Electron. Commun. Probab. **21** (2016), no. 7, 1–18.
DOI: 10.1214/16-ECP3795
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Recovering the pathwise Itô solution from averaged Stratonovich solutions*

Terry Lyons[†] Danyu Yang[†]

Abstract

We recover the pathwise Itô solution (the solution to a rough differential equation driven by the Itô signature) by concatenating averaged Stratonovich solutions on small intervals and by letting the mesh of the partition in the approximations tend to zero. More specifically, on a fixed small interval, we consider two Stratonovich solutions: one is driven by the original process and the other is driven by the original process plus a selected independent noise. Then by taking the expectation with respect to the selected noise, we can recover the increment of the bracket process and so recover the leading order approximation of the Itô solution up to a small error. By concatenating averaged increments and by letting the mesh tend to zero, the error tends to zero and we recover the Itô solution.

Keywords: rough paths theory; pathwise Itô solution; Stratonovich solution.
AMS MSC 2010: 60H10; 60G17.
Submitted to ECP on September 14, 2014, final version accepted on August 10, 2015.

1 Introduction

Itô calculus [11, 12] can be seen as a transformation between semi-martingales (i.e. the map which sends the driving process to the solution of a stochastic differential equation) and is widely used in various mathematical models. It is well-known that the classic Itô calculus is not stable under pointwise approximations. Indeed, the Wong-Zakai theorem ([23, 22] see also [5]) shows that, when controlled ordinary differential equations are driven by piecewise-linear approximations to Brownian motion, their solutions converge uniformly in probability to the Stratonovich solution as the mesh of the partition in the approximations tends to zero. In contrast to the Stratonovich solution, the Itô solution is not stable with respect to perturbations of the driving process even when the perturbations are very natural.

There has been a long interest trying to develop a pathwise Itô calculus [1, 13, 4, 21], but these attempts have their limitations. For example, the null set depends on the integrand function, or the integral is only defined for closed one-forms (but closed one-forms are rare in high dimensional spaces), or the convergence is in probability (so not truly pathwise). The theory of rough paths [15, 16, 18, 8, 7] is close in spirit to Föllmer's approach [4], but it is a far more systematical methodology that can deal with closed and non-closed one-forms, and applies but is not restricted to semi-martingales

*The authors would like to acknowledge the support of the Oxford-Man Institute and the support provided by ERC advanced grant ESig (agreement no. 291244).

†The Oxford-Man Institute, University of Oxford, United Kingdom. E-mail: terry.lyons@maths.ox.ac.uk; yangd@maths.ox.ac.uk

[19, 2, 6, 10, 20]. The methodology provides a robust pathwise solution which is continuous with respect to the driving path. It is known that the pathwise Itô resp. Stratonovich solution in the theory of rough paths coincides almost surely with the classical Itô resp. Stratonovich solution [15, 18, 14, 3, 7].

Unlike the Stratonovich integral, the Itô integral can not be approximated by a sequence of classical integrals. The pathwise Itô solution is generally defined as the Stratonovich solution to a modified equation with an additional drift term, see e.g. Lyons and Qian [17], Lejay and Victoir [14], Friz and Victoir [7], Hairer and Kelly [9]. Other than defining the pathwise Itô solution as the Stratonovich solution to a modified differential equation, we would like to demonstrate that the Itô solution is almost a Stratonovich solution, in the sense that the Itô solution can be expressed as the limit (as the mesh of the partition tends to zero) of concatenated averaged Stratonovich solutions. More specifically, we would need two Stratonovich solutions on a small time interval: one is the Stratonovich solution driven by the Stratonovich signature of the original process, and the other is the Stratonovich solution driven by the *joint* Stratonovich signature of the original process plus a selected independent noise. Then by taking the expectation of the second Stratonovich solution with respect to the selected noise, we recover the bracket process, and by working with a chosen functional of these two Stratonovich solutions, we get the leading order approximation of the increment of the pathwise Itô solution with a small error. By letting the mesh of the partition tend to zero, the error tends to zero and we recover the pathwise Itô solution. We would like to recover the Itô solution from averaged Stratonovich solutions mainly because the Stratonovich solution fits more naturally into the rough paths framework than the Itô solution. The averaging effect is also related to the reverse situation where any player in a market interacts with a random sub-sample from the stream and the actual effect on the market is the volume weighted average. Based on our result the random interactions will generate an Itô type correction to the equation for the aggregate behavior.

To convey the idea more explicitly, we illustrate it with a simple example. Suppose B is a one-dimensional Brownian motion and $f:\mathbb{R}\to\mathbb{R}$ is sufficiently regular. We want to recover the solution to the Itô stochastic differential equation

$$dy = f\left(y\right)dB,\, y_0=\xi\in\mathbb{R}.$$

Suppose W is another one-dimensional Brownian motion which is independent from B. We define a family of Stratonovich solutions $y^{1,s,t}$ and $y^{2,s,t}$ indexed by the time intervals $\{[s,t]\}_{s<t}$ that are defined to be the Stratonovich solution on $[s,t]$ to the stochastic differential equations (with y_s denoting the value of y at time s)

$$\begin{aligned} dy_u^{1,s,t} &= f\left(y_u^{1,s,t}\right)\circ dB_u,\, y_s^{1,s,t}=y_s,\, u\in[s,t]\,, \\ dy_u^{2,s,t} &= f\left(y_u^{2,s,t}\right)\circ d\left(B_u+W_u\right),\, y_s^{2,s,t}=y_s,\, u\in[s,t]\,. \end{aligned}$$

We would like to identify a function $F:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ such that $y_t-y_s\approx F(y_t^{1,s,t}-y_s^{1,s,t},y_t^{2,s,t}-y_s^{2,s,t})$ for every small time interval $[s,t]$. Then by concatenating $F(y_t^{1,s,t}-y_s^{1,s,t},y_t^{2,s,t}-y_s^{2,s,t})$ on small intervals and by letting the mesh of the partition tend to zero, we recover y in the limit. In the real construction, the initial values of $y^{1,s,t}$ and $y^{2,s,t}$ are not y_s but the value obtained from the last step of concatenation. Here we use y_s to give an intuitive explanation.

For a small time interval $[s,t]$, by using Euler's approximation, we have

$$y_t-y_s\approx f\left(y_s\right)\left(B_t-B_s\right)+f'\left(y_s\right)f\left(y_s\right)\frac{1}{2}\left(\left(B_t-B_s\right)^2-\left(t-s\right)\right), \tag{1.1}$$

and

$$\begin{aligned} y_t^{1,s,t} - y_s^{1,s,t} &\approx f(y_s)(B_t - B_s) + f'(y_s) f(y_s) \frac{1}{2} (B_t - B_s)^2, \\ y_t^{2,s,t} - y_s^{2,s,t} &\approx f(y_s)(B_t - B_s + W_t - W_s) + f'(y_s) f(y_s) \frac{1}{2} (B_t - B_s + W_t - W_s)^2. \end{aligned} \tag{1.2}$$

Since W is independent from B, if we take the expectation of $y_t^{2,s,t} - y_s^{2,s,t}$ w.r.t. W, then the expectation simulates the required continuous martingale correction $t-s$ in (1.1) and we get

$$\mathbb{E}^W \left(y_t^{2,s,t} - y_s^{2,s,t} \right) \approx f(y_s)(B_t - B_s) + f'(y_s) f(y_s) \frac{1}{2} \left((B_t - B_s)^2 + (t-s) \right). \tag{1.3}$$

Then combining (1.1), (1.2) and (1.3), we have

$$y_t - y_s \approx 2\left(y_t^{1,s,t} - y_s^{1,s,t} \right) - \mathbb{E}^W \left(y_t^{2,s,t} - y_s^{2,s,t} \right).$$

Hence, we may take $F(x,y) := 2x - \mathbb{E}^W(y)$, $\forall x, y \in \mathbb{R}$, (since B and W are independent, W is fixed once and for all for almost every sample path of B). Then it can be proved that

$$y_t - y_s = \lim_{|D| \to 0, D = \{t_k\}_{k=0}^n \subset [s,t]} \sum_{k, t_k \in D} F(y_{t_{k+1}}^{1,t_k,t_{k+1}} - y_{t_k}^{1,t_k,t_{k+1}}, y_{t_{k+1}}^{2,t_k,t_{k+1}} - y_{t_k}^{2,t_k,t_{k+1}}), \ \forall s < t,$$

where $D = \{t_k\}_{k=0}^n$ is a finite partition of $[s,t]$ with $s = t_0 < t_1 < \cdots < t_n = t$ and $|D| := \max_k |t_{k+1} - t_k|$ is the mesh of D. By taking the expectation with respect to the selected independent noise W and by working with a chosen functional of the Stratonovich solutions on a small interval, we obtain the leading order approximation of the increment of the Itô solution y, and recover y as the limit of discrete concatenations when the mesh tends to zero. More generally, we can replace B with a d-dimensional continuous martingale (or even a Gaussian process, provided the joint signature of the Gaussian process and the selected noise is well defined), and we have to estimate $\int y dy$ as well because the pathwise regularity of a continuous martingale is just above the threshold of having finite 2-variation a.s.. While the idea is similar and captured in this example.

2 Definitions and notations

We recall some notations in the theory of rough paths. Let $T^{(2)}(\mathbb{R}^d)$ denote the group $1 \oplus \mathbb{R}^d \oplus (\mathbb{R}^d)^{\otimes 2}$ with the multiplication and the inverse defined by (π_k denoting the projection to $(\mathbb{R}^d)^{\otimes k}$)

$$\begin{aligned} g \otimes h &:= (1, \pi_1(g) + \pi_1(h), \pi_2(g) + \pi_2(h) + \pi_1(g) \otimes \pi_1(h)), \forall g, h \in T^{(2)}(\mathbb{R}^d), \\ g^{-1} &:= \left(1, -\pi_1(g), (\pi_1(g))^{\otimes 2} - \pi_2(g)\right), \forall g \in T^{(2)}(\mathbb{R}^d). \end{aligned}$$

(In the definition of the multiplication, the "$\otimes$" on the l.h.s. denotes a group multiplication and the "$\otimes$" on the r.h.s. denotes the tensor product between two $\mathbb{R}^d$s.) We equip $T^{(2)}(\mathbb{R}^d)$ with[1]

$$\|g\| := |\pi_1(g)| + |\pi_2(g)|^{\frac{1}{2}}, \forall g \in T^{(2)}(\mathbb{R}^d). \tag{2.1}$$

[1] $\|\cdot\|$ is not a subadditive homogenous norm in the sense of Definition 7.34 [7] as it is not sub-additive with respect to the multiplication, but $\|\cdot\|$ is equivalent to a subadditive homogenous norm e.g. the Carnot-Carathéodory norm (Theorem 7.32, Theorem 7.44 [7]).

Then $(T^{(2)}(\mathbb{R}^d), \|\cdot\|)$ is a topological group. For $p \in [1,\infty)$ and a continuous path γ defined on $[0,T]$ taking values in $T^{(2)}(\mathbb{R}^d)$, the p-variation of γ is defined by

$$\|\gamma\|_{p-var,[0,T]} := \left(\sup_{D\subset[0,T]} \sum_{k,t_k\in D} \left\|\gamma_{t_k}^{-1}\otimes\gamma_{t_{k+1}}\right\|^p\right)^{\frac{1}{p}},$$

where the supremum is taken over all finite partitions $D=\{t_k\}_{k=0}^n$ of $[0,T]$, $0=t_0<t_1<\cdots<t_n=T$, $n\geq 1$.

Definition 2.1 (p-Rough Path, $p\in[2,3)$)**.** *Suppose γ is a continuous path on $[0,T]$ taking values in $T^{(2)}(\mathbb{R}^d)$. We say γ is a p-rough path for some $p\in[2,3)$ if $\|\gamma\|_{p-var,[0,T]}<\infty$.*

Based on Lejay and Victoir [14], any p-rough path, $p\in[2,3)$, can be interpreted as the product of a weak geometric p-rough path and another continuous path with finite $2^{-1}p$-variation. We will use this equivalence and define the solution to a rough differential equation driven by a p-rough path, $p\in[2,3)$, as the solution to a rough differential equation driven by a $(p,2^{-1}p)$-rough path as in Friz and Victoir [7].

Notation 2.2. *Suppose $\gamma:[0,T]\to T^{(2)}(\mathbb{R}^d)$ is a p-rough path for some $p\in[2,3)$. Then we denote $\gamma=\gamma^A+\gamma^S$ with continuous paths $\gamma^A:[0,T]\to T^{(2)}(\mathbb{R}^d)$ and $\gamma^S:[0,T]\to(\mathbb{R}^d)^{\otimes 2}$ defined by*

$$\begin{aligned}\gamma_t^A &:= \left(1,\pi_1(\gamma_t),\operatorname{Anti}\left(\pi_2(\gamma_t)-\frac{1}{2}(\pi_1(\gamma_t))^{\otimes 2}\right)+\frac{1}{2}(\pi_1(\gamma_t))^{\otimes 2}\right),\ t\in[0,T],\\ \gamma_t^S &:= \operatorname{Sym}\left(\pi_2(\gamma_t)-\frac{1}{2}(\pi_1(\gamma_t))^{\otimes 2}\right),\ t\in[0,T],\end{aligned}$$

where $\operatorname{Anti}(\cdot)$ denotes the projection of $(\mathbb{R}^d)^{\otimes 2}$ to $\operatorname{span}\{e_i\otimes e_j-e_j\otimes e_i|i,j=1,\ldots,d\}$ and $\operatorname{Sym}(\cdot)$ denotes the projection of $(\mathbb{R}^d)^{\otimes 2}$ to $\operatorname{span}\{e_i\otimes e_j+e_j\otimes e_i|i,j=1,\ldots,d\}$.

Then γ^A is a weak geometric p-rough path[2] (a normal driving path in rough paths theory) and γ^S is a continuous path with finite $2^{-1}p$-variation. The cross integrals between $\pi_1(\gamma^A)$ (which is equal to $\pi_1(\gamma)$) and γ^S are well-defined as Young integrals [24] because $p^{-1}+2p^{-1}=3p^{-1}>1$, see [14] for details.

Denote by $L(\mathbb{R}^d,\mathbb{R}^e)$ the set of linear mappings from $\mathbb{R}^d$ to $\mathbb{R}^e$.

Definition 2.3. *$f:\mathbb{R}^e\to L(\mathbb{R}^d,\mathbb{R}^e)$ is said to be $Lip(\beta)$ for $\beta>1$, if f is $\lfloor\beta\rfloor$-times Fréchet differentiable (where $\lfloor\beta\rfloor$ denotes the largest integer which is strictly less than β) and*

$$|f|_{Lip(\beta)} := \max_{k=0,1,\ldots,\lfloor\beta\rfloor}\left\|D^k f\right\|_\infty \vee \left\|D^{\lfloor\beta\rfloor}f\right\|_{(\beta-\lfloor\beta\rfloor)-H\ddot{o}l}<\infty,$$

where $\|\cdot\|_\infty$ denotes the uniform norm and $\|\cdot\|_{(\beta-\lfloor\beta\rfloor)-H\ddot{o}l}$ denotes the $(\beta-\lfloor\beta\rfloor)$-Hölder norm.

Let $C^{1-var}([0,T],\mathbb{R}^d)$ denote the set of continuous paths of bounded variation on $[0,T]$ taking values in $\mathbb{R}^d$.

Definition 2.4. *For $x\in C^{1-var}([0,T],\mathbb{R}^d)$, we define $S_2(x):[0,T]\to T^{(2)}(\mathbb{R}^d)$ by*

$$S_2(x)_t := \left(1,x_t-x_0,\iint_{0<u_1<u_2<t} dx_{u_1}\otimes dx_{u_2}\right),\ \forall t\in[0,T]. \tag{2.2}$$

The following definition is based on Definition 12.2 in [7].

[2] A weak geometric p-rough path is a continuous path of finite p-variation taking values in the step-$[p]$ nilpotent Lie group.

Definition 2.5 (RDE Solution, $p\in[2,3)$)**.** *Suppose $\gamma:[0,T]\to T^{(2)}\left(\mathbb{R}^d\right)$ is a p-rough path for some $p\in[2,3)$ with the decomposition $\gamma=\gamma^A+\gamma^S$ (as in Notation 2.2), and $f:\mathbb{R}^e\to L\left(\mathbb{R}^d,\mathbb{R}^e\right)$ is $Lip\,(\beta)$ for some $\beta>p-1$. Then $Y:[0,T]\to T^{(2)}(\mathbb{R}^e)$ is a solution to the rough differential equation (RDE)*

$$dY=f\left(Y\right)d\gamma,\ Y_0=\xi\in T^{(2)}\left(\mathbb{R}^e\right), \tag{2.3}$$

if there exist two sequences of continuous bounded variation paths $x^{A,m}\in C^{1-var}\left([0,T],\mathbb{R}^d\right)$ and $x^{S,m}\in C^{1-var}\left([0,T],(\mathbb{R}^d)^{\otimes 2}\right)$, $m\geq 1$, such that

$$\sup_{m\geq 1}\left(\left\|x^{A,m}\right\|_{p-var,[0,T]}+\left\|x^{S,m}\right\|_{\frac{p}{2}-var,[0,T]}\right)<\infty,$$

$$\lim_{m\to\infty}\max_{k=1,2}\sup_{0\leq s\leq t\leq T}\left|\pi_k\left(S_2\left(x^{A,m}\right)_s^{-1}\otimes S_2\left(x^{A,m}\right)_t\right)-\pi_k\left(\left(\gamma_s^A\right)^{-1}\otimes\gamma_t^A\right)\right|=0,$$

$$\lim_{m\to\infty}\sup_{0\leq s\leq t\leq T}\left|\left(x_t^{S,m}-x_s^{S,m}\right)-\left(\gamma_t^S-\gamma_s^S\right)\right|=0,$$

and the ODE solutions $y^{1,m}:[0,T]\to\mathbb{R}^e$ and $y^{2,m}:[0,T]\to(\mathbb{R}^e)^{\otimes 2}$:

$$\begin{aligned} dy^{1,m} &= f\left(y^{1,m}\right)dx^{A,m}+\left(Df\right)\left(f\right)\left(y^{1,m}\right)dx^{S,m},\ y_0^{1,m}=\pi_1\left(\xi\right)\in\mathbb{R}^e,\\ dy^{2,m} &= y^{1,m}\otimes dy^{1,m}+f\left(y^{1,m}\right)^{\otimes 2}dx^{S,m},\ y_0^{2,m}=\pi_2\left(\xi\right)\in\left(\mathbb{R}^e\right)^{\otimes 2}. \end{aligned} \tag{2.4}$$

such that

$$\lim_{m\to\infty}\max_{k=1,2}\sup_{0\leq t\leq T}\left|y_t^{k,m}-\pi_k\left(Y_t\right)\right|=0.$$

Theorem 2.6 (Existence and Uniqueness)**.** *There exists a solution to (2.3) when f is $Lip\,(\beta)$ for $\beta>p-1$, and the solution is unique when $\beta>p$.*

Theorem 2.6 follows from Theorem 12.6 and Theorem 12.10 in [7]. Comparing with Definition 12.2 in [7], we add in an extra term $f(y^{1,m})^{\otimes 2}dx^{S,m}$ in (2.4) so that the second level of the pathwise Itô solution coincides almost surely with the iterated Itô integral of the stochastic Itô solution.

The modification we made in (2.4) will not affect this existence and uniqueness result. Indeed, based on Theorem 12.6 [7], when f is $Lip\,(\beta)$ for $\beta>p-1$, $\{y^{1,m}\}_m$ are uniformly bounded in p-variation. When $y^{1,m}$ converge uniformly as $m\to\infty$ to $\pi_1\left(Y\right)$, by interpolating between the p-variation norm and the uniform norm, we have that $y^{1,m}$ converge to $\pi_1\left(Y\right)$ in p'-variation for any $p'>p$ as $m\to\infty$. Similarly, by interpolating between the $2^{-1}p$-variation norm and the uniform norm, we have that $x^{S,m}$ converge to γ^S in $2^{-1}p'$-variation for any $p'>p$ as $m\to\infty$. We choose $p'\in(p,3)$ so that $(p')^{-1}+2(p')^{-1}>1$. Then by using Young integral (Theorem 1.16 [16]) the additional term $\int_0^\cdot f(y^{1,m})^{\otimes 2}dx^{S,m}$ in (2.4) converge uniformly to $\int_0^\cdot f(\pi_1\left(Y\right))^{\otimes 2}d\gamma^S$ as $m\to\infty$. Hence, when f is $Lip\,(\beta)$ for $\beta>p-1$, if Y is a solution to (2.3) in the sense of Definition 12.2 in [7], then $Y+\int_0^\cdot f(\pi_1\left(Y\right))^{\otimes 2}d\gamma^S$ is a solution to (2.3) in the sense of Definition 2.5 (i.e. with the additional term $f(y^{1,m})^{\otimes 2}dx^{S,m}$ in (2.4)). When $\beta>p$, based on Theorem 12.10 in [7], the solution in the sense of Definition 12.2 in [7] is unique, so the path $\int_0^\cdot f(\pi_1\left(Y\right))^{\otimes 2}d\gamma^S$ is unique, and we have the uniqueness of the solution to (2.3) in the sense of Definition 2.5.

3 Recovering the pathwise Itô solution

As mentioned in the introduction, we would like to recover the pathwise Itô solution by taking the average of Stratonovich solutions. The idea is simple, but the concrete formulation needs some care. Here we try to give a sensible explanation of our formulation.

Suppose Z is a d-dimensional continuous martingale on $[0,T]$. We denote the step-2 Stratonovich signature of Z by $S_2(Z)_t := (1, Z_t - Z_0, \int_0^t (Z_u - Z_0) \otimes \circ dZ_u)$, $t \in [0,T]$, and denote the step-2 Itô signature of Z by $\mathcal{I}_2(Z)_t := (1, Z_t - Z_0, \int_0^t (Z_u - Z_0) \otimes dZ_u)$, $t \in [0,T]$. The difference between them is in the definition of the iterated integral of Z, where they are defined as Stratonovich resp. Itô integral. Both $S_2(Z)$ and $\mathcal{I}_2(Z)$ are almost surely a p-rough path for any $p \in (2,3)$ (Theorem 14.9 [7]). Usually, the pathwise Stratonovich resp. Itô solution is the solution to a rough differential equation driven by the Stratonovich resp. Itô signature.

Definition 3.1 (Perturbed Rough Path)**.** *Suppose* $\gamma : [0,T] \to \left(T^{(2)}\left(\mathbb{R}^d\right), \|\cdot\|\right)$ *is a fixed* p*-rough path for some* $p \in [2,3)$*,* $\phi = (\phi^{i,j})_{i,j=1,\ldots d}$ *is a fixed path on* $[0,T]$ *taking value in* $d \times d$ *matrices satisfying* $\max_{i,j} \int_0^T \left(\phi_u^{i,j}\right)^2 du < \infty$*, and* B *is a* d*-dimensional Brownian motion. Define a continuous* d*-dimensional martingale* M *by the Itô integral:*

$$M_t := \int_0^t \phi_u dB_u, \ \forall t \in [0,T]. \tag{3.1}$$

We define $\gamma^{(M,R)} : [0,T] \to \left(T^{(2)}\left(\mathbb{R}^d\right), \|\cdot\|\right)$ *as a perturbed rough path, if* $\gamma^{(M,R)}$ *is almost surely a* p*-rough path for some* $p \in (2,3)$*, and*

$$\gamma_t^{(M,R)} = \left(1, \pi_1(\gamma_t) + M_t, \pi_2(\gamma_t) + \iint_{0<u_1<u_2<t} \circ dM_{u_1} \otimes \circ dM_{u_2} + R_t\right), \ \forall t \in [0,T], \text{a.s.,}$$

for some process $R : [0,T] \to (\mathbb{R}^d)^{\otimes 2}$ *satisfying*

$$\mathbb{E}R_t = 0, \ \forall t \in [0,T]. \tag{3.2}$$

Since γ is fixed, the condition (3.2) is satisfied e.g. when the cross integrals between $\pi_1(\gamma)$ and M (i.e. the process R) are defined as the L^1 limit of piecewise linear approximations.

Suppose Z is a d-dimensional square integrable martingale such that its bracket process $\langle Z\rangle$ has the expression $\int \psi_u^T \psi_u du$ for some matrix-valued process ψ, and B is a d-dimensional Brownian motion independent from Z. We let $\gamma = S_2(Z)$ and define M to be the Itô integral $\int \psi_u dB_u$. In this case, the process R could be defined by (and there are other possible choices)

$$R_t := \int_0^t (M_u - M_0) \otimes \circ dZ_u + \int_0^t (Z_u - Z_0) \otimes \circ dM_u. \tag{3.3}$$

The Stratonovich integrals in (3.3) are well-defined because the $2d$-dimensional process (Z,M) is a continuous martingale w.r.t. the filtration generated by Z and B (Proposition 14.9 [7]). Then condition (3.2) is satisfied for this particular choice of R for almost every γ because the Stratonovich integrals in (3.3) can be expressed as the L^1 limit of piecewise linear approximations and Z and B are independent. For this selection of R, $\gamma^{(M,R)}$ is almost surely a p-rough path for any $p \in (2,3)$ for almost every γ because $\gamma^{(M,R)} = S_2(Z+M)$ and $S_2(Z+M)$ is almost surely a p-rough path for any $p \in (2,3)$ for almost every sample path of Z (Theorem 14.12 [7]). We did not require that γ is a geometric rough path, so we also could let $\gamma = \mathcal{I}_2(Z)$. Then without changing the definitions of M and of R, the conditions in Definition 3.1 are satisfied. Indeed, condition (3.2) is satisfied as the definition of R stays unchanged, and $\gamma^{(M,R)}$ is almost surely a p-rough path for $p \in (2,3)$ because in this case we have $\gamma^{(M,R)} = S_2(Z+M) + 2^{-1}\langle Z\rangle$ and $\left\|\gamma^{(M,R)}\right\|_{p-var,[0,T]} \le \|S_2(Z+M)\|_{p-var,[0,T]} + \|\langle Z\rangle\|_{1-var,[0,T]}^{2^{-1}} < \infty$ a.s..

As a specific example when γ is not a sample path of a martingale, suppose B is a d-dimensional Brownian motion and (X, B) is a $2d$-dimensional continuous Gaussian process with independent components. When the covariance function of (X, B) has finite ρ-variation for some $\rho \in [1, \frac{3}{2})$, the process (X, B) can be lifted to a p-rough process for any $p \in (2\rho, 3)$, and the lifted rough process is the L^1-limit of the signatures of the piecewise linear approximations (Theorem 15.33 [7]). Then we could let γ be a sample path of the rough process above X (e.g. fractional Brownian motion with Hurst parameter $H > 3^{-1}$) and let M be the Brownian motion B. Then condition (3.2) holds because the integral between $\pi_1(\gamma)$ and M is the L^1 limit of the piecewise linear approximations, and $\gamma^{(M,R)}$ is almost surely a p-rough path for some $p \in (2, 3)$ based on Theorem 15.33 [7].

As mentioned before, we have two Stratonovich solutions on a small interval: one is driven by the signature of the original process and the other is driven by the joint signature of the original process plus a noise. Here the rough path γ is (a sample path of) the signature of the original process, and $\gamma^{(M,R)}$ is the joint signature of the original process plus a noise. Suppose $f : \mathbb{R}^e \to L\left(\mathbb{R}^d, \mathbb{R}^e\right)$ is $Lip(\beta)$ for $\beta > p$ and let $\mathcal{I}_2(\gamma, M)$ denote the p-rough path for some $p \in [2, 3)$:

$$\mathcal{I}_2(\gamma, M)_t := \left(1, \pi_1(\gamma_t), \pi_2(\gamma_t) - \frac{1}{2}\langle M\rangle_t\right), \ t \in [0, T].$$

($\mathcal{I}_2(\gamma, M)$ is deterministic because $\langle M\rangle_\cdot = \int_0^\cdot \phi_u^T \phi_u du$ is deterministic as we assumed.) We would like to express the increment on $[s, t]$ of the solution to the RDE

$$dy = f(y)\, d\mathcal{I}_2(\gamma, M), \ y_0 = \xi \in T^{(2)}(\mathbb{R}^e), \tag{3.4}$$

in term of the increments on $[s, t]$ of $y^{1,s,t}$ and $y^{2,s,t}$, where $y^{i,s,t} : [0, T] \to T^{(2)}(\mathbb{R}^e)$, $i = 1, 2$, is the solution to rough differential equations (with y_s denoting the value of y in (3.4) at time s),

$$dy_u^{1,s,t} = f\left(y_u^{1,s,t}\right) d\gamma_u, \ y_s^{1,s,t} = y_s, \ u \in [s, t], \tag{3.5}$$

$$dy_u^{2,s,t} = f\left(y_u^{2,s,t}\right) d\gamma_u^{(M,R)}, \ y_s^{2,s,t} = y_s, \ u \in [s, t]. \tag{3.6}$$

Hence, we have a global solution y on $[0, T]$ and a family of solutions $y^{1,s,t}$ and $y^{2,s,t}$ indexed by the time intervals $\{[s, t]\}_{s<t}$. We would like to identify a function F such that $y_{s,t} \approx F(y_{s,t}^{1,s,t}, y_{s,t}^{2,s,t})$ (with $y_{s,t} := y_s^{-1} \otimes y_t$ and $y_{s,t}^{i,s,t} := (y_s^{i,s,t})^{-1} \otimes y_t^{i,s,t}$) for every small interval $[s, t]$. Then by concatenating $F(y_{s,t}^{1,s,t}, y_{s,t}^{2,s,t})$ on small intervals and by letting the mesh of the partition tend to zero, we can recover y in the limit. Yet in the real construction, the initial values in (3.5) and (3.6) are actually not y_s (which is the pathwise Itô solution we would like to recover) but the value obtained from the last step of discrete concatenations of $\{F(y_{s,t}^{1,s,t}, y_{s,t}^{2,s,t})\}_{[s,t]}$. Here we use y_s for illustration purposes, but discrete concatenations will create an error which propagates and the analysis will need some care.

Here $y^{1,s,t}$ and $y^{2,s,t}$ are what we call the Stratonovich solutions (on the small time interval $[s, t]$), and y is called the Itô solution (on the large time interval $[0, T]$). They are not necessarily the usual pathwise Stratonovich resp. Itô solution (e.g. γ could be a Gaussian rough path as in the example given above), and the convergence holds as long as the conditions of Theorem 3.3 below are satisfied. To recover the usual pathwise Itô solution (the RDE solution driven by the Itô signature of a continuous martingale), suppose Z is a square integrable continuous martingale such that its bracket process $\langle Z\rangle$ has the expression $\int \psi_u^T \psi_u du$ for a matrices-valued process ψ, and B is a Brownian motion independent from Z. We let $\gamma = S_2(Z)$, $M_\cdot = \int_0^\cdot \psi_s dB_s$, and define $\gamma^{(M,R)} := S_2(Z+M)$. In this case, y^1 and y^2 are pathwise Stratonovich solutions driven by the Stratonovich signature $S_2(Z)$ and $S_2(Z+M)$ respectively, and $\mathcal{I}_2(\gamma, M)$ coincides with $\mathcal{I}_2(Z)$ (the Itô signature of Z) so y is the pathwise Itô solution driven by the Itô signature $\mathcal{I}_2(Z)$.

In the following we try to give a sketch of our idea which helps to motivate and clarify our arguments and is also useful for picking apart the proof of Theorem 3.3. For a fixed interval $[s,t]$, we would like to represent the increment of y on $[s,t]$ in (3.4) in terms of the increment of $y^{1,s,t}$ and $y^{2,s,t}$ on $[s,t]$ in (3.5) and (3.6). Based on Theorem 12.6 in [7], we have (denote $y^i := y^{i,s,t}$ and $y^i_{s,t} := (y^i_s)^{-1} \otimes y^i_t$)

$$\begin{aligned} \pi_1\left(y^1_{s,t}\right) &\approx f\left(\pi_1\left(y_s\right)\right)\pi_1\left(\gamma_{s,t}\right) + \left(Df\right)\left(f\right)\left(\pi_1\left(y_s\right)\right)\pi_2\left(\gamma_{s,t}\right), \\ \pi_1\left(y^2_{s,t}\right) &\approx f\left(\pi_1\left(y_s\right)\right)\pi_1\left(\gamma^{(M,R)}_{s,t}\right) + \left(Df\right)\left(f\right)\left(\pi_1\left(y_s\right)\right)\pi_2\left(\gamma^{(M,R)}_{s,t}\right). \end{aligned} \tag{3.7}$$

(The " $\approx$ " indicates that two values are close up to a small error in pathwise sense, and the error will be made explicit in the proof.) Based on Definition 3.1, we have $\mathbb{E}\left(\int_s^t \left(M_u - M_s\right) \otimes \circ dM_u\right) = 2^{-1}\left\langle M\right\rangle_{s,t}$ (since $\left\langle M\right\rangle_{s,t} = \int_s^t \phi_u^T \phi_u du$ is deterministic) and $\mathbb{E}\left(\pi_2(\gamma^{(M,R)}_{s,t})\right) = \pi_2\left(\gamma_{s,t}\right) + 2^{-1}\left\langle M\right\rangle_{s,t}$. Hence,

$$\mathbb{E}\left(\pi_1\left(y^2_{s,t}\right)\right) \approx f\left(\pi_1\left(y_s\right)\right)\pi_1\left(\gamma_{s,t}\right) + \left(Df\right)\left(f\right)\left(\pi_1\left(y_s\right)\right)\left(\pi_2\left(\gamma_{s,t}\right) + 2^{-1}\left\langle M\right\rangle_{s,t}\right). \tag{3.8}$$

While for the increment on $[s,t]$ of the first level of y in (3.4), we have

$$\pi_1\left(y_{s,t}\right) \approx f\left(\pi_1\left(y_s\right)\right)\pi_1\left(\gamma_{s,t}\right) + \left(Df\right)\left(f\right)\left(\pi_1\left(y_s\right)\right)\left(\pi_2\left(\gamma_{s,t}\right) - 2^{-1}\left\langle M\right\rangle_{s,t}\right). \tag{3.9}$$

Then based on (3.7), (3.8) and (3.9), we have

$$\pi_1\left(y_{s,t}\right) \approx 2\pi_1\left(y^1_{s,t}\right) - \mathbb{E}\left(\pi_1\left(y^2_{s,t}\right)\right). \tag{3.10}$$

Since we work with $p \in (2,3)$, we have to consider the second level approximation as well. By following similar arguments as for the first level (again based on Theorem 12.6 in [7], but here we add in an extra term as in Definition 2.5), we have

$$\begin{aligned} \pi_2\left(y^1_{s,t}\right) &\approx f\left(\pi_1\left(y_s\right)\right) \otimes f\left(\pi_1\left(y_s\right)\right)\pi_2\left(\gamma_{s,t}\right), \\ \mathbb{E}\left(\pi_2\left(y^2_{s,t}\right)\right) &\approx f\left(\pi_1\left(y_s\right)\right) \otimes f\left(\pi_1\left(y_s\right)\right)\left(\pi_2\left(\gamma_{s,t}\right) + 2^{-1}\left\langle M\right\rangle_{s,t}\right), \\ \pi_2\left(y_{s,t}\right) &\approx f\left(\pi_1\left(y_s\right)\right) \otimes f\left(\pi_1\left(y_s\right)\right)\left(\pi_2\left(\gamma_{s,t}\right) - 2^{-1}\left\langle M\right\rangle_{s,t}\right). \end{aligned}$$

Then

$$\pi_2\left(y_{s,t}\right) \approx 2\pi_2\left(y^1_{s,t}\right) - \mathbb{E}\left(\pi_2\left(y^2_{s,t}\right)\right). \tag{3.11}$$

Combining (3.10) and (3.11), we have that the linear expression holds:

$$y_{s,t} \approx 2y^1_{s,t} - \mathbb{E}\left(y^2_{s,t}\right). \tag{3.12}$$

There are other possible expressions of $y_{s,t}$ in term of $y^1_{s,t}$ and $y^2_{s,t}$. For example,

$$y_{s,t} \approx y^1_{s,t} \otimes \mathbb{E}\left(y^2_{s,t}\right)^{-1} \otimes y^1_{s,t}, \tag{3.13}$$

which constitutes another approximation that is equivalent to (3.12) at leading order. Indeed,

$$\begin{aligned} \pi_1\left(y^1_{s,t} \otimes \mathbb{E}\left(y^2_{s,t}\right)^{-1} \otimes y^1_{s,t}\right) &= 2\pi_1\left(y^1_{s,t}\right) - \mathbb{E}\left(\pi_1\left(y^2_{s,t}\right)\right), \\ \pi_2\left(y^1_{s,t} \otimes \mathbb{E}\left(y^2_{s,t}\right)^{-1} \otimes y^1_{s,t}\right) &= 2\pi_2\left(y^1_{s,t}\right) - \mathbb{E}\left(\pi_2\left(y^2_{s,t}\right)\right) + \left(\pi_1\left(y^1_{s,t}\right) - \mathbb{E}\left(\pi_1\left(y^2_{s,t}\right)\right)\right)^{\otimes 2}, \end{aligned}$$

and (3.13) holds because $\left(\pi_1\left(y^1_{s,t}\right)-\mathbb{E}\left(\pi_1\left(y^2_{s,t}\right)\right)\right)^{\otimes 2}$ is small (based on (3.7) and (3.8)). Then it can be proved that, by concatenating the increments either in the form of (3.12) or in the form of (3.13) and by letting the mesh of the partition tend to zero, one will recover y (the solution to (3.4)) in the limit and the analysis in both cases are similar. There is some freedom to choose the expression of $y_{s,t}$ in term of $y^1_{s,t}$ and $y^2_{s,t}$, and the convergence will hold as long as the error is small. We will work with small increments in the form of (3.13).

Definition 3.2. *Suppose γ and $\gamma^{(M,R)}$ are defined as in Definition 3.1, $f:\mathbb{R}^e\to L\left(\mathbb{R}^d,\mathbb{R}^e\right)$ is $Lip\left(\beta\right)$ for $\beta>p$ and $\xi\in T^{(2)}\left(\mathbb{R}^e\right)$. For a finite partition $D=\{t_j\}_{j=0}^n$ of $[0,T]$, define the piecewise-constant process $y^D:[0,T]\to T^{(2)}\left(\mathbb{R}^e\right)$ by (with $y_{s,t}:=y_s^{-1}\otimes y_t$)*

$$y^D_0:=\xi,\ \ y^D_t:=y^D_{t_j}\otimes y^{1,j}_{t_j,t_{j+1}}\otimes\mathbb{E}\left(y^{2,j}_{t_j,t_{j+1}}\right)^{-1}\otimes y^{1,j}_{t_j,t_{j+1}},\ t\in(t_j,t_{j+1}], \tag{3.14}$$

where $y^{1,j}$ and $y^{2,j}$ denote the solution to the rough differential equations on $[t_j,t_{j+1}]$:

$$\begin{aligned} dy^{1,j}_u &= f\left(y^{1,j}_u\right)d\gamma_u,\ \ y^{1,j}_{t_j}=y^D_{t_j},\ \ u\in[t_j,t_{j+1}]\,, \\ dy^{2,j}_u &= f\left(y^{2,j}_u\right)d\gamma^{(M,R)}_u,\ \ y^{2,j}_{t_j}=y^D_{t_j},\ \ u\in[t_j,t_{j+1}]\,. \end{aligned} \tag{3.15}$$

It is worth noting that, (since γ is fixed) y^D is deterministic for each D.

Theorem 3.3. *Suppose γ and $\gamma^{(M,R)}$ are defined as in Definition 3.1 and $p\in(2,3)$. Denote p-rough path $\mathcal{I}_2\left(\gamma,M\right):[0,T]\to\left(T^{(2)}\left(\mathbb{R}^d\right),\|\cdot\|\right)$ by*

$$\mathcal{I}_2\left(\gamma,M\right)_t:=\left(1,\pi_1\left(\gamma_t\right),\pi_2\left(\gamma_t\right)-\frac{1}{2}\sum_{i,j=1}^d\left\langle M^i,M^j\right\rangle_t e_i\otimes e_j\right),\ t\in[0,T]\,.$$

Suppose $f:\mathbb{R}^e\to L\left(\mathbb{R}^d,\mathbb{R}^e\right)$ is $Lip\left(\beta\right)$ for $\beta>p$. If we assume that,

$$\mathbb{E}\left(\left\|\gamma^{(M,R)}\right\|^{2p}_{p-var,[0,T]}\right)<\infty, \tag{3.16}$$

then for $\xi\in T^{(2)}\left(\mathbb{R}^e\right)$, y^D (defined in (3.14)) converge uniformly as $|D|\to 0$ to the unique solution to the rough differential equation

$$dY=f\left(Y\right)d\mathcal{I}_2\left(\gamma,M\right),\ \ Y_0=\xi. \tag{3.17}$$

More specifically,

$$\lim_{|D|\to 0}\max_{k=1,2}\sup_{0\le t\le T}\left|\pi_k\left(y^D_t\right)-\pi_k\left(Y_t\right)\right|=0. \tag{3.18}$$

The proof of Theorem 3.3 starts from page 13.

Remark 3.4. Based on the proof of Theorem 3.3, $\mathbb{E}(\|\gamma^{(M,R)}\|^q_{p-var,[0,T]})<\infty$ for some $q>p$ is sufficient for the convergence of the first level in (3.18).

For a continuous martingale Z, let γ be a sample path of the Stratonovich signature of Z. Then by choosing a specific noise and by applying Theorem 3.3, we can recover the pathwise Itô solution.

Definition 3.5. *Suppose Z is a continuous d-dimensional martingale in L^2 on $[0,T]$ and there exists a $d\times d$-matrices-valued adapted process ψ in L^2 on $[0,T]$ such that*

$$\left\langle Z\right\rangle_t=\int_0^t\psi_s^T\psi_s ds,\ \forall t\in[0,T]\,,\ a.s..$$

Suppose B is a d-dimensional Brownian motion, independent from Z. Define a continuous martingale $M:[0,T]\to\mathbb{R}^d$ by the Itô integral:

$$M_t:=\int_0^t\psi_s dB_s,\ t\in[0,T]\,. \tag{3.19}$$

Corollary 3.6. *Suppose Z is a continuous d-dimensional martingale on $[0,T]$ in $L^{4+\epsilon}$ for some $\epsilon>0$, and $f:\mathbb{R}^e\to L\left(\mathbb{R}^d,\mathbb{R}^e\right)$ is $Lip\left(\beta\right)$ for $\beta>2$. Denote by Y the solution to the rough differential equation:*

$$dY=f\left(Y\right)d\mathcal{I}_2\left(Z\right),\ \ Y_0=\xi\in T^{(2)}\left(\mathbb{R}^e\right). \tag{3.20}$$

For almost every sample path of Z, if we let $\gamma:=S_2\left(Z\right)$ and $\gamma^{(M,R)}:=S_2(Z+M)$ (with M defined in (3.19)), then y^D (defined in (3.14)) converge to Y uniformly as $|D|\to 0$.

Corollary 3.6 follows from Theorem 3.3 and is proved on page 17.

Remark 3.7. By using the classical relationship between the Itô solution and the Stratonovich solution, it can be checked that Y in (3.20) satisfies $Y_t=\xi\otimes\mathcal{I}_2\left(y\right)_{0,t}$, $\forall t\in[0,T]$, a.s., with y denotes the unique strong continuous solution to the stochastic differential equation $dy=f\left(y\right)dZ$, $y_0=\pi_1\left(\xi\right)$.

4 Proofs

Our constants may implicitly depend on dimensions (d and e). We specify the dependence on other constants (e.g. C_p), but the exact value of constants may change from line to line.

4.1 Results from rough paths theory

The Theorem below follows from Theorem 14.12 in [7] and Doob's maximal inequality.

Theorem 4.1. *Suppose M is a d-dimensional continuous martingale. Then for $q>1$ and $p>2$, $\mathbb{E}\left(|M_T-M_0|^q\right)$, $\mathbb{E}\left(\left|\langle M\rangle_T\right|^{2^{-1}q}\right)$ and $\mathbb{E}\left(\left\|S_2\left(M\right)\right\|^q_{p-var,[0,T]}\right)$ are equivalent up to a constant depending on p,q,d.*

Suppose $\gamma=\gamma^S+\gamma^A$ (Notation 2.2) is a p-rough path on $[0,T]$ for some $p\in[2,3)$. Then, (see [14])

$$\left\|\gamma\right\|^p_{p-var,[s,t]}\leq\left\|\gamma^A\right\|^p_{p-var,[s,t]}+\left\|\gamma^S\right\|^{2^{-1}p}_{2^{-1}p-var,[s,t]}\leq C_d\left\|\gamma\right\|^p_{p-var,[s,t]},\ \forall s\leq t. \tag{4.1}$$

Theorem 4.2. *Suppose γ is a p-rough path on $[0,T]$ for some $p\in[2,3)$ taking values in $T^{(2)}\left(\mathbb{R}^d\right)$ and $f:\mathbb{R}^e\to L\left(\mathbb{R}^d,\mathbb{R}^e\right)$ is $Lip\left(\beta\right)$ for $\beta\in(p-1,2]$. If $Y:[0,T]\to T^{(2)}\left(\mathbb{R}^e\right)$ is a solution to the rough differential equation*

$$dY=f\left(Y\right)d\gamma,\ \ Y_0=\xi\in T^{(2)}\left(\mathbb{R}^e\right), \tag{4.2}$$

then with $\omega\left(s,t\right):=\left\|\gamma\right\|^p_{p-var,[s,t]}$ for any $s\leq t$ we have ($Y_{s,t}:=Y_s^{-1}\otimes Y_t$, $\gamma_{s,t}:=\gamma_s^{-1}\otimes\gamma_t$)

$$\left\|Y\right\|_{p-var,[s,t]}\leq C_{p,\beta,f}\left(\omega\left(s,t\right)^{\frac{1}{p}}\vee\omega\left(s,t\right)\right), \tag{4.3}$$

$$\left|\pi_1\left(Y_{s,t}\right)-f\left(\pi_1\left(Y_s\right)\right)\pi_1\left(\gamma_{s,t}\right)-\left(Df\right)\left(f\right)\left(\pi_1\left(Y_s\right)\right)\pi_2\left(\gamma_{s,t}\right)\right|\leq C_{p,\beta,f}\omega\left(s,t\right)^{\frac{\beta+1}{p}}, \tag{4.4}$$

$$\left|\pi_2\left(Y_{s,t}\right)-f\left(\pi_1\left(Y_s\right)\right)^{\otimes 2}\pi_2\left(\gamma_{s,t}\right)\right|\leq C_{p,\beta,f}\omega\left(s,t\right)^{\frac{\beta+1}{p}}\vee\omega\left(s,t\right)^2. \tag{4.5}$$

Theorem 4.2 follows from Theorem 12.6 in [7]. Since we added an extra term on the second level of (4.2) as in Definition 2.5, we check that the extra term can be estimated similarly. We only modified the second level, so we can use estimates of the first level. Suppose $\gamma=\gamma^S+\gamma^A$ as in Notation 2.2. For the extra term $\int f\left(\pi_1\left(Y_u\right)\right)^{\otimes 2}d\gamma_u^S$ and any $[s,t]\subseteq[0,T]$, based on estimates of Young integral in Theorem 1.16 [16] and

$\left\|\gamma^{S}\right\|_{2^{-1}p-var,[s,t]} \leq C_d \left\|\gamma\right\|_{p-var,[s,t]}^2 = C_d\omega\left(s,t\right)^{\frac{2}{p}}$ as in (4.1), we have

$$\begin{aligned} & \left\| \int_s^t f\left(\pi_1\left(Y_u\right)\right)^{\otimes 2} d\gamma_u^S - f\left(\pi_1\left(Y_s\right)\right)^{\otimes 2}\left(\gamma_t^S - \gamma_s^S\right)\right\| \\ \leq\ & C_{p,f}\left\|\pi_1\left(Y\right)\right\|_{p-var,[s,t]}\left\|\gamma^S\right\|_{2^{-1}p-var,[s,t]} \leq C_{p,\beta,f}\left(\omega\left(s,t\right)^{\frac{1}{p}}\vee\omega\left(s,t\right)\right)\omega\left(s,t\right)^{\frac{2}{p}} \\ \leq\ & C_{p,\beta,f}\omega\left(s,t\right)^{\frac{\beta+1}{p}}\vee\omega\left(s,t\right)^2. \end{aligned}$$

The Theorem below follows from Theorem 12.10 in [7].

Theorem 4.3. *Suppose γ is a p-rough path for some $p\in[2,3)$ on $[0,T]$ taking values in $\left(T^{(2)}\left(\mathbb{R}^d\right),\|\cdot\|\right)$, and $f:\mathbb{R}^e\to L\left(\mathbb{R}^d,\mathbb{R}^e\right)$ is $Lip\left(\beta\right)$ for $\beta>p$. Suppose Y^i, $i=1,2$, is the solution to the rough differential equations:*

$$dY^i = f\left(Y^i\right)d\gamma,\ Y_0^i = \xi^i \in T^{(2)}\left(\mathbb{R}^e\right). \tag{4.6}$$

Then with $\omega\left(s,t\right):=\|\gamma\|_{p-var,[s,t]}^p$ we have (with $Y_{s,t}^i := \left(Y_s^i\right)^{-1}\otimes Y_t^i$)

$$\max_{k=1,2}\sup_{0\leq s\leq t\leq T}\frac{\left|\pi_k\left(Y_{s,t}^1\right)-\pi_k\left(Y_{s,t}^2\right)\right|}{\omega\left(s,t\right)^{\frac{k}{p}}} \leq C_{p,\beta,f}\left|\pi_1\left(\xi^1\right)-\pi_1\left(\xi^2\right)\right|\exp\left(C_{p,\beta,f}\omega\left(0,T\right)\right). \tag{4.7}$$

Similar as for Theorem 4.2, we have to check that the extra term satisfies (4.7) as well. Indeed, based on the estimate of $\left|\pi_1\left(Y_{s,t}^1\right)-\pi_1\left(Y_{s,t}^2\right)\right|$ in (4.7) and that ω is a control, we have, for any $[s,t]\subseteq[0,T]$,

$$\begin{aligned} \left\|\pi_1\left(Y^1\right)-\pi_1\left(Y^2\right)\right\|_{p-var,[s,t]} & \leq C_{p,\beta,f}\omega\left(s,t\right)^{\frac{1}{p}}\left|\pi_1\left(\xi^1\right)-\pi_1\left(\xi^2\right)\right|\exp\left(C_{p,\beta,f}\omega\left(0,T\right)\right) \\ & \leq C_{p,\beta,f}\left|\pi_1\left(\xi^1\right)-\pi_1\left(\xi^2\right)\right|\exp\left(C_{p,\beta,f}\omega\left(0,T\right)\right), \end{aligned}$$

and for any $s\in[0,T]$,

$$\begin{aligned} & \left|\pi_1\left(Y_s^1\right)-\pi_1\left(Y_s^2\right)\right| \\ \leq\ & \left|\pi_1\left(\xi^1\right)-\pi_1\left(\xi^2\right)\right| + C_{p,\beta,f}\omega\left(0,s\right)^{\frac{1}{p}}\left|\pi_1\left(\xi^1\right)-\pi_1\left(\xi^2\right)\right|\exp\left(C_{p,\beta,f}\omega\left(0,T\right)\right) \\ \leq\ & C_{p,\beta,f}\left|\pi_1\left(\xi^1\right)-\pi_1\left(\xi^2\right)\right|\exp\left(C_{p,\beta,f}\omega\left(0,T\right)\right). \end{aligned}$$

Then since f is $Lip\left(\beta\right)$ for $\beta>p\geq 2$ combined with Lemma 10.22 [7] and (4.3), we have

$$\begin{aligned} & \left\|f\left(\pi_1\left(Y^1\right)\right)^{\otimes 2}-f\left(\pi_1\left(Y^2\right)\right)^{\otimes 2}\right\|_{p-var,[s,t]} \\ \leq\ & C_{p,f}\left\|\pi_1\left(Y^1\right)-\pi_1\left(Y^2\right)\right\|_{p-var,[s,t]} \\ & + C_{p,f}\left(\textstyle\sum_{i=1,2}\left\|\pi_1\left(Y^i\right)\right\|_{p-var,[s,t]}\sup_{u\in[s,t]}\left|\pi_1\left(Y_u^1\right)-\pi_1\left(Y_u^2\right)\right|\right) \\ \leq\ & C_{p,\beta,f}\left(1+\omega\left(s,t\right)^{\frac{1}{p}}\vee\omega\left(s,t\right)\right)\left|\pi_1\left(\xi^1\right)-\pi_1\left(\xi^2\right)\right|\exp\left(C_{p,\beta,f}\omega\left(0,T\right)\right) \\ \leq\ & C_{p,\beta,f}\left|\pi_1\left(\xi^1\right)-\pi_1\left(\xi^2\right)\right|\exp\left(C_{p,\beta,f}\omega\left(0,T\right)\right). \end{aligned}$$

Hence, based on Young integral and that $\left\|\gamma^S\right\|_{2^{-1}p-var,[s,t]}\leq C_d\omega\left(s,t\right)^{\frac{2}{p}}$ as in (4.1), we have

$$\begin{aligned} & \left\|\int_s^t f\left(\pi_1\left(Y_u^1\right)\right)^{\otimes 2}d\gamma_u^S - \int_s^t f\left(\pi_1\left(Y_u^2\right)\right)^{\otimes 2}d\gamma_u^S\right\| \\ \leq\ & C_p\left\|f\left(\pi_1\left(Y^1\right)\right)^{\otimes 2}-f\left(\pi_1\left(Y^2\right)\right)^{\otimes 2}\right\|_{p-var,[s,t]}\left\|\gamma^S\right\|_{\frac{p}{2}-var,[s,t]} \\ & + C_f\left|\pi_1\left(Y_s^1\right)-\pi_1\left(Y_s^2\right)\right|\left\|\gamma^S\right\|_{\frac{p}{2}-var,[s,t]} \\ \leq\ & C_{p,\beta,f}\omega\left(s,t\right)^{\frac{2}{p}}\left|\pi_1\left(\xi^1\right)-\pi_1\left(\xi^2\right)\right|\exp\left(C_{p,\beta,f}\omega\left(0,T\right)\right). \end{aligned}$$

4.2 Proofs of Theorem 3.3 and Corollary 3.6

Before proceeding to details of the proof of Theorem 3.3, we first give a sketch of the proof which may help to make the idea clearer. When f is $Lip(\beta)$ for $\beta > p$, for $\eta \in T^{(2)}(\mathbb{R}^e)$, we denote by

$$\pi_f(s, \eta)$$

the unique solution to the RDE:

$$dY = f(Y)\, d\mathcal{I}_2(\gamma, M), \ \ y_s = \eta. \tag{4.8}$$

For a finite partition $D = \{t_j\}_{j=0}^n$ of $[0, T]$, suppose y^D is defined as in (3.14). Since y^D by definition is piecewise constant and Y is continuous, to prove the uniform convergence of y^D to Y as $|D| \to 0$ it is sufficient to prove that y^D converge to Y uniformly on $\{t_j\}_{j=0}^n$. For each finite partition $D = \{t_j\}_{j=0}^n$ of $[0, T]$, we generate a sequence of RDE solutions driven by the same rough path $\mathcal{I}_2(\gamma, M)$ along the same vector field f but with different starting time t_j and with different initial value $y_{t_j}^D$, $j = 0, 1, \ldots, n$. Then Y resp. y^D is the first resp. last solution in the sequence, and if we want to compare $y_{t_j}^D$ with Y_{t_j} then we rewrite

$$y_{t_j}^D - Y_{t_j} = \sum_{i=0}^{j-1} \left(\pi_f\left(t_{i+1}, y_{t_{i+1}}^D\right)_{t_j} - \pi_f\left(t_i, y_{t_i}^D\right)_{t_j} \right). \tag{4.9}$$

Since the solution is unique, we have

$$\pi_f\left(t_i, y_{t_i}^D\right)_{t_j} = \pi_f\left(t_{i+1}, \pi_f\left(t_i, y_{t_i}^D\right)_{t_{i+1}}\right)_{t_j},$$

and the difference between two adjacent solutions can be expressed as:

$$\pi_f\left(t_{i+1}, y_{t_{i+1}}^D\right)_{t_j} - \pi_f\left(t_{i+1}, \pi_f\left(t_i, y_{t_i}^D\right)_{t_{i+1}}\right)_{t_j} \tag{4.10}$$

$$\begin{aligned} = \ & y_{t_{i+1}}^D \otimes \left(\pi_f\left(t_{i+1}, y_{t_{i+1}}^D\right)_{t_{i+1}, t_j} - \pi_f\left(t_{i+1}, \pi_f\left(t_i, y_{t_i}^D\right)_{t_{i+1}}\right)_{t_{i+1}, t_j} \right) \\ & + y_{t_i}^D \otimes \left(y_{t_i, t_{i+1}}^D - \pi_f\left(t_i, y_{t_i}^D\right)_{t_i, t_{i+1}} \right) \otimes \pi_f\left(t_{i+1}, \pi_f\left(t_i, y_{t_i}^D\right)_{t_{i+1}}\right)_{t_{i+1}, t_j}. \end{aligned} \tag{4.11}$$

Based on Theorem 4.3, the difference between the two increments in the first term in (4.11) can be relegated to the difference between their first level initial values:

$$\pi_1\left(y_{t_{i+1}}^D\right) - \pi_1\left(\pi_f\left(t_i, y_{t_i}^D\right)_{t_{i+1}}\right) = \pi_1\left(y_{t_i, t_{i+1}}^D\right) - \pi_1\left(\pi_f\left(t_i, y_{t_i}^D\right)_{t_i, t_{i+1}}\right).$$

Then combined with the expression of the second term in (4.11), we would need two elements in our proof:

(1) an estimate of $\epsilon_i := \left| y_{t_i, t_{i+1}}^D - \pi_f\left(t_i, y_{t_i}^D\right)_{t_i, t_{i+1}} \right|$ for all i,

(2) the uniform boundedness of y^D in D so that based on (4.9) and (4.11) the difference between Y and y^D can be bounded by a term comparable to $\sum_i \epsilon_i$.

The estimate of ϵ_i mainly follows from Theorem 4.2. The uniform boundedness of y^D in D can be proved by mathematical induction. The reason that we can employ induction is that based on (4.11) only the first $(k-1)$ levels of y^D contribute to the kth level difference in (4.10) because the 0th level of any solution is identically 1. Hence, by using the uniform boundedness of the first $(k-1)$ levels of y^D, we can prove the kth level convergence of y^D to Y as $|D| \to 0$, which implies the uniform boundedness of the kth level of y^D in D.

Proof of Theorem 3.3. Define $\omega_i : \{(s,t)\,|0\leq s\leq t\leq T\}\rightarrow\overline{\mathbb{R}^{+}}$, $i=1,2$, by, for any $0\leq s\leq t\leq T$,

$$\omega_1(s,t):=\|\gamma\|_{p-var,[s,t]}^{p}+\|\langle M\rangle\|_{1-var,[s,t]}^{\frac{p}{2}},\ \omega_2(s,t):=\left\|\gamma^{(M,R)}\right\|_{p-var,[s,t]}^{p}.$$

Then ω_1 is deterministic and $\omega_1(0,T)<\infty$ (since γ is a p-rough path and $\langle M\rangle$ is of bounded variation). Based on the assumption (3.16) (on p9), we have

$$\mathbb{E}\left(\omega_2(0,T)^2\right)<\infty. \tag{4.12}$$

For $0\leq s\leq t\leq T$, we denote $y_{s,t}^{D}:=\left(y_s^D\right)^{-1}\otimes y_t^D$. Recall $\left\{y^{i,j}\right\}_{i=1,2}$ in (3.15):

$$\begin{aligned} dy^{1,j} &= f\left(y^{1,j}\right)d\gamma,\ y_{t_j}^{1,j}=y_{t_j}^{D},\\ dy^{2,j} &= f\left(y^{2,j}\right)d\gamma^{(M,R)},\ y_{t_j}^{2,j}=y_{t_j}^{D},\end{aligned}$$

and we have

$$y_{t_j,t_{j+1}}^{D}=y_{t_j,t_{j+1}}^{1,j}\otimes\mathbb{E}\left(y_{t_j,t_{j+1}}^{2,j}\right)^{-1}\otimes y_{t_j,t_{j+1}}^{1,j},\ j\geq 0. \tag{4.13}$$

Based on (4.13), we have

$$\pi_1\left(y_{t_j,t_{j+1}}^{D}\right)=2\pi_1\left(y_{t_j,t_{j+1}}^{1,j}\right)-\mathbb{E}\left(\pi_1\left(y_{t_j,t_{j+1}}^{2,j}\right)\right).$$

Since f is $Lip(\beta)$ for $\beta>p\geq 2$, f is $Lip(2)$. By using the Euler estimate of solution to RDE ((4.4) in Theorem 4.2), we have, on any $[t_j,t_{j+1}]$,

$$\begin{aligned} &\left|\pi_1\left(y_{t_j,t_{j+1}}^{D}\right)-\pi_1\left(\pi_f\left(t_j,y_{t_j}^{D}\right)_{t_j,t_{j+1}}\right)\right|\\ =&\left|2\pi_1\left(y_{t_j,t_{j+1}}^{1,j}\right)-\mathbb{E}\left(\pi_1\left(y_{t_j,t_{j+1}}^{2,j}\right)\right)-\pi_1\left(\pi_f\left(t_j,y_{t_j}^{D}\right)_{t_j,t_{j+1}}\right)\right|\\ \leq& C_{p,f}\left(\omega_1(t_j,t_{j+1})^{\frac{3}{p}}+\mathbb{E}\left(\omega_2(t_j,t_{j+1})^{\frac{3}{p}}\right)\right)\\ &+\left|(Df)(f)\left(\pi_1\left(Y_{t_j}\right)\right)\left(\mathbb{E}\left(\int_{t_j}^{t_{j+1}}\left(M_u-M_{t_j}\right)\otimes\circ dM_u\right)-\frac{1}{2}\langle M\rangle_{t_j,t_{j+1}}\right)\right|.\end{aligned}$$

Since $M=\int\phi dB$ with ϕ a fixed path taking values in $d\times d$ matrices, we have

$$\mathbb{E}\left(\int_{t_j}^{t_{j+1}}\left(M_u-M_{t_j}\right)\otimes\circ dM_u\right)=\frac{1}{2}\mathbb{E}\left(\langle M\rangle_{t_j,t_{j+1}}\right)=\frac{1}{2}\langle M\rangle_{t_j,t_{j+1}}.$$

Hence, for any $t_j\in D$,

$$\left|\pi_1\left(y_{t_j,t_{j+1}}^{D}-\pi_f\left(t_j,y_{t_j}^{D}\right)_{t_j,t_{j+1}}\right)\right|\leq C_{p,f}\left(\mathbb{E}\left(\omega_2(t_j,t_{j+1})^{\frac{3}{p}}\right)+\omega_1(t_j,t_{j+1})^{\frac{3}{p}}\right). \tag{4.14}$$

For the second level, based on (4.13), we have

$$\begin{aligned} &\pi_2\left(y_{t_j,t_{j+1}}^{D}\right) \qquad (4.15)\\ =&\pi_2\left(y_{t_j,t_{j+1}}^{1,j}\otimes\mathbb{E}\left(y_{t_j,t_{j+1}}^{2,j}\right)^{-1}\otimes y_{t_j,t_{j+1}}^{1,j}\right)\\ =&2\pi_2\left(y_{t_j,t_{j+1}}^{1,j}\right)-\mathbb{E}\left(\pi_2\left(y_{t_j,t_{j+1}}^{2,j}\right)\right)+\left(\pi_1\left(y_{t_j,t_{j+1}}^{1,j}\right)-\pi_1\left(\mathbb{E}\left(y_{t_j,t_{j+1}}^{2,j}\right)\right)\right)^{\otimes 2}.\end{aligned}$$

Then, by using (4.15), combined with (4.3), (4.4) and (4.5) in Theorem 4.2, we get,

$$\begin{aligned} & \left|\pi_2\left(y^D_{t_j,t_{j+1}}\right) - f\left(\pi_1(y^D_{t_j})\right)^{\otimes 2}\left(\pi_2\left(\gamma_{t_j,t_{j+1}}\right) - \frac{1}{2}\left\langle M\right\rangle_{t_j,t_{j+1}}\right)\right| \qquad (4.16)\\ \leq\ & C_{p,f}\left(\mathbb{E}\left(\omega_2\left(t_j,t_{j+1}\right)^{\frac{3}{p}} \vee \omega_2\left(t_j,t_{j+1}\right)^2\right) + \omega_1\left(t_j,t_{j+1}\right)^{\frac{3}{p}} \vee \omega_1\left(t_j,t_{j+1}\right)^2\right) \qquad (4.17)\\ & + \left|\frac{1}{2}\left(Df\right)\left(f\right)\left(\xi_j\right)\left\langle M\right\rangle_{t_j,t_{j+1}}\right|^2 + C_{p,f}\left(\mathbb{E}\left(\omega_2\left(t_j,t_{j+1}\right)^{\frac{3}{p}}\right) + \omega_1\left(t_j,t_{j+1}\right)^{\frac{3}{p}}\right)^2\\ & + C_{p,f}\left(\mathbb{E}\left(\omega_2\left(t_j,t_{j+1}\right)^{\frac{3}{p}}\right) + \omega_1\left(t_j,t_{j+1}\right)^{\frac{3}{p}}\right)\\ & \times\left(\mathbb{E}\left(\omega_2\left(t_j,t_{j+1}\right)^{\frac{1}{p}} \vee \omega_2\left(t_j,t_{j+1}\right)\right) + \omega_1\left(t_j,t_{j+1}\right)^{\frac{1}{p}} \vee \omega_1\left(t_j,t_{j+1}\right)\right), \end{aligned}$$

where (4.17) estimates the error created by replacing $2\pi_2(y^{1,j}_{t_j,t_{j+1}}) - \mathbb{E}(\pi_2(y^{2,j}_{t_j,t_{j+1}}))$ by the corresponding Euler approximations, and the three lines after (4.17) estimate $(\pi_1(y^{1,j}_{t_j,t_{j+1}}) - \pi_1(\mathbb{E}(y^{2,j}_{t_j,t_{j+1}})))^{\otimes 2}$ based on (4.3) and (4.4) in Theorem 4.2. On the other hand, based on (4.5) in Theorem 4.2, we have,

$$\begin{aligned} & \left|\pi_2\left(\pi_f\left(t_j, y^D_{t_j}\right)_{t_j,t_{j+1}}\right) - f\left(\pi_1(y^D_{t_j})\right)^{\otimes 2}\left(\pi_2\left(\gamma_{t_j,t_{j+1}}\right) - \frac{1}{2}\left\langle M\right\rangle_{t_j,t_{j+1}}\right)\right|\\ \leq\ & C_{p,f}\ \omega_1\left(t_j,t_{j+1}\right)^{\frac{3}{p}} \vee \omega_1\left(t_j,t_{j+1}\right)^2. \end{aligned}$$

Hence, combined with (4.16), we get,

$$\begin{aligned} & \left|\pi_2\left(y^D_{t_j,t_{j+1}}\right) - \pi_2\left(\pi_f\left(t_j, y^D_{t_j}\right)_{t_j,t_{j+1}}\right)\right| \qquad (4.18)\\ \leq\ & C\left(p, f, \mathbb{E}\left(\omega_2\left(0,T\right)^2\right), \omega_1\left(0,T\right)\right)\\ & \times\left(\mathbb{E}\left(\omega_2\left(t_j,t_{j+1}\right)^{\frac{3}{p}} \vee \omega_2\left(t_j,t_{j+1}\right)^2\right) + \omega_1\left(t_j,t_{j+1}\right)^{\frac{3}{p}} \vee \omega_1\left(t_j,t_{j+1}\right)^2\right). \end{aligned}$$

Combining (4.14) and (4.18), if we define $\widetilde{\omega}_k : \{(s,t)\,|0 \leq s \leq t \leq T\} \to \overline{\mathbb{R}^+}$, $k = 1, 2$, by

$$\widetilde{\omega}_k\left(s,t\right) := \begin{cases} \mathbb{E}\left(\omega_2\left(s,t\right)^{\frac{3}{p}}\right) + \omega_1\left(s,t\right)^{\frac{3}{p}}, & k = 1\\ \mathbb{E}\left(\omega_2\left(s,t\right)^{\frac{3}{p}} \vee \omega_2\left(s,t\right)^2\right) + \omega_1\left(s,t\right)^{\frac{3}{p}} \vee \omega_1\left(s,t\right)^2, & k = 2 \end{cases}, \qquad (4.19)$$

then

$$\begin{aligned} & \left|\pi_k\left(y^D_{t_j,t_{j+1}}\right) - \pi_k\left(\pi_f\left(t_j, y^D_{t_j}\right)_{t_j,t_{j+1}}\right)\right| \qquad (4.20)\\ \leq\ & C\left(p, f, \mathbb{E}\left(\omega_2\left(0,T\right)^2\right), \omega_1\left(0,T\right)\right)\widetilde{\omega}_k\left(t_j,t_{j+1}\right), \forall j \geq 0,\ k = 1, 2. \end{aligned}$$

Based on our assumption (4.12) and that $p \in [2,3)$, we have

$$\lim_{|D|\to 0}\sum_{t_j\in D}\widetilde{\omega}_k\left(t_j,t_{j+1}\right) = 0,\ \ k = 1, 2. \qquad (4.21)$$

Since f is $Lip\left(\beta\right)$ for $\beta > p$, denote by Y the unique solution to the RDE

$$dY = f\left(Y\right)d\mathcal{I}_2\left(\gamma, M\right),\ \ Y_0 = \xi \in T^{(2)}\left(\mathbb{R}^e\right).$$

We want to prove

$$\lim_{|D|\to 0}\max_{k=1,2}\sup_{0\leq t\leq T}\left|\pi_k\left(y^D_t\right) - \pi_k\left(Y_t\right)\right| = 0. \qquad (4.22)$$

It is clear that

$$\pi_0\left(y_t^D\right)=\pi_0\left(Y_t\right)\equiv 1,$$

so (4.22) holds trivially at level 0. For integer $k=1,2$, suppose (4.22) holds for level $l\leq k-1$, we want to prove (4.22) at level k. Based on our inductive hypothesis, we have

$$\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|<\infty. \tag{4.23}$$

For $t_j\in D$, when $j=0$, $y_0^D=Y_0=\xi$. When $j=1$, based on (4.20), we have

$$\begin{aligned}\left|\pi_k\left(y_{t_1}^D-Y_{t_1}\right)\right| &= \sum_{l=0}^{k-1}\left|\pi_l\left(\xi\right)\right|\left|\pi_{k-l}\left(y_{0,t_1}^D-Y_{0,t_1}\right)\right|\\ &\leq C\left(p,f,\mathbb{E}\left(\omega_2\left(0,T\right)^2\right),\omega_1\left(0,T\right)\right)\left(\max_{0\leq l\leq k-1}\left|\pi_l\left(\xi\right)\right|\right)\sum_{l=1}^{k}\widetilde{\omega}_l\left(0,t_1\right).\end{aligned}$$

When $j\geq 2$, we have

$$\begin{aligned}&\left|\pi_k\left(y_{t_j}^D-Y_{t_j}\right)\right| \\ =&\sum_{i=0}^{j-1}\left|\pi_k\left(\pi_f\left(t_{i+1},y_{t_{i+1}}^D\right)_{t_j}-\pi_f\left(t_i,y_{t_i}^D\right)_{t_j}\right)\right|\\ \leq&\sum_{i=0}^{j-2}\left|\pi_k\left(\pi_f\left(t_{i+1},y_{t_{i+1}}^D\right)_{t_j}-\pi_f\left(t_{i+1},\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\right)_{t_j}\right)\right|\\ &+\left|\pi_k\left(y_{t_j}^D-\pi_f\left(t_{j-1},y_{t_{j-1}}^D\right)_{t_j}\right)\right|.\end{aligned} \tag{4.24}$$

Then for each $i=0,1,\ldots,j-2$,

$$\begin{aligned}&\pi_f\left(t_{i+1},y_{t_{i+1}}^D\right)_{t_j}-\pi_f\left(t_{i+1},\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\right)_{t_j}\\ =&\ y_{t_{i+1}}^D\otimes\pi_f\left(t_{i+1},y_{t_{i+1}}^D\right)_{t_{i+1},t_j}-\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\otimes\pi_f\left(t_{i+1},\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\right)_{t_{i+1},t_j}\\ =&\ y_{t_{i+1}}^D\otimes\left(\pi_f\left(t_{i+1},y_{t_{i+1}}^D\right)_{t_{i+1},t_j}-\pi_f\left(t_{i+1},\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\right)_{t_{i+1},t_j}\right)\\ &+y_{t_i}^D\otimes\left(y_{t_i,t_{i+1}}^D-\pi_f\left(t_i,y_{t_i}^D\right)_{t_i,t_{i+1}}\right)\otimes\pi_f\left(t_{i+1},\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\right)_{t_{i+1},t_j}.\end{aligned}$$

By using (4.23), Theorem 4.3 on p11 and (4.20) ($\widetilde{\omega}_1$ defined at (4.19)), we have, for $i=0,1,\ldots,j-2$,

$$\begin{aligned}&\left|\pi_k\left(y_{t_{i+1}}^D\otimes\left(\pi_f\left(t_{i+1},y_{t_{i+1}}^D\right)_{t_{i+1},t_j}-\pi_f\left(t_{i+1},\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\right)_{t_{i+1},t_j}\right)\right)\right|\\ \leq&\left(\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|\right)\\ &\times\left(\sum_{l=1}^{k}\left|\pi_l\left(\pi_f\left(t_{i+1},y_{t_{i+1}}^D\right)_{t_{i+1},t_j}-\pi_f\left(t_{i+1},\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\right)_{t_{i+1},t_j}\right)\right|\right)\\ \leq&\ C\left(p,\beta,f,\omega_1\left(0,T\right)\right)\left(\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|\right)\left|\pi_1\left(y_{t_i,t_{i+1}}^D-\pi_f\left(t_i,y_{t_i}^D\right)_{t_i,t_{i+1}}\right)\right|\\ \leq&\ C\left(p,\beta,f,\mathbb{E}\left(\omega_2\left(0,T\right)^2\right),\omega_1\left(0,T\right)\right)\left(\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|\right)\widetilde{\omega}_1\left(t_i,t_{i+1}\right)\end{aligned}$$

On the other hand, by using (4.23), (4.3) in Theorem 4.2 on p10 and (4.20) ($\widetilde{\omega}_l$ defined at (4.19)), we have

$$\begin{aligned}&\left|\pi_k\left(y_{t_i}^D\otimes\left(y_{t_i,t_{i+1}}^D-\pi_f\left(t_i,y_{t_i}^D\right)_{t_i,t_{i+1}}\right)\otimes\pi_f\left(t_{i+1},\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\right)_{t_{i+1},t_j}\right)\right|\\ \leq\ & C\left(p,\beta,f,\mathbb{E}\left(\omega_2\left(0,T\right)^2\right),\omega_1\left(0,T\right)\right)\left(\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|\right)\sum_{l=1}^k\widetilde{\omega}_l\left(t_i,t_{i+1}\right).\end{aligned}$$

Therefore, we have, for any $i=0,1,\ldots,j-2$,

$$\begin{aligned}&\left|\pi_k\left(\pi_f\left(t_{i+1},y_{t_{i+1}}^D\right)_{t_j}-\pi_f\left(t_{i+1},\pi_f\left(t_i,y_{t_i}^D\right)_{t_{i+1}}\right)_{t_j}\right)\right|\\ \leq\ & C\left(p,\beta,f,\mathbb{E}\left(\omega_2\left(0,T\right)^2\right),\omega_1\left(0,T\right)\right)\left(\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|\right)\sum_{l=1}^k\widetilde{\omega}_l\left(t_i,t_{i+1}\right).\end{aligned}$$

As a result,

$$\begin{aligned}&\sum_{i=0}^{j-2}\left|\pi_k\left(\pi_f\left(t_{i+1},y_{t_{i+1}}^D\right)_{t_j}-\pi_f\left(t_i,y_{t_i}^D\right)_{t_j}\right)\right|\\ \leq\ & C\left(p,\beta,f,\mathbb{E}\left(\omega_2\left(0,T\right)^2\right),\omega_1\left(0,T\right)\right)\\ &\times\left(\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|\right)\sum_{i=0}^{j-2}\left(\sum_{l=1}^k\widetilde{\omega}_l\left(t_i,t_{i+1}\right)\right).\end{aligned}\tag{4.25}$$

On the other hand, for the term left in (4.24),

$$\begin{aligned}&\left|\pi_k\left(y_{t_j}^D-\pi_f\left(t_{j-1},y_{t_{j-1}}^D\right)_{t_j}\right)\right|\\ =\ &\left|\pi_k\left(y_{t_{j-1}}^D\otimes\left(y_{t_{j-1},t_j}^D-\pi_f\left(t_{j-1},y_{t_{j-1}}^D\right)_{t_{j-1},t_j}\right)\right)\right|\\ \leq\ & C\left(p,\beta,f,\mathbb{E}\left(\omega_2\left(0,T\right)^2\right),\omega_1\left(0,T\right)\right)\\ &\times\left(\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|\right)\sum_{l=1}^k\widetilde{\omega}_l\left(t_{j-1},t_j\right).\end{aligned}\tag{4.26}$$

Therefore, combining (4.24), (4.25) and (4.26), we have

$$\begin{aligned}\left|\pi_k\left(y_{t_j}^D-Y_{t_j}\right)\right|\ \leq\ & C\left(p,\beta,f,\mathbb{E}\left(\omega_2\left(0,T\right)^2\right),\omega_1\left(0,T\right)\right)\\ &\times\left(\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|\right)\sum_{i=0}^{j-1}\left(\sum_{l=1}^k\widetilde{\omega}_l\left(t_i,t_{i+1}\right)\right)\end{aligned}$$

Then, based on (4.21) and the inductive assumption

$$\sup_{D\subset[0,T]}\max_{0\leq l\leq k-1}\sup_{0\leq t\leq T}\left|\pi_l\left(y_t^D\right)\right|<\infty,$$

we have

$$\lim_{|D|\to0}\max_{t_j\in D}\left|\pi_k\left(y_{t_j}^D\right)-\pi_k\left(Y_{t_j}\right)\right|=0.$$

Since Y is continuous and y^D is piecewise-constant, we have

$$\lim_{|D|\to0,D\subset[0,T]}\sup_{0\leq t\leq T}\left|\pi_k\left(y_t^D\right)-\pi_k\left(Y_t\right)\right|=0.$$

□

Proof of Corollary 3.6. (Z,M) is a $2d$-dimensional continuous martingale w.r.t. the filtration generated by Z and B, so can be enhanced by their Stratonovich integrals to a process whose sample paths are almost surely p-rough paths for any $p\in(2,3)$. Suppose Z is in $L^{4+\epsilon}$ for some $\epsilon>0$. Based on Theorem 4.1 (on page 10), we get (let $p:=2+2^{-1}\epsilon$)

$$\begin{aligned}\mathbb{E}\left(\|S_2(Z+M)\|_{p-var,[0,T]}^{2p}\right) &\leq C_{d,p}\mathbb{E}\left(|\langle Z+M\rangle_T|^p\right)\leq C_{d,p}\mathbb{E}\left(|\langle Z\rangle_T|^p\right)\\ &\leq C_{d,p}\mathbb{E}\left(|Z_T-Z_0|^{2p}\right)=C_{d,p}\mathbb{E}\left(|Z_T-Z_0|^{4+\epsilon}\right)<\infty.\end{aligned}$$

The second inequality holds because M is defined to be the Itô integral $\int\psi dB$ for the matrices-valued process ψ satisfying $\int\psi^T\psi du=\langle Z\rangle$ and the d-dimensional Brownian motion B is independent from Z so we have $\langle Z,M\rangle_T=0$ a.s. and $\langle M\rangle_T=\langle Z\rangle_T$ a.s..

As a result, we have

$$\mathbb{E}\left(\|S_2(Z+M)\|_{p-var,[0,T]}^{2p}\Big|Z\right)<\infty \text{ a.s..}$$

On the other hand, the Stratonovich integrals satisfy

$$\mathbb{E}\left(\int_0^t(Z_u-Z_0)\otimes\circ dM_u+\int_0^t(M_u-M_0)\otimes\circ dZ_u\Big|Z\right)=0,\ \forall t\in[0,T]\text{, a.s..}$$

Based on Theorem 3.3, Corollary holds. □

References

[1] Bichteler, K., Stochastic integration and L^p-theory of stochastic integration, *Ann. Prob.*, **9**, 48-89, (1981). MR-0606798

[2] Coutin, L., Qian, Z., Stochastic analysis, rough path analysis and fractional Brownian motion. *Probab. Theory Related Fields*, **122**(1),108-140, (2002). MR-1883719

[3] Davie, A. M., Differential equations driven by rough paths: an approach via discrete approximation. *Applied Mathematics Research eXpress*, abm009, (2008). MR-2387018

[4] Föllmer, H., Calcul d'Itô sans probabilités, *Seminaire de probabilites (Strasbourg)*, **15**, 143-150, (1981).

[5] Friz, P., Oberhauser, H., Rough path limits of the Wong–Zakai type with a modified drift term. *J. Funct. Anal.*, **256**(10), 3236-3256, (2009). MR-2504524

[6] Friz, P., Victoir, N., Differential equations driven by Gaussian signals. *Ann. Inst. H. Poincaré Probab. Statist.*, **46**(2), 369-413, (2010). MR-2667703

[7] Friz, P., Victoir, N., *Multidimensional Stochastic Processes as Rough Paths, Theory and Applications*, Cambridge Univ. Press, (2010). MR-2604669

[8] Gubinelli, M., Controlling rough paths, *J. Functional Analysis*, 216(1), 86-140, (2004). MR-2091358

[9] Hairer, M., Kelly, D., Geometric versus non-geometric rough paths, *Ann. Inst. H. Poincaré Probab. Statist.*, **51**(1), 207-251, (2015). MR-3300969

[10] Hairer, M., Weber, H., Rough Burgers-like equations with multiplicative noise, *Probab. Theory Related Fields*, **155**(1), 71-126, (2013). MR-3010394

[11] Itô, K., Stochastic integral, *Proc. Imp. Acad. Tokyo*, **20**, 519-524, (1944). MR-0014633

[12] Itô, K., On stochastic differential equations. *Proc. Imp. Acad. Tokyo*, **22**, 32-35, (1946).

[13] Karandikar, R. L., Pathwise solution of stochastic differential equations, *Sankhya A*, **43**, 121-132, (1981). MR-0666376

[14] Lejay, A., Victoir, N., On (p,q)-rough paths, *Journal of Differential Equations*, **225**, 103-133, (2006). MR-2228694

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

[16] Lyons, T. J., Caruana, M., Lévy, T., Picard, J., *Differential equation driven by rough paths*. Springer, (2007). MR-2314753

[17] Lyons, T. J., Qian, Z. M., Calculus for multiplicative functionals, Itô's lemma and differential equations, *Itô's Stochastic Calculus and Probability Theory*, Springer, Tokyo, 233-250, (1996). MR-1439528

[18] Lyons, T. J., Qian., Z., *System control and rough paths*, OxfordUniv. Press, (2002). MR-2036784

[19] Lyons, T. J., Stoica, L., The limits of stochastic integrals of differential forms. *Ann. Probab.*, **27**(1), 1-49, (1999). MR-1681146

[20] Lyons, T. J., Yang, D., The partial sum process of orthogonal expansion as geometric rough process with Fourier series as an example—an improvement of Menshov-Rademacher theorem, *J. Functional Analysis*, 265 (12), 3067-3103, (2013). MR-3110495

[21] Russo, F., Vallois, P., Intégrales progressive, rétrograde et symétrique de processus nonadaptés. *C. R. Acad. Sci. Paris Ser. I Math.*, **312**(8), 615-618, (1991). MR-1101044

[22] Strook, D. W., Varadhan, S. R. S., On the support of diffusion processes with applications to the strong maximum principle, *Proc. Sixth Berkeley Symp. on math. Statist. and Prob.*, Vol 3, Univ. of Calif. Press, 333-359, (1972). MR-0400425

[23] Wong, E., Zakai, M., On the relation between ordinary and stochastic differential equations, *Int. J. Engineering Sci.*, **3**(2), 213-229, (1965). MR-0183023

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

Acknowledgments. The authors are grateful to referees for constructive comments that helped improve the presentation. D. Y. thanks Horatio Boedihardjo and Harald Oberhauser for constructive comments.