

Signatures and its application to rough path theory

Isak Hammer^{*†}

August, 2024

^{*}isakhammer@gmail.com [†]Department of Mathematics, UiB

Note. This article is submitted as an examination in the course Algebraic structures for differential equations, computations and flows for the spring semester 2024 at the Department of Mathematics, UiB

Abstract—Signatures have emerged as a powerful tool in machine learning, particularly for efficiently capturing key geometric features of data. In this report, we introduce the foundational concepts of signatures, outlining their essential properties such as invariance under time reparametrization, time-reversibility, and others. We further explore various applications of signatures, highlighting their versatility and utility. Finally, we revise the signature uniqueness theorem, which underscores the theoretical significance of signatures in the context of rough path theory.

I. INTRODUCTION

Signatures have proven to be fundamental for several applications. The theory is particularly applicable in the algebraic theory of rough paths [1, 2, 3]. This has been shown to be a very interesting way to represent data due to the special properties of time-reversal and the invariance to time reparametrization when representing data via signatures [4]. In fact, in a 2013 competition focused on recognizing handwritten Chinese characters, the winning approach represented the characters as arrays based on a "signature" from rough path theory, then classified them using a convolutional neural network [5]. Recently, it has become evident that signatures can be a powerful tool for extracting the geometric shape of any path, even when data is missing.

When the space is discretized and the path is specified by the points it passes through, the resulting representation becomes large and sparse, leading to slow learning. The signature offers a more efficient way to represent a path, helping to mitigate these issues. Importantly, similar smooth curves produce similar signatures. Additionally, the signature is invariant under translations and reparametrizations of the path, making it a robust representation. This, the signature approach represent a non-parametric way of extraction of characteristics features of data.

The goal of this report is to give a short introduction to the importance of signatures and its applications. So following [1] and [4], we divide the report into two parts. In the first part we define signatures, and then state the fundamental properties that is the invariance of time reparametrization, shuffle product time-reversibility, the Chen's Identity and a short introduction to the log signature. In the

second part, we discuss machine learning application and then the applications signatures to the general rough paths theory. Lastly, we state the signature uniqueness theorem.

A. Signatures

The concept of the signature approach is to extract characteristic features from a data, that is a function or data points in a non parametric way. First we want to define the so-called path integral. Consider a two parameterized one dimensional paths $Y_t : [a, b] \rightarrow \mathbb{R}^d$ and $X_t : [a, b] \rightarrow \mathbb{R}$, then we say the path integral of Y_t against X_t is

$$\int_a^b Y_t dX_t = \int_a^b Y_t \dot{X}_t dt, \quad (1)$$

where we defined $\dot{X}_t = \frac{d}{dt}X(t)$. Following [6] and [4], will we explain the basic definition of a signature.

A fundamental piece in the definition of a signature is the so-called path integral. Lets consider the parametrized smooth path of d dimensions be $X_t : [a, b] \rightarrow \mathbb{R}^d$ such that $X_t = \{X_t^1, X_t^2, \dots, X_t^d\}$. Now since each path is $X_t^i : [a, b] \rightarrow \mathbb{R}$ for $i \in \{1, \dots, d\}$, we say define the integral

$$S(X)_{a,t}^i = \int_{a < s < t} dX^i = X_t^i - X_a^i \quad (2)$$

Similarly we define the double-iterated double integral

$$S(X)_{a,t}^{i,j} = \int_{a < s < t} S(X)_{a,s}^i dX_s^j = \int_{a < r < s < t} dX_r^i dX_s^j$$

Continuing recursively we obtain the definition

$$S(X)_{a,t}^{i_1, \dots, i_k} = \int_{a < s < t} S(X)^{i_1, \dots, i_{k-1}} dX_s^{i_k}$$

where $i_1, \dots, i_k \in \{1, \dots, d\}$. Notice that we still obtain the mapping $S(X)^{i_1, \dots, i_k} : [a, b] \rightarrow \mathbb{R}$. Finally we have the tools required to define a signature.

Definition I.1 (Signature). We say a signature of a path $X : [a, b] \rightarrow \mathbb{R}^d$, denoted by $S(X)_{a,b}$ is the collection of all the iterated integrals of X . Thus we, nor have the sequence of numbers

$$S(X)_{a,b} = (1, S(X)_{a,b}^1, \dots, S(X)_{a,b}^d, S(X)_{a,b}^{1,1}, S(X)_{a,b}^{1,2}, S(X)_{a,b}^{2,1}, \dots). \quad (3)$$

Here the first term is defined as 1. Keep in

mind that we iterate over all multi-indexes, that is the set

$$W = \left\{ \begin{array}{l} (i_1, \dots, i_k) \text{ where } k \geq 1, \\ \text{for all } i_1, \dots, i_k \in \{1, \dots, d\}. \end{array} \right\} \quad (4)$$

We denote the set W as words and $A = \{1, \dots, d\}$ as the alphabet of d letters.

One of the most fundamental properties of the signature is its invariance under time reparameterization. This can easily be demonstrated using the definitions of the path integral. Consider two paths $X, Y : [a, b] \rightarrow \mathbb{R}$, which are real-valued. Now, consider two corresponding reparameterized paths $\tilde{X}, \tilde{Y} : [a, b] \rightarrow \mathbb{R}$, where $\tilde{X}_t = X_{\psi(t)}$ and $\tilde{Y} = Y_{\psi(t)}$, with some smooth reparameterization $\psi : [a, b] \rightarrow [a, b]$. From the chain rule it is clear that

$$\frac{d}{dt} \tilde{X}_t = \tilde{X}_t \dot{\psi}(t)$$

, thus it follows that

$$\int_a^b \tilde{Y}_t d\tilde{X}_t = \int_a^b Y_{\psi(t)} \dot{X}_{\psi(t)} \dot{\psi}(t) dt = \int_a^b Y_u dX_u \quad (5)$$

Here we used the substitution $u = \psi(t)$. This is of course applicable in the multidimensional case in the case of a signature. Let $\tilde{X}, X : [a, b] \rightarrow \mathbb{R}^d$ where $\tilde{X}_t = X_{\psi(t)}$. Then we see that

$$S(\tilde{X})_{a,b}^{i_1, \dots, i_k} = S(X)_{a,b}^{i_1, \dots, i_k} \quad (6)$$

for any $i_1, \dots, i_k \in \{1, \dots, d\}$. Thus we see that the signature is, in fact, invariant under time reparameterization.

B. Shuffle product

A fundamental property of the signature is that the product of two signature terms $S(X)_{a,b}^{i_1, \dots, i_k}$ and $S(X)_{a,b}^{j_1, \dots, j_m}$ can be expressed as a sum of terms depending on shuffled multi-indexes.

To formalize this, we define the shuffle product for two multi-indexes. A permutation σ of the set $\{1, \dots, k+m\}$ is called a (k, m) -shuffle if $\sigma^{-1}(1) < \dots < \sigma^{-1}(k)$ and $\sigma^{-1}(k+1) < \dots < \sigma^{-1}(k+m)$. The list $(\sigma(1), \dots, \sigma(k+m))$ forms a shuffle of $(1, \dots, k)$ and $(k+1, \dots, k+m)$. Let $\text{Shuffles}(k, m)$ denote the collection of all such shuffles.

For multi-indexes $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_m)$ with $i_1, \dots, i_k, j_1, \dots, j_m \in \{1, \dots, d\}$, define the multi-index

$$I \sqcup J = \{(\sigma(1), \dots, \sigma(k+m)) \mid \sigma \in \text{Shuffles}(k, m)\}.$$

Thus, we have that The shuffle product $I \sqcup J$ is the set of multi-indexes of length $k+m$.

Theorem I.1. Shuffle product identity For a path $X : [a, b] \rightarrow \mathbb{R}^d$ and multi-indexes $I =$

(i_1, \dots, i_k) and $J = (j_1, \dots, j_m)$, it holds that

$$S(X)_a^b S(X)_a^b = \sum_{K \in I \sqcup J} S(X)_a^{K,b}. \quad (7)$$

Let the terms e_{i_1}, \dots, e_{i_k} be monomials. Then we can denote the representation $S(X)_{a,b}$ as a formal power series

$$S(X)_{a,b} = \sum_{k=0}^{\infty} \sum_{i_1, \dots, i_k} S(X)_{a,b}^{i_1, \dots, i_k} e_{i_1} \cdot \dots \cdot e_{i_k} \quad (8)$$

The main reason why this is fundamental is because this is necessary to state the so-called Chens identity which states the relationship between concatenation and tensor product. First we define the concatenation of paths. That is. For two paths $X : [a, b] \rightarrow \mathbb{R}^d$ and $Y : [b, c] \rightarrow \mathbb{R}^d$, we define the concatenation as the path $X * Y : [a, c] \rightarrow \mathbb{R}^d$, where the first part for $t \in [a, b]$ is $(X * Y)_t = X_t$, and for $t \in [b, c]$ we define $(X * Y)_t = X_b + (Y_t - Y_b)$. And the tensor product is simply defined joining the monomials, that is $e_{i_1} \dots e_{i_k} \otimes e_{j_1} \dots e_{j_m} = e_{i_1} \dots e_{i_k} e_{j_1} \dots e_{j_m}$. Finally we can state the relationship between these operators.

Theorem I.2. Chens Identity. Let $X : [a, b] \rightarrow \mathbb{R}^d$ and $Y : [b, c] \rightarrow \mathbb{R}^d$ be two paths. Then we have the following identity,

$$S(X * Y)_{a,c} = S(X)_{a,b} \otimes S(Y)_{b,c} \quad (9)$$

A very interesting property with the signature (8) is in fact time-reversible. For a path $X : [a, b] \rightarrow \mathbb{R}^d$ we can define the time-reverse path $\bar{X} : [a, b] \rightarrow \mathbb{R}^d$, for which $\bar{X} = X_{a+b-t}$. It can be shown that

$$S(X)_{a,b} \otimes S(\bar{X}) = 1. \quad (10)$$

This can be understood that that the time-reverse is the "tensor" inverse for the signature $S(X)_{a,b}$.

C. Log signature

For a path $X : [a, b] \rightarrow \mathbb{R}^d$, the log signature of X is defined as the formal power series $\log S(X)_{a,b}$. For two formal power series x and y , let us define their Lie bracket by

$$[x, y] = x \otimes y - y \otimes x. \quad (1.55)$$

It is clear that the first few terms of the log signature are given by

$$\begin{aligned} \log S(X)_{a,b} &= \sum_{i=1}^d S(X)_{a,b}^i e_i \\ &+ \sum_{1 \leq i < j \leq d} \frac{1}{2} \left(S(X)_{a,b}^{i,j} - S(X)_{a,b}^{j,i} \right) [e_i, e_j] \\ &+ \dots \end{aligned} \quad (11)$$

As we can see, the log signature is a compressed form of the signature, which is more complicated to

define. For the same amount (number of levels) of the signature, the log signature contains the same information in fewer numbers. The link between the two is continuous.

The log signature is a compressed version of the signature, which has a more intricate definition. For the same number of levels, the log signature retains the same information but with fewer terms. The relationship between the two is continuous.

Specifically for log signatures it is important to consider the structure of Free Lie Algebras. For examples where the log signatures is computed efficiently, see [7, 8].

D. Reconstructing Geometric Properties from the signature

One of the key properties of the shuffle product is that it enables the representation of the signature of a nonlinear function as a linear combination of iterated integrals. Moreover, we observe that the time reversal property, Chen's identity, and the invariance of the signature under time reparametrizations are preserved.

A natural question arises: what kind of information is captured by the signature? Due to the invariance of time reparametrization, it is clear that we cannot reconstruct exact speed at which the path is traversed [4]. However, another interesting property that if we have a parameterized picture of a path X which never crossing itself, then we can completely describe the image and direction of traversal of path. This implies that one can certainly reconstruct geometrical properties, see [9, 10, 2]

II. ROUGH PATHS AND THE SIGNATURE UNIQUENESS THEOREM

Recent work combines signature transforms from rough path theory with neural networks to tackle problems in data science [3, 11]. The central task is to learn the nonlinear relationship between an input data stream (a sequence of points in a vector space) and an output. For instance, this implies predicting the behavior of a dynamical system 2 hours in advance, based on a time series of data from the previous 2 hours. In this section, we will define the motivation and state the signature uniqueness theorem. A challenge with this section is that we will skip many definitions, as rough paths theory is a highly mathematical field. For further details, see [12]. An advantage of this extra theory is that it is generalized and can handle not only Brownian motion but also semimartingales, Markov processes, and Gaussian processes.

Given that an input stream is naturally ordered, it can be represented as a continuous path in some vector space. Following [1], the signature transform

$S(X)$ maps a path $X : [0, T] \rightarrow \mathbb{R}^d$ to a sequence of iterated integrals:

$$S(X) = \begin{pmatrix} X_T - X_0, \\ \int_0^T dX_s \otimes dX_t, \\ \int_0^{t_1} \cdots \int_0^{t_n} dX_{t_1} \otimes \cdots \otimes dX_{t_n}, \\ \cdots \end{pmatrix}$$

In practice, a truncated signature $S_N(x)$ is used as a feature set for learning the relationship between the path and output. This approach can leverage traditional statistical methods or deep learning techniques.

The theoretical foundation of this approach is the *signature uniqueness theorem*, which asserts that every rough path is uniquely determined by its signature, except for tree-like segments. This makes the signature a powerful tool for encoding all relevant information in the data stream. The effectiveness of this method is captured in two key properties:

- **Property 1.** The truncated signature $S_N(x)$ converges to the full signature $S(x)$ rapidly, even for moderate N , ensuring minimal information loss.
- **Property 2.** Polynomial functions on the signature space are linear, as implied by the shuffle product identity (7). Thus, lifting the problem to the signature space linearises it, making the solution robust and model-free.

It is worth mentioning that it is easy to go down the rabbit hole of defining rough paths rigorously; however, to state the theorem, we still need some technical details about three structures and weakly geometric rough paths.

Let us we define some essential tree structures. e say a *real tree* is a metric space (\mathcal{T}, ρ) in which any two distinct points can be joined by a unique continuous, non-self-intersecting path.

From the definition, it is clear that a real tree does not contain non-degenerate loops (i.e. subspaces homeomorphic to the circle). Tree-like paths are defined to be paths that can be realised as loops on a real tree (hence they must possess the aforementioned self-cancelling property).

Now, let $X : [0, T] \rightarrow V$ be a continuous path in some topological space V . We say that X is *tree-like*, if there exist a real tree \mathcal{T} together with continuous mappings

$$\phi : [0, T] \rightarrow \mathcal{T}, \quad \psi : \mathcal{T} \rightarrow V$$

such that $\phi(0) = \phi(T)$ and $x = \psi \circ \phi$. A *tree-like piece* of a continuous path x is a portion $[s, t]$ such that $X|_{[s, t]}$ is tree-like. We say that X is *tree-reduced* if it does not contain any tree-like pieces.

The signature uniqueness theorem can now be stated as follows. It was first proved by [12].

Let $\alpha \in (0, 1]$ be given fixed and denote $N_\alpha \triangleq \lfloor 1/\alpha \rfloor$ (the integer part of $1/\alpha$). In addition, let $\Delta_T = \{(s, t) : 0 \leq t \leq T\}$ be some time interval for variables s and t . Let the truncated tensor algebra over the Banach space V as $T^N(V) = \bigotimes_{n=0}^N V^{\otimes n}$. An α -Hölder rough path over V is a multiplicative functional $X : \Delta_T \rightarrow T^N(V)$ that is α -Hölder continuous in the following sense:

$$\|X^n\|_\alpha = \sup_{0 \leq s < t \leq T} \frac{|X_{s,t}^n|}{(t-s)^{n\alpha}} < \infty, \quad \text{for each } n = 1, \dots, N_\alpha. \quad [8]$$

Utilizing this definition we can define a suitable metric for rough paths. Consider X and \tilde{X} to be α -Hölder continuous, then we define the following metric

$$\rho_\alpha(X, \tilde{X}) = \sum_{n=1}^{N_\alpha} \sup_{0 \leq s < t \leq T} \frac{|X_{s,t}^n - \tilde{X}_{s,t}^n|}{(t-s)^{n\alpha}}. \quad (12)$$

Finally, the space of α -Hölder continuous rough path $X : \Delta_T \rightarrow V$ is geometric if there exists a sequence of α -Hölder continuous rough paths $X^m : \Delta_T \rightarrow V$ with total variation such that

$$\rho(X^m, X) \rightarrow 0 \text{ as } m \rightarrow \infty \quad (13)$$

Theorem II.1. Signature uniqueness theorem.

Let X be a weakly geometric rough path in a Banach space V . Then X has trivial signature if and only if $t \mapsto X_{0,t}$ is tree-like. As a result, two weakly geometric rough paths X, Y have the same signature if and only if they are equal up to tree-like pieces.

For more details, see [12, 13, 14, 2]. This theorem is quite interesting since it proves that for a bounded variation path X , then the signature is $S(X) = 1$ if and only if x is tree-like.

REFERENCES

- [1] Xi Geng. “An Introduction to the Theory of Rough Paths”. In: *Lecture Notes* (2021), p. 9.
- [2] Xi Geng. “Reconstruction for the signature of a rough path”. In: *Proceedings of the London Mathematical Society* 114.3 (2017), pp. 495–526.
- [3] Adeline Fermanian et al. “New directions in the applications of rough path theory”. In: *IEEE BITS the Information Theory Magazine* (2023).
- [4] Ilya Chevyrev and Andrey Kormilitzin. “A primer on the signature method in machine learning”. In: *arXiv preprint arXiv:1603.03788* (2016).
- [5] Fei Yin et al. “Chinese handwriting recognition competition”. In: *International Journal on Document Analysis and Recognition (ICDAR)* 12.4 (2013), pp. 1464–1470.

- [6] Jeremy Reizenstein. “Calculation of Iterated-Integral Signatures and Log Signatures”. In: *Centre for Complexity Science, University of Warwick* (2016).
- [7] Jeremy F Reizenstein and Benjamin Graham. “Algorithm 1004: The iisignature library: Efficient calculation of iterated-integral signatures and log signatures”. In: *ACM Transactions on Mathematical Software (TOMS)* 46.1 (2020), pp. 1–21.
- [8] Jeremy Reizenstein. “Calculation of iterated-integral signatures and log signatures”. In: *arXiv preprint arXiv:1712.02757* (2017).
- [9] Terry J Lyons and Weijun Xu. “Hyperbolic development and inversion of signature”. In: *Journal of Functional Analysis* 272.7 (2017), pp. 2933–2955.
- [10] Jiawei Chang and Terry Lyons. “Insertion algorithm for inverting the signature of a path”. In: *arXiv preprint arXiv:1907.08423* (2019).
- [11] Thomas Cass and Cristopher Salvi. “Lecture notes on rough paths and applications to machine learning”. In: *arXiv preprint arXiv:2404.06583* (2024).
- [12] Horatio Boedihardjo et al. “The signature of a rough path: uniqueness”. In: *Advances in Mathematics* 293 (2016), pp. 720–737.
- [13] Ben Hambly and Terry Lyons. “Uniqueness for the signature of a path of bounded variation and the reduced path group”. In: *Annals of Mathematics* (2010), pp. 109–167.
- [14] Xi Geng. “The signature of a rough path: uniqueness”. In: (2015).