Electron. Commun. Probab. **27** (2022), article no. 46, 1–12.
https://doi.org/10.1214/22-ECP473
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

A remainder estimate for branched rough differential equations*

Danyu Yang†

Abstract

Based on two isomorphisms of Hopf algebras, we provide a bound in the optimal order on the remainder of the truncated Taylor expansion for controlled differential equations driven by branched rough paths.

Keywords: rough paths theory; branched rough differential equations; Taylor expansion.
MSC2020 subject classifications: Primary 60L20, Secondary 60L70.
Submitted to ECP on November 26, 2021, final version accepted on June 5, 2022.

1 Introduction

In the seminal paper [17], Lyons builds the theory of rough paths. The theory solves rough differential equations (RDEs) of the form

$$dy_t = f(y_t)\, dx_t, y_0 = \xi,$$

where x can be highly oscillating. Under a Lipschitz condition on the vector field, Lyons proves the unique solvability of the differential equation, and the solution obtained is continuous with respect to the driving signal in rough paths metric. The theory has an embedded component in stochastic analysis, and x can be Brownian motion, continuous semi-martingales, Markov processes, Gaussian processes [11] etc.

In 1972, Butcher [4] identifies a group structure in a class of integration methods including Runge-Kutta methods and Picard iterations, where each method can be represented by a family of real-valued functions indexed by rooted trees. In [14], Grossman and Larson describe several Hopf algebras associated with families of trees. One Hopf algebra of simple rooted trees, with product [14, (3.1)] and coproduct [14, p.199], is particularly relevant to our setting, which we refer to as the Grossman Larson Hopf algebra, denoted as $\mathcal{H}$. In [6], Connes and Kreimer describe a Hopf algebra based on rooted trees [6, Section 2] to disentangle the intricate combinatorics of divergences in quantum field theory. We call this Hopf algebra the Connes Kreimer Hopf algebra, denoted as $\mathcal{H}_R$. The group identified by Butcher is the group of characters of $\mathcal{H}_R$ [7]. Based on [9, 16, 19], $\mathcal{H}$ is isomorphic to the graded dual of $\mathcal{H}_R$.

Rough differential equations are originally driven by geometric rough paths over Banach spaces [17]. Geometric rough paths satisfy an abstract integration by parts

*The author gratefully acknowledge the support of the National Natural Science Foundation of China (Young Scholar Grant 12201081), Chongqing University (Starting Grant 02080011044104) and EPSRC (Programme Grant EP/S026347/1).

†College of Mathematics and Statistics, Chongqing University, China. E-mail: danyuyang@cqu.edu.cn

formula, and take values in a nilpotent Lie group. The nilpotent Lie group can be expressed as a truncated group of characters of the shuffle Hopf algebra [20]. In [12], Gubinelli defines branched rough paths. Branched rough paths take values in a truncated group of characters of a labeled Connes Kreimer Hopf algebra. Both geometric and branched rough paths are of finite p-variation in rough paths metric, and encode information needed to construct solutions to differential equations. There exists a Hopf algebra homomorphism from the Connes Kreimer Hopf algebra onto the shuffle Hopf algebra, which induces an embedding of geometric rough paths into branched rough paths. On the other hand, the Grossman Larson algebra is freely generated by a collection of trees [5, 10]. Based on the freeness of Grossman Larson algebra, Boedihardjo and Chevyrev construct an isomorphism between branched rough paths and a class of geometric rough paths [2]. As a result, a branched RDE can be expressed as a geometric RDE driven by a Π-rough path defined by Gyurkó [13].

Based on the isomorphism between $\mathcal{H}$ and the graded dual of $\mathcal{H}_R$ [9, 16, 19], we clarify a relationship between rough paths taking values in the truncated group of characters of $\mathcal{H}_R$ and rough paths taking values in the truncated group of grouplike elements in $\mathcal{H}$ (Proposition 2.3). Based on this relationship and the freeness of the Grossman Larson algebra, sub-Riemannian geometry [11, Section 7.5] and the neo-classical inequality [15, 17], which are typical geometric rough paths tools, can be applied to branched rough paths. As an application, we provide an estimate for the remainder of the truncated Taylor expansion for the solution of a controlled differential equation driven by a branched rough path (Theorem 2.5). The remainder estimate is in the optimal order (Remark 2.7), which is pleasantly surprising noting the rapid increase of the dimension of simple rooted trees.

2 Notations and results

A rooted tree is a finite connected graph that has no cycle with a distinguished vertex called root. We call a rooted tree a tree. We assume trees are non-planar, which means that the children trees of each vertex are commutative. A forest is a commutative monomial of trees. The degree $|\rho|$ of a forest ρ is given by the number of vertices. For a given label set, a labeled forest is a forest for which each vertex is attached with a label.

Denote the label set $\mathcal{L} := \{1, 2, \ldots, d\}$. Let $F_{\mathcal{L}}$ ($T_{\mathcal{L}}$) denote the set of $\mathcal{L}$-labeled forests (trees) of degree greater or equal to 1. Let $F_{\mathcal{L}}^N$ ($T_{\mathcal{L}}^N$) denote the set of elements in $F_{\mathcal{L}}$ ($T_{\mathcal{L}}$) of degree $1, \ldots, N$.

Let $G_{\mathcal{L}}^N$ denote the set of degree-N characters of the $\mathcal{L}$-labeled Connes Kreimer Hopf algebra [6, p.214]. a is an element of $G_{\mathcal{L}}^N$, if a is an $\mathbb{R}$-linear map $\mathbb{R}F_{\mathcal{L}}^N \to \mathbb{R}$ that satisfies

$$(a, \rho_1)\,(a, \rho_2) = (a, \rho_1\rho_2)$$

for every $\rho_1, \rho_2 \in F_{\mathcal{L}}^N, |\rho_1| + |\rho_2| \leq N$, where $\rho_1\rho_2$ denotes the multiplication of commutative monomials of trees. Let $\triangle$ denote the coproduct of the Connes Kreimer Hopf algebra based on admissible cuts [6, p.215]. Then $G_{\mathcal{L}}^N$ is a group with the multiplication given by

$$(ab, \rho) := (a \otimes b, \triangle\rho)$$

for every $\rho \in F_{\mathcal{L}}^N$. $G_{\mathcal{L}}^N$ is a labeled truncated Butcher group [4]. We equip $G_{\mathcal{L}}^N$ with the norm

$$\|a\| := \max_{\rho \in F_{\mathcal{L}}^N} |(a, \rho)|^{\frac{1}{|\rho|}} . \tag{2.1}$$

With $\mathcal{L} = \{1, 2, \ldots, d\}$, let $\mathcal{H}_{\mathcal{L}}$ denote the $\mathcal{L}$-labeled Grossman Larson Hopf algebra with product [14, (3.1)] and coproduct [14, p.199]. Denote the product and coproduct of $\mathcal{H}_{\mathcal{L}}$ as $*$ and δ respectively. We consider $\mathcal{H}_{\mathcal{L}}$ as a Hopf algebra of labeled forests (by deleting the additional root in [14]). An element $a \in \mathcal{H}_{\mathcal{L}}$ is *grouplike* if $\delta a = a \otimes a$. Let $\mathcal{G}_{\mathcal{L}}$ denote the group of grouplike elements in $\mathcal{H}_{\mathcal{L}}$. For integer $N \geq 1$, the set of series $b = \sum_{\rho \in F_{\mathcal{L}}, |\rho| > N} (b, \rho) \rho$ form an ideal of $\mathcal{H}_{\mathcal{L}}$. Let $\mathcal{H}_{\mathcal{L}}^N$ denote the quotient algebra. Denote $\mathcal{G}_{\mathcal{L}}^N := \mathcal{G}_{\mathcal{L}} \cap \mathcal{H}_{\mathcal{L}}^N$. $\mathcal{G}_{\mathcal{L}}^N$ is a group. We equip $\mathcal{G}_{\mathcal{L}}^N$ with a continuous homogeneous norm.

Let $\bullet_a$ denote the labeled tree of one vertex with a label $a \in \mathcal{L}$ on the vertex. Let $[\tau_1 \cdots \tau_k]_a$ denote the labeled tree with children trees $\tau_1, \ldots, \tau_k$ on the root and a label $a \in \mathcal{L}$ on the root. Define $\sigma : F_{\mathcal{L}} \to \mathbb{N}$ as the symmetry factor given inductively by $\sigma(\bullet_a) := 1$ and

$$\sigma\left(\tau_1^{n_1} \cdots \tau_k^{n_k}\right) = \sigma\left(\left[\tau_1^{n_1} \cdots \tau_k^{n_k}\right]_a\right) := n_1! \cdots n_k! \sigma(\tau_1)^{n_1} \cdots \sigma(\tau_k)^{n_k},$$

where $\tau_i \in T_{\mathcal{L}}$ are different labeled trees (with labels counted). σ is the order of the permutation group on vertices in a tree that keeps the tree unchanged.

Let $\triangle$ denote the coproduct of the Connes Kreimer Hopf algebra, and let $*$ denote the product of the Grossman Larson Hopf algebra. Based on [9, Theorem 43] and [16, Proposition 4.4], for $\rho \in F_{\mathcal{L}}$,

$$\triangle \rho = \sum_{\rho_i \in F_{\mathcal{L}}} \frac{\sigma(\rho)}{\sigma(\rho_1) \sigma(\rho_2)} (\rho_1 * \rho_2, \rho) \rho_1 \otimes \rho_2. \tag{2.2}$$

Definition 2.1. *Suppose G is a group with norm $\|\cdot\|$. Let $X : [0, T] \to (G, \|\cdot\|)$. Denote*

$$X_{s,t} := X_s^{-1} X_t.$$

For $p \geq 1$, define

$$\|X\|_{p-var,[0,T]} := \left(\sup_{0=t_0<\cdots<t_n=T, n\geq 1} \sum_{i=0}^{n-1} \left\|X_{t_i,t_{i+1}}\right\|^p \right)^{\frac{1}{p}}.$$

For $p \geq 1$, let $[p]$ denote the largest integer that is less or equal to p.

Definition 2.2. *For $p \geq 1$, X is a branched p-rough path if $X : [0, T] \to G_{\mathcal{L}}^{[p]}$ is continuous and of finite p-variation.*

Proposition 2.3. *For $p \geq 1$, suppose $X : [0, T] \to G_{\mathcal{L}}^{[p]}$ is a branched p-rough path. Define $\bar{X} : [0, T] \to \left(F_{\mathcal{L}}^{[p]} \to \mathbb{R}\right)$ as*

$$\left(\bar{X}_t, \rho\right) := \frac{(X_t, \rho)}{\sigma(\rho)}$$

for $t \in [0, T]$ and $\rho \in F_{\mathcal{L}}^{[p]}$. Then $\bar{X}$ takes values in $\mathcal{G}_{\mathcal{L}}^{[p]}$, is continuous and of finite p-variation, and

$$\left(\bar{X}_{s,t}, \rho\right) = \frac{(X_{s,t}, \rho)}{\sigma(\rho)} \tag{2.3}$$

for $0 \leq s \leq t \leq T$ and $\rho \in F_{\mathcal{L}}^{[p]}$. For integer $N \geq [p] + 1$, there exists a unique extension of X resp. $\bar{X}$ to a continuous path of finite p-variation taking values in $G_{\mathcal{L}}^N$ resp. $\mathcal{G}_{\mathcal{L}}^N$. Still denote their extension as X resp. $\bar{X}$. Then (2.3) holds for $0 \leq s \leq t \leq T$ and $\rho \in F_{\mathcal{L}}^N$.

Remark 2.4. Since the Grossman Larson algebra is free on a collection of trees [5, 10], $\bar{X}$ acts as a bridge between X and geometric rough paths. In particular, sub-Riemannian geometry technique [11, Section 7.5] and the neo-classical inequality [15, 17] can be applied to $\bar{X}$. Then results are transferred back to X based on (2.3).

Let $L\left(\mathbb{R}^d, \mathbb{R}^e\right)$ denote the set of linear mappings from $\mathbb{R}^d$ to $\mathbb{R}^e$. For $f = (f_1, \ldots, f_d) : \mathbb{R}^e \to L\left(\mathbb{R}^d, \mathbb{R}^e\right)$ that is sufficiently smooth, define $f : T_{\mathcal{L}} \to (\mathbb{R}^e \to \mathbb{R}^e)$ inductively as

$$f(\bullet_a) := f_a \text{ and } f\left([\tau_1 \cdots \tau_k]_a\right) := \left(d^k f_a\right)\left(f(\tau_1) \cdots f(\tau_k)\right) \tag{2.4}$$

for $\tau_i \in T_{\mathcal{L}}$ and $a \in \mathcal{L}$, where $d^k f_a$ denotes the kth Fréchet derivative of f_a.

Lipschitz functions and norms are defined as in [17, Definition 1.2.4, p.230]. For $\gamma > 1$, let $\lfloor \gamma \rfloor$ denote the largest integer that is strictly less than γ.

Theorem 2.5. *For $\gamma > p \geq 1$, suppose $X : [0,T] \to G_{\mathcal{L}}^{[p]}$ is a branched p-rough path over base space $\mathbb{R}^d$, and $f : \mathbb{R}^e \to L\left(\mathbb{R}^d, \mathbb{R}^e\right)$ is $\operatorname{Lip}(\gamma)$. Let y denote the unique solution of the branched rough differential equation*

$$dy_t = f(y_t)\, dX_t, y_0 = \xi \in \mathbb{R}^e.$$

Then with $N := \lfloor \gamma \rfloor$, there exist two positive constants $c^1_{p,d,\omega(0,T)}$ and $c^2_{p,d}$ such that,

$$\left\| y_t - y_s - \sum_{\tau \in T_{\mathcal{L}}^N} f(\tau)(y_s) \frac{(X_{s,t}, \tau)}{\sigma(\tau)} \right\| \leq c^1_{p,d,\omega(0,T)} N! \frac{\omega(s,t)^{\frac{N+1}{p}}}{\left(\frac{N+1}{p}\right)!} \tag{2.5}$$

where $\omega(s,t) := c^2_{p,d} \|f\|^p_{\operatorname{Lip}(\gamma)} \|X\|^p_{p-var,[s,t]}$.

The solution to branched RDEs is defined as in [12, Section 8.1]. The existing Taylor remainder estimates for the solution of branched RDEs only deal with the case $N = [p]$ [12, Theorem 8.8]. Theorem 2.5 considers the general case $N \geq [p]$, and the estimate (2.5) is in the optimal order (Remark 2.7).

Remark 2.6. When $p = 1$ and $\|f\|_{\operatorname{Lip}(\infty)} < \infty$, suppose $\|f\|_{\operatorname{Lip}(\infty)} \|X\|_{1-var,[s,t]} < \left(c^2_{1,d}\right)^{-1}$. Then the Taylor series converges [12, Theorem 5.1].

Remark 2.7. Suppose $x : [0,T] \to \mathbb{R}^d$ is continuous and of bounded variation, and $f : \mathbb{R}^e \to L\left(\mathbb{R}^d, \mathbb{R}^e\right)$ is sufficiently smooth. Consider the ODE

$$dy_t = f(y_t)\, dx_t, y_0 = \xi.$$

Based on the fundamental theorem of calculus, for $s \leq t$,

$$y_t = y_s + \sum_{k=1}^{N} f^{\circ k}(y_s) X^k_{s,t} + \int_{s<u_1<\cdots<u_{N+1}<t} \cdots \int f^{\circ(N+1)}(y_{u_1})\, dx_{u_1} \cdots dx_{u_{N+1}}$$

where $f^{\circ 1} := f, f^{\circ(k+1)} := df^{\circ k}(f)$ and $X^k_{s,t} := \int \cdots \int_{s<u_1<\cdots<u_k<t} dx_{u_1} \otimes \cdots \otimes dx_{u_k}$.

Suppose X is a geometric p-rough path. Let $X^{N+1}_{s,t}$ denote the $(N+1)$th level element of X on $[s,t]$. Based on Lyons' extension Theorem [17, Theorem 2.2.1, p.242], there exists a positive constant β_p such that

$$\left\| X^{N+1}_{s,t} \right\| \leq \frac{\omega(s,t)^{\frac{N+1}{p}}}{\beta_p \left(\frac{N+1}{p}\right)!}$$

for every $s \leq t$ and every N. On the other hand, consider $f(t) = e^{-t}$. Then $\|f\|_{Lip(n)} = 1$ on $t \geq 0$ for $n = 1, 2, \ldots$ and

$$\left| f^{\circ(N+1)}(0) \right| = N!$$

The estimate (2.5) states that the remainder can be bounded similarly to $f^{\circ(N+1)}(y_s) X^{N+1}_{s,t}$ that is in the optimal order even in the geometric case. The dimension of trees contributes a geometric increase factor[1] that is part of the control ω.

[1] Based on [8, Section I.5.2], the number of unlabeled simple rooted trees is EIS A000081, and $H_n \sim \lambda \beta^n n^{-3/2}$ with λ approximately 0.43992 and β approximately 2.95576.

The proof of Theorem 2.5 is based on a mathematical induction that is an inhomogeneous analogue of [3]. The main estimate (2.5) is obtained by exploring the sub-Riemannian geometry of the truncated group of grouplike elements in the Grossman Larson Hopf algebra. The sub-Riemannian geometry structure is similar to that of the nilpotent Lie group [11, Theorem 7.32]. The factor $((N+1)/p)!$ is obtained by the neo-classical inequality [15, 17]. The tree neo-classical inequality is known to be false [1, Section 3]. Since the Grossman Larson algebra is free on a collection of trees, the analysis can be transferred back to the Tensor algebra where the neo-classical inequality holds. Our estimates rely critically on the simple fact that the number of words generated by a finite set of letters grows geometrically (Lemma 3.7).

3 Proofs

Proof of Proposition 2.3. Since $(X_t, \rho)/\sigma(\rho) = (\bar{X}_t, \rho)$ for $\rho \in F_{\mathcal{L}}$, $|\rho| = 1, \ldots, [p]$, it can be proved inductively based on (2.2) that for $\rho \in F_{\mathcal{L}}$, $|\rho| = 1, \ldots, [p]$, and $s \le t$,

$$\left(\bar{X}_{s,t}, \rho\right) = \frac{(X_{s,t}, \rho)}{\sigma(\rho)}.$$

The existence and uniqueness of the extension of X and $\bar{X}$ can be proved similarly to [17, Theorem 2.2.1]. Based on (2.2), when $\rho \in F_{\mathcal{L}}$, $|\rho| = n, n \ge [p]+1$,

$$\begin{aligned}
& (X_{s,t}, \rho) \\
= \;& \sigma(\rho) \lim_{D \subset [s,t], |D| \to 0} \sum_{\rho_i \in F_{\mathcal{L}}, |\rho_i| < |\rho|} \frac{(X_{t_0,t_1}, \rho_1)}{\sigma(\rho_1)} \cdots \frac{\left(X_{t_{k-1},t_k}, \rho_k\right)}{\sigma(\rho_k)} (\rho_1 * \cdots * \rho_k, \rho) \\
= \;& \sigma(\rho) \lim_{D \subset [s,t], |D| \to 0} \sum_{\rho_i \in F_{\mathcal{L}}, |\rho_i| < |\rho|} \left(\bar{X}_{t_0,t_1}, \rho_1\right) \cdots \left(\bar{X}_{t_{k-1},t_k}, \rho_k\right) (\rho_1 * \cdots * \rho_k, \rho) \\
= \;& \sigma(\rho) \left(\bar{X}_{s,t}, \rho\right). \quad \square
\end{aligned}$$

Based on [10, Section 8] and [5], the Grossman Larson algebra is freely generated by a collection of unlabeled trees. Denote this collection of trees as $\mathcal{B}$. Let $\mathcal{B}_{\mathcal{L}}$ denote the $\mathcal{L}$-labeled version of $\mathcal{B}$ with $\mathcal{L} = \{1, 2, \ldots, d\}$.

Notation 3.1. *Let $\mathcal{B}_{\mathcal{L}}^{[p]} = \{\upsilon_1, \ldots, \upsilon_K\}$ denote the set of elements in $\mathcal{B}_{\mathcal{L}}$ with degree less or equal to $[p]$.*

Definition 3.2. *For $a \in \mathcal{G}_{\mathcal{L}}^{[p]}$, define*

$$\|a\|' := \inf_x \sum_{i=1}^{K} \|x^{\upsilon_i}\|_{1-var}^{\frac{1}{|\upsilon_i|}} \tag{3.1}$$

where the infimum is taken over all continuous bounded variation paths $x = (x^{\upsilon_1}, \ldots, x^{\upsilon_K}) : [0,1] \to \mathbb{R}^K$ that satisfy

$$(a, \upsilon_{i_1} * \cdots * \upsilon_{i_k}) = \int \cdots \int_{0 < u_1 < \cdots < u_k < 1} dx_{u_1}^{\upsilon_{i_1}} \cdots dx_{u_k}^{\upsilon_{i_k}}$$

for $\upsilon_{i_j} \in \mathcal{B}_{\mathcal{L}}^{[p]}$, $|\upsilon_{i_1}| + \cdots + |\upsilon_{i_k}| \le [p]$. The infimum in (3.1) can be obtained at a continuous bounded variation path x, which is called a geodesic *associated with $a \in \mathcal{G}_{\mathcal{L}}^{[p]}$.*

Remark 3.3. Such x exists based on Chow-Rashevskii Theorem [11, Theorem 7.28]. Based on Arzelà-Ascoli Theorem and lower semi-continuity of 1-variation, the infimum can be obtained at some x that is continuous and of bounded variation.

For $a : F_{\mathcal{L}}^{N} \to \mathbb{R}$ and $c > 0$, define $\delta_c a : F_{\mathcal{L}}^{N} \to \mathbb{R}$ as $(\delta_c a, \rho) := c^{|\rho|}(a, \rho)$. A norm $\|\cdot\|$ is homogeneous if $\|\delta_c a\| = c\|a\|$ for every $c > 0$ and every a. $\|\cdot\|'$ is a continuous homogeneous norm. The continuity of $\|\cdot\|'$ can be proved similarly as [11, Proposition 7.40(v)].

Proposition 3.4. *Continuous homogeneous norms on $\mathcal{G}_{\mathcal{L}}^{[p]}$ are equivalent up to a constant depending on p and d.*

Proof. The proof is similar to [11, Theorem 7.44]. □

Lemma 3.5. *Let $x = (x^{\upsilon_1}, \ldots, x^{\upsilon_K}) : [0,1] \to \mathbb{R}^K$ be a geodesic associated with $\bar{X}_{s,t}$. Then there exists $M_{p,d} > 0$ such that*

$$\|x^{\upsilon_i}\|_{1-var} \le (M_{p,d})^{|\upsilon_i|} \|X\|_{p-var,[s,t]}^{|\upsilon_i|}$$

for every $\upsilon_i \in \mathcal{B}_{\mathcal{L}}^{[p]}$.

Proof. Define a norm on $\mathcal{G}_{\mathcal{L}}^{[p]}$ as $\|a\|_1 := \max_{\rho \in F_{\mathcal{L}}^{[p]}} |(a,\rho)|^{\frac{1}{|\rho|}}$. Based on the definition of $\|\cdot\|'$, equivalency of continuous homogeneous norms as in Proposition 3.4 and that $(\bar{X}_{s,t}, \rho) = (X_{s,t}, \rho)/\sigma(\rho)$, the proposed inequality holds. □

Notation 3.6. *Let $\mathcal{W}$ denote the set of finite sequences $t_1 \cdots t_k$ of $t_i \in \mathcal{B}_{\mathcal{L}}^{[p]}$, including the empty sequence denoted as η. The degree $|w|$ of $w = t_1 \cdots t_k$ is $|t_1| + \cdots + |t_k|$. The degree of η is 0.*

Lemma 3.7. *Let T_n denote the number of elements in $\mathcal{W}$ of degree n. Then there exists $K_p \ge 1$ such that for $n = 1, 2, 3, \ldots$*

$$T_n \le (K_p d)^n.$$

Proof. Recall that $\mathcal{B}$ denotes the collection of trees that freely generate the Grossman Larson algebra. For $i = 1, 2, \ldots, [p]$, let l_i denote the number of trees in $\mathcal{B}$ of degree i. Then $T_n \le \sum_{i=1}^{[p]} T_{n-i} l_i d^i$. Set $T_0 = 1$ and $T_{-n} = 0$ for $n = 1, 2, \ldots [p]$. For $p \ge 1$, let $K_p \ge 1$ be a number such that $\sum_{i=1}^{[p]} l_i (K_p)^{-i} \le 1$. Then it can be proved inductively $T_n \le (K_p d)^n$. □

Define $I(x) := x$ for $x \in \mathbb{R}^e$.

Notation 3.8. *For $t \in \mathcal{B}_{\mathcal{L}}^{[p]}$ and $w \in \mathcal{W}$, denote $F^{\eta} := I$, $F^t := f(t)$ as at (2.4) and*

$$F^{tw} := dF^{w}(f(t)).$$

Notation 3.9. *With $f(t_i)$ defined at (2.4), let ψ_f denote the $\mathbb{R}$-linear map from $\mathbb{R}F_{\mathcal{L}}$ to differential operators, given by $\psi_f(t_1 \cdots t_k)(\varphi) := d^k \varphi(f(t_1) \cdots f(t_k))$ for $t_i \in T_{\mathcal{L}}$ and smooth $\varphi : \mathbb{R}^e \to \mathbb{R}^e$.*

For trees t_i and a forest ρ, define $(t_1 \cdots t_k) \curvearrowright \rho$ as the sum of $|\rho|^k$ forests that are obtained by linking each of the roots of $t_i, i = 1, \ldots, k$ to a vertex of ρ by a new edge. Recall that $*$ denotes the product in the Grossman Larson Hopf algebra (we delete the additional root in [14]). Then for trees t and t_i, $t * (t_1 \cdots t_k) = t t_1 \cdots t_k + t \curvearrowright (t_1 \cdots t_k)$.

Lemma 3.10. *With $f(t)$ defined at (2.4), for $t_i \in T_{\mathcal{L}}, i = 1, \ldots, k$,*

$$F^{t_1 \cdots t_k} = f(t_1 \curvearrowright (t_2 \curvearrowright \cdots (t_{k-1} \curvearrowright t_k))) = \psi_f(t_1 * \cdots * t_k)(I).$$

Proof. Since $df(t_2) f(t_1) = f(t_1 \curvearrowright t_2)$ for $t_1, t_2 \in T_{\mathcal{L}}$, the first equality holds. For trees t_1, t_2 and a forest ρ, $t_1 \curvearrowright (\rho \curvearrowright t_2) = (t_1 * \rho) \curvearrowright t_2$. Then it can be proved inductively that, for $t_i \in T_{\mathcal{L}}$, $t_1 \curvearrowright (t_2 \curvearrowright \cdots (t_{k-1} \curvearrowright t_k)) = (t_1 * t_2 * \cdots * t_{k-1}) \curvearrowright t_k$. Then the second equality holds based on $f((t_1 * t_2 * \cdots * t_{k-1}) \curvearrowright t_k) = \psi_f(t_1 * t_2 * \cdots * t_{k-1} * t_k)(I)$. □

For $\gamma > p \geq 1$, suppose $X : [0,T] \to G_{\mathcal{L}}^{[p]}$ is a branched p-rough path and suppose $f : \mathbb{R}^e \to L\left(\mathbb{R}^d, \mathbb{R}^e\right)$ is $\mathrm{Lip}\left(\gamma\right)$. Define $\omega : \{(s,t) \,|0 \leq s \leq t \leq T\} \to [0,\infty)$ as

$$\omega\left(s,t\right) := \|f\|_{\mathrm{Lip}(\gamma)}^{p} \|X\|_{p-var,[s,t]}^{p}.$$

By rescaling $\|f\|_{\mathrm{Lip}(\gamma)}^{-1} f$ and $\delta_{\|f\|_{\mathrm{Lip}(\gamma)}} X$, we assume $\|f\|_{\mathrm{Lip}(\gamma)} = 1$.

Denote $N := \lfloor \gamma \rfloor$ and $\{\gamma\} := \gamma - \lfloor \gamma \rfloor$.

Lemma 3.11. *For $w \in \mathcal{W}, |w| \leq N$,*

$$\|F^w\left(y_1\right) - F^w\left(y_2\right)\| \leq |w|! \, \|y_1 - y_2\|$$

for $y_i \in \mathbb{R}^e$. For $w \in \mathcal{W}, |w| < N$ and $t \in \mathcal{B}_{\mathcal{L}}^{[p]}$,

$$\sup_{y \in \mathbb{R}^e} \left\|F^{tw}\left(y\right)\right\| \leq (N-1)!$$

Proof. All trees here are labeled by $\mathcal{L} = \{1,2,\ldots,d\}$. Based on Lemma 3.10, $F^{t_1 \cdots t_k} = f\left(t_1 \curvearrowright \left(t_2 \curvearrowright \cdots \left(t_{k-1} \curvearrowright t_k\right)\right)\right)$. Then $F^{t_1 \cdots t_k}$ is the sum of the image of

$$|t_k| \left(|t_k| + |t_{k-1}|\right) \cdots \left(|t_k| + |t_{k-1}| + \cdots + |t_2|\right)$$

trees. Hence, for $w \in \mathcal{W}$, the number of trees in F^w is bounded by $(|w|-1)!$. Each of these trees t is of degree $|w|$ and corresponds to $f\left(t\right) : \mathbb{R}^e \to \mathbb{R}^e$ that is at least $\mathrm{Lip}\left(1 + \{\gamma\}\right)$ as $|w| \leq N$. Then $df\left(t\right)$ is a sum of $|w|$ terms, as the differential d chooses a vertex in t. Hence, $df\left(t\right)$ is bounded by $|w|$, because f and its derivatives of order up to N are uniformly bounded by 1 (we rescaled f by $\|f\|_{\mathrm{Lip}(\gamma)}^{-1}$). As a result, for each tree t of degree $|w|$, $\|f\left(t\right)\left(y_1\right) - f\left(t\right)\left(y_2\right)\| \leq \|df\left(t\right)\|_{\infty} \|y_1 - y_2\| \leq |w| \, \|y_1 - y_2\|$. Then the first estimate follows, as there are at most $(|w|-1)!$ such trees in F^w.

For $w \in \mathcal{W}, |w| < N$ and $t \in \mathcal{B}_{\mathcal{L}}^{[p]}$, the number of trees in F^{tw} is bounded by $|w|! \leq (N-1)!$. Each tree corresponds to a map that is bounded on $\mathbb{R}^e$ by 1. ☐

Recall that $\mathcal{B}_{\mathcal{L}}^{[p]} = \{\upsilon_1, \ldots, \upsilon_K\}$. For $s \leq t$, let $x = \left(x^{\upsilon_1}, \ldots, x^{\upsilon_K}\right) : [s,t] \to \mathbb{R}^K$ be a geodesic associated with $\bar{X}_{s,t}$. With $f\left(\upsilon_i\right)$ defined at (2.4), let $y^{s,t}$ denote the unique solution of the ODE

$$dy_u^{s,t} = \sum_{i=1}^{K} f\left(\upsilon_i\right)\left(y_u^{s,t}\right) dx_u^{\upsilon_i}, y_s^{s,t} = y_s,$$

where y denotes the unique solution of the branched RDE

$$dy_t = f\left(y_t\right) dX_t, y_0 = \xi.$$

The existence and uniqueness of y is based on [18, Theorem 22].

Lemma 3.12. *For $w \in \mathcal{W}$, $|w| = N - [p], \ldots, N-1$,*

$$\begin{aligned}
&F^w\left(y_t^{s,t}\right) - F^w\left(y_s\right) \\
&- \sum_{|t_1|+\cdots+|t_k|=1,\ldots,N-|w|} F^{t_1 \cdots t_k w}\left(y_s\right) \int \cdots \int_{s<u_1<\cdots<u_k<t} dx_{u_1}^{t_1} \cdots dx_{u_k}^{t_k} \\
= &\sum_{|t_1|+\cdots+|t_k|=N-|w|} \int \cdots \int_{s<u_1<\cdots<u_k<t} \left(F^{t_1 \cdots t_k w}\left(y_{u_1}^{s,t}\right) - F^{t_1 \cdots t_k w}\left(y_s\right)\right) dx_{u_1}^{t_1} \cdots dx_{u_k}^{t_k} \\
&+ \sum_{\substack{|t_2|+\cdots+|t_k|<N-|w| \\ |t_1|+|t_2|+\cdots+|t_k|>N-|w|}} \int \cdots \int_{s<u_1<\cdots<u_k<t} F^{t_1 \cdots t_k w}\left(y_{u_1}^{s,t}\right) dx_{u_1}^{t_1} \cdots dx_{u_k}^{t_k}
\end{aligned}$$

where $t_i, i = 1,2,\ldots$ range over elements in $\mathcal{B}_{\mathcal{L}}^{[p]}$.

Proof. The equality can be obtained by iteratively applying the fundamental theorem of calculus. □

Lemma 3.13.

$$\sup_{u\in[s,t]}\left\|y_u^{s,t}-y_s\right\|\leq C\left(p,d,\omega\left(0,T\right)\right)\omega\left(s,t\right)^{\frac{1}{p}}$$

Proof. Since K is the number of elements in $\mathcal{B}_{\mathcal{L}}^{[p]}$ with $\mathcal{L}=\{1,\ldots,d\}$, K only depends on p,d. Since $\|f\|_{\mathrm{Lip}(\gamma)}=1$, based on Lemma 3.5,

$$\sup_{u\in[s,t]}\left\|y_u^{s,t}-y_s\right\|\leq\sum_{i=1}^{K}\left\|x^{\upsilon_i}\right\|_{1-var}\leq C\left(p,d,\omega\left(0,T\right)\right)\omega\left(s,t\right)^{\frac{1}{p}}. \qquad \square$$

Recall that $*$ denotes the product of the Grossman Larson Hopf algebra.

Notation 3.14. *Define T^X as*

$$\left(T_{s,t}^X,t_1\cdots t_k\right):=\left(\bar{X}_{s,t},t_1*\cdots*t_k\right)$$

for $s\leq t$ and $t_1\cdots t_k\in\mathcal{W}$ for $t_i\in\mathcal{B}_{\mathcal{L}}^{[p]},|t_1|+\cdots+|t_k|\leq[p]$.

Denote $\beta_p:=p^2\left(1+\sum_{n\geq2}\left(\frac{2}{n}\right)^{\frac{[p]+1}{p}}\right)$.

Lemma 3.15. *Denote $\tilde{\omega}:=(K_pd)^p\,\omega$. For $w\in\mathcal{W}$, $|w|=N-[p],\ldots,N-1$,*

$$\begin{aligned}&\left\|F^w\left(y_t^{s,t}\right)-F^w\left(y_s\right)-\sum_{l\in\mathcal{W},|l|=1,\ldots,N-|w|}F^{lw}\left(y_s\right)\left(T_{s,t}^X,l\right)\right\|\\ \leq\quad& C\left(p,d,\omega\left(0,T\right)\right)N!\frac{\tilde{\omega}\left(s,t\right)^{\frac{N+1-|w|}{p}}}{\beta_p\left(\frac{N+1-|w|}{p}\right)!}.\end{aligned}$$

For $w\in\mathcal{W}$, $|w|=0,\ldots,N-[p]-1$,

$$\begin{aligned}&\left\|F^w\left(y_t^{s,t}\right)-F^w\left(y_s\right)-\sum_{l\in\mathcal{W},|l|=1,\ldots,[p]}F^{lw}\left(y_s\right)\left(T_{s,t}^X,l\right)\right\|\\ \leq\quad& C\left(p,d,\omega\left(0,T\right)\right)\left(|w|+[p]\right)!\tilde{\omega}\left(s,t\right)^{\frac{[p]+1}{p}}.\end{aligned}$$

Proof. We prove the first estimate. The proof for the second estimate is similar. Recall that T_n denotes the number of elements in $\mathcal{W}$ of degree n. Based on Lemma 3.12, Lemma 3.11, Lemma 3.13, Lemma 3.5 and that $T_n\leq(K_pd)^n$ in Lemma 3.7, we have

$$\begin{aligned}&\left\|F^w\left(y_t^{s,t}\right)-F^w\left(y_s\right)-\sum_{l\in\mathcal{W},|l|=1,\ldots,N-|w|}F^{lw}\left(y_s\right)\left(T_{s,t}^X,l\right)\right\|\\ \leq\quad& C\left(p,d,\omega\left(0,T\right)\right)N!\left(T_{N-|w|}\omega\left(s,t\right)^{\frac{N+1-|w|}{p}}+\sum_{j=1}^{[p]-1}T_{N-|w|+j}\omega\left(s,t\right)^{\frac{N-|w|+j}{p}}\right)\\ \leq\quad& C\left(p,d,\omega\left(0,T\right)\right)N!\left(\left(K_pd\right)^p\omega\left(s,t\right)\right)^{\frac{N+1-|w|}{p}}\\ \leq\quad& C\left(p,d,\omega\left(0,T\right)\right)N!\frac{\tilde{\omega}\left(s,t\right)^{\frac{N+1-|w|}{p}}}{\beta_p\left(\frac{N+1-|w|}{p}\right)!}\end{aligned}$$

as $\frac{N+1-|w|}{p}\leq\frac{[p]+1}{p}\leq2$, where $\tilde{\omega}:=(K_pd)^p\,\omega$. □

Proposition 3.16. *For integer* $N \geq [p]$*,*

$$\sum_{l\in\mathcal{W},|l|=1}^{N} F^{l}\left(T_{s,t}^{X},l\right)=\sum_{\tau\in T_{\mathcal{L}}^{N}} f\left(\tau\right)\frac{\left(X_{s,t},\tau\right)}{\sigma\left(\tau\right)}.$$

Proof. According to $T_{s,t}^{X}$ in Notation 3.14,

$$\left(T_{s,t}^{X},t_{1}\cdots t_{k}\right)=\left(\bar{X}_{s,t},t_{1}*\cdots*t_{k}\right) \tag{3.2}$$

for $t_i \in \mathcal{B}_{\mathcal{L}}^{[p]}, |t_1|+\cdots+|t_k| \leq [p]$. Then based on the construction of the extension of T^X and $\bar{X}$ [17, Theorem 2.2.1], it can be proved inductively that (3.2) holds for $t_i \in \mathcal{B}_{\mathcal{L}}^{[p]}, |t_1|+\cdots+|t_n| \leq N, N \geq [p]$. Moreover, if there exists $i \in \{1,\ldots,k\}$ such that $t_i \in \mathcal{B}_{\mathcal{L}} \backslash \mathcal{B}_{\mathcal{L}}^{[p]}$, then $\left(\bar{X}_{s,t},t_1*\cdots*t_k\right)=0$.

$$\begin{aligned}
\bar{X}_{s,t} &= \sum_{t_i\in\mathcal{B}_{\mathcal{L}}^{[p]},|t_1|+\cdots+|t_k|\leq N}\left(\bar{X}_{s,t},t_1*\cdots*t_k\right)t_1*\cdots*t_k\\
&= \sum_{\rho\in F_{\mathcal{L}}^{N},|\rho|\leq N}\left(\bar{X}_{s,t},\rho\right)\rho\\
&= \sum_{\rho\in F_{\mathcal{L}}^{N},|\rho|\leq N}\frac{\left(X_{s,t},\rho\right)}{\sigma\left(\rho\right)}\rho,
\end{aligned}$$

where the last step is based on Proposition 2.3. Combined with Lemma 3.10,

$$\sum_{l\in\mathcal{W},|l|=1}^{N} F^{l}\left(T_{s,t}^{X},l\right)=\psi_{f}\left(\bar{X}_{s,t}\right)I=\sum_{\tau\in T_{\mathcal{L}}^{N}} f\left(\tau\right)\frac{\left(X_{s,t},\tau\right)}{\sigma\left(\tau\right)}.$$ □

Proposition 3.17.

$$\left\|y_t-y_t^{s,t}\right\|\leq C\left(p,d,\omega\left(0,T\right)\right)\omega\left(s,t\right)^{\frac{[p]+1}{p}}$$

Proof. Let $F^w = I$ with $|w|=0$ in the second estimate of Lemma 3.15, and combine with Proposition 3.16,

$$\left\|y_t^{s,t}-y_s-\sum_{\tau\in T_{\mathcal{L}}^{[p]}} f\left(\tau\right)\left(y_s\right)\frac{\left(X_{s,t},\tau\right)}{\sigma\left(\tau\right)}\right\|\leq C\left(p,d,\omega\left(0,T\right)\right)\omega\left(s,t\right)^{\frac{[p]+1}{p}}.$$

Based on [18, Lemma 17],

$$\left\|y_t-y_s-\sum_{\tau\in T_{\mathcal{L}}^{[p]}} f\left(\tau\right)\left(y_s\right)\frac{\left(X_{s,t},\tau\right)}{\sigma\left(\tau\right)}\right\|\leq C\left(p,d,\omega\left(0,T\right)\right)\omega\left(s,t\right)^{\frac{[p]+1}{p}}. \tag{3.3}$$

The estimate (3.3) can be proved based on the uniform bound on Picard series [18, Definition 9, Lemma 17] and that Picard series converges to the unique solution [18, Theorem 22]. □

Lemma 3.18. *Set* $\tilde{\omega} := (K_p d)^p\,\omega$*. For* $w \in \mathcal{W}$*,* $|w| = N-[p],\ldots,N$*,*

$$\begin{aligned}
&\left\|F^{w}\left(y_t\right)-F^{w}\left(y_s\right)-\sum_{l\in\mathcal{W},|l|=1,\ldots,N-|w|}F^{lw}\left(y_s\right)\left(T_{s,t}^{X},l\right)\right\| \qquad (3.4)\\
\leq\;& C\left(p,d,\omega\left(0,T\right)\right)N!\frac{\left(\tilde{\omega}\left(s,t\right)\right)^{\frac{N+1-|w|}{p}}}{\beta_p\left(\frac{N+1-|w|}{p}\right)!}.
\end{aligned}$$

For $w\in\mathcal{W}$, $|w|=0,\dots,N-[p]-1$,

$$\begin{aligned}&\left\|F^{w}\left(y_{t}\right)-F^{w}\left(y_{s}\right)-\sum_{l\in\mathcal{W},|l|=1,\dots,[p]}F^{lw}\left(y_{s}\right)\left(T_{s,t}^{X},l\right)\right\| \\ \leq\quad & C\left(p,d,\omega\left(0,T\right)\right)\left(|w|+[p]\right)!\tilde{\omega}\left(s,t\right)^{\frac{[p]+1}{p}}.\end{aligned}\tag{3.5}$$

Proof. Combine Lemma 3.11 with Proposition 3.17,

$$\left\|F^{w}\left(y_{t}^{s,t}\right)-F^{w}\left(y_{t}\right)\right\|\leq|w|!\left\|y_{t}^{s,t}-y_{t}\right\|\leq C\left(p,d,\omega\left(0,T\right)\right)|w|!\tilde{\omega}\left(s,t\right)^{\frac{[p]+1}{p}}.$$

When $|w|\leq N-1$, the results follow from Lemma 3.15. When $|w|=N$, based on Lemma 3.11,

$$\begin{aligned}\left\|F^{w}\left(y_{t}\right)-F^{w}\left(y_{s}\right)\right\| &\leq N!\left\|y_{t}-y_{s}\right\|\leq N!\left(\left\|y_{t}-y_{t}^{s,t}\right\|+\left\|y_{t}^{s,t}-y_{s}\right\|\right)\\ &\leq C\left(p,d,\omega\left(0,T\right)\right)N!\frac{\tilde{\omega}\left(s,t\right)^{\frac{1}{p}}}{\beta_{p}\left(\frac{1}{p}\right)!}\end{aligned}$$

where the last step follows from Proposition 3.17 and Lemma 3.13. □

Lemma 3.19. *For* $l\in\mathcal{W},|l|=1,2,\dots$

$$\left\|\left(T_{s,t}^{X},l\right)\right\|\leq\frac{\tilde{\omega}\left(s,t\right)^{\frac{|l|}{p}}}{\beta_{p}\left(\frac{l}{p}\right)!}$$

where $\tilde{\omega}=c_{p,d}\omega$ *for some constant* $c_{p,d}$ *depending on* p,d.

Proof. Define two norms on $\mathcal{G}_{\mathcal{L}}^{[p]}$ as $\|a\|_{1}:=\max_{\rho\in F_{\mathcal{L}}^{[p]}}|(a,\rho)|^{\frac{1}{|\rho|}}$ and

$$\|a\|_{2}:=\max_{t_{i}\in\mathcal{B}_{\mathcal{L}}^{[p]},|t_{1}|+\cdots+|t_{n}|\leq[p]}\left|\left(a,t_{1}*\cdots*t_{n}\right)\right|^{\frac{1}{|t_{1}|+\cdots+|t_{n}|}}.$$

Based on the definition of T^{X} in Notation 3.14, equivalency of continuous homogeneous norms on $\mathcal{G}_{\mathcal{L}}^{[p]}$ as in Proposition 3.4 and $\left(\bar{X}_{s,t},\rho\right)=\left(X_{s,t},\rho\right)/\sigma\left(\rho\right)$, we have, for $l\in\mathcal{W}$, $|l|=1,\dots,[p]$, with $\|X_{s,t}\|$ defined at (2.1),

$$\left\|\left(T_{s,t}^{X},l\right)\right\|^{\frac{1}{|l|}}\leq\left\|\bar{X}_{s,t}\right\|_{2}\leq c_{p,d}^{1}\left\|\bar{X}_{s,t}\right\|_{1}\leq c_{p,d}^{1}\left\|X_{s,t}\right\|\leq c_{p,d}^{1}\left\|X\right\|_{p-var,[s,t]}.$$

Then the estimate follows from [17, Theorem 2.2.1] with $c_{p,d}:=\left(c_{p,d}^{1}\beta_{p}\right)^{p}$. □

Proof of Theorem 2.5. With K_{p} in Lemma 3.7 and $c_{p,d}$ in Lemma 3.19, denote $c_{p,d}^{2}:=\left(K_{p}d\right)^{p}\left(c_{p,d}\vee1\right)$ and set $\tilde{\omega}:=c_{p,d}^{2}\omega$. Denote $Y_{t}^{w}:=F^{w}\left(y_{t}\right)$ for $w\in\mathcal{W},|w|\leq N$ and $t\in[0,T]$.

Inductive hypothesis: fix $w\in\mathcal{W}$, $|w|\leq N-[p]-1$. Suppose for every $w_{1}\in\mathcal{W}$, $|w_{1}|=|w|+1,\dots,N$ and every $s\leq t$,

$$\left\|Y_{t}^{w_{1}}-Y_{s}^{w_{1}}-\sum_{l\in\mathcal{W},|l|=1}^{N-|w_{1}|}Y_{s}^{lw_{1}}\left(T_{s,t}^{X},l\right)\right\|\leq c_{p,d,\omega(0,T)}^{1}N!\frac{\tilde{\omega}\left(s,t\right)^{\frac{N+1-|w_{1}|}{p}}}{\beta_{p}\left(\frac{N+1-|w_{1}|}{p}\right)!}.$$

The statement holds when $|w|=N-[p]-1$ based on (3.4).

Denote

$$L_{s,t} := \sum_{l\in\mathcal{W},|l|=1}^{N-|w|} Y_s^{lw}\left(T_{s,t}^X, l\right).$$

Based on Lemma 3.11, $\left\|Y_s^{lw}\right\| \le (N-1)!$ for $l,w\in\mathcal{W}$, $|l|+|w|\le N$. Combined with (3.5) and Lemma 3.19,

$$\begin{aligned}
& \left\|Y_t^w - Y_s^w - L_{s,t}\right\| \\
\le\ & \left\|Y_t^w - Y_s^w - \sum_{l\in\mathcal{W},|l|=1}^{[p]} Y_s^{lw}\left(T_{s,t}^X, l\right)\right\| + \left\|\sum_{l\in\mathcal{W},|l|=[p]+1}^{N-|w|} Y_s^{lw}\left(T_{s,t}^X, l\right)\right\| \\
\le\ & C\left(p,d,\omega\left(0,T\right),N\right)\omega\left(s,t\right)^{\frac{[p]+1}{p}}.
\end{aligned}$$

Then

$$Y_t^w - Y_s^w = \lim_{|D|\to 0, D\subset[s,t]} \sum_{i,t_i\in D} L_{t_i,t_{i+1}}.$$

For $s\le u\le t$,

$$\begin{aligned}
& L_{s,u}+L_{u,t}-L_{s,t} \\
=\ & \sum_{l\in\mathcal{W},|l|=1}^{N-|w|} Y_s^{lw}\left(T_{s,u}^X, l\right) + \sum_{l\in\mathcal{W},|l|=1}^{N-|w|} Y_u^{lw}\left(T_{u,t}^X, l\right) - \sum_{l\in\mathcal{W},|l|=1}^{N-|w|} Y_s^{lw}\left(T_{s,t}^X, l\right) \\
=\ & \sum_{l\in\mathcal{W},|l|=1}^{N-|w|}\left(Y_u^{lw} - \sum_{l_1\in\mathcal{W},|l_1|=0}^{N-|w|-|l|} Y_s^{l_1 lw}\left(T_{s,u}^X, l_1\right)\right)\left(T_{u,t}^X, l\right).
\end{aligned}$$

Combine the inductive hypothesis and Lemma 3.19,

$$\begin{aligned}
& \left\|L_{s,u}+L_{u,t}-L_{s,t}\right\| \\
\le\ & c_{p,d,\omega(0,T)}^1 N! \sum_{n=1}^{N-|w|} T_n \frac{\tilde{\omega}\left(s,u\right)^{\frac{N+1-n-|w|}{p}}}{\beta_p\left(\frac{N+1-n-|w|}{p}\right)!}\frac{\left(c_{p,d}\omega\left(u,t\right)\right)^{\frac{n}{p}}}{\beta_p\left(\frac{n}{p}\right)!}.
\end{aligned}$$

where T_n denotes the number of elements in $\mathcal{W}$ of order n, and $T_n \le (K_p d)^n$ based on Lemma 3.7.

Since $\tilde{\omega} = (K_p d)^p\left(c_{p,d}\vee 1\right)\omega$, based on the neo-classical inequality [15, 17],

$$\left\|L_{s,u}+L_{u,t}-L_{s,t}\right\| \le c_{p,d,\omega(0,T)}^1 N! \frac{p^2}{\left(\beta_p\right)^2}\frac{\tilde{\omega}\left(s,t\right)^{\frac{N+1-|w|}{p}}}{\left(\frac{N+1-|w|}{p}\right)!}$$

Since $|w|\le N-[p]-1$, $\frac{N+1-|w|}{p} > \frac{[p]+1}{p}$. Successively dropping points similarly to the proof of [17, Theorem 2.2.1],

$$\left\|Y_t^w - Y_s^w - \sum_{l\in\mathcal{W},|l|=1}^{N-|w|} Y_s^{lw}\left(T_{s,t}^X, l\right)\right\| \le c_{p,d,\omega(0,T)}^1 N! \frac{\tilde{\omega}\left(s,t\right)^{\frac{N+1-|w|}{p}}}{\beta_p\left(\frac{N+1-|w|}{p}\right)!}.$$

The induction is complete.

Let w be the empty sequence. Then $|w|=0$ and $Y_t^w = y_t$. Combined with Proposition 3.16, the proposed estimate holds. $\square$

References

[1] Horatio Boedihardjo. Decay rate of iterated integrals of branched rough paths. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 35(4):945–969, 2018. MR3795022

[2] Horatio Boedihardjo and Ilya Chevyrev. An isomorphism between branched and geometric rough paths. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 55(2):1131–1148, 2019. MR3949967

[3] Horatio Boedihardjo, Terry Lyons, and Danyu Yang. Uniform factorial decay estimates for controlled differential equations. *Electronic Communications in Probability*, 20:1–11, 2015. MR3438739

[4] John C Butcher. An algebraic theory of integration methods. *Mathematics of Computation*, 26(117):79–106, 1972. MR0305608

[5] Frédéric Chapoton. Free pre-Lie algebras are free as Lie algebras. *Canadian Mathematical Bulletin*, 53(3):425–437, 2010. MR2682539

[6] Alain Connes and Dirk Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Communications in Mathematical Physics*, 199(1):203–242, 1998. MR1660199

[7] Alain Connes and Dirk Kreimer. Lessons from quantum field theory: Hopf algebras and spacetime geometries. *Letters in Mathematical Physics*, 48(1):85–96, 1999. MR1718046

[8] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, 2009. MR2483235

[9] Loïc Foissy. Les algebres de Hopf des arbres enracinés décorés, II. *Bulletin des sciences mathematiques*, 126(4):249–288, 2002. MR1909461

[10] Loïc Foissy. Finite dimensional comodules over the Hopf algebra of rooted trees. *Journal of Algebra*, 255(1):89–120, 2002. MR1935036

[11] Peter K Friz and Nicolas B Victoir. *Multidimensional stochastic processes as rough paths: theory and applications*, volume 120. Cambridge University Press, 2010. MR2604669

[12] Massimiliano Gubinelli. Ramification of rough paths. *Journal of Differential Equations*, 248(4):693–721, 2010. MR2578445

[13] Lajos Gergely Gyurkó. Differential equations driven by Π-rough paths. *Proceedings of the Edinburgh Mathematical Society*, 59(3):741–758, 2016. MR3572768

[14] Robert Grossman and Richard G Larson. Hopf-algebraic structure of families of trees. *Journal of Algebra*, 126(1):184–210, 1989. MR1023294

[15] Keisuke Hara and Masanori Hino. Fractional order Taylor's series and the neo-classical inequality. *Bulletin of the London Mathematical Society*, 42(3):467–477, 2010. MR2651942

[16] Michael Hoffman. Combinatorics of rooted trees and Hopf algebras. *Transactions of the American Mathematical Society*, 355(9):3795–3811, 2003. MR1990174

[17] Terry J Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998. MR1654527

[18] Terry J Lyons and Danyu Yang. The theory of rough paths via one-forms and the extension of an argument of Schwartz to rough differential equations. *Journal of the Mathematical Society of Japan*, 67(4):1681–1703, 2015. MR3417509

[19] Florin Panaite. Relating the Connes–Kreimer and Grossman–Larson Hopf algebras built on rooted trees. *Letters in Mathematical Physics*, 51(3):211–219, 2000. MR1775423

[20] Christophe Reutenauer. *Free Lie algebras*. Clarendon Press, Oxford, 1993. MR1231799