Decay rate of iterated integrals of branched rough path

Article

Accepted Version

Boedihardjo, H. (2018) Decay rate of iterated integrals of branched rough path. Annales de l'Institut Henri Poincare (C) Analyse Non Linéaire, 35 (4). pp. 945-969. ISSN 0294-1449 doi: 10.1016/j.anihpc.2017.09.002 Available at https://centaur.reading.ac.uk/67377/

To link to this article DOI: http://dx.doi.org/10.1016/j.anihpc.2017.09.002

Publisher: Elsevier

DECAY RATE OF ITERATED INTEGRALS OF BRANCHED ROUGH PATHS

HORATIO BOEDIHARDJO

Abstract. Iterated integrals of paths arise frequently in the study of the Taylor's expansion for controlled differential equations. We will prove a factorial decay estimate, conjectured by M. Gubinelli, for the iterated integrals of non-geometric rough paths. We will explain, with a counter example, why the conventional approach of using the neoclassical inequality fails. Our proof involves a concavity estimate for sums over rooted trees and a non-trivial extension of T. Lyons' proof in 1994 for the factorial decay of iterated Young's integrals.

1. Introduction

The iterated integrals of a path arise naturally from the Taylor's expansion of a controlled differential equation driven by the path and play a fundamental role in the theory of rough paths [7]. Given a path x, we are interested in the behaviour of the iterated integral

$$X_{0,1}^{n}=\int_{0<s_1<\ldots<s_n<1}\mathrm{d}x_{s_1}\otimes\ldots\otimes\mathrm{d}x_{s_n} \tag{1.1}$$

as n varies. The solution for a linear controlled differential equation has a series expansion that is linear in these iterated integrals. The convergence of the series expansion is often studied using the decay of these iterated integrals. The simplest example is the case when x takes value in $\mathbb{R}^d$ and has an almost everywhere derivative in L^∞, in which case

$$\|X_{0,1}^{n}\|\leq\frac{|\dot{x}|_{L^\infty}^{n}}{n!}.$$

The problem becomes much harder when x does not have a derivative, such as in the case when the iterated integral (1.1) is defined in terms of Young's integration. It was proved by Lyons [6] that if x is γ-Hölder, $\gamma>\frac{1}{2}$, and $\|x\|_\gamma$ denotes the γ-Hölder norm of x, then

$$\|X_{0,1}^{n}\|\leq(1+\zeta(2\gamma))^{n-1}\frac{\|x\|_{\gamma}^{n}}{n!^{\gamma}}, \tag{1.2}$$

where ζ is the classical Riemann Zeta function. For $0<\gamma\leq\frac{1}{2}$ and $N=\lfloor\gamma^{-1}\rfloor$, a γ -Hölder geometric rough path takes value in the unital tensor algebra

$$T^{(N)}\left(\mathbb{R}^{d*}\right)=1\oplus\mathbb{R}^{d*}\ldots\oplus\left(\mathbb{R}^{d*}\right)^{\otimes N}$$

Key words and phrases. Branched rough paths; non-geometric rough paths; iterated integrals.
We would also like to thank M. Gubinelli and D. Yang for the useful discussions, as well as the anonymous referee for the detailed comments.

where $\mathbb{R}^{d*}$ denotes the dual of $\mathbb{R}^d$. Lyons [7] showed that a γ -Hölder rough path x can be extended uniquely to a γ -Hölder path $\mathbf{X}$ in $T^{(n)}\left(\mathbb{R}^{d*}\right)$ for any $n \geq N$. He defined the n-th order iterated integrals of x up to time 1 as the n-th tensor component of this extended path at time 1, which we will denote for latter use as $X_{0,1}^n$, and showed that

$$\left\|X_{0,1}^n\right\| \leq \gamma^{-n}\left(1+2^{(N+1)\gamma}\zeta\left(\left(N+1\right)\gamma\right)\right)^n \frac{\|x\|_\gamma^n}{\Gamma\left(n\gamma+1\right)},$$

where Γ is the Gamma function and $\|x\|_\gamma$ denotes the γ-Hölder norm of the rough path x.

Recently Gubinelli [4] proposed a *non-geometric* theory of rough path, known as the branched rough paths. The phrase "non-geometric" here refers to that the calculus with respect to branched rough paths does not have to satisfy the chain rule

$$\mathrm{d}\left(XY\right)=Y\mathrm{d}X+X\mathrm{d}Y$$

which Lyons' geometric rough paths must satisfy. For the Brownian motion B, the rough path (almost surely defined)

$$(s,t)\rightarrow\left(1,\int_s^t \mathrm{d}B_{s_1},\int_s^t\int_s^{u_1}\mathrm{d}B_{s_1}\otimes\mathrm{d}B_{s_2}\right)$$

is geometric if the integration is defined in the sense of Stratonovich and non-geometric if the integration is defined in the sense of Itô. Branched rough paths are indexed by the Connes-Kremier Hopf algebra of labelled rooted trees, which we will denote by $\mathcal{H}_\mathcal{L}$ and will recall in Section 2. The multiplication of trees in $\mathcal{H}_\mathcal{L}$ corresponds to the multiplication of the coordinate components of the path, while the operation of joining forests to a single root corresponds to integrating against the path. The theory of branched rough paths is a rough path analogue Butcher's tree-indexed series expansions of solutions to differential equations and has also been motivated by expansions in stochastic partial differential equations. We now recall an equivalent definition of branched rough path due to Hairer-Kelly [3].

Definition 1. ([4], [3])Let $0<\gamma\leq 1$. Let $(\mathcal{H}_\mathcal{L},\cdot,\triangle,S)$ be the Connes-Kremier Hopf algebra of rooted trees labelled by a finite set $\mathcal{L}$. Let $(\mathcal{H}_\mathcal{L}^*,\star,\delta,s)$ be the dual Hopf algebra of $(\mathcal{H}_\mathcal{L},\cdot,\triangle,S)$. A γ-branched rough path is a map $X:[0,1]\times[0,1]\rightarrow\mathcal{H}_\mathcal{L}^*$ such that

1. for all $s\leq t$ and all $h_1,h_2\in\mathcal{H}_\mathcal{L}$,

$$\langle X_{s,t},h_1\rangle\langle X_{s,t},h_2\rangle=\langle X_{s,t},h_1\cdot h_2\rangle. \tag{1.3}$$

2. for all $u\leq s\leq t$,

$$X_{u,s}\star X_{s,t}=X_{s,t}. \tag{1.4}$$

3.for all labelled rooted tree τ, if $|\tau|$ denote the number of vertices in τ, then

$$\|X\|_{\gamma,\tau}:=\sup_{s\neq t}\frac{|\langle X_{s,t},\tau\rangle|}{|t-s|^{\gamma|\tau|}}<\infty. \tag{1.5}$$

Remark 2. Hairer-Kelly [3] pointed out that the product $\star$ in $\mathcal{H}_\mathcal{L}^*$ is induced by the coproduct $\triangle$ on $\mathcal{H}_\mathcal{L}$ in the following sense: If $h\in\mathcal{H}_\mathcal{L}$ is a rooted tree and $\triangle h=\sum \boldsymbol{h^{(1)}}\otimes h^{(2)}$, then for $X,Y\in\mathcal{H}_\mathcal{L}^*$,

$$\langle X\star Y,h\rangle=\sum\langle X,\boldsymbol{h^{(1)}}\rangle\langle Y,h^{(2)}\rangle.$$

Hairer-Kelly also realised that condition 1. in the Definition 1 of branched rough path is equivalent to X taking value in the group of characters of the Hopf algebra, known as the *Butcher group*, analogous to the nilpotent Lie group in the geometric case.

Example 3. Let $\frac{1}{2} < \gamma \leq 1$, and $x = \left(x^1, \ldots, x^d\right) : [0,1] \to \mathbb{R}^d$ be a γ-Hölder path in the sense that

$$\sup_{s \neq t} \frac{\|x_t - x_s\|}{|t-s|^{\gamma}} < \infty.$$

The $\gamma > \frac{1}{2}$ assumption allows us to use Young's integration. According to Gubinelli [4], we may lift x to a γ-branched rough path X in the following way:

Let $\bullet_i$ be a vertex labelled by $i \in \{1, \ldots, d\}$. Let $\tau_1, \ldots, \tau_n$ be rooted trees labelled by $\{1, \ldots, d\}$, and $[\tau_1, \ldots, \tau_n]_{\bullet_i}$ denote the labelled tree obtained by connecting the roots of $\tau_1, \ldots, \tau_n$ to the labelled vertex $\bullet_i$. Then X is defined inductively by $\langle X_{s,t}, \bullet_i \rangle = x_t^i - x_s^i$ and

$$\langle X_{s,t}, [\tau_1, \ldots, \tau_n]_{\bullet_i} \rangle = \int_s^t \Pi_{j=1}^n \langle X_{s,u}, \tau_j \rangle \mathrm{d}x_u^i. \tag{1.6}$$

We now explain why the integration in (1.6) can be defined in the sense of Young: if $u \to \langle X_{s,u}, \tau_j \rangle$ is γ-Hölder, then the product $\Pi_{j=1}^n \langle X_{s,u}, \tau_j \rangle$ is also γ-Hölder. By for example Theorem 1.16 in [8], the integral in (1.6) is also γ-Hölder. More generally, in [3], Hairer-Kelly gave an explicit way of extending a γ-geometric rough path to a γ-branched rough path.

For general $0 < \gamma \leq 1$, let $N = \lfloor \gamma^{-1} \rfloor$, then by Theorem 7.3 in [4], given a family of real-valued functions $(\langle X_{\cdot,\cdot}, \tau \rangle)_{\tau \in \mathcal{H}_{\mathcal{L}}, |\tau| \leq N}$ on $[0,1] \times [0,1]$ satisfying the conditions (1),(2) and (3) in Definition 1 of branched rough path, there is a unique way of extending $(\langle X_{\cdot,\cdot}, \tau \rangle)_{|\tau| \leq N}$ to a γ-branched rough path $(\langle X_{\cdot,\cdot}, \tau \rangle)_{\tau \in \mathcal{H}_{\mathcal{L}}, |\tau| \geq 0}$. Gubinelli [4] conjectured that this extension, which can be interpreted as the iterated integrals of the truncated branched rough path $(\langle X_{\cdot,\cdot}, \tau \rangle)_{\tau \in \mathcal{H}, |\tau| \leq N}$ has a tree factorial decay. Our main result, stated below, is a proof of this conjecture.

Theorem 4. *Let $0 < \gamma \leq 1$ and $N = \lfloor \gamma^{-1} \rfloor$. Let X be a $\gamma-$branched rough path. For all rooted trees τ and all $s \leq t$,*

$$|\langle X_{s,t}, \tau \rangle| \leq \frac{\bar{c}_N^{|\tau|} (t-s)^{\gamma|\tau|}}{\tau!^{\gamma}}. \tag{1.7}$$

where

$$\bar{c}_N = 6 \exp\left(7 \sum_{i=0}^{N+1} (N+1)^{i+1}\right) \left|\mathcal{T}^N\right|^{2-2\gamma} 2^{(N+1)\gamma} \zeta\left((N+1)\gamma\right) N!^{\gamma} \max_{1 \leq |\sigma| \leq N} \|X\|_{\gamma,\sigma}^{|\sigma|^{-1}},$$

$\|X\|_{\gamma,\sigma}$ *is the Hölder norm of X as defined in (1.5) and $\mathcal{T}^N$ is the set of unlabelled rooted trees with at most N vertices.*

Remark 5. For $\gamma = 1$, Gubinelli [4] showed that the decay rate in (1.7) is attained for the identity path X, defined for all rooted trees τ by

$$\langle X_{s,t}, \tau \rangle = \frac{(t-s)^{|\tau|}}{\tau!}.$$

Remark 6. Gubinelli [5] used a similar type of factorial decay estimate to prove the convergence of his series expansion for the solution of the three-dimensional Navier-Stokes equation for sufficiently small initial data.

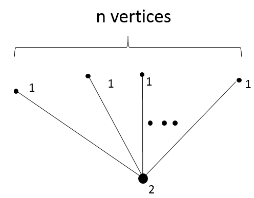


FIGURE 1.1. Tree with n branches

Theorem 4, together with the Hairer-Kelly result [3] that geometric rough paths are branched rough paths, gives another proof for the factorial decay for geometric rough paths. In some cases, our main result gives a sharper estimate for the shuffled sum of the iterated integrals than the one derived using shuffle product and the factorial decay for geometric rough paths, as the following example demonstrates.

Example 7. Let x and y be real valued $\gamma-$Hölder paths on $[0,1]$, $\frac{1}{2} < \gamma \leq 1$. Then we may estimate absolute value of the integral

$$\int_0^1 x_s^n \mathrm{d}y_s$$

by the "geometric method"

$$\begin{aligned} \left| \int_0^1 x_s^n \mathrm{d}y_s \right| &= n! \left| \int_0^1 \int_0^s \ldots \int_0^{s_2} \mathrm{d}x_{s_1} \ldots \mathrm{d}x_{s_n} \mathrm{d}y_s \right| \\ &\leq n! \left(1 + \zeta\left(2\gamma\right)\right)^{n-1} \frac{\|(x,y)\|_\gamma^{n+1}}{(n+1)!^\gamma} \end{aligned} \tag{1.8}$$

where the inequality follows from Lyons' factorial decay estimate (1.2).

By Example 3, we may extend the two-dimensional path (x,y) to a $\gamma-$branched rough path X. Let $\boldsymbol{\sigma}_n$ be the labelled forest defined inductively by $\boldsymbol{\sigma}_1 = \bullet_1$ and $\boldsymbol{\sigma}_n = \boldsymbol{\sigma}_{n-1} \cdot \bullet_1$ where the operation $\cdot$ is the formal multiplication of rooted trees. Let $\tau_n = [\boldsymbol{\sigma}_n]_{\bullet_2}$. Graphically, τ_n takes the following form:

Then

$$\langle X, \tau_n \rangle = \int_0^1 x_s^n \mathrm{d}y_s$$

and our main result Theorem 4 gives the estimate

$$\left| \int_0^1 x_s^n \mathrm{d}y_s \right| \leq \frac{\overline{c}_1^{n+1}}{\tau_n!^\gamma} = \frac{\overline{c}_1^{n+1}}{(n+1)^\gamma}, \tag{1.9}$$

where $\overline{c}_1$ is the constant (independent of n) appearing in our main result Theorem 4. As $n \to \infty$, this growth rate is a much better rate than the factorial growth given by the geometric estimate (1.8).

One reason why the geometric method fares badly here is that it bounds the slowest decaying coordinate iterated integrals. This particular integral

$$\int_0^1 \int_0^s \ldots \int_0^{s_2} \mathrm{d}x_{s_1} \ldots \mathrm{d}x_{s_n} \mathrm{d}y_s$$

decays a lot faster than the slowest decaying coordinate integral of order $n+1$. The failure of the geometric method demonstrates that the case of "fat tree" in Figure

1.1 is delicate and requires a different type of estimate, which is the purpose of Section 5. This problem of fat tree will be explored in more details in Section 3.

1.1. **The strategy of proof.** Lyons [7] proved the factorial decay for γ-geometric rough paths using the following inductive definition of X^n

$$X^n_{u,t} = \lim_{\substack{|\mathcal{P}|\to 0\\ \mathcal{P}\subset[u,t]}} \sum_{t_i\in\mathcal{P}} \sum_{k=1}^{n-1} X^{n-k}_{u,t_i} X^k_{t_i,t_{i+1}}. \tag{1.10}$$

This approach requires the use of a highly non-trivial binomial-type inequality, known as the neoclassical inequality, of the form

$$\sum_{i=0}^{n} \frac{a^{i\gamma} b^{(n-i)\gamma}}{\Gamma(i\gamma+1)\,\Gamma((n-i)\gamma+1)} \leq \gamma^{-2} \frac{(a+b)^{n\gamma}}{\Gamma(n\gamma+1)} \tag{1.11}$$

which is proved Lyons' 98 paper [7]. A sharp version of this inequality latter appeared in the work of Hara and Hino [2]. Gubinelli showed in [4] that a sufficient condition for the factorial decay for branched rough paths is a neoclassical inequality for rooted trees. Unfortunately, we are able to give a counter example for such inequality (see Lemma 10 in Section 3). Lyons' 94 approach [6], which proved the factorial decay for the $\gamma > \frac{1}{2}$ case, did not use the neoclassical inequality and use instead the equivalent definition

$$X^n_{u,t} = \lim_{\substack{|\mathcal{P}|\to 0\\ \mathcal{P}\subset[u,t]}} \sum_{t_i\in\mathcal{P}} \sum_{k=1}^{N} X^{n-k}_{u,t_i} X^k_{t_i,t_{i+1}} \tag{1.12}$$

where $N = \lfloor \gamma^{-1} \rfloor$. This approach also has its own difficulty, due to the fact that the function $(s,t) \to \omega_u(s,t)$ defined by

$$\omega_u(s,t) = \left(\sum_{k=N+1}^{m} \frac{(s-u)^{m-k}(t-s)^k}{(m-k)!k!} \right)^{\frac{1}{N+1}}$$

is not a control in the sense that $\omega_u(s,v) + \omega_u(v,t) \nleq \omega_u(s,t)$. The control property is essential in the use of Young's method of estimating (1.12) by successively removing partition points from the partition. Lyons' 94 approach [6] gets around this problem by using a control function $(s,t) \to R_u(s,t)$ which dominates $\omega_u(s,t)$ and satisfies some binomial properties similar to that of $\omega_u(s,t)$. A key difficulty in this paper is to find the right function R in the case of branched rough paths. Our strategy consists of:

(1) Proving a bound for the multiplication operator $\star$ with respect to some norm, analogous to the following bound of tensor product

$$\|a \otimes b\| \leq \|a\| \, \|b\|$$

for $a \in V^{\otimes m}$ and $b \in V^{\otimes n}$ in the geometric case.
(2) Prove that our function R is compatible with the tree multiplication.
(3) Prove that our function R is compatible with the operation of joining forests to a single root.

2. Branched Rough paths: Notation and Terminology

We first recall the setting of the Connes-Kreimer [1] Hopf algebra which indexes Branched rough paths. A rooted tree is a connected, rooted graph such that for every vertex in the graph, there exists a unique path from the root to the vertex. Let $\mathcal{T}$ denote the set of rooted trees. The empty tree will be denoted by 1. A *forest* is a finite set of rooted trees. The set of forests will be denoted by $\mathcal{F}$. We will identify two trees τ_1 and τ_2 if the forests obtained by removing the respective roots from τ_1 and τ_2 are equal. We define a commutative multiplication $\cdot$ on $\mathcal{F}$ by

$$x \cdot y = x \cup y.$$

Let $\bullet$ denote the rooted tree consisted of a single vertex. We will use a bold symbol (e.g. $\boldsymbol{\tau}$) to denote a forest while using the normal symbol (e.g. τ) to denote a rooted tree. For $\boldsymbol{\sigma} = \{\tau_1, \ldots, \tau_n\} \in \mathcal{F}$, where $\tau_1, \ldots, \tau_n$ are rooted non-empty trees, let $[\boldsymbol{\sigma}]_\bullet$ denote the rooted tree obtained by joining the roots of $\tau_1, \ldots, \tau_n$ to the vertex $\bullet$. Note that $\mathcal{F}$ is the set freely generated by elements of the form $\{\bullet\}$, through the operations of $\cdot$ and $\boldsymbol{\sigma} \to [\boldsymbol{\sigma}]_\bullet$. These two operations in fact correspond to the two fundamental operations in rough path theory, namely the multiplication between path components and the integration against a path.

To simplify our notation, we will denote the element $\{\tau_1, \ldots, \tau_n\}$ in $\mathcal{F}$ simply by $\tau_1 \ldots \tau_n$. We will let $\mathcal{H}$ denote the formal vector space spanned by $\mathcal{F}$ over $\mathbb{R}$. For a forest $\boldsymbol{\tau}$, $c(\boldsymbol{\tau})$ will denote the number of non-empty trees in $\boldsymbol{\tau}$ and $|\boldsymbol{\tau}|$ denote the total number of vertices in the forest. For each tree τ, the tree factorial is defined inductively as

$$\begin{aligned} \bullet! &= 1, \\ [\tau_1, \ldots, \tau_n]_\bullet ! &= |[\tau_1, \ldots, \tau_n]_\bullet|\, \tau_1! \ldots \tau_n!. \end{aligned}$$

The factorial of a forest $\tau_1 \ldots \tau_n$ is defined to be $\tau_1! \ldots \tau_n!$.

A coproduct of rooted trees can be inductively defined as $\triangle : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$,

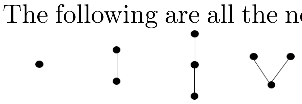
$$\begin{aligned} \triangle 1 &= 1 \otimes 1; \\ \triangle [\tau_1 \ldots \tau_n]_\bullet &= [\tau_1 \ldots \tau_n]_\bullet \otimes 1 + \sum \boldsymbol{\tau_1^{(1)}} \ldots \boldsymbol{\tau_n^{(1)}} \otimes \left[\tau_1^{(2)} \ldots \tau_n^{(2)}\right]_\bullet; \qquad (2.1) \\ \triangle (\tau_1 \ldots \tau_n) &= \triangle \tau_1 \ldots \triangle \tau_n. \end{aligned}$$

where the sum in (2.1) denotes summing over all terms $\boldsymbol{\tau_i^{(1)}}$ and $\tau_i^{(2)}$ in $\triangle \tau_i = \sum \boldsymbol{\tau_i^{(1)}} \otimes \tau_i^{(2)}$. While the coproduct was defined by Connes-Kreimer [1], this particular formulation was borrowed from Hairer-Kelly [3]. Here we define the product $\cdot$ on $\mathcal{H} \otimes \mathcal{H}$ by extending linearly the relation

$$(a \otimes b) \cdot (c \otimes d) = (a \cdot c) \otimes (b \cdot d).$$

The coproduct operator $\triangle$ is coassociative. In Connes-Kreimer's original work [1], an antipode operator S has been constructed explicitly for $\mathcal{H}$, so that the bialgebra $(\mathcal{H}, \cdot, \triangle, S)$ becomes a Hopf algebra. This Hopf algebra is called the *Connes-Kreimer Hopf algebra.*

Example 8. The following are all the non-empty rooted trees with 3 or less vertices



The coproduct $\triangle$ also has an interpretation in terms of cuts. A cut of a rooted tree is a set of edges in a rooted tree. A cut is admissible for a rooted tree τ if for any vertex in τ, the path from the root to the vertex passes through at most one element in the cut. For each admissible cuts c, let $\boldsymbol{\tau_c^{(1)}}$and $\tau_c^{(2)}$denote, respectively the components in $\tau\backslash c$ that is disconnected from the the root and the component that is connected to the root. Then

$$\triangle\tau = \sum_{\text{Admissible cuts } c} \boldsymbol{\tau_c^{(1)}} \otimes \tau_c^{(2)}. \tag{2.2}$$

Given a forest $\boldsymbol{\tau} = \tau_1 \dots \tau_n$ and $\boldsymbol{\sigma^{(1)}}$,$\boldsymbol{\sigma^{(2)}}$ in $\mathcal{F}$, we will define the counting function $c\left(\boldsymbol{\tau}, \boldsymbol{\sigma^{(1)}}, \boldsymbol{\sigma^{(2)}}\right)$ to be the number of times $\boldsymbol{\sigma^{(1)}} \otimes \boldsymbol{\sigma^{(2)}}$ appears in the sum (2.2). We will follow the notation of Gubinelli [4] and use

$$\sum f\left(\boldsymbol{\tau}, \boldsymbol{\tau^{(1)}}, \boldsymbol{\tau^{(2)}}\right)$$

to denote the summation over all $\boldsymbol{\tau^{(1)}}$ and $\boldsymbol{\tau^{(2)}}$ which appears in the sum (2.2).

Remark 9. Although the definition of branched rough paths requires the rooted trees to be labelled, the assumption and conclusion of our main result Theorem 4 are uniform estimates across all labellings. Therefore we can forget that there is any labelling and deal with only unlabelled rooted trees or forests. We will let $\mathcal{T}^n$ and $\mathcal{F}^n$ denote, respectively, the set of all (unlabelled) rooted trees and forests with n vertices.

We say $g \in \mathcal{H}^*$ lies in the group of characters of $\mathcal{H}^*$, which we will denote by $\mathcal{G}$, if for all forests $\boldsymbol{\tau}$ and $\tilde{\boldsymbol{\tau}}$,

$$\langle g, \boldsymbol{\tau} \cdot \tilde{\boldsymbol{\tau}} \rangle = \langle X, \boldsymbol{\tau} \rangle \langle X, \tilde{\boldsymbol{\tau}} \rangle.$$

In other words, $\mathcal{G}$ contains all the homomorphisms g with respect to the tree multiplication $\cdot$. This formulation can be found in, for instance Hairer-Kelly [3].

3. Counter example to the tree neoclassical inequality

We now give a counter example for a weaker version of the neoclassical inequality, which would have been sufficient in proving the factorial decay for the iterated integrals of branched rough paths. The notation $\bullet^n$ will denote the forest

$$\underbrace{\bullet\bullet\dots\bullet}_{n}.$$

Lemma 10. *Let τ_n be the tree $[\bullet^n]_\bullet$. Then for all $0 \le \gamma < 1$, for all $\beta > 0$, there exists $a, b > 0$ such that as $n \to \infty$,*

$$(a+b)^{-\gamma|\tau_n|} \sum \left(\frac{\tau_n!}{\boldsymbol{\tau_n^{(1)}}! \tau_n^{(2)}!} \right)^\gamma \frac{1}{\beta^{c(\boldsymbol{\tau_n^{(1)}}) + c(\tau_n^{(2)})}} a^{\gamma|\boldsymbol{\tau_n^{(1)}}|} b^{\gamma|\tau_n^{(2)}|} \to \infty. \tag{3.1}$$

Proof. By definition, $\tau_n! = n+1$. Observe that by the definition of coproduct $\triangle$ (see (2.1)),

$$\triangle\tau_n = \sum_{l=0}^n \binom{n}{l} \bullet^l \otimes \tau_{n-l} + \tau_n \otimes 1,$$

where $\binom{n}{l}$ denotes the binomial coefficient $\frac{n!}{(n-l)!l!}$. Therefore,

$$\begin{aligned}
& (a+b)^{-\gamma|\tau_n|} \sum \left(\frac{\tau_n!}{\boldsymbol{\tau_n^{(1)}}!\tau_n^{(2)}!}\right)^{\gamma} \frac{1}{\beta^{c(\boldsymbol{\tau_n^{(1)}})+c(\tau_n^{(2)})}} a^{\gamma|\boldsymbol{\tau_n^{(1)}}|} b^{\gamma|\tau_n^{(2)}|} \\
\geq\ & (a+b)^{-\gamma(n+1)} \sum_{l=0}^{n} \left(\frac{n+1}{n+1-l}\right)^{\gamma} \binom{n}{l} \frac{1}{\beta^{l+1}} a^{\gamma l} b^{\gamma(n+1-l)} \\
\geq\ & (a+b)^{-\gamma(n+1)}\, b^{\gamma} \frac{1}{\beta}\left(\frac{a^{\gamma}}{\beta}+b^{\gamma}\right)^{n}, \qquad (3.2)
\end{aligned}$$

where in the last line we used that for $l \leq n$, $(n+1)/(n+1-l) \geq 1$ and the binomial theorem. Since $0 \leq \gamma < 1$ and $a \geq 0$,

$$(1+a)^{\gamma} \leq 1+\gamma a.$$

Therefore, for all $0 < a < (\beta\gamma)^{\frac{1}{\gamma-1}}$,

$$(1+a)^{\gamma} < 1+\frac{a^{\gamma}}{\beta}.$$

Hence for $a < (\beta\gamma)^{\frac{1}{\gamma-1}}$ and $b=1$, (3.2) diverges as n tends to infinity. □

4. Bound for the multiplication operator $\star$

The multiplication $\star$ in the Hopf algebra $\mathcal{H}$ plays the role of the tensor product $\otimes$ in the theory of geometric rough paths. In that case, one of the key assumptions made about the tensor norms is that for all $X^n \in V^{\otimes n}$ and $Y^k \in V^{\otimes k}$,

$$\left\|X^n \otimes Y^k\right\|_{V^{\otimes(n+k)}} \leq \left\|X^n\right\|_{V^{\otimes n}} \left\|Y^k\right\|_{V^{\otimes k}}$$

so that the tensor multiplication has norm 1. We might hope that the multiplication with respect to $\star$ would also have norm 1. Unfortunately, given any numbers n and k, rooted trees in general has more than one way of being cut into two components of sizes n and k respectively. This causes the multiplication operation to have a norm that potentially depends on n and k. Fortunately, and it is a key observation in our proof, the norm can be bounded by a function of k, independently of n. Let us first describe the norm that we use.

Let $X \in \mathcal{H}^*$. Define a linear functional $X^k \in \mathcal{H}^*$ by

$$\left\langle X^k, \tau\right\rangle = \begin{cases} \langle X, \tau\rangle, & |\tau| = k; \\ 0, & |\tau| \neq k. \end{cases}$$

We define $X^n \star Y^k$ such that for all forests $\boldsymbol{\tau}$,

$$\left\langle X^n \star Y^k, \boldsymbol{\tau}\right\rangle = \left\langle X^n \otimes Y^k, \triangle\boldsymbol{\tau}\right\rangle.$$

Let

$$\left\|X^k\right\|_{\mathcal{T},\gamma,\beta} = \max_{|\tau|=k,\tau \text{ trees}} |\langle X, \tau\rangle| \frac{\beta^{c(\tau)}\tau!^{\gamma}}{|\tau|!^{\gamma}},$$

and

$$\left\|X^k\right\|_{\mathcal{F},\gamma,\beta} = \max_{|\boldsymbol{\tau}|=k,\boldsymbol{\tau} \text{ forests}} |\langle X, \boldsymbol{\tau}\rangle| \frac{\beta^{c(\boldsymbol{\tau})}\boldsymbol{\tau}!^{\gamma}}{|\boldsymbol{\tau}|!^{\gamma}}.$$

In this section we will prove a bound on the norm of the multiplication $\star$ with respect to $\|\cdot\|_{\mathcal{T},\gamma,\beta}$, which is the first of three main steps in proving our main result.

Lemma 11. *(Multiplication is bounded in tree norm) Let $\gamma \leq 1$ and that*

$$c_k := \exp\Big[\sum_{i=1}^{k} k^i(1-\gamma)\Big],\ \beta \geq c_k.$$

Let $X, Y \in \mathcal{H}^*$. Then for $n \geq 1$,

$$\left\|X^n \star Y^k\right\|_{\mathcal{T},\gamma,\beta} \leq c_k \left|\mathcal{T}^k\right|^{1-\gamma} \beta^{-1} \left\|X^n\right\|_{\mathcal{F},\gamma,\beta} \left\|Y^k\right\|_{\mathcal{T},\gamma,\beta},$$

where $\mathcal{T}^k$ denotes the set of rooted trees with k vertices.

The proof will require a series of preliminary lemmas involving the combinatorics of rooted trees. We will use $\sum_{\tau^{(2)}=\sigma}$ to denote the sum over all admissible cuts c such that $\tau_c^{(2)} = \sigma$. The following combinatorial lemma is crucial to proving our desired lemma by induction.

Lemma 12. *Let $\tau = [\tau_1 \ldots \tau_n]_\bullet$, where $\tau_1, \ldots, \tau_n$ are non-empty. Let $\sigma = [\sigma_1 \ldots \sigma_n]_\bullet \neq 1$ be a rooted tree. Let $\sim_\sigma$ be a relation on the permutation group $\mathbb{S}_n$ on $\{1, \ldots, n\}$ defined so that $\pi_1 \sim \pi_2$ if $\sigma_{\pi_1(i)} = \sigma_{\pi_2(i)}$ for all i. Let $\mathbf{P}_\sigma$ be the set of $\sim_\sigma$-equivalent classes in the permutation group $\mathbb{S}_n$. Then for all $\beta, \gamma > 0$,*

$$\sum_{\tau^{(2)}=\sigma} \frac{\beta^{-c(\boldsymbol{\tau}^{(1)})}}{\boldsymbol{\tau}^{(1)}!^\gamma} = \sum_{\pi \in \mathbf{P}_\sigma} \Pi_{i=1}^n \Big[\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}} \frac{\beta^{-c(\boldsymbol{\tau}_i^{(1)})}}{\boldsymbol{\tau}_i^{(1)}!^\gamma}\Big].$$

Proof. Note that by the definition of $\triangle$,

$$\sum \boldsymbol{\tau}^{(1)} \otimes \tau^{(2)} = \tau \otimes 1 + \sum \boldsymbol{\tau}_1^{(1)} \ldots \boldsymbol{\tau}_n^{(1)} \otimes \left[\tau_1^{(2)} \ldots \tau_n^{(2)}\right]_\bullet. \tag{4.1}$$

We define a linear functional $\boldsymbol{\sigma}$ such that for each tree a,

$$\begin{aligned} \boldsymbol{\sigma}(a) &= 1,\ \text{if } a = \sigma, \\ &= 0,\ \text{if } a \neq \sigma. \end{aligned}$$

Let $\boldsymbol{f}$ be a linear functional defined such that for each forest $\boldsymbol{\tau}$,

$$\boldsymbol{f}(\boldsymbol{\tau}) = \frac{\beta^{c(\boldsymbol{\tau})}}{\boldsymbol{\tau}!^\gamma}.$$

Note that f is a tree multiplication homomorphism. Define

$$\boldsymbol{f} \otimes \boldsymbol{\sigma}(a \otimes b) = \boldsymbol{f}(a)\, \boldsymbol{\sigma}(b).$$

By applying $\boldsymbol{f} \otimes \boldsymbol{\sigma}$ to (4.1),

$$\sum_{\tau^{(2)}=\sigma} \frac{\beta^{-c(\boldsymbol{\tau}^{(1)})}}{\boldsymbol{\tau}^{(1)}!^\gamma} = \sum_{[\tau_1^{(2)} \ldots \tau_n^{(2)}]_\bullet = \sigma} \frac{\beta^{-c(\boldsymbol{\tau}_1^{(1)}) - \ldots - c(\boldsymbol{\tau}_n^{(1)})}}{\boldsymbol{\tau}_1^{(1)}!^\gamma \ldots \boldsymbol{\tau}_n^{(1)}!^\gamma}.$$

As $[\tau_1^{(2)}, \ldots, \tau_n^{(2)}]_\bullet = [\sigma_1^{(2)}, \ldots, \sigma_n^{(2)}]_\bullet$ if and only if there exists $\pi \in \mathbf{P}_\sigma$ such that $\tau_i^{(2)} = \sigma_{\pi(i)}$ for all i,

$$\sum_{\tau^{(2)}=\sigma} \frac{\beta^{-c(\boldsymbol{\tau}^{(1)})}}{\boldsymbol{\tau}^{(1)}!^\gamma} = \sum_{\pi \in \mathbf{P}_\sigma} \Pi_{i=1}^n \sum_{\tau_i^{(2)}=\sigma_{\pi(i)}} \frac{\beta^{-c(\boldsymbol{\tau}_i^{(1)})}}{\boldsymbol{\tau}_i^{(1)}!^\gamma}.$$

$\square$

Corollary 13. *Let $\tau = [\tau_1 \dots \tau_n]_\bullet$, where $\tau_1, \dots, \tau_n$ are non-empty. Let $\sigma = [\sigma_1 \dots \sigma_n]_\bullet \neq 1$ be a rooted tree. Let $\mathbf{P}'_{\tau,\sigma}$ denote the set of all $\pi \in \mathbf{P}_\sigma$ such that $\sigma_{\pi(i)} \subseteq \tau_i$ for all i. Then for all $\beta, \gamma > 0$,*

$$\sum_{\tau^{(2)}=\sigma} \frac{\beta^{-c(\boldsymbol{\tau^{(1)}})}}{\boldsymbol{\tau^{(1)}}!^\gamma} = \sum_{\pi \in \mathbf{P}'_{\tau,\sigma}} \Pi_{i:\sigma_{\pi(i)} \subsetneq \tau_i} \Big[\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}} \frac{\beta^{-c(\boldsymbol{\tau_i^{(1)}})}}{\boldsymbol{\tau_i^{(1)}}!^\gamma} \Big].$$

Proof. We have just shown in Lemma 12 that

$$\sum_{\tau^{(2)}=\sigma} \frac{\beta^{-c(\boldsymbol{\tau^{(1)}})}}{\boldsymbol{\tau^{(1)}}!^\gamma} = \sum_{\pi \in \mathbf{P}_\sigma} \Pi_{i=1}^n \sum_{\tau_i^{(2)}=\sigma_{\pi(i)}} \frac{\beta^{-c(\boldsymbol{\tau_i^{(1)}})}}{\boldsymbol{\tau_i^{(1)}}!^\gamma}. \tag{4.2}$$

For $\pi \in \mathbf{P}_\sigma$, if an index i is such that $\sigma_{\pi(i)}$ is not a subtree of τ_i, then

$$\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}} \frac{\beta^{-c(\boldsymbol{\tau_i^{(1)}})}}{\boldsymbol{\tau_i^{(1)}}!^\gamma} = 0.$$

Therefore, in (4.2), summing over $\mathbf{P}_\sigma$ is equivalent to summing over $\mathbf{P}'_{\tau,\sigma}$. Furthermore, as $\tau_i^{(1)} = 1$ if $\tau_i^{(2)} = \tau_i$,

$$\sum_{\tau^{(2)}=\sigma} \frac{\beta^{-c(\boldsymbol{\tau^{(1)}})}}{\boldsymbol{\tau^{(1)}}!^\gamma} \quad = \quad \sum_{\pi \in \mathbf{P}'_{\tau,\sigma}} \Pi_{i:\sigma_{\pi(i)} \subsetneq \tau_i} \Big[\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}} \frac{\beta^{-c(\boldsymbol{\tau_i^{(1)}})}}{\boldsymbol{\tau_i^{(1)}}!^\gamma} \Big]. \tag{4.3}$$

□

A key step in most factorial decay estimates for rough paths is to take the fractional power γ outside a sum. In the geometric case, the job is done by the neo-classical inequality. We need the following concavity estimate in the non-geometric case.

Lemma 14. *(Concavity estimate) Let $\gamma \leq 1$. For any rooted tree σ, let*

$$c_{|\sigma|} = \exp\Big[\sum_{i=1}^{|\sigma|} |\sigma|^i (1-\gamma) \Big] \text{ and } \beta \geq c_{|\sigma|}. \tag{4.4}$$

For all rooted trees τ and $\sigma \subsetneq \tau$, we have

$$\sum_{\tau^{(2)}=\sigma} \frac{\beta^{-c(\boldsymbol{\tau^{(1)}})}}{\boldsymbol{\tau^{(1)}}!^\gamma} \quad \leq \quad c_{|\sigma|} \beta^{-1} \Big(\sum_{\tau^{(2)}=\sigma} \frac{1}{\boldsymbol{\tau^{(1)}}!} \Big)^\gamma.$$

Remark 15. The key point is that the constant we lose by taking the power γ outside the sum, $c_{|\sigma|}$, depends only on $|\sigma|$ but not $|\tau|$. To achieve this, the conventional estimate for sums

$$\sum_{i=1}^n a_i^\gamma \leq n^{1-\gamma} \big(\sum_{i=1}^n a_i \big)^\gamma \tag{4.5}$$

is insufficient by itself. We must use the tree-multiplicative property of the tree-factorial.

To prove the concavity estimate, Lemma 14, we first need a counting lemma.

Lemma 16. *Let* $\sigma = [\sigma_1, \dots, \sigma_n]_\bullet$ *and* $\tau = [\tau_1, \dots, \tau_n]_\bullet$ *be rooted trees such that* $\tau_i \neq 1$ *for all* i. Let $k_{\tau,\sigma} = \min_{\pi \in \mathbf{P}'_{\tau,\sigma}} \left|\{i : \sigma_{\pi(i)} \subsetneq \tau_i\}\right|$. Then $k_{\tau,\sigma} \geq 1$ and

$$|\mathbf{P}'_{\tau,\sigma}| \leq \exp(|\sigma|^2 k_{\tau,\sigma}).$$

Proof. As $\tau \neq \sigma$ there does not exist permutation π such that $\sigma_{\pi(i)} = \tau_i$ for all i. In particular, we have $k_{\tau,\sigma} \geq 1$, which proves the first part of the lemma.

Let m be defined by $m = |\{i : \sigma_i = 1\}|$. As $\sigma_j = \sigma_l$ for all $j, l \in \{i : \sigma_i = 1\}$, each equivalence class in $\mathbf{P}_\sigma$ must contain at least $m!$ elements. Therefore,

$$|\mathbf{P}'_{\tau,\sigma}| \leq |\mathbf{P}_\sigma| \leq \frac{n!}{m!} = n(n-1)\dots(m+1).$$

Since

$$n = m + \left|\{i : \sigma_i \neq 1\}\right| \leq m + |\sigma|,$$

we have

$$|\mathbf{P}'_{\tau,\sigma}| \quad \leq \quad (m + |\sigma|)^{|\sigma|}. \tag{4.6}$$

Note that as τ_i is assumed to be non-empty for all i,

$$m = \left|\{i : \sigma_i = 1\}\right| \leq \min_{\pi \in \mathbf{P}'_{\tau,\sigma}} \left|\{i : \sigma_{\pi(i)} \subsetneq \tau_i\}\right| = k_{\tau,\sigma}. \tag{4.7}$$

Using that for $x \geq 1$ and $b \in \mathbb{N} \cup \{0\}$,

$$(x+b)^b \leq \exp\left(b^2 x\right)$$

in combination with the estimates (4.6) and (4.7) earlier in this proof,

$$\left|\mathbf{P}'_{\tau,\sigma}\right| \leq (m + |\sigma|)^{|\sigma|} \leq (k_{\tau,\sigma} + |\sigma|)^{|\sigma|} \leq \exp\left(|\sigma|^2 k_{\tau,\sigma}\right).$$

□

Proof of concavity estimate Lemma 15. We will prove the lemma by induction on $|\sigma|$. If $|\sigma| = 0$, then $\sigma = 1$ and as $\sigma \neq \tau$,

$$\sum_{\tau^{(2)} = \sigma} \frac{\beta^{-c(\boldsymbol{\tau}^{(1)})}}{\boldsymbol{\tau}^{(1)}!^\gamma} \quad = \quad \frac{\beta^{-1}}{\tau!^\gamma}$$

which is exactly the content of the present lemma for the case $|\sigma| = 0$. Let $\tau = [\tau_1 \dots \tau_n]_\bullet$, where $\tau_1, \dots, \tau_n$ are all non-empty, and $\sigma = [\sigma_1 \dots \sigma_n]_\bullet$. Using Corollary 13 that relates the sum $\sum_{\tau^{(2)} = \sigma}$ to $\sum_{\tau_i^{(2)} = \sigma_{\pi(i)}}$,

$$\sum_{\tau^{(2)} = \sigma} \frac{\beta^{-c(\boldsymbol{\tau}^{(1)})}}{\boldsymbol{\tau}^{(1)}!^\gamma} \quad = \quad \sum_{\pi \in \mathbf{P}'_{\tau,\sigma}} \Pi_{i : \sigma_{\pi(i)} \subsetneq \tau_i} \Big[\sum_{\tau_i^{(2)} = \sigma_{\pi(i)}} \frac{\beta^{-c(\boldsymbol{\tau}_i^{(1)})}}{\boldsymbol{\tau}_i^{(1)}!^\gamma} \Big]. \tag{4.8}$$

By the induction hypothesis and that $k_{\tau,\sigma} = \min_{\pi\in\mathbf{P}'_{\tau,\sigma}} \big|\{i : \sigma_{\pi(i)} \subsetneq \tau_i\}\big|$ by Lemma 16,

$$\sum_{\pi\in\mathbf{P}'_{\tau,\sigma}} \Pi_{i:\sigma_{\pi(i)}\subsetneq\tau_i}\Big[\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}} \frac{\beta^{-c(\boldsymbol{\tau_i^{(1)}})}}{\boldsymbol{\tau_i^{(1)}}!^{\gamma}}\Big] \tag{4.9}$$

$$\leq \Big(\beta^{-1}\exp\big(\sum_{j=1}^{|\sigma|-1}(|\sigma|-1)^j(1-\gamma)\big)\Big)^{k_{\tau,\sigma}} \tag{4.10}$$

$$\times\sum_{\pi\in\mathbf{P}'_{\tau,\sigma}}\Big(\Pi_{i:\sigma_{\pi(i)}\subsetneq\tau_i}\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}}\frac{1}{\boldsymbol{\tau_i^{(1)}}!}\Big)^{\gamma}. \tag{4.11}$$

By the conventional concavity estimate for sum $\sum_{i=1}^n a_i^\gamma \leq n^{1-\gamma}(\sum_{i=1}^n a_i)^\gamma$,

$$\sum_{\pi\in\mathbf{P}'_{\tau,\sigma}}\Big(\Pi_{i:\sigma_{\pi(i)}\subsetneq\tau_i}\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}}\frac{1}{\boldsymbol{\tau_i^{(1)}}!}\Big)^{\gamma} \tag{4.12}$$

$$\leq |\mathbf{P}'_{\tau,\sigma}|^{1-\gamma}\Big(\sum_{\pi\in\mathbf{P}'_{\tau,\sigma}}\Pi_{i:\sigma_{\pi(i)}\subsetneq\tau_i}\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}}\frac{1}{\boldsymbol{\tau_i^{(1)}}!}\Big)^{\gamma}. \tag{4.13}$$

Using our estimate for $|\mathbf{P}'_{\tau,\sigma}|$ in Lemma 16,

$$\sum_{\pi\in\mathbf{P}'_{\tau,\sigma}}\big(\Pi_{i:\sigma_{\pi(i)}\subsetneq\tau_i}\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}}\frac{1}{\boldsymbol{\tau_i^{(1)}}!}\Big)^{\gamma} \tag{4.14}$$

$$\leq \exp(|\sigma|^2 k_{\tau,\sigma}(1-\gamma))\Big(\sum_{\pi\in\mathbf{P}'_{\tau,\sigma}}\Pi_{i:\sigma_{\pi(i)}\subsetneq\tau_i}\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}}\frac{1}{\boldsymbol{\tau_i^{(1)}}!}\Big)^{\gamma}. \tag{4.15}$$

Combining the identity (4.8) with all the inequalities we have so far, namely (4.10) and (4.14),

$$\begin{aligned}&\sum_{\tau^{(2)}=\sigma}\frac{\beta^{-c(\boldsymbol{\tau^{(1)}})}}{\boldsymbol{\tau^{(1)}}!^{\gamma}}\\ \leq\;& \Big(\beta^{-1}\exp\big[(1-\gamma)(|\sigma|^2+\sum_{j=1}^{|\sigma|-1}(|\sigma|-1)^j)\big]\Big)^{k_{\tau,\sigma}}\\ &\times\big(\sum_{\pi\in\mathbf{P}'_{\tau,\sigma}}\Pi_{i:\sigma_{\pi(i)}\subsetneq\tau_i}\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}}\frac{1}{\boldsymbol{\tau_i^{(1)}}!}\big)^{\gamma}.\end{aligned}$$

Finally, as $k_{\tau,\sigma}\geq 1$ (see counting lemma, Lemma 16) and $\beta \geq \exp\Big[\sum_{i=1}^{|\sigma|}|\sigma|^i(1-\gamma)\Big]$,

$$\sum_{\tau^{(2)}=\sigma}\frac{\beta^{-c(\boldsymbol{\tau^{(1)}})}}{\boldsymbol{\tau^{(1)}}!^{\gamma}} \leq \beta^{-1}\exp\big(\sum_{j=1}^{|\sigma|}|\sigma|^j(1-\gamma)\big)\big(\sum_{\pi\in\mathbf{P}'_{\tau,\sigma}}\Pi_{i:\sigma_{\pi(i)}\subsetneq\tau_i}\sum_{\tau_i^{(2)}=\sigma_{\pi(i)}}\frac{1}{\boldsymbol{\tau_i^{(1)}}!}\big)^{\gamma}.$$

Note now by the $\gamma = 1$ and $\beta = 1$ case of Corollary 13,

$$\sum_{\pi \in \mathbf{P}'_{\tau,\sigma}} \Pi_{i:\sigma_{\pi(i)} \subsetneq \tau_i} \sum_{\tau_i^{(2)} = \sigma_{\pi(i)}} \frac{1}{\boldsymbol{\tau_i^{(1)}}!} = \sum_{\tau^{(2)} = \sigma} \frac{1}{\boldsymbol{\tau^{(1)}}!}.$$

□

We now state a lemma that is equivalent to Gubinelli's tree-binomial theorem [4]. It allows us to rewrite sum over rooted trees to a sum over integers.

Lemma 17. *(Tree binomial theorem) Let τ be a rooted tree. Then*

$$\sum_{|\tau^{(2)}|=l} \frac{\tau!}{\boldsymbol{\tau^{(1)}}!\tau^{(2)}!} = \binom{|\tau|}{l}.$$

Proof. By the tree binomial theorem, Lemma 4.4 in [4], we have for $x \in \mathbb{R}$,

$$\frac{(1+x)^{|\tau|}}{\tau!} = \sum \frac{x^{|\tau^{(2)}|}}{\boldsymbol{\tau^{(1)}}!\tau^{(2)}!}.$$

The result follows by comparing the coefficients of x^l with the classical binomial theorem. □

We now prove the bound on the Hopf algebra multiplication $\star$.

Proof of boundednes of tree multiplication Lemma 11. Note first that by the definition of $\star$, if $|\tau| = n + k$,

$$\begin{aligned} & \left|\langle X^n \star Y^k, \tau\rangle\right| \\ (4.16) \qquad = & \ \Big| \sum_{|\tau^{(2)}|=k} \langle X^n, \boldsymbol{\tau^{(1)}}\rangle \langle Y^k, \tau^{(2)}\rangle \Big| \\ \leq & \sum_{|\tau^{(2)}|=k} |\langle X^n, \boldsymbol{\tau^{(1)}}\rangle| \, |\langle Y^k, \tau^{(2)}\rangle| \end{aligned}$$

By the definition of $\|\cdot\|_{\mathcal{F},\gamma,\beta}$ and $\|\cdot\|_{\mathcal{T},\gamma,\beta}$,

$$\begin{aligned} & \sum_{|\tau^{(2)}|=k} |\langle X^n, \boldsymbol{\tau^{(1)}}\rangle| \, |\langle Y^k, \tau^{(2)}\rangle| \\ (4.17) \qquad \leq & \ \|X^n\|_{\mathcal{F},\gamma,\beta} \left\|Y^k\right\|_{\mathcal{T},\gamma,\beta} (n!k!)^\gamma \sum_{|\tau^{(2)}|=k} \frac{\beta^{-c(\boldsymbol{\tau^{(1)}})-1}}{(\boldsymbol{\tau^{(1)}}!\tau^{(2)}!)^\gamma}. \end{aligned}$$

As we assumed that $n \geq 1$, we have $k < |\tau|$ and hence we may apply the concavity estimate for trees, Lemma 14, to obtain

$$\begin{aligned} \sum_{|\tau^{(2)}|=k} \frac{\beta^{-c(\boldsymbol{\tau^{(1)}})}}{(\boldsymbol{\tau^{(1)}}!\tau^{(2)}!)^\gamma} & \leq \sum_{|\sigma|=k} \frac{1}{\sigma!^\gamma} \sum_{\tau^{(2)}=\sigma} \frac{\beta^{-c(\boldsymbol{\tau^{(1)}})}}{\boldsymbol{\tau^{(1)}}!^\gamma} \\ & \leq c_k \beta^{-1} \sum_{|\sigma|=k} \frac{1}{\sigma!^\gamma} \Big(\sum_{\tau^{(2)}=\sigma} \frac{1}{\boldsymbol{\tau^{(1)}}!} \Big)^\gamma \\ & \leq c_k \left|\mathcal{T}^k\right|^{1-\gamma} \beta^{-1} \Big(\sum_{|\tau^{(2)}|=k} \frac{1}{\boldsymbol{\tau^{(1)}}!\tau^{(2)}!} \Big)^\gamma. \end{aligned}$$

We now use the tree binomial theorem (Lemma 17) to deduce that

$$\sum_{|\tau^{(2)}|=k} \frac{\beta^{-c(\boldsymbol{\tau^{(1)}})}}{\left(\boldsymbol{\tau^{(1)}}!\tau^{(2)}!\right)^{\gamma}} \leq c_k \left|\mathcal{T}^k\right|^{1-\gamma} \beta^{-1} \frac{1}{\tau!^{\gamma}} \binom{|\tau|}{k}^{\gamma}. \tag{4.18}$$

Therefore, for all rooted trees τ such that $|\tau| = n + k$, by substituting (4.18) into (4.17),

$$\left|\left\langle X^n \star Y^k, \tau\right\rangle\right| \leq c_k \left|\mathcal{T}^k\right|^{1-\gamma} \left(\frac{|\tau|!}{\tau!}\right)^{\gamma} \beta^{-2} \left\|X^n\right\|_{\mathcal{F},\gamma,\beta} \left\|Y^k\right\|_{\mathcal{T},\gamma,\beta}$$

and we have

$$\begin{aligned} \left\|X^n \star Y^k\right\|_{\mathcal{T},\gamma,\beta} &= \max_{|\tau|=n+k, \tau \text{ trees}} \beta \left|\left\langle X^n \star Y^k, \tau\right\rangle\right| \frac{\tau!^{\gamma}}{|\tau|!^{\gamma}} \\ &\leq c_k \left|\mathcal{T}^k\right|^{1-\gamma} \beta^{-1} \left\|X^n\right\|_{\mathcal{F},\gamma,\beta} \left\|Y^k\right\|_{\mathcal{T},\gamma,\beta}. \end{aligned}$$

$\square$

5. Compatibility of our estimate with tree multiplication

We showed in the last section the multiplicative bound

$$\left\|X^n \star Y^k\right\|_{\mathcal{T},\gamma,\beta} \leq c_k \left|\mathcal{T}^k\right|^{1-\gamma} \beta^{-1} \left\|X^n\right\|_{\mathcal{F},\gamma,\beta} \left\|Y^k\right\|_{\mathcal{T},\gamma,\beta} \tag{5.1}$$

with $\beta \geq c_k$. That we can choose a large β is very useful. It will help us to annihilate any constant depending on k. Suppose that $X_{\cdot,\cdot}$ is a branched rough path and $X^n_{s,t}$ denotes the restriction of the branched rough path X on trees with n vertices. Let $(t_0, \ldots, t_r)$ be a partition for $[s,t]$, then in a similar spirit to Lyons 94'[6], we have

$$X^{n+1}_{s,t} = \lim_{|t_i - t_{i+1}| \to \infty} \sum_{i=0}^{r-1} \sum_{k=1}^{\lfloor \frac{1}{\gamma} \rfloor} X^{n+1-k}_{s,t_i} \star X^k_{t_i,t_{i+1}}. \tag{5.2}$$

We would like to apply the multiplicative bound (5.1) to estimate the Riemann sum

$$\left\| \sum_{i=0}^{r-1} \sum_{k=1}^{\lfloor \frac{1}{\gamma} \rfloor} X^{n+1-k}_{s,t_i} \star X^k_{t_i,t_{i+1}} \right\|_{\mathcal{T},\gamma,\beta}. \tag{5.3}$$

A crucial point is that the biggest k can be here is $\lfloor \frac{1}{\gamma} \rfloor$ which is independent of n. The constant in front of $\left\|X^n\right\|_{\mathcal{F},\gamma,\beta} \left\|Y^k\right\|_{\mathcal{T},\gamma,\beta}$ in the multiplicative bound (5.1) is therefore independent of n. If we had use the following formula instead (as in Lyons 98'[7])

$$X^{n+1}_{s,t} = \lim_{|t_i - t_{i+1}| \to 0} \sum_{i=0}^{r-1} \sum_{k=1}^{n} X^{n+1-k}_{s,t_i} \star X^k_{t_i,t_{i+1}}, \tag{5.4}$$

then the biggest k can be is n, and the constant in (5.1) would depend on n. This is the fundamental reason why we must use the approach in Lyons 94' (5.2) instead of using (5.4) as in Lyons 98'.

After applying the bound (5.1) to estimate the Riemann sum (5.3), we should have a bound for the tree norm

$$\left\|X^{n+1}_{s,t}\right\|_{\mathcal{T},\gamma,\beta}.$$

However, to use the multiplicative bound (5.1) for estimating the tree norm $\left\|X_{s,t}^{n+2}\right\|_{\mathcal{T},\gamma,\beta}$, it is not enough to know only the tree norm $\left\|X_{s,t}^{n+1}\right\|_{\mathcal{T},\gamma,\beta}$. We must also know the forest norm $\left\|X_{s,t}^{n+1}\right\|_{\mathcal{F},\gamma,\beta}$. That is why we must find a way of bounding the forest norm $\left\|X_{s,t}^{n+1}\right\|_{\mathcal{F},\gamma,\beta}$ in terms of $\left\|X_{s,t}^{n+1}\right\|_{\mathcal{T},\gamma,\hat{\beta}}$ (with $\hat{\beta}>\beta$), which we aim to achieve in this section through a lemma. We first present the form of our estimate.

Let $\triangle_m(r,r')$ denote the m-dimensional simplex

$$\left\{(s_1,\ldots,s_m)\in\mathbb{R}^m : r<s_1<\ldots<s_m<r'\right\}.$$

For a one-dimensional path ρ, we will define

$$S^{(m)}\left(\rho\right)_{s,t}=\int_{\triangle_m(s,t)}\mathrm{d}\rho\left(s_1\right)\ldots\mathrm{d}\rho\left(s_m\right).$$

For each $a,b>0$ define a one-dimensional path ρ by

$$\rho_a^b\left(t\right)=\frac{1}{b}\left(t-a\right)^b.$$

Define the function $R_u^{n,m}(s,t)$ by

$$R_u^{n,m}(s,t)=S^{(n)}(\rho_u^{m/n})_{s,t}.$$

The construction of our estimate is based on the following lemma, which says that our estimate dominates the tail of a binomial sum. Its proof can be found in the Appendix.

Lemma 18. *Let $N\in\mathbb{N}\cup\{0\}$ and $n\geq N+1$. For all $u<s<t$*

$$\sum_{j=N+1}^{n}\frac{\left(s-u\right)^{n-j}\left(t-s\right)^{j}}{(n-j)!j!}\quad\leq\quad\frac{1}{(n-N-1)!}R_u^{N+1,n}(s,t). \tag{5.5}$$

The following lemma is the main result of this section and allows us to convert our bound from one about the tree norm to the forest norm.

Lemma 19. *(Compatibility with tree multiplication) Let $0<\gamma\leq 1$ and $N=\lfloor\gamma^{-1}\rfloor$. Let X be a γ-branched rough path. Let*

$$\hat{c}_N=3|\mathcal{T}^N|^{1-\gamma}(N+1)^{3(1-\gamma)}\exp 2\left(N+1\right),\ \beta\geq\hat{c}_N.$$

Suppose that for all $n\leq M$ and $u\leq s\leq t$,

$$\|\sum_{k\geq N+1}X_{u,s}^{n-k}\star X_{s,t}^{k}\|_{\mathcal{T},\gamma,\beta}\quad\leq\quad\left[\frac{1}{(n-N-1)!}R_u^{N+1,n}(s,t)\right]^{\gamma}. \tag{5.6}$$

Then for all $n\leq M$,

$$\|\sum_{k\geq N+1}X_{u,s}^{n-k}\star X_{s,t}^{k}\|_{\mathcal{F},\gamma,\beta\hat{c}_N^{-1}}\quad\leq\quad\left[\frac{1}{(n-N-1)!}R_u^{N+1,n}(s,t)\right]^{\gamma}.$$

We will once again need a series of lemmas. The first of which states that for factorial decay estimates the forest norm $\|\cdot\|_{\mathcal{F},\gamma,\beta}$ is the same as tree norm $\|\cdot\|_{\mathcal{T},\gamma,\beta}$.

Lemma 20. *Let $X\in\mathcal{G}$. Let $k\geq 0$ and*

$$\beta\geq\exp\big[\sum_{i=1}^{k}k^i(1-\gamma)\big].$$

If there exists $a > 0$ such that

$$\|X^n\|_{\mathcal{T},\gamma,\beta} \leq \frac{a^{\gamma n}}{n!^{\gamma}}, \tag{5.7}$$

then

$$\|X^n\|_{\mathcal{F},\gamma,\beta} \leq \frac{a^{\gamma n}}{n!^{\gamma}}. \tag{5.8}$$

Proof. By the assumption (5.7), for all rooted trees τ such that $|\tau| = n$,

$$|\langle X^n, \tau\rangle| \leq \frac{a^{\gamma n}}{\beta \tau!^{\gamma}}.$$

Therefore, for any forest $\boldsymbol{\tau} = \tau_1 \dots \tau_m$, where $|\boldsymbol{\tau}| = n$ and $\tau_1 \dots \tau_m$ are rooted trees, by the

$$\begin{aligned} |\langle X^n, \boldsymbol{\tau}\rangle| &= |\langle X, \boldsymbol{\tau}\rangle| \\ &= |\langle X, \tau_1\rangle| \dots |\langle X, \tau_m\rangle| \\ &= |\langle X^{|\tau_1|}, \tau_1\rangle| \dots |\langle X^{|\tau_m|}, \tau_m\rangle| \\ &\leq \frac{a^{\gamma n}}{\beta^m \boldsymbol{\tau}!^{\gamma}}, \end{aligned}$$

which is equivalent to our desired factorial decay estimate (5.8). □

To extend estimates about rooted trees to estimates about forests, we usually need to carry out induction on the number of components in the forest. To carry out such induction, the following algebraic identity is very useful.

Lemma 21. *(Forest factorisation lemma) Let $X, Y \in \mathcal{G}$. Then for any forests $\boldsymbol{\tau}$ and $\tilde{\boldsymbol{\tau}}$ such that $n + k = |\boldsymbol{\tau}| + |\tilde{\boldsymbol{\tau}}|$,*

$$\langle X^n \star Y^k, \boldsymbol{\tau}\tilde{\boldsymbol{\tau}}\rangle = \sum_{k_1+k_2=k} \langle X^{|\boldsymbol{\tau}|-k_1} \star Y^{k_1}, \boldsymbol{\tau}\rangle \langle X^{|\tilde{\boldsymbol{\tau}}|-k_2} \star Y^{k_2}, \tilde{\boldsymbol{\tau}}\rangle.$$

Proof. By definition of $\star$ and that the coproduct $\triangle$ is compatible with tree multiplication,

$$\begin{aligned} \langle X^n \star Y^k, \boldsymbol{\tau}\tilde{\boldsymbol{\tau}}\rangle &= \langle X^n \otimes Y^k, \triangle(\boldsymbol{\tau}\tilde{\boldsymbol{\tau}})\rangle \\ &= \langle X^n \otimes Y^k, \triangle\boldsymbol{\tau}\triangle\tilde{\boldsymbol{\tau}}\rangle \\ &= \sum_{|\boldsymbol{\tau}^{(2)}|+|\tilde{\boldsymbol{\tau}}^{(2)}|=k} \langle X^n, \boldsymbol{\tau}^{(1)}\tilde{\boldsymbol{\tau}}^{(1)}\rangle \langle Y^k, \boldsymbol{\tau}^{(2)}\tilde{\boldsymbol{\tau}}^{(2)}\rangle. \end{aligned} \tag{5.9}$$

As $n + k = |\boldsymbol{\tau}| + |\tilde{\boldsymbol{\tau}}|$ and that we are summing over $|\boldsymbol{\tau}^{(2)}| + |\tilde{\boldsymbol{\tau}}^{(2)}| = k$, we have in particular that $n = |\boldsymbol{\tau}^{(1)}| + |\tilde{\boldsymbol{\tau}}^{(1)}|$ in the sum. As $X \in \mathcal{H}^*$,

$$\begin{aligned} \langle X^n, \boldsymbol{\tau}^{(1)}\tilde{\boldsymbol{\tau}}^{(1)}\rangle &= \langle X, \boldsymbol{\tau}^{(1)}\tilde{\boldsymbol{\tau}}^{(1)}\rangle \\ &= \langle X, \boldsymbol{\tau}^{(1)}\rangle \langle X, \tilde{\boldsymbol{\tau}}^{(1)}\rangle \\ &= \langle X^{|\boldsymbol{\tau}^{(1)}|}, \boldsymbol{\tau}^{(1)}\rangle \langle X^{|\tilde{\boldsymbol{\tau}}^{(1)}|}, \tilde{\boldsymbol{\tau}}^{(1)}\rangle. \end{aligned}$$

Analogous expression also holds for $\langle Y^k, \boldsymbol{\tau}^{(2)}\tilde{\boldsymbol{\tau}}^{(2)}\rangle$ with the same proof. Therefore, by our earlier calculation (5.9) and the definition of $\star$,

$$\begin{aligned}
&\langle X^n \star Y^k, \boldsymbol{\tau}\tilde{\boldsymbol{\tau}}\rangle \\
= &\sum_{|\boldsymbol{\tau}^{(2)}|+|\tilde{\boldsymbol{\tau}}^{(2)}|=k} \langle X^{|\boldsymbol{\tau}^{(1)}|}, \boldsymbol{\tau}^{(1)}\rangle\langle X^{|\tilde{\boldsymbol{\tau}}^{(1)}|}, \tilde{\boldsymbol{\tau}}^{(1)}\rangle\langle Y^{|\boldsymbol{\tau}^{(2)}|}, \boldsymbol{\tau}^{(2)}\rangle\langle Y^{|\tilde{\boldsymbol{\tau}}^{(2)}|}, \tilde{\boldsymbol{\tau}}^{(2)}\rangle \\
= &\sum_{k_1+k_2=k} \langle X^{|\boldsymbol{\tau}|-k_1} \star Y^{k_1}, \boldsymbol{\tau}\rangle\langle X^{|\tilde{\boldsymbol{\tau}}|-k_2} \star Y^{k_2}, \tilde{\boldsymbol{\tau}}\rangle.
\end{aligned}$$

□

The following lemma states that if we assume X^n and Y^k have factorial decay estimates, then the forest norm of $X^n \star Y^k$ can be bounded by the tree norms of X^n and Y^k.

Lemma 22. *(Multiplication is bounded in forest norm) Let $X, Y \in \mathcal{H}^*$. Let $k \geq 0$ and*

$$\beta \geq \exp\Big[\sum_{i=1}^{k} k^i(1-\gamma)\Big].$$

If there exists $a > 0$ and $b > 0$ such that

$$\|X^n\|_{\mathcal{T},\gamma,\beta} \leq \frac{a^{\gamma n}}{n!^\gamma}, \ \left\|Y^k\right\|_{\mathcal{T},\gamma,\beta} \leq \frac{b^{\gamma k}}{k!^\gamma},$$

then

$$\left\|X^n \star Y^k\right\|_{\mathcal{F},\gamma,\tilde{c}_k^{-1}\beta} \leq \frac{a^{\gamma n} b^{\gamma k}}{(n!k!)^\gamma},$$

where $\tilde{c}_k = c_k((k+1)\,|\mathcal{T}_k|)^{1-\gamma}$.

Proof. We need to show that for all forests $\boldsymbol{\tau}$, and (n,k) such that $n+k = |\boldsymbol{\tau}|$,

$$\left|\langle X^n \star Y^k, \boldsymbol{\tau}\rangle\right| \leq \frac{\tilde{c}_k^{c(\boldsymbol{\tau})} a^{\gamma n} b^{\gamma k}}{\beta^{c(\boldsymbol{\tau})}\boldsymbol{\tau}!^\gamma}\binom{|\boldsymbol{\tau}|}{k}^\gamma. \tag{5.10}$$

We shall prove it by induction on $c(\boldsymbol{\tau})$. If $n = 0$, then present lemma directly follows from assumption and we will henceforth assume $n \geq 1$. For $c(\boldsymbol{\tau}) = 1$, note that in this case $\boldsymbol{\tau}$ is a tree. By the boundedness of group multiplication in tree norm, Lemma 11, and that the tree norm of X is the same as the forest norm of X (see Lemma 20),

$$\left\|X^n \star Y^k\right\|_{\mathcal{T},\gamma,\beta} \leq c_k\left|\mathcal{T}^k\right|^{1-\gamma}\beta^{-1}\frac{a^{\gamma n} b^{\gamma k}}{n!^\gamma k!^\gamma},$$

which implies our desired estimate (5.10) in the case when $\boldsymbol{\tau}$ is a rooted tree. For the induction step, let $\boldsymbol{\tau} = \tau_1\boldsymbol{\tau}_2$, where τ_1 is a non-empty tree and $\boldsymbol{\tau}_2$ is a forest. If $n + k = |\boldsymbol{\tau}|$, then by the forest factorisation lemma, Lemma 21, and the induction hypothesis,

$$\begin{aligned}
K &:= \left|\langle X^n \star Y^k, \tau_1\boldsymbol{\tau}_2\rangle\right| \\
&\leq \sum_{l+m=k} |\langle X^{|\tau_1|-l} \star Y^l, \tau_1\rangle|\,|\langle X^{|\boldsymbol{\tau}_2|-m} \star Y^m, \boldsymbol{\tau}_2\rangle| \\
&\leq \frac{\tilde{c}_k{}^{c(\boldsymbol{\tau}_2)} c_k\left|\mathcal{T}^k\right|^{1-\gamma}}{\beta^{c(\boldsymbol{\tau}_2)+1}\,(\tau_1!\boldsymbol{\tau}_2!)^\gamma} a^{\gamma n} b^{\gamma k} \sum_{l+m=k}\binom{|\tau_1|}{l}^\gamma\binom{|\boldsymbol{\tau}_2|}{m}^\gamma.
\end{aligned}$$

Using the conventional concavity estimate for sum $\sum_{i=1}^{M} a_i^{\gamma} \leq M^{1-\gamma}(\sum_{i=1}^{M} a_i)^{\gamma}$ and that $\tilde{c}_k = c_k((k+1)|\mathcal{T}^k|)^{1-\gamma}$,

$$\begin{aligned} K &\leq \frac{\tilde{c}_k{}^{c(\boldsymbol{\tau}_2)} c_k \left|\mathcal{T}^k\right|^{1-\gamma}}{\beta^{c(\boldsymbol{\tau}_2)+1}\left(\tau_1!\boldsymbol{\tau}_2!\right)^{\gamma}} a^{\gamma n} b^{\gamma k}(k+1)^{1-\gamma}\Big(\sum_{l+m=k}\binom{|\tau_1|}{l}\binom{|\boldsymbol{\tau}_2|}{m}\Big)^{\gamma} \\ &\leq \frac{\tilde{c}_k{}^{(c(\boldsymbol{\tau}_2)+1)}}{\beta^{c(\boldsymbol{\tau}_2)+1}\left(\tau_1!\boldsymbol{\tau}_2!\right)^{\gamma}}\binom{|\tau_1|+|\boldsymbol{\tau}_2|}{k}^{\gamma} a^{\gamma n} b^{\gamma k}, \end{aligned}$$

which in particular implies our desired estimate (5.10). □

We will also need the following binomial lemma that describes the product of our estimate over $[s,t]$ with the length of the overlapping interval $[u,t]$, where $u \leq s \leq t$.

Lemma 23. *(Binomial lemma for overlapping time intervals) Let $N \in \mathbb{N} \cup \{0\}$. Let $n \geq N+1$ and $m \geq 0$, and $u \leq s \leq t$,*

$$\frac{n!}{(n-N-1)!} R_u^{N+1,n}(s,t)\,(t-u)^m \tag{5.11}$$

$$\leq \frac{(n+m)!}{(n+m-N-1)!} R_u^{N+1,n+m}(s,t). \tag{5.12}$$

The proof of Lemma 23 can be found in the Appendix.

Proof of the compatibility with tree multiplication, Lemma 20. We need to show that for all forests $\boldsymbol{\tau}$,

$$\Big|\Big\langle \sum_{k\geq N+1} X_{u,s}^{n-k} \star X^k, \boldsymbol{\tau}\Big\rangle\Big| \quad \leq \quad \Big(\frac{\hat{c}_N}{\beta}\Big)^{c(\boldsymbol{\tau})}\Big[\frac{|\boldsymbol{\tau}|!}{\boldsymbol{\tau}!\,(|\boldsymbol{\tau}|-N-1)!} R_u^{N+1,|\boldsymbol{\tau}|}(s,t)\Big]^{\gamma}$$

and will do so via induction on $c(\boldsymbol{\tau})$. The case $c(\boldsymbol{\tau}) = 1$ follows directly from the assumption.

By putting $u = s$ in the assumed estimate (5.6) for trees and using $(N+1)^{N+1}\left[(N+1)!\right]^{-1} \leq \exp(N+1)$, we have that for all forests $\boldsymbol{\tau}$ with $|\boldsymbol{\tau}| \leq M$,

$$|\langle X_{s,t}, \boldsymbol{\tau}\rangle| \leq \frac{C_N^{c(\boldsymbol{\tau})}\,(t-s)^{\gamma|\boldsymbol{\tau}|}}{\beta^{c(\boldsymbol{\tau})}\boldsymbol{\tau}!^{\gamma}} \tag{5.13}$$

where $C_N = \exp(N+1)$ and hence for all $n \leq M$ and all $s \leq t$,

$$\|X_{s,t}^n\|_{\mathcal{F},\gamma,\beta C_N^{-1}} \leq \frac{(t-s)^{\gamma n}}{n!^{\gamma}}.$$

The estimate (5.13) in particular says that our estimate is a factorial decay estimate.

For the induction step, we let τ_1 be a non-empty rooted tree and $\boldsymbol{\tau}_2$ be a forest such that $\boldsymbol{\tau} = \tau_1\boldsymbol{\tau}_2$. By the forest factorisation lemma, Lemma 21,

$$\begin{aligned} &\Big\langle \sum_{k\geq N+1} X_{u,s}^{n-k} \star X_{s,t}^k, \tau_1\boldsymbol{\tau}_2\Big\rangle \\ &= \sum_{l+m\geq N+1} \langle X_{u,s}^{|\tau_1|-l} \star X_{s,t}^l, \tau_1\rangle\langle X_{u,s}^{|\boldsymbol{\tau}_2|-m} \star X_{s,t}^m, \boldsymbol{\tau}_2\rangle \\ &= \Big(\sum_{l\geq N+1} + \sum_{\substack{l\leq N\\ m\geq N+1}} + \sum_{\substack{l,m\leq N\\ N+1-l\leq m}}\Big)\langle X_{u,s}^{|\tau_1|-l} \star X_{s,t}^l, \tau_1\rangle\langle X_{u,s}^{|\boldsymbol{\tau}_2|-m} \star X_{s,t}^m, \boldsymbol{\tau}_2\rangle. \end{aligned} \tag{5.14}$$

We will denote the three terms in this decomposition (5.14) as K_1, K_2 and K_3 respectively. Using the assumed estimate (5.6) for trees and factorial decay estimate (5.13),

$$K_1 := \Big|\sum_{l\geq N+1} \langle X_{u,s}^{|\tau_1|-l}\star X_{s,t}^{l}, \tau_1\rangle\langle X_{u,s}^{|\boldsymbol{\tau}_2|-m}\star X_{s,t}^{m}, \boldsymbol{\tau}_2\rangle\Big|. \tag{5.15}$$

$$\leq \frac{C_N^{c(\boldsymbol{\tau}_2)}}{\beta^{c(\boldsymbol{\tau})}}\Big[\frac{|\tau_1|!}{\tau_1!\,(|\tau_1|-N-1)!}R_u^{N+1,|\tau_1|}(s,t)\frac{(t-u)^{|\boldsymbol{\tau}_2|}}{\boldsymbol{\tau_2}!}\Big]^{\gamma}. \tag{5.16}$$

Using the binomial lemma for overlapping intervals, Lemma 23,

$$K_1 \leq \frac{C_N^{c(\boldsymbol{\tau}_2)}}{\beta^{c(\boldsymbol{\tau})}}\Big[\frac{|\boldsymbol{\tau}|!}{\boldsymbol{\tau}!\,(|\boldsymbol{\tau}|-N-1)!}R_u^{N+1,|\boldsymbol{\tau}|}(s,t)\Big]^{\gamma}. \tag{5.17}$$

We now estimate the second term in the decomposition (5.14). Using that multiplication is bounded in tree norm (see Lemma 22, applicable as $\beta \geq \exp\big(\sum_{i=1}^{N} N^i\big)$) and the factorial decay estimate (5.13),

$$\Big|\sum_{l\leq N}\langle X_{u,s}^{|\tau_1|-l}\star X_{s,t}^{l}, \tau_1\rangle\Big| \leq \frac{\tilde{c}_N C_N^2}{\beta\tau_1!^{\gamma}}\sum_{l\leq N}\binom{|\tau_1|}{l}^{\gamma}(s-u)^{\gamma(|\tau_1|-l)}(t-s)^{\gamma l}$$

$$\leq \frac{\tilde{c}_N C_N^2 (N+1)^{1-\gamma}}{\beta\tau_1!^{\gamma}}\Big(\sum_{l\leq N}\binom{|\tau_1|}{l}(s-u)^{(|\tau_1|-l)}(t-s)^{l}\Big)^{\gamma} \tag{5.18}$$

$$\leq \frac{\tilde{c}_N C_N^2 (N+1)^{1-\gamma}}{\beta\tau_1!^{\gamma}}(t-u)^{\gamma|\tau_1|}. \tag{5.19}$$

Applying the induction hypothesis and (5.19),

$$K_2 := \Big|\sum_{l\leq N, m\geq N+1}\langle X_{u,s}^{|\tau_1|-l}\star X_{s,t}^{l}, \tau_1\rangle\langle X_{u,s}^{|\boldsymbol{\tau}_2|-m}\star X_{s,t}^{m}, \boldsymbol{\tau}_2\rangle\Big|.$$

$$\leq \frac{\tilde{c}_N C_N^2 (N+1)^{1-\gamma}\hat{c}_N^{c(\boldsymbol{\tau}_2)}}{\beta^{c(\boldsymbol{\tau})}}\Big[\frac{|\boldsymbol{\tau}_2|!\,(t-u)^{|\tau_1|}}{\boldsymbol{\tau}_2!\,(|\boldsymbol{\tau}_2|-N-1)!\tau_1!}R_u^{N+1,|\tau_2|}(s,t)\Big]^{\gamma}.$$

By the binomial lemma for overlapping intervals (5.11)

$$K_2 \leq \frac{\tilde{c}_N C_N^2 (N+1)^{1-\gamma}\hat{c}_N^{c(\boldsymbol{\tau}_2)}}{\beta^{c(\boldsymbol{\tau})}}\Big[\frac{|\boldsymbol{\tau}|!}{\boldsymbol{\tau}!\,(|\boldsymbol{\tau}|-N-1)!}R_u^{N+1,|\boldsymbol{\tau}|}(s,t)\Big]^{\gamma}. \tag{5.20}$$

Finally, we estimate the third term in the decomposition (5.14) at the beginning of this proof. By factorial decay estimate (5.13) and applying Lemma 22, which asserts that the multiplication is bounded in forest norm for factorial decay estimates,

$$K_3 := \Big|\sum_{\substack{l,m\leq N\\ N+1-l\leq m}}\langle X_{u,s}^{|\tau_1|-l}\star X_{s,t}^{l}, \tau_1\rangle\langle X_{u,s}^{|\boldsymbol{\tau}_2|-m}\star X_{s,t}^{m}, \boldsymbol{\tau}_2\rangle\Big|$$

$$\leq \sum_{\substack{l,m\leq N\\ N+1-l\leq m}}\Big(\frac{c_{1,N}}{\beta}\Big)^{c(\boldsymbol{\tau})}\Big(\binom{|\tau_1|}{l}\binom{|\boldsymbol{\tau}_2|}{m}\frac{(s-u)^{|\boldsymbol{\tau}|-l-m}(t-s)^{l+m}}{\tau_1!\boldsymbol{\tau}_2!}\Big)^{\gamma}$$

where $c_{1,N} = \tilde{c}_N C_N^2$. By the conventional concavity estimate for sum $\sum_{i=1}^M a_i^\gamma \le M^{1-\gamma}(\sum_{i=1}^m a_i)^\gamma$,

$$K_3 \le \left(\frac{c_{2,N}}{\beta}\right)^{c(\boldsymbol{\tau})} \Big(\sum_{\substack{l,m\le N\\ N+1-l\le m}} \binom{|\tau_1|}{l}\binom{|\boldsymbol{\tau}_2|}{m} \frac{(s-u)^{|\boldsymbol{\tau}|-l-m}(t-s)^{l+m}}{\tau_1!\boldsymbol{\tau}_2!}\Big)^\gamma,$$

where $c_{2,N} = (N+1)^{2(1-\gamma)}\tilde{c}_N C_N^2$. Using the binomial identity $\sum_{l+m=k}\binom{M_1}{l}\binom{M_2}{m} = \sum\binom{M_1+M_2}{k}$,

$$K_3 \le \left(\frac{c_{2,N}}{\beta}\right)^{c(\boldsymbol{\tau})} \Big(\sum_{N+1\le k} \frac{(s-u)^{|\boldsymbol{\tau}|-k}(t-s)^k}{\tau_1!\boldsymbol{\tau}_2!}\binom{|\boldsymbol{\tau}|}{k}\Big)^\gamma.$$

As our estimate dominates the tail of the binomial sum (see Lemma 18),

$$K_3 \le \left(\frac{c_{2,N}}{\beta}\right)^{c(\boldsymbol{\tau})} \left[\frac{|\boldsymbol{\tau}|!}{\boldsymbol{\tau}!\,(|\boldsymbol{\tau}|-N-1)!} R_u^{N+1,|\boldsymbol{\tau}|}(s,t)\right]^\gamma. \tag{5.21}$$

Therefore, substituting the estimates for K_1 (5.17), K_2 (5.20) and K_3 (5.21) into the decomposition (5.14), we have

$$\begin{aligned} &\big|\langle \sum_{k\ge N+1} X_{u,s}^{|\boldsymbol{\tau}|-k} \star X_{s,t}^k, \boldsymbol{\tau}\rangle\big| \\ \le\ & \beta^{-c(\boldsymbol{\tau})}\big(C_N^{c(\boldsymbol{\tau})} + \big((N+1)^{2(1-\gamma)}\tilde{c}_N C_N^2\big)^{c(\boldsymbol{\tau})} + \tilde{c}_N C_N^2 (N+1)^{1-\gamma}\hat{c}_N^{c(\boldsymbol{\tau})-1}\big) \\ & \times\big(\frac{|\boldsymbol{\tau}|!}{\boldsymbol{\tau}!\,(|\boldsymbol{\tau}|-N-1)!} R_u^{N+1,|\boldsymbol{\tau}|}(s,t)\big)^\gamma \end{aligned}$$

and the Lemma follows by $3\tilde{c}_N C_N^2 (N+1)^{2(1-\gamma)} \le \hat{c}_N$. □

6. The proof

Let X be a $\gamma-$branched rough path and let $N = \lfloor \gamma^{-1}\rfloor$. We will use the following identity that is implicit in Gubinelli's construction of iterated integrals of branched rough path (see Theorem 7.3 in [4])

$$X_{s,t}^{n+1} = \lim_{|\mathcal{P}|\to 0} \sum_{i=0}^{m-1}\sum_{k=1}^{N} X_{s,t_i}^{n+1-k} \star X_{t_i,t_{i+1}}^k, \tag{6.1}$$

for $n \ge N+1$, where the limit is taken as the mesh size $\max_{t_i\in\mathcal{P}}|t_{i+1}-t_i|$ of the partition

$$\mathcal{P} = (0 = t_0 < \ldots < t_m = 1)$$

goes to zero. Alternatively, one can check directly that the limit on the right hand side of (6.1) converges, has the multiplicative property and is γ-Hölder, which by Theorem 7.3 in [4] would imply (6.1). We will estimate the double sum on the right hand side in (6.1) by dropping points successively from the partition $\mathcal{P}$. The following lemma carries out the algebra of removing partition points from a Riemann sum.

Lemma 24. *Let X be a $\gamma-$branched rough path and let $N = \lfloor\gamma^{-1}\rfloor$. For each partition $\mathcal{P}$ of the interval $[s,t]$, define $X^{\mathcal{P},n} : [0,1]\times[0,1] \to \mathcal{H}^*$ such that*

$$X^{\mathcal{P},n}_{s,t} = \sum_{t_i\in\mathcal{P}} \sum_{1\le k\le N} X^{n-k}_{s,t_i} \star X^{k}_{t_i,t_{i+1}}$$

Then for any $t_j \in \mathcal{P}$,

$$\sum_{k\ge N+1} X^{n-k}_{u,s} \star \left(X^{\mathcal{P},k}_{s,t} - X^{\mathcal{P}\setminus\{t_j\},k}_{s,t}\right) \tag{6.2}$$

$$= \sum_{\substack{k_2+k_3\ge N+1\\ 1\le k_3\le N}} X^{n-k_2-k_3}_{u,t_{j-1}} \star X^{k_2}_{t_{j-1},t_j} \star X^{k_3}_{t_j,t_{j+1}}. \tag{6.3}$$

Proof. Note first that

$$X^{\mathcal{P},k}_{s,t} - X^{\mathcal{P}\setminus\{t_j\},k}_{s,t} = \sum_{1\le l\le N} X^{k-l}_{s,t_{j-1}} \star X^{l}_{t_{j-1},t_j} + X^{k-l}_{s,t_j} \star X^{l}_{t_j,t_{j+1}} - X^{k-l}_{s,t_{j-1}} \star X^{l}_{t_{j-1},t_{j+1}}.$$

By applying the multiplicativity of X to the third term,

$$X^{\mathcal{P},k}_{s,t} - X^{\mathcal{P}\setminus\{t_j\},k}_{s,t} = \sum_{\substack{1\le l_3\le N\\ l_2\ge N+1-l_3}} X^{k-l_2-l_3}_{s,t_{j-1}} \star X^{l_2}_{t_{j-1},t_j} \star X^{l_3}_{t_j,t_{j+1}}. \tag{6.4}$$

From this, we observe that $X^{\mathcal{P},k}_{s,t} - X^{\mathcal{P}\setminus\{t_j\},k}_{s,t}$ is nonzero only when $k \ge N+1$. Therefore,

$$\sum_{k\ge N+1} X^{n-k}_{u,s} \star \left(X^{\mathcal{P},k}_{s,t} - X^{\mathcal{P}\setminus\{t_j\},k}_{s,t}\right) = \sum_{k} X^{n-k}_{u,s} \star \left(X^{\mathcal{P},k}_{s,t} - X^{\mathcal{P}\setminus\{t_j\},k}_{s,t}\right). \tag{6.5}$$

By substituting (6.4) into (6.5) and applying the associativity of $\star$, we see that

$$\sum_{k\ge N+1} X^{n-k}_{u,s} \star \left(X^{\mathcal{P},k}_{s,t} - X^{\mathcal{P}\setminus\{t_j\},k}_{s,t}\right) = \sum_{\substack{1\le l_3\le N\\ l_2\ge N+1-l_3}} X^{n-l_2-l_3}_{u,t_{j-1}} \star X^{l_2}_{t_{j-1},t_j} X^{l_3}_{t_j,t_{j+1}}.$$

□

We now once again require some binomial-type lemmas which we will prove in the Appendix. The following says that our estimate is decreasing in some sense.

Lemma 25. *For all $0 \le k \le m \le n$ and $u \le s \le t$,*

$$\frac{1}{(n-m)!} R^{m,n}_u(s,t) \le \frac{\exp m}{(n-m+k)!} R^{m-k,n}_u(s,t). \tag{6.6}$$

The following is at the heart of the proof of our main result. It describes the product of our estimates over adjacent time intervals.

Lemma 26. *(Chen's identity for R function) For all $u \le v \le s \le t$ and $N+1 \le n$,*

$$\sum_{k=1}^{N} R^{N+1-k,n-k}_u(v,s)^\gamma \frac{(t-s)^{\gamma k}}{(k!)^\gamma} \le (N+1)^{1-\gamma} R^{N+1,n}_u(v,t)^\gamma.$$

The following result gives an estimate for the remainder of a coproduct sum of branched rough paths. It may look like we are proving more than the factorial decay result we need, but in fact such estimate provides exactly the necessary induction hypothesis to prove the factorial decay estimate.

Lemma 27. *Let $0<\gamma\leq 1$ and $N=\lfloor \gamma^{-1}\rfloor$. Let X be a $\gamma-$branched rough path. If for any $0\leq n\leq N$,*

$$\|X^n_{s,t}\|_{\mathcal{T},\gamma,\beta} \quad\leq\quad \frac{(t-s)^{n\gamma}}{n!^{\gamma}}, \tag{6.7}$$

and

$$\beta\geq 6\exp\left(7\sum_{i=1}^{N+1}(N+1)^i\right)\sum_{r=2}^{\infty}\left(\frac{2}{r-1}\wedge 1\right)^{(N+1)\gamma}|\mathcal{T}_N|^{1-\gamma} \tag{6.8}$$

then the following holds for all n,

$$\|\sum_{k\geq N+1}X^{n-k}_{u,s}\star X^k_{s,t}\|_{\mathcal{T},\gamma,\beta}\leq\left[\frac{1}{(n-N-1)!}R^{N+1,n}_u(s,t)\right]^{\gamma}1_{\{n\geq N+1\}}. \tag{6.9}$$

Proof. We shall prove this by induction on n. The base induction $n=N$ is trivial as both sides in equation (6.9) is zero. By the induction hypothesis and the compatibility of our estimate with tree multiplication (Lemma 19), for all $m\leq n$,

$$\|\sum_{l\geq N+1}X^{m-l}_{u,s}\star X^l_{s,t}\|_{\mathcal{F},\gamma,\hat{c}_N^{-1}\beta}\leq\left[\frac{1}{(m-N-1)!}R^{N+1,m}_u(s,t)\right]^{\gamma}1_{\{m\geq N+1\}},$$

and as our estimate is decreasing (see Lemma 25),

$$\|\sum_{l\geq N+1}X^{m-l}_{u,s}\star X^l_{s,t}\|_{\mathcal{F},\gamma,\hat{c}_N^{-1}\beta}\leq\left[\frac{\exp{(N+1)}}{(m-r)!}R^{r,m}_u(s,t)\right]^{\gamma}1_{\{m\geq N+1\}}.$$

for any $1\leq r\leq N$. As multiplication is bounded in forest norm (see Lemma 22) for factorial decay estimates, we have for all $1\leq r\leq N$,

$$\begin{aligned}K' &:= \|\sum_{r\leq l\leq N}X^{m-l}_{u,s}\star X^l_{s,t}\|_{\mathcal{F},\gamma,(\tilde{c}_NC_N)^{-1}\beta}\\ &\leq \sum_{r\leq l\leq N}\|X^{m-l}_{u,s}\|_{\mathcal{T},\gamma,(\tilde{c}_NC_N)^{-1}\beta}\|X^l_{s,t}\|_{\mathcal{T},\gamma,(\tilde{c}_NC_N)^{-1}\beta}.\end{aligned}$$

As our estimate is a factorial decay estimate,

$$K'\leq\sum_{r\leq l\leq N}\frac{(s-u)^{\gamma(m-l)}(t-s)^{\gamma l}}{(m-l)!^{\gamma}l!^{\gamma}}.$$

By the conventional concavity estimate for sum,

$$K'\leq(N+1)^{1-\gamma}\big(\sum_{r\leq l\leq N}\frac{(s-u)^{(m-l)}(t-s)^{l}}{(m-l)!l!}\big)^{\gamma}.$$

By the binomial Lemma which bounds the remainder of a binomial sum by our estimate (Lemma 18),

$$K'\leq(N+1)^{1-\gamma}\Big[\frac{1}{(m-r)!}R^{r,m}_u(s,t)\Big]^{\gamma}.$$

In particular, since $\hat{c}_N\geq\tilde{c}_NC_N$, for all $r\leq N$,

$$\|\sum_{r\leq l}X^{m-l}_{u,s}\star X^l_{s,t}\|_{\mathcal{F},\gamma,\hat{c}_N^{-1}\beta}\quad\leq\quad c_{5,N}\left[\frac{1}{(m-r)!}R^{r,m}_u(s,t)\right]^{\gamma}, \tag{6.10}$$

where $c_{5,N} = 2\exp(N+1)$. Note first that as multiplication $\star$ is bounded in tree norm (Lemma 11),

$$\begin{aligned} & \Big\| \sum_{\substack{1\le k\le N \\ l\ge N+1-k}} X^{n-l-k}_{u,t_{j-1}} \star X^{l}_{t_{j-1},t_j} \star X^{k}_{t_j,t_{j+1}} \Big\|_{\mathcal{T},\gamma,\hat{c}_N^{-1}\beta} \\ \le \quad & c_N \beta^{-1} \sum_{1\le k\le N} \Big\| \sum_{l\ge N+1-k} X^{n-l-k}_{u,t_{j-1}} \star X^{l}_{t_{j-1},t_j} \Big\|_{\mathcal{F},\gamma,\hat{c}_N^{-1}\beta} \| X^{k}_{t_j,t_{j+1}} \|_{\mathcal{T},\gamma,\hat{c}_N^{-1}\beta}. \end{aligned}$$

By our assumption that we have a factorial decay estimate for $X^1,\dots,X^N$ (see (6.7)) and (6.10),

$$\begin{aligned} \tilde{K} \quad := \quad & \Big\| \sum_{l\ge N+1-k} X^{n-l-k}_{u,t_{j-1}} \star X^{k}_{t_{j-1},t_j} \Big\|_{\mathcal{F},\gamma,\hat{c}_N^{-1}\beta} \| X^{k}_{t_j,t_{j+1}} \|_{\mathcal{T},\gamma,\hat{c}_N^{-1}\beta} \\ \le \quad & c_{5,N} \Big[\frac{1}{(n-N-1)!} R^{N+1-k,n-k}_u (t_{j-1},t_j) \frac{(t_{j+1}-t_j)^k}{k!} \Big]^{\gamma}. \end{aligned}$$

By Chen's identity for R function, Lemma 26,

$$\begin{aligned} & \sum_{1\le k\le N} \Big\| \sum_{l\ge N+1-k} X^{n-l-k}_{u,t_{j-1}} \star X^{l}_{t_{j-1},t_j} \Big\|_{\mathcal{F},\gamma,\hat{c}_N^{-1}\beta} \| X^{k}_{t_j,t_{j+1}} \|_{\mathcal{T},\gamma,\hat{c}_N^{-1}\beta} \\ \le \quad & \frac{c_{6,N}}{\beta} \Big[\frac{1}{(n-N-1)!} R^{N+1,n}_u (t_{j-1},t_{j+1}) \Big]^{\gamma}, \end{aligned} \tag{6.11}$$

where $c_{6,N} = c_{5,N}(N+1)^{1-\gamma}$. Note that by explicit computation, there is some constant $c_{N,n}$ independent of u, v_1, v_2 such that

$$R^{N+1,n}_u(v_1,v_2) = c_{N,n} \left[(v_2-u)^{\frac{n}{N+1}} - (v_1-u)^{\frac{n}{N+1}} \right]^{N+1}.$$

Since

$$\begin{aligned} & \sum_{j=1}^{r-1} (t_{j+1}-u)^{\frac{n}{N+1}} - (t_{j-1}-u)^{\frac{n}{N+1}} \\ \le \quad & 2\left((t-u)^{\frac{n}{N+1}} - (s-u)^{\frac{n}{N+1}} \right), \end{aligned}$$

there exists a j such that

$$\begin{aligned} & (t_{j+1}-u)^{\frac{n}{N+1}} - (t_{j-1}-u)^{\frac{n}{N+1}} \\ \le \quad & \frac{2}{r-1} \left((t-u)^{\frac{n}{N+1}} - (s-u)^{\frac{n}{N+1}} \right). \end{aligned} \tag{6.12}$$

As (6.12) would still hold if we replace $\frac{2}{r-1}$ by 1, we have

$$\begin{aligned} & (t_{j+1}-u)^{\frac{n}{N+1}} - (t_{j-1}-u)^{\frac{n}{N+1}} \\ \le \quad & \left(\frac{2}{r-1} \wedge 1 \right) \left((t-u)^{\frac{n}{N+1}} - (s-u)^{\frac{n}{N+1}} \right) \end{aligned}$$

Using this particular j in the expression (6.11) as well as the algebraic Lemma 24, we have

$$\begin{aligned} & \| \sum_{k\geq N+1} X_{u,s}^{n-k} \star \left(X_{s,t}^{\mathcal{P},k} - X_{s,t}^{\mathcal{P}\setminus(t_j),k}\right) \|_{\mathcal{T},\gamma,\hat{c}_N^{-1}\beta} \\ \leq \; & \frac{c_{6,N}}{\beta} \left(\frac{2}{r-1} \wedge 1\right)^{(N+1)\gamma} \frac{1}{(n-N-1)!^{\gamma}} R_u^{N+1,n}(s,t)^{\gamma}. \end{aligned}$$

By iteratively removing points and observing that

$$X_{s,t}^{\{s,t\},k} = 0$$

for $k \geq N+1$, we have that for all partitions $\mathcal{P}$,

$$\begin{aligned} & \| \sum_{k\geq N+1} X_{u,s}^{n-k} \star X_{s,t}^{\mathcal{P},k} \|_{\mathcal{T},\gamma,\hat{c}_N^{-1}\beta} \\ \leq \; & \frac{c_{6,N}}{\beta} \sum_{r=2}^{\infty} \left(\frac{2}{r-1} \wedge 1\right)^{(N+1)\gamma} \frac{1}{(n-N-1)!^{\gamma}} R_u^{N+1,n}(s,t)^{\gamma}. \end{aligned}$$

In particular,

$$\begin{aligned} & \| \sum_{k\geq N+1} X_{u,s}^{n-k} \star X_{s,t}^{\mathcal{P},k} \|_{\mathcal{T},\gamma,\beta} \\ = \; & \hat{c}_N \| \sum_{k\geq N+1} X_{u,s}^{n-k} \star X_{s,t}^{\mathcal{P},k} \|_{\mathcal{T},\gamma,\hat{c}_N^{-1}\beta} \\ \leq \; & \frac{c_{6,N}\hat{c}_N}{\beta} \sum_{r=2}^{\infty} \left(\frac{2}{r-1} \wedge 1\right)^{(N+1)\gamma} \frac{1}{(n-N-1)!^{\gamma}} R_u^{N+1,n}(s,t)^{\gamma}. \end{aligned}$$

We have the desired estimate if we let $|\mathcal{P}| \to 0$ and choose

$$\beta \geq c_{6,N}\hat{c}_N \sum_{r=2}^{\infty} \left(\frac{2}{r-1} \wedge 1\right)^{(N+1)\gamma}.$$

□

Proof of main result Theorem 4. Let $\hat{C}_N$ denote the right hand side of (6.8). For $X \in \mathcal{H}^*$, let $\|\cdot\|$ denote the following normalised Hölder norm of X for degrees up to N,

$$\|X\| = \max_{1\leq|\tau|\leq N, \tau \text{ trees}} \|\langle X, \tau\rangle\|_{\gamma,\tau}^{(\gamma|\tau|)^{-1}}.$$

where Hölder norm $\|\cdot\|_{\gamma,\tau}$ of each degree is define in (1.5) in the definition of branched rough path. Applying Lemma 27, which we have just proved, to the branched rough path Y defined for each rooted tree τ by

$$\langle Y_{s,t}, \tau\rangle = \frac{1}{(N!\,\|X\|)^{\gamma|\tau|}\,\hat{C}_N^{|\tau|}} \langle X_{s,t}, \tau\rangle$$

with $\beta = \hat{C}_N$, we have by taking $u = s$ that for $|\tau| \geq N+1$,

$$\begin{aligned} |\langle X_{u,t}, \tau\rangle| \; & \leq \; \frac{1}{\hat{C}_N} (N!\,\|X\|)^{\gamma|\tau|}\,\hat{C}_N^{|\tau|} \left[\frac{|\tau|!}{\tau!\,(|\tau|-N-1)!} R_u^{N+1,|\tau|}(u,t)^{\gamma}\right]^{\gamma} \\ & \leq \; \frac{2\exp(N+1)}{\hat{C}_N} (N!\,\|X\|)^{\gamma|\tau|}\,\hat{C}_N^{|\tau|} \frac{(t-u)^{\gamma|\tau|}}{\tau!^{\gamma}}. \end{aligned}$$

□

7. Appendix: Binomial-type lemmas

Lemma. *Let $N \in \mathbb{N} \cup \{0\}$ and $n \geq N+1$. For all $u < s < t$*

$$\sum_{j=N+1}^{n} \frac{(s-u)^{n-j}\,(t-s)^{j}}{(n-j)!j!} \quad \leq \quad \frac{1}{(n-N-1)!} R_u^{N+1,n}(s,t).$$

Proof. The following identity can be proved using an induction on N or Taylor's theorem,

$$J \;:=\; \sum_{j=N+1}^{n} \frac{(s-u)^{n-j}\,(t-s)^{j}}{(n-j)!j!} \tag{7.1}$$

$$= \frac{1}{(n-N-1)!} \int_{\triangle_{N+1}(s,t)} (s_1-u)^{n-N-1}\, \mathrm{d}s_1 \ldots \mathrm{d}s_{N+1}. \tag{7.2}$$

Note that as we are integrating over the domain $s_1 < s_2 \ldots < s_{N+1}$,

$$\begin{aligned} J &\leq \frac{1}{(n-N-1)!} \int_{\triangle_{N+1}(s,t)} \Pi_{i=1}^{N+1} (s_i-u)^{\frac{n-N-1}{N+1}}\, \mathrm{d}s_1 \ldots \mathrm{d}s_{N+1} \\ &= \frac{1}{(n-N-1)!} S^{(N+1)} \left(\frac{N+1}{n} (\cdot - u)^{\frac{n}{N+1}} \right)_{s,t} \\ &= \frac{1}{(n-N-1)!} R_u^{N+1,n}(s,t). \end{aligned}$$

□

Lemma. *(Binomial lemma for overlapping intervals) Let $N \in \mathbb{N} \cup \{0\}$. Let $n \geq N+1$ and $m \geq 0$, and $u \leq s \leq t$,*

$$\frac{n!}{(n-N-1)!} R_u^{N+1,n}(s,t)\,(t-u)^m \tag{7.3}$$

$$\leq \quad \frac{(n+m)!}{(n+m-N-1)!} R_u^{N+1,n+m}(s,t). \tag{7.4}$$

Proof. Using that for any $b \geq a$ and $c \geq d$, $c(b-a) \leq (bc-ad)$,

$$\begin{aligned} & R_u^{N+1,n}(s,t)\,(t-u)^m \\ = \;& \frac{\left[\,(t-u)^{\frac{n}{N+1}} - (s-u)^{\frac{n}{N+1}}\,\right]^{N+1} (t-u)^m}{(N+1)!} \\ \leq \;& \frac{\left[\,(t-u)^{\frac{n+m}{N+1}} - (s-u)^{\frac{n+m}{N+1}}\,\right]^{N+1}}{(N+1)!} \\ = \;& R_u^{N+1,n+m}(s,t). \end{aligned} \tag{7.5}$$

The lemma now follows by noting that as $n+m \geq n$,

$$\frac{n!}{(n-N-1)!} \left(\frac{N+1}{n}\right)^{N+1} \leq \frac{(n+m)!}{(n+m-N-1)!} \left(\frac{N+1}{n+m}\right)^{N+1}. \tag{7.6}$$

□

Lemma. *(The estimate is decreasing) For all $0 \le k \le m \le n$ and $u \le s \le t$,*

$$\frac{1}{(n-m)!} R_u^{m,n}(s,t) \le \frac{\exp m}{(n-m+k)!} R_u^{m-k,n}(s,t). \tag{7.7}$$

Proof. Since for any $p \ge 1$ and $a \ge b$, $(a-b)^p \le a^p - b^p$,

$$\begin{aligned}
J' &:= \frac{1}{(n-m)!} R_u^{m,n}(s,t) \\
&= \frac{1}{(n-m)!} \left(\frac{m}{n}\right)^m \frac{\left((t-u)^{\frac{n}{m}} - (s-u)^{\frac{n}{m}}\right)^m}{m!} \\
&\le \frac{1}{(n-m)!} \left(\frac{m}{n}\right)^m \frac{\left((t-u)^{\frac{n}{m-k}} - (s-u)^{\frac{n}{m-k}}\right)^{m-k}}{m!}
\end{aligned}$$

As $\frac{m^m}{m!} \le \exp(m)$, $n^m (n-m)! \ge n^{m-k}(n-m+k)!$ and $\frac{(m-k)^{m-k}}{(m-k)!} \ge 1$,

$$\begin{aligned}
J' &\le \frac{\exp(m)}{(n-m+k)!} \left(\frac{m-k}{n}\right)^{m-k} \frac{\left((t-u)^{\frac{n}{m-k}} - (s-u)^{\frac{n}{m-k}}\right)^{m-k}}{(m-k)!} \\
&= \frac{\exp(m)}{(n-m+k)!} R_u^{m-k,n}(s,t).
\end{aligned}$$

□

Lemma 28. *(Binomial lemma for adjacent intervals) Let $0 \le k \le m \le n$ and $u \le s \le t \le v$, then*

$$R_u^{m-k,n}(s,t) \frac{(v-t)^k}{k!} \le S^{(m-k)}\left(\rho_u^{\frac{n+k}{m}}\right)_{s,t} S^{(k)}\left(\rho_u^{\frac{n+k}{m}}\right)_{t,v}. \tag{7.8}$$

Proof. Recall that

$$R_u^{m-k,n}(s,t) = S^{(m-k)}\left(\frac{m-k}{n}(\cdot - u)^{\frac{n}{m-k}}\right)_{s,t}.$$

In the third line below, we used that $s_j > s_i$ for $m-k+1 \le j \le m$ and $1 \le i \le m-k$,

$$\begin{aligned}
& S^{(m-k)}\left(\frac{m-k}{n}(\cdot - u)^{\frac{n}{m-k}}\right)_{s,t} \frac{(v-t)^k}{k!} \\
=\ & \int_{s<s_1<\ldots<s_{m-k}<t} \Pi_{i=1}^{m-k} (s_i - u)^{\frac{n+k-m}{m-k}} \, \mathrm{d}s_1 \ldots \mathrm{d}s_{m-k} \\
& \times \int_{t<s_{m-k+1}<\ldots<s_m<v} \mathrm{d}s_{m-k+1} \ldots \mathrm{d}s_m \\
\le\ & \int_{s<s_1<\ldots<s_{m-k}<t} \Pi_{i=1}^{m-k} (s_i - u)^{\frac{n+k-m}{m}} \, \mathrm{d}s_1 \ldots \mathrm{d}s_{m-k} \\
& \times \int_{t<s_{m-k+1}<\ldots<s_m<v} \Pi_{i=m-k+1}^{m} (s_i - u)^{\frac{n+k-m}{m}} \, \mathrm{d}s_{m-k+1} \ldots \mathrm{d}s_m \\
=\ & S^{(m-k)}\left(\frac{m}{n+k}(\cdot - u)^{\frac{n+k}{m}}\right)_{s,t} S^{(k)}\left(\frac{m}{n+k}(\cdot - u)^{\frac{n+k}{m}}\right)_{t,v}.
\end{aligned}$$

□

Corollary 29. *(Chen's identity for R function) For all $u \le v \le s \le t$ and $N+1 \le n$,*

$$\sum_{k=1}^{N} R_u^{N+1-k,n-k}(v,s)^{\gamma}\frac{(t-s)^{\gamma k}}{(k!)^{\gamma}} \le (N+1)^{1-\gamma} R_u^{N+1,n}(v,t)^{\gamma}.$$

Proof. By the binomial Lemma for adjacent intervals, Lemma 28 ,

$$\begin{aligned}&R_u^{N+1-k,n-k}(v,s)^{\gamma}\frac{(t-s)^{\gamma k}}{(k!)^{\gamma}}\\ \le\quad &\left[S^{(N+1-k)}\left(\rho_u^{\frac{n}{N+1}}\right)_{v,s} S^{(k)}\left(\rho_u^{\frac{n}{N+1}}\right)_{s,t}\right]^{\gamma}.\end{aligned}$$

From here we use the classical concavity estimate for sums and Chen's identity (see for example Theorem 2.1.2 in [7]) to obtain

$$\begin{aligned}&\sum_{1\le k\le N} R_u^{N+1-k,n-k}(v,s)^{\gamma}\frac{(t-s)^{\gamma k}}{(k!)^{\gamma}}\\ \le\quad &(N+1)^{1-\gamma}\Big[\sum_{k=1}^{N} S^{(N+1-k)}\big(\rho_u^{\frac{n}{N+1}}\big)_{v,s} S^{(k)}\big(\rho_u^{\frac{n}{N+1}}\big)_{s,t}\Big]^{\gamma}\\ \le\quad &S^{(N+1)}\left(\rho_u^{\frac{n}{N+1}}\right)_{v,t}^{\gamma}. \qquad (7.9)\end{aligned}$$

□

References

[1] A. Connes and D. Kreimer, Hopf algebras, renormalization and noncommutative geometry, *Comm. Math. Phys.* 199(1) (1998) 203–242.

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

[3] M. Hairer, D. Kelly, Geometric versus non-geometric rough paths, *Ann. Inst. Henri Poincaré Probab. Stat.* 51(1) (2015) 207–251.

[4] M. Gubinelli: Ramification of rough paths, *J. Diff. Eq.* 248 (2010) 693–721.

[5] M. Gubinelli, Rooted trees for 3D Navier–Stokes equation, *Dyn. Partial Differ. Equ.* 3 (2) (2006) 161–172.

[6] T. Lyons, Differential equations driven by rough signals (I): an extension of an inequality of L. C. Young, *Mathematical Research Letters* 1 (1994) 451-464.

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

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

[9] L. C. Young, An inequality of Hölder type connected with Stieltjes integration, *Acta Math.*, 67 (1936) 251–282.

Department of Mathematics and Statistics, Reading University, RG6 6AX, UK.