Electron. Commun. Probab. **20** (2015), no. 94, 1–11.
DOI: 10.1214/ECP.v20-4124
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Uniform factorial decay estimates for controlled differential equations*

Horatio Boedihardjo[†] Terry Lyons[‡] Danyu Yang[§]

Abstract

We establish a uniform factorial decay estimate for the Taylor approximation of solutions to controlled differential equations in the p-variation metric. As part of the proof, we also obtain a factorial decay estimate for controlled paths which is interesting in its own right.

Keywords: Controlled differential equation ; Rough paths ; Taylor expansion ; Factorial Decay.
AMS MSC 2010: 60H10; 34H05.
Submitted to ECP on February 18, 2015, final version accepted on September 14, 2015.
Supersedes arXiv:1502.04116v1.

1 Introduction

For a controlled differential equation of the form

$$\begin{aligned} \mathrm{d}Y_t &= f(Y_t)\,\mathrm{d}X_t \\ Y_0 &= y_0. \end{aligned} \tag{1.1}$$

where $X : [0,T] \to \mathbb{R}^d$ is a path with finite 1-variation and $f : \mathbb{R}^e \to L\left(\mathbb{R}^d, \mathbb{R}^e\right)$ is a smooth vector field, we are interested in estimating the Taylor remainder

$$Y_t - Y_s - \sum_{k=1}^{N} f^{\circ k}(Y_s) \int_{s<s_1<\ldots<s_k<t} \mathrm{d}X_{s_1} \otimes \ldots \otimes \mathrm{d}X_{s_k} \tag{1.2}$$

$$\equiv \int_{s<s_1<\ldots<s_N<t} f^{\circ N}(Y_{s_1}) - f^{\circ N}(Y_s)\,\mathrm{d}X_{s_1} \otimes \ldots \otimes \mathrm{d}X_{s_N}, \tag{1.3}$$

where $f^{\circ m} : \mathbb{R}^e \to L\left(\left(\mathbb{R}^d\right)^{\otimes m}, \mathbb{R}^e\right)$ is defined inductively by

$$\begin{aligned} f^{\circ 1} &= f \\ f^{\circ k+1} &= D\left(f^{\circ k}\right) f. \end{aligned}$$

*The authors are grateful for the support of the ERC Advanced grant (grant agreement no. 291244), for which the second author is the principal investigator. The contribution of the first author had been made during his stay at Oxford-Man Institute of Quantitative Finance.

[†]Reading University, UK. E-mail: h.s.boedihardjo@reading.ac.uk
[‡]Oxford-Man Institute of Quantitative Finance, University of Oxford, UK.
E-mail: terry.lyons@oxford-man.ox.ac.uk
[§]Oxford-Man Institute of Quantitative Finance, University of Oxford, UK.
E-mail: danyu.yang@oxford-man.ox.ac.uk

The functions $f^{\circ k}$ can also be expressed in terms of iterative applications of the vector field f as differential operators [3]. The iterated integrals in (1.2) will appear numerous times and we shall use the shorthand

$$X_{s,t}^{k} := \int_{s<s_1<\ldots<s_k<t} \mathrm{d}X_{s_1} \otimes \ldots \otimes \mathrm{d}X_{s_k}. \tag{1.4}$$

Since the $1-$variation norm of X equals to the L^1 norm of the derivative of X, we have (see for example [4])

$$\left| Y_t - Y_s - \sum_{k=1}^{N} f^{\circ k}\left(Y_s\right) X_{s,t}^{k} \right| \leq \left\| f^{\circ(N+1)} \right\|_{\infty} \frac{|X|_{1-var;[s,t]}^{N+1}}{N!} \tag{1.5}$$

where

$$|X|_{1-var;[s,t]} = \sup_{s<t_1<\ldots<t_n<t} \sum_{i=0}^{n-1} \left| X_{t_{i+1}} - X_{t_i} \right|$$

and $\left\| f^{\circ N} \right\|_{\infty}$ denotes $\sup_{x\in\mathbb{R}^e} \left| f^{\circ N}(x) \right|$ with $|\cdot|$ being the operator norm

$$\left| f^{\circ N}(x) \right| = \sup_{v\in\left(\mathbb{R}^d\right)^{\otimes N}} \frac{\left| f^{\circ N}(x)(v) \right|}{\|v\|}.$$

Estimates of the form (1.5) have application both as a theoretical tool for analysing the equation (1.1) and as a practical numerical scheme for constructing the solution. The estimate (1.5), when the 1-variation metric is replaced by the p-variation metric, has been shown in [2] ($p<3$), [5] ($p<3$) and [4] (all $p \geq 1$) without the factorial decay factor. We shall prove such estimate *with* the factorial decay factor. The estimates of Davie [2], Gubinelli [5], Friz and Victoir [4] as well as our estimates below gives a numerical scheme for approximating a solution to (1.1) in $O(1)$ time steps.

Theorem 1.1. *Let $p \geq 1$. Let $X = \left(1, X^1, \ldots, X^{\lfloor p \rfloor}\right)$ be a p-weak geometric rough path. Let f be a Lip($\gamma - 1$) vector field where $\gamma > p$. Let Y be a solution to the differential equation*

$$\mathrm{d}Y_t = f\left(Y_t\right) \mathrm{d}X_t \tag{1.6}$$

defined in the sense of [3]. Then there exists a constant C_p depending only on p such that

$$\left| Y_t - Y_s - \sum_{k=1}^{\lfloor \gamma \rfloor} f^{\circ k}\left(Y_s\right) X_{s,t}^{k} \right| \leq \frac{1}{\left(\frac{\lfloor \gamma \rfloor}{p}\right)!} \beta^{\lfloor \gamma \rfloor} M_{p,\gamma} \|f\|_{\circ\gamma} \|X\|_{p-var,[s,t]}^{\gamma}, \tag{1.7}$$

where

$$\begin{aligned} M_{p,\gamma} &= 2C_p \left(|f|_{Lip((\gamma-1)\wedge\lfloor p \rfloor)} \vee 1 \right)^{\lfloor p \rfloor + 1} \left(|X|_{p-var} \vee 1 \right)^{\lfloor p \rfloor + 1}; \\ \|f\|_{\circ\gamma} &= \max_{\lfloor \gamma \rfloor - \lfloor p \rfloor + 1 \leq m \leq \lfloor \gamma \rfloor} \left| f^{\circ m} \right|_{Lip(\min(\gamma-m,1))}^{\min(\gamma-m,1)}; \qquad (1.8) \\ \beta &= p \left(1 + \sum_{r=2}^{\infty} \left(\frac{2}{r-1} \wedge 1 \right)^{\frac{\lfloor p \rfloor + 1}{p}} \right). \qquad (1.9) \end{aligned}$$

We refer the readers to Definition 9.16 and Definition 10.2 in [3] for the definition of Lip (γ) vector fields and weak geometric rough paths respectively. We shall however recall the definition of p-variation and some basic notations in Section 2.

Remark 1.2. If the equation (1.6) has more than one solution, then any solution must satisfy (1.7).

Remark 1.3. Taking the biggest γ may not yield the best estimate for the left hand side of (1.7). In general the term $\|f\|_{\circ\gamma}$ could grow factorially fast in γ. Since a Lip(γ) function is also Lip(γ') for all $\gamma' < \gamma$, we may choose γ' which optimises the estimate (1.7).

The proof for (1.5) relies heavily on the relation between the 1-variation of the path and the L^1 norm of its derivative. Proving an estimate of the form (1.5) for the p-variation metric, even without the factorial decay factor, requires the clever idea of Young[9]. The integration with respect to a path can be expressed in terms of the limit of a Riemann sum as the size of partition converges to zero. Young's idea was to estimate the Riemann sum with respect to a partition by removing points from the partition successively. This idea had been used in [6] to show that, for $p < 2$, the n-th order iterated integral of a path X is uniformly bounded by

$$\left(1 + 4^{\frac{1}{p}}\zeta\left(2/p\right)\right)^{n}\left(\frac{1}{n!}\right)^{\frac{1}{p}}\|X\|^{n}_{p-var,[0,T]}\,. \tag{1.10}$$

where ζ is the classical zeta function. T. Lyons' proof for the $p \geq 2$ case in [7] is slightly different and used the neoclassical inequality ([7],[1])

$$\sum_{k=0}^{N}\frac{1}{\Gamma\left(k/p+1\right)\Gamma\left(\left(n-k\right)/p+1\right)}a^{k/p}b^{(n-k)/p} \leq p\frac{1}{\Gamma\left(n/p+1\right)}\left(a+b\right)^{n/p} \tag{1.11}$$

to obtain an uniform bound of the form

$$\beta^{n-1}\frac{1}{\Gamma\left(n/p+1\right)}\|X\|^{n}_{p-var,[0,T]}$$

where Γ is the Gamma function and β is as defined in (1.9).

2 The Proof

2.1 Notations and basic definitions

For each $k \in \mathbb{N}$, we equip a norm on $\left(\mathbb{R}^d\right)^{\otimes k}$ by identifying it with $\mathbb{R}^{d^k}$. Let

$$T_1^N\left(\mathbb{R}^d\right) = 1 \oplus \mathbb{R}^d \oplus \ldots \oplus \left(\mathbb{R}^d\right)^N.$$

If π_k denotes the projection operator $T_1^N\left(\mathbb{R}^d\right) \to \left(\mathbb{R}^d\right)^{\otimes k}$, then we define a norm on $T_1^N\left(\mathbb{R}^d\right)$ by

$$\|x\| = \max_{1\leq k\leq N}\|\pi_k\left(x\right)\|^{\frac{1}{k}}.$$

Definition 2.1. *Let $T > 0$ and $p \geq 1$. A path $X : [0,T] \to T_1^{\lfloor p\rfloor}\left(\mathbb{R}^d\right)$ has finite p-variation if for all $0 < s < t < T$,*

$$\|X\|_{p-var,[s,t]} := \sup_{s<t_1<\ldots<t_n<t}\max_{1\leq k\leq\lfloor p\rfloor}\left(\sum_{i=0}^{n-1}\left\|\pi_k\left(X_{t_i}^{-1}X_{t_{i+1}}\right)\right\|^{\frac{p}{k}}\right)^{\frac{1}{p}} < \infty \tag{2.1}$$

where X^{-1} denote the unique multiplicative inverse of $X \in T_1^{\lfloor p\rfloor}\left(\mathbb{R}^d\right)$. We will denote $\|X\|_{p-var,[0,T]}$ by $\|X\|_{p-var}$.

We first recall Lyons' extension theorem, which will be used repeatedly in the following form:

Fact 2.2. *(Theorem 2.2.1 in [7]) Let $p \geq 1$ and $X = \left(1, X^1, \ldots, X^{\lfloor p\rfloor}\right)$ be a p-weak geometric rough path. Then for all $N \geq \lfloor p\rfloor + 1$, there exists a unique continuous* *path* $\mathbf{X} = \left(1, X^1, \ldots, X^N\right) \in T_1^N\left(\mathbb{R}^d\right)$ *which extends* X*,* $\mathbf{X}_0 = (1, 0 \ldots, 0)$ *and for all* $\lfloor p \rfloor \leq l \leq N$*,*

$$\left\|\pi_l\left(\mathbf{X}_{t_i}^{-1}\mathbf{X}_{t_{i+1}}\right)\right\| \leq \frac{\beta^{l-1}}{\left(\frac{l}{p}\right)!}\|X\|_{p-var,[s,t]}^{l}. \tag{2.2}$$

Remark 2.3. We will denote $\mathbf{X}_s^{-1}\mathbf{X}_t$ by $\mathbf{X}_{s,t}$ and $\pi_l\left(\mathbf{X}_{s,t}\right)$ by $X_{s,t}^l$. In particular, $\mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} = \mathbf{X}_{s,t}$ and so, for any $s < u < t$,

$$X_{s,t}^m = \sum_{l=0}^{m} X_{s,u}^{m-l} \otimes X_{u,t}^l. \tag{2.3}$$

Note that for paths with finite 1-variation, the $\left(X^k\right)_{k\geq 1}$ defined in this theorem are exactly the iterated integrals of X. Hence no confusion will arise by using the same notation as in (1.4).

Remark 2.4. If $r \geq \lfloor p \rfloor$, then for any $m \geq 0$,

$$X_{s,t}^m = \lim_{|\mathcal{P}|\to 0}\sum_{i=0}^{n-1}\sum_{k=1}^{r} X_{s,t_i}^{m-k} \otimes X_{t_i,t_{i+1}}^k \tag{2.4}$$

where the limit is taken as the mesh size of the partition $\mathcal{P} = (s < t_1 < \ldots < t_{n-1} < t)$ goes to zero. By convention, for any $s < t$, $X_{s,t}^0 = 1$ and $X_{s,t}^m = 0$ if $m < 0$. In the case $r = m$, (2.4) follows directly from (2.3). For $r < m$, note that the sum over k from $r+1$ to m in (2.4) vanishes after the taking of limit, due to (2.2). See [5] for details.

2.2 The proof

The following lemma is a factorial decay estimate for the Taylor remainder of a controlled path in the sense of Gubinelli [5]. This lemma is interesting in its own right. We interpret it as the dual counterpart of Fact 2.2.

Lemma 2.5. *Let* $p \geq 1$ *and* $\gamma > p$*. Let* $\left(1, X^1, \ldots, X^{\lfloor p \rfloor}\right)$ *be a* p*-weak geometric rough path. Let* $Y^{(i)}$ *be a function* $[0,T] \to L\left(\left(\mathbb{R}^d\right)^{\otimes i}, \mathbb{R}^e\right)$ *and* $\left(Y^{(0)}, Y^{(1)}, \ldots, Y^{(\lfloor\gamma\rfloor)}\right)$ *satisfies, for* $\lceil \gamma - p \rceil \leq m \leq \lfloor \gamma \rfloor$*,*

$$\left|Y_t^{(m)} - \sum_{l=0}^{\lfloor\gamma\rfloor - m} Y_s^{(l+m)} X_{s,t}^l\right| \leq \frac{1}{\left(\frac{\lfloor\gamma\rfloor - m}{p}\right)!} M\beta^{\lfloor\gamma\rfloor - m}\|X\|_{p-var,[s,t]}^{\gamma - m}, \tag{2.5}$$

for all $s \leq t$ *and for* $0 \leq m \leq \lceil \gamma - p \rceil - 1$*, the limit*

$$\lim_{|\mathcal{P}|\to 0}\sum_{i=0}^{n-1}\sum_{l=1}^{\lfloor\gamma\rfloor - m} Y_{t_i}^{(m+l)} X_{t_i,t_{i+1}}^l, \tag{2.6}$$

where $|\mathcal{P}| \to 0$ *denotes the limit as the mesh size of a partition* $\mathcal{P}$ *on* $[s,t]$ *goes to zero, exists and equals*

$$Y_t^{(m)} - Y_s^{(m)}. \tag{2.7}$$

For $l \geq \lfloor p \rfloor + 1$*, let* X^l *denote the projection to* $\left(\mathbb{R}^d\right)^{\otimes l}$ *of the unique extension of* $\left(1, X^1, \ldots, X^{\lfloor p \rfloor}\right)$ *given in Fact 2.2. Then (2.5) holds for all* $0 \leq m \leq \lfloor \gamma \rfloor$*.*

Proof. We will carry out backward induction on k starting from $\lceil \gamma - p \rceil$ and moving down to 0.

The base induction step of $k = \lceil \gamma - p \rceil$ holds because of the assumption. We will assume from now onwards that $k \leq \lceil \gamma - p \rceil - 1$. It is useful to bear in mind that

$$\lfloor \gamma \rfloor - \lfloor p \rfloor \leq \lceil \gamma - p \rceil \leq \lfloor \gamma \rfloor - \lfloor p \rfloor + 1.$$

For the induction step, note that by (2.4) and the equality of (2.6) and (2.7),

$$Y_t^{(k)} - \sum_{l=0}^{\lfloor \gamma \rfloor - k} Y_s^{(k+l)} X_{s,t}^l \tag{2.8}$$

$$= \lim_{|\mathcal{P}| \to 0} \sum_{i=0}^{n} \sum_{l_2=1}^{\lfloor \gamma \rfloor - k} \left(Y_{t_i}^{(k+l_2)} - \sum_{l_1=0}^{\lfloor \gamma \rfloor - k - l_2} Y_s^{(k+l_1+l_2)} X_{s,t_i}^{l_1} \right) X_{t_i,t_{i+1}}^{l_2}, \tag{2.9}$$

where the limit is taken as the mesh size of the partition $\mathcal{P} = (s < t_1 < \ldots < t_{n-1} < t)$ goes to zero.

We first show that the term

$$\sum_{i=0}^{n-1} \sum_{l_2=1}^{\lfloor \gamma \rfloor - k} \sum_{l_1=0}^{\lfloor \gamma \rfloor - k - l_2} Y_s^{(k+l_1+l_2)} X_{s,t_i}^{l_1} X_{t_i,t_{i+1}}^{l_2}. \tag{2.10}$$

is in fact independent of the partition $\mathcal{P}$.

$$\begin{aligned}
&\sum_{i=0}^{n-1} \sum_{l_2=1}^{\lfloor \gamma \rfloor - k} \sum_{l_1=0}^{\lfloor \gamma \rfloor - k - l_2} Y_s^{(k+l_1+l_2)} X_{s,t_i}^{l_1} X_{t_i,t_{i+1}}^{l_2} \\
= \; & \sum_{i=0}^{n-1} \left[\sum_{0 \leq l_1 + l_2 \leq \lfloor \gamma \rfloor - k} Y_s^{(k+l_1+l_2)} X_{s,t_i}^{l_1} X_{t_i,t_{i+1}}^{l_2} - \sum_{l_1=0}^{\lfloor \gamma \rfloor - k} Y_s^{(k+l_1)} X_{s,t_i}^{l_1} \right] \\
= \; & \sum_{i=0}^{n-1} \left[\sum_{r=0}^{\lfloor \gamma \rfloor - k} \sum_{l_1+l_2=r} Y_s^{(k+r)} X_{s,t_i}^{l_1} X_{t_i,t_{i+1}}^{l_2} - \sum_{l_1=0}^{\lfloor \gamma \rfloor - k} Y_s^{(k+l_1)} X_{s,t_i}^{l_1} \right] \\
= \; & \sum_{i=0}^{n-1} \left[\sum_{r=0}^{\lfloor \gamma \rfloor - k} Y_s^{(k+r)} X_{s,t_{i+1}}^{r} - \sum_{r=0}^{\lfloor \gamma \rfloor - k} Y_s^{(k+r)} X_{s,t_i}^{r} \right] \\
= \; & \sum_{r=1}^{\lfloor \gamma \rfloor - k} Y_s^{(k+r)} X_{s,t}^{r}
\end{aligned}$$

where we have used (2.3) in the third line. Let

$$\left(Y_s^{(k)} - \sum_{l=0}^{\lfloor \gamma \rfloor - k} Y_s^{(l)} X_{s,t}^l \right)^{\mathcal{P}} = \sum_{i=0}^{n-1} \sum_{l_2=1}^{\lfloor \gamma \rfloor - k} \left(Y_{t_i}^{(k+l_2)} - \sum_{l_1=0}^{\lfloor \gamma \rfloor - k - l} Y_s^{(k+l+l_1)} X_{s,t_i}^{l_1} \right) X_{t_i,t_{i+1}}^{l_2}.$$

Since (2.10) is independent of the partition,

$$\left(Y_s^{(k)} - \sum_{l=0}^{\lfloor \gamma \rfloor - k} Y_s^{(l)} X_{s,t}^l \right)^{\mathcal{P}} - \left(Y_s^{(k)} - \sum_{l=0}^{\lfloor \gamma \rfloor - k} Y_s^{(l)} X_{s,t}^l \right)^{\mathcal{P} \setminus \{t_j\}} \tag{2.11}$$

$$\begin{aligned}
&= \sum_{l'=1}^{\lfloor \gamma \rfloor - k} Y_{t_{j-1}}^{(k+l')} X_{t_{j-1},t_j}^{l'} + \sum_{l'=1}^{\lfloor \gamma \rfloor - k} Y_{t_j}^{(k+l')} X_{t_j,t_{j+1}}^{l'} - \sum_{l'=1}^{\lfloor \gamma \rfloor - k} Y_{t_{j-1}}^{(k+l')} X_{t_{j-1},t_{j+1}}^{l'} \\
&= \sum_{l_2=1}^{\lfloor \gamma \rfloor - k} \left(Y_{t_j}^{(k+l_2)} - \sum_{l_1=0}^{\lfloor \gamma \rfloor - k - l_2} Y_{t_{j-1}}^{(k+l_1+l_2)} X_{t_{j-1},t_j}^{l_1} \right) X_{t_j,t_{j+1}}^{l_2}.
\end{aligned} \tag{2.12}$$

By induction hypothesis, (2.5) which holds for $m > k$ and Theorem 2.2.1 in [7],

$$\begin{aligned}
&\left|\sum_{l_2=1}^{\lfloor\gamma\rfloor-k}\left(Y_{t_j}^{(k+l_2)}-\sum_{l_1=0}^{\lfloor\gamma\rfloor-k-l}Y_{t_{j-1}}^{(k+l_1+l_2)}X_{t_{j-1},t_j}^{l_1}\right)X_{t_j,t_{j+1}}^{l_2}\right|\\
\leq\ & \sum_{l_2=1}^{\lfloor\gamma\rfloor-k}\left[\frac{1}{\left(\frac{\lfloor\gamma\rfloor-k-l_2}{p}\right)!\left(\frac{l_2}{p}\right)!}M\beta^{\lfloor\gamma\rfloor-k-l_2}\left\|X\right\|_{p-var,[t_{j-1},t_j]}^{\gamma-k-l_2}\right.\\
&\left.\times\beta^{l_2-1}\left\|X\right\|_{p-var,[t_j,t_{j+1}]}^{l_2}\right] \qquad (2.13)\\
\leq\ & \frac{1}{\left(\frac{\lfloor\gamma\rfloor-k}{p}\right)!}\frac{p}{\beta}M\beta^{\lfloor\gamma\rfloor-k}\left\|X\right\|_{p-var,[t_{j-1},t_{j+1}]}^{\gamma-k}, \qquad (2.14)
\end{aligned}$$

where the final line is obtained by the neoclassical inequality (1.11), proved in [1].

Let $\omega(s,t)=\|X\|_{p-var,[s,t]}^{p}$. We now choose j such that, for $|\mathcal{P}|\geq 2$,

$$\omega(t_{j-1},t_{j+1})\leq\left(\frac{2}{|\mathcal{P}|-1}\wedge 1\right)\omega(s,t)$$

which exists since

$$\sum_{i=1}^{n-1}\omega(t_{i-1},t_{i+1})\leq 2\omega(s,t)$$

and also that

$$\omega(t_{j-1},t_{j+1})\leq\omega(s,t)$$

for all j. Then as $\gamma-k\geq\lfloor p\rfloor+1$, (2.14) is less than or equal to

$$\frac{1}{\left(\frac{\lfloor\gamma\rfloor-k}{p}!\right)}\frac{p}{\beta}M\beta^{\lfloor\gamma\rfloor-k}\left(\frac{2}{n-1}\wedge 1\right)^{\frac{\lfloor p\rfloor+1}{p}}\left\|X\right\|_{p-var,[s,t]}^{\gamma-k}.$$

By removing points successively from $\mathcal{P}$ and using that $\left(Y_s^{(k)}-\sum_{l=0}^{\lfloor\gamma\rfloor-k}Y_s^{(k+l)}X_{s,t}^{l}\right)^{\{s,t\}}=0$, we have

$$\begin{aligned}
\left|\left(Y_s^{(k)}-\sum_{l=0}^{\lfloor\gamma\rfloor-k}Y_s^{(k+l)}X_{s,t}^{l}\right)^{\mathcal{P}}\right| &\leq \frac{1}{\left(\frac{\lfloor\gamma\rfloor-k}{p}!\right)}\frac{p}{\beta}M\beta^{\lfloor\gamma\rfloor-k}\sum_{n=2}^{\infty}\left(\frac{2}{n-1}\wedge 1\right)^{\frac{\lfloor p\rfloor+1}{p}}\left\|X\right\|_{p-var,[s,t]}^{\gamma-k}\\
&\leq \frac{1}{\left(\frac{\lfloor\gamma\rfloor-k}{p}!\right)}M\beta^{\lfloor\gamma\rfloor-k}\left\|X\right\|_{p-var,[s,t]}^{\gamma-k},
\end{aligned}$$

where the final line follows from (1.9).

By taking limit as $|\mathcal{P}|\to 0$, (2.5) follows for $m=k$. □

For the differential equation

$$\mathrm{d}Y_t=f(Y_t)\,\mathrm{d}X_t \qquad (2.15)$$

we wish to apply Lemma 2.5 to $\left(Y,f^{\circ 1}(Y),\ldots,f^{\circ(\lfloor\gamma\rfloor)}(Y)\right)$. Using the standard estimates for rough differential equations, it turns out that it suffices to verify the assumption of Lemma 2.5 for paths with finite 1-variation. To do so, we need the following lemma.

Lemma 2.6. *Let $X : [0,T] \to \mathbb{R}^d$ be a path with finite 1-variation. Let f be a Lip($\gamma-1$) vector field. Let Y_t be a solution to the differential equation (2.15). Then*

$$\begin{aligned}&f^{\circ m}(Y_t) - f^{\circ m}(Y_s) - \sum_{k=1}^{\lfloor\gamma\rfloor - m} f^{\circ(m+k)}(Y_s)\, X^k_{s,t}\\ = &\begin{cases}\int_{s\le s_1\le\ldots\le s_{\lfloor\gamma\rfloor-m}\le t} f^{\circ\lfloor\gamma\rfloor}(Y_{s_1}) - f^{\circ\lfloor\gamma\rfloor}(Y_s)\,\mathrm{d}X_{s_1}\otimes\ldots\otimes\mathrm{d}X_{s_{\lfloor\gamma\rfloor-m}} & ,0\le m<\lfloor\gamma\rfloor\\ f^{\circ\lfloor\gamma\rfloor}(Y_t) - f^{\circ\lfloor\gamma\rfloor}(Y_s) & ,m=\lfloor\gamma\rfloor.\end{cases}\end{aligned}$$

Proof. We will prove it by backward induction, starting from $\lfloor\gamma\rfloor$.

The case $m = \lfloor\gamma\rfloor$ is trivially true.

For the induction step, note first that by the fundamental theorem of calculus,

$$\begin{aligned}&\int_s^t f^{\circ(m+1)}(Y_u)\,\mathrm{d}X_u\\ =\ &\int_s^t D\left(f^{\circ m}\right)(Y_u)\, f(Y_u)\,\mathrm{d}X_u\\ =\ &\int_s^t D\left(f^{\circ m}\right)(Y_u)\,\mathrm{d}Y_u\\ =\ &f^{\circ m}(Y_t) - f^{\circ m}(Y_s).\end{aligned} \tag{2.16}$$

Then by (2.16) and the induction hypothesis,

$$\begin{aligned}&f^{\circ m}(Y_t) - f^{\circ m}(Y_s) - \sum_{k=1}^{\lfloor\gamma\rfloor-m} f^{\circ(m+k)}(Y_s)\, X^k_{s,t}\\ =\ &\int_s^t f^{\circ m+1}\left(Y_{s_{\lfloor\gamma\rfloor-m}}\right)\mathrm{d}X_{s_{\lfloor\gamma\rfloor-m}} - \sum_{k=1}^{\lfloor\gamma\rfloor-m} f^{\circ(m+k)}(Y_s)\, X^{k-1}_{s,s_{\lfloor\gamma\rfloor-m}}\otimes\mathrm{d}X_{s_{\lfloor\gamma\rfloor-m}}\\ =\ &\int_{s\le s_1\le\ldots\le s_{\lfloor\gamma\rfloor-m}\le t} f^{\circ\lfloor\gamma\rfloor}(Y_{s_1}) - f^{\circ\lfloor\gamma\rfloor}(Y_s)\,\mathrm{d}X_{s_1}\otimes\ldots\otimes\mathrm{d}X_{s_{\lfloor\gamma\rfloor-m}}.\end{aligned}$$

□

Proof of Theorem 1. The only thing to prove is that $\left(Y, f^{\circ 1}(Y), \ldots, f^{\circ(\lfloor\gamma\rfloor)}(Y)\right)$ satisfies the assumptions of Lemma 2.5.

For each $s \le t$, let $x^{s,t} : [s,t] \to \mathbb{R}^d$ be a continuous path with finite 1-variation such that for $1 \le l \le \lfloor p\rfloor$,

$$\left(x^{s,t}\right)^l_{s,t} = X^l_{s,t}, \tag{2.17}$$

where we use the notation from (1.4) and

$$\int_s^t \left|\mathrm{d}x^{s,t}_u\right| \le c_p \left\|X\right\|_{p-var,[s,t]} \tag{2.18}$$

for a function c_p of p which is specified in [3] along with the existence of $x^{s,t}$.

Consider the differential equation

$$\begin{aligned}\mathrm{d}Y^{s,t}_u &= f\left(Y^{s,t}_u\right)\mathrm{d}x^{s,t}_u\\ Y^{s,t}_s &= Y_s.\end{aligned} \tag{2.19}$$

By Theorem 10.16 in [3], there exists a solution $Y^{s,t}$ of (2.19) such that the following estimate holds

$$\left|Y_t - Y^{s,t}_t\right| \le C_p \left|f\right|^{\gamma\wedge(\lfloor p\rfloor+1)}_{Lip((\gamma-1)\wedge\lfloor p\rfloor)} \left\|X\right\|^{\gamma\wedge(\lfloor p\rfloor+1)}_{p-var,[s,t]} \tag{2.20}$$

for some function C_p depending on p only.

Note that by (2.17) and $m \geq \lceil \gamma - p \rceil \geq \lfloor \gamma \rfloor - \lfloor p \rfloor$,

$$\begin{aligned} &\left| f^{\circ(m)}(Y_t) - \sum_{k=0}^{\lfloor\gamma\rfloor - m} f^{\circ(m+k)}(Y_s) X^k_{s,t} \right| \\ \leq \quad & \left| f^{\circ m}(Y_t) - f^{\circ m}\left(Y^{s,t}_t\right)\right| + \left| f^{\circ m}\left(Y^{s,t}_t\right) - \sum_{k=0}^{\lfloor\gamma\rfloor - m} f^{\circ(m+k)}(Y_s)\left(x^{s,t}\right)^k_{s,t}\right| \end{aligned} \tag{2.21}$$

By (2.20), for $0 \leq m \leq \lfloor\gamma\rfloor - 1$,

$$\begin{aligned} &\left| f^{\circ m}(Y_t) - f^{\circ m}\left(Y^{s,t}_t\right)\right| \\ \leq \quad & \left|f^{\circ m}\right|_{Lip(1)} \left|Y_t - Y^{s,t}_t\right| \\ \leq \quad & C_p \left|f^{\circ m}\right|_{Lip(1)} |f|^{\gamma\wedge(\lfloor p\rfloor+1)}_{Lip((\gamma-1)\wedge\lfloor p\rfloor)} \|X\|^{\gamma\wedge(\lfloor p\rfloor+1)}_{p-var,[s,t]}. \end{aligned} \tag{2.22}$$

If $\lceil \gamma - p\rceil \leq m \leq \lfloor\gamma\rfloor - 1$, then $\gamma - m \leq \lfloor p \rfloor$ and so

$$\left| f^{\circ m}(Y_t) - f^{\circ m}\left(Y^{s,t}_t\right)\right| \tag{2.23}$$

$$\leq \quad C_p \left|f^{\circ m}\right|_{Lip(1)} |f|^{\gamma\wedge(\lfloor p\rfloor+1)}_{Lip((\gamma-1)\wedge\lfloor p\rfloor)} \left(\|X\|_{p-var,[s,t]} \vee 1\right)^{(\lfloor p\rfloor+1)} \|X\|^{\gamma-m}_{p-var,[s,t]}. \tag{2.24}$$

To estimate (2.23) for $m = \lfloor\gamma\rfloor$, we note that

$$\begin{aligned} &\left| f^{\circ\lfloor\gamma\rfloor}(Y_t) - f^{\circ\lfloor\gamma\rfloor}\left(Y^{s,t}_t\right)\right| \\ \leq \quad & \left|f^{\circ\lfloor\gamma\rfloor}\right|_{Lip(\gamma-\lfloor\gamma\rfloor)} \left|Y_t - Y^{s,t}_t\right|^{\gamma-\lfloor\gamma\rfloor} \\ \leq \quad & C_p \left|f^{\circ\lfloor\gamma\rfloor}\right|_{Lip(\gamma-\lfloor\gamma\rfloor)} |f|^{\gamma\wedge(\lfloor p\rfloor+1)(\gamma-\lfloor\gamma\rfloor)}_{Lip((\gamma-1)\wedge\lfloor p\rfloor)} \|X\|^{\gamma\wedge(\lfloor p\rfloor+1)(\gamma-\lfloor\gamma\rfloor)}_{p-var,[s,t]}. \end{aligned}$$

In particular, we have

$$\begin{aligned} &\left| f^{\circ\lfloor\gamma\rfloor}(Y_t) - f^{\circ\lfloor\gamma\rfloor}\left(Y^{s,t}_t\right)\right| \\ \leq \quad & C_p \left|f^{\circ\lfloor\gamma\rfloor}\right|_{Lip(\gamma-\lfloor\gamma\rfloor)} |f|^{\gamma\wedge(\lfloor p\rfloor+1)(\gamma-\lfloor\gamma\rfloor)}_{Lip((\gamma-1)\wedge\lfloor p\rfloor)} \left(\|X\|_{p-var,[s,t]} \vee 1\right)^{(\lfloor p\rfloor+1)} \|X\|^{\gamma-\lfloor\gamma\rfloor}_{p-var,[s,t]}. \end{aligned}$$

To estimate the second term in (2.21), we use Lemma 2.6 to see that for $\lceil\gamma - p\rceil \leq m \leq \lfloor\gamma\rfloor$,

$$\left| f^{\circ m}\left(Y^{s,t}_t\right) - \sum_{k=0}^{\lfloor\gamma\rfloor - m} f^{\circ(m+k)}(Y_s)\left(x^{s,t}\right)^k_{s,t}\right|$$

$$= \quad \left| \int_{s\leq s_1\leq\ldots\leq s_{\lfloor\gamma\rfloor-m}<t} f^{\circ(\lfloor\gamma\rfloor)}\left(Y^{s,t}_{s_1}\right) - f^{\circ(\lfloor\gamma\rfloor)}(Y_s)\,\mathrm{d}x^{s,t}_{s_1}\ldots\mathrm{d}x^{s,t}_{s_{\lfloor\gamma\rfloor-m}}\right| \tag{2.25}$$

$$\leq \quad C_p^{\lfloor\gamma\rfloor-m} \left|f^{\circ\lfloor\gamma\rfloor}\right|_{Lip(\gamma-\lfloor\gamma\rfloor)} \left|Y^{s,t}_{\cdot}\right|^{\gamma-\lfloor\gamma\rfloor}_{p-var,[s,t]} \|X\|^{\lfloor\gamma\rfloor-m}_{p-var,[s,t]}$$

$$\leq \quad C'_p \left|f^{\circ\lfloor\gamma\rfloor}\right|_{Lip(\gamma-\lfloor\gamma\rfloor)} \left(|f|_{Lip((\gamma-1)\wedge\lfloor p\rfloor)} \vee 1\right)^{p(\gamma-\lfloor\gamma\rfloor)} \tag{2.26}$$

$$\times \left(\|X\|_{p-var,[s,t]} \vee 1\right)^{(p-1)(\gamma-\lfloor\gamma\rfloor)} \|X\|^{\gamma-m}_{p-var,[s,t]}, \tag{2.27}$$

where in the third line we have used the $\gamma-\lfloor\gamma\rfloor$ Hölder continuity of $f^{\circ(\lfloor\gamma\rfloor)}$ with (2.18) and in the final line we have used Theorem 10.16 in [3].

Combining (2.21), (2.23) and (2.26), we have for $\lceil\gamma-p\rceil\leq m\leq\lfloor\gamma\rfloor$,

$$\begin{aligned}
&\left|f^{\circ(m)}\left(Y_t\right)-\sum_{k=0}^{\lfloor\gamma\rfloor-m}f^{\circ(m+k)}\left(Y_s\right)X_{s,t}^k\right|\\
\leq\;& 2C_p\max_{\lfloor\gamma\rfloor-\lfloor p\rfloor+1\leq m\leq\lfloor\gamma\rfloor}\left|f^{\circ m}\right|_{Lip(\min(\gamma-m,1))}^{\min(\gamma-m,1)}\left(|f|_{Lip((\gamma-1)\wedge\lfloor p\rfloor)}\vee 1\right)^{\lfloor p\rfloor+1}\\
&\times\left(\|X\|_{p-var}\vee 1\right)^{\lfloor p\rfloor+1}\|X\|_{p-var,[s,t]}^{\gamma-m}.
\end{aligned}\tag{2.28}$$

Here since $\lceil\gamma-p\rceil\leq m\leq\lfloor\gamma\rfloor$ so $\lfloor\gamma\rfloor-m\leq\lfloor p\rfloor$ and

$$(\lfloor\gamma\rfloor-m)!\leq\lfloor p\rfloor!.$$

Therefore, by changing the constant C_p, we rewrite (2.28) in the form of the right hand side of (2.5). It now suffices to show (2.7). Note first that for $0\leq m\leq\lceil\gamma-p\rceil-1$ and $s\leq u\leq v\leq t$,

$$\left|f^{\circ m}\left(Y_v\right)-\sum_{k=0}^{\lfloor\gamma\rfloor-m}f^{\circ(m+k)}\left(Y_u\right)X_{u,v}^k\right|\tag{2.29}$$

$$\leq\left|f^{\circ m}\left(Y_v\right)-f^{\circ m}\left(Y_v^{u,v}\right)\right|+\left|f\left(Y_v^{u,v}\right)-\sum_{k=0}^{\lfloor\gamma\rfloor-m}f^{\circ(m+k)}\left(Y_u\right)\left(x^{u,v}\right)_{u,v}^k\right|\tag{2.30}$$

$$+\left|\sum_{k=\lfloor p\rfloor+1}^{\lfloor\gamma\rfloor-m}f^{\circ(m+k)}\left(Y_u\right)\left(x^{u,v}\right)_{u,v}^k-\sum_{k=\lfloor p\rfloor+1}^{\lfloor\gamma\rfloor-m}f^{\circ(m+k)}\left(Y_u\right)X_{u,v}^k\right|.\tag{2.31}$$

The estimate (2.22) still holds with (s,t) replaced by (u,v) and (2.26) would hold with the constant C_p now depending on γ as well. For the final term in (2.31),

$$\begin{aligned}
&\left|\sum_{k=\lfloor p\rfloor+1}^{\lfloor\gamma\rfloor-m}f^{\circ(m+k)}\left(Y_u\right)\left(x^{u,v}\right)_{u,v}^k-\sum_{k=\lfloor p\rfloor+1}^{\lfloor\gamma\rfloor-m}f^{\circ(m+k)}\left(Y_u\right)X_{u,v}^k\right|\\
\leq\;&\left|\sum_{k=\lfloor p\rfloor+1}^{\lfloor\gamma\rfloor-m}f^{\circ(m+k)}\left(Y_u\right)\left(x^{u,v}\right)_{u,v}^k\right|+\left|\sum_{k=\lfloor p\rfloor+1}^{\lfloor\gamma\rfloor-m}f^{\circ(m+k)}\left(Y_u\right)X_{u,v}^k\right|\\
\leq\;&2\lfloor\gamma\rfloor c_p^{\lfloor\gamma\rfloor}\max_{0\leq m\leq\lfloor\gamma\rfloor}\sup_{s\leq u\leq t}\left|f^{\circ m}\left(Y_u\right)\right|\left(\|X\|_{p-var,[s,t]}\vee 1\right)^{\lfloor\gamma\rfloor}\|X\|_{p-var,[u,v]}^{\lfloor p\rfloor+1}
\end{aligned}$$

where we used Fact 2.2 and

$$\begin{aligned}
\left|\left(x^{u,v}\right)_{u,v}^k\right|&\leq c_p^k\left(\int_u^v\left|\mathrm{d}x_r^{u,v}\right|\right)^k\\
&\leq C_p^k\|X\|_{p-var,[u,v]}^k.
\end{aligned}$$

Therefore, combining with (2.22) and (2.26), we have for some constants $C_{f,p,X,s,t\gamma},C'_{f,p,X,s,t\gamma}$

independent of u, v such that when $|u-v|$ is sufficiently small,

$$\begin{aligned}
&\left| f^{\circ m}(Y_v) - \sum_{k=0}^{\lfloor\gamma\rfloor-m} f^{\circ(m+k)}(Y_u) X_{u,v}^k \right| \\
\leq\ & C_{f,p,X,s,t\gamma}\left(\|X\|_{p-var,[u,v]}^{\gamma\wedge(\lfloor p\rfloor+1)} + \|X\|_{p-var,[u,v]}^{\gamma-m} + \|X\|_{p-var,[u,v]}^{\lfloor p\rfloor+1} \right) \\
\leq\ & C'_{f,p,X,s,t\gamma} \|X\|_{p-var,[u,v]}^{\gamma\wedge(\lfloor p\rfloor+1)}
\end{aligned}$$

Denote the expression in (2.29) as $E(u,v)$. Let $\lim_{|\mathcal{P}|\to 0}$ denote the limit as the mesh size of a partition $\mathcal{P}$ on $[s,t]$ goes to zero. Then for $m \leq \lceil\gamma-p\rceil-1$,

$$\begin{aligned}
&\lim_{|\mathcal{P}|\to 0}\sum_{i=0}^{n-1}\sum_{l=1}^{\lfloor\gamma\rfloor-m} E(t_i,t_{i+1}) \\
\leq\ & C'_{f,p,X,s,t\gamma}\lim_{|\mathcal{P}|\to 0}\sum_{i=0}^{n-1}\|X\|_{p-var,[t_i,t_{i+1}]}^{\gamma\wedge(\lfloor p\rfloor+1)} &(2.32)\\
\leq\ & C'_{f,p,X,\gamma}\lim_{|\mathcal{P}|\to 0}\max_i \|X\|_{p-var,[t_i,t_{i+1}]}^{\gamma\wedge(\lfloor p\rfloor+1)-p}\sum_{i=0}^{n-1}\|X\|_{p-var,[t_i,t_{i+1}]}^{p} &(2.33)
\end{aligned}$$

Since for $s<u<t$,

$$\|X\|_{p-var,[s,u]}^p + \|X\|_{p-var,[u,t]}^p \leq \|X\|_{p-var,[s,t]}^p ,$$

(2.33) is bounded by

$$C_{f,p,X,\gamma}\lim_{|\mathcal{P}|\to 0}\max_i \|X\|_{p-var,[t_i,t_{i+1}]}^{\gamma\wedge(\lfloor p\rfloor+1)-p}\|X\|_{p-var,[s,t]}^{p} ,$$

which equals 0 by the uniform continuity of the map $(u,v)\to\|X\|_{p-var,[u,v]}^p$ (See [8]). Finally,

$$\begin{aligned}
&\lim_{|\mathcal{P}|\to 0}\sum_{i=0}^{n-1}\sum_{l=1}^{\lfloor\gamma\rfloor-m} f^{\circ(m+l)}(Y_{t_i}) X_{t_i,t_{i+1}}^l \\
=\ & \lim_{|\mathcal{P}|\to 0}\sum_{i=0}^{n-1} f^{\circ m}\left(Y_{t_{i+1}}\right) - f^{\circ m}(Y_{t_i}) + E(t_i,t_{i+1}) \\
=\ & f^{\circ m}(Y_t) - f^{\circ m}(Y_s) .
\end{aligned}$$

□

References

[1] K. Hara and M. Hino. Fractional order Taylor's series and the neo-classical inequality, *Bull. Lond. Math. Soc.*, 42 467–477, 2010. MR-2651942

[2] A. Davie, Differential equations driven by rough paths: an approach via discrete approximation, *Appl. Math. Res. Express*, AMRX, (2):Art. ID abm009, 40, 2007. MR-2387018

[3] P. Friz, N. Victoir, *Multidimensional Stochastic Processes as Rough Paths. Theory and Applications*, Cambridge Studies of Advanced Mathematics, Vol. 120, Cambridge University Press, 2010. MR-2604669

[4] P. Friz, N. Victoir, Euler estimates for rough differential equations, *J. Differential Equations*, 244(2):388-412, 2008. MR-2376201

[5] M. Gubinelli, Controlling Rough Paths, *J. Funct. Anal.*, 216:86-140, 2004. MR-2091358

[6] T. Lyons, Differential equations driven by rough signals (I): an extension of an inequality of L. C. Young, *Mathematical Research Letters 1*, 451-464, 1994. MR-1302388

[7] T. Lyons, Differential equations driven by rough signals, *Rev. Mat. Iberoamericana.*, Vol. 14 (2), 215–310, 1998. MR-1654527

[8] P. Yam, *Analytical and Topological Aspects of Signatures*, D.Phil Thesis, available at `http://ora.ox.ac.uk/objects/uuid%3A87892930-f329-4431-bcdc-bf32b0b1a7c6/datastreams/ATTACHMENT1`, 2008.

[9] L. C. Young. An inequality of Hölder type connected with Stieltjes integration. *Acta Math.*, (67):251–282, MR-1555421 1936.

Acknowledgments. We would also like to thank H. Oberhauser for the useful discussions and the anonymous referee for the detailed suggestions.