



On the Signature of an Image

Joscha Diehl, Kurusch Ebrahimi-Fard, **Fabian N. Harang** and Samy Tindel

1. Feature extraction: finding the characteristics of time series or image data with signatures
2. From 1 to 2 input parameters: some challenges arising in simple calculus
3. The shuffle product and (a version of) Chen's relation in 2D
4. Challenges and outlook

- The signature associated with paths $x : [0, T] \rightarrow \mathbb{R}^d$ is given as the collection of iterated integrals of the form

$$S(x) = (1, \mathbf{x}^1, \dots, \mathbf{x}^n, \dots) \in T((\mathbb{R}^d)) = \{1\} \oplus \bigoplus_{n=1}^{\infty} (\mathbb{R}^d)^{\otimes n},$$

where $\mathbf{x}^n$ is defined by

$$\mathbf{x}_{s,t}^n = \int_{\Delta_{s,t}^n} \bigotimes_{i=1}^n dx_{r_i} \in (\mathbb{R}^d)^{\otimes n}.$$

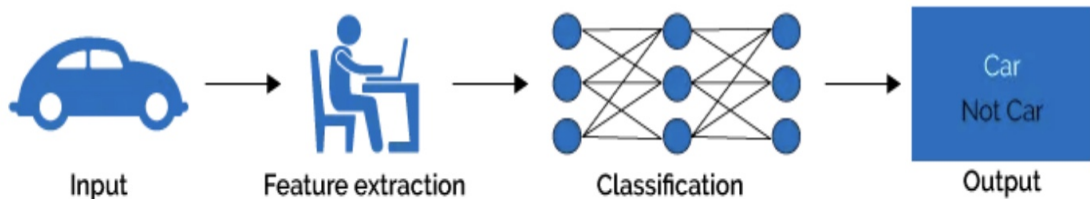
Here $\Delta_{s,t}^n$ denotes the n -th order simplex, such that $(r_1, \dots, r_n) \in \Delta_{s,t}^n$, then $s \leq r_1 \leq \dots \leq r_n \leq t$.

- We let $\mathcal{W}_n$ denote the collection of words of length n created from the alphabet $\{1, \dots, d\}$, and we use the notation

$$\langle S(x)_{s,t}, w \rangle := \int_{\Delta_{s,t}^n} \prod_{i=1}^n dx_{r_i}^{w_i} \in \mathbb{R}.$$

- The path signature was originally studied by K.T. Chen [Chen, 1957], and later proven important for Lyons' theory of rough paths [Lyons, 1998].
- In particular: In order to do pathwise construction of stochastic integrals $\int y \, dx$ when x and y are potentially irregular stochastic processes (beyond martingale setting). Turns out, one "only needs to construct components from the signature.
- The signature has in recent years proven highly successful for machine learning tasks related to data streams (time series data...).

Machine Learning



- Given input $\mathcal{X}$: The goal is to **extract information** to be fed as additional input to the machine learning algorithm.
- The more information is given, the more computationally expensive the learning task gets, and might add unnecessary noise ...
- Want to describe and understand **good features**:
 - Computationally efficient
 - Accurate description of the data distribution
 - If possible: **interpretable**.

- Want to describe and understand **good features**:
 - Computationally efficient: Algebraic structures may reduce computational complexity
 - Accurate description of the data distribution: Probabilistic/analytic understanding of the role of the features
 - If possible: **interpretable**: Understand the 'meaning' of various terms in the features we construct.

Example

Given the data set $\mathcal{X} = \{x_1, x_2, x_3\}$ we can create the features:

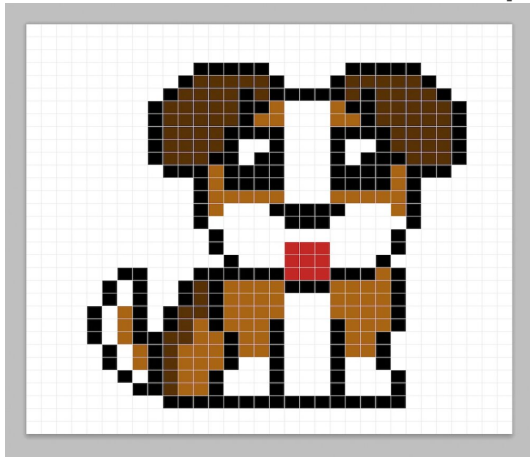
$$\mathcal{S} = \{(x_1 - x_2)^2, (x_2 - x_3)^2, (x_3 - x_1)^2\}.$$

For a machine learning task based on the input data $\mathcal{X}$ we can also include $\mathcal{S}$ as input.

The path signature has several properties which makes it suitable as a feature:

- Invariant to translation: for $a \in \mathbb{R}^d$ then $S(x + a) = S(x)$.
- Invariant to re-parameterization: For a monotonically increasing $\phi : [0, T] \rightarrow [0, T]$ with $\phi(0) = 0$ and $\phi(T) = T$ we have $S(x \circ \phi)_{0,T} = S(x)_{0,T}$.
- Has a rich algebraic structure: (Might) reduce computation time. [Lyons et al., 2007]
- Universal approximation: Linear functionals of the path signature approximate any non-linear functional $f \in C(Lip[0, T]; \mathbb{R})$ arbitrarily well. [Cuchiero et al., 2023]
- Characteristicness: The signature characterize the law of stochastic processes [Chevyrev and Oberhauser, 2022].

We can view an image as a 2-variable function $X : [0, 1]^2 \rightarrow \mathbb{R}^d$.



How do we create good features from an image?

- We believe that the construction of an iterated integral signature over fields $X : [0, T]^2 \rightarrow \mathbb{R}^d$ may be good features for image recognition tasks.
- Understanding the structure of the iterated integral signature for 2D fields is challenging; observed by Chouck/Gubinelli in their Arxiv paper "Rough Sheets" (2014).
- However, recent experimental research shows that the 2D signature have great potential [Zhang et al., 2022].

- Will use $\mathbf{t} = (t_1, t_2) \in [0, T]^2$ and write $X_{\mathbf{t}}$ to denote evaluation of a field $X : [0, T] \rightarrow \mathbb{R}^d$
- The rectangular increment over a rectangle $[\mathbf{s}, \mathbf{t}] = [s_1, t_1] \times [s_2, t_2]$ is defined by

$$\square_{\mathbf{s}, \mathbf{t}} X = X_{\mathbf{t}} + X_{\mathbf{s}} - X_{t_1, s_2} - X_{s_1, t_2}.$$

- Note that

$$\square_{\mathbf{s}, \mathbf{t}} X = \int_{\mathbf{s}}^{\mathbf{t}} \frac{\partial^2 X_{\mathbf{u}}}{\partial u_1 \partial u_2} d\mathbf{u} := \int_{s_1}^{t_1} \int_{s_2}^{t_2} \frac{\partial^2 X_{u_1, u_2}}{\partial u_1 \partial u_2} du_1 du_2 \quad (0.1)$$

- For a (cont.) differentiable field $X : [0, T]^2 \rightarrow \mathbb{R}^d$ we define $dX_{\mathbf{t}} = \frac{\partial^2}{\partial t_1 \partial t_2} X_{t_1 t_2}$. Then for a word $w = [w_1 w_2 \dots w_n] \in \mathcal{W}_n$ we may define the *ld signature* by

$$\langle \mathbf{S}_{\mathbf{st}}^{ld}(X), w \rangle = \int_{\Delta_{[\mathbf{s}, \mathbf{t}]}^n} \prod_{k=1}^n dX_{\mathbf{r}^k}^{w_k}.$$

Where we let $\Delta_{[\mathbf{s}, \mathbf{t}]}^n := \{(\mathbf{r}^1, \dots, \mathbf{r}^n) \in ([0, T]^2)^n \mid \mathbf{s} \leq \mathbf{r}^1 \leq \dots \leq \mathbf{r}^n \leq \mathbf{t}\}$.

- Letting $\mathcal{W}$ denote the collection of all possible words, we view the signature

$$\mathbf{S}_{\mathbf{st}}^{ld}(X) := \{\langle \mathbf{S}_{\mathbf{st}}^{ld}(X), w \rangle \mid w \in \mathcal{W}\}.$$

- In search of an algebraic structure of this collection of integrals, let's see what happens if we multiply two elementary components; namely, multiply the terms given by the words $w = [1]$ and $w' = [2]$. We observe that

$$\langle \mathbf{S}_{\text{st}}^{ld}(X), [1] \rangle \langle \mathbf{S}_{\text{st}}^{ld}(X), [2] \rangle = \langle \mathbf{S}_{\text{st}}^{ld}(X), [12] \rangle + \langle \mathbf{S}_{\text{st}}^{ld}(X), [21] \rangle + K$$

Where K is defined by

$$K = \int_{\Delta_{s,t}^2} dX_{r_1^1, r_2^2}^1 dX_{r_1^2, r_2^1}^2 + \int_{\Delta_{s,t}^2} dX_{r_1^1, r_2^2}^2 dX_{r_1^2, r_2^1}^1$$

- The terms in K is not contained in the ld signature $\mathbf{S}^{ld}(X)$!
- To overcome this, we must extend the signature to include more terms.

Definition

The set of extended words is defined as $\hat{\mathcal{W}} = \bigcup_{n=0}^{\infty} \hat{\mathcal{W}}_n$, where $\hat{\mathcal{W}}_0 = \{(e, \text{id})\}$ and for $n \geq 1$

$$\hat{\mathcal{W}}_n = \{(w, \nu) \mid w \in \mathcal{W}_n, \nu \in \Sigma_{\{1, \dots, n\}}\}. \quad (0.2)$$

Definition

Let X be a function in $C^2([0, T]^2; \mathbb{R}^d)$. The *extended 2D signature* $\mathbf{S}(X)$ is a function in $(C(\Delta^2))^{\hat{\mathcal{W}}}$, defined for an extended word $(w, \nu) \in \hat{\mathcal{W}}_n$ by setting

$$\langle \mathbf{S}_{\text{st}}(X), (w, \nu) \rangle := \int_{\Delta_{[s, t]}^n} \prod_{i=1}^n dX_{r_1^i, r_2^i}^{w_i, \nu_i}.$$

Consider $i_1, i_2, i_3 \in \{1, \dots, d\}$ and $132 \in \Sigma_{\{1,2,3\}}$.

$$\langle \mathbf{S}_{\text{st}}(X), ((i_1, i_2, i_3), 132) \rangle = \int_{\substack{s_1 < r_1^1 < r_1^2 < r_1^3 < t_1 \\ s_2 < r_2^1 < r_2^2 < r_2^3 < t_2}} dX_{r_1^1, r_2^1}^{i_1} dX_{r_1^2, r_2^3}^{i_2} dX_{r_1^3, r_2^2}^{i_3}.$$

Proposition (Shuffle relation)

Let X be a field in $C^2([0, T]^2; \mathbb{R}^d)$ and $\mathbf{S}(X)$ be its extended 2D signature (Definition 3). We consider two words, $(w, \nu) \in \hat{\mathcal{W}}_n$ and $(w', \nu') \in \hat{\mathcal{W}}_{n'}$ (Definition 2). Then for $(\mathbf{s}, \mathbf{t}) \in \Delta_{[0, T]^2}^2$ we have

$$\langle \mathbf{S}_{\mathbf{st}}(X), (w, \nu) \rangle \langle \mathbf{S}_{\mathbf{st}}(X), (w', \nu') \rangle = \sum_{\tau \in Sh(id_n, id_{n'})} \sum_{\lambda \in Sh(\nu, \nu')} \langle \mathbf{S}_{\mathbf{st}}(X), ([ww']_{\tau}, \lambda \circ \tau) \rangle. \quad (0.3)$$

For a word w with $|w| = n$ and a permutation $\tau \in \Sigma_{\{1, \dots, n\}}$ we write $w_{\tau} = (w_{\tau(1)}, \dots, w_{\tau(n)})$.

Proposition (Horizontal and vertical Chen's relation)

Consider a field $X \in C^2([0, T]^2)$. let $\mathbf{S}(X)$ be its associated id signature. Pick a word $w \in \mathcal{W}_n$ for $n \geq 1$. Then for $(\mathbf{s}, \mathbf{t}) \in \Delta_{[0, T]^2}^2$ the following two relations holds:

1. For any $u_1 \in [s_1, t_1]$ we have

$$\langle \delta_{u_1}^1 \mathbf{S}_{\mathbf{st}}(X), w \rangle = \sum_{k=1}^{n-1} \langle \mathbf{S}_{(u_1, s_2), \mathbf{t}}(X), (w_1 \cdots w_k) \rangle *_1 \langle \mathbf{S}_{\mathbf{s}, (u_1, \cdot)}(X), (w_{k+1} \cdots w_n) \rangle.$$

2. For any $u_2 \in [s_2, t_2]$ we have

$$\langle \delta_{u_2}^2 \mathbf{S}_{\mathbf{st}}(X), w \rangle = \sum_{k=1}^{n-1} \langle \mathbf{S}_{(s_1, u_2), \mathbf{t}}(X), (w_1 \cdots w_k) \rangle *_2 \langle \mathbf{S}_{\mathbf{s}, (\cdot, u_2)}(X), (w_{k+1} \cdots w_n) \rangle.$$

Here the convolution products $*_1$ and $*_2$ is defined as integral operations.

- **Invariant to stretching:** For two monotonically increasing functions $\phi_1, \phi_2 : [0, 1] \rightarrow [0, 1]$ with $\phi_1(0) = \pi_2(0) = 0$ and $\phi_2(1) = \phi_2(1) = 1$ then $\mathbf{S}_{0,1}(X_{\phi_1,\phi_2}) = \mathbf{S}_{0,1}(X)$
- **Path perturbation invariant:** For two paths $x_1, x_2 \in Lip([0, 1]; \mathbb{R}^d)$ consider the field $Z_{t_1,t_2} = X_{t_1,t_2} + x_1(t_1) + x_2(t_2)$ then $\mathbf{S}(Z) = \mathbf{S}(X)$.
- **Rotations are contained in the signature** (proof only over 90 degrees now).
- **Universality:** The linear span of extended signature terms is a sub algebra of the continuous functionals on $C([0, T]^2; \mathbb{R}^d)$. We have a universal approximation theorem for the extended signature.
- **The id-signature $\mathbf{S}^{id}(X)$ solves a linear Goursat type PDE:**

$$d\mathbf{S}_{0,t}^{id}(X) = \mathbf{S}_{0,t}^{id}(X) \otimes dX_t, \quad \mathbf{S}_{0,u}^{id}(X) = (1, 0, 0, \dots) \quad \text{for } u \in \delta[0, T]^2.$$

- We need a better understanding of Chen's relation, and the general algebraic structures, see the recent work of Lee-Oberhauser [Lee and Oberhauser, 2023] for some details on this problem.
- Related to the above: Find a computationally efficient recursive structure to compute

$$\langle \mathbf{S}(X), (w, \nu) \rangle.$$

- Can the extended signature characterize the law of random fields?
- How do we construct rough integration with respect to irregular fields: I.e. for two rough fields $X : [0, T]^2 \rightarrow \mathbb{R}^d$ and $Y : [0, T]^2 \rightarrow \mathcal{L}(\mathbb{R}^d; \mathbb{R}^n)$ can we construct the rough integral

$$\int_0^{t_1} \int_0^{t_2} Y_{r_1, r_2} dX_{r_1, r_2}.$$

- With a rough integral: how do we construct solutions to the non-linear hyperbolic PDE

$$dY_{t_1, t_2} = f(y_{t_1, t_2}) dX_{t_1, t_2},$$

when X is a rough field.

- This research is part of the [Signatures for Images](#) project at the Center of Advanced Studies (CAS) in Oslo, led by Kurusch Ebrahimi-Fard and FH. [▶ Link](#)
- We are very grateful to our team visiting us and contributing to this project in the year of 23/24:
 - Joscha Dhiel (Greifswald University)
 - Yvain Bruned (Université de Lorraine)
 - Frederic Patras (Université Côte d'Azur)
 - Samy Tindel (Purdue University)
 - Ilya Chevyrev (Edinburgh University)
 - Christa Cuchiero (University of Vienna).
- Two more events in 2024:
 - Workshop and winter-school: [Rough Paths with Jumps](#) 4.-8. March in Oslo.
 - Conference: [Signatures of paths and images](#) 10.-14. June in Oslo.



Chen, K.-T. (1957).

Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula.
Ann. of Math. (2), 65:163–178.



Chevyrev, I. and Oberhauser, H. (2022).

Signature moments to characterize laws of stochastic processes.
J. Mach. Learn. Res., 23(1).



Cuchiero, C., Primavera, F., and Svaluto-Ferro, S. (2023).

Universal approximation theorems for continuous functions of càdlàg paths and lévy-type signature models.



Lee, D. and Oberhauser, H. (2023).

Random surfaces and higher algebra.



Lyons, T. J. (1998).

Differential equations driven by rough signals.
Rev. Mat. Iberoamericana, 14(2):215–310.

 Lyons, T. J., Caruana, M., and Lévy, T. (2007).

Differential equations driven by rough paths, volume 1908 of *Lecture Notes in Mathematics*.

Springer, Berlin.

Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, With an introduction concerning the Summer School by Jean Picard.

 Zhang, S., Lin, G., and Tindel, S. (2022).

Two-dimensional signature of images and texture classification.

Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 478(2266):20220346.