Boedihardjo, H. and Geng, X.
Osaka J. Math.
59 (2022), 653–682

LIPSCHITZ-STABILITY OF CONTROLLED ROUGH PATHS AND ROUGH DIFFERENTIAL EQUATIONS

Horatio BOEDIHARDJO and Xi GENG

(Received November 16, 2020, revised May 31, 2021)

Abstract

We provide an account for the existence and uniqueness of solutions to rough differential equations in infinite dimensions under the framework of controlled rough paths. The case when the driving path is α-Hölder continuous for $\alpha > 1/3$ is widely available in the literature. In its extension to the case when $\alpha \leqslant 1/3$, the main challenge and missing ingredient is to show that controlled rough paths are closed under composition with Lipschitz transformations. Establishing such a property precisely, which has a strong algebraic nature, is a main purpose of the present article.

1. Introduction

Multidimensional stochastic differential equations (SDEs) of the form

$$\mathrm{d}Y_t = \sum_{j=0}^{d} V_j(Y_t)\mathrm{d}X_t^j, \quad Y_0 = y, \tag{1.1}$$

where $X_t^0 = t$, $(X_t^j)_{j=1}^d$ is a d-dimensional Brownian motion, and $(V_j)_{j=0}^d$ are smooth vector fields on $\mathbb{R}^n$, has been frequently used for modelling in mathematical physics and finance (cf. [13] and the references therein). The case when $V_j = 0$ for all $j \geqslant 0$ corresponds to ordinary differential equations (ODEs). The SDE (1.1) also has applications in pure mathematics. For instance, the distribution of its solution can be used to study some second order linear parabolic and elliptic differential equations, leading to probabilistic proofs of celebrated results in PDE theory such as Hörmander's theorem (cf. Malliavin [11]).

When using Picard's iteration to establish the existence and uniqueness of solutions to (1.1), the convergence of the iteration is established under the L^2-norm with respect to the Wiener measure. Partly inspired by the conjectures of H. Fölmer, Lyons [8] developed a *pathwise* approach to construct the integral against the "$\mathrm{d}X_t^j$'s" and showed the pathwise well-posedness of the SDE. Lyons' pathwise estimates were performed through considering the Brownian motion as an enhanced object by including the second order structure given by an iterated integral process:

$$\mathbf{X}_{s,t} = \Big(X_t - X_s, \int_{s<u_1<u_2<t} \mathrm{d}X_{u_1} \otimes \mathrm{d}X_{u_2}\Big).$$

In fact, given any function $(s,t) \to \mathbf{X}_{s,t}$ satisfying certain algebraic and analytic conditions, a unique solution Y to the equation (1.1) can be constructed in terms of $\mathbf{X}$, so that the mapping

2020 Mathematics Subject Classification. Primary 34A12, 34H99; Secondary 17B01, 60H10.

$\mathbf{X} \to Y$ is continuous. Such functions $\mathbf{X}$ are known as *weakly geometric rough paths*.

Lyons defined the solution for (1.1) effectively as

$$\mathbf{Y}_{s,t} = (Y_t - Y_s, \int_{s<u_1<u_2<t} \mathrm{d}Y_{u_1} \otimes \mathrm{d}Y_{u_2})$$

so that the solution path $\mathbf{Y}$, like $\mathbf{X}$, is also a weakly geometric rough path. Lyons' rough path theory has an analytic nature and goes way beyond the framework of Brownian motion. Later on, Gubinelli [5] proposed an alternative way to interpret the solution Y as a *controlled path*, which we will elaborate below. The monograph of Friz and Hairer [2] contains an excellent exposition of this approach. In contrast to weakly geometric rough paths, the set of controlled paths has a nice linear structure making it into a Banach space and some algebraic considerations are simplified accordingly. Both [5, 2] contain the complete theory for the case when the Hölder exponent α of $\mathbf{X}$ is greater than $1/3$.

While for most parts it is commonly believed that the extension to the case when $\alpha \leqslant 1/3$ is standard, the proofs and precise quantitative estimates under the framework of controlled paths do not seem to be readily available in the literature. There is an essential ingredient whose extension to the case when $\alpha \leqslant 1/3$ is not obvious at all. To be more specific, when formulating the differential equation

$$d\mathcal{Y} = F(\mathcal{Y})\mathrm{d}\mathbf{X} \tag{1.2}$$

in the sense of controlled paths, one needs to prove that if $\mathcal{Y}$ is controlled by $\mathbf{X}$ and F is a suitably regular function, then $F(\mathcal{Y})$ is also controlled by $\mathbf{X}$. In the case when $\alpha \leqslant 1/3$, such a stability property was first established by Gubinelli [6] and more recently by Friz-Zhang [4] in the context of branched rough paths. Correspondingly, existence and uniqueness of solutions to differential equations driven by branched rough paths was also established in these works. Since all finite dimensional geometric rough paths can be considered as branched rough paths, [6] and [4] provide a natural generalisation of the controlled rough path theory which also allow arbitrary regularity. However, the theory of branched rough path is essentially finite dimensional since branched rough paths are indexed by rooted trees over a finite set of labels. The combinatorial analysis of branched rough paths also reflects its finite dimensionality in a crucial way.

The main goal of the present article is to develop a generalisation of controlled rough path theory to the case $\alpha \leqslant 1/3$ in infinite dimensions in an intrinsic way. In other words, the underlying paths are assume to take values in Banach spaces and the current approach does not rely on a choice of basis. As we will see, the main challenge in proving the aforementioned stability property of controlled rough paths when $\alpha \leqslant 1/3$ has a strong algebraic nature that is not similar to the usual Hölder regularity estimates. The "geometric" feature of $\mathbf{X}$ plays a critical role which is not needed in the case when $\alpha > 1/3$. A major effort of the present article is to develop this algebraic component precisely based on tools from free Lie algebras (cf. Section 4 below). For completeness, we have also included a full proof towards the well-posedness (existence, uniqueness and continuity) of the equation (1.2) under the framework of controlled paths (cf. Theorem 6.1 in Section 6.2). In our modest opinion, having the controlled rough path framework properly set-up in full generality along with the key quantitative estimates might also be beneficial and convenient for the broader community. We remark that the consideration of infinite dimensional equations is needed in many applications. A notable example is the recent work of Ohashi-Russo-Shamarova [12], in which the authors established the existence and smoothness of density for path-dependent SDEs from the perspective of controlled differential equations in infinite dimensions.

Apart from Lyons' original approach and Gubinelli's controlled path approach, there are several other approaches to study differential equations driven by rough paths, some of which further develops the idea of controlled paths (see for instance Davie [1], Hairer [7], Lyons-Yang [10]).

Organization. The present article is organized as follows. In Section 2, we recall the basic notions of geometric rough paths and controlled rough paths. In Section 3, we derive a Hölder estimate for controlled rough paths in terms of the remainders. This estimate is needed for later purposes. In Section 4, we prove the stability of controlled rough paths under Lipschitz transformations. This part is a main ingredient of the present article. In Sections 5 and 6, we study rough integration and rough differential equations.

2. Preliminary notions of rough paths

We begin by recapturing some notions of geometric and controlled rough paths over Banach spaces. This provides the framework on which the present article is based.

2.1. Geometric rough paths. Let U and V denote Banach spaces. The spaces U and V will represent the space in which the paths Y and X in (1.1) take values respectively. A family of *admissible tensor norms* on $(V^{\otimes n})_{n=1}^{\infty}$ (cf. Lyons-Qian [9]) is a family of norms, one for each of $V^{\otimes n}$, such that:

For $v \in V^{\otimes n}$ and $w \in V^{\otimes k}$,

$$\|v \otimes w\|_{V^{\otimes(n+k)}} \leqslant \|v\|_{V^{\otimes n}} \|w\|_{V^{\otimes k}};$$

Given a permutation σ of order n, let P_σ denote a linear transformation on $V^{\otimes n}$ such that

$$P_\sigma(v_1 \otimes \cdots \otimes v_n) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}.$$

Then for all $v \in V^{\otimes n}$,

$$\|P_\sigma(v)\|_{V^{\otimes n}} = \|v\|_{V^{\otimes n}}.$$

Throughout the rest, whenever working with Banach tensor products, we always assume that a family of admissible tensor norms is given fixed. For simplicity, we always use $|\cdot|$ to denote norms of tensors, and use $\|\cdot\|$ to denote Hölder norms of paths.

Let $\mathcal{L}(U;V)$ denotes the space of bounded linear operators from U to V. We frequently identify spaces $\mathcal{L}(U;\mathcal{L}(U;V))$ and $\mathcal{L}(U^{\otimes 2};V)$, and similarly for more general cases $\mathcal{L}(U^{\otimes n};V)$.

Let $0 < \alpha \leqslant 1/2$ and set $N \triangleq [1/\alpha]$ to be the largest integer that does not exceed $1/\alpha$. The number α is fixed throughout this article, and all constants in the article will, without further comment, depend on α.

A continuous mapping $X^i : \Delta_T \triangleq \{(s,t) : 0 \leqslant s \leqslant t \leqslant T\} \to V^{\otimes i}$ is called *α-Hölder continuous* if

$$\|X^i\|_{i\alpha} \triangleq \sup_{0\leqslant s<t\leqslant T} \frac{|X^i_{s,t}|}{(t-s)^{i\alpha}} < \infty.$$

Let $T^{(N)}(V)$ denote the truncated tensor algebra $1 \oplus V \oplus \cdots \oplus V^{\otimes N}$. A mapping $\mathbf{X} : \Delta_T \to T^{(N)}(V)$ is called *multiplicative* if for any $s \leqslant u \leqslant t$,

$$\mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} = \mathbf{X}_{s,t}.$$

The following algebraic structure will be used in Section 4 in a crucial way. For each $k \geqslant 1$, consider the algebra

$$T^{(N)}(V)^{\boxtimes k} \triangleq \underbrace{T^{(N)}(V) \boxtimes \cdots \boxtimes T^{(N)}(V)}_{k}.$$

Here $\boxtimes$ denotes the tensor product whose notation is used to distinguish from the one $\otimes$ defined over $T^{(N)}(V)$. The product structure $*$ over $T^{(N)}(V)^{\boxtimes k}$ is induced by

$$(\xi_1 \boxtimes \cdots \boxtimes \xi_k) * (\eta_1 \boxtimes \cdots \boxtimes \eta_k) \triangleq (\xi_1 \otimes \eta_1) \boxtimes \cdots \boxtimes (\xi_k \otimes \eta_k).$$

If $f_1, \cdots, f_k \in \mathcal{L}(T^{(N)}(V); U)$, we denote

$$f_1 \boxtimes \cdots \boxtimes f_k : T^{(N)}(V)^{\boxtimes k} \to U^{\boxtimes k}$$

as the mapping induced by

$$f_1 \boxtimes \cdots \boxtimes f_k(\xi_1 \boxtimes \cdots \boxtimes \xi_k) \triangleq f_1(\xi_1) \boxtimes \cdots \boxtimes f_k(\xi_k).$$

There is an algebra homomorphism

$$\delta_k : (T^{(N)}(V), \otimes) \to (T^{(N)}(V)^{\boxtimes k}, *)$$

induced by

$$\delta_k(v) \triangleq v \boxtimes \mathbf{1} \boxtimes \cdots \boxtimes \mathbf{1} + \cdots + \mathbf{1} \boxtimes \cdots \boxtimes \mathbf{1} \boxtimes v, \quad v \in V,$$

where $\mathbf{1} \triangleq (1, 0, \cdots, 0)$ denotes the unit element of $T^{(N)}(V)$. See [14], Section 1.4 for further details about δ_k. Let $\xi = v_1 \otimes \cdots \otimes v_r \in V^{\otimes r}$. Given $I = \{i_1, \cdots, i_m\}$ with $i_1 < \cdots < i_m$, we define

$$\xi|_I \triangleq v_{i_1} \otimes \cdots \otimes v_{i_m}$$

and we adopt the convention that $\xi|_\emptyset$ is the scalar 1. One useful property of δ_k is that

$$\delta_k(\xi) = \sum_{(I_\alpha)} v|_{I_1} \boxtimes \cdots \boxtimes v|_{I_k} \tag{2.1}$$

where the above summation is taken over all partitions (I_α) of $\{1, \cdots, r\}$ into disjoint subsets $I_1, \cdots, I_k$ (some of them can be $\emptyset$).

DEFINITION 2.1. The *free nilpotent group of order N* is the subset of $T^{(N)}(V)$ defined by

$$G^N(V) = \{\xi = (\xi^0, \cdots, \xi^N) \in T^{(N)}(V) : \delta_k(\xi) = \sum_{0\leqslant l_1+\cdots+l_k\leqslant N} \xi^{l_1} \boxtimes \cdots \boxtimes \xi^{l_k}, \quad \forall k \geqslant 2\}. \tag{2.2}$$

REMARK 2.1. The above characterization of the free nilpotent group of order N is equivalent to a common definition in terms of the exponential of Lie series. Indeed, according to [14], Theorem 3.2, an element $\xi \in T^{(\infty)}(V)$ (the algebra of infinite tensor series) is the exponential of a formal Lie series if and only if

$$\tilde{\delta}_2(\xi) = \xi \boxtimes \xi,$$

where $\tilde{\delta}_2$ is the canonical extension of δ_2 onto $T^{(\infty)}(V)$. By a similar proof, this is also equivalent to

$$\tilde{\delta}_k(\xi) = \xi^{\boxtimes k}, \quad \forall k \geqslant 2. \tag{2.3}$$

To see the equivalence between (2.3) and (2.2), given r and k let us introduce the projection

$$P_r : T^{(\infty)}(V)^{\boxtimes k} \to \bigoplus_{l_1+\cdots+l_k=r} V^{\otimes l_1} \boxtimes \cdots \boxtimes V^{\otimes l_k} \subseteq T^{(N)}(V)^{\boxtimes k}.$$

If $\xi = (\xi^0, \xi^1, \xi^2, \cdots)$ with $\xi^i \in V^{\otimes i}$, then

$$P_r(\tilde{\delta}_k(\xi)) = \sum_{l_1+\cdots+l_k=r} \xi^{l_1} \boxtimes \cdots \boxtimes \xi^{l_k}.$$

As (2.1) implies that δ_k sends $V^{\otimes r}$ to $\bigoplus_{l_1+\cdots+l_k=r} V^{\otimes l_1} \boxtimes \cdots \boxtimes V^{\otimes l_k}$, we see that

$$\delta_k(\xi^r) = P_r(\tilde{\delta}_k(\xi)) = \sum_{l_1+\cdots+l_k=r} \xi^{l_1} \boxtimes \cdots \boxtimes \xi^{l_k}.$$

Therefore, if $(\xi^0, \xi^1, \cdots, \xi^N)$ is the exponential of a Lie series on $T^{(N)}(V)$, we have

$$\delta_k((\xi^0, \cdots, \xi^N)) = \sum_{0 \leqslant l_1+\cdots+l_k \leqslant N} \xi^{l_1} \boxtimes \cdots \boxtimes \xi^{l_k}.$$

REMARK 2.2. Technically, we should assume that the algebraic tensor product $T^{(N)}(V)^{\boxtimes k}$ has been completed with respect to some admissible tensor norm (e.g. the projective tensor norm). But this is not an essential point since we will only work with elements in the algebraic tensor product (i.e. linear combinations of tensors of the form $\xi_1 \boxtimes \cdots \boxtimes \xi_k$).

DEFINITION 2.2. A α-Hölder *geometric rough path* $\mathbf{X}$ is a multiplicative functional

$$\mathbf{X} = (1, X^1, \cdots, X^N) : \Delta_T \to T^{(N)}(V)$$

such that $\mathbf{X}_{s,t} \in G^{(N)}(V)$ for any $(s,t) \in \Delta_T$ and X^i is α-Hölder continuous for each $1 \leqslant i \leqslant N$.

REMARK 2.3. Here we follow the convention in [2] and call such rough paths *geometric*. In the earlier rough path literature (e.g. [3]), such paths are often called weakly geometric.

Given two α-Hölder geometric rough paths $\mathbf{X}, \tilde{\mathbf{X}}$, we define their "distance" by

$$\rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}) \triangleq \sum_{i=1}^{N} \|X^i - \tilde{X}^i\|_{i\alpha}. \tag{2.4}$$

We also denote $\|\mathbf{X}\|_\alpha \triangleq \rho_\alpha(\mathbf{X}, \mathbf{1})$.

REMARK 2.4. A typical way of constructing geometric rough paths is as follows. Let $\{X^{(m)} : m \geqslant 1\}$ be a sequence of Lipschitz continuous paths in V. Then the limit

$$\mathbf{X}_{s,t} = \lim_{m\to\infty} \left(1, \int_{s<u_1<t} \mathrm{d}X^{(m)}_{u_1}, \cdots, \int_{s<u_1<\cdots<u_N<t} \mathrm{d}X^{(m)}_{u_1} \otimes \cdots \otimes \mathrm{d}X^{(m)}_{u_N}\right) \tag{2.5}$$

yields a α-Hölder geometric rough path provided that the convergence holds under the α-Hölder metric (2.4). When V is finite dimensional, the union over $\{\alpha : \alpha < \alpha'\}$ of all functionals $\Delta_T \to T^{(N)}(V)$ that can be constructed through the procedure of (2.5) is precisely the set of α'-Hölder geometric rough paths (cf. [3], Corollary 8.24).

According to [8], when $\mathbf{X}$ is a geometric rough path, the solution to the differential equation (1.1) can be constructed in the sense of geometric rough paths.

2.2. Controlled rough paths. In this article, we take the perspective of controlled rough paths introduced by Gubinelli [5]. A benefit of this viewpoint is that the underlying path space is a Banach space which simplifies algebraic considerations to some extent. Heuristically, the solution to the rough differential equation $\mathrm{d}Y_t = F(Y_t)\mathrm{d}X_t$ can be formulated as the fixed point of the mapping $\mathcal{M} : Y_\cdot \to \int_0^\cdot F(Y_t)\,\mathrm{d}X_t$, provided that $\mathcal{M}$ is a contraction on a suitable space of controlled rough paths. We first define the notion of controlled rough paths precisely.

DEFINITION 2.3. A collection of continuous paths $\mathcal{Y}_t = (Y_t^0, Y_t^1, \cdots, Y_t^{N-1})$, where $Y_t^0 \in U$ and $Y_t^i \in \mathcal{L}(V^{\otimes i}; U)$ for $1 \leqslant i \leqslant N-1$, is called an *$\alpha$-Hölder controlled rough path* over U with respect to $\mathbf{X}$, if the "remainder" defined by

$$\mathcal{R}\mathcal{Y}^i_{s,t} \triangleq \begin{cases} Y_t^i - Y_s^i - \sum_{j=1}^{N-1-i} Y_s^{i+j} X^j_{s,t}, & \text{if } 0 \leqslant i \leqslant N-2, \\ Y_t^{N-1} - Y_s^{N-1}, & \text{if } i = N-1, \end{cases}$$

satisfies for each $0 \leqslant i \leqslant N-1$,

$$\|\mathcal{R}\mathcal{Y}^i\|_{(N-i)\alpha} \triangleq \sup_{0\leqslant s<t\leqslant T} \frac{|\mathcal{R}\mathcal{Y}^i_{s,t}|}{|t-s|^{(N-i)\alpha}} < \infty.$$

The space of controlled rough paths over U with respect to $\mathbf{X}$ is denoted as $\mathcal{D}_{\mathbf{X};\alpha}(U)$. We define a semi-norm $\|\cdot\|_{\mathbf{X};\alpha}$ on $\mathcal{D}_{\mathbf{X};\alpha}(U)$ by

$$\|\mathcal{Y}\|_{\mathbf{X};\alpha} \triangleq \sum_{i=0}^{N-1} \|\mathcal{R}\mathcal{Y}^i\|_{(N-i)\alpha}.$$

REMARK 2.5. We often use the shorthanded notation $Y^i_{s,t} \triangleq Y^i_t - Y^i_s$.

Let $\mathbf{X}, \tilde{\mathbf{X}}$ be two α-Hölder rough paths. To measure the distance between $\mathcal{Y} \in \mathcal{D}_{\mathbf{X};\alpha}(U)$ and $\tilde{\mathcal{Y}} \in \mathcal{D}_{\tilde{\mathbf{X}};\alpha}(U)$, we define the functional

$$d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}, \tilde{\mathcal{Y}}) \triangleq \sum_{i=0}^{N-1} \|\mathcal{R}\mathcal{Y}^i - \mathcal{R}\tilde{\mathcal{Y}}^i\|_{(N-i)\alpha}.$$

3. Hölder estimates for controlled rough paths

The following lemma tells us how to estimate $\|Y^i - \tilde{Y}^i\|_\alpha$ in terms of $d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y},\tilde{\mathcal{Y}})$ and the difference of the initial data. This estimate is useful in the next section when we study the stability of controlled paths under Lipschitz transformations. For simplicity, we introduce the notation

$$\delta X^i \triangleq X^i - \tilde{X}^i,\ \delta Y^i \triangleq Y^i - \tilde{Y}^i,\ \delta\mathcal{R}^i \triangleq \mathcal{R}\mathcal{Y}^i - \mathcal{R}\tilde{\mathcal{Y}}^i.$$

Lemma 3.1. *For each $2 \leqslant i \leqslant N$, there exists a universal function $M_i : [0,\infty)^5 \to [0,\infty)$ which is continuous and increasing in each variable, such that*

$$\|\delta Y^{N-i}\|_\alpha \leqslant M_i(T, \|\mathbf{X}\|_\alpha, \|\tilde{\mathbf{X}}\|_\alpha, \max_{1\leqslant j\leqslant i-1}|Y_0^{N-j}|, \max_{1\leqslant j\leqslant i-1}\|\mathcal{R}\mathcal{Y}^{N-j}\|_{j\alpha})$$
$$\times[\rho_\alpha(\mathbf{X},\tilde{\mathbf{X}}) + \|\delta\mathcal{R}^{N-i}\|_{i\alpha} + \sum_{j=1}^{i-1}(|\delta Y_0^{N-j}| + \|\delta\mathcal{R}^{N-j}\|_{j\alpha})].$$

REMARK 3.1. If $i = 1$, $\delta Y^{N-i} = \delta\mathcal{R}^{N-1}$ and hence

$$\|\delta Y^{N-i}\|_\alpha = \|\delta\mathcal{R}^{N-1}\|_\alpha. \tag{3.1}$$

Proof. We prove the lemma by induction. When $i = 2$, we have

$$\begin{aligned} Y_{s,t}^{N-2} - \tilde{Y}_{s,t}^{N-2} &= (Y_s^{N-1}X_{s,t}^1 - \tilde{Y}_s^{N-1}\tilde{X}_{s,t}^1) + (\mathcal{R}\mathcal{Y}_{s,t}^{N-2} - \mathcal{R}\tilde{\mathcal{Y}}_{s,t}^{N-2}) \\ &= Y_s^{N-1}(X_{s,t}^1 - \tilde{X}_{s,t}^1) + (Y_s^{N-1} - \tilde{Y}_s^{N-1})\tilde{X}_{s,t}^1 + (\mathcal{R}\mathcal{Y}_{s,t}^{N-2} - \mathcal{R}\tilde{\mathcal{Y}}_{s,t}^{N-2}). \end{aligned}$$

It follows from (3.1) that

$$\begin{aligned} \|\delta Y^{N-2}\|_\alpha &\leqslant (1+T^\alpha)\,(|Y_0^{N-1}| + \|Y^{N-1}\|_\alpha)\|\delta X^1\|_\alpha \\ &\quad + (1+T^\alpha)\,[(|\delta Y_0^{N-1}| + \|\delta Y^{N-1}\|_\alpha)\|\tilde{X}^1\|_\alpha + \|\delta\mathcal{R}^{N-2}\|_{2\alpha}T^\alpha] \\ &\leqslant (1+T^\alpha)\,(T^\alpha + \|\tilde{\mathbf{X}}\|_\alpha + |Y_0^{N-1}| + \|\mathcal{R}\mathcal{Y}^{N-1}\|_\alpha) \\ &\quad \times[\|\delta X^1\|_\alpha + |\delta Y_0^{N-1}| + \|\delta\mathcal{R}^{N-1}\|_\alpha + \|\delta\mathcal{R}^{N-2}\|_{2\alpha}]. \end{aligned}$$

Therefore, the claim holds in this case.

Suppose that the claim holds for $\delta Y^{N-1}, \cdots, \delta Y^{N-i}$. Using that

$$\delta Y_{s,t}^{N-(i+1)} = \sum_{j=1}^{i} Y_s^{N-j}X_{s,t}^{i+1-j} - \sum_{j=1}^{i}\tilde{Y}_s^{N-j}\tilde{X}_{s,t}^{i+1-j} + \delta\mathcal{R}^{N-(i+1)},$$

we have

$$\begin{aligned} &\|\delta Y^{N-(i+1)}\|_\alpha \\ &\leqslant (1+T^{(i+1)\alpha})[\sum_{j=1}^{i}(|Y_0^{N-j}| + \|Y^{N-j}\|_\alpha)\|\delta X^{i+1-j}\|_{(i+1-j)\alpha} \\ &\quad + \sum_{j=1}^{i}(|\delta Y_0^{N-j}| + \|\delta Y^{N-j}\|_\alpha)\|\tilde{X}^{i+j}\|_{(i+j)\alpha} + \|\delta\mathcal{R}^{N-(i+1)}\|_{(i+1)\alpha}] \\ &\leqslant (1+T^{(i+1)\alpha})(1 + \max_{1\leqslant j\leqslant i}(|Y_0^{N-j}| + \|Y^{N-j}\|_\alpha) + \|\tilde{\mathbf{X}}\|_\alpha) \end{aligned} \tag{3.2}$$

$$[\rho_\alpha(\mathbf{X},\tilde{\mathbf{X}}) + \sum_{j=1}^{i}(|\delta Y_0^{N-j}| + \|\delta Y^{N-j}\|_\alpha) + \|\delta \mathcal{R}^{N-(i+1)}\|_{(i+1)\alpha}].$$

By the induction hypothesis with $\mathbf{X} = \tilde{\mathbf{X}}$ and taking $\tilde{Y} = 0$, there is a universal function M_j that is continuous and increasing in every variable, such that

$$\|Y^{N-j}\|_\alpha \leqslant M_j(T, \|\mathbf{X}\|_\alpha, \max_{1\leqslant l\leqslant j-1} |Y_0^{N-l}|, \max_{1\leqslant l\leqslant j} \|\mathcal{R}\mathcal{Y}^{N-l}\|_\alpha) \tag{3.3}$$

and similarly for each $1 \leqslant j \leqslant i$ we have

$$\begin{aligned}\|\delta Y^{N-j}\|_\alpha \leqslant\ & \tilde{M}_j(T, \|\mathbf{X}\|_\alpha, \|\tilde{\mathbf{X}}\|_\alpha, \max_{1\leqslant l\leqslant j-1} \left|Y_0^{N-l}\right|, \max_{1\leqslant l\leqslant j-1} \|\mathcal{R}\mathcal{Y}^{N-l}\|_\alpha) \\ & \times [\rho_\alpha(\mathbf{X},\tilde{\mathbf{X}}) + \|\delta\mathcal{R}^{N-j}\|_{j\alpha} + \sum_{l=1}^{j-1}(\left|\delta Y_0^{N-l}\right| + \|\delta\mathcal{R}^{N-l}\|_{l\alpha})].\end{aligned} \tag{3.4}$$

The induction step follows by substituting (3.3) and (3.4) into (3.2). □

REMARK 3.2. An immediate consequence of Lemma 3.1 is that $\mathcal{D}_{\mathbf{X};\alpha}(U)$ is a Banach space under the norm

$$|||\mathcal{Y}|||_{\mathbf{X};\alpha} \triangleq \|\mathcal{Y}\|_{\mathbf{X};\alpha} + \sum_{i=0}^{N-1} |Y_0^i|. \tag{3.5}$$

4. Stability of controlled rough paths under Lipschitz transformations

Under the framework of controlled rough paths, an essential ingredient for solving an RDE $d\mathcal{Y} = F(\mathcal{Y})d\mathbf{X}$ (with Lipschitz vector field F) is to show that $F(\mathcal{Y})$ is also a controlled rough path. We would like to point out that the extension of this property from the case of $1/3 < \alpha \leqslant 1/2$ (which is the common setting in most of the literature) to the general case of $\alpha < 1/3$ is non-trivial. As we will see, the main challenge has an *algebraic* nature rather than just being standard regularity estimates. To point this out concisely, the Taylor expansion of F for the 0-th level function (i.e. equation (4.1) below when $j = 0$) allows us to motivate the construction of $F(\mathcal{Y}) = (F(\mathcal{Y})^0, \cdots, F(\mathcal{Y})^{N-1})$ as a controlled rough path in one go. However, justifying the remainder regularity for all the derivative paths requires deeper algebraic considerations and the geometric nature of $\mathbf{X}$ plays an essential role. For this purpose, we take the viewpoint of Reutenauer [14] and rely on the coproduct structure δ_k introduced in Section 2.1 in a crucial way.

We begin by recalling the notion of Lipschitz functions in the sense to Stein [15]. Let $\mathcal{L}_{\mathrm{sym}}(V^{\boxtimes j}; W)$ denote the set of bounded linear operators T from $V^{\boxtimes j}$ to W such that for all permutations σ over $\{1, \cdots, j\}$,

$$T(v_{\sigma(1)} \boxtimes \cdots \boxtimes v_{\sigma(n)}) = T(v_1 \boxtimes \cdots \boxtimes v_n).$$

DEFINITION 4.1. Let W, U be two Banach spaces and let K be a closed subset of W. Suppose that $\gamma \in (N, N+1]$ where N is a non-negative integer. A collection of functions $F = (F^0, F^1, \cdots, F^N)$ is said to be *γ-Lipschitz* over K, if:

(i) the functions $F^0 : K \to U$ and $F^j : K \to \mathcal{L}_{\mathrm{sym}}(W^{\boxtimes j}; U)$ $(1 \leqslant j \leqslant N)$ are bounded on K;

(ii) for each $0 \leqslant j \leqslant N$, the following Taylor expansion holds:

$$F^j(y)(\xi) = \sum_{l=0}^{N-j} \frac{1}{l!} F^{j+l}(x)((y-x)^{\boxtimes l} \boxtimes \xi) + R_j(x,y)(\xi), \quad x,y \in K, \xi \in W^{\boxtimes j}, \tag{4.1}$$

where the remainder $R_j : K \times K \to \mathcal{L}_{\text{sym}}(W^{\boxtimes j}; U)$ satisfies

$$\sup_{x \neq y \in K} \frac{|R_j(x,y)|}{|x-y|^{\gamma-j}} < \infty \quad \text{for all } 0 \leqslant j \leqslant N.$$

The Lip-γ *norm* of F, denoted as $\|F\|_{\text{Lip-}\gamma}$, is defined to be the smallest number $M > 0$ such that for all $x, y \in K$,

$$|F^j(x)| \leqslant M, \ |R_j(x,y)| \leqslant M|x-y|^{\gamma-j}$$

for all $0 \leqslant j \leqslant N$. The Banach space of all γ-Lipschitz functions $F = (F^0, \cdots, F^N)$ is denoted as $\text{Lip}(\gamma, K)$.

Now let $\mathbf{X}$ be a given α-Hölder geometric rough path over a Banach space V. Our aim in this section is to show that, if $\mathcal{Y}$ is an α-Hölder controlled rough path over W with respect to $\mathbf{X}$ and if $F = (F^0, \cdots, F^N)$ is γ-Lipschitz over W taking values in U, then $F(\mathcal{Y})$ is an α-Hölder controlled rough path over U. In addition, given another controlled rough path $\tilde{\mathcal{Y}}$, we shall establish a quantitative continuity estimate of $d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(F(\mathcal{Y}), F(\tilde{\mathcal{Y}}))$ in terms of $d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}, \tilde{\mathcal{Y}})$.

In the first place, we need to elaborate the meaning of $F(\mathcal{Y})$ as a controlled rough path, which consists of the actual path in U along with its $N-1$ derivative paths. The actual path, denoted as Z_t^0, should apparently be given by $Z_t^0 \triangleq F^0(Y_t^0)$. To motivate the derivative paths, we use the Taylor expansion of F^0 :

$$F^0(Y_t^0) - F^0(Y_s^0) \doteq \sum_{j=1}^{N-1} \frac{1}{j!} F^j(Y_s^0)((Y_{s,t}^0)^{\boxtimes j}),$$

where $\doteq$ means being equal up to a term of regularity $|t-s|^{N\alpha}$. Note that a term of such regularity is regarded as a remainder in the expansion of Z^0. To proceed further, we adopt the convention that $Y_t^i \in \mathcal{L}(V^{\otimes i}; W)$ is extended to a linear mapping from $T^{(N)}(V)$ to $T^{(N)}(W)$ by setting $Y_t^i(\xi) \triangleq 0$ if $\xi \in V^{\otimes j}$ with $j \neq i$. Using the expansion of Y^0, we have

$$(Y_{s,t}^0)^{\boxtimes j} \doteq \Big(\big(\sum_{i=1}^{N-1} Y_s^i\big)\mathbf{X}_{s,t}\Big)^{\boxtimes j} = \big(\sum_{i=1}^{N-1} Y_s^i\big)^{\boxtimes j}(\mathbf{X}_{s,t}^{\boxtimes j}).$$

Since $\mathbf{X}$ is a geometric rough path, $\mathbf{X}_{s,t}$ takes values in the free nilpotent group $G^{(N)}(V)$. By using (2.2), it is not hard to see that

$$\delta_j(\mathbf{X}_{s,t}) \doteq \mathbf{X}_{s,t} \boxtimes \cdots \boxtimes \mathbf{X}_{s,t}.$$

As a result, we have

$$F^0(Y_t^0) \doteq \sum_{j=0}^{N-1} \frac{1}{j!} F^j(Y_s^0)\Big(\big(\sum_{i=1}^{N-1} Y_s^i\big)^{\boxtimes j}(\delta_j(\mathbf{X}_{s,t}))\Big)$$

$$= F^0(Y_s^0) + \sum_{r=1}^{N-1}\sum_{j=1}^{N-1}\frac{1}{j!}F^j(Y_s^0)(\sum_{i_1+\cdots+i_j=r}(Y_s^{i_1}\boxtimes\cdots\boxtimes Y_s^{i_j})(\delta_j(X_{s,t}^r))),$$

where the summation $\sum_{i_1+\cdots+i_j=r}$ is taken over all $1 \leqslant i_1, \cdots, i_j \leqslant N-1$. It is then clear that the derivative paths $Z^1, \cdots, Z^{N-1}$ should be defined by

$$Z_s^r \triangleq \sum_{j=1}^{N-1}\frac{1}{j!}F^j(Y_s^0)(\sum_{i_1+\cdots+i_j=r}(Y_s^{i_1}\boxtimes\cdots\boxtimes Y_s^{i_j})\circ\delta_j|_{V^{\otimes j}}). \tag{4.2}$$

Note that the requirement $i_1+\cdots+i_j = r$ together with $i_1, \cdots, i_j \geqslant 1$ mean that the sum $\sum_{j=1}^{N-1}$ is in reality a sum $\sum_{j=1}^{r}$ as the terms from $j = r+1$ to $j = N-1$ are zero. We have left it as $\sum_{j=1}^{N-1}$ for the convenience of interchanging summations later on. Note that $Z_s^r \in \mathcal{L}(V^{\otimes r}; U)$.

To prove that $\mathcal{Z} = \left(Z^0, \cdots, Z^{N-1}\right)$ is controlled by $\mathbf{X}$, by Definition 2.3 we need to show that

$$Z_{s,t}^r \doteq Z_s^{r+1}X_{s,t}^1 + \cdots + Z_s^{N-1}X_{s,t}^{N-1-r} \tag{4.3}$$

for each $1 \leqslant r \leqslant N-1$, where in this case $\doteq$ means being equal up to a term of regularity $|t-s|^{(N-r)\alpha}$. The main challenge (and essence) of proving (4.3) is algebraic rather than analytic. In particular, this relies on a key algebraic lemma which we now motivate.

First of all, there is nothing to prove when $r = 0$, since the definition of Z^r guarantees the desired regularity property in this case. For $1 \leqslant r \leqslant N-1$, let $\xi \in V^{\otimes r}$ be a generic element. To simplify the notation in the computation below, we set

$$\eta_t^j \triangleq \sum_{i_1+\cdots+i_j=r}(Y_t^{i_1}\boxtimes\cdots\boxtimes Y_t^{i_j})\circ\delta_j(\xi) \in W^{\boxtimes j}. \tag{4.4}$$

Then we can write

$$\begin{aligned} Z_t^r(\xi) &= \sum_{j=1}^{N-1}\frac{1}{j!}F^j(Y_t^0)(\eta_t^j) \\ &\doteq \sum_{j=1}^{N-1}\frac{1}{j!}(\sum_{l=0}^{N-1-j}\frac{1}{l!}F^{j+l}(Y_s^0)((Y_{s,t}^0)^{\boxtimes l}\boxtimes\eta_t^j)) \\ &= \sum_{j=1}^{N-1}\frac{1}{j!}\sum_{k=j}^{N-1}\frac{1}{(k-j)!}F^k(Y_s^0)((Y_{s,t}^0)^{\boxtimes(k-j)}\boxtimes\eta_t^j) \\ &= \sum_{k=1}^{N-1}\sum_{j=1}^{k}\frac{1}{j!(k-j)!}F^k(Y_s^0)((Y_{s,t}^0)^{\boxtimes(k-j)}\boxtimes\eta_t^j). \end{aligned} \tag{4.5}$$

Let us define

$$\hat{\eta}_t^j = \sum_{i_1+\cdots+i_j=r}\Big(\sum_{l_1\geqslant i_1,\cdots,l_j\geqslant i_j}Y_s^{l_1}X_{s,t}^{l_1-i_1}\boxtimes\cdots\boxtimes Y_s^{l_j}X_{s,t}^{l_j-i_j}\Big)\circ\delta_j(\xi) \tag{4.6}$$

and

$$\hat{Y}_{s,t}^0 = \sum_{m=1}^{N-1}Y_s^mX_{s,t}^m, \tag{4.7}$$

respectively. It follows that

$$Z_t^r(\xi) \doteq \sum_{k=1}^{N-1}\sum_{j=1}^{k} \frac{1}{j!(k-j)!} F^k(Y_s^0)((Y_{s,t}^0)^{\boxtimes(k-j)} \boxtimes \eta_t^j) \tag{4.8}$$

$$\doteq \sum_{k=1}^{N-1}\sum_{j=1}^{k} \frac{1}{j!(k-j)!} F^k(Y_s^0)((\hat{Y}_{s,t}^0)^{\boxtimes(k-j)} \boxtimes \hat{\eta}_t^j).$$

On the other hand, we have

$$(Z_s^r + Z_s^{r+1}X_{s,t}^1 + \cdots + Z_s^{N-1}X_{s,t}^{N-1-r})(\xi)$$

$$= \sum_{l=r}^{N-1}(\sum_{k=1}^{N-1}\sum_{i_1+\cdots+i_k=l} \frac{1}{k!}F^k(Y_s^0)(Y_s^{i_1} \boxtimes \cdots \boxtimes Y_s^{i_k}) \circ \delta_k(\mathbf{X}_{s,t} \otimes \xi)).$$

Consequently, to prove $\mathcal{Z}$ is a controlled path, it boils down to showing that

$$\sum_{k=1}^{N-1}\sum_{j=1}^{k} \frac{1}{j!(k-j)!} F^k(Y_s^0)((\hat{Y}_{s,t}^0)^{\boxtimes(k-j)} \boxtimes \hat{\eta}_t^j) \tag{4.9}$$

$$\doteq \sum_{k=1}^{N-1}\sum_{l=1}^{N-1} \frac{F^k(Y_s^0)}{k!}(\sum_{i_1+\cdots+i_k=l} (Y_s^{i_1} \boxtimes \cdots \boxtimes Y_s^{i_k}) \circ \delta_k(\mathbf{X}_{s,t} \otimes \xi)).$$

Here an important point is that $F^k(Y_s^0)$ is a symmetric functional over $W^{\boxtimes k}$. To respect the underlying symmetry, let $S_k : W^{\boxtimes k} \to W^{\boxtimes k}$ be the symmetrization operator on homogeneous k-tensors, and let K be its kernel. We introduce the notation $\xi \overset{s}{=} \eta$ to mean that $\xi - \eta \in K$. Using the symmetry of $F^k(Y_s^0)$, it remains to establish the following algebraic lemma.

Lemma 4.1. *For each* $1 \leqslant k \leqslant N-1$*, we have*

$$\sum_{j=1}^{k} \frac{1}{j!(k-j)!} \left(\hat{Y}_{s,t}^0\right)^{\boxtimes(k-j)} \boxtimes \hat{\eta}_t^j \tag{4.10}$$

$$\overset{s}{=} \frac{1}{k!}\sum_{l=r}^{N-1}\sum_{i_1+\cdots+i_k=l} \left(Y_s^{i_1} \boxtimes \cdots \boxtimes Y_s^{i_k}\right) \circ \delta_k\left(\mathbf{X}_{s,t} \otimes \xi\right) + \Delta_{3;s,t}^k,$$

where we recall that $\hat{\eta}_t^j$ *is defined in* (4.6)*,* $\hat{Y}_{s,t}^0$ *is defined in* (4.7) *and*

$$\Delta_{3;s,t}^k \triangleq \frac{1}{k!} \sum_{\substack{1 \leqslant i_1,\cdots,i_k \leqslant N-1 \\ i_1+\cdots+i_k \geqslant N}} Y_s^{i_1} \boxtimes \cdots \boxtimes Y_s^{i_k}((\mathbf{X}_{s,t} \boxtimes \cdots \boxtimes \mathbf{X}_{s,t}) * \delta_k(\xi)). \tag{4.11}$$

Remark 4.1. The role of the term $\Delta_{3;s,t}^k$ is to compensate the difference between $\delta_k(\mathbf{X}_{s,t})$ and $\mathbf{X}_{s,t}^{\boxtimes k}$ (cf. (2.2)), which arises from tensor truncation.

Proof. By the linearity for both sides of (4.10), it is enough to consider the case $\xi = v_1 \otimes \cdots \otimes v_r$ when $v_i \in V$ for all $1 \leqslant i \leqslant r$. Recall that

$$\hat{\eta}_t^j = \sum_{i_1+\cdots+i_j=r} \left(\sum_{l_1 \geqslant i_1,\cdots,l_j \geqslant i_j} Y_s^{l_1}X_{s,t}^{l_1-i_1} \boxtimes \cdots \boxtimes Y_s^{l_j}X_{s,t}^{l_j-i_j}\right) \circ \delta_j\left(v_1 \otimes \cdots \otimes v_r\right), \tag{4.12}$$

and

$$\delta_j(v_1 \otimes \cdots \otimes v_r) = \sum_{(I_\alpha)} \xi|_{I_1} \boxtimes \cdots \boxtimes \xi|_{I_j}, \tag{4.13}$$

where the summation is taken over all disjoint subsets $I_1, \cdots, I_j$ such that $\cup_{\alpha=1}^j I_\alpha = \{1, \cdots, r\}$. Using the above formula for δ_j, equation (4.12) becomes

$$\begin{aligned} &\sum_{i_1+\cdots+i_j=r} \Big(\sum_{l_1 \geqslant i_1, \cdots, l_j \geqslant i_j} Y_s^{l_1} X_{s,t}^{l_1-i_1} \boxtimes \cdots \boxtimes Y_s^{l_j} X_{s,t}^{l_j-i_j} \Big) \sum_{(I_\alpha)} \xi|_{I_1} \boxtimes \cdots \boxtimes \xi|_{I_j} \\ &= \sum_{i_1+\cdots+i_j=r} \sum_{(I_\alpha)} \sum_{l_1 \geqslant i_1, \cdots, l_j \geqslant i_j} Y_s^{l_1}(X_{s,t}^{l_1-i_1} \otimes \xi|_{I_1}) \boxtimes \cdots \boxtimes Y_s^{l_j}(X_{s,t}^{l_j-i_j} \otimes \xi|_{I_j}). \end{aligned} \tag{4.14}$$

Since Y_s^l acts on $V^{\otimes l}$ and sends on all other elements to zero, we know that

$$Y_s^l(X_{s,t}^{l-i} \otimes \xi|_I) = 0 \qquad \text{if } |I| \neq i.$$

Therefore, the summation $\sum_{(I_\alpha)}$ in (4.14) becomes a summation over all partitions $(I_\alpha)_{\alpha=1}^j$ of $\{1, \cdots r\}$ such that $|I_\alpha| = i_\alpha$ for all α. As a result, we can write

$$\sum_{i_1+\cdots+i_j=r} \sum_{(I_\alpha)_{\alpha=1}^j : |I_\alpha| = i_\alpha \, \forall \alpha} = \sum_{(I_\alpha)_{\alpha=1}^j : |I_\alpha| \geqslant 1 \, \forall \alpha},$$

where the right hand side denotes the summation over all partitions $(I_\alpha)_{\alpha=1}^j$ of $\{1, \cdots, r\}$ such that $|I_\alpha| \geqslant 1$ for each α. Moreover, as $Y_s^l(X_{s,t}^q \otimes \xi|_I) = 0$ unless $q = l - |I|$, we have

$$Y_s^l(X_{s,t}^q \otimes \xi|_I) = Y_s^l(\mathbf{X}_{s,t} \otimes \xi|_I).$$

Note finally that as $\mathbf{X}_{s,t} \otimes \xi|_I$ has degree at least $|I|$,

$$Y_s^l(\mathbf{X}_{s,t} \otimes \xi|_I) = 0 \qquad \text{if } l < |I|.$$

Therefore, for each $1 \leqslant \alpha \leqslant j$ the summation $\sum_{l_\alpha \geqslant |I_\alpha|}$ can be replaced by the unrestricted sum $\sum_{l_\alpha=1}^{N-1}$.

Taking into account the above considerations, equation (4.14) now becomes

$$\hat{\eta}_t^j = \sum_{(I_\alpha)_{\alpha=1}^j : I_\alpha \neq \emptyset \, \forall \alpha} \sum_{l_1, \cdots, l_j=1}^{N-1} Y_s^{l_1}(\mathbf{X}_{s,t} \otimes \xi|_{I_1}) \boxtimes \cdots \boxtimes Y_s^{l_j}(\mathbf{X}_{s,t} \otimes \xi|_{I_j}).$$

It follows that

$$\begin{aligned} &\sum_{j=1}^k \frac{1}{j!(k-j)!} (\hat{Y}_{s,t}^0)^{\boxtimes(k-j)} \boxtimes \hat{\eta}_t^j \\ &= \sum_{j=1}^k \frac{1}{j!(k-j)!} \sum_{m_1, \cdots, m_{k-j}=1}^{N-1} Y_s^{m_1} \mathbf{X}_{s,t} \boxtimes \cdots \boxtimes Y_s^{m_{k-j}} \mathbf{X}_{s,t} \\ &\quad \boxtimes \sum_{(I_\alpha)_{\alpha=1}^j : I_\alpha \neq \emptyset \, \forall \alpha} \sum_{l_1, \cdots, l_j=1}^{N-1} Y_s^{l_1}(\mathbf{X}_{s,t} \otimes \xi|_{I_1}) \boxtimes \cdots \boxtimes Y_s^{l_j}(\mathbf{X}_{s,t} \otimes \xi|_{I_j}) \\ &= \sum_{j=1}^k \frac{1}{j!(k-j)!} \sum_{h_1, \cdots, h_k=1}^{N-1} Y_s^{h_1} \mathbf{X}_{s,t} \boxtimes \cdots \boxtimes Y_s^{h_{k-j}} \mathbf{X}_{s,t} \end{aligned} \tag{4.15}$$

$$\boxtimes \sum_{(I_\alpha)_{\alpha=1}^j : I_\alpha \neq \emptyset \, \forall \alpha} Y_s^{h_{k-j+1}}(\mathbf{X}_{s,t} \otimes \xi|_{I_1}) \boxtimes \cdots \boxtimes Y_s^{h_k}(\mathbf{X}_{s,t} \otimes \xi|_{I_j}).$$

As the next observation, let $(H_i)_{i=1}^k$ be a partition of $\{1, \cdots, r\}$. If there exist $\beta_1 < \cdots < \beta_j$ such that $H_{\beta_i} = I_i$ for $1 \leqslant i \leqslant j$ and $H_i = \emptyset$ for $i \notin \{\beta_1, \cdots, \beta_r\}$, then

$$\sum_{h_1,\cdots,h_k=1}^{N-1} Y_s^{h_1}\mathbf{X}_{s,t} \boxtimes \cdots \boxtimes Y_s^{h_{k-j}}\mathbf{X}_{s,t} \boxtimes Y_s^{h_{k-j+1}}(\mathbf{X}_{s,t} \otimes \xi|_{I_1}) \boxtimes \cdots \boxtimes Y_s^{h_k}(\mathbf{X}_{s,t} \otimes \xi|_{I_j})$$
$$\overset{\mathrm{s}}{=} \sum_{h_1,\cdots,h_k=1}^{N-1} Y_s^{h_1}(\mathbf{X}_{s,t} \otimes \xi|_{H_1}) \boxtimes \cdots \boxtimes Y_s^{h_k}(\mathbf{X}_{s,t} \otimes \xi|_{H_k}).$$

There are a total of $\binom{k}{j}$ such partitions $(H_i)_{i=1}^k$ for each given j-tuple $(I_1, \cdots, I_j)$. As a result, we have

$$\sum_{h_1,\cdots,h_k=1}^{N-1} \sum_{(H_i):\exists \beta_1<\cdots<\beta_j, H_{\beta_i}=I_i \, \forall i} Y_s^{h_1}(\mathbf{X}_{s,t} \otimes \xi|_{H_1}) \boxtimes \cdots \boxtimes Y_s^{h_k}\big(\mathbf{X}_{s,t} \otimes \xi|_{H_k})$$
$$\overset{\mathrm{s}}{=} \sum_{h_1,\cdots,h_k=1}^{N-1} \binom{k}{j} Y_s^{h_1}\mathbf{X}_{s,t} \boxtimes \cdots \boxtimes Y_s^{h_{k-j}}\mathbf{X}_{s,t} \boxtimes Y_s^{h_{k-j+1}}(\mathbf{X}_{s,t} \otimes \xi|_{I_1})$$
$$\boxtimes \cdots \boxtimes Y_s^{h_k}(\mathbf{X}_{s,t} \otimes \xi|_{I_j}).$$

Since the summation

$$\sum_{j=1}^{k} \sum_{(I_\alpha)_{\alpha=1}^j : I_\alpha \neq \emptyset \, \forall \alpha} \sum_{(H_i):\exists \beta_1<\cdots<\beta_j, H_{\beta_i}=I_i \, \forall i}$$

is equivalent to summing over all partitions $(H_i)_{j=1}^k$ of $\{1, \cdots, r\}$, the expression in (4.15) becomes (up to permutation symmetry with respect to $\boxtimes$)

$$\frac{1}{k!} \sum_{h_1,\cdots,h_k=1}^{N-1} \sum_{(H_i)_{i=1}^k:\text{partition of } \{1,\cdots,r\}} Y_s^{h_1}(\mathbf{X}_{s,t} \otimes \xi|_{H_1}) \boxtimes \cdots \boxtimes Y_s^{h_k}(\mathbf{X}_{s,t} \otimes \xi|_{H_k}) \tag{4.16}$$
$$= \frac{1}{k!} \sum_{h_1,\cdots,h_k=1}^{N-1} Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}((\mathbf{X}_{s,t} \boxtimes \cdots \boxtimes \mathbf{X}_{s,t}) * (\sum_{(H_i)_{i=1}^k} \xi|_{H_1} \boxtimes \cdots \boxtimes \xi|_{H_k})).$$

By using the formula (4.13) for δ_k again, the expression in (4.16) becomes

$$\frac{1}{k!} \sum_{h_1,\cdots,h_k=1}^{N-1} Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}((\mathbf{X}_{s,t} \boxtimes \cdots \boxtimes \mathbf{X}_{s,t}) * \delta_k(v_1 \otimes \cdots \otimes v_r)) \tag{4.17}$$
$$= \frac{1}{k!} \sum_{l=r}^{N-1} \sum_{h_1+\cdots+h_k=l}^{N-1} Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}((\mathbf{X}_{s,t} \boxtimes \cdots \boxtimes \mathbf{X}_{s,t}) * \delta_k(v_1 \otimes \cdots \otimes v_r)) + \Delta_{3;s,t}^k,$$

where $\Delta_{3;s,t}^k$ is defined to be the difference of the two expressions in (4.17).

Note that when $h_1 + \cdots + h_k \leqslant N$, the operator $Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}$ only acts non-trivially on elements of $\oplus_{l_1+\cdots+l_k \leqslant N} V^{\otimes l_1} \boxtimes \cdots \boxtimes V^{\otimes l_k}$. Since $\mathbf{X}_{s,t} \in G^{(N)}(V)$, according to the shuffle product formula (2.2), for such h_i's we have

$$
\begin{aligned}
& Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}((\mathbf{X}_{s,t} \boxtimes \cdots \boxtimes \mathbf{X}_{s,t}) * \delta_k(v_1 \otimes \cdots \otimes v_r)) \\
&= Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}(\sum_{l_1+\cdots+l_k \leqslant N} (X_{s,t}^{l_1} \boxtimes \cdots \boxtimes X_{s,t}^{l_k}) * \delta_k(v_1 \otimes \cdots \otimes v_r)) \\
&= Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}(\delta_k(\mathbf{X}_{s,t}) * \delta_k(v_1 \otimes \cdots \otimes v_r)).
\end{aligned}
$$

Since δ_k is a $*$-homomorphism, the expression in (4.17) becomes

$$
\frac{1}{k!} \sum_{l=r}^{N-1} \sum_{h_1+\cdots+h_k=l}^{N-1} Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}(\delta_k(\mathbf{X}_{s,t} \otimes v_1 \otimes \cdots \otimes v_r)) + \Delta_{3;s,t}^k,
$$

which is precisely the right hand side of (4.10). □

Having the above algebraic considerations, we can now prove the main result of this section. For the need in the study of RDEs in the next section, we also establish a continuity estimate for Lipschitz transformations.

Theorem 4.1. (i) *[Stability] Let $\mathcal{Y}$ be a controlled rough path over* $[0, T]$ *with respect to* $\mathbf{X}$. *Let $F = (F^0, F^1, \cdots, F^N)$ be a γ-Lipschitz function with $\gamma \in (N, N+1]$. Then the path $\mathcal{Z} = F(\mathcal{Y})$ as defined in* (4.2) *is a path controlled by* $\mathbf{X}$ *in the sense of* Definition 2.3. *In addition, we have*

$$
\|F(\mathcal{Y})\|_{\mathbf{X};\alpha} \leqslant \|F\|_{Lip\text{-}N} \cdot M(T, \max_{1\leqslant i\leqslant N-1} |Y_0^i|, \|\mathcal{Y}\|_{\mathbf{X};\alpha}, \|\mathbf{X}\|_\alpha) \tag{4.18}
$$

(ii) *[Continuity estimate] Let $\mathcal{Y}$ and $\tilde{\mathcal{Y}}$ be paths over* $[0, T]$ *controlled by* $\mathbf{X}$ *and* $\tilde{\mathbf{X}}$ *respectively, and let F be* $(N+1)$*-Lipschitz. Then we have*

$$
\begin{aligned}
& d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(F(\mathcal{Y}), F(\tilde{\mathcal{Y}})) \\
&\leqslant \|F\|_{Lip\text{-}(N+1)} \cdot M(T, \max_{1\leqslant i\leqslant N-1} |Y_0^i|, \max_{1\leqslant i\leqslant N-1} |\tilde{Y}_0^i|, \|\mathcal{Y}\|_{\mathbf{X};\alpha}, \|\tilde{\mathcal{Y}}\|_{\tilde{\mathbf{X}};\alpha}, \|\mathbf{X}\|_\alpha, \|\tilde{\mathbf{X}}\|_\alpha) \\
&\quad \times (\max_{0\leqslant i\leqslant N-1} |Y_0^i - \tilde{Y}_0^i| + d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}, \tilde{\mathcal{Y}}) + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})).
\end{aligned} \tag{4.19}
$$

In both parts, $M(\cdots)$ denotes a universal function that is continuous and increasing in every variable.

Proof. To ease our discussion, we use the notation "$\lesssim$" to denote an estimate up to a continuous increasing function M in T, $\max_{1\leqslant i\leqslant N-1} |Y_0^i|$, $\max_{1\leqslant i\leqslant N-1} |\tilde{Y}_0^i|$, $\|\mathcal{Y}\|_{\mathbf{X};\alpha}$, $\|\tilde{\mathcal{Y}}\|_{\tilde{\mathbf{X}};\alpha}$, $\|\mathbf{X}\|_\alpha$, $\|\tilde{\mathbf{X}}\|_\alpha$, which may differ from line to line. We closely follow the notation used earlier in the algebraic considerations. In particular, in order to prove the theorem, essentially we need to keep track of the remainder from each of the notions "$\doteq$" appearing earlier.

First of all, as seen before, we can write

$$
Z_t^r(\xi) = \sum_{k=0}^{N-1} F^k(Y_s^0)(C_k + \Delta_{2;s,t}^k) + \Delta_{1;s,t}, \quad EZ_{s,t}^r(\xi) = \sum_{k=0}^{N-1} F^k(Y_s^0)(D_k - \Delta_{3;s,t}^k),
$$

Here

$$
EZ_{s,t}^r(\xi) \triangleq (Z_s^r + Z_s^{r+1} X_{s,t}^1 + \cdots + Z_s^{N-1} X_{s,t}^{N-1-r})(\xi),
$$

C_k, D_k are defined by the left and right hand sides of the algebraic identity (4.10) respectively. The remainders $\Delta_{1;s,t}$, $\Delta_{2;s,t}^k$ are associated with the notions "$\doteq$" appearing earlier and

$\Delta^k_{3;,s,t}$ is defined by (4.11). To be precise, they are defined by the following equations.

(i) (cf. (4.5) and Taylor's theorem with integral form remainder)

$$\Delta_{1;s,t} \triangleq \sum_{j=1}^{r}\int_0^1 \frac{(1-\theta)^{N-1-j}}{j!(N-1-j)!}F^N(Y^0_s+\theta Y^0_{s,t})((Y^0_{s,t})^{\boxtimes(N-j)}\boxtimes\eta^j_t)\mathrm{d}\theta, \tag{4.20}$$

where we recall that $\sum_{j=1}^r = \sum_{j=1}^{N-1}$.

(ii) (cf. (4.8))

$$\Delta^k_{2;s,t} \triangleq \sum_{j=1}^{r}\frac{1}{j!(k-j)!}\sum_{i_1+\cdots+i_j=r}((Y^0_{s,t})^{\boxtimes(k-j)}\boxtimes(Y^{i_1}_t\boxtimes\cdots\boxtimes Y^{i_j}_t(\delta_j(\xi)))$$
$$-(Y^1_sX^1_{s,t}+\cdots+Y^{N-1}_sX^{N-1}_{s,t})^{\boxtimes(k-j)}\boxtimes(EY^{i_1}_{s,t}\boxtimes\cdots\boxtimes EY^{i_j}_{s,t}(\delta_j(\xi)))),$$

where

$$EY^i_{s,t} \triangleq Y^i_s + Y^{i+1}_sX^1_{s,t}+\cdots+Y^{N-1}_sX^{N-1-i}_{s,t}.$$

According to Lemma 4.1, we have

$$F^k(Y^0_s)(C_k) = F^k(Y^0_s)(D_k).$$

It follows that

$$\mathcal{R}\mathcal{Z}^r_{s,t} = \sum_{k=1}^{N-1}F^k(Y^0_s)\left(\Delta^k_{2;s,t}+\Delta^k_{3;s,t}\right)+\Delta_{1;s,t}. \tag{4.21}$$

Similar definitions and identities hold for the tilde-quantities. We need to estimate the regularity of the Δ's.

As a standard way, we frequently use the simple inequality

$$|ab-\tilde{a}\tilde{b}| \leqslant |a-\tilde{a}|\cdot|b| + |\tilde{a}|\cdot|b-\tilde{b}|. \tag{4.22}$$

Also note that $|F^k(Y^0_s)| \lesssim \|F\|_{\text{Lip-}N}$. As a consequence of Lemma 3.1 (without the presence of $\tilde{\mathcal{Y}}$), we have

$$|Y^0_{s,t}| \lesssim |t-s|^\alpha,\ |Y^i_t| \lesssim 1, \quad \forall i\geqslant 1. \tag{4.23}$$

From these considerations and the expression (4.20) of $\Delta_1; s,t$, we see that

$$|\Delta_{1;s,t}| \lesssim \|F\|_{\text{Lip-}N}\cdot|t-s|^{(N-r)\alpha}. \tag{4.24}$$

For the term $\Delta^k_{2;s,t}$, by forming a telescoping sum it boils down to estimating

$$((Y^0_{s,t})^{\boxtimes(k-j)})\boxtimes(Y^{i_1}_t\boxtimes\cdots\boxtimes Y^{i_j}_t(\delta_j(\xi))-(Y^1_sX^1_{s,t}+\cdots+Y^{N-1}_sX^{N-1}_{s,t})^{\boxtimes(k-j)}$$
$$\boxtimes(EY^{i_1}_{s,t}\boxtimes\cdots\boxtimes EY^{i_j}_{s,t}(\delta_j(\xi))))$$
$$=(Y^0_{s,t}-Y^1_sX^1_{s,t}+\cdots+Y^{N-1}_sX^{N-1}_{s,t})\boxtimes(Y^0_{s,t})^{\boxtimes(k-j-1)}\boxtimes Y^{i_1}_t\cdots\boxtimes Y^{i_j}_t(\delta_j(\xi))+\cdots$$
$$+\left(Y^1_sX^1_{s,t}+\cdots+Y^{N-1}_sX^{N-1}_{s,t}\right)^{\boxtimes(k-j)}\boxtimes\left(EY^{i_1}_{s,t}\boxtimes\cdots\boxtimes(Y^{i_j}_t-EY^{i_j}_{s,t})(\delta_j(\xi))\right). \tag{4.25}$$

Note that

$$\left|Y^0_{s,t}-Y^1_sX^1_{s,t}+\cdots+Y^{N-1}_sX^{N-1}_{s,t}\right| \lesssim |t-s|^{N\alpha}$$

and

$$\left|Y_s^1 X_{s,t}^1 + \cdots + Y_s^{N-1} X_{s,t}^{N-1}\right| = \left|Y_{s,t}^0 - \mathcal{R}\mathcal{Y}_{s,t}^0\right| \lesssim |t-s|^\alpha .$$

In addition, for each $i \geqslant 1$ we have

$$\left|EY_{s,t}^i\right| = \left|Y_s^i + Y_s^{i+1} X_{s,t}^1 + \cdots + Y_s^{N-1} X_{s,t}^{N-1-i}\right| = \left|Y_t^i - \mathcal{R}\mathcal{Y}_{s,t}^i\right| \lesssim 1,$$

and for each $i \leqslant r$ we have

$$\left|Y_t^i - EY_{s,t}^i\right| \lesssim |t-s|^{(N-r)\alpha} .$$

Consequently, we see that

$$|\Delta_{2;s,t}^k| \lesssim |t-s|^{(N-r)\alpha}. \tag{4.26}$$

We now estimate

$$\begin{aligned}\Delta_{3;s,t}^k &= \frac{1}{k!} \sum_{h_1+\cdots+h_k \geqslant N} Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}((\mathbf{X}_{s,t} \boxtimes \cdots \boxtimes \mathbf{X}_{s,t}) * \delta_k(\xi)) \\ &= \frac{1}{k!} \sum_{h_1+\cdots+h_k \geqslant N} Y_s^{h_1} \boxtimes \cdots \boxtimes Y_s^{h_k}((\sum_{l_1,\cdots,l_k=1}^{N} X_{s,t}^{l_1} \boxtimes \cdots \boxtimes X_{s,t}^{l_k}) * \delta_k(\xi)).\end{aligned}$$

Since $h_1 + \cdots + h_k \geqslant N$ and $\xi \in V^{\otimes r}$, the only non-zero terms are the ones when $l_1 + \cdots + l_k \geqslant N - r$. In this case, we see that

$$|X_{s,t}^{l_1} \boxtimes \cdots \boxtimes X_{s,t}^{l_k}| \lesssim (t-s)^{(N-r)\alpha}.$$

Therefore,

$$|\Delta_{3;s,t}^k| \lesssim \sum_{h_1+\cdots+h_k \geqslant N} |Y_s^{h_1}| \cdots |Y_s^{h_k}| \cdot (t-s)^{(N-r)\alpha} \lesssim (t-s)^{(N-r)\alpha}.$$

We particularly point out that the constant hidden within "$\lesssim$" is independent of Y_0^0. This will be important for RDE considerations later on. From (4.21), (4.24) and (4.26) we conclude that $\mathcal{Z}$ is controlled by $\mathbf{X}$ and the estimate (4.18) follows.

To prove the second part the theorem, we need to estimate

$$\begin{aligned}\mathcal{R}\mathcal{Z}_{s,t}^r - \mathcal{R}\tilde{\mathcal{Z}}_{s,t}^r &= \sum_{k=1}^{N-1} (F^k(Y_s^0)(\Delta_{2;s,t}^k) - F^k(\tilde{Y}_s^0)(\tilde{\Delta}_{2;s,t}^k)) \\ &+ \sum_{k=1}^{N-1} (F^k(Y_s^0)(\Delta_{3;s,t}^k) - F^k(\tilde{Y}_s^0)(\tilde{\Delta}_{3;s,t}^k)) + (\Delta_{1;s,t} - \tilde{\Delta}_{1;s,t}).\end{aligned}$$

For this purpose, let us introduce

$$D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}}) \triangleq \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}) + d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}, \tilde{\mathcal{Y}}) + \sum_{i=0}^{N-1} |Y_0^i - \tilde{Y}_0^i|.$$

Now it remains to establish the following set of estimates (for $1 \leqslant k \leqslant N-1$):

$$|F^k(Y_s^0) - F^k(\tilde{Y}_s^0)| \lesssim \|F\|_{\text{Lip-}N} \cdot D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}}),$$

$$|\Delta_{1;s,t} - \tilde{\Delta}_{1;s,t}| \lesssim \|F\|_{\text{Lip-}(N+1)} \cdot D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}}) \cdot |t-s|^{(N-r)\alpha}, \tag{4.27}$$

$$|\Delta^k_{2;s,t} - \tilde{\Delta}^k_{2;s,t}| \lesssim D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}}) \cdot |t-s|^{(N-r)\alpha}, \tag{4.28}$$

$$|\Delta^k_{3;s,t} - \tilde{\Delta}^k_{3;s,t}| \lesssim D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}}) \cdot |t-s|^{(N-r)\alpha}. \tag{4.29}$$

To see the first inequality, first note that

$$\begin{aligned}|F^k(Y^0_s) - F^k(\tilde{Y}^0_s)| &= \Big| \int_0^1 F^{k+1}(Y^0_s + \theta(Y^0_s - \tilde{Y}^0_s))(Y^0_s - \tilde{Y}^0_s)\mathrm{d}\theta \Big| \\ &\lesssim \|F\|_{\text{Lip-}N} \cdot \left(\left|Y^0_0 - \tilde{Y}^0_0\right| + \left\|Y^0 - \tilde{Y}^0\right\|_\alpha\right).\end{aligned}$$

The inequality then follows from Lemma 3.1.

For the inequality (4.27), according to its expression (4.20), it suffices to estimate

$$F^N(Y^0_s + \theta Y^0_{s,t})(\eta^j_t \boxtimes (Y^0_{s,t})^{\boxtimes(N-j)}) - F^N(\tilde{Y}^0_s + \theta\tilde{Y}^0_{s,t})(\tilde{\eta}^j_t \boxtimes (\tilde{Y}^0_{s,t})^{\boxtimes(N-j)}),$$

Recall from the definition (4.4) of η^j_t that

$$\eta^j_t - \tilde{\eta}^j_t = \sum_{i_1+\cdots+i_j=r} (Y^{i_1}_t \boxtimes \cdots \boxtimes Y^{i_j}_t - \tilde{Y}^{i_1}_t \boxtimes \cdots \boxtimes \tilde{Y}^{i_j}_t) \circ \delta_j(\xi)\,.$$

For each $1 \leqslant i \leqslant N-1$, we have

$$\begin{aligned}\left|Y^i_t - \tilde{Y}^i_t\right| &\leqslant \left|Y^i_0 - \tilde{Y}^i_0\right| + \left|(Y^i_t - Y^i_0) - (\tilde{Y}^i_t - \tilde{Y}^i_0)\right| \\ &\leqslant \left|Y^i_0 - \tilde{Y}^i_0\right| + \|Y^i - \tilde{Y}^i\|_\alpha T^\alpha \\ &\lesssim D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}}).\end{aligned} \tag{4.30}$$

Since $|Y^i_t| \lesssim 1$, $|\tilde{Y}^i_t| \lesssim 1$, it follows that

$$\left|\eta^j_t - \tilde{\eta}^j_t\right| \lesssim D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}}).$$

On the other hand, according to Lemma 3.1 we have

$$\left|Y^0_{s,t} - \tilde{Y}^0_{s,t}\right| \leqslant \|Y^0 - \tilde{Y}^0\|_\alpha\,(t-s)^\alpha \lesssim D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}})\,(t-s)^\alpha\,.$$

Therefore, we see that

$$\begin{aligned}&\left|F^N(Y^0_s + \theta Y^0_{s,t}) - F^N(\tilde{Y}^0_s + \theta\tilde{Y}^0_{s,t})\right| \leqslant \|F\|_{\text{Lip-}(N+1)}\left|Y^0_s + \theta Y^0_{s,t} - (\tilde{Y}^0_s + \theta\tilde{Y}^0_{s,t})\right| \\ &\leqslant \|F\|_{\text{Lip-}(N+1)}\left|Y^0_s + \theta Y^0_{s,t} - (\tilde{Y}^0_s + \theta\tilde{Y}^0_{s,t})\right| \\ &\leqslant \|F\|_{\text{Lip-}(N+1)}\left(\left|Y^0_0 - \tilde{Y}^0_0\right| + \left|Y^0_{s,t} - \tilde{Y}^0_{s,t}\right| + \left|Y^0_{0,s} - \tilde{Y}^0_{0,s}\right|\right) \\ &\leqslant \|F\|_{\text{Lip-}(N+1)}\left(\left|Y^0_0 - \tilde{Y}^0_0\right| + 2\|Y^0 - \tilde{Y}^0\|_\alpha T^\alpha\right) \\ &\lesssim \|F\|_{\text{Lip-}(N+1)} D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}}).\end{aligned}$$

The inequality (4.27) thus follows (the regularity $|t-s|^{(N-r)\alpha}$ comes from the fact that $j \leqslant r$ in the summation (4.20)).

For the inequality (4.28), to estimate $\Delta^k_{2;s,t} - \tilde{\Delta}^k_{2;s,t}$ we write this difference in the form of a telescoping sum that is similar to (4.25) but also with the tilde-quantities. We already have the required estimates for $Y^0_{s,t} - \tilde{Y}^0_{s,t}$ and $Y^i_t - \tilde{Y}^i_t$ when analyzing $\Delta_{1;s,t}$. We also have

$$\left|(Y^1_s X^1_{s,t} + \cdots + Y^{N-1}_s X^{N-1}_{s,t}) - (\tilde{Y}^1_s \tilde{X}^1_{s,t} + \cdots + \tilde{Y}^{N-1}_s \tilde{X}^{N-1}_{s,t})\right| \leqslant \left|Y^0_{s,t} - \tilde{Y}^0_{s,t}\right| + \left|\mathcal{R}\mathcal{Y}^0_{s,t} - \mathcal{R}\tilde{\mathcal{Y}}^0_{s,t}\right|$$

$\lesssim D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}})\left|t-s\right|^{\alpha},$

and

$$\begin{aligned}\left|EY^i_{s,t} - E\tilde{Y}^i_{s,t}\right| \leqslant& \left|Y^i_t - EY^i_{s,t} - (\tilde{Y}^i_t - E\tilde{Y}^i_{s,t})\right| + \left|Y^i_t - \tilde{Y}^i_t\right| \\ =&\left|\mathcal{R}\mathcal{Y}^i_{s,t} - \mathcal{R}\tilde{\mathcal{Y}}^i_{s,t}\right| + \left|Y^i_t - \tilde{Y}^i_t\right| \\ \lesssim& D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}}).\end{aligned}$$

As a result, the desired estimate for $\Delta^k_{2;s,t} - \tilde{\Delta}^k_{2;s,t}$ follows.

For the last inequality (4.29), we have

$$\begin{aligned}\Delta^k_{3;s,t} - \tilde{\Delta}^k_{3;s,t} = \frac{1}{k!}\sum_{h_1+\cdots+h_k\geqslant N} & Y^{h_1}_s \boxtimes \cdots \boxtimes Y^{h_k}_s\Big(\Big(\sum_{l_1,\cdots,l_k=1}^N X^{l_1}_{s,t} \boxtimes \cdots \boxtimes X^{l_k}_{s,t}\Big) * \delta_k(\xi)\Big) \\ & - \tilde{Y}^{h_1}_s \boxtimes \cdots \boxtimes \tilde{Y}^{h_k}_s\Big(\Big(\sum_{l_1,\cdots,l_k=1}^N \tilde{X}^{l_1}_{s,t} \boxtimes \cdots \boxtimes \tilde{X}^{l_k}_{s,t}\Big) * \delta_k(\xi)\Big).\end{aligned}$$

Note that $\left|Y^i_t\right| \lesssim 1$, $\left|\tilde{Y}^i_t\right| \lesssim 1$ and $\left|Y^i_t - \tilde{Y}^i_t\right| \lesssim D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}})$. We also have $|X^l_{s,t}| \lesssim (t-s)^{l\alpha}$, $|\tilde{X}^l_{s,t}| \lesssim (t-s)^{l\alpha}$ and $|X^l_{s,t} - \tilde{X}^l_{s,t}| \lesssim D(\mathbf{X}, \mathcal{Y}; \tilde{\mathbf{X}}, \tilde{\mathcal{Y}})(t-s)^{l\alpha}$. Therefore, we obtain the desired inequality (4.29).

Now the proof of the theorem is complete. □

5. Continuity of rough integrals

In this section, we study the integral $\int \mathcal{Z} d\mathbf{X}$ as a controlled rough path and establish a continuity estimate. Recall that $\alpha \in (0, 1/2]$ and $N \triangleq [\alpha]$, so that $\frac{1}{N+1} < \alpha \leqslant \frac{1}{N}$. Given a partition $\mathcal{P} : s = t_0 < t_1 < \cdots < t_n = t$, we set

$$|\mathcal{P}| \triangleq \max_{0\leqslant i\leqslant n-1}(t_{i+1} - t_i).$$

All paths below are defined on $[0, T]$.

Proposition 5.1. (i) *Let* $\mathbf{X}$ *be an* α*-Hölder geometric rough path over V, and let* $\mathcal{Z}$ *be an* α*-Hölder controlled rough path over* $\mathcal{L}(V; U)$ *with respect to* $\mathbf{X}$*. Then the following limit exists:*

$$\int_s^t ZdX \triangleq \lim_{|\mathcal{P}|\to 0}\sum_{t_i\in\mathcal{P}}\sum_{k=1}^N Z^{k-1}_{t_i} X^k_{t_i,t_{i+1}}. \tag{5.1}$$

In addition, if we define $\mathcal{I} \triangleq \int_0^\cdot \mathcal{Z} d\mathbf{X} \triangleq (I^0, I^1, \cdots, I^{N-1})$ *by*

$$I^0_t \triangleq \int_0^t ZdX,\ I^1_t \triangleq Z^0_t, \cdots,\ I^{N-1}_t \triangleq Z^{N-2}_t, \tag{5.2}$$

then $\mathcal{I}$ *is an* α*-Hölder controlled rough path with respect to* $\mathbf{X}$ *and the following estimate holds:*

$$\|\mathcal{I}\|_{\mathbf{X};\alpha} \leqslant C_\alpha(T^\alpha(1 + \|\mathbf{X}\|_\alpha)\|\mathcal{Z}\|_{\mathbf{X};\alpha} + |Z^{N-1}_0|\|\mathbf{X}\|_\alpha). \tag{5.3}$$

(ii) *(Continuity estimates) Let* $\mathbf{X}, \tilde{\mathbf{X}}$ *be* α*-Hölder geometric rough paths and let* $\mathcal{Z}, \tilde{\mathcal{Z}}$ *be* *paths controlled by* $\mathbf{X}, \tilde{\mathbf{X}}$ *respectively. We use* $\mathcal{I} = \int_0^\cdot \mathcal{Z} \mathrm{d}\mathbf{X}$ *and* $\tilde{\mathcal{I}} = \int_0^\cdot \tilde{\mathcal{Z}} \mathrm{d}\tilde{\mathbf{X}}$ *to denote the controlled paths obtained by integrating* $\mathcal{Z}$ *and* $\tilde{\mathcal{Z}}$ *respectively. Then the following estimate holds:*

$$d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{I}, \tilde{\mathcal{I}}) \leqslant C_\alpha (T^\alpha \max(1 + \|\mathbf{X}\|_\alpha, \|\tilde{\mathcal{Z}}\|_{\tilde{\mathbf{X}};\alpha}) \cdot (d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Z}, \tilde{\mathcal{Z}}) + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})) + \max(\|\mathbf{X}\|_\alpha, |\tilde{Z}_0^{N-1}|) \cdot (|Z_0^{N-1} - \tilde{Z}_0^{N-1}| + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}))). \tag{5.4}$$

Proof. Let $s < t$ be fixed. Given any partition $\mathcal{P}$ of $[s, t]$, we denote

$$\int_{\mathcal{P}} Z \mathrm{d}X \triangleq \sum_{t_i \in \mathcal{P}} \sum_{k=1}^{N} Z_{t_i}^{k-1} X_{t_i, t_{i+1}}^k$$

Then

$$\begin{aligned} \int_{\mathcal{P}} Z \mathrm{d}X - \int_{\mathcal{P}\setminus\{t_j\}} Z \mathrm{d}X &= \sum_{k=1}^{N} Z_{t_{j-1}}^{k-1} X_{t_{j-1},t_j}^k + \sum_{k=1}^{N} Z_{t_j}^{k-1} X_{t_j,t_{j+1}}^k - \sum_{k=1}^{N} Z_{t_{j-1}}^{k-1} X_{t_{j-1},t_{j+1}}^k \\ &= \sum_{k=1}^{N} Z_{t_{j-1}}^{k-1} X_{t_{j-1},t_j}^k + \sum_{k=1}^{N} Z_{t_j}^{k-1} X_{t_j,t_{j+1}}^k - \sum_{k=1}^{N} \sum_{l=0}^{k} Z_{t_{j-1}}^{k-1} X_{t_{j-1},t_j}^{k-l} \otimes X_{t_j,t_{j+1}}^l \\ &= \sum_{k=1}^{N} Z_{t_j}^{k-1} X_{t_j,t_{j+1}}^k - \sum_{k=1}^{N} \sum_{l=1}^{k} Z_{t_{j-1}}^{k-1} X_{t_{j-1},t_j}^{k-l} \otimes X_{t_j,t_{j+1}}^l. \end{aligned}$$

We claim that the last expression is equal to $\sum_{k=1}^N \mathcal{R}\mathcal{Z}_{t_{j-1},t_j}^{k-1} \otimes X_{t_j,t_{j+1}}^k$. Indeed, by writing

$$Z_{t_j}^{k-1} = Z_{t_{j-1}}^{k-1} + \sum_{r=1}^{N-k} Z_{t_{j-1}}^{k+r-1} X_{t_{j-1},t_j}^r + \mathcal{R}\mathcal{Z}_{t_{j-1},t_j}^{k-1},$$

it is equivalent to seeing that

$$\sum_{k=1}^{N} Z_{t_{j-1}}^{k-1} X_{t_j,t_{j+1}}^k + \sum_{k=1}^{N} \sum_{r=1}^{N-k} Z_{t_{j-1}}^{k+r-1} X_{t_{j-1},t_j}^r \otimes X_{t_j,t_{j+1}}^k - \sum_{k=1}^{N} \sum_{l=1}^{k} Z_{t_{j-1}}^{k-1} X_{t_{j-1},t_j}^{k-l} \otimes X_{t_j,t_{j+1}}^l = 0.$$

The above equation follows by interchanging the order of summation in the middle term. Consequently, we arrive at

$$\int_{\mathcal{P}} Z \mathrm{d}X - \int_{\mathcal{P}\setminus\{t_j\}} Z \mathrm{d}X = \sum_{k=1}^{N} \mathcal{R}\mathcal{Z}_{t_{j-1},t_j}^{k-1} \otimes X_{t_j,t_{j+1}}^k. \tag{5.5}$$

In the following argument, we directly consider the continuity estimate. The case of a single $\mathcal{I}$ is easier and only requires minor modification in the argument. Using (5.5), we have

$$\begin{aligned} &\Big| \int_{\mathcal{P}} Z \mathrm{d}X - \int_{\mathcal{P}\setminus\{t_j\}} Z \mathrm{d}X - \Big(\int_{\mathcal{P}} \tilde{Z} \mathrm{d}\tilde{X} - \int_{\mathcal{P}\setminus\{t_j\}} \tilde{Z} \mathrm{d}\tilde{X} \Big) \Big| \\ &= \Big| \sum_{k=1}^{N} \mathcal{R}\mathcal{Z}_{t_{j-1},t_j}^{k-1} \otimes X_{t_j,t_{j+1}}^k - \mathcal{R}\tilde{\mathcal{Z}}_{t_{j-1},t_j}^{k-1} \otimes \tilde{X}_{t_j,t_{j+1}}^k \Big| \\ &\leqslant (d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Z}, \tilde{\mathcal{Z}}) \|\mathbf{X}\|_\alpha + \|\tilde{\mathcal{Z}}\|_{\tilde{\mathbf{X}};\alpha} \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})) \cdot (t_{j+1} - t_{j-1})^{(N+1)\alpha}. \end{aligned}$$

As $\sum_{j=1}^{n-1} \left(t_{j+1} - t_{j-1}\right) \leqslant 2(t - s)$, we may choose a j such that

$$t_{j+1} - t_{j-1} \leqslant \frac{2(t-s)}{n-1}.$$

It follows that

$$\Big|\int_{\mathcal{P}} Z\mathrm{d}X - \int_{\mathcal{P}\backslash\{t_j\}} Z\mathrm{d}X - \Big(\int_{\mathcal{P}} \tilde{Z}\mathrm{d}\tilde{X} - \int_{\mathcal{P}\backslash\{t_j\}} \tilde{Z}\mathrm{d}\tilde{X}\Big)\Big|$$
$$\leqslant \Big(\frac{2}{n-1}\Big)^{(N+1)\alpha} (t-s)^{(N+1)\alpha} \big(d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Z},\tilde{\mathcal{Z}})\,\|\mathbf{X}\|_\alpha + \big\|\tilde{\mathcal{Z}}\big\|_{\tilde{\mathbf{X}};\alpha}\rho_\alpha(\mathbf{X},\tilde{\mathbf{X}})\big).$$

By successively removing partition points from $\mathcal{P}$, we arrive at

$$\Big|\int_{\mathcal{P}} Z\mathrm{d}X - \int_{\{s,t\}} Z\mathrm{d}X - \Big(\int_{\mathcal{P}} \tilde{Z}\mathrm{d}\tilde{X} - \int_{\{s,t\}} \tilde{Z}\mathrm{d}\tilde{X}\Big)\Big|$$
$$\leqslant C_\alpha (t-s)^{(N+1)\alpha} \big(d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Z},\tilde{\mathcal{Z}})\,\|\mathbf{X}\|_\alpha + \big\|\tilde{\mathcal{Z}}\big\|_{\tilde{\mathbf{X}};\alpha}\rho_\alpha(\mathbf{X},\tilde{\mathbf{X}})\big), \tag{5.6}$$

where

$$C_\alpha \triangleq \sum_{n=3}^{\infty} \Big(\frac{2}{n-1}\Big)^{(N+1)\alpha}.$$

The version of the inequality (5.6) without the tilde-paths is easily seen to be

$$\Big|\int_{\mathcal{P}} Z\mathrm{d}X - \int_{\{s,t\}} Z\mathrm{d}X\Big| \leqslant C_\alpha \|\mathcal{Z}\|_{\mathbf{X};\alpha}\|\mathbf{X}\|_\alpha (t-s)^{(N+1)\alpha}. \tag{5.7}$$

We now use the inequality (5.7) to show that the limit

$$\lim_{|\mathcal{P}|\to 0} \int_{\mathcal{P}} Z\mathrm{d}X$$

exists. Let $\hat{\mathcal{P}}$ and $\tilde{\mathcal{P}}$ be partitions over $[s,t]$, and let $\hat{\mathcal{P}} \vee \tilde{\mathcal{P}}$ be the partition obtained by taking a union of the partition points from $\hat{\mathcal{P}}$ and $\tilde{\mathcal{P}}$. For each pair (s_l, s_{l+1}) of adjacent points in $\hat{\mathcal{P}}$, by applying the estimate (5.7) to the partition $\hat{\mathcal{P}} \vee \tilde{\mathcal{P}} \cap [s_l, s_{l+1}]$, we obtain that

$$\Big|\int_{\hat{\mathcal{P}}\vee\tilde{\mathcal{P}}\cap[s_l,s_{l+1}]} Z\mathrm{d}X - \int_{\{s_l,s_{l+1}\}} Z\mathrm{d}X\Big| \leqslant C_\alpha (s_{l+1}-s_l)^{(N+1)\alpha} \|\mathcal{Z}\|_{\mathbf{X};\alpha} \|\mathbf{X}\|_\alpha .$$

By summing over l, we have

$$\Big|\int_{\hat{\mathcal{P}}\vee\tilde{\mathcal{P}}} Z\mathrm{d}X - \int_{\hat{\mathcal{P}}} Z\mathrm{d}X\Big| \leqslant \sum_l \Big|\int_{\hat{\mathcal{P}}\vee\tilde{\mathcal{P}}\cap[s_l,s_{l+1}]} Z\mathrm{d}X - \int_{\{s_l,s_{l+1}\}} Z\mathrm{d}X\Big|$$
$$\leqslant C_\alpha |\hat{\mathcal{P}}|^{(N+1)\alpha-1} \|\mathcal{Z}\|_{\mathbf{X};\alpha} \|\mathbf{X}\|_\alpha (t-s).$$

A similar inequality holds with $\hat{\mathcal{P}}$ replaced by $\tilde{\mathcal{P}}$. Using the triangle inequality, we end up with an estimate for $\int_{\hat{\mathcal{P}}} Z\mathrm{d}X - \int_{\tilde{\mathcal{P}}} Z\mathrm{d}X$, from which we can deduce the convergence of (5.1) using the Cauchy criterion.

Next, we establish the continuity estimate (5.4). By taking $|\mathcal{P}| \to 0$ in (5.6) and using the definition of $\mathcal{I}, \tilde{\mathcal{I}}$, we have

$$\Big|\int_s^t Z\mathrm{d}X - \sum_{k=1}^{N} Z_s^{k-1} X_{s,t}^k - \int_s^t \tilde{Z}\mathrm{d}\tilde{X} - \sum_{k=1}^{N} \tilde{Z}_s^{k-1}\tilde{X}_{s,t}^k\Big|$$

$$= \left| I_t^0 - \sum_{k=1}^{N-1} I_s^k X_{s,t}^k - (\tilde{I}_t^0 - \sum_{k=1}^{N-1} \tilde{I}_s^k \tilde{X}_{s,t}^k) - (Z_s^{N-1} X_{s,t}^N - \tilde{Z}_s^{N-1} \tilde{X}_{s,t}^N) \right|$$
$$\leqslant C_\alpha (t-s)^{(N+1)\alpha} (d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Z}, \tilde{\mathcal{Z}}) \|\mathbf{X}\|_\alpha + \|\tilde{\mathcal{Z}}\|_{\tilde{\mathbf{X}};\alpha} \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})).$$

In addition, note that

$$\left|Z_s^{N-1} - \tilde{Z}_s^{N-1}\right| \cdot \left|X_{s,t}^N\right| + \left|\tilde{Z}_s^{N-1}\right| \cdot \left|X_{s,t}^N - \tilde{X}_{s,t}^N\right|$$
$$\leqslant \left\|\mathcal{R}\mathcal{Z}^{N-1} - \mathcal{R}\tilde{\mathcal{Z}}^{N-1}\right\|_\alpha \|\mathbf{X}\|_\alpha (t-s)^{N\alpha} T^\alpha + \left|Z_0^{N-1} - \tilde{Z}_0^{N-1}\right| \|\mathbf{X}\|_\alpha (t-s)^{N\alpha}$$
$$+ \left\|\tilde{\mathcal{Z}}\right\|_{\tilde{\mathbf{X}};\alpha} \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}) (t-s)^{N\alpha} T^\alpha + \left|\tilde{Z}_0^{N-1}\right| \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}) (t-s)^{N\alpha}.$$

Therefore, we obtain that

$$\left|\mathcal{R}\mathcal{I}_{s,t}^0 - \mathcal{R}\tilde{\mathcal{I}}_{s,t}^0\right| = \left|(I_{s,t}^0 - \sum_{k=1}^{N-1} I_s^k X_{s,t}^k) - (\tilde{I}_{s,t}^0 - \sum_{k=1}^{N-1} \tilde{I}_s^k \tilde{X}_{s,t}^k)\right|$$
$$\leqslant 2C_\alpha (t-s)^{N\alpha} T^\alpha (d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Z}, \tilde{\mathcal{Z}}) \|\mathbf{X}\|_\alpha + \left\|\tilde{\mathcal{Z}}\right\|_{\tilde{\mathbf{X}};\alpha} \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}))$$
$$+ \left|Z_0^{N-1} - \tilde{Z}_0^{N-1}\right| \|\mathbf{X}\|_\alpha (t-s)^{N\alpha} + \left|\tilde{Z}_0^{N-1}\right| \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}) (t-s)^{N\alpha}.$$

In a similar way, the difference $\mathcal{R}\mathcal{I}_{s,t}^i - \mathcal{R}\tilde{\mathcal{I}}_{s,t}^i$ $(1 \leqslant i \leqslant N-1)$ is estimated as

$$|\mathcal{R}\mathcal{I}_{s,t}^i - \mathcal{R}\tilde{\mathcal{I}}_{s,t}^i| \leqslant |\mathcal{R}\mathcal{Z}_{s,t}^{i-1} - \mathcal{R}\tilde{\mathcal{Z}}_{s,t}^{i-1}| + |Z_s^{N-1} X_{s,t}^{N-i} - \tilde{Z}_s^{N-1} \tilde{X}_{s,t}^{N-i}|$$
$$\leqslant (T^\alpha((1 + \|\mathbf{X}\|_\alpha) d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Z}, \tilde{\mathcal{Z}}) + \|\tilde{\mathcal{Z}}\|_{\tilde{\mathbf{X}};\alpha} \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}))$$
$$+ \left|Z_0^{N-1} - \tilde{Z}_0^{N-1}\right| \|\mathbf{X}\|_\alpha + \left|\tilde{Z}_0^{N-1}\right| \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})) \cdot |t-s|^{(N-i)\alpha}.$$

The desired continuity estimate (5.4) thus follows. The estimate (5.3) is a special case with $\tilde{\mathcal{Z}} = 0$ and $\tilde{\mathbf{X}} = \mathbf{X}$, from which it is also clear that $\mathcal{I}$ is a controlled rough path (i.e. the remainders all have the required regularity). □

6. Rough differential equations

We now proceed to establish existence, uniqueness and continuity of solutions for the RDE

$$d\mathcal{Y}_t = F(\mathcal{Y}_t)\mathrm{d}\mathbf{X}_t \tag{6.1}$$

in the space of controlled rough paths. As a standard idea, the RDE (6.1) is formulated as a fixed point problem for the transformation

$$\mathcal{M} : \mathcal{Y} \mapsto Y_0 + \int F(\mathcal{Y})\mathrm{d}\mathbf{X}.$$

We first derive a continuity estimate for $\mathcal{M}$. Using such continuity estimate, we then show that $\mathcal{M}$ is a contraction on a small time interval. The general case follows from a patching argument.

Throughout the rest, let $\frac{1}{N+1} \leqslant \alpha < \frac{1}{N} \leqslant \frac{1}{2}$ be fixed ($N = [1/\alpha]$). Let $F = (F^0, \cdots, F^N)$ be a given $(N+1)$-Lipschitz function defined on U and taking values in $\mathcal{L}(V; U)$. We always use $M(\cdots)$ to denote a universal function that is continuous and increasing in every variable.

6.1. Composition of Lipschitz transformation and rough integration. The following lemma is a direct application of Proposition 5.1 and Theorem 4.1. All paths are assumed to be defined on $[0,\tau]$ with $\tau > 0$ given fixed.

Lemma 6.1. *Let* $\mathbf{X}$ *and* $\tilde{\mathbf{X}}$ *be* α*-Hölder geometric rough paths over* V*. Let* $\mathcal{Y}$ *and* $\tilde{\mathcal{Y}}$ *be* U*-valued paths controlled by* $\mathbf{X}$ *and* $\tilde{\mathbf{X}}$ *respectively. Define the controlled rough path* $\mathcal{J} = (J^0, \cdots, J^{N-1})$ *with respect to* $\mathbf{X}$ *in the following way*

$$J_t^0 \triangleq Y_0^0 + \Big(\int_0^{\cdot} F(\mathcal{Y})\mathrm{d}\mathbf{X}\Big)_t^0,\ J_t^i = \Big(\int_0^{\cdot} F(\mathcal{Y})\,\mathrm{d}\mathbf{X}\Big)_t^i \quad \text{for } i \geqslant 1. \tag{6.2}$$

Define $\tilde{\mathcal{J}}$ *controlled by* $\tilde{\mathbf{X}}$ *in a similar way. Then the following estimates hold true:*

$$\|\mathcal{J}\|_{\mathbf{X};\alpha} \leqslant C_\alpha\big(\tau^\alpha M(\tau, \|F\|_{Lip\text{-}N}, \|\mathbf{X}\|_\alpha, \max_{1\leqslant i\leqslant N-1}|Y_0^i|, \|\mathcal{Y}\|_{\mathbf{X};\alpha}) + \big|F(\mathcal{Y})_0^{N-1}\big| \cdot \|\mathbf{X}\|_\alpha\big) \tag{6.3}$$

and

$$\begin{aligned} d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{J},\tilde{\mathcal{J}}) \leqslant{}& C_\alpha M(\tau, \|F\|_{Lip\text{-}(N+1)}, \max_{1\leqslant i\leqslant N-1}(|Y_0^i| \vee |\tilde{Y}_0^i|), \|\mathcal{Y}\|_{\mathbf{X};\alpha}, \|\tilde{\mathcal{Y}}\|_{\tilde{\mathbf{X}};\alpha}, \|\mathbf{X}\|_\alpha, \\ & \|\tilde{\mathbf{X}}\|_\alpha) \times \big(\tau^\alpha d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y},\tilde{\mathcal{Y}}) + \big|F(\mathcal{Y})_0^{N-1} - F(\tilde{\mathcal{Y}})_0^{N-1}\big| + \rho_\alpha(\mathbf{X},\tilde{\mathbf{X}})\big). \end{aligned} \tag{6.4}$$

REMARK 6.1. The factor τ^α and the independence of Y_0^0 in the function M are both important for the patching argument in the RDE context.

6.2. Existence, uniqueness and continuity of RDE solutions. We first define the notion of solution for the RDE (6.1). The paths are now assumed to be defined on $[0,T]$ ($T > 0$ is given fixed).

DEFINITION 6.1. Let $\mathbf{X}$ be a α-Hölder geometric rough path over V, and let $Y_0 \in U$. We say that $\mathcal{Y} \in \mathcal{D}_{\mathbf{X};\alpha}(U)$ is a *solution* to the RDE (6.1) with initial condition Y_0, if

$$Y_t^0 = Y_0 + \Big(\int_0^{\cdot} F(\mathcal{Y})\,\mathrm{d}\mathbf{X}\Big)_t^0,\ Y_t^i = \Big(\int_0^{\cdot} F(\mathcal{Y})\,\mathrm{d}\mathbf{X}\Big)_t^i \qquad \text{for } 1 \leqslant i \leqslant N-1.$$

The main theorem in this part is stated as follows.

Theorem 6.1. (i) *[Existence and uniqueness] Let* $\mathbf{X}$ *be a given* α*-Hölder geometric rough path over* V*. For each* $Y_0 \in U$*, there exists a unique solution* $\mathcal{Y} \in \mathcal{D}_{\mathbf{X};\alpha}(U)$ *to the* RDE (6.1) *in the sense of* Definition 6.1.
(ii) *[Continuity estimate] Let* $\mathbf{X}$ *and* $\tilde{\mathbf{X}}$ *be* α*-Hölder geometric rough paths over* V*, and let* $Y_0, \tilde{Y}_0 \in U$*. Suppose that*

$$\|\mathbf{X}\|_\alpha \vee \|\tilde{\mathbf{X}}\|_\alpha \vee |Y_0| \vee |\tilde{Y}_0| \leqslant B$$

with some constant $B > 0$*. Let* $\mathcal{Y}$ *and* $\tilde{\mathcal{Y}}$ *be the solutions to (6.1) driven by* $\mathbf{X}$ *and* $\tilde{\mathbf{X}}$ *with initial conditions* Y_0 *and* $\tilde{Y}_0$ *respectively. Then the following estimate holds true:*

$$d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y},\tilde{\mathcal{Y}}) \leqslant M(T, \|F\|_{Lip\text{-}(N+1)}, B)\big(\rho_\alpha(\mathbf{X},\tilde{\mathbf{X}}) + \big|Y_0 - \tilde{Y}_0\big|\big). \tag{6.5}$$

The rest of this subsection is devoted to the proof of Theorem 6.1.

6.2.1. Local contraction. We prove existence and uniqueness by using the Banach fixed point theorem. Note that the "constants" appearing in the rough integration and Lipschitz transformation estimates depend on $\|\mathcal{Y}\|_{\mathbf{X};\alpha}$. As a result, the mapping $\mathcal{M} : \mathcal{Y} \mapsto Y_0 + \int F(\mathcal{Y})d\mathbf{X}$ can only be a contraction if we restrict $\mathcal{M}$ on a bounded subset, say a metric ball. To determine the center $\mathcal{W}$ (as a controlled rough path) of such a ball, it is natural to require $W_0^0 = Y_0$ as this is the given initial condition. The higher order terms W^i ($i \geqslant 1$) are chosen such that $\mathcal{R}\mathcal{W}^i_{s,t} = 0$. This is formulated precisely in the following lemma.

Lemma 6.2. *Let $Y_0 \in U$ be given. We set*

$$W_0^0 \triangleq Y_0,\ W_0^1 \triangleq F^0(Y_0),$$

and inductively

$$W_0^{r+1} \triangleq \sum_{j=0}^{N-1} \frac{F^j(Y_0)}{j!}\Big(\sum_{i_1+\cdots+i_j=r} (W_0^{i_1} \boxtimes \cdots \boxtimes W_0^{i_j}) \circ \delta_j\Big) \in \mathcal{L}(V^{\otimes(r+1)}; U). \tag{6.6}$$

Define the path $\mathcal{W} = (W^0, W^1, W^2, \cdots, W^{N-1})$ by

$$W_t^i(\xi) = W_0^i(\xi) + W_0^{i+1}(X_{0,t}^1 \otimes \xi) + \cdots + W_0^{N-1}(X_{0,t}^{N-1-i} \otimes \xi).$$

Then $\mathcal{W}$ is a controlled rough path with respect to $\mathbf{X}$. *More specifically, we have $\mathcal{R}\mathcal{W}^i \equiv 0$ for each* $0 \leqslant i \leqslant N-1$.

REMARK 6.2. The initial value $\mathcal{W}_0$ is canonically determined by Y_0 and F.

Proof. Note that

$$\begin{aligned} W_t^i(\xi) &= \sum_{j=i}^{N-1} W_0^j(X_{0,t}^{j-i} \otimes \xi) = \sum_{j=i}^{N-1} W_0^j\Big(\sum_{k=0}^{j-i} X_{0,s}^{j-i-k} \otimes X_{s,t}^k \otimes \xi\Big) \\ &= \sum_{k=0}^{N-1-i} \sum_{j=i+k}^{N-1} W_0^j(X_{0,s}^{j-i-k} \otimes X_{s,t}^k \otimes \xi) = \sum_{k=0}^{N-1-i} W_s^{i+k}(X_{s,t}^k \otimes \xi) \end{aligned}$$

As a result, we have $\mathcal{R}\mathcal{W}^i_{s,t} = 0$ for all $s \leqslant t$ and $0 \leqslant i \leqslant N-1$. □

The definition (4.2) of $F(\mathcal{W})_0^{N-1}$ together with the inductive definition (6.6) of W_0^i ($1 \leqslant i \leqslant N-1$) imply the existence of a continuous, increasing function M such that

$$\left|F(\mathcal{W})_0^{N-1}\right| \leqslant M(\|F\|_{\text{Lip-}(N-1)}).$$

Therefore, for any controlled path $\mathcal{Y}$ with $\mathcal{Y}_0 = \mathcal{W}_0$, we have

$$C_\alpha \left|F(\mathcal{Y})_0^{N-1}\right| \|\mathbf{X}\|_\alpha \leqslant M(\|F\|_{\text{Lip-}(N-1)}) \|\mathbf{X}\|_\alpha =: \Lambda. \tag{6.7}$$

where C_α is the constant appearing in the estimate (6.3). The following lemma gives the local existence and uniqueness for the RDE (6.1).

Lemma 6.3. *For each $\tau > 0$, we set*

$$\mathcal{B}_\tau \triangleq \{\mathcal{Y} \in \mathcal{D}_{\mathbf{X};\alpha}(U) : \|\mathcal{Y} - \mathcal{W}\|_{\mathbf{X};\alpha} \leqslant 2\Lambda, \mathcal{Y}_0 = \mathcal{W}_0\}.$$

Then there exists a small $\tau > 0$, which is independent of Y_0 and depends only on α, $\mathbf{X}$ *and $\|F\|_{Lip\text{-}(N+1)}$, such that:*

(i) *the mapping $\mathcal{M} : \mathcal{Y} \mapsto \mathcal{J}$ sends $\mathcal{B}_\tau$ to $\mathcal{B}_\tau$, where $\mathcal{J}$ is the controlled rough path defined*

by (6.2)*;*
(ii) *the mapping* $\mathcal{M}$ *is a contraction on* $\mathcal{B}_\tau$ *with respect to the norm* $|||\cdot|||_{\mathbf{X};\alpha}$ *defined by* (3.5).
(iii) *The* RDE (6.1) *has a unique solution* $\mathcal{Y}$ *on* $[0,\tau]$ *that satisfies*

$$\|\mathcal{Y}\|_{\mathbf{X};\alpha} \leqslant 2\Lambda. \tag{6.8}$$

Proof. (i) We first prove by induction that $W_0^i = J_0^i$ for all i. The $i = 0, 1$ cases follow directly from the definition of $\mathcal{J}$ and $\mathcal{W}$. For the induction step, note that by the definition of $\mathcal{J}$, if $\xi \in V^{\otimes(r+1)}$, then

$$\begin{aligned}
J_0^{r+1}(\xi) &= (F(\mathcal{Y}))_0^r(\xi) \\
&= \sum_{j=0}^{N-1} \frac{F^j(Y_0)}{j!} \sum_{i_1+\cdots+i_j=r} (Y_0^{i_1} \boxtimes \cdots \boxtimes Y_0^{i_j}) \circ \delta_j(\xi) \\
&= \sum_{j=0}^{N-1} \frac{F^j(Y_0)}{j!} \sum_{i_1+\cdots+i_j=r} (W_0^{i_1} \boxtimes \cdots \boxtimes W_0^{i_j}) \circ \delta_j(\xi) \qquad (\text{since } \mathcal{Y}_0 = \mathcal{W}_0) \\
&= W_0^{r+1} \qquad (\text{by definition of } W_0^{r+1}).
\end{aligned}$$

Therefore, $\mathcal{W}_0 = \mathcal{J}_0$.

Next, we recall from (6.3) that

$$\|\mathcal{J}\|_{\mathbf{X};\alpha} \leqslant C_\alpha \tau^\alpha M(\tau, \|F\|_{\text{Lip-}N}, \|\mathbf{X}\|_\alpha, \max_{1\leqslant i\leqslant N-1} |Y_0^i|, \|\mathcal{Y}\|_{\mathbf{X};\alpha}) + \Lambda.$$

Since $\mathcal{R}\mathcal{W}_{s,t}^i = 0$ by Lemma 6.2 and $\mathcal{Y} \in \mathcal{B}_\tau$, we know that

$$\|\mathcal{Y}\|_{\mathbf{X};\alpha} = d_{\mathbf{X},\mathbf{X};\alpha}(\mathcal{Y}, \mathcal{W}) \leqslant 2\Lambda.$$

The inductive definition of $\mathcal{W}_0$ in (6.6) implies that there is a continuous increasing function M such that

$$\left|Y_0^i\right| = \left|W_0^i\right| \leqslant M(\|F\|_{\text{Lip-}(N-1)}), \quad 1 \leqslant i \leqslant N-1. \tag{6.9}$$

As a consequence, we can choose τ to be sufficiently small (depending on $\|\mathbf{X}\|_\alpha$ and $\|F\|_{\text{Lip-}N}$), such that

$$C_\alpha \tau^\alpha M(\tau, \|F\|_{\text{Lip-}N}, \|\mathbf{X}\|_\alpha, \max_{1\leqslant i\leqslant N-1} |Y_0^i|, \|\mathcal{Y}\|_{\mathbf{X};\alpha}) < \Lambda. \tag{6.10}$$

This ensures that $\mathcal{J} = \mathcal{M}(\mathcal{Y}) \in \mathcal{B}_\tau$. Note that the choice of τ is independent of Y_0^0.

(ii) Let $\mathcal{Y}, \tilde{\mathcal{Y}} \in \mathcal{B}_\tau$. Note that $\mathcal{Y}_0 = \tilde{\mathcal{Y}}_0$ and thus $F(\mathcal{Y})_0^{N-1} = F(\tilde{\mathcal{Y}})_0^{N-1}$. By applying (6.4) with $\tilde{\mathbf{X}} = \mathbf{X}$, we have

$$\begin{aligned}
d_{\mathbf{X},\mathbf{X};\alpha}(\mathcal{M}(\mathcal{Y}), \mathcal{M}(\tilde{\mathcal{Y}})) \leqslant C_\alpha \tau^\alpha M(\tau, \|F\|_{\text{Lip-}(N+1)}, \max_{1\leqslant i\leqslant N-1} |Y_0^i|, \\
\|\mathcal{Y}\|_{\mathbf{X};\alpha}, \|\tilde{\mathcal{Y}}\|_{\mathbf{X};\alpha}, \|\mathbf{X}\|_\alpha) d_{\mathbf{X},\mathbf{X};\alpha}(\mathcal{Y}, \tilde{\mathcal{Y}}).
\end{aligned}$$

According to (6.9) and the fact that $\mathcal{Y}, \tilde{\mathcal{Y}} \in \mathcal{B}_\tau$, we may further reduce τ (independent of Y_0^0) such that

$$C_\alpha \tau^\alpha M(\tau, \|F\|_{\text{Lip-}(N+1)}, \max_{1\leqslant i\leqslant N-1} |Y_0^i|, \|\mathcal{Y}\|_{\mathbf{X};\alpha}, \|\tilde{\mathcal{Y}}\|_{\mathbf{X};\alpha}, \|\mathbf{X}\|_\alpha) < \frac{1}{2}.$$

In this way, we arrive at

$$|||\mathcal{M}(\mathcal{Y}) - \mathcal{M}(\tilde{\mathcal{Y}})|||_{\mathbf{X};\alpha} \leqslant \frac{1}{2} ||| \mathcal{Y} - \tilde{\mathcal{Y}}|||_{\mathbf{X};\alpha},$$

which shows that the mapping $\mathcal{M} : \mathcal{B}_\tau \to \mathcal{B}_\tau$ is a contraction for such a choice of τ.

(iii) Let τ be chosen as in Part (ii). Note that a solution to the RDE (6.1) is a fixed point of the mapping $\mathcal{M}$. Since $\mathcal{B}_\tau$ is a closed subset of the Banach space $(\mathcal{D}_{\mathbf{X},\alpha}(U), ||| \cdot |||_{\mathbf{X};\alpha})$, according to Part (ii) and the Banach fixed point theorem, we conclude that the RDE (6.1) admits a unique solution $\mathcal{Y} \in \mathcal{B}_\tau$ as a controlled rough path on $[0, \tau]$. The inequality (6.8) is just a consequence of $\mathcal{Y} \in \mathcal{B}_\tau$. □

Remark 6.3. It is interesting to point out that, if $\mathcal{Y} = (Y^0, \cdots, Y^{N-1})$ is a solution to the RDE (6.1), then at each t the values Y^i_t $(1 \leqslant i \leqslant N-1)$ are all canonically determined by the value Y^0_t of the 0-th level path. Indeed, by Definition 6.1 we have $Y^i_t = F(\mathcal{Y})^{i-1}_t$ for all $i \geqslant 1$. The determination of Y^i_t from Y^0_t is through the same relation as (6.6). This observation is used in the later patching argument.

6.2.2. A patching lemma. In order to obtain global existence, we need to patch local solutions in the sense of controlled rough paths. The lemma below justifies the patching of controlled rough paths in general.

Lemma 6.4. (i) *Let* $\mathbf{X}$ *be an* α*-Hölder geometric rough path on* $[a, b]$ *and let* $u \in (a, b)$ *be fixed. Let* $\mathcal{Y}$ *be a continuous path on* $[a, b]$ *such that* $\mathcal{Y}|_{[a,u]}$ *(respectively,* $\mathcal{Y}|_{[u,b]}$*) is controlled by* $\mathbf{X}|_{[a,u]}$ *(respectively, by* $\mathbf{X}|_{[u,b]}$*). Then* $\mathcal{Y}$ *is controlled by* $\mathbf{X}$ *on* $[a, b]$*.*
(ii) *Let* $\mathbf{X}, \tilde{\mathbf{X}}$ *be* α*-Hölder geometric rough paths on* $[a, b]$ *and let* $\mathcal{Y}, \tilde{\mathcal{Y}}$ *be controlled by* $\mathbf{X}, \tilde{\mathbf{X}}$ *respectively. Let* $u \in (a, b)$ *be fixed. Then we have*

$$d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}, \tilde{\mathcal{Y}}) \leqslant d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}|_{[u,b]}, \tilde{\mathcal{Y}}|_{[u,b]}) + d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}|_{[a,u]}, \tilde{\mathcal{Y}}|_{[a,u]})(1 + \|\mathbf{X}\|_\alpha) + \|\tilde{\mathcal{Y}}|_{[a,u]}\|_{\tilde{\mathbf{X}};\alpha}\rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}). \tag{6.11}$$

Proof. (i) It is enough to consider the remainder $\mathcal{R}\mathcal{Y}^k_{s,t}$ when $s < u < t$. Note that

$$\begin{aligned}
\sum_{i=k}^{N-1} Y^i_s X^{i-k}_{s,t} &= \sum_{i=k}^{N-1} Y^i_s \sum_{j=0}^{i-k} X^{i-k-j}_{s,u} X^j_{u,t} = \sum_{i=k}^{N-1} Y^i_s \sum_{j=k}^{i} X^{i-j}_{s,u} X^{j-k}_{u,t} \\
&= \sum_{j=k}^{N-1} (\sum_{i=j}^{N-1} Y^i_s X^{i-j}_{s,u}) X^{j-k}_{u,t} \\
&= \sum_{j=k}^{N-1} Y^j_u X^{j-k}_{u,t} - \sum_{j=k}^{N-1} (Y^j_u - \sum_{i=j}^{N-1} Y^i_s X^{i-j}_{s,u}) X^{j-k}_{u,t}.
\end{aligned}$$

Therefore,

$$Y^k_t - \sum_{i=k}^{N-1} Y^i_s X^{i-k}_{s,t} = Y^k_t - \sum_{j=k}^{N-1} Y^j_u X^{j-k}_{u,t} + \sum_{j=k}^{N-1} (Y^j_u - \sum_{i=j}^{N-1} Y^i_s X^{i-j}_{s,u}) X^{j-k}_{u,t},$$

or equivalently

$$\mathcal{RY}^k_{s,t}=\mathcal{RY}^k_{u,t}+\sum_{j=k}^{N-1}\mathcal{RY}^j_{s,u}X^{j-k}_{u,t}. \tag{6.12}$$

From (6.12), it is clear that the Hölder regularity of $\mathcal{RY}^k_{s,t}$ is $|t-s|^{(N-k)\alpha}$.

(ii) According to (6.12), for $s<u<t$ we also have

$$\begin{aligned}\left|\mathcal{RY}^k_{s,t}-\mathcal{R\tilde{Y}}^k_{s,t}\right| &\leqslant \left|\mathcal{RY}^k_{u,t}-\mathcal{R\tilde{Y}}^k_{u,t}\right|+\sum_{j=k}^{N-1}\left|\mathcal{RY}^j_{s,u}-\mathcal{R\tilde{Y}}^j_{s,u}\right|\|X^{j-k}\|_\alpha\,(t-u)^{(j-k)\alpha}\\ &\quad+\sum_{j=k}^{N-1}\left|\mathcal{R\tilde{Y}}^j_{s,u}\right|\|X^{j-k}-\tilde{X}^{j-k}\|_\alpha(t-u)^{(j-k)\alpha}\\ &\leqslant d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}|_{[u,b]},\tilde{\mathcal{Y}}|_{[u,b]})(t-u)^{(N-k)\alpha}\\ &\quad+\sum_{j=k}^{N-1}\|\mathcal{RY}^j|_{[a,u]}-\mathcal{R\tilde{Y}}^j|_{[a,u]}\|_{(N-j)\alpha}(u-s)^{(N-j)\alpha}\|\mathbf{X}\|_\alpha(t-u)^{(j-k)\alpha}\\ &\quad+\sum_{j=k}^{N-1}\|\mathcal{R\tilde{Y}}^j|_{[a,u]}\|_{(N-j)\alpha}\rho_\alpha(\mathbf{X},\tilde{\mathbf{X}})(u-s)^{(N-j)\alpha}(t-u)^{(j-k)\alpha}.\end{aligned}$$

The inequality (6.11) thus follows. □

6.2.3. Global existence, uniqueness and continuity. By patching local solutions and local estimates, we are now able to establish the global well-posedness of the RDE (6.1) in the space of controlled rough paths. Let α, N, F be given as before.

Proof of Theorem 6.1. *Existence*. Let τ be given by Lemma 6.3. According to that lemma, we have a solution $\mathcal{Y}[1]$ on $[0,\tau]$ satisfying

$$Y[1]^0_t=Y_0+\int_0^t F(Y[1])\mathrm{d}X,\ \ Y[1]^i_t=[F(\mathcal{Y}[1])]^{i-1}_t\qquad\forall t\in[0,\tau].$$

We define a sequence of controlled paths $\{\mathcal{Y}[n]:n\geqslant 1\}$ on $[0,\tau]$ inductively in the following way. By applying Lemma 6.3 with $Y_0=\mathcal{Y}[n-1]_\tau$ and $\mathbf{X}_t=\mathbf{X}_{(n-1)\tau+t}$, we obtain a controlled rough path $\mathcal{Y}[n]$ on $[0,\tau]$ satisfying

$$Y[n]^0_t=Y[n-1]^0_\tau+\int_0^t F(Y[n])\mathrm{d}X,\ \ Y[n]^i_t=[F(\mathcal{Y}[n])]^{i-1}_t\quad\forall t\in[0,\tau].$$

We now define $\mathcal{Y}=(Y^0,\cdots,Y^{N-1})$ as a path on $[0,\infty)$ by concatenating all the $\mathcal{Y}[n]$'s, namely

$$\mathcal{Y}_{(n-1)\tau+t}=\mathcal{Y}[n]_t,\qquad t\in[0,\tau].$$

Note from Remark 6.3 that $\mathcal{Y}$ is well defined. By Lemma 6.4, $\mathcal{Y}$ is a controlled rough path with respect to $\mathbf{X}$.

For any $t\geqslant 0$, if $t\in[(n-1)\tau,n\tau]$ we have

$$Y^i_t=Y[n]^i_{t-(n-1)\tau}=F(\mathcal{Y}[n])^{i-1}_{t-(n-1)\tau}=F(\mathcal{Y})^{i-1}_t.$$

It remains to show that,

(6.13) $$Y_t^0 = Y_0 + \int_0^t F(Y)\mathrm{d}X, \qquad \forall t \geqslant 0.$$

We use induction on n. If $t \in [(n-1)\tau, n\tau]$, then

$$\begin{aligned} Y_t^0 &= Y[n-1]^0_{t-(n-1)T} \\ &= Y[n-1]^0_0 + \int_0^{t-(n-1)\tau} F(Y[n-1]_\cdot)\mathrm{d}X_{\cdot+(n-1)\tau} \\ &= Y_0 + \int_0^{(n-1)\tau} F(Y)\mathrm{d}X + \int_0^{t-(n-1)\tau} F(Y[n-1]_\cdot)\mathrm{d}X_{\cdot+(n-1)\tau} \\ &= Y_0 + \int_0^{(n-1)\tau} F(Y)\mathrm{d}X + \int_{(n-1)\tau}^{t} F(Y)\mathrm{d}X \\ &= Y_0 + \int_0^t F(Y)\mathrm{d}X, \end{aligned}$$

where the third equality follows from induction hypothesis. Therefore, (6.13) holds. We have thus obtained the existence of solution on $[0,\infty)$.

Uniqueness. Let $\tilde{\mathcal{Y}}$ be another solution to the RDE (6.1). Suppose that

$$\sigma \triangleq \sup\{t \in [0,\infty) : \tilde{\mathcal{Y}}_s = \mathcal{Y}_s \text{ for all } s \in [0,t]\} < \infty.$$

Then $\tilde{\mathcal{Y}}_\sigma = \mathcal{Y}_\sigma$. According to Lemma 6.1 with $\mathbf{X} = \tilde{\mathbf{X}}$, for all τ sufficiently small we have

$$\begin{aligned} d_{\mathbf{X},\mathbf{X};\alpha}(\tilde{\mathcal{Y}}|_{[\sigma,\sigma+\tau]}, \mathcal{Y}|_{[\sigma,\sigma+\tau]}) \leqslant C_\alpha \tau^\alpha M(\tau, \|F\|_{\text{Lip-}(N+1)}, \max_{1\leqslant i\leqslant N-1} |Y_\sigma^i|, \|\mathcal{Y}\|_{\mathbf{X};\alpha}, \\ \|\tilde{\mathcal{Y}}\|_{\mathbf{X};\alpha}, \|\mathbf{X}\|_\alpha) \cdot d_{\mathbf{X},\mathbf{X};\alpha}(\tilde{\mathcal{Y}}|_{[\sigma,\sigma+\tau]}, \mathcal{Y}|_{[\sigma,\sigma+\tau]}). \end{aligned}$$

If we choose τ to be such that

$$C_\alpha \tau^\alpha M(\tau, \|F\|_{\text{Lip-}(N+1)}, \max_{1\leqslant i\leqslant N-1} |Y_\sigma^i|, \|\mathcal{Y}\|_{\mathbf{X};\alpha}, \|\tilde{\mathcal{Y}}\|_{\mathbf{X};\alpha}, \|\mathbf{X}\|_\alpha) < 1,$$

then

$$d_{\mathbf{X},\mathbf{X};\alpha}(\tilde{\mathcal{Y}}|_{[\sigma,\sigma+\tau]}, \mathcal{Y}|_{[\sigma,\sigma+\tau]}) = 0.$$

This implies that $\tilde{\mathcal{Y}}|_{[\sigma,\sigma+\tau]} = \mathcal{Y}|_{[\sigma,\sigma+\tau]}$ as $\tilde{\mathcal{Y}}_\sigma = \mathcal{Y}_\sigma$, which contradicts the definition of σ. Therefore, $\tilde{\mathcal{Y}} = \mathcal{Y}$ on $[0,\infty)$.

Continuity estimate. We now assume that all the underlying paths are defined on a given fixed interval $[0,T]$. According to Lemma 6.1 and the fact that $\mathcal{Y}, \tilde{\mathcal{Y}}$ are RDE solutions, when restricted on any sub-interval $[0,\tau]$ we have

$$d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y},\tilde{\mathcal{Y}}) \leqslant C_\alpha M \cdot (\tau^\alpha d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y},\tilde{\mathcal{Y}}) + |Y_0 - \tilde{Y}_0| + \rho_\alpha(\mathbf{X},\tilde{\mathbf{X}})),$$

where we have set

$$M \triangleq M(\tau, \|F\|_{\text{Lip-}(N+1)}, \|\mathcal{Y}\|_{\mathbf{X};\alpha}, \|\tilde{\mathcal{Y}}\|_{\tilde{\mathbf{X}};\alpha}, \|\mathbf{X}\|_\alpha, \|\tilde{\mathbf{X}}\|_\alpha)$$

and used the observation that (cf. (6.9))

(6.14) $$\max_{1\leqslant i\leqslant N-1} \left|Y_0^i\right| \vee \left|\tilde{Y}_0^i\right| \leqslant M(\|F\|_{\text{Lip-}(N-1)}).$$

Let Λ be defined in (6.7) with $\|\mathbf{X}\|_\alpha$ replaced by B. By choosing τ to be small enough, as a result of (6.8) we can ensure that

(6.15) $$\|\mathcal{Y}|_{[0,\tau]}\|_{\mathbf{X};\alpha} \vee \|\tilde{\mathcal{Y}}|_{[0,\tau]}\|_{\mathbf{X};\alpha} \leqslant 2\Lambda$$

and also that

$$C_\alpha \tau^\alpha M' \triangleq C_\alpha \tau^\alpha M(\tau, \|F\|_{\text{Lip-}(N+1)}, 2\Lambda, 2\Lambda, B, B) = \frac{1}{2}.$$

It follows that

(6.16) $$d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}, \tilde{\mathcal{Y}}) \leqslant 2C_\alpha M' \cdot (|Y_0 - \tilde{Y}_0| + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})).$$

on $[0,\tau]$. It is important to note that the choice of τ is independent of Y_0 and $\tilde{Y}_0$. By applying (6.16) on each sub-interval $[(n-1)\tau, n\tau]$ $(1 \leqslant n \leqslant T/\tau)$, we arrive at

(6.17) $$d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}, \tilde{\mathcal{Y}}) \leqslant 2C_\alpha M' \cdot \left(\left|Y^0_{(n-1)\tau} - \tilde{Y}^0_{(n-1)\tau}\right| + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})\right)$$

when restricted on $[(n-1)\tau, n\tau]$.

From (6.15) it is clear that

$$\|\mathcal{Y}|_{[(n-1)\tau,n\tau]}\|_{\mathbf{X};\alpha} \vee \|\tilde{\mathcal{Y}}|_{[(n-1)\tau,n\tau]}\|_{\tilde{\mathbf{X}};\alpha} \leqslant 2\Lambda \quad \forall n.$$

In addition, according to Lemma 3.1, when restricted on $[0,\tau]$ we have

$$\begin{aligned}\left|Y^i_\tau - \tilde{Y}^i_\tau\right| \leqslant& \left|Y^i_0 - \tilde{Y}^i_0\right| + \tau^\alpha \|Y^i - \tilde{Y}^i\|_\alpha \\ \leqslant& \tau^\alpha M(\tau, \|\mathbf{X}\|_\alpha, \|\tilde{\mathbf{X}}\|_\alpha, \max_{i+1\leqslant j\leqslant N-1} \left|Y^j_0\right|, \max_{i+1\leqslant j\leqslant N-1} \|\mathcal{R}\mathcal{Y}^{N-j}\|_{j\alpha}) \\ &\times (\rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}) + \sum_{i=0}^{N-1} \left|Y^i_0 - \tilde{Y}^i_0\right| + d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}, \tilde{\mathcal{Y}})), \quad 0 \leqslant i \leqslant N-2.\end{aligned}$$

Observe that

$$\sum_{i=0}^{N-1} \left|Y^i_0 - \tilde{Y}^i_0\right| \leqslant M(\|F\|_{\text{Lip-}N}) \left|Y_0 - \tilde{Y}_0\right|,$$

which is clear since all the Y^i_0's and $\tilde{Y}^i_0$'s are canonically determined by Y^0_0 and $\tilde{Y}^0_0$ via the relation (6.6). In view of (6.15) and (6.16), we can further write

(6.18) $$\left|Y^i_\tau - \tilde{Y}^i_\tau\right| \leqslant M(\tau, \|F\|_{\text{Lip-}(N+1)}, B)(\left|Y_0 - \tilde{Y}_0\right| + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})).$$

By applying (6.18) iteratively, we obtain that

(6.19) $$\left|Y^i_{n\tau} - \tilde{Y}^i_{n\tau}\right| \leqslant M_n(\tau, \|F\|_{\text{Lip-}(N+1)}, B)(\left|Y_0 - \tilde{Y}_0\right| + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}))$$

for all n, where the increasing function M_n can depend on n.

To proceed further, we show by induction that

(6.20) $$d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}|_{[0,n\tau]}, \tilde{\mathcal{Y}}|_{[0,n\tau]}) \leqslant M_n(\tau, \|F\|_{\text{Lip-}(N+1)}, B)(\left|Y_0 - \tilde{Y}_0\right| + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}))$$

for each $1 \leqslant n \leqslant T/\tau$. Suppose that (6.20) is true on $[0,(n-1)\tau]$. According to (6.17) and (6.19), we have

(6.21) $$\begin{aligned}d_{\mathbf{X},\tilde{\mathbf{X}};\alpha}(\mathcal{Y}|_{[(n-1)\tau,n\tau]}, \tilde{\mathcal{Y}}|_{[(n-1)\tau,n\tau]}) &\leqslant M(\tau, \|F\|_{\text{Lip-}(N+1)}, B)\rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}}) \\ &+ M_{n-1}(\tau, \|F\|_{\text{Lip-}(N+1)}, B)(\left|Y_0 - \tilde{Y}_0\right| + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})) \\ &\leqslant M_n(\tau, \|F\|_{\text{Lip-}(N+1)}, B)(\left|Y_0 - \tilde{Y}_0\right| + \rho_\alpha(\mathbf{X}, \tilde{\mathbf{X}})).\end{aligned}$$

We can then apply Lemma 6.4 to patch the estimate on $[0,(n-1)\tau]$ with the one on $[(n-1)\tau, n\tau]$ given by (6.21). This completes the induction step. The desired continuity estimate (6.5) follows by taking $n = T/\tau$.

Now the proof of Theorem 6.1 is complete. □

To conclude the discussion, we give a few remarks on several possible extensions.

In the first place, if the vector field F and its derivatives are not uniformly bounded, the solution to the RDE (6.1) may explode in finite time. Similar discussion gives existence and uniqueness up to the explosion time.

In addition, in the continuity estimate (6.5), one can also take into account the perturbation of the vector field F. In this case, an extra term of $\|F-\tilde{F}\|_{\text{Lip-}(N+1)}$ will appear on the right hand side of (6.5).

Finally, it is possible to reduce the Lipschitz condition $F \in \text{Lip}(N+1)$ to $F \in \text{Lip}(\gamma)$ with $\gamma > \alpha^{-1}$ by sacrificing the Hölder regularity of the remainders of the controlled rough path $F(\mathcal{Y})$. More specifically, $\mathcal{R}F(\mathcal{Y})^i_{s,t}$ should have regularity $|t-s|^{(\gamma-1-i)\alpha}$ instead of $|t-s|^{(N-i)\alpha}$. Correspondingly, the definition of control rough paths needs to be relaxed to allow more flexible Hölder exponents for the remainders. This point is essentially clear in [5] for the case of $\alpha > 1/3$ and the extension to the general case is only a technical matter. Nonetheless, we should point out that proving existence of solutions under the optimal assumption of $F \in \text{Lip}(\gamma)$ with $\gamma > \alpha^{-1} - 1$ seems to be only possible in finite dimensions, as one has to rely on the Leray-Schauder fixed point theorem in that case, which requires the compactness of $\mathcal{M}$ and it is only true in finite dimensions.

Acknowledgements. The author XG is supported by the ARC grant DE210101352.

References

[1] A.M. Davie: *Differential equations driven by rough paths: an approach via discrete approximation*, Appl. Math. Res. Express. AMRX 2008, Art. ID abm009, 40pp.
[2] P.K. Friz and M. Hairer: A Course on Rough Paths, Universitext, Springer, Cham, 2014.
[3] P.K. Friz and N.B. Victoir: Multidimensional Stochastic Processes as Rough Paths: Theory and Applications, Cambridge University Press, Cambridge, 2010.
[4] P.K. Friz and H. Zhang: *Differential equations driven by rough paths with jumps*, J. Differential Equations **264** (2018), 6226–6301.
[5] M. Gubinelli: *Controlling rough paths*, J. Funct. Anal. **216** (2004), 86–140.
[6] M. Gubinelli: *Ramification of rough paths*, J. Differential Equations **248** (2010), 693–721.
[7] M. Hairer: *A theory of regularity structures*, Invent. Math. **198** (2014), 269–504.
[8] T.J. Lyons: *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana **14** (1998), 215–310.
[9] T.J. Lyons and Z. Qian: System Control and Rough Paths, Oxford Mathematical Monographs, Oxford University Press, Oxford, 2002.
[10] T.J. Lyons and D. Yang: *Integration of time-varying cocyclic one-forms against rough paths*, arXiv:1408.2785, 2014.
[11] P. Malliavin: *Stochastic calculus of variation and hypoelliptic operators*; in Proc. Internat. Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), Wiley, New York-Chichester-Brisbane, 1978, 195–263.
[12] A. Ohashi, F. Russo and E. Shamarova: *Smoothness of densities for path-dependent SDEs under Hörmander's condition*, arXiv:2011.04089, 2020.

[13] B. Oksendal: Stochastic Differential Equations: An Introduction with Applications, 6th Edition, Springer-Verlag, 2013.
[14] C. Reutenauer: Free Lie Algebras, London Mathematical Society Monographs, New Series, **7**, The Clarendon Press, New York, 1993.
[15] E.M. Stein: Singular Integrals and Differentiability Properties of Functions, Princeton Mathematical Series **30**, Princeton University Press, Princeton, N.J., 1970.

Horatio Boedihardjo
Department of Statistics, University of Warwick
Coventry, CV4 7AL
United Kingdom
e-mail: horatio.boedihardjo@warwick.ac.uk

Xi Geng
School of Mathematics and Statistics, University of Melbourne
Parkville, VIC 3010
Australia
e-mail: xi.geng@unimelb.edu.au