Numerical methods for approximating solutions to Rough Differential Equations

Lajos Gergely Gyurkó

St Hugh's College
University of Oxford

DPhil Thesis
Michaelmas 2008

Acknowledgments

I would like to express my gratitude to Prof Terry Lyons for being my supervisor, mentor and friend and also for the never-ending and illuminating discussions and debates on SDEs, RDEs, C++ and in general on maths and life.

I am grateful to Péter Medvegyev for his long-term inspiration.

Among my fellow DPhil students, I owe many thanks to Christian Litterer and Michael Coulon for reading and commenting on this thesis. I cannot mention here all the students and other researchers I have had the chance to talk to and learn about maths. I am grateful to all of them.

I owe many thanks to Sam Howison, Glenys Luke, Ben Hambly, Pierre Tarrès, Andrew Dancer and Christoph Reisinger for the teaching opportunities at the Mathematical Institute, St Catherine's, St Hugh's and Jesus Colleges. I am also thankful to Matthias Arnsdorf, Régis Guichard, Ian Dowker, Rob Baston and many other former colleagues at UBS and JP Morgan for the internship opportunities and support.

I thankfully acknowledge the financial support of the EPSRC and the Oxford-Man Institute.

Finally, I am deeply grateful to my wife, parents, sisters and in-laws for their many forms of devoted support and in particular to my wife Tündi for her love and (almost) unlimited patience.

Dedicated to my sons Domonkos and Félix
and to the loving memory of my late grandparents
Boldizsár, Borbála and Lajos

CONTENTS

List of Figures

.

LIST OF NOTATIONS

In the following list, $R = (r_1, \dots, r_l)$ and $Q = (q_1, \dots, q_m)$ are a k-multi-indices and Π is a positive real k-tuple (see Definition 2.1.1). V and $V^1, \dots, V^k$ are Banach spaces such that $V = V^1 \oplus \cdots \oplus V^k$.

- $OS(k_1, \dots, k_n)$ see Definition 1.2.3
- $T((V))$ the space of formal series of tensors of V (Definition 1.1.1)
- $\mathcal{A}^\Pi = \mathcal{A}^k$ set of all k-multi-indices of finite length (Definition 2.1.1)
- ϵ the empty k-multi-index (Definition 2.1.1)
- $-R = -(r_1, r_2, \dots, r_{l-1}, r_l) = (r_2, \dots, r_{l-1}, r_l)$ (Definition 2.1.1)
- $R- = (r_1, r_2, \dots, r_{l-1}, r_l)- = (r_1, r_2, \dots, r_{l-1})$ (Definition 2.1.1)
- $R * Q = (r_1, \dots, r_l) * (q_1, \dots, q_m) = (r_1, \dots, r_l, q_1, \dots, q_m)$ (Definition 2.1.1)
- $V^{\otimes R} = V^{r_1} \otimes \cdots \otimes V^{r_l}$ (Definition 2.1.2)
- $\pi_A : A \oplus B \to A$ general form of the canonical projection, special case $\pi_V : T((V)) \to V$ (Definition 2.1.2)
- $\pi_R = \pi_{V^{r_1} \otimes \cdots \otimes V^{r_l}} : T((V)) \to V^{\otimes R}$ canonical projection (Definition 2.1.2)
- $\pi_{T((V^i))} : T((V)) \to T((V^i))$ canonical projection (Definition 2.1.2)
- $\mathbf{a}^R = \pi_R \mathbf{a}$ for $\mathbf{a} \in T((V))$ (Definition 2.1.2)
- $v_R := (\pi_{(r_1)} v) \otimes \cdots \otimes (\pi_{(r_l)} v) \in V^{\otimes R}$ for $v \in V$ (Definition 2.1.2)
- $n_j(R) := \text{card}\{i | r_i = j, r_i \in R\}$ (Definition 2.1.3)
- $deg_\Pi(R) = \sum_{i=1}^k \frac{n_{j,\Pi}(R)}{p_j}$ (Definition 2.1.3)
- $\Gamma_\Pi(R) = \left(\frac{n_1(R)}{p_1}\right)! \cdots \left(\frac{n_k(R)}{p_k}\right)!$ for $R \in \mathcal{A}^k$ (Definition 2.1.3)
- $\mathcal{A}_s^\Pi := \left\{R = (r_1, \dots, r_l) \middle| \, l \geq 1, \, deg_\Pi(R) \leq s\right\}$ (Definition 2.1.3)

- $B_s^\Pi = \left\{ \mathbf{a} \in T((V)) | \, \forall R \in \mathcal{A}_s^\Pi, \; \mathbf{a}^R := 0 \right\}$ (Definition 2.1.3)
- $T^{(\Pi,s)}(V) := T((V))/B_s^\Pi$ (Definition 2.1.3)
- $\pi_{\Pi,s} : T((V)) \to \bigoplus_{R\in\mathcal{A}_s^\Pi} V^R$ canonical projection (Remark 2.1.1)
- $\mathbf{a}^{\Pi,s} = \pi_{\Pi,s}\mathbf{a}$ (Remark 2.1.1)
- the restriction of π_V to $T^{(\Pi,s)}(V)$ is also denoted by π_V
- $S^\Pi = \{0 = s_0 < s_1 < s_2 < \cdots\}$ (Remark 2.1.2)
- $V^{\otimes(\Pi,s_m)} = \left\{ \left(\pi_{\Pi,s_m} - \pi_{\Pi,s_{m-1}}\right)\mathbf{a} \mid \mathbf{a} \in T((V)) \right\}$ (Remark 2.1.2)
- $d_\Pi(\mathbf{X},\mathbf{Y}) := \max_{R\in\mathcal{A}_1^k} \sup_{\mathcal{D}\in\mathcal{P}([0,T])} \left(\sum_{\mathcal{D}} \left\| \mathbf{X}^R_{t_{l-1},t_l} - \mathbf{Y}^R_{t_{l-1},t_l} \right\|^{1/deg_\Pi(R)} \right)^{deg_\Pi(R)}$ the Π-variational metric (Definition 2.1.6)
- $d_\Pi^\alpha(\mathbf{X},\mathbf{Y}) := \max_{R\in\mathcal{A}_\alpha^k} \sup_{\mathcal{D}\in\mathcal{P}([0,T])} \left(\sum_{\mathcal{D}} \left\| \mathbf{X}^R_{t_{l-1},t_l} - \mathbf{Y}^R_{t_{l-1},t_l} \right\|^{1/deg_\Pi(R)} \right)^{deg_\Pi(R)}$ (Definition 2.1.6)
- $\mathbf{a}^{\langle s_m \rangle} = \sum_{deg_{\Pi,1}(R)=s_m} \mathbf{a}^R$ for $\mathbf{a} \in T((V))$ (Definition 2.3.1)
- $\mathbf{a}^V = \pi_V \mathbf{a}$ for $\mathbf{a} \in T((V))$ (Definition 2.3.1)
- $\pi_{(s_{m_1},\ldots,s_{m_n})}(\mathbf{a}) = \sum_{\substack{R_1,\ldots,R_n\in\mathcal{A}^k \\ deg_\Pi(R_i-)=s_{m_i}, i=1,\ldots,n}} \pi_{R_1*\cdots*R_n}\mathbf{a}.$ for $s_{m_1},\ldots,s_{m_n} \in S^{(\Pi,1)}$ and $\mathbf{a} \in T((V))$ (Definition 2.3.2)
- $\Delta_T^{\mathcal{D}} = \{(s,t) | \, s,t \in \mathcal{D}, \; s \le t\}$. for $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$ (Definition 3.1.1)
- $d_{\Pi_Y}^{\alpha,\widehat{\mathcal{D}}}(\mathbf{Y},\widehat{\mathbf{Y}}) := \max_{\substack{R\in\mathcal{A}_\alpha^k \\ \overline{\mathcal{D}}\subseteq\widehat{\mathcal{D}}}} \left(\sum_{\overline{\mathcal{D}}} \left\| \mathbf{Y}^R_{t_{l-1},t_l} - \widehat{\mathbf{Y}}^R_{t_{l-1},t_l} \right\|^{1/deg_{\Pi_Y}(R)} \right)^{deg_{\Pi_Y}(R)}$ (Definition 3.1.4)
- $(\mathcal{B}_n)_{n>0}$ a sequence of tree-like sets (Definition 3.2.4)
- W_i^y, $i = 1,\ldots,k$ vector fields (Definition 3.3.8)
- Φ^y an algebra homomorphism (Definition 3.3.8)
- $W_i^{y,\mathcal{B}}$, $i = 1,\ldots,k$ vector fields for a tree-like set $\mathcal{B}$ (Definition 3.3.9)
- $\Phi^{y,\mathcal{B}}$ an algebra homomorphism (Definition 3.3.9)
- $\widehat{X}^R_{s,t}\left\{ g(\widehat{\mathbf{Y}}^U_{0,\cdot}) \right\} := \int_{s<u_1<\cdots<u_l<t} g(\widehat{\mathbf{Y}}^U_{0,u_1}) \mathrm{d}X^{(r_1)}_{0,u_1} \cdots \mathrm{d}X^{(r_l)}_{0,u_l} \in \mathbb{R}$ iterated integrals (see after Lemma 3.3.4)
- $\mathrm{Exp}(tF)(\xi)$ solution to the ODE driven by the vector field F and started at ξ (section 3.4.1)

- $B^R_{s,t} := \int_{s<t_1<\cdots<t_k<t} \circ \mathrm{d}B^{r_1}_{t_1} \circ \cdots \circ \mathrm{d}B^{r_k}_{t_k}$ Stratonovich iterated integrals (Definition 4.2.1)
- $B^R_{s,t}(Y.) := \int_{s<t_1<\cdots<t_k<t} Y_{t_1} \circ \mathrm{d}B^{r_1}_{t_1} \circ \cdots \circ \mathrm{d}B^{r_k}_{t_k}$ Stratonovich iterated integrals (Definition 4.2.1)
- $D^R_{s,t} := \int_{s<t_1<\cdots<t_k<t} \mathrm{d}B^{r_1}_{t_1} \cdots \mathrm{d}B^{r_k}_{t_k}$ Itô iterated integrals (Definition 4.2.1)
- $D^R_{s,t}(Y.) := \int_{s<t_1<\cdots<t_k<t} Y_{t_1} \mathrm{d}B^{r_1}_{t_1} \cdots \mathrm{d}B^{r_k}_{t_k}$ Itô iterated integrals (Definition 4.2.1)
- $\mathcal{H}_R$ a class of Itô integrable functions (introduced in Definition 4.2.3)
- $\mathcal{H}_R(\xi.)$ a class of Stratonovich integrable functions corresponding to the process $\xi.$ (introduced in Lemma 4.2.2)

Introduction

The core ideas for this thesis arose while attempting to generalize the method referred to as Cubature on Wiener space or as the Kusuoka-Lyons-Victoir (KLV) family of numerical methods (Kusuoka [17], Kusuoka [18], Lyons & Victoir [28]) to boundary problems. This generalization has not been completed yet, however we believe that some results have been achieved, which are crucial for the extension of KLV. In particular, we constructed numerical methods for approximating the solution to rough differential equations jointly with the truncated signature of the solution.

Stochastic differential equations (SDEs) under the KLV consideration are driven by Brownian motion and time. The paths of Brownian motion coupled with time and lifted to the appropriate tensor algebra are almost surely path-wise rough paths of inhomogeneous degree of smoothness. Another natural example of rough paths of inhomogeneous degree of smoothness arises when the truncated signature of a rough path itself is regarded as a rough path. When aiming to construct numerical approximations of SDEs and in general of rough differential equations (RDEs), we considered the general case, i.e. RDEs driven by rough paths of inhomogeneous degree of smoothness.

After sketching the core ideas of the Rough Paths Theory (ref. Lyons [26]) in Chapter 1, the versions of the core theorems corresponding to the inhomogeneous degree of smoothness case are stated and proved in Chapter 2 (Theorems 2.1.1, 2.2.1, 2.3.1 and 2.5.1) along with some auxiliary claims on the continuity of the solution in a certain sense, including an RDE-version of Gronwall's lemma (Proposition 2.6.3). It is clear, that rough paths of inhomogeneous degree of smoothness can be extended and regarded as p-rough paths for an appropriate p, and hence the integration theorem and the Universal Limit Theorem of [26] hold for the extended rough paths. The importance of Chapter 2 lies in showing that under the truly inhomogeneous treatment similar results hold under weaker conditions.

In Chapter 3, numerical schemes for approximating solutions to differential equations driven by rough paths of inhomogeneous degree of smoothness are constructed. We start with setting up some principles of approximations. Then a general class of local approximations is introduced. This class is used to construct global approximations by pasting together the local ones. A general sufficient condition on the local approximations implying global convergence is given and proved in Theorems 3.2.1 and 3.2.2. The next step is to construct particular local approximations in finite dimensions based on solutions to ordinary differential equations derived locally and satisfying the sufficient condition for global convergence. The core ingredients are the Log-signature Theorems 3.4.1 and 3.4.2, which can be regarded as generalizations of the results of Strichartz [35] as well as the rough paths versions of Castell [4] or Hu [13]. These local approximations require strong conditions on the one-form defining the rough differential equation. Finally, we show (mainly in Proposition 3.5.2) that when the local ODE-based schemes are applied in combination with rough polynomial approximations (justified in Proposition 3.5.1), the conditions on the one-form can be weakened. We note that the core ideas of polynomial approximations appeared in the paper by Davie [8] and were generalized by Caruana [3].

In Chapter 4, the results of Gyurkó & Lyons [12] on path-wise approximation of solutions to stochastic differential equations are recalled and extended to the truncated signature level of the solution. Furthermore, some practical considerations related to the implementation of high order schemes are described. The effectiveness of the derived schemes is demonstrated on numerical examples.

In Chapter 5, the background theory of the Kusuoka-Lyons-Victoir (KLV) family of weak approximations is recalled and linked to the results of Chapter 4. We highlight how the different versions of the KLV family are related. Moreover, some practical considerations are described, including the right choice of numerical ODE solvers. Finally, a numerical evaluation of the autonomous ODE-based versions of the family is carried out, focusing on SDEs in dimensions from 1 up to 4, using degree 3/2, 5/2 and 7/2 cubature formulas[1] and high order Runge-Kutta numerical solvers. We demonstrate the effectiveness and the occasional non-effectiveness of the numerical approximations in cases when the KLV family is used in its original version and also when used in combination with partial sampling methods (Monte-Carlo, TBBA) and Romberg extrapolation. Furthermore, we demonstrate the role of high order numerical ODE solvers within the KLV family.

[1] In [28], these cubature formulas are referred to as degree 3, degree 5 and degree 7 respectively.

CHAPTER 1

ROUGH DIFFERENTIAL EQUATIONS

In this section, an introduction to rough paths and rough differential equations is presented. We mainly follow Lyons [25], Lyons & Qian [26] and in some cases Lyons et al. [27].

1.1 INTRODUCTION TO ROUGH PATHS

In this chapter, V is a Banach space with norm $\|\cdot\|$ together with a sequence of tensor norms $\|\cdot\|_k$ satisfying the following compatibility condition

$$\|a \otimes b\|_{k+l} \leq \|a\|_k \|b\|_l, \ \forall a \in V^{\otimes k}, \ \forall b \in V^{\otimes l}.$$

We use the convention $V^{\otimes 0} = \mathbb{R}$.

Definition 1.1.1. *The space of formal series of tensors of V is defined to be the following space of sequences:*

$$T((V)) := \mathbb{R} \oplus V \oplus V^{\otimes 2} \oplus \cdots = \left\{ \mathbf{a} = (a_0, a_1, \dots) | \forall n \geq 0, a_n \in V^{\otimes n} \right\}$$

Given $\lambda \in \mathbb{R}$, $\mathbf{a}, \mathbf{b} \in T((V))$ as $\mathbf{a} = (a_0, a_1, \dots)$ and $\mathbf{b} = (b_0, b_1, \dots)$, we define the operations

$$\cdot : \mathbb{R} \times T((V)) \to T((V)) \text{ by } \lambda \cdot \mathbf{a} = (\lambda a_0, \lambda a_1, \dots)$$

$$+ : T((V)) \times T((V)) \to T((V)) \text{ by } \mathbf{a} + \mathbf{b} = (a_0 + b_0, a_1 + b_1, \dots)$$

and

$$\otimes : T((V)) \times T((V)) \to T((V)) \text{ by } \mathbf{a} \otimes \mathbf{b} = (c_0, c_1, \dots)$$

where for $n \geq 0$

$$c_n = \sum_{k=0}^{n} a_k \otimes b_{n-k}$$

$T((V))$ equipped with these operations is a real non-commutative algebra with unit

$$\mathbf{1} = (1, 0, 0, \dots).$$

Definition 1.1.2. *For $n \geq 1$, set $B_n = \{\mathbf{a} = (a_0, a_1, \dots) | a_0 = a_1 = \cdots = a_n = 0\}$ is an ideal in $T((V))$. The* truncated tensor algebra of V of order n *is defined as the quotient algebra*

$$T^{(n)}(V) = T((V))/B_n. \tag{1.1}$$

The canonical homomorphism $T((V)) \to T^{(n)}(V)$ *is denoted by* π_n.

$T^{(n)}(V)$ is canonically isomorphic to $\mathbb{R} \oplus V \oplus V^{\otimes 2} \oplus \cdots \oplus V^{\otimes n}$ equipped with the product $\otimes_n : T^{(n)}(V) \times T^{(n)}(V) \to T^{(n)}(V)$

$$(a_0, a_1, \dots, a_n) \otimes (b_0, b_1, \dots, b_n) = (c_0, c_1, \dots, c_n)$$

with

$$c_k = \sum_{i=0}^{k} a_i \otimes b_{k-i} \text{ for } 0 \leq k \leq n.$$

Definition 1.1.3 (Multiplicative functional). *Let Δ_T denote the set $\{(s,t) \in [0,T] \times [0,T] | s \leq t\}$ and let $n \geq 1$ be an integer. The continuous map $\mathbf{X} : \Delta_t \to T^{(n)}(V)$ is* multiplicative *if*

$$\begin{aligned} \mathbf{X}^0_{s,t} &= 1 \;\forall (s,t) \in \Delta_T \\ \mathbf{X}_{s,t} &= \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} \;\forall s,u,t \in [0,T],\; s \leq u \leq t \end{aligned}$$

where $\mathbf{X}_{s,t} := \mathbf{X}(s,t)$ for $(s,t) \in \Delta_T$ and

$$\mathbf{X}_{s,t} = (\mathbf{X}^0_{s,t}, \mathbf{X}^1_{s,t}, \dots, \mathbf{X}^n_{s,t}) \in \mathbb{R} \oplus V \oplus V^{\otimes 2} \oplus \cdots \oplus V^{\otimes n}.$$

Note that given a multiplicative functional $\mathbf{X} : \Delta_T \to T^{(n)}(V)$ and a $\mathbf{x} \in T^{(n)}(V)$, then

$$\mathbf{X}_t = \mathbf{x} \otimes \mathbf{X}_{0,t} = (\mathbf{X}^0_t, \mathbf{X}^1_t, \dots, \mathbf{X}^n_t) \in \mathbb{R} \oplus V \oplus V^{\otimes 2} \oplus \cdots \oplus V^{\otimes n} \tag{1.2}$$

defines a continuous $[0,T] \to T^{(n)}(V)$ path and $\mathbf{X}^1_t$ defines a continuous $[0,T] \to V$ path.

Example 1.1.1. Given a continuous path $X : [0,T] \to V$ of bounded variation, then for any $n \geq 1$ integer the following formula

$$\begin{aligned} \mathbf{X}_{s,t} &:= (1, \mathbf{X}^1_{s,t}, \dots, \mathbf{X}^n_{s,t}) \\ \mathbf{X}^i_{s,t} &:= \int_{s<u_1<\cdots<u_i<t} \mathrm{d}X_{u_1} \otimes \cdots \otimes \mathrm{d}X_{u_i},\; i = 1, \dots, n \end{aligned}$$

defines a $\mathbf{X} : \Delta_t \to T^{(n)}(V)$ multiplicative functional (ref.: Theorem 2.1.2 by Lyons [25]).

Note that in general the continuous path $X : [0,T] \to V$ does not uniquely determine $\mathbf{X} : \Delta_T \to T^{(n)}(V)$.

Definition 1.1.4 (Control function). *A* control function, *or* control, *on $[0,T]$ is a uniformly continuous non-negative function $\omega : \Delta_T \to [0,+\infty)$ which is super-additive, i.e.*

$$\omega(s,u) + \omega(u,t) \leq \omega(s,t) \;\forall s,u,t \in [0,T],\; s \leq u \leq t$$

and for which $\omega(t,t) = 0$ for all $t \in [0,T]$.

Example 1.1.2. Let C be a positive real number. The function $\omega(s,t) := C|t-s|$ satisfy all the conditions of Definition 1.1.4, hence giving us a particular example for control functions.

Definition 1.1.5 (Finite p-variation). *Let $p \geq 1$ be a real number and $n \geq 1$ be an integer. Let $\omega : [0,T] \to [0,+\infty)$ be a control. Let $\mathbf{X} : \Delta_T \to T^{(n)}(V)$ be a multiplicative functional.*

We say that $\mathbf{X}$ has finite p-variation on Δ_T controlled by ω *if*

$$\left\|\mathbf{X}^i_{s,t}\right\| \leq \frac{\omega(s,t)^{\frac{i}{p}}}{\beta\left(\frac{i}{p}\right)!} \;\forall i = 1,\dots,n,\; \forall (s,t) \in \Delta_T$$

where β is some real constant depending only on p and $x! := \Gamma(x+1)$.

We say that $\mathbf{X}$ has a finite p-variation *if there exists a control ω such that the conditions above are satisfied.*

Theorem 1.1.1 (Extension theorem). *Let $p \geq 1$ be a real number and $n \geq 1$ be an integer. Let $\mathbf{X} : \Delta_T \to T^{(n)}(V)$ be a multiplicative functional with finite p-variation controlled by a control ω. Assume that $n \geq \lfloor p \rfloor$.*

Then for every $m \geq \lfloor p \rfloor + 1$, there exists a unique continuous function $\mathbf{X}^m : \Delta_T \to V^{\otimes m}$ such that

$$(s,t) \to \mathbf{X}_{s,t} = (1, \mathbf{X}^1_{s,t}, \dots, \mathbf{X}^{\lfloor p \rfloor}_{s,t}, \dots, \mathbf{X}^m_{s,t}, \dots) \in T((V))$$

is a multiplicative functional with finite p-variation controlled by ω in the following sense:

$$\left\|\mathbf{X}^i_{s,t}\right\| \leq \frac{\omega(s,t)^{\frac{i}{p}}}{\beta\left(\frac{i}{p}\right)!} \;\forall i \geq 1,\; \forall (s,t) \in \Delta_T,$$

where

$$\beta \geq 2p^2\left(1 + \sum_{r=3}^{\infty}\left(\frac{2}{r-2}\right)^{\frac{\lfloor p \rfloor + 1}{p}}\right).$$

See Theorem 2.1.1 and Remark 2.1.3 for proof.

The above theorem tells us, that a $T((V))$-valued multiplicative functional of finite p-variation is determined by the first $\lfloor p \rfloor$ terms. This fact motivated the following definition.

Definition 1.1.6 (p-rough path). *Let $p \geq 1$ be a real number. A* p-rough path *in V is a multiplicative functional of degree $\lfloor p \rfloor$ in V with finite p-variation. The space of p-rough paths is denoted by $\Omega_p(V)$.*

Definition 1.1.7 (p-variation metric). *Let $C_{0,p}\left(\Delta_T, T^{(\lfloor p \rfloor)}(V)\right)$ denote the space of all continuous functions from the simplex Δ_T into the truncated tensor algebra $T^{(\lfloor p \rfloor)}(V)$ with finite p-variation. The p-variation metric on this linear space if defined as follows:*

$$d_p(\mathbf{X},\mathbf{Y}) := \max_{1 \leq i \leq \lfloor p \rfloor} \sup_{\mathcal{D} \in \mathcal{P}([0,T])} \left(\sum_{\mathcal{D}} \left\|\mathbf{X}^i_{t_{l-1},t_l} - \mathbf{Y}^i_{t_{l-1},t_l}\right\|^{\frac{p}{i}}\right)$$

where $\mathcal{P}([a,b])$ denotes the set of all finite partitions of the $[a,b]$ interval.

The space Ω_p equipped with the p-variation metric is a complete metric space (ref.: Lemma 3.3.3 of Lyons [26]).

Definition 1.1.8 (Geometric p-rough path)**.** *A* geometric p-rough path *is a p-rough path which can be expressed as a limit of* 1*-rough paths in the p-variation distance. The space of geometric p-rough paths in V is denoted by $G\Omega_p(V)$.*

In the rest of the paper, we will restrict our focus to differential equations driven by geometric p-rough paths.

1.2 Integration along rough paths

The following functional in the truncated tensor algebra is required for the introduction of integrals along rough paths.

Definition 1.2.1 (Almost p-rough path)**.** *Let $p \geq 1$ be a real number and ω a control. A function $\mathbf{Y} : \Delta_T \to T^{(\lfloor p \rfloor)}(V)$ as an* almost p-rough path *if*

(i) $\mathbf{Y}$ *has finite p-variation controlled by ω, i.e.*

$$\left\| \mathbf{Y}^i_{s,t} \right\| \leq \frac{\omega(s,t)^{\frac{i}{p}}}{\beta\left(\frac{i}{p}\right)!} \; \forall i = 1, \ldots, \lfloor p \rfloor, \; \forall (s,t) \in \Delta_T$$

(ii) $\mathbf{Y}$ *is* almost multiplicative *in the sense*

$$\left\| (\mathbf{Y}_{s,u} \otimes \mathbf{Y}_{u,t})^i - \mathbf{Y}^i_{s,t} \right\| \leq \omega(s,t)^{\theta} \; \forall i = 1, \ldots, \lfloor p \rfloor, \; \forall s,u,t \in [0,T], \; s \leq u \leq t$$

and for some $\theta > 1$.

Each almost rough path determines a rough path in the following sense.

Theorem 1.2.1. *Let $p \geq 1$ be a real number and ω be a control. Let $\mathbf{Y} : \Delta_T \to T^{(\lfloor p \rfloor)}(V)$ be an almost p-rough path with p-variation controlled by ω as in Definition 1.2.1. Then there exists a unique p-rough path $\mathbf{X} : \Delta_T \to T^{(\lfloor p \rfloor)}(V)$ such that*

$$\sup_{\substack{0 \leq s < t \leq T \\ i = 0, \ldots, \lfloor p \rfloor}} \frac{\left\| \mathbf{X}^i_{s,t} - \mathbf{Y}^i_{s,t} \right\|}{\omega(s,t)^{\theta}} < +\infty.$$

Moreover, there exists a constant K depending only on p,θ and $\omega(0,T)$, such that the supremum is smaller than K, and the p-variation of $\mathbf{X}$ is controlled by $K\omega$.

The reader is referred to [25] and [26] for proof.

We now recall the definition of $Lip(\gamma)$ functions.

Definition 1.2.2. *Let V and W be two Banach spaces, and let $k \geq 1$ be an integer. Let $\gamma \in (k, k+1]$ be a real number and suppose that F is a closed subset of V. Let $f : F \to W$ be a function and for each integer $j = 1, \ldots, k$, let $f^j : F \to L(V^{\otimes j}, W)$ be functions taking values in the space of symmetric j-linear mappings from V to W. The collection $(f = f^0, f^1, \ldots, f^k)$ is an element of $Lip(\gamma, F)$, if the following condition holds: there exists a constant M such that for each $j = 0, 1, \ldots, k$*

$$\sup_{x \in F} |f^j(x)| \leq M$$

and there exist functions $R_j : V \times V \to L(V^{\otimes j}, W)$, $j = 0, 1, \ldots, k$, such that for each $x, y \in F$ and each $v \in V^{\otimes j}$

$$f^j(y)(v) = \sum_{l=0}^{k-j} \frac{1}{l!} f^{j+l}(x)(v \otimes (y-x)^{\otimes l}) + R_j(x, y)(v)$$

and

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

The smallest constant M for which the inequalities hold for all j, is called the $Lip(\gamma, F)$-norm of f and is denoted by $\|f\|_{Lip(\gamma)}$.

Example 1.2.1. Let k be a positive integer and let the function $f : [0, 1] \to \mathbb{R}$ be defined as $f(x) := x^{k+1/2}$. Deriving the Taylor expansion of f up to $k+1$ terms, it's easy to see that f is $Lip(\gamma)$ for $0 < \gamma \leq k + 3/2$.

Definition 1.2.3. *Let $K \geq 1$ be an integer and let Σ_K denote the symmetric group on $V^{\otimes K}$. For all elements $\sigma \in \Sigma_K$ the image of $x_1 \otimes \cdots \otimes x_K$ under π is denoted by $x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(K)}$. Given integers $k_1, \ldots, k_n$ such that $k_1 + \cdots + k_n = K$, the set consisting of elements $\sigma \in \Sigma_K$ for which*

$$\sigma(1) < \sigma(2) < \cdots < \sigma(k_1),\ \sigma(k_1+1) < \sigma(k_1+2) < \cdots < \sigma(k_1+k_2), \ldots,$$

$$\sigma(K-k_n+1) < \sigma(K-k_n+2) < \cdots < \sigma(K), \text{ and } \sigma(k_1) < \sigma(k_2) < \cdots < \sigma(k_n)$$

is denoted by $OS(k_1, \ldots, k_n)$ and referred to as ordered shuffles.

The definition of the integral along rough paths is based on the following theorem.

Theorem 1.2.2. *Let $\mathbf{Z} : \Delta_T \to T^{(\lfloor p \rfloor)}(V)$ be a geometric p-rough path and let $\gamma > p$ be a real number. Furthermore, let $\alpha : V \to L(V, W)$ be a $Lip(\gamma - 1)$ function. Then $\mathbf{Y} : \Delta_T \to T^{(\lfloor p \rfloor)}(W)$ defined for all $(s, t) \in \Delta_T$ and $n \geq 1$ by*

$$\mathbf{Y}^n_{s,t} = \sum_{k_1, \ldots, k_n = 1}^{\lfloor p \rfloor} \alpha^{k_1 - 1}(\mathbf{Z}^1_{0,s}) \otimes \cdots \otimes \alpha^{k_n - 1}(\mathbf{Z}^1_{0,s}) \sum_{\pi \in OS(k_1, \ldots, k_n)} \pi^{-1} \mathbf{Z}^{k_1 + \cdots + k_n}_{s,t}$$

is an almost p-rough path.

The reader is referred to [25] and [26] for proof.

Definition 1.2.4 (Integral of a one-form). *Let $\mathbf{Z} : \Delta_T \to T^{(\lfloor p \rfloor)}(V)$ be a geometric p-rough path and let $\gamma > p$ be a real number. Let $\alpha : V \to L(V, W)$ be a Lip$(\gamma - 1)$ function. Let $\mathbf{Y} : \Delta_T \to T^{(\lfloor p \rfloor)}(W)$ be the almost p-rough path introduced in Theorem 1.2.2. The unique p-rough path associated to $\mathbf{Y}$ by Theorem 1.2.1 is called the integral of α along $\mathbf{Z}_t$ and it is denoted by $\int_\cdot^\cdot \alpha(\mathbf{Z}^1_{0,u})\mathrm{d}\mathbf{Z}_{0,u} : \Delta_T \to T^{(\lfloor p \rfloor)}(W)$. The notation $\int_\cdot^\cdot \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z}$ is also used.*

We use the notation $\int_s^t \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z}^n$ to denote the nth term of $\int_s^t \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z}$.

Theorem 1.2.3. *Let $\mathbf{Z} : \Delta_T \to T^{(\lfloor p \rfloor)}(V)$ be a geometric p-rough path and let $\gamma > p$ be a real number. Let $\alpha : V \to L(V, W)$ be a Lip$(\gamma - 1)$ function. The mapping $\mathbf{Z} \to \int \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z}$ is continuous from $G\Omega_p(V)$ to $G\Omega_p(W)$.*

Moreover, if the p-variation of $\mathbf{Z}$ is controlled by ω, then there exists a constant K depending only on γ, p, $\omega(0, T)$ and $\|\alpha\|_{Lip(\gamma-1)}$, such that

$$\left\| \int_s^t \alpha(\mathbf{Z}^1_{0,u})\mathrm{d}\mathbf{Z}^i_{0,u} \right\| \leq K\omega(s,t)^{\frac{i}{p}}.$$

The reader is referred to [25] and [26] for proof.

1.3 Differential equations driven by p-rough paths

Definition 1.3.1 (Rough differential equations). *Let V and W denote Banach spaces. Let $\gamma > p \geq 1$ be real numbers and let $f : W \to L(V, W)$ be a Lip$(\gamma - 1)$ function. Finally, let $\mathbf{X} \in G\Omega_p(V)$ and $\xi \in W$.*

Then we will say that $\mathbf{Z} \in G\Omega_p(V \oplus W)$ is a solution of the differential equation

$$\mathrm{d}\mathbf{Y}_t = f(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_0 = \xi$$

if $\pi_V(\mathbf{Z}) = \mathbf{X}$ and

$$\mathbf{Z} = \int h(\mathbf{Z})\mathrm{d}\mathbf{Z}$$

where $h : V \oplus W \to End(V \oplus W)$ defined by

$$h(x,y) = \begin{pmatrix} Id_V & 0 \\ f(y+\xi) & 0 \end{pmatrix}.$$

Theorem 1.3.1 (Universal Limit Theorem). *Let V and W denote Banach spaces. Let $\gamma > p \geq 1$ be real numbers and let $f : W \to L(V, W)$ be a Lip(γ) function. For all $\mathbf{X} \in G\Omega_p(V)$ and $\xi \in W$, the equation*

$$\mathrm{d}\mathbf{Y}_t = f(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_0 = \xi$$

admits a unique solution $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) \in G\Omega_p(V \oplus W)$ in the sense of Definition 1.3.1. Furthermore, this solution depends continuously on $\mathbf{X}$ and ξ in the p-variation topology.

The reader is referred to [25] or [26] for proof.

CHAPTER 2

DIFFERENTIAL EQUATIONS DRIVEN BY Π-ROUGH PATHS

In this chapter, we recall from Lyons [25] the definition of rough paths of inhomogeneous degrees of smoothness (referred to as Π-rough paths in this paper). We prove the extension theorem sketched in [25]. Then we present a slightly generalized form of the integration and differential equation theory of [25]. We show that each Π-rough path is also a p-rough path for some p, however the integral along Π-rough paths and solutions to differential equations driven by Π-rough paths exist under weaker conditions (less regularities) compared to the homogeneous smoothness case.

2.1 Π-ROUGH PATHS

When solving differential equations driven by rough paths, we will consider driving paths of inhomogeneous degrees of smoothness. A simple example is a stochastic differential equation (almost surely) regarded path-wise driven by a Brownian motion and time.

Definition 2.1.1. *Let k be a positive integer and let $\Pi = (p_1, \ldots, p_k)$ be a* real k-tuple, *such that $p_i \geq 1$ is a real number for all $i \in \{1, \ldots, k\}$. We say that $R = (r_1, \ldots, r_l)$ is a* k-multi-index *if $1 \leq r_j \leq k$ is an integer for all $j \in \{1, \ldots, l\}$. The empty multi-index is denoted by ϵ and the set of all k-multi-indexes of finite length is denoted by $\mathcal{A}^k$. Since the k-tuple Π determines the integer k, in some cases we will use the notation $\mathcal{A}^\Pi$ instead of $\mathcal{A}^k$.*

We define the k-multi-indexes $-R$ and $R-$ by

$$\begin{aligned} -R &= -(r_1, r_2, \ldots, r_{l-1}, r_l) = (r_2, \ldots, r_{l-1}, r_l) \\ R- &= (r_1, r_2, \ldots, r_{l-1}, r_l)- = (r_1, r_2, \ldots, r_{l-1}). \end{aligned}$$

The concatenation of the multi-indices $R = (r_1, \ldots, r_l)$ and $Q = (q_1, \ldots, q_m)$ is denoted by

$$R * Q = (r_1, \ldots, r_l) * (q_1, \ldots, q_m) = (r_1, \ldots, r_l, q_1, \ldots, q_m).$$

Definition 2.1.2. *Let k be a positive integer and let the Banach space V be of the form $V = V^1 \oplus \cdots \oplus V^k$ for some Banach spaces $V^1, \ldots, V^k$. In this case, the space of formal series of tensors of V is equivalently represented by*

$$T((V)) = \bigoplus_{n=0}^{\infty} V^{\otimes n} = \bigoplus_{(r_1,\ldots,r_l)\in\mathcal{A}_k} V^{r_1} \otimes \cdots \otimes V^{r_l}. \tag{2.1}$$

For a k-multi-index $R = (r_1, \ldots, r_l)$, we introduce the notation

$$V^{\otimes R} = V^{r_1} \otimes \cdots \otimes V^{r_l}. \tag{2.2}$$

In general, for a vector space $U = A \oplus B$, π_A and π_B denote the canonical projection onto A and B respectively, i.e. for $\mathbf{u} = \mathbf{a} + \mathbf{b} \in U$, such that $\mathbf{a} \in A$ and $\mathbf{b} \in B$, $\pi_A \mathbf{u} = \mathbf{a}$ and $\pi_B \mathbf{u} = \mathbf{b}$. We extensively use the projection π_V onto the V component of $T((V))$.

Let $\pi_R := \pi_{V^{r_1}\otimes\cdots\otimes V^{r_l}}$ and $\pi_{T((V^i))}$ for $i \in \{1, \ldots, k\}$ denote the canonical projections

$$\pi_R := \pi_{V^{r_1}\otimes\cdots\otimes V^{r_l}} : T((V)) \to V^{\otimes R}$$
$$\pi_{T((V^i))} : T((V)) \to T((V^i)).$$

We also use the notation

$$\mathbf{a}^R := \mathbf{a}^{(r_1,\ldots,r_l)} := \pi_{V^{r_1}\otimes\cdots\otimes V^{r_l}}(\mathbf{a}).$$

for $\mathbf{a} \in T((V))$, i.e.

$$\mathbf{a} = \sum_{R\in\mathcal{A}^k} \mathbf{a}^R. \tag{2.3}$$

Given an element $v \in V$ and a multi-index $R = (r_1, \ldots, r_l)$, we introduce the element v_R as follows:

$$v_R := (\pi_{(r_1)} v) \otimes \cdots \otimes (\pi_{(r_l)} v) \in V^{\otimes R}.$$

In this chapter we will assume the existence of tensor norms $\| \cdot \|_R$ for all $R \in \mathcal{A}^k$ satisfying

$$\|\mathbf{a} \otimes \mathbf{b}\|_{R*Q} \le \|\mathbf{a}\|_R \|\mathbf{b}\|_Q, \forall \mathbf{a} \in V^{\otimes R}, \forall \mathbf{b} \in V^{\otimes Q}.$$

Definition 2.1.3. *Let k be a positive integer and $\Pi = (p_1, \ldots, p_k)$ a real k-tuple as in Definition 2.1.1. Let V be a Banach space as in Definition 2.1.2. For the k-multi-index $R = (r_1, \ldots, r_l)$ we denote the* length *by $\|R\| = l$. Furthermore, we define the function n_j for $j \in \{1, \ldots, k\}$ by*

$$n_j(R) := card\{i | r_i = j, r_i \in R\}.$$

We introduce the Π-degree of R as

$$deg_\Pi(R) = \sum_{i=1}^{k} \frac{n_j(R)}{p_j}. \tag{2.4}$$

Note that $deg_\Pi(\epsilon) = 0$. *We also introduce the function* $\Gamma_\Pi : \mathcal{A}^k \to [0, \infty)$ *by*

$$\Gamma_\Pi(R) = \left(\frac{n_1(R)}{p_1}\right)! \cdots \left(\frac{n_k(R)}{p_k}\right)!, \text{ for } R \in \mathcal{A}^k.$$

Let $s \geq 0$ *be real. We introduce the set of k-multi-indices*

$$\mathcal{A}_s^\Pi := \left\{ R = (r_1, \ldots, r_l) \middle| \, l \geq 1, \, deg_\Pi(R) \leq s \right\}.$$

The set B_s^Π *defined by*

$$B_s^\Pi := \left\{ \mathbf{a} \in T((V)) | \; \forall R \in \mathcal{A}_s^\Pi, \; \mathbf{a}^R = 0 \right\}$$

is an ideal in $T((V))$. *The* truncated tensor algebra of order (k, s) *is defined as the quotient algebra*

$$T^{(\Pi,s)}(V) := T((V))/B_s^\Pi.$$

Remark 2.1.1. Note that $T^{(\Pi,s)}(V)$ is isomorphic to $\bigoplus_{R \in \mathcal{A}_s^\Pi} V^{\otimes R}$ equipped with the product $\otimes_{\Pi,s} : T^{(\Pi,s)}(V) \times T^{(\Pi,s)}(V) \to T^{(\Pi,s)}(V)$

$$\mathbf{a} \otimes_{\Pi,s} \mathbf{b} = \left(\sum_{Q \in \mathcal{A}_s^\Pi} \mathbf{a}^Q \right) \otimes_{\Pi,s} \left(\sum_{R \in \mathcal{A}_s^\Pi} \mathbf{b}^R \right) := \sum_{Q * R \in \mathcal{A}_s^\Pi} \mathbf{a}^Q \otimes \mathbf{b}^R$$

where

$$(q_1, \ldots, q_l) * (r_1, \ldots, r_m) := (q_1, \ldots, q_l, r_1, \ldots, r_m).$$

Given this isomorphism of $T^{(\Pi,s)}(V)$, for each $\mathbf{a} \in T((V))$, we will identify

$$\mathbf{a}^{\Pi,s} := \pi_{\Pi,s}(\mathbf{a}) := \left(\sum_{R \in \mathcal{A}_s^\Pi} \pi_R \right)(\mathbf{a}) = \sum_{R \in \mathcal{A}_s^\Pi} \mathbf{a}^R \tag{2.5}$$

with a corresponding element in $T^{(\Pi,s)}(V)$ and furthermore the multiplication

$$\otimes_{\Pi,s} : T^{(\Pi,s)}(V) \times T^{(\Pi,s)}(V) \to T^{(\Pi,s)}(V)$$

will also be denoted by $\otimes$.

Remark 2.1.2. Let S denote the set

$$S = \left\{ s = deg_\Pi(R) \mid R \in \mathcal{A}^k \right\}. \tag{2.6}$$

Note that S is closed for addition.

Also note that since for any $R \in \mathcal{A}^k$,

$$deg_\Pi(R) > \frac{\|R\|}{\max_{1 \leq i \leq k} p_i}$$

the set $\{R \in \mathcal{A}^k | deg_{\Pi}(R) \leq s\}$ is finite for all $s \geq 0$. This implies, that the elements of S can be listed in ascending order. We denote this list by S^{Π} and the ith element of the list is denoted by s_i^{Π} or s_i:

$$S^{\Pi} = \{0 = s_0 < s_1 < s_2 < \cdots\}.$$

Furthermore, we introduce the notation

$$V^{\otimes(\Pi,s_m)} = \left\{ \left(\pi_{\Pi,s_m} - \pi_{\Pi,s_{m-1}}\right)\mathbf{a} \mid \mathbf{a} \in T((V)) \right\}$$

for $m = 1, 2 \ldots$.

Definition 2.1.4 (Finite Π-variation). *Let the Banach space V be of the form $V = V^1 \oplus \cdots \oplus V^k$ for some Banach spaces $V^1, \ldots, V^k$. Let $\Pi = (p_1, \ldots, p_k)$ denote a k-tuple as in Definition 2.1.1 and let ω be a control. For a positive real q, the map $\mathbf{X} : \Delta_T \to T^{(\Pi,q)}$ is* multiplicative *if for all $0 \leq s < t \leq T$, $\pi_\varepsilon \mathbf{X}_{s,t} = \mathbf{1}$ and for all $0 \leq s < t < u \leq T$,*

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

Furthermore, $\mathbf{X}$ is of finite Π variation controlled by ω if

$$\left\| \mathbf{X}_{s,t}^R \right\| \leq \frac{\omega(s,t)^{deg_{\Pi}(R)}}{\beta^k \Gamma_{\Pi}(R)}$$

for all $(s,t) \in \Delta_T$ and for all k-multi-index $R \in \mathcal{A}_q^{\Pi}$.

The following lemma is used in the proof of Theorem 2.1.1.

Lemma 2.1.1 (Neo-classical Inequality). *For any $p \in [1, \infty)$, $n \in \mathbb{N}$ and $s, t, \geq 0$,*

$$\frac{1}{p^2} \sum_{i=0}^{n} \frac{s^{\frac{i}{p}} t^{\frac{n-i}{p}}}{\left(\frac{i}{p}\right)! \left(\frac{n-i}{p}\right)!} \leq \frac{(s+t)^{\frac{n}{p}}}{\left(\frac{n}{p}\right)!}.$$

The reader is referred to [26] for proof.

The following multinomial version of the Neo-classical Inequality can be proved by induction on the number of terms r of the product on the left-hand-side under the summation:

$$\frac{1}{p^{2r-2}} \sum_{\substack{k_1, \ldots, k_r \in \mathbb{N} \\ k_1 + \cdots + k_r = n}} \frac{x_1^{\frac{k_1}{p}}}{\left(\frac{k_1}{p}\right)!} \cdots \frac{x_r^{\frac{k_r}{p}}}{\left(\frac{k_r}{p}\right)!} \leq \frac{(x_1 + \cdots + x_r)^{\frac{n}{p}}}{\left(\frac{n}{p}\right)!}$$

for any $p \in [0, \infty)$, $r, n \in \mathbb{N}^*$ and $x_1, \ldots, x_r \geq 0$.

Theorem 2.1.1 (Extension theorem of multiplicative functionals of finite Π-variation). *Let k be a positive integer and let the Banach space V be of the form $V = V^1 \oplus \cdots \oplus V^k$ for some Banach spaces $V^1, \ldots, V^k$. Let $\Pi = (p_1, \ldots, p_k)$ be a real k-tuple as in Definition 2.1.1 and ω be a control. Let $\mathbf{X} : \Delta_T \to T^{(\Pi,1)}(V)$ be a multiplicative functional of finite Π-variation controlled by ω. Then* *for every k-multi-index $R \in \mathcal{A}^k / \mathcal{A}_1^\Pi$, there exists a unique continuous function $\mathbf{X}^R : \Delta_t \to V^{\otimes R}$ such that*

$$(s,t) \mapsto \mathbf{X}_{s,t} = \sum_{R \in \mathcal{A}^k} \mathbf{X}_{s,t}^R \in T((V))$$

is a multiplicative functional of finite Π-variation controlled by ω in the following sense:

$$\left\| \mathbf{X}_{s,t}^R \right\| \leq \frac{\omega(s,t)^{\frac{n_1(R)}{p_1} + \cdots + \frac{n_k(R)}{p_k}}}{\beta^k \left(\frac{n_1(R)}{p_1} \right)! \cdots \left(\frac{n_k(R)}{p_k} \right)!} = \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}$$

for all $R \in \mathcal{A}^k$, where

$$\beta \geq \left(p_1^2 \cdots p_k^2 \left(1 + \sum_{r=3}^{\infty} \left(\frac{2}{r-2} \right)^{s_{m^*+1}} \right) \right)^{1/k} \tag{2.7}$$

and s_{m^*} and s_{m^*+1} are the unique pair of adjacent elements of the list S^Π for which $s_{m^*} \leq 1 < s_{m^*+1}$. (This pair exists by Remark 2.1.2.)

Proof. Note that $\mathbf{X}(s_{m^*}) := \mathbf{X}$ is also a $\Delta_T \to T^{(\Pi, s_{m^*})}(V)$ multiplicative functional of finite Π-variation controlled by ω.

Now, we adjust the proof of Theorem 1.1.1 presented in [25]. The proof is an induction on m as follows. Let $m \geq m^*$ be an integer. Let us assume that $\mathbf{X}(s_m) \in T^{(\Pi, s_m)}$, $\mathbf{X}(s_m)$ is multiplicative and

$$\|\pi_R \mathbf{X}_{s,t}(s_m)\| \leq \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)} \tag{2.8}$$

for all $R \in \mathcal{A}_{s_m}^\Pi$. Then we show the existence and uniqueness of a $\mathbf{X}(s_{m+1}) : \Delta_T \to T^{(\Pi, s_{m+1})}(V)$ multiplicative functional with the same properties.

Let the map $\widehat{\mathbf{X}}_{\cdot,\cdot} : \Delta_T \to T^{(\Pi, s_{m+1})}(V)$ be defined by

$$\pi_R \widehat{\mathbf{X}}_{s,t} = \begin{cases} \pi_R \mathbf{X}_{s,t}(s_m) & \text{if } R \in \mathcal{A}_{s_m}^\Pi \\ 0_R & \text{if } R \in \mathcal{A}^k \backslash \mathcal{A}_{s_m}^\Pi \end{cases}$$

We claim that the multiplicative functional we want to construct is given by:

$$\mathbf{X}_{s,t}(s_{m+1}) = \lim_{\substack{\mathcal{D} \in \mathcal{P}([s,t]) \\ |\mathcal{D}| \to 0}} \widehat{\mathbf{X}}_{s,t}^{\mathcal{D}} \tag{2.9}$$

where for a partition $\mathcal{D} = \{s = t_0 \leq t_1 \leq \cdots \leq t_r = t\} \in \mathcal{P}([s,t])$ we define

$$\widehat{\mathbf{X}}_{s,t}^{\mathcal{D}} = \widehat{\mathbf{X}}_{s,t_1} \otimes \cdots \otimes \widehat{\mathbf{X}}_{t_{r-1},t}$$

with respect to the $T^{(\Pi, s_{m+1})}(V)$-multiplication.

Step 1

Firstly, we estimate the magnitude of $\pi_R \widehat{\mathbf{X}}^{\mathcal{D}}_{s,t}$ for $R \in \mathcal{A}^{\Pi}_{s_{m+1}}$. By the induction hypothesis (i.e by the multiplicative property of $\mathbf{X}(s_m)$ in $T^{(\Pi,(s_m))}(V)$ and by the inequality (2.8)), for all $R \in \mathcal{A}^{\Pi}_{s_m}$ we have:

$$\left\|\pi_R \widehat{\mathbf{X}}^{\mathcal{D}}_{s,t}\right\| \leq \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)} \tag{2.10}$$

By the super-additivity of the control ω there exists a $t_j \in \mathcal{D}$ such that

$$\omega(t_{j-1}, t_{j+1}) \leq \frac{2}{r-1}\omega(s,t).$$

Let $\mathcal{D}'$ be the partition $\mathcal{D}\backslash\{t_j\}$. Then

$$\begin{aligned}\widehat{\mathbf{X}}^{\mathcal{D}}_{s,t} - \widehat{\mathbf{X}}^{\mathcal{D}'}_{s,t} &= \widehat{\mathbf{X}}_{s,t_1} \otimes \cdots \otimes \widehat{\mathbf{X}}_{t_{j-2},t_{j-1}} \otimes \left(\widehat{\mathbf{X}}_{t_{j-1},t_j} \otimes \widehat{\mathbf{X}}_{t_j,t_{j+1}} - \widehat{\mathbf{X}}_{t_{j-1},t_{j+1}}\right) \\ &\quad \otimes \widehat{\mathbf{X}}_{t_{j+1},t_{j+2}} \otimes \cdots \otimes \widehat{\mathbf{X}}_{t_{r-1},t}\end{aligned} \tag{2.11}$$

By the multiplicative property of $\mathbf{X}(s_m)$,

$$\pi_{\Pi,s_m}\left(\widehat{\mathbf{X}}_{t_{j-1},t_j} \otimes \widehat{\mathbf{X}}_{t_j,t_{j+1}} - \widehat{\mathbf{X}}_{t_{j-1},t_{j+1}}\right) = 0$$

hence

$$\widehat{\mathbf{X}}_{t_{j-1},t_j} \otimes \widehat{\mathbf{X}}_{t_j,t_{j+1}} - \widehat{\mathbf{X}}_{t_{j-1},t_{j+1}} = \sum_{R*Q \in \mathcal{A}^{\Pi}_{s_{m+1}} \backslash \mathcal{A}^{\Pi}_{s_m}} \left(\pi_R \widehat{\mathbf{X}}_{t_{j-1},t_j}\right) \otimes \left(\pi_Q \widehat{\mathbf{X}}_{t_j,t_{j+1}}\right)$$

This representation implies, that

$$(\pi_{\Pi,s_{m+1}} - \pi_{\Pi,s_m})\left(\widehat{\mathbf{X}}_{t_{j-1},t_j} \otimes \widehat{\mathbf{X}}_{t_j,t_{j+1}} - \widehat{\mathbf{X}}_{t_{j-1},t_{j+1}}\right) = \widehat{\mathbf{X}}^{\mathcal{D}}_{s,t} - \widehat{\mathbf{X}}^{\mathcal{D}'}_{s,t}. \tag{2.12}$$

Since (2.11) is understood in the $T^{(\Pi,s_{m+1})}(V)$-multiplication, the equality (2.12) implies

$$\widehat{\mathbf{X}}^{\mathcal{D}}_{s,t} - \widehat{\mathbf{X}}^{\mathcal{D}'}_{s,t} = \sum_{R*Q \in \mathcal{A}^{\Pi}_{s_{m+1}} \backslash \mathcal{A}^{\Pi}_{s_m}} \left(\pi_R \widehat{\mathbf{X}}_{t_{j-1},t_j}\right) \otimes \left(\pi_Q \widehat{\mathbf{X}}_{t_j,t_{j+1}}\right),$$

and in particular

$$\begin{aligned}\left\|\pi_R\left(\widehat{\mathbf{X}}_{t_{j-1},t_j} \otimes \widehat{\mathbf{X}}_{t_j,t_{j+1}} - \widehat{\mathbf{X}}_{t_{j-1},t_{j+1}}\right)\right\| &= \left\|\sum_{P*Q=R}\left(\pi_P \widehat{\mathbf{X}}_{t_{j-1},t_j}\right) \otimes \left(\pi_Q \widehat{\mathbf{X}}_{t_j,t_{j+1}}\right)\right\| \\ &\leq \sum_{P*Q=R}\left\|\pi_P \widehat{\mathbf{X}}_{t_{j-1},t_j}\right\| \left\|\pi_Q \widehat{\mathbf{X}}_{t_j,t_{j+1}}\right\|.\end{aligned}$$

By the induction hypothesis,

$$\sum_{P*Q=R}\left\|\pi_P \widehat{\mathbf{X}}_{t_{j-1},t_j}\right\| \left\|\pi_Q \widehat{\mathbf{X}}_{t_j,t_{j+1}}\right\| \leq \sum_{P*Q=R} \frac{\omega(t_{j-1},t_j)^{deg_\Pi(P)}\omega(t_j,t_{j+1})^{deg_\Pi(Q)}}{\beta^k\left(\frac{n_1(P)}{p_1}\right)!\cdots\left(\frac{n_k(P)}{p_k}\right)!\beta^k\left(\frac{n_1(Q)}{p_1}\right)!\cdots\left(\frac{n_k(Q)}{p_k}\right)!}$$

Note that

$$\sum_{P*Q=R} \frac{\omega(t_{j-1},t_j)^{\frac{n_1(P)}{p_1}+\cdots+\frac{n_k(P)}{p_k}}\omega(t_j,t_{j+1})^{\frac{n_1(Q)}{p_1}+\cdots+\frac{n_k(Q)}{p_k}}}{\beta^k\left(\frac{n_1(P)}{p_1}\right)!\cdots\left(\frac{n_k(P)}{p_k}\right)!\beta^k\left(\frac{n_1(Q)}{p_1}\right)!\cdots\left(\frac{n_k(Q)}{p_k}\right)!}$$

$$\leq \prod_{l=1}^{k} \sum_{i_l=0}^{n_l(R)} \frac{\omega(t_{j-1},t_j)^{\frac{i_l}{p_l}} \omega(t_j,t_{j+1})^{\frac{n_l(R)-i_l}{p_l}}}{\beta^2 \left(\frac{i_l}{p_l}\right)! \left(\frac{n_l(R)-i}{p_l}\right)!}$$

Hence by Lemma 2.1.1

$$\begin{aligned} \sum_{P*Q=R} \left\|\pi_P \widehat{\mathbf{X}}_{t_{j-1},t_j}\right\| \left\|\pi_Q \widehat{\mathbf{X}}_{t_j,t_{j+1}}\right\| &\leq \prod_{l=1}^{k} \frac{p_l^2 \omega(t_{j-1},t_{j+1})^{\frac{n_l(R)}{p_l}}}{\beta^2 \left(\frac{n_l(R)}{p_l}\right)!} \\ &\leq \frac{p_1^2 \cdots p_k^2}{\beta^k} \left(\frac{2}{r-1}\right)^{deg_\Pi(R)} \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)} \end{aligned} \tag{2.13}$$

By successively dropping points from $\mathcal{D}$ until $\mathcal{D} = \{s,t\}$, we get

$$\left\|\pi_R\left(\widehat{\mathbf{X}}_{s,t}^{\mathcal{D}} - \widehat{\mathbf{X}}_{s,t}\right)\right\| \leq \left\|\pi_R\left(\widehat{\mathbf{X}}_{s,t}^{\mathcal{D}} - \widehat{\mathbf{X}}_{s,t}^{\mathcal{D}'}\right)\right\| + \left\|\pi_R\left(\widehat{\mathbf{X}}_{s,t}^{\mathcal{D}'} - \widehat{\mathbf{X}}_{s,t}^{\mathcal{D}''}\right)\right\| + \cdots + \left\|\pi_R \widehat{\mathbf{X}}_{s,t}\right\|$$

$$\leq \frac{p_1^2 \cdots p_k^2}{\beta^k} \left(1 + \sum_{r=3}^{\infty} \left(\frac{2}{r-2}\right)^{s_{m^*+1}}\right) \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}. \tag{2.14}$$

Hence if β satisfies

$$\beta^k \geq p_1^2 \cdots p_k^2 \left(1 + \sum_{r=3}^{\infty} \left(\frac{2}{r-2}\right)^{s_{m^*+1}}\right), \tag{2.15}$$

then the inequality (2.10) holds for all partitions $\mathcal{D}$ of $[s,t]$ and for all $R \in \mathcal{A}_{s_{m+1}}^{\Pi}$.

Step 2

We now prove the existence of the limit $\lim_{|\mathcal{D}|\to 0} \widehat{\mathbf{X}}_{s,t}^{\mathcal{D}}$, by showing that $\widehat{\mathbf{X}}_{s,t}^{\mathcal{D}}$ satisfies the Cauchy property. Let us suppose that $\mathcal{D}$ and $\mathcal{D}'$ are two partitions of $[s,t]$ such that $|\mathcal{D}|, |\mathcal{D}'| \leq \delta$ for some positive mesh size δ. Let $[t_j, t_{j+1}]$ be an interval in $\mathcal{D}$. The common refinement $\widehat{\mathcal{D}}$ of $\mathcal{D}$ and $\mathcal{D}'$ breaks up $[t_j, t_{j+1}]$ in a partition of $[t_j, t_{j+1}]$, which we denote by $\widehat{\mathcal{D}}_j$. Note that $\widehat{\mathbf{X}}_{t_j,t_{j+1}}^{\widehat{\mathcal{D}}} = \widehat{\mathbf{X}}_{t_j,t_{j+1}}$ up to multi-indices in $\mathcal{A}_{s_m}^{\Pi}$.

Since

$$\widehat{\mathbf{X}}_{t_j,t_{j+1}}^{\widehat{\mathcal{D}}_j} - \widehat{\mathbf{X}}_{t_j,t_{j+1}} = \pi_{\Pi,s_{m+1}} \widehat{\mathbf{X}}_{t_j,t_{j+1}}^{\widehat{\mathcal{D}}_j}, \tag{2.16}$$

then by the inequality (2.14) we have

$$\left\|\widehat{\mathbf{X}}_{t_j,t_{j+1}}^{\widehat{\mathcal{D}}_j} - \widehat{\mathbf{X}}_{t_j,t_{j+1}}\right\| \leq \omega(t_j,t_{j+1})^{s_{m+1}} \sum_{R \in \mathcal{A}_{s_{m+1}}^{\Pi} \setminus \mathcal{A}_{s_m}^{\Pi}} \frac{1}{\beta^k \Gamma_\Pi(R)}.$$

On the interval $[s,t]$

$$\begin{aligned} \widehat{\mathbf{X}}_{s,t}^{\widehat{\mathcal{D}}} - \widehat{\mathbf{X}}_{s,t} &= \widehat{\mathbf{X}}_{s,t_1}^{\widehat{\mathcal{D}}_0} \otimes \cdots \otimes \widehat{\mathbf{X}}_{t_{r-1},t}^{\widehat{\mathcal{D}}_{r-1}} - \widehat{\mathbf{X}}_{s,t_1} \otimes \cdots \otimes \widehat{\mathbf{X}}_{t_{r-1},t} \\ &= \sum_{j=0}^{r-1} \widehat{\mathbf{X}}_{s,t_1}^{\widehat{\mathcal{D}}_0} \otimes \cdots \otimes \widehat{\mathbf{X}}_{t_{j-2},t_{j-1}}^{\widehat{\mathcal{D}}_{j-1}} \otimes \left(\widehat{\mathbf{X}}_{t_j,t_{j+1}}^{\widehat{\mathcal{D}}_j} - \widehat{\mathbf{X}}_{t_j,t_{j+1}}\right) \otimes \widehat{\mathbf{X}}_{t_{j+1},t_{j+2}} \otimes \cdots \otimes \widehat{\mathbf{X}}_{t_{r-1},t} \\ &= \sum_{j=0}^{r-1} \pi_{\Pi,s_{m+1}} \widehat{\mathbf{X}}_{t_j,t_{j+1}}^{\widehat{\mathcal{D}}_j} \end{aligned}$$

where the last equality is implied by (2.16). Thus

$$\begin{aligned}\left\|\widehat{\mathbf{X}}_{s,t}^{\widehat{\mathcal{D}}}-\widehat{\mathbf{X}}_{s,t}^{\mathcal{D}}\right\| &\leq \sum_{\mathcal{D}}\omega(t_j,t_{j+1})^{s_{m+1}}\sum_{R\in\mathcal{A}^{\Pi}_{s_{m+1}}\backslash\mathcal{A}^{\Pi}_{s_m}}\frac{1}{\beta^k\Gamma_{\Pi}(R)}\\ &\leq \left(\max_{|u-v|<\delta}\omega(u,v)^{s_{m+1}-1}\right)\omega(s,t)\sum_{R\in\mathcal{A}^{\Pi}_{s_{m+1}}\backslash\mathcal{A}^{\Pi}_{s_m}}\frac{1}{\beta^k\Gamma_{\Pi}(R)}.\end{aligned}$$

This upper bound is independent of the refinement $\widehat{\mathcal{D}}$ and since $s_{m+1}>1$ and ω is uniformly continuous, the upper bound can be made arbitrarily small, uniformly over all partitions with mesh size smaller than δ. By the triangle inequality, we get a uniform bound on $\widehat{\mathbf{X}}^{\mathcal{D}}-\widehat{\mathbf{X}}^{\mathcal{D}'}$, and hence $\lim_{|\mathcal{D}|\to 0}\widehat{\mathbf{X}}^{\mathcal{D}}_{s,t}$ exists.

Step 3

Finally, we prove the multiplicative property of the limit $\mathbf{X}_{s,t}(s_{m+1})=\lim_{|\mathcal{D}|\to 0}\widehat{\mathbf{X}}^{\mathcal{D}}_{s,t}$. Let $u\in(s,t)$ and $\{\mathcal{D}_n\}_{n\geq 0}$ be a sequence of partitions such that $|\mathcal{D}_n|\to 0$ as n tends to infinity. Define $\widetilde{\mathcal{D}}_n=\mathcal{D}_n\cup\{u\}$. Note that $|\widetilde{\mathcal{D}}_n|\to 0$ as n tends to infinity. Then we have

$$\begin{aligned}\mathbf{X}^{(s_{m+1})}_{s,t} &= \lim_{|\mathcal{D}|\to 0}\widehat{\mathbf{X}}^{\mathcal{D}}_{s,t}=\lim_{n\to\infty}\widehat{\mathbf{X}}^{\widetilde{\mathcal{D}}_n}_{s,t}=\lim_{n\to\infty}\widehat{\mathbf{X}}^{\widetilde{\mathcal{D}}_n\cap[s,u]}_{s,t}\otimes\widehat{\mathbf{X}}^{\widetilde{\mathcal{D}}_n\cap[u,t]}_{s,t}\\ &= \left(\lim_{n\to\infty}\widehat{\mathbf{X}}^{\widetilde{\mathcal{D}}_n\cap[s,u]}_{s,t}\right)\otimes\left(\lim_{n\to\infty}\widehat{\mathbf{X}}^{\widetilde{\mathcal{D}}_n\cap[u,t]}_{s,t}\right)=\mathbf{X}^{(s_{m+1})}_{s,u}\otimes\mathbf{X}^{(s_{m+1})}_{u,t}.\end{aligned}$$

The proof of uniqueness is based on the fact that if $\mathbf{Y}$ and $\mathbf{Z}$ are two multiplicative functionals in $T^{(\Pi,s_{m+1})}(V)$ such that

$$\pi_{\Pi,s_m}(\mathbf{Y}_{s,t}-\mathbf{Z}_{s,t})=0,\ \forall s,t \text{ such that } 0\leq s\leq t\leq T$$

then the $\Delta_T\to T^{(\Pi,s_{m+1})}(V)$ functional $\Psi_{\cdot,\cdot}:=\mathbf{Y}_{\cdot,\cdot}-\mathbf{Z}_{\cdot,\cdot}$ is additive. If the above $\mathbf{Y}$ and $\mathbf{Z}$ were two choices for $\mathbf{X}^{(s_{m+1})}$, then

$$\|\Psi_{s,t}\|\leq C\omega(s,t)^{s_{m+1}}$$

for some constant C. Therefore $\Psi_{\cdot,\cdot}$ is a path in $T^{(\Pi,s_{m+1})}(V)\backslash T^{(\Pi,s_m)}(V)$ with finite $s_{m+1}>1$ variation starting at zero, implying that $\Psi_{\cdot,\cdot}$ is identically equal to zero. The assertion follows. □

Remark 2.1.3. Note that in the case when $\Pi=(p)$, $T^{(n)}(V)$ coincides with $T^{(\Pi,n/p)}$ and so Theorem 1.1.1 is a special case of Theorem 2.1.1.

Definition 2.1.5 (Π-rough paths)**.** *Let the Banach space V be of the form $V=V^1\oplus\cdots\oplus V^k$ for some Banach spaces $V^1,\ldots,V^k$. Let $\Pi=(p_1,\ldots,p_k)$ be a k-tuple as in Definition 2.1.1 and ω be a control. A* Π-rough path *in V is a continuous $\Delta_T\to T^{(\Pi,1)}(V)$ multiplicative functional $\mathbf{X}$ with finite Π-variation controlled by ω.*

The space of Π-rough paths is denoted by $\Omega_{\Pi}(V)$.

Remark 2.1.4. The extension theorem guarantees that a Π-rough path with Π-variation controlled by a control ω can be uniquely extended from a $\Delta_T \to T^{(\Pi,1)}(V)$ multiplicative functional to a $\Delta_T \to T^{(\lfloor p_{\max} \rfloor)}(V)$ multiplicative functional where $p_{\max} = \max_{1 \le i \le k} p_i$. Note that the extended functional is also a $p_{\max}$-rough path with $p_{\max}$-variation controlled by $C\omega$ where the constant C only depends on Π and $\omega(0,T)$. This implies that a Π-rough path also can be regarded as a $p_{\max}$-rough path. This fact also highlights why Theorem 2.1.1 is a stronger version of Theorem 1.1.1. I.e. in the inhomogeneous degree of smoothness case, fewer components determine uniquely the higher order components. For example, for $2 < p_1 < 3$ and $3 < p_2 < 4$, let a (p_1)-rough path and a (p_2)-rough path be given with (p_1) and (p_2) variation controlled by a control ω on Δ_T. In this case, Theorem 1.1.1 requires all cross terms corresponding to multi-indices up to length 3 for the existence of the coupled rough path. However, Theorem 2.1.1 requires only the terms corresponding to the following indices $(1,2)$, $(2,1)$, $(1,2,2)$, $(2,1,2)$ and $(2,2,1)$.

Remark 2.1.5. In the later chapters, we will work with rough paths extended to $T^{(\Pi,\alpha)}(V)$ for some $\alpha > 1$. Note that although the condition (2.7) on β might not hold in some of the practical cases, the Extension Theorem restricted to $T^{(\Pi,\alpha)}(V)$ is still valid with some modification, i.e. the extended rough-path exists and is unique, moreover its Π-variation is controlled by $C\omega$, where the constant C depends on α, β and Π.

Definition 2.1.6 (Π-variation metric)**.** *Let $C_{0,\Pi}\left(\Delta_T, T^{(\Pi,1)}(V)\right)$ denote the space of all continuous functions from the simplex Δ_T into the truncated tensor algebra $T^{(\Pi,1)}(V)$ with finite Π-variation. The Π-variation metric on this linear space if defined as follows*

$$d_\Pi(\mathbf{X},\mathbf{Y}) := \max_{R \in \mathcal{A}_1^\Pi} \sup_{\mathcal{D} \in \mathcal{P}([0,T])} \left(\sum_{\mathcal{D}} \left\| \mathbf{X}^R_{t_{l-1},t_l} - \mathbf{Y}^R_{t_{l-1},t_l} \right\|^{1/deg_\Pi(R)} \right)^{deg_\Pi(R)}$$

Let $\alpha > 0$. With analogy, we define the space $C_{0,\Pi}\left(\Delta_T, T^{(\Pi,\alpha)}(V)\right)$ and the metric on this space

$$d^\alpha_\Pi(\mathbf{X},\mathbf{Y}) := \max_{R \in \mathcal{A}_\alpha^\Pi} \sup_{\mathcal{D} \in \mathcal{P}([0,T])} \left(\sum_{\mathcal{D}} \left\| \mathbf{X}^R_{t_{l-1},t_l} - \mathbf{Y}^R_{t_{l-1},t_l} \right\|^{1/deg_\Pi(R)} \right)^{deg_\Pi(R)}$$

The proof of Lemma 3.3.3 in [26] can be adapted using the notation and terminology of this section to show that the space Ω_Π equipped with the Π-variation metric is a complete metric space.

With analogy to [26], we introduce the Π-variational topology.

Definition 2.1.7 (Π-variational topology)**.** *A sequence $\{\mathbf{X}(n)\}$ in $C_{0,\Pi}\left(\Delta_T, T^{(\Pi,1)}(V)\right)$ converges to $\mathbf{X} \in C_{0,\Pi}\left(\Delta_T, T^{(\Pi,1)}(V)\right)$ in the* Π-variational topology*, if there is a control ω such that for all $n \in \mathbb{N}$*

$$\left\| \mathbf{X}^R_{s,t}(n) \right\|, \left\| \mathbf{X}^R_{s,t} \right\| \le \omega(s,t)^{deg_\Pi(R)},\ \forall R \in \mathcal{A}_1^\Pi,\ \forall (s,t) \in \Delta_T \tag{2.17}$$

and there exists a function $a : \mathbb{N} \to \mathbb{R}$ *depending on the control* ω, $\mathbf{X}$ *and the sequence* $\{\mathbf{X}(n)\}$, *such that* $\lim_{n\to\infty} a(n) = 0$ *and*

$$\left\|\mathbf{X}^R_{s,t}(n) - \mathbf{X}^R_{s,t}\right\| \leq a(n)\omega(s,t)^{deg_\Pi(R)}, \forall R \in \mathcal{A}^\Pi_1, \forall (s,t) \in \Delta_T. \tag{2.18}$$

Remark 2.1.6. The convergence with respect to the metric d^α_Π for $\alpha \geq 1$ implies the convergence in the Π-variational topology. The converse statement is that any sequence convergent in the Π-variational topology has a subsequence convergent with respect to the metric d^1_Π. The proof of the latter statement is analogous to the proof Proposition 3.3.3 in [26].

By the extension theorem, a $T^{(\Pi,1)}$-valued rough path determines all the higher order terms. However, slight extension of the arguments gives extra information on the Π-variational topology.

Lemma 2.1.2. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two* Π*-rough paths and let* β *satisfy the inequality* (2.7)*. If* ω *is a control, such that*

$$\left\|\mathbf{X}^R_{s,t}\right\|, \left\|\mathbf{Y}^R_{s,t}\right\| \leq \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}, \ \forall R \in \mathcal{A}^\Pi_1, \forall (s,t) \in \Delta_T,$$

and

$$\left\|\mathbf{X}^R_{s,t} - \mathbf{Y}^R_{s,t}\right\| \leq \varepsilon \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}, \ \forall R \in \mathcal{A}^\Pi_1, \forall (s,t) \in \Delta_T, \tag{2.19}$$

then the inequality (2.19) *holds for all* $R \in \mathcal{A}^k$*, i.e. for all the extended coordinates of* $\mathbf{X}$ *and* $\mathbf{Y}$*.*

The proof of the lemma is analogous to the proof of Theorem 3.1.3 in [26]. Note that in the case when β does not satisfy the inequality (2.7), the Π-rough path still can be extended from $T^{(\Pi,1)}$ to $T^{(\Pi,\alpha)}$ with ω replaced by $C\omega$ where C depends on β and Π using the same arguments as in the Extension Theorem.

Remark 2.1.7. Lemma 2.1.2 implies that in the case of a suitable β, if we replace $\mathcal{A}^\Pi_1$ with any subset of $\mathcal{A}^k$ including $\mathcal{A}^\Pi_1$ in the inequalities (2.17) and (2.18), then we get an equivalent definition of the Π-variational topology. Furthermore, considering Remark 2.1.6, a sequence of Π-rough paths $\{\mathbf{X}(n)\}$ controlled by a control ω converging to $\mathbf{X}$ with respect to the metric d_Π also converges to $\mathbf{X}$ in the Π-variational topology. The first part of this remark implies that there exists a sub-sequence of $\mathbf{X}(n)$ converging to $\mathbf{X}$ with respect to the metric d^α_Π for any $\alpha \geq 1$ and also with respect to the metric $d_{p_{\max}}$ where $p_{\max} = \max_{p_i \in \Pi} p_i$.

The following subset of the space of Π-rough paths is crucial for our further analysis.

Definition 2.1.8 (Geometric Π-rough path)**.** *A* geometric Π-rough path *is a* Π*-rough path which can be expressed as a limit of* (1)*-rough paths (or* smooth *rough paths) in the* Π*-variation distance. The space of geometric* Π*-rough paths in* V *is denoted by* $G\Omega_\Pi(V)$*.*

Remark 2.1.8. In the remaining chapters, we will work with geometric Π-rough paths. A useful feature of this class of rough paths is that any property of smooth paths preserved by the Π-variation metric continuity is also a property of geometric rough paths.

One such important property of smooth rough paths is that

$$\mathrm{d}\mathbf{X}_{s,t}^{R} = \mathbf{X}_{s,t}^{R-} \otimes \mathrm{d}\mathbf{X}_{s,t}^{(r_{\|R\|})} \tag{2.20}$$

for all $R = (r_1, \ldots, r_{\|R\|}) \in \mathcal{A}^k$ and $(s,t) \in \Delta_T$. The signature of a smooth path, i.e. the multiplicative map constructed from the iterated integrals of the path (recall Example 1.1.1) clearly satisfy the ODE (2.20). The integration theory introduced in the next sections is constructed in a way which preserves this property for Π-rough paths.

2.2 Almost Π-rough paths

Definition 2.2.1 (Almost Π-rough path). *Let the Banach space V be of the form $V = V^1 \oplus \cdots \oplus V^k$ for some Banach spaces $V^1, \ldots, V^k$. Let $\Pi = (p_1, \ldots, p_k)$ denote a k-tuple as in Definition 2.1.1 and let ω be a control. The functional $\mathbf{X} : \Delta_T \to T^{(\Pi,1)}(V)$ is called an* almost Π-rough path *if*

(i) it has a finite Π-variation controlled by ω:

$$\left\|\mathbf{X}_{s,t}^{R}\right\| \leq \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}$$

for all $(s,t) \in \Delta_T$ and for all multi-index $R \in \mathcal{A}_1^\Pi$.

(ii) it is almost-multiplicative*, i.e. there exists $\theta > 1$ such that*

$$\|\pi_R(\mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} - \mathbf{X}_{s,t})\| \leq \omega(s,t)^\theta \ \forall s < u < t \in [0,T],\ \forall R \in \mathcal{A}_1^\Pi.$$

We will also say that $\mathbf{X}$ is a θ-almost Π-rough path controlled by ω.

Theorem 2.2.1. *Let the Banach space V be of the form $V = V^1 \oplus \cdots \oplus V^k$ for some Banach spaces $V^1, \ldots, V^k$. Let $\Pi = (p_1, \ldots, p_k)$ denote a k-tuple as in Definition 2.1.1 and let ω be a control. Let the functional $\mathbf{X} : \Delta_T \to T^{(\Pi,1)}(V)$ be a θ-almost Π-rough path controlled by ω.*

Then there exists a unique Π-rough path $\widehat{\mathbf{X}} : \Delta_T \to T^{(\Pi,1)}(V)$ such that

$$\sup_{\substack{0 \leq s < t \leq T \\ R \in \mathcal{A}_1^\Pi}} \frac{\left\|\pi_R\left(\widehat{\mathbf{X}}_{s,t} - \mathbf{X}_{s,t}\right)\right\|}{\omega(s,t)^\theta} < +\infty. \tag{2.21}$$

Moreover, there exists a constant K which depends only on Π, θ and $\omega(0,T)$ such that the supremum (2.21) *is smaller than K and the Π-variation of $\widehat{\mathbf{X}}$ is controlled by $K\omega$.*

Sketch of the proof. The proof is analogous to the proof of Theorem 1.2.1 as it is presented in [27]. The only difference to the referred proof is the replacement of the m^{th} *level truncation* by the truncation π_{Π,s_m}. In [27], induction arguments are used as follows. If $\mathbf{Y} : \Delta_T \to T^{(\Pi,1)}(V)$ is a θ-almost Π-rough path controlled by ω such that $\pi_{\Pi,s_m}\mathbf{Y}$ is multiplicative in $T^{(\Pi,s_m)}(V)$ for $s_m < 1$, where $\{s_m\}$ is the sequence introduced in the proof of Theorem 2.1.1, then there exists a θ-almost Π-rough path $\mathbf{Z} : \Delta_T \to T^{(\Pi,1)}(V)$ such that

(i) $\mathbf{Z}$ controlled in Π-variation by $K_{m+1}\omega$

(ii) $\|\pi_R(\mathbf{Y}_{s,t} - \mathbf{Z}_{s,t})\| \leq K_{m+1}\omega(s,t)$ for all $(s,t) \in \Delta_T$ and for all $R \in \mathcal{A}^{\Pi}_{s_{m+1}}$

(iii) and furthermore $\pi_{\Pi,s_{m+1}}\mathbf{Z}$ is multiplicative in $T^{(\Pi,s_{m+1})}(V)$.

The constants K_m for $m = 0, 1, \ldots$ depend only on m, θ, Π and $\omega(0,T)$.

Since the lowest layer $\pi_{\Pi,0}\mathbf{X}$ satisfies the above conditions in $T^{(\Pi,s_0)}(V)$, successive application of the induction step will result in a Π-rough path with the required properties.

2.3 Integration along Π-rough paths

We introduce some new notation.

Definition 2.3.1. *Let* $\mathbf{a}$ *be an element of* $T((V))$*. Then the canonical projection of* $\mathbf{a}$ *onto the tensor of level s is denoted by* $\mathbf{a}^{\langle s\rangle}$*, i.e.*

$$\begin{aligned} \mathbf{a}^{\langle s\rangle} &= \sum_{deg_\Pi(R)=s} \mathbf{a}^R, \\ \mathbf{a}^V &= \pi_V \mathbf{a}. \end{aligned}$$

Definition 2.3.2. *Let* $s_{m_1}, \ldots, s_{m_n}$ *be elements in* S^Π*. The projection*

$$\pi_{(s_{m_1},\ldots,s_{m_n})} : T((V)) \to \bigoplus_{\substack{R_1,\ldots,R_n \in \mathcal{A}^k \\ deg_\Pi(R_i-)=s_{m_i}, i=1,\ldots,n}} V^{\otimes R_1 * \cdot * R_n}$$

is defined as follows:

$$\pi_{(s_{m_1},\ldots,s_{m_n})}(\mathbf{a}) = \sum_{\substack{R_1,\ldots,R_n \in \mathcal{A}^k \\ deg_\Pi(R_i-)=s_{m_i}, i=1,\ldots,n}} \pi_{R_1*\cdots*R_n}\mathbf{a},\ \mathbf{a} \in T((V)).$$

Definition 2.3.3. *We will say that an element* $\mathbf{a} = \sum_{deg_\Pi(R)=s_m} \mathbf{a}^R \in V^{\otimes(\Pi,s_m)}$ *is an* s_m-symmetric element *(*$s_m \in S^\Pi$*) if each term of the form*

$$\sum_{\substack{deg_\Pi(R)=s_m \\ \|R\|=i}} \pi_R \mathbf{a} \tag{2.22}$$

is a projection of a symmetric element in $V^{\otimes i}$.

Let W *be a Banach space. A map* $f \in L\big(V^{\otimes(\Pi,s_m)}, W\big)$ *is an* s_m-symmetric linear map, *if there exists* $V^{\otimes i}$*-symmetric* $L\big(V^{\otimes i}, W\big)$ *maps* f_i*, such that*

$$f(\mathbf{a}) = \sum_{\substack{deg_{\Pi}(R)=s_m \\ \|R\|=i}} f_i \circ \pi_{V^{\otimes i}}(\mathbf{a})$$

for all $\mathbf{a} \in V^{\otimes(\Pi,s_m)}$.

Definition 2.3.4 ((Π,Γ)-Lipschitz function)**.** *Let* U*,* V *and* W *be Banach spaces, such that* $U = U^1 \oplus \cdots \oplus U^N$ *for some Banach spaces* $U^1, \dots, U^N$ *and* $V = V^1 \oplus \cdots \oplus V^k$ *for some Banach spaces* $V^1, \dots, V^k$*. Let* $\Pi_U = (q_1, \dots, q_N)$ *and* $\Gamma = (\gamma_1, \dots, \gamma_k)$ *denote a real* N*-tuple and* k*-tuple as in Definition 2.1.1. Let* $q_{\max} = \max_{1 \le i \le N} q_i$*. Let* F *be a closed subset in* U*.*

The function $\alpha : F \to L(V, W)$ *is a* (Π_U,Γ)-Lipschitz one-form *on* F*, if for all* $s_m < \gamma_i$*,* $s_m \in S^{\Pi_U}$ *there exist functions* $\alpha_i^{s_m} : F \to L(V^i, W)$ *for* $i = 1, \dots, k$ *such that*

(i) $\alpha(u) = \sum_{i=1}^k \alpha_i^{s_0}(u) \circ \pi_{V^i}$ *for all* $u \in F$.

(ii) $\alpha_i^{s_m} : F \to L\left(U^{\otimes(\Pi_U,s_m)}, L(V^i, W)\right)$ *for* $i = 1, \dots, k$*, taking values in the space of* s_m*-symmetric linear maps, satisfies*

$$\alpha_i^{s_m}(y)(u) = \sum_{s_m \le s_n < \gamma_i} \alpha_i^{s_n}(x) \left(u \otimes \sum_{deg_{\Pi_U}(R)=s_n-s_m} \frac{(y-x)_R}{\|R\|!} \right) + R_i^{s_m}(x,y)(u)$$

for all $x, y \in F$ *and* $u \in U^{\otimes(\Pi_U,s_m)}$*, where* $R_i^{s_m} : F \times F \to L\left(U^{\otimes(\Pi_U,s_m)}, L(V^i, W)\right)$ *and*

$$\left\| R_i^{s_m}(x,y) \right\| < M \left\| x-y \right\|^{(\gamma_i - s_m)q_{\max}}.$$

The least non-negative constant M *satisfying the above properties is denoted by* $\|\alpha\|_{Lip(\Pi_U,\Gamma)}$ *and is referred to as the* (Π_U,Γ)*-Lipschitz norm of* α *on* F*.*

In addition to the above definition and for practical reasons we introduce the function $\alpha^{s_m} : F \to L\left(U^{\otimes(\Pi_U,s_m)}, L(V,W)\right)$ for $s_m < \max_{1 \le i \le k} \gamma_i = \gamma_{\max}$ defined by

$$\alpha^{s_m}(v)(\mathbf{u}) = \sum_{i | s_m < \gamma_i} \alpha_i^{s_m}(v)(\mathbf{u}) \circ \pi_{V^i}, \ \forall v \in F, \forall \mathbf{u} \in U^{\otimes(\Pi_U,s_m)}. \tag{2.23}$$

Note that α^{s_m} takes s_m-symmetric linear maps as values. Furthermore we introduce the functions $R_{s_m} : F \times F \to L\left(U^{\otimes(\Pi_U,s_m)}, L(V,W)\right)$ defined by

$$R_{s_m}(x,y)(\mathbf{u}) = \sum_{i | s_m < \gamma_i} R_i^{s_m}(x,y)(\mathbf{u}) \circ \pi_{V^i}, \ \forall x, y \in F, \forall \mathbf{u} \in U^{\otimes(\Pi_U,s_m)}. \tag{2.24}$$

Example 2.3.1. Let $U = V = W = \mathbb{R}$, furthermore let n be positive integer and q a positive real. Let us define $\alpha : [0,1] \to L(V,W)$ by $\alpha(x) = x^{n+1/2}$. Recalling Example 1.2.1, α is a *Lip*$(n+3/2)$ function and by Definition 2.3.4 also a *Lip*(Π,Γ) function for $\Pi = (q)$ and $\Gamma = (\gamma)$, where $\gamma q = n+3/2$. In general, for positive γ and q, a *Lip*(γq) function in the sense of Definition 1.2.2 is also a *Lip*(Π,Γ) function in the sense of Definition 2.3.4, for $\Pi = (q)$ and $\Gamma = (\gamma)$.

Example 2.3.2. Now, let $U = W = \mathbb{R}$ and $V = \mathbb{R} \oplus \mathbb{R}$. Let n_1 and n_2 be positive integers. Define $\alpha : [0,1] \to L(V,W)$ as $\alpha(x) = (\alpha_1(x), \alpha_2(x))$ where $\alpha_i(x) := x^{n_i+1/2}$ for $i = 1,2$. Then α is a *Lip*(Π,Γ) for $\Pi = (q)$ and $\Gamma = \left(\frac{n_1+3/2}{q}, \frac{n_2+3/2}{q}\right)$. In the case when $V = V^1 \oplus \cdots V^k$, $\Pi = (q)$ and $\Gamma = (\gamma_1, \ldots, \gamma_k)$, a *Lip*$(\Pi,\Gamma)$ funtion α is in general of the form:

$$\alpha(x) = \big(\alpha_1(x), \ldots, \alpha_k(x)\big) = \sum_{i=1}^{k} \alpha_i(x) \circ \pi_{V^i}$$

where each α_i is a *Lip*$(\gamma_i q)$ function from F to $L(V^i, W)$ in the sense of Definition 1.2.2.

Remark 2.3.1. Note that the s_m-symmetric part of $\sum_{deg_{\Pi_V}(R)=s_m} \mathbf{X}^R_{s,t}$ is $\sum_{deg_{\Pi_V}(R)=s_m} \frac{(\mathbf{X}^V_{s,t})_R}{\|R\|!}$.

Theorem 2.3.1 (Integration of a One-form)**.** *Let V and W be Banach spaces, such that $V = V^1 \oplus \cdots \oplus V^k$ for some Banach spaces $V^1, \ldots, V^k$. Let $\Pi = (p_1, \ldots, p_k)$ denote a k-tuple as in Definition 2.1.1 with $p_{\max} = \max_{1 \le i \le k} p_i$ and let ω be a control. Let $\mathbf{Z} : \Delta_T \to T^{(\Pi,1)}(V)$ be a geometric Π-rough path controlled by ω. Let $\Gamma = (\gamma_1, \ldots, \gamma_k)$ be a real k-tuple such that $\gamma_i > 1 - 1/p_i$ for $i = 1, \ldots, k$. And finally let $\alpha : V \to L(V,W)$ be a Lip(Π,Γ) function as in Definition 2.3.4.*

Then $\mathbf{Y} : \Delta_T \to T^{((p_{\max}),1)}(W)$ defined for all $(s,t) \in \Delta_T$ by

$$\mathbf{Y}^n_{s,t} = \pi_{W^{\otimes n}} \mathbf{Y}_{s,t} = \sum_{s_{m_1}+\cdots+s_{m_n} < \gamma_{\max}} \alpha^{s_{m_1}} \left(\mathbf{Z}^V_{0,s}\right) \otimes \cdots \otimes \alpha^{s_{m_n}} \left(\mathbf{Z}^V_{0,s}\right) \pi_{(s_{m_1},\ldots,s_{m_n})} \sum_{\substack{R_1,\ldots,R_n \in \mathcal{A}^k \\ \sigma \in OS(\|R_1\|,\ldots,\|R_n\|)}} \sigma^{-1} \mathbf{Z}^{R_1 * \cdots * R_n}_{s,t} \tag{2.25}$$

is an almost $(p_{\max})$-rough path.

Remark 2.3.2. Note that the second sum in (2.25) is in fact finite due to the $\pi_{(s_{m_1},\ldots,s_{m_n})}$ projection.

Remark 2.3.3. By Example 2.3.1, in the case when $k = 1$, Theorem 2.3.1 defines the same almost rough path as Theorem 1.2.2. Also note that in the general $k > 1$ case, the conditions of the above theorem are weaker than the conditions we would get by regarding the driving rough path as a homogeneous $(p_{\max})$-rough path and defining the integral by Theorem 1.2.2.

We prove Theorem 2.3.1 using the following lemmas.

Lemma 2.3.1. *If $\mathbf{Z} \in \Omega_\Pi$ is a smooth rough-path, then for any $s < u < t$ in $[0, T]$,*

$$\sum_{s_m<\gamma} \alpha^{s_m}\left(\mathbf{Z}^V_{0,s}\right)\left(\mathbf{Z}^{\langle s_m\rangle}_{s,t}\right)\left(\mathrm{d}\mathbf{Z}^V_{0,t}\right) = \sum_{s_m<\gamma}\left(\alpha^{s_m}\left(\mathbf{Z}^V_{0,u}\right) - R_{s_m}\left(\mathbf{Z}^V_{0,s}, \mathbf{Z}^V_{0,u}\right)\right)\left(\mathbf{Z}^{\langle s_m\rangle}_{u,t}\right)\left(\mathrm{d}\mathbf{Z}^V_{0,t}\right) \tag{2.26}$$

The proof of the lemma is analogous to the proof of Lemma 5.5.2 in [26].

Lemma 2.3.2. *Let $\mathbf{Z} \in \Omega_\Pi$ be a smooth rough-path and let the function $\widehat{\mathbf{Y}}^1 : \Delta_T \to W$ be defined for all $(s, t) \in \Delta_T$ by*

$$\widehat{\mathbf{Y}}^1_{s,t} = \sum_{s_m<\gamma_{\max}} \alpha^{s_m}\left(\mathbf{Z}^V_{0,s}\right) \sum_{\substack{R\in\mathcal{A}^k \\ deg_\Pi(R-)=s_m}} \mathbf{Z}^R_{s,t}. \tag{2.27}$$

Then

$$\begin{aligned}
\widehat{\mathbf{Y}}^n_{s,t} &:= \int_{s<u_1<\cdots<u_n<t} \mathrm{d}\widehat{\mathbf{Y}}^1_{s,u_1} \otimes \cdots \otimes \mathrm{d}\widehat{\mathbf{Y}}^1_{s,u_n} \\
&= \int_{s<u_1<\cdots<u_n<t} \sum_{s_{m_1}<\gamma_{\max}} \alpha^{s_{m_1}}\left(\mathbf{Z}^V_{0,s}\right) \sum_{\substack{R\in\mathcal{A}^k \\ deg_\Pi(R-)=s_{m_1}}} \mathbf{Z}^R_{s,u_1}\mathrm{d}\mathbf{Z}^V_{s,u_1} \otimes \\
&\qquad \cdots \otimes \sum_{s_{m_n}<\gamma_{\max}} \alpha^{s_{m_n}}\left(\mathbf{Z}^V_{0,s}\right) \sum_{\substack{R\in\mathcal{A}^k \\ deg_\Pi(R-)=s_{m_n}}} \mathbf{Z}^R_{s,u_n}\mathrm{d}\mathbf{Z}^V_{s,u_n} \qquad (2.28) \\
&= \sum_{s_{m_1},\cdots,s_{m_n}<\gamma_{\max}} \alpha^{s_{m_1}}\left(\mathbf{Z}^V_{0,s}\right) \otimes \cdots \otimes \alpha^{s_{m_n}}\left(\mathbf{Z}^V_{0,s}\right) \pi_{(s_{m_1},\ldots,s_{m_n})} \sum_{\substack{R_1,\ldots,R_n\in\mathcal{A}^k \\ \sigma\in OS(\|R_1\|,\ldots,\|R_n\|)}} \sigma^{-1}\mathbf{Z}^{R_1*\cdots*R_n}_{s,t} \qquad (2.29)
\end{aligned}$$

for $n = 2, \ldots, \lfloor p_{\max} \rfloor$

The proof of the lemma is analogous to what is described in Section 4.2. of [27].

Remark 2.3.4. The right-hand side of (2.27) is regarded as the term in the integral approximating sum of $\int_0^T \alpha\left(\mathbf{Z}^V_{0,s}\right) \mathrm{d}\mathbf{Z}_{0,s}$ corresponding to the subinterval $[s, t]$. When $\mathbf{Z}$ is a smooth path, the integral defined using the first term of the right-hand side of (2.27) is equal to the integral defined by the entire right-hand side. Hence for smooth rough paths, the rough path integral is consistent with the Riemann integral.

Lemma 2.3.3. *Let $\widehat{\mathbf{Y}} : \Delta_T \to T^{((p_{\max}),1)}(W)$ be defined as in (2.29). Then for all $s < u < t$ in $[0, T]$,*

$$\widehat{\mathbf{Y}}_{s,u} \otimes \widehat{\mathbf{Y}}_{u,t} - \widehat{\mathbf{Y}}_{s,t} = \widehat{\mathbf{Y}}_{s,u} \otimes \mathbf{N}_{s,u,t} \tag{2.30}$$

where

$$\mathbf{N}^i_{s,u,t} = \pi_{W^{\otimes i}}\mathbf{N}_{s,u,t} = \sum_{\substack{s_{m_1},\dots,s_{m_i}<\gamma_{\max}\\ \varepsilon_1,\dots,\varepsilon_i\in\{0,1\}\\ \varepsilon_1\cdots\varepsilon_i=0}} \beta^{\varepsilon_1}_{s_{m_1}}\left(\mathbf{Z}^V_{0,s},\mathbf{Z}^V_{0,u}\right)\cdots\beta^{\varepsilon_i}_{s_{m_i}}\left(\mathbf{Z}^V_{0,s},\mathbf{Z}^V_{0,u}\right)\pi_{(s_{m_1},\dots,s_{m_i})}\sum_{\substack{R_1,\dots,R_i\in\mathcal{A}^k\\ \sigma\in OS(\|R_1\|,\dots,\|R_i\|)}}\sigma^{-1}\mathbf{Z}^{R_1*\cdots*R_i}_{s,t} \tag{2.31}$$

with

$$\beta^{\varepsilon}_{s_m}\left(\mathbf{Z}^V_{0,s},\mathbf{Z}^V_{0,u}\right) = \begin{cases} R_{s_m}\left(\mathbf{Z}^V_{0,s},\mathbf{Z}^V_{0,u}\right) & \text{if } \varepsilon=0,\\ -\alpha^{s_m}\left(\mathbf{Z}^V_{0,s}\right) & \text{if } \varepsilon=1.\end{cases}$$

The proof is based on Lemmas 2.3.1 and 2.3.2, and is analogous to the proof of Lemma 5.5.3 in [26].

Remark 2.3.5. The Lemmas 2.3.1, 2.3.2 and 2.3.3 are stated for a smooth rough path $\mathbf{Z}$. However for each of the equalities (2.26), (2.29) and (2.30), both the right-hand side and the left-hand side are continuous in the Π-variation norm. This fact extends the lemmas for geometric Π-rough paths and this is the key to the next proof.

We now prove Theorem 2.3.1.

Proof. (Theorem 2.3.1) First we prove that $\widehat{\mathbf{Y}} : \Delta_T \to T^{((p_{\max}),1}(W)$, defined by

$$\begin{aligned}\widehat{\mathbf{Y}}^n_{s,t} &= \sum_{s_{m_1},\cdots,s_{m_n}<\gamma_{\max}} \alpha^{s_{m_1}}\left(\mathbf{Z}^V_{0,s}\right)\otimes\cdots\otimes\alpha^{s_{m_n}}\left(\mathbf{Z}^V_{0,s}\right)\\ &\quad\times\pi_{(s_{m_1},\dots,s_{m_n})}\sum_{\substack{R_1,\dots,R_n\in\mathcal{A}^k\\ \sigma\in OS(\|R_1\|,\dots,\|R_n\|)}}\sigma^{-1}\mathbf{Z}^{R_1*\cdots*R_n}_{s,t}\end{aligned}$$

is an almost $(p_{\max})$-rough path. Each term in the above sum is of the form

$$\alpha^{s_{m_1}}\left(\mathbf{Z}^V_{0,s}\right)\otimes\cdots\otimes\alpha^{s_{m_n}}\left(\mathbf{Z}^V_{0,s}\right)\mathbf{Z}^{R_1*\cdots*R_n}_{s,t} \tag{2.32}$$

where $deg_{\Pi}(R_i-) = s_{m_i}$. Since such a term is bounded by $C_0\|\alpha\|^n_{\mathrm{Lip}(\Pi,\Gamma)}\omega(s,t)^{n/p_{\max}}$ where C_0 only depends on Γ, Π and $\omega(0,T)$, this implies that condition *i*) of Definition 2.2.1 is satisfied.

We prove condition *ii*) by giving a bound on the norm of

$$(\widehat{\mathbf{Y}}_{s,u}\otimes\widehat{\mathbf{Y}}_{u,t})^n - \widehat{\mathbf{Y}}^n_{s,t} = \sum_{i=0}^n \widehat{\mathbf{Y}}^i_{s,u}\otimes\mathbf{N}^{n-i}_{s,u,t}.$$

In the (2.31) (Lemma 2.3.3) representation of $\mathbf{N}^{n-i}_{s,u,t}$, there is at least one factor of the form $R_{s_m}\left(\mathbf{Z}^V_{0,s},\mathbf{Z}^V_{0,u}\right)$. Considering the representation (2.24) of R_{s_m}, the following bound is implied:

$$M\sum_{i=1}^k\left\|\mathbf{Z}^V_{s,u}\right\|^{(\gamma_i-s_m)p_{\max}} \le M\sum_{i=1}^k\omega(s,t)^{\gamma_i-s_m}.$$

Moreover considering that $R_i^{s_m}(x,y)(\mathbf{u})$ only acts on elements of V^i, there exists a constant C_1 depending only on $\|\alpha\|_{\mathrm{Lip}(\Pi,\Gamma)}$, Γ, Π and $\omega(0,T)$ such that

$$\|(\widehat{\mathbf{Y}}_{s,u} \otimes \widehat{\mathbf{Y}}_{u,t})^n - \widehat{\mathbf{Y}}^n_{s,t}\| \leq C_1 \sum_{i=1}^{k} \omega(s,t)^{\gamma_i + (1/p_i)}.$$

By the choice of Γ, $\theta := \min_{1\leq i\leq k}(\gamma_i + (1/p_i)) \geq 1$, which implies that there exists a constant C depending only on $\|\alpha\|_{\mathrm{Lip}(\Pi,\Gamma)}$, Γ, Π and $\omega(0,T)$ such that

$$\|(\widehat{\mathbf{Y}}_{s,u} \otimes \widehat{\mathbf{Y}}_{u,t})^n - \widehat{\mathbf{Y}}^n_{s,t}\| \leq C\omega(s,t)^{\theta}$$

and hence $\mathbf{Y}$ is a θ-almost $(p_{\max})$-rough path.

Arguments analogous to Proposition 4.10 in [27] prove that $\mathbf{Y}$ is also a θ-almost $(p_{\max})$-rough path and furthermore that the $(p_{\max})$-rough associated to $\mathbf{Y}$ by Theorem 1.2.1 coincides with the $(p_{\max})$-rough path associated to $\widehat{\mathbf{Y}}$. □

Definition 2.3.5 (Integration of a one-form)**.** *Let V and W be Banach spaces, such that $V = V^1 \oplus \cdots \oplus V^k$ for some Banach spaces $V^1, \ldots, V^k$. Let $\Pi = (p_1, \ldots, p_k)$ denote a k-tuple as in Definition 2.1.1 with $p_{\max} = \max_{1\leq i\leq k} p_i$ and let ω be a control. Let $\mathbf{Z} : \Delta_T \to T^{(\Pi,1)}(V)$ be a geometric Π-rough path controlled by ω. Let $\Gamma = (\gamma_1, \ldots, \gamma_k)$ be a real k-tuple such that $\gamma_i > 1 - 1/p_i$ for $i = 1, \ldots, k$. And finally let $\alpha : V \to L(V,W)$ be a $Lip(\Pi,\Gamma)$ function.*

Let $\mathbf{Y} : \Delta_T \to T^{((p_{\max}),1)}(W)$ be the almost $(p_{\max})$-rough path defined by Theorem 2.3.1. The unique $(p_{\max})$-rough path associated to $\mathbf{Y}$ by Theorem 1.2.1 is called the integral of α along $\mathbf{Z}$ and it is denoted by

$$\int_{\cdot}^{\cdot} \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z} : \Delta_T \to T^{((p_{\max}),1)}(W).$$

Theorem 2.3.2. *Under the conditions of Definition 2.3.5, there exists a constant K depending only on Γ, Π and $\omega(0,T)$, such that*

$$\left\|\pi_{W^{\otimes i}} \int_s^t \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z}\right\| \leq K\|\alpha\|^i_{Lip(\Pi,\Gamma)}\omega(s,t)^{\frac{i}{p_{\max}}}.$$

The proof is analogous to the proof of Theorem 4.12. of [27].

2.4 Integral of inhomogeneous degree of smoothness

Without loss of generality, we can assume, that $p_1 \leq \cdots \leq p_k$ are ordered in the k-tuple Π. In this setting, the integral of a Π-rough path introduced in the previous section resulted in a (p_k)-rough path. The reason is that in the general case the roughest component had effect on the full W. In this section we present an example in which the integral itself is of inhomogeneous degree of smoothness.

Proposition 2.4.1. *Let V and W be Banach spaces, such that $V = V^1 \oplus \cdots \oplus V^k$ and $W = W^1 \oplus \cdots \oplus W^N$ for some Banach spaces $V^1, \ldots, V^k$ and $W^1, \ldots, W^N$. Let $\Pi_V = (p_1, \ldots, p_k)$ denote a k-tuple and $\Pi_W = (q_1, \ldots, q_N)$ denote an N-tuple both as in Definition 2.1.1 such that $p_1 \leq \cdots \leq p_k$ and $q_i = p_{j_i}$, $j_i \in \{1, \ldots, k\}$, $i = 1, \ldots, N$. Let ω be a control. Let $\mathbf{Z} : \Delta_T \to T^{(\Pi_V,1)}(V)$ be a geometric Π_V-rough path controlled by ω.*

Let $\Gamma = (\gamma_1, \ldots, \gamma_k)$ be a real k-tuple such that $\gamma_i > 1 - 1/p_i$ for $i = 1, \ldots, k$. And finally let $\alpha : V \to L(V, W)$ be a $Lip(\Pi, \Gamma)$ function of the form

$$\alpha(x)(y) = \begin{pmatrix} \sum_{j=1}^{j_1} f_{1,j}(x)(\pi_{V^j} y) \\ \vdots \\ \sum_{j=1}^{j_i} f_{i,j}(x)(\pi_{V^j} y) \\ \vdots \\ \sum_{j=1}^{j_N} f_{N,j}(x)(\pi_{V^j} y) \end{pmatrix}$$

for all $x, y \in V$, where $f_{i,j} : V \to L(V^j, W^i)$ for $1 \leq j \leq k$, $1 \leq i \leq N$.

Then the integral $\int_\cdot^\cdot \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z}$ is a Π_W-rough path controlled by $K\omega$, where K depends on $\|\alpha\|_{Lip(\Pi_V,\Gamma)}$, Γ, Π_V and $\omega(0,T)$.

Proof. Firstly, we prove that the $\pi_{\Pi_W,1}$ projection of the almost $(p_{\max})$-rough path $\mathbf{Y}_{s,t}$ given by (2.25) is in fact an almost Π-rough path. Theorem 2.3.1 implies that condition ii) of Definition 2.2.1 is satisfied. For condition i), we derive which terms in (2.25) contribute to the $\mathbf{Y}_{s,t}^R$ for all $R \in \mathcal{A}_1^\Pi$.

A general (non-zero) term in the sum on the right-hand side of (2.25) is of the form

$$f_{i_1,j_1}^{s_{m_1}}\left(\mathbf{Z}_{0,s}^V\right) \otimes \cdots \otimes f_{i_n,j_n}^{s_{m_n}}\left(\mathbf{Z}_{0,s}^V\right) \mathbf{Z}_{s,t}^{R_1*(j_1)*\cdots R_n*(j_n)} \tag{2.33}$$

where $deg_{\Pi_V}(R_1 * \cdots * R_n) = s_{m_1} + \cdots + s_{m_n}$. This term contributes to $\mathbf{Y}_{s,t}^{(i_1,\ldots,i_n)}$. Note that since $\mathbf{Z}$ is a Π_V-rough path and $q_{i_l} \geq p_{j_l}$ for $l = 1, \ldots, n$ implying

$$deg_{\Pi_W}(i_1, \cdots, i_n) \leq deg_{\Pi_V}(j_1, \cdots, j_n),$$

$\mathbf{Y}_{s,t}^{(i_1,\ldots,i_n)}$ is bounded by

$$C \max\left\{\|\alpha\|_{\mathrm{Lip}(\Pi_V,\Gamma)}, 1\right\}^n \frac{\omega(s,t)^{deg_{\Pi_W}(i_1,\cdots,i_n)}}{\beta^k \Gamma_\Pi\Big((i_1, \ldots, i_n)\Big)},$$

where the constant C only depends on Γ, Π_V and $\omega(0,T)$. This implies condition (i).

Since $\pi_{\Pi_W,1}\mathbf{Y}$ is an almost Π_W-rough path, by Theorem 2.2.1 there exists a unique Π_W-rough path $\mathbf{X}$ associated to $\mathbf{Y}$. The proof of this theorem implies that $\mathbf{X}$ coincides with the $\pi_{\Pi_W,1}$-truncated $\int \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z}$. The Extension Theorem 2.1.1 guarantees that the unique extension of $\mathbf{X}$ coincides with the extension of $\int \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z}$.

Considering the norm of the terms (2.33) and using Theorem 2.2.1, one can derive the control on the integral. The technical steps are analogous to the steps of the proof of Theorem 4.12. of [27]. □

Corollary 2.4.1. *Under the conditions of Proposition 2.4.1, there exists a constant K depending on* Γ, Π_V *and* $\omega(0,T)$, *such that the* Π_W*-variation of the integral* $\int \alpha(\mathbf{Z})\mathrm{d}\mathbf{Z}$ *is controlled by*

$$K \max\left\{\|\alpha\|_{Lip(\Pi_V,\gamma)}, 1\right\}^{p_{\max}} \omega.$$

2.5 Differential equations driven by Π-rough paths

We start the section by defining differential equations driven by Π-rough paths (RDEs).

Definition 2.5.1 (Differential equations driven by Π-rough paths)**.** *Let V and W be Banach spaces, such that* $V = V^1 \oplus \cdots \oplus V^k$ *for some Banach spaces* $V^1, \ldots, V^k$ *and* $W = W^1 \oplus \cdots \oplus W^N$ *for some Banach spaces* $W^1, \ldots, W^N$*. Let* $\Pi_V = (p_1, \ldots, p_k)$ *denote a k-tuple and* $\Pi_W = (q_1, \ldots, q_N)$ *denote a N-tuple both as in Definition 2.1.1. We also introduce the* $(k+N)$*-tuple* $\widehat{\Pi} = \Pi_V * \Pi_W$*. Let* $f : W \to L(V, W)$ *be a function. Finally let* $\mathbf{X} \in G\Omega_{\Pi_V}(V)$ *be a geometric* Π*-rough path and* ξ *an element in W.*

We will say that $\mathbf{Z} \in G\Omega_{\widehat{\Pi}}(V \oplus W)$ *is a solution of the differential equation*

$$\mathrm{d}\mathbf{Y}_t = f(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_0 = \xi$$

if $\pi_{T^{(\Pi_V,1)}(V)}(\mathbf{Z}) = \mathbf{X}$ *and*

$$\mathbf{Z} = \int h(\mathbf{Z})\mathrm{d}\mathbf{Z} \tag{2.34}$$

where $h : V \oplus W \to End(V \oplus W)$ *is defined by*

$$h(x,y) = \begin{pmatrix} Id_V & 0 \\ f(y+\xi) & 0 \end{pmatrix}$$

provided the integral (2.34) *is well defined.*

In the following theorem we give sufficient conditions for the existence and uniqueness of solutions to RDEs driven by Π-rough paths.

Theorem 2.5.1 (Universal Limit Theorem, inhomogeneous case)**.** *Let V and W be Banach spaces, such that* $V = V^1 \oplus \cdots \oplus V^k$ *for some Banach spaces* $V^1, \ldots, V^k$ *and* $W = W^1 \oplus \cdots \oplus W^N$ *for some Banach spaces* $W^1, \ldots, W^N$*. Let* $\Pi_V = (p_1, \ldots, p_k)$ *denote a k-tuple and* $\Pi_W = (q_1, \ldots, q_N)$ *denote a N-tuple both as in Definition 2.1.1 and satisfying the conditions of Proposition 2.4.1. We also introduce the* $(k+N)$*-tuple* $\widehat{\Pi} = \Pi_V * \Pi_W$*.*

Let the function $f : W \to L(V, W)$ *possess the following properties:*

i) there exist real k-tuples $\Gamma = (\gamma_1, \ldots, \gamma_k)$ *and* $\hat{\Gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_k)$ *such that* $\gamma_i \geq \hat{\gamma}_i > 1 - 1/p_i$ *for* $i = 1, \ldots, k$,

ii) f *is a* (Π_W, Γ)*-Lipschitz function,*

iii) there exists a $(\Pi_W * \Pi_W, \hat{\Gamma})$*-Lipschitz function* $g : W \times W \to L(W, L(V, W))$ *such that*

$$f(x) - f(y) = g(x, y)(x - y).$$

Then for all $\mathbf{X} \in G\Omega_{\Pi_V}(V)$ *and* $\xi \in W$*, the equation*

$$\mathrm{d}\mathbf{Y}_t = f(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_0 = \xi \tag{2.35}$$

admits a unique solution $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) \in G\Omega_{\hat{\Pi}}(V \oplus W)$ *in the sense of Definition 2.5.1. Furthermore, this solution depends continuously on* $\mathbf{X}$ *and* ξ *in the* $\Pi - \hat{\Pi}$*-variation topology.*

Example 2.5.1. Recall the general statement of Example 2.3.2. Let $V = \mathbb{R} \oplus \mathbb{R}$, $\mathbf{X} = (\mathbf{X}(1), \mathbf{X}(2))$ be a Π-rough path in $T^{(\Pi,1)}$ for $\Pi = (p_1, p_2)$ and let $f = (f_1, f_2)$ be an $\mathbb{R} \to L(V, \mathbb{R})$ function. Then by the inhomogeneous smoothness version of the Universal Limit Theorem, the RDE

$$\mathrm{d}\mathbf{Y}_t = f_1(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t(1) + f_2(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t(2), \mathbf{Y}_0 = \xi$$

has a unique solution if for $i = 1, 2$, the function f_i is *Lip*$(\gamma_i p_{\max})$ for some $\gamma_i \geq 1 - 1/p_i$ and the function

$$g_i(x, y) = \frac{f_i(x) - f_i(y)}{x - y},$$

is *Lip*$(\hat{\gamma}_i p_{\max})$ for $\hat{\gamma}_i \geq 1 - 1/p_i$. I.e. if $p_1 < p_2$, then the component f_1 (and g_1) of the one form f corresponding to the less rough component $\mathbf{X}(1)$ of the driving noise $\mathbf{X}$ can be less regular then the component f_2 (and g_2) corresponding to the rougher term $\mathbf{X}(2)$. This demonstrates in what sense Theorem 2.5.1 is stronger than Theorem 1.3.1.

2.5.1 Proof of the Universal Limit Theorem

We prove the theorem in several steps. We follow the proof of the Universal Limit Theorem in [27]. Technically each step is analogous to the corresponding original step, however the integral is defined according to Definition 2.3.5.

We start with stating and adapting some lemmas of the original proof.

Lemma 2.5.1. *Let the Banach space* V *be of the form* $V = V^1 \oplus \cdots \oplus V^k$ *for some Banach spaces* $V^1, \ldots, V^k$*. Let* $\Pi = (p_1, \ldots, p_k)$ *denote a k-tuple as in Definition 2.1.1,* $\varepsilon > 0$*, and let* ω *be a control.*

Consider $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) \in G\Omega_{\Pi*\Pi}(V \oplus V)$ *and* $\mathbf{W} = \left(\mathbf{X}, \frac{\mathbf{Y}-\mathbf{X}}{\varepsilon}\right) \in G\Omega_{\Pi*\Pi}(V \oplus V)$. *Assume that the* $\Pi * \Pi$*-variation of* $\mathbf{W}$ *is controlled by* $\omega : \Delta_T \to [0, \infty)$. *Then there exists a constant* C *depending only on* Π, $\omega(0,T)$ *and* β, *such that*

$$\left\|\mathbf{X}^R_{s,t} - \mathbf{Y}^R_{s,t}\right\| \leq C(\varepsilon + \varepsilon^{\|R\|})\omega(s,t)^{\|R\|/p_{max}}, \ \forall (s,t) \in \Delta_T, \ \forall R \in \mathcal{A}^{\Pi}_1.$$

Proof. The claim is equivalent to Lemma 5.6 of [27] adapted to the inhomogeneous smoothness case and the proof is analogous to the proof of the referred lemma.

Let $R = (r_1, \ldots, r_l) \in \mathcal{A}^{\Pi}_1$ and $(s,t) \in \Delta_T$. First, assuming that $\mathbf{W}$ has bounded variation and writing $\mathbf{Y}^V = \mathbf{X}^V + \varepsilon \frac{\mathbf{Y}^V - \mathbf{X}^V}{\varepsilon}$, we get

$$\mathbf{Y}^R_{s,t} = \mathbf{X}^R_{s,t} + \sum_{\substack{k_1,\ldots,k_l \in \{0,1\} \\ k_1+\cdots+k_l>0}} \varepsilon^{k_1+\cdots+k_l} \mathbf{W}^{(r_1+k_1*l,\ldots,r_l+k_l*l)}_{s,t}.$$

The assertion is implied by the continuity in the $\Pi * \Pi$-variation topology and by the control on $\mathbf{W}$. □

Lemma 2.5.2 (Scaling Lemma, inhomogeneous version). *Let the Banach space* V *be of the form* $V = V^1 \oplus \cdots \oplus V^k$ *for some Banach spaces* $V^1, \ldots, V^k$. *Let* $\Pi = (p_1, \ldots, p_k)$ *denote a* k*-tuple as in Definition 2.1.1, let* ω *be a control and let* $M \geq 1$ *be a real number. Let* $E = V^1 \oplus \cdots \oplus V^l$ *and* $F = V^{l+1} \oplus \cdots \oplus V^k$ *be Banach spaces. Let* $\Pi_1 = (p_1, \ldots, p_l)$ *and* $\Pi_2 = (p_{l+1}, \ldots, p_k)$ *denote the corresponding* l *and* $(k-l)$*-tuples.*

Let $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) : \Delta_T \to T^{(\Pi,1)}(V)$ *be a geometric* Π*-rough path such that*

(i) the Π*-variation of* $\mathbf{Z}$ *is controlled by* $M\omega$,

(ii) the Π_1*-variation of* $\mathbf{X} = \pi_{T^{(\Pi_1,1)}(E)}\mathbf{Z}$ *is controlled by* ω,

(iii) $\mathbf{Y} = \pi_{T^{(\Pi_2,1)}(F)}\mathbf{Z}$.

Then, for all $0 \leq \varepsilon \leq M^{-s_{m^*}}$, *the* Π*-variation of* $(\mathbf{X}, \varepsilon\mathbf{Y})$ *is controlled by* ω, *where*

$$s_{m^*} = \max_{s_m \leq 1} \left\{ s_m \in S^{\Pi} \right\} = \max_{R \in \mathcal{A}^{\Pi}_1} \deg_{\Pi}(R).$$

Proof. This lemma is analogous to Lemma 5.8 of [27], adapted to the inhomogeneous smoothness case.

Let $\mathbf{W}$ denote the Π-rough path $(\mathbf{X}, \varepsilon\mathbf{Y})$. For a multi-index $R = (r_1, \cdots, r_m)$, $|R|_F$ denotes the cardinality of the set $\{r |\ r \in R, r > l\}$. Then if $\mathbf{Z}$ has bounded variation, by simple rescaling arguments we get

$$\mathbf{W}^R_{s,t} = \varepsilon^{|R|_F}\mathbf{Z}^R_{s,t}.$$

By continuity, the last equality holds for general geometric Π-rough path $\mathbf{Z}$. This following inequality is now implied and completes the proof:

$$\left\|\mathbf{W}^R_{s,t}\right\| \leq \varepsilon^{|R|_F} M^{\deg_\Pi(R)} \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}.$$

□

We define the one-forms $h_1 : V \oplus W \to \mathrm{End}(V \oplus W)$ and $h_2 : V \oplus W \oplus W \oplus W \to \mathrm{End}(V \oplus W \oplus W \oplus W)$ by

$$h_1(x,y) = \begin{pmatrix} Id_V & 0 \\ f(y+\xi) & 0 \end{pmatrix}$$

$$h_2(x,y_1,y_2,d) = \begin{pmatrix} Id_V & 0 & 0 & 0 \\ 0 & 0 & Id_W & 0 \\ f(y_2+\xi) & 0 & 0 & 0 \\ \rho g(y_1+\xi, y_2+\xi)(d) & 0 & 0 & 0 \end{pmatrix}$$

for a fixed $\rho > 1$.

We denote the $k_1 = (k+N)$-tuple $\hat{\Pi} = \Pi_V * \Pi_W$ by Π_1 and we introduce the $k_2 = (k+3N)$-tuple $\Pi_2 = \Pi * \Pi_W * \Pi_W * \Pi_W$. Applying Proposition 2.4.1, we define $\mathbf{Z}_1(n) \in G\Omega_{\Pi_1}(V \oplus W)$ and $\mathbf{Z}_2(n) \in G\Omega_{\Pi_2}(V \oplus W \oplus W \oplus W)$ by

$$\mathbf{Z}_1(0) = (\mathbf{X}, 0), \text{ and } \mathbf{Z}_1(n+1) = \int h_1(\mathbf{Z}_1(n))\mathrm{d}\mathbf{Z}_1(n)$$

$$\mathbf{Z}_2(0) = (\mathbf{X}, 0, \mathbf{Y}(1), \mathbf{Y}(1)), \text{ and } \mathbf{Z}_2(n+1) = \int h_2(\mathbf{Z}_2(n))\mathrm{d}\mathbf{Z}_2(n)$$

where $\mathbf{Y}(n) = \pi_W \mathbf{Z}_1(n)$.

Lemma 2.5.3. *For all $n \geq 0$,*

$$\begin{aligned} \mathbf{Z}_1(n) &= (\mathbf{X}, \mathbf{Y}(n)) \\ \mathbf{Z}_2(n) &= (\mathbf{X}, \mathbf{Y}(n), \mathbf{Y}(n+1), \rho^n(\mathbf{Y}(n+1) - \mathbf{Y}(n))). \end{aligned}$$

Furthermore, if the Π_V-variation of $\mathbf{X}$ is controlled by ω, then the Π_i-variation of $\mathbf{Z}_i(0)$ is controlled by $M\omega$ for $i = 1, 2$ respectively on $[0, T_\rho]$, where M and T_ρ are defined below.

By Corollary 2.4.1, there exists a constant M_i depending only on Π_i, Γ, $\hat{\Gamma}$ and polynomially on the *Lip*-norm of h_i, such that if $\mathbf{Z}_i$ is a rough path in the appropriate space with Π_i-variation controlled by some control ω such that $\omega(0,T) < 1$, then the Π_i-variation of $\int h_i(\mathbf{Z}_i)\mathrm{d}\mathbf{Z}_i$ is controlled by ω for $i = 1, 2$ respectively. We define $M = \max(M_1, M_2)$, and without loss of generality we assume that $M \geq 1$. We chose $\varepsilon = M^{-s_{m^*}}$ Let ω_0 be a control of the Π-variation of $\mathbf{X}$. Let $T_\rho > 0$ be chosen to satisfy $\omega_0(0, T_\rho) = \varepsilon^{p_{\max}}$. Note that for $R \in \mathcal{A}_1^\Pi$,

$$1 \geq deg_\Pi(R) = \sum_{i=1}^{k} \frac{n_j(R)}{p_i} \geq \sum_{i=1}^{k} \frac{n_j(R)}{p_{\max}} = \frac{\|R\|}{p_{\max}}.$$

This implies that by setting $\omega = \varepsilon^{-p_{\max}}\omega_0$, $\varepsilon^{-1}\mathbf{X}$ is controlled by ω and $\omega(0, T_\rho) \leq 1$.

Lemma 2.5.4. *For all $n \geq 0$, the Π_1 and Π_2-variation of the following rough paths respectively*

$$\begin{array}{ll} & (\varepsilon^{-1}\mathbf{X}, \mathbf{Y}(n)) \\ \textit{and} & (\varepsilon^{-1}\mathbf{X}, \mathbf{Y}(n), \mathbf{Y}(n+1), \rho^n(\mathbf{Y}(n+1) - \mathbf{Y}(n))) \end{array}$$

are controlled by ω on $[0, T_\rho]$.

The proof is based on the Scaling lemma and analogous to the proof of Proposition 5.9 in [27].

Now we prove the main theorem. We follow the proof of the Universal Limit Theorem corresponding to the homogeneous case presented in [27].

Proof. By Lemma 2.5.4, the Π_2-variation of $\mathbf{Z}_2(n)$ for all $n \geq 0$ is controlled by ω on $[0, T_\rho]$. We define the linear map $A : V \oplus W \oplus W \oplus W \to (V \oplus W) \oplus (V \oplus W)$ by

$$A(x, y_1, y_2, d) = ((x, y_1), (0, d)).$$

This linear map has norm 1. Note that

$$((\mathbf{X}, \mathbf{Y}(n)), \rho^n(0, \mathbf{Y}(n+1) - \mathbf{Y}(n))) = ((\mathbf{X}, \mathbf{Y}(n)), \rho^n[(\mathbf{X}, \mathbf{Y}(n+1)) - (\mathbf{X}, \mathbf{Y}(n))])$$

is controlled by ω on $[0, T_\rho]$. Then Lemma 2.5.1 implies the existence of a constant C depending only on $\hat{\Pi}$, $\omega(0, T)$ and β, such that for all $(s, t) \in \Delta_T$

$$\left\| (\mathbf{X}, \mathbf{Y}(n))^R_{s,t} - (\mathbf{X}, \mathbf{Y}(n+1))^R_{s,t} \right\| \leq C\left(\rho^{-n} + \rho^{-n\|R\|}\right) \omega(s,t)^{\|R\|/p_{\max}}, \ \forall R \in \mathcal{A}_1^{\hat{\Pi}}. \tag{2.36}$$

The inequality implies that $(\mathbf{X}, \mathbf{Y}(n))$ converges in the $\hat{\Pi}$-variational topology on the interval $[0, T_\rho]$ to a rough path $(\mathbf{X}, \mathbf{Y}) \in G\Omega_{\hat{\Pi}}$, which is also a solution to the RDE (2.35).

Note that once ρ is chosen, T_ρ is bounded from below where the bound only depends on the *Lip*-norm of f, g, Π, Γ and $\hat{\Gamma}$ and the modulus of continuity of ω on $[0, T]$. This implies that one can paste together local solutions in order to get a solution on the whole interval $[0, T]$.

In order to prove uniqueness, we assume that $\widehat{\mathbf{Z}} = (\mathbf{X}, \widehat{\mathbf{Y}})$ is also a solution to the RDE (2.35). We compare $\mathbf{Y}(n)$ and $\widehat{\mathbf{Y}}$ by defining the function $h_3 : V \oplus W \oplus W \oplus W \to$

$\mathrm{End}(V \oplus W \oplus W \oplus W)$ by

$$h_3(x, y, \widehat{y}, \widehat{d}) = \begin{pmatrix} Id_V & 0 & 0 & 0 \\ f(y+\xi) & 0 & 0 & 0 \\ 0 & 0 & Id_W & 0 \\ \rho g(y, \widehat{y})(\widehat{d}) & 0 & 0 & 0 \end{pmatrix}$$

and defining $\mathbf{Z}_3(n)$ by

$$\mathbf{Z}_3(0) = (\mathbf{X}, 0, \widehat{\mathbf{Y}}, \widehat{\mathbf{Y}}), \text{and } \mathbf{Z}_3(n+1) = \int h_3(\mathbf{Z}_3(n)).$$

Arguments analogous to the proof of Lemma 2.5.3 imply that

$$\mathbf{Z}_3(n) = (\mathbf{X}, \mathbf{Y}(n), \widehat{\mathbf{Y}}, \rho^n(\widehat{\mathbf{Y}} - \mathbf{Y}(n))).$$

Now analogously to Lemma 2.5.4, the $\hat{\Pi}$-variation of $\mathbf{Z}_3(n)$ is controlled by ω on a small enough interval. Then by Lemma 2.5.1, $\mathbf{Y} = \widehat{\mathbf{Y}}$ on the same interval. The uniqueness of $\mathbf{Y}$ is implied by the uniform continuity of ω.

Define $I_f(\mathbf{X}, \xi) = (\mathbf{X}, \mathbf{Y})$. Analogous arguments to the proof of the Universal Limit Theorem in [27] imply that I_f is continuous from $G\Omega_\Pi(V) \times W \to G\Omega_{\hat{\Pi}}(V \oplus W)$ in the Π-$\hat{\Pi}$-variation topology.

□

2.6 Results on continuity and global control

In this section we collect some results related to the previous sections and useful for the later sections.

2.6.1 Control of the solution

Remark 2.6.1. The Universal Limit Theorem and the uniform continuity of the control ω imply that there exist constants M and τ depending only on the Lipschitz norm of f, Π, Γ and $\hat{\Gamma}$, such that on any interval $[s, t] \subseteq [0, T]$ satisfying $\omega(s, t) < \tau$ the $\Pi := \Pi_V * \Pi_W$-variation of the solution $\mathbf{Z}$ to the RDE (2.35) and the Π_W-variation of $\mathbf{Y}$ are controlled by $M\omega$.

The next result provides a global control of the solution.

Proposition 2.6.1. *Let $\mathbf{Z}$ be a $\Delta_T \to T^{(\Pi,1)}(V)$ continuous multiplicative functional and let ω be a control function. Let us assume that there exists a positive constant τ, such that $\omega(s, t) < \tau$ and $(s, t) \in \Delta_T$ imply that*

$$\left\| \mathbf{Z}_{s,t}^R \right\| \leq \frac{\omega(s, t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}$$

for each $R \in \mathcal{A}_1^{\Pi}$.

Then there exists a constant C depending only on τ, Π *and T such that*

$$\left\|\mathbf{Z}_{s,t}^R\right\| \leq \frac{C\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}$$

for all $(s,t) \in \Delta_T$ *and each* $R \in \mathcal{A}_1^{\Pi}$.

Proof. Let $[s,t]$ be a subinterval of $[0,T]$. Due to the super-additivity of the control ω, there exists a partition of $\mathcal{D} = \{s = t_0 < \cdots < t_n = t\}$ of $[s,t]$ such that $\omega(t_i, t_{i+1}) < \tau$ for $i = 0, 1, \ldots, n-1$ and $n \leq 1 + \omega(s,t)/\tau$.

By the multiplicative property,

$$\mathbf{Z}_{s,t}^R \quad = \sum_{R_1 * \cdots * R_n = R} \mathbf{Z}_{t_0,t_1}^{R_1} \otimes \cdots \otimes \mathbf{Z}_{t_{n-1},t_n}^{R_n}.$$

Hence

$$\begin{aligned}
\left\|\mathbf{Z}_{s,t}^R\right\| \quad &\leq \sum_{R_1 * \cdots * R_n = R} \frac{\omega(t_0,t_1)^{deg_\Pi(R_1)}}{\beta^k \Gamma_\Pi(R_1)} \cdots \frac{\omega(t_{n-1},t_n)^{deg_\Pi(R_l)}}{\beta^k \Gamma_\Pi(R_n)} \\
&\leq \prod_{j=1}^{k} \sum_{i_{j,1}+\cdots+i_{j,n}=n_j(R)} \frac{\omega(t_0,t_1)^{\frac{i_{j,1}}{p_j}} \cdots \omega(t_{n-1},t_n)^{\frac{i_{j,n}}{p_j}}}{\beta^n \left(\frac{i_{j,1}}{p_j}\right)! \cdots \left(\frac{i_{j,n}}{p_j}\right)!} \\
&\leq \prod_{j=1}^{k} p_j^{2n-2} \frac{\omega(t_0,t_n)^{\frac{n_j(R)}{p_j}}}{\beta^n \left(\frac{n_j(R)}{p_j}\right)!} \\
&\leq \left(\prod_{j=1}^{k} \frac{p_j^{2n-2}}{\beta^{n-1}}\right) \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}
\end{aligned}$$

where the third inequality is due to the multinomial version of the Neo-classical inequality and the super-additivity of the control ω. The assertion is implied by the last inequality. □

2.6.2 Continuity of the solution

Let us regard the following rough differential equations

$$\mathrm{d}\mathbf{Y}_t(1) \quad = \quad f_1(\mathbf{Y}_t(1))\mathrm{d}\mathbf{X}_t, \; \mathbf{Y}_s(1) = \xi \tag{2.37}$$

$$\mathrm{d}\mathbf{Y}_t(2) \quad = \quad f_2(\mathbf{Y}_t(2))\mathrm{d}\mathbf{X}_t, \; \mathbf{Y}_s(2) = \xi \tag{2.38}$$

where the Π_V-rough path $\mathbf{X}$ satisfies the condition of Definition 2.5.1 with a control ω and the one-forms $f_1, f_2 : W \to L(V,W)$ satisfy the conditions of Theorem 2.5.1. The existence and uniqueness of solutions to the RDEs (2.37) and (2.38) are implied by the assumed conditions.

Lemma 2.6.1. *Let $\mathbf{X} \in G\Omega_{\Pi}(V)$ be a geometric Π_V-rough path with Π_V-variation controlled by ω. Let $\mathbf{Z}(1) = (\mathbf{X}, \mathbf{Y}(1))$ and $\mathbf{Z}(2) = (\mathbf{X}, \mathbf{Y}(2))$ be the Π_W-rough path solutions on $[0,T]$ to the RDEs* (2.37) *and* (2.38) *respectively. Then there exist constants C_R ($R \in \mathcal{A}_1^{\hat{\Pi}}$) and $\tau > 0$ such that for any $\hat{T} \in [0,T]$ satisfying $\omega(0,\hat{T}) < \tau$ and for all $(s,t) \in \Delta_{\hat{T}}$,*

$$\left\| \mathbf{Z}^R_{s,t}(1) - \mathbf{Z}^R_{s,t}(2) \right\| \leq C_R \left(\|f_1 - f_2\|_{Lip(\Pi_W,\Gamma)} + \|f_1 - f_2\|^{\|R\|}_{Lip(\Pi_W,\Gamma)} \right) \frac{\omega(s,t)^{deg_{\Pi}(R)}}{\beta^k \Gamma_{\Pi}(R)}$$

where C_R and τ depend only on $p_{\max}$ and polynomially on $\|f_1\|_{Lip(\Pi_W,\Gamma)}$ and $\|f_2\|_{Lip(\Pi_W,\Gamma)}$.

Proof. The lemma is a version of Proposition 3.3.5 of [3] and the proof is an adaptation of that proposition's proof to the case of inhomogeneous degree of smoothness. The proposition in [3] is restricted to the $\|f_1 - f_2\|_{Lip(\Pi_W,\Gamma)} < 1$ case. Here we do not make this restriction.

The proof is based on the following Picard iterations:

$$\begin{aligned}
\mathbf{U}(1,0) &= (\mathbf{X},0), \text{ and } \mathbf{U}(1,n) = \int h_1(\mathbf{U}(1,n))d\mathbf{U}(1,n) \\
\mathbf{U}(2,0) &= (\mathbf{X},0), \text{ and } \mathbf{U}(2,n) = \int h_2(\mathbf{U}(2,n))d\mathbf{U}(2,n) \\
\mathbf{U}(3,0) &= (\mathbf{X},0,0), \text{ and } \mathbf{U}(3,n) = \int h_3(\mathbf{U}(3,n))d\mathbf{U}(3,n) \\
\mathbf{U}(4,0) &= (\mathbf{X},0,0,0), \text{ and } \mathbf{U}(4,n) = \int h_4(\mathbf{U}(4,n))d\mathbf{U}(4,n)
\end{aligned}$$

where the one-forms h_i, $i = 1,\ldots,4$ are defined as

$$h_1(x,y) = \begin{pmatrix} Id_V & 0 \\ f_1(y+\xi) & 0 \end{pmatrix}$$

$$h_2(x,y) = \begin{pmatrix} Id_V & 0 \\ f_2(y+\xi) & 0 \end{pmatrix}$$

$$h_3(x,y_1,y_2) = \begin{pmatrix} Id_V & 0 & 0 \\ f_1(y_1+\xi) & 0 & 0 \\ f_2(y_2+\xi) & 0 & 0 \end{pmatrix}$$

$$h_4(x,y_1,y_2,d) = \begin{pmatrix} Id_V & 0 & 0 & 0 \\ f_1(y_1+\xi) & 0 & 0 & 0 \\ f_2(y_2+\xi) & 0 & 0 & 0 \\ \phi(f_1(y_2+\xi) - f_2(y_2+\xi)) + g_1(y_1+\xi, y_2+\xi)d & 0 & 0 & 0 \end{pmatrix}$$

and $g_1(x,y)(x-y) = f_1(x) - f_1(y)$ is specified in Theorem 2.5.1.

We set $\phi = \|f_1 - f_2\|^{-1}_{Lip(\Pi_W,\Gamma)}$. Then we can bound the appropriate Lipschitz norm of h_i for $i = 1, \ldots, 4$ and $\Pi_1 = \Pi_2 = \Pi_V * \Pi_W$, $\Pi_3 = \Pi_2 * \Pi_W$, $\Pi_4 = \Pi_3 * \Pi_W$ in terms of $\|f_1\|_{Lip(\Pi_W,\Gamma)}$ and $\|f_2\|_{Lip(\Pi_W,\Gamma)}$ as follows:

$$\begin{aligned}
\|h_1\|_{Lip(\Pi_1,\Gamma)} &\leq \|f_1\|_{Lip(\Pi_W,\Gamma)} \\
\|h_2\|_{Lip(\Pi_2,\Gamma)} &\leq \|f_2\|_{Lip(\Pi_W,\Gamma)} \\
\|h_3\|_{Lip(\Pi_3,\Gamma)} &\leq \|f_1\|_{Lip(\Pi_W,\Gamma)} + \|f_2\|_{Lip(\Pi_W,\Gamma)} \\
\|h_4\|_{Lip(\Pi_4,\hat{\Gamma})} &\leq 1 + C\left(\|f_1\|_{Lip(\Pi_W,\Gamma)} + \|f_2\|_{Lip(\Pi_W,\Gamma)}\right)
\end{aligned}$$

Note, that

$$\mathbf{U}(4,n) = (\mathbf{X}, \mathbf{Y}(1,n+1), \mathbf{Y}(2,n+1), \phi(\mathbf{Y}(2,n+1) - \mathbf{Y}(2,n+1)))$$

where $\mathbf{Y}(1,n) = \pi_W \mathbf{U}(1,n)$ and $\mathbf{Y}(2,n) = \pi_W \mathbf{U}(2,n)$.

Let M be the constant depending only on a polynomial of $\|h_i\|_{Lip(\Pi_i,\Gamma)}$, $i = 1,2,3$ and $\|h_4\|_{Lip(\Pi_i,\hat{\Gamma})}$, such that whenever $\mathbf{Z}$ is a geometric Π-rough path controlled by ω_0 such that $\omega_0(0,T) \leq 1$, then $\int h_i(\mathbf{Z})\mathrm{d}\mathbf{Z}$ is controlled by $M\omega_0$. Let $\varepsilon > 0$ be fixed and let τ be defined as $\tau := \max(\varepsilon, M)^{-\lfloor p_{\max} \rfloor}$. Furthermore, let $\widehat{T}$ satisfy $\omega(0,\widehat{T}) = \min(\omega(0,T), \tau)$ and let $\omega_1 := (1/\tau)\omega$.

Then by the Scaling Lemma and a similar technique to that used in the proof of Theorem 2.5.1, one can show that the Π-variation of

$$((\mathbf{X}, \mathbf{Y}(1,n)), \phi\left[(\mathbf{X}, \mathbf{Y}(1,n)) - (\mathbf{X}, \mathbf{Y}(2,n))\right])$$

is controlled by ω_1 on $[0,\widehat{T}]$. Finally, by Lemma 2.5.1 there exists a constant C_R uniform in n such that

$$\left\|(\mathbf{X}, \mathbf{Y}(1,n))^R_{s,t} - (\mathbf{X}, \mathbf{Y}(2,n))^R_{s,t}\right\| \leq C\left(\phi^{-1} + \phi^{-\|R\|}\right) \frac{\omega_1(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}$$

for all $R \in \mathcal{A}_1^{\hat{\Pi}}$ and $(s,t) \in \Delta_{\widehat{T}}$. This implies the assertion. □

The ideas in the proof of Lemma 2.6.1 can be used to give a bound on the difference between the solutions to RDEs driven by the same noise and the same one-form but by different initial conditions. I.e. let us regard the following RDEs

$$\mathrm{d}\mathbf{Y}_t(1) = f(\mathbf{Y}_t(1))\mathrm{d}\mathbf{X}_t, \ \mathbf{Y}_s(1) = \xi \tag{2.39}$$

$$\mathrm{d}\mathbf{Y}_t(2) = f(\mathbf{Y}_t(2))\mathrm{d}\mathbf{X}_t, \ \mathbf{Y}_s(2) = \eta \tag{2.40}$$

where the Π-rough path $\mathbf{X}$ satisfies the condition of Definition 2.5.1 with a control ω and the one-forms $f : W \to L(V, W)$ satisfy the conditions of Theorem 2.5.1.

Lemma 2.6.2. *Let $\mathbf{X} \in G\Omega_{\Pi}(V)$ be a geometric Π-rough path with Π-variation controlled by ω. Let $\mathbf{Z}(1) = (\mathbf{X}, \mathbf{Y}(1))$ and $\mathbf{Z}(2) = (\mathbf{X}, \mathbf{Y}(2))$ be the solutions on $[0, T]$ to the RDEs* (2.39) *and* (2.40) *respetively. Then for each $R \in \mathcal{A}_1^{\hat{\Pi}}$ there exist constants C_R and $\tau > 0$ such that for any $\hat{T} \in [0, T]$ satisfying $\omega(0, \hat{T}) < \tau$ and for all $(s, t) \in \Delta_{\hat{T}}$*

$$\left\| \mathbf{Z}_{s,t}^R(1) - \mathbf{Z}_{s,t}^R(2) \right\| \leq C_R \left(\|\xi - \eta\| + \|\xi - \eta\|^{\|R\|} \right) \frac{\omega(s,t)^{deg_{\Pi}(R)}}{\beta^k \Gamma_{\Pi}(R)}$$

where C_R depends only on $p_{\max}$ and polynomially on $\|f\|_{Lip(\Pi,\Gamma)}$.

Proof. Firstly, note that the solutions to the RDEs

$$\begin{aligned} \mathrm{d}\mathbf{Y}_t(1) &= f(\xi + \mathbf{Y}_t(1))\mathrm{d}\mathbf{X}_t, \ \mathbf{Y}_s(1) = \mathbf{1} \\ \mathrm{d}\mathbf{Y}_t(2) &= f(\eta + \mathbf{Y}_t(2))\mathrm{d}\mathbf{X}_t, \ \mathbf{Y}_s(2) = \mathbf{1} \end{aligned}$$

coincide with the solutions to (2.39) and (2.40) respectively.

Let the functions $f_1, f_2 : W \to L(V, W)$ be defined by $f_1(x) := f(x + \xi)$ and $f_2(x) := f(x + \eta)$. Note that

$$f_1(x) - f_2(x) = g(x + \xi, x + \eta)(\eta - \xi).$$

Then the proof is implied by Lemma 2.6.1.

□

The above lemmas are restricted to a subinterval of $[0, T]$. Now we give an extension of Lemma 2.6.2 to the entire $[0, T]$.

Proposition 2.6.2. *Under the conditions of Lemma 2.6.2, for each $R \in \mathcal{A}_1^{\hat{\Pi}}$ there exists a constant Q_R depending on $p_{\max}$ and polynomially on $\|f\|_{Lip(\Pi_W,\Gamma)}$, such that*

$$\left\| \mathbf{Z}_{s,t}^R(1) - \mathbf{Z}_{s,t}^R(2) \right\| \leq Q_R \left(\|\xi - \eta\| + \|\xi - \eta\|^{\|R\|} \right) \frac{\omega(s,t)^{deg_{\Pi}(R)}}{\beta^k \Gamma_{\Pi}(R)}$$

for all $(s, t) \in [0, T]$, where $\mathbf{Z}(1) = (\mathbf{X}, \mathbf{Y}(1))$ and $\mathbf{Z}(2) = (\mathbf{X}, \mathbf{Y}(2))$ are solutions to the RDEs (2.39) *and* (2.40) *respectively.*

Proof. **Step 1**

By Lemma 2.6.2, there exist constants C and τ depending only on $p_{\max}$ and polynomially on $\|f\|_{Lip(\Pi,\Gamma)}$, such that

$$\left\| \mathbf{Z}_{0,\hat{T}}^W(1) - \mathbf{Z}_{0,\hat{T}}^W(2) \right\| \leq 2C\|\xi - \eta\| \frac{\omega(0,\hat{T})^{\frac{1}{p_{\max}}}}{\beta^k \left(\frac{1}{p_{\max}}\right)!}$$

whenever $\omega(0, \hat{T}) < \tau$. Following the proof of Lemma 2.6.1, τ can be chosen to satisfy

$$\frac{\tau^{\frac{1}{p_{\max}}}}{\beta^k \left(\frac{1}{p_{\max}}\right)!} \leq 1.$$

Let us define the sequence $\{s_n\}$ recursively by $s_0 = 0$ and

$$s_{i+1} = \sup\{t \in (s_i, T] \mid \omega(s_i, t) \leq \tau/2\}.$$

Due to the super-additivity of the control ω, the sequence s_n takes finitely many values. More precisely, there exists a positive integer n, such that $s_l = s_n = T$ for all $l \geq n$.

On $[s_1, T]$ the solutions to the RDEs

$$\begin{aligned} d\mathbf{Y}_t(1) &= f(\xi + \mathbf{Z}^W_{0,s_1}(1) + \mathbf{Y}_t(1))d\mathbf{X}_t, \ \mathbf{Y}_{s_1}(1) = \mathbf{1} \\ d\mathbf{Y}_t(2) &= f(\eta + \mathbf{Z}^W_{0,s_1}(2) + \mathbf{Y}_t(2))d\mathbf{X}_t, \ \mathbf{Y}_{s_1}(2) = \mathbf{1} \end{aligned}$$

coincide with the solutions to (2.39) and (2.40) respectively. Therefore by Lemma 2.6.2 we have for all $(s,t) \in [s_1, s_2]$

$$\left\|\mathbf{Z}^R_{s,t}(1) - \mathbf{Z}^R_{s,t}(2)\right\| \leq C_R\left(2C\|\xi - \eta\| + (2C)^{\|R\|}\|\xi - \eta\|^{\|R\|}\right)\frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)},$$

implying

$$\left\|\mathbf{Z}^W_{s_1,s_2}(1) - \mathbf{Z}^W_{s_1,s_2}(2)\right\| \leq (2C)^2\|\xi - \eta\|$$

and

$$\left\|\mathbf{Z}^W_{0,s_2}(1) - \mathbf{Z}^W_{0,s_2}(2)\right\| \leq \left(2C + (2C)^2\right)\|\xi - \eta\|.$$

Recursively, we derive for $s_q < T$ and $q \leq n$ that

$$\left\|\mathbf{Z}^W_{0,s_q}(1) - \mathbf{Z}^W_{0,s_q}(2)\right\| \leq C_q\|\xi - \eta\|$$

and for all $(s,t) \in [s_q, s_{q+1}]$

$$\left\|\mathbf{Z}^R_{s,t}(1) - \mathbf{Z}^R_{s,t}(2)\right\| \leq C_R\left(C_q\|\xi - \eta\| + C_q^{\|R\|}\|\xi - \eta\|^{\|R\|}\right)\frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}$$

where $C_q = 2C + \cdots + (2C)^q$.

Step 2

In the case when $[s,t]$ is a subinterval of $[0,T]$ such that

$$s_p \leq t_0 = s < t_1 = s_{p+1} < \cdots < t_{q-1} = s_{p+q-1} < t_q = t \leq s_{p+q},$$

we write

$$
\begin{aligned}
&\left\| \mathbf{Z}^R_{s,t}(1) - \mathbf{Z}^R_{s,t}(2) \right\| = \\
= &\left\| \sum_{R_1 * \cdots * R_q = R} \mathbf{Z}^{R_1}_{t_0,t_1}(1) \otimes \cdots \otimes \mathbf{Z}^{R_q}_{t_{q-1},t_q}(1) - \mathbf{Z}^{R_1}_{t_0,t_1}(2) \otimes \cdots \otimes \mathbf{Z}^{R_q}_{t_{q-1},t_q}(2) \right\| \\
= &\left\| \sum_{R_1 * \cdots * R_q = R} \sum_{j=0}^{q-1} \mathbf{Z}^{R_1}_{t_0,t_1}(1) \otimes \cdots \otimes \left(\mathbf{Z}^{R_j}_{t_j,t_{j+1}}(1) - \mathbf{Z}^{R_j}_{t_j,t_{j+1}}(2) \right) \otimes \cdots \otimes \mathbf{Z}^{R_q}_{t_{q-1},t_q}(2) \right\| \\
\leq &\sum_{R_1 * \cdots * R_q = R} \sum_{j=0}^{q-1} \left\| \mathbf{Z}^{R_1}_{t_0,t_1}(1) \right\| \cdots \left\| \left(\mathbf{Z}^{R_j}_{t_j,t_{j+1}}(1) - \mathbf{Z}^{R_j}_{t_j,t_{j+1}}(2) \right) \right\| \cdots \left\| \mathbf{Z}^{R_q}_{t_{q-1},t_q}(2) \right\| \\
\leq &Q_R \left(\|\xi - \eta\| + \|\xi - \eta\|^{\|R\|} \right) \sum_{R_1 * \cdots * R_q = R} \frac{\omega(t_0,t_1)^{deg_\Pi(R_1)} \cdots \omega(t_{q-1},t_q)^{deg_\Pi(R_q)}}{\beta^k \Gamma_\Pi(R_0) \cdots \beta^k \Gamma_\Pi(R_q)} \\
\leq &Q_R \left(\|\xi - \eta\| + \|\xi - \eta\|^{\|R\|} \right) \prod_{j=1}^{k} \sum_{i_{j,1} + \cdots + i_{j,q} = n_j(R)} \frac{\omega(t_0,t_1)^{\frac{i_{j,1}}{p_j}} \cdots \omega(t_{q-1},t_q)^{\frac{i_{j,q}}{p_j}}}{\beta^q \left(\frac{i_{j,1}}{p_j} \right)! \cdots \left(\frac{i_{j,q}}{p_j} \right)!} \\
\leq &Q_R \left(\|\xi - \eta\| + \|\xi - \eta\|^{\|R\|} \right) \prod_{j=1}^{k} p_j^{2q-2} \frac{\omega(t_0,t_q)^{\frac{n_j(R)}{p_j}}}{\beta^q \left(\frac{n_j(R)}{p_j} \right)!} \\
\leq &Q_R \left(\|\xi - \eta\| + \|\xi - \eta\|^{\|R\|} \right) \left(\prod_{j=1}^{k} \frac{p_j^{2q-2}}{\beta^{q-1}} \right) \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}
\end{aligned}
$$

where the next-to-last inequality is due to the multinomial version of the Neo-classical Inequality and Q_R is a polynomial function of C and the C_{R_i}'s $R_i \in \mathcal{A}_1^{\hat{\Pi}}$ with coefficients depending on $\omega(0,T)/(\tau/2)$. The polynomial Q_R has a degree at most $\|R\|\omega(0,T)/(\tau/2)$ when regarded as a polynomial of C, and degree at most $\|R\|$ when regarded as a polynomial of C_{R_i}. □

Finally we extend Lemma 2.6.1 to the interval $[0,T]$. This proposition is an RDE version of Gronwall's Lemma.

Proposition 2.6.3. *Under the conditions of Lemma 2.6.1 there exist polynomials Q_R of*

$$\|f_1 - f_2\|_{Lip(\Pi_W,\Gamma)}$$

with coefficients depending on $p_{\max}$, $\omega(0,T)$ and polynomially on $\|f_1\|_{Lip(\Pi_W,\Gamma)}$ and $\|f_2\|_{Lip(\Pi_W,\Gamma)}$, such that

$$\left\| \mathbf{Z}^R_{s,t}(1) - \mathbf{Z}^R_{s,t}(2) \right\| \leq Q_R \left(\|f_1 - f_2\|_{Lip(\Pi,\Gamma)} + \|f_1 - f_2\|^{\|R\|}_{Lip(\Pi,\Gamma)} \right) \frac{\omega(s,t)^{deg_\Pi(R)}}{\beta^k \Gamma_\Pi(R)}$$

for all $(s,t) \in [0,T]$ and $R \in \mathcal{A}_1^{\hat{\Pi}}$.

Proof. **Step 1**

Let x be an element in W and $i \in \{1,2\}$. Let $\mathbf{Z}(x,u,i) = (\mathbf{X}(x,u,i), \mathbf{Y}(x,u,i))$ denote the solution to the RDE

$$d\mathbf{Y}_s(x,u,i) = f_i(\mathbf{Y}_s(x,u,i))d\mathbf{X}_s,\ \mathbf{Y}_u(x,u,i) = \xi + x.$$

For $(u,v) \in \Delta_T$ and $i \in \{1,2\}$, we introduce the map $D^i_{u,v} : T^{(\Pi_V * \Pi_W,1)}(V \oplus W) \to T^{(\Pi_V * \Pi_W,1)}(V \oplus W)$ defined by

$$D^i_{u,v}(z) = z \otimes \mathbf{Z}_{u,v}(\Pi_W z, u, i),\ z \in T^{(\Pi_V * \Pi_W,1)}(V \oplus W).$$

Let $\mathcal{D} = \{s = t_0 < \cdots < t_n = t\}$ be a partition of $[s,t]$, such that $\omega(t_i,t_{i+1}) < \tau$ for $i = 0,\ldots,n-1$ and $n \le 1 + \omega(s,t)/\tau \le 1 + \omega(0,T)/\tau$, where τ is the constant introduced in 2.6.1. Then $\mathbf{Z}_{s,t}(1) - \mathbf{Z}_{s,t}(2)$ can be represented as follows:

$$\begin{aligned}
\mathbf{Z}_{s,t}(1) - \mathbf{Z}_{s,t}(2) &= D^1_{s,t}(\mathbf{Z}_{0,s}(1)) - D^2_{s,t}(\mathbf{Z}_{0,s}(2)) \\
&= D^1_{s,t}(\mathbf{Z}_{0,s}(1)) - D^1_{s,t}(\mathbf{Z}_{0,s}(2)) + \qquad (2.41)\\
&\quad + \sum_{i=0}^{n-1} \Big(D^1_{t_{i+1},t_n} \circ D^1_{t_i,t_{i+1}} \circ D^2_{t_0,t_i}(\mathbf{Z}_{0,s}(2)) \\
&\qquad - D^1_{t_{i+1},t_n} \circ D^2_{t_i,t_{i+1}} \circ D^2_{t_0,t_i}(\mathbf{Z}_{0,s}(2)) \Big) \qquad (2.42)
\end{aligned}$$

Step 2

Let us consider a term of the sum (2.42). Using the identity

$$abc - ade = ab(c-e) + a(b-d)e$$

and the multiplicative property of rough paths, for $Q \in \mathcal{A}_1^{\hat{\Pi}}$, we have

$$\begin{aligned}
&\Big\| \pi_Q D^1_{t_{i+1},t_n} \circ D^1_{t_i,t_{i+1}} \circ D^2_{t_0,t_i}(\mathbf{Z}_{0,s}(2)) - \pi_Q D^1_{t_{i+1},t_n} \circ D^2_{t_i,t_{i+1}} \circ D^2_{t_0,t_i}(\mathbf{Z}_{0,s}(2)) \Big\| \\
&= \Bigg\| \sum_{\substack{Q_1,Q_2,Q_3 \in \mathcal{A}_1^{\hat{\Pi}} \\ Q_1 * Q_2 * Q_3 = Q}} \Big[\mathbf{Z}^{Q_1}_{t_0,t_i}\Big(\mathbf{Z}^W_{0,s}(2), t_0, 2\Big)\Big] \otimes \Big[\mathbf{Z}^{Q_2}_{t_i,t_{i+1}}\Big(\Pi_W y, t_i, 1\Big)\Big] \\
&\qquad \otimes \Big[\mathbf{Z}^{Q_3}_{t_{i+1},t_n}\Big(\Pi_W x_1, t_{i+1}, 1\Big) - \mathbf{Z}^{Q_3}_{t_{i+1},t_n}\Big(\Pi_W x_2, t_{i+1}, 1\Big)\Big] \\
&\quad + \Big[\mathbf{Z}^{Q_1}_{t_0,t_i}\Big(\mathbf{Z}^W_{0,s}(2), t_0, 2\Big)\Big] \\
&\qquad \otimes \Big[\mathbf{Z}^{Q_2}_{t_i,t_{i+1}}\Big(\Pi_W y, t_i, 1\Big) - \mathbf{Z}^{Q_2}_{t_i,t_{i+1}}\Big(\Pi_W y, t_i, 2\Big)\Big] \\
&\qquad \otimes \Big[\mathbf{Z}^{Q_3}_{t_{i+1},t_n}\Big(\Pi_W x_2, t_{i+1}, 1\Big)\Big] \Bigg\|
\end{aligned}$$

$$
\begin{aligned}
\leq \sum_{\substack{Q_1,Q_2,Q_3\in\mathcal{A}_1^{\hat\Pi}\\ Q_1*Q_2*Q_3=Q}} & \left\|\mathbf{Z}^{Q_1}_{t_0,t_i}\left(\mathbf{Z}^W_{0,s}(2),t_0,2\right)\right\|\cdot\left\|\mathbf{Z}^{Q_2}_{t_i,t_{i+1}}\left(\Pi_W y,t_i,1\right)\right\| \\
& \cdot\left\|\mathbf{Z}^{Q_3}_{t_{i+1},t_n}\left(\Pi_W x_1,t_{i+1},1\right)-\mathbf{Z}^{Q_3}_{t_{i+1},t_n}\left(\Pi_W x_2,t_{i+1},1\right)\right\| \\
+ & \left\|\mathbf{Z}^{Q_1}_{t_0,t_i}\left(\mathbf{Z}^W_{0,s}(2),t_0,2\right)\right\| \\
& \cdot\left\|\mathbf{Z}^{Q_2}_{t_i,t_{i+1}}\left(\Pi_W y,t_i,1\right)-\mathbf{Z}^{Q_2}_{t_i,t_{i+1}}\left(\Pi_W y,t_i,2\right)\right\| \\
& \cdot\left\|\mathbf{Z}^{Q_3}_{t_{i+1},t_n}\left(\Pi_W x_2,t_{i+1},1\right)\right\| \qquad (2.43)
\end{aligned}
$$

where

$$
\begin{aligned}
x_j &= D^j_{t_i,t_{i+1}}\circ D^2_{t_0,t_i}(\mathbf{Z}_{0,s}(2)),\ \text{for } j=1,2\\
y &= D^2_{t_0,t_i}(\mathbf{Z}_{0,s}(2)).
\end{aligned}
$$

Applying Proposition 2.6.1, there exist constants C_1 and C_2 depending on Π and T and polynomially on $\|f_1\|_{Lip(\Pi_W,\Gamma)}$ and $\|f_2\|_{Lip(\Pi_W,\Gamma)}$ respectively, such that the following inequalities are satisfied:

$$
\left\|\mathbf{Z}^{Q_1}_{t_0,t_i}\left(\mathbf{Z}_W 0,s(2),t_0,2\right)\right\| \leq \frac{(C_2\omega(t_0,t_i))^{deg_{\hat\Pi}(Q_1)}}{\beta^k\Gamma_{\hat\Pi(Q_1)}} \qquad (2.44)
$$

$$
\left\|\mathbf{Z}^{Q_2}_{t_i,t_{i+1}}\left(\Pi_W y_1,t_i,1\right)\right\| \leq \frac{(C_1\omega(t_i,t_{i+1}))^{deg_{\hat\Pi}(Q_2)}}{\beta^k\Gamma_{\hat\Pi(Q_2)}} \qquad (2.45)
$$

$$
\left\|\mathbf{Z}^{Q_3}_{t_{i+1},t_n}\left(\Pi_W x_2,t_{i+1},1\right)\right\| \leq \frac{(C_1\omega(t_i,t_{i+1}))^{deg_{\hat\Pi}(Q_3)}}{\beta^k\Gamma_{\hat\Pi(Q_3)}}. \qquad (2.46)
$$

Applying 2.6.1, there exists a constant C_3 depending on $p_{\max}$ and polynomially on $\|f_1\|_{Lip(\Pi_W,\Gamma)}$ and $\|f_2\|_{Lip(\Pi_W,\Gamma)}$ satisfying

$$
\begin{aligned}
\|x_1-x_2\| &= \left\|\mathbf{Z}^{Q_2}_{t_i,t_{i+1}}\left(\Pi_W y,t_i,1\right)-\mathbf{Z}^{Q_2}_{t_i,t_{i+1}}\left(\Pi_W y,t_i,2\right)\right\| \\
&\leq C_3\left(\|f_1-f_2\|_{Lip(\Pi_W,\Gamma)}+\|f_1-f_2\|^{\|Q_2\|}_{Lip(\Pi_W,\Gamma)}\right)\frac{\omega(t_i,t_{i+1})^{deg_\Pi(Q_2)}}{\beta^k\Gamma_\Pi(Q_2)}. \qquad (2.47)
\end{aligned}
$$

Proposition 2.6.2 implies the existence of a constant C_4 depending on $p_{\max}$ and polynomially on $\|f_1\|_{Lip(\Pi,\Gamma)}$, such that

$$
\begin{aligned}
&\left\|\mathbf{Z}^{Q_3}_{t_{i+1},t_n}\left(\Pi_W x_1,t_{i+1},1\right)-\mathbf{Z}^{Q_3}_{t_{i+1},t_n}\left(\Pi_W x_2,t_{i+1},1\right)\right\| \\
&\qquad\leq C_4\left(\|\Pi_W(x_1-x_2)\|+\|\Pi_W(x_1-x_2)\|^{\|Q_3\|}\right)\frac{\omega(t_{i+1},t_n)^{deg_\Pi(Q_3)}}{\beta^k\Gamma_\Pi(Q_3)}. \qquad (2.48)
\end{aligned}
$$

Step 3

One can bound the expression (2.43) using the inequalities (2.44)-(2.48). Then the multinomial version of the Neo-classical Inequality implies the assertion for all $(0,t)\in\Delta_T$.

Step 4

To prove the claim for all $(s,t) \in \Delta_T$, an appropriate bound on the term (2.41) is required. This can be achieved by again applying Proposition 2.6.2, i.e. for any $Q \in \mathcal{A}_1^{\hat{\Pi}}$

$$\begin{aligned} &\left\| \mathbf{Z}^Q_{s,t}(\mathbf{Z}^W_{0,s}(1), s, 1) - \mathbf{Z}^Q_{s,t}(\mathbf{Z}^W_{0,s}(2), s, 1) \right\| \\ &\leq C_4 \left(\|\mathbf{Z}^W_{0,s}(1) - \mathbf{Z}^W_{0,s}(2)\| + \|\mathbf{Z}^W_{0,s}(1) - \mathbf{Z}^W_{0,s}(2)\|^{\|Q\|} \right) \frac{\omega(s,t)^{deg_\Pi(Q)}}{\beta^k \Gamma_\Pi(Q)}. \end{aligned} \tag{2.49}$$

The general case is implied by the special case of Step 3 and the inequality (2.49).

□

2.6.3 A remark on simple functionals of rough paths

Let $\Pi = (p_1, \ldots, p_k)$ be a k-tuple and let $\mathbf{X} \in G\Omega_\Pi(V)$. Let us consider the space $W = \bigoplus_{R \in \mathcal{A}_\alpha^\Pi} V^R$ for some $\alpha \geq 1$ as well as the RDE

$$d\mathbf{Y}_t = \mathbf{Y}_t \otimes d\mathbf{X}_t, \ \mathbf{Y}_0 = \mathbf{1}, \tag{2.50}$$

This RDE is the same as (2.20) in the sense of Definition 2.5.1. Equivalently, indexing the components of W and $\mathbf{Y}^W$ by multi-indices (putting the index into $\{\cdot\}$ in order to distinguish from projection)

$$d\mathbf{Y}_t^{\{R\}} = \mathbf{Y}_t^{\{R-\}} \otimes d\mathbf{X}_t^{r_l}, \ l \in \mathbb{N}, R = (r_1, \ldots, r_l) \in \mathcal{A}_\alpha^\Pi, \mathbf{Y}_0 = \mathbf{1}.$$

By the Universal Limit Theorem, the RDE (2.50) has a unique solution and by the special form of the RDE, $\mathbf{Y}_{0,t}^{\{R\}} = \mathbf{X}_{0,t}^R$ ($R \in \mathcal{A}_\alpha^\Pi$) represents the W-level of $\mathbf{Y}$ and $\mathbf{Y}_{s,t}^{\{R\}} = \mathbf{X}_{0,t}^R - \mathbf{X}_{0,s}^R$ for all $t \in [0,T]$. Furthermore, $\mathbf{Y}$ is in $G\Omega_{p_{\max}}(W)$. However, due to the special form of the one-form defining the RDE, $\mathbf{Y}$ is also a Π_α-rough path corresponding to the decomposition $W = \bigoplus_{0 \leq m \leq C_\alpha} W^m$, where $C_\alpha = \text{card}\{s_m \in S^\Pi, s_m \leq \alpha\}$,

$$\Pi_\alpha = (q_1, \ldots q_{c_\alpha})$$

with $q_i = \max\left\{\frac{1}{s_i}, 1\right\}$ and

$$W_i = \bigoplus_{deg_\Pi(R) = s_i} V^R.$$

By Proposition 2.6.1, there exists a constant C, depending only on Π and α, such that $\mathbf{Y}$ is controlled by $C\omega$.

CHAPTER 3

NUMERICAL SOLUTION OF ROUGH DIFFERENTIAL EQUATIONS

The main aim of this chapter is to construct numerical methods to approximate solutions to rough differential equations (RDEs). Firstly, some principles of approximations are set up. Then a general class of local approximations is introduced. This class is used to construct global approximations by pasting together the local ones. A general sufficient condition on the local approximations implying global convergence is given and proved. The next step is to construct particular local approximations in finite dimensions based on solutions to ordinary differential equations derived locally and satisfying the sufficient condition for global convergence. These local approximations require strong conditions on the one-form defining the rough differential equation. Finally, we show that when the local ODE based schemes are applied in combination with rough polynomial approximations, the conditions on the one-form can be weakened.

Throughout the chapter, we assume that $V = V^1 \oplus \cdots \oplus V^k$ and $W = W^1 \oplus \cdots \oplus W^N$ are direct sums of Banach spaces. In some of the sections we assume finite dimensionality of V and W.

Our primary focus is the rough differential equation

$$\mathrm{d}\mathbf{Y}_t = f(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_0 = \xi \tag{3.1}$$

driven by a geometric Π_X-rough path $\mathbf{X}$ controlled by a control ω, where $\Pi_X = (p_1, \dots, p_k)$ is a real k-tuple and the one-form $f : W \to L(V, W)$ is of the form

$$f(y) = \begin{pmatrix} V_1(y) & \dots & V_k(y) \end{pmatrix} \tag{3.2}$$

where V_i, $i = 1, \dots, k$ are $W \to W$ functions. The jth coordinate function of V_i is denoted by V_i^j.

In general, we assume that the one-form f satisfies (i.e. the functions V_i, $i = 1, \dots, k$ satisfy) the conditions of the Universal Limit Theorem implying that the RDE (3.1) possesses a unique solution.

Without loss of generality, we can assume that the solution $\mathbf{Y}$ is a geometric Π_Y-rough path in W such that $\Pi_Y = (q_1, \dots, q_N)$ and each q_i is one of the p_i's in Π_X.

We will consider the extensions of the rough paths **X** and **Y** to

$$T^{(\Pi_X,\alpha_X)}(V) \text{ and } T^{(\Pi_Y,\alpha_Y)}(W)$$

respectively for some $\alpha_X, \alpha_Y \geq 1$.

3.1 Approximations of rough paths

In this section, a concept of approximations of Π-rough paths is introduced. We aim to construct an approximation of the solution based on some information on the driving noise. There are some issues which we must decide on. Firstly, the input information must be determined. Secondly, the object we approximate the solution with must be defined. Finally, we must decide in which sense we wish to approximate the solution; i.e. how to measure the error.

In this chapter we assume that we are provided with the *local signature* of the driving noise corresponding to a certain partition of $[0, T]$ in the following sense.

Definition 3.1.1. *Let $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$ be a partition of $[0, T]$. We define the discrete simplex corresponding to the partition $\mathcal{D}$ by*

$$\Delta_T^{\mathcal{D}} = \{(s,t)|\, s,t \in \mathcal{D},\ s \leq t\}.$$

Definition 3.1.2. *Let $\alpha_X \geq 1$ be a real number. Let $\mathbf{X} : \Delta_T \to T^{(\Pi_X,\alpha_X)}(V)$ be a continuous multiplicative functional. The tensor algebra element $\mathbf{X}_{s,t}$ will be referred to as the* signature corresponding to the interval $[s, t]$.

A $(\alpha_X, \mathcal{D})$-multiplicative functional is a $\mathbf{X} : \Delta_T^{\mathcal{D}} \to T^{(\Pi_X,\alpha_X)}(V)$ function such that

$$\mathbf{X}_{s,u} = \mathbf{X}_{s,t} \otimes \mathbf{X}_{t,u},\ \forall (s,t) \in \Delta_t^{\mathcal{D}}.$$

We define the local signature of the $\mathbf{X} : \Delta_T \to T^{(\Pi_X,\alpha_X)}(V)$ multiplicative functional corresponding to the partition $\mathcal{D}$ *as a restriction to a $\widehat{\mathbf{X}} : \Delta_T^{\mathcal{D}} \to T^{(\Pi_X,\alpha_X)}(V)$ function. This function is trivially $(\alpha_X, \mathcal{D})$-multiplicative.*

We will consider numerical schemes where the input is the local signature of the driving noise corresponding to a partition $\mathcal{D}$ of $[0, T]$ and the output is an $\left(\alpha_Y, \widehat{\mathcal{D}}\right)$-multiplicative and $T^{(\Pi_Y,\alpha_Y)}(W)$-valued functional for a real number $\alpha_Y \geq 1$ and a partition $\widehat{\mathcal{D}} \subseteq \mathcal{D}$; i.e. an approximation of the solution's local signature corresponding to the partition $\widehat{\mathcal{D}}$.

Definition 3.1.3. *Let the real number $\alpha_Y \geq 1$ and the partition $\widehat{\mathcal{D}}$ of $[0, T]$ be fixed. Then we define an* approximation *of* **Y** *as a function F assigning an $\left(\alpha_Y, \widehat{\mathcal{D}}\right)$-multiplicative functional to each partition $\mathcal{D} \supseteq \widehat{\mathcal{D}}$ of the interval $[0, T]$.*

Definition 3.1.4. *Let F be an approximation of $\mathbf{Y}$. Let $\mathcal{D}$ be a partition of $[0,T]$ and $\widehat{\mathbf{Y}} = F(\mathcal{D})$. For a positive real α and a partition $\widehat{\mathcal{D}} \subseteq \mathcal{D}$ of $[0,T]$, we define the distance between $\widehat{\mathbf{Y}}$ and $\mathbf{Y}$ by*

$$d_{\Pi_Y}^{\alpha,\widehat{\mathcal{D}}}(\mathbf{Y},\widehat{\mathbf{Y}}) := \max_{R \in \mathcal{A}_\alpha^k} \left(\sum_{\widehat{\mathcal{D}}} \left\| \mathbf{Y}^R_{t_{l-1},t_l} - \widehat{\mathbf{Y}}^R_{t_{l-1},t_l} \right\|^{1/deg_{\Pi_Y}(R)} \right)^{deg_{\Pi_Y}(R)}$$

Remark 3.1.1. The distance function $d_{\Pi_Y}^{\alpha,\widehat{\mathcal{D}}}(\cdot,\cdot)$ is a restriction of (induced by) the $d_{\Pi_Y}(\cdot,\cdot)$ distance to $\mathcal{D}$-multiplicative functions where $\widehat{\mathcal{D}} \subseteq \mathcal{D}$.

The $d_{\Pi_Y}(\cdot,\cdot)$ distance will play a crucial role in the construction of approximations. The error of these approximations can be measured by any distance function which is continuous with respect to $d_{\Pi_Y}^{\alpha,\widehat{\mathcal{D}}}(\cdot,\cdot)$.

Definition 3.1.5. *We say that an approximation F is of order $\gamma > 0$ if there exist a positive constant C depending only on α, $\widehat{\mathcal{D}}$ and the one-form defining the RDE* (3.1) *and a real $\delta > 0$, such that*

$$d_{\Pi_Y}^{\alpha,\widehat{\mathcal{D}}}(\mathbf{Y}, F(\mathcal{D})) < C \max_{\mathcal{D}} \omega(t_{l-1},t_l)^\gamma$$

for each partition $\mathcal{D}$ of the interval $[0,T]$ with mesh size at most δ.

3.2 Global convergence theorems

The purpose of this section is to give a sufficient condition on local approximations which implies the convergence of the global approximation constructed by pasting together the local ones. In the later sections, we construct particular local schemes and prove that those schemes satisfy the sufficient condition for a global convergence of certain order.

Let us fix $\alpha_X \geq 1$ and $\alpha_Y > 0$ and introduce the notation $U = T^{(\Pi_Y,\alpha_Y)}(W)$. We assume that α_Y is chosen such that each $\mathcal{A}^N$-multi-index of length 1 is contained in $\mathcal{A}_{\alpha_Y}^{\Pi_Y}$.

In this section, we will use sets of multi-indices satisfying the *tree-like conditions*.

Definition 3.2.1. *We will say that a set of multi-indices $\mathcal{B} \subseteq \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ satisfies the* tree-like conditions, *or in other words $\mathcal{B}$ is a* tree-like set, *if*

(i) $Q \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ *and* $\|Q\| = 1$ *imply* $Q \in \mathcal{B}$

(ii) $Q \in \mathcal{B}$ *implies* $Q- \in \mathcal{B}$

Example 3.2.1. Since the multi-indices of length 1 are contained in $\mathcal{A}_{\alpha_Y}^{\Pi_Y}$, the following set for any integer $i \geq 1$ satisfies the tree-like condition

$$\mathcal{B} = \{R \in \mathcal{A}_{\alpha_Y}^{\Pi_Y} \mid \|R\| \leq i\}.$$

Tree-like sets of this type are extensively used in this paper.

For a tree-like set $\mathcal{B}$, we use the notation

$$U^{\mathcal{B}} = \sum_{R \in \mathcal{B}} W^{\otimes R}.$$

In this paper, the global approximations are constructed by pasting together local approximations. When locally approximating the solution or the solution's truncated signature corresponding to the pair $(s,t) \in \Delta_T$, we consider local approximations taking the truncated signature of the driving noise corresponding to (s,t) as input. The solution corresponding to (s,t) also depends on the initial condition of the equivalent RDE started at s. In the following definition a notation is created for local approximations of this kind.

Definition 3.2.2. *Functions from $\Delta_T \times W \times G\Omega_{\Pi_X}(V)$ to $U^{\mathcal{B}}$ will be referred to as* local approximating functions *or* local approximations *on $U^{\mathcal{B}}$. In this study, unless otherwise stated, the geometric Π_X-rough path argument is the driving noise $\mathbf{X}_{\cdot,\cdot}$. In this case we will omit the argument in the notation.*

Let $(\mathcal{B}_n)_{n>0}$, be an increasing sequence of tree-like sets in $\mathcal{A}_{\alpha_Y}^{\Pi_Y}$, and let $(\widehat{D}_n)_{n>0}$ be a sequence of local approximations corresponding to $(\mathcal{B}_n)_{n>0}$, i.e. the $\widehat{D}_i$ is defined on $U^{\mathcal{B}_i}$ for $i = 1,2,\ldots$. We say, that the local approximation sequence is consistent *if for all $0 < i < j$*

$$\widehat{D}_i\big[(s,t),z\big] = \pi_{U^{\mathcal{B}_i}} \widehat{D}_j\big[(s,t),z\big].$$

Finally, we introduce the notation

$$\widehat{D}^z_{s,t}(y) := y \otimes \widehat{D}\big[(s,t), z + \Pi_W y\big].$$

Given a tree-like set $\mathcal{B} \subseteq \mathcal{A}_{\alpha_Y}^{\Pi_Y}$, and a local-approximation $\widehat{D}$ we introduce a corresponding global approximation of the solution to the RDE (3.1) as follows.

Definition 3.2.3 (Global approximation)**.** *Let $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$ be a partition of $[0,T]$, $\mathcal{B} \subseteq \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ a tree-like set and $\widehat{D}$ a local approximation on $U^{\mathcal{B}}$. The* global approximation generated by *$\widehat{D}$ is defined to be the function $\widehat{\mathbf{Y}} = \widehat{\mathbf{Y}}(\mathcal{D},\mathcal{B}) : \Delta_T^{\mathcal{D}} \to U^{\mathcal{B}}$ given by the following recursion:*

$$\begin{aligned}
\widehat{\mathbf{Y}}_{t_i,t_i} &= \mathbf{1} \\
\widehat{\mathbf{Y}}_{t_0,t_1} &= \widehat{D}\big[(t_0,t_1),\xi\big] \\
\widehat{\mathbf{Y}}_{t_i,t_{i+1}} &= \widehat{D}\big[(t_i,t_{i+1}),\xi + \widehat{\mathbf{Y}}^W_{t_0,t_i}\big] \\
\widehat{\mathbf{Y}}_{t_i,t_j} &= \widehat{\mathbf{Y}}_{t_i,t_{i+1}} \otimes \cdots \otimes \widehat{\mathbf{Y}}_{t_{j-1},t_j} = \widehat{D}^{\xi+\widehat{\mathbf{Y}}^W_{t_0,t_i}}_{t_{j-1},t_j} \circ \cdots \circ \widehat{D}^{\xi+\widehat{\mathbf{Y}}^W_{t_0,t_i}}_{t_i,t_{i+1}}(\mathbf{1})
\end{aligned}$$

for $0 \le i \le j \le n$.

Remark 3.2.1. A local approximation D can be defined by $D\big[(s,t),z\big] = \mathbf{Y}^{U^{\mathcal{B}}}_{s,t}(z,s)$, where $\mathbf{Y}_{\cdot,\cdot}(z,s)$ is the solution to the RDE

$$\mathrm{d}\mathbf{Y}_t(z,s) = f(\mathbf{Y}_t(z,s))\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_s(z,s) = z.$$

Note that, given any partition $\mathcal{D}$ of $[0,T]$, if $z=\xi$, then the global approximation corresponding to D coincides with the solution to the RDE (3.1) restricted to $\Delta_T^{\mathcal{D}}$ and $U^{\mathcal{B}}$.

Let us fix a partition $\mathcal{D}=\{0=t_0<\cdots<t_n=T\}$ together with a tree-like set $\mathcal{B}\subseteq\mathcal{A}_{\alpha_Y}^{\Pi_Y}$ and consider the global approximation $\widehat{\mathbf{Y}}=\widehat{\mathbf{Y}}(\mathcal{D},\mathcal{B})$ corresponding to $\mathcal{D}$ and $\mathcal{B}$. Note that $\mathbf{Y}_{t_i,t_j}^{U^{\mathcal{B}}}-\widehat{\mathbf{Y}}_{t_i,t_j}$ can be represented as follows:

$$
\begin{aligned}
\mathbf{Y}_{t_i,t_j}^{U^{\mathcal{B}}}-\widehat{\mathbf{Y}}_{t_i,t_j} &= D^{\eta}_{t_{j-1},t_j}\circ\cdots\circ D^{\eta}_{t_i,t_{i+1}}(\mathbf{1})-\widehat{D}^{\zeta}_{t_{j-1},t_j}\circ\cdots\circ\widehat{D}^{\zeta}_{t_i,t_{i+1}}(\mathbf{1})\\
&= D^{\eta}_{t_{j-1},t_j}\circ\cdots\circ D^{\eta}_{t_i,t_{i+1}}(\mathbf{1})-D^{\zeta}_{t_{j-1},t_j}\circ\cdots\circ D^{\zeta}_{t_i,t_{i+1}}(\mathbf{1}) \qquad (3.3)\\
&\quad+\sum_{k=i}^{j-1}\Big(D^{\zeta}_{t_{j-1},t_j}\circ\cdots\circ D^{\zeta}_{t_k,t_{k+1}}\circ\widehat{D}^{\zeta}_{t_{k-1},t_k}\circ\cdots\circ\widehat{D}^{\zeta}_{t_i,t_{i+1}}(\mathbf{1})\\
&\qquad\qquad-D^{\zeta}_{t_{j-1},t_j}\circ\cdots\circ\widehat{D}^{\zeta}_{t_k,t_{k+1}}\circ\widehat{D}^{\zeta}_{t_{k-1},t_k}\circ\cdots\circ\widehat{D}^{\zeta}_{t_i,t_{i+1}}(\mathbf{1})\Big)\\
&= \sum_{k=i}^{j-1}\Big(\widehat{D}^{\eta}_{t_{j-1},t_j}\circ\cdots\circ\widehat{D}^{\eta}_{t_k,t_{k+1}}\circ D^{\eta}_{t_{k-1},t_k}\circ\cdots\circ D^{\eta}_{t_i,t_{i+1}}(\mathbf{1}) \qquad (3.4)\\
&\qquad\qquad-\widehat{D}^{\eta}_{t_{j-1},t_j}\circ\cdots\circ\widehat{D}^{\eta}_{t_k,t_{k+1}}\circ\widehat{D}^{\eta}_{t_{k-1},t_k}\circ\cdots\circ D^{\eta}_{t_i,t_{i+1}}(\mathbf{1})\Big)\\
&\quad+\widehat{D}^{\eta}_{t_{j-1},t_j}\circ\cdots\circ\widehat{D}^{\eta}_{t_i,t_{i+1}}(\mathbf{1})-\widehat{D}^{\zeta}_{t_{j-1},t_j}\circ\cdots\circ\widehat{D}^{\zeta}_{t_i,t_{i+1}}(\mathbf{1})
\end{aligned}
$$

where $\eta=\mathbf{Y}^W_{0,t_i}+\xi$ and $\zeta=\widehat{\mathbf{Y}}^W_{0,t_i}+\xi$.

The sums (3.3) and (3.4) represent two different breakdowns of the global error. In both cases, local errors in some sense corresponding to the subintervals of the partition $\mathcal{D}$ are propagated. In (3.3), the local errors propagated through solutions to RDEs started with different initial conditions. In (3.4), the local errors are propagated through the pasted local ODEs. See Figure 3.1 for an illustration.

In either representation of the global error on $[t_i,t_j]$, it is obvious that the W-level of the solution has a special role, since this component of the approximative solution determines the initial error at t_i. Therefore, we firstly focus on approximations of the solution projected on W, i.e. the *increment-level* of the solution.

In the next sections, we derive the global convergence theorems under the following condition.

Condition 3.2.1. *Let $\widehat{D}$ be a local approximation on U (i.e. corresponding to $\mathcal{B}=\mathcal{A}_{\alpha_Y}^{\Pi_Y}$). Furthermore let $\tau>0$ and $\gamma>1$ be real numbers. We will say that $\widehat{D}$ satisfies the (τ,γ)-*condition*, if for each $R\in\mathcal{A}_{\alpha_Y}^{\Pi_Y}$ there exists a positive constant C_R not depending on (s,t) or z such that*

$$\Big\|\pi_R D\big[(s,t),z\big]-\pi_R\widehat{D}\big[(s,t),z\big]\Big\|\le C_R\,\omega(s,t)^{\gamma}$$

for all $(s,t)\in[0,T]$, and such that $\omega(s,t)\le\tau$ and $z\in W$. In case $\tau\ge\omega(0,T)$, we omit τ from the notation.

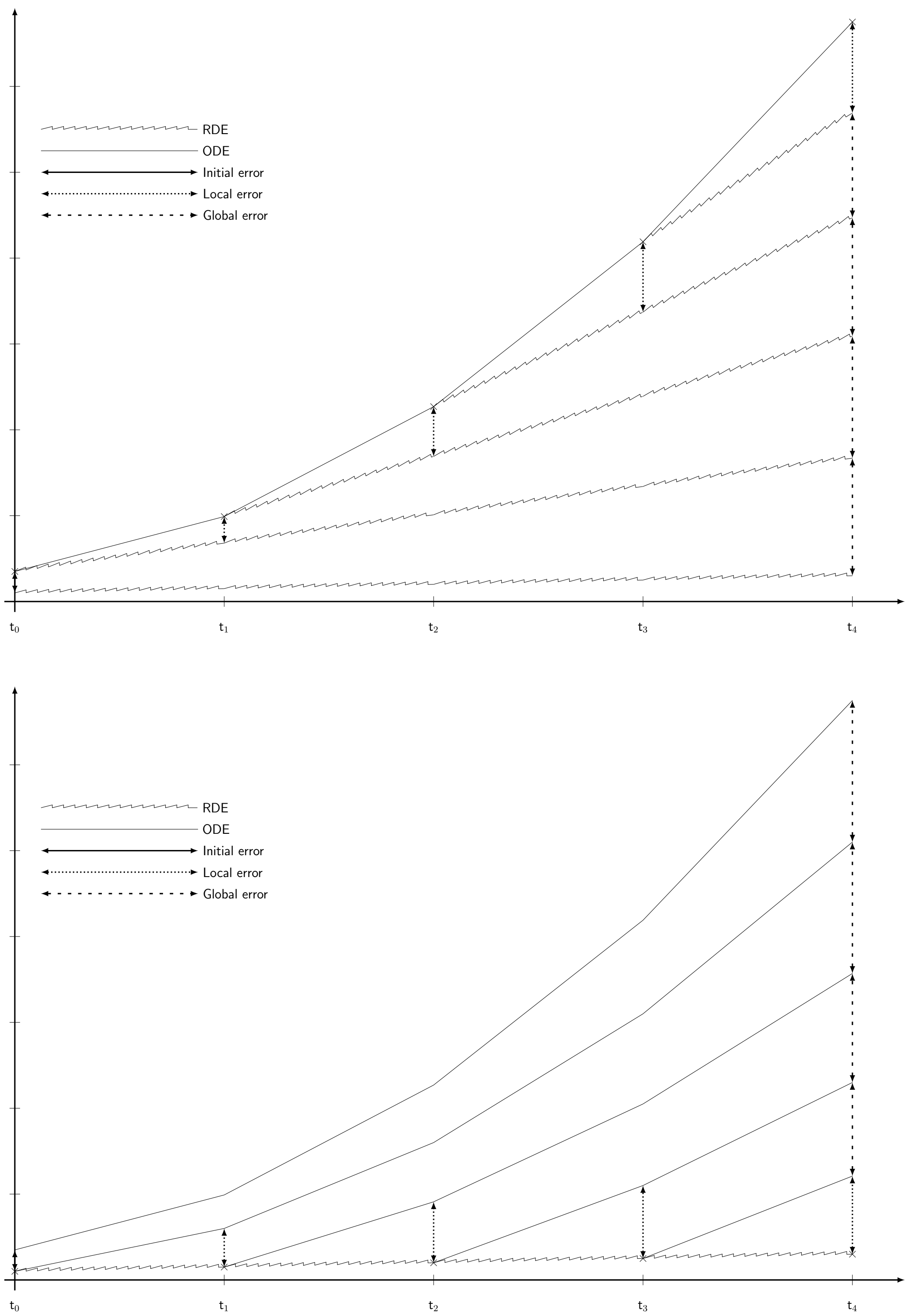

Figure 3.1: Propagation of the local errors through the RDEs and the ODEs respectively

Definition 3.2.4. *We define the sequence of tree-like sets $(\mathcal{B}_n)_{n>0}$ by*

$$\mathcal{B}_i = \{R \in \mathcal{A}_{\alpha_Y}^{\Pi_Y} \mid \|R\| \leq i\}.$$

Assuming that $\widehat{D}$ is a local approximation on U satisfying Condition 3.2.1 for a suitable pair (τ, γ), we define the approximation $\widehat{D}^i$ (and D^i) by restricting $\widehat{D}$ (and D) to $U^{\mathcal{B}_i}$.

Note that there exists a positive integer $n(\Pi_Y, \alpha_Y)$, such that $\mathcal{B}_{n(\Pi_Y,\alpha_Y)} = \mathcal{A}_{\alpha_Y}^{\Pi_Y}$.

3.2.1 Approximating terms corresponding to multi-indices in $\mathcal{B}_1$

The $\|R\| = 1$ case In this section, we consider a local approximation $\widehat{D} := \widehat{D}^1$ on $U^{\mathcal{B}_1} = W$ (or equivalently on $\mathbb{R} \oplus W$) and the corresponding global approximations $\widehat{\mathbf{Y}}_{\cdot,\cdot}(\mathcal{D}, \mathcal{B}_1)$ for admissible partitions of $[0, T]$. We derive a bound on the global error $\left\|\mathbf{Y}_{t_i,t_j}^{U^{\mathcal{B}_1}} - \widehat{\mathbf{Y}}_{t_i,t_j}\right\|$ for all $(t_i, t_j) \in \Delta_T^{\mathcal{D}}$.

Theorem 3.2.1 (Global convergence theorem, the case of $\mathcal{B}_1$). *Let $\widehat{D}$ be an approximation on $U^{\mathcal{B}_1}$ satisfying Condition 3.2.1 for some $\tau > 0$ and $\gamma > 1$. Then there exists a constant C depending only on $\omega(0, T)$, the constants $C_{(i)}$, $i = 1, \ldots, N$ defined in Condition 3.2.1 and polynomially on $\|f\|_{Lip(\Pi_Y,\Gamma)}$, such that for any partition $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$ of $[0, T]$ satisfying $\max_{\mathcal{D}} \omega(t_i, t_{i+1}) < \tau$, the following inequality holds*

$$\left\|\mathbf{Y}_{t_i,t_j}^{(r)} - \widehat{\mathbf{Y}}_{t_i,t_j}^{(r)}\right\| \quad \leq \quad C \sum_{k=0}^{j-1} \omega(t_k, t_{k+1})^{\gamma} \tag{3.5}$$

for $r = 1, \ldots, N$.

Proof. The proof is based on the representation (3.3), i.e. when the local errors are propagated using solutions to RDEs with different initial conditions. Note that at the W-level, the multiplicative property is equivalent to the additivity.

Step 1

Firstly, consider the representation (3.3) with $i = 0$. In this case $\eta = \zeta = \xi$, making the first term zero. Then we consider the differences under the summation.

The kth term is of the form

$$\begin{aligned}
D_{t_{j-1},t_j}^{\xi} \circ \cdots \circ D_{t_k,t_{k+1}}^{\xi} \circ \widehat{D}_{t_{k-1},t_k}^{\xi} \circ \cdots \circ \widehat{D}_{t_0,t_1}^{\xi}(\pi_W \mathbf{1}) \\
- D_{t_{j-1},t_j}^{\xi} \circ \cdots \circ \widehat{D}_{t_k,t_{k+1}}^{\xi} \circ \widehat{D}_{t_{k-1},t_k}^{\xi} \circ \cdots \circ \widehat{D}_{t_0,t_1}^{\xi}(\pi_W \mathbf{1}) \\
= \widehat{\mathbf{Y}}_{0,t_k} + \pi_W D\left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}_{0,t_k}\right] \\
+ \pi_W D\left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}_{0,t_k} + \pi_W D\left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}_{0,t_k}\right]\right] \\
- \left[\widehat{\mathbf{Y}}_{0,t_k} + \widehat{\mathbf{Y}}_{t_k,t_{k+1}} + \pi_W D\left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}_{0,t_k} + \widehat{\mathbf{Y}}_{t_k,t_{k+1}}\right]\right]
\end{aligned}$$

$$= \pi_W D\left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}_{0,t_k}\right] - \widehat{\mathbf{Y}}_{t_k,t_{k+1}} \tag{3.6}$$

$$\begin{aligned} &+\pi_W D\left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}_{0,t_k} + \pi_W D\left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}_{0,t_k}\right]\right] \\ &-\pi_W D\left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}_{0,t_k} + \widehat{\mathbf{Y}}_{t_k,t_{k+1}}\right] \end{aligned} \tag{3.7}$$

The term (3.6) is referred to as the *local error* over $[t_k, t_{k+1}]$ and can be bounded using Condition 3.2.1 as follows:

$$\left\|\pi_W D\left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}_{0,t_k}\right] - \widehat{\mathbf{Y}}_{t_k,t_{k+1}}\right\| \leq K_1 \omega(t_k, t_{k+1})^\gamma \tag{3.8}$$

where K_1 depends only on N and $\max_{1\leq l\leq N} C_{(l)}$ determined by Condition 3.2.1.

We can give a bound on (3.7) using (3.8) and Proposition 2.6.2 in the case of terms corresponding to multi-indices in $\mathcal{B}_1$:

$$\begin{aligned} &\Bigg\|\pi_W D\left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}_{0,t_k} + \pi_W D\left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}_{0,t_k}\right]\right] \\ &\qquad\qquad -\pi_W D\left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}_{0,t_k} + \widehat{\mathbf{Y}}_{t_k,t_{k+1}}\right]\Bigg\| \\ &\leq C_1 \left\|\pi_W D\left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}_{0,t_k}\right] - \widehat{\mathbf{Y}}_{t_k,t_{k+1}}\right\| \frac{\omega(t_{k+1}, t_j)^{\frac{1}{q_r}}}{\beta^N \left(\frac{1}{q_r}\right)!} \end{aligned} \tag{3.9}$$

for $r = 1, \ldots, N$, where C_1 depends only on K_1, $\omega(0, T)$, and polynomially on $\|f\|_{Lip(\Pi_Y,\Gamma)}$.

Summing the terms (3.9) for $k = 0, \ldots, j-1$, we have

$$\left\|\mathbf{Y}^{(r)}_{0,t_j} - \widehat{\mathbf{Y}}^{(r)}_{0,t_j}\right\| \leq K_1 \sum_{k=0}^{j-1} \omega(t_k, t_{k+1})^\gamma \left\{C_1 \frac{\omega(t_{k+1}, t_j)^{\frac{1}{q_r}}}{\beta^N \left(\frac{1}{q_r}\right)!} + 1\right\}.$$

Step 2

In the case when $i > 0$, one can work out a bound on the initial error $\|\eta - \zeta\|$ using Step 1. Then one can derive a bound on the first line of (3.3) using Proposition 2.6.2. Finally, we derive a bound on the terms under the summation in (3.3) using the techniques in Step 1. This is the way we have derived the error bound in Section 3.2.2. However, in the case of terms corresponding to multi-indices in $\mathcal{B}_1$, the resulted error bound is essentially equivalent to the following one.

$$\begin{aligned} &\left\|\mathbf{Y}^{(r)}_{t_i,t_j} - \widehat{\mathbf{Y}}^{(r)}_{t_i,t_j}\right\| = \left\|\mathbf{Y}^{(r)}_{0,t_j} - \mathbf{Y}^{(r)}_{0,t_i} - (\widehat{\mathbf{Y}}^{(r)}_{0,t_j} - \widehat{\mathbf{Y}}^{(r)}_{0,t_i})\right\| \\ &\leq \left\|\mathbf{Y}^{(r)}_{0,t_j} - \widehat{\mathbf{Y}}^{(r)}_{0,t_j}\right\| + \left\|\mathbf{Y}^{(r)}_{0,t_i} - \widehat{\mathbf{Y}}^{(r)}_{0,t_i}\right\| \end{aligned}$$

$$
\begin{aligned}
&\leq K_1 \sum_{k=0}^{i-1} \omega(t_k, t_{k+1})^{\gamma} \left\{ C_1 \frac{\omega(t_{k+1}, t_i)^{\frac{1}{q_r}}}{\beta^N \left(\frac{1}{q_r}\right)!} + 1 \right\} \\
&+K_1 \sum_{k=0}^{j-1} \omega(t_k, t_{k+1})^{\gamma} \left\{ C_1 \frac{\omega(t_{k+1}, t_j)^{\frac{1}{q_r}}}{\beta^N \left(\frac{1}{q_r}\right)!} + 1 \right\} \\
&\leq 2K_1 \sum_{k=0}^{j-1} \omega(t_k, t_{k+1})^{\gamma} \left\{ C_1 \frac{\omega(t_{k+1}, t_j)^{\frac{1}{q_r}}}{\beta^N \left(\frac{1}{q_r}\right)!} + 1 \right\}
\end{aligned}
$$

for $r = 1, \dots, N$. □

3.2.2 Approximating terms corresponding to multi-indices in $\mathcal{B}_m$ for $m > 1$

In this section we consider local approximations on U. Using the results of the previous section, we extend the global error estimate.

Recall Definition 3.2.4 for the sequence of tree-like sets $(\mathcal{B}_n)_{n>0}$ and the corresponding sequences of local approximations $(\widehat{D}_n)_{n>0}$ and $(D_n)_{n>0}$. Note that both the sequences $(\widehat{D}_n)_{n>0}$ and $(D_n)_{n>0}$ are trivially consistent in the sense of Definition 3.2.2.

Note that applying Theorem 3.2.1 on $\widehat{D}_1$, we immediately have a global error estimate on the W-level terms.

Let us define the following hypothesis.

Hypothesis 3.2.1. *Let m be a positive integer. Let $\widehat{D}$ be a local approximation on U satisfying Condition 3.2.1 for some $\tau > 0$ and $\gamma > 1$. Given any partition $\mathcal{D}$ of $[0, T]$ satisfying $\max_{\mathcal{D}} \omega(t_i, t_{i+1}) < \tau$, then for each $R \in \mathcal{B}_m$ and $\{t_i, \cdots, t_j\} \in \mathcal{D}$, the global error*

$$
\left\| \mathbf{Y}^R_{t_i,t_j} - \widehat{\mathbf{Y}}^R_{t_i,t_j} \right\| \leq K_R \sum_{k=0}^{j-1} \omega(t_k, t_{k+1})^{\gamma}
$$

where $\widehat{\mathbf{Y}}^R_{\cdot,\cdot} = \widehat{\mathbf{Y}}^R_{\cdot,\cdot}(\mathcal{D}, \mathcal{B}_m)$ denotes the global approximation corresponding to $\widehat{D}_m$, and K_R depends only on C_R introduced in Condition 3.2.1, on R, γ, τ, $\omega(0, T)$ and polynomially on $\|f\|_{Lip(\Pi_Y, \Gamma)}$.

Note that for $m = 1$, Hypothesis 3.2.1 is true by Theorem 3.2.1. The following theorem verifies the hypothesis for integers greater than 1.

Theorem 3.2.2 (Global convergence theorem, the case of $\mathcal{B}_m$ for $m > 1$)**.** *Suppose that for a local approximation $\widehat{D}$ on U, Hypothesis 3.2.1 is true for an $m > 1$ and for some $\gamma > 1$ and $\tau > 0$. Then the hypothesis is also true for $m + 1$ and for the same γ and τ.*

Proof. The proof is again based on the representation (3.3).

Step 1

Firstly, we derive a bound on $\left\| \mathbf{Y}^{Q}_{0,t_j} - \widehat{\mathbf{Y}}^{Q}_{0,t_j} \right\|$ for $Q \in \mathcal{B}_{m+1}$. In this case,

$$\begin{aligned}
&\pi_Q \Big(D^{\xi}_{t_{j-1},t_j} \circ \cdots \circ D^{\xi}_{t_k,t_{k+1}} \circ \widehat{D}^{\xi}_{t_{k-1},t_k} \circ \cdots \circ \widehat{D}^{\xi}_{t_0,t_1} (\pi_W \mathbf{1}) \\
&\qquad\qquad - D^{\xi}_{t_{j-1},t_j} \circ \cdots \circ \widehat{D}^{\xi}_{t_k,t_{k+1}} \circ \widehat{D}^{\xi}_{t_{k-1},t_k} \circ \cdots \circ \widehat{D}^{\xi}_{t_0,t_1} (\pi_W \mathbf{1}) \Big) \\
&= \sum_{Q_1 * Q_2 * Q_3 = Q} \widehat{\mathbf{Y}}^{Q_1}_{0,t_k} \otimes \pi_{Q_2} D \left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} \right] \\
&\qquad\qquad \otimes \pi_{Q_3} D \left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} + \pi_W D \left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} \right] \right] \\
&\qquad - \widehat{\mathbf{Y}}^{Q_1}_{0,t_k} \otimes \widehat{\mathbf{Y}}^{Q_2}_{t_k,t_{k+1}} \otimes \pi_{Q_3} D \left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} + \widehat{\mathbf{Y}}^{W}_{t_k,t_{k+1}} \right].
\end{aligned}$$

Furthermore, for $Q_1, Q_2, Q_3 \in \mathcal{A}^{\Pi_Y}_{\alpha_Y}$, such that $Q_1 * Q_2 * Q_3 = Q \in \mathcal{B}_{m+1}$, we have

$$\begin{aligned}
&\widehat{\mathbf{Y}}^{Q_1}_{0,t_k} \otimes \pi_{Q_2} D \left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} \right] \\
&\qquad\qquad \otimes \pi_{Q_3} D \left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} + \pi_W D \left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} \right] \right] \\
&- \widehat{\mathbf{Y}}^{Q_1}_{0,t_k} \otimes \widehat{\mathbf{Y}}^{Q_2}_{t_k,t_{k+1}} \otimes \pi_{Q_3} D \left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} + \widehat{\mathbf{Y}}^{W}_{t_k,t_{k+1}} \right] \\
&= \widehat{\mathbf{Y}}^{Q_1}_{0,t_k} \otimes \left\{ \pi_{Q_2} D \left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} \right] - \widehat{\mathbf{Y}}^{Q_2}_{t_k,t_{k+1}} \right\} \\
&\qquad\qquad \otimes \pi_{Q_3} D \left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} + \pi_W D \left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} \right] \right] \\
&\quad + \widehat{\mathbf{Y}}^{Q_1}_{0,t_k} \otimes \widehat{\mathbf{Y}}^{Q_2}_{t_k,t_{k+1}} \otimes \left\{ \pi_{Q_3} D \left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} + \pi_W D \left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} \right] \right] \right. \\
&\qquad\qquad\qquad \left. - \pi_{Q_3} D \left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}^{W}_{0,t_k} + \widehat{\mathbf{Y}}^{W}_{t_k,t_{k+1}} \right] \right\}.
\end{aligned} \tag{3.10}$$

Note that if $Q_1 = Q$, then $Q_2 = Q_3 = \epsilon$, implying that the expression (3.10) equals zero. If $Q_3 = Q$, then $Q_1 = Q_2 = \epsilon$, implying that the first triple tensor product in (3.10) is zero. Finally, if $Q_1 * Q_2 = Q$, then $Q_3 = \epsilon$, making the second triple tensor product equal to zero.

Consider the first triple tensor product of (3.10):

$$\begin{aligned}
&\left\| \widehat{\mathbf{Y}}^{Q_1}_{0,t_k} \otimes \left\{ \pi_{Q_2} D\left[(t_k,t_{k+1}),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}\right] - \widehat{\mathbf{Y}}^{Q_2}_{t_k,t_{k+1}} \right\} \right. \\
&\qquad\qquad \left. \otimes \pi_{Q_3} D\left[(t_{k+1},t_j),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}+\pi_W D\left[(t_k,t_{k+1}),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}\right]\right] \right\| \\
&\le \left\| \widehat{\mathbf{Y}}^{Q_1}_{0,t_k} \right\| \cdot \left\| \pi_{Q_2} D\left[(t_k,t_{k+1}),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}\right] - \widehat{\mathbf{Y}}^{Q_2}_{t_k,t_{k+1}} \right\| \\
&\qquad\qquad \cdot \left\| \pi_{Q_3} D\left[(t_{k+1},t_j),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}+\pi_W D\left[(t_k,t_{k+1}),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}\right]\right] \right\| \\
&\le \left\{ C_1 \frac{\omega(0,t_k)^{deg_{\Pi_y}(Q_1)}}{\beta^N \Gamma_{\Pi_Y}(Q_1)} + K_{Q_1} \sum_{l=0}^{k-1} \omega(t_l,t_{l+1})^{\gamma} \right\} \left\{ C_{Q_2} \omega(t_k,t_{k+1})^{\gamma} \right\} \\
&\qquad\qquad \cdot \left\{ C_1 \frac{\omega(t_{k+1},t_j)^{deg_{\Pi_y}(Q_3)}}{\beta^N \Gamma_{\Pi_Y}(Q_3)} \right\}
\end{aligned} \tag{3.11}$$

where

(i) C_1 exists by Proposition 2.6.1,

(ii) K_{Q_1} exists by Hypothesis 3.2.1 (we can assume that $Q_1 \neq Q$, i.e. $Q_1 \in \mathcal{B}_m$, otherwise the triple product equals zero),

(iii) C_{Q_2} exists by Condition 3.2.1,

C_1 and K_{Q_1} both depend only on Q_1, $\omega(0,T)$ and polynomially on $\|f\|_{Lip(\Pi_Y,\Gamma)}$, and C_{Q_2} depends only on Q_2, $\omega(0,T)$ and polynomially on $\|f\|_{Lip(\Pi_Y,\Gamma)}$.

Now, consider the second triple tensor product:

$$\begin{aligned}
&\left\| \widehat{\mathbf{Y}}^{Q_1}_{0,t_k} \otimes \widehat{\mathbf{Y}}^{Q_2}_{t_k,t_{k+1}} \otimes \left\{ \pi_{Q_3} D\left[(t_{k+1},t_j),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}+\pi_W D\left[(t_k,t_{k+1}),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}\right]\right] \right. \right. \\
&\qquad\qquad \left. \left. -\pi_{Q_3} D\left[(t_{k+1},t_j),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}+\widehat{\mathbf{Y}}^{W}_{t_k,t_{k+1}}\right] \right\} \right\| \\
&= \left\| \widehat{\mathbf{Y}}^{Q_1 * Q_2}_{0,t_{k+1}} \otimes \left\{ \pi_{Q_3} D\left[(t_{k+1},t_j),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}+\pi_W D\left[(t_k,t_{k+1}),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}\right]\right] \right. \right. \\
&\qquad\qquad \left. \left. -\pi_{Q_3} D\left[(t_{k+1},t_j),\xi+\widehat{\mathbf{Y}}^{W}_{0,t_k}+\widehat{\mathbf{Y}}^{W}_{t_k,t_{k+1}}\right] \right\} \right\|
\end{aligned}$$

$$\le \left\| \widehat{\mathbf{Y}}^{Q_1 * Q_2}_{0,t_{k+1}} \right\| \cdot \left\| \pi_{Q_3} D\left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}^W_{0,t_k} + \pi_W D\left[(t_k, t_{k+1}), \xi + \widehat{\mathbf{Y}}^W_{0,t_k}\right]\right] \right.$$
$$\left. - \pi_{Q_3} D\left[(t_{k+1}, t_j), \xi + \widehat{\mathbf{Y}}^W_{0,t_k} + \widehat{\mathbf{Y}}^W_{t_k,t_{k+1}}\right] \right\|$$
$$\le \left\{ C_1 \frac{\omega(0, t_{k+1})^{deg_{\Pi_y}(Q_1 * Q_2)}}{\beta^N \Gamma_{\Pi_Y}(Q_1 * Q_2)} + K_{Q_1 * Q_2} \sum_{l=0}^{k} \omega(t_l, t_{l+1})^\gamma \right\}$$
$$\cdot \left\{ L_{Q_3} C_W \omega(t_k, t_{k+1})^\gamma \frac{\omega(t_{k+1}, t_j)^{deg_{\Pi_y}(Q_3)}}{\beta^N \Gamma_{\Pi_Y}(Q_3)} \right\} \tag{3.12}$$

where C_1 is the same as above, C_W and $K_{Q_1 * Q_2}$ exist with analogy to the (3.11) case (we can assume that $Q_1 * Q_2 \neq Q$ otherwise the triple tensor product would be zero), and furthermore L_{Q_3} exists by Theorem 3.2.1 (the case of $\mathcal{B}_1$) and Proposition 2.6.2, and depends only on Q_3, $\omega(0, T)$ and polynomially on $\|f\|_{Lip(\Pi_Y, \Gamma)}$.

Step 2

Summing the terms (3.11) and (3.12) for all multi-index triplets Q_1, Q_2, Q_3 such that $Q_1 * Q_2 * Q_3 = Q$ and then for all $k \in \{0, 1, \ldots, j-1\}$, we get

$$\left\| \mathbf{Y}^Q_{0,t_j} - \widehat{\mathbf{Y}}^Q_{0,t_j} \right\| \le \mathbf{C}_Q \sum_{k=0}^{j-1} \omega(t_k, t_{k+1})^\gamma$$

where $\mathbf{C}_Q$ depends only on Q, γ, $\omega(0, T)$ and polynomially on $\|f\|_{Lip(\Pi_Y, \Gamma)}$.

Step 3

Finally, we consider $\left\| \mathbf{Y}^Q_{t_i,t_j} - \widehat{\mathbf{Y}}^Q_{t_i,t_j} \right\|$ in the case when $0 < i \le j$. Considering the representation (3.3) of the global error, the terms under the summation can be bounded independently from ζ and with analogy to the previous steps. The only difference arises due to the first two terms in (3.3) (outside the summation). The difference of these two terms can be bounded by using a bound on $\eta - \zeta$ provided by the previous steps and using Proposition 2.6.2 to propagate the initial error up to t_j. □

Corollary 3.2.1. *Assume that $\widehat{D}$ is a local approximation on U satisfying Condition 3.2.1 for some $\tau > 0$ and $\gamma > 1$. Let us fix a partition $\widehat{\mathcal{D}}$ of $[0, T]$. Let $(\mathcal{D}_n)_{n>0}$ be a sequence of partitions of $[0, T]$ for each $n > 0$ satisfying $\max_{\mathcal{D}_n}(t_i, t_{i+1}) < \tau$ and $\widehat{\mathcal{D}} \subseteq \mathcal{D}_n$ for all $n > 0$. Moreover the mesh size of $\mathcal{D}_n$ tends to zero as n tends to infinity. Let $\mathbf{Y}_{\cdot,\cdot}(\mathcal{D}_n, \mathcal{A}^{\Pi_Y}_{\alpha_Y})$ denote the corresponding sequence of global approximation of $\mathbf{Y}_{\cdot,\cdot}$ generated by $\widehat{D}$.*

Recall the definition of the following distance function:

$$d^{\alpha_Y, \widehat{\mathcal{D}}}_{\Pi_Y}\left(\mathbf{Y}, \widehat{\mathbf{Y}}(\mathcal{D}_n, \mathcal{A}^{\Pi_Y}_{\alpha_Y})\right) := \max_{\substack{R \in \mathcal{A}^{\Pi_Y}_{\alpha_Y} \\ \overline{\mathcal{D}} \subseteq \widehat{\mathcal{D}}}} \left(\sum_{\overline{\mathcal{D}}} \left\| \mathbf{Y}^R_{t_{l-1},t_l} - \widehat{\mathbf{Y}}^R_{t_{l-1},t_l}(\mathcal{D}_n, \mathcal{A}^{\Pi_Y}_{\alpha_Y}) \right\|^{1/deg_{\Pi_Y}(R)} \right)^{deg_{\Pi_Y}(R)} . \tag{3.13}$$

Then the distance (3.13) *tends to zero as n tends to infinity. In particular, the order of convergence is $\gamma - 1$, i.e. there exists a constant C, not depending on the partitions $\mathcal{D}_n$ or on n, such that*

$$d^{\alpha_Y,\widehat{\mathcal{D}}}_{\Pi_Y}\big(\mathbf{Y}, \widehat{\mathbf{Y}}(\mathcal{D}_n, \mathcal{A}^{\Pi_Y}_{\alpha_Y})\big) \leq C \max_{\mathcal{D}_n} \omega(t_{l-1}, t_l)^{\gamma-1}. \tag{3.14}$$

Proof. By the Global Convergence Theorem (3.2.2), there exists a constant K not depending on n or $\mathcal{D}_n$, such that

$$d^{\alpha_Y,\widehat{\mathcal{D}}}_{\Pi_Y}\big(\mathbf{Y}, \widehat{\mathbf{Y}}(\mathcal{D}_n, \mathcal{A}^{\Pi_Y}_{\alpha_Y})\big) \leq K \sum_{\mathcal{D}_n} \omega(t_{l-1}, t_l)^{\gamma}.$$

By the super-additive property of ω, we have the following inequality

$$\sum_{\mathcal{D}_n} \omega(t_{l-1}, t_l)^{\gamma} \leq \left(\sum_{\mathcal{D}_n} \omega(t_{l-1}, t_l)\right) \max_{\mathcal{D}_n} \omega(t_{l-1}, t_l)^{\gamma-1} \leq \omega(0,T) \max_{\mathcal{D}_n} \omega(t_{l-1}, t_l)^{\gamma-1}.$$

Furthermore, the fact that ω is continuous and tends to zero on the diagonal completes the proof. □

Remark 3.2.2. The global error bound (3.14) is derived without considering the mutual cancellation of the local errors propagated through the RDE solutions. In Chapter (4), we will show some examples, in which the (τ, γ)-condition on the local approximation implies a global convergence order higher than $\gamma - 1$.

3.3 Local behavior

In this section we set the preliminaries for the Log-signature Theorems 3.4.1 and 3.4.2 by investigating the behavior of the solution to a rough differential equation on short time intervals. This analysis is valid in finite dimensions, and therefore we assume that the Banach spaces V and W are in fact $\mathbb{R}^k$ and $\mathbb{R}^N$ respectively.

3.3.1 Algebraic setting

In this section the concept of Lyons [25] and Lyons & Victoir [28] is followed and adapted to the $T^{(\Pi_X,\alpha_X)}(V)$-valued geometric rough paths.

Definition 3.3.1. *Let U be a Banach space and $T((U))$ be the space of formal series of tensors of U. Recall that $T((U))$ is a non-commutative algebra with respect to $+$, $\otimes$ and unit element $\mathbf{1}$. We define the* power function *recursively by*

$$\begin{aligned} \mathbf{a}^{\otimes 0} &:= \mathbf{1} \\ \mathbf{a}^{\otimes 1} &:= \mathbf{a} \\ \mathbf{a}^{\otimes i} &:= \mathbf{a}^{\otimes i-1} \otimes \mathbf{a} \end{aligned}$$

for $\mathbf{a} \in T((U))$.

We define the exponential function $\exp : T((U)) \to T((U))$ *by the power series*

$$\exp(\mathbf{a}) := \sum_{i \geq 0} \frac{1}{i!} \mathbf{a}^{\otimes i}$$

for $\mathbf{a} \in T((U))$.

In the case when $\mathbf{a}_\epsilon \neq 0$ *in the representation* $\mathbf{a} = a_\epsilon(\mathbf{1} + \mathbf{b})$*, we define the* multiplicative inverse *by the power series*

$$\mathbf{a}^{\otimes(-1)} := \frac{\sum_{i \geq 0} (-1)^i \mathbf{b}^{\otimes i}}{a_\epsilon},$$

and in the case when $\mathbf{a}_\epsilon > 0$*, we define the* log function *by*

$$\log(\mathbf{a}) := \log(a_\epsilon) + \sum_{i \geq 1} \frac{\mathbf{b}^{\otimes i}}{i} (-1)^{i-1}.$$

Definition 3.3.2. *Let* U *be a Banach space and* $T((U))$ *be the space of formal series of tensors of* U*. The* Lie bracket $[\cdot, \cdot] : T((U)) \times T((U)) \to T((U))$ *is defined by*

$$[\mathbf{a}, \mathbf{b}] := \mathbf{a} \otimes \mathbf{b} - \mathbf{b} \otimes \mathbf{a}$$

for $\mathbf{a}, \mathbf{b} \in T((U))$.

Let the set $\mathcal{L}_i(U)$ *be defined by the recursion*

$$\begin{aligned} \mathcal{L}_1(U) &:= U \\ \mathcal{L}_2(U) &:= [U, U] := \text{Span}\left\{ [\mathbf{a}, \mathbf{b}] \mid \mathbf{a}, \mathbf{b} \in U \right\} \\ \mathcal{L}_{i+1}(U) &:= [\mathcal{L}_i(U), U] := \text{Span}\left\{ [\mathbf{a}, \mathbf{b}] \mid \mathbf{a} \in \mathcal{L}_i(U), \mathbf{b} \in U \right\}. \end{aligned}$$

The Lie algebra $\mathcal{L}(U)$ *is defined by*

$$\mathcal{L}(U) = \text{Span}\left\{ \mathcal{L}_i(U) | i \in \mathbb{N} \right\}.$$

The elements of $\mathcal{L}(U)$ *are referred to as* Lie polynomials*. The closure of* $\mathcal{L}(U)$*, i.e. the space of formal Lie series, is denoted by* $\mathcal{L}((U))$*.*

Definition 3.3.3. *Let* U *be a Banach space such that* $U = U^1 \oplus \cdots \oplus U^k$ *for some Banach spaces* $U^1, \ldots, U^k$*. Let* $\Pi = (p_1, \ldots, p_k)$ *be a positive real k-tuple and* $\alpha > 1$ *a real number. We recall that the algebra* $T^{(\Pi,\alpha)}(U)$ *is identified with* $\bigoplus_{R \in \mathcal{A}_\alpha^\Pi} V^R$ *and the multiplication in* $T^{(\Pi,\alpha)}(U)$ *is denoted by* $\otimes$*. The* power, exponential, multiplicative inverse and logarithm functions *in* $T^{(\Pi,\alpha)}(U)$ *are defined analogously to Definition 3.3.1. To distinguish between the functions on the different domains, we will use the notation*

$$\begin{aligned} \exp^{(\Pi,\alpha)} &: T^{(\Pi,\alpha)}(U) \to T^{(\Pi,\alpha)}(U) \\ \log^{(\Pi,\alpha)} &: T^{(\Pi,\alpha)}(U) \to T^{(\Pi,\alpha)}(U). \end{aligned}$$

The Lie bracket $[\cdot,\cdot] : T^{(\Pi,\alpha)}(U) \times T^{(\Pi,\alpha)}(U) \to T^{(\Pi,\alpha)}(U)$ *and the* Lie algebra $\mathcal{L}^{(\Pi,\alpha)}(U)$ *are defined with respect to the multiplication* $\otimes : T^{(\Pi,\alpha)}(U) \times T^{(\Pi,\alpha)}(U) \to T^{(\Pi,\alpha)}(U)$ *by analogy to Definition 3.3.2.*

Definition 3.3.4. *Let* $\varepsilon_1, \dots, \varepsilon_d$ *denote a normalized basis of* $V = \mathbb{R}^d$. *Letting* $R = (r_1, \dots, r_l) \in \mathcal{A}^{\Pi_X}$, *we introduce the notation*

$$\varepsilon_R = \varepsilon_{r_1} \otimes \cdots \otimes \varepsilon_{r_l}.$$

Note that for any positive real α_X, $T^{(\Pi_X,\alpha_X)}(V)$ *is spanned by the set* $\left\{\varepsilon_R,\ R \in \mathcal{A}^{\Pi_X}_{\alpha_X}\right\}$.

Example 3.3.1. Let $\widehat{\mathbf{X}} : \Delta_T \to T^{(\Pi_X,\alpha_X)}(V)$ be a (1)-rough path, and let

$$X_t = (X^1_t, \dots, X^d_t)^T = \widehat{\mathbf{X}}^V_{0,t}.$$

Then the signature of $\widehat{\mathbf{X}}$ corresponding to the interval $[s,t]$ is

$$\widehat{\mathbf{X}}_{s,t} = \sum_{\substack{(r_1,\dots,r_l)\in\mathcal{A}^{\Pi_X}_{\alpha_X}\\ l\in\mathbb{N}}} \left(\int_{s<u_1\cdots<u_n<t} \mathrm{d}X^{r_1}_{u_1}\cdots \mathrm{d}X^{r_l}_{u_l}\right) \varepsilon_{r_1}\otimes\cdots\otimes\varepsilon_{r_l} =: \sum_{R\in\mathcal{A}^{\Pi_X}_{\alpha_X}} X^R_{s,t}\varepsilon_R. \tag{3.15}$$

where the integral is in the Riemann sense. The multiplicative property of $\widehat{\mathbf{X}}$ is due to Chen [5] (see also Theorem 4.4 of Lyons & Victoir [25]).

In the rest of the chapter, we will distinguish between the tensor element $\widehat{\mathbf{X}}^R_{s,t} \in V^R \subset T((V))$ and its real coefficient $X^R_{s,t} \in \mathbb{R}$ as it is defined in equation (3.15).

Definition 3.3.5. *Let* $\widehat{\mathbf{X}} : \Delta_T \to T^{(\Pi_X,\alpha_X)}(V)$ *be a* Π_X*-rough path (or more precisely, the extension of a* Π_X*-rough path). The* log-signature *of* $\widehat{\mathbf{X}}$ *corresponding to the interval* $[s,t] \subseteq [0,T]$ *is defined by*

$$L^{(\Pi_X,\alpha_X)}_{s,t}(\widehat{\mathbf{X}}) := \log^{(\Pi_X,\alpha_X)}(\widehat{\mathbf{X}}_{s,t}).$$

The local log-signature *of* $\widehat{\mathbf{X}}$ *corresponding to the partition* $\mathcal{D}$ *of* $[0,T]$ *is defined accordingly.*

The following theorem and in particular (3.16) is crucial for the derivation of the numerical methods.

Theorem 3.3.1 (Chen's Theorem)**.** *Let* $\widehat{\mathbf{X}} : \Delta_T \to T^{(\Pi_X,\alpha_X)}(V)$ *be a geometric* Π_X*-rough path controlled by a control function* ω. *The log-signature* $L^{(\Pi_X,\alpha_X)}_{s,t}(\widehat{\mathbf{X}})$ *corresponding to the interval* $[s,t]$ *lies in* $\mathcal{L}^{(\Pi,\alpha)}(V)$.

Furthermore, the log-signature satisfies

$$\exp^{(\Pi_X,\alpha_X)}\left\{L^{(\Pi_X,\alpha_X)}_{s,t}(\widehat{\mathbf{X}})\right\} = \pi_{\Pi_X,\alpha_X}\widehat{\mathbf{X}}_{s,t}. \tag{3.16}$$

Finally, if $L \in \mathcal{L}^{(\Pi,\alpha)}(U)$ *is a Lie polynomial, then there exists a path* $X : [0,T] \to V$ *of bounded variation, such that*

$$\log^{(\Pi_X,\alpha_X)}(\widehat{\mathbf{X}}_{s,t}) = L$$

where $\widehat{\mathbf{X}}_{s,t}$ *is defined as in Example 3.3.1.*

Proof. If $\widehat{\mathbf{X}}$ is a (1)-rough path as in Example 3.3.1 and $\alpha_X = \infty$, the assertion is proved in [17], [25], although the original proof can be traced back to [5]. Since $T^{(\Pi,\alpha)}(U)$ is identified with $\bigoplus_{R\in\mathcal{A}^{\Pi}_{\alpha}} V^R$, the $\alpha_X < \infty$ case is implied by

$$L^{(\Pi_X,\alpha_X)}_{s,t}(\widehat{\mathbf{X}}) = \pi_{T^{(\Pi_X,\alpha_X)}(V)} L_{s,t}(\widehat{\mathbf{X}}).$$

In the case when $\widehat{\mathbf{X}}$ is a Π_X-rough path, it can be approximated by (1)-rough paths in the $d^{\alpha_X}_{\Pi_X}$-distance. In the same topology, the $\log^{(\Pi_X,\alpha_X)}$ function is continuous and $\mathcal{L}^{(\Pi,\alpha)}(U)$ is closed. For $\alpha > \alpha_X$,

$$\pi_{\Pi,\alpha_X}\left(\log^{(\Pi_X,\alpha)}(\widehat{\mathbf{X}}_{s,t}) - \log^{(\Pi_X,\alpha_X)}(\widehat{\mathbf{X}}_{s,t})\right) = 0.$$

This completes the proof of the first statement. The reader is referred to Theorem 4.4 of [28] for the proof of the last statement. □

We also define an algebra homomorphism as follows.

Definition 3.3.6. *Let $\mathcal{W} = \{W_1, \ldots, W_k\}$ be a set of vectors fields on a Banach space U. The algebra homomorphism $\Phi_{\mathcal{W}}$ from $T((V))$ into the space of differential operators is generated by $\Phi_{\mathcal{W}}(\epsilon) = Id$, $\Phi_{\mathcal{W}}(\varepsilon_i) = W_i$, $i = 1, \ldots, k$.*

For the multi-index $R = (r_1, \ldots, r_l) \in \mathcal{A}^{\Pi_X}$, we will use the notation

$$\Phi_{\mathcal{W}}(\varepsilon_R) = W_{r_1} \circ \cdots \circ W_{r_l} =: W_R.$$

Remark 3.3.1. Note that $\Phi_{\mathcal{W}}$ assigns first order differential operators to the elements of $\mathcal{L}(V)$.

Definition 3.3.7. *Let Π be a k-tuple and α a positive real number. A set $\mathcal{W}$ of vector fields $\{W_1, \ldots, W_k\}$ is called* nilpotent of Π-degree α *if any (possibly multiple nested) non-trivial composition of Lie brackets is zero when applied on the vector fields $W_{r_1}, \ldots, W_{r_l}$, where*

$$deg_{\Pi}\{(r_1, \ldots, r_l)\} > \alpha.$$

The set $\mathcal{W}$ is called nilpotent, *if there exist a k-tuple Π and a positive real α, such that $\mathcal{W}$ is nilpotent of Π-degree α.*

3.3.2 Motivation

Let us recall that by Theorem 2.5.1, the Itô map is continuous with respect to the $d_{\Pi_X} - d_{\Pi_Y}$ topology. Remark 2.1.6 implies, that for any $\varepsilon > 0$ there exists a positive δ and a Π_X-rough path $\widehat{\mathbf{X}}$ such that $d_{\Pi_X}(\mathbf{X}, \widehat{\mathbf{X}}) < \delta$ and the solution $\widehat{\mathbf{Y}}$ of the RDE

$$\mathrm{d}\widehat{\mathbf{Y}}_t = f(\widehat{\mathbf{Y}}_t)\mathrm{d}\widehat{\mathbf{X}}_t,\ \widehat{\mathbf{Y}}_0 = \xi \tag{3.17}$$

satisfies the inequality

$$d_{\Pi_Y}(\mathbf{Y}, \widehat{\mathbf{Y}}) < \varepsilon. \tag{3.18}$$

In particular, the inequality (3.18) holds if $\widehat{\mathbf{X}}$ is a (1)-rough path. In this case, the RDE (3.17) is in fact an ordinary differential equation (ODE) in W, equivalent to

$$\mathrm{d}\widehat{\mathbf{Y}}^W_{0,t} = \sum_{i=1}^{k} V_i(\widehat{\mathbf{Y}}^W_{0,t} + \xi)\mathrm{d}X^{(i)}_{0,t},\ \widehat{\mathbf{Y}}^W_{0,0} = 0 \tag{3.19}$$

or

$$\mathrm{d}\widehat{\mathbf{Y}}^W_{t} = \sum_{i=1}^{k} V_i(\widehat{\mathbf{Y}}^W_{t})\mathrm{d}X^{(i)}_{0,t},\ \widehat{\mathbf{Y}}^W_{0} = \xi \tag{3.20}$$

in the sense that the rough path $\widehat{\mathbf{Y}}$ with values in U is determined by ξ and the W-level component $t \mapsto \widehat{\mathbf{Y}}^W_{0,t} \in W = \mathbb{R}^N$ since it is itself a (1)-rough path, and furthermore $\widehat{\mathbf{Y}}$ is also determined by $t \mapsto \widehat{\mathbf{Y}}^W_t = \xi + \widehat{\mathbf{Y}}^W_{0,t}$.

For the following example, let us assume that this is the case; i.e. $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$ are (1)-rough paths. Furthermore, let us assume that the set $\widehat{\mathcal{W}}$ of the vector fields $\widehat{W}_1, \ldots, \widehat{W}_k$ on $\mathbb{R}^N$ defined by

$$\widehat{W}_i(y) = \sum_{j=1}^{N} V_i^j(y)\frac{\partial}{\partial y^j} \tag{3.21}$$

is nilpotent of Π_X-degree γ, for some positive γ. This implies that

$$\Phi_{\widehat{\mathcal{W}}}\left(\log^{(\Pi_X,\gamma)}(\widehat{\mathbf{X}}_{s,t})\right) = \Phi_{\widehat{\mathcal{W}}}\left(\log^{(\Pi_X,\hat{\gamma})}(\widehat{\mathbf{X}}_{s,t})\right),\ \forall (s,t) \in \Delta_t, \forall \hat{\gamma} \geq \gamma.$$

Finally, let us assume, that the ODE

$$\mathrm{d}\hat{y}_s = \Phi_{\widehat{\mathcal{W}}}\left(\log^{(\Pi_X,\gamma)}(\widehat{\mathbf{X}}_{0,t})\right)(\hat{y}_s)\mathrm{d}s,\ \hat{y}_0 = \xi \tag{3.22}$$

possesses a unique solution on the interval $[0,1]$.

The following result is well known, and for example is implied by the results of Strichartz [35].

Lemma 3.3.1. *The solution to the non-autonomous ODE* (3.20) $\widehat{\mathbf{Y}}_t$ *at* t *coincides with the solution to the autonomous ODE* (3.22) $\hat{y}_1$ *at* 1.

As $\widehat{\mathbf{X}}$ tends to $\mathbf{X}$ in the d_{Π_X}-topology, $\Phi_{\widehat{\mathcal{W}}}\left(\log^{(\Pi_X,\gamma)}(\widehat{\mathbf{X}}_{0,t})\right)$ tends to $\Phi_{\widehat{\mathcal{W}}}\left(\log^{(\Pi_X,\gamma)}(\mathbf{X}_{0,t})\right)$ uniformly in a compact set including $\mathbf{Y}^W_{0,s}$ for all $s \in [0,t]$ and y_u for all $u \in [0,1]$, where y is the solution to the ODE

$$\mathrm{d}y_s = \Phi_{\widehat{\mathcal{W}}}\left(\log^{(\Pi_X,\gamma)}(\mathbf{X}_{0,t})\right)(y_s)\mathrm{d}s,\ y_0 = \xi. \tag{3.23}$$

Hence $\xi + \mathbf{Y}^W_{0,t}$ at t coincides with y_1 at 1.

In this particular case, the γ-truncated log-signature of the driving noise corresponding to the interval $[0,t]$ determines $\mathbf{Y}^W_{0,t}$. This is an important fact providing us a representation of the solution on which one can base a numerical scheme (i.e. by approximating the solution to the ODE (3.23)).

Example 3.3.2. Let us regard the following Stratonovich SDE

$$\mathrm{d}\xi_t = (r - \sigma^2/2)\xi_t \mathrm{d}t + \sigma\xi_t \circ \mathrm{d}W_t.$$

The solution is known as

$$\xi_t = \xi_s \exp\left\{(r - \sigma^2/2)(t-s) + \sigma(W_t - W_s)\right\}.$$

The time and Brownian increment determine the solution. For almost all Brownian path, the above SDE can be regarded as RDE (ref.: [26]). For a fixed Brownian path ω, regard the ODE

$$\mathrm{d}y_u = \left\{(r - \sigma^2/2)(t-s) + \sigma(W_t(\omega) - W_s(\omega))\right\} y_u \mathrm{d}u,\ y_0 = \xi_s.$$

Note that $y_1 = \xi_t(\omega)$, i.e. by solving the ODE, we find the exact value of $\xi_t(\omega)$. However for intermediate time $u \in (0,1)$, y_u and ξ_{u^*} for $u^* = u(t-s) + s$ can be any far depending on the actual RDE solution path. Similarly, the local signature of the path $y_\cdot$ can differ significantly from the local log signature of the RDE solution path $\xi_\cdot$.

This example highlights the following general facts:

(i) y_u for $u \in (0,1)$ provides little information on $\mathbf{Y}^W_{0,s}$ for $s \in (0,t)$.

(ii) The iterated integrals of y provide little information on the higher level terms of $\mathbf{Y}_{0,t}$.

The first problem is not going to be relaxed, however it is somewhat natural. I.e. there could be many driving signals with the same truncated log-signature corresponding to the interval $[0,t]$ determining different solutions to the RDE, so no different behavior from the above first point is expected.

We address the second problem by replacing the ODE (3.19) with an extended system in the next section. However, in the extended case, the set of vector fields in general are not nilpotent.

3.3.3 The rough-Taylor expansion

In the following Lemma, we replace the ODE (3.19) of the previous section by an extended system of ODEs. In the rest of the chapter unless otherwise indicated, we restrict the letter U to denote the space $T^{(\Pi_Y,\alpha_Y)}(W)$.

Lemma 3.3.2. *Let $\widehat{\mathbf{X}}$ be of bounded variation, i.e. $\widehat{\mathbf{X}}$ is a (1)-rough path. Then the RDE (3.17) is equivalent to the following ODE in U:*

$$\mathrm{d}\widehat{\mathbf{Y}}_{0,t} = \sum_{i=1}^{k} \widehat{V}_i^{\xi}(\widehat{\mathbf{Y}}_{0,t})\mathrm{d}X^{(i)}_{0,t},\ \widehat{\mathbf{Y}}_{0,0} = \mathbf{1} \qquad (3.24)$$

where $\widehat{\mathbf{Y}}_{0,\cdot}$ is a $[0,T] \to U$ function, and the functions

$$\widehat{V}_i^y = \left(\widehat{V}_i^{\{(1)\},y}, \widehat{V}_i^{\{(2)\},y} \dots\right)^T : U \to U,\ (i = 1, \dots, d)$$

are defined coordinate-wise by

$$\widehat{V}_i^{\{R\},y}(x) = \begin{cases} V_i^j(x^W + y) & \text{if } R = (j), \|R\| = 1 \\ x^{R-} V_i^{r_l}(x^W + y) & \text{if } R = (r_1, \dots, r_l) \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}, l > 1 \end{cases} \tag{3.25}$$

for $x \in U$ and $y \in W$, where the coordinates are indexed by the multi-indices of $\mathcal{A}_{\alpha_Y}^{\Pi_Y}$.

Proof. The system of ODEs:

$$\begin{aligned} \mathrm{d}Y_{0,t} &= \sum_{i=1}^{k} V_i(Y_{0,t} + \xi)\mathrm{d}X_{0,t}^{(i)},\ Y_{0,0} = 0 \\ \mathrm{d}\widehat{\mathbf{Y}}_{0,t}^R &= \widehat{\mathbf{Y}}_{0,t}^{R-}\mathrm{d}Y_{0,t}^{r_l},\ \widehat{\mathbf{Y}}_{0,0}^R = 0,\ R = (r_1, \dots, r_l) \in \mathcal{A}_{\alpha_Y}^{\Pi_Y},\ l \in \mathbb{N} \end{aligned} \tag{3.26}$$

where $Y_{0,t} = (Y_{0,t}^1, \dots, Y_{0,t}^N)^T \in \mathbb{R}^N$, $\widehat{\mathbf{Y}}_{0,t}^{\epsilon} = \mathbf{1}$, $\widehat{\mathbf{Y}}_{0,t}^W = Y_{0,t}$ for $t \in [0.T]$ and $i = 1, \dots, N$ is equivalent to the RDE (3.17). This equivalence implies the Lemma. □

Remark 3.3.2. Note that for $y \in W$ the one-form $\widehat{f}^y : U \to L(V, U)$ defined by

$$\widehat{f}^y(x) = \left(\widehat{V}_1^y(x), \dots, \widehat{V}_k^y(x)\right),\ x \in U$$

is no longer global $Lip(\Pi, \Gamma)$, however it is still $Lip(\Pi, \Gamma)$ on any bounded closed set for an appropriate Π. Note that $\mathbf{Y}$ is controlled by $M\omega$ for an M depending on the Lipschitz norm of the one-form f on W and $\omega(0, T)$ but not depending on the initial condition. Therefore, the U-extended solution to the RDEs considered in this chapter, do not leave the ball in U centered at $\mathbf{1}$ with radius $M\omega(0, T)$. We restrict the one-form $\widehat{f}^y$ to this ball where it is $Lip(\Pi, \Gamma)$. This remark is a crucial ingredient for the local approximations introduced in this section.

Remark 3.3.3. Let $R = (r_1, \dots, r_l)$ be a multi-index. In some applications, one might be interested in only finding the component $\widehat{\mathbf{Y}}_{0,t}^R$ of $\widehat{\mathbf{Y}}$ solving the ODE (3.24). In this case, it is sufficient to solve the ODE (3.25) for multi-indices $Q \in \{S | \|S\| = 1\} \bigcup \{S = (r_1, \dots, r_m) | m \le l\}$, i.e. on the minimal tree-like set including the multi-index R.

Definition 3.3.8. *Let y be an element in W. We define the first order differential operators (vector fields on U) W_i^y, $i = 1, \dots k$ by*

$$W_i^y(x) = \sum_{R \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}} \widehat{V}_i^{\{R\},y}(x) \frac{\partial}{\partial x^R} \tag{3.27}$$

for $x \in U$, where $\widehat{V}_i^R$ is defined by (3.25) and $\partial / \partial x^R$ denotes the partial differentiation with respect to the coordinate of index R. Let $\mathcal{W}^y$ denote the set $\{W_1^y, \dots, W_k^y\}$. In the rest of the chapter, we use the notation $\Phi^y = \Phi_{\mathcal{W}^y}$ for the algebra homomorphism of Definition 3.3.6.

In some of the cases, we need to restrict W_i^y to a subspace of U. Therefore we introduce the following notation.

Definition 3.3.9. *Let y be an element in W and let $\mathcal{B}$ be a tree-like subset of $\mathcal{A}_{\alpha_Y}^{\Pi_Y}$. We define the vector fields $W_1^{y,\mathcal{B}}, \ldots, W_k^{y,\mathcal{B}}$ on $U^{\mathcal{B}} = \sum_{R\in\mathcal{B}} W^{\otimes R}$ by*

$$W_i^{y,\mathcal{B}}(x) = \sum_{R\in\mathcal{B}} V_i^{\{R\},y}(x)\frac{\partial}{\partial x^R}. \tag{3.28}$$

The set of vector fields $\left\{W_1^{y,\mathcal{B}}, \ldots, W_k^{y,\mathcal{B}}\right\}$ is denoted by $\mathcal{W}^{y,\mathcal{B}}$. The corresponding algebra homomorphism of Definition 3.3.6 is denoted by $\Phi^{y,\mathcal{B}} = \Phi_{\mathcal{W}^{y,\mathcal{B}}}$.

Let us recall some basic facts about ordinary differential equations driven by paths of bounded variation.

Lemma 3.3.3. *Let $\widehat{\mathbf{Z}} = (\widehat{\mathbf{X}}, \widehat{\mathbf{Y}})$ be the solution to the RDE* (3.17) *where $\widehat{\mathbf{X}}$ is of bounded variation, and let $g : U \to \mathbb{R}$ be a differentiable function. Then for $t > s > 0$,*

$$g(\widehat{\mathbf{Y}}_{0,t}^U) = g(\widehat{\mathbf{Y}}_{0,s}^U) + \int_s^t \sum_{i=1}^k W_i^\xi g(\widehat{\mathbf{Y}}_{0,u}^U)\mathrm{d}X_{0,u}^{(i)}. \tag{3.29}$$

Let m be the maximum length of the multi-indicies of $\mathcal{A}_{\alpha_X}^{\Pi_X}$, and let g be differentiable $(m+1)$-times. Then

$$\begin{aligned} g(\widehat{\mathbf{Y}}_{0,t}^U) &= \sum_{R=(r_1,\ldots,r_l)\in\mathcal{A}_{\alpha_X}^{\Pi_X}} W_R^\xi g(\widehat{\mathbf{Y}}_{0,s}^U)X_{s,t}^R \\ &\quad + \sum_{\substack{R=(r_1,\ldots,r_l)\notin\mathcal{A}_{\alpha_X}^{\Pi_X} \\ -R\in\mathcal{A}_{\alpha_X}^{\Pi_X}}} \int_{s<u_1<\cdots<u_l<t} W_R^\xi g(\widehat{\mathbf{Y}}_{0,u_1}^U)\mathrm{d}X_{0,u_1}^{(r_1)}\cdots\mathrm{d}X_{0,u_l}^{(r_l)}. \end{aligned} \tag{3.30}$$

The last term in (3.30) *is referred to as the* remainder term.

Proof. Equation (3.29) is equivalent to

$$g(\widehat{\mathbf{Y}}_{0,t}^U) = g(\widehat{\mathbf{Y}}_{0,s}^U) + \int_s^t \mathrm{d}g(\widehat{\mathbf{Y}}_{0,u}^U)\mathrm{d}Y_{0,u}.$$

Applying (3.29) recursively on the inner most term of each integral results in (3.30).

□

Remark 3.3.4. Let $\mathcal{B}$ be a tree-like subset of $\mathcal{A}_{\alpha_Y}^{\Pi_Y}$. Assume that g depends only on the components of $\widehat{\mathbf{Y}}^U$ corresponding to $\mathcal{B}$. Then Lemma 3.3.3 implies:

$$
\begin{aligned}
g(\widehat{\mathbf{Y}}^{U^{\mathcal{B}}}_{0,t}) &= \sum_{R=(r_1,\ldots,r_l)\in\mathcal{A}^{\Pi_X}_{\alpha_X}} W^{\xi}_{R} g(\widehat{\mathbf{Y}}^{U^{\mathcal{B}}}_{0,s}) X^R_{s,t} \\
&\quad + \sum_{\substack{R=(r_1,\ldots,r_l)\notin\mathcal{A}^{\Pi_X}_{\alpha_X} \\ -R\in\mathcal{A}^{\Pi_X}_{\alpha_X}}} \int_{s<u_1<\cdots<u_l<t} W^{\xi}_{R} g(\widehat{\mathbf{Y}}^{U^{\mathcal{B}}}_{0,u_1}) \mathrm{d}X^{(r_1)}_{0,u_1}\cdots \mathrm{d}X^{(r_l)}_{0,u_l} \\
&= \sum_{R=(r_1,\ldots,r_l)\in\mathcal{A}^{\Pi_X}_{\alpha_X}} W^{\xi,\mathcal{B}}_{R} g(\widehat{\mathbf{Y}}^{U^{\mathcal{B}}}_{0,s}) X^R_{s,t} \\
&\quad + \sum_{\substack{R=(r_1,\ldots,r_l)\notin\mathcal{A}^{\Pi_X}_{\alpha_X} \\ -R\in\mathcal{A}^{\Pi_X}_{\alpha_X}}} \int_{s<u_1<\cdots<u_l<t} W^{\xi,\mathcal{B}}_{R} g(\widehat{\mathbf{Y}}^{U^{\mathcal{B}}}_{0,u_1}) \mathrm{d}X^{(r_1)}_{0,u_1}\cdots \mathrm{d}X^{(r_l)}_{0,u_l}
\end{aligned}
$$

Now, we give a bound on the remainder term. Recall Section 2.6.3, and let

$$\widehat{\mathbf{Y}} \in G\Omega_{\Pi_{\alpha_Y}}(U)$$

denote the geometric Π_{α_Y}-rough path such that $\widehat{\mathbf{Y}}^U_{0,t} = \mathbf{Y}_{0,t}$ for all $t \in [0,T]$, i.e. the (Π_Y, α_Y)-truncated signature regarded as a Π_{α_Y}-rough path. The Universal Limit Theorem and Proposition 2.6.1 imply that the Π_{α_Y}-variation of $\widehat{\mathbf{Y}}$ is controlled by $M_1\omega$, where M_1 depends on $\Pi_X, \Gamma, \omega(0,T)$ and polynomially on $\|f\|_{Lip(\Pi_Y,\Gamma)}$. By analogy, one can prove that $\widehat{\mathbf{Z}} = \left(\mathbf{X}, \widehat{\mathbf{Y}}\right)$ is a $\Pi_X * \Pi_{\alpha_Y}$-rough path, with $\Pi_X * \Pi_{\alpha_Y}$-variation controlled by $M_1\omega$.

Lemma 3.3.4. *Set $\alpha > 0$ and $R \in \mathcal{A}^{\Pi_X}_{\alpha}$, such that $\|R\| \leq m$. Let $g : U \to \mathbb{R}$ be an m-times differentiable function, such that the one-form $h^{\xi}_R : V \oplus U \to L(V \oplus U, V \oplus \mathbb{R})$ defined by*

$$h^{\xi}_R(u,v) = \begin{pmatrix} Id_V & 0 \\ W^{\xi}_R g(v) & 0 \end{pmatrix} \tag{3.31}$$

is $(\Pi_X * \Pi_{\alpha_Y}, \overline{\Gamma})$-Lipschitz for a $\overline{\Gamma}$ making h^{ξ}_R integrable; i.e. $\bar{\gamma}_i > 1 - 1/\bar{p}_i$ for $i = 1, \ldots, \|\Pi_X * \Pi_{\alpha_Y}\|$, where $\bar{p}_i$ and $\bar{\gamma}_i$ denote the i^{th} elements of $\Pi_X * \Pi_{\alpha_Y}$ and $\overline{\Gamma}$ respectively.

Then there is a constant C_α depending on $\Pi_X * \Pi_{\alpha_Y}$, $\omega(0,T)$ and polynomially on $\max\{1, M_1\}$ and the Lipschitz norm of h^{ξ}_R, such that

$$\left\| \int_{s<u_1<\cdots<u_l<t} W^{\xi}_R g(\mathbf{Y}^U_{0,u_1}) \mathrm{d}X^{(r_1)}_{0,u_1} \cdots \mathrm{d}X^{(r_l)}_{0,u_l} \right\| \leq C_\alpha \omega(s,t)^{deg_{\Pi_x}(R)}$$

for all $[s,t] \subset [0,T]$.

Proof. Let $\mathbf{U}$ denote the integral $\int h^{\xi}_R \left(\widehat{\mathbf{Z}}\right) \mathrm{d}\widehat{\mathbf{Z}}$. Then, by Propositions 2.4.1 and 2.6.1, there exists a constant M_2 depending on $\Pi_X * \Pi_{\alpha_Y}$, $\omega(0,T)$ and polynomially on the Lipschitz norm of h^{ξ}_R, such that $\mathbf{U}$ is controlled by $\max\{1, M_1\}M_2\omega$. In particular for each $R =$

$(r_1, \ldots, r_l) \in \mathcal{A}_{\alpha_X}^{\Pi_X}$, there exists a constant C_R depending on $\Pi_X * \Pi_{\alpha_Y}$, $\omega(0,T)$ and polynomially on the Lipschitz norm of h_R^ξ, such that

$$\left\| \int_{s<u_1<\cdots<u_l<t} W_R^\xi g(\mathbf{Y}_{0,u_1}^U) \mathrm{d}X_{0,u_1}^{(r_1)} \cdots \mathrm{d}X_{0,u_l}^{(r_l)} \right\| \leq C_R \omega(s,t)^{deg_{\Pi_X}(R)},$$

which implies the Lemma.

□

In Lemma 3.3.4, it is shown, that the remainder term is the sum of certain cross-terms of $\int h_R^\xi \left(\widehat{\mathbf{Z}}\right) \mathrm{d}\widehat{\mathbf{Z}}$. These particular cross-terms will be denoted by

$$\begin{aligned} \widehat{X}_{s,t}^R \left\{ g(\widehat{\mathbf{Y}}_{0,\cdot}^U) \right\} &:= \int_{s<u_1<\cdots<u_l<t} g(\widehat{\mathbf{Y}}_{0,u_1}^U) \mathrm{d}X_{0,u_1}^{(r_1)} \cdots \mathrm{d}X_{0,u_l}^{(r_l)} \in \mathbb{R}, \\ \widehat{\mathbf{X}}_{s,t}^R \left\{ g(\widehat{\mathbf{Y}}_{0,\cdot}^U) \right\} &:= \varepsilon_R \widehat{X}_{s,t}^R \left\{ g(\widehat{\mathbf{Y}}_{0,\cdot}^U) \right\} \in V^R \end{aligned}$$

where $R = (r_r, \ldots, r_l) \in \mathcal{A}^k$ and $g : U \to \mathbb{R}$ is a suitable integrable function. For functions $g = (g^1, \ldots, g^n)^T : U \to \mathbb{R}^n$ we define

$$\widehat{X}_{s,t}^R \left\{ g(\widehat{\mathbf{Y}}_{0,\cdot}^U) \right\} := \begin{pmatrix} \widehat{X}_{s,t}^R \left\{ g^1(\widehat{\mathbf{Y}}_{0,\cdot}^U) \right\} \\ \vdots \\ \widehat{X}_{s,t}^R \left\{ g^n(\widehat{\mathbf{Y}}_{0,\cdot}^U) \right\} \end{pmatrix}.$$

The following corollary of Lemmas 3.3.3 and 3.3.4 is due to the continuity in the $\Pi_X * \Pi_Y$-variational topology.

Corollary 3.3.1 (Taylor expansion along rough paths). *Let $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ be the solution to the RDE* (3.1). *Let $g : U \to \mathbb{R}$ be a function such that for all multi-indices R satisfying $R = (r_1, \ldots, r_l) \notin \mathcal{A}_{\alpha_X}^{\Pi_X}$ and $-R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$, the function $h_R^\xi : V \oplus U \to L(V \oplus U, V \oplus \mathbb{R})$ defined in* (3.31) *is integrable along $\mathbf{Z}$; i.e. the one-form h_R^ξ is $Lip(\Pi_X * \Pi_{\alpha_Y}, \overline{\Gamma})$ for an appropriate $\overline{\Gamma}$. Then*

$$\begin{aligned} g(\mathbf{Y}_{0,t}^U) &= \sum_{R \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}} W_R^\xi g(\mathbf{Y}_{0,s}^U) X_{s,t}^R \\ &\quad + \sum_{\substack{R \notin \mathcal{A}_{\alpha_X}^{\Pi_X} \\ -R \in \mathcal{A}_{\alpha_X}^{\Pi_X}}} X_{s,t}^R \left\{ W_R^\xi g(\mathbf{Y}_{0,\cdot}^U) \right\}. \end{aligned} \tag{3.32}$$

Furthermore, there exists a constant C depending on $\Pi = \Pi_X * \Pi_{\alpha_Y}$, $\omega(0,T)$ and polynomially on M_1 and the Lipschitz norm of h_R^ξ for $R \notin \mathcal{A}_{\alpha_X}^{\Pi_X}$ and $-R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$, such that

$$\left\| \mathbf{X}_{s,t}^R \left\{ W_R^\xi g(\mathbf{Y}_{0,\cdot}^U) \right\} \right\| \leq C \omega(s,t)^{deg_{\Pi_X}(R)}$$

for all $R \in \mathcal{A}^{\Pi_X}$ satisfying $-R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$.

Let Id_U denote the identity function on U. Let the differential operator $W_i, i \in \{1,\dots,k\}$ act on Id_U coordinate-wise. For $u \in U$ we use the notation

$$W_R^{\xi}(u) := W_R^{\xi}(Id_U)(u) := \begin{pmatrix} W_R^{\xi}(Id_U^1(u)) \\ \vdots \\ W_R^{\xi}(Id_U^{\dim(U)}(u)) \end{pmatrix}.$$

The following expansion of the solution the RDE (3.1) is a special case of Corollary 3.3.1.

Note that when $g = Id_U$ and $R \in \mathcal{A}^{\Pi_X}$, the $(\Pi_X * \Pi_{\alpha_Y}, \overline{\Gamma})$-Lipschitz norm of the one-form $h_R^{\xi} : V \oplus U \to L(V \oplus U, V \oplus \mathbb{R})$ defined in (3.31) depends only on the $(\Pi_{\alpha_Y}, \overline{\Gamma}')$-Lipschitz norm of $W_R^{\xi}(Id_U)$, where $\overline{\Gamma}'$ is the $\|\Pi_{\alpha_Y}\|$-tuple resulted from dropping the first k elements of $\overline{\Gamma}$, assuming that h_R^{ξ} is integrable.

Corollary 3.3.2. *Let $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ be the solution to the RDE* (3.1)*. Let us assume that for Π_X and α_X the function $g = Id_U$ satisfies the conditions of Corollary 3.3.1. Then*

$$\begin{aligned} \mathbf{Y}_{0,t}^U &= \mathbf{Y}_{0,s}^U + \sum_{R=(r_1,\dots,r_l)\in\mathcal{A}_{\alpha_Y}^{\Pi_Y}\setminus\{\epsilon\}} W_R^{\xi}(\mathbf{Y}_{0,s}^U) X_{s,t}^R \\ &\quad + \sum_{\substack{R=(r_1,\dots,r_l)\notin\mathcal{A}_{\alpha_X}^{\Pi_X} \\ -R\in\mathcal{A}_{\alpha_X}^{\Pi_X}}} X_{s,t}^R \left\{ W_R^{\xi}(\mathbf{Y}_{0,\cdot}^U) \right\} \end{aligned} \tag{3.33}$$

Furthermore, there exists a constant C depending only on $\Pi = \Pi_X * \Pi_{\alpha_Y}$, $\omega(0,T)$ and polynomially on M_1 and $\|W_R^{\xi}(Id_U)\|_{Lip(\Pi_{\alpha_Y},\overline{\Gamma}')}$ for $-R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$, such that

$$\left\| \mathbf{X}_{s,t}^R \left\{ W_R^{\xi}(\mathbf{Y}_{0,\cdot}^U) \right\} \right\| \le C\omega(s,t)^{deg_{\Pi_X}(R)}$$

for all $[s,t] \subset [0,T]$ and for $-R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$.

The expansion (3.33) will be referred to as *the rough-Taylor expansion* of $\mathbf{Y}$.

Remark 3.3.5. In the case when the function g of Lemma 3.3.4 Corollary 3.3.1 depends only on the components of $\mathbf{Y}^U$ corresponding to a tree-like subset of $\mathcal{A}_{\alpha_Y}^{\Pi_Y}$, the lemma and the corollary are valid if W_R^{ξ} is replaced by $W_R^{\xi,\mathcal{B}}$. Analogously, Corollary 3.3.2 implies the expansion

$$\begin{aligned} \mathbf{Y}_{0,t}^{U^{\mathcal{B}}} &= \mathbf{Y}_{0,s}^{U^{\mathcal{B}}} + \sum_{R=(r_1,\dots,r_l)\in\mathcal{A}_{\alpha_Y}^{\Pi_Y}\setminus\{\epsilon\}} W_R^{\xi,\mathcal{B}}(\mathbf{Y}_{0,s}^{U^{\mathcal{B}}}) X_{s,t}^R \\ &\quad + \sum_{\substack{R=(r_1,\dots,r_l)\notin\mathcal{A}_{\alpha_X}^{\Pi_X} \\ -R\in\mathcal{A}_{\alpha_X}^{\Pi_X}}} X_{s,t}^R \left\{ W_R^{\xi,\mathcal{B}}(\mathbf{Y}_{0,\cdot}^{U^{\mathcal{B}}}) \right\}. \end{aligned} \tag{3.34}$$

furthermore with the constant C of Corollary 3.3.2,

$$\left\| \mathbf{X}_{s,t}^R \left\{ W_R^{\xi,\mathcal{B}}(\mathbf{Y}_{0,\cdot}^{U^{\mathcal{B}}}) \right\} \right\| \le C\omega(s,t)^{deg_{\Pi_X}(R)}$$

for all $[s,t] \subset [0,T]$ and for $-R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$.

3.3.4 Properties of the extended vector fields

Recall that $\mathbf{X}$ was a Π_X-rough path in $V = \mathbb{R}^k$ and $\mathbf{Y}$ was a Π_Y rough path in $W = \mathbb{R}^N$ where $\Pi_X = (p_1, \ldots, p_k)$ and $\Pi_Y = (q_1, \ldots, q_N)$ such that each q_i is one of the p_i-s. This is only possible, if the one-form $f = (V_1, \ldots, V_k)$ satisfies $V_i^j = 0$ for each pair of (i, j) where $p_i > q_j$.

Setting $z \in W$ and $R \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}$, we introduce the notation

$$W_i^{\{R\},z}(y) := \widehat{V}_i^{\{R\},z}(y)\frac{\partial}{\partial y^R} = \left\{ \begin{array}{ll} V_i^{r_l}(y^W + z) & \text{if } \|R\| = 1 \text{ and } R = r_l \\ y^{R-}V_i^{r_l}(y^W + z) & \text{if } \|R\| > 1 \text{ and } R = (R-) * r_l \end{array} \right\} \frac{\partial}{\partial y_R} \tag{3.35}$$

To prove the *Log-signature* theorems, we will use the special structure of

$$\Phi^z\left(\log^{(\Pi_Y,\alpha_Y)}(\mathbf{X})\right).$$

Therefore we firstly prove the following lemmas.

Lemma 3.3.5. *Let $z \in W$ and $R = (r_1, \ldots, r_l) \in \mathcal{A}_{\alpha_X}^{\Pi_X}$. The coordinate of $W_R^z \circ Id(y)$ corresponding to a multi-index $Q = (i_1, \ldots, i_m) \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ is a sum of terms of the form*

$$y^{\{(i_1,\ldots,i_{m-n})\}}U_{r_l}^{i_m}(y^W + z) \cdots U_{r_{l-n+1}}^{i_{m-n+1}}(y^W + z)U_{r_{l-n}}^{j_1}(y^W + z) \cdots U_{r_1}^{j_{l-n}}(y^W + z) \tag{3.36}$$

where $j_1, \ldots, j_{l-n} \in \{1, \ldots, k\}$, and the function U_v^u is either the u^{th} coordinate function V_v^u of V_v or one of its (multiple) partial derivatives, and $y^{\{\epsilon\}} = 1$.

Furthermore if any of the following inequalities hold

$$p_{r_l} > q_{i_m}, \; \cdots, \; p_{r_{l-n+1}} > q_{i_{m-n+1}}$$
$$p_{r_{l-n}} > q_{j_1}, \; \cdots, \; p_{r_1} > q_{j_{l-n}},$$

then the above product is 0.

Proof. Let $R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$ such that $\|R\| = 1$. Then the assertion on W_R is true by definition (see (3.35)). If $R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$ and $\|R\| = 2$, i.e. $R = (i, j)$, then

$$\begin{aligned} W_i^z \circ W_j^z \circ \mathrm{Id}(y) &= \sum_{P,Q \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}} W_i^{\{P\},z} \circ W_j^{\{Q\},z} \circ \mathrm{Id}(y) \\ &= \sum_{P,Q \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}} \left(y^{\{P-\}}V_i^{l(P)}(y^W + z)\right)\frac{\partial}{\partial y^{\{P\}}}\left(y^{\{Q-\}}V_j^{l(Q)}\right)(y^W + z)\frac{\partial}{\partial y^{\{Q\}}}Id(y) \\ &= \sum_{Q \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}} \sum_{l=1}^{N} y^{\{Q-\}}V_i^l(y^W + z)\left(\frac{\partial}{\partial y^{\{(l)\}}}V_j^{l(Q)}(y^W + z)\right) \\ &\quad + \sum_{Q \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}} y^{\{Q--\}}V_i^{l(Q-)}(y^W + z)V_j^{l(Q)}(y^W + z) \end{aligned}$$

where $l : \mathcal{A}_{\alpha_Y}^{\Pi_Y} \to \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ is determined by $(R-) * l(R) = R$ and we used the convention $y^{\{\epsilon\}} = 1$ and $y^{\{\epsilon-\}} = 0$. Recall that $V_i^j = 0$ for each pair of (i,j) where $p_i > q_j$, and hence the coefficients of $y^{\{Q-\}}$ and $y^{\{Q--\}}$ are products of terms of the type V_u^v or partial derivatives of V_u^v satisfying $u \in \{i,j\}$ and $p_u \leq q_v$.

By induction on the length of $R = (r_1, \ldots, r_l) \in \mathcal{A}_{\alpha_X}^{\Pi_X}$, one can complete the proof. □

Remark 3.3.6. Note that if $\mathcal{B} \subseteq \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ satisfies the tree-like conditions of Definition 3.2.1, then for each $Q \in \mathcal{B}$ the coordinate of $W_R^z \circ \mathrm{Id}$ corresponding to Q coincides with the coordinate of $W_R^{z,\mathcal{B}} \circ \mathrm{Id}$ corresponding to Q, where $W^{z,\mathcal{B}} \in \mathcal{W}^{z,\mathcal{B}}$ is defined in Definition 3.3.9.

Remark 3.3.7. Lemma 3.3.5 implies, that if $W_R^\xi \circ Id$ is Lipschitz on a closed set contained in a ball centered at $\mathbf{1}$ with radius r, then the Lipschitz norm of $W_R^\xi \circ Id$ depends only on the Lipschitz norm of $W^{\xi,\mathcal{B}_1} \circ Id$ and r. Note that the Lipschitz norm of $W^{\xi,\mathcal{B}_1} \circ Id$ does not depend on ξ.

Lemma 3.3.6. *Let $Q = \{i_1, \ldots, i_m\}$ be a multi-index in $\mathcal{A}_{\alpha_Y}^{\Pi_Y}$ and $[s,t] \in [0,T]$. Consider the coordinate function of the vector field $\Phi^z \left(\log^{(\Pi_Y,\alpha_Y)}(\mathbf{X}_{s,t})\right)$ corresponding to Q. The multiplying coefficient of the term including $y^{\{(i_1,\cdots,i_{m-n})\}}$ is bounded by $C\omega(s,t)^\delta$ where*

$$\delta \geq \deg_{\Pi_Y}\left((i_{m-n+1}, \cdots, i_m)\right)$$

and C depends on the $Lip(\Pi_Y, \Gamma)$-norm of the vector functions $V_1, \ldots, V_k$, and $\omega(0,T)$.

Proof. The log-signature of $\mathbf{X}$ corresponding to the interval $[s,t]$ is of the form

$$\Phi^\xi \left\{ L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X}) \right\} = \sum_{R \in \mathcal{A}_{\alpha_X}^{\Pi_X}} \left(\sum_{\substack{R_1 * \cdots * R_l = R \\ l \in \mathbb{N}}} a_{R_1,\cdots,R_l} X_{s,t}^{R_1} \cdots X_{s,t}^{R_l} \right) W_R^\xi \circ \mathrm{Id}, \tag{3.37}$$

where each $a_{R_1,\cdots,R_l}$ is a real coefficient depending only on R. Let us introduce the notation

$$A_{s,t}^R = \sum_{\substack{R_1 * \cdots * R_l = R \\ l \in \mathbb{N}}} a_{R_1,\cdots,R_l} X_{s,t}^{R_1} \cdots X_{s,t}^{R_l} \tag{3.38}$$

for $R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$.

Since $\mathbf{X}$ is controlled by ω, there exists a constant C_R depending only on R, such that

$$\left| A_{s,t}^R \right| \leq C_R \omega(s,t)^{deg_{\Pi_X}(R)}. \tag{3.39}$$

By Lemma 3.3.5 and (3.37), the multiplying coefficient of $y^{(i_{m-n+1},\cdots,i_m)}$ is a sum of terms of the type

$$A_{s,t}^R U_{r_l}^{i_m}(y^W + z) \cdots U_{r_{l-n+1}}^{i_{m-n+1}}(y^W + z) U_{r_{l-n}}^{j_1}(y^W + z) \cdots U_{r_1}^{j_{l-n}}(y^W + z)$$

for $R = (r_1, \ldots, r_l)$ or zero if any of the following inequalities hold

$$p_{r_l} > q_{i_m}, \ \cdots \ , \ p_{r_{l-n+1}} > q_{i_{m-n+1}}$$
$$p_{r_{l-n}} > q_{j_1}, \ \cdots \ , \ p_{r_1} > q_{j_{l-n}}$$

In the non-zero case, $p_{r_l} \leq q_{i_m}, \ \cdots \ , \ p_{r_{l-n+1}} \leq q_{i_{m-n+1}}$, and hence

$$\left| A^R_{s,t} \right| \leq C_R \omega(s,t)^{deg_{\Pi_X}(R)} = C_R \omega(s,t)^{\gamma_1} \omega(s,t)^{\gamma_2}$$

where

$$\begin{aligned} \gamma_1 &= \frac{1}{q_{i_m}} + \cdots + \frac{1}{q_{i_{m-n+1}}} \\ \gamma_2 &= \frac{1}{p_{r_1}} + \cdots + \frac{1}{p_{r_{l-n}}} + \frac{1}{p_{r_l}} - \frac{1}{q_{i_m}} + \cdots + \frac{1}{p_{r_{l-n+1}}} - \frac{1}{q_{i_{m-n+1}}} \geq 0 \end{aligned}$$

The inequality

$$\frac{1}{q_{i_m}} + \cdots + \frac{1}{q_{i_{m-n+1}}} \leq deg_{\Pi_Y}\left((i_{m-n+1}, \cdots, i_m)\right)$$

implies the assertion. □

3.4 Local approximation based on locally derived ODEs

In this section, we show that solutions to certain ordinary differential equations along vector fields depending on the local log-signature of the driving noise lead to a local approximation satisfying Condition 3.2.1 for some pair (τ, γ). The proof is based on the *Log-signature* theorems. These theorems are based in a finite dimensional setting, therefore we assume that the Banach spaces V and W are in fact $\mathbb{R}^k$ and $\mathbb{R}^N$ respectively.

3.4.1 Approximating terms corresponding to multi-indices in $\mathcal{B}_1$

In this section we prove the first version of the *Log-signature* Theorem, which provides us with a bound on the one-step error, when the W-level solution to the rough differential equation is locally approximated by a certain autonomous ordinary differential equation.

In this subsection, $\mathcal{B}_1$ denotes the set $\{R \in \mathcal{A}^{\Pi_Y}_{\alpha_Y}, \|R\| = 1\}$, and α_X is a positive real number. We will use $\mathcal{W}^{z,\mathcal{B}}$ and $\Phi^{z,\mathcal{B}}$ as defined in Definition 3.3.9 for $z \in W$. Finally we introduce new notation. Let F be a vector field. Then $\text{Exp}(tF)(\xi)$ denotes the time t solution y_t (if exists) to the ODE

$$\mathrm{d}y_s = F(y_s)\mathrm{d}s, \ y_0 = \xi.$$

Theorem 3.4.1 (Log-signature, the case of $\mathcal{B}_1$)**.** *Let ξ be an element in W and $s \in [0, T]$. Assume that $\mathcal{W}$ satisfies the following conditions* *(i) the RDE*

$$\mathrm{d}\mathbf{Y}_t = f(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_s = \xi$$

possess a unique solution $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}(\xi))$ for all $s \in [0, T]$

(ii) for each $R \in \mathcal{A}^{\Pi_X} \setminus \mathcal{A}^{\Pi_X}_{\alpha_X}$ such that $R = P * Q$ with $P, Q \in \mathcal{A}^{\Pi_X}_{\alpha_X}$ the function $h^{\xi}_R : V \oplus U \to L(V \oplus U, V \oplus \mathbb{R})$ defined in (3.31) for $g = Id_U$ is integrable.

Let $y_{s,t} \in \mathbb{R} \oplus W$ be defined by

$$y_{s,t} = \mathrm{Exp}\left(\Phi^{\xi,\mathcal{B}_1}\left\{L^{(\Pi_X,\alpha_X)}_{s,t}(\mathbf{X})\right\}\right)(\mathbf{1}).$$

and let $\mathcal{R}_{s,t} := \mathbf{Y}^W_{s,t} - y_{s,t}$

Then there exists a constant C such that

$$\|\mathcal{R}_{s,t}\| \leq C \sum_{\substack{Q,R \in \mathcal{A}^{\Pi_X}_{\alpha_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X}} \omega(s,t)^{deg_{\Pi_x}(Q*R)} \tag{3.40}$$

where C depends polynomially on M_1 and

$$\max_{\substack{Q,R \in \mathcal{A}^{\Pi_X}_{\alpha_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X}} \|W^{\xi,\mathcal{B}_1}_{Q*R} \circ Id_U\|_{Lip(\Pi_Y,\Gamma)}.$$

Proof. Let us fix an $s \in [0, T]$.

Step 1

The proof of the theorem is based on the comparison of the Taylor expansion of y_t and the rough-Taylor expansion of $\mathbf{Y}$. Firstly, we derive a particular form of the Taylor expansion. Similarly to the proof of Lemma 3.3.6 restricted to $\mathcal{B}_1$ (also using Remark 3.3.6)

$$\Phi^{\xi,\mathcal{B}_1}\left\{L^{(\Pi_X,\alpha_X)}_{s,t}(\mathbf{X})\right\} = \sum_{R \in \mathcal{A}^{\Pi_X}_{\alpha_X}} A^R_{s,t} W^{\xi,\mathcal{B}_1}_R. \tag{3.41}$$

Let $v_u,\ u \in [0,1]$ denote $\mathrm{Exp}\left(u\Phi^{\xi,\mathcal{B}_1}\left\{L^{(\Pi_X,\alpha_X)}_{s,t}(\mathbf{X})\right\}\right)(\mathbf{1})$. By Lemma 3.3.5 and the representation (3.37), the solution v_u exists.

By the simple differentiation rule, for a sufficiently many times differentiable function $f : U^{\mathcal{B}_1} \to \mathbb{R}$,

$$\mathrm{d}f(v_u) = \sum_{R \in \mathcal{A}^{\Pi_X}_{\alpha_X}} A^R_{s,t} W^{\xi,\mathcal{B}_1}_R f(v_u)\mathrm{d}u. \tag{3.42}$$

Applying this rule coordinate-wise on the functions $W_R \circ Id_U$, we have

$$v_1 = v_0 + \sum_{R \in \mathcal{A}^{\Pi_X}_{\alpha_X} \setminus \{\epsilon\}} A^R_{s,t} W^{\xi,\mathcal{B}_1}_R(v_0) + \sum_{Q,R \in \mathcal{A}^{\Pi_X}_{\alpha_X}} A^R_{s,t} A^Q_{s,t} \int_0^1 W^{\xi,\mathcal{B}_1}_Q(W^{\xi,\mathcal{B}_1}_R(v_u))\mathrm{d}u.$$

We say that the degree of a term $\int_0^1 W_{R_1}^{\xi,\mathcal{B}_1} \circ \cdots \circ W_{R_l}^{\xi,\mathcal{B}_1}(y_u)\mathrm{d}u$ for $R_1,\ldots,R_l \in \mathcal{A}^k$ is $deg_{\Pi_X}(R_1 * \cdots * R_l)$. Repeatedly applying the differentiation rule (3.42) to the functions inside the remainder integrals of degree at most α_X, one can derive the following expansion:

$$
\begin{aligned}
v_1 &= v_0 + \sum_{\substack{R_1,\ldots,R_l \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ 0 < deg_{\Pi_X}(R_1 * \cdots * R_l) \le \alpha_X}} \left(\prod_{i=1}^{l} A_{s,t}^{R_i} \right) W_{R_1}^{\xi,\mathcal{B}_1} \circ \cdots \circ W_{R_l}^{\xi,\mathcal{B}_1}(v_0) \int_{0<u_1<\cdots<u_l<1} \mathrm{d}u_1 \cdots \mathrm{d}u_l \\
&\quad + \sum_{\substack{R_1,\ldots,R_l \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(R_2 * \cdots * R_l) \le \alpha_X \\ deg_{\Pi_X}(R_1 * R_2 * \cdots * R_l) > \alpha_X}} \left(\prod_{i=1}^{l} A_{s,t}^{R_i} \right) \int_{0<u_1<\cdots<u_l<1} W_{R_1}^{\xi,\mathcal{B}_1} \circ \cdots \circ W_{R_l}^{\xi,\mathcal{B}_1}(v_{u_1}) \mathrm{d}u_1 \cdots \mathrm{d}u_l
\end{aligned}
\tag{3.43}
$$

We will refer to the last term on the right-hand-side of (3.43) as the *$\mathcal{B}_1$-level Taylor remainder term* $\mathcal{R}_{s,t}^{\mathrm{Tay},\mathcal{B}_1}$.

Step 2

Next, we give a bound on the Taylor remainder term. By the conditions of the theorem, the function $W_{R_1}^{\xi,\mathcal{B}_1} \circ \cdots \circ W_{R_l}^{\xi,\mathcal{B}_1} \circ \mathrm{Id}$ under the integral term corresponding to the set of multi-indices $\{R_1,\ldots,R_l\}$ in (3.43) is bounded by

$$
C_1 = \max_{\substack{Q,R \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X}} \| W_{Q*R}^{\xi,\mathcal{B}_1} \circ Id_U \|_{Lip(\Pi,\Gamma)}.
$$

Hence the integral itself is bounded by $C_1^{\|R_1 * \cdots R_l\|} \frac{1}{l!}$. Furthermore, the multiplying coefficient of the integral can be bounded using (3.39).

Therefore, the remainder term is bounded by

$$
\left\| \mathcal{R}_{s,t}^{\mathrm{Tay},\mathcal{B}_1} \right\| \le C_2 \left(\sum_{\substack{Q,R \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X}} \omega(s,t)^{deg_{\Pi_X}(Q*R)} \right)
$$

where C_2 depends on C_R, $R \in \mathcal{A}_{\alpha_X}^{\Pi_X}$, C_1, Π_X and α_X.

Step 3

Next, we compare the above Taylor expansion to the rough-Taylor expansion. Let $\mathbf{Y}$ be the solution to the RDE

$$
\mathrm{d}\mathbf{Y}_t = f(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_s = \xi.
$$

Rewriting (3.33) and restricting to $\mathcal{B}_1$, we get

$$\begin{aligned}\mathbf{Y}^W_{s,t} &= \sum_{R=(r_1,\dots,r_l)\in\mathcal{A}^{\Pi_Y}_{\alpha_Y}\setminus\{\epsilon\}} X^R_{s,t}\Phi^{\xi,\mathcal{B}_1}(\varepsilon_R)(\mathbf{Y}^W_{0,s}) \\ &+ \sum_{\substack{R=(r_1,\dots,r_l)\notin\mathcal{A}^{\Pi_X}_{\alpha_X}\\ -R\in\mathcal{A}^{\Pi_X}_{\alpha_X}}} X^R_{s,t}\left\{W^{\xi,\mathcal{B}_1}_R(\mathbf{Y}^W_{0,\cdot})\right\}. \end{aligned} \tag{3.44}$$

We will refer to the last term on the right-hand-side of (3.44) as the $\mathcal{B}_1$*-level rough-Taylor remainder term* $\mathcal{R}^{\mathrm{rTay},\mathcal{B}_1}_{s,t}$.

By Theorem 3.3.1, the exponential of the log-signature is the signature, and therefore

$$\begin{aligned}\Phi^{\xi,\mathcal{B}_1}\left[\pi_{\Pi_X,\alpha_X}\mathbf{X}_{s,t}\right] &= \Phi^{\xi,\mathcal{B}_1}\left[\exp^{(\Pi_X,\alpha_X)}\left(L^{(\Pi_X,\alpha_X)}_{s,t}(\mathbf{X})\right)\right] \\ &= \Phi^{\xi,\mathcal{B}_1}\left[\pi_{\Pi_X,\alpha_X}\left[\sum_{i=1}^{\infty}\frac{1}{i!}\left(L^{(\Pi_X,\alpha_X)}_{s,t}(\mathbf{X})\right)^{\otimes i}\right]\right] \\ &= \sum_{\substack{R_1,\dots,R_l\in\mathcal{A}^{\Pi_X}_{\alpha_X}\\ deg_{\Pi_X}(R_1*\cdots*R_l)\le\alpha_X\\ l\in\{0,1,2,\dots\}}} \frac{1}{l!}\left(\prod_{i=1}^{l}A^{R_i}_{s,t}\right)\Phi^{\xi,\mathcal{B}_1}\left[\varepsilon_{R_1}\otimes\cdots\otimes\varepsilon_{R_l}\right] \end{aligned} \tag{3.45}$$

The equality (3.45) connects (3.43) and (3.44) as follows:

$$\mathbf{Y}^W_{s,t} - y_{s,t} = \mathcal{R}^{\mathrm{rTay},\mathcal{B}_1}_{s,t} - \mathcal{R}^{\mathrm{Tay},\mathcal{B}_1}_{s,t}, \tag{3.46}$$

hence

$$\left\|\mathbf{Y}^W_{s,t} - y_{s,t}\right\| \le C_3 \sum_{\substack{Q,R\in\mathcal{A}^{\Pi_X}_{\alpha_X}\\ deg_{\Pi_X}(Q*R)>\alpha_X}} \omega(s,t)^{deg_{\Pi_X}(Q*R)}$$

where C_3 depends on polynomially on M_1, C_2 and

$$\max_{\substack{Q,R\in\mathcal{A}^{\Pi_X}_{\alpha_X}\\ deg_{\Pi_X}(Q*R)>\alpha_X}} \|W^{\xi,\mathcal{B}_1}_{Q*R}\circ Id_U\|_{Lip(\Pi,\Gamma)}.$$

This implies the existence of the required constant C. □

Corollary 3.4.1. *Assume the conditions of Theorem 3.4.1 and consider the constant C determined in the theorem. Then the local error $\mathcal{R}_{s,t}$ satisfies the inequality*

$$\|\mathcal{R}_{s,t}\| \le CK\omega(s,t)^{\gamma}$$

where $\gamma = \min_{R\in\mathcal{A}^k\setminus\mathcal{A}^{\Pi_X}_{\alpha_X}} deg_{\Pi_X}(R)$ *and K depends only on* $\omega(0,T)$ *as follows:*

$$K = \sum_{\substack{Q,R\in\mathcal{A}^{\Pi_X}_{\alpha_X}\\ deg_{\Pi_X}(Q*R)>\alpha_X}} \omega(0,T)^{deg_{\Pi_X}(Q*R)-\gamma}$$

Remark 3.4.1. Let $[s,t]$ be a subinterval of $[0,T]$. Consider

$$v_u = \mathrm{Exp}\left(u\Phi^{\xi,\mathcal{B}_1}\left\{L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X})\right\}\right)(\mathbf{1}), \text{ for } u \in [0,1].$$

Lemma 3.3.6 implies the existence of a constant C depending on $\omega(0,T)$ and polynomially on the $Lip(\Pi_Y,\Gamma)$-norm of the functions $V_1,\ldots,V_k$, such that

$$\left\|\pi_{(i)}v_u\right\| \leq uC\omega(s,t)^{\frac{1}{q_i}}.$$

Assuming that the exact solutions to the locally derived ordinary differential equations are known, we paste them together and construct a global approximation for the W-level solution of the rough differential equation. In order to give a precise construction, we introduce two operators.

Definition 3.4.1. *Let $\alpha_X,\alpha_Y \geq 1$ be fixed. Let f be the one-form defining the RDE* (3.1) *and let the driving noise $\mathbf{X}$ be given. Let the set $\mathcal{B} \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ be tree-like (i.e. satisfying the conditions of Definition 3.2.1). Let $U^{\mathcal{B}}$ denote the subspace of U defined as*

$$U^{\mathcal{B}} = \sum_{R\in\mathcal{B}} W^{\otimes R}.$$

For $(s,t) \in \Delta$, we define $D_{s,t}^{y} : U^{\mathcal{B}} \to U^{\mathcal{B}}$ to be the function assigning the $\pi_{U^{\mathcal{B}}}$ projection of $z \otimes \mathbf{Y}_{s,t}^{U^{\mathcal{B}}}(z,s)$ to $z \in U^{\mathcal{B}}$, where $\mathbf{Y}(z,s)$ is the solution (if it exists) to

$$\mathrm{d}\mathbf{Y}_t(z,s) = f(\mathbf{Y}_t(z,s) + y + \pi_W z)\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_s(z,s) = \mathbf{1}.$$

Furthermore, for $(s,t) \in \Delta$, let $\widehat{D}_{s,t}^{y} : U^{\mathcal{B}} \to U^{\mathcal{B}}$ be a shorthand notation for

$$\widehat{D}_{s,t}^{y}(z) = \pi_{U^{\mathcal{B}}}\left\{z \otimes \left\{\mathrm{Exp}\left(\Phi^{(y+\pi_W z),\mathcal{B}}\left\{L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X})\right\}\right)(\mathbf{1})\right\}\right\}.$$

When the mesh size of $\mathcal{D}$ tends to zero, the following simplified global bound proves itself to be useful.

Corollary 3.4.2. *Assume the conditions of Theorem 3.2.1 hold. Then there exist constants $K_{(r)}$, $r = 1,\ldots,N$ such that*

$$\left\|\mathbf{Y}_{t_i,t_j}^{(r)} - \widehat{\mathbf{Y}}_{t_i,t_j}^{(r)}\right\| \quad \leq \quad CK_{(r)}\sum_{k=0}^{j-1}\omega(t_k,t_{k+1})^{\gamma} \tag{3.47}$$

where $\gamma = \min_{R\in\mathcal{A}^k\setminus\mathcal{A}_{\alpha_X}^{\Pi_X}} deg_{\Pi_X}(R)$ and $K_{(r)}$ depend only on $\omega(0,T)$ Π_Y and α_X as follows:

$$K_{(r)} = \frac{1}{\beta^N\left(\frac{1}{q_r}\right)!}\sum_{\substack{Q,R\in\mathcal{A}_{\alpha_X}^{\Pi_X}\\ deg_{\Pi_X}(Q*R)>\alpha_X}}\omega(0,T)^{deg_{\Pi_X}(Q*R)+\frac{1}{q_r}-\gamma}$$

3.4.2 APPROXIMATING TERMS CORRESPONDING TO MULTI-INDICES IN $\mathcal{B}_m$ FOR $m > 1$

When extending the numerical schemes based on locally derived ordinary differential equations to provide an approximation of the higher level terms of the solution's signature, we find that the terms corresponding to multi-indices with length greater than 1 are iterated integrals in some sense of the W-level approximation. More precisely the $\|R\| = i$ level solution does not depend on the $j > i$ level components but on the $(i-1)$-level and 1st level components. Furthermore, to extend the Log-signature theorem to the multi-indices of length greater than 1, we need to prove that the each local ODE possess a unique solution, and the error term is bounded in some way, which takes us closer to satisfying Condition 3.2.1. In this section, we achieve these steps by induction on the length of the multi-indices using techniques similar to those in the previous section.

Recall the definition of the tree-like sets $\mathcal{B}_i$:

$$\mathcal{B}_i = \{R \in \mathcal{A}^{\Pi_Y,\alpha_Y}_{,} \ \|R\| \leq i\},\ i = 1, 2, \ldots.$$

Hypothesis 3.4.1. *We make the following assumption: If $R \in \mathcal{B}_m$, $[s,t] \in [0,T]$ and $u \in [0,1]$, then $\left\| v_u^{\{R\}} \right\| < uC_R\omega(s,t)^{deg_{\Pi_Y}(R)}$ where v_u is the solution to the following ODE on $U^{\mathcal{B}_m}$:*

$$\mathrm{d}v_u = \Phi^{(\xi+\pi_W z),\mathcal{B}_m}\left\{L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X})\right\}(v_u)\mathrm{d}u,\ v_0 = \mathbf{1}$$

for $z \in U^{\mathcal{B}_m}$where C_R only depends on R, $\omega(0,T)$, and polynomially on M_1 and

$$\max_{\substack{Q,R \in \mathcal{A}^{\Pi_X}_{\alpha_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X}} \|W^{\xi,\mathcal{B}_1}_{Q*R} \circ Id_U\|_{Lip(\Pi_Y,\Gamma)}.$$

Note that in the case when $m = 1$, Hypothesis 3.4.1 is verified by Remark 3.4.1. By the following theorem we extend the validity of the hypothesis to larger integers.

Theorem 3.4.2 (Log-signature, the case of $\mathcal{B}_m$, $m > 1$)**.** *Let ξ be an element in W and $s \in [0,T]$. Assume that $\mathcal{W}$ satisfies the following conditions:*

(i) the RDE

$$\mathrm{d}\mathbf{Y}_t = f(\mathbf{Y}_t)\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_s = \xi$$

possesses a unique solution $\mathbf{Z} = (\mathbf{X}, \mathbf{Y}(\xi))$ for all $s \in [0,T]$,

(ii) for each $R \in \mathcal{A}^{\Pi_X} \setminus \mathcal{A}^{\Pi_X}_{\alpha_X}$ such that $R = P * Q$ with $P, Q \in \mathcal{A}^{\Pi_X}_{\alpha_X}$, the function $h^{\xi}_R : V \oplus U \to L(V \oplus U, V \oplus \mathbb{R})$ defined in (3.31) for $g = Id_U$ is integrable.

Let $y_{s,t} \in U^{\mathcal{B}_{m+1}}$ be defined by

$$y_{s,t} = \mathrm{Exp}\left(\Phi^{\xi,\mathcal{B}_{m+1}}\left\{L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X})\right\}\right)(\mathbf{1})$$

and let $\mathcal{R}_{s,t} := \mathbf{Y}_{s,t}^{U^{\mathcal{B}_{m+1}}} - y_{s,t}$. *Let* $m > 1$ *be an integer, and assume that Hypothesis 3.4.1 hold for* $\mathcal{B}_m$.

Then for each $R \in \mathcal{B}_{m+1} \setminus \mathcal{B}_m$ *there exist constants* C_R *and* K_R *such that*

$$\|\pi_R \mathcal{R}_{s,t}\| \leq C_R \sum_{\substack{Q,R \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X}} \omega(s,t)^{deg_{\Pi_x}(Q*R)}, \tag{3.48}$$

where C_R *depends on* R *and polynomially on* M_1 *and*

$$\max_{\substack{Q,R \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X}} \|W_{Q*R}^{\xi,\mathcal{B}_1} \circ Id_U\|_{Lip(\Pi,\Gamma)}.$$

Proof. Theorem 3.4.1 covers the special $\mathcal{B}_1$-level case. The $\mathcal{B}_m$-case is proved by induction on m. The proof is analogous to the proof of Theorem 3.4.1, with the exception that some additional considerations are required on the bound on the Taylor remainder term $\mathcal{R}_{s,t}^{\mathrm{Tay},\mathcal{B}_{m+1}}$.

Note that by the special structure (Lemma 3.3.5 and Remark 3.3.6) of the vector fields in $\mathcal{W}^{\xi}$, $\pi_R \mathcal{R}_{s,t}^{\mathrm{Tay},\mathcal{B}_m} = \pi_R \mathcal{R}_{s,t}^{\mathrm{Tay},\mathcal{B}_{m+1}}$ for all $R \in \mathcal{B}_m$, hence the components corresponding to multi-indices of length $m+1$ are to be considered.

Recall Remark 3.3.7 and the following facts

(i) By the Taylor expansion (3.43) derived on $U^{\mathcal{B}_{m+1}}$, the remainder term $\mathcal{R}_{s,t}^{\mathrm{Tay},\mathcal{B}_{m+1}}$ equals

$$\sum_{\substack{R_1,\ldots,R_l \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(R_2 * \cdots * R_l) \leq \alpha_X \\ deg_{\Pi_X}(R_1 * R_2 * \cdots * R_l) > \alpha_X}} \left(\prod_{i=1}^{l} A_{s,t}^{R_i} \right) \int_{0<u_1<\cdots<u_l<1} W_{R_1}^{\xi,\mathcal{B}_{m+1}} \circ \cdots \circ W_{R_l}^{\xi,\mathcal{B}_{m+1}}(v_{u_1}) \mathrm{d}u_1 \cdots \mathrm{d}u_l.$$

(ii) By Lemma 3.3.5 the coordinate of $W_{R_1}^{\xi,\mathcal{B}_{m+1}} \circ \cdots \circ W_{R_l}^{\xi,\mathcal{B}_{m+1}}(y)$ corresponding to the multi-index $Q \in \mathcal{B}_{m+1}$ is a sum of terms of the form $y^{\widehat{Q}} U_{\widehat{Q}}(y^W)$, where $\widehat{Q}$ is a multi-index of length at most m for which there exists a multi-index $\overline{Q}$, such that $Q = \widehat{Q} * \overline{Q}$, and $U_{\widehat{Q}}$ is a product built of coordinate functions of the V_i's and their partial derivatives.

(iii) By Lemma 3.3.6, the multiplying coefficient $\left(\prod_{i=1}^{l} A_{s,t}^{R_i} \right)$ is bounded by

$$C_1 \omega(s,t)^{deg_{\Pi_Y}(\overline{Q})},$$

where C_1 depends only on Q, $R = R_1 * \cdots * R_l$ and l.

(iv) Finally, by Hypothesis 3.4.1, for all $Q \in \mathcal{B}_m$, $\left\| v_u^{\{Q\}} \right\| < uC_2\omega(s,t)^{deg_{\Pi_Y}(Q)}$ where C_2 only depends on R, $\omega(0,T)$ and polynomially on M_1 and

$$\max_{\substack{Q,R \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X}} \|W_{Q*R}^{\xi,\mathcal{B}_1} \circ Id_U\|_{Lip(\Pi_Y,\Gamma)}.$$

These facts imply the existence of a constant C_3 such that for $Q \in \mathcal{B}_{m+1} \setminus \mathcal{B}_m$ and $R_1, \ldots, R_l \in \mathcal{A}_{\alpha_X}^{\Pi_X}$ satisfying $deg_{\Pi_X}(R_2 * \cdots * R_l) \leq \alpha_X$ and $deg_{\Pi_X}(R_1 * R_2 * \cdots * R_l) > \alpha_X$, the following inequality holds:

$$\begin{aligned}
&\pi_Q \left\{ \left(\prod_{i=1}^{l} A_{s,t}^{R_i} \right) \int_{0<u_1<\cdots<u_l<1} W_{R_1}^{\xi,\mathcal{B}_{m+1}} \circ \cdots \circ W_{R_l}^{\xi,\mathcal{B}_{m+1}}(v_{u_1}) \mathrm{d}u_1 \cdots \mathrm{d}u_l \right\} \\
&\leq C_3 \sum_{\substack{R=R_1,\ldots,R_l \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(R_2*\cdots*R_l) \leq \alpha_X \\ deg_{\Pi_X}(R_1*R_2*\cdots*R_l) > \alpha_X}} \omega(s,t)^{deg_{\Pi_X}(R)} \leq C_4\omega(s,t)^{\gamma}
\end{aligned}$$

where C_3 and C_4 depend only on Q, $\sum_{i=1}^{l} \|R_i\|$, l and polynomially on C_2.

By definition, for $Q \in \mathcal{B}_{m+1} \setminus \mathcal{B}_m$

$$v_u^{\{Q\}} = \pi_Q \int_0^u \Phi^{\xi,\mathcal{B}_1} \left\{ L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X}) \right\} (v_z) \mathrm{d}z.$$

Then by Lemmas 3.3.5 and 3.3.6 and by Hypothesis 3.4.1, there exists a constant C_Q depending only on Q, $\omega(0,T)$ and polynomially on C_4, such that

$$\|\pi_Q v_u\| \leq uC_Q\omega(s,t)^{deg_{\Pi_Y}(Q)},$$

implying that Hypothesis 3.4.1 holds for $m+1$. □

The following corollary links the locally derived ODEs to global numerical schemes.

Corollary 3.4.3. *Let α_X be chosen to satisfy $\min_{R \in \mathcal{A}^k \setminus \mathcal{A}_{\alpha_X}^{\Pi_X}} deg_{\Pi_X}(R) > 1$. Let the local approximation $\widehat{D}^{ODE}$ be defined by*

$$\widehat{D}^{ODE}\left[(s,t),z\right] = \operatorname{Exp}\left(\Phi^z \left\{ L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X}) \right\}\right)(\mathbf{1}).$$

Then $\widehat{D}^{ODE}$ satisfies Condition 3.2.1 with $\tau = \omega(0,T)$ and $\gamma = \min_{R \in \mathcal{A}^{\Pi_X} \setminus \mathcal{A}_{\alpha_X}^{\Pi_X}} deg_{\Pi_X}(R)$:

$$\left\| \pi_R D\left[(s,t),z\right] - \pi_R \widehat{D}^{ODE}\left[(s,t),z\right] \right\| \leq C_R\omega(s,t)^{\gamma}$$

for all $(s,t) \in [0,T]$ and $z \in W$ where C_R depends only on R, $\omega(0,T)$, and polynomially

$$\max_{\substack{Q,R \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X \\ z \in W}} \|W_{Q*R}^{z} \circ Id_U\|_{Lip(\Pi_Y,\Gamma)}.$$

The following corollary opens the door to practical numerical schemes.

Corollary 3.4.4. *Let* α_X *be chosen to satisfy* $\min_{R\in\mathcal{A}^{\Pi_Y}\setminus\mathcal{A}^{\Pi_X}_{\alpha_X}} deg_{\Pi_X}(R) > 1$. *Let the local approximation* $\widehat{D}^{N.ODE}$ *be defined by a (numerical) approximation of the solutions to the ODE determined by* $\widehat{D}^{ODE}$, *such that for each* $R \in \mathcal{A}^{\Pi_Y}_{\alpha_Y}$ *there exists a constant* C_R *not depending on* (s,t) *or* z, *such that*

$$\left\|\pi_R\widehat{D}^{ODE}\left[(s,t),z\right] - \pi_R\widehat{D}^{N.ODE}\left[(s,t),z\right]\right\| \leq C_R\omega(s,t)^{\gamma}$$

for all $(s,t) \in [0,T]$ *and* $z \in W$.

Then $\widehat{D}^{ODE}$ *also satisfies Condition 3.2.1 with* $\tau = \omega(0,T)$ *and* $\gamma = \min_{R\in\mathcal{A}^k\setminus\mathcal{A}^{\Pi_X}_{\alpha_X}} deg_{\Pi_X}(R)$.

Remark 3.4.2. Note that $\pi_R\widehat{D}^{\text{ODE}}\left[(s,t),z\right]$ and $\pi_R\widehat{D}^{\text{N.ODE}}\left[(s,t),z\right]$ do not depend on the geometric Π_X-rough path driving noise but only on the local log-signature $L^{(\Pi_X,\alpha_X)}_{s,t}(\mathbf{X})$. Hence we can construct numerical schemes using the local log-signature as an input.

Remark 3.4.3. One advantage of approximating the solution of the RDEs in terms of solutions to ODEs driven by vector fields derived through the algebra homomorphism Φ, is that the approximation will not leave the reachability manifold determined by the one-form f. We will refer to this property as *geometric stability*. When replacing the solutions of the ODEs with numerical approximations, one might lose this property unless the ODE-numerical scheme has some built-in guarantees. However, in some of the cases, it is not known if the reachability manifold is a proper subset of $\mathbb{R}^N$. Then the higher the order convergence the ODE-approximating numerical scheme has, the closer the RDE-approximating scheme stays to the reachability manifold as well as to geometric stability. In the next section we abandon the need for strict geometric stability in order to weaken the conditions on the one-form f while constructing a convergent numerical approximation to the RDE.

3.4.3 Notes on the Log-signature theorems

Approximation of solutions to differential equations in terms of exponential series can be traced back to Magnus [29] or even earlier. We have already referred to Strichartz [35] where asymptotic results on the approximation of non-autonomous ODEs in terms of autonomous ODEs are derived. In the papers by Ben Arous [1], Castell [4], Hu [13] etc., there are analogous results for local ODE-based approximations of stochastic differential equations, some cases articulated as stochastic generalizations of [35]. Global approximations based on a local stochastic Taylor expansion are standard methods in numerical SDEs, and we refer the reader to Gard [11] and Kloeden & Platen [15] for further details, examples and references. Numerical solutions of SDEs based on explicit ODEs can be traced back to Newton [30] and probably even earlier. We refer the reader to Gaines & Lyons [10] for further references.

3.4.4 Notes on a special case

In the previous sections, we showed that a sufficient condition for the global convergence, when applying the local approximation based on solutions to ODEs derived from the log-signature $L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X})$ using the algebra homomorphism Φ, is

$$\gamma = \min_{R \in \mathcal{A}^k \setminus \mathcal{A}_{\alpha_X}^{\Pi_X}} deg_{\Pi_X}(R) > 1.$$

Recall that the log-signature is in general of the form

$$\log(\mathbf{X}_{s,t}) = \sum_{i=1}^{k} \mathbf{X}_{s,t}^{(i)} \varepsilon_i + \frac{1}{2} \sum_{1 \le i < j \le k} \left(\mathbf{X}_{s,t}^{(i,j)} - \mathbf{X}_{s,t}^{(j,i)} \right) [\varepsilon_i, \varepsilon_j] + \dots$$

Assuming that the terms $\mathbf{X}_{s,t}^{R}$ corresponding to shorter multi-indices R are easier (cheaper) to generate, we might be interested in finding the smallest α_X for which the global scheme is still convergent. The increments are part of the scheme by definition. For a particular special case, the terms corresponding to multi-indices of length at least two can be omitted from the scheme. This case occurs when for each pair $(i,j) \in \{1,\dots,k\}$ $i \neq j$, the inequality

$$deg_{\Pi_X}(i,j) = \frac{1}{p_i} + \frac{1}{p_j} > 1$$

holds. It is a slight generalization of the (p,q) rough paths introduced by Lejay & Victoir [21].

3.5 Local approximation under weaker conditions

One crucial ingredient to the proof of the Log-signature Theorems was the bound on the rough-Taylor remainder term. Unfortunately, this bound requires rather strong conditions on the one-form f. By the Universal Limit Theorem, solutions to RDEs exist under much weaker conditions, and one might be interested in determining if a tractable numerical approximation can be constructed under these conditions.

In this section we apply some results of Caruana [3] which are generalizations of the results of Davie [8]. In [3], the one-form f is locally approximated by polynomial functions and a sequence of *rough polynomial approximations* is constructed. His main reason for using polynomials is that the polynomial one-forms satisfy more regularities, implying the existence and uniqueness of solutions (locally).

In this section, we only assume that the one-form f satisfies the conditions of the Universal Limit Theorem 2.5.1, implying the existence and uniqueness of the solution to the RDE (3.1). In the first step, we show how to construct local approximations using local polynomial approximations of f, then we apply the Log-signature Theorems to the rough polynomial approximations in order to construct tractable schemes.

3.5.1 Rough polynomial approximation of RDEs

Consider the Banach spaces V and W, such that $V = V^1 \oplus \cdots \oplus V^k$ for some Banach spaces $V^1, \ldots, V^k$. Let $\Pi_X = (p_1, \ldots, p_k)$ and $\Gamma = (\gamma_1, \ldots, \gamma_k)$ denote real k-tuples as in Definition 2.1.1. Let $p_{\max} = \max_{1 \le i \le k} p_i$.

Recall that a one-form $f : W \to L(V, W)$ is a (Π_X, Γ)-*Lipschitz one form*, if for all $s_m < \gamma_i$, it implies the existence of the functions $f_i^{s_m}$ such that

(i) $f = \sum_{i=1}^k f_i^{s_0}$ where $f_i^{s_0} : W \to L(V^i, W)$ defined by $f(v) = f_i^{s_0}(v) \circ \pi_{V^i}$ for all $v \in W$,

(ii) $f_i^{s_m} : W \to L\left(W^{\otimes(\Pi, s_m)}, L(V^i, W)\right)$ for $i = 1, \ldots, k$ taking values in the space of s_m-symmetric linear maps, satisfies

$$f_i^{s_m}(y)(u) = \sum_{s_m \le s_n < \gamma_i} f_i^{s_n}(x) \left(u \otimes \sum_{deg_\Pi(R) = s_n - s_m} \frac{(y-x)_R}{\|R\|!} \right) + R_i^{s_m}(x, y)(u)$$

for all $x, y \in W$ and $u \in W^{\otimes(\Pi, s_m)}$ where $R_i^{s_m} : W \times W \to L\left(W^{\otimes(\Pi, s_m)}, L(V^i, W)\right)$ and

$$\left\| R_i^{s_m}(x, y) \right\| < \|f\|_{Lip(\Pi_X, \Gamma)} \|x - y\|^{(\gamma_i - s_m) p_{\max}}.$$

The definition of approximating polynomial one-form is adapted to the case of inhomogeneous smoothness from [3].

Definition 3.5.1. *Let $f : W \to L(V, W)$ be a (Π_X, Γ)-Lipschitz one-form, Y an element in W and ρ a positive real. We define the* polynomial one-form approximating *f in the ball $B_\rho(Y)$ of radius ρ centered at Y by*

$$p_Y(y) = \sum_{i=1}^k \sum_{s_n = s_0 + s_l < \gamma_i} f_i^{s_n}(Y) \left(\sum_{deg_\Pi(R) = s_n} (y - Y)_R \right)$$

for all $y \in B_\rho(Y)$. Let the function $q_Y : W \times W \to L(W, L(V, W))$ be defined by

$$p_Y(x) - p_Y(y) = q_Y(x, y)(x - y).$$

Definition 3.5.2. *Let $z \in W$ and $(s, t) \in \Delta_T$. Consider the RDE*

$$\mathrm{d}\mathbf{Y}_t(z, s) = p_z(\mathbf{Y}_t(z, s))\mathrm{d}\mathbf{X}_t, \ \mathbf{Y}_s(z, s) = z. \tag{3.49}$$

Let $\widehat{D}^p\left[(s, t), z\right]$ denote the map assigning $\mathbf{Y}_{s,t}(z, s)$ to $\left[(s, t), z\right]$.

Proposition 3.5.1. *The operator $\widehat{D}^p\left[(s, t), z\right]$ is a local approximation (as defined in Definition 3.2.2) satisfying Condition 3.2.1, i.e. there exist $\hat{\gamma} > 1$ and $\tau > 0$ and C_R, such that*

$$\left\| \pi_R D\left[(s, t), z\right] - \pi_R \widehat{D}^p\left[(s, t), z\right] \right\| \le C_R \omega(s, t)^{\hat{\gamma}}$$

for all $R \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}$, for all $(s, t) \in [0, T]$, $\omega(s, t) \le \tau$ and $z \in W$, where C_R does not depend on (s, t) or z but depends on R, $\hat{\gamma}$ and $\|f\|_{Lip(\Pi_X, \Gamma)}$.

Proof. The proof is a version of some results of Caruana [3] adapted to the case of inhomogeneous degree of smoothness.

Since p_z is not globally Lipschitz, the Universal Limit Theorem does not guarantee the existence and uniqueness of the solution to the RDE (3.49) on the whole interval $[0, T]$. However, our aim is to use the solutions on short enough intervals. Let us choose a positive real ρ.

Step 1 Let us restrict p_z to $B_\rho(z)$. Then

$$\|p_z\|_{Lip(\Pi_X,\Gamma)} \leq \|f\|_{Lip(\Pi_X,\Gamma)}\left(1+\rho+\cdots+\rho^l\right)$$

where

$$l = \max\left\{\|R\| \,|R \in \mathcal{A}^k, deg_\Pi(R) < \max_{1\leq i\leq k}\gamma_i\right\}.$$

Let M be the constant depending on $\|p_z\|_{Lip(\Pi_X,\Gamma)}$ (and thus on $\|f\|_{Lip(\Pi_X,\Gamma)}$, ρ and γ), such that whenever $\mathbf{Z}$ is a Π-rough path in $B_\rho(z)$ controlled by $\widehat{\omega}$, $\widehat{\omega}(0, T) < 1$, then the Π-variation of $\int h_i(\mathbf{Z})\mathrm{d}\mathbf{Z}$, $i = 1, 2$ is controlled by $M\widehat{\omega}$, where h_1 and h_2 are the one-forms defined in the proof of the Scaling Lemma corresponding to p_z and q_z. Define $\tau_1 = \min\left\{1, M^{-\lfloor p_{\max}\rfloor}\right\}$ and let τ_2 be a positive real, such that

$$\sum_{i=1}^{N}\frac{M\tau_2^{\frac{1}{q_i}}}{\beta^k\left(\frac{1}{q_i}\right)!} \leq \rho$$

whenever $\omega(s, t) < \tau_2$.

Then the solution $\mathbf{Y}_{\cdot,\cdot}(z, s)$ exists on $[s, \hat{T}(s)]$ for $\omega(s, \hat{T}(s)) < \min\{\tau_1, \tau_2\}$; i.e. $\mathbf{Y}(z, s)$ does not leave $B_\rho(z)$ in the sense

$$\left\|\mathbf{Y}^W_{s,t}(z,s)\right\| \leq \sum_{i=1}^{N}\frac{M\omega(s,t)^{\frac{1}{q_i}}}{\beta^k\left(\frac{1}{q_i}\right)!} \leq \rho$$

whenever $t \in [s, \hat{T}(s)]$.

Step 2

The definition of p_z implies that for any positive $\hat{\rho}$,

$$\|f - p_z\|_{Lip(\Pi_X,\Gamma)} \leq \max_i \|f\|_{Lip(\Pi_X,\Gamma)}\widehat{\rho}^{\gamma_i p_{\max}}$$

on $B_{\hat{\rho}}(z)$.

By Step 1, if $t \in [s, \hat{T}(s)]$, then $\mathbf{Y}_{s,t}(z, s)$ is in the ball $B_{\hat{\rho}}(z)$ for

$$\hat{\rho} = \sum_{i=1}^{N}\frac{M\omega(s,t)^{\frac{1}{q_i}}}{\beta^k\left(\frac{1}{q_i}\right)!}.$$

Then Proposition 2.6.1 implies

$$\begin{aligned}
&\left\|\pi_R D\left[(s,t),z\right] - \pi_R \widehat{D}^p\left[(s,t),z\right]\right\| \\
&\leq C_R \max_{R\in\mathcal{A}^{\Pi_Y}_{\alpha_Y}} \left\{ \left(\max_i \|f\|_{Lip(\Pi_X,\Gamma)} \hat{\rho}^{\gamma_i p_{\max}}\right)\right. \\
&\qquad\qquad\left. + \left(\max_i \|f\|_{Lip(\Pi_X,\Gamma)} \hat{\rho}^{\gamma_i p_{\max}}\right)^{\|R\|} \right\} \frac{\omega(s,t)^{deg_{\Pi_Y}(R)}}{\beta^k \Gamma_{\Pi_Y}(R)} \\
&\leq C_R K_R \omega(s,t)^{\hat{\gamma}}
\end{aligned}$$

whenever $t \in [s, \hat{T}(s)]$, where $\hat{\gamma} = \min_i(\gamma_i + 1/p_{\max}) > 1$ and K_R depends on R, $\omega(0,T)$, $\|f\|_{Lip(\Pi_X,\Gamma)}$ and Γ. □

Remark 3.5.1. In the above theorem we only assumed that V and W were direct sums of Banach spaces, and we made no assumption of finite dimensionality. This implies that the rough polynomial approximations lead to local approximations in general.

3.5.2 ODE-based approximation of RDEs determined by polynomial vector fields

In the previous section, we showed that one can construct local approximations using local rough polynomial approximations on short time intervals. In this section, we show how one can construct numerical approximations of the solution to the local rough polynomial approximations such that the resulting scheme is a local approximation satisfying Condition 3.2.1, implying that the corresponding global scheme is convergent.

Our aim is to show that the Log-signature Theorems hold for the polynomial one-forms locally in the following sense.

Proposition 3.5.2 (Log-signature, rough-polynomial case)**.** *Let Y be an element in W and $s \in [0,T]$. Assume that the function f satisfies the conditions of the Universal Limit Theorem 2.5.1, implying that the RDE*

$$\mathrm{d}\mathbf{Y}_t(\xi,s) = f(\mathbf{Y}_t(\xi,s))\mathrm{d}\mathbf{X}_t,\ \mathbf{Y}_s(\xi,s) = \xi$$

possesses a unique solution for all $\xi \in W$ and $s \in [0,T]$.

Let $\widehat{V}_1, \ldots, \widehat{V}_k$ be the $W \to W$ functions satisfying $p_Y(y) = \left(\widehat{V}_1(y), \ldots, \widehat{V}_k(y)\right)$ for $y \in W$. Note that the coordinate functions of each $\widehat{V}_i(y)$ are polynomial functions of the coordinates of $(y-Y)$. For any tree-like set $\mathcal{B}$, let $\widehat{W}^{z,\mathcal{B}}_i$ denote the vector field on $U^{\mathcal{B}}$ generated by $\widehat{V}_i$, let $\widehat{\mathcal{W}}^{z,\mathcal{B}}$ denote the set $\left\{\widehat{W}^{z,\mathcal{B}}_1, \ldots, \widehat{W}^{z,\mathcal{B}}_k\right\}$ and let $\widehat{\Phi}^{z,\mathcal{B}}$ denote the algebra homomorphism $\Phi_{\widehat{\mathcal{W}}^{z,\mathcal{B}}}$.

There exists a positive $\hat{\tau}$ such that for all $(s,t) \in \Delta_T$ satisfying $\omega(s,t) < \hat{\tau}$ the element $y_{s,t} \in U$ defined by

$$y_{s,t} = \mathrm{Exp}\left(\widehat{\Phi}^{Y,\mathcal{A}^{\Pi_Y}_{\alpha_Y}}\left\{L^{(\Pi_X,\alpha_X)}_{s,t}(\mathbf{X})\right\}\right)(\mathbf{1}). \tag{3.50}$$

exists. Moreover each $R \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ there exist constants C_R such that

$$\left\| \mathbf{Y}_{s,t}^{R}(Y,s) - \pi_R y_{s,t} \right\| \leq C_R \sum_{\substack{Q,R \in \mathcal{A}_{\alpha_X}^{\Pi_X} \\ deg_{\Pi_X}(Q*R) > \alpha_X}} \omega(s,t)^{deg_{\Pi_x}(Q*R)} \tag{3.51}$$

where C_R depends on R and $\|f\|_{Lip(\Pi_X,\alpha_X)}$.

Proof. Let us choose a positive real ρ and let τ denote the corresponding positive real number determined in the proof of Proposition 3.5.1.

The proof of the Log-signature Theorem 3.4.1 is based on the comparison of the Taylor remainder term and the rough-Taylor remainder term. Considering intervals of length at most τ, we get the required bounds on the rough-Taylor remainder term just as in the original case.

A non-trivial difference arises when bounding the Taylor remainder term, as we have to be sure that the solution to the ODE exists and does not get to far. Thus one can bound the coordinate functions of the vector fields. Let us consider the case $\|R\| = 1$, i.e. when $\mathcal{B} = \mathcal{B}_1$.

We show that the function

$$\widehat{\Phi}^{Y,\mathcal{B}_1} \left\{ L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X}) \right\} \tag{3.52}$$

is bounded in $B_\rho(Y)$ for any positive ρ. For this purpose, recall the representation (3.3.5) of the coordinate functions of $\widehat{W}_R^{z,\mathcal{B}_1} \circ Id$, $R \in \mathcal{A}^{\Pi_X}$ proved in Lemma 3.3.5 and recall the representation (3.37)-(3.38) of (3.52).

In the previous section, we have seen that the coordinate functions of the $\widehat{V}_i$'s are polynomials of degree at most l, where

$$l = \max \left\{ \|R\| \, | R \in \mathcal{A}^k, deg_\Pi(R) < \max_{1 \leq i \leq k} \gamma_i \right\}.$$

Then by Lemma 3.3.5, the coordinate functions of $\widehat{W}^{Y,\mathcal{B}_1} \circ Id$ are polynomials of degree at most $L := l \max_{R \in \mathcal{A}_{\alpha_X}^{\Pi_X}} \|R\|$. Hence

$$\max_{x \in B_\rho(Y)} \left\| \widehat{W}^{Y,\mathcal{B}_1} \circ Id(x) \right\| \leq C_1(1+\rho^L)$$

where C_2 depends only on $\|f\|_{Lip(\Pi_X,\Gamma)}$.

By equations (3.37)-(3.38), we have

$$\max_{x \in B_\rho(Y)} \left\| \widehat{\Phi}^{Y,\mathcal{B}_1} \left\{ L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X}) \right\}(x) \right\| \leq C_1(1+\rho^L) C_2 \left(\sum_{R \in \mathcal{A}_{\alpha_X}^{\Pi_X}} \omega(s,t)^{deg_{\Pi_X}(R)} \right) \tag{3.53}$$

where C_1 depends only on Π_X and α_X.

Let $\hat{v}_{\cdot}$ be the solution to the following ODE in $B_\rho(Y)$:

$$\mathrm{d}v_u = \widehat{\Phi}^{Y,\mathcal{B}_1}\left\{L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X})\right\}(v_u)\mathrm{d}u,\ v_0 = 0.$$

The inequality (3.53) implies the existence of a solution to the above ODE, however the solution might leave $B_\rho(Y)$ for $0 < u < 1$.

Let us choose τ_3 to satisfy

$$\left(\sum_{R\in\mathcal{A}_{\alpha_X}^{\Pi_X}} \tau_3^{deg_{\Pi_X}(R)}\right) \leq \frac{\rho}{C_1(1+\rho^L)C_2}.$$

Such τ_3 trivially exists.

Then for $(s,t)\in\Delta_t$, $\omega(s,t) < \tau_3$, we can determine a bound on v_u, $u\in[0,1]$ to be

$$\|v_u\| \leq u\rho$$

implying that the solution path on the interval $[0,1]$ does not leave $B_\rho(Y)$.

Define $\hat{\tau} = \min\{\tau,\tau_3\}$. Then using the above estimates, and the exact same arguments as in the proofs of the Log-signature Theorems (i.e. induction on the parameter m of $U^{\mathcal{B}_m}$) on subintervals $[s,t]$ of $[0,T]$ satisfying $\omega(s,t) < \hat{\tau}$, the assertion is proved. □

Corollary 3.5.1. *Assume that the one-form f satisfies the conditions of the Universal Limit Theorem 2.5.1. Let α_X satisfy*

$$\gamma = \min_{R\in\mathcal{A}^k\backslash\mathcal{A}_{\alpha_X}^{\Pi_X}} deg_{\Pi_X}(R) > 1.$$

Let $\widehat{D}^{p,ODE}$ denote the local approximation determined by (3.50), *i.e. the one based on the solution to the ODE driven by the vector field derived from the log-signature $L_{s,t}^{(\Pi_X,\alpha_X)}(\mathbf{X})$ using the algebra homomorphism $\widehat{\Phi}$ corresponding to the local polynomial approximations of the vector fields $V_1,\ldots,V_k$.*

Then $\widehat{D}^{p,ODE}$ satisfies Condition 3.2.1 with $\min\{\gamma,\hat{\gamma}\} > 1$ *and* $\hat{\tau} > 0$, *where $\hat{\gamma}$ is determined in Proposition 3.5.1 and $\hat{\tau}$ is determined in Proposition 3.5.2.*

Finally, Corollary 3.5.1 and Theorem 3.2.2 imply the global convergence of the global scheme based on the local approximation $\widehat{D}^{p,\mathrm{ODE}}$.

Remark 3.5.2. One might prefer to use polynomial vector fields in practice because

(i) polynomial vector fields are easy to implement,

(ii) it is easy to implement the Lie-bracket operation for polynomial vector fields,

(iii) and specialized numerical solvers exists for ODEs driven by polynomial vector fields.

CHAPTER 4

NUMERICAL SOLUTIONS OF STOCHASTIC DIFFERENTIAL EQUATIONS

In this chapter we consider Stratonovich stochastic differential equations on $\mathbb{R}^N$ of the type

$$\left.\begin{array}{rcl} d\xi_t & = & V_0(\xi_t)dt + \sum_{i=1}^k V_i(\xi_t) \circ dB_t^i \\ \xi_0 & = & x_0 \end{array}\right\} \tag{4.1}$$

where $B_t = (B_t^1, \dots, B_t^k)^T$ denotes a k-dimensional Brownian motion and the coefficient functions V_i, $i = 0, \dots, d$ are smooth and bounded $\mathbb{R}^N \to \mathbb{R}^N$ functions with bounded derivatives up to a certain order (to be specified later).

We consider (a.s.) path-wise approximations of the solution to the above SDE on an interval $[0, T]$ (assigning an element $\widehat{X}_T(\mathcal{D}) \in \mathbb{R}^N$ to each partition $\mathcal{D}$ of $[0, T]$). An approximation is said to have *strong order* γ if

$$\mathbb{E}\left[\left\|\xi_T - \widehat{X}_T(\mathcal{D})\right\|^q\right]^{1/q} < C|\mathcal{D}|^\gamma$$

for some norm $\|\cdot\|$ on $\mathbb{R}^n$ and integer q where $|\mathcal{D}|$ denotes the mesh size of the partition $\mathcal{D}$.

The aim of this chapter is to develop high order numerical schemes path-wise approximating the solution's signature on $[0, T]$. The construction of the schemes is based on techniques somewhat similar to the ones used in Chapter 3.

4.1 PRELIMINARIES

In this chapter, we will use the notation $V = \mathbb{R}^{k+1}$ and $W = \mathbb{R}^N$. The objects considered in this chapter are indexed with multi-indices from the set $\mathcal{A} = \bigcup_{k\geq 0} \{0, 1, \dots, k\}$. If $R \in \mathcal{A}$, then the length of R is denoted by $\|R\|$. We define the degree of $R = (r_1, \dots, r_l) \in \mathcal{A}$ by

$$deg(R) := deg(r_1, \dots, r_l) := \frac{\|R\| + \mathrm{card}\{r_i = 0 \,|1 \leq i \leq l\}}{2}.$$

We work with objects similar to those introduced in Chapter 3, formally corresponding to the $(k+1)$-tuple $\Pi_X = (p_0, \dots, p_1) = (1, 1/2, \dots, 1/2)$ and N-tuple $\Pi_Y = (q_1, \dots, q_N) = (1/2, \dots, 1/2)$. We will also use the notation $\mathcal{A}^{\Pi_X} = \mathcal{A}$ and $\mathcal{A}^{\Pi_Y}$.

We fix the positive reals α_X and α_Y, and recall the definition of the truncated tensor algebras $T^{(\Pi_X,\alpha_X)}(V)$ and $T^{(\Pi_Y,\alpha_Y)}(W)$.

Let $\mathcal{B}_i$ denote the set $\{R \in \mathcal{A}_{\alpha_Y}^{\Pi_Y} \mid \|R\| = i\}$. Let $\mathcal{B} \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ be a tree-like set (ref. Definition 3.2.1). Let us recall Definition 3.3.9 of the set of vector fields $\mathcal{W}^{z,\mathcal{B}} := \left\{W_0^{z,\mathcal{B}}, \ldots, W_k^{z,\mathcal{B}}\right\}$ where $z \in W$. Let $\Phi^{z,\mathcal{B}}$ denote the corresponding algebra homomorphism.

4.2 Estimates of Stratonovich integrals

In this section we recall some properties of the Stratonovich and the Itô integrals. Let $\xi_\cdot$ be the solution to the SDE (4.1).

Now recall, that given two continuous semi-martingales X_t and Y_t, the Stratonovich integral satisfies

$$\int^\cdot X \circ \mathrm{d}Y = \int^\cdot X\mathrm{d}Y + \frac{1}{2}\langle X, Y\rangle_\cdot, \tag{4.2}$$

where $\int^\cdot X\mathrm{d}Y$ denotes the Itô integral and $\langle X, Y\rangle_\cdot$ is the cross-variation process.

To simplify the notation, we introduce $B_t^0 := t$; i.e. the 0th coordinate of the Brownian motion is the time. One collection of basic objects in the paper is given by the Stratonovich *iterated integrals*.

Definition 4.2.1. *Given a multi-index $R = (r_1, \ldots, r_k) \in \mathcal{A}_{\alpha_X}^{\Pi_X}$ we define the* Stratonovich iterated integrals *as*

$$B_{s,t}^R := \int_{s<t_1<\cdots<t_k<t} \circ\mathrm{d}B_{t_1}^{r_1} \circ \cdots \circ \mathrm{d}B_{t_k}^{r_k} \tag{4.3}$$

$$B_{s,t}^R(Y_\cdot) := \int_{s<t_1<\cdots<t_k<t} Y_{t_1} \circ \mathrm{d}B_{t_1}^{r_1} \circ \cdots \circ \mathrm{d}B_{t_k}^{r_k} \tag{4.4}$$

for an integrable process Y_t.

The analogous Itô iterated integrals *are denoted by $D_{t,s}^R$ and $D_{t,s}^R(Y_\cdot)$ respectively.*

We will mainly work with Stratonovich iterated integrals, but for some cases we will need the Itô forms. To derive a general Stratonovich-Itô transformation we need the following relation between multi-indices.

Definition 4.2.2. *Let us regard a partition $(R_1, \ldots, R_l) \in \mathcal{A}_{\alpha_X}^{\Pi_X}$ of a multi-index R, i.e. $R = R_1 * \cdots * R_l$ for some l such that $deg(R_i) \leq 1$ for $i = 1, \ldots, l$. and suppose there exists a multi-index Q which can be partitioned into sub-indices $Q_1, \ldots, Q_l$ such that $Q = Q_1 * \cdots * Q_l$ and for each $i = 1, \ldots, l$ either $R_i = Q_i$ with both length 1 or $R_i = (j, j)$ and $Q_i = (0)$ for some $j \in \{1, \ldots, k\}$. Then we will say that Q is* related to R through the partitions $(Q_1, \ldots, Q_l)$ and $(R_1, \ldots, R_l)$ *and denote this relationship by*

$$R \sim Q.$$

Note that if $R \sim Q$, it is only possible through one particular pair of partitions $(Q_1, \ldots, Q_l)$ and $(R_1, \ldots, R_l)$.

We define the function $\nu :\in \mathcal{A}^{\Pi_X} \times \mathcal{A}^{\Pi_X} \to \mathbb{N}$ by

$$\nu(R,Q) := \begin{cases} card\{i\,|1 \le i \le l,\ R_i \neq Q_i\} & \text{if } R \sim Q \text{ through } (Q_1,\ldots,Q_l) \text{ and } (R_1,\ldots,R_l) \\ 0 & \text{otherwise} \end{cases}$$

Example 4.2.1. Let $R = (0,0,1,2,2,2)$, the following multi-indices are related to R:

$$(0,0,1,2,0),(0,0,1,0,2),(0,0,1,2,2,2).$$

Applying the definition of the Stratonovich integral, we can derive

$$\int_0^t \int_0^s \circ \mathrm{d}B_u^i \circ \mathrm{d}B_s^j = \int_0^t \int_0^s \mathrm{d}B_u^i \mathrm{d}B_s^j + \delta_{i,j}\frac{1}{2}\int_0^t \mathrm{d}B_s^0$$

where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. Using this assertion repeatedly, we get the following general formula:

Lemma 4.2.1. *For any multi-index $R \in \mathcal{A}^{\Pi_X}$*

$$B_{s,t}^R = \sum_{R\sim Q} \frac{1}{2^{\nu(R,Q)}} D_{s,t}^Q \tag{4.5}$$

Following Kloeden & Platen [15], we introduce the classes of functions $\mathcal{H}_R$ for $R \in \mathcal{A}^{\Pi_X}$.

Definition 4.2.3. *Let $\mathcal{H}_\epsilon$ denote the set of adapted RCLL almost surely bounded functions f. For $i \in \{0,1,\ldots,k\}$, let $\mathcal{H}_{(i)}$ denote the set of adapted RCLL functions f almost surely satisfying*

$$\int_0^t |f(s,\omega)|^{2-\delta_{i,0}} \mathrm{d}s < \infty, \text{ for } t \in [0,T].$$

For $R = (r_1,\ldots,r_l)$ and $\|R\| > 1$, $\mathcal{H}_R$ is defined recursively as the set of adapted RCLL functions satisfying

$$D_{0,t}^{R-}\{f(\cdot,\omega)\} \in \mathcal{H}_{(r_l)}.$$

Again, by (4.2), for the integrable process ξ_t and smooth function $f : \mathbb{R}^N \to \mathbb{R}$ such that $f(\xi_\cdot) \in \mathcal{H}_{(i)}$ for $i \in \{0,1,\ldots,k\}$ and $W_i^{\xi,\mathcal{B}_1} f(\xi_\cdot) \in \mathcal{H}_{(0)}$, we get

$$\int_0^t f(\xi_u) \circ \mathrm{d}B_u^i \quad = \quad \int_0^t f(\xi_u)\mathrm{d}B_u^i + (1-\delta_{i,0})\frac{1}{2}\int_0^t W_i^{\xi,\mathcal{B}_1} f(\xi_u)\mathrm{d}B_u^0.$$

This implies

Lemma 4.2.2. *Let $\xi_\cdot(z)$ denote the solution to the SDE* (4.1) *started at $z \in W$. For any multi-index $R \in \mathcal{A}^{\Pi_X}$ and a function $f : W \to \mathbb{R}$ such that $f(\xi_\cdot(z)) \in \mathcal{H}_Q$ for each $Q : R \sim Q$ and $W^{\xi,\mathcal{B}_1}_{r_1} f(\xi_\cdot) \in \mathcal{H}_{Q*(0)}$ for each $Q : -R \sim Q$, we have*

$$\begin{aligned} B^R_{s,t}(f(\xi_\cdot(z))) &= \sum_{R\sim Q} \frac{1}{2^{\nu(R,Q)}} D^Q_{s,t}(f(\xi_\cdot(z))) \\ &\quad +(1-\delta_{j_1,0})\frac{1}{2}\sum_{-R\sim Q} \frac{1}{2^{\nu(-R,Q)}} D^Q_{s,t}\Big(D^0_{s,\cdot}(W^{\xi,\mathcal{B}_1}_{r_1} f(\xi_\cdot(z)))\Big). \end{aligned} \tag{4.6}$$

The set of function satisfying the above conditions for all $z \in W$ is denoted by $\mathcal{H}_R(\xi_\cdot)$.

Remark 4.2.1. Note that the Itô form of a Stratonovich iterated integral of the type (4.6) corresponding to a multi-index R has a term corresponding to an all-zero multi-index of degree $\deg(R)$ if and only if $R = R_1 * \cdots * R_l$ such that either $R_i = (0)$ or $R_i = (j,j)$ for all $i = 1, \dots, l$ and for some $1 \le j \le k$. Note that in this case, *deg*(R) is an integer.

We will base our estimates on the following lemma.

Lemma 4.2.3. *Let $q \ge 1$ be an integer. There exists constants C depending only on q and T such that for $i \in \{0, \dots, k\}$ and $f(\xi_\cdot) \in \mathcal{H}_{(i)}$*

$$\mathbb{E}\left[\left\|\int_s^t f(\xi_u)\mathrm{d}B_i\right\|^{2q}\right]^{1/q} \le C(t-s)^{\delta_{i,0}} \int_s^t \mathbb{E}\left[\|f(\xi_u))\|^{2q}\right]^{1/q} \mathrm{d}u. \tag{4.7}$$

There exists a constant C_R depending only on the multi-index R, q and T such that

$$\begin{aligned} \mathbb{E}\left[|D^R_{s,t}|^{2q}\right]^{1/2q} &\le C_R(t-s)^{deg(R)} \\ \mathbb{E}\left[|D^R_{s,t}(f(\xi_\cdot))|^{2q}\right]^{1/2q} &\le C_R(t-s)^{deg(R)} \sup_{t\in[0,T]} \mathbb{E}\left[\|f(\xi_t)\|^{2q}\right]^{1/2q} \\ &\le C_R(t-s)^{deg(R)}\|f\|_\infty \end{aligned}$$

assuming that $f(\xi_\cdot)$ is a function in $\mathcal{H}_R$.

We refer the reader to Lemma 5.7.5 of [15] for proof.

The error of the numerical schemes we consider in this paper can be represented in terms of iterated integrals. The following lemma is the key step for determining an upper bound on the global error.

Lemma 4.2.4. *Let $q \ge 1$ be an integer. Let $\mathcal{D} = \{0 = t_0 < t_1 < \cdots < t_n \le T\}$ be a partition of $[0,T]$. For each multi-index $R \in \mathcal{A}^{\Pi_X}_{\alpha_X}$ where $\alpha_X > 1$, there exists a constant C_R depending only on R, q and T such that for any $f(\xi_\cdot)$ in $\mathcal{H}_R$, the following inequality holds*

$$\begin{aligned} \mathbb{E}\left[\sup_{1\le j\le l}\left(\sum_{i=1}^{j} D^R_{t_{i-1},t_i}(f(\xi_\cdot))\right)^{2q}\right]^{1/2q} &< C_R|\mathcal{D}|^m \sup_{t\in[0,T]} \mathbb{E}\left[\|f(\xi_t)\|^{2q}\right]^{1/2q} \\ &< C_R|\mathcal{D}|^m\|f\|_\infty \end{aligned} \tag{4.8}$$

where

(i) $m = deg(R) - 1$ *if* $deg(R)$ *is an integer*

(ii) $m = deg(R) - 1/2$ *otherwise.*

One can derive the proof for the above Lemma by applying Lemma 4.2.3 and Doob's inequality, exploiting the fact that in the case $deg(R)$ is not an integer, the terms of the sum on the left-hand-side of the expected value (4.8) are orthogonal. A detailed proof of the $q = 1$ case can be found in Chapter 10 of [15]. The proof of the $q > 1$ case is analogous.

The Stratonovich version of Lemma 4.2.4 is as follows.

Lemma 4.2.5. *Let $q \geq 1$ be an integer. Let $\mathcal{D} = \{0 = t_0 < t_1 < \cdots < t_n \leq T\}$ be a partition of $[0, T]$. For each multi-index $R \in \mathcal{A}$ there exists a constant C_R depending only on R, q, k and T, such that for all f in $\mathcal{H}_R(\xi_\cdot)$, the following inequality holds*

$$\begin{aligned}\mathbb{E}\left[\sup_{1\leq j\leq l}\left(\sum_{i=1}^{j} B^R_{t_{i-1},t_i}(f(\xi_\cdot))\right)^{2q}\right]^{1/2q} &< C_R|\mathcal{D}|^m \max_{deg(Q)\leq 1/2}\left\{\sup_{t\in[0,T]}\mathbb{E}\left[\|W_Q^{\xi,\mathcal{B}_1} f(\xi_t)\|^{2q}\right]^{1/2q}\right\}\\ &\leq C_R|\mathcal{D}|^m \max_{deg(Q)\leq 1/2}\left\{\|W_Q^{\xi,\mathcal{B}_1} f\|_\infty\right\} \qquad (4.9)\end{aligned}$$

where

(i) $m = deg(R) - 1$ *if an all zero multi-index is related to* R

(ii) $m = deg(R) - 1/2$ *otherwise.*

Proof. By the triangle inequality and Lemma 4.2.2

$$\begin{aligned}&\mathbb{E}\left[\left(\sup_{1\leq j\leq l}\sum_{i=1}^{l} B^R_{t_{i-1},t_i}(f(\xi_\cdot))\right)^{2q}\right]^{1/2q}\\ &\leq \sum_{R\sim Q}\mathbb{E}\left[\sup_{1\leq j\leq l}\left(\sum_{i=1}^{j}\frac{1}{2^{\nu(R,Q)}} D^Q_{t_{i-1},t_i}(f(\xi_\cdot))\right)^{2q}\right]^{1/2q}\\ &+(1-\delta_{0,r_1})\sum_{-R\sim Q}\mathbb{E}\left[\sup_{1\leq j\leq l}\left(\sum_{i=1}^{j}\frac{1}{2^{\nu(R-,Q)}} D^Q_{t_{i-1},t_i}\left(D^0_{t_{i-1},\cdot}(W^{\xi,\mathcal{B}_1}_{r_1} f(\xi_\cdot))\right)\right)^{2q}\right]^{1/2q}.\end{aligned}$$

Then Lemma 4.2.4 and Remark 4.2.1 imply the assertion. □

The following lemma is a special case of Lemma 4.2.5.

Lemma 4.2.6. *Let $q \geq 1$ be an integer. Let $\mathcal{D} = \{0 = t_0 < t_1 < \cdots < t_n \leq T\}$ be a partition of $[0, T]$. For each multi-index $R \in \mathcal{A}$ there exists a constant C_R depending only on R, q, k and T, such that for all f in $\mathcal{H}_R(\xi_\cdot)$, the following inequality holds*

$$\begin{aligned}\mathbb{E}\left[\left(\sup_{1\leq j\leq l}\sum_{i=1}^{j} f(\xi_{t_{i-1}})B^R_{t_{i-1},t_i}\right)^{2q}\right]^{1/2q} &< C_R|\mathcal{D}|^m \max_{deg(Q)\leq 1/2}\left\{\sup_{t\in[0,T]}\mathbb{E}\left[\|W_Q^{\xi,\mathcal{B}_1}f(\xi_t)\|^{2q}\right]^{1/2q}\right\}\\ &\leq C_R|\mathcal{D}|^m \max_{deg(Q)\leq 1/2}\left\{\|W_Q^{\xi,\mathcal{B}_1}f\|_\infty\right\}\end{aligned}$$

where

(i) $m = deg(R) - 1$ if an all zero multi-index is related to R

(ii) $m = deg(R) - 1/2$ otherwise.

Lemma 4.2.7. *Let $q \geq 1$ be an integer. Given the multi-indices $R_1, \ldots, R_l \in \mathcal{A}$, there exists a constant C depending only on $R_1 * \cdots * R_l$, T, l, k and q, such that*

$$\mathbb{E}\left[\left|\prod_{i=1}^{l} B^{R_i}_{s,t}\right|^{2q}\right]^{1/2q} < C(t-s)^{\sum_{i=1}^{l} deg(R_i)}$$

The lemma is a corollary of Lemma 4.2.2, Lemma 4.2.3 and the fact that for the multi-indices $Q = (q_1, \ldots, q_m)$ and $R = (r_1, \ldots, r_n)$

$$\begin{aligned}B^Q_{s,t}B^R_{s,t} &= \int_{s<u_1<\cdots<u_m<t} \circ dB^{q_1}_{u_1}\cdots\circ dB^{q_m}_{u_m}\int_{s<v_1<\cdots<v_n<t}\circ dB^{r_1}_{v_1}\cdots\circ dB^{r_n}_{v_n}\\ &= \sum_{\sigma\in\text{Shuffles}(m,n)}\int_{s<w_1<\cdots<w_{m+n}<t}\circ dB^{\sigma^{-1}(1)}_{w_1}\cdots\circ dB^{\sigma^{-1}(n+m)}_{w_{n+m}}\end{aligned}$$

where Shuffles(m, n) is the subset of all permutations of $\{1, 2, \ldots, m+n\}$ such that $\sigma \in$ Shuffles(m, n) if and only if

$$\sigma(1) < \sigma(2) < \cdots < \sigma(m) \text{ and } \sigma(m+1) < \sigma(m+2) < \cdots < \sigma(m+n).$$

4.3 Approximating terms corresponding to multi-indices in $\mathcal{B}_1$

In this section, we connect the asymptotic estimates of Ben Arous [1], Castell [4] and Hu [13] with the Rough Paths perspective. The integration theory used is the one by Itô (ref. e.g. Ikeda & Watanabe [14]) and the existence of the solution for the SDE (4.1) is implied by the well known results. The results are somewhat standard, and we refer the reader to Gard [11] and Kloeden & Platen [15] for further details and references.

Firstly, we construct approximation of the terms of the signature indexed by multi-indices of length 1. Then, in the next section we extend the approximations to the higher order terms of the signature by induction.

4.3.1 GLOBAL CONVERGENCE THEOREM

We consider numerical schemes on $[0, T]$ based on local approximations derived from the α_X-truncated log-signature of the driving noise. Let us denote $T^{(\Pi_Y,\alpha_Y)}(W)$ by U and recall the definition of $U^{\mathcal{B}}$ for a tree-like set $\mathcal{B}$.

Definition 4.3.1. *Let $\mathcal{B} \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ be a tree-like set.*

(i) The functions from $\Delta_T \times W \times U^{\mathcal{B}}$ to a $U^{\mathcal{B}}$-valued probability space are referred to as local approximations on $U^{\mathcal{B}}$.

(ii) For a positive integer i, let D_i denote the local approximation assigning $\mathbf{X}_{0,t}^{U^{\mathcal{B}_i}} - \mathbf{X}_{0,s}^{U^{\mathcal{B}_i}}$ to $((s,t), z, \mathbf{X}_{0,s}^{U^{\mathcal{B}_i}})$, where $\mathbf{X}_{0,\cdot}^{U^{\mathcal{B}_i}}$ is the solution to

$$\mathrm{d}\mathbf{X}_{0,t}^{U^{\mathcal{B}_i}} = \sum_{j=0}^{k} W_j^{z,\mathcal{B}_i} \circ Id(\mathbf{X}_{0,t}^{U^{\mathcal{B}_j}}) \circ \mathrm{d}B_t^i,\ \mathbf{X}_{0,s}^{U^{\mathcal{B}_i}} = x.$$

(iii) For a positive integer i, let $\widehat{D}_i$ denote the local approximation, assigning

$$\mathrm{Exp}\left(\Phi^{z,\mathcal{B}_i}\left\{L_{s,t}^{(\Pi_X,\alpha_X)}(B)\right\}\right)(x)$$

to $((s,t), z, x) \in \Delta_T \times W \times U^{\mathcal{B}_i}$

(iv) Given a local approximation $\overline{D}$ on $U^{\mathcal{B}}$ and a partition $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$, the corresponding global approximation scheme $\widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}}} = \widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}}}(\mathcal{D})$ is defined recursively by

- *(a)* $\widehat{\mathbf{X}}_{t_0,t_0}^{U^{\mathcal{B}}} = 0$
- *(b)* $\widehat{\mathbf{X}}_{t_0,t_1}^{U^{\mathcal{B}}} = \overline{D}[(t_0,t_1), \xi_{t_0}, 0]$
- *(c)* $\widehat{\mathbf{X}}_{t_i,t_{i+1}}^{U^{\mathcal{B}}} = \widehat{\mathbf{X}}_{t_0,t_i}^{U^{\mathcal{B}}} + \overline{D}[(t_i,t_{i+1}), \xi, \widehat{\mathbf{X}}_{t_0,t_i}^{U^{\mathcal{B}}}]$

(v) The local error *corresponding to the ith step of the scheme $\widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}}}(\mathcal{D})$ generated by the local approximation $\overline{D}$ is defined by*

$$E_{t_{i-1},t_i} = D[(t_i,t_{i+1}), \xi, \widehat{\mathbf{X}}_{t_0,t_i}^{U^{\mathcal{B}}}] - \overline{D}[(t_i,t_{i+1}), \xi, \widehat{\mathbf{X}}_{t_0,t_i}^{U^{\mathcal{B}}}].$$

(vi) The auxiliary process $\mathbf{Y} = \mathbf{Y}(\mathcal{D}) : [0,T] \to U^{\mathcal{B}}$ *associated with $\widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}}}(\mathcal{D})$ is introduced as follows:*

$$\mathbf{Y}_t = \begin{cases} \widehat{\mathbf{X}}_{t_0,t_i}^{U^{\mathcal{B}}} + \sum_{j=0}^{k} \int_{t_i}^{t} V^{\xi,\mathcal{B}}(\mathbf{Y}_u) \circ \mathrm{d}B_u^j & \text{if } t \in [t_i, t_{i+1}) \\ \widehat{\mathbf{X}}_{t_0,t_i}^{U^{\mathcal{B}}} + \sum_{j=0}^{k} \int_{t_i}^{t} V^{\xi,\mathcal{B}}(\mathbf{Y}_u) \circ \mathrm{d}B_u^j + E_{t_i,t_{i+1}} & \text{if } t = t_{i+1} \end{cases}$$

i.e. $\mathbf{Y}_t$ satisfies an SDE on each sub-interval $[t_i, t_{i+1})$.

We recall a version of Gronwall's inequality.

Lemma 4.3.1. *Let $s,t \in \mathbb{R}$, $s < t$ and suppose that b,c and r are $\mathbb{R} \to \mathbb{R}$ functions such that b is continuous, c is continuously differentiable and r is piece-wise continuous on $[s,t]$, and furthermore*

$$r(u) \leq c(u) + \int_s^u b(v)r(v)\mathrm{d}v.$$

Then

$$r(u) \leq c(u) + \int_s^u b(v)c(v)e^{[\int_v^u b]}\mathrm{d}v.$$

This version of the inequality is proved in the textbook [34].

Now we give a bound on the global error provided by a numerical scheme generated by pasting together local approximations.

Theorem 4.3.1 (Global convergence of strong approximations, the case of $\mathcal{B}_1$)**.** *Let $q \geq 1$ be an integer and $m \geq 1$ be a real number. Let $\overline{D}$ be a local approximation on $U^{\mathcal{B}_1}$ such that there exists a constant $C(q)$, not depending on the partition $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$ or n, satisfying*

$$\mathbb{E}\left[\sup_{0\leq j\leq n-1}\left\|\sum_{i=0}^{j} E_{t_i,t_{i+1}}\right\|^{2q}\right]^{1/2q} \leq C(q)|\mathcal{D}|^m. \tag{4.10}$$

Then there exists a constant C_1 depending only on k, q and linearly on $\|\nabla W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ and $\|\nabla W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $i = 0,\dots k$ and $j = 1,\dots,k$ such that the global error is bounded by

$$\mathbb{E}\left[\left\|\mathbf{X}_{0,T}^{U^{\mathcal{B}_1}} - \widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}_1}}(\mathcal{D})\right\|^{2q}\right]^{1/2q} \leq C(q)e^{TC_1}|\mathcal{D}|^m$$

for each partition $\mathcal{D}$ of $[0,T]$.

Proof. Note that the conditions on the vector fields imply that the SDE (4.1) possesses a unique strong solution. Consider the auxiliary process $\mathbf{Y}_\cdot = \mathbf{Y}_\cdot(\mathcal{D})$ introduced in Definition 4.3.1 and corresponding to the partition $\mathcal{D}$, the tree-like set $\mathcal{B}_1$ and the local approximation $\overline{D}$. Note that $\mathbf{X}_{0,T}^{U^{\mathcal{B}_1}} - \widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}_1}}(\mathcal{D}) = \xi_T - \xi - \mathbf{Y}_T$.

By definition, we have

$$\begin{aligned}\mathbb{E}\left[\|\xi_t - \xi - \mathbf{Y}_t\|^{2q}\right]^{1/2q} \leq\ & \mathbb{E}\left[\left\|\sum_{j=0}^{k}\int_0^t V_j^{\xi,\mathcal{B}_1}(\xi_s - \xi) - V_j^{\xi,\mathcal{B}_1}(\mathbf{Y}_s)\circ \mathrm{d}B_s^j\right\|^{2q}\right]^{1/2q} \\ & + \mathbb{E}\left[\sup_{0\leq j\leq n-1}\left\|\sum_{i=0}^{j} E_{t_i,t_{i+1}}\right\|^{2q}\right]^{1/2q}.\end{aligned} \tag{4.11}$$

Let us introduce the notation

$$r(t) := \mathbb{E}\left[\|\xi_t - \xi - \mathbf{Y}_t\|^{2q}\right]^{1/2q}.$$

By Lemma 4.2.2, Lemma 4.2.3 and the boundedness assumption on the coefficient functions and their derivatives, there exists a constant C_1 depending only on k, q and linearly on $\|\nabla W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ and $\|\nabla W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $i = 0, \dots k$ and $j = 1, \dots, k$ such that

$$\mathbb{E}\left[\left\|\sum_{j=0}^{k} \int_0^t V_j^{\xi,\mathcal{B}_1}(\xi_s - \xi) - V_j^{\xi,\mathcal{B}_1}(\mathbf{Y}_s) \circ \mathrm{d}B_s^j\right\|^{2q}\right]^{1/2q} \leq C_1 \int_0^t r(u)\mathrm{d}u.$$

Then applying Gronwall's lemma on the functions r, $b \equiv K$ and

$$c(t) := \mathbb{E}\left[\sup_{0\leq j\leq n-1,\ t_j\leq t}\left\|\sum_{i=0}^{j} E_{t_i,t_{i+1}}\right\|^{2q}\right]^{1/2q},$$

we get

$$\mathbb{E}\left[\|\xi_T - \xi - \mathbf{Y}_T\|^2\right] \leq c(T)e^{TC_1}.$$

The assertion is thus proved.

□

Corollary 4.3.1. *Let $q \geq 1$ be an integer and $m \geq 1$ a real number. Let $\overline{D}_1$ be a local approximation on $U^{\mathcal{B}_1}$ such that there exists a constant C not depending on (s,t) or z, such that*

$$\overline{D}_1\left[(s,t),z,x\right] - D_1\left[(s,t),z,x\right] = M_{s,t}(z,x) + N_{s,t}(z,x)$$

and

$$\begin{aligned}
\mathbb{E}\left[M_{s,t}(z,x)\middle|\mathcal{F}_s\right] &= 0 \\
\mathbb{E}\left[\|M_{s,t}(z,x)\|^{2q}\right]^{1/2q} &\leq C(t-s)^{m+1/2} \\
\mathbb{E}\left[\|N_{s,t}(z,x)\|^{2q}\right]^{1/2q} &\leq C(t-s)^{m+1}
\end{aligned}$$

for all $(s,t) \in \Delta_T$, $z \in W$ and $x \in U^{\mathcal{B}_1}$, where $(\mathcal{F}_s)_{s\geq 0}$ is the completed filtration generated by the Brownian motion B.

Then there exists a constant K depending only on C, k, q and $\|\nabla W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ and $\|\nabla W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $i = 0, \dots k$ and $j = 1, \dots, k$, such that the global error is bounded by

$$\mathbb{E}\left[\|\mathbf{X}_{0,T}^{U^{\mathcal{B}_1}} - \widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}_1}}(\mathcal{D})\|^{2q}\right]^{1/2q} \leq K|\mathcal{D}|^m,$$

for each partition $\mathcal{D}$ of $[0,T]$.

Proof. Note that

$$\left\|\sum_{j=0}^{l} E_{t_i,t_{i+1}}\right\|^{2q} \leq C_q\left\|\sum_{i=0}^{l} M_{t_i,t_{i+1}}\right\|^{2q} + C_q\left\|\sum_{i=0}^{l} N_{t_i,t_{i+1}}\right\|^{2q} \tag{4.12}$$

for a constant C_q depending only on q. Given the assumptions on $M_{s,t}$, the first term on the right hand side of (4.12) is a discrete time sub-martingale indexed by l. Then using the conditions of the Corollary and applying Doob's inequality and the orthogonality of the terms in the first sum on the right-hand-side of (4.12), the condition (4.10) of Theorem 4.3.1 is satisfied, and hence global convergence of order m is implied. □

4.3.2 THE STOCHASTIC LOG-SIGNATURE THEOREM

Recall the stochastic Taylor expansion (ref. [15]).

Lemma 4.3.2 (Stochastic-Taylor expansion). *Let $q \geq 1$ be an integer and $m \geq 1$ be a real number such that $2m$ is an integer. Let $\mathcal{B}$ be a tree-like set, let $\xi_.$ be the solution to the SDE* (4.1) *and let $X_t = \xi_t - \xi$. Assume that f is an $\mathbb{R}^N \to \mathbb{R}$ function, twice continuously differentiable. Then*

$$f(X_t) = f(X_s) + \sum_{deg(R)\leq m} B^R_{s,t} W^{\xi,\mathcal{B}}_R f(X_s) + \sum_{\substack{deg(-R)\leq m\\ deg(R)>m}} B^R_{s,t}\left\{W^{\xi,\mathcal{B}}_R f(X_.)\right\}. \tag{4.13}$$

In particular, applying (4.13) *coordinate-wise on the identity function,*

$$X_t = X_s + \sum_{deg(R)\leq m} B^R_{s,t} W^{\xi,\mathcal{B}}_R (X_s) + \sum_{\substack{deg(-R)\leq m\\ deg(R)>m}} B^R_{s,t}\left\{W^{\xi,\mathcal{B}}_R (X_.)\right\} \tag{4.14}$$

where the last term is referred to as the remainder and denoted by $\mathcal{R}^{stoch}_{s,t}(m)$.

Moreover, there exists a constant C depending only on k, m, q, and T, such that

$$\begin{aligned}
\mathbb{E}\left[\|\mathcal{R}^{stoch}_{s,t}(m)\|^{2q}\right]^{1/2q} &\leq C \sum_{\substack{deg(-R)\leq m\\ deg(R)=m+1/2}} (t-s)^{m+1/2}\mathbb{E}\left[\sup_{s\leq u\leq t}\|W^{\xi,\mathcal{B}}_R \circ Id(X_u)\|^{2q}\right]^{1/2q}\\
&\quad +C \sum_{\substack{deg(-R)\leq m\\ deg(R)=m+1}} (t-s)^{m+1}\mathbb{E}\left[\sup_{s\leq u\leq t}\|W^{\xi,\mathcal{B}}_R \circ Id(X_u)\|^{2q}\right]^{1/2q} &(4.15)\\
&\leq C \sum_{\substack{deg(-R)\leq m\\ deg(R)=m+1/2}} (t-s)^{m+1/2}\|W^{\xi,\mathcal{B}}_R \circ Id\|_\infty\\
&\quad +C \sum_{\substack{deg(-R)\leq m\\ deg(R)=m+1}} (t-s)^{m+1}\|W^{\xi,\mathcal{B}}_R \circ Id\|_\infty &(4.16)
\end{aligned}$$

The proof is based on the repeated application of Itô's lemma:

$$\mathrm{d}f(Y_t) = \sum_{i=0}^{d} W^{\xi,\mathcal{B}}_i(f(Y_t)) \circ \mathrm{d}B^i_t,$$

for any $\mathbb{R}^N \to \mathbb{R}$ function f twice continuously differentiable on $[s,t]$.

The numerical schemes introduced later in this chapter are based on ordinary differential equations driven by a vector field derived from the log-signature of the driving path. For this purpose, we introduce the following notation.

Definition 4.3.2. *Let n be a positive integer, α a positive real number, Π an n-tuple (either Π_X with $n=k$ or Π_Y with $n=N$), and let $X:[0,T]\to\mathbb{R}^n$ be an integrable semi-martingale. Let $\{\varepsilon_i, i\in I\}$ denote the canonical basis of $\mathbb{R}^n$. Let $T^{(\Pi,\alpha)}(\mathbb{R}^n)$ denote the truncated tensor algebra:*

$$T^{(\Pi,\alpha)}(\mathbb{R}^n) = \text{Span}\left\{\varepsilon_R = \varepsilon_{r_1}\otimes\cdots\otimes\varepsilon_{r_l} \;\middle|\; R=(r_1,\dots,r_l),\, deg(R)\le\alpha,\, l\in\mathbb{N}\right\}$$

We define the (random) Stratonovich signature *of X corresponding to the interval $[s,t]\subseteq[0,T]$ by*

$$\mathbf{X}^{\alpha}_{s,t} = \sum_{\substack{R=(r_1,\dots,r_l)\\ deg(R)\le\alpha\\ l\in\mathbb{N}}} \varepsilon_R \int_{s<u_1<\cdots<u_l<t} \circ \mathrm{d}X^{r_1}_{u_1}\cdots\circ\mathrm{d}X^{r_l}_{u_l}. \tag{4.17}$$

Moreover, let the (random) Stratonovich log-signature *corresponding to the interval $[s,t]\subseteq[0,T]$ be defined by*

$$L^{(\Pi,\alpha)}_{s,t}(X) = \log^{(\Pi,\alpha)}(\mathbf{X}^{\alpha}_{s,t})$$

where $\log^{(\Pi,\alpha)}$ denotes the logarithm function on $T^{(\Pi,\alpha)}(\mathbb{R}^n)$.

Remark 4.3.1. The Stratonovich log-signature is almost surely an element of the free Lie algebra generated by $\{\varepsilon_i, i\in I\}$ (ref.: Remark 4.6. of Lyons & Victoir [28]).

Theorem 4.3.2 (Stochastic log-signature theorem, the case of $\mathcal{B}_1$)**.** *Let $q\ge 1$ be an integer. Let z be an element of $\mathbb{R}^N$ and let m be a positive real number such that $2m$ is integer. Assume that $W^{\xi,\mathcal{B}_1}_R\circ Id$ are in $\mathcal{H}_R(\xi_\cdot)$ for each multi-index R of degree at most $2m$. Let $X_\cdot$ be the solution to the SDE*

$$\mathrm{d}X_u = \sum_{i=0}^{k} V_i(X_u+\xi)\circ\mathrm{d}B^i_u,\; X_s = z. \tag{4.18}$$

Let $y_{s,t}$ be defined by

$$y_{s,t} = \text{Exp}\left(\Phi^{\xi,\mathcal{B}_1}\left\{L^{(\Pi_X,m)}_{s,t}(B)\right\}\right)(z),$$

and let the local error be defined by $\mathcal{R}_{s,t}(m) = X_t - y_{s,t}$. Then there exists a constant C depending only on k, m, q, T and linearly on

$$\max_{deg(R)\le 2m}\left\|W^{\xi,\mathcal{B}_1}_R\right\|_\infty$$

such that

$$\mathbb{E}\left[\|\mathcal{R}_{s,t}(m)\|^{2q}\right]^{1/2q} \le C(t-s)^{m+1/2}. \tag{4.19}$$

Moreover, if m is integer, then

$$\mathbb{E}\left[\sum_{deg(R)=m+1/2}\pi_R\mathcal{R}_{s,t}(m)\middle|\mathcal{F}_s\right] = 0.$$

where $(\mathcal{F}_s)_{s\geq 0}$ denotes the completed filtration generated by the Brownian motion B.

Proof. **Step 1**

The first step is analogous to the proof of Theorem 3.4.1. Note that the conditions on the vector fields imply that the SDE (4.18) possesses a unique strong solution.

Recall, that the log-signature $L_{s,t}^{(\Pi_X,m)}$ can be represented in the form

$$L_{s,t}^{(\Pi_X,m)}(B) = \sum_{deg(R)\leq m} \left(\sum_{\substack{R_1*\cdots*R_l=R \\ l\in\mathbb{N}}} a_{R_1,\ldots,R_l} B_{s,t}^{R_1} \cdots B_{s,t}^{R_l} \right)$$

where each constant $a_{R_1,\ldots,R_l}$ depends only on $R_1,\ldots,R_l$ and k.

We introduce the notation

$$A_{s,t}^R = \sum_{\substack{R_1*\cdots*R_l=R \\ l\in\mathbb{N}}} a_{R_1,\ldots,R_l} B_{s,t}^{R_1} \cdots B_{s,t}^{R_l}.$$

Then we get

$$\Phi^{\xi,\mathcal{B}_1}\left\{ L_{s,t}^{(\Pi_X,m)}(B) \right\} = \sum_{deg(R)\leq m} A_{s,t}^R W_R^{\mathcal{B}_1} \circ Id_W.$$

Let v_u denote $\mathrm{Exp}\left(u\Phi^{\xi,\mathcal{B}_1}\left\{ L_{s,t}^{(\Pi_X,m)}(B) \right\}\right)(z)$, which exists and is unique due to the conditions of the theorem. With analogy to the proof of Theorem 3.4.1, we expand v_1 around v_0 up to remainder terms of degree at least $m+1/2$ and at most $2m$. The ODE-Taylor remainder term of the expansion is denoted by $\mathcal{R}_{s,t}^{\mathrm{Tay},\mathcal{B}_1}(m)$.

Recalling the definition of the stochastic Taylor expansion's remainder term $\mathcal{R}_{s,t}^{\mathrm{stoch},\mathcal{B}_1}(m)$, we get

$$X_t - y_{s,t} = \mathcal{R}_{s,t}^{\mathrm{stoch},\mathcal{B}_1}(m) - \mathcal{R}_{s,t}^{\mathrm{Tay},\mathcal{B}_1}(m).$$

Then Lemmas 4.2.5 and 4.2.7 imply the L^2 bound (4.19) on the local error.

Step 2

The last part is implied by the representation

$$\sum_{deg(R)=m+1/2} \pi_R(X_t - y_{s,t}) = \sum_{deg(R)=m+1/2} \Phi^{\xi,\mathcal{B}_1}\left\{ \pi_R(L_{s,t}(B)) \right\} \circ Id(X_s)$$

i.e. the degree $m+1/2$ level of the local error is the sum of Stratonovich iterated integrals and the fact that no all-zero multi-index is related to R if $deg(R)$ is not integer (ref. Remark 4.2.1) and by Lemma 4.2.7. □

Now we link the stochastic log-signature theorem to Corollary 4.3.1.

Corollary 4.3.2. *Let $q\geq 1$ be an integer. Assume that $\alpha_X \geq 1$. Let $\mathcal{D}$ be a partition of $[0,T]$ and let $\widehat{\mathbf{X}}_{\cdot,\cdot}^{U^{\mathcal{B}_1}}(\mathcal{D})$ denote the numerical scheme generated by the local approximation $\widehat{\widehat{D}}_1$ on $U^{\mathcal{B}_1}$, such that there exists a constant C not depending on (s,t), z or x, satisfying*

$$\mathbb{E}\left[\left\| \widehat{D}_1\left[(s,t),z,x\right] - \widehat{\widehat{D}}_1\left[(s,t),z,x\right] \right\|^{2q} \right]^{1/2q} \leq C(t-s)^{\lfloor \alpha_X \rfloor + 1}. \tag{4.20}$$

Furthermore, assume that $W_R^{\xi,\mathcal{B}_1} \circ Id$ *are in* $\mathcal{H}_R(\xi_\cdot)$ *for each multi-index R of degree at most* $2\lfloor \alpha_X \rfloor$.

Then there exists a constant C_1 *depending only on k, q, T and linearly on*

$$\max_{deg(R)\leq 2\lfloor \alpha_X \rfloor} \|W_R^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty,$$

and $\|\nabla W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ *and* $\|\nabla W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ *for* $i = 0, \ldots k$ *and* $j = 1, \ldots, k$, *such that the global error is bounded by*

$$\mathbb{E}\left[\|\mathbf{X}_{0,T}^{U^{\mathcal{B}_1}} - \widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}_1}}(\mathcal{D})\|^{2q}\right]^{1/2q} \leq K|\mathcal{D}|^{\lfloor \alpha_X \rfloor}.$$

Proof. Theorem 4.3.2 implies that the conditions of Corollary 4.3.1 are satisfied. □

Remark 4.3.2. When the scheme based on the local approximation $\widehat{D}$ is implemented using an ODE numerical solver, the inequality (4.20) determines a sufficient condition on the accuracy of the ODE scheme to preserve the global rate of convergence $\lfloor \alpha_X \rfloor$.

4.4 Approximating terms corresponding to multi-indices in $\mathcal{B}_h$ for $h > 1$

In this section, we construct approximations of the solution's signature by extending both the log-signature theorem and the global convergence theorem to the tree-like sets $\mathcal{B}_i$ $(i = 1, \ldots)$.

Theorem 4.4.1 (Global convergence of strong approximations, the case of $\mathcal{B}_h$, $h > 1$)**.** *Let* $q \geq 1$ *be an integer. Let* $m \geq 1$ *be a real number and h be a positive integer. Let* $\overline{D}_h$ *be a local approximation on* $U^{\mathcal{B}_h}$ *such that there exist a constant* $C(q)$ *not depending on the partition* $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$ *or n, satisfying*

$$\mathbb{E}\left[\sup_{0\leq j\leq n-1}\left\|\sum_{i=0}^{j} E_{t_i,t_{i+1}}\right\|^{4q}\right]^{1/4q} \leq C(q)|\mathcal{D}|^m. \tag{4.21}$$

Then there exists a constant $L(q)$ *depending only on C, k, q and on* $\|W_R^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ *for* $\|R\| \leq h+1$ *and* $\|\nabla W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ *and* $\|\nabla W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ *for* $i = 0, \ldots k$ *and* $j = 1, \ldots, k$, *such that the global error is bounded by*

$$\mathbb{E}\left[\left\|\mathbf{X}_{0,T}^{U^{\mathcal{B}_h}} - \widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}_h}}(\mathcal{D})\right\|^{2q}\right]^{1/2q} \leq L(q)|\mathcal{D}|^m. \tag{4.22}$$

for each partition $\mathcal{D}$ *of* $[0, T]$.

Remark 4.4.1. Note that the on the left-hand-side of inequality (4.21) the norm is the L^{4q} norm, whereas on the left-hand-side of (4.22) the norm is the L^{2q} norm. The extra regularity in (4.21) required due to technical reasons (see the proof of the theorem).

Proof. **Step 1**

Let $p \geq 1$ be an integer. Lemma 4.2.3 implies the existence of a constant $K_1(p)$ depending only on k, p, T and linearly on $\|W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ and $\|W_{(j,j)}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $i = 0, \ldots k$ and $j = 1, \ldots, k$ such that

$$\begin{aligned}
\mathbb{E}\left[\left\|\mathbf{X}_{0,s}^W\right\|^{2p}\right]^{1/2p} &= \mathbb{E}\left[\left\|\sum_{i=0}^k \int_0^s V_i\left(\mathbf{X}_{0,u}^W + \xi\right) \circ \mathrm{d}B_u^i\right\|^{2p}\right]^{1/2p} \\
&= \mathbb{E}\left[\left\|\sum_{i=0}^k \int_0^s W_i^{\xi,\mathcal{B}_1} \circ Id\left(\mathbf{X}_{0,u}^W\right) \mathrm{d}B_u^i + \frac{1}{2}\sum_{i=1}^k \int_0^s W_{(i,i)}^{\xi,\mathcal{B}_1} \circ Id\left(\mathbf{X}_{0,u}^W\right) \mathrm{d}u\right\|^{2p}\right]^{1/2p} \\
&\leq \sum_{i=0}^k \left\{Cs^{\delta_{i,0}} \int_0^s \mathbb{E}\left[\left\|W_i^{\xi,\mathcal{B}_1} \circ Id\left(\mathbf{X}_{0,u}^W\right)\right\|^{2p}\right]^{1/p} \mathrm{d}u\right\}^{1/2} \\
&\qquad + \sum_{i=1}^k \left\{Cs \int_0^s \mathbb{E}\left[\left\|W_{(i,i)}^{\xi,\mathcal{B}_1} \circ Id\left(\mathbf{X}_{0,u}^W\right)\right\|\right]^{1/p} \mathrm{d}u\right\}^{1/2} \\
&\leq K_1(p) s^{1/2}
\end{aligned}$$

for all $s \in [0, T]$, where C is the constant introduced in Lemma 4.2.3.

Using Lemma 4.2.3 recursively, one can prove the existence of a constant $K_h(p)$ depending only on k, p, T, and $\|W_R^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $\|R\| \leq h+1$, such that

$$\begin{aligned}
\mathbb{E}\left[\left\|\mathbf{X}_{0,s}^Q\right\|^{2p}\right]^{1/2p} &= \mathbb{E}\left[\left\|\sum_{i=0}^k \int_0^s \mathbf{X}_{0,s}^{Q-} V_i^{q_h}\left(\mathbf{X}_{0,u}^W + \xi\right) \circ \mathrm{d}B_u^i\right\|^{2p}\right]^{1/2p} \\
&= \mathbb{E}\Bigg[\Bigg\|\sum_{i=0}^k \int_0^s \mathbf{X}_{0,s}^{Q-} V_i^{q_h}\left(\mathbf{X}_{0,u}^W + \xi\right) \mathrm{d}B_u^i \\
&\qquad + (1-\delta_{0,q_{h-1}})\frac{1}{2}\int_0^s \mathbf{X}_{0,s}^{Q--} V_i^{q_{h-1}}\left(\mathbf{X}_{0,u}^W + \xi\right) V_i^{q_h}\left(\mathbf{X}_{0,u}^W + \xi\right) \mathrm{d}u \\
&\qquad\qquad + \frac{1}{2}\sum_{i=1}^k \int_0^s \mathbf{X}_{0,s}^{Q-} W_i^{(\xi,\mathcal{B}_1)}\left(V_i^{q_h}\left(\mathbf{X}_{0,u}^W + \xi\right)\right) \mathrm{d}u\Bigg\|^{2p}\Bigg]^{1/2p} \\
&\leq \sum_{i=0}^k \left\{Cs^{\delta_{i,0}} \int_0^s \mathbb{E}\left[\left\|\mathbf{X}_{0,s}^{Q-} V_i^{q_h}\left(\mathbf{X}_{0,u}^W + \xi\right)\right\|^{2p}\right]^{1/p} \mathrm{d}u\right\}^{1/2} \\
&\quad + (1-\delta_{0,q_{h-1}})\left\{Cs \int_0^s \mathbb{E}\left[\left\|\mathbf{X}_{0,s}^{Q--} V_i^{q_{h-1}}\left(\mathbf{X}_{0,u}^W + \xi\right) V_i^{q_h}\left(\mathbf{X}_{0,u}^W + \xi\right)\right\|^{2p}\right]^{1/p} \mathrm{d}u\right\}^{1/2} \\
&\quad + \sum_{i=1}^k \left\{Cs \int_0^s \mathbb{E}\left[\left\|\mathbf{X}_{0,s}^{Q-} W_i^{(\xi,\mathcal{B}_1)}\left(V_i^{q_h}\left(\mathbf{X}_{0,u}^W + \xi\right)\right)\right\|^{2p}\right]^{1/p} \mathrm{d}u\right\}^{1/2} \\
&\leq K_h(p) s^{h/2}
\end{aligned}$$

for all $s \in [0, T]$ and multi-index $Q = (q_1, \ldots, q_h) \in \mathcal{A}_{\alpha_Y}^{\Pi_Y}$ of length h (formally defining $\mathbf{X}_{0,s}^{Q--} = 0$ if $\|Q\| < 2$).

Step 2

Let us fix a partition $\mathcal{D}$ of $[0, T]$. Let $\mathbf{Y}$ be the auxiliary process corresponding to the tree-like set $\mathcal{B}_h$ and the partition $\mathcal{D}$. Then the following inequality is implied for any multi-index $R = (r_1, \ldots, r_l)$ with length l at least 2:

$$\begin{aligned}
&\mathbb{E}\left[\left\|\sum_{j=0}^{k}\int_0^t \mathbf{X}_{0,s}^{R-} V_j^{r_l}(\mathbf{X}_{0,s}^W + \xi) - \mathbf{Y}_{0,s}^{R-} V_j^{r_l}(\mathbf{Y}_{0,s}^W + \xi) \circ \mathrm{d}B_s^j\right\|^{2q}\right]^{1/2q} \\
&\qquad \le \mathbb{E}\left[\left\|\sum_{j=0}^{k}\int_0^t \mathbf{X}_{0,s}^{R-} \left(V_j^{r_l}(\mathbf{X}_{0,s}^W + \xi) - V_j^{r_l}(\mathbf{Y}_{0,s}^W + \xi)\right) \circ \mathrm{d}B_s^j\right\|^{2q}\right]^{1/2q} \\
&\qquad +\mathbb{E}\left[\left\|\sum_{j=0}^{k}\int_0^t \left(\mathbf{X}_{0,s}^{R-} - \mathbf{Y}_{0,s}^{R-}\right) V_j^{r_l}(\mathbf{Y}_{0,s}^W + \xi) \circ \mathrm{d}B_s^j\right\|^{2q}\right]^{1/2q} \qquad (4.23)
\end{aligned}$$

The first term on the right-hand side of (4.23) is bounded above as follows:

$$\begin{aligned}
&\mathbb{E}\left[\left\|\sum_{j=0}^{k}\int_0^t \mathbf{X}_{0,s}^{R-} \left(V_j^{r_l}(\mathbf{X}_{0,s}^W + \xi) - V_j^{r_l}(\mathbf{Y}_{0,s}^W + \xi)\right) \circ \mathrm{d}B_s^j\right\|^{2q}\right]^{1/2q} \\
&= \mathbf{E}\left[\left\|\sum_{j=0}^{k}\int_0^t \mathbf{X}_{0,s}^{R-} \left(V_j^{r_l}(\mathbf{X}_{0,s}^W + \xi) - V_j^{r_l}(\mathbf{Y}_{0,s}^W + \xi)\right) \mathrm{d}B_s^j\right.\right. \\
&\quad +(1-\delta_{0,r_{l-1}})\frac{1}{2}\int_0^t \mathbf{X}_{0,s}^{R--} V_j^{r_{l-1}}(\mathbf{X}_{0,s}^W + \xi)\left(V_j^{r_l}(\mathbf{X}_{0,s}^W + \xi) - V_j^{r_l}(\mathbf{Y}_{0,s}^W + \xi)\right) \mathrm{d}s \\
&\quad \left.\left. + \frac{1}{2}\sum_{j=1}^{k}\int_0^t \mathbf{X}_{0,s}^{R-} W_j^{\xi,\mathcal{B}_1}(\mathbf{X}_{0,s}^W + \xi)\left(V_j^{r_l}(\mathbf{X}_{0,s}^W + \xi) - V_j^{r_l}(\mathbf{Y}_{0,s}^W + \xi)\right) \mathrm{d}s\right\|^{2q}\right]^{1/2q} \\
&\le C_1\left\{\int_0^t \mathbb{E}\left[\left\|\mathbf{X}_{0,s}^{R-}\right\|^{4q}\right]^{1/2q} \mathbb{E}\left[\left\|\mathbf{X}_{0,s}^W - \mathbf{Y}_{0,s}^W\right\|^{4q}\right]^{1/2q} \mathrm{d}s\right\}^{1/2} \\
&\quad +C_1\left\{\int_0^t \mathbb{E}\left[\left\|\mathbf{X}_{0,s}^{R--}\right\|^{4q}\right]^{1/2q} \mathbb{E}\left[\left\|\mathbf{X}_{0,s}^W - \mathbf{Y}_{0,s}^W\right\|^{4q}\right]^{1/2q} \mathrm{d}s\right\}^{1/2}
\end{aligned} \qquad (4.24)$$

where C_1 depends only on q, k, T and on $\|W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$, $\|\nabla W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ and $\|\nabla W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $i = 0, \ldots k$ and $j = 1, \ldots, k$.

Similarly, a bound on the second term can be derived as follows:

$$\mathbb{E}\left[\left\|\sum_{j=0}^{k}\int_0^t \left(\mathbf{X}_{0,s}^{R-} - \mathbf{Y}_{0,s}^{R-}\right) V_j^{r_l}(\mathbf{Y}_{0,s}^{W} + \xi) \circ \mathrm{d}B_s^j\right\|^{2q}\right]^{1/2q}$$
$$\leq C_2 \left\{\int_0^t \mathbb{E}\left[\left\|\mathbf{X}_{0,s}^{R-} - \mathbf{Y}_{0,s}^{R-}\right\|^{2q}\right]^{1/q} \mathrm{d}s\right\}^{1/2}$$
$$+ C_2 \left\{\int_0^t \mathbb{E}\left[\left\|\mathbf{X}_{0,s}^{R--} - \mathbf{Y}_{0,s}^{R--}\right\|^{2q}\right]^{1/q} \mathrm{d}s\right\}^{1/2} \tag{4.25}$$

where C_2 depends only on q, k, T and on $\|W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ and $\|W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $i = 0, \ldots k$ and $j = 1, \ldots, k$.

Step 3

Now we prove the assertion by induction. The $h = 1$ case is equivalent to Theorem 4.3.1. Assume that $h \geq 2$ and for each $l \in \{1, \ldots, h-1\}$ there exists a constant L_l depending only on C, k and on $\|W_R^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $\|R\| \leq l+1$ and $\|\nabla W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ and $\|\nabla W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $i = 0, \ldots k$ and $j = 1, \ldots, k$, such that

$$\mathbb{E}\left[\left\|\mathbf{X}_{0,T}^{U^{\mathcal{B}_j}} - \widehat{\mathbf{X}}_{0,T}^{U^{\mathcal{B}_j}}(\mathcal{D})\right\|^{2q}\right]^{1/2q} \leq L_j |\mathcal{D}|^m$$

for each partition $\mathcal{D}$ of $[0, T]$. Then Step 1 and Step 2 imply

$$\mathbb{E}\left[\left\|\mathbf{X}_{0,T}^{R} - \mathbf{Y}_{0,T}^{R}(\mathcal{D})\right\|^{2q}\right]^{1/2q}$$
$$\leq \{C_1 K(2q)\,(K_{h-1}(2q) + K_{h-2}(2q)) + C_2(L_{h-1} + L_{h-2}) + C(q)\} \sqrt{T} |\mathcal{D}|^m,$$

where $L_0 := K_0 := 0$ and $K(2q)$ is the constant introduced in Theorem 4.3.1. □

Theorem 4.4.2 (Stochastic log-signature theorem, the case of $\mathcal{B}_h$, $h > 1$). *Let $q, h \geq 1$ be integers. Let $\mathbf{z}$ be a $U^{\mathcal{B}_h}$-valued $\mathcal{F}_s$-measurable random variable such that $\mathbb{E}\left[\|\mathbf{z}\|^{2q}\right] < \infty$ and $\pi_\epsilon \mathbf{z} = 1$. Let m be a positive real number such that $2m$ is integer. Assume that $W_R^{\xi,\mathcal{B}_1} \circ Id$ are in $\mathcal{H}_R(\xi_\cdot)$ for each multi-index R of degree at most $2m$. Let $\mathbf{X}_\cdot$ be the solution to the SDE*

$$\mathrm{d}\mathbf{X}_u = \sum_{i=0}^{k} V_i^{\xi,\mathcal{B}_h}(\mathbf{X}_u) \circ \mathrm{d}B_u^i,\ \mathbf{X}_s = \mathbf{z}. \tag{4.26}$$

Let $y_{s,t}$ be defined by

$$y_{s,t} = \operatorname{Exp}\left(\Phi^{\xi,\mathcal{B}_h}\left\{L_{s,t}^{(\Pi_X,m)}(B)\right\}\right)(\mathbf{z}),$$

and let the local error be defined by $\mathcal{R}_{s,t}(m) = \mathbf{X}_t - y_{s,t}$. Then there exists a constant C depending only on q, k, m, h, T and $\|W_R^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$, $deg(R) \leq 2\alpha_X$, such that

$$\mathbb{E}\left[\|\mathcal{R}_{s,t}(m)\|^{2q}\right]^{1/2q} \leq C(t-s)^{m+1/2}\left[\mathbb{E}\left[\left\|\pi_{U^{\mathcal{B}_{h-1}}}\mathbf{z}\right\|^{2q}\right]^{1/2q} + 1\right]. \tag{4.27}$$

Moreover, if m is integer, then

$$\mathbb{E}\left[\sum_{deg(R)\leq m+1/2}\pi_R\mathcal{R}_{s,t}(m)\big|\mathcal{F}_s\right]=0, \tag{4.28}$$

where $(\mathcal{F}_s)_{s\geq 0}$ denotes the completed filtration generated by the Brownian motion B.

Proof. The proof is again based on comparing the remainder term of the stochastic-Taylor expansion to the remainder term of the ODE-Taylor expansion.

Step 1

Recall that

$$\Phi^{\xi,\mathcal{B}_h}\left\{L_{s,t}^{(\Pi_X,m)}(B)\right\}=\sum_{R\in\mathcal{A}_m^{\Pi_X}}A_{s,t}^R W_R^{\xi,\mathcal{B}_h}\circ Id.$$

Let $v_u\in U^{\mathcal{B}_h}$ denote Exp $\left(u\Phi^{\xi,\mathcal{B}_h}\left\{L_{s,t}^{(\Pi_X,m)}(B)\right\}\right)(\mathbf{z})$ i.e.

$$v_u=\mathbf{z}+\int_0^u\Phi^{\xi,\mathcal{B}_h}\left\{L_{s,t}^{(\Pi_X,m)}(B)\right\}(v_z)\mathrm{d}z$$

$\pi_W v_u$ exists almost surely since $W_R^{\xi,\mathcal{B}_1}\circ Id$ are are in $\mathcal{H}_R(\xi_\cdot)$ for each multi-index R of degree at most m and since $\mathbf{z}$ is square integrable. Moreover, there exists a constant C_1 depending only on k, T, q and $\left\|W_R^{\xi,\mathcal{B}_1}\circ Id\right\|_\infty$ for $deg(R)\leq m$ such that

$$\mathbb{E}\left[\|v_u^W\|^{2q}\right]^{1/2q}\leq\mathbb{E}\left[\|\pi_W\mathbf{z}\|^{2q}\right]^{1/2q}+C_1u$$

By induction on l using Lemmas 3.3.5 and 4.2.7, one can prove that for any multi-index $R\in\mathcal{B}_l$

$$v_u^{\{R\}}=\mathbf{z}^{\{R\}}+\sum_{\|Q\|<\|R\|}\sum_{P\in\mathcal{A}_m^{\Pi_X}}A_{s,t}^P\int_0^u v_z^{\{Q\}}U(P,R,Q,v_u^W)\mathrm{d}z$$

where $U(P,R,Q,v_u^W)$ is a product of the coordinate functions of V_i^ξ for $i\in\{0,\ldots,k\}$ and their (multiple) partial derivatives. This implies the inequality for each $R\in\mathcal{B}_l$

$$\mathbb{E}\left[\left\|v_u^{\{R\}}\right\|^{2q}\right]^{1/2q}\leq C_l\left[\mathbb{E}\left[\|\pi_{U^{\mathcal{B}_l}}\mathbf{z}\|^{2q}\right]^{1/2q}+u\right]$$

where C_l is a constant depending only on k, q, T, l and $\left\|W_P^{\xi,\mathcal{B}_1}\circ Id\right\|_\infty$ for each $deg(P)\leq m$.

Step 2

Recall from the proof of Theorem 3.4.2 that the remainder term of the ODE-Taylor expansion is of the form

$$\mathcal{R}_{s,t}^{\mathrm{Tay},\mathcal{B}_h}(m)=$$

$$\sum_{\substack{R_1,\ldots,R_l\in\mathcal{A}_m^{\Pi_X}\\ deg(R_2*\cdots*R_l)\leq m\\ deg(R_1*R_2*\cdots*R_l)>m}}\left(\prod_{i=1}^l A_{s,t}^{R_i}\right)\int_{0<u_1<\cdots<u_l<1}W_{R_1}^{\xi,\mathcal{B}_h}\circ\cdots\circ W_{R_l}^{\xi,\mathcal{B}_h}\circ Id(v_{u_1})\mathrm{d}u_1\cdots\mathrm{d}u_l.$$

The combinations of Step 1 and Lemmas 3.3.5 and 4.2.7 imply that for any $Q \in \mathcal{B}_h$,

$$\mathbb{E}\left[\left\|\left(\prod_{i=1}^{l} A_{s,t}^{R_i}\right)\pi_Q \int_{0<u_1<\cdots<u_l<1} W_{R_1*\cdots*R_l}^{\xi,\mathcal{B}_h}\circ Id(v_{u_1})\mathrm{d}u_1\cdots\mathrm{d}u_l\right\|^2\right]^{1/2q} \leq C_Q(t-s)^{deg(R_1*\cdots*R_l)}\sum_{l<\|Q\|} C_l\left[\mathbb{E}\left[\left\|\pi_{U^{\mathcal{B}_{h-1}}}\mathbf{z}\right\|^{2q}\right]^{1/2q}+1\right]$$

where C_Q depends only on k, q, T, Q and $\left\|W_R^{\xi,\mathcal{B}_1}\circ Id\right\|_\infty$ for $deg(R)\leq 2m$.

Step 3

Similar arguments as in the previous step and the combination of Lemmas 3.3.5 and 4.2.7 and Step 1 of the proof of Theorem 4.4.1 imply that

$$\begin{aligned}\mathbb{E}\left[\left\|\mathcal{R}_{s,t}^{\mathrm{stoch},\mathcal{B}_h}(m)\right\|^{2q}\right]^{1/2q} &= \mathbb{E}\left[\left\|\sum_{\substack{deg(-R)\leq m\\ deg(R)>m}} B_{s,t}^R\left\{W_R^{\xi,\mathcal{B}_h}(\mathbf{X}_\cdot)\right\}\right\|^{2q}\right]^{1/2q}\\ &\leq K_h(t-s)^{(m+1/2)}\left[\mathbb{E}\left[\left\|\pi_{U^{\mathcal{B}_{h-1}}}\mathbf{z}\right\|^{2q}\right]^{1/2q}+1\right]\end{aligned}$$

where K_h only depends on k, q, T, and $\left\|W_R^{\xi,\mathcal{B}_1}\circ Id\right\|_\infty$ for $deg(R)\leq m+1$.

The inequality (4.27) is implied by Steps 2 and 3.

Step 4

The equality (4.28) is implied by the representation

$$\sum_{deg(R)=m+1/2}\pi_R(\mathbf{X}_t-y_{s,t}) = \sum_{deg(R)=m+1/2}\Phi^{\xi}\left\{\pi_R(L_{s,t}(B))\right\}\circ Id(\mathbf{X}_s)$$

i.e. by the fact that the degree $m+1/2$ level of the local error is the sum of Stratonovich iterated integrals and the fact that no all-zero multi-index is related to R if $deg(R)$ is not integer (ref. Remark 4.2.1) and by Lemma 4.2.7.

□

Now we connect Theorems 4.4.1 and 4.4.2.

Corollary 4.4.1. *Let $q\geq 1$ be an integer. Assume that $\alpha_X\geq 1$. Let $\mathcal{D}$ be a partition of $[0,T]$ and let $\widehat{\mathbf{X}}^{U}_{\cdot,\cdot}(\mathcal{D})$ denote the numerical scheme generated by the local approximation $\widehat{\widehat{D}}_\infty$ on U, such that there exists a constant C_D not depending on (s,t), z or x, satisfying*

$$\mathbb{E}\left[\left\|\widehat{D}_\infty[(s,t),z,x]-\widehat{\widehat{D}}_\infty[(s,t),z,x]\right\|^{2q}\right]^{1/2q}\leq C_D(t-s)^{\lfloor\alpha_X\rfloor+1} \tag{4.29}$$

Furthermore, assume that $W_R^{\xi,\mathcal{B}_1}\circ Id$ are in $\mathcal{H}_R(\xi_\cdot)$ for each multi-index R of degree at most $2\alpha_X$.

Then there exists a constant K depending only on C_D, k, q, T and on $\|W_R^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$, $deg(R) \leq 2\alpha_X$ and on $\|\nabla W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ and $\|\nabla W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $i = 0, \ldots k$ and $j = 1, \ldots, k$, such that the global error is bounded by

$$\mathbb{E}\left[\|\mathbf{X}_{0,T}^U - \widehat{\mathbf{X}}_{0,T}^U(\mathcal{D})\|^2\right]^{1/2} \leq K|\mathcal{D}|^{\lfloor \alpha_X \rfloor}.$$

Proof. We only need to prove that the local approximation $\widehat{\widehat{D}}_\infty$ satisfies the inequality (4.21) for an appropriate constant C. Theorem 4.4.2 implies (i.e. a combination of (4.27) and (4.28) imply) the inequality (4.21) if we can give a bound on the initial values of each step, i.e. on $\mathbb{E}\left[\left\|\widehat{\mathbf{X}}_{0,t_i}^{U_{h-1}^{\mathcal{B}}}\right\|^{4q}\right]^{1/4q}$ for each $i \in \{0, \ldots, n-1\}$. This can be done by induction on h. Let $h \geq 1$ be an integer. Assume that (4.21) is satisfied by the local approximation $\widehat{\widehat{D}}_h$ on $U^{\mathcal{B}_h}$. Then (4.29) and Theorems 4.4.1 and 4.4.2 imply the existence of a constant $K_h(2q)$ not depending on the partition $\mathcal{D}$ or n but only on C_D, k, q, T and $\|W_R^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$, $deg(R) \leq 2\alpha_X$ and on $\|W_R^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$, $deg(R) \leq 2\alpha_X$ and on $\|\nabla W_i^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ and $\|\nabla W_{j,j}^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$ for $i = 0, \ldots k$ and $j = 1, \ldots, k$, such that

$$\mathbb{E}\left[\left\|\mathbf{X}_{0,t_i}^{U_h^{\mathcal{B}}} - \widehat{\mathbf{X}}_{0,t_i}^{U_h^{\mathcal{B}}}\right\|^{4q}\right]^{1/4q} \leq K_h(2q)|\mathcal{D}|^m. \tag{4.30}$$

The $h = 1$ case is verified by Theorem 4.3.1.

Moving to $U^{\mathcal{B}_{h+1}}$, the required bound on the initial condition at each step results from the combination of the inequality

$$\mathbb{E}\left[\left\|\widehat{\mathbf{X}}_{0,t_i}^{U_h^{\mathcal{B}}}\right\|^{4q}\right]^{1/4q} \leq \mathbb{E}\left[\left\|\mathbf{X}_{0,t_i}^{U_h^{\mathcal{B}}}\right\|^{4q}\right]^{1/4q} + \mathbb{E}\left[\left\|\mathbf{X}_{0,t_i}^{U_h^{\mathcal{B}}} - \widehat{\mathbf{X}}_{0,t_i}^{U_h^{\mathcal{B}}}\right\|^{4q}\right]^{1/4q},$$

Step 1 of the proof of Theorem 4.4.1 and (4.30).

□

4.5 Examples

In this section, some numerical schemes are collected. Some of the schemes based on particular numerical solutions of locally derived ODEs are standard ones, appearing in monographs [11] and [15]. However, a wide range of schemes can be derived by choosing different ODE solvers.

Example 4.5.1. Consider the numerical schemes based on the following local approximations:

(i) $\overline{D}_1$: assigning

$$\mathrm{Exp}\left(\Phi^z\left\{\sum_{i=0}^{k}\varepsilon_i B_{s,t}^i\right\}\right)(x) \tag{4.31}$$

to $((s,t),z,x) \in \Delta_T \times W \times U$. This scheme is referred to as the *lowest order ODE-approach based scheme*.

(ii) $\overline{D}_1^{\mathrm{PC}}$: assigning the predictor-corrector scheme based numerical solution of the ODE (4.31) to $((s,t),z,x) \in \Delta_T \times W \times U$, in particular:

$$\begin{aligned} x^{\mathrm{pred}} &= z + \sum_{i=0}^{k} \mathbf{X}_{s,t}^{(i)} W_i^{\xi} \circ Id(x) \\ x^{\mathrm{new}} &= \frac{1}{2}\left[z + \sum_{i=0}^{k} \mathbf{X}_{s,t}^{(i)} W_i^{\xi} \circ Id(x^{\mathrm{pred}}) + x^{\mathrm{pred}} \right]. \end{aligned}$$

(iii) $\overline{D}_1^{\mathrm{Taylor}}$: assigning the two-step Taylor expansion based numerical solution of the ODE (4.31) to $((s,t),z,x) \in \Delta_T \times W \times U$, in particular

$$x^{\mathrm{new}} = z + \sum_{i=0}^{k} \mathbf{X}_{s,t}^{(i)} W_i^{\xi} \circ Id(x) + \frac{1}{2}\left(\sum_{i=1}^{k} \mathbf{X}_{s,t}^{(i)} W_i^{\xi} \right) \circ \left(\sum_{i=1}^{k} \mathbf{X}_{s,t}^{(i)} W_i^{\xi} \right) \circ Id(x).$$

Scheme (ii) is also referred to as the Heun scheme (ref. [11]). Scheme (iii) is equivalent to the 1-dimensional strong Milstein scheme when $k = 1$.

These schemes have global order 1 if $k = 1$ and global order $1/2$ if $k > 1$. This case is somewhat special because

(a) the terms $\frac{1}{2}(B_{s,t}^i)^2 W_{(i,i)}^{\xi} \circ Id(z) = B_{s,t}^{(i,i)} W_{(i,i)}^{\xi} \circ Id(z)$ are not part of the remainder term of the stochastic Taylor expansion since these are also included in the truncated Taylor expansion based representation of the ODE solution (these terms appear explicitly in scheme (iii)),

(b) moreover the rest of the degree 1 remainder terms (which correspond to the Lévy area) have conditional expectation equal to zero.

implying the inequality (4.21).

Remark 4.5.1. Recall from [6], that the maximum order of global convergence of the schemes depending only on the increment in the general case is 1 when $k = 1$ and $1/2$ when $k > 1$. However, if the vector fields commute, the global convergence rate is $1/2$ in the $k > 1$ case too.

Example 4.5.2. The general dimensional strong Milstein scheme corresponds to the $\alpha_X = 1$ case, when the ODE appearing in the local approximation $\widehat{D}_\infty$ is approximated by a truncated 1-step Taylor expansion based ODE numerical solver.

4.6 IMPLEMENTING HIGH ORDER SCHEMES

In practice, path-wise approximation of solution paths is often used in combination with weak approximations. For example, in case of path-dependent terminal condition, a standard approach is to randomly generate the Brownian input information, then approximate the solution paths along the simulated driving noise. Repeating this step for different input information, Monte-Carlo type methods can be constructed. To implement high order schemes in this case, the simulation of high degree Brownian iterated integrals is required. However, the exact simulation is complicated and numerically expensive. In this section, we investigate how to generate objects close enough to the truncated Brownian log-signature, such that the numerical schemes based on the ODEs derived from the *replacement* objects result in high order global convergence.

Since the simulation of Brownian increments is relatively cheap, the construction of the replacement object is based on the following definition.

Definition 4.6.1. *Let $\mathcal{D} = \{s = t_0 < \cdots < t_l = t\}$ be a partition of the interval $[s,t] \subseteq [0,T]$. The discretized Brownian log-signature corresponding to the partition $\mathcal{D}$ is defined by*

$$L_{s,t}(\mathcal{D}, B) = \log\left[\exp\left(\sum_{j=0}^{k} B^j_{t_0,t_1}\varepsilon_j\right) \otimes \cdots \otimes \exp\left(\sum_{j=0}^{k} B^j_{t_{l-1},t_l}\varepsilon_j\right)\right].$$

Note that the discretized log-signature is almost surely the signature of the random piece-wise linear path $\omega_t = (\omega^0_t, \ldots, \omega^k_t)^T$ of the form

$$\omega^j(\mathcal{D}, t) = \begin{cases} 0 & \text{if } t = 0 \\ \omega^j(\mathcal{D}, t_i) + \frac{t - t_i}{t_{i+1} - t_i} B^j_{t_i, t_{i+1}} & \text{if } t \in [t_i, t_{i+1}) \end{cases}$$

for $j = 0, 1, \ldots, k$.

Let $\mathbf{z}$ be an element in U such that $\pi_\epsilon \mathbf{z} = \mathbf{1}$, let ξ be an element in W and let $\mathcal{D}$ be a partition of $[s,t] \subseteq [0,T]$. Consider the random ODE

$$\mathrm{d}\mathbf{X}_{0,u}(\mathcal{D}) = \sum_{i=0}^{k} W^\xi_i \circ Id(\mathbf{X}_{0,u}(\mathcal{D}))\mathrm{d}\omega^i(\mathcal{D}, u),\ \mathbf{X}_{0,s}(\mathcal{D}) = \mathbf{z}. \tag{4.32}$$

and the SDE

$$\mathrm{d}\mathbf{X}_{0,u} = \sum_{i=0}^{k} W^\xi_i \circ Id(\mathbf{X}_{0,u}) \circ \mathrm{d}B^i_t,\ \mathbf{X}_{0,s} = \mathbf{z}. \tag{4.33}$$

Section 4.4 and in particular Example 4.5.1 imply that $\mathbf{X}_{0,t}(\mathcal{D})$ is the lowest order ODE-based approximation of $\mathbf{X}_{0,t}$; i.e. for each positive integer q, there exists a constant C_q not depending on $\mathcal{D}$ such that

$$\mathbb{E}\left[\|\mathbf{X}_{0,t} - \mathbf{X}_{0,t}(\mathcal{D})\|^{2q}\right]^{1/2q} \leq C_q|\mathcal{D}|^\gamma \tag{4.34}$$

where $\gamma = 1$ if $k = 1$ and otherwise $\gamma = 1/2$.

We will construct high order methods based on the following lemma.

Lemma 4.6.1. *Let $[s,t]$ be a subinterval of $[0,T]$ and $\alpha_X \geq 1$ real. Let $\widehat{D}^{\mathcal{D}}_\infty$ denote the local approximation on U depending on the partition $\mathcal{D}$ of $[s,t]$ assigning*

$$\mathrm{Exp}\left(\Phi^x\left\{L^{(\Pi_X,\alpha_X)}_{s,t}(\mathcal{D},B)\right\}\right)(z)$$

to $((s,t),x,z) \in \Delta_T \times W \times U$. Assume that $W^{x,\mathcal{B}_1}_R \circ Id$ are in $\mathcal{H}_R(\xi.)$ for each multi-index R of degree at most $2\lfloor\alpha_X\rfloor$. Let $\mathbf{z}$ be a $U^{\mathcal{B}_h}$-valued $\mathcal{F}_s$-measurable random variable such that $\mathbb{E}\left[\|\mathbf{z}\|^{2q}\right] < \infty$ and $\pi_\epsilon \mathbf{z} = 1$.

Then for each positive integer q, there exist a constant C depending on q, k, T, $\|W^{\xi,\mathcal{B}_1}_R \circ Id_W\|_\infty$, $deg(R) \leq 2\alpha_X$ and a positive real δ depending on $t-s$, such that

$$D\big[(s,t),x,\mathbf{z}\big] - \widehat{D}(\mathcal{D})\big[(s,t),x,\mathbf{z}\big] = M_{s,t}(x,\mathbf{z}) + N_{s,t}(x,\mathbf{z})$$

and

$$\begin{aligned}
\mathbb{E}\left[M_{s,t}(x,\mathbf{z})\middle|\mathcal{F}_s\right] &= 0\\
\mathbb{E}\left[\|M_{s,t}(x,\mathbf{z})\|^{2q}\right]^{1/2q} &\leq C(t-s)^{m+1/2}\\
\mathbb{E}\left[\|N_{s,t}(x,\mathbf{z})\|^{2q}\right]^{1/2q} &\leq C(t-s)^{m+1}
\end{aligned}$$

for all $z \in W$ and $x \in U$ whenever $|\mathcal{D}| \leq \delta$.

Proof. Let $D^{\mathcal{D}}_\infty$ denote the local approximation on U depending on the partition $\mathcal{D}$ of $[s,t]$ assigning the $\mathbf{X}_{0,t} - z$ to $((s,t),x,z) \in \Delta_T \times W \times U$. Considering the equality

$$D_\infty - \widehat{D}^{\mathcal{D}}_\infty = \left(D_\infty - D^{\mathcal{D}}_\infty\right) + \left(D^{\mathcal{D}}_\infty - \widehat{D}^{\mathcal{D}}_\infty\right),$$

we prove the assertion term-wise.

Step 1

Firstly, we derive a bound on $\left(D^{\mathcal{D}}_\infty - \widehat{D}^{\mathcal{D}}_\infty\right)$.

With analogy to the arguments leading to the stochastic Taylor expansion (4.14), one can prove that

$$\begin{aligned}
&\mathbf{X}_{0,t}(\mathcal{D}) =\\
&= \mathbf{X}_{0,s}(\mathcal{D}) + \sum_{deg(R)\leq m}\left\{\int_{s<u_1<\cdots<u_{\|R\|}<t} \mathrm{d}\omega^{r_1}(\mathcal{D},u_1)\cdots\mathrm{d}\omega^{r_{\|R\|}}(\mathcal{D},u_{\|R\|})\right\} W^{\xi}_R(\mathbf{X}_{0,s}(\mathcal{D}))\\
&\qquad + \sum_{\substack{deg(-R)\leq m\\ deg(R)>m}} \int_{s<u_1<\cdots<u_{\|R\|}<t} W^{\xi}_R(\mathbf{X}_{0,u_1})\mathrm{d}\omega^{r_1}(\mathcal{D},u_1)\cdots\mathrm{d}\omega^{r_{\|R\|}}(\mathcal{D},u_{\|R\|})
\end{aligned}$$

We refer to last term as the $\mathcal{D}$-discretized stochastic-Taylor expansion remainder, and use the notation $\mathcal{R}^{\mathrm{stoch}}_{s,t}(\mathcal{D},m)$.

Note that for $R = (r_1,\ldots,r_l)$,

$$\int_{s<u_1<\cdots<u_l<t} W^{\xi}_R(\mathbf{X}_{0,u_1})\mathrm{d}\omega^{r_1}(\mathcal{D},u_1)\cdots\mathrm{d}\omega^{r_l}(\mathcal{D},u_l)$$

is a lowest order ODE-based approximation of $B^R_{s,t}\left\{W^\xi_R(\mathbf{X}_\cdot)\right\}$; i.e. there exists a constant $C_1(q,m)$ not depending on $\mathcal{D}$, such that

$$\mathbb{E}\left[\left\|\mathcal{R}^{\text{stoch}}_{s,t}(m)-\mathcal{R}^{\text{stoch}}_{s,t}(\mathcal{D},m)\right\|^{2q}\right]^{1/2q}\leq C_1(q,m)|\mathcal{D}|^\gamma$$

and similarly for any R such that $deg(R)\leq 2\lfloor\alpha_X\rfloor$

$$\mathbb{E}\left[\left\|B^R_{s,t}-\int_{s<u_1<\cdots<u_{\|R\|}<t}\mathrm{d}\omega^{r_1}(\mathcal{D},u_1)\cdots\mathrm{d}\omega^{r_{\|R\|}}(\mathcal{D},u_{\|R\|})\right\|^{2q}\right]^{1/2q}\leq C_1(q,m)|\mathcal{D}|^\gamma$$

where $\gamma=1$ if $k=1$ otherwise $\gamma=1/2$.

Step 2

Adapting the arguments of Steps 1 and 2 of the proof of Theorem 4.4.2, one can prove that the remainder term of the Taylor expansion of $\operatorname{Exp}\left(\Phi^\xi\left\{L^{(\Pi_X,\lfloor\alpha_X\rfloor)}_{s,t}(\mathcal{D},B)\right\}\right)(\mathbf{z})$ satisfies

$$\mathbb{E}\left[\|\mathcal{R}^{\text{Tay}}_{s,t}(m)\|^{2q}\right]^{1/2q}\leq C_2\left((t-s)^{m+1/2}+|\mathcal{D}|^\gamma\right)\left[\mathbb{E}\left[\|\pi_{U^{\mathcal{B}_{h-1}}}\mathbf{z}\|^{2q}\right]^{1/2q}+1\right]$$

where C_2 depends on q,k,α_X,h,T and $\|W^{\xi,\mathcal{B}_1}_R\circ Id_W\|_\infty$, $deg(R)\leq 2\alpha_X$.

Step 3

Let $v_u(\mathcal{D})\in U$ denote $\operatorname{Exp}\left(u\Phi^\xi\left\{L^{(\Pi_X,\lfloor\alpha_X\rfloor)}_{s,t}(\mathcal{D},B)\right\}\right)(\mathbf{z})$. Then the degree $m+1/2$ level of the local error is represented by

$$\sum_{deg(R)=m+1/2}\pi_R(\mathbf{X}_t(\mathcal{D})-\mathbf{z}-v_1(\mathcal{D}))=\sum_{deg(R)=m+1/2}\Phi^\xi\left\{\pi_R(L_{s,t}(\mathcal{D},B))\right\}\circ Id(\mathbf{z}).$$

There exists a constant C_3 depending only on k,q,T and $\|W^{\xi,\mathcal{B}_1}_R\circ Id_W\|_\infty$, $\lfloor\alpha_X\rfloor+1/2\leq deg(R)\leq\lfloor\alpha_X\rfloor+1$, such that

$$\mathbb{E}\left[\left\|\sum_{deg(R)=m+1/2}\Phi^\xi\left\{\pi_R(L_{s,t}(B)-L_{s,t}(\mathcal{D},B))\right\}\circ Id(\mathbf{z})\right\|^{2q}\right]^{1/2q}\leq C_3|\mathcal{D}|^\gamma,$$

implying

$$\mathbb{E}\left[\left\|\sum_{deg(R)=m+1/2}\Phi^\xi\left\{L_{s,t}(\mathcal{D},B))\right\}\circ Id(\mathbf{z})\right\|^{2q}\right]^{1/2q}\leq C_4\left(|\mathcal{D}|^\gamma+(t-s)^{\lfloor\alpha_X\rfloor+1/2}\right),$$

where C_4 depends on k,q,T and $\|W^{\xi,\mathcal{B}_1}_R\circ Id_W\|_\infty$, $\lfloor\alpha_X\rfloor+1/2\leq deg(R)\leq\lfloor\alpha_X\rfloor+1$.

Step 4

Since $D_\infty^{\mathcal{D}}$ is the lowest order ODE-approach based approximation of D_∞, there exists a constant C_5 on k, q, T and $\|W_R^{\xi,\mathcal{B}_1} \circ Id_W\|_\infty$, $\lfloor \alpha_X \rfloor + 1/2 \leq deg(R) \leq 2$, such that

$$\mathbb{E}\left[\left\| D_\infty((s,t),x,z) - D_\infty^{\mathcal{D}}((s,t),x,z)\right\|^{2q}\right]^{1/2q} \leq C_5|\mathcal{D}|^\gamma$$

for all $(s,t) \in \Delta_T$, $x \in W$ and $z \in U$.

Step 5

By defining

$$\begin{aligned} M_{s,t}(x,\mathbf{z}) &= \sum_{deg(R)=m+1/2} \Phi^\xi\left\{\pi_R(L_{s,t}(B))\right\} \circ Id(\mathbf{z}) \\ N_{s,t}(x,\mathbf{z}) &= \left(D_\infty - D_\infty^{\mathcal{D}}\right) + \left(D_\infty^{\mathcal{D}} - \widehat{D}_\infty^{\mathcal{D}}\right) - M_{s,t}(x,\mathbf{z}) \end{aligned}$$

and by choosing the partition $\mathcal{D}$ to satisfy

$$\delta \leq (t-s)^{\frac{\lfloor \alpha_X \rfloor + 1}{\gamma}}, \tag{4.35}$$

Steps 1-4 imply the assertion.

□

4.6.1 Practical considerations

The implementation of the high order schemes based on the piece-wise linear approximation of the Brownian paths on a fine sub-scale requires the repeated computation of $\Phi^\xi\left\{L_{s,t}^{(\Pi_X,\alpha_X)}(\mathcal{D},B)\right\}$. To work out $L_{s,t}^{(\Pi_X,\alpha_X)}(\mathcal{D},B)$, one can use the Campbell-Baker-Hausdorff formula at each discretization step, however this is not efficient.

Furthermore, the principle given by (4.35) implies that shorter time intervals require a relatively finer sub-scale to ensure the high order convergence of our scheme. However, for a finer sub-scale the computation of the CBH formula increases rapidly and one loses the linear growth of computational expense with the linear growth of the number of steps.

Lemma 4.6.2. *Let $\mathcal{D} = \{s = t_0 < \cdots < t_l = t\}$ be a partition of $[s,t] \subseteq [0,T]$. The truncated discretized log-signature $L_{s,t}^{(\Pi_X,\alpha_X)}(\mathcal{D},B)$ written in a Lie basis $\{\ell_i,\ i \in \mathbb{N}\}$ is of the form*

$$L_{s,t}^{(\Pi_X,\alpha_X)}(\mathcal{D},B) = \sum p_i \ell_i \tag{4.36}$$

where each p_i is a polynomial in the variables

$$\{B_{t_{i-1},t_i}^j \mid j = 0,\ldots,k,\ i = 1,\ldots,l.\}$$

Note that the structure of (4.36) does not change if k is increased but the polynomials p_i have more variables. Hence to preserve efficiency, one can pre-compute either this formula or the polynomial coefficients therein for a number of independent Brownian paths. The pre-computation might take some time, however it only has to be done once. Ultimately, the reuse of the pre-computed (pseudo) random Lie elements results in a fast and high order numerical algorithm.

Since the formula (4.36) does not depend on the SDE but on the dimension and the choice of the sub-scale, one could create a universal database of pre-computed Lie elements usable for many SDEs.

4.6.2 A note on the computational expense

The implementation of the high order scheme presented in this section is after all a high order approximation of a lowest order scheme corresponding to a discretization on a finer sub-scale. One might ask if one can ever benefit from working out the high order scheme instead of just solving the lowest order scheme on the sub-scale. Let us do a brief cost-analysis to answer the question.

Let us assume that the computational cost of evaluating one step corresponding to the lowest order scheme is E_1 and the cost of evaluating one step of an m-truncated ODE based scheme is E_2 assuming that the truncated discretized log-signature is pre-computed, i.e. its computational cost is not part of E_2. In general E_1 is much smaller than E_2. If we apply the lowest order scheme on an interval with an l-substep fine scale, then the cost is lE_1. As long as lE_1 is smaller than E_2, the lowest order scheme is recommended. However, if one needs a more accurate approximation and wants to choose a finer sub-scale, then l will increase according to (4.35) and eventually the implementation of the higher order scheme will become more efficient than the corresponding low order scheme; i.e. $lE_1 > E_2$.

4.7 Numerical examples

In this section, some numerical examples for the ODE approach are presented. We demonstrate the efficiency of different ODE solvers and the high order methods. The primary aim is to estimate and compare the order of the strong convergence, however we also present some weak approximation results.

4.7.1 Estimating the order of convergence

Let us recall some standard results.

Definition 4.7.1. *Let us regard a discretization method assigning an approximation of* $\mathbf{X}_{0,T}$ *to each partition of* $[0, T]$. *If the discretization scheme is based on the partition* $\mathcal{D}$ *of* $[0, T]$, *then the resulting approximation is denoted by* $\widehat{X}_T(\Delta t)$.

Lemma 4.7.1. *Let $q \geq 1$ be an integer. Let $\mathcal{D}_n$ denote the partition $\{0 < \frac{T}{n} < \frac{2T}{n} < \cdots < T\}$. Let us suppose that a discretization scheme $\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)$ corresponding to the partition $\mathcal{D}_n$ approximates $\mathbf{X}_{0,T}$:*

$$\mathbb{E}\left[\left\|\mathbf{X}_{0,T} - \widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)\right\|^{2q}\right]^{1/2q} \to 0 \ \text{ as } n \to \infty$$

Furthermore, let us assume the existence of constants C_2 and γ not depending on $|\mathcal{D}_n|$ such that for $|\mathcal{D}_n| < 1$,

$$\mathbb{E}\left[\left\|\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_{2n}) - \widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)\right\|^{2q}\right]^{1/2q} \leq C_2|\mathcal{D}_n|^\gamma \tag{4.37}$$

is satisfied. Then for $|\mathcal{D}_n| < 1$

$$\mathbb{E}\left[\left\|\mathbf{X}_{0,T} - \widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)\right\|^{2q}\right]^{1/2q} \leq C_2|\mathcal{D}_n|^\gamma \frac{1}{1-(1/2)^\gamma}. \tag{4.38}$$

The proof of the lemma is based on the triangle inequality.

We apply Lemma 4.7.1 to estimate the order of convergence γ and an upper bound on the strong approximation error as follows. Estimating the expected value on the left hand side of (4.37) for a sequence of step lengths $T/2^n$ for $n = l, l+1, \ldots$ where $T/2^l < 1$ and fitting a log-regression, one can estimate C_2 and γ. If C_2 is estimated, then by (4.38) we get an estimate for an upper bound of the error at step size $|\mathcal{D}_n|$. When γ is known, we only estimate C_2.

One can observe the expected value on the left hand side of (4.37) very accurately, if the schemes corresponding to the different step sizes $|\mathcal{D}_n|$ and $|\mathcal{D}_{2n}|$ are run on the same Brownian paths. In that case, the random variables $\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)$ and $\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_{2n})$ are correlated and a simple Monte Carlo method results in a low variance and unbiased estimation of the expected value of their distance. Running schemes corresponding to different step lengths on the same paths can be achieved by using Lévy's construction of the Brownian paths (see e.g. [10] for details).

The idea of running schemes on the same Brownian paths is also useful when comparing two different discretization methods generating $\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)$ and $\overline{\mathbf{X}}_{0,T}(\mathcal{D}_n)$ respectively as approximations of $\mathbf{X}_{0,T}$. The high correlation between $\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)$ and $\overline{\mathbf{X}}_{0,T}(\mathcal{D}_n)$ results in a low variance for the Monte Carlo estimate of

$$\mathbb{E}\left[\left\|\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n) - \overline{\mathbf{X}}_{0,T}(\mathcal{D}_n)\right\|^{2q}\right]^{1/2q} \tag{4.39}$$

This comparison can be further applied as the triangle inequality implies.

Lemma 4.7.2. *Let two discretization methods producing $\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)$ and $\overline{\mathbf{X}}_{0,T}(\mathcal{D}_n)$ respectively as approximations of $\mathbf{X}_{0,T}$ be given. Let us suppose that there exists a constant C such that*

$$\mathbb{E}\left[\left\|\mathbf{X}_{0,T} - \overline{\mathbf{X}}_{0,T}(\mathcal{D}_n)\right\|^{2q}\right]^{1/2q} < C.$$

Then

$$\mathbb{E}\left[\left\|\mathbf{X}_{0,T}-\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)\right\|^{2q}\right]^{1/2q} \geq \mathbb{E}\left[\left\|\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)-\overline{\mathbf{X}}_{0,T}(\mathcal{D}_n)\right\|^{2q}\right]^{1/2q} - C. \tag{4.40}$$

Lemma 4.7.2 implies that if one can estimate a small enough upper bound C on the global error of $\overline{\mathbf{X}}_{0,T}(\mathcal{D}_n)$, then (4.40) gives an accurate estimate for the lower bound on the error of $\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)$. On the other hand, if no small enough bound C can be derived, the lower estimate (4.40) is meaningless.

In the following sections, for different SDEs and discretization methods we run simple Monte Carlo simulation based estimations of the following quantities:

(i) the estimated upper bound of the global error, i.e. (4.38)

(ii) the L^2 distance of different schemes, i.e. (4.39)

(iii) in some cases the estimated lower bound on the global error, i.e. (4.40)

(iv) and $\mathbb{E}\left[\widehat{\mathbf{X}}_{0,T}(\mathcal{D}_n)\right]$

for different step sizes. Each method was run on 10^6 paths.

Since each Monte Carlo simulation is based on sampling, the resulting estimate is a random variable with positive variance. Using this variance, when estimating (iv), we fit a 99% confidence interval centered at the resulting realization of the random variable. The length of these confidence intervals is proportional to the square root of the number of runs. In case of (i), (ii) and (iii), the calculated confidence intervals were very small.

The algorithm was implemented in C++. For the Lie algebra-level computations we used the libraries of the *CoRoPa*[1] project developed by D. Chafaï, T.J. Lyons et. al. For random number generation, we used the Boost[2] implementation of the Mersenne Twister generator, in particular the *mt19937* generator. We used our own code for transforming Lie algebra elements to vector fields, for manipulating vector fields and for solving ODEs. The fourth order Runge-Kutta scheme were taken from [2].

4.7.2 First order approximation of the CIR process

Firstly, we regarded simple SDEs, driven by one dimensional Brownian motion and tested the first order ODE approach with different ODE solvers as well as comparing it to the Euler-Maruyama scheme. The test results, presented in Figures 4.1 and 4.2, are the weak and strong approximation results respectively for the CIR process, i.e. in the Itô form

$$\mathrm{d}r_t = a(b-r)\mathrm{d}t + \sigma\sqrt{r}\mathrm{d}B \tag{4.41}$$

[1] Computational Rough Paths, website: http://coropa.sourceforge.net/

[2] Website: http://www.boost.org/

where a, b and σ are positive constants satisfying $ab/2 > \sigma^2$, which ensures that r_t is a.s. positive (ref. [24]).

We implemented the ODE approach with three different ODE solvers, namely the predictor-corrector, Runge Kutta order 4 and splitting. By splitting we mean the ODE solver in which W is written as $W := W_1 + W_2$ where

$$\begin{aligned} W_1(x) &:= B^0 \left[a(b-x) - \sigma^2/4\right] \\ W_2(x) &:= B^1 \sigma \sqrt{x} \end{aligned}$$

and at each step (starting at $\widehat{X}_{t_i}$), three ODEs are solved as follows:

$$\begin{aligned} x_1 &:= \operatorname{Exp}\left[\frac{1}{2}W_1\right](\widehat{X}_{t_i}) \\ x_2 &:= \operatorname{Exp}\left[W_2\right](x_1) \\ \widehat{X}_{t_{i+1}} &:= \operatorname{Exp}\left[\frac{1}{2}W_1\right](x_2) \end{aligned}$$

In case of the CIR SDE, the exact solution for the ODEs appearing in the splitting is known, and the computation is very fast. The splitting method is recommended for example in [33] and some nice weak approximation properties of it are presented as well.

The first order ODE approach implemented with the predictor-corrector ODE solver is referred to as the Heun scheme.

Remark 4.7.1. Note, that despite the guaranteed positivity of r_t, all of the tested schemes except the Splitting version can result in negative interest rates. In the case of the ODE-based method, the exact solution to the derived random ODEs preserves the positivity, but their numerical approximations might not. One possible way to overcome this difficulty is to adaptively reduce the ODE numerical solver's step size when the solver results in negative interest rates. In the case of splitting, the chosen combination of parameters guarantees positive solutions and no extra care is required. The Euler-Maruyama method is fixed by taking a positive value of the resulting interest rate at each step. We refer the reader to [24] for a review of simulation schemes approximating the CIR process.

Observation 4.7.1. In Figure 4.1, the weak approximation results, i.e. the confidence intervals corresponding to Monte Carlo estimations for $\mathbb{E}\left[\widehat{X}_T\right]$, are presented for different schemes implemented with different numbers of steps. The horizontal dashed line is the exact value of $\mathbb{E}\left[r_T\right]$. Given the number of runs (10^6), the variance of the evaluated estimate is relatively large, and at the weak level one cannot make a difference between the Runge Kutta 4 and the splitting versions. The weak error of the Euler method at 64 steps seems smaller than the weak error of the Runge Kutta 4, but the calculated values are samples from random variables with positive variance. Note that the Euler-Maruyama method results in a first order weak approximation; i.e. at the weak level there is no difference in the convergence order of the tested methods.

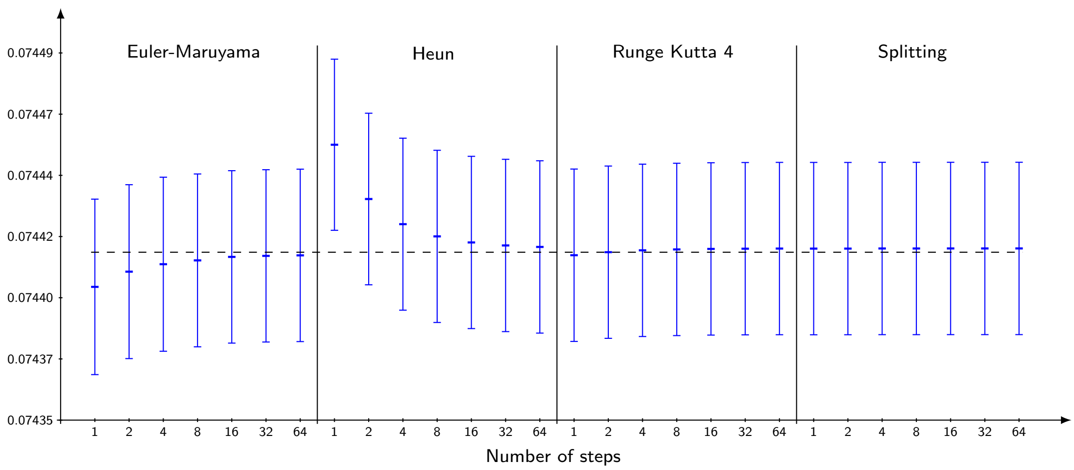

Figure 4.1: Weak approximation results (CIR)

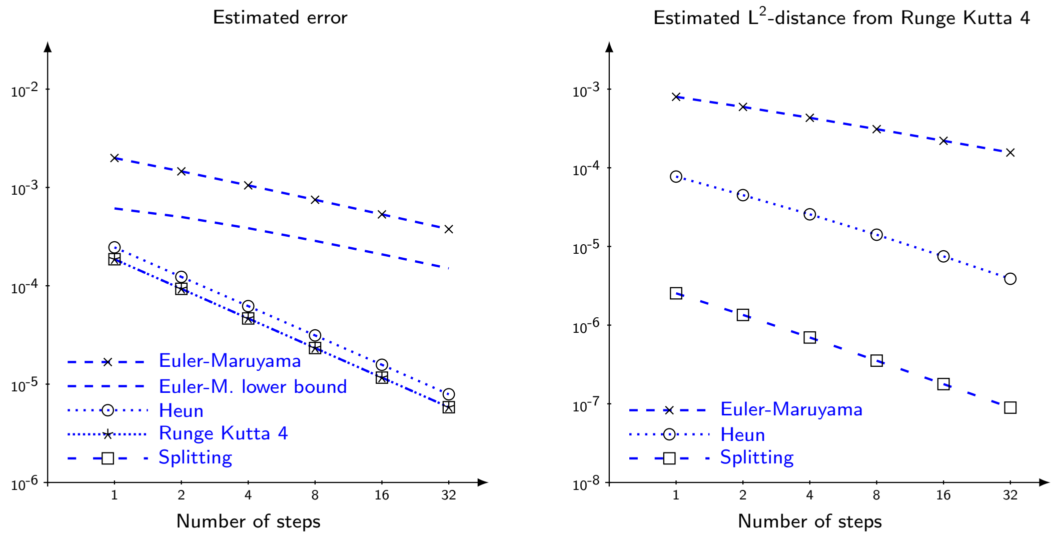

Figure 4.2: Estimated error, comparison of solvers (CIR)

Observation 4.7.2. In Figure 4.2, the accuracy of the path-wise approximations is presented. The graph on the left-hand side presents the estimated upper bound on the error based on (4.38). In the case of the Euler method, we can apply (4.40) and a lower bound on the error is estimated. This estimated lower bound demonstrates, that despite the nice weak approximation properties, the Euler-Maruyama scheme is less accurate in the path-wise approximations when compared to the other schemes. The estimated orders of convergence (i.e. the slopes of the curves) are as expected. In the case of the Euler-Maruyama scheme it's close to $1/2$ and in the case of the ODE-based methods, it is close to 1.

Observation 4.7.3. The graph on the right-hand side of Figure 4.2 compares each method with the Runge Kutta 4 version, estimating the L^2-distance based on (4.39). Note that the L^2 distance of the splitting version from the Runge Kutta 4 based method has an order of magnitude 10^{-6}-10^{-7}, whereas the estimated global error of both methods has order 10^{-5}. So the two versions are equivalent numerically. However the splitting is a bit faster than the Heun scheme and more than twice as faster than Runge Kutta 4, so in this particular case the Splitting version is recommended.

Each ODE solver tested here can be easily extended to handle the term $\int_0^t r_u \mathrm{d}u$. The simulation of this term is required when pricing or hedging bonds and derivatives on bonds. In the case of Splitting, the exact solution to the new ODEs appearing in this extension are known.

4.7.3 Second order approximation implemented

In this section we present some numerical results of some tests run with the second order ODE-based method. The SDE chosen here has no natural financial interpretation. However due to its nice properties, it has proved to be a useful test case.

The SDE is given by

$$\begin{aligned} \mathrm{d}x_t^1 &= \sin(x_t^1) \circ \mathrm{d}B_t^1 \\ \mathrm{d}x_t^2 &= \sin(x_t^1)\mathrm{d}t \end{aligned}$$

The left-hand side of Figure 4.3 presents the weak approximation results in the form of confidence intervals. According to the rule (4.35), for the 1-step second order version we used a linear interpolation of the Brownian motion on a two step sub-scale, i.e. $k = 2$. In the case of the 2-step version, $k = 4$ was chosen, whereas the 4-step version was run with $k = 8$.

Observation 4.7.4. The estimated upper bound on the strong global error is presented on the right-hand side of Figure 4.3. The second order method, implemented as described in section 4.6 has an estimated order close to 2.

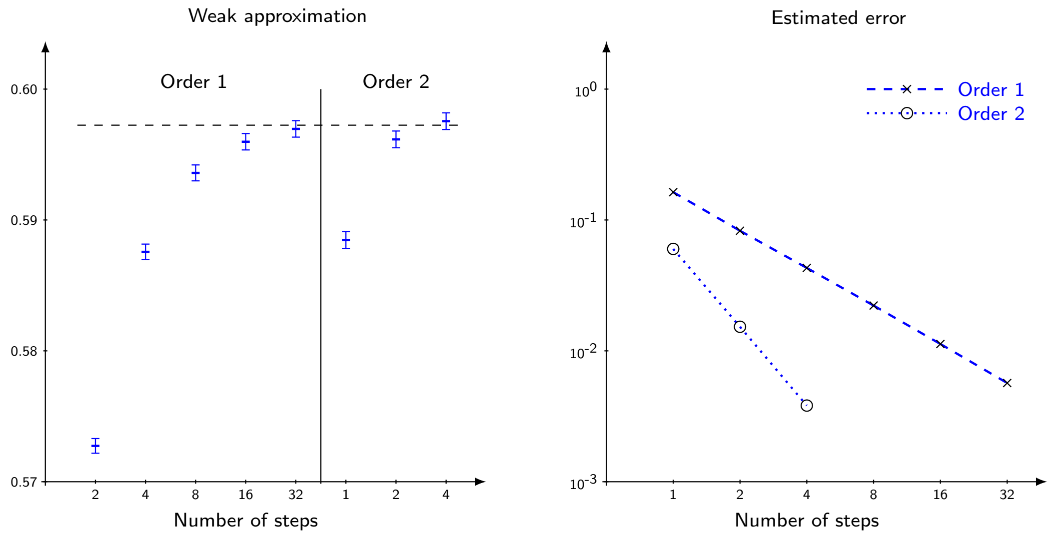

Figure 4.3: Implementing high order approximations

We analyzed the computational expense for this particular SDE as described in Section 4.6.2. In this case $E_2/E_1 \approx 3/2$, so the second order scheme is computationally more efficient than the lowest order at every step. For higher order schemes and in higher dimensions, the lower step high order versions might be relatively less efficient.

CHAPTER 5

NUMERICAL EVALUATION OF THE KUSUOKA-LYONS-VICTOIR METHOD

The first variant of the family of high order weak approximations we refer to as the Kusuoka-Lyons-Victoir (KLV) family appeared in [17] by Kusuoka. The second and third variants were presented in [18] by Kusuoka and [28] by Lyons & Victoir. The latter two papers had mutual influences on each other and can be regarded as interpretations of a wide class of weak approximations based on different approaches. In this section, we give a general description of this wide class of approximations and present a numerical evaluation of the different variants highlighting and demonstrating the crucial differences of the variants, the effectiveness of the KLV family as well as some of the drawbacks.

5.1 DESCRIPTION OF THE KLV FAMILY

Consider the SDE on $\mathbb{R}^N$

$$\xi_t(\xi) = \sum_{i=0}^{k} V_i(\xi_t) \circ \mathrm{d}B_t^i,\ \xi_0(\xi) = \xi$$

The KLV family provides high order approximations of $\mathbb{E}\left[f(\xi_T(\xi))\right]$ when $V_0, \ldots, V_k$ are smooth functions with bounded derivatives up to a certain degree and f is a smooth or Lipschitz $\mathbb{R}^N \to \mathbb{R}$ function.

5.1.1 CUBATURE ON WIENER SPACE

Consider the space $\Omega = C_0^0([0,T], \mathbb{R}^k)$ of $\mathbb{R}^k$-valued continuous functions defined on the closed interval $[0,T]$ starting at zero. Let $\mathcal{F}$ denote the Borel σ-field of $C_0^0([0,T], \mathbb{R}^d)$ and $\mathbb{P}$ the Wiener measure. That is, $(\Omega, \mathcal{F}, \mathbb{P})$ is the Wiener space of k-dimensions. By convention, for all paths $\omega(\cdot) \in \Omega$, we set $\omega^0(t) = t$. We define the coordinate mapping process $B_t^i(\omega) = \omega^i(t)$ for $t \in [0,T]$ and $\omega \in \Omega$. Under $\mathbb{P}$, the process $(B_t^1, \ldots, B_t^k)_{t\in[0,T]}$ is a k-dimensional Brownian motion. By definition, $B_t^0(\omega) = t$ for all $\omega \in \Omega$. We continue using the notation $B_t = (B_t^0, B_t^1, \ldots, B_t^k) \in \mathbb{R}^{k+1}$. Let $W_0, \ldots, W_k$ be the vector fields on $\mathbb{R}^N$

derived from the $\mathbb{R}^N \to \mathbb{R}^N$ functions $V_0, \ldots, V_k$ (in the sense of (3.21)). Let Φ denote the corresponding algebra homomorphism (ref. Definition 3.3.6). Throughout the chapter, we assume that α_X is a positive real such that $2\alpha_X \in \mathbb{Z}$.

The approach of Lyons & Victoir [28] is based on constructing cubature formulas on the infinite dimensional Wiener space. A cubature formula on the Wiener space is a set of pairs $(\omega_i(\cdot), \lambda_i)$ of paths on $\mathbb{R}^{k+1}$ and positive real weights for $i = 1, \ldots, n$, i.e. a discrete measure $\mathbb{Q}$ on Ω. A cubature formula of degree m on a finite dimensional space accurately integrates polynomials up to degree m. Analogously, a cubature formula of degree m on the Wiener space $(\Omega, \mathcal{F}, \mathbb{P})$ and on $[0,1]$ accurately integrates the iterated integrals up to degree m, i.e.

$$\mathbb{E}_{\mathbb{P}}\left[B^R_{0,1}\right] = \mathbb{E}_{\mathbb{Q}}\left[B^R_{0,1}\right] = \sum_{i=1}^n \lambda_i \int_{0<u_1<\cdots<u_l<1} \mathrm{d}\omega_i^{r_1}(u_1)\cdots \mathrm{d}\omega_i^{r_l}(u_l)$$

for each $R = (r_1, \ldots, r_l)$ of degree at most m, or equivalently (recall (4.17) for the path-wise definition $\mathbf{B}^m_{0,1}$)

$$\mathbb{E}_{\mathbb{P}}\left[\mathbf{B}^m_{0,1}\right] = \sum_{i=1}^n \lambda_i S(\omega_i)^m_{0,1} \tag{5.1}$$

where $S(\omega_i)^m_{s,t}$ denotes the signature of the path ω_i truncated at degree m and corresponding to the interval $[s,t]$.

Given that $B^R_{0,t}$ and $t^{deg(R)} B^R_{0,1}$ are equal in law, any cubature on $[0,1]$ can be rescaled to the interval $[0,t]$ using

$$\omega^j_{t,i}(u) = \begin{cases} u & \text{if } j = 0 \\ \sqrt{t}\,\omega^j_i(u/t) & \text{if } j \in \{1, \ldots, k\} \end{cases} \tag{5.2}$$

where the rescaled paths of the cubature support are denoted by $\omega_{t,i}(\cdot)$, $i = 1, \ldots, n$. The measure described by the pairs $(\omega_{t,i}(\cdot), \lambda_i)$ for $i = 1, \ldots, n$ is denoted by $\mathbb{Q}_t$.

Recalling the stochastic Taylor expansion (4.13) for smooth $U \to \mathbb{R}$ functions, we have

$$\begin{aligned}(\mathbb{E}_{\mathbb{P}} - \mathbb{E}_{\mathbb{Q}_t})\left[f(\xi_t(\xi))\right] &\le (\mathbb{E}_{\mathbb{P}} + \mathbb{E}_{\mathbb{Q}_t})\left[\mathcal{R}^{\text{stoch}}_{0,t}(m)\right] \\ &\le C(q,m,\mathbb{Q}_1)t^{m+1/2} \sup_{\substack{0\le u\le t \\ m+1/2\le deg(R)\le m+1}} (\mathbb{E}_{\mathbb{P}} + \mathbb{E}_{\mathbb{Q}_t})\left[\|W_R f(\xi_t(\xi))\|^{2q}\right]^{1/2q}\end{aligned} \tag{5.3}$$

where $C(q,m,\mathbb{Q}_1)$ depends on q, m, $\mathbb{Q}_1$, k and T.

Let $\mathcal{D} = \{0 = t_0 < \cdots < t_l = T\}$ be a partition of $[0,T]$. Let $I = (i_1, \ldots, i_l) \in \{1, \ldots, n\}^l$. Let $\mathbb{Q}_{\mathcal{D}}$ denote the discrete measure on Ω, defined by

$$\mathbb{Q}_{\mathcal{D}}\left[\omega_{\mathcal{D},I}\right] = \prod_{j=1}^l \lambda_{i_j}$$

where $\omega_{\mathcal{D},I}$ is the concatenation of the paths $\omega_{t_{j+1}-t_j, i_j}$ for $j = 0, \ldots, l$.

A simple case, when $f : W \to \mathbb{R}$ is a smooth function with bounded derivatives, is given by the following theorem.

Theorem 5.1.1. *There exists a function $C(k,m)$ not depending on $\mathcal{D}$ or ξ, such that for any smooth function $f : W \to \mathbb{R}$ the following holds:*

$$(\mathbb{E}_{\mathbb{P}} - \mathbb{E}_{\mathbb{Q}_{\mathcal{D}}})\left[f(\xi_t(\xi))\right] \leq C(k,m)\sum_{j=0}^{l-1}(t_{j+1}-t_j)^{(m+1)/2} \sup_{m+1/2\leq deg(R)\leq m+1}\left\|W_R P_{T-t_j} f\right\|_{\infty}, \tag{5.4}$$

where $\{P_t\}_{t\geq 0}$ denotes the semi-group of linear operators defined by

$$(P_t f)(x) = \mathbb{E}\left[f(\xi_t(x))\right]. \tag{5.5}$$

The proof is based on the combination of (5.3) and the following lemma:

Lemma 5.1.1. *Let $\{Q_t | t \geq 0\}$ be a set of linear operators on Lipschitz $W \to \mathbb{R}$ functions. Assume that for each Lipschitz f there exist constants $C(f)$ and $\gamma(f)$ not depending on t such that*

$$\sup_{x\in\mathbb{R}^N}\left\|P_t f(x) - Q_t f(x)\right\| \leq C(f) t^{\gamma(f)}. \tag{5.6}$$

Let $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$ be a partition of $[0,T]$.

Then

$$\sup_{x\in\mathbb{R}^N}\left\|(P_T f)(x) - \left(Q_{(t_1-t_0)}\cdots Q_{(t_n-t_{n-1})}f\right)(x)\right\| \leq \sum_{i=1}^{n} C(P_{(t_n-t_i)}f)(t_i-t_{i-1})^{\gamma(P_{(t_n-t_i)}f)}.$$

The lemma is implied by the following representation:

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^n}\left\|P_T f(x) - Q_{(t_1-t_0)}\cdots Q_{(t_n-t_{n-1})}f(x)\right\| \\
&\quad = \sup_{x\in\mathbb{R}^N}\left\|\sum_{i=1}^{n} Q_{(t_1-t_0)}\cdots Q_{(t_{i-1}-t_{i-2})}\left(P_{(t_i-t_{i-1})} - Q_{(t_i-t_{i-1})}\right)P_{(T-t_i)}f\right\| \\
&\quad \leq \sum_{i=1}^{n}\sup_{x\in\mathbb{R}^N}\left\|\left(P_{(t_i-t_{i-1})} - Q_{(t_i-t_{i-1})}\right)P_{(T-t_i)}f(x)\right\|
\end{aligned} \tag{5.7}$$

The case, in which $f : \mathbb{R}^N \to \mathbb{R}$ is only Lipschitz, the results of Kusuoka, & Stroock [16] and Kusuoka [19] and some ideas of Kusuoka [17] (sketched in the next sections in Lemma 5.1.6 and Theorem 5.1.2) imply that the error formula (equivalent to the formula of Corollary 5.1.1) can be expressed in terms of $\|\nabla f\|_\infty$ under the UFG condition (ref. Definition 5.1.4).

In [28], cubature formulas on the Wiener space are constructed (i.e. (5.1) is solved) in two steps. First a set of Lie elements $L = \{L_1,\ldots,L_n\} \subset \mathcal{L}^{(\Pi_x,m)}(\mathbb{R}^{k+1})$ and a set of positive weights $\Lambda = \{\lambda_1,\ldots,\lambda_n\}$ are found satisfying

$$\mathbb{E}_{\mathbb{P}}\left[\mathbf{B}^m_{0,1}\right] = \sum_{i=1}^{n}\lambda_i \exp^{(\Pi_x,m)}(L_i) \tag{5.8}$$

Any set of pairs $\{(L_i, \lambda_i) \mid i = 1, \ldots, n\}$ satisfying (5.8) for m will be referred to as a *degree m cubature formula*. Note that in the literature ([22], [23], [28] etc.), the degree m cubature formula is referred to as *degree* $2m$. This difference is due to our slightly different definition of degrees of multi-indices. We use our version to be consistent with Chapters 2 and 3.

The existence of L and Λ with size at most $n = \operatorname{card}\left(\mathcal{A}_m^{\Pi_X}\right)$ is guaranteed by Tchakaloff's Theorem (ref. [28]). Note that L and Λ also determine a discrete measure on $\mathcal{L}^{(\Pi_x,m)}(\mathbb{R}^{k+1})$.

In the second step, piece-wise linear $[0,1] \to \mathbb{R}^{k+1}$ paths $\omega_1, \ldots, \omega_n$ of finite and minimal length are constructed satisfying $S(\omega_i)_{0,1}^m = \exp^{(\Pi_x,m)}(L_i)$. The existence of such paths is guaranteed by Chen's theorem 3.3.1.

Degree 3/2 and 5/2 solutions for (5.8) are given by the following lemmas (proved in [28]).

Lemma 5.1.2. *Let (x_i, λ_i) for $i = 1, \ldots, n$ be a cubature formula of degree* 3 *with respect to the k-dimensional Gaussian measure. Then (L_i, λ_i) for $i = 1, \ldots, n$ where*

$$L_i = \varepsilon_0 + \sum_{j=1}^{k} x_j^i \varepsilon_j$$

solve (5.8) *when $m = 3/2$.*

Lemma 5.1.3. *Let (x_l, λ_l) for $l = 1, \ldots, n$ be a cubature formula of degree* 5 *with respect to the k-dimensional Gaussian measure. Then $(L_l, \lambda_{|l|}/2)$ for $l = \pm 1, \ldots, \pm n$ and any real constant y where*

$$\begin{aligned} L_l \quad = \quad & \varepsilon_0 + \sum_{i=1}^{k} \frac{1}{12}(x_{|l|}^i)^2[[\varepsilon_0, \varepsilon_i], \varepsilon_i] + \sum_{i=1}^{k} x_{|l|}^i \varepsilon_i + sign(l) \sum_{1 \le i < j \le k} \frac{1}{2} x_{|l|}^i x_{|l|}^j [\varepsilon_i, \varepsilon_j] \\ & + \sum_{1 \le i < j \le k} \frac{1}{6} \left(y x_{|l|}^i (x_{|l|}^j)^2 [[\varepsilon_i, \varepsilon_j], \varepsilon_j] + (1-y) x_l^j (x_l^i)^2 [[\varepsilon_j, \varepsilon_i], \varepsilon_i] \right) \end{aligned}$$

solve (5.8) *when $m = 5/2$.*

Lemma 5.1.4. *Let (x_i, λ_i) for $i = 1, \ldots, n$ be a cubature formula of degree* 7 *with respect to the* 1*-dimensional Gaussian measure. Then $(L_i, \lambda_{|i|}/2)$ for $i = \pm 1, \ldots, \pm n$ where*

$$L_i = \varepsilon_0 + x_i \varepsilon_1 + sign(i) \frac{x_i}{\sqrt{12}} [\varepsilon_0, \varepsilon_1] + \frac{1}{12}[[\varepsilon_0, \varepsilon_1], \varepsilon_1] + \frac{1}{360}[[[[\varepsilon_0, \varepsilon_1]\varepsilon_1]\varepsilon_1]\varepsilon_1]$$

(5.8) *for $m = 7/2$ on the one-dimensional Wiener space.*

Degree 7/2 cubature formulas on the 2-dimensional Wiener space are introduced by Litterer [22]. Degree 3,5 and 7 cubature formulas with respect to the Gaussian measure are collected by Stroud [36].

5.1.2 Kusuoka's approach

Kusuoka's interpretation is related to the approximations based on autonomous random ODEs locally derived from the log-signature of the driving noise. Recall from Chapter 4 the strong global approximation $X_T(\xi, \mathcal{D})$ of $\xi_t(\xi)$ on $\mathbb{R}^N$ corresponding to the partition $\mathcal{D}$ of $[0, T]$ determined by the local approximation $\widehat{D}$, where

$$\widehat{D}\big[(s,t),x\big] := \operatorname{Exp}\left(\Phi\left\{L_{s,t}^{(\Pi_X,\alpha_X)}(B)\right\}\right)(x). \tag{5.9}$$

Let $f : \mathbb{R}^N \to \mathbb{R}^N$ be a Lipschitz function. For any positive integer q, Corollary 4.3.2 implies

$$\begin{aligned}&\|\mathbb{E}\left[f(\xi_T(\xi)) - f(X_T(\xi,\mathcal{D}))\right]\| \\ &\quad \le \|\nabla f\|_\infty \mathbb{E}\left[\|\xi_T(\xi) - X_T(\xi,\mathcal{D})\|^{2q}\right]^{1/2q} \le C\|\nabla f\|_\infty |\mathcal{D}|^{\lfloor \alpha_X \rfloor},\end{aligned} \tag{5.10}$$

where C is a constant not depending on $\mathcal{D}$ or ξ. However, there are at least two difficulties in directly exploiting the above inequality in practice. Firstly, to make $\mathbb{E}\left[f(X_T(\xi,\mathcal{D}))\right]$ into a high order approximation of $\mathbb{E}\left[f(\xi_T(\xi))\right]$, one needs to simulate high degree Brownian iterated integrals. This difficulty has been addressed in Section 4.6. The other difficulty lies in the actual computation of the expectation. A standard approach is to use Monte-Carlo simulation which introduces another source of error due to the variance of sampling.

One of the key ingredients in Kusuoka's approach is the fact that the truncated log-signature $L_{s,t}^{(\Pi_X,\alpha_X)}(B)$ in (5.9) can be replaced with certain $\mathcal{L}^{(\Pi_X,\alpha_X)}(\mathbb{R}^k)$-valued random variables preserving the weak approximation property. To make this precise, we introduce two definitions.

Definition 5.1.1. *Let* $\mathbf{a} \in T^{(\Pi_X,\alpha_X)}(\mathbb{R}^k)$*. We introduce the* rescaling operator

$$\Psi : \mathbb{R}^+ \times T^{(\Pi_X,\alpha_X)}(\mathbb{R}^k) \to T^{(\Pi_X,\alpha_X)}(\mathbb{R}^k)$$

defined by

$$\Psi(\lambda,\mathbf{a}) = \Psi_\lambda(\mathbf{a}) = \sum_{deg(R)\le\alpha_X} \lambda^{deg(R)}\mathbf{a}^R$$

for $\lambda \in \mathbb{R}+$*. Note that* $\mathcal{L}^{(\Pi_X,\alpha_X)}(\mathbb{R}^k)$ *is invariant under* Ψ_λ *for any* $\lambda \in \mathbb{R}$*.*

Definition 5.1.2. *Two* $\mathcal{L}^{(\Pi_X,\alpha_X)}(\mathbb{R}^k)$*-valued random variables* η_1 *and* η_2 *under the measures* $\mathbb{P}_1$ *and* $\mathbb{P}_2$ *respectively are called* m-moment equivalent *for* $m \le \alpha_X$ *if* η_1 *and* η_2 *have finite moments up to* $2m$*,* $\eta_1^{(0)} = \eta_2^{(0)} = 1$ *and*

$$\mathbb{E}_{\mathbb{P}_1}\left[\exp^{(\Pi_X,\alpha_X)}(\eta_1)\right] = \mathbb{E}_{\mathbb{P}_2}\left[\exp^{(\Pi_X,\alpha_X)}(\eta_2)\right]. \tag{5.11}$$

Lemma 5.1.5. *Let λ be a positive real number in $(0,1)$. Let η_1, η_2 be two α_X-moment equivalent $\mathcal{L}^{(\Pi_X,\alpha_X)}(\mathbb{R}^k)$-valued random variables and let f be a smooth $\mathbb{R}^N \to \mathbb{R}$ function with bounded derivatives up to order $2m$.*

Then

$$\sup_{x\in\mathbb{R}^k} \left\| \mathbb{E}\left[f\left(\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_1)\right\})(x)\right)\right] - \mathbb{E}\left[f\left(\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_2)\right\})(x)\right)\right] \right\| \leq C \sum_{\alpha_X+1/2\leq deg(R)\leq 2\alpha_X} \lambda^{deg(R)} \|W_R f\|_\infty$$

where C depends on the moments of η_1 and η_2 up to $2\alpha_X$.

The proof is based on comparing the remainder terms of the stochastic Taylor expansion, i.e. on techniques analogous to the proof of Theorem 4.3.2.

Remark 5.1.1. Let η be a $\mathcal{L}^{(\Pi_X,\alpha_X)}(\mathbb{R}^k)$-valued random variable α_X-moment equivalent to

$$\log^{(\Pi_X,\alpha_X)}(\mathbf{B}_{0,1}).$$

Let $\{Q_t | t \geq 0\}$ be the set of linear operators defined by

$$(Q_s f)(x) = \mathbb{E}\left[f\big(\mathrm{Exp}(\Phi\{\Psi_s(\eta)\})(x)\big)\right]. \tag{5.12}$$

Then inequality (5.10), Lemmas 5.1.5 and 5.1.1 determine a class of weak approximations based on the set $\{Q_t | t \geq 0\}$ for smooth terminal conditions with bounded derivatives up to a certain order. However, the case when the terminal condition is only Lipschitz requires some further considerations.

Definition 5.1.3. *We introduce the vector fields $W_{[R]}$, $R \in \mathcal{A}^{\Pi_X}$ recursively using*

$$\begin{aligned} W_{[(i)]} &= W_i,\ i \in \{0,\dots,k\} \\ W_{[R]} &= \left[W_{r_1}, W_{[-R]}\right] \end{aligned}$$

For $R \in \mathcal{A}^{\Pi_X}$ and the set of vector fields $\mathcal{W} = \{W_0,\dots,W_k\}$, we introduce the norm $\|f\|_{\mathcal{W},R}$ on smooth functions with bounded derivatives as follows:

$$\|f\|_{\mathcal{W},R} = \sum_{i=1}^{\|R\|} \sum_{\substack{R_1,\dots,R_i \in \mathcal{A}^{\Pi_X}\setminus\{\epsilon,(0)\} \\ R_1*\cdots*R_i = R}} \left\| W_{[R_1]} \circ \cdots W_{[R_i]} \circ f \right\|_\infty$$

We present a stronger version of Lemma 5.1.5.

Lemma 5.1.6. *Let λ be a positive real number in $(0,1)$. Let η_1, η_2 be two α_X-moment equivalent $\mathcal{L}^{(\Pi_X,\alpha_X)}(\mathbb{R}^k)$-valued random variables and let f be a smooth $\mathbb{R}^N \to \mathbb{R}$ function with bounded derivatives up to order $2m$.*

Then

$$\sup_{x\in\mathbb{R}^k}\left\|\mathbb{E}\left[f\left(\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_1)\right\})(x)\right)\right]-\mathbb{E}\left[f\left(\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_2)\right\})(x)\right)\right]\right\|$$
$$\leq C\sum_{\alpha_X+1/2\leq deg(R)\leq 2\alpha_X}\lambda^{deg(R)}\left(\|f\|_{\mathcal{W},R}+\|\nabla f\|_\infty\right)$$

where C depends on the moments of η_1 and η_2 up to $2\alpha_X$.

Proof. We adapt the results of Kusuoka ([17] and also in [18]).

Step 1

The fact that for $i=1,2$, $L_{\lambda,i}:=\log(\exp(\Psi_\lambda(\eta_i))\otimes\exp(-\lambda\varepsilon_0))$ is an element of $\mathcal{L}(\mathbb{R}^k)$ is implied by Proposition 9 of [17]. Moreover, Proposition 18 of [17] implies the existence of a constant C_1 depending on α_X, q and the moments of η_i up to $2\alpha_X$ such that

$$\sup_{x\in\mathbb{R}^N}\mathbb{E}\left[\left\|\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_i)\right\})\mathrm{Exp}(\Phi\left\{-\lambda\varepsilon_0\right\})(x)-\mathrm{Exp}(\Phi\left\{\pi_{\alpha_X}L_{\lambda,i}\right\})(x)\right\|^{2q}\right]^{1/2q}$$
$$\leq C_1\sum_{\alpha_X\leq degR\leq 2\alpha_X}\lambda^{deg(R)}\left\|W_R\circ Id\right\|_\infty.$$

In the proof of Proposition 18 of [17], terms of degree up to α_X^2 are included. As opposed to that proof, the above inequality is derived using a partial Taylor expansion (i.e. in each step, we expanded only the terms of degree at most α_X) as we did in the proof of Theorem 3.4.1.

The inequality implies

$$\sup_{x\in\mathbb{R}^N}\left\|\mathbb{E}\left[f\left(\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_i)\right\})\mathrm{Exp}(\Phi\left\{-\lambda\varepsilon_0\right\})(x)\right)-f\left(\mathrm{Exp}(\Phi\left\{\pi_{\alpha_X}L_{\lambda,i}\right\})(x)\right)\right]\right\|$$
$$\leq C_1\|\nabla f\|_\infty\sum_{\alpha_X\leq degR\leq 2\alpha_X}\lambda^{deg(R)}\left\|W_R\circ Id\right\|_\infty.$$

Step 2

The fact that $\pi_{(0)}L_{\lambda,i}=0$ implies the inequality

$$\sup_{x\in\mathbb{R}^N}\left\|\mathbb{E}\left[f\left(\mathrm{Exp}(\Phi\left\{\pi_{\alpha_X}L_{\lambda,i}\right\})(x)\right)-\Phi\left\{\pi_{\alpha_X}\exp(L_{\lambda,i})\right\}f(x)\right]\right\|$$
$$\leq C_2\sum_{\alpha_X+1/2\leq deg(R)\leq 2\alpha_X}\lambda^{deg(R)}\|f\|_{\mathcal{W},R}$$

where C_2 depends only on α_X, q and the moments of η_i up to $2\alpha_X$. The proof of the inequality is based on Proposition 11 of [17] simplified using a partial Taylor expansion.

Step 3

Noting that

$$\sup_{x\in\mathbb{R}^k}\left\|\mathbb{E}\left[f\Big(\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_1)\right\})(x)\Big)\right]-\mathbb{E}\left[f\Big(\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_2)\right\})(x)\Big)\right]\right\|$$
$$=\sup_{x\in\mathbb{R}^k}\Big\|\mathbb{E}\left[f\Big(\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_1)\right\})\Big(\mathrm{Exp}(\Phi\left\{-\lambda\varepsilon_0\right\})(x)\Big)\Big)\right]$$
$$-\mathbb{E}\left[f\Big(\mathrm{Exp}(\Phi\left\{\Psi_\lambda(\eta_2)\right\})\Big(\mathrm{Exp}(\Phi\left\{-\lambda\varepsilon_0\right\})(x)\Big)\Big)\right]\Big\|,$$

Steps 1 and 2 and the α_X-moment similarity of η_1 and η_2 imply the assertion.

□

We will link Lemma 5.1.1 and 5.1.6 using the following results.

Definition 5.1.4 (UFG condition). *Let $\mathcal{W} = \{W_0, \ldots, W_k\}$ be a set of vector fields on $\mathbb{R}^N$. We say that $\mathcal{W}$ satisfies the* UFG condition, *if there exists an integer l such that for each $R, Q \in \mathcal{A}^{\Pi_X} \setminus \{\epsilon, (0)\}$ where $\|Q\| \le l$, there exist smooth $\mathbb{R}^N \to \mathbb{R}$ functions $\phi_{R,Q}$ with bounded derivatives of any order satisfying*

$$W_{[R]} = \sum_{\|Q\| \le l} \phi_{R,Q} W_{[Q]}.$$

The following Theorem is a special case of the main result of Kusuoka [19].

Theorem 5.1.2. *Let n be a positive integer, let $R_1, \ldots, R_n$ be multi-indices in $\mathcal{A}^{\Pi_X} \setminus \{\epsilon, (0)\}$ and let $\mathcal{W}$ satisfy the UFG condition. Then there exists a constant C not depending on $t > 0$ or n such that*

$$\left\| W_{[R_1]} \circ \cdots \circ W_{[R_n]} \circ P_t f \right\|_\infty \le \frac{C t^{1/2}}{t^{deg(R_1 * \cdots * R_n)}} \|\nabla f\|_\infty .$$

for any smooth function f with bounded derivatives of all order.

Corollary 5.1.1. *Let η be a $\mathcal{L}^{(\Pi_X, \alpha_X)}(\mathbb{R}^k)$-valued random variable and α_X-moment equivalent to $\log^{(\Pi_X, \alpha_X)}(\mathbf{B}_{0,1})$, $\{Q_t | t \ge 0\}$ the set of linear operators defined by*

$$(Q_s f)(x) = \mathbb{E}\left[f\big(\mathrm{Exp}(\Phi\{\Psi_s(\eta)\})(x)\big)\right], \tag{5.13}$$

and $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$ a partition of $[0, T]$. Then there exists a constant C not depending on $\mathcal{D}$ or f such that

$$\begin{aligned}
&\sup_{x \in \mathbb{R}^N} \left\| (P_T f)(x) - \left(Q_{(t_1 - t_0)} \cdots Q_{(t_n - t_{n-1})} f \right)(x) \right\| \\
&\qquad \le C \|\nabla f\|_\infty \left((t_n - t_{n-1})^{1/2} + \sum_{i=1}^{n-1} \frac{(t_i - t_{i-1})^{\alpha_X + 1/2}}{(T - t_i)^{\alpha_X}} \right).
\end{aligned}$$

Proof. For $i = 1, \ldots, n-1$, the constants $C(P_{(t_n - t_i)} f)$ and $\gamma(P_{(t_n - t_i)} f)$ of Lemma 5.1.1 are determined by Theorem 5.1.2 applied to Lemma 5.1.6 and implied by the fact that even if f is not smooth but Lipschitz, $P_t f$ is smooth.

The i=n case is given by

$$\begin{aligned}
&\sup_{x \in \mathbb{R}^N} \left\| P_{(t_n - t_{n-1})} f(x) - Q_{(t_n - t_{n-1})} f(x) \right\| \\
&\quad \le \sup_{x \in \mathbb{R}^N} \left\| P_{(t_n - t_{n-1})} f(x) - f(x) \right\| + \sup_{x \in \mathbb{R}^N} \left\| f(x) - Q_{(t_n - t_{n-1})} f(x) \right\| \\
&\quad \le C \|\nabla f\|_\infty (t_n - t_{n-1})^{1/2}
\end{aligned}$$

□

5.2 Practical considerations

In this section, we review some practical considerations arising when implementing any version of the KLV family.

5.2.1 Solving the ODEs

When implementing the version of KLV described in Section 5.1.2, the exact solution to the random ODEs which arise is rarely known. One can apply ODE numerical schemes $\widehat{\widehat{D}}$ assigning a numerical solution of $\operatorname{Exp}(\Phi\{\Psi_t(\eta)\})(x)$ to (t, x) and the realization of η. If we assume that in addition to the condition (5.6), $\widehat{\widehat{D}}$ satisfies the inequality

$$\sup_{x\in\mathbb{R}^N} \mathbb{E}\left[\left\|\widehat{\widehat{D}}(t,x) - \operatorname{Exp}(\Phi\{\Psi_\lambda(\eta)\})(x)\right\|^{2q}\right]^{1/2q} \leq C(q)t^{\gamma(f)}, \tag{5.14}$$

then the following condition is implied

$$\sup_{x\in\mathbb{R}^N} \left\|\mathbb{E}\left[f\left(\xi_t(x)\right)\right] - \mathbb{E}\left[f\left(\widehat{\widehat{D}}(t,x)\right)\right]\right\| \leq (C(f) + C(q)\|\nabla f\|_\infty)t^{\gamma(f)} \tag{5.15}$$

and arguments analogous to the proof of Lemma 5.1.1 imply global convergence.

A special choice for $\widehat{\widehat{D}}(t, x)$ is given by factorizing η in the form :

$$\eta = \log^{(\Pi_X,\alpha_X)}\left(\exp\left(\sum_{i=0}^k a_{i,1}\varepsilon_i\right) \otimes \cdots \otimes \exp\left(\sum_{i=0}^k a_{i,l}\varepsilon_i\right)\right) \tag{5.16}$$

i.e. when η is written as the truncated log-signature of a piece-wise linear (random) $[0,1] \to \mathbb{R}^k$ path $\omega(\cdot)$ and $\widehat{\widehat{D}}(t, x)$ is defined as $\widehat{X}_1(t, x)$ where $\widehat{X}_\cdot(t, x)$ denotes the solution to the non-autonomous ODE

$$\mathrm{d}\widehat{X}_u(t,x) = \sum_{i=0}^k V_i(\widehat{X}_u(t,x))\mathrm{d}\omega_t(u)$$

and $\omega_t(u)$ denotes the rescaled path as defined in (5.2). Note that the factorization problem (5.16) is equivalent to the second step in the construction of cubature formulas on the Wiener space as described in Section 5.1.1. The fact that this particular choice of $\widehat{\widehat{D}}$ satisfies the condition (5.15) is implied by repeated application of Proposition 13 in [17]. This fact also highlights the equivalence between Kusuoka's version and the Cubature on Wiener space approach. In [28], the high order Runge-Kutta method is suggested for approximating the solution to the above non-autonomous ODE. Ninomiya & Ninomiya [32] investigated in detail the use of particular Runge-Kutta schemes and a condition similar to (5.14).

5.2.2 Uneven partitions

It's obvious from Corollary 5.1.1 that when f is only Lipschitz, the local error of the last step is of order $1/2$. Hence for a better global order of convergence one should consider uneven partitions. Kusuoka [17] constructed the following example.

Example 5.2.1. Let $\mathcal{D} = \{0 = t_0 < \cdots < t_n = T\}$ be defined by

$$t_j = T\left(1 - \left(1 - \frac{j}{n}\right)^{\gamma}\right),$$

where γ is a positive real number. Then

(i) for $0 < \gamma < 2\alpha_X - 1$

$$\sup_{x\in\mathbb{R}^N} \left\| (P_T f)(x) - \left(Q_{(t_1-t_0)} \cdots Q_{(t_n-t_{n-1})} f\right)(x) \right\| \leq K \|\nabla f\|_\infty n^{-\gamma/2}$$

(ii) for $\gamma = 2\alpha_X - 1$

$$\sup_{x\in\mathbb{R}^N} \left\| (P_T f)(x) - \left(Q_{(t_1-t_0)} \cdots Q_{(t_n-t_{n-1})} f\right)(x) \right\| \leq K \|\nabla f\|_\infty n^{-(\alpha_X-1/2)} \log(n)$$

(iii) for $\gamma > 2\alpha_X - 1$

$$\sup_{x\in\mathbb{R}^N} \left\| (P_T f)(x) - \left(Q_{(t_1-t_0)} \cdots Q_{(t_n-t_{n-1})} f\right)(x) \right\| \leq K \|\nabla f\|_\infty n^{-(\alpha_X-1/2)}$$

5.2.3 The case when the support of the global cubature measure explodes

In our interpretation, the weak approximations of the KLV family involve solving random ODEs derived locally as in Chapter 4. However, in the weak case, the Brownian log-signature can be replaced with α_X-moment equivalent random variables; i.e. the schemes leading to strong approximations are taken under alternative measures referred to as *global cubature measures*.

In the case when the locally chosen random Lie elements are finitely supported, the corresponding global cubature measure is finitely supported, i.e. the computation of the corresponding expectation is practically feasible assuming sufficient computational capacity. However, as the partition $\mathcal{D}$ splits $[0, T]$ into more subintervals, the support of the global cubature measure growths exponentially resulting in a tractability problem. This problem has been addressed by Litterer & Lyons [23] and Litterer [22], where an algorithm to reduce the measure's support is described.

Note that many of the standard weak approximation schemes (ref. [15]) in which the Brownian iterated integrals are replaced with simpler random variables satisfying certain moment conditions present particular cases of the moment equivalent random variables, although Lie algebra valued random variables are not mentioned explicitly. I.e. many of the standard schemes fall into the KLV perspective (e.g. the weak Euler scheme, the weak Milstein scheme etc.). The moment equivalent random variables of the standard schemes however have infinite support. The standard approach is to use Monte-Carlo simulations or other partial sampling methods resulting in an extra source of approximation error due to the variance. More precisely, let $X(n)$ denote the random variable corresponding to a partial sampling method based on n particles, i.e. when a sample of size n is taken from the global cubature measure. Then the L^2 error is given by

$$\mathbb{E}\left[(\mathbb{E}[f(\xi_T)] - X(n))^2\right] = (\mathbb{E}[f(\xi_T)] - \mathbb{E}[X(n)])^2 + \mathrm{Var}[X(n)], \tag{5.17}$$

where the first term is referred to as squared *discretization error* (same as in Corollary 5.1.1 if the sampling is unbiased) and the second term is referred to as squared *sampling error* due to the variance. In the case of the pure Monte-Carlo sampling, the sampling error is of order $1/n$.

Ninomiya [31] and Kusuoka & Ninomiya [20] restrict their analysis to the case of finitely supported moment equivalent random variables and in the case of huge support they considered alternative partial sampling methods, namely the tree-based branching algorithm or TBBA (see Appendix A for a brief description) of Crisan and Lyons [7]. The TBBA is used to construct a random and finitely supported probability measure with the minimal entropy relative to the distribution of the moment equivalent random variable. The TBBA results in lower variance compared to the standard Monet-Carlo based sampling.

5.3 Numerical results

In this section numerical results are presented demonstrating some properties of the KLV family.

We implemented the KLV version based on autonomous ODEs; i.e. we worked at the Lie algebra and the vector field level. The code was implemented in C++. For the Lie algebra-level computations we used the libraries of the *CoRoPa*[1] project developed by D. Chafaï, T.J. Lyons et. al. For random number generation, we used the Boost[2] implementation of the Mersenne Twister generator, in particular the *mt19937* generator. We used our own code for transforming Lie algebra elements to vector fields, for manipulating vector fields and for solving ODEs. The high order Runge-Kutta schemes (of order $2, 3, \ldots, 7$) were taken from [2]. We used our own implementation of the tree-based branching algorithm.

[1] Computational Rough Paths, website: http://coropa.sourceforge.net/

[2] Website: http://www.boost.org/

When implementing the different modules, our aim was to create a general and flexible environment in which SDEs can be set up easily. In this way, we could test many SDEs conveniently, however our implementation was not optimized for the individual SDEs and the run time of the different tests provided us only partial information on the algorithm's efficiency. In each case we recorded the number of ODEs solved during the run in order to get some further information on the computational cost. Note that solving the ODEs derived from cubature formulas of different degree corresponds to different level of computational complexity.

When testing the version in which we sampled from the global or local cubature measure, we repeated the sampling while starting the random number generator at different seeds and constructed confidence intervals on the results. These confidence intervals contribute to the description of the accuracy as well as to the computational needs.

We tested the pure cubature methods (i.e. evaluating the exact global cubature measure) on a linear and a non-linear problem in 1 dimension, on the Heston volatility model in 2 dimensions and finally on a basket option of three shares modeled with joint Heston stochastic volatility. In each case we tested different (both smooth and non-smooth but Lipschitz) terminal conditions. Furthermore we tested the efficiency of the TBBA and Monte-Carlo sampling, the Romberg extrapolation, the effect of using different numerical ODE solvers and the effect of using an uneven partition on some of these pricing problems. In 1 and 2 dimensions we tested cubature formulas of degree 3/2, 5/2 and 7/2. The degree 7/2 cubature formulas are due to Litterer [22]. In 4 dimensions, we focused on the degree 5/2 case.

The results of some numerical tests of other versions of the KLV family run on SDEs set in 1 and 2 dimensions has been published by Kusuoka & Ninomiya [20], Ninomiya [31], Ninomiya & Ninomiya [32], Ninomiya & Victoir [33] and Litterer [22]. Each of these but one was restricted to degree 3/2 and degree 5/2 cubature formulas. Litterer tested the degree 7/2 formulas, however his analysis focused on the SABR model.

The purpose of this section is to demonstrate the efficiency of the KLV family through several examples considering and comparing more KLV versions than the ones tested in the above sources. We also spot some cases with certain irregularities.

5.3.1 LINEAR SDE IN ONE DIMENSION

The first test case is based on the SDE

$$\begin{aligned} \mathrm{d}S_t &= aS_t\mathrm{d}t + bS_t\mathrm{d}B_t \\ \mathrm{d}A_t &= S_t\mathrm{d}t \end{aligned} \tag{5.18}$$

run with $a = 0.08$, $b = 0.08$, $S_0 = 100$ and $A_0 = 0$ on $[0, T]$ for $T = 1.5$. The following terminal conditions were implemented:

$$f_1(x,y) = x,\ f_2(x,y) = y,\ f_3(x,y) = \max(x - 75, 0),$$
$$f_4(x,y) = \max(x - 125, 0),\ f_5(x,y) = \max(x - y/T, 0)$$

where f_3 corresponds to an ITM European call option, f_4 to an OTM European call option and f_5 to an Asian option.

The degree 3/2, 5/2 and 7/2 cubature formulas we implemented had 2, 3 and 4 elements in their support respectively. Due to the small support, we ran the degree 3/2, 5/2 and 7/2 cubature formulas up to 24, 16 and 8 steps respectively. We ran the algorithm using Taylor expansion based numerical solvers as well as different Runge-Kutta schemes of order $2, 3, \ldots, 7$.

Definition 5.3.1. *Let us fix an SDE and a terminal condition and let e_1 and e_2 denote the weak approximation error corresponding to two ODE solvers. We say that the algorithms corresponding to the two ODE solvers are* essentially equivalent *on the given SDE and terminal condition, if* $|e_1 - e_2| \leq \max(|e_1|, |e_2|)/100$*; i.e. if the relative difference between the errors is at most* $1/100$*. We will say that the algorithm with error e_1 is* essentially more accurate *than the one with error e_2 if $|e_1| < |e_2|$ and the algorithms are not essentially equivalent.*

For each cubature formula we plotted the error corresponding to the lowest order ODE solver among the *essentially most accurate* ones.

In the case of non-smooth but Lipschitz terminal conditions, we considered the use of the uneven partitions of Example 5.2.1. We ran the degree 5/2 and 7/2 cubature formulas on uneven partitions determined by the parameter $\gamma = 1, 1.5, 2, 2.5$ and the most accurate results were plotted.

Observation 5.3.1. When estimating $\mathbb{E}[S_T]$ (figure 5.1), the essentially most accurate algorithms correspond to the Runge-Kutta order 4, Runge-Kutta order 6 and Runge-Kutta order 7 ODE solvers with the degree 3/2, degree 5/2 and degree 7/2 cubature formulas respectively. The convergence orders corresponding to the different degree cubature formulas are no worse than the theoretical order (smooth terminal condition case) and in particular for the degree 7/2 case the convergence order is even better. The degree 7/2 formula results in a very accurate approximation (8-12 digit accuracy), and at 2 steps is already more accurate than the other two cubature formulas at any of the observed steps.

Observation 5.3.2. In the case of the approximation of $\mathbb{E}[A_T]$ (figure 5.2), the essentially most accurate algorithms correspond to the Runge-Kutta order 4, Runge-Kutta order 5 and Runge-Kutta order 5 solvers with the degree 3/2, degree 5/2 and degree 7/2 cubature formulas respectively. The convergence orders corresponding to the different degree cubature formulas match the theoretical order (smooth terminal condition case).

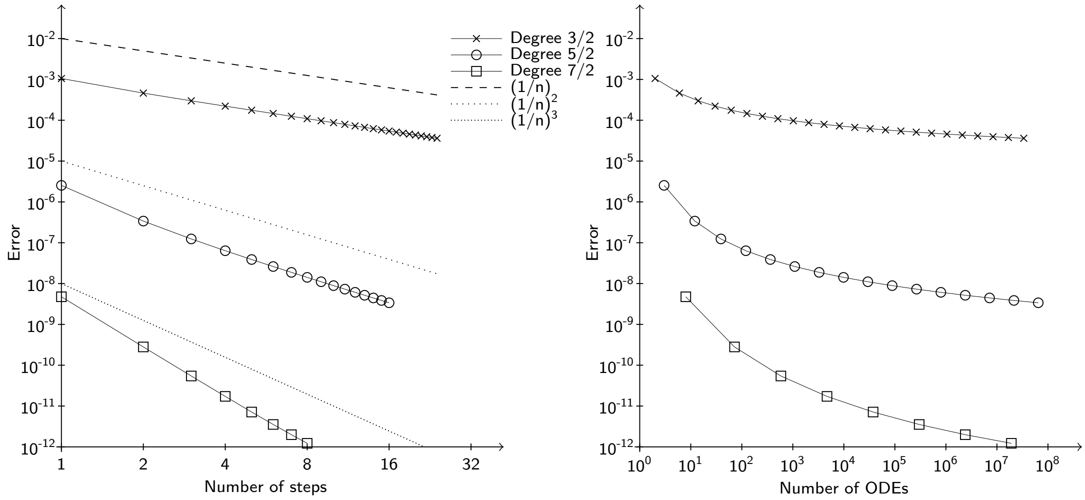

Figure 5.1: Error of approximating $\mathbb{E}\left[S_T\right]$

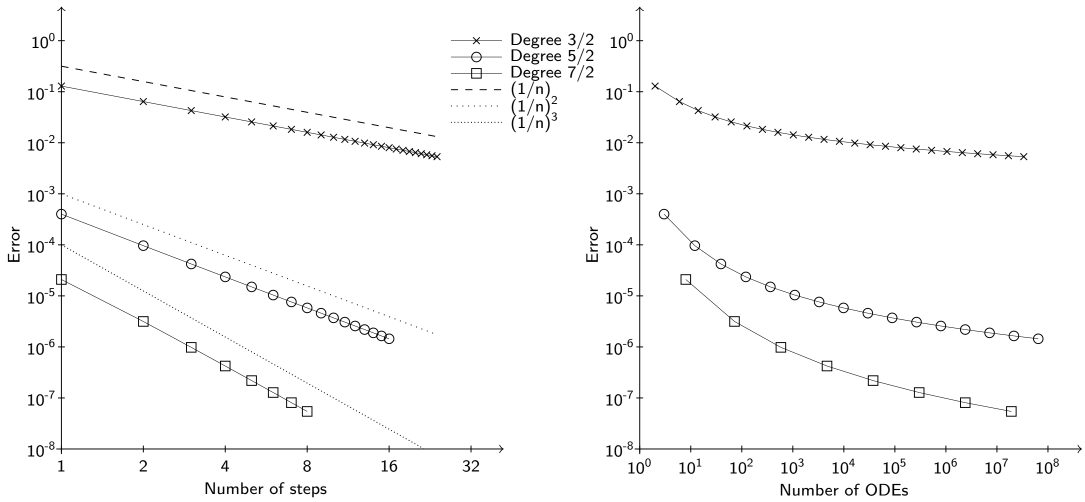

Figure 5.2: Error of approximating $\mathbb{E}\left[A_T\right]$

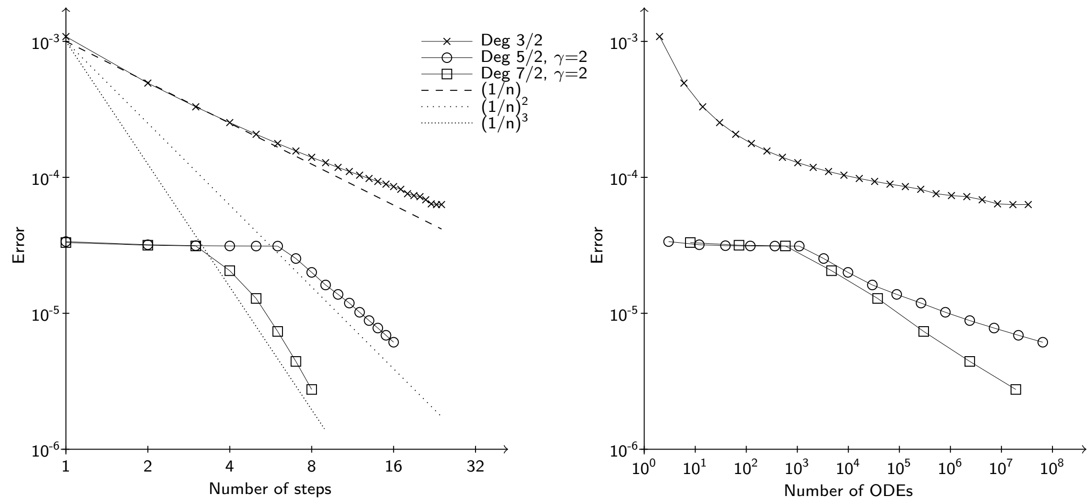

Figure 5.3: Error of valuing the ITM option

Observation 5.3.3. In the case of the approximation of the ITM option (figure 5.3), the essentially most accurate algorithms correspond to the Runge-Kutta order 4, Runge-Kutta order 6 and Runge-Kutta order 6 solvers respectively. The ITM payoff is not smooth, which might explain why the global order of convergence of the degree 3/2 cubature formula is slower than 1. In the case of the degree 5/2 and 7/2 cubature formulas, the uneven partition corresponding to the $\gamma = 2$ parameter proved to be the most accurate among the observed ones. The theoretical order of convergence is reached after a few steps.

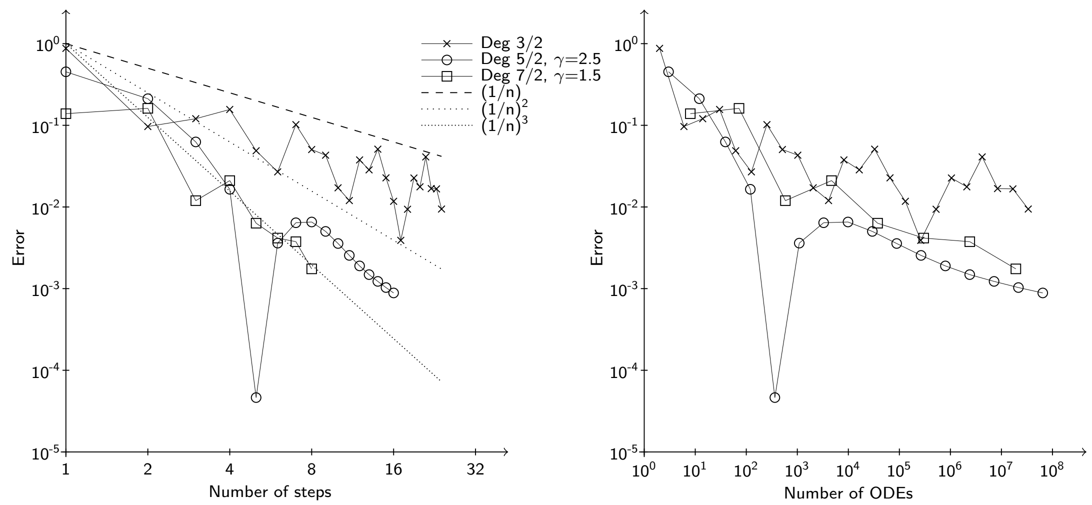

Figure 5.4: Error of valuing the OTM option

Observation 5.3.4. The OTM option (figure 5.4) again corresponds to a non-smooth terminal condition and an uneven partition is required to find the (near) optimal order of convergence. In the case of the degree 5/2 and 7/2 cubature formulas, we could observe the theoretical order of convergence after a couple of steps. Interestingly, if we consider the accuracy expressed in terms of the number of ODEs solved, we find that the degree 5/2 cubature formula is more efficient than the degree 7/2 one.

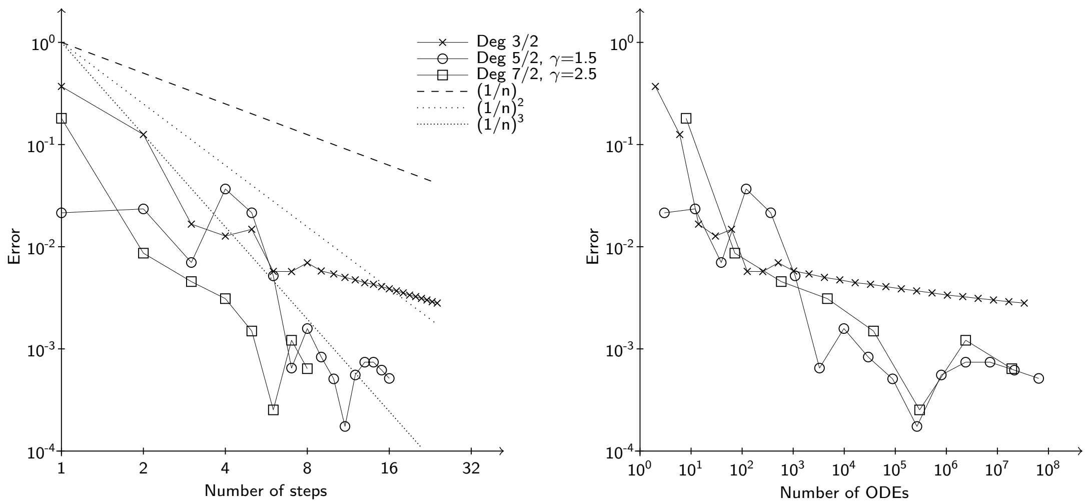

Figure 5.5: Estimated error of valuing the Asian option

Observation 5.3.5. The case of the Asian option (figure 5.5) is very similar to the case of the OTM option. The degree 5/2 and 7/2 formulas behave better when run on uneven partitions. Expressing the accuracy in terms of number of ODEs solved, in the observed cases the degree 5/2 cubature formula outperforms the degree 7/2 cubature formula.

5.3.2 NON-LINEAR SDE IN ONE DIMENSION

The second test case set in 1 dimension is based on the CIR SDE

$$\begin{aligned} \mathrm{d}r_t &= a(b - r_t)\mathrm{d}t + \sigma\sqrt{r_t}\mathrm{d}B_t \\ \mathrm{d}q_t &= r_t\mathrm{d}t \end{aligned} \tag{5.19}$$

run with $a = 0.08$, $b = 0.06$, $\sigma = 0.08$, $r_0 = 0.075$ and $q_0 = 0$ on $[0, T]$ for $T = 1.5$. Under these parameters, $r_t > 0$ with probability one. To ensure that the numerical solutions have the same property, we truncated the numerical approximations from below at a small $\varepsilon > 0$.

The following terminal conditions were implemented:

$$g_1(x, y) = x,\ g_2(x, y) = y,\ g_3(x, y) = \max(x, 0.09),\ g_4(x, y) = \max(x, 0.045)$$

At the Lie algebra level, we used the same degree 3/2, 5/2 and 7/2 cubature formulas as in the case of the linear SDE.

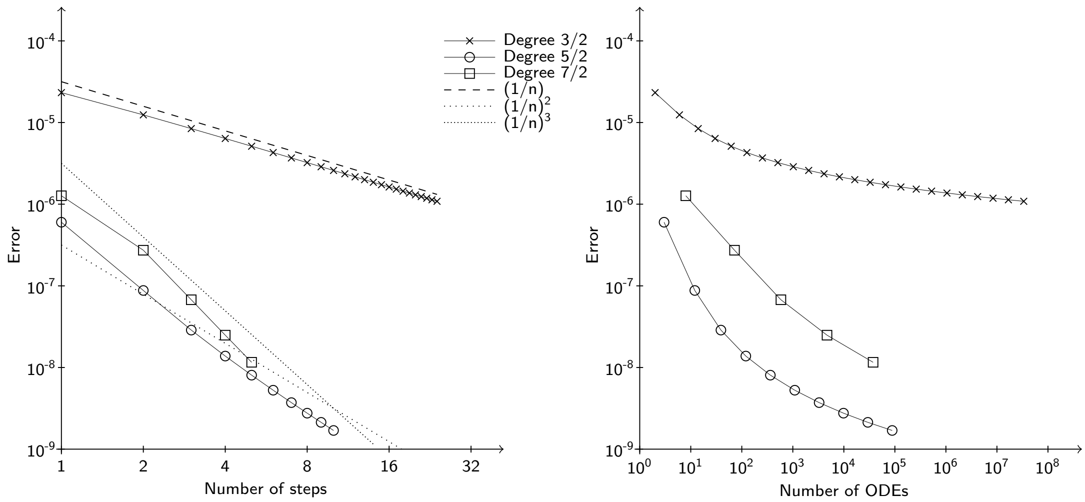

Figure 5.6: Error of approximating $\mathbb{E}\left[r_T\right]$

Observation 5.3.6. When approximating $\mathbb{E}\left[r_T\right]$ (figure 5.6), the global orders of convergence of the degree 3/2, 5/2 and 7/2 formulas match the theoretical orders. However, in terms of the number of ODEs solved, the degree 7/2 formula is less efficient compared to the degree 5/2 cubature formula.

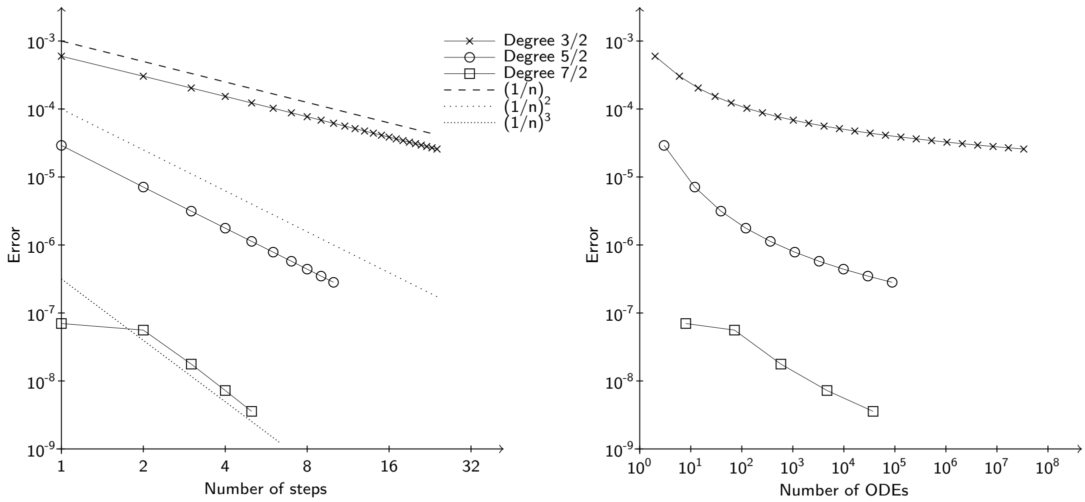

Figure 5.7: Error of approximating $\mathbb{E}\left[q_T\right]$

Observation 5.3.7. When approximating $\mathbb{E}\left[q_T\right]$ (figure 5.7), the global orders of convergence of the degree 3/2, 5/2 and 7/2 formulas match the theoretical orders. Moreover the degree 7/2 cubature formula outperforms the degree 5/2 one.

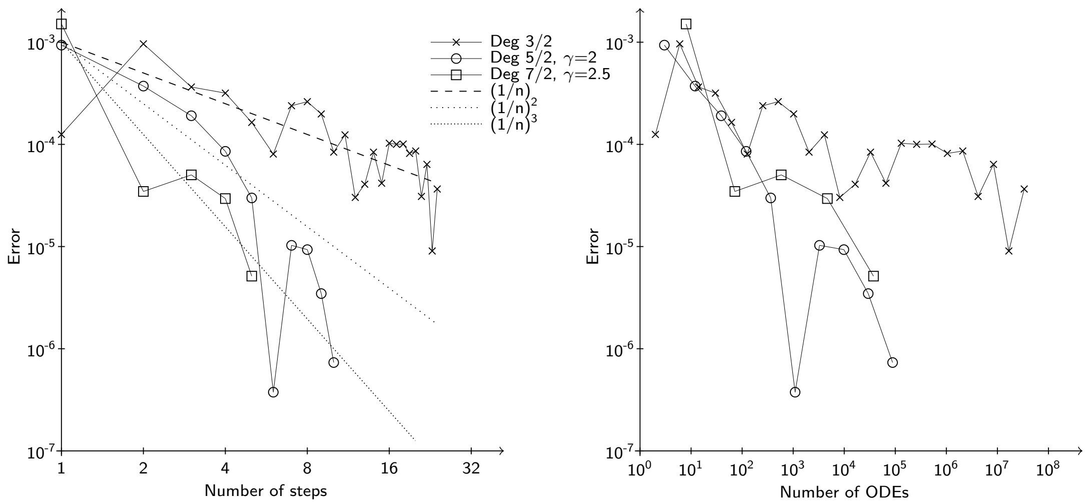

Figure 5.8: Estimated error of $\mathbb{E}\left[g_3(r_T)\right]$

Observation 5.3.8. In the cases of the terminal conditions g_3 and g_4 (figures 5.8 and 5.9 respectively) uneven partitions were required to find a near optimal performance. As the performance of the cubature formulas of different degrees were diverse, it is not clear whether the asymptotic behavior has been reached in the observed cases.

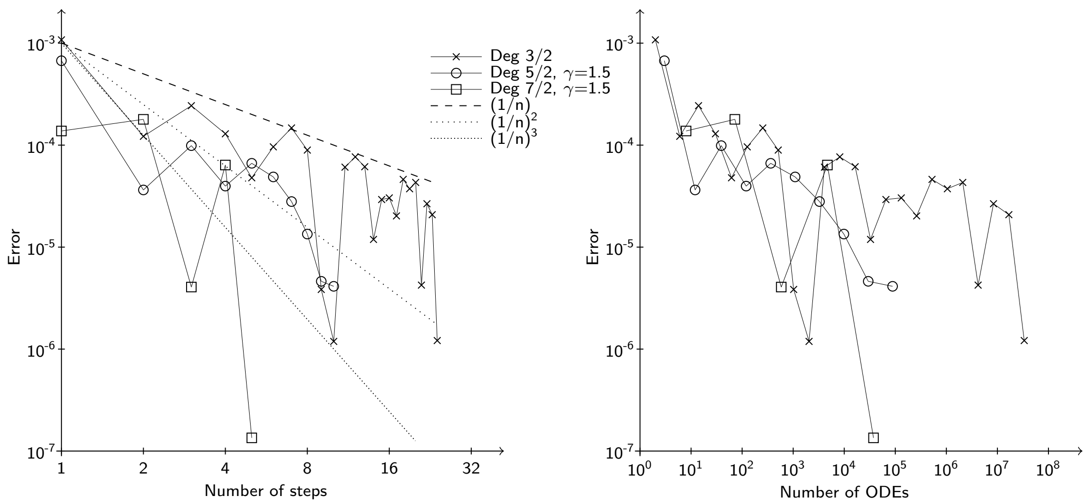

Figure 5.9: Estimated error of $\mathbb{E}\left[g_4(r_T)\right]$

5.3.3 STOCK WITH HESTON VOLATILITY

The 2-dimensional test problem is based on the stock price process with stochastic Heston volatility:

$$\begin{aligned}
\mathrm{d}v_t &= \alpha(\theta - v_t)\mathrm{d}t + \beta\sqrt{v_t}\mathrm{d}B_t^1 \\
\mathrm{d}S_t &= aS_t + \rho S_t\sqrt{v_t}\mathrm{d}B_t^1 + \sqrt{1-\rho^2}S_t\sqrt{v_t}\mathrm{d}B_t^2 \\
\mathrm{d}A_t &= S_t\mathrm{d}t \\
\mathrm{d}W_t &= v_t\mathrm{d}t
\end{aligned}$$

run with $a = 0.05$, $\alpha = 1$, $\theta = 0.3$, $\beta = 0.3$ and $\rho = 0.8$ on $[0, T]$ for $T = 0.5$. We tested the following terminal conditions:

$$\begin{aligned}
&h_1(v,x,y,z) = v,\ h_2(v,x,y,z) = x, \\
&h_3(v,x,y,z) = \max(x - 1.1y/T, 0),\ h_4(v,x,y,z) = \max(y/T - 51, 0)
\end{aligned}$$

where the payoff functions h_3 and h_4 correspond to two types of Asian options.

The implemented degree 3/2, 5/2 and 7/2 cubature formulas have 4, 14 and 48 elements in their support respectively. The referred cubature formulas were run up to 12, 6 and 4 steps respectively.

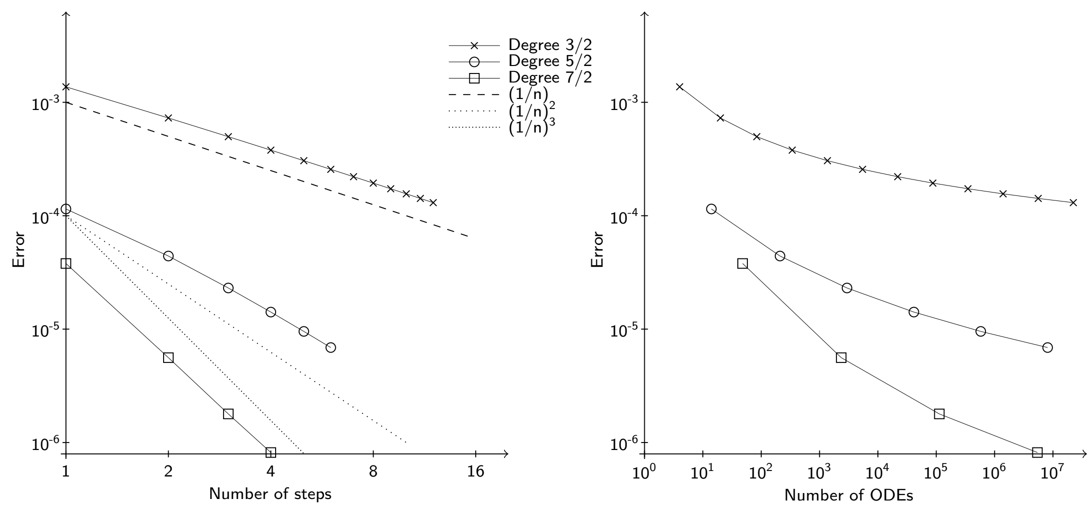

Figure 5.10: Error of approximating $\mathbb{E}[v_T]$

Observation 5.3.9. Since the volatility process is of the form of the CIR process, the approximation of $\mathbb{E}[v_T]$ (figure 5.10) shows similar performance to the CIR approximation although the degree 7/2 formula seems to have a convergence order smaller than the expected 3. This might be because the asymptotic region has not been reached yet. The approximation of $\mathbb{E}[S_T]$ (figure 5.11) performs as expected.

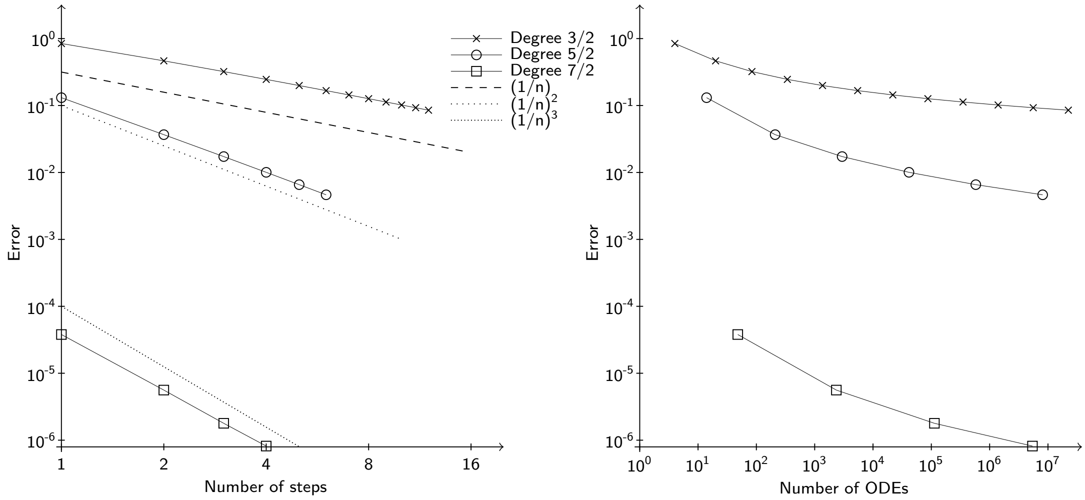

Figure 5.11: Error of approximating $\mathbb{E}\left[S_T\right]$

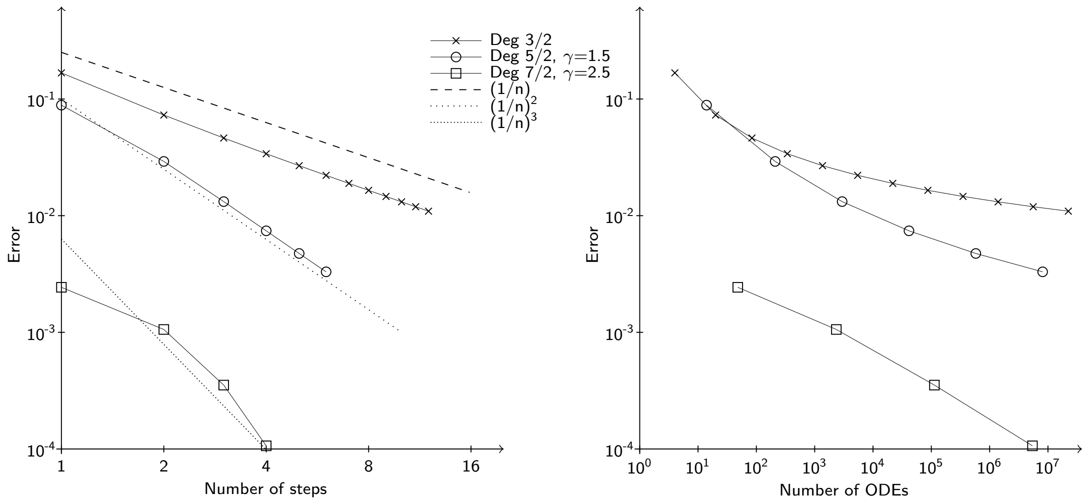

Figure 5.12: Estimated error of the Asian option 1

Observation 5.3.10. The approximation of the Asian options (figures 5.12 and 5.13) resulted in smooth error curves and in some cases convergence order higher than the theoretical ones. However the asymptotic regions might have not been reached due to the huge support of the global cubature measures of degree 7/2 at higher steps. In the observed cases, the degree 7/2 cubature formula proved to be the most efficient, especially when the accuracy was considered in terms of number of ODEs solved.

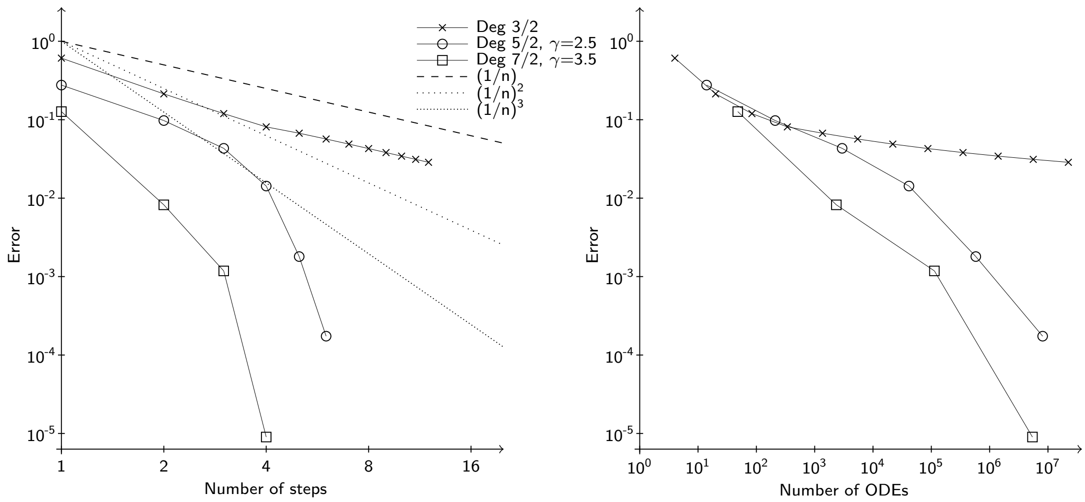

Figure 5.13: Estimated error of the Asian option 2

5.3.4 Tree based branching algorithm applied in 4 dimensions

The four dimensional test problem is defined by the following SDE:

$$\begin{aligned} \mathrm{d}v_t &= \alpha(\theta - v_t)\mathrm{d}t + \beta\sqrt{v_t}\mathrm{d}B_t^1 \\ \mathrm{d}S_t^1 &= a_1 S_t^1 + \sigma_1\rho S_t^1\sqrt{v_t}\mathrm{d}B_t^1 + \sqrt{1-\rho^2}\sigma_1 S_t^1\sqrt{v_t}\mathrm{d}B_t^2 \\ \mathrm{d}S_t^2 &= a_2 S_t^2 + \sigma_2\rho S_t^2\sqrt{v_t}\mathrm{d}B_t^1 + \sqrt{1-\rho^2}\sigma_2 S_t^2\sqrt{v_t}\mathrm{d}B_t^3 \\ \mathrm{d}S_t^3 &= a_3 S_t^3 + \sigma_3\rho S_t^3\sqrt{v_t}\mathrm{d}B_t^1 + \sqrt{1-\rho^2}\sigma_3 S_t^3\sqrt{v_t}\mathrm{d}B_t^4 \end{aligned}$$

run with $\alpha = 1$, $\theta = 0.3$, $\beta = 0.3$, $\rho = 0.25$, $a_1 = 0.1$, $\sigma_1 = 2$, $a_2 = 0.5$, $\sigma_2 = 1$, $a_3 = 0.15$, $\sigma_3 = 3$, $v_0 = 0.1$, $S_0^1 = S_0^2 = S_0^3 = 100$ on $[0, T]$ for $T = 0.5$.

The results corresponding to the following terminal conditions are presented below:

$$\begin{aligned} j_1(v, x, y, z) &= \max(x, y, z) \\ j_2(v, x, y, z) &= \min(\max(\max(x, y, z) - 1.25K, 0), 0.75K) \end{aligned}$$

where $k = 102$. The payoff j_1 corresponds to a basket option, and the payoff j_2 corresponds to a spread option on the maximum of the three stock prices.

We tested a degree $5/2$ cubature formula with 50 elements in the support. We had the computational capacity to work out the exact global cubature measure up to 4 steps. Since we aimed to observe the asymptotic behavior, i.e. the behavior at more steps, we combined the KLV method with the tree-based branching algorithm (TBBA). The global cubature measure can be regarded and constructed from a tree whose nodes hold the *local cubature measures*. The TBBA was applied at each node; i.e. we sampled at each node from the local cubature measures. We ran the TBBA on a $1, 2, 4$ and 8 step cubature trees with $1000, 10000, 100000$ and 1250000 particles starting at the root of the tree.

Figures 5.14 and 5.15 demonstrate the approximation results of the pure cubature version and the TBBA run with 100000 and 1250000 particles. In the case of the partial sampling versions, confidence intervals are plotted and should be considered when evaluating the efficiency of those versions. In Figure 5.16, the estimated variance of the partial sampling based version is plotted as a function of number of particles.

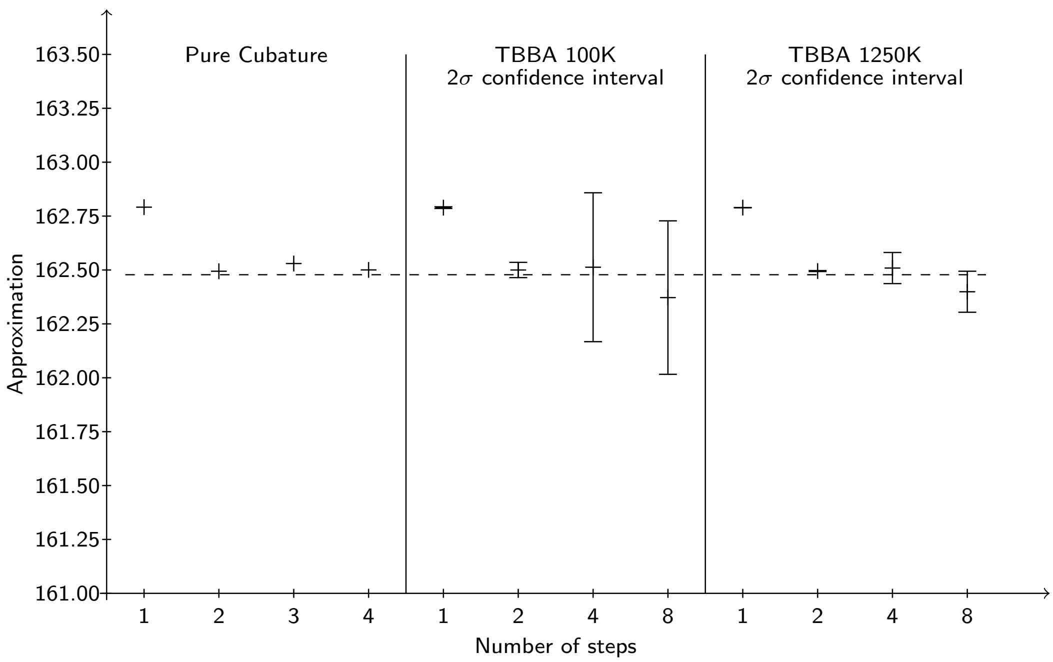

Figure 5.14: Estimating the max

Observation 5.3.11. Figure 5.16 provides important information on the efficiency of the TBBA and Monte-Carlo based partial sampling versions. The variance of the Monte-Carlo based partial sampling is expected to be of order $1/\sqrt{n}$ where n denotes the number of particles. Our estimates do not contradict this expectation. The variance of the TBBA based partial sampling is expected to decay at a faster rate. This faster rate is clearly verified

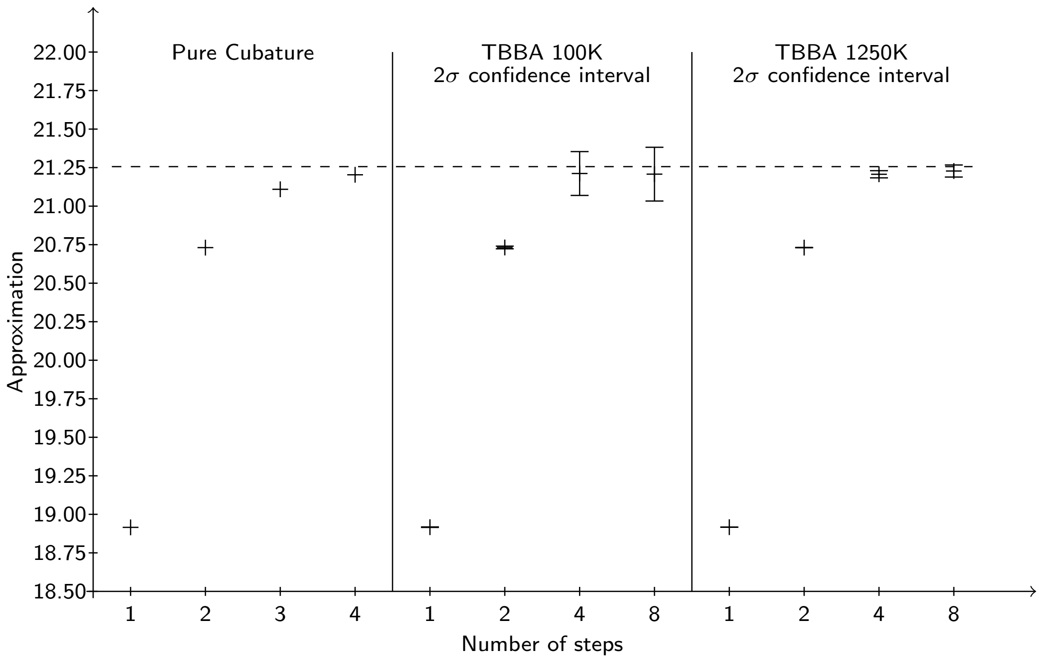

Figure 5.15: Estimating the value of a basket spread option

when the sampling is done from the 1 and 2 step cubature tree. However when sampling from the 4 and 8 step tree, the region of faster decay is shifted towards higher numbers of particles.

We can also observe that the variance of both of the TBBA and the MC run with fixed number of particles increases with the number of steps in the cubature tree. Even when the KLV algorithm approximates the actual variance of the payoff, the convergence might be slow (although the curves corresponding to the 4-step and 8-step versions are fairly close to each other, there is a huge gap between the 1-step and 2-step curves as well as the 2-step and 4-step curves). Therefore the small variance of the estimates at a lower number of steps does not imply similarly low variance of the estimates at a higher number of steps. This is especially critical if the computational capacity only allows runs at steps where the variance is sharply increasing. This observation can be regarded as a addendum to Remark 4 of [33].

5.3.5 Some special issues

In this section we discuss some *tricks* which significantly improve the accuracy of the approximations in certain cases and do not improve or even weaken the accuracy of the approximations in some other cases.

One such trick is referred to as *Romberg extrapolation*. Heuristically, the Romberg extrap-

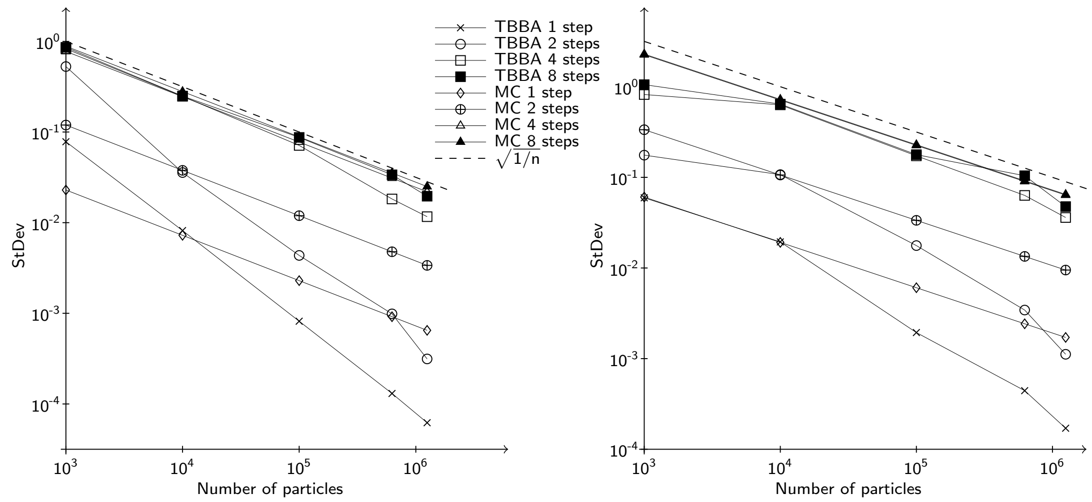

Figure 5.16: Variance of TBBA and Monte-Carlo estimates corresponding to the basket option (left-hand-side graph) and the spread option (left-hand-side graph)

olation is aimed at exploiting the case in which the global error $e(\mathcal{D})$ of an approximation $\widehat{X}_T(\mathcal{D})$ corresponding to the partition $\mathcal{D}$ of $[0, T]$ (asymptotically) satisfies

$$e(\mathcal{D}) \sim C_1|\mathcal{D}|^p + C_2|\mathcal{D}|^{p+\varepsilon} \tag{5.20}$$

for some $p, \varepsilon > 0$ and real numbers C_1, C_2 not depending on $\mathcal{D}$. Then due to the cancellation of the order p terms, the formula

$$\widehat{X}_T^{\text{Romberg}}(\mathcal{D}) := \frac{2^p}{2^p - 1}\widehat{X}_T(\widehat{\mathcal{D}}) - \frac{1}{2^p - 1}\widehat{X}_T(\mathcal{D}) \tag{5.21}$$

results in an approximation of order $p + \varepsilon$, where $\widehat{\mathcal{D}}$ denotes the partition constructed from $\mathcal{D}$ by halving each subinterval in $\mathcal{D}$. If the approximation oscillates around the exact value, the formula (5.21) can be replaced with

$$\widehat{X}_T^{\text{Romberg}}(\mathcal{D}) := \frac{2^p}{2^p + 1}\widehat{X}_T(\widehat{\mathcal{D}}) + \frac{1}{2^p + 1}\widehat{X}_T(\mathcal{D}). \tag{5.22}$$

We highlight two potential difficulties when attempting to apply the Romberg extrapolation. Firstly, the constant C_1 in (5.20) might be bounded but not stable enough, implying that the required cancellation does not happen in general. Secondly, the formula (5.21) requires the knowledge of the constant p, which is hard to estimate in some cases (e.g. when the KLV is run on an uneven partition).

We applied the Romberg extrapolation to some of the approximations presented above. We found that the extrapolation in general did not improve the convergence order or the efficiency of the KLV algorithm. We demonstrate this observation for the following three examples (figures 5.17, 5.18 and 5.19).

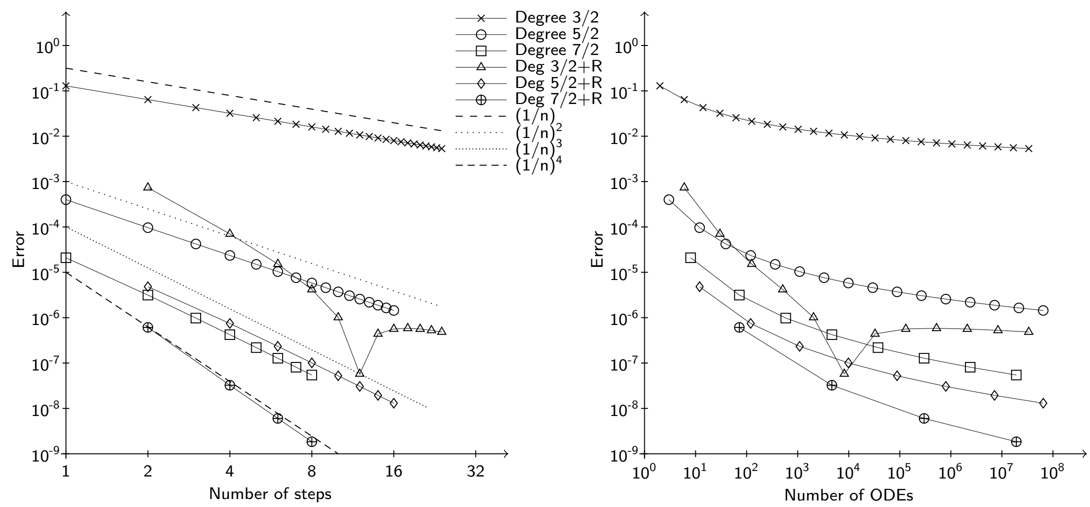

Figure 5.17: Romberg extrapolation applied to the 1D linear problem ($\mathbb{E}\left[A_T\right]$)

Observation 5.3.12. We applied the Romberg extrapolation to the linear SDE-based 1-dimensional problem (in Figure 5.17, the results corresponding to the extrapolation are labeled with Deg 3/2+R, Deg 5/2+R and Deg 7/2+R). The case of $\mathbb{E}\left[A_T\right]$ reflects the ideal case. In particular the convergence orders of the degree 5/2 and 7/2 cubature formulas are improved to 3 and 4 respectively. Considering the accuracy expressed in terms of number of ODEs solved (right-hand-side graph of figure 5.17), the Romberg extrapolation shows a huge improvement since both the Deg 5/2+R and Deg 7/2+R curves are below the Degree 5/2 and Degree 7/2 curves.

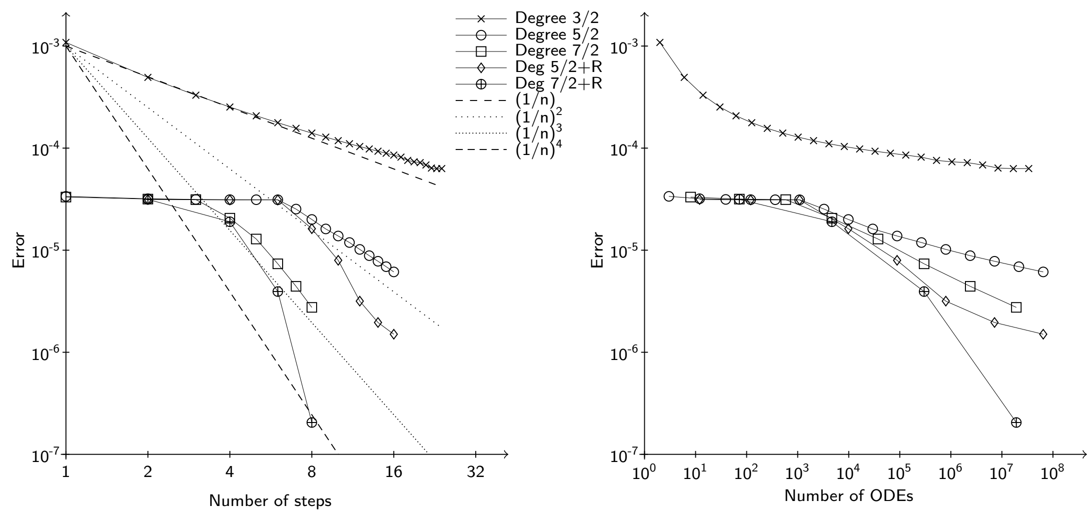

Figure 5.18: Romberg extrapolation applied to the 1D linear problem (ITM option)

Observation 5.3.13. In the case of the ITM option (Figure 5.18), the theoretical order of convergence of the pure KLV method is reached after a couple of steps, implying that the Romberg extrapolation reflects the expected property at even higher steps. Considering the accuracy expressed in terms of number of ODEs solved, the Romberg extrapolation provides a moderate improvement.

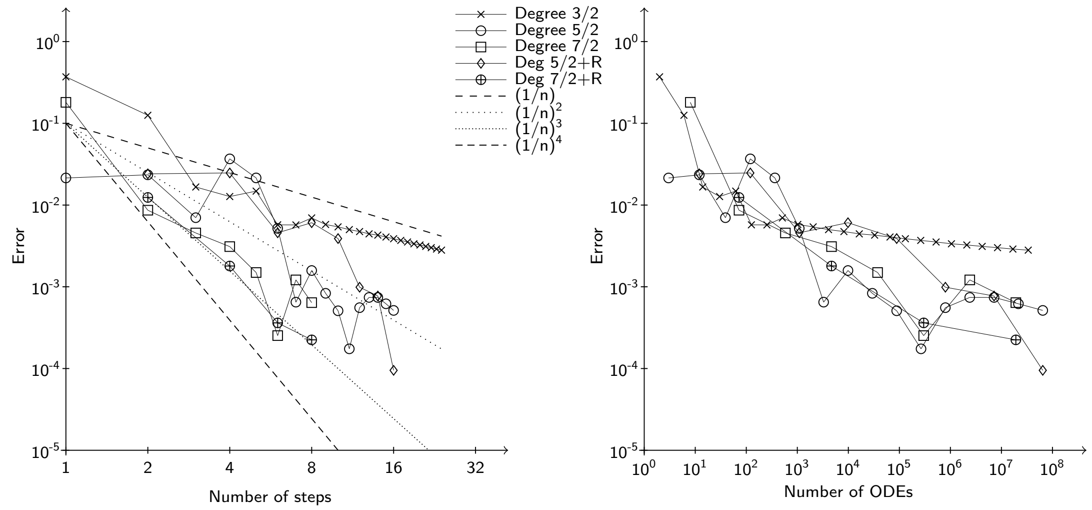

Figure 5.19: Romberg extrapolation applied on the 1D linear problem (Asian option)

Observation 5.3.14. The case of the Asian option (figure 5.19) presents a situation in which the Romberg extrapolation does not provide an obvious improvement. This is probably implied by the unknown rate of global convergence (an uneven partition is used due to the non-smooth terminal condition). On the other hand, the instability of the constant C might also be an issue.

We have already considered the use of uneven partitions, and for each case presented the approximation corresponding to an uneven partition resulting in a near optimal accuracy. We demonstrate the diverse effect of the uneven partitions in the following two examples (figure 5.20).

Observation 5.3.15. In general we observed that the KLV run on uneven partitions corresponding to higher γ parameters resulted in smoother error curves (as a function of number of steps). Too high γ parameters decrease the order of convergence. In the case of the ATM option (left graph of Figure 5.20), among the observed partition the one corresponding to $\gamma = 2.5$ resulted in the most accurate approximation at a higher number of steps. In the case of the ITM option (right graph of the figure) $\gamma = 1.5$ performed well. This might be explained by the fact that the pay-off of a deep in the money option technically can be regarded as smooth payoff.

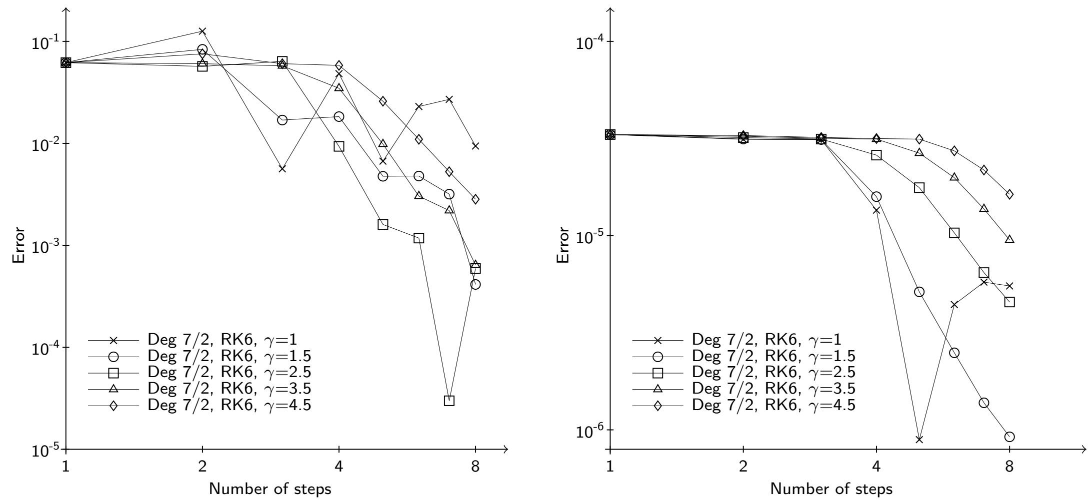

Figure 5.20: Testing uneven partitions on the 1D linear problem (ATM option on the left hand side and ITM option on the right hand side)

Finally we tested how sensitive the KLV approximation could be to the choice of numerical ODE solver. We know that when using ODE solvers of low order we might lose the global order of convergence of the KLV method (ref. Section 5.2.1). However higher order and geometrically stable (in the sense of Remark 3.4.3) solvers do not just ensure the theoretical rate of convergence but might improve the accuracy. Technically, we tested the effect of conditions (5.14) and (5.15). We demonstrate the observed properties on the 1-dimensional degree 3/2 and degree 7/2 cubature formulas (Figure 5.21).

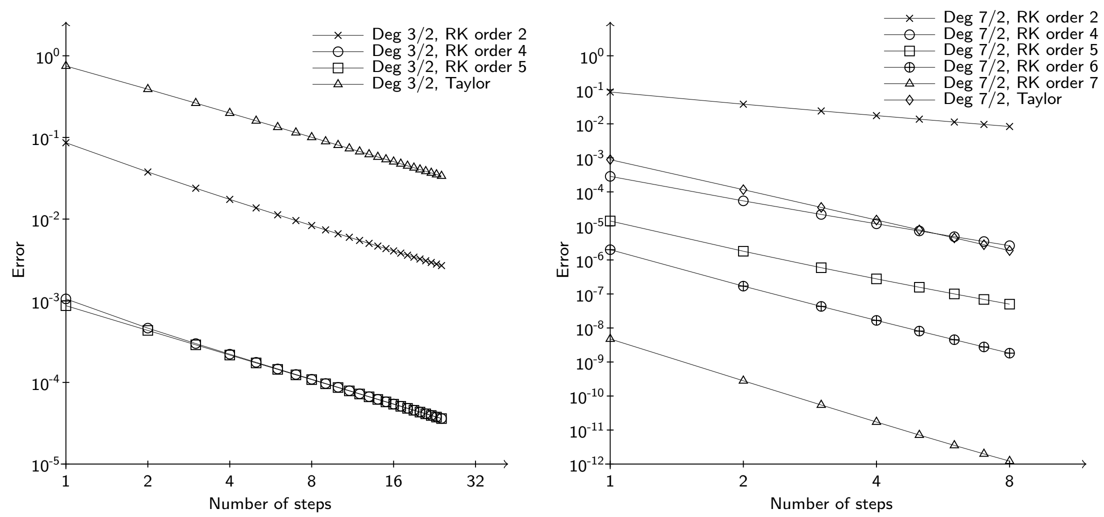

Figure 5.21: KLV with different ODE solvers applied to the 1D linear problem ($\mathbb{E}[S_T]$)

Observation 5.3.16. The degree 3/2 cubature formula (left-hand-side graph of Figure 5.21) combined with second order ODE solvers resulted in global convergence order 1. Higher order ODE solvers improved the accuracy but not the rate of convergence. The fourth order Runge-Kutta solver resulted in an approximation essentially equivalent to the fifth order Runge-Kutta solver based approximation.

In the case of the degree 7/2 cubature formula (right graph of figure 5.21), the ODE solvers of order at least 5 resulted in a third order global convergence. The approximations based on lower order ODE solvers had lower global order (first and second).

In both of the above two examples as well as all the other examples, which we tested with several ODE solvers, we found that typically the high order Taylor expansion based ODE solver resulted in approximations matching the theoretical global order, just like the Runge-Kutta solvers of the same order, however the latter ones had better accuracy, in some cases even 5 digit better accuracy. This observation motivates the use of more stable ODE solver than the high order Taylor expansion based ones (the only ODE solvers which were considered in the first versions of Kusuoka's approach).

APPENDIX A

TREE-BASED BRANCHING ALGORITHM

In the present section, we give a brief overview of the *Tree-based Branching Algorithm* (TBBA) (following Crisan & Lyons [7]).

Let I be a countable set and $\mathcal{P}(I)$ the set of probability measures on I:

$$\mathcal{P}(I) = \left\{ (x_i)_{i\in I} \in \mathbb{R}^I \Big| \sum_{i\in I} x_i = 1,\ x_i \geq 0\ \forall i \in I \right\}. \tag{A.1}$$

The Borel σ-field $\mathcal{B}(\mathcal{P}(I))$ is generated by the cylinder sets. We introduce a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathcal{F}$ is a σ-field with no atoms. Let $a \in \mathcal{P}$ be a probability measure with finite entropy $H(a) = \sum_{i\in I} a(i) \log a(i) < \infty$ and let $\tilde{a} : \Omega \to \mathcal{P}(I)$ be a $\mathcal{B}(\mathcal{P}(I))/\mathcal{F}$ measurable map, i.e. a random probability measure. We define the set $\mathcal{S}_n^a$ as

$$\mathcal{S}_n^a = \left\{ \tilde{a} : \Omega \to \mathcal{P}(I) \Big| \tilde{a} = \frac{1}{n} \sum_{k=1}^{n} \delta_{Z_k},\ Z_k : \Omega \to I,\ \text{measurable},\ \mathbb{E}\left[\tilde{a}\right] = a \right\},$$

where δ_x denotes the Dirac distribution concentrated on x.

Example 1. If we define Z_k $k = 1, \ldots, n$ to be independent random variables having the same distribution as a, then we obtain the naive Monte Carlo sampling measure $\tilde{a}_n^{MC}$.

Our aim is to find an $\tilde{a} \in \mathcal{S}_n^a$ which has only a small discrepancy with respect to a. As a measure of such discrepancy, [7] defines the relative entropy $H(\tilde{a}|a)$ of $\tilde{a}$ with respect to a as

$$H(\tilde{a}|a) = \mathbb{E}\left[\sum_{i\in I} \tilde{a}(i) \log \frac{\tilde{a}(i)}{a(i)} \right].$$

To describe the optimal random measures, [7] introduces the minimal variance property as follows.

Definition A.1. *$\tilde{a} \in \mathcal{S}_n^a$ has the property $MV(n)$, if*

$$\tilde{a}(i) = \begin{cases} \frac{\lfloor na(i) \rfloor}{n} & \textit{with probability } 1 - \{na(i)\} \\ \frac{\lfloor na(i) \rfloor + 1}{n} & \textit{with probability } \{na(i)\} \end{cases} \quad \forall i \in I$$

where for a real x, $\lfloor x \rfloor$ denotes the largest integer less than or equal to x and $\{x\} = x - \lfloor x \rfloor$.

Theorem A.1. *i) Let $\tilde{a} \in \mathcal{S}_n^a$. Then $H(\tilde{a}|a)$ is minimal if and only if $\tilde{a}$ has the property $MV(n)$.*

ii) Random measures $\tilde{a} \in \mathcal{S}_n^a$ with property $MV(n)$ exist and can be constructed using a tree-based branching algorithm.

See [7] for proof and details of the construction. The random measure obtained by the TBBA is denoted by $\tilde{a}_n^{TB}$. [31] also gives an error estimate as follows.

Theorem A.2. *Let f be a function defined on I and $\sum_{i\in I}(f(i))^2 < \infty$. Then*

$$\mathbb{E}_\Omega^\mathbb{P}\left[\left(\sum_{i\in I} f(i)\left(a(i) - \tilde{a}_n^{TB}(i)\right)\right)^2\right] \le \frac{1}{4n^2}\sum_{i\in I}(f(i))^2.$$

A similar estimate for $\tilde{a}_n^{MC}$ is given as

$$\mathbb{E}_\Omega^\mathbb{P}\left[\left(\sum_{i\in I} f(i)\left(a(i) - \tilde{a}_n^{MC}(i)\right)\right)^2\right] \le \frac{1}{n}\left(1 - \sum_{i\in I}(a(i))^2\right)\left(\sum_{i\in I}(f(i))^2\right).$$

Bibliography

[1] Gérard Ben Arous, 1989, *Flots et séries de Taylor stochastiques*, Probab. Theory Related Fields, 81(1):29-77

[2] J.C. Butcher, 2008, *Numerical Methods for Ordinary Differential Equations*, Second Edition, John Wiley and Sons Ltd

[3] M. Caruana, 2007, *The Existence of Solutions for Rough Differential Equations*, DPhil Thesis, Mathematical Institute, University of Oxford

[4] F. Castell, 1993, *Asymptotic expansion of stochastic flows*, Probability Theory and Related Fields, Springer-Verlag

[5] K. T. Chen, 1957, *Integration of paths, geometric invariant and generalized Campbell-Baker-Hausdorff formula*, Ann. Math. II 65, 163-178

[6] J.M.C. Clark & R.J. Cameron, 1980, *The maximum rate of convergence of discrete approximations for stochastic differential equations*, Lecture Notes in Control and Information Sciences, Springer Berlin/Heidelberg, Volume 25, 162-171

[7] D Crisan & T Lyons, 2002, *Minimal Entropy Approximations and Optimal Algorithms*, Monte Carlo methods and Appl., Vol. 8, No. 4, pp.343-355

[8] A. M. Davie, 2007, *Differential Equations Driven by Rough Paths: An Approach via Discrete Approximation*, Appl Math Res Express, Volume 2007, article ID abm009

[9] P. Friz & N. Victoir, 2006, *Euler estimates of rough differential equations*, Preprint, http://arxiv.org/abs/math/0602345

[10] J. Gaines & T.J. Lyons, 1997, *Variable step size control in the numerical solution of stochastic differential equations*, SIAM J. Appl. Math 57

[11] T. Gard, 1988, *Introduction to Stochastic Differential Equations*, Marcel Dekker Ltd

[12] L.G. Gyurkó & T.J. Lyons, 2008, *Rough Paths based Numerical Algorithms in Computational Finance*, preprint

[13] Y. Hu, 1990, *Calculs formels sur les e.d.s. de Stratonovich*, Séminaire de Probabilités XXIV 1988-89

[14] N. Ikeda, & S. Watanabe, 1981, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam

[15] P.E. Kloeden, & E. Platen, 1999, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag Berlin Heidelberg New York

[16] S. Kusuoka, & D. Stroock, 1987, *Application of the Malliavin calculus III* J. Fac. Sci. Univ. Tokyo, 1A 34, 391-442

[17] S. Kusuoka, 2001, *Approximation of expectation of diffusion process and mathematical finance*, Adv. Stud. Pure Math., 31, 147-165

[18] S. Kusuoka, 2003, *Approximation of expectation of diffusion process based on Lie algebra and Malliavin calculus*, preprint

http://kyokan.ms.u-tokyo.ac.jp/users/preprint/pdf/2003-34.pdf

[19] S. Kusuoka, 2003, *Malliavin calculus Revisited*, J. Math. Sci. Univ. Tokyo 10, 261-277

[20] S. Kusuoka & S Ninomiya, 2004, *A New Simulation Method of Diffusion Processes Applied to Finance*, appeared in Stochastic Processes and Applications to Mathematical Finance: Proceedings of the Ritsumeikan International Symposium, Kusatsu, Shiga, Japan, 5-9 March 2003

[21] A. Lejay & N.Victoir, 2006, *On (p,q)-rough paths*, J. Differential Equations 225, no. 1, 103–133.

[22] C. Litterer, 2008, *The Signature in Numerical Algorithms*, DPhil Thesis, Mathematical Institute, University of Oxford

[23] C. Litterer & T.J. Lyons, 2006, *Cubature on Wiener space continued*, appeared in Stochastic Processes and Applications to Mathematical Finance, Proceedings of the 6th Ritsumeikan International Symposium Ritsumeikan University, Japan, 6 - 10 March 2006

[24] R. Lord, R. Koekkoek & D. van Dijk, 2006, *A Comparison of Biased Simulation Schemes for Stochastic Volatility Models*, Tinbergen Institute Discussion Paper

http://www.tinbergen.nl/discussionpapers/06046.pdf

[25] T.J. Lyons, 1998, *Differential equations driven by rough signals*, Revista Mathematica Iber. Vol 14, Nr. 2, 215-310

[26] T.J. Lyons & Z. Qian, 2002, *System Control and Rough Paths*, Oxford mathematical monographs, Clarendon Press, Oxford

[27] T.J. Lyons, M. Caruana, T. Lévy, 2007, *Differential Equations Driven by Rough Paths*, Ecole d'Eté de Porbabilités de Saint-Flour XXXIV - 2004, Lecture Notes in Mathematics, Springer

[28] T.J. Lyons & N. Victoir, 2004, *Cubature on Wiener Space*, Proc. R. Soc. Lond. A 460, 169-198

[29] W. Magnus, 1954, *On the exponential solution of differential equations for a linear operator*, Commun. Pure Appl. Math. 7

[30] N. Newton, 1986, *An asymptotically efficient difference formula for solving stochastic differential equations*, Stochastics An International Journal of Probability and Stochastic Processes, Volume 19, Issue 3, 175-206

[31] S. Ninomiya, 2003, *A partial sampling method applied to the Kusuoka approximation*, Monte Carlo Methods and Appl., Vol. 9, No. 1, pp.27-38

[32] M. Ninomiya & S. Ninomiya, 2008, *A new weak approximation scheme of stochastic differential equations by using the Runge-Kutta method*, preprint

http://arxiv.org/PS_cache/arxiv/pdf/0709/0709.2434v2.pdf

[33] S. Ninomiya & N. Victoir, 2006, *Weak approximation of stochastic differential equations and application to derivative pricing*, preprint

http://arxiv.org/PS_cache/math/pdf/0605/0605361v3.pdf

[34] P. Simon & J. Tóth, 2005, *Differenciálegyenletek, Bevezetés az elméletbe és az alkalmazásokba*, Typotex Kiadó, Budapest

[35] R.S. Strichartz, 1987, *The Campbell-Baker-Hausdorff-Dynkin Formula and Solutions of Differential Equations*, Journal of Functional Analysis 72, 320-345

[36] A. H. Stroud, 1971, *Approximate Calculation of Multiple Integrals*, Prentice Hall

[37] S.C.P. Yam, 2008, *Analytical and topological aspects of signatures*, DPhil Thesis, Mathematical Institute, University of Oxford