm{RSym}(\underbrace{l_a, \cdots, l_a}_{x \text{ times}}, \underbrace{l_b, \cdots, l_b}_{y \text{ times}}) \\
&= \sum_{\xi_2 \in S} \left(\frac{b}{r} + y - y(\xi_2)!\right) \cdot \mathrm{RSym}(\underbrace{l_a, \cdots, l_a}_{b/r \text{ times}}, \underbrace{l_b, \cdots, l_b}_{y - y(\xi_2) \text{ times}}) \otimes \xi_2.
\end{aligned}$$

For each $\xi_2 \in S$, by writing $k_1 \triangleq a/r + y - y(\xi_2)$, Lemma 5.1 again implies that

$$\langle l_b^{\otimes k_1}, \mathrm{RSym}(\underbrace{l_a, \cdots, l_a}_{b/r \text{ times}}, \underbrace{l_b, \cdots, l_b}_{y - y(\xi_2) \text{ times}})\rangle \cdot \langle l_b^{\otimes (k-k_1)}, \xi_2\rangle = 0,$$

since

$$k_1 + \frac{b}{r} + y - y(\xi_2) = 2k_1 + \frac{b-a}{r}$$

is an odd integer. Therefore, (5.3) holds for the reduced symmetrized product corresponding to the given x.

It follows that $l_b^{\otimes k}$ is orthogonal to Q provided $k \geqslant a/r$, and the proof of the theorem is now complete.

$\square$

Acknowledgement

HB and NS are supported by EPSRC grant EP/R008205/1. XG is supported in part by NSF grant DMS1814147. The authors wish to thank Dr Wenzhe Yang for his valuable comments and suggestions from the viewpoint of algebraic geometry.

Appendix A Some properties of pure rough paths

In this section, we prove the two properties of pure rough paths stated in Proposition 2.10 and Proposition 2.11 respectively in Section 2.3.

Proof of Proposition 2.10. Let $\mathbf{X}_t = \exp(tl)$ $(0 \leqslant t \leqslant 1)$ be a pure m-rough path, where $l \in \mathcal{L}^{(m)}(V)$ with $l_m \triangleq \pi_m(l) \neq 0$. For each $1 \leqslant k \leqslant m$, the degree k component of $\mathbf{X}_{s,t} \triangleq \mathbf{X}_s^{-1} \otimes \mathbf{X}_t$ has the form

$$X_{s,t}^k = \sum_{r=0}^{m} \frac{(t-s)^r}{r!} \pi_k(l^{\otimes r}) = \sum_{r=1}^{k} (t-s)^r \xi_r^{(k)},$$

where $\xi_r^{(k)} \in V^{\otimes k}$ are tensors constructed from $\pi_1(l), \cdots, \pi_m(l)$. It follows that

$$\|X_{s,t}^k\| \leqslant \sum_{r=1}^{k} |t-s|^r \|\xi_r^{(k)}\| \leqslant C_{\mathbf{X}} \cdot |t-s|, \tag{A.1}$$

where $C_{\mathbf{X}}$ denotes a constant depending only on $\mathbf{X}$. This implies that $\mathbf{X}$ is an m-rough path in the sense of Definition 2.2.

Now if $k < m$, from (A.1) we have

$$\|X_{s,t}^k\|^{\frac{m}{k}} \leqslant C_{\mathbf{X}}^{\frac{m}{k}} \cdot |t-s|^{\frac{m}{k}},$$

showing that

$$\lim_{\mathrm{mesh}(\mathcal{P}) \to 0} \sum_{t_i \in \mathcal{P}} \|X_{t_{i-1},t_i}^k\|^{\frac{m}{k}} = 0.$$

If $k = m$, notice that $\xi_1^{(m)} = l_m$. Therefore, given a finite partition $\mathcal{P}$ of $[0,1]$, we have

$$\|X_{t_{i-1},t_i}^m\| = (t_i - t_{i-1}) \cdot \left\| l_m + (t_i - t_{i-1})\xi_2^{(m)} + \cdots + (t_i - t_{i-1})^{m-1}\xi_m^{(m)} \right\|.$$

It is now elementary to see that

$$\lim_{\mathrm{mesh}(\mathcal{P}) \to 0} \sum_{t_i \in \mathcal{P}} \|X_{t_{i-1},t_i}^m\| = \|l_m\|.$$

Consequently, we conclude that the local m-variation of $\mathbf{X}$ equals $\|l_m\|$.

$\square$

Remark A.1. This property apparently extends to the non-geometric setting, i.e. for the case when $l \in T^{(m)}(V)$. Indeed, even more holds true with essentially the same proof. Let $\mathbf{X}_t = \exp(L(t)) \in T^{(m)}(V)$, where $L(t)$ is a bounded variation path in $T^{(m)}(V)$. Then

$$\lim_{\delta \to 0} \sum_{k=1}^{m} \left(\inf_{\mathrm{mesh}(\mathcal{P}) \leqslant \delta} \sum_{t_i \in \mathcal{P}} \|X_{t_{i-1},t_i}^k\|^{\frac{m}{k}} \right)^{\frac{k}{m}} = \|\pi_m(L)\|_{1\text{-var}}.$$

Proof of Proposition 2.11. Let $\mathbf{X}_t = \exp(tl) \in G^{(m)}(V)$ be a pure m-rough path. For any $n \geqslant m$, it is not hard to see that the multiplicative functional $\mathbf{X}_{s,t}^{(n)} \triangleq \exp((t-s)l) \in T^{(n)}(V)$ has finite total m-variation, where the exponential is taken over $T^{(n)}(V)$. Therefore, $\mathbf{X}^{(n)}$ is the unique extension of $\mathbf{X}$ to $T^{(n)}(V)$ given by Theorem 2.5. By the definition of signature, $\exp(l)$ is the signature of $\mathbf{X}$ where the exponential is now taken over $T((V))$. The second part of the proposition is a direct consequence of the uniqueness result for signature in [2]. $\square$

References

[1] H. Boedihardjo and X. Geng, Tail asymptotics of the Brownian signature, *Trans. Amer. Math. Soc.* 372 (2019): 585–614.

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

[3] J. Chang, *Ph.D. Thesis*, Oxford, 2018.

[4] M. Chesney, M. Jeanblanc and M. Yor, *Mathematical methods for financial markets*, Springer-Verlag, 2009.

[5] J. Chang, T. Lyons and H. Ni, Super-multiplicativity and a lower bound for the decay of the signature of a path of finite length, *C. R. Acad. Sci. Paris* 356 (1) (2018): 720–724.

[6] I. Chevyrev, A characteristic function on the space of signatures of geometric rough paths, *arXiv: 1307.3580v2*, 2013.

[7] I. Chevyrev and T. Lyons, Characteristic functions of measures on geometric rough paths, *Ann. Probab.* 44 (6) (2016): 4049–4082.

[8] A.M. Davie, Differential equations driven by rough paths: an approach via discrete approximation, *Appl. Math. Res. Express.* (2) (2007): Art. ID abm009.

[9] J. Diestel and J.J. Uhl, *Vector measures*, American Mathematical Society, 1977.

[10] P. Friz and S. Riedel, Integrability of linear rough differential equations and moment estimates for iterated integrals of Gaussian processes, *arXiv: 1104.0577v2*, 2011.

[11] P. Friz and N. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge University Press, 2010.

[12] X. Geng, Reconstruction for the signature of a rough path, *Proc. Lond. Math. Soc.* 114 (3) (2017): 495–526.

[13] M. Gubinelli, Controlling rough paths, *J. Funct. Anal.* 216 (1) (2004): 86– 140.

[14] M. Gubinelli, Ramifications of rough paths, *J. Different. Equations* 248 (4) (2010): 693–721.

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

[16] K. Hara and M. Hino, Fractional order Taylor's series and the neo-classical inequality, *Bull. London Math. Soc.* 42 (2010): 467–477.

[17] S. Helgason, Differential geometry, Lie groups, and symmetric spaces, Academic Press, 1978.

[18] J. E. Humphreys, *Introduction to Lie algebras and representation theory*, Springer-Verlag, 1972.

[19] G. Kallianpur, *Stochastic filtering theory*, Springer-Verlag, 1980.

[20] T. Kato, *A short introduction to perturbation theory for linear operators*, Springer-Verlag, 1982.

[21] T. Lyons, Differential equations driven by rough signals, *Rev. Mat. Iberoamericana* 14 (2) (1998): 215–310.

[22] T. Lyons and Z. Qian, *System control and rough paths*, Oxford University Press, 2002.

[23] T. Lyons and N. Sidorova, On the radius of convergence of the logarithmic signature, *Illinois J. Math.* 50 (4) (2006): 763–790.

[24] T. Lyons and W. Xu, Hyperbolic development and the inversion of signature, *J. Funct. Anal.* 272 (7) (2015): 2933–2955.

[25] T. Lyons and W. Xu, Inverting the signature of a path, *J. Eur. Math. Soc.* 20 (7) (2018): 1655–1687.

[26] B. Øksendal, *Stochastic differential equations: an introduction with applications*, Springer Science and Business Media, 2013.

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

[28] R.A. Ryan, *Introduction to tensor products of Banach spaces*, Springer, 2002.

[29] N.N. Vorob'ev, Estimates of real roots of a system of algebraic equations, *J. Sov. Math.* 34 (4) (1986): 1754–1762.

[30] J. Yong and X. Zhou, *Stochastic controls*, Springer-Verlag, 1999.