

# Signatures of paths transformed by polynomial maps

Laura Colmenarejo<sup>\*</sup> and Rosa Preiß<sup>†</sup>

## Abstract

We characterize the signature of piecewise continuously differentiable paths transformed by a polynomial map in terms of the signature of the original path. For this aim, we define recursively an algebra homomorphism between two shuffle algebras on words. This homomorphism does not depend on the path and behaves well with respect to composition and homogeneous maps. It allows us to describe the relation between the signature of a piecewise continuously differentiable path and the signature of the path obtained by transforming it under a polynomial map. We also study this map as a half-shuffle homomorphism and give a generalization of our main theorem in terms of Zinbiel algebras.

*Keywords:* Signature tensors, iterated integrals, tensor algebra, shuffle product, polynomial maps.

## 1 Introduction

In the 1950s, K. T. Chen introduced the *iterated-integral signature* of a piecewise continuously differentiable path, which up to a natural equivalence relation, determines the initial path. In general, the signature of a path can be seen as a multidimensional time series. When the terminal time is fixed, the signature of a path can be seen as tensors and the calculation of the signature becomes a standard problem in data science. In [PSS19], M. Pfeffer, A. Seigal, and B. Sturmfels study the inverse problem: given partial information from a signature, can we recover the path? They consider signature tensors of order three under linear transformations and establish identifiability results and recovery algorithms for piecewise linear paths, polynomial paths, and generic dictionaries.

Coming from stochastic analysis, the signatures are becoming more relevant in other areas, such as algebraic geometry and combinatorics, and we would like to highlight some recent work. For instance, in [DR19], J. Diehl and J. Reizenstein offer a combinatorial approach to the understanding of invariants of multidimensional time series based on their signature. Another reference is [AFS19], in which C. Améndola, P. Friz, and B. Sturmfels look at the varieties of signatures of tensors for both deterministic and random paths, focusing on piecewise linear paths and polynomial paths, among others. Answering one of their questions, in [Gal19], F. Galuppi looks at rough paths, for which their signature variety shows surprising analogies with the Veronese variety.

In stochastic analysis, the study of the signatures of paths arises in the theory of rough paths, where [FV10, FH14] are textbook references. Iterated integrals and the non-commutative series that encode them have also arisen in a variety of contexts in geometry and arithmetic, including the work of R. Hain in [Hai02], M. Kapranov in [Kap09], and J. Balakrishnan in [Bal13]. The results we derive in this paper have the potential for future applications in all of these contexts.

---

<sup>\*</sup>Laura Colmenarejo, UMass Amherst (USA), [laura.colmenarejo.hernando@gmail.com](mailto:laura.colmenarejo.hernando@gmail.com)

<sup>†</sup>Rosa Preiß, Technical University Berlin (Berlin, Germany), [preiss@math.tu-berlin.de](mailto:preiss@math.tu-berlin.de)

Let us now present our problem and our main two results, Theorems 1.1 and 1.2.

A piecewise continuously differentiable path  $X$  in  $\mathbb{R}^d$  is a map defined by  $d$  piecewise smooth functions  $X^i(t)$  in a parameter  $t \in [0, L]$ , for  $i = 1, \dots, d$ . Its signature stores the collection of all the iterated integrals of the path  $X$ , which are of the form

$$\int_0^L \int_0^{r_n} \cdots \int_0^{r_2} dX_{r_1}^{i_1} \cdots dX_{r_n}^{i_n}, \quad (1)$$

where  $X_r^i := X^i(r)$ . The iterated integral (1) is a real number and it is associated to the sequence  $(i_1, i_2, \dots, i_n)$ , for which the order is relevant. Therefore, we consider the signature  $\sigma(X)$  as an element of  $T(\mathbb{R}^d)$ , the space of formal power series in words in the alphabet  $\{1, 2, \dots, d\}$ . This space becomes an algebra with the concatenation product, denoted by the symbol  $\bullet$ . Its algebraic dual, denoted by  $T(\mathbb{R}^d)$ , is the space of non-commutative polynomials in the same set of words. It is a commutative algebra with the shuffle product, which is denoted by  $\sqcup$  and interleaves two words in all order-preserving ways, [Reu93].

We also consider the following duality pairing in  $T(\mathbb{R}^d) \times T(\mathbb{R}^d)$ :

$$\left\langle \sum_{\mathbf{w} \in \mathcal{W}_d} a_{\mathbf{w}} \mathbf{w}, \mathbf{v} \right\rangle = a_{\mathbf{v}}, \quad (2)$$

where  $\mathcal{W}_d$  denotes the set of words in the alphabet  $\{1, \dots, d\}$ , together with the empty word  $\mathbf{e}$ .

Let  $X$  be a piecewise continuously differentiable path in  $\mathbb{R}^d$  and  $\sigma(X)$  be its signature. Consider a polynomial map  $p$  from  $\mathbb{R}^d$  to  $\mathbb{R}^m$ . One can compute the image path  $p(X)$  and ask for its signature,  $\sigma(p(X))$ . Then, the following question comes up:

*How are both signatures,  $\sigma(X)$  and  $\sigma(p(X))$ , related?*

We approach this question from an algebraic point of view. We consider the dual map  $p^* : \mathbb{R}[x_1, \dots, x_m] \longrightarrow \mathbb{R}[x_1, \dots, x_d]$ , where both sets of variables are commutative. It is natural and common to embed the polynomial ring  $\mathbb{R}[x_1, \dots, x_m]$  into the tensor algebra  $(T(\mathbb{R}^m), \sqcup)$ . For that we identify the variable  $x_i$  with the letter  $\mathbf{i}$  and we define the embedding, denoted by  $\varphi_m$  (or  $\varphi$ ), by sending the monomial  $x_{i_1} \cdots x_{i_l}$  to the shuffle product  $\mathbf{i}_1 \sqcup \cdots \sqcup \mathbf{i}_l$ , for  $1 \leq i_1, \dots, i_l \leq m$ , and extending by linearity. By construction, this map is a morphism of commutative algebras, and it is injective but not surjective. For instance, for  $t \geq 2$ ,  $\varphi(x_1 \cdot x_2) = \mathbf{1} \sqcup \mathbf{2} = \mathbf{12} + \mathbf{21}$  and there is no other way to obtain the words  $\mathbf{12}$  and  $\mathbf{21}$  as  $\varphi_d(h)$ , for any polynomial  $h \in \mathbb{R}[x_1, \dots, x_d]$ . Therefore, we cannot find a polynomial in  $\mathbb{R}[x_1, \dots, x_d]$  with image  $\mathbf{12}$ .

Our first step is to define a map  $M_p : (T(\mathbb{R}^m), \sqcup) \longrightarrow (T(\mathbb{R}^d), \sqcup)$ , which is an algebra homomorphism and that is unique in the following sense.

**Theorem 1.1.** *There exists an algebra homomorphism  $M_p : (T(\mathbb{R}^m), \sqcup) \longrightarrow (T(\mathbb{R}^d), \sqcup)$  such that its restriction  $M_p|_{\text{Im}(\varphi_m)}$  is the unique algebra homomorphism that makes the*

following diagram commute:

$$\begin{array}{ccc}
\mathbb{R}[x_1, \dots, x_m] & \xrightarrow{p^*} & \mathbb{R}[x_1, \dots, x_d] \\
\downarrow \varphi_m & & \downarrow \varphi_d \\
\text{Im}(\varphi_m) & \xrightarrow{M_p|_{\text{Im}(\varphi_m)}} & \text{Im}(\varphi_d) \\
\cap & & \cap \\
(\mathbb{T}(\mathbb{R}^m), \sqcup) & & (\mathbb{T}(\mathbb{R}^d), \sqcup)
\end{array} \tag{3}$$

The map  $M_p$  has some further interesting properties and is the key to relate the signature of a path with the signature of its transformation under a polynomial map.

**Theorem 1.2.** *Let  $X : [0, L] \longrightarrow \mathbb{R}^d$  be a piecewise continuously differentiable path with  $X(0) = 0$  and let  $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$  be a polynomial map with  $p(0) = 0$ . Then, for all  $\mathbf{w} \in \mathbb{T}(\mathbb{R}^m)$ ,*

$$\langle \sigma(p(X)), \mathbf{w} \rangle = \langle \sigma(X), M_p(\mathbf{w}) \rangle.$$

Equivalently,  $\sigma(p(X)) = M_p^*(\sigma(X))$ .

The paper is organized as follows. In Section 2, we introduce briefly the framework of signatures of paths, as well as the basic notions that we need for our key combinatorial objects, the words. In Section 3, we define the map  $M_p$  and prove its properties in Proposition 3.2. Moreover, we present our main theorems, Theorems 1.1 and 1.2, together with a generalization of the last one, Corollary 3.3. In Section 3.1 we look at  $M_p$  as a half-shuffle homomorphism and give a generalization of Theorem 1.2 in terms of Zinbiel algebras. In Section 4, we also present two examples and a few consequences, Corollaries 4.4–4.7. Finally, Section 5 is dedicated to applications and future work.

## 2 Signatures of paths and words

Given a piecewise continuously differentiable path  $X : [0, L] \longrightarrow \mathbb{R}^d$ , for any  $i_1, \dots, i_n \in \{1, 2, \dots, d\}$  the following integral is classically well-defined

$$\int_0^L dX^{i_1} \dots dX^{i_n} := \int_0^L \int_0^{r_n} \dots \int_0^{r_2} dX_{r_1}^{i_1} \dots dX_{r_n}^{i_n} = \int_0^L \int_0^{r_n} \dots \int_0^{r_2} \dot{X}_{r_1}^{i_1} \dots \dot{X}_{r_n}^{i_n} dr_1 \dots dr_n.$$

We would like to store the collection of all these integrals.

**Definition 2.1.** *The signature of  $X$  is defined as the following formal power series*

$$\sigma(X) = \sum_{n \geq 0} \sum_{i_1 \dots i_n} \underbrace{\int_0^L \int_0^{r_n} \dots \int_0^{r_2} dX_{r_1}^{i_1} \dots dX_{r_n}^{i_n}}_{\in \mathbb{R}} \cdot \mathbf{i}_1 \dots \mathbf{i}_n \in \mathbb{T}(\mathbb{R}^d).$$

As we mention in the introduction,  $\mathbb{T}(\mathbb{R}^d)$  is the space of formal power series in words in the alphabet  $\{1, \dots, d\}$ , and we denote by  $\mathbf{e}$  the empty word. It is an algebra with the concatenation product, denoted by  $\mathbf{w} \bullet \mathbf{v}$  (or simply  $\mathbf{wv}$ ), which is well-defined since it respects the grading given by the number of letters appearing in each word. We also consider its algebraic dual  $\mathbb{T}(\mathbb{R}^d)$ , which is the set of polynomials in words in the same alphabet. The algebra  $\mathbb{T}(\mathbb{R}^d)$  has the concatenation product, which is the same as for  $\mathbb{T}(\mathbb{R}^d)$  if we multiply two finite power series. However, we consider  $\mathbb{T}(\mathbb{R}^d)$  as an algebra with the shuffle product, which we define recursively as follows.

**Definition 2.2.** Let  $\mathbf{w}$ ,  $\mathbf{w}_1$  and  $\mathbf{w}_2$  be three words and  $\mathbf{a}$  and  $\mathbf{b}$  two letters. We define the shuffle product of two words recursively by

$$\begin{aligned} \mathbf{e} \sqcup \mathbf{w} &= \mathbf{w} \sqcup \mathbf{e} = \mathbf{w}, \text{ and} \\ (\mathbf{w}_1 \bullet \mathbf{a}) \sqcup (\mathbf{w}_2 \bullet \mathbf{b}) &= (\mathbf{w}_1 \sqcup (\mathbf{w}_2 \bullet \mathbf{b})) \bullet \mathbf{a} + ((\mathbf{w}_1 \bullet \mathbf{a}) \sqcup \mathbf{w}_2) \bullet \mathbf{b}. \end{aligned}$$

Note that in the shuffle product, we distinguish duplicated letters. For instance, for a letter  $\mathbf{a}$ , we have  $\mathbf{a} \sqcup \mathbf{a} = 2 \cdot \mathbf{aa}$ . Notice that the concatenation is a non-commutative operation, whereas the shuffle product is commutative.

We also need a few notions on words. The *length of a word*  $\mathbf{w}$  is denoted by  $\ell(\mathbf{w})$  and counts the number of letters in  $\mathbf{w}$ . We extend this definition by linearity, defining  $\ell(\mathbf{w}_1 + \mathbf{w}_2) := \max\{\ell(\mathbf{w}_1), \ell(\mathbf{w}_2)\}$ , for any words  $\mathbf{w}_1$  and  $\mathbf{w}_2$ . Therefore,  $\ell(\mathbf{w}_1 \bullet \mathbf{w}_2) = \ell(\mathbf{w}_1) + \ell(\mathbf{w}_2) = \ell(\mathbf{w}_1 \sqcup \mathbf{w}_2)$ . As an  $\mathbb{R}$ -vector space,  $T(\mathbb{R}^d)$  is graded by the length of the words:

$$T(\mathbb{R}^d) = \bigoplus_{n \geq 0} T^n(\mathbb{R}^d),$$

where  $T^n(\mathbb{R}^d)$  is the vector space spanned by the words of length  $n$ . We also denote by  $T^{\leq n}(\mathbb{R}^d)$  the partial direct sum  $\bigoplus_{k \leq n} T^k(\mathbb{R}^d)$ . This notation extends to  $T((\mathbb{R}^d))$ . The same

way,  $\sigma^{(n)}(X)$  denotes the partial sum of  $\sigma(X)$  for which all the appearing words have length exactly  $n$ . We are ready to prove the following result.

**Proposition 2.3.** *The shuffle product is associative.*

*Proof.* The associativity is clear for the empty word since  $(\mathbf{e} \sqcup \mathbf{e}) \sqcup \mathbf{e} = \mathbf{e} = \mathbf{e} \sqcup (\mathbf{e} \sqcup \mathbf{e})$ . Now, we proceed by induction. Assume that for any words  $\mathbf{w}_1$ ,  $\mathbf{v}_1$ , and  $\mathbf{u}_1$  such that  $\ell(\mathbf{w}_1) + \ell(\mathbf{v}_1) + \ell(\mathbf{u}_1) = n$ , for some  $n \in \mathbb{N}_0$ , we have that  $(\mathbf{w}_1 \sqcup \mathbf{v}_1) \sqcup \mathbf{u}_1 = \mathbf{w}_1 \sqcup (\mathbf{v}_1 \sqcup \mathbf{u}_1)$ . This is our inductive hypothesis.

Let  $\mathbf{w}_2, \mathbf{v}_2, \mathbf{u}_2$  be arbitrary words with the property that  $\ell(\mathbf{w}_2) + \ell(\mathbf{v}_2) + \ell(\mathbf{u}_2) = n + 1$ . At least one of those words must thus be non-empty. If exactly two of the words are empty, both  $(\mathbf{w}_2 \sqcup \mathbf{v}_2) \sqcup \mathbf{u}_2$  and  $\mathbf{w}_2 \sqcup (\mathbf{v}_2 \sqcup \mathbf{u}_2)$  are obviously equal to the non-empty word. If exactly one of the words is empty, both  $(\mathbf{w}_2 \sqcup \mathbf{v}_2) \sqcup \mathbf{u}_2$  and  $\mathbf{w}_2 \sqcup (\mathbf{v}_2 \sqcup \mathbf{u}_2)$  are obviously equal to the shuffle product of the two non-empty words. In the remaining case, if  $\mathbf{w}_2, \mathbf{v}_2, \mathbf{u}_2$  are all non-empty, there are words  $\mathbf{w}, \mathbf{v}, \mathbf{u}$  and letters  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  such that  $\mathbf{w}_2 = \mathbf{wi}$ ,  $\mathbf{v}_2 = \mathbf{vj}$  and  $\mathbf{u}_2 = \mathbf{uk}$ . Then,

$$\begin{aligned} (\mathbf{w}_2 \sqcup \mathbf{v}_2) \sqcup \mathbf{u}_2 &= (\mathbf{wi} \sqcup \mathbf{vj}) \sqcup \mathbf{uk} = ((\mathbf{w} \sqcup \mathbf{vj}) \bullet \mathbf{i} + (\mathbf{wi} \sqcup \mathbf{v}) \bullet \mathbf{j}) \sqcup \mathbf{uk} \\ &= ((\mathbf{w} \sqcup \mathbf{vj}) \sqcup \mathbf{uk}) \bullet \mathbf{i} + ((\mathbf{wi} \sqcup \mathbf{v}) \sqcup \mathbf{uk}) \bullet \mathbf{j} + (((\mathbf{w} \sqcup \mathbf{vj})\mathbf{i} + (\mathbf{wi} \sqcup \mathbf{v})\mathbf{j}) \sqcup \mathbf{u}) \bullet \mathbf{k} \\ &= ((\mathbf{w} \sqcup \mathbf{vj}) \sqcup \mathbf{uk}) \bullet \mathbf{i} + ((\mathbf{wi} \sqcup \mathbf{v}) \sqcup \mathbf{uk}) \bullet \mathbf{j} + ((\mathbf{wi} \sqcup \mathbf{vj}) \sqcup \mathbf{u}) \bullet \mathbf{k} \end{aligned}$$

Analogously,

$$\mathbf{w}_2 \sqcup (\mathbf{v}_2 \sqcup \mathbf{u}_2) = (\mathbf{w} \sqcup (\mathbf{vj} \sqcup \mathbf{uk})) \bullet \mathbf{i} + (\mathbf{wi} \sqcup (\mathbf{v} \sqcup \mathbf{uk})) \bullet \mathbf{j} + (\mathbf{wi} \sqcup (\mathbf{vj} \sqcup \mathbf{u})) \bullet \mathbf{k}.$$

Thus, since

$$\ell(\mathbf{w}) + \ell(\mathbf{vj}) + \ell(\mathbf{uk}) = \ell(\mathbf{wi}) + \ell(\mathbf{v}) + \ell(\mathbf{uk}) = \ell(\mathbf{wi}) + \ell(\mathbf{vj}) + \ell(\mathbf{u}) = n,$$

we again get  $(\mathbf{w}_2 \sqcup \mathbf{v}_2) \sqcup \mathbf{u}_2 = \mathbf{w}_2 \sqcup (\mathbf{v}_2 \sqcup \mathbf{u}_2)$  due to the induction hypothesis.  $\square$

Going back to the signatures, the dual pairing (2) in  $\mathbb{T}(\mathbb{R}^d) \times \mathbb{T}(\mathbb{R}^d)$  allows us to extract the coefficient of a word in the signature of a path in the following way:

$$\langle \sigma(X), \mathbf{i}_1 \mathbf{i}_2 \dots \mathbf{i}_n \rangle = \int_0^L dX^{i_1} \dots dX^{i_n}.$$

Both operations, the concatenation and the shuffle products, behave nicely with respect to the signature, as the following two known results describe. The first result, known as the *shuffle identity*, relates the signature of a path with the shuffle product.

**Proposition 2.4** (Shuffle identity, [Ree58]). *Let  $X : [0, L] \rightarrow \mathbb{R}^d$  be a piecewise continuously differentiable path. Then, for every  $\mathbf{u}, \mathbf{v} \in \mathbb{T}(\mathbb{R}^d)$ ,*

$$\langle \sigma(X), \mathbf{u} \rangle \langle \sigma(X), \mathbf{v} \rangle = \langle \sigma(X), \mathbf{u} \sqcup \mathbf{v} \rangle.$$

Another important result, known as *Chen's relation*, describes the signature when we concatenate paths. Let us see how the concatenation path is defined.

**Definition 2.5.** *Let  $X, Y : [0, L] \rightarrow \mathbb{R}^d$  be two piecewise continuously differentiable paths. We define the concatenation of  $X$  and  $Y$  as the path  $X \sqcup Y : [0, 2L] \rightarrow \mathbb{R}^d$  given by  $X$  on  $[0, L]$  and by  $Y_{-L} - Y_0 + X_L$  on  $[L, 2L]$  (i.e. take  $Y$ , move it back to 0 and then move it to the end of  $X$ ).*

The concatenation product interplays nicely with the concatenation of paths, as the following proposition shows.

**Proposition 2.6** (Chen's identity, [Che57]). *Let  $X, Y : [0, L] \rightarrow \mathbb{R}^d$  be two piecewise continuously differentiable paths and consider their concatenation  $X \sqcup Y : [0, 2L] \rightarrow \mathbb{R}^d$ . Then,*

$$\sigma(X \sqcup Y) = \sigma(X) \bullet \sigma(Y).$$

We finish this section with an example on how to compute the first terms of the signature of a path.

**Example 2.7.** *Consider the path  $X : [0, 1] \rightarrow \mathbb{R}^2$  given by  $X^1(t) = t$  and  $X^2(t) = t^2$ . We compute a few terms of its signature.*

$$\begin{aligned} \langle \sigma(X), 1 \rangle &= \int_0^1 dX_{r_1}^1 = \int_0^1 1 dt = 1 & \langle \sigma(X), 2 \rangle &= \int_0^1 dX_{r_1}^2 = \int_0^1 2t dt = 1 \\ \langle \sigma(X), 11 \rangle &= \int_0^1 \int_0^{r_2} dX_{r_1}^1 dX_{r_2}^1 = \int_0^1 r_2 dX_{r_2}^1 = \int_0^1 r_2 dr_2 = \frac{1}{2} \\ \langle \sigma(X), 12 \rangle &= \int_0^1 \int_0^{r_2} dX_{r_1}^1 dX_{r_2}^2 = \int_0^1 r_2 dX_{r_2}^2 = \int_0^1 2r_2^2 dr_2 = \frac{2}{3} \\ \langle \sigma(X), 21 \rangle &= \int_0^1 \int_0^{r_2} dX_{r_1}^2 dX_{r_2}^1 = \int_0^1 r_2^2 dX_{r_2}^1 = \int_0^1 r_2^2 dr_2 = \frac{1}{3} \\ \langle \sigma(X), 22 \rangle &= \int_0^1 \int_0^{r_2} dX_{r_1}^2 dX_{r_2}^2 = \int_0^1 r_2^2 dX_{r_2}^2 = \int_0^1 2r_2^3 dr_2 = \frac{2}{4} = \frac{1}{2} \\ \langle \sigma(X), 222 \rangle &= \int_0^1 \int_0^{r_3} \int_0^{r_2} dX_{r_1}^2 dX_{r_2}^2 dX_{r_3}^2 = \int_0^1 \int_0^{r_3} r_2^2 dX_{r_2}^2 dX_{r_3}^2 = \int_0^1 \int_0^{r_3} 2r_2^3 dr_2 dX_{r_3}^2 = \\ &= \int_0^1 \frac{r_3^4}{2} dX_{r_3}^2 = \int_0^1 r_3^5 dr_3 = \frac{r_3^6}{6} \Big|_0^1 = \frac{1}{6} \end{aligned}$$

Therefore, the signature of  $X$  is of the form

$$\sigma(X) = 1 + 2 + \frac{1}{2} \cdot (11 + 22) + \frac{1}{3} \cdot (2 \cdot 12 + 21) + \frac{1}{6} \cdot 222 + \dots$$

### 3 Signatures under the action of polynomial maps

Let  $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$  be a polynomial map given by the polynomials  $p_i(x_1, x_2, \dots, x_d)$ , for  $i = 1, 2, \dots, m$ , with the property that  $p(0) = 0$ . The *degree of the polynomial map*  $p$  is the maximum of the degree of the polynomials that define it,  $\deg(p) = \max_i \deg(p_i)$ . Moreover, we say that a polynomial map is *homogeneous* if the polynomials  $p_i$  are homogeneous of the same degree. Finally, we denote by  $J_p$  the Jacobian matrix of format  $m \times d$  with entries  $J_p^{ij} = \partial_j p_i$ , for  $i \in \{1, 2, \dots, m\}$  and  $j \in \{1, 2, \dots, d\}$ .

Recall the algebra homomorphism  $\varphi_d$  defined by:

$$\begin{aligned} \varphi_d : \mathbb{R}[x_1, x_2, \dots, x_d] &\longrightarrow (\mathbb{T}(\mathbb{R}^d), \sqcup) \\ x_i &\longmapsto \mathbf{i} \\ x_{i_1} \cdots x_{i_l} &\longmapsto \mathbf{i}_1 \sqcup \cdots \sqcup \mathbf{i}_l \end{aligned}$$

As we mention above, by the properties of the shuffle product in  $\mathbb{T}(\mathbb{R}^d)$ , this map is injective but is not surjective.

We now consider maps  $M_p$ ,  $\varphi$ , and  $p^*$  such that they complete the diagram (3) that arises in our main question in the following way:

$$\begin{array}{ccc} \mathbb{R}[x_1, \dots, x_m] & \xrightarrow{p^*} & \mathbb{R}[x_1, \dots, x_d] \\ \varphi_m \downarrow & & \downarrow \varphi_d \\ (\mathbb{T}(\mathbb{R}^m), \sqcup) & \xrightarrow{\exists M_p} & (\mathbb{T}(\mathbb{R}^d), \sqcup) \end{array}$$

Notice that the map  $M_p$  is unique when restricted to the image of  $\varphi_m$ , but not the full  $M_p$  on the tensor algebra  $\mathbb{T}(\mathbb{R}^m)$ .

**Definition 3.1.** For any polynomial map  $p : \mathbb{R}^d \rightarrow \mathbb{R}^m$  such that  $p(0) = 0$ , let  $k_p^{ij} = \varphi_d(J_p^{ij}) \in \mathbb{T}(\mathbb{R}^d)$ , where  $J_p^{ij}$  is the  $(i, j)$ -entry of the Jacobian matrix of  $p$ . We define the map  $M_p : \mathbb{T}(\mathbb{R}^m) \longrightarrow \mathbb{T}(\mathbb{R}^d)$  recursively as follows:

$M_p(\mathbf{e}) = \mathbf{e}$ , for  $\mathbf{e}$  the empty word, and

$$M_p(\mathbf{wi}) = \sum_{j=1}^d \left( M_p(\mathbf{w}) \sqcup k_p^{ij} \right) \bullet \mathbf{j}, \text{ for any word } \mathbf{w} \text{ and any letter } \mathbf{i} \in \{1 \dots, m\}.$$

The following result summarizes a few properties of the map  $M_p$ . We will use these properties to show that the map that  $M_p$  as we construct it restricts according to what we need.

**Proposition 3.2.** Consider two polynomial maps  $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$  and  $q : \mathbb{R}^m \longrightarrow \mathbb{R}^s$ , with  $p(0) = 0$  and  $q(0) = 0$ , and the algebra homomorphisms  $\varphi_m$  and  $\varphi_d$ . Then, we have the following list of properties:

- (I)  $M_p : (\mathbb{T}(\mathbb{R}^m), \sqcup) \longrightarrow (\mathbb{T}(\mathbb{R}^d), \sqcup)$  is an algebra homomorphism.
- (II) For  $i = 1, \dots, s$ ,  $M_p(\varphi_m(q_i)) = \varphi_d(q_i \circ p)$ , where the  $q_i$ 's are the polynomials defining the polynomial map  $q$ .
- (III)  $k_{q \circ p}^{ij} = \sum_{l=1}^m M_p(k_q^{il}) \sqcup k_p^{lj}$ .

(IV)  $M_{q \circ p} = M_p M_q$ .

(V) If  $p$  is a polynomial map of degree  $n$ , then  $M_p(\mathbb{T}^k(\mathbb{R}^m)) \subseteq \mathbb{T}^{\leq nk}(\mathbb{R}^d)$ .

(VI) If  $p$  is an homogeneous polynomial map of degree  $n$ ,  $M_p(\mathbb{T}^k(\mathbb{R}^m)) \subseteq \mathbb{T}^{nk}(\mathbb{R}^d)$

*Proof.* (I) We need to show that, for any words  $\mathbf{w}_1$  and  $\mathbf{w}_2$  in  $\mathbb{T}(\mathbb{R}^m)$ ,

$$M_p(\mathbf{w}_1 \sqcup \mathbf{w}_2) = M_p(\mathbf{w}_1) \sqcup M_p(\mathbf{w}_2).$$

We proceed by induction on  $\ell(\mathbf{w}_1) + \ell(\mathbf{w}_2)$ . For  $\ell(\mathbf{w}_1) + \ell(\mathbf{w}_2) \leq 1$ , at least one of the two words is the empty word  $\mathbf{e}$ , and so we assume that  $\mathbf{w}_2 = \mathbf{e}$ . Therefore,

$$M_p(\mathbf{w}_1 \sqcup \mathbf{w}_2) = M_p(\mathbf{w}_1 \sqcup \mathbf{e}) = M_p(\mathbf{w}_1) = M_p(\mathbf{w}_1) \sqcup \mathbf{e} = M_p(\mathbf{w}_1) \sqcup M_p(\mathbf{w}_2).$$

Assume now that the statement is true for any pair of words with sum of lengths at most  $n - 1$ . Let  $\mathbf{u}$  and  $\mathbf{v}$  be two words such that  $\ell(\mathbf{u}) + \ell(\mathbf{v}) = n - 1$  and  $\mathbf{a}$  and  $\mathbf{b}$  two arbitrary letters. Then, using Definitions 2.2 and 3.1, and the inductive hypothesis (IH),

$$\begin{aligned} M_p(\mathbf{ua} \sqcup \mathbf{vb}) &\stackrel{Def.}{=} M_p((\mathbf{u} \sqcup \mathbf{vb}) \bullet \mathbf{a} + (\mathbf{ua} \sqcup \mathbf{v}) \bullet \mathbf{b}) \stackrel{Def.}{=} \\ &\sum_{i=1}^d [M_p(\mathbf{u} \sqcup \mathbf{vb}) \sqcup k_p^{ai} + M_p(\mathbf{ua} \sqcup \mathbf{v}) \sqcup k_p^{bi}] \bullet \mathbf{i} \stackrel{Def.}{=} \\ &\sum_{i=1}^d \sum_{j=1}^d [M_p(\mathbf{u}) \sqcup ((M_p(\mathbf{v}) \sqcup k_p^{bj}) \bullet \mathbf{i}) \sqcup k_p^{ai} + ((M_p(\mathbf{u}) \sqcup k_p^{aj}) \bullet \mathbf{j}) \sqcup M_p(\mathbf{v}) \sqcup k_p^{bi}] \bullet \mathbf{i} \stackrel{(IH)}{=} \\ &\left( \sum_{i=1}^d (M_p(\mathbf{u}) \sqcup k_p^{ai}) \bullet \mathbf{i} \right) \sqcup \left( \sum_{i=1}^d (M_p(\mathbf{v}) \sqcup k_p^{bi}) \bullet \mathbf{i} \right) \stackrel{Def.}{=} M_p(\mathbf{ua}) \sqcup M_p(\mathbf{vb}). \end{aligned}$$

(II) Assume that  $M_p(\mathbf{i}) = \varphi_d(p_i)$ , for all  $i$ . Then, for the monomial  $h(x_1, \dots, x_m) = x_1^{n_1} \cdot x_2^{n_2} \cdots x_m^{n_m}$ ,  $\varphi_m(h) = \mathbf{1}^{\sqcup n_1} \sqcup \cdots \sqcup \mathbf{m}^{\sqcup n_m}$ . Therefore, by the property (I),

$$\begin{aligned} M_p(\varphi_m(h)) &= M_p(\mathbf{1})^{\sqcup n_1} \sqcup \cdots \sqcup M_p(\mathbf{m})^{\sqcup n_m} = \varphi_d(p_1)^{\sqcup n_1} \sqcup \cdots \sqcup \varphi_d(p_m)^{\sqcup n_m} = \\ &\varphi_d(p_1^{n_1} \cdots p_m^{n_m}) = \varphi_d(h \circ p), \end{aligned}$$

and the property (II) follows by linearity.

We prove now the claim  $M_p(\mathbf{i}) = \varphi_d(p_i)$ , for all  $i$ . Since the two maps  $p \mapsto M_p(\mathbf{i})$  and  $p \mapsto \varphi_d(p_i)$  are linear, it is enough to prove the claim for the case when  $p_i$  is a monomial of the form  $p_i = x_1^{n_1} \cdot x_2^{n_2} \cdots x_d^{n_d}$ , with at least one of the  $n_i$ 's non-zero. In this case,

$$\begin{aligned} M_p(\mathbf{i}) &= \sum_{j=1}^d (M_p(\mathbf{e}) \sqcup k_p^{ij}) \bullet \mathbf{j} = \sum_{j=1}^d k_p^{ij} \bullet \mathbf{j} \stackrel{*}{=} \\ &\sum_{\substack{j=1 \\ n_j \neq 0}}^d n_j \left( \mathbf{1}^{\sqcup(n_1 - \delta_{1j})} \sqcup \cdots \sqcup \mathbf{d}^{\sqcup(n_d - \delta_{dj})} \right) \bullet \mathbf{j} = \mathbf{1}^{\sqcup n_1} \sqcup \cdots \sqcup \mathbf{d}^{\sqcup n_d} = \varphi_d(p_i), \end{aligned}$$

where  $*$  follows by applying enough iterations of the recursive definition of the shuffle product and the fact that  $\mathbf{i}^{\sqcup n+1} = (n+1)\mathbf{i}^{\sqcup n} \bullet \mathbf{i}$ .

(III) By the *chain rule*,  $J_{q \circ p}^{ij} = \sum_{l=1}^m (J_q^{il} \circ p) \cdot J_p^{lj}$ . For one term of that sum, by (II),

$$\varphi_d(J_q^{il} \circ p) = M_p(\varphi_m(J_q^{il})) = M_p(k_q^{il}). \text{ Thus,}$$

$$k_{q \circ p}^{ij} = \varphi_d(J_{q \circ p}^{ij}) = \varphi_d\left(\sum_{l=1}^m (J_q^{il} \circ p) \cdot J_p^{lj}\right) = \sum_{l=1}^m \varphi_d(J_q^{il} \circ p) \sqcup \varphi_d(J_p^{lj}) = \sum_{l=1}^m M_p(k_q^{il}) \sqcup k_p^{lj}.$$

(IV) We proceed by induction on  $\ell(\mathbf{w})$ . For a letter  $\mathbf{i}$ ,

$$\begin{aligned} M_p \circ M_q(\mathbf{i}) &= M_p(M_q(\mathbf{i})) = M_p\left(\sum_{j=1}^m k_q^{ij} \cdot \mathbf{j}\right) = \sum_{j=1}^m M_p(k_q^{ij} \cdot \mathbf{j}) = \\ &\sum_{j=1}^m \left(\sum_{l=1}^d M_p(k_q^{ij} \sqcup k_p^{jl})\right) \cdot \mathbf{1} = \sum_{l=1}^d \left(\sum_{j=1}^m M_p(k_q^{ij} \sqcup k_p^{jl})\right) \cdot \mathbf{1} \stackrel{\text{(III)}}{=} \sum_{l=1}^d k_{q \circ p}^{il} \cdot \mathbf{1} = M_{q \circ p}(\mathbf{i}). \end{aligned}$$

Now, we assume that the statement is true for all the words of length at most  $n$  and we refer to it as (IH). Let  $\mathbf{w}$  be one of these words and  $\mathbf{i}$  any letter. Then,

$$\begin{aligned} M_p \circ M_q(\mathbf{wi}) &= M_p(M_q(\mathbf{wi})) = M_p\left(\left(\sum_{j=1}^m M_q(\mathbf{w}) \sqcup k_q^{ij}\right) \cdot \mathbf{j}\right) = \\ &\sum_{j=1}^m M_p\left((M_q(\mathbf{w}) \sqcup k_q^{ij}) \cdot \mathbf{j}\right) = \sum_{j=1}^m \sum_{l=1}^d [M_p(M_q(\mathbf{w}) \sqcup k_q^{ij}) \sqcup k_p^{jl}] \cdot \mathbf{1} \stackrel{\text{(I)}}{=} \\ &\sum_{j=1}^m \sum_{l=1}^d [M_p(M_q(\mathbf{w})) \sqcup M_p(k_q^{ij}) \sqcup k_p^{jl}] \cdot \mathbf{1} \stackrel{\text{(IH)}}{=} \sum_{j=1}^m \sum_{l=1}^d [M_{q \circ p}(\mathbf{w}) \sqcup M_p(k_q^{ij}) \sqcup k_p^{jl}] \cdot \mathbf{1} = \\ &\sum_{l=1}^d \left[M_{q \circ p}(\mathbf{w}) \sqcup \left(\sum_{j=1}^m M_p(k_q^{ij}) \sqcup k_p^{jl}\right)\right] \cdot \mathbf{1} \stackrel{\text{(III)}}{=} \sum_{l=1}^d [M_{q \circ p}(\mathbf{w}) \sqcup k_{q \circ p}^{il}] \cdot \mathbf{1} = M_{q \circ p}(\mathbf{wi}). \end{aligned}$$

(V) We start by noticing that since the polynomial map  $p$  has degree  $n$ , then  $\deg(p_i) \leq n$ , for all  $i$ . Thus,  $\deg(J_p^{ij}) \leq n - 1$  and  $\ell(\varphi_d(J_p^{ij})) \leq n - 1$ , for all  $i$  and  $j$ .

Now, we proceed by induction on  $k$ . For  $k = 1$ ,  $\mathbb{T}^1(\mathbb{R}^m)$  is the set of letters  $\{\mathbf{1}, \dots, \mathbf{m}\}$ .

Since for a letter  $\mathbf{i}$  in this set  $M_p(\mathbf{i}) = \sum_{j=1}^d k_p^{ij} \cdot \mathbf{j}$ , then  $\ell(M_p(\mathbf{i})) = \max_j \{\ell(k_p^{ij} \cdot \mathbf{j})\} \leq n$ .

Thus,  $M_p(\mathbf{i}) \in \mathbb{T}^{\leq n}(\mathbb{R}^d)$ .

Assume that the statement is true for  $k$ . Any word  $\mathbf{w}' \in \mathbb{T}^{k+1}(\mathbb{R}^m)$  can be written as  $\mathbf{w}' = \mathbf{w} \cdot \mathbf{i}$ , with  $\mathbf{w} \in \mathbb{T}^k(\mathbb{R}^m)$  and  $\mathbf{i}$  a letter. We analyze the length of  $M_p(\mathbf{w} \cdot \mathbf{i})$ .

Since  $M_p(\mathbf{w} \cdot \mathbf{i}) = \sum_{j=1}^d (M_p(\mathbf{w}) \sqcup k_p^{ij}) \cdot \mathbf{j}$ , it is enough to upper bound the length of the terms appearing in the sum. By the inductive hypothesis,  $\ell(M_p(\mathbf{w})) \leq nk$ , and since  $\ell(k_p^{ij}) \leq n - 1$ ,  $\ell(M_p(\mathbf{w}) \sqcup k_p^{ij}) \leq nk + n - 1$ . Therefore,  $\ell(M_p(\mathbf{w} \cdot \mathbf{i})) \leq nk + n - 1 + 1 = n(k + 1)$ .

(VI) In this case, since  $p$  is homogeneous of degree  $n$ , then  $\deg(p_i) = n$ , for all  $i$ . Moreover,  $\deg(J_p^{ij}) = n - 1$ , if the variable  $x_j$  appears in  $p_i$ , or zero, otherwise.



We proceed by induction on  $k$ . For  $k = 1$ , let  $\mathbf{i}$  be a letter in  $\{1, \dots, \mathbf{m}\}$ . Then,  $M_p(\mathbf{i}) = \sum_{j=1}^d k_p^{ij} \bullet \mathbf{j}$ . This sum contains only terms  $k_p^{ij} \bullet \mathbf{j}$ , which has length exactly  $n$ , otherwise  $k_p^{ij}$  is zero according to our observation about the Jacobian entries above. Therefore, the statement follows.

Now, assume the statement is true for  $k$ . Let  $\mathbf{w} \in \mathbb{T}^k(\mathbb{R}^m)$  be a word and  $\mathbf{i}$  a letter. In this case,  $M_p(\mathbf{w} \bullet \mathbf{i}) = \sum_{j=1}^d (M_p(\mathbf{w}) \sqcup k_p^{ij}) \bullet \mathbf{j}$ . Again, the terms appearing in this sum have length  $nk + n - 1 + 1 = n(k + 1)$ , which concludes the proof.  $\square$

Once we have these properties, we recall Theorem 1.1 and prove it.

**Theorem 1.1.** *There exists an algebra homomorphism  $M_p : (\mathbb{T}(\mathbb{R}^m), \sqcup) \longrightarrow (\mathbb{T}(\mathbb{R}^d), \sqcup)$  such that its restriction  $M_p|_{Im(\varphi_m)}$  is the unique algebra homomorphism that makes the following diagram commute:*

$$\begin{array}{ccc} \mathbb{R}[x_1, \dots, x_m] & \xrightarrow{p^*} & \mathbb{R}[x_1, \dots, x_d] \\ \varphi_m \downarrow & & \downarrow \varphi_d \\ Im(\varphi_m) & \xrightarrow{\exists! M_p|_{Im(\varphi_m)}} & Im(\varphi_d) \end{array}$$

*Proof.* By (I) in Proposition 3.2, the restriction of  $M_p$  to the image  $Im(\varphi_m)$  is an algebra homomorphism. Moreover, due to (II), we have that  $M_p(Im(\varphi_m)) \subseteq Im(\varphi_d)$ . Since we restrict to their images,  $\varphi_m$  and  $\varphi_d$  are isomorphisms and the map  $M_p$  is the unique one making the diagram commute.  $\square$

Let us see now the answer to our main question, which is stated as Theorem 1.2

**Theorem 1.2.** *Let  $X : [0, L] \longrightarrow \mathbb{R}^d$  be a piecewise continuously differentiable path with  $X_0 = 0$  and let  $p : \mathbb{R}^d \longrightarrow \mathbb{R}^m$  be a polynomial map with  $p(0) = 0$ . Then, for all  $\mathbf{w} \in \mathbb{T}(\mathbb{R}^m)$ ,*

$$\langle \sigma(p(X)), \mathbf{w} \rangle = \langle \sigma(X), M_p(\mathbf{w}) \rangle.$$

*Equivalently,  $\sigma(p(X)) = M_p^*(\sigma(X))$ .*

*Proof.* Denote by  $Y = p(X)$  and by  $\dot{Y} = \sum_{j=1}^d J_p^{ij}(X) \cdot \dot{X}^j$ . Notice that each component  $X^j$  of the path equals the first signature component,  $\sigma^1(X) = X$ , and therefore, the entries of the Jacobian matrix can be seen as coefficients of the signature of  $X$ ,

$$J_p^{ij}(X_t) = \langle \sigma(X|_{[0,t]}), k_p^{ij} \rangle. \quad (4)$$

We proceed by induction on the length of the word  $\mathbf{w}$ . For a letter  $\mathbf{i}$ , we have that

$$\begin{aligned} \langle \sigma(Y), \mathbf{i} \rangle &= \int_0^L \langle \sigma(Y|_{[0,t]}), \mathbf{e} \rangle d\dot{Y}_t^i = \int_0^L d\dot{Y}_t^i = \sum_{j=1}^d \int_0^L J_p^{ij}(X_t) \dot{X}_t^j dt = \sum_{j=1}^d \int_0^L J_p^{ij}(X_t) dX_t^j \stackrel{(4)}{=} \\ &= \sum_{j=1}^d \int_0^L \langle \sigma(X|_{[0,t]}), k_p^{ij} \rangle dX_t^j = \sum_{j=1}^d \langle \sigma(X), k_p^{ij} \bullet \mathbf{j} \rangle = \langle \sigma(X), M_p(\mathbf{i}) \rangle. \end{aligned}$$

Now, assume that the statement is true for all the words of length at most  $n$ . Let  $\mathbf{w}$  be any of these words and  $\mathbf{i}$  any letter. By the definition of the signature,

$$\begin{aligned} \langle \sigma(Y), \mathbf{wi} \rangle &= \int_0^L \langle \sigma(Y|_{[0,t]}), \mathbf{w} \rangle dY_t^i = \int_0^L \langle \sigma(Y|_{[0,t]}), \mathbf{w} \rangle \dot{Y}_t^i dt = \\ &= \sum_{j=1}^d \int_0^L \langle \sigma(Y|_{[0,t]}), \mathbf{w} \rangle J_p^{ij}(X_t) \dot{X}_t^j dt = \sum_{j=1}^d \int_0^L \langle \sigma(Y|_{[0,t]}), \mathbf{w} \rangle J_p^{ij}(X_t) dX_t^j \stackrel{(4)}{=} \\ &= \sum_{j=1}^d \int_0^L \langle \sigma(Y|_{[0,t]}), \mathbf{w} \rangle \langle \sigma(X|_{[0,t]}), k_p^{ij} \rangle dX_t^j. \quad (5) \end{aligned}$$

Now, apply the inductive hypothesis to  $\langle \sigma(Y|_{[0,t]}), \mathbf{w} \rangle$  in (5), and then by Chen's identity, Proposition 2.6,

$$\begin{aligned} \langle \sigma(Y), \mathbf{wi} \rangle &= \sum_{j=1}^d \int_0^L \langle \sigma(X|_{[0,t]}), M_p(\mathbf{w}) \rangle \langle \sigma(X|_{[0,t]}), k_p^{ij} \rangle dX_t^j = \\ &= \sum_{j=1}^d \int_0^L \langle \sigma(X|_{[0,t]}), M_p(\mathbf{w}) \sqcup k_p^{ij} \rangle dX_t^j = \sum_{j=1}^d \langle \sigma(X_t), (M_p(\mathbf{w}) \sqcup k_p^{ij}) \bullet \mathbf{j} \rangle = \langle \sigma(X), M_p(\mathbf{wi}) \rangle. \end{aligned}$$

□

We finish this section with a generalization of Theorem 1.2 to polynomial maps that do not satisfy the condition  $p(0) = 0$  and paths that do not start at the origin.

**Corollary 3.3.** *Let  $X : [0, L] \rightarrow \mathbb{R}^d$  be a piecewise continuously differentiable path and let  $p : \mathbb{R}^d \rightarrow \mathbb{R}^m$  be a polynomial map. Consider the map  $\tilde{p}$  given by  $\tilde{p}(y) = p(y + X_0) - p(X_0)$ . Then, for all  $\mathbf{w} \in \mathbb{T}(\mathbb{R}^m)$ ,*

$$\langle \sigma(p(X)), \mathbf{w} \rangle = \langle \sigma(X), M_{\tilde{p}}(\mathbf{w}) \rangle.$$

*Proof.* The statement follows using the same argument as in the proof of Theorem 1.2 if we take into account that in this case, at the end of the step (4),

$$J_p^{ij}(X_t) \dot{X}_t^j dt = J_{\tilde{p}}^{ij}(X_t - X_0) dX_t^j = \langle \sigma(X|_{[0,t]}), k_{\tilde{p}}^{ij} \rangle.$$

□

### 3.1 $M_p$ as a half-shuffle homomorphism

The shuffle product can be seen as the symmetrization of the right half-shuffle, which we define in the following way. Let  $\mathbb{T}^{\geq 1}(\mathbb{R}^d) = \bigoplus_{n \geq 1} \mathbb{T}^n(\mathbb{R}^d)$  denote the vector space spanned by the non-empty words built from  $d$  letters.

**Definition 3.4.** *[Sch58, Eq. (S2)], [FP13, Def. 1], [EM53, Sect. 18] The right half-shuffle  $\succ : \mathbb{T}^{\geq 1}(\mathbb{R}^d) \times \mathbb{T}^{\geq 1}(\mathbb{R}^d) \rightarrow \mathbb{T}^{\geq 1}(\mathbb{R}^d)$  is recursively given on words as*

$$\begin{aligned} \mathbf{w} \succ \mathbf{i} &:= \mathbf{wi}, \\ \mathbf{w} \succ \mathbf{vi} &:= (\mathbf{w} \succ \mathbf{v} + \mathbf{v} \succ \mathbf{w}) \bullet \mathbf{i}, \end{aligned}$$

where  $\mathbf{w}, \mathbf{v}$  are words and  $\mathbf{i}$  is a letter.

Therefore, for any non-empty words  $w, v$

$$w \sqcup v = w > v + v > w.$$

Indeed, for non-empty words  $w, v$  and letters  $i, j$ , we have

$$\begin{aligned} i > j + j > i &= ij + ji, \\ wi > j + j > wi &= wij + (j > w + w > j) \bullet i, \quad \text{and} \\ wi > vj + vj > wi &= (wi > v + v > wi) \bullet j + (vj > w + w > vj) \bullet i, \end{aligned}$$

in accordance with Definition 2.2. Thus, the second equation in Definition 3.4 can be rewritten as

$$w > vi = (w \sqcup v) \bullet i. \quad (6)$$

It turns out that the right half-shuffle is an example of a more general type of algebras, the Zinbiel algebras.

**Definition 3.5.** [Sch58, Eq. (S0)], [FP13, Eq. (4)] A (right) Zinbiel algebra is a vector space  $Z$  together with a bilinear map  $\succ : Z \times Z \rightarrow Z$  such that, for all  $a, b, c \in Z$ ,

$$a \succ (b \succ c) = (a \succ b + b \succ a) \succ c.$$

We include here the proof of the next result since it is interesting.

**Theorem 3.6.** [Lod95, Proposition 1.8]  $(T^{\geq 1}(\mathbb{R}^d), >)$  is a Zinbiel algebra, i.e. for any non-empty words  $w, v$ , and  $u$ ,

$$w > (v > u) = (w \sqcup v) > u. \quad (7)$$

*Proof.* Let  $w, v$ , and  $u$  be non-empty words and  $i$  be an arbitrary letter. By the definition of the half-shuffle and Equation (6), we have

$$w > (v > i) = w > vi = (w \sqcup v) \bullet i = (w \sqcup v) > i.$$

Using Equation (6) and associativity of the shuffle product, Proposition 2.3, we obtain that

$$w > (v > ui) = w > ((v \sqcup u) \bullet i) = (w \sqcup (v \sqcup u)) \bullet i = ((w \sqcup v) \sqcup u) \bullet i = (w \sqcup v) > ui.$$

□

In fact, it is known that  $(T^{\geq 1}(\mathbb{R}^d), >)$  is free in this case.

**Theorem 3.7.** [Sch58, page 19][Lod95, Proposition 1.8] Indeed,  $(T^{\geq 1}(\mathbb{R}^d), >)$  is the free Zinbiel algebra over  $\mathbb{R}^d$ .

This means that for any Zinbiel algebra  $(Z, \succ)$  and any linear map  $B : \mathbb{R}^d \rightarrow Z$ , there is a unique homomorphism  $\Lambda_B : (T^{\geq 1}(\mathbb{R}^d), >) \rightarrow (Z, \succ)$  such that  $B = \Lambda_B \circ \iota$ , where  $\iota : \mathbb{R}^d \rightarrow T^{\geq 1}(\mathbb{R}^d)$  is the canonical embedding. This is known as the *universal property of the free Zinbiel algebra* and is described in Diagram 3.1. We call  $\Lambda_B$  the *unique extension of  $B$  to a Zinbiel homomorphism*.

$$\begin{array}{ccc}
\mathbb{R}^d & \xrightarrow{\iota} & (\mathbb{T}^{\geq 1}(\mathbb{R}^d), \succ) \\
& \searrow B & \downarrow \Lambda_B \\
& & (Z, \succ)
\end{array}$$

Diagram 3.1: Universal property of the free Zinbiel algebra

*Proof.* Define the linear map  $\Lambda_B : \mathbb{T}^{\geq 1}(\mathbb{R}^d) \rightarrow Z$  recursively by

$$\Lambda_B \mathbf{i} := L\mathbf{i}, \quad \Lambda_B \mathbf{v}\mathbf{i} := \Lambda_B \mathbf{v} \succ \Lambda_B \mathbf{i},$$

where we identified  $\mathbb{R}^d$  with the letters in  $\mathbb{T}^{\geq 1}(\mathbb{R}^d)$ . Since  $\mathbf{v}\mathbf{i} = \mathbf{v} \succ \mathbf{i}$ , this is the only candidate for a map with the desired properties. It remains to show that it is indeed a homomorphism of Zinbiel algebras.

By definition, it holds that  $\Lambda_B \mathbf{i}\mathbf{j} = \Lambda_B \mathbf{i} \succ \Lambda_B \mathbf{j}$ . Thus, assume that  $\Lambda_B \mathbf{x}\mathbf{y} = \Lambda_B \mathbf{x} \succ \Lambda_B \mathbf{y}$  holds for all nonempty words  $\mathbf{x}, \mathbf{y}$  such that  $|\mathbf{x}| + |\mathbf{y}| = n$ . Then, for all nonempty words  $\mathbf{w}$  and  $\mathbf{v}$  such that  $|\mathbf{w}| + |\mathbf{v}| = n$ , we have

$$\begin{aligned}
\Lambda_B \mathbf{w} \succ \Lambda_B \mathbf{v}\mathbf{i} &= \Lambda_B \mathbf{w} \succ (\Lambda_B \mathbf{v} \succ \Lambda_B \mathbf{i}) = (\Lambda_B \mathbf{w} \succ \Lambda_B \mathbf{v} + \Lambda_B \mathbf{v} \succ \Lambda_B \mathbf{w}) \succ \Lambda_B \mathbf{i} = \\
\Lambda_B (\mathbf{w} \succ \mathbf{v} + \mathbf{v} \succ \mathbf{w}) &\succ \Lambda_B \mathbf{i} = \Lambda_B ((\mathbf{w} \succ \mathbf{v} + \mathbf{v} \succ \mathbf{w}) \succ \mathbf{i}) = \Lambda_B (\mathbf{w} \succ (\mathbf{v} \succ \mathbf{i})) = \Lambda_B (\mathbf{w} \succ \mathbf{v}\mathbf{i}),
\end{aligned}$$

and  $\Lambda_B \mathbf{u}\mathbf{i} := \Lambda_B \mathbf{u} \succ \Lambda_B \mathbf{i}$ , for all nonempty words  $\mathbf{u}$  with  $|\mathbf{u}| = n$  by definition. The claim follows by induction over  $n$ .  $\square$

The following result describes the relation between the map  $M_p$  and the half-shuffle.

**Theorem 3.8.** *The restriction of  $M_p$  to  $\mathbb{T}^{\geq 1}(\mathbb{R}^d)$ ,  $M_p|_{\mathbb{T}^{\geq 1}(\mathbb{R}^d)}$ , is the unique half-shuffle homomorphism such that  $M_p(\mathbf{i}) = \varphi_d(p_i)$ .*

*Proof.* Using (6) and the definition of  $M_p$ , we get

$$M_p(\mathbf{w}\mathbf{i}) = \sum_{j=1}^d (M_p(\mathbf{w}) \sqcup k_p^{ij}) \bullet \mathbf{j} = M_p(\mathbf{w}) \succ \left( \sum_{j=1}^d k_p^{ij} \bullet \mathbf{j} \right) = M_p(\mathbf{w}) \succ M_p(\mathbf{i}),$$

and thus the statement follows immediately from the proof of Theorem 3.7.  $\square$

We finish this section with a generalization of Theorem 1.2 in terms of Zinbiel algebras stated in the following result.

**Theorem 3.9.** *Let  $X : [0, L] \rightarrow \mathbb{R}^d$  be a piecewise continuously differentiable path and  $B : \mathbb{R}^m \rightarrow (\mathbb{T}^{\geq 1}(\mathbb{R}^d), \succ)$  be a linear map. Then, the signature of the path*

$$Y : [0, L] \rightarrow \mathbb{R}^m, \quad Y_t^i := \langle \sigma(X|_{[0,t]}), B\mathbf{i} \rangle,$$

*is a linear transformation of the signature of  $X$ , namely*

$$\langle \sigma(Y), \mathbf{w} \rangle = \langle \sigma(X), \Lambda_B \mathbf{w} \rangle,$$

*where  $\Lambda_B$  is the unique extension of  $B$  to a Zinbiel homomorphism.*

Before proving this result, we introduce some notation. We denote by  $X_{0s}^z$  the coefficient of  $z$  in the signature of the path  $X$  restricted to the interval  $[0, s]$ . That is,

$$X_{0s}^z := \langle \sigma(X \upharpoonright_{[0,s]}), z \rangle.$$

Note that  $X_{0s}^i = X_s^i - X_0^i$ . Then, the path  $Y$  introduced in Theorem 3.9 is given by  $Y_s^i = X_{0s}^{B1}$ . Moreover, for any letter  $i$ , we define the maps  $T_i^-, T_i^+ : T(\mathbb{R}^d) \rightarrow T(\mathbb{R}^d)$  to be the unique linear maps given recursively by  $T_i^+ \mathbf{w} = \mathbf{wi}$  and by  $T_i^- \mathbf{wj} = \delta_{ij} \mathbf{w}$  with  $T_i^- \mathbf{e} = 0$ , respectively, for any word  $\mathbf{w}$ .

These two maps allows us to define the right half-shuffle as shown in the following technical result.

**Lemma 3.10.** *For any  $x, y \in T^{\geq 1}(\mathbb{R}^d)$ , we have  $x \succ y = \sum_{i=1}^d T_i^+(x \sqcup T_i^- y)$ .*

*Proof.* This is just a reformulation of (6) in the following way. For any word  $\mathbf{v}$ , any non-empty word  $\mathbf{w}$ , and any letter  $j$ , we have that

$$\sum_{i=1}^d T_i^+(\mathbf{w} \sqcup T_i^- \mathbf{vj}) = T_j^+(\mathbf{w} \sqcup \mathbf{v}) = (\mathbf{w} \sqcup \mathbf{v}) \bullet j = \mathbf{w} \succ \mathbf{vj},$$

where the last equality follows from (6). Then, the general statement for any  $x, y \in T^{\geq 1}(\mathbb{R}^d)$  follows from (bi)linearity.  $\square$

Now we are ready to prove Theorem 3.9.

*Proof of Theorem 3.9.* For better readability, we put  $\Lambda := \Lambda_B$ . First note that by the definition of the signature and the fact that  $X$  is continuously differentiable almost everywhere, we have

$$Y_s^i = X_{0s}^{B1} = X_{0s}^{\Lambda i} = \sum_{i=1}^d \int_0^s X_{0t}^{T_i^- \Lambda i} dX_t^i = \sum_{i=1}^d \int_0^s X_{0t}^{T_i^- \Lambda i} \dot{X}_t^i dt$$

and thus, for almost all  $s \in [0, T]$ ,

$$\dot{Y}_s^i = \sum_{i=1}^d X_{0s}^{T_i^- \Lambda i} \dot{X}_s^i.$$

Following an inductive argument, assume now that  $X_{0s}^{\Lambda \mathbf{w}} = Y_{0s}^{\mathbf{w}}$  holds for some word  $\mathbf{w}$ . Then,

$$\begin{aligned} Y_{0s}^{\mathbf{wi}} &= \int_0^s Y_{0t}^{\mathbf{w}} dY_t^i = \int_0^s Y_{0t}^{\mathbf{w}} \dot{Y}_t^i dt = \sum_{l=1}^d \int_0^s X_{0t}^{\Lambda \mathbf{w}} X_{0t}^{T_l^- \Lambda i} \dot{X}_t^l dt \\ &= \int_0^s X_{0t}^{\sum_{l=1}^d \Lambda \mathbf{w} \sqcup T_l^- \Lambda i} dX_{0t}^1 = X_{0t}^{\sum_{l=1}^d T_l^+(\Lambda \mathbf{w} \sqcup T_l^- \Lambda i)} = X_{0t}^{\Lambda \mathbf{w} \succ \Lambda i} = X_{0t}^{\Lambda \mathbf{wi}}, \end{aligned}$$

where we used Lemma 3.10 and the fact that  $\Lambda$  is a homomorphism of Zinbiel algebras.  $\square$

Theorem 3.9 can also be directly shown using

$$\int_0^s X_{0t}^x dX_{0t}^y = X_{0s}^{x \succ y} \quad (8)$$

for any  $x, y \in T^{\geq 1}(\mathbb{R}^d)$ , a relation which is quite fundamental for an algebraic understanding of the signature and was mentioned already for example in [GK08] right after equation (6). Conversely, starting from Theorem 3.9, equation (8) is immediate with the choice  $B1 = x, B2 = y$ .

The following example illustrates the results presented in this section.

**Example 3.11.** For a given path  $X : [0, L] \rightarrow \mathbb{R}^3$ , let

$$Y = \left( \text{Area}(X^2, X^3), \text{Area}(X^3, X^1), \text{Area}(X^1, X^2) \right)$$

denote the area path of  $X$ , where  $\text{Area}(X^i, X^j)_t = \int_0^t \int_0^s dX_u^i dX_s^j - \int_0^t \int_0^s dX_u^j dX_s^i$ . Let us compute  $\langle \sigma(Y), 12 - 21 \rangle$  in the special case that  $X : [0, 1] \rightarrow \mathbb{R}^3$ ,  $X(t) = (t, t^2, t^3)$ . We have  $Y_s^i = X_{0s}^{B1}$  where

$$B1 = 23 - 32, \quad B2 = 31 - 13, \quad B3 = 12 - 21$$

We need to compute  $\Lambda_B(12 - 21) = B1 \succ B2 - B2 \succ B1$ . To this end,

$$\begin{aligned} (23 - 32) \succ (31 - 13) &= 2 \cdot 2331 - 2 \cdot 3321 - 2313 - 2133 - 1233 + 3213 + 3123 + 1323, \\ (31 - 13) \succ (23 - 32) &= 3123 + 3213 + 2313 - 1323 - 1233 - 2133 - 3132 - 2 \cdot 3312 \\ &\quad + 2 \cdot 1332 + 3132. \end{aligned}$$

Thus,  $\Lambda_B(12 - 21) = 2 \cdot (-1323 + 1332 + 2313 - 2331 + 3312 - 3321)$ . Since in our special case of  $X(t) = (t, t^2, t^3)$  it holds that [AFS19, Example 2.2]

$$\int_0^1 \int_0^u \int_0^t \int_0^s dX_r^i dX_s^j dX_t^k dX_u^l = \frac{j \cdot k \cdot l}{(j+i)(k+j+i)(l+k+j+i)},$$

we get that

$$\begin{aligned} \langle \sigma(Y), 12 - 21 \rangle &= 2 \cdot \left( -\frac{3 \cdot 2 \cdot 3}{4 \cdot 6 \cdot 9} + \frac{3 \cdot 3 \cdot 2}{4 \cdot 7 \cdot 9} + \frac{3 \cdot 1 \cdot 3}{5 \cdot 6 \cdot 9} - \frac{3 \cdot 3 \cdot 1}{5 \cdot 8 \cdot 9} + \frac{3 \cdot 1 \cdot 2}{6 \cdot 7 \cdot 9} - \frac{3 \cdot 2 \cdot 1}{6 \cdot 8 \cdot 9} \right) \\ &= -\frac{1}{315} \approx -0.00317 \end{aligned}$$

as the desired result.

## 4 Examples and consequences

Let us start with an *easy example* to illustrate how Theorem 1.2 works, and also the property (V) in Proposition 3.2.

**Example 4.1.** Consider the polynomial map  $p : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  given by the polynomials  $p_1 = x^2$ ,  $p_2 = y^3$  and  $p_3 = x - y$ . Consider also the path in Example 2.7,  $X : [0, 1] \rightarrow \mathbb{R}^2$  given by  $X^1(t) = t$  and  $X^2(t) = t^2$ . We want to compute a few terms in the signature of the path  $p(X)$ .

We start computing the Jacobian matrix and its image under  $\varphi$ :

$$\left( J_p^{ij} \right)_{i,j} = \begin{pmatrix} 2x & 0 \\ 0 & 3y^2 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad \left( k_p^{ij} \right)_{i,j} = \left( \varphi \left( J_p^{ij} \right) \right)_{i,j} = \begin{pmatrix} 2 \cdot 1 & 0 \\ 0 & 6 \cdot 22 \\ \mathbf{e} & -\mathbf{e} \end{pmatrix}.$$

Notice that  $\varphi(3y^2) = 3 \cdot 2 \sqcup 2 = 6 \cdot 22$ . We use Definition 3.1 to compute the image of a few words:

$$\begin{aligned} M_p(1) &= 2 \cdot 11, & M_p(2) &= 6 \cdot 222, & M_p(3) &= 1 - 2, \\ M_p(33) &= (1 - 2) \bullet 1 - (1 - 2) \bullet 2 = 11 + 22 - 12 - 21. \end{aligned}$$

We observe that for any word  $\mathbf{w}$  above,  $\ell(M_p(\mathbf{w})) \leq 3 \cdot \ell(\mathbf{w})$ . This is due to property (V) in Proposition 3.2 since  $\deg(p) = 3$ . Now, applying Theorem 1.2 and looking at the signature terms computed in Example 2.7, we obtain that

$$\begin{aligned} \langle \sigma(p(X)), 1 \rangle &= \langle \sigma(X), M_p(1) \rangle = \langle \sigma(X), 2 \cdot 11 \rangle = 2 \cdot \frac{1}{2} = 1, \\ \langle \sigma(p(X)), 2 \rangle &= \langle \sigma(X), M_p(2) \rangle = \langle \sigma(X), 6 \cdot 222 \rangle = 6 \cdot \frac{1}{6} = 1, \\ \langle \sigma(p(X)), 3 \rangle &= \langle \sigma(X), M_p(3) \rangle = \langle \sigma(X), 1 - 2 \rangle = 1 - 1 = 0, \text{ and} \\ \langle \sigma(p(X)), 33 \rangle &= \langle \sigma(X), M_p(33) \rangle = \langle \sigma(X), 11 + 22 - 12 - 21 \rangle = 0. \end{aligned}$$

This second example is more generic and shows the property (VI) in Proposition 3.2.

**Example 4.2.** Let  $X$  be any piecewise continuously differentiable path in  $\mathbb{R}^2$ . Consider the polynomial map  $p: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  given by  $p(x, y) = (x^2, xy, y^2)$ , and fix  $k = 2$ .

By Theorem 1.2, the coefficient of  $\mathbf{w}$  in  $\sigma(p(X))$  is given by the coefficient of  $M_p(\mathbf{w})$  in  $\sigma(X)$ . Moreover, by Proposition 3.2 (VI), the words appearing in  $M_p(\mathbf{w})$  have length exactly 4. One way of storing  $\sigma^{(2)}(p(X))$  is using a matrix that encodes the coefficients in  $M_p(\mathbf{w})$ ,

$$\sigma^{(2)}(p(X)) \quad \leftrightarrow \quad \begin{bmatrix} 11 & 12 & 13 & 21 & 22 & 23 & 31 & 32 & 33 \\ 12 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 13 & 0 & 0 & 0 & 4 & 0 & 4 & 0 & 0 \\ 21 & 0 & 0 & 4 & 0 & 4 & 0 & 0 & 4 \\ 22 & 0 & 0 & 0 & 2 & 0 & 2 & 2 & 0 \\ 23 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ 31 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\ 32 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 33 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{matrix} \sigma^{(4)}(X) \\ 1111 \\ 1112 \\ 1121 \\ 1122 \\ 1211 \\ 1212 \\ 1221 \\ 1222 \\ 2111 \\ 2112 \\ 2121 \\ 2122 \\ 2211 \\ 2212 \\ 2221 \\ 2222 \end{matrix}.$$

More generally, by Proposition 3.2 (VI), for any word  $\mathbf{w}$  in the alphabet  $\{1, 2, 3\}$  with  $\ell(\mathbf{w}) = k$ ,  $M_p(\mathbf{w})$  is a sum of words in the alphabet  $\{1, 2\}$  of length  $2k$ . Applying Theorem 1.2, the information of  $\sigma^{(k)}(p(X))$  can be stored in terms of  $\sigma^{(2k)}(X)$ . In fact, there exists a matrix  $M$  of format  $3^k \times 2^{2k}$  that describes the change of coordinates in the following sense. Fix an order on the words, for instance, lexicographic order as above.

The coefficients of  $\sigma^{(k)}(p(X))$  and the coefficients of  $\sigma^{(2k)}(X)$  are related as

$$\sigma^{(k)}(p(X)) \quad \sigma^{(2k)}(X)$$

$$\begin{bmatrix} \overbrace{11\dots 1}^k \\ 1\dots 12 \\ \vdots \\ 3\dots 32 \\ \underbrace{3\dots 33}_k \end{bmatrix} = M \begin{bmatrix} \overbrace{11\dots 1}^{2k} \\ 1\dots 12 \\ \vdots \\ 2\dots 21 \\ \underbrace{2\dots 22}_{2k} \end{bmatrix},$$

where the row indexed by  $\mathbf{w}$  in  $M$  is given by the coefficients of  $M_p(\mathbf{w})$ . Moreover, we want to point out the following properties of this homomorphism  $M_p$ :

- For  $\mathbf{w} = \underbrace{1\dots 1}_k$ ,  $M_p(\mathbf{w}) = \frac{(2k)!}{k!} \cdot \mathbf{w} \bullet \mathbf{w}$ .
- For  $\mathbf{w} = \underbrace{2\dots 2}_k$ ,  $M_p(\mathbf{w}) = k! \cdot \underbrace{1\dots 1}_k \sqcup \underbrace{2\dots 2}_k$ .
- For  $\mathbf{w} = \underbrace{3\dots 3}_k$ ,  $M_p(\mathbf{w}) = \frac{(2k)!}{k!} \underbrace{2\dots 2}_{2k}$ .

The rest of this section is dedicated to analyze the consequences of Theorems 1.1 and 1.2 in several particular cases. We start looking at the case in which  $X$  is itself a polynomial map and we need the following definition.

**Definition 4.3.** For an element  $\mathbf{a} \in \mathbb{T}((\mathbb{R}^d))$ , with zero coefficient for the empty word  $\mathbf{e}$ , we define the concatenation product exponential of  $\mathbf{a}$  as

$$\exp_{\bullet}(\mathbf{a}) := \sum_{n \geq 0} \frac{\mathbf{a}^{\bullet n}}{n!}.$$

More information on this exponential map and its inverse, the *logarithm*, can be found in [Reu93].

**Corollary 4.4.** Let  $X : [0, L] \longrightarrow \mathbb{R}^d$  be a polynomial map, with  $L \in \mathbb{R}$ ,  $L \geq 1$ . Then, for any  $\mathbf{w} \in \mathbb{T}(\mathbb{R}^d)$ ,

$$\langle \sigma(X), \mathbf{w} \rangle = \langle \exp_{\bullet}(L \cdot 1), M_{\tilde{X}}(\mathbf{w}) \rangle,$$

where  $\tilde{X}(y) = X(y) - X_0$ . Equivalently,  $\sigma(X) = M_{\tilde{X}}^*(\exp_{\bullet}(L \cdot 1))$ .

*Proof.* Let  $Y : [0, L] \longrightarrow \mathbb{R}$  be the path given by  $Y(t) = t$ . Then,  $\sigma(Y) = \exp_{\bullet}(L \cdot 1)$ . The statement follows by Corollary 3.3 applied to the path  $Y$  and the polynomial map  $X$ .  $\square$

The next result looks at the case when  $M_p^*(\sigma(X)) = 0$  from the perspective of the polynomial map and of the piecewise continuously differentiable path. We introduce first two concepts. Given a polynomial map  $p$ , we define the *ideal generated by  $p$*  as the ideal  $I_p$  generated by the polynomials that define the map  $p$ , i.e.  $I_p = \langle p_1, \dots, p_m \rangle$ . Moreover, we define a *tree-like path* as a path  $X$  such that  $\sigma(X) = 0$ . This definition is the characterization given by B.M. Hambly and T. J. Lyons, [HL10], and a more general topological definition can be found in [BGLY16, Definition 1.1].



**Remark 4.5.** A very simple example of a tree-like path is a concatenation of paths  $A \sqcup B \sqcup C \sqcup D$  such that the paths  $A$  and  $B$  (resp.  $C$  and  $D$ ) are of the same shape, but parametrized in the opposite direction. When we compute the integrals on such a path, we get cancellations and the signature of the path does not see the  $A \sqcup B \sqcup C \sqcup D$  loop, i.e.  $\sigma(A \sqcup B \sqcup C \sqcup D) = \mathbf{e}$ .

The following result describes the situation in which the (image of the) path lies in the zeros of the ideal  $I_p$ , for some polynomial map  $p$ .

**Corollary 4.6.** Let  $p : \mathbb{R}^d \rightarrow \mathbb{R}^m$  be a polynomial map with associated ideal  $I_p$ . We define the polynomial map  $\tilde{p}(y) = p(y + X_0) - p(X_0)$ , for which  $\tilde{p}(0) = 0$ .

- If  $X : [0, L] \rightarrow \mathbb{R}^d$  is a piecewise continuously differentiable path such that  $X(t) \in \mathcal{V}(I_p)$  for all  $t \in [0, L]$ , then  $M_{\tilde{p}}^*(\sigma(X)) = \mathbf{e}$ .
- Conversely, if  $M_{\tilde{p}}^*(\sigma(X)) = \mathbf{e}$ , for some piecewise continuously differentiable path  $X : [0, L] \rightarrow \mathbb{R}^d$ , then  $p(X)$  is tree-like.

The last consequence is that the dual map  $M_p^*$  behaves nicely with respect to the concatenation of signatures.

**Corollary 4.7.** Let  $p : \mathbb{R}^d \rightarrow \mathbb{R}^m$  be a polynomial map with  $p(0) = 0$  and let  $X, Y : [0, L] \rightarrow \mathbb{R}^d$  be two piecewise continuously differentiable paths with  $X_0 = Y_0 = 0$ . Then,

$$M_p^*(\sigma(X) \bullet \sigma(Y)) = M_p^*(\sigma(X)) \bullet M_q^*(\sigma(Y)),$$

where  $q(y) = p(y + X_L - X_0) - p(X_L - X_0)$ .

## 5 Applications and future work

The results presented in this paper solve an algebraic question that arises from the signatures of paths, an object commonly studied in stochastic analysis. There are several interesting problems that do not fit on the algebraic flavour of this paper. We summarize them in the following list.

- (A) In comparison with the results presented in [PSS19], we would like to explore a non-linear version of their approach using dictionaries. The idea is that if we have a family of generic paths,  $\chi$ , for which we know the signature and a polynomial map  $p$ , then Theorem 1.2 allows us to compute the signature of all the paths in  $p(\chi)$ .

For instance, Example 4.2 shows that we can compute the signature of  $p(X)$  for any piecewise continuously differentiable path  $X$  by multiplying the signature of  $X$  by a matrix at each level. Therefore, we have the following question:

*Is it possible to understand  $\sigma(p(X))$  in terms of  $\sigma(X)$  and  $M_p$  in the language of tensors?*

- (B) Another line of future research is focused on the map  $M_p$ . Since it is defined from a polynomial map without involving any piecewise continuously differentiable path and gives us the commuting diagram (3), we intuit that it is worth to look for more interesting properties. For that, we should look to the *big picture* involving the Hopf algebra structures, as well as other constructions.

In this direction, at the end of Example 4.2 we describe combinatorially  $M_p(\mathbf{w})$  for some particular words  $\mathbf{w}$ . A more general question would be the following:

For which words  $w$  and polynomial maps  $p$  is there a *non-recursive* combinatorial formula for  $M_p(w)$ ?

Answering this question could be very useful from the computational perspective.

## Acknowledgements

The authors would like to thank Bernd Sturmfels, Francesco Galuppi, Joscha Diehl, Mateusz Michałek, and Max Pfeffer for their fruitful conversations and observations. The collaboration between the co-authors would not have been possible without the financial support from the research institute MPI-MiS Leipzig (Germany). R.P. is currently supported by European Research Council through CoG-683164.

## References

- [AFS19] C. Améndola, P. Friz, and B. Sturmfels. Varieties of signature tensors. *Forum of Mathematics, Sigma*, 7:e10, 2019.
- [Bal13] J. S. Balakrishnan. Iterated Coleman integration for hyperelliptic curves. In *ANTS X—Proceedings of the Tenth Algorithmic Number Theory Symposium*, volume 1 of *Open Book Ser.*, pages 41–61. Math. Sci. Publ., Berkeley, CA, 2013.
- [BGLY16] H. Boedihardjo, X. Geng, T. Lyons, and D. Yang. The signature of a rough path: Uniqueness. *Advances in Mathematics*, 293:720–737, April 2016.
- [Che57] K.-T. Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Ann. of Math. (2)*, 65:163–178, 1957.
- [DR19] J. Diehl and J. Reizenstein. Invariants of multidimensional time series based on their iterated-integral signature. *Acta Applicandae Mathematicae*, 164(1):83–122, 2019.
- [EM53] S. Eilenberg and S. Mac Lane. On the Groups  $H(\Pi, n)$ , I. *Annals of Mathematics*, 58(1):55–106, 1953.
- [FH14] P. K. Friz and M. Hairer. *A course on rough paths*. Universitext. Springer, Cham, 2014. With an introduction to regularity structures.
- [FP13] L. Foissy and F. Patras. Natural endomorphisms of shuffle algebras. *Internat. J. Algebra Comput.*, 23(4):989–1009, 2013.
- [FV10] P. K. Friz and N. B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Theory and applications.
- [Gal19] Francesco Galuppi. The rough veronese variety. *Linear algebra and its applications*, 583:282–299, 2019.
- [GK08] E. Gehrig and M. Kawski. A hopf-algebraic formula for compositions of non-commuting flows. 2008.
- [Hai02] R. Hain. Iterated integrals and algebraic cycles: examples and prospects. In *Contemporary trends in algebraic geometry and algebraic topology (Tianjin, 2000)*, volume 5 of *Nankai Tracts Math.*, pages 55–118. World Sci. Publ., River Edge, NJ, 2002.

- [HL10] B. Hambly and T. Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Ann. of Math. (2)*, 171(1):109–167, 2010.
- [Kap09] M. Kapranov. Noncommutative geometry and path integrals. In *Algebra, arithmetic, and geometry: in honor of Yu. I. Manin. Vol. II*, volume 270 of *Progr. Math.*, pages 49–87. Birkhäuser Boston, Inc., Boston, MA, 2009.
- [Lod95] J.-L. Loday. Cup-product for leibniz cohomology and dual leibniz algebras. *Mathematica Scandinavica*, 77(2):189–196, 1995.
- [PSS19] M. Pfeffer, A. Seigal, and B. Sturmfels. Learning paths from signature tensors. *SIAM Journal on Matrix Analysis and Applications*, 40(2):394–416, 2019.
- [Ree58] R. Ree. Lie elements and an algebra associated with shuffles. *Ann. of Math. (2)*, 68:210–220, 1958.
- [Reu93] C. Reutenauer. *Free Lie Algebras*, volume 7 of *London Mathematical Society monographs, new series*. Clarendon Press, Oxford University Press, Oxford, 1993.
- [Sch58] M. P. Schützenberger. Sur une propriété combinatoire des algèbres de lie libres pouvant être utilisée dans un problème de mathématiques appliquées. Séminaire Dubreil. Algèbres et théorie des nombres, Volume 12 (1958-1959) no. 1, Talk no. 1, p. 1-23, 1958.