Order book models, signatures and numerical approximations of rough differential equations

Arend Janssen
Magdalen College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity Term 2011

Order book models, signatures and numerical approximations of rough differential equations

Arend Janssen
Magdalen College
University of Oxford

A thesis submitted for the
degree of Doctor of Philosophy,
Trinity Term 2011

Abstract

We construct a mathematical model of an order driven market where traders can submit limit orders and market orders to buy and sell securities. We adapt the notion of no free lunch of Harrison and Kreps and Jouini and Kallal to our setting and we prove a no-arbitrage theorem for the model of the order driven market.

Furthermore, we compute signatures of order books of different financial markets. Signatures, i.e. the full sequence of definite iterated integrals of a path, are one of the fundamental elements of the theory of rough paths. The theory of rough paths provides a framework to describe the evolution of dynamical systems that are driven by rough signals, including rough paths based on Brownian motion and fractional Brownian motion (see the work of Lyons). We show how we can obtain the solution of a polynomial differential equation and its (truncated) signature from the signature of the driving signal and the initial value.

We also present and analyse an ODE method for the numerical solution of rough differential equations. We derive error estimates and we prove that it achieves the same rate of convergence as the corresponding higher order Euler schemes studied by Davie and Friz and Victoir. At the same time, it enhances stability. The method has been implemented for the case of polynomial vector fields as part of the *CoRoPa* software package which is available at `http://coropa.sourceforge.net`. We describe both the algorithm and the implementation and we show by giving examples how it can be used to compute the pathwise solution of stochastic rough differential equations driven by Brownian rough paths and fractional Brownian rough paths.

Acknowledgements

I would like to thank my supervisor, Professor Terry Lyons, for his invaluable support and guidance. In addition, I would like to thank Jonathan Field and Dr Anthony Ledford from Man Investments for their advice and the provision of market data of order books.

I gratefully acknowledge the support of the German Academic Exchange Service (DAAD) and the German National Academic Foundation (Studienstiftung des deutschen Volkes) without which it would not have been possible to write this thesis. Furthermore, I am very grateful to Magdalen College and the Oxford-Man Institute of Quantitative Finance.

At this point, I would also like to thank the organisers of the Conference on Combinatorics and Control in Madrid in 2010 for giving me the opportunity to present my work.

Finally, special thanks are due to my parents and Amy, Al and Markus.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis covers a range of topics related to order book models, signatures and numerical approximations of rough differential equations.

Financial markets function like auctions (see Smith et al. [44]). Traders submit limit orders and market orders to buy and sell securities. The outstanding orders are stored in an order book until execution or withdrawel and the transaction prices arise as a consequence of supply and demand. In chapter 2, we construct a mathematical model of such an order driven market which replicates the real trading mechanisms (see also [44]). We adapt the notion of no free lunch of Harrison and Kreps [29] and Jouini and Kallal [30] to our setting and we prove a no-arbitrage theorem for the model of the order driven market.

The theory of rough paths provides a framework to describe the evolution of dynamical systems that are driven by rough signals, including rough paths based on Brownian motion and fractional Brownian motion (see Lyons [36], Lyons et al. [37], Lyons and Qian [39] or Friz and Victoir [22]). Signatures, i.e. the full sequence of definite iterated integrals of a path, are one of the fundamental elements of the theory of rough paths. As an example, we compute the (log-) signatures of data of order books from different financial markets, including the FTSE 100 stock index future contract which is traded on the London International Financial Futures and

Options Exchange (LIFFE). The order book data has kindly been provided by *Man Investments*. The examples for signatures of order books are presented in chapter 3. In addition, chapter 3 deals with probability measures on signatures. A compactly supported probability measure on signatures is uniquely determined by its expectation. This result has already been known to Chen [12] and Fawcett [20]. We give a different proof which is based on the Stone-Weierstrass Theorem.

In chapter 4, we study the relationship between polynomial differential equations driven by paths with bounded variation and their signatures. It follows from Hambly and Lyons [28] that the value y_t of the solution of a polynomial differential equation

$$dy_t = P(y_t)\, d\gamma_t, \qquad y_0 = z$$

is uniquely determined by the signature $\Gamma_{0,t}$ of the driving signal γ over the interval $[0, t]$ and the initial value z. For small t, we derive series expansions that show how we can actually obtain the solution y_t and its truncated signature $Y_{0,t}^{[M]}$ for arbitrary integers $M \geq 1$ from the signature $\Gamma_{0,t}$ of the driving signal and the initial value z. For linear differential equations, such series expansions can be found in [36]. It turns out that every polynomial differential equation can be reduced to a quadratic differential equation on (truncated) signatures. Furthermore, we provide an algorithm for the numerical computation of the solution of a polynomial differential equation and its signature.

In chapter 5, we present and analyse an ODE method for the numerical solution of rough differential equations which has also been studied by Gyurkó [26]. It is an extension of a method described by Gaines and Lyons [24] and Gyurkó and Lyons [27] for stochastic differential equations driven by Brownian motion. The basic idea is to approximate the solution of the rough differential equation over a small time step by the solution of an ordinary differential equation (see Castell [10] or Castell and Gaines [11] for the case of stochastic differential equations driven by Brownian motion). We derive error estimates for this ODE method. Similar results have been obtained by Gyurkó [26] for the case of γ-Lipschitz vector fields and polynomial vector fields. We use a different set-up and we work with vector fields that are only locally γ-Lipschitz. It turns out that the ODE method achieves in general the same rate of convergence as the corresponding higher order Euler schemes for rough differential equations in Davie [17] and Friz and Victoir [23] or [22], chapter 10. At the same time, it enhances stability (see Gyurkó and Lyons [27]).

The numerical method has been implemented for the case of polynomial vector fields as part of the *CoRoPa* software package (see `http://coropa.sourceforge.net`) which has been jointly developed by various members of the Stochastic Analysis Group of the University of Oxford under the guidance of Terry Lyons, including Stephen Buckley, Djalil Chafai, Greg Gyurkó, Christian Litterer, Chang Liang Xu, Rahul Raghuram and the author of this thesis. We also describe the implementation and we show by giving examples how it can be used to compute the pathwise solution of stochastic rough differential equations driven by Brownian rough paths and fractional Brownian rough paths.

Chapter 2

Arbitrage in order driven markets

In the classical securities market model of Harrison and Kreps (see [29]), the asset price is described by a real-valued process $(S_t)_{t\in[0,T]}$. Traders can buy and sell assets at this price. Harrison and Kreps show that, under certain assumptions, their model is arbitrage-free if and only if there exists an equivalent probability measure under which the price process S is a martingale (see Harrison and Kreps [29], section 3). Delbaen and Schachermayer [18] have generalised this no-arbitrage result to semi-martingale models. For securities market models with transaction costs, a detailed study of consistent price systems can be found in Guasoni, Rásonyi and Schachermayer [25].

The work of Harrison and Kreps has been extended by Jouini and Kallal [30]. In their model, they assign to an asset a bid price process and an ask price process for which the asset can be sold or bought respectively. Then, the absence of arbitrage is equivalent to the existence of an equivalent probability measure and a process between the bid price process and the ask price process which is a martingale under the new probability measure (see Jouini and Kallal [30], Theorem 3.2).

Most financial markets operate in a different way than the above models. They function like certain types of auctions (see Smith et al. [44]). In this chapter we construct a mathematical model for a market which replicates the real trading mechanisms (see also [44]) and we extend the no-arbitrage results of Harrison and Kreps [29] and Jouini and Kallal [30] to this model of an order driven market. For simplicity, we model a financial market which comprises only one security. In our model traders can submit buy and sell orders for shares of this security. The orders that cannot be executed immediately are stored in an order book until execution or withdrawal. We allow two different types of orders: market orders and limit orders. Market orders are requests to buy or sell a specified number of shares immediately at the best available price. Limit orders are also requests to buy or sell shares but they include a limit price. The limit price is the worst price at which the trader is willing to buy or sell. In many cases, limit orders do not result in immediate transactions and they remain in the order book until they are executed or withdrawn. We assume that all traders can see the entire order book at any time. In our model, they know exactly which orders are outstanding. In reality, there is a time lag with which information is accessible and there are some further restrictions on the availability of information.

For the model, we assume that at most one order can be placed at a time and that orders can also be withdrawn. Thus, there is a natural order of the outstanding buy and sell limit orders in the order book according to their limit prices and arrival times. Orders which are placed by traders are matched against the outstanding orders in the order book according to this order. For instance, a buy order which is submitted by a trader is matched against the sell limit orders in the order book in order of their limit prices. The sell limit order with the lowest limit price is executed first. If there is more than one limit order in the order book with the same limit price, priority will be given according to the arrival times. Sell orders that are placed by traders are treated respectively. It is possible that limit orders can be executed in parts. If that happens, they will remain in the order book with a reduced size.

There exist financial markets with other matching rules, too. For example in some short-term interest rate futures markets, limit orders with the same limit price are filled proportionally to their size (see Field and Large [21] for an extensive study of these markets).

Limit prices at real financial markets are integer multiples of a tick, i.e. they are elements of the discrete set $\{1\Delta_p, 2\Delta_p, \dots\}$, where Δ_p denotes the tick size. In this chapter, we assume that the limit prices can take any positive real value. Furthermore, we assume that there are no transaction costs and that limit orders can be submitted and withdrawn without any fee.

The highest limit price of all the buy limit orders in the order book is called the best bid. Likewise, the lowest limit price of all the sell limit orders in the order book is called the best ask. The distance between the best ask and the best bid is called the spread. We call a buy limit order with limit price greater than (or equal to) the best ask a crossing limit order. The same name is used for a sell limit order with limit price less than (or equal) to the best bid. Crossing limit orders lead to immediate transactions. For example, Kühn and Stroh [32] model the best bid and best ask price processes as geometric Brownian motions and they assume that the bid-ask spread is proportional to the best bid price. With these assumptions, they find optimal trading strategies in their rather simplistic model of an order driven market.

In this chapter, we mainly focus on arbitrage in a general model. We define arbitrage as free lunch in a similar way as Harrison and Kreps [29] and Jouini and Kallal [30]. In section 2.1, we revisit Jouini and Kallal's model of securities markets with bid ask spread (see [30]) and we generalise their result in such a way that we can use it for our study of order driven markets. In section 2.2, we construct the mathematical model of the order driven market. We adapt the notion of no free lunch of Harrison and Kreps [29] and Jouini and Kallal [30] to our setting and we prove a no-arbitrage theorem.

The main difficulty is to deal with the discontinuities that arise from the use of both limit and market orders. Whenever a market order is submitted, a trader, who has submitted a matching limit order beforehand, buys or sells shares for less than the best ask price or more than the best bid price respectively. This opportunity exists only at the moment of submission of the market order. Therefore, price processes occur which are neither right-continuous nor left-continuous. Furthermore, the trading strategies need to handle asymmetric information. A trader who submits a market order buys/sells shares immediately and knows the share price before he submits his order. On the other hand, a trader who submits a limit order needs to wait until another trader submits a matching order that executes his limit order. At the time of submission he does not know when (or if at all) his order will be executed. In exchange, he can potentially achieve a better price.

2.1 Securities markets with bid-ask spread

Jouini and Kallal have studied securities markets with bid-ask spread in [30]. In this section we generalise one of their results, which is part of [30], Theorem 3.2, to a more general setting. In particular, we do not assume that the bid price process and the ask price process are right-continuous.

We consider a market which consists of one security and a bank account which we use as the numeraire. Assume that shares of the security can be bought for $Z_1(t)$ and sold for $Z'_1(t) \leq Z_1(t)$ at every time t. We assume that agents can trade a finite number of shares at each time t and that short selling is possible. For simplicity, we neglect interest. A unit of money in the bank account has the value $Z_2(t) = Z'_2(t) = 1$ at every time t. The results of this section can easily be extended to securities markets with more than one security.

Let $T > 0$ be a finite time horizon and let $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a filtered probability space which satisfies the usual conditions, i.e. the filtration is right-continuous and complete. We assume that Z_1 and Z'_1 are positive adapted processes which are progressively measurable and have second moments

$$E(Z_1(t)^2) < \infty \quad \text{and} \quad E(Z_1'(t)^2) < \infty$$

for all $t \in [0, T]$.

Let $Z = \binom{Z_1}{Z_2}$ and $Z' = \binom{Z_1'}{Z_2'}$. Let $\mathcal{G}$ denote the set of all pairs (W', W) of $\mathbb{R}^2$-valued stochastic processes $W' = \binom{W_1'}{W_2'}$ and $W = \binom{W_1}{W_2}$ that satisfy the same assumptions as Z' and Z. Note that in particular $W_2'(t) = W_2(t) = 1$ for all $t \in [0, T]$.

From now on, we call the model, which we have defined above, the securities market model (Z', Z). The processes Z' and Z indicate which bid price process and which ask price process underlie the model.

We define simple trading strategies in a slightly more general way than Jouini and Kallal [30], Definition 3.1. Let $\theta_1(t)$ denote the total number of shares which we have bought before time t and let $\theta_1'(t)$ denote the total number of shares which we have sold before time t respectively. Define $\theta_2(t)$ and $\theta_2'(t)$ analogously for the amount of money in the bank account.

Definition 2.1 *A simple trading strategy is a pair (θ, θ') of $\mathbb{R}^2$-valued processes $\theta = \binom{\theta_1}{\theta_2}$ and $\theta' = \binom{\theta_1'}{\theta_2'}$ such that*

(i) (θ, θ') is adapted to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$,

(ii) θ_k and θ_k' are non-negative and non-decreasing processes for every $k \in \{1, 2\}$,

(iii) the second moments

$$E((\theta_k(t)\, Z_k(t))^2) < \infty, \quad E((\theta_k'(t)\, Z_k'(t))^2) < \infty,$$

$$E((\theta_k(t)\, Z_k'(t))^2) < \infty \quad \textit{and} \quad E((\theta_k'(t)\, Z_k(t))^2) < \infty$$

exist for every $t \in [0, T]$ and each $k \in \{1, 2\}$,

(iv) *there exists an integer* N *and stopping times* $0 < \tau_1 < \cdots < \tau_N < T$ *such that* $(\theta(\omega), \theta'(\omega))$ *is constant in the intervals* $[0, \tau_1(\omega)]$, $(\tau_N(\omega), T]$ *and* $(\tau_n(\omega), \tau_{n+1}(\omega)]$ *for all* $n = 1, \ldots, N-1$ *for every* $\omega \in \Omega$.

This definition differs from Jouini and Kallal [30], Definition 3.1, in the fact that we allow the times $\tau_1, \ldots, \tau_N$ to be stopping times. Jouini and Kallal assume that they are prespecified, fixed dates. While pursuing a simple trading strategy, we buy and sell shares only at the times $\tau_1, \ldots, \tau_N$. Note that we hold $\theta_1(\tau_{k+1}) - \theta_1'(\tau_{k+1})$ shares at time τ_k and $\theta_1(T) - \theta_1'(T)$ shares at time T.

Now, we adapt the definition of self-financing trading strategies (see [30], Definition 3.2) to our setting.

Definition 2.2 *Let* $\tau_{N+1} = T$. *A simple trading strategy* (θ, θ') *is self-financing if*

$$(\theta(\tau_{n+1}(\omega), \omega) - \theta(\tau_n(\omega), \omega)) \cdot Z(\tau_n(\omega), \omega) \leq (\theta'(\tau_{n+1}(\omega), \omega) - \theta'(\tau_n(\omega), \omega)) \cdot Z'(\tau_n(\omega), \omega)$$

holds for every $\omega \in \Omega$ *and for every* $n \in \{1, \ldots, N\}$.

This means that we pay for the shares we buy with money from the bank account and that we put the proceeds from the sale of shares back into it. Also, we allow that money can be taken out of the bank account for consumption.

Not every contingent claim is necessarily attainable by simple self-financing trading strategies. Let $\mathcal{R}$ denote the set of square-integrable, $\mathcal{F}_T$-measurable random variables. We define the set of claims which can be hedged or dominated by simple self-financing trading strategies in the following way (see also [30], Definition 3.3).

Definition 2.3 *A contingent claim* $X \in \mathcal{R}$ *is called marketed if there exists a simple self-financing trading strategy* (θ, θ') *such that the hedging portfolio has the value*

$$(\theta(T, \omega) - \theta'(T, \omega))^+ \cdot Z'(T, \omega) - (\theta(T, \omega) - \theta'(T, \omega))^- \cdot Z(T, \omega) \geq X(\omega) \qquad (2.1)$$

at time T *for every* $\omega \in \Omega$.

Let $\mathcal{M} \subset \mathcal{R}$ denote the set of marketed contingent claims and let $\mathcal{R}^+$ denote the set of random variables $Y \in \mathcal{R}$ which satisfy $P(Y \geq 0) = 1$ and $P(Y > 0) > 0$. Now, we can define the notion of free lunch (see also [30], Definition 3.4).

Definition 2.4 *A multiperiod free lunch in the securities market model (Z', Z) is a sequence of marketed contingent claims $X_n \in \mathcal{M}$, $n \in \mathbb{N}$, and a contingent claim $X \in \mathcal{R}^+$ such that*

(i) X_n converges to X P-almost surely,

(ii) there exists an integrable, non-positive minorant g such that $X_n(\omega) \geq g(\omega)$ for all $\omega \in \Omega$ and all $n \in \mathbb{N}$,

(iii) there exists a sequence of self-financing simple trading strategies $(\theta^{(n)}, \theta'^{(n)})$, $n \in \mathbb{N}$, such that the values of the hedging portfolios satisfy

$$(\theta^{(n)}(T,\omega)-\theta'^{(n)}(T,\omega))^+ \cdot Z'(T,\omega)-(\theta^{(n)}(T,\omega)-\theta'^{(n)}(T,\omega))^- \cdot Z(T,\omega) \geq X_n(\omega) \tag{2.2}$$

at time T for every $\omega \in \Omega$ and for every $n \in \mathbb{N}$ and the costs of setting up these portfolios satisfy

$$\limsup_{n\to\infty} \left(\theta^{(n)}(0,\omega) \cdot Z(0,\omega) - \theta'^{(n)}(0,\omega) \cdot Z'(0,\omega)\right) \leq 0 \tag{2.3}$$

at time 0 for every $\omega \in \Omega$.

This means that it is possible to get arbitrarily close to a risk-free profit by using self-financing simple trading strategies.

Note that condition (ii), i.e. the existence of the integrable minorant, says that the trading strategy needs to be admissible. For example, traders may have a finite credit line so that the value of the portfolio must always satisfy

$$(\theta^{(n)}(t,\omega) - \theta'^{(n)}(t,\omega))^+ \cdot Z'(t,\omega) - (\theta^{(n)}(t,\omega) - \theta'^{(n)}(t,\omega))^- \cdot Z(t,\omega) \geq -c$$

for every $t \in [0,T]$ and every $\omega \in \Omega$ for some constant $c \geq 0$.

The following theorem extends the result of Jouini and Kallal [30], Theorem 3.2. We consider a wider class of trading strategies and in part (i), we do not assume that the bid price process and the ask price process are right-continuous.

Theorem 2.5 *Consider bid and ask price processes $(Z', Z) \in \mathcal{G}$.*

(i) The securities market model (Z', Z) admits no multiperiod free lunch if there exists a probability measure Q equivalent to P with $E_P\left(\left(\frac{dQ}{dP}\right)^2\right) < \infty$ and an adapted process $(Z_1^*(t))_{t\in[0,T]}$ satisfying $Z_1'(t,\omega) \leq Z_1^*(t,\omega) \leq Z_1(t,\omega)$ for all $\omega \in \Omega$ and all $t \in [0,T]$ such that Z^* is a martingale under Q.

(ii) If the processes Z' and Z are right-continuous, then the converse of (i) does also hold.

Bion-Nadal proves a similar result in the context of dynamic pricing procedures (see [2], Theorem 2.1).

Note that part (ii) of Theorem 2.5 is Jouini and Kallal's result (see [30], Theorem 3.2). For their proof, the right-continuity of Z' and Z is required for the construction of the martingale (Z_1^*) as the limit of discretisations.

We need the following lemmas for the proof of Theorem 2.5.

Lemma 2.6 *Let $(Z', Z) \in \mathcal{G}$ and $(W', W) \in \mathcal{G}$ be stochastic processes which satisfy*

$$Z_1'(t,\omega) \leq W_1'(t,\omega) \leq W_1(t,\omega) \leq Z_1(t,\omega) \tag{2.4}$$

for every $\omega \in \Omega$ and every $t \in [0,T]$. If the securities market model (Z', Z) admits a multiperiod free lunch, then the securities market model (W', W) also admits a multiperiod free lunch.

Proof: It follows from (2.4) that the same trading strategies generate bigger profits in the model (W', W) than in the model (Z', Z). This implies that a contingent claim which is marketed in the model (Z', Z) is also marketed in the model (W', W). Thus, contingent claims and trading strategies which lead to a multiperiod free lunch in the model (Z, Z') form a multiperiod free lunch in the model (W, W') as well. $\square$

Lemma 2.7 *Let $(W', W) \in \mathcal{G}$ be stochastic processes such that $W' = W$. Then, the securities market model (W', W) admits no multiperiod free lunch if there exists a probability measure Q equivalent to P with $E_P\left(\left(\frac{dQ}{dP}\right)^2\right) < \infty$ such that W_1 is a martingale under Q with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$.*

Proof: By contradiction. Assume that there exists a probability measure Q equivalent to P such that W_1 is a martingale under Q and that there is a multiperiod free lunch in the securities market model $(W', W) = \left(\binom{W_1}{1}, \binom{W_1}{1}\right)$. Then, there exists a sequence of contingent claims $X_k \in \mathcal{M}$, $k \in \mathbb{N}$, and $X \in \mathcal{R}^+$ such that X_k converges to X P-almost surely. Furthermore, there exists a sequence of self-financing simple trading strategies $(\theta^{(k)}, \theta'^{(k)})$ such that (2.2) and (2.3) are satisfied.

By Definition 2.1, there are stopping times $0 < \tau_1^{(k)} < \cdots < \tau_{N_k}^{(k)} < T$ for every $k \in \mathbb{N}$ at which shares are bought or sold in course of the trading strategy $(\theta^{(k)}, \theta'^{(k)})$. In between, no trading takes places.

For $k \in \mathbb{N}$ and $n \in \{1, \dots, N_k\}$, define

$$b_n^{(k)} = \theta_1^{(k)}(\tau_{n+1}^{(k)}) - \theta_1^{(k)}(\tau_n^{(k)}), \quad s_n^{(k)} = \theta_1'^{(k)}(\tau_{n+1}^{(k)}) - \theta_1'^{(k)}(\tau_n^{(k)}),$$

$$p_n^{(k)} = \theta_2^{(k)}(\tau_{n+1}^{(k)}) - \theta_2^{(k)}(\tau_n^{(k)}) \quad \text{and} \quad w_n^{(k)} = \theta_2'^{(k)}(\tau_{n+1}^{(k)}) - \theta_2'^{(k)}(\tau_n^{(k)}),$$

where $\tau_{N_k+1} = T$.

While pursuing the simple trading strategy $(\theta^{(k)}, \theta'^{(k)})$, $b_n^{(k)} \geq 0$ shares are bought and $s_n^{(k)} \geq 0$ shares are sold at time $\tau_n^{(k)}$. Furthermore, $p_n^{(k)} \geq 0$ units of money are put into the bank account and $w_n^{(k)} \geq 0$ units are withdrawn at time $\tau_n^{(k)}$.

By Definition 2.1, ($\theta^{(k)}$ and $\theta'^{(k)}$) are left-continuous and piecewise constant. Hence, $b_n^{(k)}$, $s_n^{(k)}$, $p_n^{(k)}$ and $w_n^{(k)}$ are $\mathcal{F}_{\tau_n^{(k)}+}$-measurable. Since the filtration is right-continuous, they are also $\mathcal{F}_{\tau_n^{(k)}}$-measurable.

Since the simple trading strategy $(\theta^{(k)}, \theta'^{(k)})$ is self-financing, it follows from Definition 2.2 that

$$\begin{pmatrix} b_n^{(k)}(\omega) - s_n^{(k)}(\omega) \\ p_n^{(k)}(\omega) - w_n^{(k)}(\omega) \end{pmatrix} \cdot \begin{pmatrix} W_1(\tau_n^{(k)}(\omega), \omega) \\ 1 \end{pmatrix} \leq 0$$

holds for every $\omega \in \Omega$ and every $n \in \{1, \ldots, N_k\}$ as $W_1 = W_1'$.

At time T, the hedging portfolio for X_k consists of $\theta_1^{(k)}(0) - \theta_1'^{(k)}(0) + \sum_{n=1}^{N_k}(b_n^{(k)} - s_n^{(k)})$ shares and $\theta_2^{(k)}(0) - \theta_2'^{(k)}(0) + \sum_{n=1}^{N_k}(p_n^{(k)} - w_n^{(k)})$ units of money in the bank account. For each $k \in \mathbb{N}$, let $V_t^{(k)}$ denote the value of the hedging portfolio for X_k at time $t \in [0, T]$. The initial costs $\theta^{(k)}(0) \cdot W(0) - \theta'^{(k)}(0) \cdot W'(0)$ of setting up the portfolio satisfy (2.3) and we have

$$V_0^{(k)} \leq \theta^{(k)}(0) \cdot W(0) - \theta'^{(k)}(0) \cdot W'(0). \tag{2.5}$$

The value of the hedging portfolio for X_k at time T is

$$V_T^{(k)} = V_0^{(k)} + \sum_{n=1}^{N_k} \begin{pmatrix} b_n^{(k)} - s_n^{(k)} \\ p_n^{(k)} - w_n^{(k)} \end{pmatrix} \cdot \begin{pmatrix} W_1(T) \\ 1 \end{pmatrix}$$

where $\binom{b_n^{(k)} - s_n^{(k)}}{p_n^{(k)} - w_n^{(k)}}$ is $\mathcal{F}_{\tau_n^{(k)}}$-measurable. Recall that W_1 and $b_n^{(k)}$ are non-negative by definition. The expectation of $b_n^{(k)} W_1(T)$ under Q exists for every $n \in \{1, \ldots, N_k\}$ because of Definition 2.1 and the Cauchy-Schwarz inequality. It follows in the same way that the expectations of $s_n^{(k)} W_1(T)$, $p_n^{(k)}$, $w_n^{(k)}$ and $V_0^{(k)}$ exist under Q. Therefore, the expectation of $V_T^{(k)}$ exists under Q for every $k \in \mathbb{N}$.

For every $k \in \mathbb{N}$, define

$$A_k = \left\{ \omega \in \Omega \,\middle|\, \sup_{j \geq k} V_0^{(j)} \leq 1 \right\}.$$

The set $A_k \subset \Omega$ is an element of $\mathcal{F}_0$. Hence, the indicator function $\mathbf{1}_{A_k}$ is $\mathcal{F}_{\tau_n}$-measurable for every $n \in \{1, \ldots, N_k\}$. The expectation of $V_T^{(k)} \mathbf{1}_{A_k}$ under Q does also exist and it satisfies

$$
\begin{aligned}
E_Q(V_T^{(k)}\mathbf{1}_{A_k}) &= E_Q(V_0^{(k)}\mathbf{1}_{A_k}) \\
&\quad + \sum_{n=1}^{N_k} E_Q\left(E_Q\left(\mathbf{1}_{A_k}\begin{pmatrix} b_n^{(k)} - s_n^{(k)} \\ p_n^{(k)} - w_n^{(k)} \end{pmatrix} \cdot \begin{pmatrix} W_1(T) \\ 1 \end{pmatrix} \middle| \mathcal{F}_{\tau_n^{(k)}}\right)\right) \\
&= E_Q(V_0^{(k)}\mathbf{1}_{A_k}) + \sum_{n=1}^{N_k} E_Q\left(\mathbf{1}_{A_k}\begin{pmatrix} b_n^{(k)} - s_n^{(k)} \\ p_n^{(k)} - w_n^{(k)} \end{pmatrix} \cdot \begin{pmatrix} W_1(\tau_n^{(k)}) \\ 1 \end{pmatrix}\right) \\
&\leq E_Q(V_0^{(k)}\mathbf{1}_{A_k})
\end{aligned}
$$

as W_1 is a martingale under Q. By inequality (2.2), we have $X_k(\omega) \leq V_T^{(k)}(\omega)$ for every $\omega \in \Omega$ and every integer $k \in \mathbb{N}$. Hence, the expectations under Q satisfy $E_Q(X_k\mathbf{1}_{A_k}) \leq E_Q(V_T^{(k)}\mathbf{1}_{A_k}) \leq E_Q(V_0^{(k)}\mathbf{1}_{A_k})$ for every $k \in \mathbb{N}$.

The random variables X_k converge to X P-almost surely and $\mathbf{1}_{A_k}$ converges to 1 P-almost surely. Therefore, $X_k\mathbf{1}_{A_k}$ converges to X P-almost surely. Since P and Q are equivalent, $X_k\mathbf{1}_{A_k}$ also converges to X Q-almost surely.

Recall that, by Definition 2.4, there exists an integrable, non-positive minorant g. Since it is non-positive, it satisfies $g(\omega) \leq X_k(\omega)\mathbf{1}_{A_k}(\omega)$ for all $\omega \in \Omega$ and all $k \in \mathbb{N}$. Now, Fatou's lemma shows that

$$
\begin{aligned}
E_Q(X) &= E_Q(\lim_{k\to\infty} X_k\mathbf{1}_{A_k}) = E_Q(\liminf_{k\to\infty} X_k\mathbf{1}_{A_k}) \leq \liminf_{k\to\infty} E_Q(X_k\mathbf{1}_{A_k}) \\
&\leq \limsup_{k\to\infty} E_Q(X_k\mathbf{1}_{A_k}) \leq \limsup_{k\to\infty} E_Q(V_0^{(k)}\mathbf{1}_{A_k}) \qquad (2.6) \\
&\leq E_Q(\limsup_{k\to\infty} V_0^{(k)}\mathbf{1}_{A_k}) = E_Q(\limsup_{k\to\infty} V_0^{(k)}) \leq 0
\end{aligned}
$$

as $\limsup_{k\to\infty} V_0^{(k)}(\omega) \leq 0$ holds for every $\omega \in \Omega$ by (2.3) and (2.5).

On the other hand, we have $X \in \mathcal{R}^+$. We get $Q(X \geq 0) = 1$ and $Q(X > 0) > 0$, since P and Q are equivalent. This implies $E_Q(X) > 0$ which contradicts (2.6).

Thus, Lemma 2.7 holds. □

Proof of Theorem 2.5: Assume that there exist Q and Z^* as described in Theorem 2.5 (i). By Lemma 2.7, the securities market model $\left(\binom{Z_1^*}{1}, \binom{Z_1^*}{1}\right)$ admits no multiperiod free lunch. Since $Z_1'(t,\omega) \leq Z_1^*(t,\omega) \leq Z_1(t,\omega)$ holds for every $\omega \in \Omega$ and every $t \in [0,T]$, it follows from Lemma 2.6 that the securities market model (Z', Z) admits no multiperiod free lunch either. This proves part (i).

Part (ii) follows from [30], Theorem 3.2, since no multiperiod free lunch in our model implies no multiperiod free lunch in their setting. □

Remark 2.8 *Theorem 2.5 can be extended to a securities market with $m > 1$ securities and the same proof works for such a market with more than one security.*

2.2 Free lunch in order driven markets

In this section, we prove a no-arbitrage result for a model of an order driven market for a small trader. First, we define the mathematical model.

Consider a market with a single security as described in the introduction to this chapter. We assume that the market comprises a group of noise traders and a small trader. At every time t, the traders can submit buy/sell limit and market orders which are stored in the order book and matched according to the matching rule explained at the beginning of this chapter. Outstanding limit orders can be withdrawn from the order book at every time t.

We assume liquidity. By this, we mean that there are outstanding buy limit orders and sell limit orders in the order book at every time t which have been submitted by noise traders. Let $\hat{b}_t$ be the best bid price of the noise traders and let $\hat{a}_t$ be the best ask price of the noise traders at time t. We assume that the matching of orders is carried out immediately whenever an order is submitted and that $\hat{a}_t$ and $\hat{b}_t$ are the best ask and the best bid prices directly after the orders have been executed. Thus, we have $\hat{a}_t > \hat{b}_t > 0$ for every time t. Note that $\hat{a}$ and $\hat{b}$ stem from the limit orders of the noise traders only. They do not necessarily represent the best bid b and best ask a in the market since they do not take the outstanding limit orders into account which have been submitted by the small trader.

Assume that the sizes of the orders of the small trader are considerably smaller than the sizes of the orders of the noise traders. Therefore, we can assume that an order of the small trader never completely fills an order of one of the noise traders. So the small trader cannot alter $\hat{a}$ and $\hat{b}$.

For the same reason, we assume that every limit order of the small trader is executed either in full or not at all by the orders of the noise traders. In our model, we do not allow that the orders of the small trader can be partially executed by orders of the noise traders. Furthermore, we assume that the limit orders of the small trader in the order book get priority over the limit orders of the noise traders with the same limit price no matter when they have been submitted. This means that if there are several limit orders in the order book with the same limit price, the limit orders of the small trader will be executed first. A similar assumption is implicitly made by Kühn and Stroh [32] who assume that the limit orders of the small trader are always executed when a market order (of the opposite kind) arrives.

For every time t, define

$$\alpha_t = \begin{cases} \hat{b}_t & \begin{array}{l}\text{if a noise trader submits a sell order at time } t \text{ which immediately} \\ \text{executes an outstanding buy limit order of another noise trader}\end{array} \\ \\ \hat{a}_t & \text{otherwise} \end{cases}$$

and

$$\beta_t = \begin{cases} \hat{a}_t & \begin{array}{l}\text{if a noise trader submits a buy order at time } t \text{ which immediately} \\ \text{executes an outstanding sell limit order of another noise trader}\end{array} \\ \\ \hat{b}_t & \text{otherwise.} \end{cases}$$

By definition, we have

$$\hat{a}_t \geq \alpha_t \geq \beta_t \geq \hat{b}_t > 0 \tag{2.7}$$

for every time t. If the small trader buys shares at time t, the best price he can possibly achieve is α_t. He can always buy shares for $\hat{a}_t$ and if he has a buy limit order in the order book which gets executed by a market sell order or a crossing sell limit order of a noise trader, he will buy for $\hat{b}_t$ in the best case. In the same way, β_t is the best price the small trader can possibly obtain if he sells shares at time t.

We have assumed that the sizes of the orders of the small trader are considerably smaller than the sizes of the orders of the noise traders. Moreover, we assume that the behaviour of the small trader has no influence on the noise traders. They do not change the orders, which they submit, the limit prices or the order sizes in response to any action taken by the small trader. Therefore, we can assume that $\hat{a}$, $\hat{b}$, α and β are independent of the behaviour of the small trader. They are solely determined by the behaviour of the noise traders which we regard as random in our model. Thus, $\hat{a}$, $\hat{b}$, α and β can be modelled as exogenously given stochastic processes. They contain all the information about the order book we need and they fully characterise our model.

Let $T > 0$ be a finite time horizon and let $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a filtered probability space which satisfies the usual conditions, i.e. the filtration is right-continuous and complete. We assume that the noise traders submit their orders in such a way that the following assumptions are satisfied:

(A1) The processes $\hat{a}$, $\hat{b}$, α and β are adapted to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$.

(A2) The processes $\hat{a}$ and $\hat{b}$ are right-continuous with left limits.

(A3) There exist finitely many stopping times $\tilde{\tau}_1, \ldots, \tilde{\tau}_{\tilde{N}}$ and $\hat{\tau}_1, \ldots, \hat{\tau}_{\hat{N}}$ such that the processes α and β satisfy

$$\alpha_t = \begin{cases} \hat{b}_t & \text{if } t = \tilde{\tau}_i \text{ for some } i \in \{1, \dots, \tilde{N}\} \\ \hat{a}_t & \text{otherwise} \end{cases}$$

and

$$\beta_t = \begin{cases} \hat{a}_t & \text{if } t = \hat{\tau}_j \text{ for some } j \in \{1, \dots, \hat{N}\} \\ \hat{b}_t & \text{otherwise.} \end{cases}$$

We assume that the stopping times $\tilde{\tau}_1, \dots, \tilde{\tau}_{\tilde{N}}$ and $\hat{\tau}_1, \dots, \hat{\tau}_{\hat{N}}$ differ so that $\tilde{\tau}_i(\omega) \neq \tilde{\tau}_j(\omega)$ for all $i \neq j$, $\hat{\tau}_i(\omega) \neq \hat{\tau}_j(\omega)$ for all $i \neq j$ and $\tilde{\tau}_i(\omega) \neq \hat{\tau}_j(\omega)$ for all i, j and all $\omega \in \Omega$.

(A4) The second moment $E(\hat{a}_t^2) < \infty$ exists for every $t \in [0, T]$.

The stopping times $\tilde{\tau}_1, \dots, \tilde{\tau}_{\tilde{N}}$ and $\hat{\tau}_1, \dots, \hat{\tau}_{\hat{N}}$ are the times when the noise traders submit sell and buy orders respectively which cross the spread and execute at least one of the outstanding limit orders of a noise trader in the order book.

Note that $\hat{a}$, $\hat{b}$, α and β are progressively measurable and the processes α and β are in general not right-continuous. It follows from assumption (A4) and (2.7) that the second moments $E(\alpha_t^2) < \infty$, $E(\beta_t^2) < \infty$ and $E(\hat{b}_t^2) < \infty$ also exist for every $t \in [0, T]$.

For example, Smith et al. [44] model the order flows as Poisson processes which satisfies our assumptions.

Furthermore, we assume that the small trader has access to a bank account where he can invest and borrow money. For simplicity, we neglect interest. We use this bank account as the numeraire. The small trader can also place market and limit orders to buy or sell shares and withdraw orders from the order book. We assume that short selling is possible.

Consider $t \in [0, T]$. The outstanding limit orders of the small trader in the order book just before all activities at time t can be described by the signed measure

$$O_t = \sum_{i=1}^{n_t} \lambda_i(t)\delta_{p_i}, \tag{2.8}$$

where $n_t \in \mathbb{N}$, $p_i \in \mathbb{R}^{>0}$, $\lambda_i(t) \in \mathbb{R} \setminus \{0\}$ and δ_{p_i} is the Dirac measure at p_i. Each summand $\lambda_i(t)\delta_{p_i}$ describes an outstanding limit order with limit price p_i and order size $|\lambda_i(t)|$. If $\lambda_i(t) > 0$, the order is a buy order and if $\lambda_i(t) < 0$, the order is a sell order.

Let $(L_t)_{t\in[0,T]}$ and $(M_t)_{t\in[0,T]}$ describe the limit and market orders that the small trader submits. If he submits a buy limit order at time t with limit price p and order size λ, set $L_t = \lambda\delta_p$. If he submits a sell limit order with limit price p and order size λ, set $L_t = -\lambda\delta_p$. If no limit orders are submitted, set $L_t = 0$. In the same way, set $M_t = \lambda\delta_{\hat{a}_t}$ if the small trader submits a buy market order with order size λ at time t. And set $M_t = -\lambda\delta_{\hat{b}_t}$ if the small trader submits a sell market order with order size λ. If no market orders are submitted, set $M_t = 0$.

Furthermore, let $(W_t)_{t\in[0,T]}$ describe the orders that the small trader withdraws. Note that it is possible to withdraw orders in parts. Set $W_t = \sum_{i=1}^{n_t} w_i(t)\lambda_i(t)\delta_{p_i}$ with $w_i(t) \in [0,1]$ if he withdraws $w_i(t)\lambda_i(t)$ lots of the limit order with limit price p_i at time t for $i = 1, \ldots, n_t$. This order has been in the order book with the order size $\lambda_i(t)$ beforehand. Set $W_t = 0$ if no orders are withdrawn at time t.

Let C_0 denote the initial amount of money that the small trader has in the bank account at time 0 before he starts trading. In the same way, let S_0 be the initial number of shares that the small trader owns at time 0 before he starts trading. Assume that C_0 and S_0 are real-valued random variables which are $\mathcal{F}_0$-measurable. Furthermore, assume that the second moments

$$E(S_0^2) < \infty, \quad E(C_0^2) < \infty \quad \text{and} \quad E((S_0\hat{a}_t)^2) < \infty$$

exist for all $t \in [0, t]$. We define a simple trading strategy for the small trader in the following way.

Definition 2.9 *A simple trading strategy for the small trader is a triplet of measure valued stochastic processes $Y = (L, M, W)$ such that:*

(i) $L_t = \lambda^L(t)\delta_{p^L(t)}$ *and* $M_t = \lambda^M(t)\delta_{p^M(t)}$ *where* $\lambda^L(t)$, $p^L(t)$, $\lambda^M(t)$ *and* $p^M(t)$ *are real-valued and* $\mathcal{F}_t$*-measurable for every* $t \in [0, T]$.

(ii) $W_t = \sum_{i=1}^{n_t} w_i(t)\lambda_i(t)\delta_{p_i}$ *where the outstanding limit orders at time* t *are given by (2.8) and* $w_i(t) \in [0, 1]$ *is* $\mathcal{F}_t$*-measurable for* $i = 1, \ldots, n_t$.

(iii) For every $t \in [0, T]$, *at most one of* L_t, M_t *and* W_t *differs from zero.*

(iv) The second moments

$$E((|\lambda^M(t)|\hat{a}_s)^2) < \infty \quad \textit{and} \quad E((|\lambda^L(t)|\hat{a}_s)^2) < \infty$$

exist for all $s, t \in [0, T]$ *with* $s \geq t$.

(v) There exists an integer N *and stopping times* $0 < \tau_1 < \cdots < \tau_N < T$ *such that* Y *differs from* $(0, 0, 0)$ *only if* $t = \tau_k$ *for some* $k \in \{1, \ldots, N\}$.

Consider a simple trading strategy $Y = (L, M, W)$. Set $\tau_{N+1} = T$. For integers $k \in \{1, \ldots, N\}$ and real numbers $t \in (\tau_k, \tau_{k+1}]$, define $E_k^B(t) \subset \mathbb{R}$ and $E_k^S(t) \subset \mathbb{R}$ by

$$E_k^B(t) = \begin{cases} [\min_{s \in [\tau_k, t)} \alpha_s, \infty) & \text{if the minimum exists} \\ (\inf_{s \in [\tau_k, t)} \alpha_s, \infty) & \text{otherwise} \end{cases}$$

and

$$E_k^S(t) = \begin{cases} (-\infty, \max_{s \in [\tau_k, t)} \beta_s] & \text{if the maximum exists} \\ (-\infty, \sup_{s \in [\tau_k, t)} \beta_s) & \text{otherwise.} \end{cases}$$

We assume that the order sizes of the small trader are considerably smaller than the order sizes of the noise traders and that the orders of the small trader get priority over the orders of the noise traders. Thus, between τ_k and t, all the buy limit orders of the small trader with limit prices $p \in E_k^B(t)$ and all his sell limit orders with limit prices $p \in E_k^S(t)$ are executed by orders of the noise traders. Define

$$D_k^B(t) = \{i \in \{1, \dots, n_{\tau_k}\} : p_i \in E_k^B(t) \text{ and } \lambda_i(\tau_k) > 0\}$$

and

$$D_k^S(t) = \{i \in \{1, \dots, n_{\tau_k}\} : p_i \in E_k^S(t) \text{ and } \lambda_i(\tau_k) < 0\}.$$

The sets $D_k^B(t)$ and $D_k^S(t)$ are the sets of indices which characterise the buy limit orders of the small trader with limit prices $p \in E_k^B(t)$ and his sell limit orders with limit prices $p \in E_k^S(t)$ respectively. As mentioned above, these limit orders are executed by orders of the noise traders between τ_k and t.

The outstanding orders O_t of the small trader immediately before t are given recursively by

$$O_t = 0 \quad \text{for } t \in [0, \tau_1]$$

and

$$\begin{aligned} O_t &= O_{\tau_k} + L_{\tau_k} - W_{\tau_k} - \lambda^L(\tau_k)^+ \mathbf{1}_{\{p^L(\tau_k) \in E_k^B(t)\}} \delta_{p^L(\tau_k)} \\ &\quad + \lambda^L(\tau_k)^- \mathbf{1}_{\{p^L(\tau_k) \in E_k^S(t)\}} \delta_{p^L(\tau_k)} - \sum_{i \in D_k^B(t) \cup D_k^S(t)} (1 - w_i(\tau_k)) \lambda_i(\tau_k) \delta_{p_i} \end{aligned}$$

for $t \in (\tau_k, \tau_{k+1}]$. For $i = 1, \dots, N$, let

$$\tilde{\rho}_i = \tau_i \mathbf{1}_{\{\lambda^M(\tau_i) \neq 0\}} + \mathbf{1}_{\{\lambda^L(\tau_i) \neq 0\}} \inf\{t \in (\tau_i, T] : O_t(\{p^L(\tau_i)\}) = 0 \text{ and } W_t(\{p^L(\tau_i)\}) = 0\}$$

denote the time at which the order, that the small trader has submitted at time τ_i, is executed. Market orders execute immediately and limit orders execute when a noise trader submits a matching order of the opposite kind. Note that $\tilde{\rho}_i = 0$ if no order has been submitted at time τ_i and that $\tilde{\rho}_i = \infty$ if a limit order has been submitted at τ_i and it has been withdrawn or has not been executed by T. The times $\tilde{\rho}_i$ are not necessarily different. Let $M \in \mathbb{N}$ denote the number of times $\tilde{\rho}_i$ which are pairwise different and which also differ from 0 and ∞. The times at which the small trader actually buys or sells shares are

$$\begin{aligned}
\rho_1 &= \inf\{t > 0 : \exists i \in \{1, \dots, N\} \text{ such that } t = \tilde{\rho}_i\} \\
\rho_2 &= \inf\{t > \rho_1 : \exists i \in \{1, \dots, N\} \text{ such that } t = \tilde{\rho}_i\} \\
&\vdots \\
\rho_M &= \inf\{t > \rho_{M-1} : \exists i \in \{1, \dots, N\} \text{ such that } t = \tilde{\rho}_i\}.
\end{aligned}$$

Note that the times $\rho_1, \rho_2, \dots, \rho_M$ are stopping times. Set $\rho_0 = 0$ and $\rho_{M+1} = T$. For $i = 1, \dots, M$, the transaction at time ρ_i is described by the (signed) measure

$$\begin{aligned}
T_i &= O_{\rho_{i+1}} - O_{\rho_i} - \sum_{j \in \{1, \dots N\} : \tau_j \in [\rho_i, \rho_{i+1})} (L_{\tau_j} + W_{\tau_j}) \\
&= \sum_{k=1}^{n_{\rho_{i+1}}} \lambda_k(\rho_{i+1}) \delta_{p_k} - \sum_{k=1}^{n_{\rho_i}} \lambda_k(\rho_i) \delta_{p_k} \\
&\quad - \sum_{j \in \{1, \dots N\} : \tau_j \in [\rho_i, \rho_{i+1})} \left(\lambda^L(\tau_j) \delta_{p^L(\tau_j)} + \sum_{l=1}^{n_{\tau_j}} w_l(\tau_j) \lambda_l(\tau_j) \delta_{p_l} \right).
\end{aligned}$$

The (signed) measure T_i contains all the information about the numbers of shares that are bought and sold at time ρ_i and about the respective prices. At time ρ_i the small trader buys/sells $T_i(\mathbb{R})$ shares and pays/receives

$$\begin{aligned} P_i &= \int_{\mathbb{R}} x\, T_i(dx) \\ &= \sum_{k=1}^{n_{\rho_{i+1}}} \lambda_k(\rho_{i+1}) p_k - \sum_{k=1}^{n_{\rho_i}} \lambda_k(\rho_i) p_k \\ &\quad - \sum_{j\in\{1,\dots N\}:\tau_j\in[\rho_i,\rho_{i+1})} \left(\lambda^L(\tau_j) p^L(\tau_j) + \sum_{l=1}^{n_{\tau_j}} w_l(\tau_j)\lambda_l(\tau_j) p_l \right) \end{aligned}$$

units of money in exchange. If $T_i(\mathbb{R})$ is positive, shares are bought and if $T_i(\mathbb{R})$ is negative, shares are sold.

Let S_t be the number of shares that the small trader owns at time t after his orders have been processed. We have

$$S_t = S_0 + \sum_{i\in\{1,\dots,M\}:\rho_i\le t} T_i(\mathbb{R}).$$

In the same way, let C_t denote the amount of money which the small trader has in the bank account at time t after his orders have been processed. Then,

$$C_t = C_0 + \sum_{i\in\{1,\dots,M\}:\rho_i\le t} \int_{\mathbb{R}} x\, T_i(dx)$$

holds.

Definition 2.10 *We call the simple trading strategy Y self-financing if*

$$C_{\rho_i(\omega)}(\omega) \le C_{\rho_{i-1}(\omega)}(\omega) + P_i(\omega) \tag{2.9}$$

holds for every $i \in \{1,\dots,M\}$ and every $\omega \in \Omega$.

Note that we have equality in (2.9) if the small trader pays for the purchase of shares solely with money from the bank account, puts all the proceeds from the sale of shares back into it and does not take any money out of it for consumption.

We define the value V_t of the portfolio of the small trader at time $t \in [0, T]$ as the value in case of immediate liquidation. The portfolio consists of S_t shares and C_t units of money in the bank account. If $S_t > 0$, the value of the shares is the amount of money which the small trader would obtain if he sold all his shares immediately at the best available price, i.e. if he placed a sell market order for all S_t shares. We can assume that it is always possible to sell all S_t shares immediately since we have assumed liquidity in the market and we consider a small trader. If $S_t < 0$, the value of the shares is the cost of closing out the short position immediately at time t. For this purpose, the small trader would have to put in a buy market order for S_t shares. In total, the value V_t of the portfolio is

$$\begin{aligned} V_t &= S_t^+ \hat{b}_t - S_t^- \hat{a}_t + C_t \\ &\leq S_t \beta_t + C_t \end{aligned}$$

for every $t \in [0, T]$. As in the previous section, let $\mathcal{R}$ denote the set of square-integrable, $\mathcal{F}_T$-measurable random variables. Not every contingent claim $X \in \mathcal{R}$ can necessarily be attained by simple self-financing trading strategies.

Definition 2.11 *We call a contingent claim $X \in \mathcal{R}$ marketed if there exists a simple self-financing trading strategy such that the value of the hedging portfolio at time T satisfies $V_T(\omega) \geq X(\omega)$ for every $\omega \in \Omega$.*

We denote the set of marketed contingent claims in the model of the order driven market by $\hat{\mathcal{M}} \subset \mathcal{R}$. As in the previous section, let $\mathcal{R}^+$ denote the set of random variables $Z \in \mathcal{R}$ which satisfy $P(Z \geq 0) = 1$ and $P(Z > 0) > 0$.
We define the notion of free lunch in the same way as in the previous section.

Definition 2.12 *A multiperiod free lunch in the order book model for the small trader is a sequence of marketed contingent claims $X_n \in \hat{\mathcal{M}}$, $n \in \mathbb{N}$, and a contingent claim $X \in \mathcal{R}^+$ such that* *(i)* *X_n converges to X P-almost surely,*

(ii) *there exists an integrable, non-positive minorant g such that $X_n(\omega) \geq g(\omega)$ for all $\omega \in \Omega$ and all $n \in \mathbb{N}$,*

(iii) *there exists a sequence of self-financing simple trading strategies $Y^{(n)}$, $n \in \mathbb{N}$, such that the values of the hedging portfolios satisfy $V_T^{(n)}(\omega) \geq X_n(\omega)$ for every $\omega \in \Omega$ and every $n \in \mathbb{N}$ and the costs to set up these portfolios satisfy*

$$\limsup_{n\to\infty} \Big((S_0^{(n)}(\omega))^+ \hat{a}_0(\omega) - (S_0^{(n)}(\omega))^- \hat{b}_0(\omega) + C_0(\omega) \Big) \leq 0$$

for every $\omega \in \Omega$.

Now, we can prove the following no-arbitrage result for the model of the order driven market for the small trader.

Theorem 2.13 *(i) The model of the order driven market for the small trader admits no multiperiod free lunch if there exists a probability measure Q equivalent to P with $E_P\left(\left(\frac{dQ}{dP}\right)^2\right) < \infty$ and an adapted process $(\gamma_t)_{t\in[0,T]}$ satisfying $\beta_t \leq \gamma_t \leq \alpha_t$ for all $t \in [0,T]$ such that γ is a martingale under Q.*

(ii) If there exists no multiperiod free lunch, then there is a probability measure Q equivalent to P with $E_P\left(\left(\frac{dQ}{dP}\right)^2\right) < \infty$ and an adapted process $(\gamma_t)_{t\in[0,T]}$ satisfying $\hat{b}_t \leq \gamma_t \leq \hat{a}_t$ for all $t \in [0,T]$ such that γ is a martingale under Q.

Note that part (ii) of Theorem 2.13 is not quite the converse of part (i). In fact, it is a slightly weaker result. Recall that $\hat{a}_t \geq \alpha_t \geq \beta_t \geq \hat{b}_t$ holds for every $t \in [0,T]$ and that the processes $\hat{a}$ and α and the processes β and $\hat{b}$ respectively differ only in finitely many points which are the points where α or β are neither left- nor right-continuous.

Proof of Theorem 2.13: By contradiction. Assume that there exists a probability measure Q equivalent to P and an adapted process $(\gamma_t)_{t\in[0,T]}$ satisfying $\beta_t \leq \gamma_t \leq \alpha_t$

for all $t \in [0, T]$ such that γ is a martingale under Q and assume that there is a multiperiod free lunch in the order book model. Then there exists a sequence of marketed contingent claims $X_n \in \hat{\mathcal{M}}$, $n \in \mathbb{N}$, and a contingent claim $X \in \mathcal{R}^+$ such that X_n converges to X P-almost surely. Furthermore, there exist simple self-financing trading strategies $Y^{(n)}$, $n \in \mathbb{N}$, such that $Y^{(n)}$ leads to a portfolio with value $V_T^{(n)} \geq X_n$ at time T for every $n \in \mathbb{N}$. As in Definition 2.10, let $\rho_1^{(n)}, \ldots, \rho_{N_n}^{(n)}$ be the times at which the orders belonging to the trading strategy $Y^{(n)}$ are executed. These are the times at which the small trader buys and sells shares if he follows the trading strategy $Y^{(n)}$. For $i \in \{1, \ldots, N_n\}$, let $\lambda_i^{(n)}$ be the number of shares which he buys at time $\rho_i^{(n)}$ and let $\mu_i^{(n)}$ be the number of shares which he sells at time $\rho_i^{(n)}$. By construction, the price that he has to pay for the shares which he buys at time $\rho_i^{(n)}$ is $\alpha_{\rho_i^{(n)}}$ apiece or worse. The price that he receives for the shares which he sells at time $\rho_i^{(n)}$ is $\beta_{\rho_i^{(n)}}$ apiece or worse. Recall that $S_0^{(n)}$ denotes the number of shares which the small trader holds at time 0 and that $C_0^{(n)}$ is the amount of money which he has in the bank account at time 0.

Now, consider the securities market model $\left(\binom{\beta}{1}, \binom{\alpha}{1}\right)$ as defined in section 2.1. Define trading strategies $(\theta^{(n)}, \theta'^{(n)})$, $n \in \mathbb{N}$, for this model in the following way: Start at time 0 with $S_0^{(n)}$ shares and $C_0^{(n)}$ units of money in the bank account. Then, buy $\lambda_i^{(n)}$ shares at time $\rho_i^{(n)} > 0$ for $\alpha_{\rho_i^{(n)}}$ apiece and sell $\mu_i^{(n)}$ shares at time $\rho_i^{(n)} > 0$ for $\beta_{\rho_i^{(n)}}$ apiece (for $i = 1, \ldots, m_n$). The trading strategy $(\theta^{(n)}, \theta'^{(n)})$ is simple and self-financing for every $n \in \mathbb{N}$ and it leads to a portfolio with value $\hat{V}_T^{(n)}$ satisfying

$$\hat{V}_T^{(n)}(\omega) \geq S_t(\omega)\beta_t(\omega) + C_t(\omega) \geq V_T^{(n)}(\omega) \geq X_n(\omega)$$

at time T for every $\omega \in \Omega$ since we buy and sell the same amount of shares at the same times as in the model of the order driven market at the same prices (or better). Hence, the contingent claims X_n are marketed in the securities market model $\left(\binom{\beta}{1}, \binom{\alpha}{1}\right)$ for every $n \in \mathbb{N}$. Therefore, the sequence $X_n \in \mathcal{M}$, which converges to $X \in \mathcal{R}^+$ P-almost surely, is a multiperiod free lunch in the securities market model $\left(\binom{\beta}{1}, \binom{\alpha}{1}\right)$.

This contradicts Theorem 2.5 (i) which shows that the securities market model $\left(\binom{\beta}{1}, \binom{\alpha}{1}\right)$ admits no multiperiod free lunch. Thus, part (i) holds.

Part (ii) follows from the fact that no multiperiod free lunch in the model of the order driven market implies no multiperiod free lunch in the securities market model $(\hat{b}, \hat{a})$ with bid-ask spread (as defined in section 2.1). This can be seen as follows: If the model of the order driven market allows no free multiperiod lunch, then the small trader cannot obtain a multiperiod free lunch by using market orders only. We have assumed that the order sizes of the small trader are considerably smaller than the order sizes of the noise traders and that his orders never completely fill one of the orders of the noise traders. Thus, submitting market orders means for the small trader that he buys and sells shares for $\hat{a}$ and $\hat{b}$ respectively. Therefore, he cannot obtain a multiperiod free lunch by buying for $\hat{a}$ and selling for $\hat{b}$.

This means that the securities market model $(\hat{b}, \hat{a})$ with bid-ask spread admits no multiperiod free lunch. Note that $\hat{a}$ and $\hat{b}$ and the trading strategies are defined in such a way that all the assumptions from section 2.1 are satisfied. Theorem 2.5 (ii) now shows the existence of Q and γ as described in Theorem 2.13. □

Remark 2.14 *Theorem 2.13 can be extended to models of order driven markets with $m > 1$ securities. In fact, the same proof works for a market with more than one security. Also, non-zero interest rates can easily be incorporated into the model.*

2.3 Simulation

We have implemented a version of the model of the order driven market which is based on an idea of Maslov [41]. We assume that at every discrete time $t = 0, 1, \ldots$ an order is submitted randomly. The order is a buy order or a sell order with equal probability $\frac{1}{2}$ independently of the previous orders. Also with probability $\frac{1}{2}$, it is a market order to buy/sell one share or it is a limit order with order size one. The limit prices are defined in the following way: Let Δ be a random variable which takes its values in $\{1, 2, 3, 4\}$ uniformly. If the limit order is a buy order, the limit price is the last transaction price minus Δ. If the limit order is a sell order, the limit price is the last transaction price plus Δ.

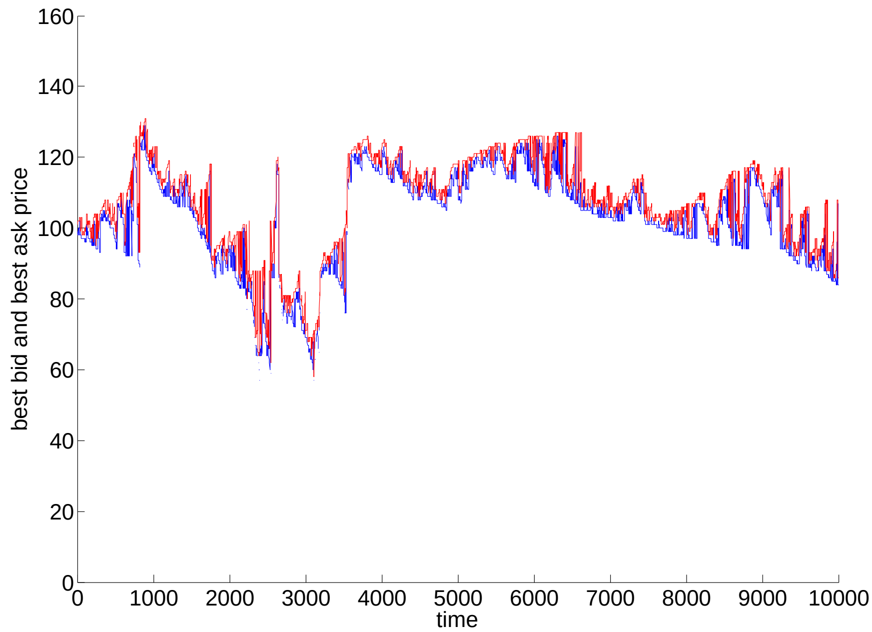

Figure 2.1: Best bid (blue line) and best ask (red line) in the model of the order driven market (over a time interval with length 10000)

Figures 2.1 and 2.2 show the best ask price process (red line) and the best bid price process (blue line) in simulations. In figure 2.1, the evolution over a time interval with length 10000 is depicted, whereas Figure 2.2 shows the fluctuations over a shorter time interval with length 300.

Moreover, we have studied the following trading strategy in this model: Whenever the spread, i.e. the difference between the best ask and the best bid, is greater than two, we place limit orders in the gap. We submit both a buy limit order with limit price equal to the best bid plus one and a sell limit order with limit price equal to the best ask minus one. Then, we wait until the next five orders have arrived. If

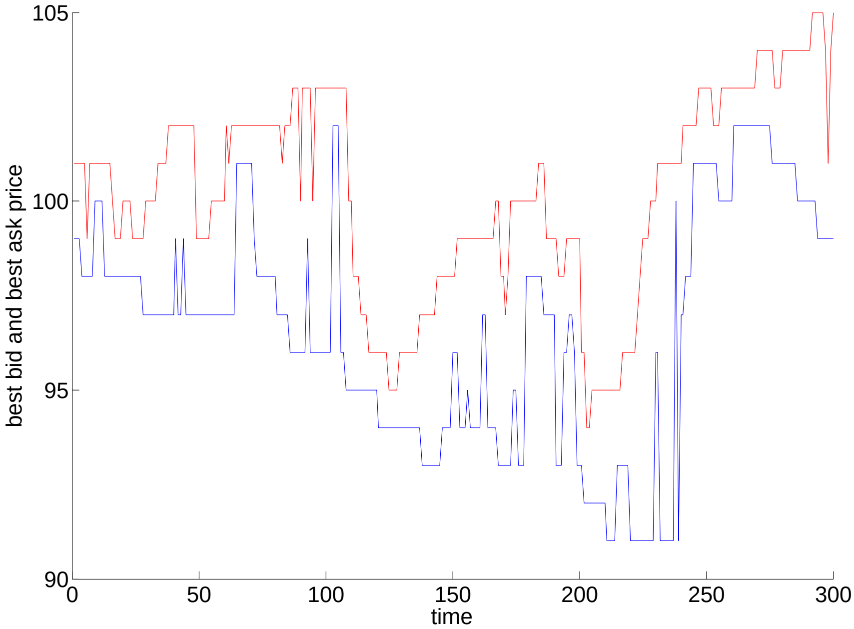

Figure 2.2: Best bid (blue line) and best ask (red line) in the model of the order driven market (over a shorter time interval with length 300)

both of our orders have been filled, we will have earned the difference of the limit prices. Otherwise, we close out our position and possibly incur a loss. Simulations have shown that, in this model, this strategy generates a profit with a positive expectation. Figure 2.3 shows an example of the profit over time.

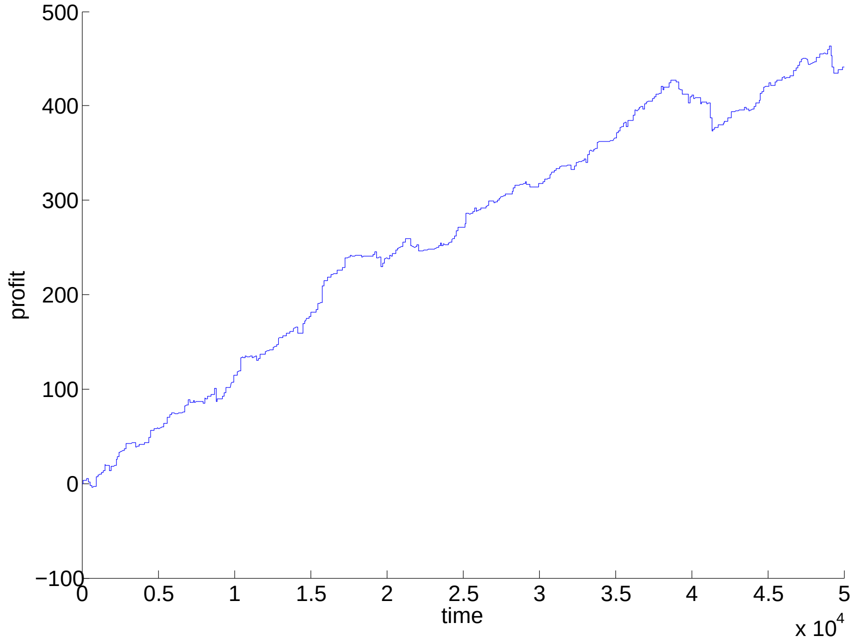

Figure 2.3: Profit generated by the trading strategy in Maslov's model

Chapter 3

Probability measures on signatures and signatures of order books

The signature of a path is the full sequence of its definite iterated integrals. It is a meaningful object which contains a lot of information. In particular, it determines the path up to what is known as tree-like equivalence (see Hambly and Lyons [28]). Moreover, a compactly supported probability measure on signatures is uniquely determined by its expectation. This result has already been known to Chen [12] and Fawcett [20]. In this chapter, we give a different proof which is based on the Stone-Weierstrass Theorem (see for example [34]).

Signatures are also one of the fundamental elements of the theory of rough paths. The theory of rough paths extends classical calculus in such a way that it allows us to describe the dynamical evolution of systems driven by rough signals, including rough paths based on Brownian motion and fractional Brownian motion. A detailed study of signatures and rough paths can be found in Lyons [36], Lyons et al. [37], Lyons and Qian [39] or Friz and Victoir [22].

Seeing the importance of signatures, we compute signatures of data from financial markets. As an example, we look at the (log-) signatures of data from order books of four different futures markets. The data sets have kindly been provided by *Man Investments*.

3.1 Signatures of paths with bounded variation

At first, we give a brief introduction to signatures following [37]. We define the space of formal series of tensors of a finite dimensional Banach space as in [37], Definition 2.4.

Definition 3.1 *Let E be a finite dimensional Banach space. The space of formal series of tensors of E is defined as the space of sequences*

$$T((E)) = \{\mathbf{a} = (a_0, a_1, \dots)|a_n \in E^{\otimes n}\,\forall n \in \mathbb{N}_0\},$$

where $E^{\otimes 0} = \mathbb{R}$.
It is endowed with a scalar multiplication and two other intrinsic operations, an addition + and a product $\otimes$. Let $\mathbf{a} = (a_0, a_1, \dots)$ and $\mathbf{b} = (b_0, b_1, \dots)$ be two elements of $T((E))$ and $\lambda \in \mathbb{R}$. Then, we define

$$\begin{aligned} \lambda\mathbf{a} &= (\lambda a_0, \lambda a_1, \dots), \\ \mathbf{a} + \mathbf{b} &= (a_0 + b_0, a_1 + b_1, \dots) \end{aligned}$$

and

$$\mathbf{a} \otimes \mathbf{b} = (c_0, c_1, \dots),$$

where for each $n \in \mathbb{N}$,

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

We equip $T((E))$ with the supremum norm: Let $\mathbf{a} = (a_0, a_1, \dots)$ be an element of $T((E))$. Let $d < \infty$ denote the dimension of E and let $\{e_1, \dots, e_d\}$ be a basis of E. Let $k \geq 1$ be an integer. Then, the elements $e_{i_1} \otimes \cdots \otimes e_{i_k}$, $(i_1, \dots, i_k) \in \{1, \dots, d\}^k$, form a basis of $E^{\otimes k}$. Therefore, the element $a_k \in E^{\otimes k}$ can be uniquely written as

$$a_k = \sum_{(i_1,\dots,i_k)\in\{1,\dots,d\}^k} a_k^{(i_1,\dots,i_k)} e_{i_1} \otimes \cdots \otimes e_{i_k} \tag{3.1}$$

with coefficients $a_k^{(i_1,\dots,i_k)} \in \mathbb{R}$. Let $\|\cdot\|_{E^{\otimes k},\infty}$ be the maximum norm on $E^{\otimes k}$, i.e.

$$\|a_k\|_{E^{\otimes k},\infty} = \max_{(i_1,\dots,i_k)\in\{1,\dots,d\}^k} |a_k^{(i_1,\dots,i_k)}|.$$

For $k = 0$, we use the notation $E^{\otimes 0} = \mathbb{R}$ and we write $\|a_0\|_{E^{\otimes 0},\infty} = |a_0|$. Then, set

$$\|\mathbf{a}\| = \sup_{k\in\mathbb{N}_0} \|a_k\|_{E^{\otimes k},\infty},$$

where $\mathbb{N}_0$ denotes the set of non-negative integers $\{0, 1, 2, \dots\}$.

In the same way, we define the space of truncated formal series of tensors over the d-dimensional space $\mathbb{R}^d$ (see also [37]).

Definition 3.2 *Let $n \geq 1$ be an integer. The space of truncated formal series of tensors over $\mathbb{R}^d$ is defined as the space*

$$T^{(n)}(\mathbb{R}^d) = \bigoplus_{i=0}^{n} (\mathbb{R}^d)^{\otimes i} = \big\{\mathbf{a} = (a_0, a_1, \dots, a_n) : a_i \in (\mathbb{R}^d)^{\otimes i}\ \forall i \in \{0, 1, \dots, n\}\big\}$$

where $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$.

It is endowed with a scalar multiplication and two other intrinsic operations, an addition $+$ and a product $\otimes$. Let $\mathbf{a} = (a_0, a_1, \dots, a_n)$ and $\mathbf{b} = (b_0, b_1, \dots, b_n)$ be two *elements of* $T^{(n)}(\mathbb{R}^d)$ *and* $\lambda \in \mathbb{R}$. *Then, we define*

$$\begin{aligned} \lambda \mathbf{a} &= (\lambda a_0, \lambda a_1, \ldots, \lambda a_n), \\ \mathbf{a} + \mathbf{b} &= (a_0 + b_0, a_1 + b_1, \ldots, a_n + b_n) \end{aligned}$$

and

$$\mathbf{a} \otimes \mathbf{b} = (c_0, c_1, \ldots, c_n),$$

where

$$c_i = \sum_{j=0}^{i} a_j \otimes b_{i-j}.$$

for each $i \in \{0, 1, \ldots, n\}$.

We also equip $T^{(n)}(\mathbb{R}^d)$ with the supremum norm: Let $\mathbf{a} = (a_0, a_1, \ldots, a_n)$ be an element of $T^{(n)}(\mathbb{R}^d)$ and let $\{e_1, \ldots, e_d\}$ be the standard basis of $\mathbb{R}^d$. For every $k \in \{1, \ldots, n\}$, a_k is an element of $(\mathbb{R}^d)^{\otimes k}$ and it can be written uniquely as

$$a_k = \sum_{(i_1,\ldots,i_k) \in \{1,\ldots,d\}^k} a_k^{(i_1,\ldots,i_k)} e_{i_1} \otimes \cdots \otimes e_{i_k}$$

with coefficients $a_k^{(i_1,\ldots,i_k)} \in \mathbb{R}$. Let $\|\cdot\|$ be the maximum-norm on $(\mathbb{R}^d)^{\otimes k}$, i.e.

$$\|a_k\| = \max_{(i_1,\ldots,i_k) \in \{1,\ldots,d\}^k} |a_k^{(i_1,\ldots,i_k)}|.$$

Define

$$\|\mathbf{a}\| = \max\{|a_0|, \|a_1\|, \ldots, \|a_n\|\}.$$

We can now define the signature of a path (as in [37], Definition 2.6).

Definition 3.3 *Let* $[s,t] \subset \mathbb{R}$ *denote a compact interval. Let* $X : [s,t] \to E$ *be a continuous path with bounded variation. The signature of the path* X *is the element*

$X_{s,t} \in T((E))$ defined by

$$X_{s,t} = (1, X^1_{s,t}, X^2_{s,t}, \dots),$$

where

$$X^n_{s,t} = \int_{s<u_1<\dots<u_n<t} dX_{u_1} \otimes \dots \otimes dX_{u_n}$$

for each integer $n \geq 1$.

We write $S(\mathcal{V}^1([s,t], E)) \subset T((E))$ for the set of signatures of continuous paths $X : [s,t] \to E$ with bounded variation.

Note that

$$\int_{s<u_1<\dots<u_n<t} dX_{u_1} \otimes \dots \otimes dX_{u_n} = \sum_{(i_1,\dots,i_n)\in\{1,\dots,d\}^n} X^{(i_1,\dots,i_n)}_{s,t} e_{i_1} \otimes \dots \otimes e_{i_n},$$

where $\{e_1, \dots, e_d\}$ is the basis of E and

$$X^{(i_1,\dots,i_n)}_{s,t} = \int_{s<u_1<\dots<u_n<t} dX^{i_1}_{u_1} \dots dX^{i_n}_{u_n} \tag{3.2}$$

are the coordinate iterated integrals.

Remark 3.4 *Hambly and Lyons [28] show that two $\mathbb{R}^d$-valued continuous paths X and Y with bounded variation have the same signatures if and only if the concatenation of X and 'Y run backwards' is a Lipschitz tree-like path. They call a path $Z : [0,T] \to \mathbb{R}^d$ Lipschitz tree-like, if there exists a continuous function $h : [0,T] \to [0,+\infty)$ with bounded variation such that $h(0) = h(T) = 0$ and*

$$\|X_t - X_s\| \leq h(s) + h(t) - 2 \inf_{u\in[s,t]} h(u)$$

for all $s,t \in [0,T]$ with $s \leq t$.

We can define the exponential map $\exp : T((E)) \to T((E)))$ by

$$\exp(\mathbf{a}) = \sum_{k=0}^{\infty} \frac{\mathbf{a}^k}{k!}$$

for all $\mathbf{a} \in T((E))$ (see [37], Lemma 2.19 and Definition 2.20). For all elements $\mathbf{a} = (a_0, a_1, \dots) \in T((E))$ with $a_0 > 0$, we can define the logarithm (see also [37], Definition 2.20) by

$$\log \mathbf{a} = \log(a_0) + \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \left(\mathbf{1} - \frac{\mathbf{a}}{a_0} \right)^k.$$

Again, we can also define the exponential map $\exp : T^{(n)}(\mathbb{R}^d) \to T^{(n)}(\mathbb{R}^d)$ on the space of truncated formal series of tensors in the same way by

$$\exp(\mathbf{a}) = \sum_{k=0}^{\infty} \frac{\mathbf{a}^k}{k!}$$

for all $\mathbf{a} \in T^{(n)}(\mathbb{R}^d)$ (see [37], Lemma 2.19 and Definition 2.20). Also, for all elements $\mathbf{a} = (a_0, a_1, \dots, a_n) \in T^{(n)}(\mathbb{R}^d)$ with $a_0 > 0$, the logarithm (see also [37], Definition 2.20) is defined by

$$\log \mathbf{a} = \log(a_0) + \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \left(\mathbf{1} - \frac{\mathbf{a}}{a_0} \right)^k.$$

Lemma 3.5 *Let α and β be elements of $\mathbb{R}^d$ and let $X : [0,1] \to \mathbb{R}^d$, $X_t = \alpha t + \beta$ be the straight line. Then, the signature of X is*

$$X_{0,1} = \exp(\alpha),$$

where we write $\exp(\alpha)$ *for* $\exp((0, \alpha, 0, 0, \dots))$.

Proof: By induction, we get

$$\int\limits_{0<u_1<\cdots<u_n<t} du_1 \dots du_n = \frac{1}{n!} t^n$$

for every $t > 0$ and every integer $n \geq 1$ since $\int_0^t du_1 = t$ holds ($n = 1$) and since we have

$$\begin{aligned} \int\limits_{0<u_1<\cdots<u_n<t} du_1 \dots du_n &= \int_0^t \left(\int\limits_{0<u_1<\cdots<u_{n-1}<u_n} du_1 \dots du_{n-1} \right) du_n \\ &= \int_0^t \frac{1}{(n-1)!} u_n^{n-1} \, du_n = \frac{1}{n!} t^n \end{aligned}$$

where we have assumed that the assertion holds for $n - 1$.

Therefore, we have

$$\int\limits_{0<u_1<\cdots<u_n<1} dX_{u_1}^{i_1} \dots dX_{u_n}^{i_n} = \alpha_{i_1} \dots \alpha_{i_n} \int\limits_{0<u_1<\cdots<u_n<1} du_1 \dots du_n = \frac{1}{n!} \alpha_{i_1} \dots \alpha_{i_n}$$

for every $(i_1, \dots, i_n) \in \{1, \dots, d\}^n$ and thus the signature of the path X is $X_{0,1} = (1, X_{0,1}^1, X_{0,1}^2, \dots)$, where

$$X_{0,1}^n = \frac{1}{n!} \sum_{(i_1,\dots,i_n)\in\{1,\dots,d\}^n} \alpha_{i_1} \dots \alpha_{i_n} \, e_{i_1} \otimes \dots \otimes e_{i_n}$$

for every integer $n \geq 1$.

Let $\mathbf{a} = (a_0, a_1, \dots) \in T((\mathbb{R}^d))$ be given by $\mathbf{a} = \exp(\alpha)$. By the definition of the exponential and the multilinearity of the tensor product, we have $a_0 = 1$ and

$$a_n = \frac{\alpha^{\otimes n}}{n!} = \frac{1}{n!} \sum_{(i_1,\dots,i_n)\in\{1,\dots,d\}^n} \alpha_{i_1} \dots \alpha_{i_n} \, e_{i_1} \otimes \dots \otimes e_{i_n}$$

for every $n \geq 1$.

Thus, we have indeed $X_{0,1} = \exp(\alpha)$. □

The logarithm of signatures of continuous paths with bounded variation lives in an important subspace of $T((E))$ (see also [37]). Define a Lie bracket $[\cdot,\cdot]$ on $T((E))$ by

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

For linear subspaces $F \subset T((E))$ and $\tilde{F} \subset T((E))$, we write $[F,\tilde{F}]$ for the linear space that is spanned by all elements of the form $[\mathbf{a},\mathbf{b}]$, where $\mathbf{a}\in F$ and $\mathbf{b}\in\tilde{F}$. Using this notation, we now give the definition of the space of Lie series as in [37], Definition 2.22.

Definition 3.6 *The space of Lie series $\mathcal{L}((E))$ is the space*

$$\mathcal{L}((E)) = \{(l_0, l_1, l_2, \dots)| l_k \in F_k \forall k \in \mathbb{N}_0\} \subset T((E)),$$

where $F_0 = \{0\}$ and for $k = 1, 2, \dots,$ the linear spaces $F_k \subset E^{\otimes k}$ are defined recursively by $F_{k+1} = [E, F_k]$ starting from $F_1 = E$.

Then, let $X : [0,T] \to E$ be a continuous path with bounded variation and let $X_{0,T}$ be its signature. It follows from [37] that $\log X_{0,T}$ is an element of $\mathcal{L}((E))$. Indeed, $X_{0,T}$ is a group-like element (see [37], Definition 2.18) because of [37], equation (2.6), and the logarithm of a group-like element is a Lie series (see [37], Theorem 2.23). We also define a Lie bracket $[\cdot,\cdot]$ on the space $T^{(n)}(\mathbb{R}^d)$ of truncated formal series of tensors by

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

Again, for linear subspaces $F \subset T^{(n)}(\mathbb{R}^d)$ and $\tilde{F} \subset T^{(n)}(\mathbb{R}^d)$, we write $[F,\tilde{F}]$ for the linear space that is spanned by all elements of the form $[\mathbf{a},\mathbf{b}]$, where $\mathbf{a}\in F$ and $\mathbf{b}\in\tilde{F}$. Using this notation, we now give the definition of the space of truncated Lie series as in [37].

Definition 3.7 *Let $n \geq 1$ be an integer. The space of truncated Lie series $\mathcal{L}^{(n)}(\mathbb{R}^d)$*

is the space

$$\mathcal{L}^{(n)}(\mathbb{R}^d) = \{(l_0, l_1, \ldots, l_n) : l_k \in F_k \forall k \in \{0, 1, \ldots, n\}\} \subset T^{(n)}(\mathbb{R}^d),$$

where $F_0 = \{0\}$ *and for* $k = 1, \ldots, n$, *the spaces* $F_k \subset (\mathbb{R}^d)^{\otimes k}$ *are defined recursively by* $F_{k+1} = [\mathbb{R}^d, F_k]$ *starting from* $F_1 = \mathbb{R}^d$.

Example 3.8 (Brownian motion) *Let* $(B_t)_{t\in[0,1]}$ *be a* 2*-dimensional standard Brownian motion and let* $\{\varepsilon_1, \varepsilon_2\}$ *be the standard basis of* $\mathbb{R}^2$. *Then, the Stratonovich signature of* B *is*

$$B_{0,1} = \left(1, B_1 - B_0, \sum_{i_1,i_2=1}^{2} \left(\int_{0<u_1<u_2<1} \circ dB^{i_1}_{u_1} \circ dB^{i_2}_{u_2} \right) \varepsilon_{i_1} \otimes \varepsilon_{i_2}, \ldots \right),$$

where the integrals are Stratonovich integrals. The expectation of the Stratonovich signature is

$$\mathrm{E}(B_{0,1}) = \exp\left(\frac{1}{2}(\varepsilon_1 \otimes \varepsilon_1 + \varepsilon_2 \otimes \varepsilon_2)\right)$$

(see Fawcett [20], Lyons and Victoir [40] or Levin and Wildon [35]), where we identify $\frac{1}{2}(\varepsilon_1 \otimes \varepsilon_1 + \varepsilon_2 \otimes \varepsilon_2)$ *with the element* $(0, 0, \frac{1}{2}(\varepsilon_1 \otimes \varepsilon_1 + \varepsilon_2 \otimes \varepsilon_2), 0, \ldots) \in T((\mathbb{R}^2))$. *Lyons and Ni [38] study the expected Stratonovich signature of Brownian motion stopped at the first exit time of a regular domain. Let* $D \subset \mathbb{R}^2$ *be a compact domain which is regular in the sense of [38]. Let* $(B^{[x]}_t)_{t\geq 0}$ *be a* 2*-dimensional standard Brownian motion starting at* $x \in D$. *Let* $\tau^{[x]}_D = \min\{t \geq 0 | B^{[x]}_t \in \partial D\}$ *be the first exit time of the domain* D *and set* $\Phi_D(x) = \mathrm{E}(B^{[x]}_{0,\tau^{[x]}_D})$. *Then, Lyons and Ni prove that* Φ_D *satisfies the partial differential equation*

$$\Delta \Phi_D(x) = -\left(\sum_{i=1}^{2} \varepsilon_i \otimes \varepsilon_i\right) \otimes \Phi_D(x) - 2 \sum_{i=1}^{2} \varepsilon_i \otimes \frac{\partial \Phi_D(x)}{\partial x_i}$$

with boundary condition

$$\Phi_D(x) = \mathbf{1} \quad \text{for } x \in \partial D$$

(see [38], Theorem 3.2). Using this partial differential equation, they show how every term of the expected Stratonovich signature $\mathrm{E}(B^{[x]}_{0,\tau^{[x]}_D})$ *can be obtained recursively.*

3.2 Probability measures on signatures

Throughout this section, we assume again that E is a finite dimensional Banach space as in the previous section. Recall that we have endowed $T((E))$ with the supremum norm $\|\cdot\|$ and that $\|\cdot\|_{E^{\otimes k},\infty}$ is the maximum norm on $E^{\otimes k}$.
For every $k \in \mathbb{N}$, the norm $\|\cdot\|_{E^{\otimes k},\infty}$ induces a metric $d_{E^{\otimes k}}$ on the space $E^{\otimes k}$ where $d_{E^{\otimes k}}(x,y) = \|x-y\|_{E^{\otimes k},\infty}$ for $x, y \in E^{\otimes k}$. The space $E^{\otimes k}$ equipped with $d_{E^{\otimes k}}$ forms a complete metric space. The metric $d_{E^{\otimes k}}$ induces a topology on $E^{\otimes k}$, which shall be denoted by τ_k. Let p_k be the projection of $T((E))$ on $E^{\otimes k}$. Furthermore, let $\tau_{T((E))}$ be the product topology of $\tau_0, \tau_1, \ldots$ on $T((E))$ and let Σ be the Borel σ-algebra on $T((E))$ associated with $\tau_{T((E))}$.

Lemma 3.9 *Let μ be a probability measure on $(T((E)), \Sigma)$ with compact support M. Then, the Bochner integral $\int_{T((E))} p_k(s)\, \mu(ds) \in E^{\otimes k}$ exists for every $k \in \mathbb{N}$, where the Bochner integral is defined as in Diestel [19], Chapter IV.*

For example, finite sets of signatures are compact. They occur for example when studying self-avoiding paths on a bounded subset of a lattice.

Proof of Lemma 3.9: Consider the probability space $(T((E)), \Sigma, \mu)$ and the Banach space $(E^{\otimes k}, d_{E^{\otimes k}})$. Since the space $E^{\otimes k}$ is finite dimensional, every linear form $F \in (E^{\otimes k})^*$ is continuous. The projection p_k is also continuous. Hence, $F \circ p_k : T((E)) \to \mathbb{R}$ is continuous which shows that it is Borel-measurable.
In the notation of [19], the projection $p_k : T((E)) \to E^{\otimes k}$ is scalarly μ-measurable since $F \circ p_k : T((E)) \to \mathbb{R}$ is measurable for every $F \in (E^{\otimes k})^*$. The space $E^{\otimes k}$ is separable. Therefore, it follows from Pettis' Measurability Theorem (see Diestel [19], p. 25) that the projection $p_k : T((E)) \to E^{\otimes k}$ is μ-measurable.

Furthermore, we have $\int_{T((E))} \|p_k(s)\|_{E^{\otimes k},\infty}\, \mu(ds) < \infty$ as the support of μ is compact. Hence, Bochner's Characterization of Integrable Functions (see Diestel [19], p. 26) shows that the projection $p_k : T((E)) \to E^{\otimes k}$ is Bochner integrable. Thus, the Bochner integral $\int_{T((E))} p_k(s)\, \mu(ds)$ exists. □

From now on, we equip the support $M \subset T((E))$ with the relative topology which shall be denoted by τ_M. Then, Σ induces a σ-algebra Σ_M on M which is the Borel σ-algebra associated with τ_M. The probability measure μ can also be regarded as a probability measure on (M, Σ_M).
The following proposition states that compactly supported probability measures on signatures are uniquely determined by their expectation. This has already been known to Chen [12] and Fawcett [20]. Here, we provide a different proof, which is based on the Stone-Weierstrass Theorem (see for example Lax [34], page 126, Theorem 4) and the Riesz Representation Theorem (see for example Rudin [43], Theorem 2.14).

Proposition 3.10 *Let μ be a probability measure on $(T((E)), \Sigma)$ with compact support $M \subset S(\mathcal{V}^1([0,T], E)) \subset T((E))$. Then, μ is uniquely determined by the sequence of expectations $\int_M p_n(s)\, \mu(ds) \in E^{\otimes n}$, $n \in \{1, 2, \dots\}$.*

Expected signatures also play an important role for applications like cubature on Wiener space (see Lyons and Victoir [40]).
For the proof of Proposition 3.10, we need shuffle products and linear forms on $T((E))$, restricted to the signatures $S(\mathcal{V}^1([0,T], E))$. In the following paragraphs, we introduce these notions as in [37], section 2.2.3.
Let $n \geq 1$ be an integer. Recall that $\{e_1, \dots, e_d\}$ denotes the basis of E. Let $\{e_1^*, \dots, e_d^*\}$ be the associated dual basis of E^* such that

$$e_i^*(e_j) = \delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} .$$

Then, the elements $e_{i_1} \otimes \cdots \otimes e_{i_n}$, $(i_1, \ldots, i_n) \in \{1, \ldots, d\}^n$, form a basis of $E^{\otimes n}$ and the elements $e^*_{i_1} \otimes \cdots \otimes e^*_{i_n}$, $(i_1, \ldots, i_n) \in \{1, \ldots, d\}^n$, form a basis of $(E^*)^{\otimes n}$.
For every $(i_1, \ldots, i_n) \in \{1, \ldots, d\}^n$, define a linear form $e^*_{i_1,\ldots,i_n} \in (E^{\otimes n})^*$ by

$$e^*_{i_1,\ldots,i_n}(e_{j_1} \otimes \cdots \otimes e_{j_n}) = \delta_{i_1,j_1} \ldots \delta_{i_n,j_n} \tag{3.3}$$

for all $(j_1, \ldots, j_n) \in \{1, \ldots, d\}^n$. Then, the elements $e^*_{i_1,\ldots,i_n}$, $(i_1, \ldots, i_n) \in \{1, \ldots, d\}^n$, form a basis of $(E^{\otimes n})^*$. The canonical isomorphism $\psi_n : (E^*)^{\otimes n} \to (E^{\otimes n})^*$ is determined by $\psi_n(e^*_{i_1} \otimes \cdots \otimes e^*_{i_n}) = e^*_{i_1,\ldots,i_n}$ for every $(i_1, \ldots, i_n) \in \{1, \ldots, d\}^n$.
Every linear form g^* on $E^{\otimes n}$ has a canonical extension to a linear form $\theta_n(g^*)$ on $T((E))$, defined by

$$(\theta_n(g^*))(\mathbf{a}) = g^*(a_n) \tag{3.4}$$

for all $\mathbf{a} = (a_0, a_1, \ldots) \in T((E))$. Each element of $T(E^*) = \bigoplus_{k=0}^{\infty}(E^*)^{\otimes k}$ can be regarded as a linear form on $T((E))$: There is a natural linear map

$$\varphi : T(E^*) \longrightarrow T((E))^*,$$

defined by

$$\varphi(f^*_0, f^*_1, f^*_2, \ldots) = \sum_{k \in \mathbb{N}_0} \theta_k(\psi_k(f^*_k))$$

where $f^*_k \in (E^*)^{\otimes k}$ for every $k \in \mathbb{N}$ and only finitely many elements of the sequence $(f^*_0, f^*_1, f^*_2, \ldots) \in \bigoplus_{k=0}^{\infty}(E^*)^{\otimes k}$ differ from 0.

Lemma 3.11 *Every linear form on $T((E))$, which is induced by an element of $T(E^*)$, is continuous.*

Proof: Let $\mathbf{f}^*$ be a linear form on $T((E))$, which is induced by an element of $T(E^*)$. Then, there exists $g^* \in T(E^*)$ such that $\varphi(g^*) = \mathbf{f}^*$. By the linearity of φ, it is sufficient to prove the assertion for $g^* = (g^*_0, g^*_1, g^*_2, \ldots) \in T(E^*)$ where only $g^*_n \in (E^*)^{\otimes n}$ differs from 0 for some $n \in \mathbb{N}$ and all the other $g^*_k \in (E^*)^{\otimes k}$, $k \neq n$, are equal to 0. Recall that $p_n : T((E)) \to E^{\otimes n}$ denotes the projection of $T((E))$ onto $E^{\otimes n}$. Then, $\mathbf{f}^* = g_n^* \circ p_n$. By the definition of the product topology, p_n is continuous. The linear form g_n^* on $E^{\otimes n}$ is also continuous as $E^{\otimes n}$ is finite dimesional. Therefore, $\mathbf{f}^*$ is continuous. □

We can define a product on $T(E^*)$, called the shuffle product. Let e^* and f^* be two elements of $T(E^*)$. By linearity, it is sufficient to define the shuffle product $e^* \sqcup f^*$ for $e^* = (0, \dots, 0, e_{i_1}^* \otimes \dots \otimes e_{i_r}^*, 0, 0, \dots)$ and $f^* = (0, \dots, 0, e_{j_1}^* \otimes \dots \otimes e_{j_s}^*, 0, 0, \dots)$ where $(i_1, \dots, i_r) \in \{1, \dots, d\}^r$ and $(j_1, \dots, j_s) \in \{1, \dots, d\}^s$. All the components of e^* equal 0 bar the r-th component and all the components of f^* equal 0 bar the s-th component.

A permutation σ of $\{1, \dots, r+s\}$ is called a shuffle of $\{1, \dots, r\}$ and $\{r+1, \dots, r+s\}$ if $\sigma(1) < \dots < \sigma(r)$ and $\sigma(r+1) < \dots < \sigma(r+s)$. We denote the set of all such shuffles by *Shuffles*(r, s). Set $(k_1, \dots, k_{r+s}) = (i_1, \dots, i_r, j_1, \dots, j_s)$ and define $e^* \sqcup f^* \in T(E^*)$ by

$$e^* \sqcup f^* = (0, \dots, 0, g_{r+s}^*, 0, 0, \dots)$$

where the $(r+s)$-th component is

$$g_{r+s}^* = \sum_{\sigma \in \mathit{Shuffles}(r,s)} e_{k_{\sigma^{-1}(1)}}^* \otimes \dots \otimes e_{k_{\sigma^{-1}(r+s)}}^* \in (E^*)^{\otimes(r+s)}$$

and all the other components of $e^* \sqcup f^*$ equal 0.

The shuffle product on $T(E^*)$ defines a product on the linear forms on $T((E))$ which are induced by elements of $T(E^*)$. This product is also called shuffle product. Let $\mathbf{e}^*$ and $\mathbf{f}^*$ be two elements of $T((E))^*$ which are induced by elements of $T(E^*)$. Then, there exist $g^* \in T(E^*)$ and $h^* \in T(E^*)$ such that $\mathbf{e}^* = \varphi(g^*)$ and $\mathbf{f}^* = \varphi(h^*)$. Define the shuffle product $\mathbf{e}^* \sqcup \mathbf{f}^*$ by

$$\mathbf{e}^* \sqcup \mathbf{f}^* = \varphi(g^* \sqcup h^*) \in T((E))^*.$$

Let $\mathbf{g}^*$ and $\mathbf{h}^*$ be two linear forms on $T((E))$, which are induced by elements of $T(E^*)$, and let $A \subset T((E))$. Then, we define the shuffle product of the restricted linear maps $\mathbf{g}^*_{|A}, \mathbf{h}^*_{|A} : A \to \mathbb{R}$ by

$$\mathbf{g}^*_{|A} \sqcup \mathbf{h}^*_{|A} = (\mathbf{g}^* \sqcup \mathbf{h}^*)_{|A}.$$

Let $\mathcal{B} = \{\varphi(f^*)_{|S(\mathcal{V}^1([0,T],E))} | f^* \in T(E^*)\}$ be the set of linear maps from the signatures $S(\mathcal{V}^1([0,T],E))$ to $\mathbb{R}$ which are linear forms on $T((E))$, induced by elements of $T(E^*)$ and restricted to the signatures $S(\mathcal{V}^1([0,T],E))$. Then, [37], Theorem 2.15, shows that $\mathcal{B}$ equipped with the shuffle product is an algebra of real-valued functions. Moreover,

$$(\mathbf{e}^* \sqcup \mathbf{f}^*)(\mathbf{a}) = \mathbf{e}^*(\mathbf{a})\, \mathbf{f}^*(\mathbf{a}) \tag{3.5}$$

holds for all signatures $\mathbf{a} \in S(\mathcal{V}^1([0,T],E))$ and all linear forms $\mathbf{e}^*$ and $\mathbf{f}^*$ on $T((E))$, which are induced by elements of $T(E^*)$ (see [37], equation (2.6)).

Let $\mathcal{A}$ denote the vector space of affine maps from $S(\mathcal{V}^1([0,T],E))$ to $\mathbb{R}$ of the form $s \mapsto f^*(s) + b$ where $b \in \mathbb{R}$ and $f^* \in \mathcal{B}$ is a linear form on $T((E))$ induced by an element of $T(E^*)$ and restricted to the signatures $S(\mathcal{V}^1([0,T],E))$. Then, we can define the shuffle product on $\mathcal{A}$ in the natural way: Let $g_1(s) = f_1^*(s) + b_1$ and $g_2(s) = f_2^*(s) + b_2$ be two elements of $\mathcal{A}$. Define

$$g_1 \sqcup g_2 = f_1^* \sqcup f_2^* + b_2 f_1^* + b_1 f_2^* + b_1 b_2 \in \mathcal{A}. \tag{3.6}$$

Lemma 3.12 *The vector space $\mathcal{A}$ equipped with the shuffle product is an algebra.*

Proof: It follows from (3.6) and (3.5) that the shuffle product on $\mathcal{A}$ is bilinear. Therefore, $\mathcal{A}$ is an algebra. □

Now we can prove Proposition 3.10.

Proof of Proposition 3.10: The Riesz Representation Theorem (see for example

Rudin [43], Theorem 2.14) implies that μ is uniquely determined by all integrals

$$\int_M f(s)\,\mu(ds), \qquad f \in C(M),$$

where $C(M)$ denotes the set of real-valued continuous functions on M.

Now, recall the Stone-Weierstrass Theorem (see for example Lax [34], p. 126, Theorem 4): Let H be a compact Hausdorff space and let $C(H)$ denote the algebra of real-valued continuous functions on H, equipped with the topology of uniform convergence. Let D be a subalgebra of $C(H)$. Then, the Stone-Weierstrass Theorem says that D is dense in $C(H)$ if it seperates the points of H and contains the constant functions.

For every $k \in \mathbb{N}$, the space $E^{\otimes k}$ with the topology τ_k is a Hausdorff space as τ_k is induced by the metric $d_{E^{\otimes k}}$. Therefore, the space $T((E))$ equipped with the product topology $\tau_{T((E))}$ is also a Hausdorff space. By assumption, the support $M \subset S(\mathcal{V}^1([0,T],E)) \subset T((E))$ of μ is compact. Thus, M equipped with the relative topology is a compact Hausdorff space.

Let $\mathcal{A}_M = \{g_{|M} |\, g \in \mathcal{A}\}$ be the set of affine maps $g \in \mathcal{A}$, restricted to M, where $\mathcal{A}$ is defined as above. Lemma 3.12 implies that $\mathcal{A}_M$ is an algebra and it follows from Lemma 3.11 that every function $f \in \mathcal{A}_M$ is continuous. Therefore, $\mathcal{A}_M$ is a subalgebra of $C(M)$.

Let $\mathbf{a} = (a_0, a_1, \dots)$ and $\mathbf{b} = (b_0, b_1, \dots)$ be two different elements of M. Then, there exists $m \in \mathbb{N}$ such that $a_m \neq b_m$. The elements a_m and b_m of $E^{\otimes m}$ can be uniquely written as

$$a_m = \sum_{(i_1,\dots,i_m)\in\{1,\dots,d\}^m} \alpha_{i_1,\dots,i_m} e_{i_1} \otimes \cdots \otimes e_{i_m}$$

and

$$b_m = \sum_{(j_1,\dots,j_m)\in\{1,\dots,d\}^m} \beta_{j_1,\dots,j_m} e_{j_1} \otimes \cdots \otimes e_{j_m}$$

with coefficients $\alpha_{i_1,\dots,i_m}, \beta_{j_1,\dots,j_m} \in \mathbb{R}$ for all $(i_1,\dots,i_m),(j_1,\dots,j_m) \in \{1,\dots,d\}^m$. Since $a_m \neq b_m$, there exists $(l_1,\dots,l_m) \in \{1,\dots,d\}^m$ such that $\alpha_{l_1,\dots,l_m} \neq \beta_{l_1,\dots,l_m}$. Consider the function

$$h = \varphi((0,\dots,0,e^*_{l_1} \otimes \dots \otimes e^*_{l_m},0,0,\dots))_{|M}.$$

Then, $h \in \mathcal{A}_M$ and we have $h(\mathbf{a}) = \alpha_{l_1,\dots,l_m}$ and $h(\mathbf{b}) = \beta_{l_1,\dots,l_m}$. Hence, $h(\mathbf{a}) \neq h(\mathbf{b})$. Thus, $\mathcal{A}_M$ separates the points of M.

By definition, $\mathcal{A}_M$ contains all constant functions from M to $\mathbb{R}$.

The Stone-Weierstrass Theorem (see for example Lax [34], p. 126, Theorem 4) now shows that $\mathcal{A}_M$ is dense in $C(M)$ in the maximum norm. Therefore, μ is uniquely determined by all integrals

$$\int_M f(s)\,\mu(ds), \qquad f \in \mathcal{A}_M.$$

For every integer $n \geq 1$, the expectation $\int_M p_n(s)\,\mu(ds) \in E^{\otimes n}$ can be uniquely written as

$$\int_M p_n(s)\,\mu(ds) = \sum_{(i_1,\dots,i_n)\in\{1,\dots,d\}^n} \gamma_{i_1,\dots,i_n} e_{i_1} \otimes \dots \otimes e_{i_n}$$

with $\gamma_{i_1,\dots,i_n} \in \mathbb{R}$ for every $(i_1,\dots,i_n) \in \{1,\dots,d\}^n$.

Now, we show that all the integrals $\int_M f(s)\,\mu(ds)$, $f \in \mathcal{A}_M$, are known from the sequence of expectations $\int_M p_n(s)\,\mu(ds)$, $n \in \{1,2,\dots\}$.

By linearity, it is sufficient to consider only the functions f of the form

$$f = \varphi((0,\dots,0,e^*_{i_1} \otimes \dots \otimes e^*_{i_n},0,0,\dots))_{|M} \in \mathcal{A}_M \tag{3.7}$$

for $n \in \{1, 2, \dots\}$ and $(i_1, \dots, i_n) \in \{1, \dots, d\}^n$ and the constant functions. Let f be a function as in (3.7). Then,

$$f(s) = (\psi_n(e_{i_1}^* \otimes \cdots \otimes e_{i_n}^*))(p_n(s))$$

holds for all $s \in M$. Recall that $\psi_n(e_{i_1}^* \otimes \cdots \otimes e_{i_n}^*) = e_{i_1,\dots,i_n}^* \in (E^{\otimes n})^*$ as defined in (3.3). By linearity, we have

$$\int\limits_M f(s)\,\mu(ds) = \int\limits_M e_{i_1,\dots,i_n}^*(p_n(s))\,\mu(ds) = e_{i_1,\dots,i_n}^*\left(\int_M p_n(s)\,\mu(ds)\right) = \gamma_{i_1,\dots,i_n}.$$

The integrals of the constant functions $f \equiv c$, $c \in \mathbb{R}$, are given by $\int_M f(s)\,\mu(ds) = c$. Thus, all integrals $\int_M f(s)\,\mu(ds)$, $f \in \mathcal{A}_M$, are known from the sequence of expectations $\int_M p_n(s)\,\mu(ds)$, $n \in \{1, 2, \dots\}$ and this sequence determines μ. □

Remark 3.13 *The proof of Proposition 3.10 also shows that μ is uniquely determined by all integrals $\int_M e_{i_1,\dots,i_n}^*(p_n(s))\,\mu(ds)$, $n \in \{1, 2, \dots\}$ and $(i_1, \dots, i_n) \in \{1, \dots, d\}^n$.*

3.3 Signatures of order books of futures markets

As an example, we have computed the (log-) signatures of order books of four futures markets. The data has kindly been provided by *Man Investments*. It comprises order book data for the entire trading week from 27 October 2008 to 31 October 2008 of the following four markets:

- EDC: Eurodollar STIR (short term interest rate) future contract for December 2009, traded on the Chicago Mercantile Exchange (CME)

- EUL: Euribor STIR future contract for December 2009, traded on the London International Financial Futures and Options Exchange (LIFFE)

- FTL: FTSE 100 stock index future contract for December 2008, traded on LIFFE
- SSL: Short-Sterling STIR future contract for December 2009, traded on LIFFE

In particular, we have looked at the 4-dimensional signal containing the best bid price, the best ask price and the respective depths, i.e. the volumes available in the order book to trade at these price levels. By linear interpolation, we turn each of these signals into a continuous path X with values in $\mathbb{R}^4$ which has bounded variation. It follows from Lemma 3.5 and Chen's identity (see for example [37], Theorem 2.9) that the signature of X over the entire trading week is given by

$$\mathbf{X}_{t_0,t_N} = \exp(X_{t_1} - X_{t_0}) \otimes \exp(X_{t_2} - X_{t_1}) \otimes \cdots \otimes \exp(X_{t_N} - X_{t_{N-1}}),$$

where $t_0, t_1, \ldots, t_N$ are the nodes of the piecewise-linear path X.
The logarithm of the signature $\mathbf{X}_{t_0,t_N}$ is an element of the space of Lie series $\mathcal{L}((\mathbb{R}^4))$ (see section 3.1). For integers $k \geq 1$, let $F_k \subset (\mathbb{R}^4)^{\otimes k}$ be the linear spaces that are defined recursively by $F_{k+1} = [\mathbb{R}^4, F_k]$, starting from $F_1 = \mathbb{R}^4$ (as in Definition 3.6). Let $\{e_1, e_2, e_3, e_4\}$ be the standard basis of $\mathbb{R}^4$ and let $\mathcal{B}$ be a basis of $\mathcal{L}((\mathbb{R}^4))$. Then, the log-signature $\log \mathbf{X}_{t_0,t_N}$ can be uniquely written as

$$\log \mathbf{X}_{t_0,t_N} = \sum_{\ell \in \mathcal{B}} L^{\ell}_{t_0,t_N} \ell,$$

where the coefficients $L^{\ell}_{t_0,t_N}$ are real numbers.
We have computed the log-signatures $\log \mathbf{X}_{t_0,t_N}$ for the four futures markets EDC, EUL, FTL and SSL. For the computations, we have written a C++ programme which uses a templated tool called *libalgebra* that offers all the functionalities for working with tensors and (truncated) Lie series. It has originally been developed by Djalil Chafai and Terry Lyons and it is part of the *CoRoPa* software package (see `http://coropa.sourceforge.net`).

The table 3.1 gives the log-signatures $\log X_{t_0,t_N}$ over the entire trading week up to order 3 for all four futures markets EDC, EUL, FTL and SSL and the tables 3.2 and 3.3 give the ratios of the coefficients $\frac{L^{(i_1,i_2)}_{t_0,t_N}}{|L^{(i_1)}_{t_0,t_N}||L^{(i_2)}_{t_0,t_N}|}$ and $\frac{L^{(i_1,i_2,i_3)}_{t_0,t_N}}{|L^{(i_1)}_{t_0,t_N}||L^{(i_2,i_3)}_{t_0,t_N}|}$ respectively.
The log-signatures show the differences of the futures markets EDC, EUL, FTL and SSL. The first order terms $L^{(i_1)}_{t_0,t_N}$ give the increments over the trading week. The second order terms $L^{(i_1,i_2)}_{t_0,t_N}$ determine the area that the 2-dimensional paths $\binom{X^{(i_1)}}{X^{(i_2)}}$ enclose. If there is such an area, the sign of the coefficients $L^{(i_1,i_2)}_{t_0,t_N}$ indicates which of the processes $X^{(i_1)}$ or $X^{(i_2)}$ has been following the other one. For example, we have $L^{(1,2)}_{t_0,t_N} > 0$ for the futures market EUL which indicates that the bid price process has been leading the ask price process in this market during this trading week. On the other hand, we have $L^{(1,2)}_{t_0,t_N} < 0$ for the futures market SSL and this indicates that the bid price process has been following the ask price process in this market during this trading week.
The ratios $\frac{L^{(i_1,i_2)}_{t_0,t_N}}{|L^{(i_1)}_{t_0,t_N}||L^{(i_2)}_{t_0,t_N}|}$ and $\frac{L^{(i_1,i_2,i_3)}_{t_0,t_N}}{|L^{(i_1)}_{t_0,t_N}||L^{(i_2,i_3)}_{t_0,t_N}|}$ are forms of normalisations of the coefficients of the log-signatures. They allow us to compare markets regardless of the different price levels and trading volumes. For example, they show differences between the futures markets FTL and SSL. We have $\frac{L^{(3,4)}_{t_0,t_N}}{|L^{(3)}_{t_0,t_N}||L^{(4)}_{t_0,t_N}|} = -234.617$ for the futures market FTL and $\frac{L^{(3,4)}_{t_0,t_N}}{|L^{(3)}_{t_0,t_N}||L^{(4)}_{t_0,t_N}|} = 28859.2$ for the futures market SSL. These ratios differ both in the order of magnitude and in the sign. We also see differences in the higher order terms. We have $\frac{L^{(3,3,4)}_{t_0,t_N}}{|L^{(3)}_{t_0,t_N}||L^{(3,4)}_{t_0,t_N}|} = -269.282$ for the futures market FTL and $\frac{L^{(3,3,4)}_{t_0,t_N}}{|L^{(3)}_{t_0,t_N}||L^{(3,4)}_{t_0,t_N}|} = 10.9849$ for the futures market SSL.
Moreover, the figures 3.1 – 3.8 depict the coefficient functions $(s,t) \mapsto L^{(i_1,i_2)}_{s,t}$ and $(s,t) \mapsto L^{(i_1,i_2,i_3)}_{s,t}$ for some of the basis elements $[e_{i_1}, e_{i_2}]$ and $[e_{i_1}, [e_{i_2}, e_{i_3}]]$ and some of the futures markets EDC, EUL, FTL and SSL as visualised examples. For these figures, we have rescaled time such that $t_n = n$, i.e. such that the orders are submitted or withdrawn at times $0, 1, \ldots, N$. Note that in some of the figures, the different trading days can be seen, whereas others just show the fluctuations in the markets.

basis element	EDC	EUL	FTL	SSL
e_1	-0.225	0.215	368	-0.07
e_2	-0.14	0.21	368.5	-0.075
e_3	4	83	-12	8
e_4	7	156	-10	3
$[e_1, e_2]$	-5.36374	9577.66	-2.008×10^6	-78.8753
$[e_1, e_3]$	3087.23	-5.42616×10^6	263467	-29849.9
$[e_1, e_4]$	-522.645	-8.90032×10^6	-40941.8	-23952.4
$[e_2, e_3]$	3090.77	-48379.6	-382567	-29740.8
$[e_2, e_4]$	-532.238	12512.2	-185264	-23823.3
$[e_3, e_4]$	6.59092×10^6	2.32456×10^7	-28154	692620
$[e_1, [e_1, e_2]]$	173.329	2.83737×10^6	2.80605×10^9	2543.54
$[e_1, [e_1, e_3]]$	-85012.3	-6.42141×10^{10}	-4.50874×10^7	956157
$[e_1, [e_1, e_4]]$	9577.04	-1.05077×10^{11}	-7.52809×10^7	775432
$[e_2, [e_1, e_2]]$	173.45	-611275	2.6226×10^9	2534.51
$[e_2, [e_1, e_3]]$	-170130	3.50342×10^8	-2.60536×10^8	1.90824×10^6
$[e_2, [e_1, e_4]]$	19024.8	5.74211×10^8	-1.21718×10^8	1.54861×10^6
$[e_2, [e_2, e_3]]$	-85118.4	-7.0303×10^6	-6.12474×10^7	952084
$[e_2, [e_2, e_4]]$	9447.72	-2.77182×10^6	-1.39305×10^7	773176
$[e_3, [e_1, e_2]]$	85031.1	-1.7731×10^8	1.6388×10^8	-960059
$[e_3, [e_1, e_3]]$	63304.7	-5.77108×10^8	1.16882×10^8	-2.09858×10^6
$[e_3, [e_1, e_4]]$	85210.7	-6.21376×10^8	-2.79801×10^7	-563736
$[e_3, [e_2, e_3]]$	65152.4	-1.05004×10^7	-1.6316×10^8	-2.09365×10^6
$[e_3, [e_2, e_4]]$	-243908	-6.67908×10^6	-7.15569×10^7	-548682
$[e_3, [e_3, e_4]]$	-7.40087×10^9	1.54565×10^{10}	-9.09765×10^7	6.08669×10^7
$[e_4, [e_1, e_2]]$	-9614.46	-2.88756×10^8	1.04246×10^8	-777505
$[e_4, [e_1, e_3]]$	-25183.4	-1.41516×10^9	1.52182×10^8	-526685
$[e_4, [e_1, e_4]]$	-211699	-1.79009×10^9	-6.48562×10^7	-1.26365×10^6
$[e_4, [e_2, e_3]]$	225689	-3.24278×10^6	-1.50248×10^8	-513700
$[e_4, [e_2, e_4]]$	-155682	2.24578×10^6	-8.50847×10^7	-1.25207×10^6
$[e_4, [e_3, e_4]]$	1.82583×10^{10}	4.60899×10^9	-1.61758×10^8	7.59536×10^7

Table 3.1: Log-signatures of futures markets

	EDC	EUL	FTL	SSL
$\frac{L^{(1,2)}_{t_0,t_N}}{\lvert L^{(1)}_{t_0,t_N}\rvert\lvert L^{(2)}_{t_0,t_N}\rvert}$	-170.277	212130	-14.8074	-15023.9
$\frac{L^{(1,3)}_{t_0,t_N}}{\lvert L^{(1)}_{t_0,t_N}\rvert\lvert L^{(3)}_{t_0,t_N}\rvert}$	-3430.26	-304072	-59.6619	53303.4
$\frac{L^{(1,4)}_{t_0,t_N}}{\lvert L^{(1)}_{t_0,t_N}\rvert\lvert L^{(4)}_{t_0,t_N}\rvert}$	331.838	265364	11.1255	114059
$\frac{L^{(2,3)}_{t_0,t_N}}{\lvert L^{(2)}_{t_0,t_N}\rvert\lvert L^{(3)}_{t_0,t_N}\rvert}$	-5519.23	-2775.65	86.5145	49568
$\frac{L^{(2,4)}_{t_0,t_N}}{\lvert L^{(2)}_{t_0,t_N}\rvert\lvert L^{(4)}_{t_0,t_N}\rvert}$	543.1	381.935	50.2752	105881
$\frac{L^{(3,4)}_{t_0,t_N}}{\lvert L^{(3)}_{t_0,t_N}\rvert\lvert L^{(4)}_{t_0,t_N}\rvert}$	235390	1795.3	-234.617	28859.2
$\frac{L^{(1,1,2)}_{t_0,t_N}}{\lvert L^{(1)}_{t_0,t_N}\rvert\lvert L^{(1,2)}_{t_0,t_N}\rvert}$	-143.622	1377.9	-3.79738	460.68
$\frac{L^{(1,1,3)}_{t_0,t_N}}{\lvert L^{(1)}_{t_0,t_N}\rvert\lvert L^{(1,3)}_{t_0,t_N}\rvert}$	122.386	55042.7	-0.46503	457.602
$\frac{L^{(1,1,4)}_{t_0,t_N}}{\lvert L^{(1)}_{t_0,t_N}\rvert\lvert L^{(1,4)}_{t_0,t_N}\rvert}$	81.4408	-54911.5	4.99655	462.484
$\frac{L^{(2,1,2)}_{t_0,t_N}}{\lvert L^{(2)}_{t_0,t_N}\rvert\lvert L^{(1,2)}_{t_0,t_N}\rvert}$	230.982	-303.919	-3.5443	428.442
$\frac{L^{(2,1,3)}_{t_0,t_N}}{\lvert L^{(2)}_{t_0,t_N}\rvert\lvert L^{(1,3)}_{t_0,t_N}\rvert}$	393.626	-307.454	-2.68351	852.371
$\frac{L^{(2,1,4)}_{t_0,t_N}}{\lvert L^{(2)}_{t_0,t_N}\rvert\lvert L^{(1,4)}_{t_0,t_N}\rvert}$	260.007	307.218	8.06771	862.049
$\frac{L^{(2,2,3)}_{t_0,t_N}}{\lvert L^{(2)}_{t_0,t_N}\rvert\lvert L^{(2,3)}_{t_0,t_N}\rvert}$	196.711	691.978	0.43445	426.836
$\frac{L^{(2,2,4)}_{t_0,t_N}}{\lvert L^{(2)}_{t_0,t_N}\rvert\lvert L^{(2,4)}_{t_0,t_N}\rvert}$	126.792	-1054.9	0.20405	432.728
$\frac{L^{(3,1,2)}_{t_0,t_N}}{\lvert L^{(2)}_{t_0,t_N}\rvert\lvert L^{(1,4)}_{t_0,t_N}\rvert}$	-3963.24	-223.047	6.80113	1521.48
$\frac{L^{(3,1,3)}_{t_0,t_N}}{\lvert L^{(3)}_{t_0,t_N}\rvert\lvert L^{(1,3)}_{t_0,t_N}\rvert}$	5.12633	1.2814	-36.9692	8.78805
$\frac{L^{(3,1,4)}_{t_0,t_N}}{\lvert L^{(3)}_{t_0,t_N}\rvert\lvert L^{(1,4)}_{t_0,t_N}\rvert}$	-40.7594	0.84114	-56.951	2.94196
$\frac{L^{(3,2,3)}_{t_0,t_N}}{\lvert L^{(3)}_{t_0,t_N}\rvert\lvert L^{(2,3)}_{t_0,t_N}\rvert}$	5.26992	582.887	-35.5406	8.79957
$\frac{L^{(3,2,4)}_{t_0,t_N}}{\lvert L^{(3)}_{t_0,t_N}\rvert\lvert L^{(2,4)}_{t_0,t_N}\rvert}$	114.567	-6.43139	-32.1869	2.87891
$\frac{L^{(3,3,4)}_{t_0,t_N}}{\lvert L^{(3)}_{t_0,t_N}\rvert\lvert L^{(3,4)}_{t_0,t_N}\rvert}$	280.722	8.0111	-269.282	10.9849

Table 3.2: Ratios of the coefficients of the log-signatures of futures markets

	EDC	EUL	FTL	SSL
$\frac{L^{(4,1,2)}_{t_0,t_N}}{\lvert L^{(4)}_{t_0,t_N}\rvert\lvert L^{(1,2)}_{t_0,t_N}\rvert}$	256.07	-193.262	5.19153	3285.8
$\frac{L^{(4,1,3)}_{t_0,t_N}}{\lvert L^{(4)}_{t_0,t_N}\rvert\lvert L^{(1,3)}_{t_0,t_N}\rvert}$	-1.16533	1.67182	-57.7613	5.88148
$\frac{L^{(4,1,4)}_{t_0,t_N}}{\lvert L^{(4)}_{t_0,t_N}\rvert\lvert L^{(1,4)}_{t_0,t_N}\rvert}$	57.8647	-1.28927	-158.412	17.5856
$\frac{L^{(4,2,3)}_{t_0,t_N}}{\lvert L^{(4)}_{t_0,t_N}\rvert\lvert L^{(2,3)}_{t_0,t_N}\rvert}$	10.4315	0.42967	-39.857	5.75752
$\frac{L^{(4,2,4)}_{t_0,t_N}}{\lvert L^{(4)}_{t_0,t_N}\rvert\lvert L^{(2,4)}_{t_0,t_N}\rvert}$	-41.7864	1.15056	-45.9262	17.5188
$\frac{L^{(4,3,4)}_{t_0,t_N}}{\lvert L^{(4)}_{t_0,t_N}\rvert\lvert L^{(3,4)}_{t_0,t_N}\rvert}$	395.746	1.27098	-574.547	36.5538

Table 3.3: Ratios of the coefficients of the log-signatures of futures markets (continued)

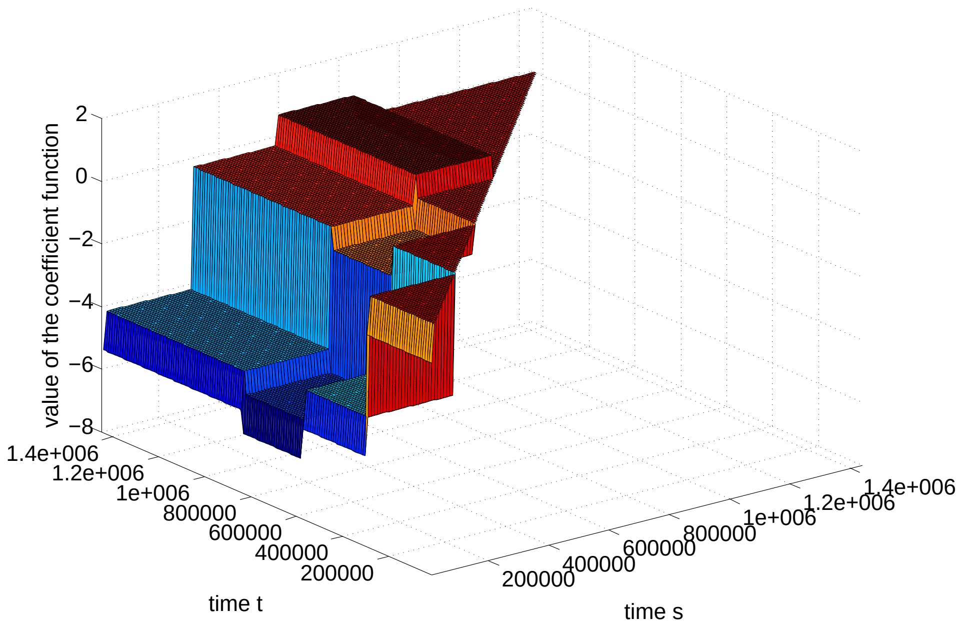

Figure 3.1: Coefficient function $(s,t) \mapsto L^{(1,2)}_{s,t}$ of the log-signature for the futures market EDC over the trading week

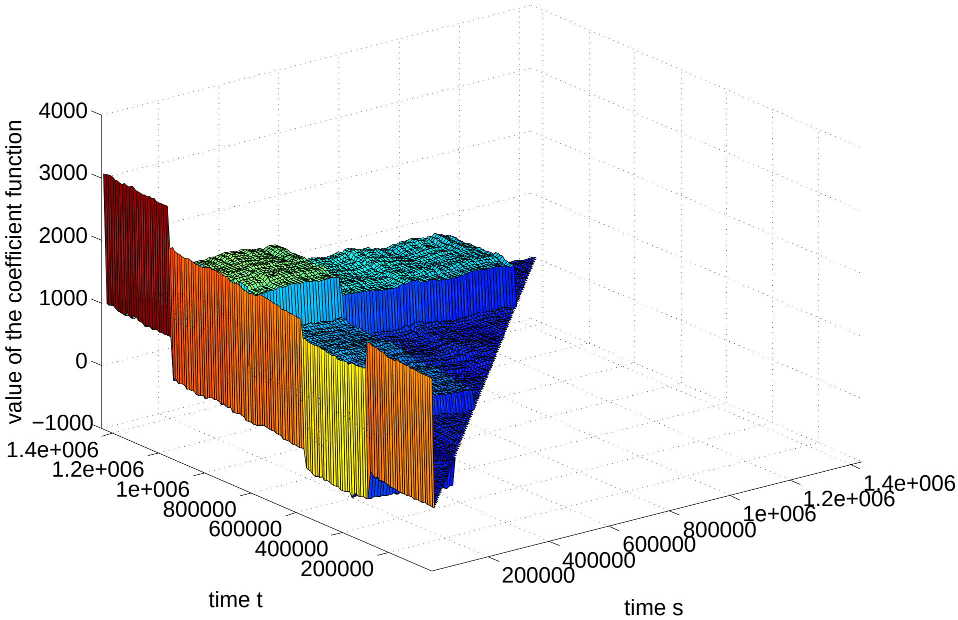

Figure 3.2: Coefficient function $(s,t) \mapsto L_{s,t}^{(2,3)}$ of the log-signature for the futures market EDC over the trading week

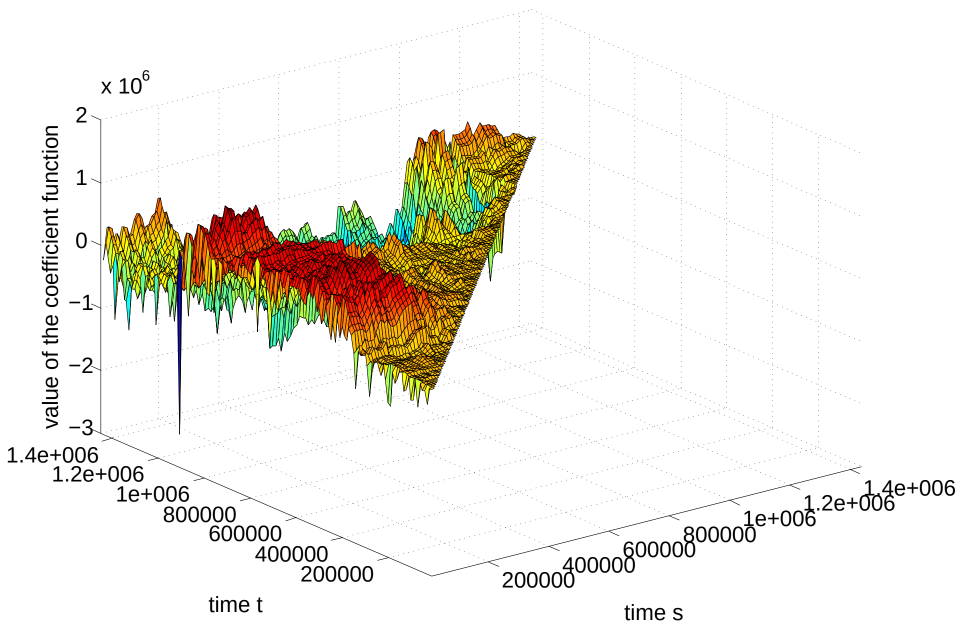

Figure 3.3: Coefficient function $(s,t) \mapsto L_{s,t}^{(4,2,3)}$ of the log-signature for the futures market EDC over the trading week

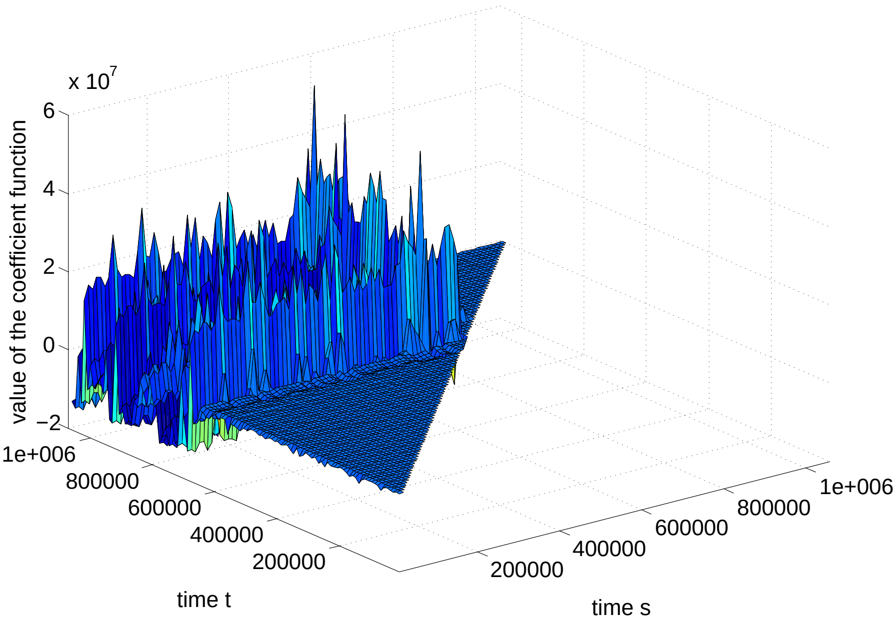

Figure 3.4: Coefficient function $(s,t) \mapsto L_{s,t}^{(4,2,3)}$ of the log-signature for the futures market EUL over the trading week

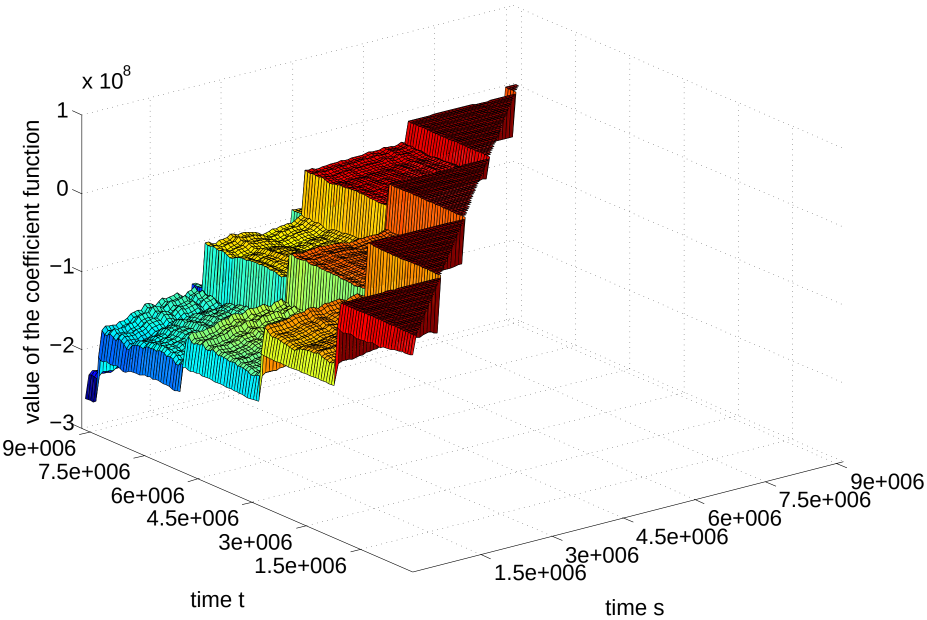

Figure 3.5: Coefficient function $(s,t) \mapsto L_{s,t}^{(2,1,3)}$ of the log-signature for the futures market FTL over the trading week

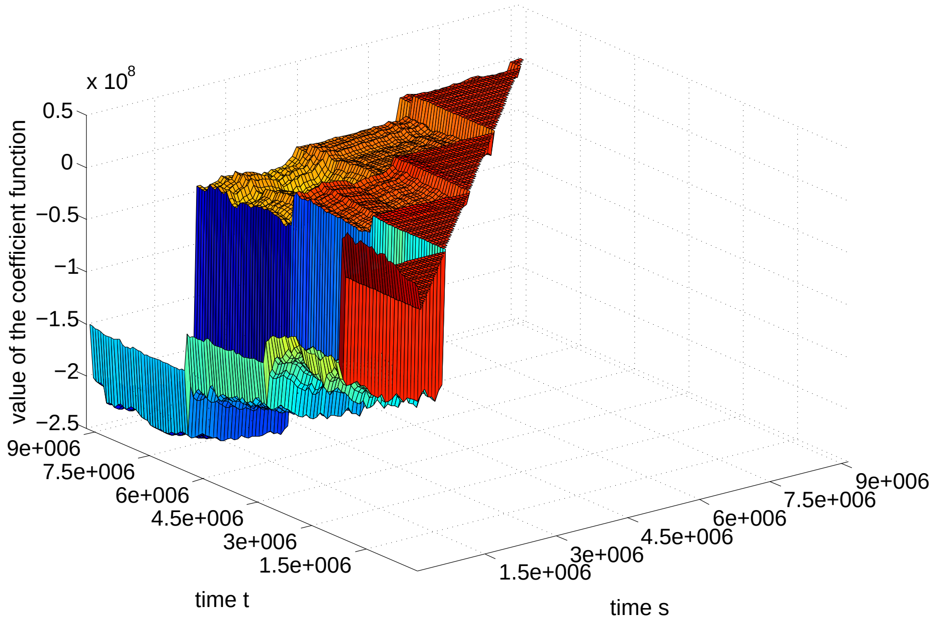

Figure 3.6: Coefficient function $(s,t) \mapsto L_{s,t}^{(4,2,3)}$ of the log-signature for the futures market FTL over the trading week

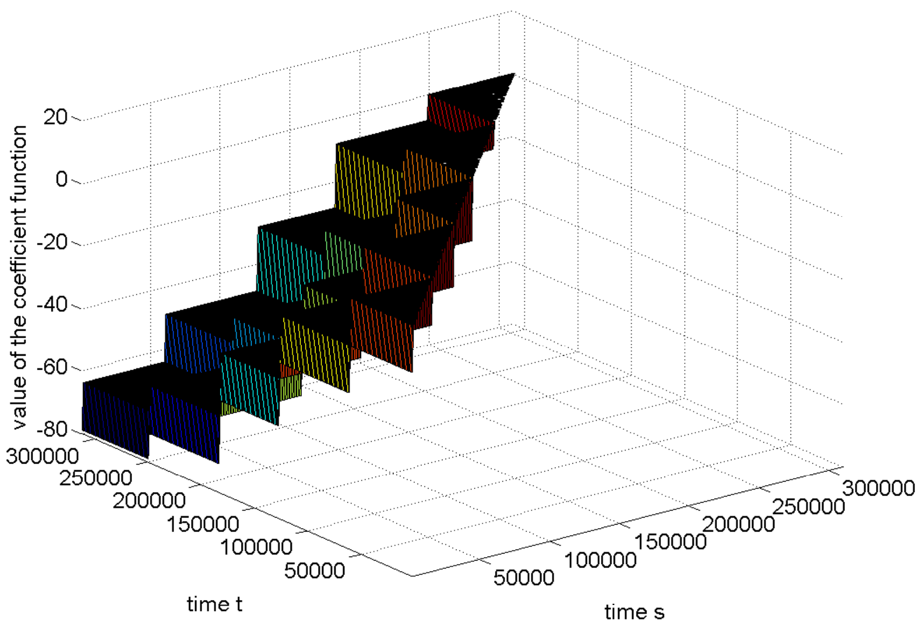

Figure 3.7: Coefficient function $(s,t) \mapsto L_{s,t}^{(1,2)}$ of the log-signature for the futures market SSL over the trading week

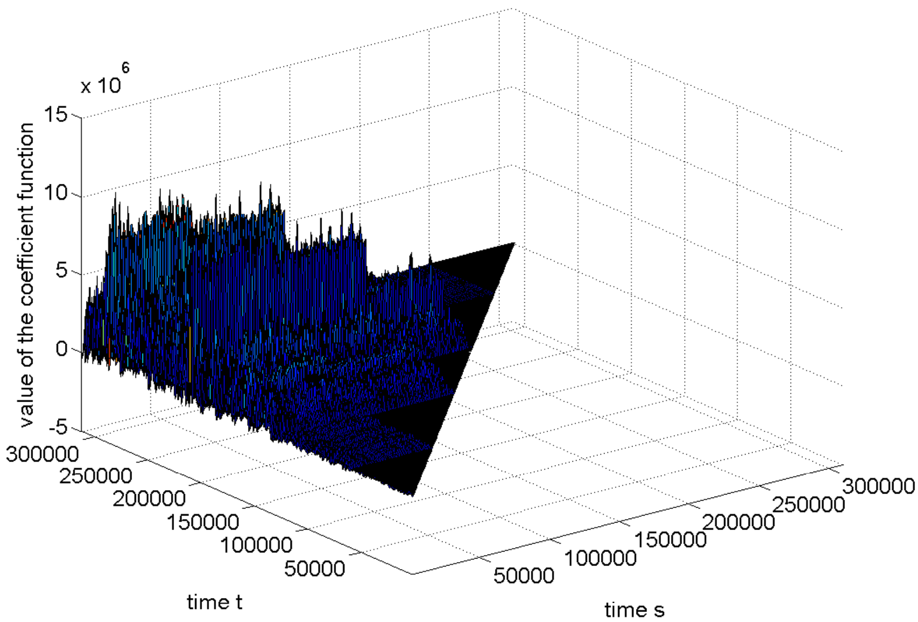

Figure 3.8: Coefficient function $(s,t) \mapsto L_{s,t}^{(4,2,3)}$ of the log-signature for the futures market SSL over the trading week

Chapter 4

Polynomial differential equations and signatures

In this chapter, we study polynomial differential equations driven by a path with bounded variation. Fix a finite time horizon $T > 0$ and consider the spaces $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$. Let $\gamma : [0,T] \to \mathbb{R}^{d_2}$ be a continuous path with bounded variation. For $(i,j) \in \{1,\dots,d_1\} \times \{1,\dots,d_2\}$, let $p_{ij} : \mathbb{R}^{d_1} \to \mathbb{R}$ be polynomials and let z be an element of $\mathbb{R}^{d_1}$.

Consider the differential equation

$$dy_t = P(y_t)\, d\gamma_t, \qquad y_0 = z, \tag{4.1}$$

where $P : \mathbb{R}^{d_1} \to M(d_1, d_2, \mathbb{R})$ is a continuous function, which takes values in the $d_1 \times d_2$-matrices, given by

$$P(y_t) = \begin{pmatrix} p_{11}(y_t) & \cdots & p_{1d_2}(y_t) \\ \vdots & \ddots & \vdots \\ p_{d_1 1}(y_t) & \cdots & p_{d_1 d_2}(y_t) \end{pmatrix}.$$

The path $y : [0,T] \to \mathbb{R}^{d_1}$ is a solution of (4.1) if it satisfies

$$y_t = z + \int_0^t P(y_s)\, d\gamma_s$$

for all $t \in [0,T]$. Note that polynomial differential equations may explode in finite time. But at least for small t, there exists a unique solution y_t of (4.1). It follows from Hambly and Lyons [28] that this solution y_t is uniquely determined by the signature $\Gamma_{0,t}$ of the signal γ and the initial value $y_0 = z$. Recall that the signature of a path is the full series of its definite iterated integrals (see chapter 3). In this chapter, we show how we can actually obtain y_t from the signature of γ and y_0 for small t. We work out a series expansion of the solution y_t in terms of the signature of γ and y_0. Furthermore, it turns out that the truncated signature of y is itself the solution of a quadratic differential equation driven by γ, no matter what the degree of the polynomial is, and again for small t, we give a serious expansion for the truncated signature $Y_{0,t}^{[M]}$ of y for arbitrary integers $M \geq 1$ in terms of the signature of γ and y_0.

For linear differential equations, a series expansion of the solution and its signature in terms of the signature of the driving signal and the initial value can be found in Lyons [36]. We generalise this result to polynomial differential equations. Polynomial differential equations are a rich class of differential equations which have various applications. For instance, a solution of a rough differential equation can be well approximated by the successive solution of polynomial differential equations (see Caruana [9] and Davie [17]). Our approach leads to a purely algebraic way to solve polynomial differential equations which fits into the framework of rough paths. Butcher [6] studies the algebraic properties of a class of methods to solve differential equations, including Runge-Kutta methods and Picard iteration. He provides power series expansions of the solutions of differential equations using trees (see [5] and [8]). Crouch and Grossman [15] apply this concept to polynomial differential equations and Lamnabhi-Lagarrigue and Lamnabhi [33] show how forced polynomial differential equations can be solved algebraically in one dimension. They work out a functional expansion of the solution using non-commutative generating power series.

4.1 Signatures and the shuffle product

The shuffle product is one of the key tools for this chapter. We have already introduced it in chapter 3 in detail, but given its importance, we briefly restate its definition and key property here.

Let $X : [0,T] \to \mathbb{R}^d$ be a continuous path with bounded variation. Recall that the signature of X over the interval $[s,t] \subset [0,T]$ is

$$X_{s,t} = (1, X^1_{s,t}, X^2_{s,t}, \dots) \in T((\mathbb{R}^d)),$$

where

$$X^n_{s,t} = \int\limits_{s<u_1<\dots<u_n<t} dX_{u_1} \otimes \dots \otimes dX_{u_n}$$

for each integer $n \geq 1$ (see also Definition 3.3).

Moreover, let $\left((\mathbb{R}^d)^{\otimes m}\right)^*$ be the dual space of $(\mathbb{R}^d)^{\otimes m}$, i.e. the space of bounded linear maps from $(\mathbb{R}^d)^{\otimes m}$ to $\mathbb{R}$.

The linear forms

$$e^*_{i_1,\dots,i_m} : (\mathbb{R}^d)^{\otimes m} \to \mathbb{R},$$

defined by

$$e^*_{i_1,\dots,i_m}(e_{j_1} \otimes \dots \otimes e_{j_m}) = \begin{cases} 1 & \text{if } j_1 = i_1, \dots, j_m = i_m \\ 0 & \text{otherwise,} \end{cases} \tag{4.2}$$

for $i_1, \ldots, i_m \in \{1, \ldots, d\}$, form a basis of $\left((\mathbb{R}^d)^{\otimes m}\right)^*$. A shuffle σ of $\{1, \ldots, m\}$ and $\{m+1, \ldots, m+n\}$ is a permutation of $\{1, \ldots, m+n\}$ such that

$$\sigma(1) < \cdots < \sigma(m)$$

and

$$\sigma(m+1) < \cdots < \sigma(m+n).$$

We denote the set of all shuffles of $\{1, \ldots, m\}$ and $\{m+1, \ldots, m+n\}$ by $\text{Shuffles}(m, n)$. The shuffle product of the linear forms $e^*_{i_1,\ldots,i_m}$ and $e^*_{j_1,\ldots,j_n}$ is defined in the following way (see also [37], section 2.2.3).

Definition 4.1 *The shuffle product of the linear forms*

$$e^*_{i_1,\ldots,i_m} : (\mathbb{R}^d)^{\otimes m} \to \mathbb{R}$$

and

$$e^*_{j_1,\ldots,j_n} : (\mathbb{R}^d)^{\otimes n} \to \mathbb{R}$$

is the linear form

$$e^*_{i_1,\ldots,i_m} \sqcup e^*_{j_1,\ldots,j_n} : (\mathbb{R}^d)^{\otimes(m+n)} \to \mathbb{R}$$

defined by

$$e^*_{i_1,\ldots,i_m} \sqcup e^*_{j_1,\ldots,j_n} = \sum_{\sigma \in \text{Shuffles}(m,n)} e^*_{k_{\sigma^{-1}(1)},\ldots,k_{\sigma^{-1}(m+n)}},$$

where $(k_1, \ldots, k_{m+n}) = (i_1, \ldots, i_m, j_1, \ldots, j_{m+n})$.

By linearity, the shuffle product can be extended to the linear forms on $T((\mathbb{R}^d))$ that are induced by elements of $\bigoplus_{l=0}^{\infty} \left((\mathbb{R}^d)^{\otimes l}\right)^*$.
Let g and h be such linear forms on $T((\mathbb{R}^d))$. Then,

$$g(X_{s,t})h(X_{s,t}) = (g \sqcup h)(X_{s,t}) \tag{4.3}$$

holds for the signature $X_{s,t}$ of every continuous path $X : [s,t] \to \mathbb{R}^d$ with bounded variation (see [37], Theorem 2.15 and equation (2.6)).

4.2 Polynomials as linear functionals on signatures

A polynomial $p : \mathbb{R}^d \to \mathbb{R}$ is a function of the form

$$p(x_1, \ldots, x_d) = \lambda_0 + \sum_{|\alpha|=1}^{N} \lambda_\alpha \, x_1^{\alpha_1} \ldots x_d^{\alpha_d} \tag{4.4}$$

where $N \in \mathbb{N}$, $\lambda_0 \in \mathbb{R}$ and $\lambda_\alpha \in \mathbb{R}$ for every multi-index α.

The multi-indeces $\alpha = (\alpha_1, \ldots, \alpha_d)$ are d-tupels of non-negative integers and we define $|\alpha| = \alpha_1 + \cdots + \alpha_d$. Set

$$i_\alpha = (\underbrace{1, \ldots, 1}_{\alpha_1 \text{ times}}, \underbrace{2, \ldots, 2}_{\alpha_2 \text{ times}}, \ldots, \underbrace{d, \ldots, d}_{\alpha_d \text{ times}})$$

and define $\mathcal{M}_\alpha \subset \{1, \ldots, d\}^{|\alpha|}$ as the set of all permutations of i_α. This means that $\mathcal{M}_\alpha$ is the set of all $|\alpha|$-tupels $(i_1, \ldots, i_{|\alpha|}) \in \{1, \ldots, d\}^{|\alpha|}$ which contain the integer j exactly α_j times for all $j = 1, \ldots, d$.

The polynomial p can be represented as a linear functional of signatures in the following way.

Theorem 4.2 *Let $p : \mathbb{R}^d \to \mathbb{R}$ be the polynomial (4.4). Then, there exists a linear map $\Phi : T((\mathbb{R}^d)) \to \mathbb{R}$ such that for every continuous path $X : [0,T] \to \mathbb{R}^d$ with bounded variation*

$$p(X_t - X_0) = \Phi(X_{0,t})$$

holds for all $t \in [0,T]$.

This linear map Φ *is given by*

$$\Phi = \lambda_0 \Psi_0 + \sum_{|\alpha|=1}^{N} \lambda_\alpha \Psi_\alpha, \tag{4.5}$$

where the linear map $\Psi_0 : T((\mathbb{R}^d)) \to \mathbb{R}$ *is defined by*

$$\Psi_0(\mathbf{a}) = a_0$$

and the linear maps $\Psi_\alpha : T((\mathbb{R}^d)) \to \mathbb{R}$ *are defined by*

$$\Psi_\alpha(\mathbf{a}) = \alpha_1! \dots \alpha_d! \sum_{(i_1,\dots,i_{|\alpha|}) \in \mathcal{M}_\alpha} a_{|\alpha|}^{(i_1,\dots,i_{|\alpha|})}$$

for every $\mathbf{a} = (a_0, a_1, \dots) \in T((\mathbb{R}^d))$, *where* $a_{|\alpha|}$ *can be written uniquely as in (3.1).*

Recall that $X_{0,t}^{(i_1,\dots,i_{|\alpha|})}$ is the coordinate iterated integral (3.2). Thus, the map Ψ_α, which corresponds to the monomial $x_1^{\alpha_1} \dots x_d^{\alpha_d}$, reads

$$\Psi_\alpha(X_{0,t}) = \alpha_1! \dots \alpha_d! \sum_{(i_1,\dots,i_{|\alpha|}) \in \mathcal{M}_\alpha} X_{0,t}^{(i_1,\dots,i_{|\alpha|})}.$$

Proof of Theorem 4.2: We prove Theorem 4.2 using shuffle products. Let $X : [0,T] \to \mathbb{R}^d$ be a continuous path with bounded variation and let $t \in [0,T]$. By the definition of the signature, we have

$$\Psi_0(X_{0,t}) = 1.$$

Furthermore, we have

$$(X_t^i - X_0^i)^{\alpha_i} = (X_{0,t}^{(i)})^{\alpha_i} = (e_i^*(X_{0,t}^1))^{\alpha_i} = (\underbrace{e_i^* \sqcup \cdots \sqcup e_i^*}_{\alpha_i\text{-times}})(X_{0,t}^{\alpha_i}) = \alpha_i! X_{0,t}^{(i,\dots,i)}$$

for $i = 1, \dots, d$ and every multi-index α, and thus

$$\begin{aligned}(X_t^1 - X_0^1)^{\alpha_1} \dots (X_t^d - X_0^d)^{\alpha_d} &= (X_{0,t}^{(1)})^{\alpha_1} \dots (X_{0,t}^{(d)})^{\alpha_d} \\ &= \prod_{i=1}^{d} \alpha_i! X_{0,t}^{(i,\dots,i)} \\ &= \left(\bigsqcup_{i=1}^{d} (\alpha_i!\, e^*_{i,\dots,i}) \right) (X_{0,t}^{|\alpha|}) \\ &= \alpha_1! \dots \alpha_d! \sum_{(i_1,\dots,i_{|\alpha|}) \in \mathcal{M}_\alpha} X_{0,t}^{(i_1,\dots,i_{|\alpha|})} \\ &= \Psi_\alpha(X_{0,t}),\end{aligned}$$

where the linear forms e_i^* and $e^*_{i,\dots,i}$ are defined as in (4.2). Now, linearity completes the proof. □

Remark 4.3 *For Theorem 4.2, we do not need the full signature. It is sufficient to work with truncated signatures $X_{0,t}^{[k]} = (1, X_{0,t}^1, \dots, X_{0,t}^k)$ and on the space of truncated tensors $T(\mathbb{R}^{d_1})^{(k)} = \bigoplus_{i=0}^{k} (\mathbb{R}^d)^{\otimes i}$ if k is greater than the degree of p.*

4.3 Picard iteration

Picard iteration is a way to construct a solution of a differential equation and to prove its uniqueness (see for example Arnold [1] or Coddington and Levinson [13]). Let V and W be finite dimensional Banach spaces with norms $\|\cdot\|_V$ and $\|\cdot\|_W$ respectively. Let $L(V, W)$ denote the set of bounded linear maps from V to W, equipped with the operator norm $\|\cdot\|_{L(V,W)}$ which is defined by

$$\|A\|_{L(V,W)} = \sup_{\|v\|_V \leq 1} \|A(v)\|_W.$$

Let $X : [0,T] \to V$ be a continuous path with bounded variation. Consider the differential equation

$$dY_t = f(Y_t)\,dX_t, \quad Y_0 = y, \tag{4.6}$$

where $f : W \to L(V,W)$ is locally Lipschitz continuous. Recall that f is locally Lipschitz continuous if for every $w \in W$ there exists a neighbourhood U of w such that there exists a constant $L_U > 0$ such that

$$\|f(u) - f(\tilde{u})\|_{L(V,W)} \leq L_U \|u - \tilde{u}\|_W$$

holds for all $u, \tilde{u} \in U$. It follows from the local Lipschitz continuity that for every compact set $K \subset W$, there exists a constant $L_K > 0$ such that

$$\|f(w) - f(\tilde{w})\|_{L(V,W)} \leq L_K \|w - \tilde{w}\|_W$$

holds for all $w, \tilde{w} \in K$. The constant L_K is called a Lipschitz constant for f on K.

Definition 4.4 *Let $X : [s,t] \to V$ be a continuous path with bounded variation. The 1-variation of X on the interval $[s,t]$ is defined by*

$$\|X\|_{1,[s,t]} = \sup_{\mathcal{D}\subset[s,t]} \sum_{i=1}^{n} \|X_{t_i} - X_{t_{i-1}}\|_V,$$

where $\mathcal{D} = \{t_0, t_1, \ldots, t_n\}$ is a partition $s = t_0 < t_1 < \cdots < t_n = t$ of $[s,t]$.

Now, we prove a local version of Picard's Theorem that is adapted to our setting and that gives us the local solution of the differential equation (4.6). Note that we choose the constants in such a way that we can later construct the maximal interval of existence of the solution.

Theorem 4.5 *Consider the differential equation (4.6), i.e.*

$$dY_t = f(Y_t)\,dX_t, \quad Y_0 = y,$$

where $f : W \to L(V,W)$ is locally Lipschitz continuous and $X : [0,T] \to V$ is a continuous path with bounded variation.

(i) Set

$$M_1 = \max_{w \in \bar{B}_1(y)} \|f(w)\|_{L(V,W)},$$

where $\bar{B}_1(y)$ is the closed ball $\bar{B}_1(y) = \{u \in W : \|u-y\|_W \le 1\}$. Fix $t_1 \in [0,T]$ so that

$$\|X\|_{1,[0,t_1]} = \frac{1}{M_1}$$

if $M_1^{-1} < \|X\|_{1,[0,T]}$. Otherwise, set $t_1 = T$.
Then, there exists a unique solution $Y : [0,t_1] \to W$ of the differential equation (4.6) on $[0,t_1]$.

(ii) Define paths $Y^n : [0,t_1] \to W$ recursively by

$$Y_t^{n+1} = y + \int_0^t f(Y_s^n)\, dX_s$$

starting from

$$Y_t^0 = y$$

for every $t \in [0,t_1]$. Let $L_{\bar{B}_1(y)}$ be a Lipschitz constant for f on $\bar{B}_1(y)$. Then,

$$\|Y_t^n - Y_t\|_W \le \frac{(L_{\bar{B}_1(y)}/M_1)^n}{(n+1)!} e^{L_{\bar{B}_1(y)}/M_1}$$

holds for every $n \in \mathbb{N}$ and every $t \in [0,t_1]$. Furthermore, Y^n converges uniformly to Y on $[0,t_1]$ in the sense that

$$\lim_{n\to\infty} \sup_{t \in [0,t_1]} \|Y_t^n - Y_t\|_W = 0$$

and Y is continuous and has bounded variation.

The proof requires the following Lemma.

Lemma 4.6 *Let $h : [0,T] \to L(V,W)$ be a continuous function and let $X : [0,T] \to V$ be a continuous path with bounded variation. Then,*

$$\left\| \int_0^t h_s \, dX_s \right\|_W \leq \int_0^t \|h_s\|_{L(V,W)} \, d\|X\|_{1,[0,s]}$$

holds for every $t \in [0,T]$.

Proof: Fix a time $t \in [0,T]$. Let $0 = s_0 < s_1 < \cdots < s_n = t$ be a partition $\mathcal{D} = \{s_0, s_1, \ldots, s_n\}$ of the interval $[0,t]$ and let $|\mathcal{D}| = \max\{|s_i - s_{i-1}| : i = 1, \ldots, n\}$ denote the mesh of the partition $\mathcal{D}$. Then, we have

$$\begin{aligned}
\left\| \int_0^t h_s \, dX_s \right\|_W &= \lim_{|\mathcal{D}| \to 0} \left\| \sum_{i=0}^{n-1} h_{s_i}(X_{s_{i+1}} - X_{s_i}) \right\|_W \\
&\leq \lim_{|\mathcal{D}| \to 0} \sum_{i=0}^{n-1} \|h_{s_i}\|_{L(V,W)} \|X_{s_{i+1}} - X_{s_i}\|_V \\
&\leq \lim_{|\mathcal{D}| \to 0} \sum_{i=0}^{n-1} \|h_{s_i}\|_{L(V,W)} \|X\|_{1,[s_i,s_{i+1}]} \\
&= \lim_{|\mathcal{D}| \to 0} \sum_{i=0}^{n-1} \|h_{s_i}\|_{L(V,W)} \left(\|X\|_{1,[0,s_{i+1}]} - \|X\|_{1,[0,s_i]} \right) \\
&= \int_0^t \|h_s\|_{L(V,W)} \, d\|X\|_{1,[0,s]}.
\end{aligned}$$

□

Now we can prove Theorem 4.5.

Proof of Theorem 4.5: At first, we show that

$$Y_t^n \in \bar{B}_1(y) \tag{4.7}$$

for all $t \in [0, t_1]$ and all $n \in \mathbb{N}$ via induction over n. For $n = 0$, we have

$$Y_t^0 = y \in \bar{B}_1(y)$$

for all $t \in [0, t_1]$. Assume that (4.7) holds for n. Then,

$$\begin{aligned}
\|Y_t^{n+1} - y\|_W &= \left\| \int_0^t f(Y_s^n)\, dX_s \right\|_W \leq \|X\|_{1,[0,t]} \sup_{s \in [0,t]} \|f(Y_s^n)\|_{L(V,W)} \\
&\leq \|X\|_{1,[0,t]} M_1 \leq 1
\end{aligned}$$

holds for all $t \in [0, t_1]$ and hence we have $Y_t^{n+1} \in \bar{B}_1(y)$ for all $t \in [0, t_1]$, which completes the induction.

Furthermore, we show that

$$\|Y_t^{n+1} - Y_t^n\|_W \leq \frac{M_1}{L_{\bar{B}_1(y)}} \frac{(L_{\bar{B}_1(y)} \|X\|_{1,[0,t]})^{n+1}}{(n+1)!} \tag{4.8}$$

for every $t \in [0, t_1]$ and every $n \in \mathbb{N}$ via induction over n. For $n = 0$, we have

$$\|Y_t^1 - Y_t^0\|_W = \|Y_t^1 - y\|_W \leq M_1 \|X\|_{1,[0,t]}$$

for every $t \in [0, t_1]$ by the same argument as above. Assume that (4.8) holds for n. Then, it follows from Lemma 4.6 and the local Lipschitz continuity of f that

$$\begin{aligned}
\|Y_t^{n+2} - Y_t^{n+1}\|_W &= \left\| y + \int_0^t f(Y_s^{n+1})\, dX_s - y - \int_0^t f(Y_s^n)\, dX_s \right\|_W \\
&= \left\| \int_0^t (f(Y_s^{n+1}) - f(Y_s^n))\, dX_s \right\|_W \\
&\leq \int_0^t \|f(Y_s^{n+1}) - f(Y_s^n)\|_{L(V,W)}\, d\|X\|_{1,[0,s]} \\
&\leq \int_0^t L_{\bar{B}_1(y)} \|Y_s^{n+1} - Y_s^n\|_W\, d\|X\|_{1,[0,s]} \\
&\leq L_{\bar{B}_1(y)} \int_0^t \frac{M_1}{L_{\bar{B}_1(y)}} \frac{(L_{\bar{B}_1(y)} \|X\|_{1,[0,s]})^{n+1}}{(n+1)!}\, d\|X\|_{1,[0,s]} \\
&= M_1 \frac{L_{\bar{B}_1(y)}^{n+1}}{(n+1)!} \int_0^{\|X\|_{1,[0,t]}} u^{n+1}\, du \\
&= \frac{M_1}{L_{\bar{B}_1(y)}} \frac{(L_{\bar{B}_1(y)} \|X\|_{1,[0,t]})^{n+2}}{(n+2)!}
\end{aligned}$$

holds for every $t \in [0, t_1]$ by the induction hypothesis. This completes the induction and thus shows (4.8).

It follows from (4.8) that

$$\begin{aligned}
\|Y_t^{n+k} - Y_t^n\|_W &\leq \sum_{i=0}^{k-1} \|Y_t^{n+i+1} - Y_t^{n+i}\|_W \leq \sum_{i=0}^{k-1} \frac{M_1}{L_{\bar{B}_1(y)}} \frac{(L_{\bar{B}_1(y)}\|X\|_{1,[0,t]})^{n+i+1}}{(n+i+1)!} \\
&\leq \frac{M_1}{L_{\bar{B}_1(y)}} \frac{(L_{\bar{B}_1(y)}\|X\|_{1,[0,t]})^{n+1}}{(n+1)!} \sum_{i=0}^{k-1} \frac{(L_{\bar{B}_1(y)}\|X\|_{1,[0,t]})^{i}}{i!} \\
&\leq \frac{M_1}{L_{\bar{B}_1(y)}} \frac{(L_{\bar{B}_1(y)}\|X\|_{1,[0,t]})^{n+1}}{(n+1)!} \sum_{i=0}^{\infty} \frac{(L_{\bar{B}_1(y)}\|X\|_{1,[0,t]})^{i}}{i!} \\
&= \frac{M_1}{L_{\bar{B}_1(y)}} \frac{(L_{\bar{B}_1(y)}\|X\|_{1,[0,t]})^{n+1}}{(n+1)!} e^{L_{\bar{B}_1(y)}\|X\|_{1,[0,t]}} \\
&\leq \frac{M_1}{L_{\bar{B}_1(y)}} \frac{(L_{\bar{B}_1(y)}\|X\|_{1,[0,t_1]})^{n+1}}{(n+1)!} e^{L_{\bar{B}_1(y)}\|X\|_{1,[0,t_1]}} \\
&\leq \frac{(L_{\bar{B}_1(y)}/M_1)^{n}}{(n+1)!} e^{L_{\bar{B}_1(y)}/M_1}
\end{aligned}$$

holds for all $n \in \mathbb{N}$, all integers $k \geq 1$ and all $t \in [0, t_1]$ since $\|X\|_{1,[0,t_1]} \leq \frac{1}{M_1}$ by definition. Therefore, $(Y_t^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in the Banach space W and hence converges pointwise. Thus, the limit $\lim_{n\to\infty} Y_t^n$ exists for every $t \in [0, t_1]$. Define the function $Y : [0, t_1] \to W$ by

$$Y_t = \lim_{n\to\infty} Y_t^n \in \bar{B}_1(y).$$

Then, we have

$$\|Y_t - Y_t^n\|_W = \lim_{k\to\infty} \|Y_t^{n+k} - Y_t^n\|_W \leq \frac{(L_{\bar{B}_1(y)}/M_1)^{n}}{(n+1)!} e^{L_{\bar{B}_1(y)}/M_1}$$

for every $t \in [0, t_1]$ and hence

$$\lim_{n\to\infty} \sup_{t\in[0,t_1]} \|Y_t - Y_t^n\|_W = 0.$$

Thus, Y^n converges uniformly to Y on $[0, t_1]$. It follows from the uniform convergence that Y is continuous as all the paths Y^n are continuous by definition.

Moreover, Y satisfies

$$Y_t = y + \int_0^t f(Y_s)\, dX_s \tag{4.9}$$

for all $t \in [0, t_1]$ by construction and therefore, Y is a solution of the differential equation (4.6) on $[0, t_1]$. It also follows from (4.9) that the path $Y : [0, t_1] \to W$ has bounded variation since $Y_s \in \bar{B}_1(y)$ for all $s \in [0, t_1]$. Hence, the function $f \circ Y : [0, t_1] \to L(V, W)$, $s \mapsto f(Y_s)$, is bounded by M_1.

Now, it just remains to show the uniqueness of the solution. Assume that there is another solution $\tilde{Y}$ of the differential equation (4.6) on $[0, t_1]$. Then, $\tilde{Y} : [0, t_1] \to W$ satisfies

$$\tilde{Y}_t = y + \int_0^t f(\tilde{Y}_s)\, dX_s$$

for every $t \in [0, t_1]$. Hence, the path $\tilde{Y}$ is continuous. Let $\bar{B}_2(y)$ be the closed ball $\bar{B}_2(y) = \{u \in W : \|u - y\|_W \leq 2\}$. The path $\tilde{Y}$ starts from $\tilde{Y}_0 = y \in \bar{B}_2(y)$. Set $\tau_1 = \inf\{t > 0 : \|\tilde{Y}_t - y\|_W = 2\}$ if there exists $t \in [0, t_1]$ such that $\|\tilde{Y}_t - y\|_W = 2$. If $\|\tilde{Y}_t - y\|_W < 2$ for all $t \in [0, t_1]$, set $\tau_1 = t_1$. Then, we have $\tilde{Y}_t \in \bar{B}_2(y)$ and $Y_t \in \bar{B}_1(y) \subset \bar{B}_2(y)$ for all $t \in [0, \tau_1]$. Let $L_{\bar{B}_2(y)}$ be a Lipschitz constant for f on $\bar{B}_2(y)$. If $\|X\|_{1,[0,\tau_1]} > (2L_{\bar{B}_2(y)})^{-1}$, fix $u_1 \in [0, \tau_1]$ so that $\|X\|_{1,[0,u_1]} = (2L_{\bar{B}_2(y)})^{-1}$.

Otherwise, set $u_1 = \tau_1$. Then,

$$\begin{aligned}
\sup_{t\in[0,u_1]} \|Y_t - \tilde{Y}_t\|_W &= \sup_{t\in[0,u_1]} \left\| y + \int_0^t f(Y_s)\, dX_s - y - \int_0^t f(\tilde{Y}_s)\, dX_s \right\|_W \\
&= \sup_{t\in[0,u_1]} \left\| \int_0^t (f(Y_s) - f(\tilde{Y}_s))\, dX_s \right\|_W \\
&\leq \sup_{t\in[0,u_1]} \left(\|X\|_{1,[0,t]} \sup_{s\in[0,t]} \|f(Y_s) - f(\tilde{Y}_s)\|_{L(V,W)} \right) \\
&\leq \sup_{t\in[0,u_1]} \left(\|X\|_{1,[0,t]} \sup_{s\in[0,t]} L_{\bar{B}_2(y)} \|Y_s - \tilde{Y}_s\|_W \right) \\
&\leq \frac{1}{2} \sup_{t\in[0,u_1]} \|Y_t - \tilde{Y}_t\|_W
\end{aligned}$$

holds because of the Lipschitz continuity of f. Hence,

$$\sup_{t\in[0,u_1]} \|Y_t - \tilde{Y}_t\|_W = 0$$

and $Y_t = \tilde{Y}_t$ holds for all $t \in [0, u_1]$.

If $\|X\|_{1,[0,\tau_1]} > (2L_{\bar{B}_2(y)})^{-1}$, and therefore $u_1 < \tau_1$, we need to iterate the above argument to show that $Y_t = \tilde{Y}_t$ holds for all $t \in [0, \tau_1]$. So, if $\|X\|_{1,[0,\tau_1]} > (2L_{\bar{B}_2(y)})^{-1}$, let m be the smallest integer such that

$$\frac{\|X\|_{1,[0,\tau_1]}}{m} \leq \frac{1}{2L_{\bar{B}_2(y)}}.$$

Now, define $u_2, \dots, u_{m-1} \in [0, \tau_1]$ recursively, starting from u_1 as defined above, so that $\|X\|_{1,[u_{k-1},u_k]} = (2L_{\bar{B}_2(y)})^{-1}$ for $k = 2, \dots, m-1$. Set $u_0 = 0$ and $u_m = \tau_1$. Then, we have constructed a partition $0 = u_0 < u_1 < \cdots < u_m = \tau_1$ of $[0, \tau_1]$ such that $\|X\|_{1,[u_k,u_{k+1}]} \leq (2L_{\bar{B}_2(y)})^{-1}$ for all $k = 0, 1, \dots, m-1$. Now, we show by induction that

$$\sup_{t\in[u_k,u_{k+1}]} \|Y_t - \tilde{Y}_t\|_W = 0 \tag{4.10}$$

holds for all $k = 0, 1, \dots, m-1$. We have already shown above that

$$\sup_{t \in [u_0, u_1]} \|Y_t - \tilde{Y}_t\|_W = 0.$$

Assume that the induction hypothesis holds for some $k \in \{0, 1, \dots, m-2\}$. In particular, this implies $Y_{u_{k+1}} = \tilde{Y}_{u_{k+1}}$. Then,

$$\begin{aligned}
&\sup_{t \in [u_{k+1}, u_{k+2}]} \|Y_t - \tilde{Y}_t\|_W \\
= &\sup_{t \in [u_{k+1}, u_{k+2}]} \left\| Y_{u_{k+1}} + \int_{u_{k+1}}^t f(Y_s)\, dX_s - \tilde{Y}_{u_{k+1}} - \int_{u_{k+1}}^t f(\tilde{Y}_s)\, dX_s \right\|_W \\
= &\sup_{t \in [u_{k+1}, u_{k+2}]} \left\| \int_{u_{k+1}}^t (f(Y_s) - f(\tilde{Y}_s))\, dX_s \right\|_W \\
\leq &\sup_{t \in [u_{k+1}, u_{k+2}]} \left(\|X\|_{1,[u_{k+1},t]} \sup_{s \in [u_{k+1}, t]} \|f(Y_s) - f(\tilde{Y}_s)\|_{L(V,W)} \right) \\
\leq &\sup_{t \in [u_{k+1}, u_{k+2}]} \left(\|X\|_{1,[u_{k+1},t]} \sup_{s \in [u_{k+1}, t]} L_{\bar{B}_2(y)} \|Y_s - \tilde{Y}_s\|_W \right) \\
\leq &\frac{1}{2} \sup_{t \in [u_{k+1}, u_{k+2}]} \|Y_t - \tilde{Y}_t\|_W
\end{aligned}$$

holds because of the Lipschitz continuity of f and hence

$$\sup_{t \in [u_{k+1}, u_{k+2}]} \|Y_t - \tilde{Y}_t\|_W = 0,$$

which completes the induction.

Now, it follows immediately from (4.10) that $Y_t = \tilde{Y}_t$ for every $t \in [0, \tau_1]$.

Finally, we show by contradiction that $\tau_1 = t_1$. Assume that $\tau_1 < t_1$. Then,

$$\|\tilde{Y}_{\tau_1} - y\|_W = 2$$

and

$$Y_{\tau_1} = \tilde{Y}_{\tau_1}$$

holds as shown above. Thus,

$$\|Y_{\tau_1} - y\|_W = 2,$$

but this is a contradiction to the fact that $Y_t \in \bar{B}_1(y)$ for all $t \in [0, t_1]$. Therefore, $\tau_1 = t_1$ and $Y_t = \tilde{Y}_t$ holds for all $t \in [0, t]$.
Thus, Y is the unique solution of the differential equation (4.6) on $[0, t_1]$. □

We can extend the local solution $Y : [0, t_1] \to W$ of the differential equation (4.6) to bigger time intervals $[0, t_n]$ by successively applying the construction from Theorem 4.5 to small time intervals $[t_n, t_{n+1}]$ and then concatenating the paths.
Given the local solution on $[0, t_1]$ from Theorem 4.5, we know, in particular, Y_{t_1}. So now, we can solve the initial value problem

$$dY_t = f(Y_t)\, dX_t \tag{4.11}$$

starting from Y_{t_1} at time t_1. In the same way as in Theorem 4.5, we obtain $t_2 \in [t_1, T]$ and a unique solution of (4.11) on $[t_1, t_2]$. We can concatenate these solutions and together, they form a unique solution of the initial value problem (4.6) on $[0, t_2]$.
More precisely, we construct a sequence of times $(t_n)_{n\in\mathbb{N}} \subset [0, T]$ and a unique solution of (4.6) on every interval $[0, t_n]$ in the following way.

Lemma 4.7 *Consider again the differential equation (4.6), i.e.*

$$dY_t = f(Y_t)\, dX_t, \quad Y_0 = y,$$

where $f : W \to L(V, W)$ *is locally Lipschitz continuous and* $X : [0, T] \to V$ *is a continuous path with bounded variation. Define a sequence of times* $(t_n)_{n\in\mathbb{N}} \subset [0, T]$ *and construct a sequence of paths* $Y : [0, t_n] \to W$ *recursively: Start with the time* t_1

and the local solution $Y : [0, t_1] \to W$ from Theorem 4.5. Given t_n and the path $Y : [0, t_n] \to W$, we define

$$M_{n+1} = \max_{w \in \bar{B}_1(Y_{t_n})} \|f(w)\|_{L(V,W)},$$

where $\bar{B}_1(Y_{t_n})$ is the closed ball $\bar{B}_1(Y_{t_n}) = \{u \in W : \|u - Y_{t_n}\|_W \leq 1\}$. Fix the time $t_{n+1} \in [t_n, T]$ so that

$$\|X\|_{1,[t_n,t_{n+1}]} = \frac{1}{M_{n+1}}$$

if $M_{n+1}^{-1} < \|X\|_{1,[t_n,T]}$. Otherwise, set $t_{n+1} = T$.
Then, there exists a unique solution $Y : [t_n, t_{n+1}] \to W$ of the initial value problem

$$dY_t = f(Y_t)\, dX_t$$

starting from Y_{t_n} at time t_n. Define the path $Y : [0, t_{n+1}] \to W$ as the concatenation of $Y : [0, t_n] \to W$ and this solution $Y : [t_n, t_{n+1}] \to W$.
Then, the path $Y : [0, t_n] \to W$ is the unique solution of the differential equation (4.6) on $[0, t_n]$ for every $n \in \mathbb{N}$ and it is continuous and has bounded variation.

Proof: We prove Lemma 4.7 by induction over n. For $n = 1$, Theorem 4.5 gives us t_1 and the unique solution $Y : [0, t_1] \to W$ of (4.6) on $[0, t_1]$ which is continuous and has bounded variation.
Given t_n and $Y : [0, t_n] \to W$, we know Y_{t_n} and it follows from Theorem 4.5 by reparametrisation that there exists a unique solution $Y : [t_n, t_{n+1}] \to W$ of the initial value problem

$$dY_t = f(Y_t)\, dX_t$$

starting from Y_{t_n} at time t_n, which is continuous and has bounded variation. Now, we can concatenate the unique solutions Y on $[0, t_n]$ and $[t_n, t_{n+1}]$ as they have the same value Y_{t_n} at time t_n. This path $Y : [0, t_{n+1}] \to W$ is the unique solution of the differential equation (4.6) on $[0, t_{n+1}]$. It is continuous and has bounded variation. This completes the induction and hence the proof. □

Note that, if the sequence (t_n) reaches T in finitely many steps, i.e. there is $n^* \in \mathbb{N}$ such that $t_{n^*} = T$, then the times t_{n^*+i}, $i \in \mathbb{N}$, coincide with T and the solution on $[t_{n^*+i}, t_{n^*+i+1}] = \{T\}$ is trivially $Y_{t_{n^*}} = Y_T$. But it is not guaranteed that the sequence (t_n) reaches T in finitely many steps. In fact, the solution of the differential equation (4.6) may explode in finite time.

Definition 4.8 *We say that a solution Y of the differential equation (4.6) explodes in finite time if there exists $\tau^* \in [0, T]$ such that the solution $Y_t \in W$ exists for all $t \in [0, \tau^*)$ and that, for every $R > 0$, there exists $t_R \in [0, \tau^*)$ such that $\|Y_{t_R}\|_W > R$.*

Note that the time of explosion is

$$\lim_{R\to\infty} \inf\{t > 0 : \|Y_t\|_W \geq R\} = \tau^*.$$

This follows from the fact that every solution of (4.6) has to be continuous and therefore attains its maximum on every compact time interval $[0, t]$.

Now, reconsider the sequence (t_n) from Lemma 4.7. If the times t_n reach T in a finite number of steps, i.e. there exists $n^* \in \mathbb{N}$ such that $t_{n^*} = T$, the construction from Lemma 4.7 gives us the unique solution Y of the differential equation (4.6) on the entire time interval $[0, T]$.

The following Lemma explains what happens, if the sequence (t_n) does not reach T in a finite number of steps.

Lemma 4.9 *Consider the sequence (t_n) from Lemma 4.7. If the times t_n satisfy*

$$t_n < T$$

for all $n \in \mathbb{N}$, then the sequence (t_n) converges and the construction from Lemma 4.7 gives us the unique solution Y of the differential equation (4.6) on the time *interval* $[0, \lim_{n\to\infty} t_n)$*. Furthermore, the solution* Y *explodes in finite time and the time of explosion is*

$$\tau^* = \lim_{n\to\infty} t_n \in [0, T].$$

Thus, the construction from Lemma 4.7 gives us the unique solution of (4.6) on the maximal interval of existence.

Proof of Lemma 4.9: By definition, the sequence (t_n) is monotonically nondecreasing and it is bounded from above by T. Therefore, its limit $\lim_{n\to\infty} t_n$ exists in $[0, T]$. Set $t^* = \lim_{n\to\infty} t_n$.

The construction from Lemma 4.7 gives us the unique solution Y of (4.6) on every interval $[0, t_n]$, $n \in \mathbb{N}$. For every $t \in [0, t^*)$, there exists $n \in \mathbb{N}$ such that $t \le t_n$. Hence, we get Y_t for every $t \in [0, t^*)$ and the solution is unique on $[0, t^*)$.

We have $t_n < T$ for all $n \in \mathbb{N}$ and thus, t_n is defined so that

$$\|X\|_{1,[t_n,t_{n+1}]} = \frac{1}{M_{n+1}}.$$

Set $t_0 = 0$ and define the sequence $(x_n)_{n\in\mathbb{N}}$ by

$$x_n = \frac{1}{M_{n+1}}.$$

Then, there exists a subsequence $(x_{n_j})_{j\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that x_{n_j} converges to 0. This can be shown by contradiction. Assume that there is no such subsequence. Then, there exists $\varepsilon > 0$ such that $x_n > \varepsilon$ for all $n \in \mathbb{N}$. Fix $m \in \mathbb{N}$ such that $m > \frac{1}{\varepsilon}\|X\|_{1,[0,T]}$. We get

$$\|X\|_{1,[0,T]} \ge \sum_{j=1}^{m} \|X\|_{1,[t_{n_j},t_{n_j+1}]} = \sum_{j=1}^{m} x_{n_j} > \sum_{j=1}^{m} \varepsilon = m\varepsilon > \|X\|_{1,[0,T]},$$

which is a contradiction. Hence, there exists a subsequence (x_{n_j}) of $(x_n)_{n\in\mathbb{N}}$ that converges to 0.

Then, the sequence $(M_{n_j+1})_{j\in\mathbb{N}}$ diverges to infinity since $M_{n_j+1} = x_{n_j}^{-1}$. It follows that

$$\left\|Y_{t_{n_j}}\right\|_W \stackrel{j\to\infty}{\longrightarrow} \infty$$

as $M_{n_j+1} = \max_{w\in B_1(\bar{Y}_{t_{n_j}})} \|f(w)\|_{L(V,W)}$ and f is continuous. The subsequence (t_{n_j}) converges to the same limit as (t_n) because of the monotonicity. Thus, Y explodes by time t^*.

It cannot explode any earlier since $Y_t \in W$ exists for every $t \in [0, t^*)$ and

$$\max_{s\in[0,t]} \|Y_s\|_W < \infty$$

because of the continuity.

Hence, we have $\tau^* = t^* = \lim_{n\to\infty} t_n$ and the solution Y explodes in finite time. □

Remark 4.10 *If f is Lipschitz continuous and not just locally Lipschitz continuous, there exists a unique solution of the differential equation (4.6) on the entire interval $[0,T]$. In this case, Picard iteration also works on the entire interval $[0,T]$ and the Picard approximations converge uniformly to the solution.*

4.4 Solutions of polynomial differential equations

4.4.1 Existence and uniqueness

Polynomials are locally Lipschitz continuous. Thus, we can apply the results from section 4.3 to prove the existence and uniqueness of solutions of polynomial differential equations.

Note that solutions of polynomial differential equations may explode in finite time. See for example Csikja and Tóth [16] for a study of blow-ups in polynomial differential equations.

Theorem 4.11 *Consider the polynomial differential equation (4.1), i.e.*

$$dy_t = P(y_t)\, d\gamma_t, \qquad y_0 = z,$$

where $\gamma : [0,T] \to \mathbb{R}^{d_2}$ *is a continuous path with bounded variation,* $z \in \mathbb{R}^{d_1}$*, and* $P : \mathbb{R}^{d_1} \to M(d_1, d_2, \mathbb{R})$ *is a continuous function, which takes values in the* $d_1 \times d_2$*-matrices, and is given by*

$$P(y_t) = \begin{pmatrix} p_{11}(y_t) & \cdots & p_{1d_2}(y_t) \\ \vdots & \ddots & \vdots \\ p_{d_1 1}(y_t) & \cdots & p_{d_1 d_2}(y_t) \end{pmatrix},$$

where $p_{ij} : \mathbb{R}^{d_1} \to \mathbb{R}$ *is a polynomial for every* $(i,j) \in \{1, \ldots, d_1\} \times \{1, \ldots, d_2\}$*. Then, there exists a unique solution* $y : \mathcal{I} \to \mathbb{R}^{d_1}$*, where* $\mathcal{I} \subset [0,T]$ *is the maximal interval of existence. If the solution explodes within* $[0,T]$*, we have* $\mathcal{I} = [0, \tau^*)$*, where* τ^* *is the time of explosion. Otherwise, we have* $\mathcal{I} = [0,T]$*.*

Proof: The function P is locally Lipschitz continuous. Therefore, Theorem 4.11 follows directly from Lemma 4.7 and Lemma 4.9. □

4.4.2 Value of the solution

In this section, we show how to obtain the value y_t of the solution of the polynomial differential equation (4.1) for small t from the signature $\Gamma_{0,t}$ of the signal γ, restricted to $[0,t]$. For linear differential equations, such a result can be found in [37]. Here, we generalise it to polynomial differential equations.

At first, we prove the following lemmas. Recall that we use the notation $L(V,W)$ for the space of bounded linear maps from a Banach space V to a Banach space W.

Lemma 4.12 *Let* $X : [0,t] \to \mathbb{R}^{d_2}$ *be a continuous path with bounded variation and let* $X_{0,t} = (1, X^1_{0,t}, X^2_{0,t}, \ldots)$ *be its signature.*

Define linear maps

$$L^j_{k,l} : L\left(\bigoplus_{i=k}^{l}(\mathbb{R}^{d_2})^{\otimes i}, \mathbb{R}\right) \to L\left(\bigoplus_{i=k+1}^{l+1}(\mathbb{R}^{d_2})^{\otimes i}, \mathbb{R}\right)$$

for $j \in \{1, \ldots, d_2\}$ *and integers* $k \le l$ *by*

$$L^j_{k,l}(\psi)(e_{i_1} \otimes \cdots \otimes e_{i_h}) = \begin{cases} \psi(e_{i_1} \otimes \cdots \otimes e_{i_{h-1}}) & \textit{if } i_h = j \\ 0 & \textit{if } i_h \neq j \end{cases}$$

for the basis elements $e_{i_1} \otimes \cdots \otimes e_{i_h}$, $h = k+1, \ldots, l+1$, *of* $\bigoplus_{i=k+1}^{l+1}(\mathbb{R}^{d_2})^{\otimes i}$.

Then,

$$L^j_{k,l}(\psi)((X^{k+1}_{0,t}, \ldots, X^{l+1}_{0,t})) = \int_0^t \psi((X^k_{0,s}, \ldots, X^l_{0,s}))\, dX^j_s$$

holds for every $\psi \in L\left(\bigoplus_{i=k}^{l}(\mathbb{R}^{d_2})^{\otimes i}, \mathbb{R}\right)$ *and every* $j \in \{1, \ldots, d_2\}$.

Proof: Consider $\psi \in L\left(\bigoplus_{i=k}^{l}(\mathbb{R}^{d_2})^{\otimes i}, \mathbb{R}\right)$. Let $e^*_{i_1,\ldots,i_h}$ be the element of the space $L\left((\mathbb{R}^{d_2})^{\otimes h}, \mathbb{R}\right)$ that is defined by

$$e^*_{i_1,\ldots,i_h}(e_{j_1} \otimes \cdots \otimes e_{j_h}) = \begin{cases} 1 & \text{if } j_1 = i_1, \ldots, j_h = i_h \\ 0 & \text{otherwise.} \end{cases}$$

The linear maps $e^*_{i_1,\ldots,i_h}$, $(i_1, \ldots, i_h) \in \{1, \ldots, d_2\}^h$, form a basis of $L\left((\mathbb{R}^{d_2})^{\otimes h}, \mathbb{R}\right)$. Furthermore, the linear maps $e^*_{i_1,\ldots,i_h}$, $(i_1, \ldots, i_h) \in \{1, \ldots, d_2\}^h$ for $h = k, k+1, \ldots, l$, form a basis of $L\left(\bigoplus_{i=k}^{l}(\mathbb{R}^{d_2})^{\otimes i}, \mathbb{R}\right)$. Therefore, there exist elements $\mu_{i_1,\ldots,i_h} \in \mathbb{R}$ such that

$$\psi = \sum_{\substack{(i_1,\ldots,i_h)\in\{1,\ldots,d_2\}^h,\\ h=k,k+1,\ldots,l}} \mu_{i_1,\ldots,i_h} e^*_{i_1,\ldots,i_h}$$

and we get

$$
\begin{aligned}
& \int_0^t \psi((X_{0,s}^k, \dots, X_{0,s}^l))\, dX_s^j \\
= & \int_0^t \left(\sum_{\substack{(i_1,\dots,i_h)\in\{1,\dots,d_2\}^h,\\ h=k,k+1,\dots,l}} \mu_{i_1,\dots,i_h} e^*_{i_1,\dots,i_h}((X_{0,s}^k, \dots, X_{0,s}^l)) \right) dX_s^j \\
= & \sum_{\substack{(i_1,\dots,i_h)\in\{1,\dots,d_2\}^h,\\ h=k,k+1,\dots,l}} \mu_{i_1,\dots,i_h} \int_0^t X_{0,s}^{(i_1,\dots,i_h)}\, dX_s^j \\
= & \sum_{\substack{(i_1,\dots,i_h)\in\{1,\dots,d_2\}^h,\\ h=k,k+1,\dots,l}} \mu_{i_1,\dots,i_h} \int_0^t \int_{0<u_1<\cdots<u_h<s} dX_{u_1}^{i_1} \dots dX_{u_h}^{i_h}\, dX_s^j \\
= & \sum_{\substack{(i_1,\dots,i_h)\in\{1,\dots,d_2\}^h,\\ h=k,k+1,\dots,l}} \mu_{i_1,\dots,i_h} X_{0,t}^{(i_1,\dots,i_h,j)} \\
= & L_{k,l}^j(\psi)((X_{0,t}^{k+1}, \dots, X_{0,t}^{l+1})).
\end{aligned}
$$

□

Let $P : \mathbb{R}^{d_1} \to M(d_1, d_2, \mathbb{R})$ be the function from the polynomial differential equation (4.1). It has the form

$$P = (p_{ij})_{\substack{i=1,\dots,d_1\\ j=1,\dots,d_2}}$$

where $p_{ij} : \mathbb{R}^{d_1} \to \mathbb{R}$ is a polynomial for every $(i,j) \in \{1,\dots,d_1\} \times \{1,\dots,d_2\}$. Suppose that

$$p_{ij}(x_1, \dots, x_{d_1}) = \lambda_0^{(i,j)} + \sum_{|\alpha|=1}^{N^{(i,j)}} \lambda_\alpha^{(i,j)} x_1^{\alpha_1} \dots x_d^{\alpha_d}. \tag{4.12}$$

Set

$$N = \max_{(i,j)\in\{1,\dots,d_1\}\times\{1,\dots,d_2\}} N^{(i,j)}. \tag{4.13}$$

and

$$\mathcal{S}_\alpha = (\{0,1,\dots,\alpha_1\} \times \cdots \times \{0,1,\dots,\alpha_{d_1}\}) \setminus \{(\alpha_1,\dots,\alpha_{d_1})\}$$

for multi-indeces $\alpha = (\alpha_1, \ldots, \alpha_{d_1})$. Define polynomials

$$\hat{p}_{ij} : \mathbb{R}^{d_1} \times \mathbb{R}^{d_1} \to \mathbb{R}$$

by

$$\hat{p}_{ij}(a,b) = \sum_{|\alpha|=1}^{N^{(i,j)}} \lambda_\alpha^{(i,j)} \sum_{(k_1,\ldots,k_{d_1}) \in \mathcal{S}_\alpha} \prod_{h=1}^{d_1} \left(\binom{\alpha_h}{k_h} a_h^{k_h} b_h^{\alpha_h - k_h} \right). \tag{4.14}$$

Furthermore, define the function

$$\hat{P} : \mathbb{R}^{d_1} \times \mathbb{R}^{d_1} \to M(d_1, d_2, \mathbb{R})$$

by

$$\hat{P}(a,b) = \left(\hat{p}_{ij}(a,b)\right)_{\substack{i=1,\ldots,d_1 \\ j=1,\ldots,d_2}}.$$

Lemma 4.13 *With the above definitions,*

$$\hat{p}_{ij}(a,b) = p_{ij}(a+b) - p_{ij}(a) \tag{4.15}$$

and

$$\hat{P}(a,b) = P(a+b) - P(a) \tag{4.16}$$

hold for every $a, b \in \mathbb{R}^{d_1}$.

Proof: By the binomial theorem, we get

$$(a_l + b_l)^{\alpha_l} = \sum_{k_l=0}^{\alpha_l} \binom{\alpha_l}{k_l} a_l^{k_l} b_l^{\alpha_l - k_l}$$

for $\alpha_l \in \mathbb{N}$ and $a_l, b_l \in \mathbb{R}$. Therefore, we have

$$\begin{aligned}
\prod_{l=1}^{d_1}(a_l+b_l)^{\alpha_l} &= \prod_{l=1}^{d_1}\left(\sum_{k_l=0}^{\alpha_l}\binom{\alpha_l}{k_l}a_l^{k_l}b_l^{\alpha_l-k_l}\right)\\
&= \sum_{k_1=0}^{\alpha_1}\cdots\sum_{k_{d_1}=0}^{\alpha_{d_1}}\binom{\alpha_1}{k_1}\cdots\binom{\alpha_{d_1}}{k_{d_1}}a_1^{k_1}\dots a_{d_1}^{k_{d_1}}b_1^{\alpha_1-k_1}\dots b_{d_1}^{\alpha_{d_1}-k_{d_1}}\\
&= a_1^{\alpha_1}\dots a_{d_1}^{\alpha_{d_1}}+\sum_{|\alpha|=1}^{N^{(i,j)}}\sum_{(k_1,\dots,k_{d_1})\in\mathcal{S}_\alpha}\prod_{h=1}^{d_1}\left(\binom{\alpha_h}{k_h}a_h^{k_h}b_h^{\alpha_h-k_h}\right).
\end{aligned}$$

Thus, (4.15) follows from (4.12) and (4.14). Equation (4.16) is a direct consequence of (4.15). □

Now, we are in the position to show how to obtain y_t from the signature of γ. Set

$$m_0 = 0$$

and

$$m_n = 1 + N + N^2 + \cdots + N^{n-1} \tag{4.17}$$

for integers $n \geq 1$, where N is the degree of P as given by (4.13). Define linear maps

$$A^{n,z} : \bigoplus_{l=n}^{m_n}(\mathbb{R}^{d_2})^{\otimes l} \to \mathbb{R}^{d_1}$$

recursively for integers $n \geq 0$ and $z \in \mathbb{R}^{d_1}$, starting from

$$A^{0,z} : \mathbb{R} \to \mathbb{R}^{d_1}, \quad A^{0,z}(x) = xz \tag{4.18}$$

and

$$A^{1,z} : \mathbb{R}^{d_2} \to \mathbb{R}^{d_1}, \quad A^{1,z}(x) = P(z)x. \tag{4.19}$$

Given $A^{0,z}, A^{1,z}, \ldots, A^{n,z}$, we construct $A^{n+1,z}$ as follows: Define the linear map

$$\varphi^{n-1,z} : \bigoplus_{l=0}^{m_{n-1}} (\mathbb{R}^{d_2})^{\otimes l} \to \mathbb{R}^{d_1}$$

by

$$\varphi^{n-1,z}((x_0, x_1, \ldots, x_{m_{n-1}})) = \sum_{l=0}^{n-1} A^{l,z}((x_l, \ldots, x_{m_l})),$$

where $x_h \in (\mathbb{R}^{d_2})^{\otimes h}$ for $h = 1, \ldots, m_{n-1}$.

For $k = 1, \ldots, d_1$, let $e_k^* : \mathbb{R}^{d_1} \to \mathbb{R}$ be the projection onto the k-th coordinate. Note that e_k^* is linear and set $\varphi_k^{n-1,z} = e_k^* \circ \varphi^{n-1,z}$ and $A_k^{n,z} = e_k^* \circ A^{n,z}$.

For $i = 1, \ldots, d_1$ and $j = 1, \ldots, d_2$, define linear maps

$$\Phi_{ij}^{n,z} : \bigoplus_{l=n}^{m_{n+1}-1} (\mathbb{R}^{d_2})^{\otimes l} \to \mathbb{R}$$

by

$$\Phi_{ij}^{n,z} = \sum_{|\alpha|=1}^{N^{(i,j)}} \lambda_\alpha^{(i,j)} \sum_{(k_1,\ldots,k_{d_1}) \in \mathcal{S}_\alpha} \binom{\alpha_1}{k_1} \cdots \binom{\alpha_{d_1}}{k_{d_1}} (\varphi_1^{n-1,z})^{\sqcup k_1} \sqcup \cdots \sqcup (\varphi_{d_1}^{n-1,z})^{\sqcup k_{d_1}} \sqcup (A_1^{n,z})^{\sqcup(\alpha_1 - k_1)} \sqcup \cdots \sqcup (A_{d_1}^{n,z})^{\sqcup(\alpha_{d_1} - k_{d_1})}, \tag{4.20}$$

where

$$(\varphi_l^{n-1,z})^{\sqcup k_l} = \underbrace{\varphi_l^{n-1,z} \sqcup \cdots \sqcup \varphi_l^{n-1,z}}_{k_l\text{-times}}$$

and

$$(A_l^{n,z})^{\sqcup(\alpha_l - k_l)} = \underbrace{A_l^{n,z} \sqcup \cdots \sqcup A_l^{n,z}}_{(\alpha_l - k_l)\text{-times}}$$

for $l=1,\dots,d_1$. Note that $\Phi_{ij}^{n,z}$ is well defined as a map from $\bigoplus_{l=n}^{m_{n+1}-1}(\mathbb{R}^{d_2})^{\otimes l}$ to $\mathbb{R}$. Indeed, we have

$$\begin{aligned} & m_{n-1}\sum_{l=1}^{d_1} k_l + m_n\sum_{l=1}^{d_1}(\alpha_l-k_l) \le m_n\sum_{l=1}^{d_1}(k_l+(\alpha_l-k_l)) \\ = \; & m_n\sum_{l=1}^{d_1}\alpha_l \le Nm_n = N+N^2+\cdots+N^n = m_{n+1}-1 \end{aligned}$$

for every integer $n \ge 1$. In addition, for every α such that $\mathcal{S}_\alpha \neq \emptyset$ and every $(k_1,\dots,k_{d_1}) \in \mathcal{S}_\alpha$, there exists $\tilde{l} \in \{1,\dots,d_1\}$ such that $\alpha_{\tilde{l}} - k_{\tilde{l}} \ge 1$. Thus, we have

$$n\sum_{l=1}^{d_1}(\alpha_l-k_l) \ge n.$$

Finally, define

$$A_i^{n+1,z} = \sum_{j=1}^{d_2} L^j_{n,m_{n+1}-1}(\Phi_{ij}^{n,z}) \tag{4.21}$$

for $i=1,\dots,d_1$ and set

$$A^{n+1,z} = \begin{pmatrix} A_1^{n+1,z} \\ \vdots \\ A_{d_1}^{n+1,z} \end{pmatrix}. \tag{4.22}$$

Theorem 4.14 *Consider the polynomial differential equation (4.1), i.e.*

$$dy_t = P(y_t)\,d\gamma_t, \qquad y_0 = z,$$

where $\gamma : [0,T] \to \mathbb{R}^{d_2}$ is a continuous path with bounded variation, $z \in \mathbb{R}^{d_1}$, and $P : \mathbb{R}^{d_1} \to M(d_1,d_2,\mathbb{R})$ is a continuous function, which takes values in the

$d_1 \times d_2$-matrices, and is given by

$$P(y_t) = \begin{pmatrix} p_{11}(y_t) & \cdots & p_{1d_2}(y_t) \\ \vdots & \ddots & \vdots \\ p_{d_1 1}(y_t) & \cdots & p_{d_1 d_2}(y_t) \end{pmatrix},$$

where $p_{ij} : \mathbb{R}^{d_1} \to \mathbb{R}$ is a polynomial for every $(i,j) \in \{1,\dots,d_1\} \times \{1,\dots,d_2\}$. Fix $t \in [0,T]$ so that

$$\|\gamma\|_{1,[0,t]} \leq \left(\max_{w \in \bar{B}_1(z)} \|P(w)\| \right)^{-1}.$$

Then, the following holds:

(i) There exists a unique solution $y : [0,t] \to \mathbb{R}^{d_1}$ of the polynomial differential equation (4.1) on $[0,t]$ and we can obtain the solution y_t via Picard iteration. For every $n \in \mathbb{N}$, define paths $y^{(n)} : [0,t] \to \mathbb{R}^{d_1}$ recursively by

$$y_s^{(n+1)} = z + \int_0^s P(y_u^{(n)})\, d\gamma_u \tag{4.23}$$

starting from

$$y_s^{(0)} = z$$

for all $s \in [0,t]$. Then, the sequence $(y_t^{(n)})$ converges and

$$\lim_{n\to\infty} y_t^{(n)} = y_t. \tag{4.24}$$

(ii) Let

$$\Gamma_{0,t} = (1, \Gamma_{0,t}^1, \Gamma_{0,t}^2, \dots) \in T((\mathbb{R}^{d_2}))$$

be the signature of the signal γ, restricted to $[0,t]$, and let

$$A^{k,z} : \bigoplus_{i=k}^{m_k} (\mathbb{R}^{d_2})^{\otimes i} \to \mathbb{R}^{d_1}$$

be the linear maps that are constructed in (4.18), (4.19), (4.21) and (4.22), where m_k is given by (4.17). Then,

$$y_t^{(n)} = \sum_{k=0}^{n} A^{k,z}((\Gamma_{0,t}^{k}, \ldots, \Gamma_{0,t}^{m_k})) \tag{4.25}$$

holds for every integer $n \geq 0$ and we have

$$y_t = \sum_{k=0}^{\infty} A^{k,z}((\Gamma_{0,t}^{k}, \ldots, \Gamma_{0,t}^{m_k})). \tag{4.26}$$

The rate of convergence is

$$\left\| y_t - \left(\sum_{k=0}^{n} A^{k,z}((\Gamma_{0,t}^{k}, \ldots, \Gamma_{0,t}^{m_k})) \right) \right\| \leq \frac{(L/M)^n}{(n+1)!} e^{L/M}, \tag{4.27}$$

where $\bar{B}_1(z) \subset \mathbb{R}^{d_1}$ is the closed ball $\bar{B}_1(z) = \{u \in \mathbb{R}^{d_1} : \|u - z\| \leq 1\}$, L is a Lipschitz constant for P on $\bar{B}_1(z)$ and

$$M = \max_{w \in \bar{B}_1(z)} \|P(w)\|.$$

Proof: Part (i) follows from Theorem 4.5 since P is locally Lipschitz continuous. In order to show part (ii), we first prove (4.25) by induction over n. For $n = 0$, we get

$$y_t^{(0)} = z = A^{0,z}(1) = A^{1,z}(\Gamma_{0,t}^{0})$$

and

$$\begin{aligned} y_t^{(1)} &= z + \int_0^t P(y_s^{(0)})\, d\gamma_s = z + \int_0^t P(z)\, d\gamma_s = z + P(z) \int_0^t d\gamma_s \\ &= z + P(z)\Gamma_{0,t}^{1} = z + A^{1,z}(\Gamma_{0,t}^{1}) \end{aligned}$$

for $n = 1$ by (4.23) and the definition of $A^{1,z}$. Now, assume that (4.25) holds for integers $0, 1, \ldots, n$. Then, we get

$$
\begin{aligned}
y_t^{(n+1)} &= z + \int_0^t P(y_s^{(n)})\, d\gamma_s = z + \int_0^t P\left(\sum_{k=0}^{n} A^{k,z}((\Gamma_{0,s}^k, \ldots, \Gamma_{0,s}^{m_k}))\right) d\gamma_s \\
&= z + \int_0^t P\left(\sum_{k=0}^{n-1} A^{k,z}((\Gamma_{0,s}^k, \ldots, \Gamma_{0,s}^{m_k}))\right) d\gamma_s \\
&\quad + \int_0^t \hat{P}\left(\sum_{k=0}^{n-1} A^{k,z}(\Gamma_{0,s}^k, \ldots, \Gamma_{0,s}^{m_k}), A^{n,z}((\Gamma_{0,s}^n, \ldots, \Gamma_{0,s}^{m_n}))\right) d\gamma_s
\end{aligned}
$$

by (4.23), (4.16) and the induction hypothesis. Again by the induction hypothesis and (4.23), we have

$$
\begin{aligned}
z + \int_0^t P\left(\sum_{k=0}^{n-1} A^{k,z}((\Gamma_{0,s}^k, \ldots, \Gamma_{0,s}^{m_k}))\right) d\gamma_s &= z + \int_0^t P(y_s^{(n-1)})\, d\gamma_s = y_t^{(n)} \\
&= \sum_{k=0}^{n} A^{k,z}((\Gamma_{0,t}^k, \ldots, \Gamma_{0,t}^{m_k})).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
y_t^{(n+1)} = &\sum_{k=0}^{n} A^{k,z}((\Gamma_{0,t}^k, \ldots, \Gamma_{0,t}^{m_k})) \\
&+ \int_0^t \hat{P}\left(\sum_{k=0}^{n-1} A^{k,z}((\Gamma_{0,s}^k, \ldots, \Gamma_{0,s}^{m_k})), A^{n,z}((\Gamma_{0,s}^n, \ldots, \Gamma_{0,s}^{m_n}))\right) d\gamma_s
\end{aligned}
\tag{4.28}
$$

holds. By (4.14), (4.20) and the property (4.3) of the shuffle product, we get

$$
\Phi_{ij}^{n,z}\left((\Gamma_{0,s}^n, \ldots, \Gamma_{0,s}^{m_{n+1}-1})\right) = \hat{p}_{ij}\left(\sum_{k=0}^{n-1} A^{k,z}((\Gamma_{0,s}^k, \ldots, \Gamma_{0,s}^{m_k})), A^{n,z}((\Gamma_{0,s}^n, \ldots, \Gamma_{0,s}^{m_n}))\right)
$$

for all $s \in [0,T]$ and all $(i,j) \in \{1,\ldots,d_1\} \times \{1,\ldots,d_2\}$. Therefore, we have

$$
\begin{aligned}
& \int_0^t \hat{P}\left(\sum_{k=0}^{n-1} A^{k,z}((\Gamma^k_{0,s},\ldots,\Gamma^{m_k}_{0,s})), A^{n,z}((\Gamma^n_{0,s},\ldots,\Gamma^{m_n}_{0,s}))\right) d\gamma_s \\
= & \sum_{i=1}^{d_1}\left(\sum_{j=1}^{d_2}\int_0^t \hat{p}_{ij}\left(\sum_{k=0}^{n-1} A^{k,z}((\Gamma^k_{0,s},\ldots,\Gamma^{m_k}_{0,s})), A^{n,z}((\Gamma^n_{0,s},\ldots,\Gamma^{m_n}_{0,s}))\right) d\gamma_s\right) e_i \\
= & \sum_{i=1}^{d_1}\left(\sum_{j=1}^{d_2}\int_0^t \Phi^{n,z}_{ij}\left((\Gamma^n_{0,s},\ldots,\Gamma^{m_{n+1}-1}_{0,s})\right) d\gamma_s\right) e_i \\
= & \sum_{i=1}^{d_1}\left(\sum_{j=1}^{d_2} L^j_{n,m_{n+1}-1}(\Phi^{n,z}_{ij})((\Gamma^{n+1}_{0,t},\ldots,\Gamma^{m_{n+1}}_{0,t}))\right) e_i \\
= & \sum_{i=1}^{d_1} A^{n+1,z}_i((\Gamma^{n+1}_{0,t},\ldots,\Gamma^{m_{n+1}}_{0,t}))e_i \\
= & A^{n+1,z}_i((\Gamma^{n+1}_{0,t},\ldots,\Gamma^{m_{n+1}}_{0,t}))
\end{aligned}
$$

by Lemma 4.12, (4.21) and (4.22). Thus, it follows from (4.28) that

$$
y^{(n+1)}_t = z + \sum_{k=1}^{n+1} A^{k,z}((\Gamma^k_{0,t},\ldots,\Gamma^{m_k}_{0,t}))
$$

which completes the induction and hence shows (4.25).

Now, (4.26) follows from (4.25) and (4.24), and (4.27) follows from (4.25) and Theorem 4.5. □

4.4.3 Signature of the solution

Now, we show how to obtain the truncated signature of the solution of the polynomial differential equation (4.1) over small time intervals from the signature of the signal γ and not just its value y_t as in the previous section. Lyons et al. [37] demonstrate this for linear differential equations (see [37], section 4.2). They use the series expansion for the value of the solution in terms of the signature of γ (analogue of Theorem 4.14(ii) for linear differential equations) and then compute the iterated integrals and identify them as functions of the signature of γ. Here, we use a different approach. The truncated signature is itself the solution of a differential equation driven by the path γ. We solve this differential equation by Picard iteration.

Definition 4.15 *Let M be an element of $\mathbb{N}$. The truncated tensor algebra $T^{(M)}(\mathbb{R}^{d_1})$ is defined as*

$$T^{(M)}(\mathbb{R}^{d_1}) = \bigoplus_{i=0}^{M} (\mathbb{R}^{d_1})^{\otimes i}$$

equipped with the product

$$(a_0, a_1, \ldots, a_M)(b_0, b_1, \ldots, b_M) = (c_0, c_1, \ldots, c_M),$$

where

$$c_i = \sum_{j=0}^{i} a_j \otimes b_{i-j}$$

for $i = 0, 1, \ldots, M$.

We equip $T^{(M)}(\mathbb{R}^{d_1})$ with the maximum norm. Let $\mathbf{a} = (a_0, a_1, \ldots, a_M)$ be an element of $T^{(M)}(\mathbb{R}^{d_1})$. Then, set

$$\|\mathbf{a}\| = \max\{|a_0|, \|a_1\|_{\mathbb{R}^{d_1},\infty}, \|a_2\|_{(\mathbb{R}^{d_1})^{\otimes 2},\infty}, \ldots, \|a_M\|_{(\mathbb{R}^{d_1})^{\otimes M},\infty}\},$$

where $\|\cdot\|_{(\mathbb{R}^{d_1})^{\otimes k},\infty}$, $k = 1, \ldots, M$, is the maximum norm on $(\mathbb{R}^{d_1})^{\otimes k}$ as in section 3.1.

Theorem 4.16 *Consider the polynomial differential equation (4.1), i.e.*

$$dy_t = P(y_t)\, d\gamma_t, \qquad y_0 = z,$$

where $\gamma : [0, T] \to \mathbb{R}^{d_2}$ is a continuous path with bounded variation, $z \in \mathbb{R}^{d_1}$, and $P : \mathbb{R}^{d_1} \to M(d_1, d_2, \mathbb{R})$ is a continuous function, which takes values in the

$d_1 \times d_2$-matrices, and is given by

$$P(y_t) = \begin{pmatrix} p_{11}(y_t) & \cdots & p_{1d_2}(y_t) \\ \vdots & \ddots & \vdots \\ p_{d_11}(y_t) & \cdots & p_{d_1d_2}(y_t) \end{pmatrix},$$

where $p_{ij} : \mathbb{R}^{d_1} \to \mathbb{R}$ is a polynomial for every $(i,j) \in \{1,\dots,d_1\} \times \{1,\dots,d_2\}$. Let $y : \mathcal{I} \to \mathbb{R}^{d_1}$ be its unique solution as discussed in Theorem 4.11, where $\mathcal{I} \subset [0,T]$ is the maximal interval of existence. Fix $\tilde{T} \in \mathcal{I}$ and let $M \geq N$ be an integer, where N is the degree of P as given by (4.13). Let

$$Y_{0,t}^{[M]} = (1, Y_{0,t}^1, \dots, Y_{0,t}^M) \in T^{(M)}(\mathbb{R}^{d_1})$$

be the truncated signature of y, restricted to $[0,t]$.
Then, there exists a linear map

$$\Psi^{M,z} : T^{(M)}(\mathbb{R}^{d_1}) \to M(d_1, d_2, \mathbb{R})$$

such that

$$\Psi^{M,z}(Y_{0,t}^{[M]}) = P(y_t) \tag{4.29}$$

holds for all $t \in [0, \tilde{T}]$.
Furthermore, the truncated signature $Y_{0,\cdot}^{[M]}$ is the unique solution of the differential equation

$$dY_{0,t}^{[M]} = Y_{0,t}^{[M]} \otimes \Psi^{M,z}(Y_{0,t}^{[M]})\, d\gamma_t, \qquad Y_{0,0}^{[M]} = (1, 0, \dots, 0) \tag{4.30}$$

and the solution exists on $[0, \tilde{T}]$.

Note that we have reduced the polynomial differential equation (4.1) to the quadratic differential equation (4.30) on $T^{(M)}(\mathbb{R}^{d_1})$.
In (4.30), the notation $Y_{0,t}^{[M]} \otimes \Psi^{M,z}(Y_{0,t}^{[M]})$ has the following meaning. Define the

map $g^{M,z} : T^{(M)}(\mathbb{R}^{d_1}) \to L\left(\mathbb{R}^{d_2}, T^{(M)}(\mathbb{R}^{d_1})\right)$ by

$$g^{M,z}(\mathbf{a})x = \left(0, a_0 \otimes \left(\Psi^{M,z}(\mathbf{a})x\right), \ldots, a_{M-1} \otimes \left(\Psi^{M,z}(\mathbf{a})x\right)\right) \tag{4.31}$$

for $\mathbf{a} = (a_0, a_1, \ldots, a_M) \in T^{(M)}(\mathbb{R}^{d_1})$ and $x \in \mathbb{R}^{d_2}$. Then,

$$Y_{0,t}^{[M]} \otimes \Psi^{M,z}(Y_{0,t}^{[M]}) = g^{M,z}(Y_{0,t}^{[M]})$$

and with this notation, the differential equation (4.30) reads

$$dY_{0,t}^{[M]} = g^{M,z}(Y_{0,t}^{[M]})\, d\gamma_t, \qquad Y_{0,0}^{[M]} = (1, 0, \ldots, 0). \tag{4.32}$$

Proof of Theorem 4.16: Define polynomials $\tilde{p}_{ij}^z : \mathbb{R}^{d_1} \to \mathbb{R}$ by

$$\tilde{p}_{ij}^z(x) = p_{ij}(x + z).$$

It follows from Theorem 4.2 that there exist linear maps $\Psi_{ij}^{M,z} : T^{(M)}(\mathbb{R}^{d_1}) \to \mathbb{R}$ such that

$$\Psi_{ij}^{M,z}(Y_{0,t}^{[M]}) = \tilde{p}_{ij}^z(y_t - y_0)$$

holds for all $t \in [0, \tilde{T}]$ and this implies

$$\Psi_{ij}^{M,z}(Y_{0,t}^{[M]}) = \tilde{p}_{ij}^z(y_t - y_0) = p_{ij}(y_t).$$

Define the linear map $\Psi^{M,z} : T^{(M)}(\mathbb{R}^{d_1}) \to M(d_1, d_2, \mathbb{R})$ by

$$\Psi^{M,z}(\mathbf{a}) = \left(\Psi_{ij}^{M,z}(\mathbf{a})\right)_{\substack{i=1,\ldots,d_1 \\ j=1,\ldots,d_2}}$$

for all $\mathbf{a} \in T^{(M)}(\mathbb{R}^{d_1})$. Then, we have

$$\Psi^{M,z}(Y_{0,t}^{[M]}) = P(y_t)$$

for all $t \in [0, \tilde{T}]$, which proves (4.29). It follows from [37], Lemma 2.10, that the truncated signature $Y_{0,\cdot}^{[M]}$ is the unique solution of the differential equation

$$dY_{0,t}^{[M]} = Y_{0,t}^{[M]} \otimes dy_t, \qquad Y_{0,0}^{[M]} = (1, 0, \dots, 0),$$

and that the solution exists on $[0, \tilde{T}]$. By (4.1), this differential equation is

$$dY_{0,t}^{[M]} = Y_{0,t}^{[M]} \otimes P(y_t)\, d\gamma_t$$

and it becomes

$$dY_{0,t}^{[M]} = Y_{0,t}^{[M]} \otimes \Psi^{M,z}(Y_{0,t}^{[M]})\, d\gamma_t$$

by (4.29). □

Now, we show how to obtain the truncated signature $Y_{0,\tilde{T}}^{[M]}$ from the signature $\Gamma_{0,\tilde{T}}$ of the signal γ for small $\tilde{T}$. Define linear maps

$$B^{n,z} : \bigoplus_{i=n}^{2^n-1} (\mathbb{R}^{d_2})^{\otimes i} \to T^{(M)}(\mathbb{R}^{d_1})$$

recursively for $n \in \mathbb{N}$, starting from

$$B^{0,z} : \mathbb{R} \to T^{(M)}(\mathbb{R}^{d_1}), \qquad B^{0,z}(x_0) = (x_0, 0, \dots, 0) \tag{4.33}$$

and

$$B^{1,z} : \mathbb{R}^{d_2} \to T^{(M)}(\mathbb{R}^{d_1}), \qquad B^{1,z}(x_1) = g^{M,z}((1, 0, \dots, 0))(x_1), \tag{4.34}$$

where the map $g^{M,z}$ is defined in (4.31). Given $B^{0,z}, B^{1,z}, \dots, B^{n,z}$, we construct $B^{n+1,z}$ in the following way: For integers $m = 0, 1, \dots, n$, define the linear maps

$$\hat{\varphi}^{m,z} : \bigoplus_{i=0}^{2^m-1} (\mathbb{R}^{d_2})^{\otimes i} \to T^{(M)}(\mathbb{R}^{d_1})$$

by

$$\hat{\varphi}^{m,z}((a_0, a_1, \ldots, a_{2^m-1})) = \sum_{i=0}^{m} B^{i,z}((a_i, a_{i+1}, \ldots, a_{2^i-1})).$$

Recall that the linear maps $\Psi_{lj}^{M,z}$ are defined such that

$$\Psi_{lj}^{M,z}(Y_{0,\tilde{T}}^{[M]}) = p_{lj}(y_t)$$

for $(l,j) \in \{1,\ldots,d_1\} \times \{1,\ldots,d_2\}$ as in Theorem 4.16. For $m = 1,\ldots,M$, define the linear maps

$$\tilde{f}_m^{n,z} : \bigoplus_{i=n}^{2^{n+1}-2} (\mathbb{R}^{d_2})^{\otimes i} \to L\left(\mathbb{R}^{d_2}, (\mathbb{R}^{d_1})^{\otimes m}\right)$$

by

$$\tilde{f}_m^{n,z}((a_n,\ldots,a_{2^{n+1}-2}))(x) = \sum_{i_1,\ldots,i_{m-1}}^{d_1} \sum_{l=1}^{d_1} \left(\sum_{j=1}^{d_2} \left((e^*_{i_1,\ldots,i_{m-1}} \circ B^{n,z}) \sqcup (\Psi_{lj}^{M,z} \circ \hat{\varphi}^{n,z}) \right) ((a_n,\ldots,a_{2^{n+1}-2}))x_j \right) e_{i_1} \otimes \cdots \otimes e_{i_{m-1}} \otimes e_l,$$

the linear maps

$$\hat{f}_m^{n,z} : \bigoplus_{i=n}^{3\times 2^{n-1}-2} (\mathbb{R}^{d_2})^{\otimes i} \to L\left(\mathbb{R}^{d_2}, (\mathbb{R}^{d_1})^{\otimes m}\right)$$

by

$$\hat{f}_m^{n,z}((a_n,\ldots,a_{3\times 2^{n-1}-2}))(x) = \sum_{i_1,\ldots,i_{m-1}}^{d_1} \sum_{l=1}^{d_1} \left(\sum_{j=1}^{d_2} \left((e^*_{i_1,\ldots,i_{m-1}} \circ \hat{\varphi}^{n-1,z}) \sqcup (\Psi_{lj}^{M,z} \circ B^{n,z}) \right) ((a_n,\ldots,a_{3\times 2^{n-1}-2}))x_j \right) e_{i_1} \otimes \cdots \otimes e_{i_{m-1}} \otimes e_l$$

and the linear maps

$$f_m^{n,z} : \bigoplus_{i=n}^{2^{n+1}-2} (\mathbb{R}^{d_2})^{\otimes i} \to L\left(\mathbb{R}^{d_2}, (\mathbb{R}^{d_1})^{\otimes m}\right)$$

by

$$f_m^{n,z}((a_n,\dots,a_{2^{n+1}-2}))(x) = \tilde{f}_m^{n,z}((a_n,\dots,a_{2^{n+1}-2}))(x) + \hat{f}_m^{n,z}((a_n,\dots,a_{3\times 2^{n-1}-2}))(x).$$

Furthermore, define the linear maps

$$\hat{\Phi}_{(i_1,\dots,i_m),j} : \bigoplus_{k=n}^{2^{n+1}-2} (\mathbb{R}^{d_2})^{\otimes k} \to \mathbb{R}$$

by

$$\hat{\Phi}_{(i_1,\dots,i_m),j}((a_n,\dots,a_{2^{n+1}-2})) = e^*_{i_1,\dots,i_m}\left(f_m^{n,z}((a_n,\dots,a_{2^{n+1}-2}))(e_j)\right)$$

for $i_1,\dots,i_m \in \{1,\dots,d_1\}$, $m = 1,\dots,M$.

Now, we can finally define $B^{n+1,z}$. Define the linear map

$$B^{n+1,z} : \bigoplus_{i=n+1}^{2^{n+1}-1} (\mathbb{R}^{d_2})^{\otimes i} \to T^{(M)}(\mathbb{R}^{d_1})$$

by

$$B^{n+1,z}((a_{n+1},\dots,a_{2^{n+1}-1})) = \sum_{\substack{i_1,\dots,i_m=1,\\ m=1,\dots,M}}^{d_1} \sum_{j=1}^{d_2} \left(L^j_{n,2^{n+1}-2}(\hat{\Phi}_{(i_1,\dots,i_m),j})\right) ((a_{n+1},\dots,a_{2^{n+1}-1}))e_{i_1}\otimes\cdots\otimes e_{i_m}, \tag{4.35}$$

where $L^j_{n,2^{n+1}-2}$ is defined in Lemma 4.12.

The following theorem shows how we can obtain the truncated signature $Y^{[M]}_{0,\tilde{T}}$ from the signature $\Gamma_{0,\tilde{T}}$ of the signal γ for small $\tilde{T}$.

Theorem 4.17 *Consider the polynomial differential equation (4.1), i.e.*

$$dy_t = P(y_t)\,d\gamma_t, \qquad y_0 = z,$$

where $\gamma : [0,T] \to \mathbb{R}^{d_2}$ *is a continuous path with bounded variation,* $z \in \mathbb{R}^{d_1}$, *and $P : \mathbb{R}^{d_1} \to M(d_1, d_2, \mathbb{R})$ is a continuous function, which takes values in the $d_1 \times d_2$-matrices, and is given by*

$$P(y_t) = \begin{pmatrix} p_{11}(y_t) & \cdots & p_{1d_2}(y_t) \\ \vdots & \ddots & \vdots \\ p_{d_11}(y_t) & \cdots & p_{d_1d_2}(y_t) \end{pmatrix},$$

where $p_{ij} : \mathbb{R}^{d_1} \to \mathbb{R}$ is a polynomial for every $(i,j) \in \{1,\dots,d_1\} \times \{1,\dots,d_2\}$. Let $y : \mathcal{I} \to \mathbb{R}^{d_1}$ be its unique solution as discussed in Theorem 4.11, where $\mathcal{I} \subset [0,T]$ is the maximal interval of existence. Fix $\tilde{T} \in \mathcal{I}$ small enough so that

$$\|\gamma\|_{1,[0,\tilde{T}]} \leq \left(\max_{\mathbf{a} \in \bar{B}_1((1,0,\dots,0))} \|g^{M,z}(\mathbf{a})\| \right)^{-1}$$

and let $M \geq N$ be an integer, where N is the degree of P as given by (4.13). Let

$$\Gamma_{0,\tilde{T}} = (1, \Gamma^1_{0,\tilde{T}}, \Gamma^2_{0,\tilde{T}}, \dots) \in T((\mathbb{R}^{d_2}))$$

be the signature of the signal γ, restricted to $[0,\tilde{T}]$. Let

$$B^{i,z} : \bigoplus_{j=i}^{2^i-1} (\mathbb{R}^{d_2})^{\otimes j} \to T^{(M)}(\mathbb{R}^{d_1})$$

be the linear maps that are constructed in (4.33), (4.34) and (4.35). Then, the truncated signature $Y^{[M]}_{0,\tilde{T}}$ of the solution y, restricted to $[0,\tilde{T}]$, is

$$Y^{[M]}_{0,\tilde{T}} = \sum_{i=0}^{\infty} B^{i,z}((\Gamma^i_{0,\tilde{T}}, \Gamma^{i+1}_{0,\tilde{T}}, \dots, \Gamma^{2^i-1}_{0,\tilde{T}})). \tag{4.36}$$

The rate of convergence is

$$\left\| Y^{[M]}_{0,\tilde{T}} - \sum_{i=0}^{n} B^{i,z}((\Gamma^i_{0,\tilde{T}}, \Gamma^{i+1}_{0,\tilde{T}}, \dots, \Gamma^{2^i-1}_{0,\tilde{T}})) \right\| \leq \frac{(L/M)^n}{(n+1)!} e^{L/M}, \tag{4.37}$$

where $g^{M,z}$ is the map defined in (4.31), $\bar{B}_1((1,0,\ldots,0)) \subset T^{(M)}(\mathbb{R}^{d_1})$ is the closed ball $\bar{B}_1((1,0,\ldots,0)) = \{\mathbf{a} \in T^{(M)}(\mathbb{R}^{d_1}) : \|\mathbf{a} - (1,0,\ldots,0)\| \leq 1\}$, L is a Lipschitz constant for $g^{M,z}$ on $\bar{B}_1((1,0,\ldots,0))$ and

$$M = \max_{\mathbf{a} \in \bar{B}_1((1,0,\ldots,0))} \|g^{M,z}(\mathbf{a})\|.$$

Proof: It follows from Theorem 4.16 that $Y^{[M]}_{0,\tilde{T}}$ is the unique solution of the differential equation (4.30). With the notation (4.31), the differential equation (4.30) becomes (4.32) and it follows that $Y^{[M]}_{0,\tilde{T}}$ is the unique solution of the differential equation (4.32).

The map $g^{M,z}$, as defined in (4.31), is locally Lipschitz continuous by definition. Therefore, we can apply the results from section 4.3 and it follows from Theorem 4.5 that we can solve the differential equation (4.32) via Picard iteration.

For every $n \in \mathbb{N}$, define $Y^{[M,n]}_{0,\cdot} : [0,\tilde{T}] \to T^{(M)}(\mathbb{R}^{d_1})$ recursively by

$$Y^{[M,n+1]}_{0,t} = (1,0,\ldots,0) + \int_0^t g^{M,z}(Y^{[M,n]}_{0,s})\, d\gamma_s \tag{4.38}$$

starting from

$$Y^{[M,0]}_{0,t} = (1,0,\ldots,0)$$

for all $t \in [0,\tilde{T}]$. Note that the integral $\int_0^t g^{M,z}(Y^{[M,n]}_{0,s})\, d\gamma_s$ is defined component-wise

$$\int_0^t g^{M,z}(Y^{[M,n]}_{0,s})\, d\gamma_s = \left(0, \int_0^t g_1^{M,z}(Y^{[M,n]}_{0,s})\, d\gamma_s, \ldots, \int_0^t g_M^{M,z}(Y^{[M,n]}_{0,s})\, d\gamma_s\right).$$

Then, it follows from Theorem 4.5 that

$$\lim_{n\to\infty} Y^{[M,n]}_{0,\tilde{T}} = Y^{[M]}_{0,\tilde{T}}. \tag{4.39}$$

Now, we show that

$$Y_{0,t}^{[M,n]} = \sum_{i=0}^{n} B^{i,z}((\Gamma_{0,t}^{i}, \Gamma_{0,t}^{i+1}, \ldots, \Gamma_{0,t}^{2^i-1})) \tag{4.40}$$

holds for every $t \in [0, \tilde{T}]$ and every $n \in \mathbb{N}$.

For the brevity of notation, set $\mathbf{1} = (1, 0, \ldots, 0) \in T^{(M)}(\mathbb{R}^{d_1})$. We show (4.40) by induction over n. For $n = 0$, we have

$$Y_{0,t}^{[M,0]} = \mathbf{1} = B^{0,z}(1) = B^{0,z}(\Gamma_{0,t}^{0}),$$

and for $n = 1$, we have

$$\begin{aligned} Y_{0,t}^{[M,1]} &= \mathbf{1} + \int_0^t g^{[M],z}(Y_{0,s}^{[M,0]})\, d\gamma_s = \mathbf{1} + \int_0^t g^{[M],z}(\mathbf{1})\, d\gamma_s \\ &= \mathbf{1} + g^{[M],z}(\mathbf{1}) \left(\int_0^t d\gamma_s \right) = \mathbf{1} + g^{[M],z}(\mathbf{1})(\Gamma_{0,t}^{1}) \\ &= B^{0,z}(\Gamma_{0,t}^{0}) + B^{1,z}(\Gamma_{0,t}^{1}) \end{aligned}$$

because of (4.38) and the definition of $B^{1,z}$.

Now, suppose that (4.40) holds for integers $0, 1, \ldots, n$. Then,

$$Y_{0,t}^{[M,n+1]} = \mathbf{1} + \int_0^t g^{M,z}(Y_{0,s}^{[M,n]})\, d\gamma_s = \mathbf{1} + \int_0^t Y_{0,s}^{[M]} \otimes \Psi^{M,z}(Y_{0,s}^{[M]})\, d\gamma_s$$

follows from (4.38) and the definition of $g^{M,z}$ and

$$\begin{aligned}
&\int_0^t Y_{0,s}^{[M,n]} \otimes \Psi^{M,z}(Y_{0,s}^{[M,n]})\, d\gamma_s \\
= &\int_0^t \left(\sum_{i=0}^{n} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) \otimes \Psi^{M,z}\left(\sum_{i=0}^{n} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) d\gamma_s \\
= &\int_0^t \left(\sum_{i=0}^{n-1} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) \otimes \Psi^{M,z}\left(\sum_{i=0}^{n-1} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) d\gamma_s \\
&\quad + \int_0^t B^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^n-1})) \otimes \Psi^{M,z}\left(\sum_{i=0}^{n} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) d\gamma_s \\
&\quad + \int_0^t \left(\sum_{i=0}^{n-1} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) \otimes \Psi^{M,z}\left(B^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^n-1}))\right) d\gamma_s
\end{aligned}$$

holds because of the induction hypothesis, the linearity of $\Psi^{M,z}$ and the bilinearity of the tensor product. Furthermore,

$$\begin{aligned}
&\int_0^t \left(\sum_{i=0}^{n-1} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) \otimes \Psi^{M,z}\left(\sum_{i=0}^{n-1} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) d\gamma_s \\
= &\int_0^t Y_{0,s}^{[M,n-1]} \otimes \Psi^{M,z}(Y_{0,s}^{[M,n-1]})\, d\gamma_s = Y_{0,t}^{[M,n]} - \mathbf{1}
\end{aligned}$$

follows from the induction hypothesis and (4.38). Thus, we have

$$\begin{aligned}
&Y_{0,t}^{[M,n+1]} \\
= &\; Y_{0,t}^{[M,n]} + \int_0^t B^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^n-1})) \otimes \Psi^{M,z}\left(\sum_{i=0}^{n} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) d\gamma_s \\
&+ \int_0^t \left(\sum_{i=0}^{n-1} B^{i,z}((\Gamma_{0,s}^i, \dots, \Gamma_{0,s}^{2^i-1}))\right) \otimes \Psi^{M,z}\left(B^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^n-1}))\right) d\gamma_s.
\end{aligned}$$

We get

$$
\begin{aligned}
& \left(B^{n,z}((\Gamma^n_{0,s},\ldots,\Gamma^{2^n-1}_{0,s}))\otimes\Psi^{M,z}\left(\sum_{i=0}^{n}B^{i,z}((\Gamma^i_{0,s},\ldots,\Gamma^{2^i-1}_{0,s}))\right)\right)(x)\\
= & \left(B^{n,z}((\Gamma^n_{0,s},\ldots,\Gamma^{2^n-1}_{0,s}))\otimes\Psi^{M,z}\left(\hat{\varphi}^{n,z}((\Gamma^0_{0,s},\ldots,\Gamma^{2^n-1}_{0,s}))\right)\right)(x)\\
= & \left(\sum_{m=1}^{M}\sum_{i_1,\ldots,i_{m-1}=1}^{d_1}e^*_{i_1,\ldots,i_{m-1}}\left(B^{n,z}((\Gamma^n_{0,s},\ldots,\Gamma^{2^n-1}_{0,s}))\right)e_{i_1}\otimes\cdots\otimes e_{i_{m-1}}\right)\\
& \otimes\left(\sum_{l=1}^{d_1}\sum_{j=1}^{d_2}\Psi^{M,z}_{lj}\left(\hat{\varphi}^{n,z}((\Gamma^0_{0,s},\ldots,\Gamma^{2^n-1}_{0,s}))\right)x_je_l\right)\\
= & \sum_{m=1}^{M}\sum_{i_1,\ldots,i_{m-1}=1}^{d_1}\sum_{l=1}^{d_1}\left(\sum_{j=1}^{d_2}e^*_{i_1,\ldots,i_{m-1}}\left(B^{n,z}((\Gamma^n_{0,s},\ldots,\Gamma^{2^n-1}_{0,s}))\right)\right.\\
& \left.\times\Psi^{M,z}_{lj}\left(\hat{\varphi}^{n,z}((\Gamma^0_{0,s},\ldots,\Gamma^{2^n-1}_{0,s}))\right)x_j\right)e_{i_1}\otimes\cdots\otimes e_{i_{m-1}}\otimes e_l\\
= & \sum_{m=1}^{M}\sum_{i_1,\ldots,i_{m-1}=1}^{d_1}\sum_{l=1}^{d_1}\left(\sum_{j=1}^{d_2}\left((e^*_{i_1,\ldots,i_{m-1}}\circ B^{n,z})((\Gamma^n_{0,s},\ldots,\Gamma^{2^n-1}_{0,s}))\right)\right.\\
& \left.\times\left((\Psi^{M,z}_{lj}\circ\hat{\varphi}^{n,z})((\Gamma^0_{0,s},\ldots,\Gamma^{2^n-1}_{0,s}))\right)x_j\right)e_{i_1}\otimes\cdots\otimes e_{i_{m-1}}\otimes e_l\\
= & \sum_{m=1}^{M}\sum_{i_1,\ldots,i_{m-1}=1}^{d_1}\sum_{l=1}^{d_1}\left(\sum_{j=1}^{d_2}\Big((e^*_{i_1,\ldots,i_{m-1}}\circ B^{n,z})\right.\\
& \left.\sqcup(\Psi^{M,z}_{lj}\circ\hat{\varphi}^{n,z})((\Gamma^n_{0,s},\ldots,\Gamma^{2^{n+1}-2}_{0,s}))\Big)x_j\right)e_{i_1}\otimes\cdots\otimes e_{i_{m-1}}\otimes e_l\\
= & \sum_{m=1}^{M}\tilde{f}^{n,z}_m((\Gamma^n_{0,s},\ldots,\Gamma^{2^{n+1}-2}_{0,s}))(x)
\end{aligned}
$$

for every $x \in \mathbb{R}^{d_2}$ by the bilinearity of the tensor product and the property (4.3) of the shuffle product. In the same way, we get

$$
\begin{aligned}
&\left(\sum_{i=0}^{n-1} B^{i,z}((\Gamma^i_{0,s},\dots,\Gamma^{2^i-1}_{0,s}))\right)\otimes\Psi^{M,z}\left(B^{n,z}((\Gamma^n_{0,s},\dots,\Gamma^{2^n-1}_{0,s}))\right)(x)\\
=\ &\left(\hat{\varphi}^{n-1,z}((\Gamma^0_{0,s},\dots,\Gamma^{2^{n-1}-1}_{0,s}))\otimes\Psi^{M,z}\left(B^{n,z}((\Gamma^n_{0,s},\dots,\Gamma^{2^n-1}_{0,s}))\right)\right)(x)\\
=\ &\left(\sum_{m=1}^{M}\sum_{i_1,\dots,i_{m-1}=1}^{d_1} e^*_{i_1,\dots,i_{m-1}}\left(\hat{\varphi}^{n-1,z}((\Gamma^0_{0,s},\dots,\Gamma^{2^{n-1}-1}_{0,s}))\right)e_{i_1}\otimes\cdots\otimes e_{i_{m-1}}\right)\\
&\otimes\left(\sum_{l=1}^{d_1}\sum_{j=1}^{d_2}\Psi^{M,z}_{lj}\left(B^{n,z}((\Gamma^n_{0,s},\dots,\Gamma^{2^n-1}_{0,s}))\right)x_j e_l\right)\\
=\ &\sum_{m=1}^{M}\sum_{i_1,\dots,i_{m-1}=1}^{d_1}\sum_{l=1}^{d_1}\left(\sum_{j=1}^{d_2} e^*_{i_1,\dots,i_{m-1}}\left(\hat{\varphi}^{n-1,z}((\Gamma^0_{0,s},\dots,\Gamma^{2^{n-1}-1}_{0,s}))\right)\right.\\
&\left.\times\Psi^{M,z}_{lj}\left(B^{n,z}((\Gamma^n_{0,s},\dots,\Gamma^{2^n-1}_{0,s}))\right)x_j\right)e_{i_1}\otimes\cdots\otimes e_{i_{m-1}}\otimes e_l\\
=\ &\sum_{m=1}^{M}\sum_{i_1,\dots,i_{m-1}=1}^{d_1}\sum_{l=1}^{d_1}\left(\sum_{j=1}^{d_2}\left((e^*_{i_1,\dots,i_{m-1}}\circ\hat{\varphi}^{n-1,z})((\Gamma^0_{0,s},\dots,\Gamma^{2^{n-1}-1}_{0,s}))\right)\right.\\
&\left.\times\left((\Psi^{M,z}_{lj}\circ B^{n,z})((\Gamma^n_{0,s},\dots,\Gamma^{2^n-1}_{0,s}))\right)x_j\right)e_{i_1}\otimes\cdots\otimes e_{i_{m-1}}\otimes e_l\\
=\ &\sum_{m=1}^{M}\sum_{i_1,\dots,i_{m-1}=1}^{d_1}\sum_{l=1}^{d_1}\left(\sum_{j=1}^{d_2}\left((e^*_{i_1,\dots,i_{m-1}}\circ\hat{\varphi}^{n-1,z})\right.\right.\\
&\left.\left.\sqcup(\Psi^{M,z}_{lj}\circ B^{n,z})((\Gamma^n_{0,s},\dots,\Gamma^{3\times 2^{n-1}-2}_{0,s}))\right)x_j\right)e_{i_1}\otimes\cdots\otimes e_{i_{m-1}}\otimes e_l\\
=\ &\sum_{m=1}^{M}\hat{f}^{n,z}_m((\Gamma^n_{0,s},\dots,\Gamma^{3\times 2^{n-1}-2}_{0,s}))(x)
\end{aligned}
$$

for every $x \in \mathbb{R}^{d_2}$. Thus, we have

$$\begin{aligned}
Y_{0,t}^{[M,n+1]} &= Y_{0,t}^{[M,n]} + \int_0^t \left(\sum_{m=1}^{M} \tilde{f}_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^{n+1}-2})) \right) d\gamma_s \\
&\quad + \int_0^t \left(\sum_{m=1}^{M} \hat{f}_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{3\times 2^{n-1}-2})) \right) d\gamma_s \\
&= Y_{0,t}^{[M,n]} + \sum_{m=1}^{M} \int_0^t \Big(\tilde{f}_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^{n+1}-2})) \\
&\qquad\qquad + \hat{f}_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{3\times 2^{n-1}-2})) \Big) \, d\gamma_s \\
&= Y_{0,t}^{[M,n]} + \sum_{m=1}^{M} \int_0^t f_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^{n+1}-2})) \, d\gamma_s \\
&= Y_{0,t}^{[M,n]} + \sum_{m=1}^{M} \sum_{i_1,\dots,i_m}^{d_1} e_{i_1,\dots,i_m}^{*} \left(\int_0^t f_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^{n+1}-2})) \, d\gamma_s \right) \\
&\qquad\qquad \times e_{i_1} \otimes \cdots \otimes e_{i_m}
\end{aligned}$$

and

$$\begin{aligned}
& e_{i_1,\dots,i_m}^{*} \left(\int_0^t f_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^{n+1}-2})) \, d\gamma_s \right) \\
&= e_{i_1,\dots,i_m}^{*} \left(\int_0^t f_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^{n+1}-2})) \, d\left(\sum_{j=1}^{d_2} \gamma_s^j e_j \right) \right) \\
&= \sum_{j=1}^{d_2} e_{i_1,\dots,i_m}^{*} \left(\int_0^t f_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^{n+1}-2}))(e_j) \, d\gamma_s^j \right) \\
&= \sum_{j=1}^{d_2} \int_0^t e_{i_1,\dots,i_m}^{*} \left(f_m^{n,z}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^{n+1}-2}))(e_j) \right) d\gamma_s^j \\
&= \sum_{j=1}^{d_2} \int_0^t \hat{\Phi}_{(i_1,\dots,i_m),j}((\Gamma_{0,s}^n, \dots, \Gamma_{0,s}^{2^{n+1}-2})) \, d\gamma_s^j \\
&= \sum_{j=1}^{d_2} \left(L_{n,2^{n+1}-2}^{j}(\hat{\Phi}_{(i_1,\dots,i_m),j}) \right) ((\Gamma_{0,s}^{n+1}, \dots, \Gamma_{0,s}^{2^{n+1}-1})).
\end{aligned}$$

Hence, we get

$$
\begin{aligned}
Y_{0,t}^{[M,n+1]} &= Y_{0,t}^{[M,n]} + \sum_{m=1}^{M} \sum_{i_1,\dots,i_m}^{d_1} \sum_{j=1}^{d_2} \left(L^{j}_{n,2^{n+1}-2}(\hat{\Phi}_{(i_1,\dots,i_m),j}) \right) ((\Gamma_{0,s}^{n+1}, \dots, \Gamma_{0,s}^{2^{n+1}-1})) \\
&\quad \times (e_{i_1} \otimes \cdots \otimes e_{i_m}) \\
&= Y_{0,t}^{[M,n]} + B^{n+1,z}((\Gamma_{0,t}^{n+1}, \dots, \Gamma_{0,t}^{2^{n+1}-1})) \\
&= \sum_{i=0}^{n+1} B^{i,z}((\Gamma_{0,t}^{i}, \Gamma_{0,t}^{i+1}, \dots, \Gamma_{0,t}^{2^{i}-1}))
\end{aligned}
$$

which completes the induction and thus shows (4.40).

Now, (4.36) follows from (4.40) and (4.39), and (4.37) follows from (4.40) and Theorem 4.5. □

4.5 Numerical computation of signatures of solutions of polynomial differential equations

If we do already know the signature of the driving signal, for example from precomputation, we can use the result of Theorem 4.17 to compute the (truncated) signature of the solution of the polynomial differential equation (4.1). In this section, we show how we can compute the signature of the solution numerically without knowing the signature of the signal.

We can solve polynomial differential equations by Picard iteration. The following theorem shows that over a short time intervall, the signatures of the Picard approximations converge to the signature of the solution.

Recall that the supremum norm on $T((\mathbb{R}^{d_1}))$ is

$$
\|\mathbf{a}\| = \sup\{|a_0|, \|a_1\|, \|a_2\|, \dots\}
$$

for $\mathbf{a} = (a_0, a_1, a_2, \dots) \in T((\mathbb{R}^{d_1}))$.

Theorem 4.18 *Consider the polynomial differential equation (4.1), i.e.*

$$dy_t = P(y_t)\, d\gamma_t, \qquad y_0 = z,$$

where $\gamma : [0,T] \to \mathbb{R}^{d_2}$ *is a continuous path with bounded variation,* $z \in \mathbb{R}^{d_1}$*, and* $P : \mathbb{R}^{d_1} \to M(d_1, d_2, \mathbb{R})$ *is a continuous function, which takes values in the* $d_1 \times d_2$*-matrices, and is given by*

$$P(y_t) = \begin{pmatrix} p_{11}(y_t) & \cdots & p_{1d_2}(y_t) \\ \vdots & \ddots & \vdots \\ p_{d_11}(y_t) & \cdots & p_{d_1d_2}(y_t) \end{pmatrix},$$

where $p_{ij} : \mathbb{R}^{d_1} \to \mathbb{R}$ *is a polynomial for every* $(i,j) \in \{1, \dots, d_1\} \times \{1, \dots, d_2\}$*.*
Let $y : \mathcal{I} \to \mathbb{R}^{d_1}$ *be its unique solution as discussed in Theorem 4.11, where* $\mathcal{I} \subset [0,T]$ *is the maximal interval of existence. Fix* $\tilde{T} \in \mathcal{I}$ *such that*

$$\|\gamma\|_{1,[0,\tilde{T}]} \leq \left(\max_{w \in \bar{B}_1(z)} \|P(w)\| \right)^{-1},$$

where $\bar{B}_1(z) \subset \mathbb{R}^{d_1}$ *is the closed ball* $\bar{B}_1(z) = \{u \in \mathbb{R}^{d_1} : \|u - z\| \leq 1\}$*.*
We can obtain the solution y *via Picard iteration. For every* $n \in \mathbb{N}$*, define paths* $y^{(n)} : [0, \tilde{T}] \to \mathbb{R}^{d_1}$ *recursively by*

$$y_t^{(n+1)} = z + \int_0^t P(y_s^{(n)})\, d\gamma_s$$

starting from

$$y_t^{(0)} = z$$

for all $t \in [0, \tilde{T}]$*. Then, the sequence* $(y_t^{(n)})_{n \in \mathbb{N}}$ *converges for every* $t \in [0, \tilde{T}]$ *and*

$$\lim_{n \to \infty} y_t^{(n)} = y_t.$$

Furthermore, let $Y_{0,\tilde{T}}$ *be the signature of the solution* y *on* $[0,\tilde{T}]$ *and let* $Y^{[(n)]}_{0,\tilde{T}}$ *be the full signature of the Picard approximation* $y^{(n)}$ *on* $[0,\tilde{T}]$*. Then,* $(Y^{[(n)]}_{0,\tilde{T}})_{n\in\mathbb{N}}$ *converges to* $Y_{0,\tilde{T}}$ *in the supremum norm. The rate of convergence is*

$$\|Y^{[(n)]}_{0,\tilde{T}} - Y_{0,\tilde{T}}\| \leq \frac{(L/M)^n}{\beta n!} e^{\beta + L/M}$$

for all integers $n > \frac{L}{M} e^{2L/M}$*, where* $\beta = 2\left(1 + 4 \times \frac{\pi^2}{6}\right)$,

$$M = \max_{w \in \bar{B}_1(z)} \|P(w)\|$$

and L *is a Lipschitz constant for* P *on* $\bar{B}_1(z)$*.*

The proof requires the following lemma.

Lemma 4.19 *With the notation and the definitions from Theorem 4.18, let* $Y_{s,t} = (1, Y^1_{s,t}, Y^2_{s,t}, \dots)$ *be the signature of the path* y*, restricted to* $[s,t]$*, and let* $Y^{[(n)]}_{s,t} = (1, Y^{[(n)],1}_{s,t}, Y^{[(n)],2}_{s,t}, \dots)$ *be the signature of* $y^{(n)}$*, restricted to* $[s,t]$*. Then,*

$$\|Y^{[(n)],1}_{s,t} - Y^1_{s,t}\| \leq M \frac{(L/M)^n}{n!} e^{L/M} \|\gamma\|_{1,[s,t]}$$

holds for all $n \in \mathbb{N}$ *and all* s *and* t *such that* $0 \leq s < t \leq \tilde{T}$*.*

Proof: Fix s and t such that $0 \leq s < t \leq \tilde{T}$. Note that $Y^1_{s,t} = y_t - y_s$ and $Y^{[(n)],1}_{s,t} = y^{(n)}_t - y^{(n)}_s$. It follows from (4.7) and Theorem 4.5 that $y^{(n)}_u \in \bar{B}_1(z)$ and $y_u \in \bar{B}_1(z)$ for all $u \in [s,t]$ and all $n \in \mathbb{N}$. Thus, for $n = 0$, we have

$$\begin{aligned} \|Y^{[(0)],1}_{s,t} - Y^1_{s,t}\| &= \|y^{(0)}_t - y^{(0)}_s - (y_t - y_s)\| = \left\| z - z - \int_s^t P(y^0_u)\, d\gamma_u \right\| \\ &\leq \|\gamma\|_{1,[s,t]} \sup_{u\in[s,t]} \|P(y_u)\| \leq M \|\gamma\|_{1,[s,t]} \leq M e^{L/M} \|\gamma\|_{1,[s,t]} \end{aligned}$$

since $L > 0$ and $M > 0$. It also follows from Theorem 4.5 that $(y_u^{(n)})_{n\in\mathbb{N}}$ converges to y_u for every $u \in [s,t]$ and that

$$\|y_u^{(n)} - y_u\| \leq \frac{(L/M)^n}{(n+1)!} e^{L/M}$$

holds for every $u \in [s,t]$ and every $n \in \mathbb{N}$. Therefore, we get

$$\begin{aligned}
\|Y_{s,t}^{[(n)],1} - Y_{s,t}^1\| &= \|y_t^{(n)} - y_s^{(n)} - (y_t - y_s)\| \\
&= \left\| \int_s^t P(y_u^{(n-1)})\, d\gamma_u - \int_s^t P(y_u)\, d\gamma_u \right\| \\
&= \left\| \int_s^t (P(y_u^{(n-1)}) - P(y_u))\, d\gamma_u \right\| \\
&\leq \|\gamma\|_{1,[s,t]} \sup_{u\in[s,t]} \|P(y_u^{(n-1)}) - P(y_u)\| \\
&\leq \|\gamma\|_{1,[s,t]} \sup_{u\in[s,t]} L\|y_u^{(n-1)} - y_u\| \\
&\leq \|\gamma\|_{1,[s,t]} L \frac{(L/M)^{n-1}}{n!} e^{L/M} \\
&= M \frac{(L/M)^n}{n!} e^{L/M} \|\gamma\|_{1,[s,t]}
\end{aligned}$$

for $n \geq 1$ because of the recursive definition of $y^{(n)}$ and the local Lipschitz continuity of P, which completes the proof. □

Now, we can prove Theorem 4.18.

Proof of Theorem 4.18: The proof is based on results in Lyons et al. [37] and we use their notation. Set $\Delta_{\tilde{T}} = \{(s,t) \in [0,\tilde{T}]^2 : 0 \leq s \leq t \leq \tilde{T}\}$ and define $\omega : \Delta_{\tilde{T}} \to \mathbb{R}$ by

$$\omega(s,t) = \beta M \|\gamma\|_{1,[s,t]}.$$

Then, ω is a control function.

It follows from Theorem 4.5 that there exists a unique solution y of the polynomial differential equation (4.1) and that the Picard approximations $(y_t^{(n)})_{n\in\mathbb{N}}$ converge to y_t for every $t \in [0, \tilde{T}]$. The solution y is continuous and has bounded variation (see also Theorem 4.5). By definition, $y^{(n)}$ is also continuous and has bounded variation for every $n \in \mathbb{N}$.

Fix $n \in \mathbb{N}$ such that $n > \frac{L}{M} e^{2L/M}$. Note that this implies $n > 0$. For all $(s,t) \in \Delta_{\tilde{T}}$, let $Y_{s,t} = (1, Y_{s,t}^1, Y_{s,t}^2, \dots)$ denote the signature of y, restricted to $[s,t]$, and let $Y_{s,t}^{[(n)]} = (1, Y_{s,t}^{[(n)],1}, Y_{s,t}^{[(n)],2}, \dots)$ denote the signature of $y^{(n)}$, restricted to $[s,t]$. Then,

$$Y_{s,t}^1 = y_t - y_s$$

and

$$Y_{s,t}^{[(n)],1} = y_t^{(n)} - y_s^{(n)}$$

hold for all $(s,t) \in \Delta_{\tilde{T}}$, and $(1, Y_{s,t}^1)_{(s,t)\in\Delta_{\tilde{T}}}$ and $(1, Y_{s,t}^{[(n)],1})_{(s,t)\in\Delta_{\tilde{T}}}$ are both multiplicative functionals in $T^{(1)}(\mathbb{R}^{d_1})$ with finite 1-variation.

Furthermore, it follows from (4.7) and Theorem 4.5 that $y_u^{(n)} \in \bar{B}_1(z)$ and $y_u \in \bar{B}_1(z)$ for all $u \in [0, \tilde{T}]$. Thus, we have

$$\|y_t - y_s\| = \left\| \int_s^t P(y_u)\, d\gamma_u \right\| \leq \|\gamma\|_{1,[s,t]} \sup_{u\in[s,t]} \|P(y_u)\| \leq M \|\gamma\|_{1,[s,t]} = \frac{1}{\beta}\omega(s,t)$$

and

$$\begin{aligned} \|y_t^{(n)} - y_s^{(n)}\| &= \left\| \int_s^t P(y_u^{(n-1)})\, d\gamma_u \right\| \leq \|\gamma\|_{1,[s,t]} \sup_{u\in[s,t]} \|P(y_u^{(n-1)})\| \\ &\leq M \|\gamma\|_{1,[s,t]} = \frac{1}{\beta}\omega(s,t) \end{aligned}$$

for all $(s,t) \in \Delta_{\tilde{T}}$. Therefore, both multiplicative functionals $(1, Y_{s,t}^1)_{(s,t)\in\Delta_{\tilde{T}}}$ and $(1, Y_{s,t}^{[(n)],1})_{(s,t)\in\Delta_{\tilde{T}}}$ have finite 1-variation controlled by ω in the sense of [37], Definition 3.6.

The full signatures $(Y_{s,t})_{(s,t)\in\Delta_{\tilde{T}}}$ and $(Y_{s,t}^{[(n)]})_{(s,t)\in\Delta_{\tilde{T}}}$ are their unique extensions to multiplicative functionals in $T((\mathbb{R}^{d_1}))$ and they also have finite 1-variation controlled by ω (see [37], Theorem 3.7).

Set

$$\varepsilon = \frac{(L/M)^n}{n!} e^{L/M}.$$

Then, we have

$$2\left(1 + \sum_{r=3}^{\infty}\left(\frac{2}{r-2}\right)^2\right) = 2\left(1 + 4\sum_{r=1}^{\infty}\frac{1}{r^2}\right) = 2\left(1 + 4 \times \frac{\pi^2}{6}\right) = \beta$$

as $\sum_{r=1}^{\infty} \frac{1}{r^2} = \frac{\pi^2}{6}$ (see for example [3]). In addition,

$$0 < \varepsilon < 1$$

holds since

$$0 < \varepsilon = \frac{L/M}{n}\frac{(L/M)^{n-1}}{(n-1)!}e^{L/M} < \frac{L/M}{n}e^{L/M}e^{L/M} = \frac{L/M}{n}e^{2L/M} < 1$$

because

$$n > \frac{L}{M}e^{2L/M}.$$

By Lemma 4.19, we have

$$\|Y_{s,t}^{[(n)],1} - Y_{s,t}^{1}\| \leq \varepsilon\frac{\omega(s,t)}{\beta}$$

for all $(s,t) \in \Delta_{\tilde{T}}$. Thus, it follows from [37], Theorem 3.10, that

$$\|Y_{s,t}^{[(n)],i} - Y_{s,t}^{i}\| \leq \varepsilon\frac{\omega(s,t)^i}{\beta i!}$$

holds for every $i \in \mathbb{N}$ and every $(s,t) \in \Delta_{\tilde{T}}$. Therefore, we get

$$\begin{aligned} \|Y_{0,\tilde{T}}^{[(n)]} - Y_{0,\tilde{T}}\| &= \sup_{i\geq 0} \varepsilon \frac{\omega(0,\tilde{T})^i}{\beta i!} = \frac{\varepsilon}{\beta} \sup_{i\geq 0} \frac{(\beta M \|\gamma\|_{1,[0,\tilde{T}]})^i}{i!} \\ &\leq \frac{\varepsilon}{\beta} \sup_{i\geq 0} \frac{\beta^i}{i!} \leq \frac{\varepsilon}{\beta} e^{\beta} = \frac{(L/M)^n}{\beta n!} e^{(L/M)+\beta}. \end{aligned}$$

Now, it is an immediate consequence that $(Y_{0,\tilde{T}}^{[(n)]})_{n\in\mathbb{N}}$ converges to $Y_{0,\tilde{T}}$ in the supremum norm. □

So, Theorem 4.18 shows that the signatures of the Picard approximations converge to the signature of the solution. We can use this result to compute an approximation for the signature of the solution. As an example, the following theorem provides a numerical algorithm for the case that the driving signal is $\gamma_t = (t,\ldots,t)^\top$. The basic idea is to approximate the Picard approximations with piecewise linear paths and to compute the signatures of these piecewise linear approximations.

Theorem 4.20 *Consider now the polynomial differential equation*

$$dy_t = P(y_t)\,dt, \qquad y_0 = z, \tag{4.41}$$

where $z \in \mathbb{R}^d$ for some integer $d \geq 1$ and $P : \mathbb{R}^d \to \mathbb{R}^d$ is a continuous function, which is given by

$$P(y_t) = \begin{pmatrix} P_1(y_t) \\ \vdots \\ P_d(y_t) \end{pmatrix},$$

where $P_i : \mathbb{R}^d \to \mathbb{R}$ is a polynomial for every $i \in \{1,\ldots,d\}$.
Let $y : \mathcal{I} \to \mathbb{R}^d$ be its unique solution as discussed in Theorem 4.11, where $\mathcal{I} \subset [0,T]$

is the maximal interval of existence. Set

$$M = \max_{x \in \bar{B}_2(z)} \|P(x)\|$$

$$M' = \max_{j \in \{1,\dots,d\}} \max_{x \in \bar{B}_2(z)} \sum_{k=1}^{d} \left| \frac{dP_j}{dx_k}(x) \right|$$

and

$$M'' = \max_{j \in \{1,\dots,d\}} \max_{x \in \bar{B}_2(z)} \sum_{k=1}^{d} \sum_{l=1}^{d} \left| \frac{d^2 P_j}{dx_k\, dx_l}(x) \right|,$$

where $\bar{B}_2(z) \subset \mathbb{R}^d$ *is the closed ball* $\bar{B}_2(z) = \{u \in \mathbb{R}^d : \|u - z\| \leq 2\}$.

Fix $\tilde{T} \in \mathcal{I}$ *such that*

$$\tilde{T} \leq \frac{1}{M} \qquad \text{and} \qquad \tilde{T} < \frac{1}{L},$$

where L *is a Lipschitz constant for* P *on* $\bar{B}_2(z)$.

Let $Y_{0,\tilde{T}}$ *be the signature of the solution* y*, restricted to* $[0, \tilde{T}]$*. Define the polynomial* $\hat{P}$ *by* $\hat{P}(x) = P(x + z)$ *for all* $x \in \mathbb{R}^d$*. Let* $\Phi : T((\mathbb{R}^d)) \to \mathbb{R}^d$ *be the linear map associated with* $\hat{P}$ *(as in Theorem 4.2).*

Let $m \in \mathbb{N}$ *and let* $0 = t_0 < t_1 < \dots < t_m = \tilde{T}$ *be an equally spaced partition of* $[0, \tilde{T}]$*, i.e.* $t_k = k\frac{\tilde{T}}{m}$*. Define sequences* $(\hat{Y}^{[(m,n)]}_{0,t_k})_{n \in \mathbb{N}} \subset T((\mathbb{R}^d))$ *for* $k = 0, 1, \dots, m$ *recursively in the following way:*

- *Start with* $\hat{Y}^{[(m,0)]}_{0,t_k} = (1, 0, 0, \dots)$ *for every* $k \in \{0, 1, \dots, m\}$.
- *Given* $\hat{Y}^{[(m,n)]}_{0,t_0}$, $\hat{Y}^{[(m,n)]}_{0,t_1}$, $\dots$, $\hat{Y}^{[(m,n)]}_{0,t_m}$*, define*

$$\hat{y}^{m,n+1}_{t_k} = \hat{y}^{m,n+1}_{t_{k-1}} + \frac{1}{2m}(\Phi(\hat{Y}^{[(m,n)]}_{0,t_k}) + \Phi(\hat{Y}^{[(m,n)]}_{0,t_{k-1}}))$$

  recursively for $k \in \{1, \dots, m\}$ *starting from* $\hat{y}^{m,n+1}_{t_0} = z$.

  Then, set $\hat{Y}^{[(m,n+1)]}_{0,t_0} = (1, 0, 0, \dots)$ *and*

$$\hat{Y}^{[(m,n+1)]}_{0,t_k} = \exp(\hat{y}^{m,n+1}_{t_1} - \hat{y}^{m,n+1}_{t_0}) \otimes \dots \otimes \exp(\hat{y}^{m,n+1}_{t_k} - \hat{y}^{m,n+1}_{t_{k-1}})$$

for $k = 1, \ldots, m$*, where we write* $\exp(\hat{y}^{m,n+1}_{t_i} - \hat{y}^{m,n+1}_{t_{i-1}})$ *for the expression* $\exp((0, \hat{y}^{m,n+1}_{t_i} - \hat{y}^{m,n+1}_{t_{i-1}}, 0, 0, \ldots))$.

Then, $\hat{Y}^{[(m,n)]}_{0,\tilde{T}}$ *converges to* $Y_{0,\tilde{T}}$ *in the sense that*

$$\lim_{n\to\infty} \lim_{m\to\infty} \|\hat{Y}^{[(m,n)]}_{0,\tilde{T}} - Y_{0,\tilde{T}}\| = 0. \tag{4.42}$$

Define constants $c_0 = M^2 M''$*,* $c_1 = M'M$ *and* $c = c_0\tilde{T} + 2c_1$*. Set* $\beta = 2\left(1 + 4 \times \frac{\pi^2}{6}\right)$*. Then, the rate of convergence is*

$$\|\hat{Y}^{[(m,n)]}_{0,\tilde{T}} - Y_{0,\tilde{T}}\| \leq \frac{e^{\beta(c\tilde{T}^2+1)} - 1}{\beta(c\tilde{T} + M)} \left(\frac{c\tilde{T}}{m} \frac{1}{1 - L\tilde{T}} + M_1 \frac{(L_1/M_1)^n}{n!} e^{L_1/M_1} \right) \tag{4.43}$$

for all integers $m > \frac{1}{1-L\tilde{T}} \max\left\{2\frac{c\tilde{T}}{c\tilde{T}+M}, c\tilde{T}^2\right\}$ *and* $n > 2\frac{L_1}{c\tilde{T}+M} e^{2L_1/M_1}$*, where*

$$M_1 = \max_{x \in \bar{B}_1(z)} \|P(x)\|,$$

L_1 *is a Lipschitz constant for* P *on* $\bar{B}_1(z)$ *and* $\bar{B}_1(z) \subset \mathbb{R}^d$ *is the closed ball* $\bar{B}_1(z) = \{u \in \mathbb{R}^d : \|u - z\| \leq 1\}$.

Recall that the map Φ is defined in such a way that

$$\Phi(X_{0,t}) = \hat{P}(X_t - X_0)$$

holds for every continuous path $X : [0, T] \to \mathbb{R}^d$ with bounded variation and every $t \in [0, T]$, where $X_{0,t}$ denotes the signature of X, restricted to $[0, t]$.

Proof of Theorem 4.20: Fix $m > \frac{1}{1-L\tilde{T}} \max\left\{2\frac{c\tilde{T}}{c\tilde{T}+M}, c\tilde{T}^2\right\}$. Define paths

$$y^{(n)},\ \tilde{y}^{(m,n)},\ \bar{y}^{(m,n)} \text{ and } \hat{y}^{(m,n)} : [0, \tilde{T}] \to \mathbb{R}^d$$

recursively for every $n \in \mathbb{N}$ in the following way: Start from

$$y_t^{(0)} = \tilde{y}_t^{(m,0)} = \bar{y}_t^{(m,0)} = \hat{y}_t^{(m,0)} = z$$

for every $t \in [0, \tilde{T}]$. Given $y^{(n)}$, $\tilde{y}^{(m,n)}$, $\bar{y}^{(m,n)}$ and $\hat{y}^{(m,n)}$, set

$$\begin{aligned} y_t^{(n+1)} &= z + \int_0^t P(y_s^{(n)})\, ds, \\ \tilde{y}_t^{(m,n+1)} &= z + \int_0^t P(\hat{y}_s^{(m,n)})\, ds \end{aligned}$$

and let $\bar{y}^{(m,n+1)}$ be the following approximation of $\tilde{y}^{(m,n+1)}$: Set

$$g^{(m,n,i)}(u) = P(\hat{y}_{t_i}^{(m,n)}) + \frac{P(\hat{y}_{t_{i+1}}^{(m,n)}) - P(\hat{y}_{t_i}^{(m,n)})}{t_{i+1} - t_i}(u - t_i)$$

and

$$\begin{aligned} \bar{y}_t^{(m,n+1)} &= z + \left(\sum_{k=0}^{i-1} \frac{1}{2} \left(P(\tilde{y}_{t_k}^{(m,n+1)}) + P(\tilde{y}_{t_{k+1}}^{(m,n+1)}) \right) (t_{k+1} - t_k) \right) \\ &\quad + \int_{t_i}^t g^{(m,n,i)}(u)\, du \end{aligned}$$

for $t \in (t_i, t_{i+1}]$, $i = 0, 1, \ldots, m-1$. Note that

$$\int_{t_i}^{t_{i+1}} g^{(m,n,i)}(u)\, du = \frac{1}{2} \left(P(\tilde{y}_{t_i}^{(m,n+1)}) + P(\tilde{y}_{t_{i+1}}^{(m,n+1)}) \right) (t_{i+1} - t_i)$$

and if $t = t_i$ for some $i \in \{1, \ldots, m\}$, we have

$$\bar{y}_{t_i}^{(m,n+1)} = z + \left(\sum_{k=0}^{i-1} \frac{1}{2} \left(P(\tilde{y}_{t_k}^{(m,n+1)}) + P(\tilde{y}_{t_{k+1}}^{(m,n+1)}) \right) (t_{k+1} - t_k) \right),$$

which corresponds to the approximation of $\tilde{y}^{(m,n+1)}$ that we get by applying the compound trapezoidal rule to the integration (see for example [42]).

Furthermore, let $\hat{y}^{(m,n+1)}$ be the piecewise linear approximation of $\bar{y}^{(m,n+1)}$. Set

$$\hat{y}_{t_i}^{(m,n+1)} = \bar{y}_{t_i}^{(m,n+1)}$$

for all $i = 0, 1, \ldots, m$ and

$$\hat{y}_t^{(m,n+1)} = \hat{y}_{t_i}^{(m,n+1)} + \frac{\hat{y}_{t_{i+1}}^{(m,n+1)} - \hat{y}_{t_i}^{(m,n+1)}}{t_{i+1} - t_i}(t - t_i)$$

for $t \in (t_i, t_{i+1})$, $i = 0, 1, \ldots, m-1$.

Then, $y^{(n)}$ are the Picard approximations for the solution y and $\hat{y}^{(m,n)}$ are their piecewise linear approximations that we actually compute. Note that

$$\Phi(\hat{Y}_{0,t}^{[(m,n)]}) = \hat{P}(\hat{y}_t^{(m,n)} - \hat{y}_0^{(m,n)}) = P(\hat{y}_t^{(m,n)} - \hat{y}_0^{(m,n)} + z) = P(\hat{y}_t^{(m,n)})$$

holds for all $t \in [0, \tilde{T}]$ and all $n \in \mathbb{N}$.

At first, we show by induction over n that

$$|(\hat{y}_t^{(m,n),j} - \hat{y}_s^{(m,n),j}) - (y_t^{(n),j} - y_s^{(n),j})| \leq (t-s)\frac{c\tilde{T}}{m}\sum_{k=0}^{n-1} L^k (t-s)^k \tag{4.44}$$

holds for all integers $n \geq 1$, all $(s,t) \in \Delta_{\tilde{T}}$ and all components $j \in \{1, \ldots, d\}$. Recall that $\Delta_{\tilde{T}} = \{(s,t) \in [0,\tilde{T}]^2 : 0 \leq s \leq t \leq \tilde{T}\}$.

Set

$$r_n = 1 + \frac{c\tilde{T}^2}{m}\sum_{k=0}^{n-1} L^k \tilde{T}^k.$$

Note that (4.44) implies that

$$\hat{y}_t^{(m,n)} \in \bar{B}_{r_n}(z) \subset \bar{B}_2(z)$$

for all integers $n \geq 1$ and all $t \in [0, \tilde{T}]$ since $y_t^{(n)} \in \bar{B}_1(z)$ for every $t \in [0, \tilde{T}]$ and

$$\begin{aligned}
|\hat{y}_t^{(m,n),j} - y_t^{(n),j}| &= |(\hat{y}_t^{(m,n),j} - z_j) - (y_t^{(n),j} - z_j)| \\
&= |(\hat{y}_t^{(m,n),j} - \hat{y}_0^{(m,n),j}) - (y_t^{(n),j} - y_0^{(n),j})| \\
&\leq r_n - 1
\end{aligned}$$

and

$$r_n < 1 + \frac{c\tilde{T}^2}{m} \sum_{k=0}^{\infty} L^k \tilde{T}^k = 1 + \frac{c\tilde{T}^2}{m} \frac{1}{1 - L\tilde{T}} \leq 2$$

as $\tilde{T} < \frac{1}{L}$ and $m > \frac{1}{1-L\tilde{T}} c\tilde{T}^2$.

For $n = 1$, we have

$$y_t^{(1)} = z + \int_0^t P(y_u^{(0)})\, du = z + \int_0^t P(z)\, du = z + tP(z)$$

and

$$\tilde{y}_t^{(m,1)} = z + \int_0^t P(\hat{y}_u^{(m,0)})\, du = z + \int_0^t P(z)\, du = z + tP(z)$$

as $y_u^{(0)} = \hat{y}_u^{(m,0)} = z$ for all $u \in [0, \tilde{T}]$. Since $\hat{y}_u^{(m,0)} = z$ is constant, we also get

$$\bar{y}_t^{(m,1)} = \tilde{y}_t^{(m,0)} = z + tP(z)$$

and

$$\hat{y}_t^{(m,1)} = \bar{y}_t^{(m,0)} = z + tP(z).$$

Therefore, we have

$$\begin{aligned}
|(\hat{y}_t^{(m,1),j} - \hat{y}_s^{(m,1),j}) - (y_t^{(1),j} - y_s^{(1),j})| &= |(t-s)P_j(z) - (t-s)P_j(z)| \\
&= 0 \leq (t-s)\frac{c\tilde{T}}{m}
\end{aligned}$$

for every component $j \in \{1, \ldots, d\}$ and every $(s,t) \in \Delta_{\tilde{T}}$.

Now, assume that (4.44) holds for all integers $1, \ldots, n$. Recall that

$$g_j^{(m,n,i)}(u) = P_j(\hat{y}_{t_i}^{(m,n)}) + \frac{P_j(\hat{y}_{t_{i+1}}^{(m,n)}) - P_j(\hat{y}_{t_i}^{(m,n)})}{t_{i+1} - t_i}(u - t_i).$$

Then, for $(s,t) \in \Delta_{\tilde{T}}$ such that there exists an integer $i \in \{0, 1, \ldots, m-1\}$ such that $t_i \le s \le t \le t_{i+1}$, we get

$$\begin{aligned}
& |(\tilde{y}_t^{(m,n+1),j} - \tilde{y}_s^{(m,n+1),j}) - (\bar{y}_t^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j})| \\
= & \left| \int_s^t P_j(\hat{y}_u^{(m,n)})\, du - \int_s^t g_j^{(m,n,i)}(u)\, du \right| \\
= & \left| \int_s^t (P_j(\hat{y}_u^{(m,n)}) - g_j^{(m,n,i)}(u))\, du \right| \\
\le & \int_s^t |P_j(\hat{y}_u^{(m,n)}) - g_j^{(m,n,i)}(u)|\, du \\
\le & (t-s) \max_{u \in [s,t]} |P_j(\hat{y}_u^{(m,n)}) - g_j^{(m,n,i)}(u)|
\end{aligned}$$

for every component $j \in \{1, \ldots, d\}$.

For the simplicity of notation, set $f_j^{(m,n)}(u) = P_j(\hat{y}_u^{(m,n)})$. By Taylor's formula, there exist $\xi_1 \in (t_i, u)$ and $\xi_2 \in (t_i, t_{i+1})$ such that

$$f_j^{(m,n)}(u) = f_j^{(m,n)}(t_i) + f_j'^{(m,n)}(\xi_1)(u - t_i)$$

and

$$\frac{P_j(\hat{y}_{t_{i+1}}^{(m,n)}) - P_j(\hat{y}_{t_i}^{(m,n)})}{t_{i+1} - t_i} = \frac{f_j^{(m,n)}(t_{i+1}) - f_j^{(m,n)}(t_i)}{t_{i+1} - t_i} = f_j'^{(m,n)}(\xi_2), \tag{4.45}$$

and thus

$$g_j^{(m,n,i)}(u) = f_j^{(m,n)}(t_i) + f_j'^{(m,n)}(\xi_2)(u - t_i).$$

Again by Taylor's formula, there exists ξ_3 between ξ_1 and ξ_2 such that

$$f_j'^{(m,n)}(\xi_1) = f_j'^{(m,n)}(\xi_2) + f_j''^{(m,n)}(\xi_3)(\xi_1 - \xi_2).$$

Therefore, we get

$$\begin{aligned} |P_j(\hat{y}_u^{(m,n)}) - g_j^{(m,n,i)}(u)| &= |f_j^{(m,n)}(u) - g_j^{(m,n,i)}(u)| \\ &= (u - t_i)\,|f_j'^{(m,n)}(\xi_1) - f_j'^{(m,n)}(\xi_2)| \\ &= (u - t_i)(\xi_1 - \xi_2)\,|f_j''^{(m,n)}(\xi_3)| \\ &\leq \frac{\tilde{T}^2}{m^2}|f_j''^{(m,n)}(\xi_3)| \end{aligned}$$

for $u \in (t_i, t_{i+1})$. Note that $P_j(\hat{y}_{t_i}^{(m,n)}) = g_j^{(m,n,i)}(t_i)$ and $P_j(\hat{y}_{t_{i+1}}^{(m,n)}) = g_j^{(m,n,i)}(t_{i+1})$. We have

$$f_j'^{(m,n)}(\xi_3) = \sum_{k=1}^{d} \frac{dP_j}{dx_k}(\hat{y}_{\xi_3}^{(m,n)})\,\dot{\hat{y}}_{\xi_3}^{(m,n),k}. \tag{4.46}$$

Note that $\hat{y}^{(m,n)}$ is piecewise linear and thus $\dot{\hat{y}}^{(m,n),k}$ is constant within (t_i, t_{i+1}). Therefore, we have

$$f_j''^{(m,n)}(\xi_3) = \sum_{k=1}^{d}\sum_{l=1}^{d} \frac{d^2 P_j}{dx_k\,dx_l}(\hat{y}_{\xi_3}^{(m,n)})\,\dot{\hat{y}}_{\xi_3}^{(m,n),l}\,\dot{\hat{y}}_{\xi_3}^{(m,n),k}.$$

By definition, we have

$$\begin{aligned} \dot{\hat{y}}_u^{(m,n),h} &= \frac{\hat{y}_{t_{i+1}}^{(m,n),h} - \hat{y}_{t_i}^{(m,n),h}}{t_{i+1} - t_i} = \frac{\bar{y}_{t_{i+1}}^{(m,n),h} - \bar{y}_{t_i}^{(m,n),h}}{t_{i+1} - t_i} \\ &= \frac{1}{t_{i+1} - t_i}\left(\frac{1}{2}\left(P_h(\tilde{y}_{t_{i+1}}^{(m,n)}) + P_h(\tilde{y}_{t_i}^{(m,n)})\right)(t_{i+1} - t_i)\right) \\ &= \frac{1}{2}\left(P_h(\tilde{y}_{t_{i+1}}^{(m,n)}) + P_h(\tilde{y}_{t_i}^{(m,n)})\right) \end{aligned}$$

and thus

$$|\dot{\hat{y}}_u^{(m,n),h}| \leq M \tag{4.47}$$

for every $u \in (t_i, t_{i+1})$ and every $h \in \{1, \dots, d\}$, since $\tilde{y}^{(m,n)}_{t_i}, \tilde{y}^{(m,n)}_{t_{i+1}} \in \bar{B}_2(z)$, which follows immediately from the fact that

$$\|\tilde{y}^{(m,n)}_u - z\| = \left\| \int_0^u P(\hat{y}^{(m,n-1)}_v)\, dv \right\| \le u \max_{v \in [0,u]} \|P(\hat{y}^{(m,n-1)}_v)\| \le \tilde{T} M \le 1$$

for every time $u \in [0, \tilde{T}]$ as $\hat{y}^{(m,n-1)}_v \in \bar{B}_2(z)$ for every $v \in [0, \tilde{T}]$. The fact that $\hat{y}^{(m,n-1)}_v \in \bar{B}_2(z)$ is a consequence of the induction hypothesis if $n \ge 2$ (as shown above) or it follows from $\hat{y}^{(m,0)}_v = z \in \bar{B}_2(z)$ if $n = 1$.

Therefore, we get

$$|f''^{(m,n)}_j(\xi_3)| \le M^2 \left(\max_{x \in \bar{B}_2(z)} \sum_{k=1}^{d} \sum_{l=1}^{d} \left| \frac{d^2 P_j}{dx_k\, dx_l}(x) \right| \right) \le c_0.$$

Hence, it follows that

$$|(\tilde{y}^{(m,n+1),j}_t - \tilde{y}^{(m,n+1),j}_s) - (\bar{y}^{(m,n+1),j}_t - \bar{y}^{(m,n+1),j}_s)| \le (t-s) c_0 \frac{\tilde{T}^2}{m^2}. \tag{4.48}$$

for all $(s,t) \in \Delta_{\tilde{T}}$ such that there exists an integer $i \in \{0, 1, \dots, m-1\}$ such that $t_i \le s < t \le t_{i+1}$.

For the other $(s,t) \in \Delta_{\tilde{T}}$, i.e. those ones for which there exist $l_1, l_2 \in \{1, \dots, m-1\}$

such that $t_{l_1-1} \leq s < t_{l_1} \leq t_{l_2} < t \leq t_{l_2+1}$, it follows from (4.48) that

$$
\begin{aligned}
& |(\tilde{y}_t^{(m,n+1),j} - \tilde{y}_s^{(m,n+1),j}) - (\bar{y}_t^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j})| \\
\leq\ & |(\tilde{y}_{t_{l_1}}^{(m,n+1),j} - \tilde{y}_s^{(m,n+1),j}) - (\bar{y}_{t_{l_1}}^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j})| \\
& + \sum_{i=l_1}^{l_2-1} |(\tilde{y}_{t_{i+1}}^{(m,n+1),j} - \tilde{y}_{t_i}^{(m,n+1),j}) - (\bar{y}_{t_{i+1}}^{(m,n+1),j} - \bar{y}_{t_i}^{(m,n+1),j})| \\
& + |(\tilde{y}_t^{(m,n+1),j} - \tilde{y}_{t_{l_2}}^{(m,n+1),j}) - (\bar{y}_t^{(m,n+1),j} - \bar{y}_{t_{l_2}}^{(m,n+1),j})| \\
\leq\ & (t_{l_1} - s)c_0 \frac{\tilde{T}^2}{m^2} + \left(\sum_{i=l_1}^{l_2-1} (t_{i+1} - t_i) c_0 \frac{\tilde{T}^2}{m^2} \right) + (t - t_{l_2}) c_0 \frac{\tilde{T}^2}{m^2} \\
\leq\ & \frac{\tilde{T}}{m} c_0 \frac{\tilde{T}^2}{m^2} + \left(\sum_{i=l_1}^{l_2-1} \frac{\tilde{T}}{m} c_0 \frac{\tilde{T}^2}{m^2} \right) + \frac{\tilde{T}}{m} c_0 \frac{\tilde{T}^2}{m^2} \leq m \frac{\tilde{T}}{m} c_0 \frac{\tilde{T}^2}{m^2} \\
\leq\ & (t - s) c_0 \frac{\tilde{T}^2}{m}
\end{aligned}
$$

since $t - s \geq \frac{\tilde{T}}{m}$. Hence, we have

$$
|(\tilde{y}_t^{(m,n+1),j} - \tilde{y}_s^{(m,n+1),j}) - (\bar{y}_t^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j})| \leq (t - s) c_0 \frac{\tilde{T}^2}{m} \tag{4.49}
$$

for every $(s,t) \in \Delta_{\tilde{T}}$ as $m \geq 1$.

Moreover, we show that

$$
|(\bar{y}_t^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j}) - (\hat{y}_t^{(m,n+1),j} - \hat{y}_s^{(m,n+1),j})| \leq 2(t - s) c_1 \frac{\tilde{T}}{m} \tag{4.50}
$$

for all $(s,t) \in \Delta_{\tilde{T}}$ in the following way: First, consider again $(s,t) \in \Delta_{\tilde{T}}$ such that there exists $i \in \{0, 1, \ldots, m-1\}$ such that $t_i \leq s < t \leq t_{i+1}$. By Taylor's Theorem, there exist $\xi_4 \in (s,t)$ and $\xi_5 \in (t_i, t_{i+1})$ such that

$$
\bar{y}_t^{(m,n+1),j} = \bar{y}_s^{(m,n+1),j} + \dot{\bar{y}}_{\xi_4}^{(m,n+1),j} (t - s) \tag{4.51}
$$

and

$$\bar{y}_{t_{i+1}}^{(m,n+1),j} = \bar{y}_{t_i}^{(m,n+1),j} + \dot{\bar{y}}_{\xi_5}^{(m,n+1),j}(t_{i+1} - t_i). \tag{4.52}$$

Then, again by Taylor's Theorem, there exists ξ_6 between ξ_4 and ξ_5 such that

$$\dot{\bar{y}}_{\xi_4}^{(m,n+1),j} = \dot{\bar{y}}_{\xi_5}^{(m,n+1),j} + \ddot{\bar{y}}_{\xi_6}^{(m,n+1),j}(\xi_4 - \xi_5). \tag{4.53}$$

From the definitions of $\bar{y}^{(m,n+1)}$ and $\tilde{y}^{(m,n+1)}$, we get

$$\dot{\bar{y}}_u^{(m,n+1),j} = g_j^{(m,n,i)}(u)$$

and

$$\ddot{\bar{y}}_u^{(m,n+1),j} = \frac{P_j(\hat{y}_{t_{i+1}}^{(m,n)}) - P_j(\hat{y}_{t_i}^{(m,n)})}{t_{i+1} - t_i}$$

for every $u \in (t_i, t_{i+1})$.

It follows from (4.45), (4.46) and (4.47) that

$$|\ddot{\bar{y}}_u^{(m,n+1),j}| \leq M'M = c_1 \tag{4.54}$$

for every $u \in (t_i, t_{i+1})$ and every $j \in \{1, \ldots, d\}$.

Hence, we obtain

$$\begin{aligned}
& |(\bar{y}_t^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j}) - (\hat{y}_t^{(m,n+1),j} - \hat{y}_s^{(m,n+1),j})| \\
= & \left|(\bar{y}_t^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j}) - \frac{\bar{y}_{t_{i+1}}^{(m,n+1),j} - \bar{y}_{t_i}^{(m,n+1),j}}{t_{i+1} - t_i}(t - s)\right| \\
= & (t-s)|\dot{\bar{y}}_{\xi_1}^{(m,n+1),j} - \dot{\bar{y}}_{\xi_2}^{(m,n+1),j}| \\
= & (t-s)\ddot{\bar{y}}_{\xi_3}^{(m,n+1),j}(\xi_1 - \xi_2) \\
\leq & (t-s)c_1\frac{\tilde{T}}{m}
\end{aligned}$$

by (4.51), (4.52), (4.53) and (4.54) for all $(s,t) \in \Delta_{\tilde{T}}$ such that there exists an integer $i \in \{0,1,\dots,m-1\}$ such that $t_i \le s < t \le t_{i+1}$.

Note that

$$\bar{y}_{t_i}^{(m,n+1),j} = \hat{y}_{t_i}^{(m,n+1),j}$$

holds for all $i \in \{0,1,\dots,m\}$. Thus, it follows for all the other $(s,t) \in \Delta_{\tilde{T}}$, i.e. those ones for which there exist $l_1, l_2 \in \{1,\dots,m-1\}$ such that

$$t_{l_1-1} \le s < t_{l_1} \le t_{l_2} < t \le t_{l_2+1},$$

that

$$\begin{aligned}
& |(\bar{y}_t^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j}) - (\hat{y}_t^{(m,n+1),j} - \hat{y}_s^{(m,n+1),j})| \\
\le\ & |(\bar{y}_{t_{l_1}}^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j}) - (\hat{y}_{t_{l_1}}^{(m,n+1),j} - \hat{y}_s^{(m,n+1),j})| \\
& + \sum_{i=l_1}^{l_2-1} |(\bar{y}_{t_{i+1}}^{(m,n+1),j} - \bar{y}_{t_i}^{(m,n+1),j}) - (\hat{y}_{t_{i+1}}^{(m,n+1),j} - \hat{y}_{t_i}^{(m,n+1),j})| \\
& + |(\bar{y}_t^{(m,n+1),j} - \bar{y}_{t_{l_2}}^{(m,n+1),j}) - (\hat{y}_t^{(m,n+1),j} - \hat{y}_{t_{l_2}}^{(m,n+1),j})| \\
=\ & |(\bar{y}_{t_{l_1}}^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j}) - (\hat{y}_{t_{l_1}}^{(m,n+1),j} - \hat{y}_s^{(m,n+1),j})| \\
& + |(\bar{y}_t^{(m,n+1),j} - \bar{y}_{t_{l_2}}^{(m,n+1),j}) - (\hat{y}_t^{(m,n+1),j} - \hat{y}_{t_{l_2}}^{(m,n+1),j})| \\
\le\ & 2(t-s)c_1\frac{\tilde{T}}{m},
\end{aligned}$$

which completes the proof of (4.50). By the induction hypothesis and the local Lipschitz continuity of P, we now get

$$
\begin{aligned}
& |(y_t^{(n+1),j} - y_s^{(n+1),j}) - (\tilde{y}_t^{(m,n+1),j} - \tilde{y}_s^{(m,n+1),j})| \\
= \quad & \left| \int_s^t P_j(y_u^{(n)})\, du - \int_s^t P_j(\hat{y}_u^{(m,n),j})\, du \right| \\
= \quad & \left| \int_s^t (P_j(y_u^{(n)}) - P_j(\hat{y}_u^{(m,n),j}))\, du \right| \\
\leq \quad & \int_s^t |P_j(y_u^{(n)}) - P_j(\hat{y}_u^{(m,n),j})|\, du \\
\leq \quad & \int_s^t L\|y_u^{(n)} - \hat{y}_u^{(m,n),j}\|\, du \\
\leq \quad & (t-s)L \max_{u\in[s,t]} \|y_u^{(n)} - \hat{y}_u^{(m,n),j}\| \\
\leq \quad & (t-s)L(t-s)\frac{c\tilde{T}}{m}\sum_{k=0}^{n-1} L^k(t-s)^k \\
= \quad & (t-s)\frac{c\tilde{T}}{m}\sum_{k=1}^{n} L^k(t-s)^k.
\end{aligned}
$$

Therefore, it follows with (4.49) and (4.50) that

$$
\begin{aligned}
& |(\hat{y}_t^{(m,n+1),j} - \hat{y}_s^{(m,n+1),j}) - (y_t^{(n+1),j} - y_s^{(n+1),j})| \\
\leq \quad & |(y_t^{(n+1),j} - y_s^{(n+1),j}) - (\tilde{y}_t^{(m,n+1),j} - \tilde{y}_s^{(m,n+1),j})| \\
& +|(\tilde{y}_t^{(m,n+1),j} - \tilde{y}_s^{(m,n+1),j}) - (\bar{y}_t^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j})| \\
& +|(\bar{y}_t^{(m,n+1),j} - \bar{y}_s^{(m,n+1),j}) - (\hat{y}_t^{(m,n+1),j} - \hat{y}_s^{(m,n+1),j})| \\
\leq \quad & (t-s)c_0\frac{\tilde{T}^2}{m} + 2(t-s)c_1\frac{\tilde{T}}{m} + (t-s)\frac{c\tilde{T}}{m}\sum_{k=1}^{n} L^k(t-s)^k \\
= \quad & (t-s)\frac{c\tilde{T}}{m}\sum_{k=0}^{n} L^k(t-s)^k
\end{aligned}
$$

since $c = c_0\tilde{T} + 2c_1$, which completes the induction and hence shows (4.44).

Thus, we have

$$\|(\hat{y}_t^{(m,n)} - \hat{y}_s^{(m,n)}) - (y_t^{(n)} - y_s^{(n)})\| \leq (t-s)\frac{c\tilde{T}}{m}\frac{1}{1-LT} \tag{4.55}$$

for all $(s,t) \in \Delta_{\tilde{T}}$ and all integers $n \geq 0$ because of (4.44) and the fact that

$$\sum_{k=0}^{n-1} L^k(t-s)^k < \sum_{k=0}^{\infty} L^k(t-s)^k = \frac{1}{1-LT}$$

as $\tilde{T} < \frac{1}{L}$.

The rest of the proof works in the same way as the proof of Theorem 4.18. Define a control $\omega : \Delta_{\tilde{T}} \to \mathbb{R}$ by

$$\omega(s,t) = \beta(c\tilde{T} + M)(t-s).$$

It follows from Theorem 4.5 that there exists a unique solution y of the polynomial differential equation (4.41) on $[0, \tilde{T}]$ and that the Picard approximations $(y_t^{(n)})_{n\in\mathbb{N}}$ converge to y_t for every $t \in [0, \tilde{T}]$. The solution y is continuous and has bounded variation (see also Theorem 4.5). By definition, the paths $y^{(n)}$ are also continuous and have bounded variation for every $n \in \mathbb{N}$.

Fix integers m and n such that $m > \frac{1}{1-L\tilde{T}} \max\left\{2\frac{c\tilde{T}}{c\tilde{T}+M}, c\tilde{T}^2\right\}$ and $n > 2\frac{L_1}{c\tilde{T}+M}e^{2L_1/M_1}$. For all $(s,t) \in \Delta_{\tilde{T}}$, let $Y_{s,t} = (1, Y_{s,t}^1, Y_{s,t}^2, \ldots)$ denote the signature of y, restricted to $[s,t]$, and let $\hat{Y}_{s,t}^{[(m,n)]} = (1, \hat{Y}_{s,t}^{[(m,n)],1}, \hat{Y}_{s,t}^{[(m,n)],2}, \ldots)$ denote the signature of $\hat{y}^{(m,n)}$, restricted to $[s,t]$. Then,

$$Y_{s,t}^1 = y_t - y_s$$

and

$$\hat{Y}_{s,t}^{[(m,n)],1} = \hat{y}_t^{(m,n)} - \hat{y}_s^{(m,n)}$$

hold for all $(s,t) \in \Delta_{\tilde{T}}$. Moreover, $(1, Y_{s,t}^1)_{(s,t)\in\Delta_{\tilde{T}}}$ and $(1, \hat{Y}_{s,t}^{[(m,n)],1})_{(s,t)\in\Delta_{\tilde{T}}}$ are both multiplicative functionals in $T^{(1)}(\mathbb{R}^d)$ with finite 1-variation.

Furthermore, it follows from (4.7) and Theorem 4.5 that $y_u \in \bar{B}_1(z)$ for every time $u \in [0, \tilde{T}]$ and we have shown above that $\hat{y}_u^{(m,n)} \in \bar{B}_2(z)$ for every $u \in [0, \tilde{T}]$. Therefore, we have

$$\|y_t - y_s\| = \left\| \int_s^t P(y_u)\, du \right\| \leq (t-s) \max_{u \in [s,t]} \|P(y_u)\| \leq M(t-s) \leq \frac{1}{\beta}\omega(s,t)$$

and

$$\|\tilde{y}_t^{(m,n)} - \tilde{y}_s^{(m,n)}\| = \left\| \int_s^t P(\hat{y}_u^{(m,n-1)})\, du \right\| \leq (t-s) \max_{u \in [s,t]} \|P(\hat{y}_u^{(m,n-1)})\| \leq M(t-s).$$

Thus,

$$\begin{aligned}
\|\hat{y}_t^{(m,n)} - \hat{y}_s^{(m,n)}\| &\leq \|(\hat{y}_t^{(m,n)} - \hat{y}_s^{(m,n)}) - (\bar{y}_t^{(m,n)} - \bar{y}_s^{(m,n)})\| \\
&\quad + \|(\bar{y}_t^{(m,n)} - \bar{y}_s^{(m,n)}) - (\tilde{y}_t^{(m,n)} - \tilde{y}_s^{(m,n)})\| \\
&\quad + \|\tilde{y}_t^{(m,n)} - \tilde{y}_s^{(m,n)}\| \\
&\leq (t-s)\left(c_0 \frac{\tilde{T}^2}{m} + 2c_1 \frac{\tilde{T}}{m} + M \right) \\
&= (t-s)\left(c\frac{\tilde{T}}{m} + M \right) \\
&\leq \frac{1}{\beta}\omega(s,t)
\end{aligned}$$

holds for all $(s,t) \in \Delta_{\tilde{T}}$ because of (4.49) and (4.50). Hence, both multiplicative functionals $(1, Y_{s,t}^1)_{(s,t)\in\Delta_{\tilde{T}}}$ and $(1, \hat{Y}_{s,t}^{[(m,n)],1})_{(s,t)\in\Delta_{\tilde{T}}}$ have finite 1-variation controlled by ω. Then, the full signatures $(Y_{s,t})_{(s,t)\in\Delta_{\tilde{T}}}$ and $(\hat{Y}_{s,t}^{[(m,n)]})_{(s,t)\in\Delta_{\tilde{T}}}$ are their unique extensions to multiplicative functionals in $T((\mathbb{R}^d))$ and they also have finite 1-variation controlled by ω (see [37], Theorem 3.7).

It follows from Lemma 3.5 and Chen's identity (see [37], Theorem 2.9) that the signature $\hat{Y}^{[(m,n)]}_{0,\tilde{T}}$ is given by

$$\hat{Y}^{[(m,n)]}_{0,\tilde{T}} = \exp(\hat{y}^{(m,n)}_{t_1} - \hat{y}^{(m,n)}_{t_0}) \otimes \cdots \otimes \exp(\hat{y}^{(m,n)}_{t_m} - \hat{y}^{(m,n)}_{t_{m-1}})$$

since $\hat{y}^{(m,n)}$ is piecewise linear.

Set

$$\varepsilon = \frac{1}{c\tilde{T} + M}\left(\frac{c\tilde{T}}{m}\frac{1}{1 - L\tilde{T}} + \frac{M_1 (L_1/M_1)^n}{n!} e^{L_1/M_1}\right).$$

Then, we have

$$2\left(1 + \sum_{r=3}^{\infty}\left(\frac{2}{r-2}\right)^2\right) = 2\left(1 + 4\sum_{r=1}^{\infty}\frac{1}{r^2}\right) = 2\left(1 + 4 \times \frac{\pi^2}{6}\right) = \beta$$

as $\sum_{r=1}^{\infty}\frac{1}{r^2} = \frac{\pi^2}{6}$ (see for example [3]). In addition,

$$0 < \varepsilon < 1$$

holds since

$$\begin{aligned}
\frac{M_1}{c\tilde{T} + M}\frac{(L_1/M_1)^n}{n!}e^{L_1/M_1} &= \frac{M_1}{c\tilde{T} + M}\frac{L_1/M_1}{n}\frac{(L_1/M_1)^{n-1}}{(n-1)!}e^{L_1/M_1} \\
&< \frac{M_1}{c\tilde{T} + M}\frac{L_1/M_1}{n}e^{L_1/M_1}e^{L_1/M_1} \\
&< \frac{1}{2}
\end{aligned}$$

because of the choice of m and n.

By (4.55) and Lemma 4.19, we have

$$\begin{aligned}
\|\hat{Y}^{[(m,n)],1}_{s,t} - Y^1_{s,t}\| &\leq \|(\hat{y}^{(m,n)}_t - \hat{y}^{(m,n)}_s) - (y^{(n)}_t - y^{(n)}_s)\| + \|(y^{(n)}_t - y^{(n)}_s) - Y^1_{s,t}\| \\
&\leq (t-s)\frac{c\tilde{T}}{m}\frac{1}{1 - LT} + M_1\frac{(L_1/M_1)^n}{n!}e^{L_1/M_1}(t-s) \\
&= (t-s)\varepsilon(c\tilde{T} + M) = \varepsilon\frac{\omega(s,t)}{\beta}
\end{aligned}$$

for all $(s,t) \in \Delta_{\tilde{T}}$ since here $\gamma_t = (t,\dots,t)^\top$ and hence $\|\gamma\|_{1,[s,t]} = t-s$. Thus, it follows from [37], Theorem 3.10, that

$$\|\hat{Y}^{[(m,n)],i}_{s,t} - Y^i_{s,t}\| \le \varepsilon \frac{\omega(s,t)^i}{\beta i!}$$

holds for every $i \in \mathbb{N}$ and every $(s,t) \in \Delta_{\tilde{T}}$. Therefore, we get

$$\begin{aligned} \|\hat{Y}^{[(m,n)]}_{0,\tilde{T}} - Y_{0,\tilde{T}}\| &= \sup_{i\ge 1} \varepsilon \frac{\omega(0,\tilde{T})^i}{\beta i!} \le \frac{\varepsilon}{\beta}\left(e^{\omega(0,\tilde{T})} - 1\right) = \frac{\varepsilon}{\beta}\left(e^{\beta \tilde{T}(c\tilde{T}+M)} - 1\right) \\ &\le \frac{\varepsilon}{\beta}\left(e^{\beta(c\tilde{T}^2+1)} - 1\right) \end{aligned}$$

for the full signatures, where we have used the fact that $Y^{[(m,n)],0}_{0,\tilde{T}} = Y^0_{0,\tilde{T}} = 1$. This shows (4.43).

Finally, (4.42) follows immediately from the inequality (4.43), which completes the proof. □

Chapter 5

An ODE method for the numerical solution of rough differential equations

In this chapter, we present and analyse an ODE method for the numerical solution of the rough differential equation

$$dY_t = F(Y_t)\,dX_t, \quad Y_0 = \xi, \tag{5.1}$$

where X is a geometric p-rough path, $\xi \in \mathbb{R}^{d_1}$ and $F : \mathbb{R}^{d_2} \to \mathrm{Lip}^{\gamma}_{\mathrm{loc}}(\mathbb{R}^{d_1}, \mathbb{R}^{d_1})$ is a linear map into locally γ-Lipschitz vector fields for some $\gamma > p$. Note that the expression $F(Y_t)\,dX_t$ has to be interpreted as $F(dX_t)(Y_t)$.

We can rewrite (5.1) in the form

$$dY_t = \sum_{i=1}^{d_2} F_i(Y_t)\,dX^i_t, \quad Y_0 = \xi, \tag{5.2}$$

where the vector fields $F_i : \mathbb{R}^{d_1} \to \mathbb{R}^{d_1}$ are the maps $F_i = F(e_i) \in \mathrm{Lip}^{\gamma}_{\mathrm{loc}}(\mathbb{R}^{d_1}, \mathbb{R}^{d_1})$ for $i = 1, \ldots, d_2$ and $\{e_1, \ldots, e_{d_2}\}$ denotes the standard basis of $\mathbb{R}^{d_2}$. Here, $\mathrm{Lip}^{\gamma}_{\mathrm{loc}}(\mathbb{R}^{d_1}, \mathbb{R}^{d_1})$ is the space of vector fields $f : \mathbb{R}^{d_1} \to \mathbb{R}^{d_1}$ that are $(\lceil\gamma\rceil - 1)$-times continuously differentiable and whose $(\lceil\gamma\rceil - 1)$-th derivatives are locally $(\gamma - (\lceil\gamma\rceil - 1))$-Hölder continuous.

Rough differential equations describe the evolution of dynamical systems driven by rough signals. For example, stochastic differential equations driven by Brownian motion or fractional Brownian motion with Hurst coefficient $H > \frac{1}{4}$ can be interpreted as rough differential equations. It follows from Lyons' Universal Limit Theorem (see Lyons [36], Theorem 4.1.1), that there exists at least locally a unique solution of the rough differential equation (5.1) over a short time interval. Note that, however, polynomial rough differential equations may explode in finite time.

The ODE method that we use to compute the numerical solution of rough differential equations has also been studied by Gyurkó [26]. It is an extension of a method described by Gaines and Lyons [24] and Gyurkó and Lyons [27] for stochastic differential equations driven by Brownian motion. The basic idea is to approximate the solution of the rough differential equation over a small time step by the solution of an ordinary differential equation (see Castell [10] or Castell and Gaines [11] for the case of stochastic differential equations driven by Brownian motion). We derive error estimates for this ODE method. Similar results have been obtained by Gyurkó [26] for the case of γ-Lipschitz vector fields and polynomial vector fields. We use a different set-up and we work with vector fields that are only locally γ-Lipschitz. It turns out that the ODE method achieves in general the same rate of convergence as the corresponding higher order Euler schemes for rough differential equations in Davie [17] for the case $p < 3$ and Friz and Victoir [23] or [22], chapter 10, for the general case. At the same time, it enhances stability (see Gyurkó and Lyons [27]).

For the implementation, we assume that the vector fields F_i are polynomials. In general, a solution of a rough differential equation can be well approximated by the successive solution of polynomial rough differential equations (see Caruana [9] and Davie [17]). Moreover, polynomial rough differential equations are themselves a rich class of differential equations which have numerous applications. For example, they include the stochastic Landau-Lifshitz-Gilbert equation (see Brzeźniak and Goldys [4]) which we solve numerically as an example in section 5.5.

The numerical method, that we describe here, has been implemented as part of the *CoRoPa* software package (see `http://coropa.sourceforge.net`) which has been jointly developed by various members of the Stochastic Analysis Group of the University of Oxford under the guidance of Terry Lyons, including Stephen Buckley, Djalil Chafai, Greg Gyurkó, Christian Litterer, Chang Liang Xu, Rahul Raghuram and the author of this thesis.

5.1 Rough paths

The theory of rough paths extends classical calculus in such a way that it allows us to describe the dynamical evolution of systems driven by rough signals, including rough paths based on Brownian motion and fractional Brownian motion. A detailed study of rough paths can be found in Lyons [36], Lyons et al. [37], Lyons and Qian [39] or Friz and Victoir [22]. In this section, we give a brief introduction following [37].

5.1.1 General definition

At first, we define controls as in [37], Definition 1.9. Recall that $T^{(n)}(\mathbb{R}^d)$ is the space of truncated formal series of tensors over $\mathbb{R}^d$ (see Definition 3.2).

Definition 5.1 *Let $T > 0$ be a real number and let Δ_T denote*

$$\Delta_T = \{(s,t) \in [0,T] \times [0,T] : 0 \leq s \leq t \leq T\}.$$

A control on $[0,T]$ is a non-negative function $\omega : \Delta_T \to \mathbb{R}$ which is super-additive *in the sense that*

$$\omega(s,u)+\omega(u,t)\le\omega(s,t)\quad \textit{for all } s,u,t\in[0,T] \textit{ satisfying } s\le u\le t$$

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

Before we can define rough paths, we need to introduce the notion of multiplicative functionals (see also [37], Definition 3.1).

Definition 5.2 (Multiplicative functionals) *Let* $n\ge 1$ *be an integer and let* $X:\Delta_T\to T^{(n)}(\mathbb{R}^d)$ *be a continuous map. For each* $(s,t)\in\Delta_T$*, let*

$$X_{s,t}=(X^0_{s,t},X^1_{s,t},\dots,X^n_{s,t})\in T^{(n)}(\mathbb{R}^d)$$

denote the image of (s,t) *under* X*. The function* X *is called a multiplicative functional of degree* n *over* $\mathbb{R}^d$ *if* $X^0_{s,t}=1$ *holds for all* $(s,t)\in\Delta_T$ *and*

$$X_{s,u}\otimes X_{u,t}=X_{s,t}\tag{5.3}$$

holds for all $s,u,t\in\Delta_T$ *satisfying* $s\le u\le t$.

For any positive real number x, we write $x!=\Gamma(x+1)$ for the Gamma function and $\lfloor x\rfloor$ for the biggest integer smaller than (or equal to) x, i.e. $\lfloor x\rfloor$ is the integer given by $\lfloor x\rfloor\le x<\lfloor x\rfloor+1$. Using this notation, we make the following definition (see also [37], Definition 3.6).

Definition 5.3 *Let* $p\ge 1$ *be a real number and* $n\ge 1$ *be an integer. Let* ω *be a control on* $[0,T]$ *and let* $X:\Delta_T\to T^{(n)}(\mathbb{R}^d)$ *be a multiplicative functional.*
We say that X *has finite* p*-variation on* Δ_T *controlled by* ω *if*

$$\|X^i_{s,t}\|\le\frac{\omega(s,t)^{\frac{i}{p}}}{\beta_p\left(\frac{i}{p}\right)!}$$

holds for all $i \in \{1, \ldots, n\}$ *and all* $(s,t) \in \Delta_T$*, where*

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

In general, we say that X *has finite* p*-variation if there exists a control* ω *such that the conditions above are satisfied. Then, the* p*-variation of* X *over the interval* $[s,t]$ *is defined as*

$$\|X\|_{p,[s,t]} = \max_{k \in \{1,\ldots,n\}} \sup_{\mathcal{D} \subset [s,t]} \left(\sum_{i=0}^{m-1} \|X^k_{t_i,t_{i+1}}\|^{\frac{p}{k}} \right)^{\frac{1}{p}},$$

where $\mathcal{D}$ *is the partition* $s = t_0 < t_1 < \cdots < t_m = t$ *of* $[s,t]$*.*

Note that the function $(s,t) \mapsto \|X\|^p_{p,[s,t]}$ is a control if X has finite p-variation (see for example [37], section 1.2.2).

The multiplicative functional X is $\frac{1}{p}$-Hölder continuous on $[s,t]$ if there exists a constant c, such that

$$\|X^k_{\rho,\tau}\|^{\frac{1}{k}} \leq c|\tau - \rho|^{\frac{1}{p}}$$

holds for all ρ and τ with $s \leq \rho < \tau \leq t$ and all $k \in \{1, \ldots, n\}$. Then, the $\frac{1}{p}$-Hölder norm of X on $[s,t]$ is defined as

$$\|X\|_{1/p-\text{Höl},[s,t]} = \max_{k \in \{1,\ldots,n\}} \sup_{s \leq \rho < \tau \leq t} \frac{\|X^k_{\rho,\tau}\|^{\frac{1}{k}}}{|\tau - \rho|^{\frac{1}{p}}}.$$

Now, we can finally define rough paths (as in [37], Definition 3.11).

Definition 5.4 (Rough paths) *Let* $p \geq 1$ *be a real number. A* p*-rough path over* $\mathbb{R}^d$ *is a multiplicative functional of degree* $\lfloor p \rfloor$ *over* $\mathbb{R}^d$ *with finite* p*-variation.*

5.1.2 Paths with bounded variation and signatures

Continuous paths with bounded variation correspond to 1-rough paths in a natural way: Let $x : [0,T] \to \mathbb{R}^d$ be a continuous path with bounded variation. Define $X : \Delta_T \to T^{(1)}(\mathbb{R}^d)$ by $X_{s,t} = (1, x_t - x_s)$. Then, X is a 1-rough path since the multiplicativity property (5.3) on $T^{(1)}(\mathbb{R}^d)$ simply means additivity of increments. It is controlled by the length of x, i.e. $\omega(s,t) = \beta_1 \|x\|_{1,[s,t]}$, where $\|x\|_{1,[s,t]}$ is the 1-variation

$$\|x\|_{1,[s,t]} = \sup_{\mathcal{D} \subset [s,t]} \sum_{i=1}^{n} \|x_{t_i} - x_{t_{i-1}}\|$$

of x. Note that the supremum is being taken over all partitions $\mathcal{D} = \{t_0, t_1, \dots, t_n\}$ of $[s,t]$ with $s = t_0 < t_1 < \dots < t_n = t$. It is possible to construct a multiplicative functional of arbitrary degree above the path x (see for example [37]). Let

$$X_{s,t} = (1, X^1_{s,t}, X^2_{s,t}, \dots)$$

be the signature of x over the interval $[s,t]$, where

$$X^k_{s,t} = \sum_{(i_1,\dots,i_k) \in \{1,\dots,d\}^k} \left(\int_{s<u_1<\dots<u_k<t} dX^{i_1}_{u_1} \dots dX^{i_k}_{u_k} \right) e_{i_1} \otimes \dots \otimes e_{i_k} \in (\mathbb{R}^d)^{\otimes k}$$

for each integer $k \geq 1$ and $\{e_1, \dots, e_d\}$ denotes the standard basis of $\mathbb{R}^d$.
Let $n \geq 1$ be an integer and define $X^{[n]} : \Delta_T \to T^{(n)}(\mathbb{R}^d)$ by $X^{[n]}_{s,t} = (1, X^1_{s,t}, \dots, X^n_{s,t})$, where $(1, X^1_{s,t}, \dots, X^n_{s,t})$ is the truncated signature of x, restricted to $[s,t]$. Then, $X^{[n]}$ is continuous and Chen's identity (see for example [37], Theorem 2.9) shows that $X^{[n]}$ is multiplicative in the sense of (5.3). Moreover, it follows form [37], Proposition 2.2, that the multiplicative functional $X^{[n]}$ of degree n has finite 1-variation. The Extension Theorem (see [37], Theorem 3.7) ensures that this extension of the 1-rough path $(s,t) \mapsto (1, x_t - x_s)$ to a multiplicative functional of degree n is unique.

5.1.3 Geometric rough paths

Geometric rough paths are an important class of rough paths (see for example [37], section 3.2.2). They are the drivers of the rough differential equations in the following sections.

Let $p \geq 1$ be a real number and let $X, Y : \Delta_T \to T^{(\lfloor p \rfloor)}(\mathbb{R}^d)$ be continuous (multiplicative) functionals with finite p-variation. We define their p-variation distance by

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

(as in [37], section 3.2.1), where the supremum is being taken over all partitions $\mathcal{D} = \{t_0, t_1, \dots, t_N\}$ of $[0,T]$ with $0 = t_0 < t_1 < \cdots < t_N = T$.

Now, we define geometric rough paths as in [37], Definition 3.13.

Definition 5.5 (Geometric rough paths) *Let $p \geq 1$ be a real number. A geometric p-rough path is a p-rough path that can be expressed as the limit of* 1*-rough paths in the p-variation distance, where the* 1*-rough paths are identified with their unique extensions to multiplicative functionals of degree $\lfloor p \rfloor$ as discussed in the previous section.*

The logarithm of the values of geometric p-rough paths lives in the space of truncated Lie series $\mathcal{L}^{(\lfloor p \rfloor)}(\mathbb{R}^d) \subset T^{(\lfloor p \rfloor)}(\mathbb{R}^d)$ (as defined in Definition 3.7): Let X be a geometric p-rough path. It follows from [37] that $\log X_{s,t}$ is an element of the space $\mathcal{L}^{(\lfloor p \rfloor)}(\mathbb{R}^d)$ for every $(s,t) \in \Delta_T$. Indeed, $X_{s,t}$ is a group-like element (see [37], Definition 2.18) because X is a geometric rough path (see [37], section 3.2.2) and the logarithm of a group-like element is an element of $\mathcal{L}^{(\lfloor p \rfloor)}(\mathbb{R}^d)$ (see [37], Theorem 2.23).

5.1.4 Example: the Brownian rough path

Brownian motion, together with its iterated Stratonovich integrals, forms a geometric rough path (see for example [37], section 3.3.2): Let B be a d-dimensional standard Brownian motion. Define

$$S_{s,t} = \left(1, B_t - B_s, \sum_{i_1,i_2=1}^{d} \left(\int_{s<u_1<u_2<t} \circ dB^{i_1}_{u_1} \circ dB^{i_2}_{u_2} \right) e_{i_1} \otimes e_{i_2} \right) \in T^{(2)}(\mathbb{R}^d),$$

where the integrals are Stratonovich integrals. Then, $S : (s,t) \mapsto S_{s,t}$ is a geometric p-rough path for any $p \in (2,3)$.

Note that, if we used Itô integrals instead of Stratonovich integrals, S would still be a p-rough path for any $p \in (2,3)$, but not a geometric rough path anymore (see [37], section 3.3.1).

5.1.5 Example: the fractional Brownian rough path

It is also possible to construct a geometric rough path above fractional Brownian motion with Hurst coefficient $H > \frac{1}{4}$ using dyadic polygonal approximations (see Coutin and Qian [14]): Let B^H be a d-dimensional fractional Brownian motion with Hurst parameter $H > \frac{1}{4}$. Then, the dyadic polygonal approximations provide a lift to a multiplicative functional W of degree 3, which takes values in $T^{(3)}(\mathbb{R}^d)$ and has finite p-variation for any $p \in (\frac{1}{H}, 4)$, such that W is a geometric p-rough path for any $p \in (\frac{1}{H}, 4)$ and $W^1_{s,t} = B^H_t - B^H_s$.

5.2 Rough differential equations

Core to the theory of rough paths is an integration of one-forms against geometric rough paths and the continuity of the integral for smooth one-forms (see [37], section 4). With this notion of an integral, it is possible to interpret the rough differential equation (5.1) as an integral equation, but it needs reformulation: Consider a rough path (X, Y) extending X to a rough path that takes its values in $T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2} \oplus \mathbb{R}^{d_1})$ and think of the rough differential equation (5.1) as the system

$$\begin{aligned} dX_t &= dX_t \\ dY_t &= F(Y_t)\, dX_t. \end{aligned}$$

Define the map $\alpha_\xi : \mathbb{R}^{d_2} \oplus \mathbb{R}^{d_1} \to L(\mathbb{R}^{d_2} \oplus \mathbb{R}^{d_1}, \mathbb{R}^{d_2} \oplus \mathbb{R}^{d_1})$ by

$$\alpha_\xi((x,y))(w,v) = (w, F(w)(y+\xi)).$$

Recall that X is a geometric p-rough path, $\xi \in \mathbb{R}^{d_1}$ and $F : \mathbb{R}^{d_2} \to \mathrm{Lip}^{\gamma}_{\mathrm{loc}}(\mathbb{R}^{d_1}, \mathbb{R}^{d_1})$ is a linear map into locally γ-Lipschitz vector fields for some $\gamma > p$. Note that α_ξ is locally Lip(γ) in the sense of [37], Definition 1.21. Then, a geometric p-rough path Z, which takes its values in $T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2} \oplus \mathbb{R}^{d_1})$, is called a solution of the rough differential equation (5.1) if

$$Z = \int \alpha_\xi(Z)\, dZ$$

and

$$\pi_{\mathbb{R}^{d_2}}(Z) = X$$

hold, where the integral is defined as in [37], Definition 4.9, and $\pi_{\mathbb{R}^{d_2}}$ denotes the projection onto $\mathbb{R}^{d_2}$. This is Lyons' definition of the solution of rough differential equations (see [36] or [37], section 5).
Given a solution Z, we define

$$Y_t = \xi + \pi_{\mathbb{R}^{d_1}}(Z^1_{0,t}).$$

Note that this definition of Y coincides with the classical definition of solutions of differential equations for the case that X is a continuous path with bounded variation, i.e. a 1-rough path. Thus, Y is also an RDE solution in the sense of Friz and Victoir [22], Definition 10.17, because of the continuity of the extension of the Itô map (see [36], Theorem 4.1.1, or [37], Theorem 5.3). The following theorem is an immediate consequence of [22], Theorem 10.21, and it shows that the rough differential equation (5.1) admits at least a local solution which is unique. This result goes back to Lyons' Universal Limit Theorem ([36], Theorem 4.1.1).

Theorem 5.6 *Let $T > 0$, $p \geq 1$ and $\gamma > p$ be real numbers and let $X : \Delta_T \to T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2})$ be a geometric p-rough path. Consider the rough differential equation (5.1), which reads*

$$dY_t = \sum_{i=1}^{d_2} F_i(Y_t)\, dX_t^i, \quad Y_0 = \xi,$$

in the form (5.2), where $\xi \in \mathbb{R}^{d_1}$ and $F_i \in \mathrm{Lip}^{\gamma}_{\mathrm{loc}}(\mathbb{R}^{d_1}, \mathbb{R}^{d_1})$ for $i = 1, \ldots, d_2$. Then:

(i) There exists a time $\varepsilon > 0$ such that there exists a unique solution $Z : \Delta_\varepsilon \to T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2} \oplus \mathbb{R}^{d_1})$ over $[0, \varepsilon] \subset [0, T]$, which is a geometric p-rough path.

(ii) There exists either a unique solution $Z : \Delta_T \to T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2} \oplus \mathbb{R}^{d_1})$ over the entire interval $[0, T]$, which is a geometric p-rough path, or the solution explodes within $[0, T]$. In the latter case, there exists $\tau^* \in (0, T]$ such that for every $\tau \in [0, \tau^*)$ there exists a unique solution $Z : \Delta_\tau \to T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2} \oplus \mathbb{R}^{d_1})$ over $[0, \tau]$, which is a geometric p-rough path, and

$$\lim_{\tau \nearrow \tau^*} \|Z_{0,\tau}\| = +\infty,$$

i.e. τ^* is the time of explosion.

Proof: This Theorem follows immediately from [22], Theorem 10.21 and sections 10.3.4 and 10.4.1. □

For instance, polynomial rough differential equations may explode in finite time. Note that the time of explosion depends on the vector fields F_i, the initial value ξ and the signal X.

5.3 The numerical method

In this section, we present and analyse the ODE method that we use to compute numerical approximations for the solution of the rough differential equation (5.1) (see also Gyurkó [26]). It is an extension of the method that Gaines and Lyons [24] and Gyurkó and Lyons [27] use for the numerical solution of stochastic differential equations driven by Brownian motion.

Recall that we can rewrite the rough differential equation (5.1) in the form (5.2), i.e.

$$dY_t = \sum_{i=1}^{d_2} F_i(Y_t)\, dX_t^i, \quad Y_0 = \xi,$$

where the vector fields

$$F_i = \begin{pmatrix} F_i^1 \\ \vdots \\ F_i^{d_1} \end{pmatrix} : \mathbb{R}^{d_1} \to \mathbb{R}^{d_1}$$

are locally γ-Lipschitz. For $i = 1, \ldots, d_2$, define differential operators

$$V^i = \sum_{j=1}^{d_1} F_i^j \frac{\partial}{\partial x_j}.$$

Furthermore, recall that $\mathrm{Lip}^\gamma_{\mathrm{loc}}(\mathbb{R}^{d_1}, \mathbb{R}^{d_1})$ is the space of vector fields that are $(\lceil\gamma\rceil - 1)$-times continuously differentiable and whose $(\lceil\gamma\rceil - 1)$-th derivatives are locally $(\gamma - (\lceil\gamma\rceil - 1))$-Hölder continuous. Here, $\lceil\gamma\rceil$ is the integer which is uniquely defined by $\lceil\gamma\rceil \geq \gamma > \lceil\gamma\rceil - 1$. Note that 1-Lipschitz corresponds to the Lipschitz continuity that has been used in chapter 4. Let $f \in \mathrm{Lip}^\gamma_{\mathrm{loc}}(\mathbb{R}^{d_1}, \mathbb{R}^{d_1})$ and let $M \subset \mathbb{R}^{d_1}$ be a compact set. Then, there exists a constant c_M such that the supremum norm of f, all its i-th derivatives for $i = 1, \ldots, \lceil\gamma\rceil - 1$ and the $(\gamma - (\lceil\gamma\rceil - 1))$-Hölder norm of its $(\lceil\gamma\rceil - 1)$-th derivative are bounded by c_M on M. The smallest such constant c_M is called the γ-Lipschitz norm of f on M and denoted by $|f|_{\mathrm{Lip}^\gamma_M}$ (see [22], Definition 10.2).

The geometric p-rough path X can be uniquely written in the form

$$X_{s,t} = 1 + \sum_{i_1=1}^{d_2} X_{s,t}^{(i_1)} e_{i_1} + \sum_{j=2}^{\lfloor p \rfloor} \sum_{(i_1,\ldots,i_j)\in\{1,\ldots,d_2\}^j} X_{s,t}^{(i_1,\ldots,i_j)} e_{i_1} \otimes \cdots \otimes e_{i_j}, \tag{5.4}$$

where the coefficients $X_{s,t}^{(i_1,\ldots,i_j)}$ are real numbers.

Recall that $\log X_{s,t}$ is an element of the space of truncated Lie series $\mathcal{L}^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2})$ for every $(s,t) \in \Delta_T$ because X is a geometric p-rough path (see section 5.1.3). Let $\mathcal{B}$ be a basis of $\mathcal{L}^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2})$. Then, $\log X_{s,t}$ allows a unique representation as

$$\log X_{s,t} = \sum_{\ell \in \mathcal{B}} L_{s,t}^{\ell} \ell, \tag{5.5}$$

where the coefficients $L_{s,t}^{\ell}$ are real numbers.

Note that for numerical purposes it is advantageous to work with $\log X_{s,t} \in \mathcal{L}^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2})$ instead of $X_{s,t} \in T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2})$ because this reduces the dimension of the space.

We map elements of $T((\mathbb{R}^{d_2}))$ to differential operators in the following way (see also Gyurkó [26], Definition 3.3.6).

Definition 5.7 *Let $\mathcal{F} = \{f_1, \ldots, f_{d_2}\}$ be a set of smooth vector fields on $\mathbb{R}^{d_1}$ and let $\mathcal{V} = \{v^1, \ldots, v^{d_2}\}$ be the set of differential operators that are associated with $f_1, \ldots, f_{d_2}$. Then, the algebra homomorphism $\Phi_{\mathcal{V}}$ from $T((\mathbb{R}^{d_2}))$ into the space of differential operators is generated by*

$$\Phi_{\mathcal{V}}(\mathbf{1}) = \mathrm{Id}_{\mathbb{R}^{d_1}} \qquad \textit{and} \qquad \Phi_{\mathcal{V}}(e_i) = v^i$$

for $i = 1, \ldots, d_2$.

For elements of $T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2})$, we define the map $\Phi_{\mathcal{V}}$ in the respective way and in this case, it is sufficient to assume that the vector fields $f_1, \ldots, f_{d_2}$ are $\lfloor p \rfloor$-times continuously differentiable. Note that $\Phi_{\mathcal{V}}$ maps elements of $\mathcal{L}^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2})$ to first-order differential operators which can be associated with vector fields.

The following theorem describes the numerical method for the solution of rough differential equations. In every time step, we solve an ordinary differential equation to approximate the solution of the rough differential equation (5.1) as in Gaines and Lyons [24], Gyurkó and Lyons [27], Castell [10] or Castell and Gaines [11] for the case of stochastic differential equations driven by Brownian motion. Note that similar results have been obtained by Gyurkó [26] for the case of γ-Lipschitz vector fields and polynomial vector fields. Here, we use a different set-up and we work with vector fields that are only locally γ-Lipschitz.

Theorem 5.8 *Let $T > 0$, $p \geq 1$ and $\gamma > p$ be real numbers. Consider the rough differential equation (5.1), i.e.*

$$dY_t = F(Y_t)\,dX_t, \quad Y_0 = \xi,$$

where $X : \Delta_T \to T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2})$ is a geometric p-rough path, $\xi \in \mathbb{R}^{d_1}$ and $F : \mathbb{R}^{d_2} \to \mathrm{Lip}^{\gamma}_{\mathrm{loc}}(\mathbb{R}^{d_1}, \mathbb{R}^{d_1})$ is a linear map into locally γ-Lipschitz vector fields. Theorem 5.6 shows that either there exists a unique solution on the entire interval $[0,T]$ or it explodes within $[0,T]$. In the latter case, there exists $\tau^* \in (0,T]$ such that there exists a unique solution on $[0,\tau^*)$ and τ^* is the time of explosion. Fix a time $\tilde{T}$ within the maximal interval of existence, i.e. $\tilde{T} \in [0,T]$ if the solution exists on the entire interval $[0,T]$ or otherwise $\tilde{T} < \tau^*$ before the time of explosion. Consider the solution $Z : \Delta_{\tilde{T}} \to T^{(\lfloor p \rfloor)}(\mathbb{R}^{d_2} \oplus \mathbb{R}^{d_1})$ and $Y : [0,\tilde{T}] \to \mathbb{R}^{d_1}$ given by $Y_t = \xi + \pi_{\mathbb{R}^{d_1}}(Z^1_{0,t})$. Then, there exists a radius $R > 0$ such that Y_t remains in the closed ball $\bar{B}_R(\xi)$ around ξ with radius R for all $t \in [0,\tilde{T}]$. Let $\tilde{K} = \max\{|F_1|_{\mathrm{Lip}^{\gamma}_{\bar{B}_{R+1}(\xi)}}, \ldots, |F_{d_2}|_{\mathrm{Lip}^{\gamma}_{\bar{B}_{R+1}(\xi)}}\}$ be the maximum of the γ-Lipschitz norms of $F_1, \ldots, F_{d_2}$ on $\bar{B}_{R+1}(\xi)$. Set $K = \max\{1, \tilde{K}, \tilde{K}^{\lceil \gamma \rceil}\}$ and let $\mathcal{V} = \{V^1, \ldots, V^{d_2}\}$ be the set of differential operators that are associated with $F_1, \ldots, F_{d_2}$. Furthermore, let $\mathcal{B}$ be a basis of the space of truncated Lie series $\mathcal{L}^{(\lceil \gamma \rceil - 1)}(\mathbb{R}^{d_2})$.

(i) Let $N \geq 1$ be an integer and let the partition $0 = t^{[N]}_0 < t^{[N]}_1 < \cdots < t^{[N]}_N = \tilde{T}$

of $[0,\tilde{T}]$ *be defined by*

$$t_k^{[N]} = \min\{\rho \in [0,\tilde{T}] : \|X\|_{p,[0,\rho]}^p = \frac{k}{N}\|X\|_{p,[0,\tilde{T}]}^p\}$$

for $k = 1,\ldots,N$. *Let* $\log X$ *be given in the form (5.5). Define approximations* $\bar{y}_k^{[N]}$ *for* $Y_{t_k^{[N]}}$ *in the following way: Start with* $y_0^{[N]} = \xi$ *and for every step, solve the ordinary differential equation*

$$\frac{du_{t_k^{[N]},t_{k+1}^{[N]}}}{dt} = \sum_{\ell\in\mathcal{B}} \Phi_{\mathcal{V}}(\ell)(\mathrm{Id}_{\mathbb{R}^{d_1}})(u_{t_k^{[N]},t_{k+1}^{[N]}})L^{\ell}_{t_k^{[N]},t_{k+1}^{[N]}} \tag{5.6}$$

with initial condition $u_{t_k^{[N]},t_{k+1}^{[N]}}(0) = \bar{y}_k^{[N]}$ *and set*

$$\bar{y}_{k+1}^{[N]} = u_{t_k^{[N]},t_{k+1}^{[N]}}(1).$$

Then, $(\bar{y}_N^{[N]})_{N\geq 1}$ *converges to* $Y_{\tilde{T}}$ *and there exists a constant* c*, which only depends on* γ*,* d_1*,* d_2 *and* p*, such that*

$$\|Y_{\tilde{T}} - \bar{y}_N^{[N]}\| \leq cK^{p+1}(\max\{1,\|X\|_{p,[0,\tilde{T}]}^{p+1}\})\exp(cK^p\|X\|_{p,[0,\tilde{T}]}^p)\left(\frac{K^p\|X\|_{p,[0,\tilde{T}]}^p}{N}\right)^{\frac{\lceil\gamma\rceil}{p}-1}$$

holds for all integers

$$N \geq \max\Big\{\Big(c^2K^{\lceil\gamma\rceil+2}(\max\{1,\|X\|_{p,[0,\tilde{T}]}^{\lceil\gamma\rceil+1}\})\exp(2cK^p\|X\|_{p,[0,\tilde{T}]}^p)\Big)^{\frac{p}{\lceil\gamma\rceil-p}},$$
$$\Big(cK\|X\|_{p,[0,\tilde{T}]}\Big)^p\Big\}.$$

(ii) If the geometric rough path X *is* $\frac{1}{p}$*-Hölder continuous, let* $N \geq 1$ *be an integer and let* $0 = t_0^{[N]} < t_1^{[N]} < \cdots < t_N^{[N]} = \tilde{T}$ *be the partition of* $[0,\tilde{T}]$ *into* N *equal parts, i.e.* $t_k^{[N]} = \frac{k}{N}\tilde{T}$. *Define approximations* $\bar{y}_k^{[N]}$ *for* $Y_{t_k^{[N]}}$ *as above. Then,*

$(\bar{y}_N^{[N]})_{N\geq 1}$ *also converges to* $Y_{\tilde{T}}$ *and*

$$\|Y_{\tilde{T}} - \bar{y}_N^{[N]}\| \leq cK^{p+1}(\max\{1, \|X\|^{p+1}_{1/p-\text{Höl},[0,\tilde{T}]}\tilde{T}^{\frac{p+1}{p}}\}) \left(\frac{K^p\|X\|^p_{1/p-\text{Höl},[0,\tilde{T}]}\tilde{T}}{N}\right)^{\frac{\lceil\gamma\rceil}{p}-1} \times \exp(cK^p\|X\|^p_{1/p-\text{Höl},[0,\tilde{T}]}\tilde{T})$$

holds for all integers

$$\begin{aligned} N \quad \geq \quad & \max\Big\{\Big(c^2K^{\lceil\gamma\rceil+2}(\max\{1, \|X\|^{\lceil\gamma\rceil+1}_{1/p-\text{Höl},[0,\tilde{T}]}\tilde{T}^{\frac{\lceil\gamma\rceil+1}{p}}\}) \\ & \times \exp(2cK^p\|X\|^p_{1/p-\text{Höl},[0,\tilde{T}]}\tilde{T})\Big)^{\frac{p}{\lceil\gamma\rceil-p}}, \Big(cK\|X\|_{1/p-\text{Höl},[0,\tilde{T}]}\tilde{T}^{\frac{1}{p}}\Big)^p\Big\}, \end{aligned}$$

where c is the same constant as in part (i).

Note that, if $\gamma > \lfloor p \rfloor + 1$, we use the unique extension of X to a multiplicative functional of order $\lceil\gamma\rceil - 1$ (see [37], Theorem 3.7) to obtain all the coefficients $L^\ell_{t_k^{[N]},t_{k+1}^{[N]}}$. Also, note that there exists a unique solution $u_{t_k^{[N]},t_{k+1}^{[N]}} : [0,1] \to \mathbb{R}^{d_1}$ of (5.6) for each $k \in \{0, 1, \ldots, N-1\}$, which lives in the closed ball $\bar{B}_{R+1}(\xi)$. It is sufficient to solve these ordinary differential equations (5.6) up to an error of

$$\tilde{c}\|X\|^{\lceil\gamma\rceil}_{p,[t_k^{[N]},t_{k+1}^{[N]}]} \leq \tilde{c}\left(\frac{\|X\|^p_{p,[0,\tilde{T}]}}{N}\right)^{\frac{\lceil\gamma\rceil}{p}}$$

for some constant $\tilde{c}$ in order to maintain the overall order of convergence of Theorem 5.8.

For example, the Brownian rough path (as defined in section 2.4) is a geometric p-rough path for any $p \in (2, 3)$, which is $\frac{1}{p}$-Hölder continuous. And the fractional Brownian rough path with Hurst coefficient $H > \frac{1}{4}$ (see section 2.5) is a geometric p-rough path for any $p \in (\frac{1}{H}, 4)$, which is also $\frac{1}{p}$-Hölder continuous.

The proof of Theorem 5.8 requires the following lemma. Again, note that a similar result has been obtained by Gyurkó [26] for the case of γ-Lipschitz vector fields and polynomial vector fields. We adapt it to a different set-up and we work with vector fields that are only locally γ-Lipschitz.

Lemma 5.9 *Under the assumptions of Theorem 5.8, fix* $(\rho, \tau) \in \Delta_{\tilde{T}}$. *Let* $u_{\rho,\tau} : [0,1] \to \mathbb{R}^{d_1}$ *be the solution of the differential equation*

$$\frac{du_{\rho,\tau}}{dt} = \sum_{\ell \in \mathcal{B}} \Phi_{\mathcal{V}}(\ell)(\mathrm{Id}_{\mathbb{R}^{d_1}})(u_{\rho,\tau}) L^{\ell}_{\rho,\tau} \tag{5.7}$$

with initial condition $u_{\rho,\tau}(0) = Y_\rho$. *Then, there exists a constant* c*, which only depends on* γ, d_1, d_2 *and* p*, such that*

$$\|Y_\tau - u_{\rho,\tau}(1)\| \leq c \left(K \|X\|_{p,[\rho,\tau]}\right)^{\lceil \gamma \rceil} \tag{5.8}$$

if the times ρ *and* τ *have been chosen close enough to each other so that* $\|X\|_{p,[\rho,\tau]} \leq \frac{1}{cK}$ *holds.*

Note that there exists a unique solution $u_{\rho,\tau} : [0,1] \to \mathbb{R}^{d_1}$ of (5.7) which lives in the closed ball $\bar{B}_{R+1}(\xi)$.

Proof of Lemma 5.9: For this proof, we use the notation of Friz and Victoir [22]. Let c_1 be the constant from [22], Corollary 10.15 and let c_2 be the constant from [22], Proposition 10.3. Note that c_1 and c_2 only depend on γ and p. Set

$$W = \sum_{\ell \in \mathcal{B}} L^{\ell}_{\rho,\tau} \Phi_{\mathcal{V}}(\ell).$$

Then, W is a first-order differential operator which corresponds to the vector field $W(\mathrm{Id}_{\mathbb{R}^{d_1}})$. Note that this vector field $W(\mathrm{Id}_{\mathbb{R}^{d_1}})$ is $(\lceil \gamma \rceil - 1)$-Lipschitz with $(\lceil \gamma \rceil - 1)$-Lipschitz norm $|W(\mathrm{Id}_{\mathbb{R}^{d_1}})|_{\mathrm{Lip}^{\lceil \gamma \rceil - 1}_{\bar{B}_{R+1}(\xi)}} \leq c_3 K \|X\|_{p,[\rho,\tau]}$ on $\bar{B}_{R+1}(\xi)$ for some constant c_3 which only depends on γ, d_1 and d_2.

Set $c = \max\{1, c_3, c_1 + c_2 c_3^{\lceil\gamma\rceil}\}$. Assume that ρ and τ have been chosen so that

$$\|X\|_{p,[\rho,\tau]} \leq \frac{1}{cK}. \tag{5.9}$$

Set

$$\mathcal{E}_{F,\lceil\gamma\rceil-1}(Y_\rho, X_{s,t}) = \sum_{l=1}^{\lceil\gamma\rceil-1} \sum_{(i_1,\ldots,i_l)\in\{1,\ldots,d_2\}^l} V^{i_1}\ldots V^{i_l}(\mathrm{Id}_{\mathbb{R}^{d_1}})(Y_\rho)\, X_{s,t}^{(i_1,\ldots,i_l)}. \tag{5.10}$$

Note that this corresponds to [22], Definition 10.1. The vector fields $F_1,\ldots,F_{d_2}$ are locally γ-Lipschitz. Hence, we also have $F_1,\ldots,F_{d_2} \in \mathrm{Lip}_{\mathrm{loc}}^{\lceil\gamma\rceil-1}$. Thus, it follows from [22], Corollary 10.15, that

$$\|(Y_\tau - Y_\rho) - \mathcal{E}_{F,\lceil\gamma\rceil-1}(Y_\rho, X_{\rho,\tau})\| \leq c_1(K\|X\|_{p,[\rho,\tau]})^{\lceil\gamma\rceil}.$$

Furthermore, we have

$$\begin{aligned}
\mathcal{E}_{F,\lceil\gamma\rceil-1}(Y_\rho, X_{\rho,\tau}) &= \Phi_{\mathcal{V}}(X_{\rho,\tau})(\mathrm{Id}_{\mathbb{R}^{d_1}})(Y_\rho) \\
&= \Phi_{\mathcal{V}}\left(\sum_{k=1}^{\infty}\frac{1}{k!}(\log X_{\rho,\tau})^{\otimes k}\right)(\mathrm{Id}_{\mathbb{R}^{d_1}})(Y_\rho) \\
&= \Phi_{\mathcal{V}}\left(\sum_{k=1}^{\lceil\gamma\rceil-1}\frac{1}{k!}(\log X_{\rho,\tau})^{\otimes k}\right)(\mathrm{Id}_{\mathbb{R}^{d_1}})(Y_\rho) \\
&= \sum_{k=1}^{\lceil\gamma\rceil-1}\frac{1}{k!}\underbrace{\Phi_{\mathcal{V}}(\log X_{\rho,\tau})\ldots\Phi_{\mathcal{V}}(\log X_{\rho,\tau})}_{k\text{-times}}(\mathrm{Id}_{\mathbb{R}^{d_1}})(Y_\rho) \\
&= \sum_{k=1}^{\lceil\gamma\rceil-1}\frac{1}{k!}\underbrace{W\ldots W}_{k\text{-times}}(\mathrm{Id}_{\mathbb{R}^{d_1}})(Y_\rho).
\end{aligned}$$

It follows from [22], Proposition 10.3, that

$$\left\|(u_{\rho,\tau}(1) - u_{\rho,\tau}(0)) - \sum_{k=1}^{\lceil\gamma\rceil-1}\frac{1}{k!}\underbrace{W\ldots W}_{k\text{-times}}(\mathrm{Id}_{\mathbb{R}^{d_1}})(Y_\rho)\right\| \leq c_2(c_3K\|X\|_{p,[\rho,\tau]})^{\lceil\gamma\rceil}.$$

Thus, we have

$$\|Y_\tau - u_{\rho,\tau}(1)\| \le c\left(K\|X\|_{p,[\rho,\tau]}\right)^{\lceil\gamma\rceil}.$$

$\square$

Finally, we can now prove Theorem 5.8.

Proof of Theorem 5.8: The proof works in a similar way as the proofs of [22], Theorem 10.30, or [22], Proposition 10.33. But we need to be careful that we apply all the arguments locally and stay within the closed ball $\bar{B}_{R+1}(\xi)$ at all times.
Set $c = \max\{c_1, c_2, c_1c_2\}$, where c_1 is the constant from [22], Theorem 10.26, and c_2 is the constant from Lemma 5.9. Note that c only depends on γ, d_1, d_2 and p. Fix

$$N \ge \max\Big\{ \Big(c^2K^{\lceil\gamma\rceil+2}(\max\{1, \|X\|_{p,[0,\tilde{T}]}^{\lceil\gamma\rceil+1}\})\exp(2cK^p\|X\|_{p,[0,\tilde{T}]}^p)\Big)^{\frac{p}{\lceil\gamma\rceil-p}}, \Big(cK\|X\|_{p,[0,\tilde{T}]}\Big)^p\Big\}. \tag{5.11}$$

At first, we show by induction over k that $\bar{y}_k^{[N]} \in \bar{B}_{R+1}(\xi)$ and

$$\|Y_{t_k^{[N]}} - \bar{y}_k^{[N]}\| \le cK^{\lceil\gamma\rceil+1}(\max\{1, \|X\|_{p,[0,\tilde{T}]}\})\exp(cK^p\|X\|_{p,[0,\tilde{T}]}^p)\sum_{l=1}^{k}\|X\|_{p,[t_{l-1}^{[N]},t_l^{[N]}]}^{\lceil\gamma\rceil} \tag{5.12}$$

for all $k = 0, 1, \ldots, N$.
For $k = 0$, we have $\bar{y}_0^{[N]} = Y_{t_0^{[N]}} = \xi$. For $k = 1$, the inequality (5.12) and the fact $\bar{y}_1^{[N]} \in \bar{B}_{R+1}(\xi)$ follow immediately from Lemma 5.9.
Assume that, for some $k \in \{1, \ldots, N-1\}$, we have $\bar{y}_0^{[N]}, \bar{y}_1^{[N]}, \ldots, \bar{y}_k^{[N]} \in \bar{B}_{R+1}(\xi)$ and that (5.12) holds for all integers $0, 1, \ldots, k$. Then, consider the rough differential equation

$$d\tilde{Y}_t = F(\tilde{Y}_t)\, dX_t$$

over the time interval $[t_l^{[N]}, t_{k+1}^{[N]}]$ with initial value $\tilde{Y}_{t_l^{[N]}} = \bar{y}_l^{[N]}$. For $l = 0, 1, \ldots, k$, there exists a unique solution, which we denote by $\varphi(\cdot; t_l^{[N]}, \bar{y}_l^{[N]}) : [t_l^{[N]}, t_{k+1}^{[N]}] \to \mathbb{R}^{d_1}$.

It lives inside the closed ball $\bar{B}_{R+1}(\xi)$. This can be seen in the following way: By [22], Theorem 10.26, we have

$$\|Y_t - \varphi(t; t_l^{[N]}, \bar{y}_l^{[N]})\| \leq c_1 K \|X\|_{p,[t_l^{[N]}, t_{k+1}^{[N]}]} \|Y_{t_l^{[N]}} - \bar{y}_l^{[N]}\| \exp(c_1 K^p \|X\|^p_{p,[t_l^{[N]}, t_{k+1}^{[N]}]})$$

for every $t \in [t_l^{[N]}, t_{k+1}^{[N]}]$. In addition, we have

$$\sum_{l=1}^{k} \|X\|^{\lceil \gamma \rceil}_{p,[t_{l-1}^{[N]}, t_l^{[N]}]} \leq \sum_{l=1}^{k} \left(\frac{\|X\|^p_{p,[0,\tilde{T}]}}{N} \right)^{\frac{\lceil \gamma \rceil}{p}} \leq \|X\|^{\lceil \gamma \rceil}_{p,[0,\tilde{T}]} \left(\frac{1}{N} \right)^{\frac{\lceil \gamma \rceil}{p} - 1}$$

and this, together with the induction hypothesis (5.12) and (5.11), shows that

$$\|Y_t - \varphi(t; t_l^{[N]}, \bar{y}_l^{[N]})\| \leq 1$$

for every $t \in [t_l^{[N]}, t_{k+1}^{[N]}]$. Note that $Y_t \in \bar{B}_R(\xi)$ and hence $\varphi(t; t_l^{[N]}, \bar{y}_l^{[N]}) \in \bar{B}_{R+1}(\xi)$. For $l = k+1$, we trivially have $\varphi(t_{k+1}^{[N]}; t_{k+1}^{[N]}, \bar{y}_{k+1}^{[N]}) = \bar{y}_{k+1}^{[N]}$ and for $l = 0$, we have $\varphi(t_{k+1}^{[N]}; t_0^{[N]}, \bar{y}_0^{[N]}) = Y_{t_{k+1}^{[N]}}$ since $t_0^{[N]} = 0$ and $\bar{y}_0^{[N]} = \xi$.
Therefore, we get

$$\|Y_{t_{k+1}^{[N]}} - \bar{y}_{k+1}^{[N]}\| \leq \sum_{l=1}^{k+1} \|\varphi(t_{k+1}^{[N]}; t_l^{[N]}, \bar{y}_l^{[N]}) - \varphi(t_{k+1}^{[N]}; t_{l-1}^{[N]}, \bar{y}_{l-1}^{[N]})\|. \tag{5.13}$$

Moreover, we have

$$\varphi(t_{k+1}^{[N]}; t_l^{[N]}, \bar{y}_l^{[N]}) = \varphi(t_{k+1}^{[N]}; t_l^{[N]}, u_{t_{l-1}^{[N]}, t_l^{[N]}}(1))$$

and

$$\varphi(t_{k+1}^{[N]}; t_{l-1}^{[N]}, \bar{y}_{l-1}^{[N]}) = \varphi(t_{k+1}^{[N]}; t_l^{[N]}, \varphi(t_l^{[N]}; t_{l-1}^{[N]}, \bar{y}_{l-1}^{[N]})).$$

Then, it follows from [22], Theorem 10.26, that there exists a constant c_1, which only depends on γ and p, such that

$$\|\varphi(t_{k+1}^{[N]}; t_l^{[N]}, \bar{y}_l^{[N]}) - \varphi(t_{k+1}^{[N]}; t_{l-1}^{[N]}, \bar{y}_{l-1}^{[N]})\| \leq c_1 K \|u_{t_{l-1}^{[N]}, t_l^{[N]}}(1) - \varphi(t_l^{[N]}; t_{l-1}^{[N]}, \bar{y}_{l-1}^{[N]})\| \\ \times \|X\|_{p,[t_l^{[N]}, t_{k+1}^{[N]}]} \exp(c_1 K^p \|X\|^p_{p,[t_l^{[N]}, t_{k+1}^{[N]}]}). \tag{5.14}$$

Furthermore, it follows from Lemma 5.9 that

$$\|u_{t_{l-1}^{[N]}, t_l^{[N]}}(1) - \varphi(t_l^{[N]}; t_{l-1}^{[N]}, \bar{y}_{l-1}^{[N]})\| \leq c_2 (K \|X\|_{p,[t_{l-1}^{[N]}, t_l^{[N]}]})^{\lceil \gamma \rceil}. \tag{5.15}$$

Note that $\|X\|_{p,[t_l^{[N]}, t_{k+1}^{[N]}]} \leq \|X\|_{p,[0,\tilde{T}]}$. Then, it follows from (5.13), (5.14) and (5.15) that

$$\|Y_{t_{k+1}^{[N]}} - \bar{y}_{k+1}^{[N]}\| \leq c K^{\lceil \gamma \rceil + 1} \|X\|_{p,[0,\tilde{T}]} \exp(c K^p \|X\|^p_{p,[0,\tilde{T}]}) \sum_{l=1}^{k+1} \|X\|^{\lceil \gamma \rceil}_{p,[t_{l-1}^{[N]}, t_l^{[N]}]}$$

and it is an immediate consequence that $\bar{y}_{k+1}^{[N]} \in \bar{B}_{R+1}(\xi)$ because of (5.11), the fact that $Y_{t_{k+1}^{[N]}} \in \bar{B}_R(\xi)$ and

$$\sum_{l=1}^{k+1} \|X\|^{\lceil \gamma \rceil}_{p,[t_{l-1}^{[N]}, t_l^{[N]}]} \leq \sum_{l=1}^{k+1} \left(\frac{\|X\|^p_{p,[0,\tilde{T}]}}{N} \right)^{\frac{\lceil \gamma \rceil}{p}} \leq \|X\|^{\lceil \gamma \rceil}_{p,[0,\tilde{T}]} \left(\frac{1}{N} \right)^{\frac{\lceil \gamma \rceil}{p} - 1}.$$

This finishes the induction and hence shows (5.12).

Now, it follows from (5.12) that

$$\begin{aligned}
\|Y_{\tilde{T}} - \bar{y}_N^{[N]}\| &= \|Y_{t_N^{[N]}} - \bar{y}_N^{[N]}\| \\
&\leq cK^{\lceil\gamma\rceil+1}(\max\{1, \|X\|_{p,[0,\tilde{T}]}\})\exp(cK^p\|X\|^p_{p,[0,\tilde{T}]})\sum_{l=1}^{N}\|X\|^{\lceil\gamma\rceil}_{p,[t_{l-1}^{[N]},t_l^{[N]}]} \\
&\leq cK^{\lceil\gamma\rceil+1}(\max\{1, \|X\|_{p,[0,\tilde{T}]}\})\exp(cK^p\|X\|^p_{p,[0,\tilde{T}]})\sum_{l=1}^{N}\left(\frac{\|X\|^p_{p,[0,\tilde{T}]}}{N}\right)^{\frac{\lceil\gamma\rceil}{p}} \\
&\leq cK^{\lceil\gamma\rceil+1}(\max\{1, \|X\|^{p+1}_{p,[0,\tilde{T}]}\})\exp(cK^p\|X\|^p_{p,[0,\tilde{T}]})\left(\frac{\|X\|^p_{p,[0,\tilde{T}]}}{N}\right)^{\frac{\lceil\gamma\rceil}{p}-1}
\end{aligned}$$

since

$$\|X\|^p_{p,[t_{l-1}^{[N]},t_l^{[N]}]} \leq \|X\|^p_{p,[0,t_l^{[N]}]} - \|X\|^p_{p,[0,t_{l-1}^{[N]}]} = \frac{\|X\|^p_{p,[0,\tilde{T}]}}{N}.$$

This also shows the convergence of $(\bar{y}_N^{[N]})_{N\geq 1}$ to $Y_{\tilde{T}}$ and thus completes the proof of part (i).

Part (ii) follows in exactly the same way as part (i), where we use the fact that

$$\|X\|^p_{p,[s,t]} \leq \|X\|^p_{1/p-\text{Höl},[s,t]}(t-s)$$

holds for every $(s,t) \in \Delta_{\tilde{T}}$ if X is $\frac{1}{p}$-Hölder continuous. □

Remark 5.10 *If the vector fields $F_1, \ldots, F_{d_2}$ are γ-Lipschitz and not just locally γ-Lipschitz, the results from Theorem 5.8 and Lemma 5.9 hold without the restrictions on N or ρ and τ respectively.*

Remark 5.11 *An alternative way to solve rough differential equations numerically are higher order Euler schemes (see Davie [17] for the case $p < 3$ and Friz and Victoir [23] and [22], chapter 10, for the general case). They correspond to the Taylor approximation method described in Kloeden and Platen [31] for stochastic differential equations driven by Brownian motion.*

Using the notation of Theorem 5.8, such a higher order Euler scheme defines ap-

proximations $\tilde{y}_k^{[N]}$ *for* $Y_{t_k^{[N]}}$ *recursively by*

$$\tilde{y}_{k+1}^{[N]} = \tilde{y}_k^{[N]} + \sum_{l=1}^{\lceil\gamma\rceil-1} \sum_{(i_1,\ldots,i_l)\in\{1,\ldots,d_2\}^l} V^{i_1}\ldots V^{i_l}(\mathrm{Id}_{\mathbb{R}^{d_1}})(\tilde{y}_k^{[N]})\, X^{(i_1,\ldots,i_l)}_{t_k^{[N]},t_{k+1}^{[N]}}$$

starting from $\tilde{y}_0^{[N]} = \xi$. *In general, this higher order Euler scheme achieves the same rate of convergence as the ODE method presented in Theorem 5.8. But the ODE method enhances stability (see for example Gyurkó and Lyons [27] and Castell and Gaines [10]).*

5.4 The implementation

The ODE method for the numerical solution of rough differential equations, that has been described in the previous section, has been implemented in C++ and is available in the form of a Microsoft Visual Studio 2008 project on *SourceForge*[1]. It is part of the *CoRoPa* software package that has been developed jointly by various members of the Stochastic Analysis Group of the University of Oxford under the guidance of Terry Lyons.

The *CoRoPa* software package contains a templated tool, called *libalgebra*, for working with tensors, (truncated) Lie series and polynomial vector fields. It has initially been written by Djalil Chafai, while working with Terry Lyons (supported by EPSRC grant GR/R2962/8/01 [HBKBU]), and has recently been substantially updated and rewritten with improved algorithms. The functionality has also been extended. The current version of the *CoRoPa* package provides the C++ programme for the numerical solution of linear and polynomial rough differential equations. Greg Gyurkó, Christian Litterer and Terry Lyons wrote the code for the case of linear rough differential equations, including the first OpenGL visualisations. The author of this thesis, together with Stephen Buckley, developed the code for the polynomial

[1] `http://coropa.sourceforge.net`

case. Chang Liang Xu has helped to upgrade the visualisations and the author has turned them into the web application for Internet Explorer that can be accessed from `http://coropa.sourceforge.net/rde`. The *CoRoPa* package now also includes the implementation of the Brownian rough path and the fractional Brownian rough path which has been written by Terry Lyons and Rahul Raghuram.

The following description is the author's explanation on how to use the *CoRoPa* software package for the numerical solution of polynomial rough differential equations. For simplicity, we demonstrate it by looking at a generic example. The method can then be modified easily so that it is applicable for other rough differential equations. The full source code of this example can be found in the project *RDE_Solver(example)* which is part of the *RDE* package. It can be downloaded at `http://sourceforge.net/projects/coropa`.

Consider the rough differential equation

$$dY_t = \sum_{i=1}^{3} F_i(Y_t)\, dB_t^i, \quad Y_0 = \xi, \tag{5.16}$$

where B is a 3-dimensional Brownian rough path and the vector fields $F_i : \mathbb{R}^3 \to \mathbb{R}^3$ are polynomials for $i = 1, 2, 3$.

At first, we need to define the spaces for the driving signal and the solution. The object

```
matrix_alg_types<2,3,DPReal>
```

identifies the algebra $T^{(2)}(\mathbb{R}^3)$ of truncated tensors up to degree 2 in 3 dimensions, where the coefficients are double precision real numbers. Using this object `matrix_alg_types`, we construct the spaces for the driving signal (`my_alg_type_IN`) and the solution (`my_alg_type_OUT`). The parameter `2` defines the depth and the parameter `3` the dimension. The option `DPReal` indicates that we are working up to double precision, i.e. the coefficients are of type `double`. Alternatively, we can use `SPReal` to work with single precision, i.e. coefficients of type `float`, or the programme allows to work with arbitrary precision arithmetic[2] by choosing the option `Rational`.

Then, we construct a generator `rngen` for normally distributed random variables with a random seed

```
RandomSeed seed;
NormalRandomNumberGenerator rngen(seed.SeedArray,
   seed.SeedArray_size);
```

and the Brownian rough path `BM`.

```
Path_IN BM = MakeBrownianPath<my_alg_type_IN>(rngen);
```

We associate the vector fields F_i with the differential operators

$$V^i = \sum_{j=1}^{3} F_i^j \frac{\partial}{\partial x_j}$$

for $i = 1, 2, 3$. Each of them is constructed as an object of type

```
POLYLIE<my_alg_type_OUT>
```

where the constructor

```
POLYLIE<my_alg_type_OUT>(i,j,k)
```

creates $x_j^k \frac{\partial}{\partial x_i}$. We store the vector fields F_1, F_2, and F_3 altogether in a `vector`, called `theVectorFields`.

Next, we create the class `Solution`, which contains the functionality to compute the first order terms of the solution of the rough differential equation (5.16), i.e. the path Y as defined in section 5.2.

```
Trajectory_OUT Solution = MakeNonLinearSolutionTrajectory
   <my_alg_type_IN,my_alg_type_OUT>(BM,theVectorFields,
   InitialValue,Inf);
```

[2]The *CoRoPa* software package uses the MPIR library for arbitrary precision arithmetic (see `http://www.mpir.org`).

The parameters, that we pass to the constructor, are the Brownian rough path `BM`, the vector fields `theVectorFields`, the initial Value `InitialValue` and the start time `Inf`, which is 0 in this example. The class is templated, so we also need to provide the spaces for the driving signal and for the solution, i.e. `my_alg_type_IN` and `my_alg_type_OUT` respectively.
Now, the command

```
Solution.Value(1)
```

returns the solution Y_1 at time 1.

5.5 Numerical examples

Finally, we present two numerical examples[3]. The first one, a quadratic rough differential equation, is the Landau-Lifshitz-Gilbert Equation and the second one is a linear rough differential equation that characterises fractional Brownian motion on hyperbolic space. Animated versions of these examples can be found online at `http://coropa.sourceforge.net/rde`.

5.5.1 The Landau-Lifshitz-Gilbert Equation

The Landau-Lifshitz-Gilbert Equation is "fundamental for the theory of magnetic memories" and describes the "evolution of spins in ferromagnetic materials under the influence of thermal noise. The necessity to include the thermal noise in the equation was observed by Physicists in the early fifties but the rigorous mathematical theory was missing."[4] Further details on stochastic Landau-Lifshitz-Gilbert Equations can be found in Brzeźniak and Goldys [4], for instance.

[3]They are part of the *CoRoPa* software package (see section 5.4) which has been jointly developed by various members of the Stochastic Analysis Group of the University of Oxford under the guidance of Terry Lyons, including Stephen Buckley, Djalil Chafai, Greg Gyurkó, Christian Litterer, Chang Liang Xu, Rahul Raghuram and the author of this thesis.

[4]Zdzislaw Brzeźniak, University of York, on `http://www.maths.usyd.edu.au/s/scnitm/carberry-Colloquium-Brzezniak-Stoc` (as on 28 April 2011)

In a simple case, the Landau-Lifshitz-Gilbert Equation, perturbed by white noise, reads

$$dy_t = a\,y_t \times dB_t + b\,y_t \times (y_t \times dB_t),$$

where B is a 3-dimensional Brownian rough path and $a \in \mathbb{R}$ and $b > 0$ are real-valued parameters. Figure 5.1 shows a sample path of the solution for parameters $a = b = 1$ and initial value $(0, 0, 1)$. It lives on the sphere.

The Brownian rough path consists of terms of order up to 2. For this example, we solve the ordinary differential equations (5.6), that occur as part of the ODE method, with the fifth order Runge-Kutta-Fehlberg method RKF 54 as described in Butcher [7] or Press et al. [42]. Note that this method has a sufficiently high order of convergence so that the ODE method works.

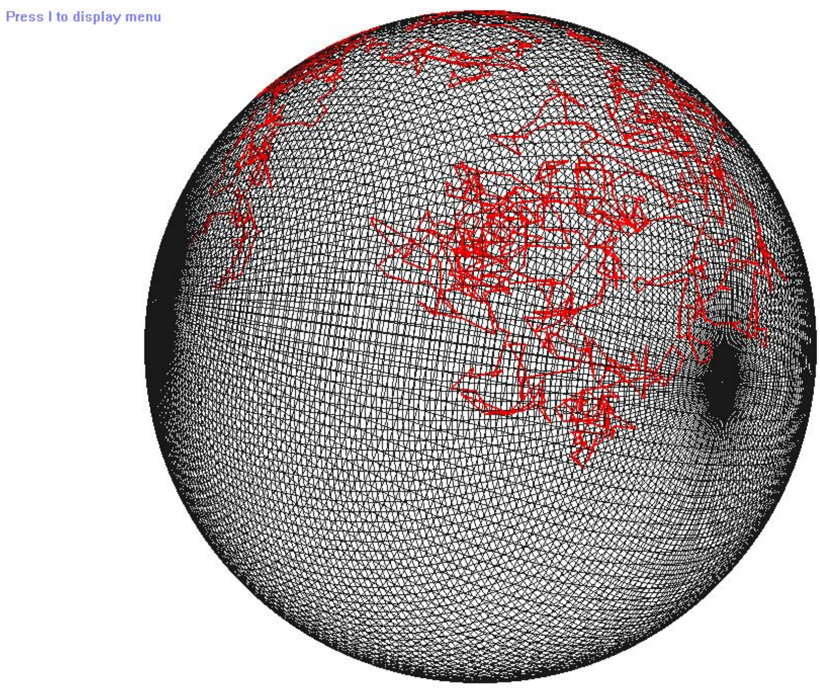

Figure 5.1: Sample path of the solution of the Landau-Lifshitz-Gilbert Equation for parameters $a = b = 1$ and initial value $(0, 0, 1)$

5.5.2 Fractional Brownian motion on hyperbolic space

Consider the linear rough differential equation

$$dy_t = A^1 y_t \, dW_t^1 + A^2 y_t \, dW_t^2,$$

where

$$A^1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

and W is a 2-dimensional fractional Brownian rough path with Hurst coefficient $H > \frac{1}{4}$. Let the initial value be $(0, 0, 1)$. Figures 5.2, 5.3 and 5.4 show sample paths of the solution of this linear rough differential equation for different degrees of roughness with Hurst coefficients 0.7, 0.5 and 0.3 respectively. Note that it lives on hyperbolic space.

In this example, the vector fields are linear. Therefore, we can solve the ordinary differential equations (5.6), that occur as part of the ODE method, by exponentiating the vector fields.

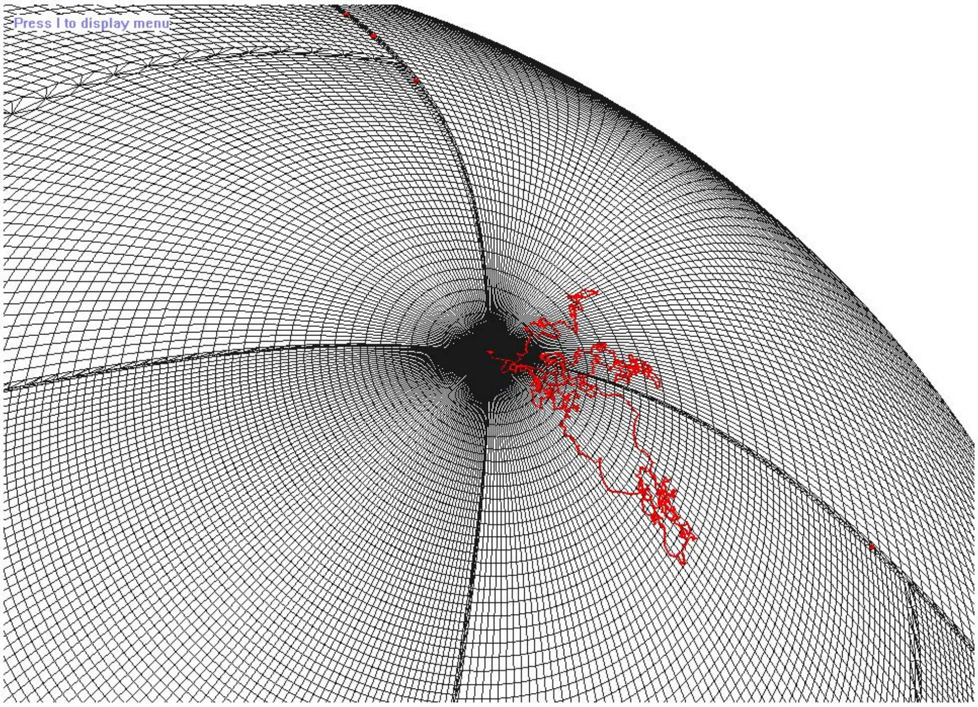

Figure 5.2: Fractional Brownian motion on hyperbolic space with Hurst coefficient 0.7

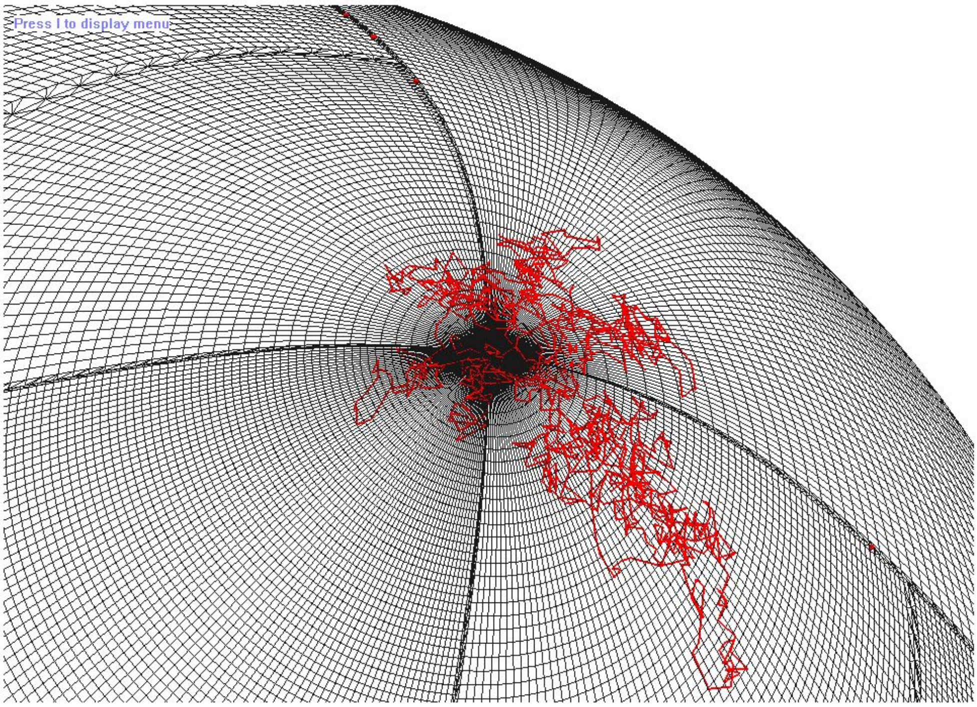

Figure 5.3: Fractional Brownian motion on hyperbolic space with Hurst coefficient 0.5

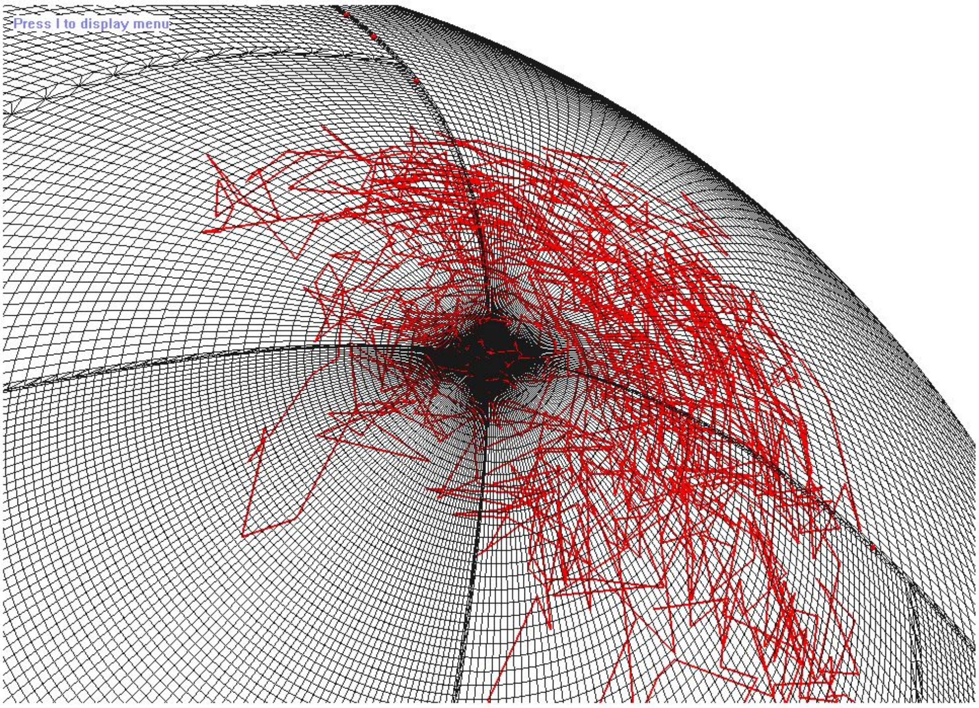

Figure 5.4: Fractional Brownian motion on hyperbolic space with Hurst coefficient 0.3

Chapter 6

Conclusion and outlook

In this thesis, we have constructed a mathematical model of an order driven market where traders can submit limit orders and market orders to buy and sell securities. We have adapted the notion of no free lunch of Harrison and Kreps [29] and Jouini and Kallal [30] to our setting and we have proved a no-arbitrage theorem for the model of the order driven market.

Furthermore, we have computed signatures of order books of different financial markets. Also, we have given a proof which shows that a compactly supported probability measure on signatures is uniquely determined by its expectation. In addition, we have shown how we can obtain the solution of a polynomial differential equation and its (truncated) signature from the signature of the driving signal and the initial value.

We have also presented and analysed an ODE method for the numerical solution of rough differential equations. We have derived error estimates and we have proved that it achieves the same rate of convergence as the corresponding higher order Euler schemes studied by Davie [17] and Friz and Victoir [23] or [22], chapter 10. At the same time, it enhances stability (see Gyurkó and Lyons [27]). The ODE method has been implemented for the case of polynomial vector fields as part of the *CoRoPa* software package which is available at `http://coropa.sourceforge.net`. We have described both the algorithm and the implementation and we have shown by giving examples how it can be used to compute the pathwise solution of stochastic rough differential equations driven by Brownian rough paths and fractional Brownian rough paths.

For future work, we believe that it is interesting to further explore applications of rough paths and signatures in finance. In particular, we think that it can be advantageous to use the ODE method in financial modelling. We expect that it can potentially outperform the corresponding higher order Euler schemes in certain models because of the enhanced stability. For example, Gyurkó and Lyons [27] use it for the Cox-Ingersoll-Ross model.

Also, it is possible to extend the *CoRoPa* software package, so that it can be used to solve rough differential equations with all types of Lipschitz vector fields with sufficient regularity and not just linear or polynomial vector fields. This can be achieved using rough polynomial approximations as described in Caruana [9]. The implementation of the ODE method for polynomial rough differential equations already provides the basis for such an extension.

Bibliography

[1] Arnold, V. I.: Ordinary differential equations, MIT Press, 1973

[2] Bion-Nadal, J.: Bid-ask dynamic pricing in financial markets with transaction costs and liquidity risk, Journal of Mathematical Economics, Volume 45 (2009), pp. 738-750

[3] Bronshtein, I. N., and Semendyayev, K. A.: Handbook of Mathematics, Reprint of the 3rd edition, Springer, 1998

[4] Brzeźniak, Z., and Goldys, B.: Weak Solutions of the Stochastic Landau-Lifshitz-Gilbert Equation, arXiv:0901.0039v1, 2008

[5] Butcher, J. C.: Coefficients for the study of Runge-Kutta integration processes, Journal of the Australian Mathematical Society (1963), 3: 185-201

[6] Butcher, J. C.: An Algebraic Theory of Integration Methods, Mathematics of Computation, 26(117):79-106, 1972

[7] Butcher, J. C.: Numerical methods for ordinary differential equations, 2nd ed., Wiley, 2008

[8] Butcher, J. C.: Trees, B-series and exponential integrators, IMA Journal of Numerical Analysis (2010) **30**, 131-140

[9] Caruana, M.: Peano's theorem for rough differential equations in infinite-dimensional Banach spaces, Proc. London Math. Soc. 100 (1):177-215, 2010

[10] Castell, F.: Asymptotic expansion of stochastic flows, Probability Theory and Related Fields, Vol. 96, No. 2, pp. 225-239, 1993

[11] Castell, F., and Gaines, J.: The ordinary differential equation approach to asymptotically efficient schemes for solution of stochastic differential equations, Ann. Inst. H. Poincaré, Probab. Stat., Vol. 32, No. 2, pp. 231-250, 1996

[12] Chen, K.-T.: Linear independence of exponential of Lie elements, An. Acad. Brasil Ci., 31 (1959), pp. 507-509, as reprinted in Tondeur, P.: Collected papers of K.-T. Chen, Birkhäuser, Boston 2001

[13] Coddington, E. A., Levinson, N.: Theory of ordinary differential equations, Tata McGraw-Hill, 1972

[14] Coutin, L., and Qian, Z.: Stochastic analysis, rough path analysis and fractional Brownian motions, Probability Theory and Related Fields 122, 108-140 (2002)

[15] Crouch, P. E., and Grossman, R. L.: The explicit computation of integration algorithms and first integrals for ordinary differential equations with polynomial coefficients using trees, Proceedings of the 1992 International Symposium on Symbolic and Algebraic Computation, ACM Press, pp. 89-94, 1992

[16] Csikja, R., and Tóth, J.: Blow up in polynomial differential equations, World Academy of Science, Engineering and Technology, 26, pp. 226-231, 2007

[17] Davie, A.M.: Differential Equations Driven by Rough Paths: An Approach via Discrete Approximation, Applied Mathematics Research eXpress, Article ID abm009, 2007

[18] Delbaen, F., and Schachermayer, W.: A general version of the fundamental theorem of asset pricing, Mathematische Annalen, Volume 300 (1994), pages 463-520

[19] Diestel, J.: Sequences and Series in Banach Spaces, Springer, 1984

[20] Fawcett, T.: Problems in Stochastic Analysis. Connections between Rough paths and non-commutative harmonic analysis, DPhil Thesis, University of Oxford, 2003

[21] Field, J., and Large, J.: Pro-Rata Matching and One-Tick Futures Markets, working paper, 2008 (as on 6 May 2011) (http://www.economics.ox.ac.uk/members/jeremy.large/ProRataApril08.pdf)

[22] Friz, P. K., and Victoir, N. B.: Multidimensional Stochastic Processes as Rough Paths. Theory and Applications, Cambridge University Press, 2010

[23] Friz, P., and Victoir, N.: Euler estimates for rough differential equations, Journal of Differential Equations, Vol. 244 (2), pp. 388-412, 2008

[24] Gaines, J. G., and Lyons, T. J.: Variable Step Size Control in the Numerical Solution of Stochastic Differential Equations, SIAM Journal on Applied Mathematics, Vol. 57, No. 5, pp. 1455-1484, 1997

[25] Guasoni, P., Rásonyi, M., and Schachermayer, W.: Consistent price systems and face-lifting pricing under transaction costs, Annals of Applied Probability, Vol. 18, No. 2, pp. 491-520, 2008

[26] Gyurkó, L. G.: Numerical methods for approximating solutions to rough differential equations, DPhil Thesis, University of Oxford, 2008

[27] Gyurkó, L. G., and Lyons, T.: Rough Paths based Numerical Algorithms in Computational Finance, Contemporary Mathematics, Vol. 515: Mathematics in Finance, pp. 17-46, AMS, 2010

[28] Hambly, B., and Lyons, T.: Uniqueness for the signature of a path of bounded variation and the reduced path group, Annals of Mathematics, **171** (2010), pp. 109-167

[29] Harrison, J. M., and Kreps, D. M.: Martingales and Arbitrage in Multiperiod Securities Markets, Journal of Economic Theory, Volume 20 (1979), pp. 381-408

[30] Jouini, E., and Kallal, H.: Martingales and Arbitrage in Securities Markets with Transaction Costs, Journal of Economic Theory, Volume 66 (1995), pp. 178-197

[31] Kloeden , P. E., and Platen, E.: Numerical Solution of Stochastic Differential Equations, corr. 3rd print, Springer, 1999

[32] Kühn, C., and Stroh, M.: Optimal portfolios of a small investor in a limit order market: a shadow price approach, Mathematics and Financial Economics, Volume 3 (2010), pp. 45-72

[33] Lamnabhi-Lagarrigue, F., and Lamnabhi, M.: Algebraic computation of the solution of some nonlinear differential equations, Lecture Notes in Comput. Sci., 162:55-67, 1983

[34] Lax, P. D.: Functional Analysis, Wiley, 2002

[35] Levin, D., and Wildon, M.: A combinatorial method for calculating the moments of Lévy area, Transactions of the American Mathematical Society, Volume 360, Number 12, pp. 6695-6709

[36] Lyons, T.: Differential equations driven by rough signals, Revista Matemática Iberoamericana, 14(2):215-310, 1998

[37] Lyons, T. J., Caruana, M., and Lévy, T.: Differential Equations Driven by Rough Paths, Lecture Notes in Mathematics, École d'Été de Probabilités de Saint-Flour XXXIV-2004, Springer, 2007

[38] Lyons, T., and Ni, H.: Expected signature of two dimensional Brownian Motion up to the first exit time of the domain, arXiv:1101.5902v2

[39] Lyons, T., and Qian, Z.: System Control and Rough Paths, Oxford University Press, 2002

[40] Lyons, T., and Victoir, N.: Cubature on Wiener space, Proc. R. Soc. Lond. A (2004), 460:169-198

[41] Maslov, S.: Simple model of a limit order-driven market, Physica A, Volume 278 (2000), pages 571-578

[42] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: Numerical recipes, 3rd ed., Cambridge University Press, 2007

[43] Rudin, W.: Real and Complex Analysis, second edition, McGraw-Hill, 1974

[44] Smith, E., Farmer, J. D., Gillemot, L., and Krishnamurthy, S.: Statistical theory of the continuous double auction, Quantitative Finance, Volume 3 (2003), pp. 481-514