Tail Asymptotics of the Brownian Signature

H. Boedihardjo[*], X. Geng[†]

Abstract

In the groundbreaking work of B. Hambly and T. Lyons (Uniqueness for the signature of a path of bounded variation and the reduced path group, *Ann. of Math.* 2010), it has been conjectured that the geometry of a tree-reduced bounded variation path can be recovered from the tail asymptotics of its associated sequence of iterated path integrals. While this conjecture is still remaining open in the general deterministic case, in the present article we investigate a similar problem in the probabilistic setting for Brownian motion. It turns out that a martingale approach applied to the hyperbolic development of Brownian motion allows us to extract useful information from the tail asymptotics of Brownian iterated integrals, which can be used to determined the Brownian rough path along with its natural parametrization uniquely. This in particular strengthens the existing uniqueness results in the literature.

1 Introduction

Every continuous path $\gamma : [0,T] \to \mathbb{R}^d$ with bounded variation is naturally associated with a sequence

$$g \triangleq \left\{ \int_{0<t_1<\cdots<t_n<T} d\gamma_{t_1} \otimes \cdots \otimes d\gamma_{t_n} : n \in \mathbb{N} \right\} \tag{1.1}$$

of iterated path integrals, which is known as the signature of γ. In the renowned work of Hambly and Lyons [10], it was proved that every continuous path with

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

[†]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213, United States. Email: xig@andrew.cmu.edu.

bounded variation is uniquely determined by its signature up to a tree-like equivalence. This result was extended to the general rough path case in [3].

On the one hand, from the uniqueness results, it is seen that every tree-like equivalence class contains a unique representative path γ with given signature g, which does not contain any tree-like pieces. This representative path is called the tree-reduced path. On the other hand, from the algebraic shuffle product structure of signatures, every finite degree component of a signature element g can be recovered from looking at the component of g at degree "infinite" (i.e. the tail of g). Therefore, it is natural and reasonable to expect that some intrinsic geometric properties associated with a tree-reduced rough path can be explicitly recovered from the tail behavior of its signature.

In the bounded variation case, it was proved that the length of a path γ can be recovered from the tail asymptotics of its signature g in the following way:

$$\|\gamma\|_{1-\text{var}} = \lim_{n\to\infty} \left(n!\|g_n\|_{\text{proj}}\right)^{\frac{1}{n}}, \tag{1.2}$$

provided that $\gamma \in C^1$ when parametrized by unit speed and the modulus of continuity $\delta_{\gamma'}$ for γ' satisfies $\delta_{\gamma'}(\varepsilon) = o(\varepsilon^{3/4})$ as $\varepsilon \downarrow 0$. Here g_n is the homogeneous component in degree n of the signature g and the tensor norm is the projective norm induced by the Euclidean norm on $\mathbb{R}^d$. The same result can also be proved for piecewise linear paths and monotonely increasing paths. It has been conjectured that the same result should hold for all tree-reduced continuous path with bounded variation. However, very little progress has been made towards a complete solution.

For an arbitrary continuous path with bounded variation, one can easily see that

$$\|g_n\|_{\text{proj}} \leqslant \frac{\|\gamma\|_{1-\text{var}}^n}{n!}, \quad \forall n \in \mathbb{N}.$$

So the length conjecture (1.2) is about establishing a matching lower bound. If proved to be true in general, it will indicate that for a tree-reduced path, the signature components decay in an exact factorial rate. The original idea of Hambly and Lyons for proving (1.2) in the C^1-case is to look at the lifting X^λ of $\lambda \cdot \gamma$ (rescaling γ by a large constant λ) to the hyperbolic manifold of constant curvature -1 (the hyperbolic development). It turns out that when $\lambda \to \infty$, X^λ becomes more and more like a hyperbolic geodesic in the sense that the hyperbolic distance between the two endpoints of X^λ is asymptotically comparable to its hyperbolic length. As a simple consequence of the nature of hyperbolic development, the first quantity is related to the signature of γ is a fairly explicit way, while the second quantity is the same as the original length. In this way, one sees a lower bound for the signature in terms of the length. It seems to us that in the deterministic setting, the technique of hyperbolic development is essentially a C^1-technique which requires major modification in the general bounded variation case in a quite fundamental way.

In parallel, we could certainly ask a similar question in the rough path context. According to Lyons [13], for a rough path $\mathbf{X}$ with roughness $p \geqslant 1$, the signature estimate takes the form

$$\|g_n\|_{\text{proj}} \leqslant \frac{\omega(\mathbf{X})^{\frac{n}{p}}}{\left(\frac{n}{p}\right)!}, \quad \forall n \in \mathbb{N},$$

where $\omega(\mathbf{X})$ is a constant depending on the p-variation of $\mathbf{X}$. To expect an analogue of (1.2) which is not even clear at this point, it is natural to search lower bounds for g_n of the same form and look at the quantity

$$\widetilde{L}_p \triangleq \limsup_{n\to\infty} \left(\left(\frac{n}{p}\right)! \|g_n\|_{\text{proj}}\right)^{\frac{p}{n}}.$$

On the one hand, the reason of looking at the "limsup" instead of an actual limit is that unlike the bounded variation case, one could easily find examples of tree-reduced geometric rough paths with infinitely many zero signature components (for instance $\mathbf{X}_t \triangleq \exp(t[v,w])$ for certain vectors $v, w \in \mathbb{R}^d$). One might expect that $\widetilde{L}_p$ recovers the p-variation of the underlying rough path. However, if the length conjecture (1.2) is true for bounded variation paths, this cannot be the case since $\widetilde{L}_p = 0$ for a bounded variation path when $p > 1$. On the other hand, if we define the local p-variation of a rough path in the same way as the usual p-variation but additionally by requiring that the mesh size of partitions goes to zero instead of taking supremum over all partitions, it is easy to see that the local p-variation of a bounded variation path is also zero when $p > 1$. Therefore, it is not entirely unreasonable to expect that the quantity $\widetilde{L}_p$ recovers the local p-variation of $\mathbf{X}$.

In the present article, we investigate a similar problem for the Brownian rough path $\mathbf{B}_t$, which is the canonical lifting of the Brownian motion B_t as geometric p-rough paths for $2 < p < 3$. One can equivalently view it as the Brownian motion coupled with the Lévy area process. It is well known that B_t has a quadratic variation process, which can be viewed as the local 2-variation of Brownian motion in certain probabilistic sense. In view of the previous discussion, if we define the normalized "limsup"

$$\widetilde{L}_{s,t} \triangleq \limsup_{n\to\infty} \left(\left(\frac{n}{2}\right)! \left\| \int_{s<t_1<\cdots<t_n<t} \circ dB_{t_1} \otimes \cdots \otimes \circ dB_{t_n} \right\| \right)^{\frac{2}{n}} \tag{1.3}$$

for the Brownian signature path under suitable tensor norms, one might expect that $\widetilde{L}_{s,t}$ recovers some sort of quadratic variation of the Brownian rough path. The aim of the present article is to establish a result of this kind. It is a priori unclear that $\widetilde{L}_{s,t}$ is even finite since Brownian motion has infinite 2-variation almost surely.

We are going to show that $\widetilde{L}_{s,t}$ is a deterministic multiple of $t-s$: $\widetilde{L}_{s,t} = \kappa(t-s)$ for some deterministic constant κ. This implies that the natural speed of Brownian motion (i.e. its quadratic variation) can be recovered from the tail asymptotics of its signature. In addition, we establish upper and lower bounds on the constant κ.

On the one hand, the upper estimate is shown by using general rough path arguments and does not reflect the tree-reduced nature of the Brownian rough path at all. The deterministic nature of $\widetilde{L}_{s,t}$ comes from the fact that Brownian motion has independent increments. The result holds under a wide choice of tensor norms.

On the other hand, the lower estimate is obtained by considering the hyperbolic development of Brownian motion. Although we also work in the hyperbolic situation, our calculation diverges early on from the work of Hambly and Lyons [10], which makes use of martingale arguments instead of deterministic hyperbolic analysis. Our lower estimate allows us to conclude that the Brownian rough path is tree-reduced with probability one and also its natural parametrization can be recovered from the tail asymptotics of the Brownian signature. In particular, with probability one, every Brownian rough path is uniquely determined by its signature. This result is stronger than the existing uniqueness results for Brownian motion in the literature (c.f. [2], [12]), since it was only known that the signature determines the Brownian rough path up to reparametrization.

Our main result on the upper and lower estimates of $\widetilde{L}_{s,t}$ can be summarized as follows.

Theorem 1.1. *Let $B_t = (B_t^1, \cdots, B_t^d)$ be a d-dimensional Brownian motion ($d \geqslant 2$). Define $\widetilde{L}_{s,t}$ as in (1.3) under suitable tensor norms to be specified in the following.*

(1) (upper estimate) Given arbitrary admissible tensor norms under which each element of the canonical basis $\{\mathrm{e}_1, \cdots, \mathrm{e}_d\}$ of $\mathbb{R}^d$ has norm one, there exists

a deterministic constant $\kappa_d \leqslant d^2$ depending on the choice of tensor norms, such that with probability one

$$\widetilde{L}_{s,t} = \kappa_d(t-s) \quad \forall s < t.$$

(2) (lower estimate) Under the l^p-norm ($1 \leqslant p \leqslant 2$) on $\mathbb{R}^d$ and the associated projective tensor norms on the tensor products, we have

$$\kappa_d \geqslant \frac{d-1}{2}.$$

Remark 1.1. The one dimensional case ($d = 1$) is uninteresting and the result of Theorem 1.1 holds trivially since $\widetilde{L}_{s,t} \equiv 0$ in this case.

It worths mentioning that the idea of hyperbolic development was also used by Lyons and Xu [15] to recover the derivative of a unit speed C^1-path at the endpoint. There are also some recent works on recovering the full trajectory of a path from its signature. See for instance [6], [9], [16].

Our article is organized in the following way. In Section 2, we present some basic notions from rough path theory which are needed for our analysis. In Section 3, we prove the first part of Theorem 1.1. In Section 4, we prove the second part of Theorem 1.1. In that section we also present some crucial details for understanding the hyperbolic development which seems to be incomplete or missing in the literature. In Section 5, we present some interesting applications of our main result to the Brownian rough path itself. In Section 6, we give some concluding remarks and discuss a few related further problems.

2 Notions from rough path theory

In this section, we present some basic notions from rough path theory which are needed for our study. We refer the reader to the monographs [5], [8], [14] for a systematic introduction.

Suppose that V is a finite dimensional normed vector space. For each $n \in \mathbb{N}$, define $T^{(n)}(V) \triangleq \oplus_{i=0}^{n} V^{\otimes i}$, and let $T((V))$ be the algebra of formal sequences of homogeneous tensors $a = (a_0, a_1, a_2, \cdots)$ with $a_n \in V^{\otimes n}$ for each n.

Definition 2.1. A family of tensor norms $\{\|\cdot\|_{V^{\otimes n}} : \ n \geqslant 1\}$ on the tensor products is called *admissible* if

(1) for any $a \in V^{\otimes m}$ and $b \in V^{\otimes m}$,

$$\|a \otimes b\|_{V^{\otimes(m+n)}} \leqslant \|a\|_{V^{\otimes m}} \|b\|_{V^{\otimes n}}; \tag{2.1}$$

(2) for any permutation $\sigma \in \mathcal{S}_n$ and $a \in V^{\otimes n}$,

$$\|\mathcal{P}^{\sigma}(a)\|_{V^{\otimes n}} = \|a\|_{V^{\otimes n}},$$

where $\mathcal{P}^{\sigma}$ is the linear operator on $V^{\otimes n}$ induced by $a_1 \otimes \cdots \otimes a_n \mapsto a_{\sigma(1)} \otimes \cdots \otimes a_{\sigma(n)}$ for $a_1, \cdots, a_n \in V$.

We call it a family of *cross-norms* if the inequality in (2.1) is an equality.

Definition 2.2. The *projective tensor norm* on $V^{\otimes n}$ is defined to be

$$\|a\|_{\mathrm{proj}} \triangleq \inf\left\{\sum_l |a_1^{(l)}| \cdots |a_n^{(l)}| : \text{ if } a = \sum_l a_1^{(l)} \otimes \cdots \otimes a_n^{(l)}\right\}.$$

It is known that the projective tensor norm is the largest cross-norm on $V^{\otimes n}$. In the case when $V = \mathbb{R}^d$ is equipped with the l^1-norm, one can see by definition that the projective tensor norm on $V^{\otimes n}$ is just the l^1-norm under the canonical tensor basis induced from the one on $\mathbb{R}^d$.

We assume that V is equipped with a family of admissible tensor norms. Define $\triangle \triangleq \{(s,t) : 0 \leqslant s \leqslant t \leqslant 1\}$. Given $p \geqslant 1$, we denote $\lfloor p \rfloor$ as the largest integer not exceeding p.

Definition 2.3. A *multiplicative functional of degree* $n \in \mathbb{N}$ is a continuous map $\mathbf{X}_{\cdot,\cdot} = \left(1, \mathbb{X}^1_{\cdot,\cdot}, \cdots, \mathbb{X}^n_{\cdot,\cdot}\right) : \triangle \to T^n(V)$ which satisfies

$$\mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} = \mathbf{X}_{s,t}, \text{ for } 0 \leqslant s \leqslant u \leqslant t \leqslant 1.$$

Let $\mathbf{X}, \mathbf{Y}$ be two multiplicative functionals of degree n. Define

$$d_p(\mathbf{X}, \mathbf{Y}) \triangleq \max_{1 \leqslant i \leqslant n} \sup_{\mathcal{P}} \left(\sum_l \left\|\mathbb{X}^i_{t_{l-1},t_l} - \mathbb{Y}^i_{t_{l-1},t_l}\right\|_{V^{\otimes i}}^{\frac{p}{i}}\right)^{\frac{i}{p}},$$

where the supremum is taken over all possible finite partitions of $[0,1]$. d_p is called the *p-variation metric*. If $d_p(\mathbf{X}, \mathbf{1}) < \infty$ where $\mathbf{1} = (1, 0, \cdots, 0)$, we say that $\mathbf{X}$ has *finite p-variation*. A multiplicative functional of degree $\lfloor p \rfloor$ with finite p-variation is called a *p-rough path*.

The following important result, proved by Lyons [13], asserts that "iterated path integrals" for a rough path are also well defined.

Theorem 2.1 (Lyons' extension theorem)**.** *Let* $\mathbf{X} = (1, \mathbb{X}^1, \cdots, \mathbb{X}^{\lfloor p \rfloor})$ *be a p-rough path. Then for any* $n \geqslant \lfloor p \rfloor + 1$*, there exists a unique continuous map* $\mathbb{X}^n : \ \Delta \to V^{\otimes n}$*, such that*

$$\mathbb{X}_{\cdot,\cdot} \triangleq \left(1, \mathbb{X}^1_{\cdot,\cdot}, \cdots, \mathbb{X}^{\lfloor p \rfloor}_{\cdot,\cdot}, \cdots, \mathbb{X}^n_{\cdot,\cdot}, \cdots\right)$$

is a multiplicative functional in $T((V))$ *whose projection onto* $T^{(n)}(V)$ *has finite p-variation for every n.*

Remark 2.1. Due to the multiplicative structure, when we consider a rough path, one could simply look at the path $t \mapsto \mathbf{X}_{0,t}$ whose increments are defined to be $\mathbf{X}_s^{-1} \otimes \mathbf{X}_t$.

Definition 2.4. Let $\mathbf{X}$ be a p-rough path. The path $t \mapsto \mathbb{X}_{0,t} \in T((V))$ defined by Lyons' extension theorem is called the *signature path* of $\mathbf{X}$. The quantity $\mathbb{X}_{0,1}$ is called the *signature* of $\mathbf{X}$.

It was also proved by Lyons [13] in his extension theorem that the signature $\mathbb{X}_{0,1}$ satisfies the following factorial decay estimate:

$$\|\mathbb{X}^n\| \leqslant \frac{\omega(\mathbf{X})^{\frac{n}{p}}}{\left(\frac{n}{p}\right)!}, \quad \forall n \geqslant 1,$$

where $\omega(\mathbf{X})$ is a constant depending on the p-variation of $\mathbf{X}$.

When $p = 1$ and $\mathbf{X}$ is a continuous path with bounded variation, all the previous notions reduces to the classical iterated path integrals defined in the sense of Lebesgue-Stieltjes.

Among general rough paths there is a fundamental class of paths called geometric rough paths.

Definition 2.5. For a continuous path $\gamma : \ [0,1] \to V$, define

$$\mathbb{X}^n_{s,t} = \int_{s<u_1<\cdots<u_n<t} d\gamma_{u_1} \otimes \cdots \otimes d\gamma_{u_n}, \quad n \geqslant 1, \ s \leqslant t.$$

The closure of the space

$$\{(1, \mathbb{X}^1_{s,t}, \ldots, \mathbb{X}^{\lfloor p \rfloor}_{s,t}) : \ \gamma \text{ is a continuous path with bounded variation}\}$$

under the p-variation metric d_p is called the space of *geometric p-rough paths*.

The space of geometric rough paths plays a fundamental role in rough path theory. In particular, a complete integration and differential equation theory with respect to geometric rough paths has been established by Lyons [13]. The rough path theory has significant applications in probability theory, mainly due to the fact that a wide class of interesting stochastic processes can be regarded as geometric rough paths in a canonical way in the sense of natural approximations.

In particular, it is known that (c.f. [18]) a multidimensional Brownian motion B_t admits a canonical lifting as geometric p-rough path $\mathbf{B}_t$ with $p \in (2,3)$. $\mathbf{B}_t$ is called the *Brownian rough path.* The corresponding *Brownian signature path*, determined by Lyons' extension theorem, is denoted as

$$\mathbb{B}_{s,t} = (1, \mathbb{B}^1_{s,t}, \mathbb{B}^2_{s,t}, \cdots), \quad s \leqslant t.$$

Under the canonical tensor basis on tensor products over $V \triangleq \mathbb{R}^d$, for each word $(i_1, \cdots, i_n)$ over $\{1, \cdots, d\}$, the coefficient of $\mathbb{B}^n_{s,t}$ with respect to $\mathrm{e}_{i_1} \otimes \cdots \otimes \mathrm{e}_{i_n}$ coincides with the iterated Stratonovich integral (c.f. [5]):

$$\mathbb{B}^{n;i_1,\cdots,i_n}_{s,t} = \int_{s<u_1<\cdots<u_n<t} \circ dB^{i_1}_{u_1} \cdots \circ dB^{i_n}_{u_n}.$$

For a given family of admissible tensor norms, we define

$$\widetilde{L}_{s,t} \triangleq \limsup_{n\to\infty} \left(\left(\frac{n}{2}\right)! \left\|\mathbb{B}^n_{s,t}\right\| \right)^{\frac{2}{n}}, \quad s \leqslant t.$$

3 First part of the main result: the upper estimate

In this section, we develop the proof of the first part of Theorem 1.1.

Lemma 3.1. *The signature coefficients $\mathbb{B}^{n;i_1,\cdots,i_n}_{s,t}$ satisfy the following estimate:*

$$\mathbb{E}\left[\sup_{s\leqslant u\leqslant t} \left|\mathbb{B}^{n;i_1,\cdots,i_n}_{s,u}\right|\right] \leqslant \left(\frac{1}{2} + \sqrt{2}\right) \left(\frac{\mathrm{e}}{\sqrt{2\pi}}\right)^{\frac{1}{2}} \frac{2^{\frac{n}{2}}}{(n-2)^{\frac{1}{4}}\sqrt{n!}} (t-s)^{\frac{n}{2}}$$

for all $s < t$, $n \geqslant 1$ and $1 \leqslant i_1, \cdots, i_n \leqslant d$.

Proof. By translation, it suffices to consider the case when $s = 0$.

We first estimate the second moment of $\mathbb{B}^{n;i_1,\cdots,i_n}_{0,u}$. According to the shuffle product formula for signatures (c.f. [5]),

$$\mathbb{E}\left[\left|\mathbb{B}^{n;i_1,\cdots,i_n}_{0,u}\right|^2\right] = \sum_{\sigma\in\mathcal{S}(n,n)} \mathbb{E}\left[\mathbb{B}^{2n;j_{\sigma^{-1}(1)},\cdots,j_{\sigma^{-1}(2n)}}_{0,u}\right],$$

where $(j_1, \cdots, j_{2n}) \triangleq (i_1, \cdots, i_n, i_1, \cdots, i_n)$ and $\mathcal{S}(n,n)$ is the set of (n,n)-shuffles in the permutation group of order $2n$. On the other hand, by the formula of the Brownian expected signature (c.f. [7]), we know that

$$\mathbb{E}[\mathbb{B}_{0,u}] = \exp\left(\frac{u}{2}\sum_{i=1}^{d} \mathrm{e}_i \otimes \mathrm{e}_i\right).$$

It follows that

$$\mathbb{E}\left[\mathbb{B}_{0,u}^{2n}\right] = \frac{u^n}{n!2^n}\left(\sum_{i=1}^{d} \mathrm{e}_i \otimes \mathrm{e}_i\right)^{\otimes n}.$$

In particular, every coefficient of $\mathbb{E}\left[\mathbb{B}_{0,u}^{2n}\right]$ is either zero or $\frac{u^n}{n!2^n}$. Therefore,

$$\begin{aligned}
\mathbb{E}\left[\left|\mathbb{B}_{0,u}^{n;i_1,\cdots,i_n}\right|^2\right] &\leqslant \frac{(2n)!}{(n!)^2}\cdot\frac{u^n}{n!2^n} \\
&\leqslant \frac{\mathrm{e}(2n)^{2n+\frac{1}{2}}\mathrm{e}^{-2n}}{2\pi n^{2n+1}\mathrm{e}^{-2n}}\cdot\frac{u^n}{n!2^n} \\
&= \frac{\mathrm{e}}{\sqrt{2\pi}}\frac{2^n}{\sqrt{n}n!}u^n, \qquad (3.1)
\end{aligned}$$

where in the second inequality we have used Stirling's approximation.

Secondly, according to the differential equation for the signature path, we have

$$\begin{aligned}
d\mathbb{B}_{0,u}^{n;i_1,\cdots,i_n} &= \mathbb{B}_{0,u}^{n-1;i_1,\cdots,i_{n-1}} \circ dB_u^{i_n} \\
&= \mathbb{B}_{0,u}^{n-1;i_1,\cdots,i_{n-1}} \cdot dB_u^{i_n} + \frac{1}{2} d\mathbb{B}_{0,u}^{n-1;i_1,\cdots,i_{n-1}} \cdot dB_u^{i_n} \\
&= \mathbb{B}_{0,u}^{n-1;i_1,\cdots,i_{n-1}} \cdot dB_u^{i_n} + \frac{1}{2}\left(\mathbb{B}_{0,u}^{n-2;i_1,\cdots,i_{n-2}} \circ dB_u^{i_{n-1}}\right) \cdot dB_u^{i_n} \\
&= \mathbb{B}_{0,u}^{n-1;i_1,\cdots,i_{n-1}} \cdot dB_u^{i_n} + \frac{1}{2}\delta_{i_{n-1},i_n}\mathbb{B}_{0,u}^{n-2;i_1,\cdots,i_{n-2}} du.
\end{aligned}$$

It follows from (3.1) that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0\leqslant u\leqslant t}\left|\mathbb{B}_{0,u}^{n;i_1,\cdots,i_n}\right|\right]\\
&\leqslant \mathbb{E}\left[\sup_{0\leqslant u\leqslant t}\left|\int_0^u \mathbb{B}_{0,v}^{n-1;i_1,\cdots,i_{n-1}}\cdot dB_v^{i_n}\right|\right]+\frac{1}{2}\int_0^t\sqrt{\mathbb{E}\left[\left|\mathbb{B}_{0,u}^{n-2;i_1,\cdots,i_{n-2}}\right|^2\right]}du\\
&\leqslant \mathbb{E}\left[\sup_{0\leqslant u\leqslant t}\left|\int_0^u \mathbb{B}_{0,v}^{n-1;i_1,\cdots,i_{n-1}}\cdot dB_v^{i_n}\right|\right]+\frac{1}{2}\int_0^t\left(\frac{\mathrm{e}}{\sqrt{2}\pi}\frac{2^{n-2}}{\sqrt{n-2}(n-2)!}u^{n-2}\right)^{\frac{1}{2}}du\\
&= \mathbb{E}\left[\sup_{0\leqslant u\leqslant t}\left|\int_0^u \mathbb{B}_{0,v}^{n-1;i_1,\cdots,i_{n-1}}\cdot dB_v^{i_n}\right|\right]+\frac{1}{2}\left(\frac{\mathrm{e}}{\sqrt{2}\pi}\right)^{\frac{1}{2}}\frac{2^{\frac{n}{2}}}{(n-2)^{\frac{1}{4}}\sqrt{(n-2)!n}}t^{\frac{n}{2}}\\
&\leqslant \mathbb{E}\left[\sup_{0\leqslant u\leqslant t}\left|\int_0^u \mathbb{B}_{0,v}^{n-1;i_1,\cdots,i_{n-1}}\cdot dB_v^{i_n}\right|\right]+\frac{1}{2}\left(\frac{\mathrm{e}}{\sqrt{2}\pi}\right)^{\frac{1}{2}}\frac{2^{\frac{n}{2}}}{(n-2)^{\frac{1}{4}}\sqrt{n!}}t^{\frac{n}{2}}
\end{aligned}
$$

The first term can be estimated easily by using Doob's L^p-inequality:

$$
\begin{aligned}
\mathbb{E}\left[\sup_{0\leqslant u\leqslant t}\left|\int_0^u \mathbb{B}_{0,v}^{n-1;i_1,\cdots,i_{n-1}}\cdot dB_v^{i_n}\right|\right] &\leqslant \left\|\sup_{0\leqslant u\leqslant t}\left|\int_0^u \mathbb{B}_{0,v}^{n-1;i_1,\cdots,i_{n-1}}\cdot dB_v^{i_n}\right|\right\|_2\\
&\leqslant 2\left\|\int_0^t \mathbb{B}_{0,v}^{n-1;i_1,\cdots,i_{n-1}}\cdot dB_v^{i_n}\right\|_2\\
&= 2\left(\int_0^t \mathbb{E}\left[\left|\mathbb{B}_{0,v}^{n-1;i_1,\cdots,i_{n-1}}\right|^2\right]dv\right)^{\frac{1}{2}}\\
&\leqslant 2\left(\int_0^t \frac{\mathrm{e}}{\sqrt{2}\pi}\frac{2^{n-1}}{\sqrt{n-1}(n-1)!}v^{n-1}dv\right)^{\frac{1}{2}}\\
&= \sqrt{2}\left(\frac{\mathrm{e}}{\sqrt{2}\pi}\right)^{\frac{1}{2}}\frac{2^{\frac{n}{2}}}{(n-2)^{\frac{1}{4}}\sqrt{n!}}t^{\frac{n}{2}}.
\end{aligned}
$$

Now the desired estimate follows immediately. □

Remark 3.1. Second moment estimate on iterated Stratonovich's integrals was studied by Ben Arous [1] through iterated Itô's integrals. Here the estimate (3.1) we obtained through the shuffle product formula and the Brownian expected signature is sharper in the exponential factor.

Now we are able to establish the following main upper estimate.

Proposition 3.1. *Suppose that the tensor products* $(\mathbb{R}^d)^{\otimes n}$ *are equipped with given admissible norms, under which each element of the standard basis* $\{\mathrm{e}_1, \cdots, \mathrm{e}_d\}$ *of* $\mathbb{R}^d$ *has norm one. Then for each* $s < t$*, with probability one, we have*

$$\max\left\{\limsup_{n\to\infty}\left(\left(\frac{n}{2}\right)!\sup_{s\leqslant u\leqslant t}\left\|\mathbb{B}^n_{s,u}\right\|\right)^{\frac{2}{n}}, \limsup_{n\to\infty}\left(\left(\frac{n}{2}\right)!\sup_{s\leqslant u\leqslant t}\left\|\mathbb{B}^n_{u,t}\right\|\right)^{\frac{2}{n}}\right\} \leqslant d^2(t-s).$$

Proof. Since the tensor norms are admissible, we have

$$\begin{aligned}\left\|\mathbb{B}^n_{s,u}\right\| &= \left\|\sum_{i_1,\cdots,i_n=1}^{d}\mathbb{B}^{n;i_1,\cdots,i_n}_{s,u}\mathrm{e}_{i_1}\otimes\cdots\otimes\mathrm{e}_{i_n}\right\| \\ &\leqslant \sum_{i_1,\cdots,i_n=1}^{d}\left|\mathbb{B}^{n;i_1,\cdots,i_n}_{s,u}\right|,\end{aligned}$$

and thus

$$\mathbb{E}\left[\sup_{s\leqslant u\leqslant t}\|\mathbb{B}^n_{s,t}\|\right] \leqslant \sum_{i_1,\cdots,i_n=1}^{d}\mathbb{E}\left[\sup_{s\leqslant u\leqslant t}\left|\mathbb{B}^{n;i_1,\cdots,i_n}_{s,u}\right|\right].$$

According to Lemma 3.1, we arrive at

$$\mathbb{E}\left[\sup_{s\leqslant u\leqslant t}\|\mathbb{B}^n_{s,u}\|\right] \leqslant d^n\cdot\frac{C2^{\frac{n}{2}}}{(n-2)^{\frac{1}{4}}\sqrt{n!}}(t-s)^{\frac{n}{2}},$$

where

$$C \triangleq \left(\frac{1}{2}+\sqrt{2}\right)\left(\frac{\mathrm{e}}{\sqrt{2\pi}}\right)^{\frac{1}{2}}.$$

Now for each $r > (t-s)$, we have

$$\mathbb{P}\left(\sup_{s\leqslant u\leqslant t}\|\mathbb{B}^n_{s,u}\| > \frac{Cd^n2^{\frac{n}{2}}}{(n-2)^{\frac{1}{4}}\sqrt{n!}}r^{\frac{n}{2}}\right) \leqslant \left(\frac{t-s}{r}\right)^{\frac{n}{2}}.$$

By the Borel-Cantelli lemma, with probability one (with null set depending on s and t),

$$\sup_{s\leqslant u\leqslant t}\|\mathbb{B}^n_{s,u}\| \leqslant \frac{Cd^n2^{\frac{n}{2}}}{(n-2)^{\frac{1}{4}}\sqrt{n!}}r^{\frac{n}{2}}$$

for all sufficiently large n. It follows from Stirling's approximation that with probability one,

$$\limsup_{n\to\infty}\left(\left(\frac{n}{2}\right)!\sup_{s\leqslant u\leqslant t}\left\|\mathbb{B}^n_{s,u}\right\|\right)^{\frac{2}{n}} \leqslant \lim_{n\to\infty}\left(\left(\frac{n}{2}\right)!\frac{Cd^n2^{\frac{n}{2}}}{(n-2)^{\frac{1}{4}}\sqrt{n!}}r^{\frac{n}{2}}\right)^{\frac{2}{n}} = d^2r.$$

By taking a rational sequence $r \downarrow (t-s)$, we conclude that with probability one (with null set depending on t),

$$\limsup_{n\to\infty}\left(\left(\frac{n}{2}\right)!\sup_{s\leqslant u\leqslant t}\left\|\mathbb{B}^n_{s,u}\right\|\right)^{\frac{2}{n}} \leqslant d^2(t-s). \tag{3.2}$$

For the estimate involving $\mathbb{B}^n_{u,t}$, observe that

$$\begin{aligned}\mathbb{B}^n_{u,t} &= \int_{u<v_1<\cdots<v_n<t} dB_{v_1}\otimes\cdots\otimes dB_{v_n}\\ &= \int_{0<r_n<\cdots<r_1<t-u} dB_{t-r_1}\otimes\cdots\otimes dB_{t-r_n}\\ &= \mathcal{P}^\tau\left(\int_{0<r_1<\cdots<r_n<t-u} dW_{r_1}\otimes\cdots\otimes dW_{r_n}\right),\end{aligned}$$

where $W_r \triangleq B_{t-r}-B_t$ $(0\leqslant r\leqslant t-u)$ is again a Brownian motion and $\mathcal{P}^\tau$ is the linear transformation on $\left(\mathbb{R}^d\right)^{\otimes n}$ determined by $\xi_1\otimes\cdots\otimes\xi_n \mapsto \xi_n\otimes\cdots\otimes\xi_1$. It follows that

$$\|\mathbb{B}^n_{u,t}\| = \left\|\mathbb{W}^n_{0,t-u}\right\|.$$

Therefore,

$$\sup_{s\leqslant u\leqslant t}\|\mathbb{B}^n_{u,t}\| = \sup_{0\leqslant v\leqslant t-s}\|\mathbb{W}^n_{0,v}\|.$$

Therefore, what we have proven before shows that

$$\begin{aligned}\limsup_{n\to\infty}\left(\left(\frac{n}{2}\right)!\sup_{s\leqslant u\leqslant t}\left\|\mathbb{B}^n_{u,t}\right\|\right)^{\frac{2}{n}} &= \limsup_{n\to\infty}\left(\left(\frac{n}{2}\right)!\sup_{0\leqslant v\leqslant t-s}\left\|\mathbb{W}^n_{0,v}\right\|\right)^{\frac{2}{n}}\\ &\leqslant d^2(t-s)\end{aligned}$$

for almost surely. □

Recall that $\widetilde{L}_{s,t}$ is defined by (1.3) under given admissible tensor norms. It is immediate from Proposition 3.1 that $\widetilde{L}_{s,t}\leqslant d^2(t-s)$ for almost surely.

Now we are going to show that $\widetilde{L}_{s,t}$ is almost surely a deterministic constant.

Recall that $g\in T((\mathbb{R}^d))$ is a group-like element if and only if g satisfies the shuffle product formula. In particular, the signature of a geometric rough path is always a group-like element.

Lemma 3.2. *Let $g=(1,g^1,g^2,\cdots)$ be a non-trivial group-like element in the tensor algebra $T((\mathbb{R}^d))$, where the tensor products are equipped with given admissible norms. Then g has infinitely many non-zero components.*

Proof. Suppose that $g^k \neq 0$ for some $k \geqslant 1$. According to the shuffle product formula, for each $n \geqslant 1$,

$$\left(g^k\right)^{\otimes n} = \sum_{\sigma \in \mathcal{S}(k,\cdots,k)} \mathcal{P}^\sigma\left(g^{nk}\right).$$

Since the tensor norms are admissible, we have

$$\begin{aligned} \|g^k\|^n &\leqslant \sum_{\sigma \in \mathcal{S}(k,\cdots,k)} \left\|\mathcal{P}^\sigma\left(g^{nk}\right)\right\| \\ &= \frac{(nk)!}{(k!)^n}\|g^{nk}\|. \end{aligned}$$

In particular, $g^{nk} \neq 0$ for all n. □

Lemma 3.3. *Given $\alpha > 0$, there exists a constant $C > 0$, such that*

$$\frac{\left(\frac{n}{p}\right)!}{\left(\frac{n-\alpha}{p}\right)!} \leqslant C n^{\frac{\alpha}{p}}, \quad \forall n > 2\alpha,\ p \geqslant 1.$$

Proof. According to Stirling's approximation, there exist constants $C_1, C_2 > 0$, such that

$$C_1 \lambda^{\lambda+\frac{1}{2}} \mathrm{e}^{-\lambda} \leqslant \lambda! \leqslant C_2 \lambda^{\lambda+\frac{1}{2}} \mathrm{e}^{-\lambda}, \quad \forall \lambda > 0.$$

Therefore,

$$\begin{aligned} \frac{\left(\frac{n}{p}\right)!}{\left(\frac{n-\alpha}{p}\right)!} &\leqslant \frac{C_2\left(\frac{n}{p}\right)^{\frac{n}{p}+\frac{1}{2}} \mathrm{e}^{-\frac{n}{p}}}{C_1\left(\frac{n-\alpha}{p}\right)^{\frac{n-\alpha}{p}+\frac{1}{2}} \mathrm{e}^{-\frac{n-\alpha}{p}}} \\ &= \frac{C_2}{C_1\,(p\mathrm{e})^{\frac{\alpha}{p}}}\left(1+\frac{\alpha}{n-\alpha}\right)^{\frac{n-\alpha}{p}+\frac{1}{2}} n^{\frac{\alpha}{p}} \\ &\leqslant \frac{\sqrt{2}C_2\mathrm{e}^\alpha}{C_1} n^{\frac{\alpha}{p}}. \end{aligned}$$

Choosing $C \triangleq \sqrt{2}C_2\mathrm{e}^\alpha/C_1$ suffices. □

The following deterministic sub-additivity property is essential for us.

Proposition 3.2. *Suppose that* $\mathbf{X}$ *is a rough path, where the tensor products are equipped with given admissible norms. Let* $p \geqslant 1$ *be a given constant. Define*

$$\widetilde{l}_{s,t} \triangleq \limsup_{n\to\infty} \left\| \left(\frac{n}{p}\right)! \mathbb{X}^n_{s,t} \right\|^{\frac{p}{n}}, \quad s \leqslant t.$$

Then $(s,t) \mapsto \widetilde{l}_{s,t}$ *is sub-additive, i.e.*

$$\widetilde{l}_{s,t} \leqslant \widetilde{l}_{s,u} + \widetilde{l}_{u,t}$$

for $s \leqslant u \leqslant t$.

Proof. We may assume that $\widetilde{l}_{s,u}, \widetilde{l}_{u,t}$ are both finite. Moreover, we may also assume that both of $\mathbb{X}_{s,u}$ and $\mathbb{X}_{u,t}$ are non-trivial, otherwise the desired inequality is trivial due to Chen's identity. From Lemma 3.2, $\mathbb{X}_{s,u}$ and $\mathbb{X}_{u,t}$ have infinitely many non-zero components.

Given integers $\alpha > 2p$ and $n > 2\alpha$, according to Chen's identity, we have

$$\begin{aligned} \left\| \mathbb{X}^n_{s,t} \right\| &= \left\| \sum_{k=0}^{n} \mathbb{X}^k_{s,u} \otimes \mathbb{X}^{n-k}_{u,t} \right\| \\ &\leqslant \sum_{k=0}^{\alpha-1} \left\| \mathbb{X}^{n-k}_{s,u} \right\| \cdot \left\| \mathbb{X}^k_{u,t} \right\| + \sum_{k=n-\alpha+1}^{n} \left\| \mathbb{X}^{n-k}_{s,u} \right\| \cdot \left\| \mathbb{X}^k_{u,t} \right\| \\ &\quad + \sum_{k=\alpha}^{n-\alpha} \left\| \mathbb{X}^{n-k}_{s,u} \right\| \cdot \left\| \mathbb{X}^k_{u,t} \right\| . \end{aligned}$$

Define $(s,t) \mapsto \widetilde{l}^{\alpha}_{s,t} \triangleq \sup_{k \geqslant \alpha} \left\| (k/p)! \mathbb{X}^k_{s,t} \right\|^{p/k}$. It follows that

$$\begin{aligned}
\left\| \mathbb{X}^n_{s,t} \right\| &\leqslant \sum_{k=0}^{\alpha-1} \frac{\left(\widetilde{l}^{\alpha}_{s,u}\right)^{\frac{n-k}{p}}}{\left(\frac{n-k}{p}\right)!} \cdot \left\| \mathbb{X}^k_{u,t} \right\| + \sum_{k=n-\alpha+1}^{n} \frac{\left(\widetilde{l}^{\alpha}_{u,t}\right)^{\frac{k}{p}}}{\left(\frac{k}{p}\right)!} \cdot \left\| \mathbb{X}^{n-k}_{s,u} \right\| \\
&\quad + \sum_{k=\alpha}^{n-\alpha} \frac{\left(\widetilde{l}^{\alpha}_{s,u}\right)^{\frac{n-k}{p}}}{\left(\frac{n-k}{p}\right)!} \cdot \frac{\left(\widetilde{l}^{\alpha}_{u,t}\right)^{\frac{k}{p}}}{\left(\frac{k}{p}\right)!} \\
&\leqslant \sum_{k=0}^{\alpha-1} \frac{\left(\widetilde{l}^{\alpha}_{s,u}\right)^{\frac{n-k}{p}}}{\left(\frac{n-k}{p}\right)!} \cdot \left\| \mathbb{X}^k_{u,t} \right\| + \sum_{k=n-\alpha+1}^{n} \frac{\left(\widetilde{l}^{\alpha}_{u,t}\right)^{\frac{k}{p}}}{\left(\frac{k}{p}\right)!} \cdot \left\| \mathbb{X}^{n-k}_{s,u} \right\| \\
&\quad + p \frac{\left(\widetilde{l}^{\alpha}_{s,u} + \widetilde{l}^{\alpha}_{u,t}\right)^{\frac{n}{p}}}{\left(\frac{n}{p}\right)!},
\end{aligned}$$

where we have used the neo-classical inequality (c.f. [11]), which states that

$$\sum_{i=0}^{N} \frac{a^{\frac{i}{p}} b^{\frac{N-i}{p}}}{\left(\frac{i}{p}\right)! \left(\frac{N-i}{p}\right)!} \leqslant p \frac{(a+b)^{\frac{N}{p}}}{\left(\frac{N}{p}\right)!}, \quad \forall a, b \geqslant 0, p \geqslant 1, N \in \mathbb{N}.$$

According to Lemma 3.3,

$$\begin{aligned}
&\left(\frac{n}{p}\right)! \left\| \mathbb{X}^n_{s,t} \right\| \\
&\leqslant \frac{\left(\frac{n}{p}\right)!}{\left(\frac{n-\alpha}{p}\right)!} \left(\sum_{k=0}^{\alpha-1} \left(\widetilde{l}^{\alpha}_{s,u}\right)^{\frac{n-k}{p}} \cdot \left\| \mathbb{X}^k_{u,t} \right\| + \sum_{k=n-\alpha+1}^{n} \left(\widetilde{l}^{\alpha}_{u,t}\right)^{\frac{k}{p}} \cdot \left\| \mathbb{X}_{s,u} \right\|^{n-k} \right) \\
&\quad + p \left(\widetilde{l}^{\alpha}_{s,u} + \widetilde{l}^{\alpha}_{u,t}\right)^{\frac{n}{p}} \\
&\leqslant C n^{\frac{\alpha}{p}} \left(\sum_{k=0}^{\alpha-1} \left(\widetilde{l}^{\alpha}_{s,u}\right)^{\frac{n-k}{p}} \cdot \left\| \mathbb{X}^k_{u,t} \right\| + \sum_{k=n-\alpha+1}^{n} \left(\widetilde{l}^{\alpha}_{u,t}\right)^{\frac{k}{p}} \cdot \left\| \mathbb{X}_{s,u} \right\|^{n-k} \right) \\
&\quad + p \left(\widetilde{l}^{\alpha}_{s,u} + \widetilde{l}^{\alpha}_{u,t}\right)^{\frac{n}{p}}.
\end{aligned}$$

Therefore,

$$\widetilde{l}_{s,t} = \limsup_{n \to \infty} \left(\left(\frac{n}{p}\right)! \left\| \mathbb{X}^n_{s,t} \right\| \right)^{\frac{p}{n}} \leqslant \widetilde{l}^{\alpha}_{s,u} + \widetilde{l}^{\alpha}_{u,t},$$

where we have used the simple fact that

$$\lim_{n\to\infty}\left(\left(\lambda a^{\frac{n}{p}}+\mu b^{\frac{n}{p}}\right)n^{\nu}+(a+b)^{\frac{n}{p}}\right)^{\frac{p}{n}}=a+b$$

for any $\lambda,\mu,\nu,a,b,p>0$ (note that $\widetilde{l}^{\alpha}_{s,u},\widetilde{l}^{\alpha}_{u,t}>0$ according to our assumption).

Now the result follows from taking $\alpha\to\infty$. □

Remark 3.2. Typically if $p\geqslant 1$ is the roughness of $\mathbf{X}$, then from Lyons' extension theorem we know that $\widetilde{L}_{s,t}$ is finite.

Theorem 3.1. *Let the tensor products over $\mathbb{R}^d$ be equipped with given admissible norms, under which each element of the standard basis $\{\mathrm{e}_1,\cdots,\mathrm{e}_d\}$ of $\mathbb{R}^d$ has norm one. Then $\widetilde{L}_{s,t}$ is almost surely a deterministic constant which is bounded above by $d^2(t-s)$.*

Proof. For $m\geqslant 1$, consider the dyadic partition

$$t_i^m\triangleq s+\frac{i}{2^m}(t-s),\quad i=0,\cdots,2^m.$$

According to Proposition 3.2, we know that pathwisely

$$\widetilde{L}_{s,t}\leqslant\sum_{i=1}^{2^m}\widetilde{L}_{t_{i-1}^m,t_i^m}=2^{-m}\sum_{i=1}^{2^m}2^m\widetilde{L}_{t_{i-1}^m,t_i^m}.$$

On the one hand, by the Brownian scaling, for each i, $2^m\widetilde{L}_{t_{i-1}^m,t_i^m}$ has the same distribution as $\widetilde{L}_{s,t}$. In particular, by Proposition 3.1, it is bounded above by $d^2(t-s)$ almost surely. On the other hand, the family $\{2^m\widetilde{L}_{t_{i-1}^m,t_i^m}:\ 1\leqslant i\leqslant 2^m\}$ are independent. According to the weak law of large numbers, we conclude that

$$2^{-m}\sum_{i=1}^{2^m}2^m\widetilde{L}_{t_{i-1}^m,t_i^m}\to\mathbb{E}\left[\widetilde{L}_{s,t}\right]$$

in probability. By taking an almost surely convergent subsequence, we obtain that

$$\widetilde{L}_{s,t}\leqslant\mathbb{E}\left[\widetilde{L}_{s,t}\right]$$

almost surely. This certainly implies that $\widetilde{L}_{s,t}=\mathbb{E}\left[\widetilde{L}_{s,t}\right]$ almost surely. □

Remark 3.3. Although the fact of $\widetilde{L}_{s,t}$ being a deterministic constant is a result of independent increments for Brownian motion, it is not clear that any simple type of 0-1 law argument could apply.

Corollary 3.1. *Under the assumption of Theorem 3.1, there exists a constant κ_d depending on d, such that for each pair of $s < t$, with probability one we have $\widetilde{L}_{s,t} = \kappa_d(t-s)$.*

Proof. The result follows immediately from Theorem 3.1 and Brownian scaling. □

Remark 3.4. We should emphasize that the constant κ_d depends on the choice of given admissible norms on the tensor products.

We can further show that the $\mathbb{P}$-null set arising from Corollary 3.1 associated with each pair of $s < t$ can be chosen to be universal. This point will be very useful for applications to the level of the Brownian rough path (c.f. Section 6 below).

Proposition 3.3. *With probability one, we have*

$$\widetilde{L}_{s,t} = \kappa_d(t-s) \quad \text{for all } s < t.$$

Proof. According to Proposition 3.1 and Corollary 3.1, there exists a $\mathbb{P}$-null set $\mathcal{N}$, such that for all $\omega \notin \mathcal{N}$, we have

$$\max\left\{L'_{r_1,r_2}(\omega),\ L''_{r_1,r_2}(\omega)\right\} \leqslant d^2(r_2 - r_1)$$

and

$$\widetilde{L}_{r_1,r_2}(\omega) = \kappa_d(r_2 - r_1)$$

for all $r_1, r_2 \in \mathbb{Q}$ with $r_1 < r_2$, where

$$L'_{r_1,r_2} \triangleq \limsup_{n\to\infty}\left(\left(\frac{n}{2}\right)!\sup_{r_1\leqslant u\leqslant r_2}\left\|\mathbb{B}^n_{r_1,u}\right\|\right)^{\frac{2}{n}},$$

$$L''_{r_1,r_2} \triangleq \limsup_{n\to\infty}\left(\left(\frac{n}{2}\right)!\sup_{r_1\leqslant u\leqslant r_2}\left\|\mathbb{B}^n_{u,r_2}\right\|\right)^{\frac{2}{n}}.$$

Now fix $\omega \notin \mathcal{N}$ and let $s < r$ with $r \in \mathbb{Q}$. For arbitrary $r_1, r_2 \in \mathbb{Q}$ with $r_1 < s < r_2$, we know that

$$\begin{aligned}\kappa_d(r - r_1) &= \widetilde{L}_{r_1,r}(\omega)\\ &\leqslant \widetilde{L}_{r_1,s}(\omega) + \widetilde{L}_{s,r}(\omega)\\ &\leqslant L'_{r_1,r_2}(\omega) + \widetilde{L}_{s,r}(\omega)\\ &\leqslant d^2(r_2 - r_1) + \widetilde{L}_{s,r}(\omega).\end{aligned}$$

By letting $r_1 \uparrow s$ and $r_2 \downarrow s$ along rational times, we obtain that

$$\widetilde{L}_{s,r}(\omega) \geqslant \kappa_d (r-s).$$

Similarly, from

$$\begin{aligned}\widetilde{L}_{s,r}(\omega) &\leqslant \widetilde{L}_{s,r_2}(\omega) + \widetilde{L}_{r_2,r}(\omega) \\ &\leqslant L''_{r_1,r_2}(\omega) + \widetilde{L}_{r_2,r}(\omega) \\ &\leqslant d^2(r_2 - r_1) + \kappa_d(r - r_2),\end{aligned}$$

we conclude that

$$\widetilde{L}_{s,r}(\omega) \leqslant \kappa_d (r-s).$$

Therefore,

$$\widetilde{L}_{s,r}(\omega) = \kappa_d (r-s).$$

By repeating the same argument, we conclude that for all $s < t$,

$$\widetilde{L}_{s,t}(\omega) = \kappa_d (t-s).$$

□

4 The second part of the main result: the lower estimate

For given admissible tensor norms, from the last section we know that with probability one,

$$\widetilde{L}_{s,t} = \kappa_d (t-s),$$

where κ_d is a deterministic constant depending only on the dimension d of Brownian motion, which is bounded above by d^2. It is not even clear that κ_d should be strictly positive. In this section, we are going to establish a lower estimate of κ_d under the projective tensor norm by applying the technique of hyperbolic development which was introduced by Hambly and Lyons in their remarkable paper [10]. In the next section, we shall see that the positivity of κ_d reflects certain non-degeneracy properties of the Brownian rough path.

4.1 The hyperbolic development of a regular path

Before study the Brownian signature, let us first summarize the fundamental idea of hyperbolic development in the deterministic context for regular paths. We present proofs of a few results which seems not appearing in the literature. For an expository review on hyperbolic geometry, we refer the reader to the wonderful survey [4].

Let $\mathbb{H}^d$ ($d \geqslant 2$) be the complete, connected and simply-connected d-dimensional Riemannian manifold with constant sectional curvature -1. For computational convenience, we choose the hyperboloid model. In particular, $\mathbb{H}^d$ is defined to be the submanifold $\{x \in \mathbb{R}^{d+1} : \ x * x = -1, \ x^{d+1} > 0\}$, where $*$ is the Minkowski metric on $\mathbb{R}^{d+1}$ given by

$$x * y \triangleq \sum_{i=1}^{d} x^i y^i - x^{d+1} y^{d+1}.$$

The Minkowski metric induces a Riemannian metric on $\mathbb{H}^d$ which gives it the desired hyperbolic structure. For $x, y \in \mathbb{H}^d$, one can show that

$$\cosh \rho(x, y) = -x * y, \tag{4.1}$$

where $\rho(x, y)$ is the hyperbolic distance between x and y.

It is known that the isometry group $\mathrm{SO}(d, 1)$ of $\mathbb{H}^d$ is the space of $(d+1) \times (d+1)$-invertible matrices Γ such that $\Gamma^{-1} = J\Gamma^* J$ and $\Gamma_{d+1}^{d+1} > 0$, where $J \triangleq \mathrm{diag}(1, \cdots, 1, -1)$. The Lie algebra $\mathrm{so}(d, 1)$ of $\mathrm{SO}(d, 1)$ is the space of $(d+1) \times (d+1)$-matrices A of the form

$$A = \begin{pmatrix} A_0 & b \\ b^* & 0 \end{pmatrix}$$

where A_0 is an antisymmetric $d \times d$-matrix and $b \in \mathbb{R}^d$.

Define a linear map $F : \ \mathbb{R}^d \to \mathrm{so}(d, 1)$ by

$$F(x) \triangleq \begin{pmatrix} 0 & \cdots & 0 & x^1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & x^d \\ x^1 & \cdots & x^d & 0 \end{pmatrix}, \quad x = (x^1, \cdots, x^d) \in \mathbb{R}^d.$$

Given a continuous path $\gamma : \ [0, 1] \to \mathbb{R}^d$ with bounded variation, consider the linear ordinary differential equation

$$\begin{cases} d\Gamma_t = \Gamma_t F(d\gamma_t), & t \in [0, 1], \\ \Gamma_0 = \mathrm{I}_{d+1}. \end{cases}$$

The solution Γ_t defines a continuous path with bounded variation in the isometry group $\mathrm{SO}(d,1)$. Explicitly, by Picard's iteration, we see that

$$\Gamma_t = \sum_{n=0}^{\infty} \int_{0<t_1<\cdots<t_n<t} F(d\gamma_{t_1})\cdots F(d\gamma_{t_n}) = \sum_{n=0}^{\infty} F^{\otimes n}\left(\gamma_{0,t}^n\right). \tag{4.2}$$

Define $X_t \triangleq \Gamma_t o$, where $o = (0,\cdots,0,1)^* \in \mathbb{H}^d$.

Definition 4.1. Γ_t is called the *Cartan development* of γ_t onto $\mathrm{SO}(d,1)$. X_t is called the *hyperbolic development* of γ_t onto $\mathbb{H}^d$.

The reason of expecting a lower estimate of κ_d in our Brownian setting from the hyperbolic development is quite related to the philosophy in the deterministic setting. To be precise, define

$$\widetilde{l} \triangleq \sup_{n\geqslant 1} \left(n!\|g_n\|_{\mathrm{proj}}\right)^{\frac{1}{n}} \leqslant \|\gamma\|_{1-\mathrm{var}}, \tag{4.3}$$

where $g_n \triangleq \int_{0<t_1<\cdots<t_n<1} d\gamma_{t_1}\otimes\cdots\otimes d\gamma_{t_n}$ is the n-th component of the signature of γ, and $\|\cdot\|_{\mathrm{proj}}$ is the projective tensor norm induced by the Euclidean norm on $\mathbb{R}^d$.

Remark 4.1. Let $g = (1, g_1, g_2, \cdots)$ be a group-like element. From the shuffle product formula,

$$g_k^{\otimes n} = \sum_{\sigma\in\mathcal{S}(k,\cdots,k)} \mathcal{P}^{\sigma}(g_{nk}).$$

Therefore,

$$\|g_k\|_{\mathrm{proj}}^n \leqslant \frac{(nk)!}{(k!)^n}\|g_{nk}\|_{\mathrm{proj}}.$$

It follows that

$$\left(k!\|g_k\|_{\mathrm{proj}}\right)^{\frac{1}{k}} \leqslant \left((nk)!\|g_{nk}\|_{\mathrm{proj}}\right)^{\frac{1}{nk}}, \quad \forall n,k\geqslant 1.$$

Therefore, we conclude that

$$\sup_{n\geqslant 1}\left(n!\|g_n\|_{\mathrm{proj}}\right)^{\frac{1}{n}} = \limsup_{n\to\infty}\left(n!\|g_n\|_{\mathrm{proj}}\right)^{\frac{1}{n}}.$$

This is indeed true for any given admissible norms. A similar statement with a fractional factorial normalization (which naturally corresponds to the rough path situation) is *not* true. Indeed, considering the Brownian motion case, we have

$$\sup_{n\geqslant 1}\left(\left(\frac{n}{2}\right)!\|\mathbb{B}_{0,1}^n\|_{\mathrm{proj}}\right)^{\frac{2}{n}} \geqslant \left(\left(\frac{1}{2}\right)!\|B_1 - B_0\|_{\mathbb{R}^d}\right)^2,$$

while on the other hand, by Theorem 3.1,

$$\limsup_{n\to\infty}\left(\left(\frac{n}{2}\right)!\|\mathbb{B}^n_{0,1}\|_{\mathrm{proj}}\right)^{\frac{2}{n}}=\kappa_d$$

for almost surely. Therefore, with positive probability the “sup” is not equal to the “limsup” for the Brownian signature.

Now suppose that γ is tree-reduced. There are essentially two cases in which the length conjecture $\widetilde{l}=\|\gamma\|_{1-\mathrm{var}}$ is known to be true: piecewise linear paths or C^1-paths in constant speed parametrization with suitable modulus of continuity for the derivative.

The fundamental reason that the hyperbolic development yields the lower bound $\widetilde{l}\geqslant\|\gamma\|_{1-\mathrm{var}}$ is hidden in the following two key facts.

Fact 1. The hyperbolic development is length preserving. Moreover, if γ_t is piecewise linear , then its hyperbolic development X_t is piecewise geodesic with the same intersection angles as those of γ_t.

Proof. We first show that the Cartan development is length preserving.

If γ_t is smooth, then the equation for Γ_t becomes

$$\dot{\Gamma}_t=\Gamma_tF(\dot{\gamma}_t),$$

and thus

$$\dot{X}_t=\dot{\Gamma}_t\,o=\Gamma_t\begin{pmatrix}\dot{\gamma}_t\\0\end{pmatrix}.$$

Since Γ_t is an isometry of $\mathbb{H}^d$, by identifying $T_o\mathbb{H}^d\cong\mathbb{R}^d$, we conclude that

$$\|\dot{X}_t\|_*=\left\|\begin{pmatrix}\dot{\gamma}_t\\0\end{pmatrix}\right\|_*=\|\dot{\gamma}_t\|_{\mathrm{Euclidean}},$$

It follows that the hyperbolic length of X_t is the same as the Euclidean length of γ_t. The general bounded variation case can be proved by smooth approximation.

Next we show that the Cartan development of a piecewise linear path is a piecewise geodesic with the same intersection angles.

If $\gamma_t=tv$ is a linear path, according to (4.2), it is easily seen that

$$X_1^{d+1}=(\Gamma_1o)^{d+1}=\sum_{n=0}^{\infty}\frac{\|v\|^{2n}_{\mathrm{Euclidean}}}{(2n)!}=\cosh\|v\|_{\mathrm{Euclidean}}.$$

From the identity (4.1), we know that

$$\cosh \rho(X_1, o) = -X_1 * o = X_1^{d+1} = \cosh \|v\|_{\text{Euclidean}},$$

which implies that

$$\rho(X_1, o) = \|v\|_{\text{Euclidean}} = \|\gamma\|_{1-\text{var}}.$$

Therefore, X is a geodesic in $\mathbb{H}_d$.

Now suppose that γ_t is piecewise linear over a partition $\mathcal{P}:\ 0 = t_0 < t_1 < \cdots < t_{n+1} = 1$, where $\dot{\gamma_t} = v_k \in \mathbb{R}^d$ for $t \in [t_{k-1}, t_k]$. Apparently, the Cartan development X_t of γ_t is a piecewise geodesic. Given $1 \leqslant k \leqslant n$, we have

$$\dot{X}_{t_k-} = \Gamma_{t_{k-1}} \Gamma_{t_{k-1}}^{-1} \Gamma_{t_k} \begin{pmatrix} v_k \\ 0 \end{pmatrix} = \Gamma_{t_k} \begin{pmatrix} v_k \\ 0 \end{pmatrix},$$

and

$$\dot{X}_{t_k+} = \Gamma_{t_k} \begin{pmatrix} v_{k+1} \\ 0 \end{pmatrix}.$$

Therefore,

$$\langle v_k, v_{k+1} \rangle_{\text{Euclidean}} = \left\langle \begin{pmatrix} v_k \\ 0 \end{pmatrix}, \begin{pmatrix} v_{k+1} \\ 0 \end{pmatrix} \right\rangle_* = \left\langle \dot{X}_{t_k-}, \dot{X}_{t_k+} \right\rangle_*,$$

This implies that the Cartan development preserves intersection angles. □

Fact 2. In a hyperbolic triangle with edges $a, b, c > 0$, we have $a \geqslant b + c - \log \frac{2}{1-\cos\theta_A}$, where θ_A is the angle opposite a.

Proof. The only point which requires some attention is the following fact: for $\lambda > 0$, if we consider triangles with the same angle θ_A, and edges $\lambda b, \lambda c$ with the corresponding $a(\lambda)$, then

$$f(\lambda) \triangleq \lambda b + \lambda c - a(\lambda)$$

is monotonely increasing in λ. Based on this fact, one finds the upper bound of $b + c - a$ to be $\lim_{\lambda\to\infty}(\lambda b + \lambda c - a(\lambda))$, which can be computed by using the hyperbolic cosine law (c.f. Proof of Lemma 3.4 in [10])

To this end, it suffices to show that $f'(\lambda) = b + c - a'(\lambda) \geqslant 0$. By the first hyperbolic cosine law, we have

$$\cosh a(\lambda) = \cosh \lambda b \cosh \lambda c - \sinh \lambda b \sinh \lambda c \cos \theta_A.$$

Differentiating with respect to λ, we obtain that

$$
\begin{aligned}
a'(\lambda)\sinh a(\lambda) = b\,(\sinh\lambda b\cosh\lambda c - r\cosh\lambda b\sinh\lambda c) \\
+ c\,(\cosh\lambda b\sinh\lambda c - r\sinh\lambda b\cosh\lambda c)
\end{aligned}
$$

where $r \triangleq \cos\theta_A$. For simplicity we write $\sinh = \mathrm{sh}$, $\cosh = \mathrm{ch}$. Now it suffices to show that

$$
b(\mathrm{sh}\lambda b\cdot\mathrm{ch}\lambda c - r\mathrm{ch}\lambda b\cdot\mathrm{sh}\lambda c) + c(\mathrm{ch}\lambda b\cdot\mathrm{sh}\lambda c - r\mathrm{sh}\lambda b\cdot\mathrm{ch}\lambda c) \leqslant (b+c)\mathrm{sh}a(\lambda).
$$

We use X, Y to denote the left and right hand sides respectively. From direct computation, we see that

$$
\begin{aligned}
X^2 = {} & (b-cr)^2\mathrm{sh}^2\lambda b\cdot\mathrm{ch}^2\lambda c + (c-br)^2\mathrm{ch}^2\lambda b\cdot\mathrm{sh}^2\lambda c \\
& + (2bc + 2bcr^2 - 2b^2r - 2c^2r)\mathrm{sh}\lambda b\cdot\mathrm{ch}\lambda b\cdot\mathrm{sh}\lambda c\cdot\mathrm{ch}\lambda c,
\end{aligned}
$$

and

$$
\begin{aligned}
Y^2 = {} & (b+c)^2((1+r)^2\mathrm{sh}^2\lambda b\cdot\mathrm{sh}^2\lambda c + \mathrm{sh}^2\lambda b + \mathrm{sh}^2\lambda c \\
& - 2r\mathrm{sh}\lambda b\cdot\mathrm{ch}\lambda b\cdot\mathrm{sh}\lambda c\cdot\mathrm{ch}\lambda c).
\end{aligned}
$$

By using $\cosh^2 x - \sinh^2 x = 1$, we obtain that

$$
\begin{aligned}
\frac{Y^2 - X^2}{1+r} = {} & 2bc(1+r)\mathrm{sh}^2\lambda b\cdot\mathrm{sh}^2\lambda c - 2bc(1+r)\mathrm{sh}\lambda b\cdot\mathrm{ch}\lambda b\cdot\mathrm{sh}\lambda c\cdot\mathrm{ch}\lambda c \\
& + (c^2(1-r) + 2bc)\mathrm{sh}^2\lambda b + (b^2(1-r) + 2bc)\mathrm{sh}^2\lambda c.
\end{aligned}
$$

Define $g(r)$ to be the function in r given by the right hand side of the above equality. Then

$$
g(1) = 2bc(\mathrm{sh}\lambda b\cdot\mathrm{ch}\lambda c - \mathrm{ch}\lambda b\cdot\mathrm{sh}\lambda c)^2 \geqslant 0.
$$

Moreover,

$$
\begin{aligned}
g'(r) &= -2bc\mathrm{sh}\lambda b\cdot\mathrm{sh}\lambda c\cdot\mathrm{ch}\lambda(b-c) \\
&\quad - c^2\mathrm{sh}^2\lambda b - b^2\mathrm{sh}^2\lambda c \\
&\leqslant 0.
\end{aligned}
$$

Therefore, $g(r) \geqslant 0$ for $r \in [-1, 1]$, which implies that $Y^2 \geqslant X^2$. Since $Y \geqslant 0$, we conclude that $Y \geqslant X$. □

Let $\gamma : [0,1] \to \mathbb{R}^d$ be a tree-reduced path with bounded variation. From (4.1) and the explicit formula for the Cartan development, it is not hard to see that (see also (4.6) below), for each $\lambda > 0$,

$$\cosh \rho(X_1^\lambda, o) = \sum_{n=0}^{\infty} \lambda^{2n} \int_{0<t_1<\cdots<t_{2n}<1} \langle d\gamma_{t_1}, d\gamma_{t_2}\rangle \cdots \langle d\gamma_{t_{2n-1}}, d\gamma_{t_{2n}}\rangle \leqslant \cosh \lambda \widetilde{l},$$

where $\widetilde{l}$ is defined by (4.3), $\gamma_t^\lambda \triangleq \lambda\gamma_t$ $(0 \leqslant t \leqslant)$ is the path obtained by rescaling γ by the factor λ, and X_t^λ is the hyperbolic development of γ_t^λ. In particular, we see that $\lambda\widetilde{l} \geqslant \rho(X_1^\lambda, o)$.

The previous Fact 2 tells us that for all two-edge piecewise geodesic paths $Y : [0,1] \to \mathbb{H}^d$ with fixed intersection angle $0 < \theta < \pi$, the distance between hyperbolic length of Y and $\rho(Y_1, o)$ is uniformly bounded by a constant depending on θ. Now suppose that $\gamma : [0,1] \to \mathbb{R}^d$ is a two-edge piecewise linear path with intersection angle $0 < \theta < \pi$. Fact 1 and 2 together implies that

$$0 \leqslant \lambda \|\gamma\|_{1-\mathrm{var}} - \rho\left(X_1^\lambda, o\right) \leqslant K(\theta) \triangleq \log \frac{2}{1-\cos\theta},$$

uniformly in $\lambda > 0$. In particular,

$$\lim_{\lambda\to\infty} \frac{\rho\left(X_1^\lambda, o\right)}{\lambda} = \|\gamma\|_{1-\mathrm{var}}, \tag{4.4}$$

from which we obtain the desired estimate $\widetilde{l} \geqslant \|\gamma\|_{1-\mathrm{var}}$. It is important to note that the angle θ captures the tree-reduced nature of γ in this simple case. Indeed, if $\theta = 0$, $K(\theta) = +\infty$.

With some effort, the previous argument extends to tree-reduced piecewise linear paths with minimal intersection angle given by $\theta > 0$. In this case, one can obtain an estimate of the form

$$0 \leqslant \lambda \|\gamma\|_{1-\mathrm{var}} - \rho\left(X_1^\lambda, o\right) \leqslant N \cdot \Lambda(\theta)$$

uniformly in $\lambda > 0$, where N is the number of edges of γ and $\Lambda(\theta)$ is a constant depending only θ (which explodes as $\theta \downarrow 0$). We again obtain (4.4) and thus the desired estimate. Here $\theta > 0$ captures the tree-reduced nature of γ. With some further delicate analysis, one can establish a similar estimate for a path $\gamma : [0,1] \to \mathbb{R}^d$ which is continuously differentiable when parametrized at constant speed. The estimate takes the form

$$0 \leqslant \lambda \|\gamma\|_{1-\mathrm{var}} - \rho(X_1^\lambda, o) \leqslant C_1 \lambda \|\gamma\|_{1-\mathrm{var}} \delta_\gamma \left(\frac{C_2}{\lambda}\right)^2$$

provided that λ is large, where C_1, C_2 are universal constants and $\delta_\gamma(\cdot)$ is the modulus of continuity for $\dot{\gamma}$. In particular, we again obtain (4.4) and thus the desired estimate. Here the existence of modulus of continuity for $\dot{\gamma}$ already implies that γ is tree-reduced implicitly. In any case, the fundamental reason which makes the technique of hyperbolic development work is hidden in the nature of Fact 1 and 2.

If one is attempting to attack the length conjecture $\widetilde{l} = \|\gamma\|_{1-\text{var}}$ for a general tree-reduced path with bounded variation by using the idea of hyperbolic development, it seems that a crucial point is to find a quantity ω_γ, a certain kind of "modulus of continuity", which on the one hand captures the tree-reduced nature of γ quantitatively, and on the other hand can be used to control the growth of $\lambda \mapsto \lambda\|\gamma\|_{1-\text{var}} - \rho(X_1^\lambda, o)$ (difference between hyperbolic length and hyperbolic distance for the rescaled path). Up to the current point, this fascinating and challenging problem remains unsolved.

4.2 The hyperbolic development of Brownian motion and a lower estimate for κ_d

In spite of the huge difficulty in obtaining lower estimates of the hyperbolic distance function in the general deterministic setting, it is surprising that a simple martingale argument will give us a meaningful lower estimate for the hyperbolic development of Brownian motion. In particular, we can obtain a lower estimate on the constant κ_d.

From now on, we assume that $\mathbb{R}^d$ is equipped with the l^p-norm for some given $1 \leqslant p \leqslant 2$, and the tensor products over $\mathbb{R}^d$ are equipped with the associated projective tensor norms.

The following characterization of projective tensor norms is important for us.

Lemma 4.1. *For each $\xi \in \left(\mathbb{R}^d\right)^{\otimes n}$, we have*

$$\|\xi\|_{\text{proj}} = \sup\left\{|\Phi(\xi)| : \ \Phi \in L(\mathbb{R}^d, \cdots, \mathbb{R}^d; \mathbb{R}^1), \ \|\Phi\| \leqslant 1\right\},$$

where we identify $L(\mathbb{R}^d, \cdots, \mathbb{R}^d; \mathbb{R}^1)$ with $\left((\mathbb{R}^d)^{\otimes n}\right)^*$ through the universal property, and

$$\|\Phi\| \triangleq \inf\{C \geqslant 0 : \ |\Phi(v_1, \cdots, v_n)| \leqslant C\|v_1\| \cdots \|v_n\| \ \ \forall v_1, \cdots, v_n \in \mathbb{R}^d\}.$$

Proof. See [17], Identity (2.3). □

Let $B_t = (B_t^1, \cdots, B_t^d)$ be a d-dimensional Brownian motion. We define

$$\widetilde{L}_t \triangleq \limsup_{n\to\infty} \left(\left(\frac{n}{2}\right)! \|\mathbb{B}_{0,t}^n\|_{\mathrm{proj}}\right)^{\frac{2}{n}}.$$

For each $\lambda > 0$, we consider the Cartan development

$$\begin{cases} d\Gamma_t^\lambda = \lambda \Gamma_t^\lambda F(\circ dB_t), & t \geqslant 0, \\ \Gamma_0^\lambda = \mathrm{I}_{d+1}, \end{cases}$$

of $\lambda \cdot B_t$, where the differential equation is understood in the Stratonovich sense. Let $X_t^\lambda \triangleq \Gamma_t^\lambda o$ be the hyperbolic development of $\lambda \cdot B_t$. Since B_t has a canonical lifting as geometric rough paths, from the universal limit theorem for rough differential equations, we see that Γ_t^λ defines a path on the isometry group $\mathrm{SO}(d,1)$ and hence X_t^λ is a path on $\mathbb{H}^d$ starting at o.

Picard's iteration again shows that

$$\Gamma_t^\lambda = \sum_{n=0}^\infty \lambda^n \int_{0<t_1<\cdots<t_n<t} F(\circ dB_{t_1}) \cdots F(\circ dB_{t_n}). \tag{4.5}$$

Define $h_t^\lambda \triangleq \left(X_t^\lambda\right)^{d+1}$ to be the hyperbolic height of X_t^λ (the last coordinate of X_t^λ). It follows from (4.1), (4.5) and the definition of F that

$$\begin{aligned} h_t^\lambda &= \cosh \rho(X_t^\lambda, o) \\ &= \sum_{n=0}^\infty \lambda^n \int_{0<t_1<\cdots<t_n<t} \left(F(\circ dB_{t_1}) \cdots F(\circ B_{t_n}) o\right)^{d+1} \\ &= \sum_{n=0}^\infty \lambda^{2n} \int_{0<t_1<\cdots<t_n<t} \langle \circ dB_{t_1}, \circ dB_{t_2}\rangle \cdots \langle \circ dB_{t_{2n-1}}, \circ dB_{t_{2n}}\rangle, \end{aligned} \tag{4.6}$$

The following result shows that the quantity $\widetilde{L}_t$ can be controlled from below in terms of the asymptotics of h_t^λ as $\lambda \to \infty$.

Proposition 4.1. *With probability one, we have*

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log h_t^\lambda \leqslant \widetilde{L}_t, \quad \forall t \geqslant 0.$$

Proof. For each $n \geqslant 1$, define a real-valued $2n$-linear map Φ_n over $\mathbb{R}^d$ by

$$\Phi_n(v_1, \cdots, v_{2n}) \triangleq \langle v_1, v_2\rangle \cdots \langle v_{2n-1}, v_{2n}\rangle, \quad v_1, \cdots, v_{2n} \in \mathbb{R}^d.$$

Since we are taking the l^p-norm on $\mathbb{R}^d$ for $1 \leqslant p \leqslant 2$, we see that

$$\begin{aligned}|\Phi_n(v_1, \cdots, v_{2n})| &\leqslant \|v_1\|_{l^2} \cdots \|v_{2n}\|_{l^2} \\ &\leqslant \|v_1\|_{l^p} \cdots \|v_{2n}\|_{l^p}.\end{aligned}$$

In particular, $\|\Phi_n\| \leqslant 1$. Therefore, by Lemma 4.1, we have

$$\left|\int_{0<t_1<\cdots<t_n<t} \langle \circ dB_{t_1}, \circ dB_{t_2}\rangle \cdots \langle \circ dB_{t_{2n-1}}, \circ dB_{t_{2n}}\rangle\right| = |\Phi_n(\mathbb{B}_{0,t}^{2n})| \leqslant \|\mathbb{B}_{0,t}^{2n}\|_{\mathrm{proj}}.$$

Now for each $\alpha \geqslant 1$, define

$$\widetilde{L}_t^\alpha \triangleq \sup_{n\geqslant\alpha} \left(\left(\frac{n}{2}\right)! \|\mathbb{B}_{0,t}^n\|_{\mathrm{proj}}\right)^{\frac{2}{n}}.$$

It follows that

$$\begin{aligned}h_t^\lambda &\leqslant \sum_{n=0}^{\infty} \lambda^{2n} \|\mathbb{B}_{0,t}^{2n}\|_{\mathrm{proj}} \\ &= \sum_{n=0}^{\alpha-1} \lambda^{2n} \|\mathbb{B}_{0,t}^{2n}\|_{\mathrm{proj}} + \sum_{n=\alpha}^{\infty} \lambda^{2n} \|\mathbb{B}_{0,t}^{2n}\|_{\mathrm{proj}} \\ &\leqslant \sum_{n=0}^{\alpha-1} \lambda^{2n} \|\mathbb{B}_{0,t}^{2n}\|_{\mathrm{proj}} + \sum_{n=\alpha}^{\infty} \lambda^{2n} \cdot \frac{\left(\widetilde{L}_t^{2\alpha}\right)^n}{n!} \\ &= \exp\left(\lambda^2 \widetilde{L}_t^{2\alpha}\right) + \sum_{n=0}^{\alpha-1} \lambda^{2n} \left(\|\mathbb{B}_{0,t}^{2n}\|_{\mathrm{proj}} - \frac{\left(\widetilde{L}_t^{2\alpha}\right)^n}{n!}\right).\end{aligned}$$

Therefore,

$$\begin{aligned}&\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log h_t^\lambda \\ &\leqslant \lim_{\lambda\to\infty} \frac{1}{\lambda^2} \log \left(\exp\left(\lambda^2 \widetilde{L}_t^{2\alpha}\right) + \sum_{n=0}^{\alpha-1} \lambda^{2n} \left(\|\mathbb{B}_{0,t}^{2n}\|_{\mathrm{proj}} - \frac{\left(\widetilde{L}_t^{2\alpha}\right)^n}{n!}\right)\right) \\ &= \widetilde{L}_t^{2\alpha}.\end{aligned}$$

Since α is arbitrary, we conclude that

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log h_t^\lambda \leqslant \inf_{\alpha\geqslant 1} \widetilde{L}_t^{2\alpha} = \widetilde{L}_t.$$

□

The following result is the probabilistic counterpart of a lower estimate on the hyperbolic height function h_t^λ.

Lemma 4.2. *For any* $0 < \mu < d-1$, *we have*

$$\mathbb{E}\left[\left(h_t^\lambda\right)^{-\mu}\right] \leqslant \exp\left(-\frac{\lambda^2\mu(d-1-\mu)}{2}t\right).$$

Proof. Note that

$$F(dB_t)\cdot F(dB_t) = \begin{pmatrix} 0 & dB_t \\ (dB_t)^* & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & dB_t \\ (dB_t)^* & 0 \end{pmatrix} = \begin{pmatrix} \mathrm{I}_d & 0 \\ 0 & d \end{pmatrix} dt.$$

By the Itô-Stratonovich conversion, we have

$$\begin{aligned} d\Gamma_t^\lambda &= \lambda\Gamma_t^\lambda \cdot F(dB_t) + \frac{\lambda}{2} d\Gamma_t^\lambda \cdot F(dB_t) \\ &= \lambda\Gamma_t^\lambda \cdot F(dB_t) + \frac{\lambda^2}{2}\Gamma_t^\lambda \left(F(dB_t)\cdot F(dB_t)\right) \\ &= \lambda\Gamma_t^\lambda \cdot F(dB_t) + \frac{\lambda^2}{2}\Gamma_t^\lambda \begin{pmatrix} \mathrm{I}_d & 0 \\ 0 & d \end{pmatrix} dt. \end{aligned}$$

Therefore

$$\begin{aligned} dh_t^\lambda &= d\left(\Gamma_t^\lambda\right)_{d+1}^{d+1} \\ &= \lambda \sum_{i=1}^d \left(\Gamma_t^\lambda\right)_i^{d+1} \cdot dB_t^i + \frac{\lambda^2 d}{2} h_t^\lambda dt. \end{aligned}$$

Moreover, since $\Gamma_t^\lambda \in \mathrm{SO}(d,1)$, we know that

$$\sum_{i=1}^d \left(\left(\Gamma_t^\lambda\right)_i^{d+1}\right)^2 - \left(h_t^\lambda\right)^2 = -1,$$

and hence

$$dh_t^\lambda \cdot dh_t^\lambda = \lambda^2 \sum_{i=1}^d \left(\left(\Gamma_t^\lambda\right)_i^{d+1}\right)^2 dt = \lambda^2\left(\left(h_t^\lambda\right)^2 - 1\right) dt.$$

According to Itô's formula,

$$\begin{aligned} d\left(h_t^\lambda\right)^{-\mu} &= -\mu\left(h_t^\lambda\right)^{-(\mu+1)} dh_t^\lambda + \frac{\mu(\mu+1)}{2}\left(h_t^\lambda\right)^{-(\mu+2)}\left(dh_t^\lambda \cdot dh_t^\lambda\right) \\ &= -\lambda\mu\left(h_t^\lambda\right)^{-(\mu+1)} \sum_{i=1}^{d} \left(\Gamma_t^\lambda\right)_i^{d+1} dB_t^i \\ &\quad - \left(\frac{\lambda^2\mu(d-1-\mu)}{2}\left(h_t^\lambda\right)^{-\mu} + \frac{\lambda^2\mu(\mu+1)}{2}\left(h_t^\lambda\right)^{-(\mu+2)}\right) dt. \end{aligned}$$

By taking expectation and differentiating with respect to t, we obtain that

$$\begin{aligned} \frac{d}{dt}\mathbb{E}\left[\left(h_t^\lambda\right)^{-\mu}\right] &= -\frac{\lambda^2\mu(d-1-\mu)}{2}\mathbb{E}\left[\left(h_t^\lambda\right)^{-\mu}\right] \\ &\quad - \frac{\lambda^2\mu(\mu+1)}{2}\mathbb{E}\left[\left(h_t^\lambda\right)^{-(\mu+2)}\right] \\ &\leqslant -\frac{\lambda^2\mu(d-1-\mu)}{2}\mathbb{E}\left[\left(h_t^\lambda\right)^{-\mu}\right]. \end{aligned}$$

By Gronwall's inequality, we arrive at

$$\mathbb{E}\left[\left(h_t^\lambda\right)^{-\mu}\right] \leqslant \exp\left(-\frac{\lambda^2\mu(d-1-\mu)}{2}t\right).$$

□

Now we can state our main lower estimate on κ_d.

Theorem 4.1. *Under the l^p-norm ($1 \leqslant p \leqslant 2$) on $\mathbb{R}^d$ and the associated projective tensor norms on the tensor products, we have*

$$\kappa_d \geqslant \frac{d-1}{2}.$$

Proof. Fix $t > 0$, $\lambda > 0$ and $0 < \mu < d-1$. According to Lemma 4.2, for each $K > 0$,

$$\begin{aligned} \mathbb{P}\left(h_t^\lambda \leqslant K\right) &= \mathbb{P}\left(\left(h_t^\lambda\right)^{-\mu} \geqslant K^{-\mu}\right) \\ &\leqslant K^\mu \mathbb{E}\left[\left(h_t^\lambda\right)^{-\mu}\right] \\ &\leqslant K^\mu \exp\left(-\frac{\lambda^2\mu(d-1-\mu)}{2}t\right). \end{aligned}$$

Now for each $m \geqslant 1$, define $\lambda_m \triangleq m$ and

$$K_m \triangleq \exp\left(\frac{m^2(d-1-\mu)}{2}s\right).$$

It follows that

$$\mathbb{P}\left(h_t^{\lambda_m} \leqslant K_m\right) \leqslant \exp\left(-\frac{m^2\mu(d-1-\mu)}{2}(t-s)\right).$$

In particular, $\sum_{m=1}^{\infty} \mathbb{P}\left(h_t^{\lambda_m} \leqslant K_m\right) < \infty$. By the Borel-Cantelli lemma, there exists a $\mathbb{P}$-null set $\mathcal{N}(s,t,\mu)$, such that for any $\omega \notin \mathcal{N}(s,t,\mu)$, there exists $M(\omega) \geqslant 1$ with

$$h_t^{\lambda_m}(\omega) > \exp\left(\frac{m^2(d-1-\mu)}{2}s\right), \quad \forall m \geqslant M(\omega).$$

Therefore,

$$\limsup_{m\to\infty} \frac{1}{m^2}\log h_t^{\lambda_m}(\omega) \geqslant \frac{d-1-\mu}{2}s.$$

By enlarging the $\mathbb{P}$-null set through rationals $s \uparrow t$ and $\mu \downarrow 0$, we conclude that

$$\limsup_{m\to\infty} \frac{1}{m^2}\log h_t^{\lambda_m} \geqslant \frac{d-1}{2}t$$

for almost surely.

Finally, according to Proposition 4.1, we obtain that

$$\kappa_d = \frac{\widetilde{L}_t}{t} \geqslant \frac{d-1}{2}.$$

□

5 Applications to the Brownian rough path

We present a few interesting consequences of the lower estimate on κ_d given in Theorem 4.1.

Let us consider the d-dimensional Brownian motion B_t on $[0,1]$. Recall that with probability one, B_t has a canonical lifting $\mathbf{B}_t$ as geometric p-rough path for $2 < p < 3$. As a process on $G^2(\mathbb{R}^d)$, the Brownian rough path $\mathbf{B}_t$ is canonically defined and it is independent of the choice of tensor norms on $(\mathbb{R}^d)^{\otimes 2}$.

Corollary 5.1. *For almost every ω, the path $t \mapsto \mathbf{B}_t(\omega)$ is tree-reduced.*

Proof. Let the tensor products be equipped with the projective tensor norms associated with the l^2-norm on $\mathbb{R}^d$. From Proposition 3.3, for every ω outside some $\mathbb{P}$-null set $\mathcal{N}$,

$$\widetilde{L}_{s,t}(\omega) = \kappa_d(t-s) \quad \forall s < t.$$

In addition, from Theorem 4.1 we know that the constant κ_d is strictly positive. Therefore, for every $\omega \notin \mathcal{N}$, $\mathbf{B}(\omega)$ cannot have any tree-like pieces, otherwise $\widetilde{L}_{s,t}(\omega) = 0$ for some $s < t$ which is a contradiction. □

From Corollary 5.1, we know that for any two ω_1, ω_2 outside some $\mathbb{P}$-null set, $\mathbf{B}(\omega_1)$ and $\mathbf{B}(\omega_2)$ have the same signature if and only if they differ from each other by a reparametrization. In other words, with probability one the Brownian rough path is uniquely determined by its signature up to reparametrization. This result was first proved by Le Jan and Qian [12] (see also [2]).

We can actually obtain the following stronger uniqueness result which is not implied by the work in the literature.

Corollary 5.2. *There exists a $\mathbb{P}$-null set $\mathcal{N}$, such that for any two distinct $\omega_1, \omega_2 \notin \mathcal{N}$, $\mathbf{B}(\omega_1)$ and $\mathbf{B}(\omega_2)$ cannot be equal up to a reparametrization. In particular, with probability one, the Brownian rough path is uniquely determined by its signature.*

Proof. We follow the same notation as in the proof of Corollary 5.1. Given two distinct $\omega_1, \omega_2 \notin \mathcal{N}$, suppose that

$$\mathbf{B}_t(\omega_2) = \mathbf{B}_{\sigma(t)}(\omega_1), \quad 0 \leqslant t \leqslant 1,$$

for some reparametrization $\sigma:\ [0,1] \to [0,1]$. Then we have

$$\widetilde{L}_{0,\sigma(t)}(\omega_1) = \kappa_d \sigma(t)$$

and

$$\widetilde{L}_{0,t}(\omega_2) = \kappa_d t.$$

But from assumption we know that $\widetilde{L}_{0,\sigma(t)}(\omega_1) = \widetilde{L}_{0,t}(\omega_2)$. Therefore, we must have $\sigma(t) = t$ and hence $\mathbf{B}(\omega_1) = \mathbf{B}(\omega_2)$. □

Finally, we also have the following useful consequence.

Corollary 5.3. *With probability one, Brownian rough path together with its parametrization can be recovered from its signature.*

Proof. We know from [2], [12] that for every ω outside some $\mathbb{P}$-null set, the equivalence class $[\mathbf{B}(\omega)]$ (i.e. $\mathbf{B}(\omega)$ modulo reparametrization defined in the sense of [2], Section 5.3) can be recovered from its signature. Pick an arbitrary representative $(\mathbf{X}_t)_{0\leqslant t\leqslant 1} \in [\mathbf{B}(\omega)]$. Then

$$\mathbf{X}_t = \mathbf{B}_{\sigma(t)}(\omega), \quad 0 \leqslant t \leqslant 1,$$

for some unique reparametrization σ that we want to figure out. According to Proposition 3.3 and Theorem 4.1, we have

$$\sigma(t) = \frac{1}{\kappa_d} \limsup_{n\to\infty} \left(\left(\frac{n}{2}\right)! \|\mathbb{X}_{0,t}^n\|_{\mathrm{proj}} \right)^{\frac{2}{n}},$$

where we again choose the projective tensor norms on the tensor products associated with the l^2-norm on $\mathbb{R}^d$. The underlying path $\mathbf{B}(\omega)$ is then given by

$$\mathbf{B}_t(\omega) = \mathbf{X}_{\sigma^{-1}(t)}, \quad 0 \leqslant t \leqslant 1.$$

□

Another way of understanding the previous result is the following. Since $[\mathbf{B}(\omega)]$ can be recovered from its signature, we know that the image of the signature path $\mathbb{B}(\omega)$ can be recovered from its endpoint. For every $\xi = (1, \xi_1, \xi_2, \cdots) \in \mathrm{Im}(\mathbb{B}(\omega))$, we then have

$$\mathbf{B}_{\|\xi\|/\kappa_d}(\omega) = \pi^{(2)}(\xi),$$

where

$$\|\xi\| \triangleq \limsup_{n\to\infty} \left(\left(\frac{n}{2}\right)! \|\xi_n\|_{\mathrm{proj}} \right)^{\frac{2}{n}}$$

and $\pi^{(2)}: T((\mathbb{R}^d)) \to T^{(2)}((\mathbb{R}^d))$ is the canonical projection map.

6 Further remarks and related problems

In Theorem 1.1, we considered the tail asymptotics of the Brownian signature defined in terms of iterated Stratonovich's integrals. Stratonovich's integrals arise naturally when we study Brownian motion from the rough path point of view. On the other hand, one could ask a similar question for Itô's iterated integrals. Indeed, if we define

$$\widehat{L}_{s,t} \triangleq \limsup_{n\to\infty} \left(\left(\frac{n}{2}\right)! \left\| \int_{s<u_1<\cdots<u_n<t} dB_{u_1} \otimes \cdots \otimes dB_{u_n} \right\|_{l^1} \right)^{\frac{2}{n}}$$

where the iterated integrals are defined in the sense of Itô and the tensor products are equipped with the l^1-norm, then by a similar type of arguments, one can show that

$$\frac{d(t-s)}{2} \leqslant \widehat{L}_{s,t} \leqslant \frac{d^2(t-s)}{2} \tag{6.1}$$

for almost surely. Since the lifting of Brownian motion in *Itô's* sense is not a geometric rough path, uniqueness of signature result does not apply and the intrinsic meaning of the quantity $\widehat{L}_{s,t}$ is unclear. The proof of (6.1) will not be included here since it is essentially parallel to the Stratonovich case.

Our main result of Theorem 1.1 gives rise to many interesting and related problems in the probabilistic context.

(1) The first interesting and immediate question one could come up with is the exact value of κ_d and its probabilistic meaning. In view of the length conjecture (1.2) and Theorem 1.1, if we consider the projective tensor norms on the tensor products induced by the Euclidean norm on $\mathbb{R}^d$, it is quite natural to expect that, κ_d would have a meaning related to certain kind of quadratic variation for the Brownian rough path. It also seems that there are rooms for improving the upper estimate for κ_d. The point is that in the proof of Lemma 3.1, if we shuffle an arbitrary long word $\{i_1, \cdots, i_n\}$ over $\{1, \cdots, d\}$ with itself, the chance of hitting a nonzero coefficient in the $2n$-degree component of the Brownian expected signature is quite small. But to make the analysis precise, some hard combinatorics argument for the shuffle product structure might be involved.

(2) If κ_d is related to certain kind of quadratic variation for the Brownian motion, it is reasonable to expect that our main result and corollaries apply to diffusions or even general continuous semimartingales, though there is no reason to believe that in this case the corresponding $\widetilde{L}_{s,t}$ will still be deterministic. For Gaussian processes, it is even not clear that any analogous version of $\widetilde{L}_{s,t}$ would be meaningful since for instance we know that

$$\lim_{n\to\infty}\sum_{i=1}^{n}|B_{\frac{i}{n}} - B_{\frac{i-1}{n}}|^p = 0 \quad \text{or} \quad \infty$$

in probability for a fractional Brownian motion with Hurst parameter $H \in (0,1)$, according to whether $pH > 1$ or $pH < 1$.

(3) There is a quite subtle point in the discussion of Section 6. With probability one, the lifting map $\omega \mapsto \mathbf{B}(\omega)$ is canonically well-defined. Therefore, although Corollary 5.3 (the uniqueness result) is stated at the level of the Brownian rough path, by projection to degree one, it also holds at the level of sample paths.

However, it is not at all clear if the first part of Corollary 5.2 is true at the level of Brownian sample paths. More precisely, to our best knowledge, a solution to the following classical question for Brownian motion is not known (at least not to us yet): does there exist a $\mathbb{P}$-null set $\mathcal{N}$, such that no two sample paths outside $\mathcal{N}$ can be equal up to a non-trivial reparametrization? This question is stated for Brownian sample paths and has nothing to do with the lifting of Brownian motion to rough paths.

It is a subtle point that the result of Corollary 5.2 does not yield an affirmative answer easily to the above question. Indeed, if one wants to apply Corollary 5.2, a missing point is whether the lifting operation and the reparametrization operation are commutative outside some universal $\mathbb{P}$-null set. In other words, it is not known if there exists a $\mathbb{P}$-null set $\mathcal{N}$, such that one could define a lifting map $\omega \mapsto \mathbf{B}(\omega)$ for all $\omega \notin \mathcal{N}$, which satisfies

$$\mathbf{B}_{\cdot}(\omega_\sigma) = \mathbf{B}_{\sigma(\cdot)}(\omega)$$

for all reparametrizations $\sigma : [0,1] \to [0,1]$. When defining the almost sure lifting of Brownian motion, the $\mathbb{P}$-null set comes with the given choice of approximation. It is quite subtle (and could be false) to see if the $\mathbb{P}$-null set can be chosen in a universal way.

References

[1] G. Ben Arous, Flots et series de Taylor stochastiques, *Probab. Theory Related Fields* 81 29–77, 1989.

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

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

[4] J.W. Cannon, W.J. Floyd, R. Kenyon and W.R. Parry, hyperbolic geometry, in *Flavors of Geometry, Math. Sci. Res. Inst. Publ.* 31: 59–115, Cambridge University Press, Cambridge, 1997.

[5] M. Caruana, T. Lévy and T. Lyons, *Differential equations driven by rough paths*, Springer, 2007.

[6] J. Chang, N. Duffield, H. Ni and W. Xu, Signature inversion for monotone paths, Electron. Commun. Probab. 22 (42): 1–11, 2017.

[7] T. Fawcett, *Problems in stochastic analysis: connections between rough paths and non-commutative harmonic analysis*, Ph.D. thesis, University of Oxford, 2003.

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

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

[10] 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): 109–167, 2010.

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

[12] Y. Le Jan and Z. Qian, Stratonovich's signatures of Brownian motion determine Brownian sample paths, *Probab. Theory Relat. Fields* 157: 209–223, 2013.

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

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

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

[16] T. Lyons and W. Xu, Inverting the signature of a path, to appear in *J. Eur. Math. Soc.*, 2017.

[17] R. Ryan, *Introduction to Tensor Products of Banach Spaces*, Springer, 2016.

[18] E. M. Sipiläinen, *A pathwise view of solutions of stochastic differential equations*, Ph.D. thesis, University of Edinburgh, 1993.