Manuscript version: Published Version
The version presented in WRAP is the published version (Version of Record).

Persistent WRAP URL:
http://wrap.warwick.ac.uk/138859

How to cite:
The repository item page linked to above, will contain details on accessing citation guidance from the publisher.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 3.0 International license (CC BY 3.0) and may be reused according to the conditions of the license. For more details see:http://creativecommons.org/licenses/by/3.0/.

Publisher’s statement:
Please refer to the repository item page, publisher’s statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk
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{align}
\text{d}Y_t &= f(Y_t) \text{d}X_t \\
Y_0 &= y_0.
\end{align}

(1.1)

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

\begin{align}
Y_t - Y_s &= \sum_{k=1}^N f^\circ k (Y_s) \int_{s < s_1 < \ldots < s_k < t} \text{d}X_{s_1} \otimes \ldots \otimes \text{d}X_{s_k} \\
\equiv \int_{s < s_1 < \ldots < s_N < t} f^\circ N (Y_{s_1}) - f^\circ N (Y_s) \text{d}X_{s_1} \otimes \ldots \otimes \text{d}X_{s_N},
\end{align}

(1.2)

(1.3)

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

\begin{align}
f^\circ 1 &= f \\
f^\circ k+1 &= D(f^\circ k) f.
\end{align}

*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
Factorial decay estimates for differential equations

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

\[
X_{s,t}^k := \int_{s < s_1 < \ldots < s_k < t} dX_{s_1} \otimes \ldots \otimes dX_{s_k}.
\]  

(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])

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

(1.5)

where

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

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

\[
|f^{\circ N}(x)| = \sup_{v \in \mathbb{R}^n} \frac{|f^{\circ N}(x)(v)|}{\|v\|}.
\]

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 \( X = (1, X^1, \ldots, X^{[p]}) \) 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

\[
dY_t = f(Y_t) \, dX_t
\]

(1.6)

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

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

(1.7)

where

\[
M_{p, \gamma} = 2C_p \left( |f|_{Lip((\gamma-1) \wedge [p])} \lor 1 \right)^{[p]+1} \left( |X|_{p-var} \lor 1 \right)^{[p]+1};
\]

\[
\|f\|_{\gamma} = \max_{[\gamma]-[p]+1 \leq m \leq [\gamma]} \left| f_{\gamma-m,1} \right|_{Lip(\gamma-m,1)};
\]

\[
\beta = p \left( 1 + \sum_{r=2}^{\infty} \frac{2}{r-1} \wedge 1 \right)^{\frac{[p]+1}{2}}.
\]

(1.8)

We refer the readers to Definition 9.16 and Definition 10.2 in [3] for the definition of Lip(\( \gamma \)) 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 $\gamma$ may not yield the best estimate for the left hand side of (1.7). In general the term $\|f\|_{\gamma}$ could grow factorially fast in $\gamma$. Since a Lip($\gamma$) function is also Lip($\gamma'$) for all $\gamma' < \gamma$, we may choose $\gamma'$ 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

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

where $\zeta$ 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])

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

to obtain an uniform bound of the form

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

where $\Gamma$ is the Gamma function and $\beta$ 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 $(\mathbb{R}^d)^\otimes k$ by identifying it with $\mathbb{R}^{d^k}$. Let

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

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

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

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

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

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

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 $X = (1,X^1,\ldots,X^{[p]})$ be a $p$-weak geometric rough path. Then for all $N \geq [p] + 1$, there exists a unique continuous
path $X = (1, X^1, \ldots, X^N) \in T_1^N (\mathbb{R}^d)$ which extends $X$, $X_0 = (1, 0, \ldots, 0)$ and for all $[p] \leq l \leq N$, 
\[
\|\pi_l (X_{t_l}^{-1}X_{t_{l+1}})\| \leq \frac{\beta^{l-1}}{(l)!} \|X\|_{p-\text{var},[s,t]}^{l}.
\] (2.2)

**Remark 2.3.** We will denote $X^{-1}_{s,t}$ by $X_{s,t}$ and $\pi_l (X_{s,t})$ by $X^l_{s,t}$. In particular, $X_{s,u} \otimes X_{u,t} = X_{s,t}$ and so, for any $s < u < t$,
\[
X^m_{s,t} = \sum_{i=0}^{m} X^{m-i}_{s,u} \otimes X^i_{u,t}.
\] (2.3)

Note that for paths with finite 1-variation, the $(X^k)_{k \geq 1}$ 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 $r \geq [p]$, then for any $m \geq 0$,
\[
X^{m}_{s,t} = \lim_{|P| \to 0} \sum_{j=0}^{n-1} \sum_{k=1}^{r} X^{m-k}_{s,t_j} \otimes X^{k}_{t_j,t_{j+1}}
\] (2.4)

where the limit is taken as the mesh size of the partition $P = (s < t_1 \ldots < t_{n-1} < t)$ goes to zero. By convention, for any $s < t$, $X^0_{s,t} = 1$ and $X^m_{s,t} = 0$ 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 $(1, X^1, \ldots, X^{[p]})$ be a $p$-weak geometric rough path. Let $Y^{(i)}$ be a function $[0,T] \to L \left( (\mathbb{R}^d)^i, \mathbb{R}^c \right)$ and $(Y^{(0)}, Y^{(1)}, \ldots, Y^{([\gamma])})$ satisfies, for $[\gamma - p] \leq m \leq [\gamma]$,
\[
\left| Y^{(m)}_t - \sum_{l=0}^{[\gamma]-m} Y^{(l+m)}_s X^{l}_{s,t} \right| \leq \frac{1}{(\gamma - m)!} M \beta^{[\gamma]-m} \|X\|_{p-\text{var},[s,t]}^{m},
\] (2.5)

for all $s \leq t$ and for $0 \leq m \leq [\gamma - p] - 1$, the limit
\[
\lim_{|P| \to 0} \sum_{i=0}^{n-1} \sum_{l=1}^{m} Y^{(m+l)}_{t_i} X^{l}_{t_i, t_{i+1}},
\] (2.6)

where $|P| \to 0$ denotes the limit as the mesh size of a partition $P$ on $[s,t]$ goes to zero, exists and equals
\[
Y^{(m)}_t - Y^{(m)}_s.
\] (2.7)

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

**Proof.** We will carry out backward induction on $k$ starting from $[\gamma - p]$ and moving down to 0.
The base induction step of $k = \gamma - p$ holds because of the assumption. We will assume from now onwards that $k \leq \gamma - p - 1$. It is useful to bear in mind that

$$|\gamma| - |p| \leq |\gamma - p| - 1.$$  

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

$$Y^{(k)}_t = \lim_{|P| \to 0} \sum_{i=0}^{n-1} Y^{(k+l_i)}_s X_{s,t}^{l_i} = \left( \sum_{i=0}^{n-1} Y^{(k+l_i)}_s X_{s,t}^{l_i} \right) X_{t^1, t_{t+1}}^{l_1},$$

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

We first show that the term

$$\sum_{i=0}^{n-1} \sum_{l_2=1}^{\gamma-k} \sum_{l_1=0}^{\gamma-k-l_2} Y_s^{(k+l_2)} X_{s,t}^{l_1} X_{t^1, t_{t+1}}^{l_2},$$

is in fact independent of the partition $P$.

$$\sum_{i=0}^{n-1} \sum_{l_2=1}^{\gamma-k} \sum_{l_1=0}^{\gamma-k-l_2} Y_s^{(k+l_2)} X_{s,t}^{l_1} X_{t^1, t_{t+1}}^{l_2} = \sum_{i=0}^{n-1} \left[ \sum_{r=0}^{\gamma-k} Y_s^{(r)} X_{s,t}^{r} - \sum_{r=0}^{\gamma-k} Y_s^{(r)} X_{s,t}^{r} \right]$$

$$= \sum_{i=0}^{n-1} \left[ \sum_{r=0}^{\gamma-k} Y_s^{(r)} X_{s,t}^{r} - \sum_{r=0}^{\gamma-k} Y_s^{(r)} X_{s,t}^{r} \right]$$

$$= \sum_{r=0}^{\gamma-k} Y_s^{(r)} X_{s,t}^{r}$$

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

$$\left( Y_s^{(k)} - \sum_{i=0}^{n-1} Y_s^{(l_i)} X_{s,t}^{l_i} \right)^P = \sum_{i=0}^{n-1} \sum_{l_2=1}^{\gamma-k} \sum_{l_1=0}^{\gamma-k-l_2} Y_s^{(k+l_2)} X_{s,t}^{l_1} X_{t^1, t_{t+1}}^{l_2}.$$
Factorial decay estimates for differential equations

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

\[
\left| \sum_{l_2=1}^{[\gamma] - k} \left( Y_{t_j}^{(k+l_2)} - \sum_{l_1=0}^{[\gamma] - k - l_2} Y_{t_{j-1}}^{(k+l_1+l_2)} X_{t_{j-1}, t_j} \right) X_{t_j, t_{j+1}} \right|
\leq \sum_{l_2=1}^{[\gamma] - k} \frac{1}{\binom{[\gamma] - k - l_2}{p}!} M^{\beta \cdot [\gamma] - k - l_2} \|X\|^{\gamma - k - l_2}_{p \cdot \text{var}, [t_{j-1}, t_j]} \times \beta^{l_2 - 1} \|X\|^{l_2}_{p \cdot \text{var}, [t_j, t_{j+1}]}
\leq \frac{1}{\binom{[\gamma] - k}{p}!} M^{\beta \cdot [\gamma] - k} \|X\|^{\gamma - k}_{p \cdot \text{var}, [t_{j-1}, t_{j+1}]},
\]

\[
(2.13)
\]

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

Let \( \omega(s, t) = \|X\|^{p \cdot \text{var}, [s, t]} \). 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 \( \gamma - k \geq \lceil p \rceil + 1 \), (2.14) is less than or equal to

\[
\frac{1}{\binom{[\gamma] - k}{p}!} M^{\beta \cdot [\gamma] - k} \left( \frac{2}{n - 1} \wedge 1 \right) ^{\frac{\lceil p \rceil + 1}{p}} \|X\|^{\gamma - k}_{p \cdot \text{var}, [s, t]},
\]

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

\[
\left| \left( Y_s^{(k)} - \sum_{l=0}^{[\gamma] - k} Y_s^{(k+l)} X_{s, t}^{l} \right)^{\mathcal{P}} \right| \leq \frac{1}{\binom{[\gamma] - k}{p}!} M^{\beta \cdot [\gamma] - k} \sum_{n=2}^{\infty} \left( \frac{2}{n - 1} \wedge 1 \right) ^{\frac{\lceil p \rceil + 1}{p}} \|X\|^{\gamma - k}_{p \cdot \text{var}, [s, t]}
\]

\[
\leq \frac{1}{\binom{[\gamma] - k}{p}!} M^{\beta \cdot [\gamma] - k} \|X\|^{\gamma - k}_{p \cdot \text{var}, [s, t]},
\]

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

\[
dY_t = f(Y_t) \, dX_t
\]

we wish to apply Lemma 2.5 to \( (Y, f^{(1)}(Y), \ldots, f^{([\gamma])}(Y)) \). 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

$$f^{\circ m} (Y_t) - f^{\circ m} (Y_s) - \sum_{k=1}^{[\gamma]-m} f^{\circ (m+k)} (Y_s) X^k_{s,t}$$

$$= \begin{cases} \int_{s \leq s_1 \leq \ldots \leq s_{[\gamma]-m} \leq t} f^{\circ [\gamma]} (Y_{s_1}) - f^{\circ [\gamma]} (Y_s) \, dX_{s_1} \otimes \ldots \otimes dX_{s_{[\gamma]-m}}, & 0 \leq m < [\gamma] \\ f^{\circ [\gamma]} (Y_t) - f^{\circ [\gamma]} (Y_s), & m = [\gamma]. \end{cases}$$

Proof. We will prove it by backward induction, starting from $[\gamma]$.

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

$$\int_s^t f^{\circ (m+1)} (Y_u) \, dX_u$$

$$= \int_s^t D (f^{\circ m}) (Y_u) f (Y_u) \, dX_u$$

$$= \int_s^t D (f^{\circ m}) (Y_u) \, dY_u$$

$$= f^{\circ m} (Y_t) - f^{\circ m} (Y_s).$$

(2.16)

Then by (2.16) and the induction hypothesis,

$$f^{\circ m} (Y_t) - f^{\circ m} (Y_s) - \sum_{k=1}^{[\gamma]-m} f^{\circ (m+k)} (Y_s) X^k_{s,t}$$

$$= \int_s^t f^{\circ (m+1)} (Y_{s_{[\gamma]-m}}) \, dX_{s_{[\gamma]-m}} - \sum_{k=1}^{[\gamma]-m} f^{\circ (m+k)} (Y_s) X^{k-1}_{s,s_{[\gamma]-m}} \otimes dX_{s_{[\gamma]-m}}$$

$$= \int_{s \leq s_1 \leq \ldots \leq s_{[\gamma]-m} \leq s} f^{\circ [\gamma]} (Y_{s_1}) - f^{\circ [\gamma]} (Y_s) \, dX_{s_1} \otimes \ldots \otimes dX_{s_{[\gamma]-m}}.$$

\[ \square \]

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

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

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

(2.17)

where we use the notation from (1.4) and

$$\int_s^t |dx^{s,t}_u| \leq c_p \|X\|_{p-var,[s,t]}$$

(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

$$dY^{s,t}_u = f \left( Y^{s,t}_u \right) dx^{s,t}_u$$

$$Y^{s,t}_s = Y_s.$$

(2.19)

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

$$|Y_t - Y^{s,t}_t| \leq C_p \|f\|_{\text{Lip}([\gamma-1] \wedge [p])} \|X\|_{p-var,[s,t]}^{\gamma ([p]+1)}$$

(2.20)
Factorial decay estimates for differential equations

for some function $C_p$ depending on $p$ only.

Note that by (2.17) and $m \geq |\gamma - p| \geq |\gamma| - [p],

\[
\left| f^{\circ m} (Y_t) - \sum_{k=0}^{[\gamma]-m} f^{\circ (m+k)} (Y_s) X_s^k \right| \leq \left| f^{\circ m} (Y_t) - f^{\circ m} (Y_t^{s,t}) \right| + \left| f^{\circ m} (Y_t^{s,t}) - \sum_{k=0}^{[\gamma]-m} f^{\circ (m+k)} (Y_s) (x^{s,t})_s^k \right|
\]

(2.21)

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

\[
\left| f^{\circ m} (Y_t) - f^{\circ m} (Y_t^{s,t}) \right| \leq \left| f^{\circ m} \right|_{Lip(\gamma-1)} \left| Y_t - Y_t^{s,t} \right| \leq C_p \left| f^{\circ m} \right|_{Lip(\gamma-1)} \left[ f^{\gamma \wedge ([p]+1)} \| X \|_{p-var,[s,t]} \right].
\]

(2.22)

If $|\gamma - p| \leq m \leq |\gamma| - 1,$ then $\gamma - m \leq [p]$ and so

\[
\left| f^{\circ m} (Y_t) - f^{\circ m} (Y_t^{s,t}) \right| \leq C_p \left| f^{\circ m} \right|_{Lip(\gamma-1)} \left[ f^{\gamma \wedge ([p]+1)(\gamma - [\gamma])} \| X \|_{p-var,[s,t]} \right].
\]

(2.23)

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

\[
\left| f^{\circ |\gamma|} (Y_t) - f^{\circ |\gamma|} (Y_t^{s,t}) \right| \leq \left| f^{\circ |\gamma|} \right|_{Lip(\gamma - |\gamma|)} \left| Y_t - Y_t^{s,t} \right| \leq C_p \left| f^{\circ |\gamma|} \right|_{Lip(\gamma - |\gamma|)} \left[ f^{\gamma \wedge ([p]+1)(\gamma - |\gamma|)} \||X\|_{p-var,[s,t]} \right].
\]

In particular, we have

\[
\left| f^{\circ |\gamma|} (Y_t) - f^{\circ |\gamma|} (Y_t^{s,t}) \right| \leq C_p \left| f^{\circ |\gamma|} \right|_{Lip(\gamma - |\gamma|)} \left[ f^{\gamma \wedge ([p]+1)(\gamma - |\gamma|)} \left( \| X \|_{p-var,[s,t]} \wedge 1 \right) \right] \left( [p]+1 \right) \left( |\gamma| - [\gamma] \right).\]

(2.24)

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

\[
\left| f^{\circ m} (Y_t^{s,t}) - \sum_{k=0}^{[\gamma]-m} f^{\circ (m+k)} (Y_s) (x^{s,t})_s^k \right| = \left| \int_{s \leq t \leq s + \delta \leq \cdots \leq s + \delta (\gamma - |\gamma|) \leq t} f^{\circ ([\gamma])} (Y_s^t \cdots (Y_s - Y_s) \, dx_{s,t}^1 \cdots dx_{s,\gamma - |\gamma|} \leq m \right|
\]

(2.25)

\[
\leq C_p^{[\gamma]-m} \left| f^{\circ [\gamma]} \right|_{Lip(\gamma - |\gamma|)} \left| Y_t^{s,t} \right| \left| X \|_{p-var,[s,t]} ^{[\gamma]-m} \right| \left( [p]+1 \right) \left( |\gamma| - |\gamma| \right)
\]

(2.26)

\[
\times \left( \| X \|_{p-var,[s,t]} \wedge 1 \right) \left( [p]+1 \right) \left( |\gamma| - |\gamma| \right).\]

(2.27)
where in the third line we have used the $\gamma - \lceil \gamma \rceil$ 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 $\gamma + p \leq m \leq \lfloor \gamma \rfloor$, $\epsilon > 0$ and $\delta > 0$,

$$
\begin{align*}
|f^{\circ (m)} (Y_t) - \sum_{k=0}^{\lfloor \gamma \rfloor - m} f^{\circ (m+k)} (Y_s) X^k_{s,t}| \\
\leq 2C_p \max_{p \leq t} \sup_{0 \leq \xi \leq t} \int_0^t |f| \bigl( |f^{\circ 1} (Y_s) X^k_{s,t}| \bigr) ds \\
\times \left( \|X\|_{p-var, s,t} \right)^{m-p} (|f| X^k_{p-var, s,t})^{m-p}.
\end{align*}
$$

(2.2) holds with $p \leq \gamma \leq m \leq \lfloor \gamma \rfloor$ and $\epsilon > 0$.

Here since $\gamma + p \leq m \leq \lfloor \gamma \rfloor$ so $\gamma - m \leq |p|$ and

$$
|\gamma - m|! \leq |p|!.
$$

Therefore, by changing the constant $C_p$, we rewrite (2.2) in the form of the right hand side of (2.5). It now suffices to show (2.7). Note first that for $0 \leq m \leq \lfloor \gamma - p \rfloor$ and $s \leq u \leq v \leq t$,

$$
\begin{align*}
|f^{\circ m} (Y_u) - f^{\circ m} (Y_v)| \\
&\leq |f^{\circ m} (Y_u) - f^{\circ m} (Y_{u,v})| + |f (Y_{u,v}) - f^{\circ m} (Y_{u,v}) (x^{u,v})_{u,v}^k| \\
&\quad + \left| \sum_{k=|p|+1}^{\lfloor \gamma \rfloor - m} f^{\circ (m+k)} (Y_u) (x^{u,v})_{u,v}^k - \sum_{k=|p|+1}^{\lfloor \gamma \rfloor - m} f^{\circ (m+k)} (Y_u) X^k_{u,v} \right|.
\end{align*}
$$

The estimate (2.2) still holds with $(s, t)$ replaced by $(u, v)$ and (2.26) would hold with the constant $C_p$ now depending on $\gamma$. For the final term in (2.31),

$$
\begin{align*}
\left| \sum_{k=|p|+1}^{\lfloor \gamma \rfloor - m} f^{\circ (m+k)} (Y_u) (x^{u,v})_{u,v}^k \right| \\
\leq \left| \sum_{k=|p|+1}^{\lfloor \gamma \rfloor - m} f^{\circ (m+k)} (Y_u) (x^{u,v})_{u,v}^k \right| \\
\leq 2|\gamma| C_p \max_{0 \leq \xi \leq t} \sup_{0 \leq \xi \leq t} \int_0^t |f^{\circ m} (Y_u) \left( \|X\|_{p-var, s,t} \right)^{m-p} \left( \|X\|_{p-var, s,t} \right)^{m-p} \right) \|X\|_{p-var, s,t}^{m-p}.
\end{align*}
$$

where we used Fact 2.2 and

$$
\begin{align*}
|x^{u,v}_{u,v}|^k &\leq c_p^k \left( \int_u^v |dx^{u,v}_{t,v}| \right)^k \\
&\leq c_p^k \|X\|_{p-var, u,v}^k.
\end{align*}
$$

Therefore, combining with (2.2) and (2.26), we have for some constants $C_{f,p,X,s,t}\gamma, C_{f,p,X,s,t}^p$
Factorial decay estimates for differential equations

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

$$
\left| f^{om} (Y_u) - \sum_{k=0}^{[\gamma]-m} f^{(m+k)} (Y_u) X_{u,v}^k \right|
\leq C_{f,p,X,s,t} \left( \|X\|_{p-var,[u,v]}^{\gamma([p]+1)} + \|X\|_{p-var,[u,v]}^{\gamma-m} \right)
\leq C'_{f,p,X,s,t} \|X\|_{p-var,[u,v]}^{\gamma([p]+1)}
$$

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

$$
\lim_{|P| \to 0} \sum_{i=0}^{n-1} \sum_{l=1}^{[\gamma]-m} E(t_i, t_{i+1})
\leq C'_{f,p,X,s,t} \lim_{|P| \to 0} \sum_{i=0}^{n-1} \|X\|_{p-var,[t_i,t_{i+1}]}^{\gamma([p]+1)}
\leq C'_{f,p,X,s,t} \max_i \|X\|_{p-var,[t_i,t_{i+1}]}^{\gamma([p]+1)-p} \sum_{i=0}^{n-1} \|X\|_{p-var,[t_i,t_{i+1}]}^p
$$

(2.32)

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

$$
C_{f,p,X,s,t} \max_i \|X\|_{p-var,[t_i,t_{i+1}]}^{\gamma([p]+1)-p} \|X\|_{p-var,[s,t]}^p,
$$

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

Finally,

$$
\lim_{|P| \to 0} \sum_{i=0}^{n-1} \sum_{l=1}^{[\gamma]-m} f^{(m+l)} (Y_{t_i}) X_{t_i,t_{i+1}}^l
= \lim_{|P| \to 0} \sum_{i=0}^{n-1} f^{om} (Y_{t_{i+1}}) - f^{om} (Y_{t_i}) + E(t_i, t_{i+1})
= f^{om} (Y_t) - f^{om} (Y_s).
$$

\[\square\]

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-bf32b01a7c6/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.
Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS\textsuperscript{1})
- Non profit, sponsored by IMS\textsuperscript{2}, BS\textsuperscript{3}, PKP\textsuperscript{4}
- Purely electronic and secure (LOCKSS\textsuperscript{5})

Help keep the journal free and vigorous

- Donate to the IMS open access fund\textsuperscript{6} (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

\textsuperscript{1}OJS: Open Journal Systems http://pkp.sfu.ca/ojs/
\textsuperscript{2}IMS: Institute of Mathematical Statistics http://www.imstat.org/
\textsuperscript{3}BS: Bernoulli Society http://www.bernoulli-society.org/
\textsuperscript{4}PK: Public Knowledge Project http://pkp.sfu.ca/
\textsuperscript{5}LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
\textsuperscript{6}IMS Open Access Fund: http://www.imstat.org/publications/open.htm