

## On the signature of an image<sup>☆</sup>

Joscha Diehl<sup>a</sup>, Kurusch Ebrahimi-Fard<sup>b</sup>, Fabian N. Harang<sup>c,\*,\*</sup>, Samy Tindel<sup>d</sup>

<sup>a</sup> Universität Greifswald, Institut für Mathematik und Informatik, Walther-Rathenau-Str. 47, 17489, Greifswald, Germany

<sup>b</sup> Department of Mathematical Sciences, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway,

<sup>c</sup> Department of Economics, BI Norwegian Business School, Handelshøyskolen BI, 0442, Oslo, Norway

<sup>d</sup> Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907, USA

### ARTICLE INFO

#### MSC:

primary 60L10

60L20

60L70

secondary 60L90

#### Keywords:

Signatures

Rough paths theory

Random fields

Feature extraction

Image analysis

### ABSTRACT

Over the past decade, the importance of the 1D signature which can be seen as a functional defined over a path, has been pivotal in both path-wise stochastic calculus and the analysis of time series data. By considering an image as a two-parameter function that takes values in a  $d$ -dimensional space, we introduce an extension of the path signature to images. We address numerous challenges associated with this extension and demonstrate that the 2D signature satisfies a version of Chen's relation in addition to a shuffle-type product. Furthermore, we show that specific variations of the 2D signature can be recursively defined, thereby satisfying an integral-type equation. We analyze the properties of the proposed signature, such as continuity, invariance to stretching, translation and rotation of the underlying image. Additionally, we establish that the proposed 2D signature over an image satisfies a universal approximation property.

### 1. Introduction

To enhance the accuracy of representations in data analysis, it is often important to enrich the original data using nonlinearities. In much of the current literature on data processing, neural networks are often trained without explicit guidance on determining the appropriate type of nonlinearity to be added. Nevertheless, it may be more efficient to begin with a natural and easily interpretable notion of nonlinearity within the relevant context.

With this general understanding, the concept of a 1D signature<sup>1</sup> defined over a curve has proven to be crucial in both path-wise stochastic calculus and the analysis of time series data. For a given ambient dimension  $d \geq 1$  and interval  $[0, T]$ , let us briefly recall the definition of the signature of a differentiable path

$$x = \{x_t^i \mid i = 1, \dots, d, t \in [0, T]\}.$$

To achieve this, let us initially define the set of words on a finite alphabet.

<sup>☆</sup> All authors gratefully acknowledge the support of the Center for Advanced Studies (CAS) in Oslo, which funded the Signatures for Images project during the academic year 2023/2024, enabling the writing of this article. JD was supported by a Trilateral ANR-DGF-JST AI program, DFG Project 442590525. KEF is supported by the Research Council of Norway through project 302831 Computational Dynamics and Stochastics on Manifolds (CODYSMA). He also received support from the project "Pure Mathematics in Norway", which is part of the Mathematics Programme of the Trond Mohn Foundation.

\* Corresponding author.

E-mail addresses: [joscha.diehl@uni-greifswald.de](mailto:joscha.diehl@uni-greifswald.de) (J. Diehl), [kurusch.ebrahimi-fard@ntnu.no](mailto:kurusch.ebrahimi-fard@ntnu.no) (K. Ebrahimi-Fard), [fabian.a.harang@bi.no](mailto:fabian.a.harang@bi.no) (F.N. Harang), [stindel@purdue.edu](mailto:stindel@purdue.edu) (S. Tindel).

URLs: <https://folk.ntnu.no/kurusche/> (K. Ebrahimi-Fard), <https://www.math.purdue.edu/~stindel/> (S. Tindel).

<sup>1</sup> In this work 1D and 2D will correspond to the *parameter* space. All curves or fields (or membranes) will have arbitrary ambient dimension  $d$ .

**Definition 1.1.** Let  $\mathcal{A} = \{1, \dots, d\}$  be an alphabet. Denoting the collection of words with  $n$  letters

$$\mathcal{W}_n := \{w = (i_1, \dots, i_n) \mid n \geq 0 \text{ and } i_j \in \mathcal{A} \text{ for all } j = 1, \dots, n\}, \quad (1.1)$$

where the case  $n = 0$  corresponds to the empty word  $e$ , we define the set of all words over  $\mathcal{A}$  as  $\mathcal{W} := \cup_{n \geq 0} \mathcal{W}_n$  ( $\mathcal{W}' := \mathcal{W} - \{e\}$ ). The length of a word  $w = (i_1, \dots, i_m)$  is denoted  $|w| = m$ .

Now we can define the signature  $S(x)$  as an infinite-dimensional object in  $(C(\Delta_{0,T}^2))^{\mathcal{W}}$ , where the simplex

$$\Delta_{0,T}^2 := \{(s, t) \mid 0 \leq s \leq t \leq T\},$$

by specifying recursively the projections  $\langle S(x), w \rangle$  for every word  $w \in \mathcal{W}$ . Namely, we set  $\langle S_{st}(x), e \rangle := 1$  and if  $w = (i_1, \dots, i_n) \in \mathcal{W}'$ , then we define

$$\langle S_{st}(x), w \rangle := \int_s^t \langle S_{sr}(x), (i_1, \dots, i_{n-1}) \rangle dx_r^{i_n}. \quad (1.2)$$

The rich structure enjoyed by  $S(x)$  has first been emphasized by the mathematician K. T. Chen [1] within an algebraic context. Subsequently, the concept of signature was imported by T. Lyons [2] to the realm of stochastic analysis, leading to the development of the fruitful theory of rough paths. Among the properties which have made the notion of path signature so insightful, let us mention the following ones:

- (i) Algebraic properties like Chen's relation as well as shuffle products. This yields a wealth of structures, such as Lie groups and algebras, as well as Hopf algebras, providing a convenient mathematical framework for studying the path signature  $S(x)$ .
- (ii) Regularity properties in terms of Hölder continuity and  $p$ -variations.
- (iii) Ability to expand solutions of differential equations in terms of  $S(x)$ , leading to pathwise analysis of noisy equations.
- (iv) The fact that a generic curve can be recovered from its signature, up to tree-like equivalence.

We refer e.g. to [3–5] for more details on those properties. In any case, items (i)–(iv) above have been a sufficient motivation to consider signatures as proper features for classification tasks related to curves. See e.g. Refs. [6,7]. Notice the special importance of (iv) in this regard. In the past decade, there has been a significant increase in the importance of utilizing signature methods in diverse data analysis applications. This trend extends across various domains, encompassing Chinese character recognition [8], topological data analysis [9], and the development of signature-based machine learning models for psychiatric diagnosis [10,11].

Based on those preliminary remarks, it is rather natural to ask if appropriate generalizations of the notion of signature can shed new light on the analysis of images. Indeed, even if nonlinear features have been a prominent fixture of image processing in the recent past (see, e.g. [12]), those features are often defined in a somewhat arbitrary fashion. Moreover, one may remark that a precise and transparent algebraic structure, which is useful for basic analysis as well as efficient implementation of those objects, is also lacking in [12] and related papers. Hence, there exists a compelling need for an appropriate extension of the signature concept applicable to 2D-indexed  $\mathbb{R}^d$ -valued fields. Specifically, our objective is to emulate at least some of the desirable properties (i)–(iv) mentioned earlier, bridging the gap in the current literature and fostering a more comprehensive understanding of image analysis.

We briefly mention earlier work where the concept of 2D signature is touched on. In [13], Horozov explores a generalization of Chen's iterated integrals from paths to membranes, i.e., 2D surfaces or fields in our parlance. He describes a shuffle identity that is similar to ours. However, we note two major differences. First, Horozov integrates first-order differentials. Second, he works in ambient dimension 2 which implies homotopy invariance, something that is not true in our more general setting. Lee et al. [14,15] pursue an abstract algebraic-geometric approach, which in the latter reference is built around a categorical generalization of Chen's identity to surfaces. In contrast to these works, our proposed signature is constructed from mixed partial differentials of the underlying field, motivated by multiplicative noisy controls seen in the field of SPDEs. This approach creates different properties of the resulting signature, which will be discussed in depth in the next paragraph. In [16], a two-parameter version of the iterated-sums signature [17,18] is elaborated. Finally, in Ref. [19] images are considered as time-series, by going through them “line by line”.

To extend the path signature (1.2) to surfaces or fields indexed by pairs of parameters from the square  $[0, T] \times [0, T]$ , an initial approach that naturally comes to mind is to employ concepts from rough pathwise calculus within the two-dimensional plane (as introduced in [20,21]). In Ref. [22], the fourth author, together with collaborators, discretized part of the two-parameter analog of the 1D signature considered in [20,21]. A low (that is 15-)dimensional set of features was obtained, with good classification performances for a standard set of textures. This preliminary result provided sufficient encouragement to go further in that particular direction.

In the broader context outlined, the present contribution seeks to establish a conceptually straightforward definition of the signature for 2D-indexed random fields. While the constructions and analysis presented here are naturally deterministic, much in the spirit of rough paths theory, we believe that the insight gained from this analysis has the potential to influence future analysis of random fields, in the same way the study of signatures has influenced the theory on stochastic processes and stochastic differential equations.

A remark is in order regarding item (iii) above. Our point of view is that the full power of rough calculus in the plane (as exhibited in [20,21]) may require a degree of generalization of the signature that could be unnecessarily complex for the specific goals of data analysis. Therefore, with reference to items (i)–(iv) above, we will place less emphasis on item (iii) while prioritizing items (i)–(ii) and (iv). We assert that this shift in emphasis is effectively achieved by considering the various 2D signatures described below. Indeed, we introduce three candidates for 2D signatures over a field  $X = X(s, t)$ : the 2D id-signature,  $S^{\text{id}}(X)$ , the full 2D signature,  $S(X)$ ,

and the symmetrized full 2D signature,  $S^{\text{sym}}(X)$ . However, all of these constructions provide insight into the workings of iterated integrals of 2D fields. Our hope is that this insight can be leveraged to further investigate item (iii) in the context of pathwise solutions to multi-parameter stochastic differential equations.

Let us conclude the introduction with a motivation for the coming analysis through a brief overview of some of the main properties that we will prove to hold for the proposed full signature  $S(X)$ :

- (i) The full 2D signature  $S(X)$  over a field  $X : [0, T]^2 \rightarrow \mathbb{R}^d$  is invariant to translation by constants, but also by 1-parameter paths! This shows that our proposed signature can really be seen as a complementary feature to the classical path signature, and that the 2D signature indeed captures multi-directional relative change in the signal.
- (ii) The full 2D signature of 90-degree rotations of a field  $X$  can be expressed in the 2D signature of  $X$  itself (up to multiplication by  $-1$ ). So, while we cannot state that the 2D signature is rotation invariant, all necessary information related to 90-degree rotations is contained in the signature.
- (iii) Already in the case of one-dimensional ambient space of the field, the proposed full 2D signature is non-trivial (i.e. not equivalent to polynomials in the increment). This is in contrast to the path signature as well as the two-parameter signatures presented in [14,15].
- (iv) The full 2D signature is a continuous functional.
- (v) Certain terms in the full 2D signature can be used in expressing solutions to certain hyperbolic PDEs, known as Goursat equations, with multiplicative noise.
- (vi) The full 2D signature satisfies a variant of the well-known shuffle relation, and thus the linear span of 2D signature elements form an algebra.
- (vii) The full 2D signature is universal, in the sense that linear combinations of signature terms approximate continuous functionals on the space of  $C^2$  fields arbitrarily well.
- (viii) The full 2D signature satisfies a Chen-type relation, where a convolution product is involved. Through certain symmetrization arguments in the integrand, similar to what is used in the Wiener–Itô chaos expansion, one obtains multiplicative Chen-type relations in the horizontal or vertical direction of an image.

While our work provides new results related to the construction of 2D signature, in addition to a detailed overview of the challenges involved in this extension, we also provide discussions of several open problems associated with this structure.

The paper is organized into the following Sections:

- 2 We recall some of the basic structures of the path signature.
- 3 We develop notation, discuss multiparameter integral operators, and provide an overview of calculus in the plane.
- 4 We propose a 2D signature over a two-parameter,  $\mathbb{R}^d$ -valued field. In fact, starting from the expansion of certain PDEs, we derive an object termed the id-signature. Chen’s relation is explored within this 2D framework. However, we will show that the id signature lacks some of the desired properties. This motivates the definition of a full 2D signature, defined over a larger word shuffle algebra. Computations are presented to exemplify the properties of the full 2D signature.
- 5 We symmetrize the integrand, and observe that this new signature-type feature satisfies Chen’s relation.
- 6 A proof is given, showing that the 2D signature characterizes the field.
- 7 We provide a few concluding remarks and discuss open problems and future progress.

## 2. Linear ODEs and the 1D signature

In this section, we recall how signatures for curves parametrized over an interval  $[0, T]$  can be motivated in terms of linear ordinary differential equations. We first show how the signature  $S(x)$  in (1.2) emerges from  $\mathbb{R}^e$ -valued differential equations. Then we identify the signature itself as the solution of a linear equation in an infinite dimensional space. Those notions will be generalized later to a multi-parametric setting.

### 2.1. Motivation from real-valued differential equations

Signatures of curves parametrized over  $[0, T]$  appear naturally in the context of linear fixed point equations (e.g., see [23] for this perspective in the context of Volterra equation).

**Lemma 2.1.** *Let  $x : [0, T] \rightarrow \mathbb{R}^d$  be a  $C^1$  signal and let  $\{A^i \mid i = 1, \dots, d\} \subset \mathbb{R}^{e \times e}$  be a collection of constant matrices with real entries. We consider the  $\mathbb{R}^e$ -valued solution  $y$  to the linear equation*

$$y_{st} := y_t - y_s = \sum_{i=1}^d \int_s^t A^i y_r dx_r^i, \quad (2.1)$$

for  $(s, t) \in \Delta_{0,T}^2$ , with a given initial condition  $y_s \in \mathbb{R}^d$ . Recalling (1.1), we set

$$A^{\circ w} := A^{i_n} \dots A^{i_1}, \quad \text{for } w = (i_1, \dots, i_n) \in \mathcal{W}', \quad (2.2)$$

and  $A^\circ$  is the  $e \times e$  identity matrix  $\mathbf{1}_e$ . Then Eq. (2.1) can be expanded as

$$y_t = \sum_{w \in \mathcal{W}} A^{\circ w} y_s \langle S_{st}(x), w \rangle, \quad (2.3)$$

where the quantity  $\langle S_{st}(x), w \rangle$  was introduced in (1.2).

**Proof.** We will provide a brief outline of the proof here, establishing the foundation for the forthcoming exploration of the two-parameter case. Employing Einstein's summation convention for repeated upper indices, Eq. (2.1) transforms into:

$$y_{st} = \int_s^t A^{i_1} y_{r_1} dx_{r_1}^{i_1}, \quad \text{for } (s, t) \in \Delta_{0,T}^2.$$

Replacing  $y_{r_1}$  by  $y_{sr_1} + y_s$  on the righthand side yields

$$y_{st} = A^{i_1} y_s \int_s^t dx_{r_1}^{i_1} + R_{st}^1, \quad \text{where } R_{st}^1 := \int_s^t A^{i_1} y_{sr_1} dx_{r_1}^{i_1}. \quad (2.4)$$

Repeating this step, the increment  $y_{sr_1}$  in the remainder term  $R_{st}^1$  can be further expanded

$$y_{sr_1} = A^{i_2} y_s \int_s^{r_1} dx_{r_2}^{i_2} + \int_s^{r_1} A^{i_2} y_{sr_2} dx_{r_2}^{i_2}.$$

Plugging this into (2.4), we end up with

$$y_{st} = A^{i_1} y_s \int_s^t dx_{r_1}^{i_1} + A^{i_1} A^{i_2} y_s \int_{s \leq r_2 < r_1 \leq t} dx_{r_2}^{i_2} dx_{r_1}^{i_1} + R_{st}^2. \quad (2.5)$$

The expression of the remainder  $R_{st}^2$  follows the same procedure. In addition, observe that invoking our conventions (1.2) and (2.2), one can recast (2.5) as

$$y_{st} = A^{\circ(i_1)} y_s \langle S_{st}(x), (i_1) \rangle + A^{\circ(i_2, i_1)} y_s \langle S_{st}(x), (i_2, i_1) \rangle + R_{st}^2.$$

Our claim (2.3) follows from iterating this procedure.  $\square$

## 2.2. The algebra of signatures

Following [4], we start this subsection by recalling the algebraic and geometric setting for 1D signatures indexed by a single parameter. This will allow us to provide a linear differential equation which governs the signature of a path. We will try to replicate this linear equation for images in the next sections.

We begin by recalling the definition of tensor algebra over the space  $\mathbb{R}^d$

$$\mathcal{T}(\mathbb{R}^d) := \bigoplus_{n=0}^{\infty} (\mathbb{R}^d)^{\otimes n},$$

with  $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}\mathbf{1}$ . It is equipped with a product  $\otimes$  defined for any  $g, h \in \mathcal{T}(\mathbb{R}^d)$  by

$$[g \otimes h]^n = \sum_{k=0}^n [g]^{n-k} \otimes [h]^k, \quad (2.6)$$

where  $[g]^n \in (\mathbb{R}^d)^{\otimes n}$  designates the projection onto the  $n$ th tensor level, making  $\mathcal{T}(\mathbb{R}^d)$  an associative and non-commutative algebra with unit  $\mathbf{1} \in (\mathbb{R}^d)^{\otimes 0}$ .

The space  $\mathcal{T}(\mathbb{R}^d)$  can be equipped with the commutative shuffle product defined inductively, i.e.,  $w \sqcup \mathbf{1} = \mathbf{1} \sqcup w = w$  for any  $w \in \mathcal{T}(\mathbb{R}^d)$  and for  $u \otimes w_1, v \otimes w_2, u, v \in (\mathbb{R}^d)^{\otimes 1}, w_1, w_2 \in \mathcal{T}(\mathbb{R}^d)$ , we define

$$(u \otimes w_1) \sqcup (v \otimes w_2) := u \otimes (w_1 \sqcup (v \otimes w_2)) + v \otimes ((u \otimes w_1) \sqcup w_2). \quad (2.7)$$

For example, shuffling two letters,  $u$  and  $v$ , gives

$$u \sqcup v = u \otimes v + v \otimes u.$$

See Notation 3.2 for an explicit formulation of shuffle product in terms of permutations. Further below (see Notation 4.27), we will use word notation to denote elements in  $\mathcal{T}(\mathbb{R}^d)$ , i.e., if  $e_1, \dots, e_d$  denotes the canonical basis of  $\mathbb{R}^d$ , then we use  $(i_1, \dots, i_n) \in \mathcal{W}$  to denote  $e_{i_1} \otimes \dots \otimes e_{i_n} \in \mathcal{T}(\mathbb{R}^d)$ .

The space of tensor series is denoted  $\mathcal{T}((\mathbb{R}^d))$ . It contains linear maps  $F := \sum_{w \in \mathcal{W}} f_w e_w$  from  $\mathcal{T}(\mathbb{R}^d)$  to  $\mathbb{R}$ , where

$$\left\{ e_w = e_{i_1} \otimes \dots \otimes e_{i_n} ; w = (i_1, \dots, i_n) \in \mathcal{W} \right\} \quad (2.8)$$

such that for  $v \in \mathcal{W}$ , we define  $F(v) = \sum_{w \in \mathcal{W}} f_w \langle e_w, v \rangle = f_v$ . The tensor product (2.6) can be extended to  $\mathcal{T}((\mathbb{R}^d))$  making it an unital algebra, with unit  $\mathbf{1}$  (in place for the empty word).

A continuous map  $\Phi : \Delta_{a,b}^2 \rightarrow \mathcal{T}((\mathbb{R}^d))$  is called a multiplicative functional if

$$\Phi_{st} = \Phi_{su} \otimes \Phi_{ut},$$

for  $s < u < t$ . The signature  $S(x)$  of a path  $x$  with finite variation is a particular example of such a multiplicative functional. Indeed, recalling (1.2) we can embed  $\langle S(x), w \rangle$  by setting

$$S_{st}(x) = \mathbf{1} + \sum_{w \in \mathcal{W}'} \langle S_{st}(x), w \rangle e_w \in \mathcal{T}((\mathbb{R}^d)). \quad (2.9)$$

In addition, it can be shown that  $S_{st}(x)$  lives in a subset  $\mathcal{G}(\mathbb{R}^d) \subset \mathcal{T}((\mathbb{R}^d))$  consisting of so-called group-like elements

$$\mathcal{G}(\mathbb{R}^d) = \exp^{\otimes}(\mathcal{L}((\mathbb{R}^d))),$$

where  $\mathcal{L}((\mathbb{R}^d))$  denotes the set of all Lie series over  $\mathbb{R}^d$ . Hence, every element in  $\mathcal{G}(\mathbb{R}^d)$  is the tensor-exponential of a Lie series. Regarding the latter, recall that, denoting by  $\mathcal{L}(\mathbb{R}^d)$  the free Lie algebra over  $\mathbb{R}^d$ , an element in  $\mathcal{T}((\mathbb{R}^d))$  is a Lie series if it can be written in the form  $\sum_{n \geq 1} p_n$ , where each Lie polynomial  $p_n \in \mathcal{L}(\mathbb{R}^d)$  has support in  $\mathcal{W}_n$ .

As seen above, the notion of signature can be motivated by linear differential equations. We now state a differential/integral equation in  $\mathcal{T}((\mathbb{R}^d))$  for  $S_{st}(x)$  itself. It should be seen as a simple non-truncated version of [4, Prop. 7.8].

**Proposition 2.2.** For  $(s, t) \in \Delta_{a,b}^2$ , the signature  $S_{st}(x)$  defined by (2.9) solves a linear integral fixed point equation in  $\mathcal{T}((\mathbb{R}^d))$  of the form

$$S_{st}(x) = \mathbf{1} + \int_s^t S_{sr}(x) \otimes dx_r. \quad (2.10)$$

### 3. Simplexes and integration

This section is devoted to collecting some notations as well as elementary results frequently used across the article in the context of calculus in the plane.

#### 3.1. Rectangles and simplexes

Throughout the text, rectangles in the square  $[0, T]^2 \subset \mathbb{R}^2$  will play a crucial role. We first spell out our convention and notation for those objects.

**Convention:** we refer to the horizontal and vertical axis in the plane as the first respectively second variable, indexed respectively by subscripts 1 and 2. See Fig. 1.

**Notation 3.1.** A generic rectangle  $R$  in  $\mathbb{R}^2$  is of the form  $R = [s_1, t_1] \times [s_2, t_2]$ , with  $s_1 \leq t_1$  and  $s_2 \leq t_2$ . A more compact notation follows by considering points in  $\mathbb{R}^2$ , setting  $\mathbf{s} = (s_1, s_2)$  and  $\mathbf{t} = (t_1, t_2)$ , where  $s_1 \leq t_1 \leq T$ ,  $s_2 \leq t_2 \leq T$ . We then denote by  $R = [\mathbf{s}, \mathbf{t}] := [s_1, t_1] \times [s_2, t_2]$  a sub-rectangle specified by its lower left and upper right corners, inside the standard-square  $[0, T]^2$ . We shall also write  $\mathbf{T} = (T, T)$  for any positive quantity  $T$ .

**Notation 3.2.** Denote by  $\Sigma_{\mathcal{A}}$  the set of permutations of elements of a finite set  $\mathcal{A}$ . Furthermore, for  $n, k \geq 1$  we consider the following set of permutations of the set  $\{1, \dots, n+k\}$ :

$$\text{Sh}(n, k) = \{ \rho \in \Sigma_{\{1, \dots, n+k\}} \mid \rho(i) < \rho(j) \text{ if } 1 \leq i < j \leq n \text{ or } n+1 \leq i < j \leq n+k \}. \quad (3.1)$$

Our upcoming computations employ the concept of rectangular increment, and will heavily depend on simplexes, including a variant that incorporates permutations.

**Definition 3.3.** Let  $a \leq b$  be elements in  $[0, T]$ . Then we define the following simplexes:

(i) The usual simplex of order  $n$  is given by

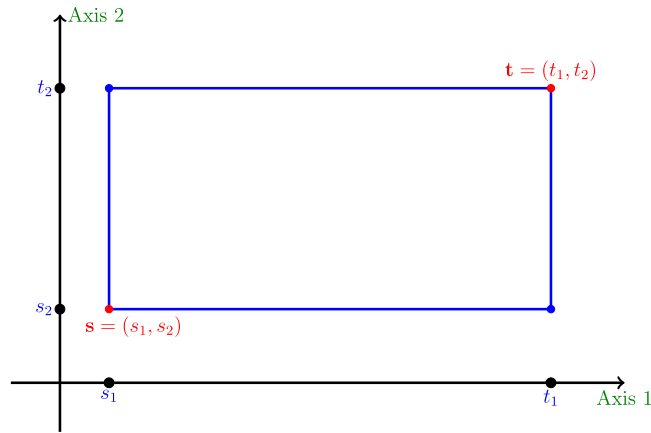
$$\Delta_{a,b}^n = \{ r \in [0, T]^n \mid a \leq r^1 \leq \dots \leq r^n \leq b \}. \quad (3.2)$$

(ii) Consider a permutation  $\sigma \in \Sigma_{\{1, \dots, n\}}$ . Then we define the  $\sigma$ -simplex

$$\Delta_{a,b}^{\sigma} = \{ r \in [0, T]^n \mid a \leq r^{\sigma(1)} \leq \dots \leq r^{\sigma(n)} \leq b \}. \quad (3.3)$$

**Remark 3.4.** Note the use of superscripts in (3.2) and (3.3) to denote components of  $r \in [0, T]^n$ .

It is well-known that permutations carry a shuffle property, which can be translated into a corresponding property for simplexes. We recall this elementary result for further use. For a proof of this statement, we refer the reader to Patras [24, Prop. 2].

Fig. 1. Typical example of rectangle  $[s, t]$  in  $[0, T]^2$ .

**Lemma 3.5.** Consider two positive integers  $n, k$ , and let  $\sigma \in \Sigma_{\{1, \dots, n\}}$  and  $\tau \in \Sigma_{\{1, \dots, k\}}$ . For  $a \leq b$  we define  $\Delta_{a,b}^n \sigma$  and  $\Delta_{a,b}^k \tau$  like in (3.3). Then the product  $\Delta_{a,b}^n \sigma \times \Delta_{a,b}^k \tau$  can be “linearized”:

$$\Delta_{a,b}^n \sigma \times \Delta_{a,b}^k \tau = \bigcup_{\rho \in \text{Sh}(n,k)} \Delta_{a,b}^{n+k} (\sigma * \tau) \circ \rho^{-1}, \quad (3.4)$$

where  $\sigma * \tau \in \Sigma_{\{1, \dots, n+k\}}$  is defined as

$$(\sigma * \tau)(i) := \begin{cases} \sigma(i) & i \leq n \\ \tau(i - n) + n & n + 1 \leq i. \end{cases}$$

The union is disjoint up to sets of zero Lebesgue measure.

We close this section by defining the main type of simplex which will be used in the context of integration in the plane.

**Notation 3.6.** Let  $s = (s_1, s_2)$  and  $t = (t_1, t_2)$  be two elements of  $[0, T]^2$ , where we are using Notation 3.1 on rectangles. Consider an integer  $n \geq 2$ . We denote the set

$$\Delta_{[s,t]}^n := \left\{ (\mathbf{r}^1, \dots, \mathbf{r}^n) \in ([0, T]^2)^n; s_1 \leq r_1^1 \leq \dots \leq r_1^n \leq t_1 \text{ and } s_2 \leq r_2^1 \leq \dots \leq r_2^n \leq t_2 \right\}. \quad (3.5)$$

### 3.2. Some properties of integral operators

To facilitate future computations associated with the construction and properties of signature over an image, this section will revisit several essential properties of iterated integral operators. Commencing with a straightforward definition, we clarify the concept of an iterated integral operator:

**Definition 3.7.** Let  $n$  be a positive integer, and consider a measurable set  $A \subset [0, T]^n$ . We write  $\mathcal{L}(V; W)$  for the set of bounded linear operators from a vector space  $V$  to a vector space  $W$ . Then we define an integral operator  $\int_A$  as an element in  $\mathcal{L}(L^1([0, T]^n); \mathbb{R})$  such that

$$\int_A : X \in L^1([0, T]^n) \mapsto \int_A X := \int \dots \int_A X(r_1, \dots, r_n) dr_1 \dots dr_n \in \mathbb{R}. \quad (3.6)$$

In the sequel, we shall compose integral operators and therefore introduce a particular notation for that purpose.

**Definition 3.8.** We define two types of compositions of integral operators as follows:

(i) Let  $A \subset [0, T]^m$  and  $B \subset [0, T]^n$  be two measurable sets. Then we define

$$\int_A \star \int_B \equiv \int_{A \times B}, \quad (3.7)$$

where the right hand side of (3.7) is understood as in Definition 3.7.

(ii) For two measurable sets  $A, B \subset \mathbb{R}^n$ , we define

$$\int_A \star' \int_B \equiv \int_{A \cap B}, \quad (3.8)$$

where the “interlacing” of  $A$  and  $B$  is defined as the set

$$A \sqcap B := \{(r_1^1, r_2^1, r_1^2, r_2^2, \dots, r_1^n, r_2^n) \in \mathbb{R}^{2n} \mid (r_1^1, \dots, r_1^n) \in A, (r_2^1, \dots, r_2^n) \in B\}.$$

Note that, if  $A, B \subset \mathbb{R}$ , then  $\int_A \star \int_B = \int_A \star' \int_B$ .

The following “interchange law” is immediate.

**Lemma 3.9.** For  $A, B \subset \mathbb{R}^m, A', B' \subset \mathbb{R}^n$ ,

$$\left( \int_A \star \int_{A'} \right) \star' \left( \int_B \star \int_{B'} \right) = \left( \int_A \star' \int_B \right) \star \left( \int_{A'} \star' \int_{B'} \right).$$

**Example 3.10.** As an illustration of the previous notation, let us consider the particular case of simplexes. Here we use [Definition 3.3](#) and [Notation 3.6](#). Consider now  $(s_1, t_1)$  and  $(s_2, t_2)$  in  $\Delta_{0,T}^2$ , where we recall the notation given in [\(3.2\)](#) for  $\Delta_{a,b}^n$ . Then

$$\int_{\Delta_{s_1,t_1}^n} \star \int_{\Delta_{s_2,t_2}^n} = \int_{\Delta_{s_1,t_1}^n \times \Delta_{s_2,t_2}^n}$$

and

$$\int_{\Delta_{s_1,t_1}^n} \star' \int_{\Delta_{s_2,t_2}^n} = \int_{\Delta_{[s,t]}^n}.$$

The main reason why we have introduced the notion of integral operator is the following: we wish to highlight the fact that the usual algebraic properties of iterated integrals can be reduced to manipulations on simplexes. Let us first rephrase Chen’s relation in this framework.

**Lemma 3.11.** Let  $T > 0$  and consider  $(s, u, t) \in \Delta_{0,T}^3$ . Recall our notation [\(3.6\)](#) for integral operators and relation [\(3.7\)](#) for compositions. Then we have that

$$\int_{\Delta_{s,t}^n} = \sum_{l+k=n} \int_{\Delta_{s,u}^l} \star \int_{\Delta_{u,t}^k}. \quad (3.9)$$

**Proof.** The following fact about simplexes is well-known and easy to prove:

$$\Delta_{s,t}^n = \bigcup_{l+k=n} \Delta_{s,u}^l \times \Delta_{u,t}^k, \quad (3.10)$$

where the union is disjoint, up to a set of measure zero, and where it is understood that  $\Delta_{s,u}^n \times \Delta_{u,t}^0 = \Delta_{s,u}^n = \Delta_{u,t}^0 \times \Delta_{s,u}^n$ . Our claim [\(3.9\)](#) then follows.  $\square$

**Remark 3.12.** As a particular case of [Lemma 3.11](#), we will later use the following relation for  $(s, u, t) \in \Delta_{0,T}^3$ :

$$\int_{\Delta_{s,u}^1} \star \int_{\Delta_{u,t}^1} = \int_{\Delta_{s,t}^2} - \int_{\Delta_{s,u}^2} - \int_{\Delta_{u,t}^2}. \quad (3.11)$$

**Remark 3.13.** We can recast relation [\(3.9\)](#) in the language of algebraic integration as described in [\[25\]](#). To this aim we define the space  $C = C_1$  of continuous functions  $x : [0, T] \rightarrow \mathbb{R}^d$ . We also set  $C_2$  as the space of continuous functions  $y : [0, T]^2 \rightarrow \mathbb{R}^d$  vanishing on the diagonal (i.e.  $y_{ss} = 0$ ). Next we define an operation  $\delta : C_i \rightarrow C_{i+1}$  by

$$\delta_t x_s = x_t - x_s, \quad \text{and} \quad \delta_u y_{s,t} = y_{s,t} - y_{s,u} - y_{u,t}. \quad (3.12)$$

Then [Lemma 3.11](#) asserts that for  $(s, u, t) \in \Delta_{0,T}^3$

$$\delta_u \int_{\Delta_{s,t}^n} \equiv \int_{\Delta_{s,t}^n} - \int_{\Delta_{s,u}^n} - \int_{\Delta_{u,t}^n} = \sum_{\substack{l+k=n \\ k,l>0}} \int_{\Delta_{s,u}^l} \star \int_{\Delta_{u,t}^k}. \quad (3.13)$$

As mentioned in the introduction, iterated integral operators satisfy the shuffle property. This relation is rather useful in both the analysis and applications of the path signature. We write it here purely in terms of the integral operator.

**Proposition 3.14** (Shuffle Relation of Permuted Integral Operators). Recall the notion of permuted simplex,  $\Delta_{s,t}^n \rho$ , introduced in [\(3.3\)](#). For  $s \leq t \in [0, T]$  and permutations  $\rho \in \Sigma_{\{1, \dots, n\}}$  and  $\nu \in \Sigma_{\{1, \dots, m\}}$ , we have that

$$\int_{\Delta_{s,t}^n \rho} \star \int_{\Delta_{s,t}^m \nu} = \sum_{\sigma \in \text{Sh}(n,m)} \int_{\Delta_{s,t}^{n+m} (\rho * \nu) \circ \sigma^{-1}}, \quad (3.14)$$

where  $\text{Sh}(n, m)$  is given in [Notation 3.2](#) and the product  $*$  on permutations is defined in [Lemma 3.5](#).

**Proof.** Owing to (3.7), we have

$$\int_{\Delta_{s,t}^n \rho} \star \int_{\Delta_{s,t}^m \nu} = \int_{\Delta_{s,t}^n \rho \times \Delta_{s,t}^m \nu}.$$

Our claim (3.14) is then a direct consequence of Lemma 3.5.  $\square$

We conclude this section with a useful symmetrization property relating iterated integral operators to integral operators over hypercubes.

**Lemma 3.15.** *Let  $(s, t) \in \Delta_{0,T}^2$ . Then*

$$\sum_{\rho \in \Sigma_{\{1, \dots, n\}}} \int_{\Delta_{s,t}^n \rho} = \int_{[s,t]^n}.$$

**Proof.** This is known to result from integration by parts. In geometric terms, it follows from the fact that the hypercube  $[s, t]^n$  can be divided into  $n!$  different simplexes defined by all possible ways of ordering  $n$  variables from  $[s, t]^n$ . This is described by all possible permutations of the set  $\{1, \dots, n\}$ , and thus we have

$$\bigcup_{\rho \in \Sigma_{\{1, \dots, n\}}} \Delta_{s,t}^n \rho = [s, t]^n. \quad \square$$

### 3.3. Increments in the plane

In the sequel, a field  $X$  is indexed by a pair of elements from the standard-square  $[0, T]^2$  (analogous to a curve  $x$  being indexed by elements from  $[0, T]$ ). We first introduce notation for the so-called rectangular increment of a field  $X$ .

**Notation 3.16.** Let  $X \in C^2([0, T]^2; \mathbb{R}^d)$  be a field. For  $\mathbf{r} = (r_1, r_2)$  we set

$$\partial_1 X_{\mathbf{r}} = \frac{\partial X_{\mathbf{r}}}{\partial r_1}, \quad \partial_2 X_{\mathbf{r}} = \frac{\partial X_{\mathbf{r}}}{\partial r_2}, \quad \text{and} \quad \partial_{12} X_{\mathbf{r}} = \frac{\partial^2 X_{\mathbf{r}}}{\partial r_1 \partial r_2}. \quad (3.15)$$

For a function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , we also use the standard notation  $\partial_i f(x) = \frac{\partial f}{\partial x^i}(x)$ .

**Definition 3.17.** Consider an  $\mathbb{R}^d$ -valued function  $X : [0, T]^2 \rightarrow \mathbb{R}^d$ , and a generic rectangle  $R = [s, \mathbf{t}]$ , specified by its lower left and upper right corners,  $s = (s_1, s_2)$  respectively  $\mathbf{t} = (t_1, t_2)$  – using Notation 3.1. Then the **rectangular increment** of  $X$  over  $R$  is defined to be

$$\square_{\mathbf{st}} X = X_{t_1; t_2} - X_{s_1; t_2} - X_{t_1; s_2} + X_{s_1; s_2}.$$

Note that, still in line with Notation 3.1, we further condense notation by writing

$$X_{\mathbf{r}} := X_{r_1; r_2},$$

for  $\mathbf{r} = (r_1, r_2) \in [0, T]^2$ .

Recall that the 1D signature is built on the concept of increments. In particular, we observe that  $\langle S_{st}(x), (i_1) \rangle = x_t^{i_1} - x_s^{i_1}$ . The rectangular increment defined above naturally extends this to the 2D setting

$$\square_{\mathbf{st}} X = \int_{\Delta_{[s, \mathbf{t}]}^1} \partial_{12} X_{s_1; s_2} ds_1 ds_2,$$

when  $X$  is twice continuously differentiable. With this in mind, using Notation 3.16, we define the following convention for the mixed partial differentials.

**Notation 3.18.** Consider a function  $X \in C^2([0, T]^2; \mathbb{R}^d)$ . For  $\mathbf{r} = (r_1, r_2)$  and  $i, j \in \{1, \dots, d\}$ , we define

$$d^i X_{\mathbf{r}} := \partial_{12} X_{\mathbf{r}}^i dr_1 dr_2, \quad \text{and} \quad \hat{d}^{ij} X_{\mathbf{r}} := \partial_1 X_{\mathbf{r}}^i \partial_2 X_{\mathbf{r}}^j dr_1 dr_2, \quad (3.16)$$

where  $X_{\mathbf{r}}^i$  denotes the  $i$ th component of  $X_{\mathbf{r}} \in \mathbb{R}^d$ . Note that in future computations we will write  $d^i X_{\mathbf{r}}$  and  $dX_{\mathbf{r}}^i$  without distinction.

This subsection is closed by stating a basic change of variables formula in the plane. We employ Einstein's summation convention.

**Proposition 3.19.** *Let  $X$  be a field in  $C^2([0, T]^2; \mathbb{R}^d)$  and consider a  $C^2$ -function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ . Recall Notation 3.6 used for 2D-simplexes. Then for  $(s, \mathbf{t}) \in \Delta_{[0, T]}^2$  we have*

$$\square_{\mathbf{st}} f(X) = \int_{[s, \mathbf{t}]} \partial_i f(X_{\mathbf{r}}) d^i X_{\mathbf{r}} + \int_{[s, \mathbf{t}]} \partial_{ij} f(X_{\mathbf{r}}) \hat{d}^{ij} X_{\mathbf{r}}, \quad (3.17)$$



where we use convention (3.16) for the differentials  $d$  and  $\hat{d}$ , and where according to notation (3.7) we set

$$\int_{[s,t]} := \int_{s_1}^{t_1} \star \int_{s_2}^{t_2} = \int_{[s_1, t_1] \times [s_2, t_2]}. \quad (3.18)$$

**Remark 3.20.** Formula (3.17) above can also be stated using Faà di Bruno's formula. Namely, by considering  $s = (s_1, s_2)$  as fixed, the field  $t \mapsto Z_t = \square_{st} f(X)$  is such that

$$\partial_{12} Z_t = \partial_i f(X_t) \partial_{12} X_t^i + \partial_{ij} f(X_t) \partial_1 X_t^i \partial_2 X_t^j, \quad (3.19)$$

where we have used Notation 3.16 for the right hand side.

#### 4. Definition of the 2D signature

In this section, we will introduce a notion of 2D signature based on a linear equation, similar to Lemma 2.1. However, we shall see that this simple definition is not enough to grant basic algebraic properties. Therefore in the second part, we shall propose a new and more fruitful notion of 2D signature. The main algebraic propositions are postponed to later sections.

##### 4.1. A first notion of 2D signature: the id-signature

With the notation of Section 3.3 at hand, we now introduce our first notion of 2D signature in terms of a linear equation which is reminiscent of (2.1). Namely for a collection of constant matrices  $A^i \in \mathbb{R}^{e \times e}$ ,  $i = 1, \dots, d$ , an initial condition  $v \in \mathbb{R}^e$  and a smooth field  $X : [0, T]^2 \rightarrow \mathbb{R}^d$ , we consider the  $\mathbb{R}^e$ -valued solution  $Y$  to the linear equation

$$Y_t = v + \sum_{i=1}^d \int_{[s,t]} A^i Y_r d^i X_r, \quad \text{for } t \in [s, T], \quad (4.1)$$

where we recall from Notation 3.1 that  $s = (s_1, s_2)$  and  $T = (T, T)$ . Note the double integral (3.18) as well as the differential (3.16) present in (4.1). The next definition proposes a 2D signature over  $X$ , seen as a function defined recursively on words similarly to (1.2). Recall  $\Delta^2$  defined in Notation 3.6.

**Definition 4.1.** Let  $X : [0, T]^2 \rightarrow \mathbb{R}^d$  be a smooth field. The id-signature  $S^{\text{id}}(X)$  is a function in  $(C(\Delta^2))^{\mathcal{W}}$ , defined for  $(s, t) \in \Delta^2$  and every  $w = (i_1, \dots, i_n) \in \mathcal{W}$  by

$$\langle S_{st}^{\text{id}}(X), w \rangle := \int_{[s,t]} \langle S_{sr}^{\text{id}}(X), (i_1, \dots, i_{n-1}) \rangle d^{i_n} X_r. \quad (4.2)$$

We illustrate the above definition of the identity signature with the following example.

**Example 4.2.** Let the field  $X : [0, T]^2 \rightarrow \mathbb{R}^d$  be given as the Hadamard product of two paths  $x^1, x^2 \in \text{Lip}([0, T]; \mathbb{R}^d)$ , such that  $X$  is given by the vector

$$X_t = \begin{bmatrix} x_{t_1}^{1,1} & x_{t_2}^{2,1} \\ \vdots & \vdots \\ x_{t_1}^{1,d} & x_{t_2}^{2,d} \end{bmatrix}.$$

Then for any word  $w \in \mathcal{W}$ , it is readily checked that the signature inherits the multiplicative structure of the field  $X$ , in the sense that

$$\langle S_{st}^{\text{id}}(X), w \rangle = \langle S_{s_1 t_1}(x^1), w \rangle \langle S_{s_2 t_2}(x^2), w \rangle.$$

In [26] the authors show that the signature kernel  $K(x^1, x^2) := \langle S(x^1), S(x^2) \rangle_{\mathcal{T}(\mathbb{R}^d)}$ , as used in machine learning [27], solves a Goursat PDE, such that the signature kernel is given as the solution of the PDE

$$\frac{\partial^2}{\partial t_1 \partial t_2} u_t = u_t \langle \dot{x}_{t_1}^1, \dot{x}_{t_2}^2 \rangle, \quad u_{0,t_2} = u_{t_1,0} = 1. \quad (4.3)$$

A formal expansion of the solution to this equation as the signature kernel can then be written as

$$u_t = 1 + \sum_{w \in \mathcal{W}} \langle S_{s_1 t_1}(x^1), w \rangle \langle S_{s_2 t_2}(x^2), w \rangle = 1 + \sum_{w \in \mathcal{W}} \langle S_{st}^{\text{id}}(X), w \rangle. \quad (4.4)$$

Investigations of this type of equation have recently been expanded upon in the setting of Fubini's theorem in [28], and for a general analysis of the non-linear versions of this Goursat PDE see [29] in the case of Young fields  $X$ , and [30] for a regularization by noise perspective.

Analogously to Lemma 2.1, the following lemma shows how the 2D signature over  $X$  defined in (4.2) permits to express the  $\mathbb{R}^e$ -valued solution  $Y$  to (4.1).

**Lemma 4.3.** Given a smooth field  $X : [0, T]^2 \rightarrow \mathbb{R}^d$ , the rectangular increment of the  $\mathbb{R}^c$ -valued solution  $Y$  to (4.1) can be expanded as

$$\square_{st} Y = \sum_{w \in \mathcal{W}} A^{ow} v \langle S_{st}^{\text{id}}(X), w \rangle. \quad (4.5)$$

Recall that the matrix  $A^{ow}$  corresponds to the products of matrices in reversed order of the word  $w$ , introduced in (2.2).

**Proof.** According to (4.1), the initial conditions for  $Y$  on the axes are

$$Y_s = Y_{s_1, s_2} = Y_{t_1, s_2} = v$$

for all  $t = (t_1, t_2) \in [s, T]$ . Therefore, for all  $t \in [s, T]$  we also have

$$\square_{st} Y = Y_t - v.$$

Therefore, (4.1) becomes

$$\square_{st} Y = \sum_{i=1}^d \int_{[s, t]} A^i Y_r d^i X_r. \quad (4.6)$$

We now proceed as in the proof of Lemma 2.1. Namely on the right hand side of (4.6) we write  $Y_r = v + \square_{sr} Y$ . Using Einstein's summation convention we get

$$\square_{st} Y = A^{i1} v \int_{[s, t]} d^{i1} X_{r^1} + R_{st}^1, \quad \text{where} \quad R_{st}^1 = \int_{[s, t]} A^{i1} (\square_{sr^1} Y) d^{i1} X_{r^1}. \quad (4.7)$$

Iterating (4.6) and (4.7) as in the proof of Lemma 2.1 confirms a “formal expansion” (4.5). After establishing Proposition 4.25, this formal expansion transforms into a rigorous expansion.  $\square$

**Remark 4.4.** The signature  $S_{st}^{\text{id}}(X)$  can be written more explicitly as a tensor series in  $\mathcal{T}((\mathbb{R}^d))$  with multiple integrals as coefficients. Indeed, using (3.5) for simplexes, recursion (4.2) yields

$$\langle S_{st}^{\text{id}}(X), w \rangle = \int_{\Delta_{[s, t]}^n} d^{i_1} X_{r^1} \cdots d^{i_n} X_{r^n}, \quad (4.8)$$

where we recall (3.5) for  $\Delta_{[s, t]}^n$ .

The signature  $S_{st}^{\text{id}}(X)$  solves a linear differential equation which is reminiscent of (2.10). We state this infinite-dimensional variant of (4.1) below.

**Proposition 4.5.** Let  $X$  be a function in  $C^2([0, T]^2; \mathbb{R}^d)$ . Recall the definition of tensor series  $\mathcal{T}((\mathbb{R}^d))$  spelled out in Section 2. Then as a  $\mathcal{T}((\mathbb{R}^d))$ -valued function indexed by  $[0, T]^2$ , the signature  $S^{\text{id}}(X)$  satisfies the following equation for all  $(s, t) \in \Delta_{[0, T]}^2$ :

$$S_{st}^{\text{id}}(X) = \mathbf{1} + \int_{[s, t]} S_{sr}^{\text{id}}(X) \otimes dX_r. \quad (4.9)$$

**Remark 4.6.** One might ask about existence and uniqueness for equations like (4.9). And indeed, Proposition 4.5 already gives the existence of a solution. As far as uniqueness is concerned, notice again that we consider  $S_{st}^{\text{id}}(X)$  as an object taking values in  $\mathcal{T}((\mathbb{R}^d))$ . Therefore one possibility is to interpret equation (4.9) as follows: denote  $S_{st}^N := \pi_{\leq N} S_{st}^{\text{id}}(X)$  the projection of  $S_{st}^{\text{id}}(X)$  onto  $\oplus_{n=0}^N (\mathbb{R}^d)^{\otimes n}$ . Moreover, equip  $\oplus_{n=0}^N (\mathbb{R}^d)^{\otimes n}$  with the proper tensor product defined by (2.6). Then (4.9) should be interpreted as a family of finite-dimensional differential equations indexed by  $N$ , similarly to what is done in [4, Prop. 7.8] for one parameter:

$$S_{st}^N = \mathbf{1} + \int_{[s, t]} S_{sr}^N \otimes dX_r. \quad (4.10)$$

Now for a fixed  $N$ , finite-dimensional differential equations like (4.10) have been shown to have a unique solution e.g in [29, 31]. This also proves uniqueness of the solution for Eq. (4.9).

**Proof.** We denote by  $\pi_n(S_{st}^{\text{id}}(X))$  the projection of  $S_{st}^{\text{id}}(X)$  onto the set of level  $n$  signature components. More specifically, using Definition 1.1 and our notation in (2.9) we set

$$\pi_n(S_{st}^{\text{id}}(X)) = \sum_{w \in \mathcal{W}_n} \langle S_{st}^{\text{id}}(X), w \rangle e_w.$$

Then resorting to our explicit representation (4.8) we have

$$\pi_n(S_{st}^{\text{id}}(X)) = \int_{\Delta_{[s, t]}^n} \bigotimes_{i=1}^n dX_{r^i}, \quad (4.11)$$

where  $dX = (dX^1, \dots, dX^d)$ . Moreover, one can recast (4.11) as

$$\pi_n(S_{st}^{\text{id}}(X)) = \int_{[s, t]} \left( \int_{\Delta_{[s, r^n]}^{n-1}} \bigotimes_{i=1}^{n-1} dX_{r^i} \right) \otimes dX_{r^n}.$$

Resorting to the explicit representation in (4.8) again, we end up with

$$\pi_n(\mathbf{S}_{\text{st}}^{\text{id}}(X)) = \int_{[s,t]} \pi_{n-1}(\mathbf{S}_{\text{sr}^n}^{\text{id}}(X)) \otimes dX_{\mathbf{r}^n}.$$

Putting together all the projections finishes the proof.  $\square$

#### 4.2. Chen's relation: Horizontal and vertical

Recall that Chen's relation for curves indexed by  $[0, T]$  has been stated in Lemma 3.11. We now investigate proper generalizations to the 2D-setting. Similarly to other contributions in the field (see [14,23]) we will only get partial versions of this relation for the 2D signature. Our restrictions will be of two types:

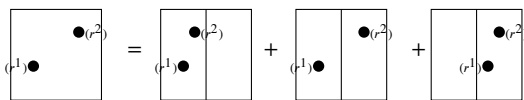
- (i) We will only treat partial Chen-type relations in directions 1 and 2 separately. One can combine the two, but will then obtain a rather complicated relation.
- (ii) Our results will be restricted to the signature  $\mathbf{S}^{\text{id}}(X)$  introduced in Section 4.1.

In subsequent sections, we shall give some hints about the way to overcome those difficulties, see in particular Remark 4.15. The following example illustrates the partial Chen's relation in the "horizontal direction".

**Example 4.7.** For any  $s_1 < u_1 < t_1$  and  $s_2 < t_2$  we have

$$\begin{aligned} \langle \mathbf{S}_{(s_1, s_2), (t_1, t_2)}^{\text{id}}(X), ab \rangle &= \int_{\substack{s_1 \leq r_1^1 \leq r_1^2 \leq t_1 \\ s_2 \leq r_2^1 \leq r_2^2 \leq t_2}} dX_{r_1^1, r_2^1}^a dX_{r_1^2, r_2^2}^b \\ &= \int_{\substack{s_1 \leq r_1^1 \leq r_1^2 \leq u_1 \\ s_2 \leq r_2^1 \leq r_2^2 \leq t_2}} dX_{r_1^1, r_2^1}^a dX_{r_1^2, r_2^2}^b + \int_{\substack{s_1 \leq r_1^1 \leq u_1 \leq r_1^2 \leq t_1 \\ s_2 \leq r_2^1 \leq r_2^2 \leq t_2}} dX_{r_1^1, r_2^1}^a dX_{r_1^2, r_2^2}^b \\ &\quad + \int_{\substack{u_1 \leq r_1^1 \leq r_1^2 \leq t_1 \\ s_2 \leq r_2^1 \leq r_2^2 \leq t_2}} dX_{r_1^1, r_2^1}^a dX_{r_1^2, r_2^2}^b \\ &= \langle \mathbf{S}_{(s_1, s_2), (u_1, t_2)}^{\text{id}}(X), ab \rangle + \int_{s_2 \leq q \leq t_2} \int_{s_1 \leq r_1^1 \leq u_1} dX_{r_1^1, q}^a \langle \mathbf{S}_{(u_1, q), (t_1, t_2)}^{\text{id}}(X), b \rangle \\ &\quad + \langle \mathbf{S}_{(u_1, s_2), (t_1, t_2)}^{\text{id}}(X), ab \rangle \\ &= \langle \mathbf{S}_{(s_1, s_2), (u_1, t_2)}^{\text{id}}(X), ab \rangle + \int_{s_2 \leq q \leq t_2} \frac{d}{dq} \langle \mathbf{S}_{(s_1, s_2), (u_1, q)}^{\text{id}}(X), b \rangle \langle \mathbf{S}_{(u_1, q), (t_1, t_2)}^{\text{id}}(X), a \rangle \\ &\quad + \langle \mathbf{S}_{(u_1, s_2), (t_1, t_2)}^{\text{id}}(X), ab \rangle \end{aligned}$$

Different from the one-parameter case, the signature does not precisely "split"; a form of "convolution" appears in the middle term. We visualize the respective points  $r^1, r^2$  in the plane, according to which side of the line  $\{u_1\} \times [s_2, t_2]$  they lie on:



In each of the three terms, the point  $r^1$  is to the south-west of  $r^2$ . Again, the middle term does not split, but is a "convolution" of the two sides.

Let  $f, g : \mathbb{A}^2 \rightarrow \mathbb{R}$  be two functions. Define for  $a_1 < b_1 < c_1, s_2 < t_2$

$$(f *_{b_1, H} g)((a_1, s_2), (c_1, t_2)) := \int_{s_2}^{t_2} f((a_1, s_2), (b_1, q)) g((b_1, q), (c_1, t_2)) dq.$$

Analogously, for  $s_1 < t_1, a_2 < b_2 < c_2$ ,

$$(f *_{b_2, V} g)((s_1, a_2), (t_1, c_2)) := \int_{s_1}^{t_1} f((s_1, a_2), (p, b_2)) g((p, b_2), (t_1, c_2)) dp.$$

For a differentiable map  $f : \mathbb{A}^2 \rightarrow \mathbb{R}$  define

$$(\mathbb{D}_{R,1} f)((s_1, s_2), (t_1, t_2)) := \frac{d}{dt_1} f((s_1, s_2), (t_1, t_2)).$$

The differential in "the right slot, in the first direction" and

$$(\mathbb{D}_{R,2} f)((s_1, s_2), (t_1, t_2)) := \frac{d}{dt_2} f((s_1, s_2), (t_1, t_2)).$$

**Proposition 4.8.** Let  $w = w_1 \cdots w_n \in \mathcal{W}$  be a word of length  $n$ . For  $a_1 < b_1 < c_1$ ,  $s_2 < t_2$  we have the following horizontal convolutional Chen's relation

$$\begin{aligned} \langle \mathbf{S}_{(a_1, s_2), (c_1, t_2)}^{\text{id}}(X), w \rangle &= \langle \mathbf{S}_{(a_1, s_2), (b_1, t_2)}^{\text{id}}(X), w \rangle + \langle \mathbf{S}_{(b_1, s_2), (c_1, t_2)}^{\text{id}}(X), w \rangle \\ &\quad + \sum_{k=1}^{n-1} \left( (\mathbb{D}_{R,2}(\mathbf{S}^{\text{id}}(X), w_1 \cdots w_k)) *_{b_1, H} \langle \mathbf{S}^{\text{id}}(X), w_{k+1} \cdots w_n \rangle \right) ((a, s_2), (c, t_2)). \end{aligned}$$

For  $s_1 < t_1$ ,  $a_2 < b_2 < c_2$ , we have the following vertical convolutional Chen's relation

$$\begin{aligned} \langle \mathbf{S}_{(s_1, a_2), (t_1, c_2)}^{\text{id}}(X), w \rangle &= \langle \mathbf{S}_{(s_1, a_2), (t_1, b_2)}^{\text{id}}(X), w \rangle + \langle \mathbf{S}_{(s_1, b_2), (t_1, c_2)}^{\text{id}}(X), w \rangle \\ &\quad + \sum_{k=1}^{n-1} \left( (\mathbb{D}_{R,1}(\mathbf{S}^{\text{id}}(X), w_1 \cdots w_k)) *_{b_2, V} \langle \mathbf{S}^{\text{id}}(X), w_{k+1} \cdots w_n \rangle \right) ((s_1, a_2), (t_1, c_2)). \end{aligned}$$

**Proof.** We prove the first statement.

$$\begin{aligned} \langle \mathbf{S}_{(a_1, s_2), (c_1, t_2)}^{\text{id}}(X), w \rangle &= \int_{\substack{a_1 \leq r_1^1 \leq \cdots \leq r_1^n \leq c_1 \\ s_2 \leq r_2^1 \leq \cdots \leq r_2^n \leq t_2}} dX_{r_1^1, r_2^1}^{w_1} \cdots dX_{r_1^n, r_2^n}^{w_n} \\ &= \sum_{k=0}^n \int_{\substack{a_1 \leq r_1^1 \leq \cdots \leq r_1^k \leq b_1 \leq r_1^{k+1} \leq \cdots \leq r_1^n \leq c_1 \\ s_2 \leq r_2^1 \leq \cdots \leq r_2^k \leq t_2}} dX_{r_1^1, r_2^1}^{w_1} \cdots dX_{r_1^n, r_2^n}^{w_n} \\ &= \sum_{k=0}^n \int_{s_2 \leq r_2^k \leq t_2} \int_{\substack{a_1 \leq r_1^1 \leq \cdots \leq r_1^k \leq b_1 \\ s_2 \leq r_2^1 \leq \cdots \leq r_2^{k-1} \leq r_2^k}} dX_{r_1^1, r_2^1}^{w_1} \cdots dX_{r_1^k, r_2^k}^{w_k} \\ &\quad \times \int_{\substack{b_1 \leq r_1^{k+1} \leq \cdots \leq r_1^n \leq c_1 \\ r_2^k \leq r_2^{k+1} \leq \cdots \leq r_2^n \leq t_2}} dX_{r_1^{k+1}, r_2^{k+1}}^{w_{k+1}} \cdots dX_{r_1^n, r_2^n}^{w_n}. \end{aligned}$$

Here we have colored the integration boundaries in gray, to distinguish them from the integration variables. Now

$$\begin{aligned} \int_{\substack{a_1 \leq r_1^1 \leq \cdots \leq r_1^k \leq b_1 \\ s_2 \leq r_2^1 \leq \cdots \leq r_2^{k-1} \leq r_2^k}} dX_{r_1^1, r_2^1}^{w_1} \cdots dX_{r_1^k, r_2^k}^{w_k} &= \frac{d}{dr_2^k} \langle \mathbf{S}_{(a_1, s_2), (b_1, r_2^k)}^{\text{id}}(X), w_1 \cdots w_k \rangle \\ \int_{\substack{b_1 \leq r_1^{k+1} \leq \cdots \leq r_1^n \leq c_1 \\ r_2^k \leq r_2^{k+1} \leq \cdots \leq r_2^n \leq t_2}} dX_{r_1^{k+1}, r_2^{k+1}}^{w_{k+1}} \cdots dX_{r_1^n, r_2^n}^{w_n} &= \langle \mathbf{S}_{(b_1, r_2^k), (c_1, t_2)}^{\text{id}}(X), w_{k+1} \cdots w_n \rangle, \end{aligned}$$

and the claim follows.  $\square$

Below we start by introducing some useful notation to state our main results.

#### 4.3. Challenges with the id-signature

The signature proposed in (4.2) is rather natural and interestingly simple. However, it lacks some of the basic properties compared to the signature of a path. Let us briefly highlight two of those shortcomings:

(i) *Lack of shuffle property.* A common requirement for a proper signature is stability under multiplication, i.e., iterated integrals (4.8) should form an algebra. However,  $\mathbf{S}^{\text{id}}(X)$  does not fulfill this property; the  $d = 1$  case, i.e., considering a real-valued field  $X \in \mathcal{C}^2([0, T]^2; \mathbb{R})$  suffices to see this. Indeed, due to (4.8), the increment  $\square X = \int dX$  is trivially seen as a component of  $\mathbf{S}^{\text{id}}(X)$ . Now define the product

$$\Pi_{\text{st}} = \left( \int_{[\text{s}, \text{t}]} dX_{\text{r}} \right)^2 = \int_{[\text{s}, \text{t}]} dX_{\text{r}} \int_{[\text{s}, \text{t}]} dX_{\text{v}},$$

where we recall that  $dX$  is given by (3.16). Then some elementary computations involving integration in  $[\text{s}, \text{t}] \times [\text{s}, \text{t}]$  reveal that  $\Pi_{\text{st}} = 2\Pi_{\text{st}}^1 + 2\Pi_{\text{st}}^2$ , with

$$\Pi_{\text{st}}^1 = \int_{\Delta_{[\text{s}, \text{t}]}} dX_{r_1^1, r_2^1} dX_{r_1^2, r_2^2} \quad \text{and} \quad \Pi_{\text{st}}^2 = \int_{\Delta_{[\text{s}, \text{t}]}} dX_{r_1^1, r_2^2} dX_{r_1^2, r_2^1}. \quad (4.12)$$

Now referring to (4.8) again, the term  $\Pi_{\text{st}}^1$  in (4.12) is easily identified with  $\langle \mathbf{S}_{\text{st}}^{\text{id}}(X), (1, 1) \rangle$ . However,  $\Pi_{\text{st}}^2$  is not part of the signature, due to the permutation of  $r_1^1$  and  $r_2^2$  in the double integrals. This permutation phenomenon will feature prominently in our future considerations.

(ii) *Lack of change of variables formula for exponential functions.* A signature should accommodate simple expansions in change of variables formulae. Here again, it is easy to find examples for which  $\mathbf{S}^{\text{id}}(X)$  fails to be appropriate for this elementary task. That is, consider the same generic real-valued field  $X \in \mathcal{C}^2([0, T]^2; \mathbb{R})$  as in item (i) above. Next, for  $\text{t} \in [\text{s}, \text{T}]$  set  $Z_{\text{t}} = \exp(\square_{\text{st}} X)$ . Hence,

applying a small variant of the Faà Di Bruno type formula stated in [Remark 3.20](#) (differentiating  $\mathbf{t} \mapsto f(\square_{\mathbf{st}} X) = Z_{\mathbf{t}}$  instead of  $\mathbf{t} \mapsto f(X_{\mathbf{t}})$ ), we have that

$$\partial_{12} Z_{\mathbf{t}} = Z_{\mathbf{t}} \partial_{12} X + Z_{\mathbf{t}} \partial_1 (X_{\mathbf{t}} - X_{t_1, s_2}) \partial_2 (X_{\mathbf{t}} - X_{s_1, t_2}).$$

Therefore, a first-order approximation for the rectangular increment of  $Z$  is

$$\square_{\mathbf{st}} Z \approx Z_{\mathbf{s}} \square_{\mathbf{st}} X + Z_{\mathbf{s}} R_{\mathbf{st}}^X, \quad (4.13)$$

where the term  $R_{\mathbf{st}}^X$  is defined by

$$R_{\mathbf{st}}^X := \int_{[\mathbf{s}, \mathbf{t}]} \partial_1 (X_{\mathbf{r}} - X_{r_1, s_2}) \partial_2 (X_{\mathbf{r}} - X_{s_1, r_2}) d\mathbf{r}.$$

Now, as mentioned in [\(i\)](#), the increment  $\square_{\mathbf{st}} X$  is part of  $\mathbf{S}^{\text{id}}(X)$ , i.e.,  $\square_{\mathbf{st}} X = \langle \mathbf{S}_{\mathbf{st}}^{\text{id}}(X), (1) \rangle$ . However, the last term  $R^X$  in [\(4.13\)](#) is not an element of the signature  $\mathbf{S}^{\text{id}}(X)$ . The above first-order approximation has thus to be expressed in terms of a larger object than the id-signature, which we will call the full signature.

#### 4.4. Definition of the full 2D signature

In [Section 4.1](#), we have seen that the id-signature  $\mathbf{S}^{\text{id}}(X)$  does not possess some of the desirable properties known to hold for the classical path signatures. As addressed in [relation \(4.12\)](#), the problem seems to arise from the fact that permutations of simplexes appear. To address this issue, we will now propose another notion of 2D signature, taking such permutations into account.

Before proceeding to the definition of the 2D signature, we will introduce the new set of word-permutation tuples. Recall that the set of words  $\mathcal{W}_n$  of length  $n \geq 0$  is defined in [Definition 1.1](#).

**Definition 4.9.** (i) 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  $n \geq 1$

$$\hat{\mathcal{W}}_n = \{(w, v) \mid w \in \mathcal{W}_n, v \in \Sigma_{\{1, \dots, n\}}\}. \quad (4.14)$$

(ii) If  $\rho \in \Sigma_{\{1, \dots, n\}}$  we set

$$w_{\rho} := (i_{\rho(1)}, \dots, i_{\rho(n)}). \quad (4.15)$$

Note that for the concatenation of words,  $ww'$ , we set  $[ww']_{\tau}$  (which should not be confused with the notation  $[ww']_j$  referring to the  $j$ th letter in  $ww'$ ).

Recall that  $\Delta^2$  is given in [Notation 3.6](#). We are now ready to define the full 2D signature.

**Definition 4.10.** Let  $X$  be a function in  $C^2([0, T]^2; \mathbb{R}^d)$ . The (full) 2D signature  $\mathbf{S}(X)$  is a function in  $(C(\Delta^2))^{\hat{\mathcal{W}}}$ , defined for an extended word  $(w, v) \in \hat{\mathcal{W}}_n$  by setting

$$\langle \mathbf{S}_{\mathbf{st}}(X), (w, v) \rangle := \int_{\Delta_{[\mathbf{s}, \mathbf{t}]}^n} \prod_{i=1}^n dX_{r_1^{i'}, r_2^{i'}}^{w_i v_i}. \quad (4.16)$$

**Remark 4.11.** According to [\(3.3\)](#) and [\(3.8\)](#), one can recast [\(4.16\)](#) as

$$\langle \mathbf{S}_{\mathbf{st}}(X), (w, v) \rangle = \int_{\Delta_{s_1, t_1}^n} \star' \int_{\Delta_{s_2, t_2}^{n-1}} \prod_{i=1}^n dX_{r_1^{i'}, r_2^{i'}}^{w_i v_i}. \quad (4.17)$$

The representation in [\(4.17\)](#) will be useful to prove some of our algebraic relations.

With the definition of the full 2D signature at hand, let us illustrate its behavior concerning the 2D signature  $\mathbf{S}^{\text{id}}(X)$  proposed in [Definition 4.1](#) and its relation to 2D iterated sums as proposed in [\[16\]](#).

**Example 4.12.** Let us illustrate the 2D signature in light of the multiplicative field discussed in [Example 4.2](#). Namely we consider a  $\mathbb{R}$ -valued field  $X_{\mathbf{t}} = x_{t_1}^1 x_{t_2}^2$  for two Lipschitz paths  $x^1$  and  $x^2$ . Choosing an extended word  $(w, v) \in \hat{\mathcal{W}}_n$  we see that

$$\langle \mathbf{S}_{\mathbf{st}}(X), (w, v) \rangle = \langle S_{s_1 t_1}(x^1), w \rangle \langle S_{s_2 t_2}(x^2), w_{v^{-1}} \rangle, \quad (4.18)$$

where  $w_{v^{-1}} = (w_{v^{-1}(1)}, \dots, w_{v^{-1}(n)})$ . Indeed, to see this observe that

$$\langle \mathbf{S}_{\mathbf{st}}(X), (w, v) \rangle = \langle S_{s_1 t_1}(x^1), w \rangle \int_{\Delta_{s_2 t_2}^n} \prod_{i=1}^n dx_{r_2^{i'}}^{2, w_i v_i}, \quad (4.19)$$

and by Fubini's theorem it follows that

$$\int_{\Delta_{s_2 t_2}^n} \prod_{i=1}^n dx_{r_2^{i'}}^{2, w_i v_i} = \int_{\Delta_{s_2 t_2}^n} \prod_{i=1}^n dx_{r_2^{i'}}^{2, w_{v^{-1}(i)}} = \langle S_{s_2 t_2}(x^2), w_{v^{-1}} \rangle.$$

We see that when  $v = \text{id}$ , then we recover the results of [Example 4.2](#).

**Example 4.13.** Consider  $i_1, i_2, i_3 \in \{1, \dots, d\}$  and the permutation  $[1\ 3\ 2] \in \Sigma_{\{1,2,3\}}$ .

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

Employing a matrix notation akin to the one used in [16]<sup>2</sup> this corresponds to the matrix

$$\left\langle S_{\text{st}}(X), \begin{bmatrix} 0 & i_2 & 0 \\ 0 & 0 & i_3 \\ i_1 & 0 & 0 \end{bmatrix} \right\rangle.$$

**Remark 4.14.** The two issues (i)–(ii) raised after Proposition 4.5 are easily resolved by using the full 2D signature  $S(X)$ . Specifically, the term  $\Pi_{\text{st}}^2$  in (4.12) can be written as

$$\Pi_{\text{st}}^2 = \int_{\Delta_{[s,t]}^2} dX_{r_1^1, r_2^1} dX_{r_1^2, r_2^2} = \langle S_{\text{st}}(X), (w, v) \rangle, \quad (4.20)$$

where  $w = (1, 1)$  and  $v \in \Sigma_{\{1,2\}}$  is defined by  $v = [2\ 1]$ . As for the term  $R_{\text{st}}^X$  in (4.13), observe that by applying the fundamental theorem of calculus to the increments in

$$\partial_1 (X_r - X_{r_1, s_2}) \partial_2 (X_r - X_{s_1, r_2}),$$

we also get that

$$R_{\text{st}}^X = \int_{\Delta_{[s,t]}^2} dX_{r_1^1, r_2^1} dX_{r_1^2, r_2^2} = \langle S_{\text{st}}(X), ((1, 1), [2\ 1]) \rangle = \Pi_{\text{st}}^2.$$

With this computation we also observe that even for fields with one ambient dimension, the full signature is not trivial (i.e. it cannot be written in terms of monomials), in contrast to what is the case for the path signature.

**Remark 4.15.** It is a non-trivial task to generalize Proposition 4.8 to the 2D signature  $S(X)$  from Definition 4.10. Indeed, consider Example 4.13, visualized as constellations of three points in the plane (as in Example 4.7):

Now, the second term on the right does not split, even when allowing for a convolution of the type in Proposition 4.8. We conjecture that an operadic structure can handle this issue, but this is beyond the scope of the current article.

We conclude this section by presenting a linear integral equation that is satisfied by the truncation of the 2D signature  $S(X)$  to words of length  $n \leq 3$ .

**Proposition 4.16.** Let  $X : [0, T]^2 \rightarrow \mathbb{R}^d$  be a  $C^2$  field. Consider the 2D signature  $S(X)$  given in Definition 4.10, seen as an element of  $\mathcal{C}(\Delta^2)^{\hat{W}}$ . Then for  $n \leq 3$  and  $(w, v) \in \hat{W}_n$ , with  $w = (i_1, \dots, i_n)$ , the coordinate  $\langle S(X), (w, v) \rangle$  satisfies an equation of the form

$$\langle S_{\text{st}}(X), (w, v) \rangle = \mathbf{1}_{\{w=e\}} + \mathbf{1}_{\{w \neq e\}} \int_{[s,t]} \mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1 \dots i_{n-1}, v)} dX_{\mathbf{v}}^{i_n}, \quad (4.21)$$

where the function  $\mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1 \dots i_{n-1}, v)}$  is given by the following expressions for various permutations  $v$ :

$$\begin{aligned} \mathcal{L}_{\mathbf{v}}^{\text{st}; (e, [1])} &= 1 \\ \mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1, [1\ 2])} &= \langle S_{s_1, s_2; v_1, v_2}(X), ((i_1), \text{id}) \rangle \\ \mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1, [2\ 1])} &= \langle S_{s_1, v_2; v_1, i_2}(X), ((i_1), \text{id}) \rangle \\ \mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1 i_2, [1\ 2\ 3])} &= \langle S_{s_1, s_2; v_1, v_2}(X), ((i_1, i_2), \text{id}) \rangle \\ \mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1 i_2, [1\ 3\ 2])} &= \langle S_{s_1, s_2; v_1, i_2}(X), ((i_1, i_2), \text{id}) \rangle - \langle S_{s_1, s_2; v_1, v_2}(X), ((i_1, i_2), \text{id}) \rangle - \langle S_{s_1, v_2; v_1, i_2}(X), ((i_1, i_2), \text{id}) \rangle \\ \mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1 i_2, [2\ 1\ 3])} &= \langle S_{s_1, s_2; v_1, v_2}(X), ((i_1, i_2), [2\ 1]) \rangle \\ \mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1 i_2, [2\ 3\ 1])} &= \langle S_{s_1, v_2; v_1, i_2}(X), ((i_1, i_2), [1\ 2]) \rangle \\ \mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1 i_2, [3\ 2\ 1])} &= \langle S_{s_1, v_2; v_1, i_2}(X), ((i_1, i_2), [2\ 1]) \rangle, \end{aligned}$$

and

$$\mathcal{L}_{\mathbf{v}}^{\text{st}; (i_1 i_2, [3\ 1\ 2])} = \langle S_{s_1, s_2; v_1, i_2}(X), ((i_1, i_2), [2\ 1]) \rangle -$$

<sup>2</sup> Contrary to what has been done in [16], we orient the plane, and correspondingly the matrices, from bottom-left to top-right.

$$\langle \mathbf{S}_{s_1, s_2; v_1, v_2}(X), ((i_1, i_2), [2\ 1]) \rangle - \langle \mathbf{S}_{s_1, v_2; v_1, i_2}(X), ((i_1, i_2), [2\ 1]) \rangle.$$

**Proof.** The cases  $n = 1, 2$  are easily handled, similarly to Proposition 4.5. The same kind of argument also holds for  $n = 3$  when  $v(3) \in \{1, 3\}$ . Therefore, in the remainder of the proof, we will focus on the case  $n = 3$  and  $v(3) = 2$ . Specifically, let us assume  $w = (i_1, i_2, i_3)$  for  $i_1, i_2, i_3 \in \{1, \dots, d\}$  and  $v = [1\ 3\ 2]$ . The other case,  $v = [3\ 1\ 2]$ , can be treated similarly. The value of the 2D signature is defined for  $(s, t) \in \Delta_{[s, t]}^2$ :

$$\langle \mathbf{S}_{st}(X), (w, v) \rangle = \int_{\Delta_{[s, t]}^3} dX_{r_1^1, r_2^1}^{i_1} dX_{r_1^2, r_2^2}^{i_2} dX_{r_1^3, r_2^3}^{i_3}.$$

Recall that  $\Delta_{[s, t]}^3$  was introduced in (3.5). To set up an induction, let us write

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

Otherwise stated, one has

$$\langle \mathbf{S}_{st}(X), (w, v) \rangle = \int_{\substack{s_1 \leq r_1^3 \leq t_1 \\ s_2 \leq r_2^3 \leq t_2}} Q_{r_1^3, r_2^3} dX_{r_1^3, r_2^3}^{i_3}, \quad (4.22)$$

where for  $\mathbf{v} = (v_1, v_2) \in [s, t]$  we have set

$$Q_{v_1, v_2} := \int_{(r_1^1, r_2^1, r_1^2, r_2^2) \in D_{v_1, v_2}} dX_{r_1^1, r_2^1}^{i_1} dX_{r_1^2, r_2^2}^{i_2}, \quad (4.23)$$

and where the domain  $D_{v_1, v_2} \subset [0, T]^4$  can be decomposed as  $D_{v_1, v_2} = D_{v_1}^1 \times D_{v_2}^2$  with

$$D_{v_1}^1 = \{(p_1, q_1) \in \Delta_{s_1, t_1}^2; s_1 \leq p_1 \leq q_1 \leq v_1\}$$

$$D_{v_2}^2 = \{(p_2, q_2) \in \Delta_{s_2, t_2}^2; s_2 \leq p_2 \leq q_2 \leq v_2, \text{ and } v_2 \leq q_2 \leq t_2\}.$$

Notice that we have several dummy integration variables above, so we will simplify relations (4.22)–(4.23) as

$$\langle \mathbf{S}_{st}(X), (w, v) \rangle = \int_{\substack{s_1 \leq v_1 \leq t_1 \\ s_2 \leq v_2 \leq t_2}} Q_{v_1, v_2} dX_{v_1, v_2}^{i_3} = \int_{[s, t]} Q_{\mathbf{v}} dX_{\mathbf{v}}^{i_3}, \quad (4.24)$$

with

$$Q_{\mathbf{v}} = \int_{(\mathbf{r}^1, \mathbf{r}^2) \in D_{\mathbf{v}}} dX_{\mathbf{r}^1}^{i_1} dX_{\mathbf{r}^2}^{i_2}.$$

Now notice that due to relation (3.11) we have

$$\int_{D_{v_2}^2} = \int_{\Delta_{s_2, v_2}^1} \star \int_{\Delta_{v_2, t_2}^1} = \int_{\Delta_{s_2, t_2}^2} - \int_{\Delta_{s_2, v_2}^2} - \int_{\Delta_{v_2, t_2}^2}.$$

This information is then inserted into the definition of  $Q$  to discover that

$$Q_{\mathbf{v}} = \int_{\Delta_{[s_1, v_1] \times [s_2, v_2]}^2} dX_{\mathbf{r}^1}^{i_1} dX_{\mathbf{r}^2}^{i_2} - \int_{\Delta_{[s_1, v_1] \times [s_2, v_2]}^2} dX_{\mathbf{r}^1}^{i_1} dX_{\mathbf{r}^2}^{i_2} - \int_{\Delta_{[s_1, v_1] \times [v_2, t_2]}^2} dX_{\mathbf{r}^1}^{i_1} dX_{\mathbf{r}^2}^{i_2} \quad (4.25)$$

$$= \langle \mathbf{S}_{s_1, s_2; v_1, v_2}(X), ((i_1, i_2), \text{id}) \rangle - \langle \mathbf{S}_{s_1, s_2; v_1, v_2}(X), ((i_1, i_2), \text{id}) \rangle - \langle \mathbf{S}_{s_1, s_2; v_1, t_2}(X), ((i_1, i_2), \text{id}) \rangle.$$

Plugging (4.25) into (4.24), we thus end up with

$$\langle \mathbf{S}_{st}(X), (w, v) \rangle = \int_{[s, t]} \langle \mathbf{S}_{s_1, s_2; v_1, t_2}(X), ((i_1, i_2), \text{id}) \rangle dX_{\mathbf{v}}^{i_3}$$

$$- \int_{[s, t]} \langle \mathbf{S}_{s_1, s_2; v_1, v_2}(X), ((i_1, i_2), \text{id}) \rangle dX_{\mathbf{v}}^{i_3} - \int_{[s, t]} \langle \mathbf{S}_{s_1, v_2; v_1, t_2}(X), ((i_1, i_2), \text{id}) \rangle dX_{\mathbf{v}}^{i_3}.$$

Summarizing this by setting

$$L_{\mathbf{v}}^{\text{st}} = \langle \mathbf{S}_{s_1, s_2; v_1, t_2}(X), ((i_1, i_2), \text{id}) \rangle - \langle \mathbf{S}_{s_1, s_2; v_1, v_2}(X), ((i_1, i_2), \text{id}) \rangle - \langle \mathbf{S}_{s_1, v_2; v_1, t_2}(X), ((i_1, i_2), \text{id}) \rangle,$$

finishes the proof.  $\square$

**Remark 4.17.** Proposition 4.16 only gives an integral equation for the signature up to level  $n = 3$ . For  $n \geq 4$ , deriving such an equation would require further insight into the algebraic structure which is beyond the scope of the current paper.

**Remark 4.18.** Even though Eq. (4.21) gives a recursive relation for computing signature terms, a simple choice for low-level terms would be to discretize the iterated integrals directly. That is, for the double integral  $\langle S_{st}(X), ((i_1, i_2), \text{id}) \rangle$  we have that

$$\int_{[s,t]} \int_{[s,r^2]} dX_{r^1}^{i_1} dX_{r^2}^{i_2} \approx \sum_{[u_1, v_1] \in \mathcal{P}[s_1, t_1]} \sum_{[u_2, v_2] \in \mathcal{P}[s_2, t_2]} \square_{(s_1, s_2; u_1, u_2)} X^{i_1} \square_{(u_1, u_2; v_1, v_2)} X^{i_2}.$$

here  $\mathcal{P}[s_i, t_i]$  is a partition of the interval  $[s_i, t_i]$ , for  $i = 1, 2$ . We refer to [22] for further details on practical computational implementation. In addition, note that there might be algebraic structures (as illustrated above) which could reduce computational complexity further. We have not delved deeper into this problem for the sake of conciseness of the current paper. We plan to report on such issues in subsequent articles.

#### 4.5. Invariance properties

Among the central and elementary properties of the path signature, let us mention the following ones:

- (i) *Translation invariance:* For  $a \in \mathbb{R}^d$  and a path  $x : [0, 1] \rightarrow \mathbb{R}^d$  one has  $S(x + a) = S(x)$ .
- (ii) *Re-parameterization invariance:* For a non-decreasing function  $\phi : [0, 1] \rightarrow [0, 1]$  with  $\phi(0) = 0$  and  $\phi(1) = 1$  one has  $S(X \circ \phi)_{0,1} = S(X)_{0,1}$ .

Heuristically, translation invariance tells us that the signature “does not see” the absolute level at which the path is operating, but only sees relative differences. Reparametrization invariance tells us that the signature does not care about different speeds at which we might run through our path. These two features have been crucial in assessing the usefulness of the signature method in machine learning and other areas, see e.g. [6,7].

We will now discuss these properties in the setting of the 2D signature. They rely on the choice of differential structure to work with for the concept of a 2D signature. Consequently, we will explore and elaborate on them in relation to the differential  $dX$  as introduced in Notation 3.18.

##### 4.5.1. Translation invariance

When working with the 2D signature over a  $\mathbb{R}^d$ -valued field  $X : [0, T]^2 \rightarrow \mathbb{R}^d$ , there are three different ways to look at translation; one is to shift the field with a constant vector  $a \in \mathbb{R}^d$ , another is to shift the field by 1D paths, only depending on either the first variable or the second. More precisely, let  $x^i : [0, T] \rightarrow \mathbb{R}^d$  for  $i = 1, 2$  be two paths, and set  $\tau X_{t_1, t_2} := X_{t_1, t_2} + x_{t_1}^1 + x_{t_2}^2 + a$ . Then, given that we work with the differential structure  $dX$  defined in Notation 3.18 for the 2D signature  $S(X)$ , we observe that for every  $i \in \{1, \dots, d\}$  we have

$$d^i(\tau X) = d^i X,$$

and so it follows that

$$S(\tau X) = S(X).$$

**Remark 4.19.** This is a rather striking property of the 2D signature and sheds light on the type of properties in a field  $X$  captured by the signature. It is complementary to the classical path signature, and for feature extraction purposes, depending on the task at hand, it might be beneficial to also include the path signature of the paths  $t_1 \mapsto X(t_1, u)$  and  $t_2 \mapsto X(u, t_2)$  for some choices of  $u$ , as the 2D signature will not capture information provided from this. In contrast, the 2D signature will provide us information about how much “small changes in one direction are affected by small changes in the other direction”, and then the iterated integrals provide us with a systematic (partially ordered) comparison of this information over a field  $X$ .

##### 4.5.2. Invariance to stretching

While there is no direct analog of the re-parameterization invariance of the path signature in the 2D setting, the closest thing would be to consider invariance to stretching.

**Definition 4.20.** We say that a continuously differentiable function  $\phi : [0, T]^2 \rightarrow [0, T]^2$  is a *stretching* if

$$\phi(t_1, t_2) = (\phi_1(t_1), \phi_2(t_2)),$$

where  $\phi_i : [0, T] \rightarrow [0, T]$  are monotone functions with the property that  $\phi_i(0) = 0$  and  $\phi_i(T) = T$  for  $i = 1, 2$ .

For a field  $X : [0, T]^2 \rightarrow \mathbb{R}^d$  and a stretching  $\phi$  we define now the stretched field  $X^\phi(t_1, t_2) = X(\phi(t_1, t_2))$ . By the chain rule, for all  $i \in \{1, \dots, d\}$  we have that

$$d^i(X^\phi) = (d^i X)(\phi) \phi'_1 \phi'_2.$$

Integrating over a rectangle  $[s, t]$  and applying change of variable, it is readily checked that

$$\int_s^t d^i X^\phi = \int_{\phi(s)}^{\phi(t)} d^i X. \quad (4.26)$$



In particular, recalling that  $\mathbf{T}$  stands for the tuple  $(T, T)$  we have  $\int_0^T d^i X^\phi = \int_0^T d^i X$ . Iterating (4.26) over the word  $w = (i_1, \dots, i_n)$  we further deduce that

$$\mathbf{S}_{\text{st}}(X^\phi) = \mathbf{S}_{\phi(\mathbf{s}), \phi(\mathbf{t})}(X),$$

and conclude that the 2D signature over  $[0, T]^2$  is invariant to stretching.

#### 4.5.3. Equivariance to rotation

We will in this section explore how rotations of the field  $X$  affect the 2D signature. To this end, we begin with a definition of rotating a field.

**Definition 4.21.** A counterclockwise coordinate rotation by angle  $\theta$  around the origin is defined by

$$\phi_\theta(t_1, t_2) = (t_1 \cos \theta - t_2 \sin \theta, t_1 \sin \theta + t_2 \cos \theta).$$

Given a continuous field  $X : [-1, 1]^2 \rightarrow \mathbb{R}^d$ , its clockwise rotation<sup>3</sup> by  $\theta$  is defined by

$$X^\theta(t_1, t_2) := X(\phi_\theta(t_1, t_2)). \quad (4.27)$$

**Remark 4.22.** We will here consider fields defined on the standard square centered at zero,  $[-1, 1]^2$ , since the rotation provided above is around  $(0, 0)$ . This guarantees that when  $\theta = \pi/2$  the field is still well defined over  $[-1, 1]^2$ . In principle, one could investigate rotations of any degree, and it would be highly interesting to see how this affects the signature. However, going outside the 90 degree cases will require more in depth analysis of the image under rotation. It also yields more complicated structures, as illustrated by the rotation operator in Definition 4.21. For conciseness we have not investigated this further here, but hope to address this in a more applied research project.

In the next proposition, we will show a certain equivariance of the id-signature with respect to rotations by multiples of  $\pi/2$  of an image over  $[-1, 1]^2$ .

**Proposition 4.23.** Let  $X$  be a field in  $C^2([-1, 1]^2; \mathbb{R}^d)$  and let  $(w, v)$  be a word in  $\hat{\mathcal{W}}_n$  for  $v \in \Sigma_{\{1, \dots, n\}}$ . Let  $\rho_n \in \Sigma_{\{1, \dots, n\}}$  denote the reversal permutation, i.e. where  $\rho_n(i) = n - i + 1$  for  $i = 1, \dots, n$ . Then the following three identities hold:

(i) The rotation by  $\theta = \pi/2$ , satisfies (recall the notation for the right-action on words (4.15))

$$\langle \mathbf{S}_{-\mathbf{II}}(X^{\pi/2}), (w, v) \rangle = (-1)^n \langle \mathbf{S}_{-\mathbf{II}}(X), (w_{v^{-1} \circ \rho_n}, v^{-1} \circ \rho_n) \rangle.$$

(ii) The rotation by  $\theta = \pi$ , satisfies

$$\langle \mathbf{S}_{-\mathbf{II}}(X^\pi), (w, v) \rangle = \langle \mathbf{S}_{-\mathbf{II}}(X), (w_{\rho_n^{-1} \circ v \circ \rho_n}, \rho_n^{-1} \circ v \circ \rho_n) \rangle.$$

(iii) The rotation by  $\theta = 3\pi/2$ , satisfies

$$\langle \mathbf{S}_{-\mathbf{II}}(X^{3\pi/2}), (w, v) \rangle = (-1)^n \langle \mathbf{S}_{-\mathbf{II}}(X), (w_{\rho_n^{-1} \circ v^{-1}}, \rho_n^{-1} \circ v^{-1}) \rangle.$$

Here,  $X^\theta$  is defined in (4.27).

**Proof.** It is readily checked from Definition 4.21 that a 90 degrees rotation is given by the coordinate change  $(t_1, t_2) \mapsto X^{\pi/2}(t_1, t_2) = X(-t_2, t_1)$ . By elementary calculus rules, in particular repeated change of variables, and using the fact that for  $a < b$  we have  $\int_a^b g = -\int_b^a g$  for some integrable function  $g$ , one can check that the following relation holds for a sequence of integrable functions  $\{g^i\}_{i=1}^n$

$$\int_{\Delta_{s,t}^n} \prod_{i=1}^n g^i(-r^i) dr^i = \int_{\Delta_{-t,-s}^n} \prod_{i=1}^n g^i(u^{n-i+1}) du^i \quad (4.28)$$

<sup>3</sup> Note that the counterclockwise rotation of the parameters induces a clockwise rotation of the field.

Furthermore, we use that  $d(X^{\pi/2})_{(r_1, r_2)}^{w_k} = -dX_{-r_2, r_1}^{w_k}$ , and find the following

$$\begin{aligned}
 \langle S_{\text{st}}(X^{\pi/2}), (w, v) \rangle &= \int_{\Delta_{s_1, t_1}^n} \int_{\Delta_{s_2, t_2}^n} v^{-1} \prod_{k=1}^n d(X^{\pi/2})_{(r_1^k, r_2^k)}^{w_k} \\
 &= (-1)^n \int_{\Delta_{s_1, t_1}^n} \int_{\Delta_{s_2, t_2}^n} \prod_{k=1}^n dX_{(-r_2^k, r_1^k)}^{w_k} \\
 &= (-1)^n \int_{\Delta_{s_1, t_1}^n} \int_{\Delta_{s_2, t_2}^n} \prod_{k=1}^n dX_{(-r_2^k, r_1^k)}^{w_{v^{-1}(k)}} \\
 &= (-1)^n \int_{\Delta_{s_1, t_1}^n} \int_{\Delta_{-t_2, -s_2}^n} \prod_{k=1}^n dX_{(u_2^{n-k+1}, r_1^{v^{-1}(k)})}^{w_{v^{-1}(k)}} \\
 &= (-1)^n \int_{\Delta_{s_1, t_1}^n} \int_{\Delta_{-t_2, -s_2}^n} \prod_{\ell=1}^n dX_{(u_2^\ell, r_1^{v^{-1}(n-\ell+1)})}^{w_{v^{-1}(n-\ell+1)}} \\
 &= (-1)^n \langle S_{(-t_2, s_1), (-s_2, t_1)}(X), (w_{v^{-1} \circ \rho_n}, v^{-1} \circ \rho_n) \rangle.
 \end{aligned}$$

Note that, the number  $-s_2$  is now inserted in the spot where  $t_1$  usually is, and similarly for the rest of the variables, reflecting the rotation of the domain. And in particular, choosing  $[s, t] = [-1, 1]^2$  we see that

$$\langle S_{-11}(X^{\pi/2}), (w, v) \rangle = (-1)^n \langle S_{-11}(X), (w_{v^{-1} \circ \rho_n}, v^{-1} \circ \rho_n) \rangle.$$

This proves item (i).

Now, the items (ii) and (iii) follow from iterated application of (i) and noting that the following two relations hold

$$(v^{-1} \circ \rho_n)^{-1} \circ \rho_n = \rho_n^{-1} \circ v \circ \rho_n, \quad (\rho_n^{-1} \circ v \circ \rho_n)^{-1} \circ \rho_n = \rho_n^{-1} \circ v^{-1}.$$

where we have used that  $\rho_n = \rho_n^{-1}$  for the second relation.  $\square$

#### 4.5.4. On homotopy invariance

As we see in Section 4.7, there are some parallels to the iterated integrals defined in [13]. The latter are shown to be homotopy invariant, which is *not* the case for the integrals defined here.

**Example 4.24.** Consider the map

$$H_t(r_1, r_2) = (t \sin(\pi r_2) r_1, t \sin(\pi r_1) r_2).$$

It is a homotopy between the zero map and the map

$$X_{r_1, r_2} := (\sin(\pi r_2) r_1, \sin(\pi r_1) r_2),$$

which preserves the boundary on  $[0, 1]^2$ . Now, the integrals of the zero map are all zero. If the signature were homotopy invariant, we would in particular have

$$\langle S_{01}(X), ((1, 2), \text{id}) \rangle = 0. \quad (4.29)$$

In order to compute the left hand side of (4.29), observe from our Notation 3.18 that

$$d^1 X_{r_1, r_2} = \pi \cos(\pi r_2) dr_1 dr_2$$

$$d^2 X_{r_1, r_2} = \pi \cos(\pi r_1) dr_1 dr_2.$$

Therefore, relation (4.8) and some elementary computations show that

$$\begin{aligned}
 \langle S_{01}(X), ((1, 2), \text{id}) \rangle &= \pi \int_0^1 ds_1 \int_0^1 ds_2 \sin(\pi s_2) s_1 \cos(\pi s_1) \\
 &= \pi \left( \int_0^1 ds_1 s_1 \cos(\pi s_1) \right) \left( \int_0^1 ds_2 \sin(\pi s_2) \right) \\
 &= \pi \left( -\frac{2}{\pi^2} \right) \left( \frac{2}{\pi} \right) \\
 &= -\frac{4}{\pi^2} \neq 0.
 \end{aligned}$$

Therefore relation (4.29) is not fulfilled, which concludes the lack of homotopy invariance.

Example 4.24 points to a fundamental difference between Horozov's approach in [13] and ours. Indeed, the integrals of Horozov differ from ours in particular in the form of the differentials being integrated: ours are second-order, his are first order. Even when using first order differentials, the integrals are, in general, not homotopy invariant. That they are in [13] stems from the fact that the ambient dimension there (our  $d$ ) is equal to 2, and in this case all homotopies are *thin* homotopies.

#### 4.6. Continuity of the 2D signature

Continuity concerning the signal is an essential part of rough paths theory as well as classical analysis (see e.g. [4, Prop. 2.8]). We derive this type of continuity property here for the 2D signature of sufficiently smooth fields.

**Proposition 4.25.** *Let  $X, \tilde{X}$  be two fields in  $C^2([0, T]^2; \mathbb{R}^d)$ . Consider the corresponding 2D signatures,  $S(X)$  and  $S(\tilde{X})$ , given in Definition 4.10. Let also  $(w, \tau) \in \hat{\mathcal{W}}_n$ , where  $\hat{\mathcal{W}}_n$  is given as in Definition 4.9. Then the following estimate holds for any  $(s, t) \in \Delta_{[0, T]^2}^2$ :*

$$|\langle S_{st}(X) - S_{st}(\tilde{X}), (w, \tau) \rangle| \leq \frac{T^{2n}}{(n!)^2} \sum_{k=1}^n \left( \prod_{i=1}^{k-1} \|X^{w_i}\|_{C^2} \prod_{j=k+1}^n \|\tilde{X}^{w_j}\|_{C^2} \right) \|X^{w_k} - \tilde{X}^{w_k}\|_{C^2}. \quad (4.30)$$

In particular, the mapping  $X \mapsto S_{st}(X)$  is locally Lipschitz.

**Proof.** Observe that for two sequences of numbers,  $\{a_i\}_{i=1}^n$  and  $\{b_i\}_{i=1}^n$ , we have that

$$\prod_{i=1}^n a_i - \prod_{i=1}^n b_i = \prod_{i=1}^{n-1} a_i (a_n - b_n) + \left( \prod_{i=1}^{n-1} a_i - \prod_{i=1}^{n-1} b_i \right) b_n. \quad (4.31)$$

Now consider a generic word  $(w, \tau) \in \hat{\mathcal{W}}_n$  and  $(r^1, \dots, r^n) \in \Delta_{[s, t]}^n$ . Recalling Notation 3.16 for partial derivatives and invoking relation (4.31), we have

$$\begin{aligned} & \left| \prod_{i=1}^n \partial_{12} X_{r_1^i, r_2^{\tau(i)}}^{w_i} - \prod_{i=1}^n \partial_{12} \tilde{X}_{r_1^i, r_2^{\tau(i)}}^{w_i} \right| \\ & \leq \sum_{k=1}^n \left| \prod_{i=1}^{k-1} \partial_{12} X_{r_1^i, r_2^{\tau(i)}}^{w_i} \prod_{j=k+1}^n \partial_{12} \tilde{X}_{r_1^j, r_2^{\tau(j)}}^{w_j} \right| \left| \partial_{12} X_{r_1^k, r_2^{\tau(k)}}^{w_k} - \partial_{12} \tilde{X}_{r_1^k, r_2^{\tau(k)}}^{w_k} \right|. \end{aligned}$$

This leads us to conclude that

$$\left| \prod_{i=1}^n \partial_{12} X_{r_1^i, r_2^{\tau(i)}}^{w_i} - \prod_{i=1}^n \partial_{12} \tilde{X}_{r_1^i, r_2^{\tau(i)}}^{w_i} \right| \leq \sum_{k=1}^n \prod_{i=1}^{k-1} \|X^{w_i}\|_{C^2} \prod_{j=k+1}^n \|\tilde{X}^{w_j}\|_{C^2} \|X^{w_k} - \tilde{X}^{w_k}\|_{C^2},$$

Inserting this into the signature definition, and using that

$$\int_{\Delta_{[0, T]}^n} 1 \, d\mathbf{r}^1 \dots d\mathbf{r}^n = \frac{T^n}{(n!)^2},$$

we obtain

$$\begin{aligned} |\langle S_{st}(X) - S_{st}(\tilde{X}), (w, \tau) \rangle| & \leq \int_{\Delta_{[0, T]}^n} \left| \prod_{i=1}^n \partial_{12} X_{r_1^i, r_2^{\tau(i)}}^{w_i} - \prod_{i=1}^n \partial_{12} \tilde{X}_{r_1^i, r_2^{\tau(i)}}^{w_i} \right| d\mathbf{r}^1 \dots d\mathbf{r}^n \\ & \leq \frac{T^{2n}}{(n!)^2} \sum_{k=1}^n \prod_{i=1}^{k-1} \|X^{w_i}\|_{C^2} \prod_{j=k+1}^n \|\tilde{X}^{w_j}\|_{C^2} \|X^{w_k} - \tilde{X}^{w_k}\|_{C^2}, \end{aligned}$$

which is our claim (4.30). This finishes the proof.  $\square$

#### 4.7. Products of iterated integrals – 2D shuffle

Our interpretation of the shuffle property is that a product of two elements in the signature should be expressed as a linear combination of elements in the signature. As already mentioned in Remark 4.14, one of the main reasons for introducing the notion of the full 2D signature is related to the shuffle property. We will now state this relation in full generality. Notice that a similar result appears in [32, Proposition 1.3] without proof. We include a complete argument here for the sake of clarity. We note that the two-dimensional shuffle algebra has also recently been investigated in [33], in particular in its relation to the sums signature of [16].

Before stating our main result on shuffle products, let us rephrase the shuffle notation in (3.1) and  $(\sigma \star \tau) \circ \rho^{-1}$  in (3.4) directly in terms of the permutations  $\sigma, \tau$ . This will shorten some of our notation in the computations throughout the section.

**Notation 4.26.** Recall Notation 3.2 about permutations. Then consider

$$\sigma \in \Sigma_{\{1, \dots, n\}}, \quad \text{and} \quad \tau \in \Sigma_{\{n+1, \dots, n+k\}} \quad (4.32)$$

We call  $\text{Sh}(\sigma, \tau)$  the following set of permutations of  $\{1, \dots, n+k\}$

$$\text{Sh}(\sigma, \tau) := \{\rho \in \Sigma_{\{1, \dots, n+k\}} \mid \rho \text{ does not change the order of } \tau \text{ and } \sigma\}.$$

Note that this can equivalently be defined as

$$\text{Sh}(\sigma, \tau) := \{(\sigma * \tau) \circ \rho^{-1} \mid \rho \in \text{Sh}(n, k)\},$$

where we have used that any  $\rho \in \text{Sh}(\sigma, \tau)$  can be written as  $\rho = (\sigma * \tau) \circ \hat{\rho}^{-1}$  for a given element  $\hat{\rho} \in \text{Sh}(n, k)$ .

**Notation 4.27.** Consider two words  $w = (i_1, \dots, i_n)$  and  $w' = (j_1, \dots, j_{n'})$  in  $\mathcal{W}$ , where  $\mathcal{W}$  was defined in Definition 1.1. The concatenation of  $w$  and  $w'$  is denoted by  $ww' \in \mathcal{W}$  and  $[ww']_j$  refers to the  $j$ th letter in  $ww'$ .

**Remark 4.28.** Thanks to Notation 4.26 one can rephrase Proposition 3.14 in a slightly more compact way. That is, for  $\sigma \in \Sigma_{\{1, \dots, n\}}$  and  $\tau \in \Sigma_{\{1, \dots, n'\}}$  we have

$$\int_{\Delta_{s,t}^n} \star \int_{\Delta_{s,t}^{n'}} = \sum_{\phi \in \text{Sh}(\sigma, \tau)} \int_{\Delta_{s,t}^{n+n'} \phi}. \quad (4.33)$$

We are now ready to state the shuffle property for the full signature  $\mathbf{S}$ .

**Proposition 4.29 (Shuffle Relation).** Let  $X$  be a field in  $C^2([0, T]^2; \mathbb{R}^d)$  and  $\mathbf{S}(X)$  be its 2D signature (Definition 4.10). We consider two words,  $(w, v) \in \hat{\mathcal{W}}_n$  and  $(w', v') \in \hat{\mathcal{W}}_{n'}$  (Definition 4.9). Then for  $(s, t) \in \Delta_{[0, T]^2}^2$  we have

$$\langle \mathbf{S}_{\text{st}}(X), (w, v) \rangle \langle \mathbf{S}_{\text{st}}(X), (w', v') \rangle = \sum_{\tau \in \text{Sh}(\text{id}_n, \text{id}_{n'})} \sum_{\lambda \in \text{Sh}(v, v')} \langle \mathbf{S}_{\text{st}}(X), ([ww']_{\tau}, \lambda \circ \tau) \rangle, \quad (4.34)$$

where we recall Notation 4.27 for the concatenation  $ww'$  of words. In (4.34), we also recall that for a word  $w$  with  $|w| = n$  and a permutation  $\tau \in \Sigma_{\{1, \dots, n\}}$  we set  $w_{\tau} = (w_{\tau(1)}, \dots, w_{\tau(n)})$ .

**Proof.** For two generic words  $(w, v) \in \hat{\mathcal{W}}_n$  and  $(w', v') \in \hat{\mathcal{W}}_{n'}$ , let us denote by  $\mathcal{G}_{\text{st}}$  the quantity of interest, i.e.

$$\mathcal{G}_{\text{st}} := \langle \mathbf{S}_{\text{st}}(X), (w, v) \rangle \langle \mathbf{S}_{\text{st}}(X), (w', v') \rangle.$$

Owing to relation (4.17) one can write

$$\mathcal{G}_{\text{st}} = \left( \int_{\Delta_{s_1, t_1}^n} \star' \int_{\Delta_{s_2, t_2}^{n'} v^{-1}} \right) \star \left( \int_{\Delta_{s_1, t_1}^{n'}} \star' \int_{\Delta_{s_2, t_2}^{n'} v'^{-1}} \right) \prod_{i=1}^n dX_{\mathbf{r}}^{w_i} \prod_{j=1}^{n'} dX_{\mathbf{r}}^{w'_j},$$

which, by Lemma 3.9, is equal to

$$\mathcal{G}_{\text{st}} = \left( \int_{\Delta_{s_1, t_1}^n} \star \int_{\Delta_{s_1, t_1}^{n'}} \right) \star' \left( \int_{\Delta_{s_2, t_2}^{n'} v^{-1}} \star \int_{\Delta_{s_2, t_2}^{n'} v'^{-1}} \right) \prod_{i=1}^n dX_{\mathbf{r}}^{w_i} \prod_{j=1}^{n'} dX_{\mathbf{r}}^{w'_j}$$

We now apply relation (4.33) separately for the integrals over  $[s_1, t_1]$  and  $[s_2, t_2]$ . This yields

$$\mathcal{G}_{\text{st}} = \sum_{\rho \in \text{Sh}(\text{id}_n, \text{id}_{n'})} \sum_{\sigma \in \text{Sh}(v^{-1}, v'^{-1})} \mathcal{G}_{\text{st}}^{\rho, \sigma}, \quad (4.35)$$

where we define

$$\mathcal{G}_{\text{st}}^{\rho, \sigma} = \int_{\Delta_{s_1, t_1}^{n+n'} \rho \times \Delta_{s_2, t_2}^{n+n'} \sigma} \prod_{i=1}^{n+n'} dX_{\mathbf{r}}^{\hat{w}_i}, \quad (4.36)$$

and we have set  $\text{id}_n = \text{id}_{\{1, \dots, n\}}$  and  $\hat{w} = ww'$ . To express every element in (4.35) as an element of the 2D signature like (4.17), we now replace the simplexes  $\Delta_{s_1, t_1}^{n+n'} \rho$  by  $\Delta_{s_1, t_1}^{n+n'}$  in the definition (4.36) of  $\mathcal{G}_{\text{st}}^{\rho, \sigma}$ . To this aim, we simply declare that  $r_1^k = r_1^{\rho k}$  and perform the corresponding change in the multi-index  $\hat{w}$ . Namely, we define a new word  $w^{\rho}$  by setting

$$\hat{w}_{\rho^{-1}} = (\hat{w}_{\rho^{-1}}, \dots, \hat{w}_{\rho^{-1}}).$$

This simple change of variables enables us to write

$$\mathcal{G}_{\text{st}}^{\rho, \sigma} = \int_{\Delta_{s_1, t_1}^{n+n'} \times \Delta_{s_2, t_2}^{n+n'} \sigma} \prod_{i=1}^{n+n'} dX_{\mathbf{r}}^{\hat{w}_i} = \int_{\Delta_{s_1, t_1}^{n+n'} \times \Delta_{s_2, t_2}^{n+n'} \rho \circ \sigma} \prod_{i=1}^{n+n'} dX_{\mathbf{r}}^{\hat{w}_{\rho^{-1}(i)}}. \quad (4.37)$$

And so we can rewrite (4.37) resorting to the representation in (4.17), in order to obtain

$$\mathcal{G}_{\text{st}}^{\rho, \sigma} = \langle \mathbf{S}_{\text{st}}(X), (\hat{w}_{\rho^{-1}}, \sigma^{-1} \circ \rho^{-1}) \rangle. \quad (4.38)$$

Gathering (4.35) and (4.38), we have thus obtained

$$\mathcal{G}_{\text{st}} = \sum_{\rho \in \text{Sh}(\text{id}_n, \text{id}_{n'})} \sum_{\sigma \in \text{Sh}(v^{-1}, v'^{-1})} \langle \mathbf{S}_{\text{st}}(X), (\hat{w}_{\rho^{-1}}, \sigma^{-1} \circ \rho^{-1}) \rangle.$$

Next, notice that  $\sigma \in \text{Sh}(v^{-1}, v'^{-1})$  if and only if  $\sigma^{-1} \in \text{Sh}(v, v')$ . Setting  $\lambda = \sigma^{-1}$  and similarly  $\tau = \rho^{-1}$  above we end up with

$$\mathcal{G}_{\text{st}} = \sum_{\tau \in \text{Sh}(\text{id}_n, \text{id}_{n'})} \sum_{\lambda \in \text{Sh}(v, v')} \langle \mathbf{S}_{\text{st}}(X), (\hat{w}_{\tau}, \lambda \circ \tau) \rangle,$$

which is exactly the relation in (4.34), and so concludes our proof.  $\square$

As an application of the shuffle relation in [Proposition 4.29](#), we will now derive a formula for products of rectangular increments. The proposition below should be seen as an integration by parts formula for 2D-increments.

**Proposition 4.30.** *Under the same conditions and using the same notation as in [Proposition 4.29](#), recall that the rectangular increments  $\square X$  are given in [Definition 3.17](#). Then for  $w = (i_1, \dots, i_n) \in \mathcal{W}_n$  and  $(s, t) \in \Delta^2_{[0,T]^2}$  we have*

$$\prod_{k=1}^n \square_{st} X^{w_k} = \sum_{v, v' \in \Sigma_{\{1, \dots, n\}}} \langle S_{st}(X), (w_v, v') \rangle. \quad (4.39)$$

**Proof.** The relation is obvious for  $n = 1$ . We now proceed by induction. Namely, assume relation (4.39) holds true for any  $n \geq 1$ . We consider a word  $w = (\hat{w}, i_{n+1})$  in  $\mathcal{W}_{n+1}$ , where  $\hat{w} = (i_1, \dots, i_n) \in \mathcal{W}_n$  and  $i_{n+1} \in \{1, \dots, d\}$ . Denote by  $U_{st}^{n+1}$  the quantity of interest for the induction procedure, that is,  $U_{st}^{n+1} = \prod_{k=1}^{n+1} \square_{st} X^{i_k}$ . Then we trivially have that

$$U_{st}^{n+1} = \left( \prod_{k=1}^n \square_{st} X^{i_k} \right) \square_{st} X^{i_{n+1}}.$$

Owing to our induction assumption and the fact that  $\square_{st} X^{i_k}$  is an element of the signature, this yields

$$U_{st}^{n+1} = \sum_{v, v' \in \Sigma_{\{1, \dots, n\}}} \langle S_{st}(X), (\hat{w}_v, v') \rangle \langle S_{st}(X), (i_{n+1}, \text{id}_1) \rangle. \quad (4.40)$$

Observe that instead of using (4.34) to express every term in (4.40), it is slightly more convenient to invoke (4.35). We get

$$\langle S_{st}(X), (\hat{w}_v, v') \rangle \langle S_{st}(X), (i_{n+1}, \text{id}_1) \rangle = \sum_{\rho \in \text{Sh}(v^{-1}, \text{id}_1)} \sum_{\sigma \in \text{Sh}(v'^{-1}, \text{id}_1)} \int_{\Delta_{s_1, t_1}^{n+1} \rho \times \Delta_{s_2, t_2}^{n+1} \sigma} \prod_{k=1}^{n+1} dX_{\mathbf{r}^k}^{i_k}.$$

Plugging this identity into (4.40) we obtain

$$U_{st}^{n+1} = \sum_{v, v' \in \Sigma_{\{1, \dots, n\}}} \sum_{\rho \in \text{Sh}(v^{-1}, \text{id}_1)} \sum_{\sigma \in \text{Sh}(v'^{-1}, \text{id}_1)} \int_{\Delta_{s_1, t_1}^{n+1} \rho \times \Delta_{s_2, t_2}^{n+1} \sigma} \prod_{k=1}^{n+1} dX_{\mathbf{r}^k}^{i_k}. \quad (4.41)$$

In addition, it is easily seen that

$$\bigcup_{v' \in \Sigma_{\{1, \dots, n\}}} \text{Sh}(v'^{-1}, \text{id}_1) = \Sigma_{\{1, \dots, n+1\}}.$$

Therefore, one can recast (4.41) as

$$U_{st}^{n+1} = \sum_{v \in \Sigma_{\{1, \dots, n\}}} \sum_{\rho \in \text{Sh}(v^{-1}, \text{id}_1)} \sum_{\tau \in \Sigma_{\{1, \dots, n+1\}}} \int_{\Delta_{s_1, t_1}^{n+1} \rho \times \Delta_{s_2, t_2}^{n+1} \tau} \prod_{k=1}^{n+1} dX_{\mathbf{r}^k}^{i_k}. \quad (4.42)$$

We now proceed as in the proof of [Proposition 4.29](#). Namely, we want to replace the integral over  $\Delta_{s_1, t_1}^{n+1} \rho$  by an integral over  $\Delta_{s_1, t_1}^{n+1}$ . To this end, we perform the same change of variable,  $r_1^k = r_1^{\rho_k}$ , which leads to a corresponding permutation  $\mu$  of the word  $(\hat{w}, i_{n+1})$ . We let the reader then check that (4.42) can then be read as

$$U_{st}^{n+1} = \sum_{\mu, v \in \Sigma_{\{1, \dots, n+1\}}} \int_{\Delta_{s_1, t_1}^{n+1} \times \Delta_{s_2, t_2}^{n+1} v} \prod_{k=1}^{n+1} dX_{\mathbf{r}^k}^{w_{(\mu^{-1} \circ v)(k)}} = \sum_{v, v' \in \Sigma_{\{1, \dots, n+1\}}} \int_{\Delta_{s_1, t_1}^{n+1} \times \Delta_{s_2, t_2}^{n+1} v'} \prod_{k=1}^{n+1} dX_{\mathbf{r}^k}^{w_{v(k)}}.$$

This proves the induction step for (4.39), which allows us to conclude.  $\square$

## 5. The symmetrized signature

Up to now, we have introduced two types of 2D signatures for a  $C^2$ -field  $X$ , namely  $\text{S}^{\text{id}}(X)$  and  $\text{S}(X)$ . The 2D signature  $\text{S}^{\text{id}}(X)$  satisfies a Chen-type relation but fails to satisfy the shuffle relation. For  $\text{S}(X)$  the situation is the opposite. In the current section, we introduce another notion of 2D signature which enjoys both Chen and shuffle relations, at the price of integrating over the whole rectangle  $[s_2, t_2]^k$  in the second variable (in contrast to the simplex  $\Delta_{s_2, t_2}^k$ ). Let us start by defining this new type of 2D signature.

Recall that for  $n \geq 0$  we defined the set  $\mathcal{W}_n$  of words of length  $n$  in 1.1 and the set  $\hat{\mathcal{W}}_n$  of extended words in [Definition 4.9](#). Furthermore, the 2D signature  $\text{S}(X)$  is introduced in [Definition 4.10](#).

**Definition 5.1.** Let  $X$  be a field in  $C^2([0, T]^2; \mathbb{R}^d)$ . For  $w \in \mathcal{W}_n$  and  $(s, t) \in \Delta^2_{[0,T]^2}$  we set

$$\langle S_{st}^{\text{Sym}}(X), w \rangle := \sum_{v \in \Sigma_{\{1, \dots, n\}}} \langle S_{st}(X), (w, v) \rangle. \quad (5.1)$$

**Remark 5.2.** Notice that for  $(s, t) \in \Delta^2_{[0,T]^2}$  one can also see  $S_{st}^{\text{Sym}}(X)$  as an element of  $\mathcal{T}(\mathbb{R}^d)$ . This kind of representation follows closely the lines of Section 2.2.

**Remark 5.3.** Combining (5.1) with the representation (4.17) of  $S(X)$  and Lemma 3.15, we find for  $(s, t) \in \Delta_{[0,T]^2}^2$

$$\langle S_{st}^{\text{Sym}}(X), w \rangle = \sum_{v \in \Sigma_{\{1, \dots, n\}}} \int_{\Delta_{s_1, t_1}^n} \star' \int_{\Delta_{s_2, t_2}^n} \prod_{i=1}^n dX_{\mathbf{r}^i}^{w_i} = \int_{\Delta_{s_1, t_1}^n} \star' \int_{[s_2, t_2]^n} \prod_{i=1}^n dX_{\mathbf{r}^i}^{w_i}. \quad (5.2)$$

This formula will be useful for future computations.

### 5.1. Chen's relation

As mentioned above, the symmetrized 2D signature enjoys important algebraic properties. We now prove partial versions of Chen's relation in this context.

**Notation 5.4.** For functions or increments defined on  $[0, T]^2$  we will denote by  $\delta^1$  the delta operator  $\delta$  introduced in Remark 3.13, restricted to act on the 1-variable only, e.g.

$$\delta_{u_1}^1 f_{(s_1, s_2), (t_1, t_2)} = f_{(s_1, s_2), (t_1, t_2)} - f_{(s_1, s_2), (u_1, t_2)} - f_{(u_1, s_2), (t_1, t_2)}.$$

Similarly, we denote by  $\delta^2$  the  $\delta$  operator restricted to the 2-variable.

**Remark 5.5.** It is readily checked that we can compose the two delta functions as  $\delta^1 \circ \delta^2$ , and that the composition is commutative. Furthermore, it is clear that, with the definition from (3.12) in mind, for every field  $X$  in  $C^2([0, T]^2)$  we have

$$\delta_{t_1}^1 \delta_{t_2}^2 X_s = \square_{st} X.$$

**Proposition 5.6.** Let  $X \in C^2([0, T]^2; \mathbb{R}^d)$  and consider  $S^{\text{Sym}}(X)$  defined by (5.1). In addition, pick a word  $w \in \mathcal{W}_n$  together with  $(s, t) \in \Delta_{[0,T]^2}^2$  and  $s_1 \leq u_1 \leq t_1$ . Then we have

$$\langle \delta_{u_1}^1 S_{st}^{\text{Sym}}(X), w \rangle = \sum_{k=1}^{n-1} \langle S_{s(u_1, t_2)}^{\text{Sym}}(X), (w_1 \dots w_k) \rangle \langle S_{(u_1, s_2)t}^{\text{Sym}}(X), (w_{k+1} \dots w_n) \rangle. \quad (5.3)$$

**Proof.** Start from relation (5.2) and apply the  $\delta^1$  operation from Notation 5.4 on both sides of the identity. Thanks to (3.13), we get

$$\langle \delta_{u_1}^1 S_{st}^{\text{Sym}}(X), w \rangle = \sum_{k=1}^{n-1} \left( \int_{\Delta_{s_1, u_1}^k} \star \int_{\Delta_{u_1, t_1}^{n-k}} \right) \star' \int_{[s_2, t_2]^n} \prod_{i=1}^n dX_{\mathbf{r}^i}^{w_i}.$$

Now we decompose  $[s_2, t_2]^n = [s_2, t_2]^k \times [s_2, t_2]^{n-k}$  and apply Fubini's theorem. We get

$$\begin{aligned} \langle \delta_{u_1}^1 S_{st}^{\text{Sym}}(X), w \rangle &= \sum_{k=1}^{n-1} \left( \int_{\Delta_{s_1, u_1}^k} \star \int_{[s_2, t_2]^k} \right) \star \left( \int_{\Delta_{u_1, t_1}^{n-k}} \star' \int_{[s_2, t_2]^{n-k}} \right) \prod_{i=1}^n dX_{\mathbf{r}^i}^{w_i} \\ &= \sum_{k=1}^{n-1} \langle S_{s(u_1, t_2)}^{\text{Sym}}(X), (w_1 \dots w_k) \rangle \langle S_{(u_1, s_2)t}^{\text{Sym}}(X), (w_{k+1} \dots w_n) \rangle, \end{aligned}$$

where the last identity is a simple consequence of (5.1). The proof of (5.3) is now complete.  $\square$

### 5.2. Shuffle relations for the symmetrized 2D signature

We have seen in Proposition 4.29 that the 2D signature  $S(X)$  satisfies a shuffle-type relation. The same is true for the symmetrized 2D signature.

**Proposition 5.7.** Under the same conditions as in Proposition 5.6, let  $w \in \mathcal{W}_n$  and  $w' \in \mathcal{W}_{n'}$  for  $n, n' \geq 1$ . Then for all  $(s, t) \in \Delta_{[0,T]^2}^2$  we have

$$\langle S_{st}^{\text{Sym}}(X), w \rangle \langle S_{st}^{\text{Sym}}(X), w' \rangle = \sum_{\sigma \in \text{Sh}(n, n')} \langle S_{st}^{\text{Sym}}(X), [ww']_{\sigma} \rangle = \langle S_{st}^{\text{Sym}}(X), w \sqcup w' \rangle, \quad (5.4)$$

where  $w \sqcup w'$  is inductively defined in (2.7).

**Proof.** For generic words  $w \in \mathcal{W}_n$  and  $w' \in \mathcal{W}_{n'}$ , we use (5.2) in order to write

$$\langle S_{st}^{\text{Sym}}(X), w \rangle \langle S_{st}^{\text{Sym}}(X), w' \rangle = \left( \int_{\Delta_{s_1, t_1}^n} \star' \int_{[s_2, t_2]^n} \prod_{i=1}^n dX_{\mathbf{r}^i}^{w_i} \right) \left( \int_{\Delta_{s_1, t_1}^{n'}} \star' \int_{[s_2, t_2]^{n'}} \prod_{i=1}^{n'} dX_{\mathbf{r}^i}^{w'_i} \right).$$

By Lemma 3.9

$$\langle S_{st}^{\text{Sym}}(X), w \rangle \langle S_{st}^{\text{Sym}}(X), w' \rangle = \left( \left( \int_{s_1, t_1}^n \star \int_{s_1, t_1}^{n'} \right) \star' \int_{[s_2, t_2]^{n+n'}} \prod_{i=1}^{n+n'} dX_{\mathbf{r}^i}^{[w w']_i} \right) \quad (5.5)$$

Now apply Proposition 3.14 to get the claimed result.  $\square$

## 6. The signature determines the field

In this section we show that the 2D signature of a (time-enhanced) field determines the field. Our proof is based on the Stone–Weierstrass theorem, and relies upon lifting the path to an extended path where the two ‘time’ variables are included. A similar trick is typically used in the classical universal approximation theorem for path signatures. We note that after establishing the “point-separating” property of the signature, together with the shuffle relation derived in Proposition 4.29, one can deduce that the linear functionals on the signature are dense in an appropriate space of continuous functionals on fields (see for example [14, Section 10]). For notational sake, the analysis in this section will be restricted to fields defined on  $[0, 1]^2$ . the trivial extension to  $[0, T]^2$  is left to the reader.

We begin this section with a technical lemma that will be central in subsequent proofs.

**Lemma 6.1.** *Given  $\delta \in (0, 1)$ ,  $\varepsilon > 0$  there exists a univariate polynomial  $\Psi$  such that*

$$\begin{aligned} \Psi(y) &\in [0, 1], & \text{for } y \in [0, 1], \\ |\Psi(y)| &< \varepsilon, & \text{for } y \in [0, \delta/2], \\ |\Psi(y) - 1| &< \varepsilon, & \text{for } y \in [\delta, 1]. \end{aligned}$$

**Proof.** The proof idea is from [34]. We recall Bernoulli’s inequality,

$$(1 + x)^n \geq 1 + nx, \quad \text{for } x \geq -1, n \in \mathbb{N}.$$

Pick  $k \in \mathbb{N}_{\geq 1}$  such that the following relation holds

$$1 < k\delta < 2.$$

Given  $m \in \mathbb{N}_{\geq 1}$  define the function

$$\Phi_m(y) := (1 - y^m)^{k^m}.$$

Then we see that

$$\Phi_m(y) \in [0, 1], \quad \text{for all } y \in [0, 1].$$

Furthermore, for any  $y \in [0, \frac{\delta}{2}]$ , using Bernoulli’s inequality we see that

$$\Phi_m(y) \geq 1 - k^m y^m > 1 - \left(\frac{k\delta}{2}\right)^m.$$

We observe that since  $k\delta/2 < 1$ , this tends to 1 (independent of  $y \in [0, \frac{\delta}{2}]$ ) as  $m \rightarrow \infty$ . Now for the case  $y \in [\delta, 1]$ , it is readily checked that

$$\begin{aligned} \Phi_m(y) &= \frac{1}{k^m y^m} (1 - y^m)^{k^m} k^m y^m \\ &\leq \frac{1}{k^m y^m} (1 - y^m)^{k^m} (1 + k^m y^m) \\ &\leq \frac{1}{k^m y^m} (1 - y^{2m})^{k^m} \leq \frac{1}{k^m y^m} \leq \frac{1}{(k\delta)^m}. \end{aligned}$$

Here we used Bernoulli’s inequality for the second inequality. We observe that since  $k\delta > 1$ , this tends to 0 (uniformly in  $y \in [\delta, 1]$ ) as  $m \rightarrow \infty$ . Combining the two cases, we see that we can choose  $m$  large enough such that  $\Phi := \Phi_m$  satisfies,

$$\Phi(y) \in [0, 1] \text{ for } y \in [0, 1] \quad (6.1)$$

$$\Phi(y) \in [1 - \varepsilon, 1] \text{ for } y \in [0, \frac{\delta}{2}] \quad (6.2)$$

$$\Phi(y) \in [0, \varepsilon] \text{ for } y \in [\delta, 1]. \quad (6.3)$$

Then  $\Psi := 1 - \Phi$  satisfies the required bounds.  $\square$

The next step is to consider a family of functions that is dense in the space of continuous functions on  $[0, 1]^2$  with zero boundary. To this end, let  $\mathcal{F}$  denote the  $\mathbb{R}$ -algebra spanned by the functions,

$$(s_1, s_2) \mapsto s_1^{2m+n} s_2^{2n+m}, \quad m, n \geq 0. \quad (6.4)$$

With this definition at hand, we prove another technical lemma that will allow us to prove that  $\mathcal{F}$  is a dense subset of  $C([0, T]^2; \mathbb{R})$  with zero boundary.

**Lemma 6.2.** *For every  $b > 0$  small enough, there exists a  $b' \in (0, b)$  such that for all  $a > 0$  there is a function  $g_{a,b',b} \in \mathcal{F}$  such that*

$$\|g\| \leq 1, \quad (6.5)$$

$$|g(s_1, s_2)| < a, \quad \text{for } s_1 \in [0, b'] \text{ or } s_2 \in [0, b'], \quad (6.6)$$

$$|g(s_1, s_2) - 1| < a, \quad \text{for } s_1, s_2 \in [b, 1], \quad (6.7)$$

where  $\|g\|$  stands for the sup norm in  $[0, 1]^2$ .

**Proof.** Consider the function

$$P(s_1, s_2) := s_1 s_2 \in \mathcal{F}.$$

Then  $P(s_1, s_2) \in [0, 1]$  for  $s_1, s_2 \in [0, 1]$ . For the given  $b$ , let  $\delta$  be the number defined by

$$\delta := \inf_{s_1, s_2 \in [b, 1]} P(s_1, s_2),$$

and observe that  $\delta = b^2$ . Now pick  $b' < b$  such that the following relation holds

$$\sup_{s_1, s_2 \in ([0, 1] \times [0, b'] \cup ([0, b'] \times [0, 1]))} P(s_1, s_2) < \delta/2.$$

Next we consider a polynomial  $\Phi$  satisfying the requirements of Lemma 6.1, for  $\varepsilon = a$ . We let  $g := \Phi \circ P$ . It is readily checked from the definition of the space proposed in (6.4) that  $g \in \mathcal{F}$ . Therefore, relations (6.1)–(6.3) translate into

$$\begin{aligned} |g(s_1, s_2)| &\in [0, 1], \quad \forall (s_1, s_2) \in [0, 1]^2 \\ |g(s_1, s_2)| &\leq a, \quad \forall (s_1, s_2) \in ([0, 1] \times [0, b']) \cup ([0, b'] \times [0, 1]) \\ |1 - g(s_1, s_2)| &\leq a, \quad \forall (s_1, s_2) \in [0, 1]^2 \end{aligned}$$

Otherwise stated, our claims (6.5)–(6.7) are proved.  $\square$

**Lemma 6.3.** *Let  $\mathcal{F}$  be the vector space defined by (6.4). Then  $\mathcal{F}$  is dense in the subspace*

$$V := \{f \in C([0, 1]^2, \mathbb{R}) \mid f(0, s_2) = f(s_1, 0) = 0, s_1, s_2 \in [0, 1]\},$$

of continuous functions vanishing at the “left-bottom” boundaries.

**Proof.** The strategy for the proof is a subtle variation on classical Stone–Weierstrass type arguments. Namely on every sub-rectangle  $[b', 1]$ , Stone–Weierstrass theorem will give the desired conclusion. Then close to the axes we shall resort to a polynomial interpolation given by Lemma 6.2.

First let us prove that  $\mathcal{F}$  separates points in  $(0, 1]^2$ . Indeed, if

$$s_1^2 s_2 = q_1^2 q_2, \quad s_1 s_2^2 = q_1 q_2^2,$$

and  $s_1, s_2, q_1, q_2 \in (0, 1)$  then

$$\frac{s_1}{q_1} = \frac{q_2^2}{s_2^2} = \frac{s_1^4}{q_1^4},$$

hence

$$s_1 = q_1, \quad s_2 = q_2.$$

Furthermore, the constant function is contained in  $\mathcal{F}$  and it is closed under multiplication. Hence, by the Stone–Weierstrass theorem, the functions are dense in  $C([\varepsilon, 1]^2, \mathbb{R})$ , for every  $\varepsilon > 0$ .

We now start our interpolation argument close to the axes. That is consider  $f \in V$  and  $\varepsilon > 0$  given. Choose  $\delta > 0$  such that

$$|f(s_1, s_2)| < \varepsilon, \quad \text{for } s_1 \in [0, \delta] \text{ or } s_2 \in [0, \delta]. \quad (6.8)$$

Let  $b' < b$  be given as in Lemma 6.2, for  $b := \delta$ . Denote  $\mathcal{F}_{b'}$  the restriction of  $\mathcal{F}$  to  $[b', 1]^2$ . By the classical Stone–Weierstrass theorem, there exists  $\hat{f}_{\varepsilon, b'} \in \mathcal{F}_{b'}$  such that

$$\|f - \hat{f}_{\varepsilon, b'}\|_{C([b', 1]^2, \mathbb{R})} < \varepsilon. \quad (6.9)$$

Fix an extension  $f_{\varepsilon, b'} \in \mathcal{F}$  of  $\hat{f}_{\varepsilon, b'}$  to  $[0, 1]^2$ . Define

$$M := \sup_{s_1, s_2 \in [0, b']} |f_{\varepsilon, b'}(s_1, s_2)| + 1.$$

Consider  $g_{\varepsilon/M, b', b} \in \mathcal{F}$  as in Lemma 6.2. Then,  $h := f_{\varepsilon, \delta} \cdot g_{\varepsilon/M, b', b} \in \mathcal{F}$  satisfies the following three properties:



(i) For  $s_1 \in [0, b']$  or  $s_2 \in [0, b']$ :

$$\begin{aligned} |h(s_1, s_2)| &= |f_{\varepsilon, \delta}(s_1, s_2)| \cdot |g_{\varepsilon/M, b', b}(s_1, s_2)| \\ &\leq M \cdot \varepsilon / M = \varepsilon. \end{aligned}$$

Since  $f$  verifies (6.8) for  $s_1 \in [0, b']$  or  $s_2 \in [0, b']$ , we get that  $|h(s_1, s_2) - f(s_1, s_2)| \leq 2\varepsilon$ .

(ii) For  $s_1 \in [b', b]$ ,  $s_2 \in [b', 1]$  or  $s_2 \in [b', b]$ ,  $s_1 \in [b', 1]$

$$\begin{aligned} |h(s_1, s_2)| &= |f_{\varepsilon, \delta}(s_1, s_2)| \cdot |g_{\varepsilon/M, b', b}(s_1, s_2)| \\ &\leq 2\varepsilon \cdot \|g_{\varepsilon/M, b', b}\|_{\infty} \leq 2\varepsilon. \end{aligned}$$

(iii) For  $s_1, s_2 \in [\delta, 1]$

$$\begin{aligned} |h(s_1, s_2) - f(s_1, s_2)| &\leq |(f_{\varepsilon, \delta}(s_1, s_2) - f(s_1, s_2)) \cdot g_{\varepsilon/M, b', b}| + |f(s_1, s_2)(g_{\varepsilon/M, b', b} - 1)| \\ &\leq \varepsilon + \|f\|\varepsilon/M. \end{aligned}$$

Now  $\varepsilon$  was arbitrary (and  $M \geq 1$ ), and so it follows that  $\mathcal{F}$  is dense in  $V$ .  $\square$

We now turn to the main result of the section, which states that an extended version of the signature characterizes a smooth field up to an equivalent class.

**Theorem 6.4.** Consider the following equivalence relation on  $C^2([0, T]^2; \mathbb{R}^d)$ :

$$X \sim Y \iff \partial_{12}X = \partial_{12}Y. \quad (6.10)$$

For  $X \in C^2([0, 1]^2; \mathbb{R}^d)$  consider the “lift” to  $\hat{X} \in C^2([0, 1]^2; \mathbb{R}^{2+d})$

$$\hat{X}_{s_1, s_2} := (s_1^2 s_2, s_1 s_2^2, X_{s_1, s_2}).$$

Then, for  $X, Y \in C^2([0, 1]^2; \mathbb{R}^d)$  we have  $X \sim Y$  if and only if

$$\mathbf{S}_{0,1}(\hat{X}) = \mathbf{S}_{0,1}(\hat{Y}).$$

**Proof.** It suffices to show the case  $d = 1$ , i.e.

$$\hat{X}_{s_1, s_2} = (\hat{X}_{s_1, s_2}^{(1)}, \hat{X}_{s_1, s_2}^{(2)}, \hat{X}_{s_1, s_2}^{(3)}) = (s_1^2 s_2, s_1 s_2^2, X_{s_1, s_2}) \in \mathbb{R}^3.$$

It is clear that

$$\langle \mathbf{S}_{0s}(\hat{X}), 1 \rangle = s_1^2 s_2, \quad \text{and} \quad \langle \mathbf{S}_{0s}(\hat{X}), 2 \rangle = s_1 s_2^2.$$

Then, by the shuffle relation it follows that

$$\langle \mathbf{S}_{0s}(\hat{X}), 1 \sqcup^m 2 \sqcup^n \rangle = s_1^{2m+n} s_2^{2n+m}.$$

Hence, by Lemma 6.3, for every  $f \in \mathcal{F}$ ,  $\varepsilon > 0$  there exists  $\phi_\varepsilon \in T(\mathbb{R}^d)$  such that with

$$\Phi_\varepsilon(s_1, s_2) := \langle \mathbf{S}_{0s}(\hat{X}), \phi_\varepsilon \rangle,$$

we have  $\|f - \Phi_\varepsilon\| < \varepsilon$ . Note that

$$\langle \mathbf{S}_{01}(\hat{X}), (1 \sqcup^m 2 \sqcup^n) \rangle = \int_0^1 ds_1 \int_0^1 ds_2 s_1^{2m+n} s_2^{2n+m} \partial_{12} X_{s_1, s_2}.$$

If  $\mathbf{S}_{01}(\hat{X}) = \mathbf{S}_{01}(\hat{Y})$  then, for every  $f \in \mathcal{F}$  we have

$$\int_0^1 ds_1 \int_0^1 ds_2 f(s_1, s_2) \partial_{12} X_{s_1, s_2} = \int_0^1 ds_1 \int_0^1 ds_2 f(s_1, s_2) \partial_{12} Y_{s_1, s_2}.$$

By the usual approximation of the Dirac delta we hence get

$$\partial_{12} X(s_1, s_2) = \partial_{12} Y(s_1, s_2), \quad \text{for } s_1, s_2 \in (0, 1].$$

By continuity of  $\partial_{12} X, \partial_{12} Y$ , we hence get

$$\partial_{12} X = \partial_{12} Y.$$

We have thus proved that whenever  $\mathbf{S}_{01}(\hat{X}) = \mathbf{S}_{01}(\hat{Y})$  we have  $\tilde{X} \sim \tilde{Y}$  according to (6.10). The other implication being trivial, this finishes the proof.  $\square$

## 7. Concluding remarks and open problems

We have proposed several 2D extensions of the path signature arising in the theory of rough paths [5]. In contrast to the multiparameter extensions proposed by [14,15], the 2D-signatures we consider here are constructed over fields  $X : [0, T]^2 \rightarrow \mathbb{R}^d$ , using the mixed partial derivative  $\frac{\partial^2}{\partial t_1 \partial t_2} X$  of the field as the driving signal. As such we have invariance to path perturbations (Section 4.5.1), and thus the 2D-signatures may serve as a complement to the classical path signature when working with image data. We show that our (full) signature satisfies shuffle relations (Proposition 4.29), and the id-signature satisfies a type of Chen relation based on convolutional products (Proposition 4.8). While the current paper only deals with sufficiently smooth, i.e., twice continuously differentiable fields  $X$ , we envision that this signature construction will in time see great potential for application in stochastic analysis for random fields, in the spirit signatures for paths has been influential for stochastic process theory. However, we have not considered rough path-type questions yet, such as an extension theorem for the signature, or investigations of probabilistic aspects of the 2D signature constructed from random fields, but see this as highly interesting ongoing and future work.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- [1] Kuo-Tsai Chen, Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula, *Ann. of Math.* (2) 65 (1957) 163–178.
- [2] Terry J. Lyons, Differential equations driven by rough signals, *Rev. Mat. Iberoam.* 14 (2) (1998) 215–310.
- [3] Peter K. Friz, Hairer Martin, *A Course on Rough Paths With an Introduction to Regularity Structures*, in: Universitext, Springer Cham, 2020.
- [4] Peter K. Friz, Nicolas B. Victoir, *Multidimensional Stochastic Processes as Rough Paths*, in: Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010, Theory and applications.
- [5] Terry J. Lyons, Michael Caruana, Thierry Lévy, *Differential Equations Driven by Rough Paths*, in: Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007, 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.
- [6] Ilya Chevyrev, Andrey Kormilitzin, *A primer on the signature method in machine learning*, 2016, arXiv:1603.03788.
- [7] Terry Lyons, Andrew D. McLeod, *Signature methods in machine learning*, 2022, arXiv:2206.14674.
- [8] Benjamin Graham, *Sparse arrays of signatures for online character recognition*, 2013, arXiv:1308.0371.
- [9] Ilya Chevyrev, Vidit Nanda, Harald Oberhauser, Persistence paths and signature features in topological data analysis, *Trans. Pattern Anal. Mach. Int.* 42 (2020) 192–202.
- [10] Imanol Perez Arribas, Guy M. Goodwin, John R. Geddes, Terry Lyons, Kate E.A. Saunders, A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder, *Transl. Psychiatry* 8 (274) (2018).
- [11] Yue Wu, Guy M. Goodwin, Terry Lyons, Kate E.A. Saunders, Identifying psychiatric diagnosis from missing mood data through the use of log-signature features, *PLoS One* 17 (11) (2022).
- [12] L. Sifre, Stéphane Mallat, Rigid-motion scattering for texture classification, 2014, ArXiv, arXiv:1403.1687abs/1403.1687.
- [13] Ivan Horozov, Noncommutative Hilbert modular symbols, *Algebra Number Theory* 9 (2) (2015) 317–370.
- [14] Chad Giusti, Darrick Lee, Vidit Nanda, Harald Oberhauser, A topological approach to mapping space signatures, 2022, arXiv preprint arXiv:2202.00491.
- [15] Darrick Lee, Harald Oberhauser, *Random surfaces and higher algebra*, 2023.
- [16] Joscha Diehl, Leonard Schmitz, Two-parameter sums signatures and corresponding quasisymmetric functions, 2022, arXiv preprint arXiv:2210.14247.
- [17] Joscha Diehl, Kurusch Ebrahimi-Fard, Nikolas Tapia, Time-warping invariants of multidimensional time series, *Acta Appl. Math.* 170 (1) (2020) 265–290.
- [18] Joscha Diehl, Kurusch Ebrahimi-Fard, Nikolas Tapia, Tropical time series, iterated-sums signatures, and quasisymmetric functions, *SIAM J. Appl. Algebra Geom.* 6 (4) (2022) 563–599.
- [19] Mohamed R. Ibrahim, Terry Lyons, Imagesig: A signature transform for ultra-lightweight image recognition, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 3649–3659.
- [20] Chouk Khalil, Gubinelli Massimiliano, *Rough sheets*, 2013, arXiv:1406.7748.
- [21] Khalil Chouk, Samy Tindel, Skorohod and Stratonovich integration in the plane, *Electron. J. Probab.* 20 (39) (2015) 39.
- [22] Sheng Zhang, Guang Lin, Samy Tindel, Two-dimensional signature of images and texture classification, *Proc. A* 478 (2266) (2022) 20220346, 13.
- [23] Fabian A. Harang, Samy Tindel, Volterra equations driven by rough signals, *Stochastic Process. Appl.* 142 (2021) 34–78.
- [24] Frédéric Patras, Construction géométrique des idempotents eulériens. filtration des groupes de polytopes et des groupes d'homologie de hochschild, *Bull. Soc. Math. France* 119 (1991) 173–198.
- [25] Massimiliano Gubinelli, Controlling rough paths, *J. Funct. Anal.* 216 (1) (2004) 86–140.
- [26] Cristopher Salvi, Thomas Cass, James Foster, Terry Lyons, Weixin Yang, The signature kernel is the solution of a goursat pde, *SIAM J. Math. Data Sci.* 3 (3) (2021) 873–899.
- [27] Franz J. Kiraly, Harald Oberhauser, Kernels for sequentially ordered data, *J. Mach. Learn. Res.* 20 (31) (2019) 1–45.
- [28] Thomas Cass, Jeffrey Pei, A Fubini type theorem for rough integration, *Rev. Mat. Iberoam.* 39 (2) (2023) 761–802.
- [29] Fabian A. Harang, An extension of the sewing lemma to hyper-cubes and hyperbolic equations driven by multi-parameter Young fields, *Stoch. Partial. Differ. Equ. Anal. Comput.* 9 (3) (2021) 746–788.
- [30] Florian Bechtold, Fabian A. Harang, Nimit Rana, Non-linear young equations in the plane and pathwise regularization by noise for the stochastic wave equation, *Stoch. Partial. Differ. Equ.: Anal. Comput.* (2022) 1–41.
- [31] Lluís Quer-Sardanyons, Samy Tindel, The 1-d stochastic wave equation driven by a fractional brownian sheet, *Stochastic Process. Appl.* 117 (10) (2007) 1448–1472.
- [32] Ivan Horozov, *Non-commutative two dimensional modular symbol*, 2006.
- [33] Leonard Schmitz, Nikolas Tapia, Free generators and Hoffman's isomorphism for the two-parameter shuffle algebra, 2024, arXiv preprint arXiv:2401.06691.
- [34] Bruno Brosowski, Frank Deutsch, An elementary proof of the Stone–Weierstrass theorem, *Proc. Amer. Math. Soc.* 81 (1) (1981) 89–92.