Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Tropical time series, iterated-sums signatures and quasisymmetric functions

Joscha Diehl[1], Kurusch Ebrahimi-Fard[2], Nikolas Tapia[3]

submitted: September 18, 2020

[1] Universität Greifswald
Institut für Mathematik und Informatik
Walther-Rathenau-Str. 47
17489 Greifswald
Germany
E-Mail: joscha.diehl@gmail.com

[2] Department of Mathematical Sciences
NTNU
7491 Trondheim
Norway
E-Mail: kurusch.ebrahimi-fard@ntnu.no

[3] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
and
Institut für Mathematik
Technische Universität Berlin
Fakultät II
Str. des 17. Juni 136
10587 Berlin
Germany
E-Mail: nikolasesteban.tapiamunoz@wias-berlin.de

No. 2760
Berlin 2020

2020 *Mathematics Subject Classification.* 60L10, 60L70, 16Y60, 93C55.

Key words and phrases. Time series analysis, time warping, tropical quasisymmetric functions, semirings.

The first author thanks Bernd Sturmfels (MPI Leipzig) for introducing him to the tropical semiring. The second author was supported by the Research Council of Norway through project 302831 “Computational Dynamics and Stochastics on Manifolds” (CODYSMA). The third author was supported the BMS MATH+ Excellence Cluster EF1, project no. 5 "On robustness of Deep Neural Networks".

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Tropical time series, iterated-sums signatures and quasisymmetric functions

Joscha Diehl, Kurusch Ebrahimi-Fard, Nikolas Tapia

ABSTRACT. Driven by the need for principled extraction of features from time series, we introduce the iterated-sums signature over any commutative semiring. The case of the tropical semiring is a central, and our motivating, example, as it leads to features of (real-valued) time series that are not easily available using existing signature-type objects.

CONTENTS

DOI 10.20347/WIAS.PREPRINT.2760 Berlin 2020

1. Introduction

Recent developments [42, 37, 38, 16, 58] have shown that various forms of *iterated-sum* and *iterated-integral* operations can form a useful component in machine learning pipelines for sequential data. Originating from the study of (discretized) controlled ordinary differential equations (ODEs) [25, 44], they are particularly apt to model input-output relations that are well-approximated by dynamical systems [18, 19]. In fact, the *iterated-integrals signature* $\mathrm{IIS}(x)$[1] of a (smooth enough) multidimensional curve $x = (x^{(1)}, \ldots, x^{(d)}) : [0, T] \to \mathbb{R}^d$, is the solution to a certain *universal* controlled ODE [27, Proposition 7.8]. It is universal in the sense that the solution to *any* other controlled ODE can be well-approximated by a linear expression of the iterated-integrals signature. On a more concrete level, the entries of the "classical" iterated-integrals signature are real numbers, indexed by words $w = w_1 \cdots w_k$ in the alphabet $A' = \{1, \ldots, d\}$, and given as follows

$$
\begin{aligned}
\left\langle \mathrm{IIS}(x)_{s,u}, w \right\rangle &= \int_{s \le t_1 \le t_2 \le \cdots \le t_k \le u} dx_{t_1}^{(w_1)} \cdots dx_{t_k}^{(w_k)} \\
&= \int_{s \le t_1 \le t_2 \le \cdots \le t_k \le u} \dot{x}_{t_1}^{(w_1)} \cdots \dot{x}_{t_k}^{(w_k)} dt_1 \cdots dt_k.
\end{aligned} \tag{1}
$$

Not all input-output relations are well-modeled by controlled ODEs, though. As an extreme example we mention that a controlled ODE does not care about so-called "tree-like" excursions of the driving signal [34]. The iterated-integrals signature can therefore, for example, not distinguish the following two one-dimensional curves, $t \in [0, 1]$,

$$
t \mapsto 0 \qquad \text{and} \qquad t \mapsto \sin(2\pi t).
$$

There are several ways to circumvent this particular problem, by e.g. "lifting" a one-dimensional curve to a two-dimensional curve [26]. The *iterated-sums signature* (ISS) introduced in [16] (see also [38, 58]) forgoes this particular problem altogether and brings the added benefit of working directly with discrete time series (in order to apply the theory of iterated-*integrals* to discrete-time sequential data, it has to be interpolated to a, say, piecewise linear curve).

But even the ISS cannot "see" all aspects of a time series. Indeed, the ISS is invariant to time warping and hence cannot distinguish time series run at different speeds. It turns out that such an invariance is often desirable. The search for such invariants was in fact the starting point of [16], where it is shown that the ISS contains all polynomial expressions in the time series entries that are invariant to time warping.

A non-polynomial, time-warping invariant, functional is the following example (this is well-defined for any time series z that is eventually constant):

$$
(z_1, z_2, \ldots) \mapsto \min_j z_j. \tag{2}
$$

Moreover, it is expected to be poorly approximated by polynomial expressions. This is related to the fact that this functional is *not* well-approximated by (discretized) ODEs.

> *The aim of the work at hand is to introduce a principled feature extraction method for time series that encompasses functionals as the one in eq.* (2).

The entry point for our investigation was the observation that (2) can be considered as a polynomial expression, if one changes the underlying field of the reals to the tropical (or min-plus) semiring. This, as well as other semirings have (a subset of) the reals as the underlying set, and only the operations of "addition" and "multiplication" have a different meaning. As a result, this opens ways to consider real-valued time series under many different lenses. In particular it allows us to consider (2) as part

[1] Also just called the *signature* and denoted with $S(X)$.

DOI 10.20347/WIAS.PREPRINT.2760

of an iterated-sums signature. Moreover, semirings whose underlying sets are not given by subsets of the real line enable one to look at time series with values in more general spaces.

After this general motivation, we now present two ways to naturally arrive at the signature we introduce in this work, where the first one makes the remarks above more concrete.

1 **Invariants of a time series.** Accommodating the discrete nature of time series, $x = (x_1, x_2, \dots, x_N)$, $x_i \in \mathbb{R}^d$, one may consider the discrete analog of (1), i.e., the so-called iterated-sums signature ISS over the reals is defined as

$$\left\langle \mathrm{ISS}(x)_{p,q}, w \right\rangle = \sum_{p<i_1<i_2<\cdots<i_k\leq q} (\delta x_{i_1})^{w_1} \cdots (\delta x_{i_k})^{w_k}. \tag{3}$$

Here $w = w_1 \cdots w_k$ is a word over a certain alphabet – larger than A' – which is adapted to the discrete nature of the summation operation. It was shown in [16] that the map $\mathrm{ISS}(x)$ stores all polynomial invariants to time warping (and translation).

Alternatively, the ISS can be defined as

$$\left\langle \mathrm{ISS}(z)_{p,q}, w \right\rangle = \sum_{p<i_1<i_2<\cdots<i_k\leq q} (z_{i_1})^{w_1} \cdots (z_{i_k})^{w_k},$$

and the former definition is obtained by evaluating at increments, $z_i = \delta x_i = x_i - x_{i-1}$. The latter definition yields an object that is invariant to insertion of $0 \in \mathbb{R}$ into the time series z, and it is this viewpoint that generalizes to arbitrary semirings. More precisely, in Section 2 we construct a signature over commutative semirings that is invariant to insertion of zeros of the semiring. The underlying mathematical object are quasisymmetric expressions over commutative semirings, which we introduce for, to the best of our knowledge, the first time in Section 3. The way back from invariants to inserting zeros to time-warping invariants is not as straightforward as in the case over a field, and we investigate it in Section 4. As we will see, expressions like (2) will be covered by the theory.

2 **Cheap chronological information of a time series.**

The importance of convolutional neural networks (CNNs) is hard to overestimate [40] Their success, in particular in image recognition, is usually attributed to two ingredients

2.1 *weight sharing* reduces, in comparison to fully connected networks, the amount of parameters and hence allows for deeper architectures

2.2 convolution and its particular *structure* (usually combined with max-pooling) leads to desirable properties with respect to image recognition (modelling of receptive fields, approximate translation invariance, etc.)

Although CNNs have been successfully applied in the context of time series data (see [23] for a recent survey), this does not seem to be based on inherent properties of sequential data. In particular, the *structure* of time series is very different from that of images. It is not clear why the receptive-field structure of CNNs captures intrinsically meaningful information of sequential data. Moreover, time series possess a characteristic that images do not: a *chronology*, that is, the order of the series' values through time. We explain this using an example.

Example 1.1. *We consider a very concrete toy example. Let the input $x \in \{2, 4, 8, 16\}^4$ consist in sequences of length four in the numbers $2, 4, 8, 16$. On this input space we consider a convolutional layer with kernel-size 2, stride 1, followed by a max-pool with kernel-size 4. For example*

$$\begin{bmatrix} 2 & 4 & 4 & 16 \end{bmatrix} \mapsto \max\{a_1 \cdot 2 + a_2 \cdot 4, a_1 \cdot 4 + a_2 \cdot 4, a_1 \cdot 4 + a_2 \cdot 16\},$$

where $a_1, a_2 \in \mathbb{R}$ are the parameters of the convolutional kernel. This network can learn to answer questions of the following type.

- *Is there a* 16 *in the sequence* somewhere*? (Just as a full-blown CNN on image data can answer the question: Is there a dog* somewhere *in the picture?)*
- *Is there a* 2 directly *followed by a* 16 *somewhere in the sequence? (indeed, with* $a_1 = -1, a_2 = 1$ *one gets output* 14 *if and only if the statement is true)*

However, it can not *answer the question*

- *Is there a* 2 *somewhere and then,* sometime after*, a* 16 *in the sequence?*

We believe that chronological questions *of this type are the relevant questions for time series. Note that the following architecture allows to answer this question*

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} \mapsto \max\{a_1 \cdot x_{i_1} + a_2 \cdot x_{i_2} : i_1 < i_2\}.$$

Indeed, again with $a_1 = -1, a_2 = 1$*, the output is* 14 *if and only if the question is answered positively.*

Abstractly we can describe functions that extract such *chronological features* of sequences $x \in \mathbb{R}^N$ in the following form

$$x \mapsto \operatorname{pool}\Big(K(x_{i_1}, \ldots, x_{i_k}) : \{i_1 < \cdots < i_k\} \subset [N]\Big),$$

where

$$K : \mathbb{R}^k \to \mathbb{R}, \quad \operatorname{pool} : \mathbb{R}^{\binom{N}{k}} \to \mathbb{R}.$$

Now, in this generality, such features are computationally intractable, even for modest values of N and k, since K has to be evaluated $\binom{N}{k}$ times. The iterated-sums signature presented in this work represents a special case of the functions K and pool that *is* tractable. The application of this structure in deep learning pipelines will addressed in subsequent work.

The central object of this work, the *iterated-sums signature* $\mathrm{ISS}^{\mathbb{S}}$ will be properly defined in (19). To get there, we need to work through some algebraic background in Section 2 first. We therefore now give a preview in the setting of the tropical semiring $\mathbb{S} = \mathbb{R}_{\min+}$. Let $z_1, z_2, z_3, \cdots \in \mathbb{R}$ be an infinite time series. Define $\mathrm{ISS}^{\mathbb{R}_{\min+}}$, indexed by words $w = w_1 \cdots w_k$ in the alphabet $A = \{1, 2, 3, \ldots\}$[2] and $1 \leq s \leq t < +\infty$, as

$$\begin{aligned}\left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), w \right\rangle &:= \bigoplus_{\substack{\min+ \\ s<j_1<\cdots<j_k\leq t}} z_{j_1}^{\odot_{\min+} w_1} \odot_{\min+} \cdots \odot_{\min+} z_{j_k}^{\odot_{\min+} w_k} \\ &= \min_{s<j_1<\cdots<j_k\leq t} \{w_1 \cdot z_{j_1} + \cdots + w_k \cdot z_{j_k}\}.\end{aligned}$$

For example

$$\begin{aligned}\left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 1 \right\rangle &= \min_{s<j\leq t} z_j \\ \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 74 \right\rangle &= \min_{s<j_1<j_2\leq t} \{7 \cdot z_{j_1} + 4 \cdot z_{j_2}\} \\ \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 714 \right\rangle &= \min_{s<j_1<j_2<j_3\leq t} \{7 \cdot z_{j_1} + z_{j_2} + 4 \cdot z_{j_3}\}.\end{aligned} \tag{4}$$

We remark two, maybe, non-obvious properties of this object. Firstly, in order to calculate $\mathrm{ISS}^{\mathbb{R}_{\min+}}$ over large intervals, it suffices to calculate it over small intervals:

[2] Caution: in the main text we write (in the one-dimensional case) the alphabet as $\{[1], [1^2], [1^3], \ldots\} \cong \{1, 2, 3, \ldots\}$, since this notation extends nicely to higher dimensions.

Example 1.2. *For* $0 \leq s < t < u$,

$$\begin{aligned}
\left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,u}(z), 74 \right\rangle &= \bigoplus_{\substack{\min+ \\ s<j_1<j_2\leq u}} z_{j_1}^{\odot_{\min+}[1^7]} \odot_{\min+} z_{j_2}^{\odot_{\min+}[1^4]} \\
&= \min_{s<j_1<j_2\leq u} \{7z_{j_1} + 4z_{j_2}\} \\
&= \min \Big\{ \min_{s<j_1<j_2\leq t} \{7z_{j_1} + 4z_{j_2}\}, \\
&\qquad \min_{t<j_1<j_2\leq u} \{7z_{j_1} + 4z_{i_2}\}, \min_{s<i\leq t} \{7z_i\} + \min_{t<j\leq u} \{4z_j\} \Big\} \\
&= \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 74 \right\rangle \oplus_{\min+} \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{t,u}(z), 74 \right\rangle \\
&\qquad \oplus_{\min+} \left(\left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 7 \right\rangle \odot_{\min+} \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{t,u}(z), 4 \right\rangle \right).
\end{aligned}$$

This is also the reason why expressions that seem to have polynomial complexity (after all, the third order iterated sum (4) takes the maximum of $O(|t-s|^3)$-terms), are in fact calculable in linear time.

Secondly, we note that the *product* (in the semiring) of any iterated sums (which are iterated minima here) can be written as the *sum* (in the semiring) of (different) iterated sums:

Example 1.3. *In the min-plus semiring* $\mathbb{S} = \mathbb{R}_{\min+}$ *we have*

$$\begin{aligned}
&\left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 1 \right\rangle \odot_{\mathbb{S}} \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 74 \right\rangle \\
&= \min_{s<i\leq t} \{z_i\} + \min_{s<j<k\leq t} \{7z_j + 4z_k\} \\
&= \min_{s<i\leq t; s<j<k\leq t} \{z_i + 7z_j + 4z_k\} \\
&= \max \Big\{ \min_{s<i<j<k\leq t} \{z_i + 7z_j + 4z_k\}, \min_{s<i=j<k\leq t} \{z_i + 7z_j + 4z_k\}, \min_{s<j<i<k\leq t} \{z_i + 7z_j + 4z_k\}, \\
&\qquad \min_{s<j<k=i\leq t} \{z_i + 7z_j + 4z_k\}, \min_{s<j<k<i\leq t} \{z_i + 7z_j + 4z_k\} \Big\} \\
&= \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 174 \right\rangle \oplus_{\mathbb{S}} \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 84 \right\rangle \oplus_{\mathbb{S}} \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 714 \right\rangle \\
&\qquad \oplus_{\mathbb{S}} \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 75 \right\rangle \oplus_{\mathbb{S}} \left\langle \mathrm{ISS}^{\mathbb{R}_{\min+}}_{s,t}(z), 741 \right\rangle.
\end{aligned}$$

Both of these facts might come as no surprise to people familiar with iterated integrals or iterated sums over fields. Indeed, the first property is a version of *Chen's identity*. It just says that the computation of iterated integrals and iterated sums can be split into calculations on subintervals. This property is usually encoded algebraically by the non-cocommutative *deconcatenation* coproduct on the unital tensor algebra over an alphabet. The general form of Chen's identity in our setting is stated in Lemma 2.8.

Integration by parts implies that linear combinations of iterated integrals are closed under multiplication. This finds its abstract algebraic formulation in terms of the commutative shuffle product on the unital tensor algebra over an alphabet [13, 53]. Analogously, its discrete counterpart, i.e., summation by parts, permits to define an algebra on iterated sums, leading to the notion of commutative quasi-shuffle algebra [28]. The general form of the quasi-shuffle identity in our setting is found in Lemma 2.11.

Maybe more interestingly, new phenomena appear when working over general semirings. As we will see in Section 4, over an *idempotent* semiring, non-strict iterated sums satisfy a *shuffle* identity and in this sense behave like iterated *integrals*. These non-strict iterated sums also give a nice way to get certain time-warping invariants of a real-valued time series, covering expression (2).

The paper is organized as follows. In Section 2 we define the iterated-sums signature over a commutative semiring. In Section 3 we take a closer look at quasisymmetric functions, which are underlying the iterated-sums signature and are of independent interest. In Section 4 we return to the question of time warping invariants in the context of iterated-sums signature over a commutative semiring. We finish with conclusions and an outlook in Section 5. In Appendix A we present a categorical view on semirings, semimodules and semialgebras. Such a categorical view is useful in highlighting the similarities to the theory of rings, modules and algebras.

Related work

We finish this introduction by mentioning related literature. We already indicated how iterated sums and integrals have, in the last decade, been successfully applied as a feature extraction method in machine learning. The relation of iterated sums to the Hopf algebra of quasisymmetric functions [47] was established in [16].

Symmetric functions form an important subspace, and the generalization of this subspace to the setting of semirings has been investigated in [11, 36, 35].

Semirings play an important role in computer science. They appear, for example, in the closely related fields of language processing [32], the theory of algorithms [14], the theory of weighted automata [57, 5], shortest-paths problems in weighted directed graphs [24, 49], and iteration theories [7].

The tropical semiring in particular has been intensely studied, for example in algebraic geometry [45], in statistics [51], in economics [6], and in biology [52, Section 2]. Its linear algebra is well-understood [10, 1].

Acknowledgements: The first author thanks Bernd Sturmfels (MPI Leipzig) for introducing him to the tropical semiring. The second author was supported by the Research Council of Norway through project 302831 “Computational Dynamics and Stochastics on Manifolds” (CODYSMA). The third author was supported the BMS MATH+ Excellence Cluster EF1, project nº 5 "On robustness of Deep Neural Networks".

2. Iterated-sums signatures over a semiring

We start by introducing basic concepts from semiring theory. Relevant references are [41, 59, 45]. The definitions and constructions recalled here are “hands on”, for a categorical view see the Appendix.

Recall the definition of a **monoid**, which consist of a non-empty set M together with an associative product and a neutral element $\mathbf{1}_M$ for this product. For example, starting from an alphabet $A = \{a_1, \ldots, a_n\}$, the set of (finite) words over A, $w = a_{i_1} \cdots a_{i_k}$, forms under concatenation the free non-commutative monoid, denoted by A^*. The empty word, $\mathbf{e}$, is the neutral element. The length of a word $w = a_{i_1} \cdots a_{i_k} \in A^*$ is denoted $|w| = k$. A **monoid morphism** is a map between monoids which is compatible with the products and take the neutral element to the neutral element. Note that for an alphabet A and any monoid $(M, \centerdot, 1_M)$, every map $\phi : A \to M$ can be uniquely extended to a monoid morphism $\hat{\phi} : A^* \to M$ by defining $\hat{\phi}(\mathbf{e}) = 1_M$ and for any word $w = a_{i_1} \cdots a_{i_k} \in A^*$, $\hat{\phi}(a_{i_1} \cdots a_{i_k}) := \phi(a_{i_1}) \centerdot \cdots \centerdot \phi(a_{i_k})$. In other words, A^* is the **free monoid** over A.

The notion of semiring combines two monoids in a compatible, i.e. distributive, way. However, contrary to rings, the invertibility under addition is not part of the data.

Definition 2.1. The tuple $(\mathbb{S}, \oplus_{\mathbb{S}}, \odot_{\mathbb{S}}, \mathbf{0}_{\mathbb{S}}, \mathbf{1}_{\mathbb{S}})$ is a **semiring** if

- $(\mathbb{S}, \oplus_{\mathbb{S}}, \mathbf{0}_{\mathbb{S}})$ is a commutative monoid with unit $\mathbf{0}_{\mathbb{S}}$
- $(\mathbb{S}, \odot_{\mathbb{S}}, \mathbf{1}_{\mathbb{S}})$ is a monoid with unit $\mathbf{1}_{\mathbb{S}}$
- $\mathbf{0}_{\mathbb{S}} \odot_{\mathbb{S}} s = s \odot_{\mathbb{S}} \mathbf{0}_{\mathbb{S}} = \mathbf{0}_{\mathbb{S}}$ for all $s \in \mathbb{S}$

- multiplication distributes over addition, i.e.

(5) $$a \odot_{\mathbb{S}} (b \oplus_{\mathbb{S}} c) = (a \odot_{\mathbb{S}} b) \oplus_{\mathbb{S}} (a \odot_{\mathbb{S}} c), \quad (a \oplus_{\mathbb{S}} b) \odot_{\mathbb{S}} c = (a \odot_{\mathbb{S}} c) \oplus_{\mathbb{S}} (b \odot_{\mathbb{S}} c).$$

Note that the parentheses on the right-hand sides of the identities in (5) can be omitted assuming the common precedence of multiplication over addition. More importantly, a semiring $\mathbb{S}$ is called **idempotent**, if for all elements $a \in \mathbb{S}$ we have that $a \oplus_{\mathbb{S}} a = a$. In what follows, we will assume that the semiring under consideration is **commutative**, i.e., that $(\mathbb{S}, \odot_{\mathbb{S}}, \mathbf{1}_{\mathbb{S}})$ is a commutative monoid.

Semirings form an essential part of the modern theory of automata and languages. We refer the reader to the introductory references [41, 12].

Example 2.2. *Any commutative ring, in particular the field of reals* $(\mathbb{R}, +, \cdot, 0, 1)$*, forms a commutative semiring.*

The paradigms of honest semirings whose underlying sets are subsets of the reals are

1 $\mathbb{R}_{\text{min+}}$: *min-plus semiring* $(\mathbb{R} \cup \{+\infty\}, \min, +, +\infty, 0)$
2 $\mathbb{R}_{\text{max+}}$: *max-plus semiring* $(\mathbb{R} \cup \{-\infty\}, \max, +, -\infty, 0)$*,*

which are also known as tropical, respectively arctic, semirings. Here, maximum respectively minimum are considered as binary operations replacing the usual additive structure on $\mathbb{R}$*, and addition becomes multiplication. This results in particular arithmetic rules, e.g.,* $3 \oplus_{\mathbb{R}_{\text{max+}}} 3 = 3$*,* $4 \oplus_{\text{max+}} 3 = 4$*, and* $3 \odot_{\text{max+}} 3 = 6$*,* $-1 \odot_{\text{max+}} -1 = -2$*.*

3 *bottleneck semiring* $(\mathbb{R} \cup \{\pm\infty\}, \max, \min, -\infty, +\infty)$
4 *possibilistic semiring*[3] $([0, 1], \max, \cdot, 0, 1)$
5 $\mathbb{N}$: *non-negative integers* $(\mathbb{N}, +, \cdot, 0, 1)$
6 $\mathbb{R}_{\text{max}-\text{min}}$: *bottleneck semiring* $(\mathbb{R} \cup \{\pm\infty\}, \max, \min, -\infty, +\infty)$
7 $\overline{\mathbb{R}}_{\text{max}}$: *completed max-plus semiring* $(\mathbb{R} \cup \{\pm\infty\}, \max, +, -\infty, 0)$
8 *expectation semiring (or gradient semiring)* [22]*,* $(\mathbb{R}_{\geq 0} \times V, \oplus, \odot, (0, 0), (1, 0))$ *with* V *an arbitrary vector space and*

$$(a, v) \oplus (a', v') := (a + a', v + v')$$
$$(a, v) \odot (a', v') := (a \cdot a', a'v + av').$$

There are also examples of semirings whose underlying sets are not *given by (subsets of) the real line.*

9 *semiring of (bounded) polytopes* [52, Proposition 2.23]*.*
10 *k-best proof semiring* [32, 30]
11 *k-tropical semiring* [49]
12 *semiring of formal languages* [17]
13 *semiring of binary relations* [17]
14 *semiring of subsets of a set* M [50, Beispiel 2.10.a)] $(2^M, \cup, \cap, \varnothing, M)$
15 $\mathbb{B}$: *Boolean semiring* $(\{\mathit{false}, \mathit{true}\}, \text{or}, \text{and}, \mathit{false}, \mathit{true})$*.*

Regarding the last two examples, in fact any distributive lattice (with minimal and maximal element) naturally yields a commutative semiring, [31, Proposition 2.25]*.*

16 *semirings constructed from t-norms* [39, Example 6]*.*

[3] Also known as Viterbi or Bayesian semiring [56].

We note that summation by parts holds in a semiring. Indeed, multiplying two finite sums over semiring elements $a_i, b_j \in \mathbb{S}$ yields

$$
\begin{aligned}
\Big(\bigoplus_{0<i\leq N}{}_{\mathbb{S}}\, a_i\Big) \odot_{\mathbb{S}} \Big(\bigoplus_{0<j\leq N}{}_{\mathbb{S}}\, b_j\Big) &= \bigoplus_{0<i,j\leq N}{}_{\mathbb{S}}\, a_i \odot_{\mathbb{S}} b_j \\
&= \bigoplus_{0<i<j\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_j) \oplus_{\mathbb{S}} \bigoplus_{0<j<i\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_j) \oplus_{\mathbb{S}} \bigoplus_{0<i\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_i).
\end{aligned} \tag{6}
$$

Using commutativity and denoting $2_{\mathbb{S}} := \mathbf{1}_{\mathbb{S}} \oplus_{\mathbb{S}} \mathbf{1}_{\mathbb{S}}$,[4] we obtain for $a_i = b_i$ that

$$
\Big(\bigoplus_{0<i\leq N}{}_{\mathbb{S}}\, a_i\Big) \odot_{\mathbb{S}} \Big(\bigoplus_{0<j\leq N}{}_{\mathbb{S}}\, a_j\Big) = 2_{\mathbb{S}} \bigoplus_{0<i<j\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} a_j) \oplus_{\mathbb{S}} \bigoplus_{0<i\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} a_i).
$$

Regarding iterated sums of depth k

$$
\bigoplus_{0<j_1<\cdots<j_k\leq N}{}_{\mathbb{S}}\, a_{j_1} \odot_{\mathbb{S}} \cdots \odot_{\mathbb{S}} a_{j_k}, \tag{7}
$$

we see that (6) allows to express the product of two iterated sums of depths k_1 and k_2 in terms of a linear combination of iterated sums of depths $\max(k_1, k_2) \leq k \leq k_1 + k_2$. The algebraic formulation of this leads to the quasi-shuffle identity (see below).

We remark on a peculiarity in the semiring setting. The lack of inverses with respect to addition turns summation by parts for non-strict iterated sums less appealing, since products of such sums do not close under semiring multiplication. Indeed, returning to (6) we observe that already at this level the product can not be expressed exclusively in terms of non-strict iterated sums (as the doubly counted diagonal term can not be subtracted in the semiring)

$$
\begin{aligned}
\Big(\bigoplus_{0<i\leq N}{}_{\mathbb{S}}\, a_i\Big) \odot_{\mathbb{S}} \Big(\bigoplus_{0<j\leq N}{}_{\mathbb{S}}\, b_j\Big) &= \bigoplus_{0<i,j\leq N}{}_{\mathbb{S}}\, a_i \odot_{\mathbb{S}} b_j \\
&= \bigoplus_{0<i\leq j\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_j) \oplus_{\mathbb{S}} \bigoplus_{0<j<i\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_j) \\
&= \bigoplus_{0<i<j\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_j) \oplus_{\mathbb{S}} \bigoplus_{0<j\leq i\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_j).
\end{aligned} \tag{8}
$$

However, if we consider (8) in an *idempotent* semiring, one observes an interesting phenomenon. The fact that $a \oplus_{\mathbb{S}} a = a$ for all elements a in such a semiring, implies the somewhat surprising identity

$$
\Big(\bigoplus_{0<i\leq N}{}_{\mathbb{S}}\, a_i\Big) \odot_{\mathbb{S}} \Big(\bigoplus_{0<j\leq N}{}_{\mathbb{S}}\, b_j\Big) = \bigoplus_{0<i\leq j\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_j) \oplus_{\mathbb{S}} \bigoplus_{0<j\leq i\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_j), \tag{9}
$$

where we used that

$$
\bigoplus_{0<i\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_i) \oplus_{\mathbb{S}} \bigoplus_{0<i\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_i) = \bigoplus_{0<i\leq N}{}_{\mathbb{S}}\, (a_i \odot_{\mathbb{S}} b_i),
$$

[4] We note that in the case of an idempotent semiring we have that $2_{\mathbb{S}} = \mathbf{1}_{\mathbb{S}} \oplus_{\mathbb{S}} \mathbf{1}_{\mathbb{S}} = \mathbf{1}_{\mathbb{S}}$.

which allows to equate (9) and (6). Let us look at the following example

$$
\begin{aligned}
&\left(\bigoplus_{s<i_1\le i_2\le N}^{\mathbb{S}} z_{i_1}^{\odot_{\mathbb{S}}7}\odot_{\mathbb{S}} z_{i_2}^{\odot_{\mathbb{S}}3}\right)\odot_S\left(\bigoplus_{s<i_1\le N}^{\mathbb{S}} z_{i_1}^{\odot_{\mathbb{S}}5}\right)\\
&=\left(\bigoplus_{s<i_1\le i_2\le i_3\le N}^{\mathbb{S}} z_{i_1}^{\odot_{\mathbb{S}}7}\odot_{\mathbb{S}} z_{i_2}^{\odot_{\mathbb{S}}3}\odot_{\mathbb{S}} z_{i_3}^{\odot_{\mathbb{S}}5}\right)\oplus_{\mathbb{S}}\left(\bigoplus_{s<i_1\le i_3\le i_2\le N}^{\mathbb{S}} z_{i_1}^{\odot_{\mathbb{S}}7}\odot_{\mathbb{S}} z_{i_2}^{\odot_{\mathbb{S}}3}\odot_{\mathbb{S}} z_{i_3}^{\odot_{\mathbb{S}}5}\right)\\
&\quad\oplus_{\mathbb{S}}\left(\bigoplus_{s<i_3<i_1\le i_2\le N}^{\mathbb{S}} z_{i_1}^{\odot_{\mathbb{S}}7}\odot_{\mathbb{S}} z_{i_2}^{\odot_{\mathbb{S}}3}\odot_{\mathbb{S}} z_{i_3}^{\odot_{\mathbb{S}}5}\right)\\
&=\left(\bigoplus_{s<i_1\le i_2\le i_3\le N}^{\mathbb{S}} z_{i_1}^{\odot_{\mathbb{S}}7}\odot_{\mathbb{S}} z_{i_2}^{\odot_{\mathbb{S}}3}\odot_{\mathbb{S}} z_{i_3}^{\odot_{\mathbb{S}}5}\right)\oplus_{\mathbb{S}}\left(\bigoplus_{s<i_1\le i_3\le\le i_2\le N}^{\mathbb{S}} z_{i_1}^{\odot_{\mathbb{S}}7}\odot_{\mathbb{S}} z_{i_2}^{\odot_{\mathbb{S}}3}\odot_{\mathbb{S}} z_{i_3}^{\odot_{\mathbb{S}}5}\right)\\
&\quad\oplus_{\mathbb{S}}\left(\bigoplus_{s<i_3\le i_1\le i_2\le t}^{\mathbb{S}} z_{i_1}^{\odot_{\mathbb{S}}7}\odot_{\mathbb{S}} z_{i_2}^{\odot_{\mathbb{S}}3}\odot_{\mathbb{S}} z_{i_3}^{\odot_{\mathbb{S}}5}\right).
\end{aligned}
\tag{10}
$$

Hence, we observe that in the idempotent case, products of non-strict iterated sums satisfy the *shuffle* relation. We return to this in Section 4.

Definition 2.3. A **semimodule** over a commutative semiring $(\mathbb{S},\oplus_{\mathbb{S}},\odot_{\mathbb{S}},\mathbf{0}_{\mathbb{S}},\mathbf{1}_{\mathbb{S}})$ consists in a commutative monoid $(M,+_M,\mathbf{0}_M)$ and a scalar multiplication $\mathbb{S}\times M\to M$, $(s,m)\mapsto sm$, satisfying for all $s,s'\in\mathbb{S}$ and $m,m'\in M$

$$
\begin{gathered}
\mathbf{1}_{\mathbb{S}}m=m \qquad \mathbf{0}_{\mathbb{S}}m=\mathbf{0}_M\\
(s\odot_{\mathbb{S}} s')m=s(s'm) \qquad (s\oplus_{\mathbb{S}} s')m=sm+_M s'm\\
s(m+_M m')=sm+_M sm'.
\end{gathered}
$$

Note that if the underlying semiring $\mathbb{S}$ is idempotent, then the semimodule M is idempotent as well.

Let $(N,+_N,\mathbf{0}_N)$ be another semimodule over $\mathbb{S}$. A map $\phi:M\to N$ is a **semimodule morphism** if for all $s,s'\in\mathbb{S}$ and $m,m'\in M$

$$\phi(sm+_M s'm')=s\phi(m)+_N s'\phi(m').$$

Example 2.4. *An incarnation of the **free** $\mathbb{S}$**-semimodule** $\mathbb{F}$ on a set D, is given by functions $f:D\to\mathbb{S}$ with finite support, i.e., $f(d)=\mathbf{0}_{\mathbb{S}}$ for all but finitely many elements $d\in D$. The action of $\mathbb{S}$ as well as the addition are defined pointwise.*

Definition 2.5. An associative **semialgebra** over a commutative semiring $\mathbb{S}$ consists of a semiring $(\mathbb{A},\oplus_{\mathbb{A}},\odot_{\mathbb{A}},\mathbf{0}_{\mathbb{A}},\mathbf{1}_{\mathbb{A}})$ such that $(\mathbb{A},\oplus_{\mathbb{A}},\mathbf{0}_{\mathbb{A}})$ is a semimodule over $\mathbb{S}$ and such that the semimodule structure is compatible with $\odot_{\mathbb{A}}$ in the following way

$$s(a\odot_{\mathbb{A}} a')=(sa)\odot_{\mathbb{A}} a'=a\odot_{\mathbb{A}}(sa').$$

Remark 2.6. *Motivated by summation by parts* (6) *and the particular property of iterated non-strict sums* (9)*, one can introduce the notion of a Rota–Baxter $\mathbb{S}$-semialgebra. Let $\mathbb{A}$ be a $\mathbb{S}$-semialgebra. A* Rota–Baxter map *of weight $\lambda\in\mathbb{S}$ is a $\mathbb{S}$-linear map $R:\mathbb{A}\to\mathbb{A}$ satisfying for any $x,y\in\mathbb{A}$*

$$R(x)\odot_{\mathbb{A}} R(y)=R\big(R(x)\odot_{\mathbb{A}} y\oplus_{\mathbb{A}} x\odot_{\mathbb{A}} R(y)\big)\oplus_{\mathbb{A}}\lambda R(x\odot_{\mathbb{A}} y). \tag{11}$$

Note that if the semiring $\mathbb{S}$ is idempotent, and therefore also the semialgebra $\mathbb{A}$, then the map $\tilde{R}:=\lambda\,\mathrm{id}_{\mathbb{A}}\oplus_{\mathbb{A}} R$ also satisfies the particular relation (11)*. In fact, we have the more surprising (weight zero)* *identity*[5] *(compare also* (10)*)*

$$\tilde{R}(x) \odot_{\mathbb{A}} \tilde{R}(y) = \tilde{R}\big(\tilde{R}(x) \odot_{\mathbb{A}} y \oplus_{\mathbb{A}} x \odot_{\mathbb{A}} \tilde{R}(y)\big). \tag{12}$$

We now consider formal series over the (possibly infinite) alphabet A with coefficients in a commutative semiring $\mathbb{S}$

$$F := \sum_{w \in A^*} c_w w, \quad c_w \in \mathbb{S}. \tag{13}$$

The set of all such series is denoted by $\mathbb{S}\langle\langle A\rangle\rangle$. For a series $F \in \mathbb{S}\langle\langle A\rangle\rangle$, the support, $\mathrm{supp}(F)$, consists of all words in $w \in A^*$ with coefficient c_w different from $\mathbf{0}_{\mathbb{S}}$. We denote with $\mathbb{S}\langle A\rangle$ the subset of series with finite support. We may view a series (13) as a map $F : A^* \to \mathbb{S}$, with

$$\langle F, w\rangle_{\mathbb{S}} := F(w) := c_w.$$

By linear extension such F become maps on $\mathbb{S}\langle A\rangle$ with image in $\mathbb{S}$, and we denote this pairing with $\langle ., .\rangle_{\mathbb{S}}$ still.

We can equip $\mathbb{S}\langle\langle A\rangle\rangle$ with a linear and multiplicative structure by defining

$$\langle sF, w\rangle_{\mathbb{S}} := s\langle F, w\rangle_{\mathbb{S}}, \quad s \in \mathbb{S}$$

$$\langle F_1 + F_2, w\rangle_{\mathbb{S}} := \langle F_1, w\rangle_{\mathbb{S}} \oplus_{\mathbb{S}} \langle F_2, w\rangle_{\mathbb{S}} \tag{14}$$

$$\langle F_1 F_2, w\rangle_{\mathbb{S}} := \bigoplus_{vu=w}{}_{\mathbb{S}} \langle F_1, v\rangle_{\mathbb{S}} \odot_{\mathbb{S}} \langle F_2, u\rangle_{\mathbb{S}}. \tag{15}$$

For instance, let $F = s_1 a_1 a_2 + s_2 a_1$ and $G = t_1 a_3 + t_2 a_2 a_3$ be elements in $\mathbb{S}\langle A\rangle$ then

$$FG = ((s_1 \odot_{\mathbb{S}} t_1) \oplus_{\mathbb{S}} (s_2 \odot_{\mathbb{S}} t_2))a_1 a_2 a_3 + (s_1 \odot_{\mathbb{S}} t_2)a_1 a_2 a_2 a_3 + (s_2 \odot_{\mathbb{S}} t_1)a_1 a_3.$$

This turns $\mathbb{S}\langle\langle A\rangle\rangle$, as well as $\mathbb{S}\langle A\rangle$, into $\mathbb{S}$-semialgebras. The constant series are given by defining for any element $s \in \mathbb{S}$ the series $F_s := s\mathbf{e}$, which include in particular the constant series 1 and 0.[6]

As is the case over rings, $\mathbb{S}\langle A\rangle$ is the **free associative $\mathbb{S}$-semialgebra** over the alphabet A. This manifests in the following universal property: for any $\mathbb{S}$-semialgebra $\mathbb{U}$ and map $\phi : A \to \mathbb{U}$ there exists a $\mathbb{S}$-semialgebra morphism uniquely defined by extending ϕ to $\hat{\phi} : A^* \to \mathbb{U}$ multiplicatively as well as $\mathbb{S}$-linearly.

We shall now equip the $\mathbb{S}$-semimodule $\mathbb{S}\langle A\rangle$ with another product. This product is a natural extension of the well-known shuffle (or Hurwitz) product commonly defined in automata theory [41]. For this we assume that the alphabet A carries a commutative semigroup product denoted by the binary bracket operation $[- -] : A \times A \to A$. Observe that commutativity and associativity permit to express iterations $[a_{i_1}[\cdots[a_{i_{n-1}}\, a_{i_n}]]\cdots] = [a_{i_1} \cdots a_{i_n}]$. The commutative **quasi-shuffle product** on $\mathbb{S}\langle A\rangle$ is defined first on words and then extended bilinearly. For words ua and vb, where $a_i, a_j \in A$, $u, v \in A^*$, we define $ua_i * \mathbf{e} = ua_i = \mathbf{e} * ua_i$ and inductively

$$ua_i * va_j := (u * va_j)a_i + (ua_i * v)a_j + (u * v)[a_i\, a_j]. \tag{16}$$

[5] Indeed, we see that by expanding the right-hand side of (12) we obtain

$$\begin{aligned}
\tilde{R}\big(\tilde{R}(x) \odot_{\mathbb{A}} y \oplus_{\mathbb{A}} x \odot_{\mathbb{A}} \tilde{R}(y)\big) &= \tilde{R}\big((\lambda \oplus_{\mathbb{S}} \lambda)x \odot_{\mathbb{A}} y \oplus_{\mathbb{A}} R(x) \odot_{\mathbb{A}} y \oplus_{\mathbb{A}} x \odot_{\mathbb{A}} R(y)\big)\\
&= (\lambda \odot_{\mathbb{S}} \lambda)x \odot_{\mathbb{A}} y \oplus_{\mathbb{A}} \lambda R(x) \odot_{\mathbb{A}} y \oplus_{\mathbb{A}} \lambda x \odot_{\mathbb{A}} R(y) \oplus_{\mathbb{A}} R\big(\lambda x \odot_{\mathbb{A}} y \oplus_{\mathbb{A}} R(x) \odot_{\mathbb{A}} y \oplus_{\mathbb{A}} x \odot_{\mathbb{A}} R(y)\big)\\
&= \lambda \odot_{\mathbb{S}} \lambda(x \odot_{\mathbb{A}} y) \oplus_{\mathbb{A}} \lambda x \odot_{\mathbb{A}} R(y) \oplus_{\mathbb{A}} \lambda R(x) \odot_{\mathbb{A}} y \oplus_{\mathbb{A}} R(x) \odot_{\mathbb{A}} R(y) = \tilde{R}(x) \odot_{\mathbb{A}} \tilde{R}(y).
\end{aligned}$$

[6] Recall that $\mathbf{e} \in A^*$ is the empty word.

For instance, $a_i * a_j = a_i a_j + a_j a_i + [a_i\ a_j]$. It is easy to observe that for a trivial bracket product on A, the quasi-shuffle product (16) reduces to the usual shuffle product on A. The latter will be denoted with $\sqcup\!\sqcup$ and satisfies the recursion [54]

$$ua_i \sqcup\!\sqcup va_j = (u \sqcup\!\sqcup va_j)a_i + (ua_i \sqcup\!\sqcup v)a_j. \tag{17}$$

It is known that an explicit expression can be defined for the shuffle product in terms of so-called shuffle permutations (bijections). In the case of the quasi-shuffle product, an analogous non-recursive formula can be given in terms of certain surjections [21, 20]. We refer the reader to Appendix B for details.

For the remainder of the paper, we specialise to a specific alphabet A. Let $A' = \{1, 2, \ldots, d\}$ and let A be the – extended – alphabet containing all formal brackets in elements of A, i.e. all formal monomials in those letters,

$$A = \{[1], [2], \ldots, [\mathtt{d}], [1^2], [12], \ldots, [\mathtt{d}^2], [1^3], \ldots\}. \tag{18}$$

Here, for consistency of notation, we write $[1] = 1, \ldots, [\mathtt{d}] = \mathtt{d}$.

We consider the space of $\mathbb{S}^d$-valued time series of infinite length that are eventually equal to $\mathbf{0}^d_{\mathbb{S}}$,

$$\mathbb{S}^{d,\mathbb{N}_{\geq 1}}_{\mathbf{0}_{\mathbb{S}}} := \{z : \mathbb{N}_{\geq 1} \to \mathbb{S}^d : \exists N \geq 1 \text{ such that } z_n = \mathbf{0}^d_{\mathbb{S}},\ \forall n > N\}.$$

It contains sequences $z = (z_1, z_2, \ldots, z_N, \mathbf{0}^d_{\mathbb{S}}, \mathbf{0}^d_{\mathbb{S}}, \ldots)$ of elements $z_i = (z_i^{(1)}, \ldots, z_i^{(d)}) \in \mathbb{S}^d$.

Example 2.7. *Let $x = (x_0, x_1, x_2, \ldots)$, $x_i \in \mathbb{R}^d$, be a time series that is eventually constant, then z with entries*

$$z_n^{(i)} := -\log|x_n^{(i)} - x_{n-1}^{(i)}|, \quad n = 1, 2, 3, \ldots$$

is in $\mathbb{S}^{d,\mathbb{N}_{\geq 1}}_{\mathbf{0}_{\mathbb{S}}}$, for $\mathbb{S}$ the tropical semiring $\mathbb{R}_{\min+} = (\mathbb{R} \cup \{+\infty\}, \min, +, +\infty, 0)$.

We now define for $z \in \mathbb{S}^{d,\mathbb{N}_{\geq 1}}_{\mathbf{0}_{\mathbb{S}}}$ the **$\mathbb{S}$-iterated-sums signature** $\mathrm{ISS}^{\mathbb{S}}_{s,t}(z) \in \mathbb{S}\langle\langle A\rangle\rangle$ as

$$\left\langle \mathrm{ISS}^{\mathbb{S}}_{s,t}(z), w\right\rangle_{\mathbb{S}} := \bigoplus_{s<j_1<\cdots<j_k<t+1}{}_{\mathbb{S}}\ z_{j_1}^{\odot_{\mathbb{S}} w_1} \odot_{\mathbb{S}} \cdots \odot_{\mathbb{S}} z_{j_k}^{\odot_{\mathbb{S}} w_k}, \quad 0 \leq s \leq t \leq +\infty. \tag{19}$$

Here $w = w_1 \cdots w_k \in A^*$, $w_i \in A$. We also write $\mathrm{ISS}^{\mathbb{S}}(z) := \mathrm{ISS}^{\mathbb{S}}_{0,\infty}(z)$. Here the notation means that for $w_i = [a_{i_1} \cdots a_{i_m}] \in A$

$$z_j^{\odot_{\mathbb{S}}[a_{i_1}\cdots a_{i_m}]} := z_j^{(a_{i_1})} \odot_{\mathbb{S}} \cdots \odot_{\mathbb{S}} z_j^{(a_{i_m})}.$$

As an example, we compute

$$\left\langle \mathrm{ISS}^{\mathbb{S}}_{s,t}(z), [1][23]\right\rangle_{\mathbb{S}} = \bigoplus_{s<j_1<j_2<t+1}{}_{\mathbb{S}}\ z_{j_1}^{\odot_{\mathbb{S}}[1]} \odot_{\mathbb{S}} z_{j_2}^{\odot_{\mathbb{S}}[23]} = \bigoplus_{s<j_1<j_2<t+1}{}_{\mathbb{S}}\ z_{j_1}^{(1)} \odot_{\mathbb{S}} z_{j_2}^{(2)} \odot_{\mathbb{S}} z_{j_2}^{(3)}.$$

Our first results concern the verification that $\mathrm{ISS}^{\mathbb{S}}$ is a proper iterated-sums signature. By this we mean that it carries the two main properties mentioned in the introduction, i.e., it satisfies Chen's identity and it is compatible with the quasi-shuffle product (16). To some extend this may be expected as both reflect basic properties of the iteration of summation operation combined with the chronological order in the time domain preserved through words. Indeed, the algebraic structure of $\mathrm{ISS}^{\mathbb{S}}$ is nicely compatible with concatenation of time series.

Lemma 2.8 (Chen's identity)**.** *For $p < r < q$, $w \in \mathbb{S}\langle A\rangle$ and $z \in \mathbb{S}^{d,\mathbb{N}_{\geq 1}}_{\mathbf{0}_{\mathbb{S}}}$*

$$\left\langle \mathrm{ISS}^{\mathbb{S}}_{p,q}(z), w\right\rangle_{\mathbb{S}} = \bigoplus_{uv=w}{}_{\mathbb{S}} \left\langle \mathrm{ISS}^{\mathbb{S}}_{p,r}(z), u\right\rangle_{\mathbb{S}} \odot_{\mathbb{S}} \left\langle \mathrm{ISS}^{\mathbb{S}}_{r,q}(z), v\right\rangle_{\mathbb{S}}, \tag{20}$$

or, equivalently, using the non-commutative concatenation product on $\mathbb{S}\langle\langle A\rangle\rangle$,

$$\mathsf{ISS}^{\mathbb{S}}_{p,r}(z)\mathsf{ISS}^{\mathbb{S}}_{r,q}(z) = \mathsf{ISS}^{\mathbb{S}}_{p,q}(z). \tag{21}$$

Remark 2.9. *Note that this, in general, only allows, for $p < r < q$, to calculate $\mathsf{ISS}^{\mathbb{S}}_{p,q}(z)$ from $\mathsf{ISS}^{\mathbb{S}}_{p,r}(z)$ and $\mathsf{ISS}^{\mathbb{S}}_{r,q}(z)$ but* not *to calculate $\mathsf{ISS}^{\mathbb{S}}_{r,q}(z)$ from $\mathsf{ISS}^{\mathbb{S}}_{p,q}(z)$ and $\mathsf{ISS}^{\mathbb{S}}_{p,r}(z)$. This is due to the fact that semiring addition, $\oplus_{\mathbb{S}}$, is not invertible.*

Example 2.10. *In the min-plus semiring (here $d = 1$ corresponding to the single letter alphabet $A' = \{1\}$) we obtain for example*

$$\begin{aligned}
\left\langle \mathsf{ISS}^{\mathbb{R}_{\min+}}_{p,q}(z), [1^7][1^4]\right\rangle &= \bigoplus_{\substack{\min+\\ p<j_1<j_2<q+1}} z_{j_1}^{\odot_{\min+}[1^7]} \odot_{\min+} z_{j_2}^{\odot_{\min+}[1^4]}\\
&= \min_{p<i_1<i_2\le q}\{7z_{i_1} + 4z_{i_2}\}\\
&= \min\Big\{ \min_{p<i_1<i_2\le r}\{7z_{i_1}+4z_{i_2}\},\\
&\qquad \min_{r<i_1<i_2\le q}\{7z_{i_1}+4z_{i_2}\}, \min_{p<i\le r}\{7z_i\} + \min_{r<i\le q}\{4z_i\}\Big\}\\
&= \left\langle \mathsf{ISS}^{\mathbb{R}_{\min+}}_{p,r}(z), [1^7][1^4]\right\rangle \oplus_{\min+} \left\langle \mathsf{ISS}^{\mathbb{R}_{\min+}}_{r,q}(z), [1^7][1^4]\right\rangle\\
&\qquad \oplus_{\min+} \left(\left\langle \mathsf{ISS}^{\mathbb{R}_{\min+}}_{p,r}(z), [1^7]\right\rangle \odot_{\min+} \left\langle \mathsf{ISS}^{\mathbb{R}_{\min+}}_{r,q}(z), [1^4]\right\rangle\right).
\end{aligned}$$

Proof. We now show (20) by a direct calculation

$$\begin{aligned}
\left\langle \mathsf{ISS}^{\mathbb{S}}_{p,q}(z), w\right\rangle &= \bigoplus_{\substack{\mathbb{S}\\ p<j_1<j_2<\cdots<j_k\le q}} z_{j_1}^{\odot_{\mathbb{S}} w_1} \odot_{\mathbb{S}} \cdots \odot_{\mathbb{S}} z_{j_k}^{\odot_{\mathbb{S}} w_k}\\
&= \bigoplus_{\substack{\mathbb{S}\\ p<r<j_1<j_2<\cdots<j_k\le q}} z_{j_1}^{\odot_{\mathbb{S}} w_1} \odot_{\mathbb{S}} \cdots \odot_{\mathbb{S}} z_{j_k}^{\odot_{\mathbb{S}} w_k}\\
&\oplus_{\mathbb{S}} \bigoplus_{\substack{\mathbb{S}\\ p<j_1\le r<j_2<\cdots<j_k\le q}} z_{j_1}^{\odot_{\mathbb{S}} w_1} \odot_{\mathbb{S}} \cdots \odot_{\mathbb{S}} z_{j_k}^{\odot_{\mathbb{S}} w_k}\\
&\oplus_{\mathbb{S}} \dots\\
&\oplus_{\mathbb{S}} \bigoplus_{\substack{\mathbb{S}\\ p<j_1<j_2<\cdots<j_{k-1}\le r<j_k\le q}} z_{j_1}^{\odot_{\mathbb{S}} w_1} \odot_{\mathbb{S}} \cdots \odot_{\mathbb{S}} z_{j_k}^{\odot_{\mathbb{S}} w_k}\\
&\oplus_{\mathbb{S}} \bigoplus_{\substack{\mathbb{S}\\ p<j_1<j_2<\cdots<j_{k-1}<j_k\le r<q}} z_{j_1}^{\odot_{\mathbb{S}} w_1} \odot_{\mathbb{S}} \cdots \odot_{\mathbb{S}} z_{j_k}^{\odot_{\mathbb{S}} w_k}\\
&= \sum_{uv=w} \left\langle \mathsf{ISS}^{\mathbb{S}}_{p,r}(z), u\right\rangle \odot_{\mathbb{S}} \left\langle \mathsf{ISS}^{\mathbb{S}}_{r,q}(z), v\right\rangle.
\end{aligned}$$

□

From summation by parts (6) extended to iterated $\mathbb{S}$-sums we deduce the multiplicativity of $\mathsf{ISS}^{\mathbb{S}}(z)$ over the quasi-shuffle product (16) on $\mathbb{S}\langle A\rangle$.

Lemma 2.11 (Multiplicativity)**.** *For $w, u \in \mathbb{S}\langle A\rangle$ and $z \in \mathbb{S}^{d,\mathbb{N}_{\ge 1}}_{\mathbf{0}_{\mathbb{S}}}$*

$$\left\langle \mathsf{ISS}^{\mathbb{S}}_{s,t}(z), w * u\right\rangle = \left\langle \mathsf{ISS}^{\mathbb{S}}_{s,t}(z), w\right\rangle \odot_{\mathbb{S}} \left\langle \mathsf{ISS}^{\mathbb{S}}_{s,t}(z), u\right\rangle.$$

Example 2.12. *In the min-plus semiring $\mathbb{S} = \mathbb{R}_{\min+}$ (again, here we consider the single letter case $d = 1$) we have*

$$\left\langle \mathsf{ISS}^{\mathbb{S}}(z), [1^1]\right\rangle \odot_{\mathbb{S}} \left\langle \mathsf{ISS}^{\mathbb{S}}(z), [1^7][1^4]\right\rangle$$

$$
\begin{aligned}
&= \max_i\{z_i\} + \max_{j<k}\{7z_j + 4z_k\} \\
&= \max_{i;j<k}\{z_i + 7z_j + 4z_k\} \\
&= \max\Big\{ \max_{i<j<k}\{z_i + 7z_j + 4z_k\}, \max_{i=j<k}\{z_i + 7z_j + 4z_k\}, \max_{j<i<k}\{z_i + 7z_j + 4z_k\}, \\
&\qquad \max_{j<k=i}\{z_i + 7z_j + 4z_k\}, \max_{j<k<i}\{z_i + 7z_j + 4z_k\}\Big\} \\
&= \Big\langle \mathrm{ISS}^{\mathbb{S}}(z), [1^1][1^7][1^4]\Big\rangle \oplus_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}(z), [1^8][1^4]\Big\rangle \oplus_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}(z), [1^7][1^1][1^4]\Big\rangle \\
&\qquad \oplus_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}(z), [1^7][1^5]\Big\rangle \oplus_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}(z), [1^7][1^4][1^1]\Big\rangle.
\end{aligned}
$$

Proof. We perform induction on the sum of length $q = |w| + |v|$ of the words. It is trivially true for $q = 0$. Let it be true up to arbitrary $q-1$ and assume $|w| + |v| = q$. For $f_i, g_i \in \mathbb{S}$, $i = s+1, \dots, t$, define

$$
\begin{aligned}
f_i &:= \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,i}(z), v_1 \cdots v_{k-1}\Big\rangle \odot_{\mathbb{S}} z_i^{v_k} \\
g_i &:= \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,i}(z), w_1 \cdots w_{\ell-1}\Big\rangle \odot_{\mathbb{S}} z_i^{w_\ell}.
\end{aligned}
$$

Then, summation by parts (6) implies

$$
\begin{aligned}
\big\langle \mathrm{ISS}^{\mathbb{S}}_{s,t}(z), v\big\rangle \odot_{\mathbb{S}} \big\langle \mathrm{ISS}^{\mathbb{S}}_{s,t}(z), w\big\rangle &= \Bigg(\bigoplus_{s<i<t+1}{}_{\mathbb{S}} f_i\Bigg) \odot_{\mathbb{S}} \Bigg(\bigoplus_{s<j<t+1}{}_{\mathbb{S}} g_i\Bigg) \\
&= \Bigg(\bigoplus_{s<i<j<t+1}{}_{\mathbb{S}} f_i \odot_{\mathbb{S}} g_j\Bigg) \oplus_{\mathbb{S}} \Bigg(\bigoplus_{s<j<i<t+1}{}_{\mathbb{S}} f_i \odot_{\mathbb{S}} g_j\Bigg) \oplus_{\mathbb{S}} \Bigg(\bigoplus_{s<i<t+1}{}_{\mathbb{S}} f_i \odot_{\mathbb{S}} g_i\Bigg). \qquad (22)
\end{aligned}
$$

Now, the first term in the last equality is equal to

$$
\begin{aligned}
&\bigoplus_{s<i<j<t+1}{}_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,i}(z), v_1 \cdots v_{k-1}\Big\rangle \odot_{\mathbb{S}} z_i^{v_k} \odot_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,j}(z), w_1 \cdots w_{\ell-1}\Big\rangle \odot_{\mathbb{S}} z_j^{w_\ell} \\
&= \bigoplus_{s<j<t+1}{}_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,j}(z), v_1 \cdots v_k\Big\rangle \odot_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,j}(z), w_1 \cdots w_{\ell-1}\Big\rangle \odot_{\mathbb{S}} z_j^{w_\ell} \\
&= \bigoplus_{s<j<t+1}{}_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,j}(z), v * (w_1 \cdots w_{\ell-1})\Big\rangle \odot_{\mathbb{S}} z_j^{w_\ell} \\
&= \big\langle \mathrm{ISS}^{\mathbb{S}}_{s,t}(z), (v * (w_1 \cdots w_{\ell-1}))w_\ell\big\rangle,
\end{aligned}
$$

where we used the induction hypothesis, since $|v| + |w_1 \cdots w_{\ell-1}| = q-1$. Analogously, we argue for the second term. The last term is equal to

$$
\begin{aligned}
&\bigoplus_{s<i<t+1}{}_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,i}(z), v_1 \cdots v_{k-1}\Big\rangle \odot_{\mathbb{S}} z_i^{v_k} \odot_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,i}(z), w_1 \cdots w_{\ell-1}\Big\rangle \odot_{\mathbb{S}} z_i^{w_\ell} \\
&= \bigoplus_{s<i<t+1}{}_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,i}(z), (v_1 \cdots v_{k-1}) * (w_1 \cdots w_{\ell-1})\Big\rangle \odot_{\mathbb{S}} z_i^{v_k} \odot_{\mathbb{S}} z_i^{w_\ell} \\
&= \bigoplus_{s<i<t+1}{}_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,i}(z), (v_1 \cdots v_{k-1}) * (w_1 \cdots w_{\ell-1})\Big\rangle \odot_{\mathbb{S}} z_i^{[v_k\ w_\ell]} \\
&= \big\langle \mathrm{ISS}^{\mathbb{S}}_{s,t}(z), (v_1 \cdots v_{k-1}) * (w_1 \cdots w_{\ell-1})[v_k\ w_\ell]\big\rangle.
\end{aligned}
$$

Combining those terms, we get

$$\begin{aligned}
&\left\langle \mathsf{ISS}^{\mathbb{S}}_{s,t}(z), v \right\rangle \odot_{\mathbb{S}} \left\langle \mathsf{ISS}^{\mathbb{S}}_{s,t}(z), w \right\rangle \\
&= \Big\langle \mathsf{ISS}^{\mathbb{S}}_{s,t}(z), (v \star (w_1 \cdots w_{\ell-1}))w_\ell + ((v_1 \cdots v_{k-1}) \star w)v_k \\
&\qquad + (v_1 \cdots v_{k-1} v_{k-1}) * (w_1 \cdots w_{\ell-1})[v_k w_\ell] \Big\rangle \\
&= \left\langle \mathsf{ISS}^{\mathbb{S}}_{s,t}(z), v * w \right\rangle.
\end{aligned}$$

□

3. Quasisymmetric expressions over a semiring

The aim of this section is to study the coefficients, i.e., iterated sums, used in the definition of the $\mathsf{ISS}^{\mathbb{S}}$, (19), as formal power series expressions. Analogous to the classical case, this results in the definition of the notion of quasisymmetric expressions defined over a semiring. These are formal series with coefficients in $\mathbb{S}$ which have a particular symmetry property defined below. When considered over a commutative ring, their siblings form the well studied Hopf algebra of quasisymmetric functions with the monomial quasisymmetric functions as one of many bases [43, 47, 29]. As we shall see, working over a semiring leads to rather minor changes compared to the classical theory of quasisymmetric functions. This stems from the fact that most properties only rely on the *index set* (i.e., the totally ordered set of integers). However, it turns out that the monomial basis is the only reasonable one, Remark 3.7.

In the following we denote by

$$\mathbb{S}[[X_1, X_2, X_3, \ldots]]$$

the commutative $\mathbb{S}$-semialgebra of **formal power series expressions** in commuting ordered indeterminates $X := \{X_1, X_2, X_3, \ldots\}$ with coefficients in $\mathbb{S}$. We write monomials in these variables in the usual form

$$X_{s_1}^{\alpha_1} \cdots X_{s_n}^{\alpha_n}, \quad n \geq 0, \quad \alpha_1, \ldots, \alpha_n \geq 1,$$

but note that this is – of course – *just a formal expression*, so that we might as well have written $X_{s_1}^{\odot_{\mathbb{S}}\alpha_1} \odot_{\mathbb{S}} \cdots \odot_s X_{s_n}^{\odot_{\mathbb{S}}\alpha_n}$. The **degree** of such a monomial is $|X_{s_1}^{\alpha_1} \cdot \cdots \cdot X_{s_n}^{\alpha_n}| := \alpha_1 + \cdots + \alpha_k$. Similar to the power series semialgebra in noncommuting variables of the previous section, elements $P \in \mathbb{S}[[X_1, X_2, X_3, \ldots]]$ can be considered as formal power series

$$P = \sum_m c_m m,$$

where $c_m \in \mathbb{S}$ and the sum is over formal commutative monomials in the indeterminates $X_1, X_2, \ldots$. The linear structure follows as for the case of noncommutative variables and the multiplicative structure is induced from the product of formal monomials (Cauchy product). We shall write $P(m) := c_m$. By small abuse of notation, we let $\mathbb{S}[[X_1, X_2, X_3, \ldots]]$ contain **only power series of bounded degree**, i.e., for $P \in \mathbb{S}[[X_1, X_2, X_3, \ldots]]$ there is $N \geq 0$ such that for all monomials m with $|m| \geq N$, $P(m) = \mathbf{0}_{\mathbb{S}}$. The subset with power series of finite support is denoted $\mathbb{S}[X_1, X_2, \ldots]$ and forms the space of **formal polynomial expressions**.

Definition 3.1. An element $P \in \mathbb{S}[[X_1, X_2, X_3, \ldots]]$ is a **quasisymmetric expression** if for all $\alpha_1, \ldots, \alpha_n$, $0 < s_1 < \cdots < s_n$ and $0 < t_1 < \cdots < t_n$ the coefficients of

$$X_{s_1}^{\alpha_1} \cdot \cdots \cdot X_{s_n}^{\alpha_n} \text{ and } X_{t_1}^{\alpha_1} \cdot \cdots \cdot X_{t_n}^{\alpha_n},$$

coincide.

Define the **monomial quasisymmetric expression** indexed by $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{N}^k_{\geq 1}$, $k \geq 0$ as

$$M_\alpha := \sum_{1 \leq t_1 < \cdots < t_k < +\infty} X_{t_1}^{\alpha_1} \cdot \cdots \cdot X_{t_k}^{\alpha_k}.$$

Lemma 3.2. *The space of all quasisymmetric expressions is a sub-semialgebra of* $\mathbb{S}[[X_1, X_2, \ldots]]$. *We denote it by* $\mathrm{QSym}_{\mathbb{S}}$.

Proof. Immediate. □

Remark 3.3. 1 *We can naturally evaluate a formal monomial at values in a commutative semiring, e.g. for* $z_1, z_2 \in \mathbb{S}$,

$$X_1^3 X_2^5 \,|_{X_1 = z_1, X_2 = z_2} = z_1^{\odot_{\mathbb{S}} 3} \odot_{\mathbb{S}} z_2^{\odot_{\mathbb{S}} 5}.$$

The iterated-sums in the definition of the $\mathrm{ISS}^{\mathbb{S}}$, (19)*, then amount, in the one-dimensional case, to evaluation of the monomial quasisymmetric functions expression,*

$$\left\langle \mathrm{ISS}^{\mathbb{S}}_{s,t}(z), [1^{\alpha_1}] \cdots [1^{\alpha_k}] \right\rangle_{\mathbb{S}} = M_\alpha \,|_{X_1 = z_1, X_2 = z_2, \ldots}.$$

2 *There is a straightforward extension to "multidimensional" quasisymmetric functions, compare* [16, Remark 3.5]*. We omit the details for brevity.*

Example 3.4. *The simplest, non-trivial quasisymmetric expression is*

$$P(m) := \begin{cases} \mathbf{1}_{\mathbb{S}} & \text{if } m = X_i \quad \text{for some } i \\ \mathbf{0}_{\mathbb{S}} & \text{else} \end{cases},$$

or, written as formal sum,

$$P = \sum_{0 < i < +\infty} X_i = \sum_{0 < i < +\infty} \mathbf{1}_{\mathbb{S}} X_i.$$

Another example is given by

$$\sum_{0 < i_1 < i_2 < +\infty} X_{i_1} X_{i_2}^2.$$

There are different concepts of linear independence in semimodules. The reader is referred to [2] for an overview. The strongest one seems to be

Definition 3.5. Let M be an $\mathbb{S}$-semimodule. A family of elements $v_i \in M$, $i \in I$, is **linearly independent** (in the Gondran–Minoux sense) if there do *not* exist non-empty finite sets $J, K \subset I$, $J \cap K = \emptyset$, and $\alpha_j, \beta_k \in \mathbb{S}, j \in J, k \in K$, all non-zero, with

$$\sum_{j \in J} \alpha_j v_j = \sum_{k \in K} \beta_k v_k.$$

We then have, as expected, that the monomial quasisymmetric functions are a basis for $\mathrm{QSym}_{\mathbb{S}}$.

Proposition 3.6. *The family* M_α *is linearly independent and is spanning* $\mathrm{QSym}_{\mathbb{S}}$.

Proof. **Linear independence**
For $\alpha \neq \beta$ the support of M_α and M_β are disjoint. This gives linear independence.

Spanning property
Let $Q \in \mathrm{QSym}_{\mathbb{S}}$. If Q is the zero power series, we are done. Otherwise, take a monomial $m =$

$X_{t_1}^{\alpha_1} \cdot \dots \cdot X_{t_k}^{\alpha_k}$ in Q with non-zero coefficient $c \in \mathbb{S}$ (i.e., considering Q as a function on monomials, $Q(m) = c$). Then

$$Q = Q' + cM_\alpha$$

with $Q' \in \mathrm{QSym}_{\mathbb{S}}$. Since Q has finite degree we can repeat this finitely many times to see that Q is a linear combination of monomial quasisymmetric functions. □

Remark 3.7. *The space of quasisymmetric functions over a commutative ring has several relevant linear bases, for example the* fundamental basis*,* [47, (2.13)].

Over a general commutative semiring the monomial basis is the only *basis. Indeed, let us work over the tropical semiring and let G_i, $i \in I$, be another basis. Let M_α be some monomial basis element. Then we can write*

$$M_\alpha = \sum_{j=1}^{n} c_j G_{i_j},$$

with $c_j \in \mathbb{S}_{\text{min+}} \setminus \{\mathbf{0}_{\text{min+}}\}$. Let m be any monomial not appearing in M_α. Then

$$\mathbf{0}_{\text{min+}} = \bigoplus_{j=1}^{n}{}_{\text{min+}}\; c_j \odot_{\text{min+}} G_{i_j}(m),$$

i.e.

$$+\infty = \min_j \{c_j G_{i_j}(m)\}.$$

Since the c_j are not equal to $+\infty$, m does not appear in any of the G_{i_j}. Hence $n = 1$ and $c_1 G_{i_1} = M_\alpha$. Hence the basis $(G_i)_i$ contains, up to multiplicative factors, the monomial basis. Since this subset already forms a basis, the basis $(G_i)_i$ is equal, up to multiplicative factors, to the monomial basis.

3.1. **Invariance to inserting zeros.**

We will now show that $\mathrm{QSym}_{\mathbb{S}}$ can be characterized by invariance to "inserting zeros". For $n \geq 1$ define the commutative $\mathbb{S}$-semialgebra morphism

$$\mathrm{zero}_n : \mathbb{S}[[X_1, X_2, \dots]] \to \mathbb{S}[[X_1, X_2, \dots]],$$

induced from the following map on $X_1, X_2, \dots$

$$\mathrm{zero}_n(X_i) = \begin{cases} X_i & i < n \\ \mathbf{0}_{\mathbb{S}} & i = n \\ X_{i-1} & i > n \end{cases}.$$

If we consider elements of $P \in \mathbb{S}[[X_1, X_2, \dots]]$ as $\mathbb{S}$-valued functions on monomials, this gives, with $m = X_{t_1}^{\odot_{\mathbb{S}} \alpha_1} \odot_{\mathbb{S}} \dots \odot_{\mathbb{S}} X_{t_n}^{\odot_{\mathbb{S}} \alpha_n}$, that $\mathrm{zero}_i(P)(m)$ is equal to

$$\begin{cases} P(m) & i > t_n \\ \mathbf{0}_{\mathbb{S}} & i \in \{t_1, \dots, t_n\} \\ P\Big(X_{t_1}^{\alpha_1} \cdot \dots \cdot X_{t_{k-1}}^{\alpha_{k-1}} X_{t_k - 1}^{\alpha_k} X_{t_{k+1}-1}^{\alpha_{k+1}} \dots X_{t_n - 1}^{\alpha_n}\Big) & t_{k-1} < i < t_k, k \in \{2, \dots, n\} \\ P\Big(X_{t_1 - 1}^{\alpha_1} \dots X_{t_n - 1}^{\alpha_n}\Big) & i < t_1. \end{cases}$$

Example 3.8.

$$\begin{aligned} \mathrm{zero}_9(X_2 X_6^7 X_8^5) &= X_2 X_6^7 X_8^5 \\ \mathrm{zero}_8(X_2 X_6^7 X_8^5) &= \mathbf{0}_{\mathbb{S}} \\ \mathrm{zero}_3(X_2 X_6^7 X_8^5) &= X_2 X_5^7 X_7^5 \\ \mathrm{zero}_1(X_2 X_6^7 X_8^5) &= X_1 X_5^7 X_7^5. \end{aligned}$$

Theorem 3.9. *A power series expression $P \in \mathbb{S}[[X_1, X_2, \dots]]$ is in* $\mathrm{QSym}_{\mathbb{S}}$ *if and only if*

$$\mathrm{zero}_n P = P \qquad \forall n \geq 1.$$

Proof. $\Rightarrow$: Immediate.

$\Leftarrow$: We begin with an example. If P is invariant in the prescribed sense, then for any monomial m

$$P(m) = \Big((\mathrm{zero}_1)^2 \mathrm{zero}_4 P\Big)(m).$$

We apply this to $m = X_3^7 X_5$ to get

$$P(X_3^7 X_5) = \Big((\mathrm{zero}_1)^2 \mathrm{zero}_4 P\Big)(X_3^7 X_5) = P(X_1^7 X_2).$$

Since the time points $3, 5$ were arbitrary, the coefficients of the monomials $X_{t_1}^7 X_{t_2}$, $1 \leq t_1 < t_2 < +\infty$, must coincide.

The general proof follows analogously: let $n \geq 1$, $\alpha_1, \dots, \alpha_n$, $0 < t_1 < \dots < t_n$ be given. We then have

$$\begin{aligned} &P\Big(X_{t_1}^{\alpha_1} \dots X_{t_n}^{\alpha_n}\Big) \\ &= \Big((\mathrm{zero}_1)^{t_1-1} (\mathrm{zero}_{t_1+1})^{t_2-t_1-1} \cdots (\mathrm{zero}_{t_{n-1}+1})^{t_n-t_{n-1}-1} P\Big)\Big(X_{t_1}^{\alpha_1} \dots X_{t_n}^{\alpha_n}\Big) \\ &= P\big(X_1^{\alpha_1} \dots X_n^{\alpha_n}\big). \end{aligned}$$

Since n, $\alpha_1, \dots, \alpha_n$ and $t_1, \dots, t_n$ were arbitrary this shows that P is quasisymmetric. □

From Theorem 3.9 we get the following consequence.

Corollary 3.10. $\mathrm{ISS}^{\mathbb{S}}(z)_{0,\infty}$ *is invariant to inserting* $\mathbf{0}_{\mathbb{S}}$ *into* z.

4. Time warping invariants in an idempotent semiring

Example 2.7 together with Corollary 3.10 shows one way to obtain time warping invariants of a real valued time series. This does not cover the invariant (2) though.

Since $\mathbb{R} \subset \mathbb{R} \cup \{+\infty\}$, and since the tropical semiring is idempotent we can also calculate $\mathrm{ISS}^{\mathbb{S}}(x)$ on a real-valued time series that is eventually constant. Recall that $\mathrm{ISS}^{\mathbb{S}}(x) = \mathrm{ISS}^{\mathbb{S}}_{0,\infty}(x)$. Since

$$\Big\langle \mathrm{ISS}^{\mathbb{S}}(x), [1] \Big\rangle = \min_i x_i,$$

this includes the invariant (2). But, as is quickly seen, most coefficients are *not* invariant to time warping. To wit,

$$\Big\langle \mathrm{ISS}^{\mathbb{S}}(x), [1][1] \Big\rangle = \min_{i_1 < i_2} \{x_{i_1} + x_{i_2}\}, \tag{23}$$

gives, for,

$$\begin{aligned} x &= (1, -3, 2, 2, \dots) \\ x' &= (1, -3, -3, 2, 2, \dots), \end{aligned}$$

the values -2 and -6 respectively.

It turns out that if we change the strict inequality over point in time in (23) into a weak, or non-strict inequality, namely

$$\min_{i_1 \leq i_2} \{x_{i_1} + x_{i_2}\},$$

then we *do* get a time warping invariant. In this section we would like to spell out how this works in general.

Assume that $\mathbb{S}$ is an idempotent semiring. Let z be a time series with values in $\mathbb{S}$, that is eventually constant. Define for $1 \leq s < t \leq +\infty$,

$$\Big\langle \mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z), w \Big\rangle := \bigoplus_{s<j_1\leq j_2\leq\cdots\leq j_k<t+1}^{\mathbb{S}} z_{j_1}^{\odot_{\mathbb{S}} w_1} \odot_{\mathbb{S}} \cdots \odot_{\mathbb{S}} z_{j_k}^{\odot_{\mathbb{S}} w_k}, \tag{24}$$

where the possibly infinite sum is well-defined, since $\mathbb{S}$ is idempotent and z is eventually constant. As before, we write $\mathrm{ISS}^{\mathbb{S},\mathrm{idem}}(z) = \mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{0,+\infty}(z)$.

The following lemma is immediate.

Lemma 4.1. $\mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z)$ *is invariant to* ***time warping****. That is, define for $n \geq 1$ the time series $\tau_n(z)$ as*

$$\tau_n(z)_j := \begin{cases} z_j & j \leq n \\ z_{j-1} & j > n. \end{cases}$$

Then, for all $n \geq 1$:

$$\mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(\tau_n(z)) = \mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z).$$

Lemma 4.2. $\mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z)$ *is a shuffle character, i.e.*

$$\Big\langle \mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z), v \Big\rangle \odot_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z), w \Big\rangle = \Big\langle \mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z), v \sqcup\!\sqcup\, w \Big\rangle.$$

Example 4.3. *Using, for example, the computation in* (10)*, we see that*

$$\begin{aligned} &\Big\langle \mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z), [1^7][1^3] \Big\rangle \odot_{\mathbb{S}} \Big\langle \mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z), [1^5] \Big\rangle \\ &= \Big\langle \mathrm{ISS}^{\mathbb{S},\mathrm{idem}}_{s,t}(z), [1^7][1^3][1^5] + [1^7][1^5][1^3] + [1^5][1^7][1^3] \Big\rangle, \end{aligned}$$

where we used idempotency of $\oplus_{\mathbb{S}}$.

Proof. The proof follows analogously to the one of Lemma 2.11. Owing to idempotency, for $f_i, g_i \in \mathbb{S}$, $i = s+1, \ldots, t$, (6) becomes

$$\Big(\bigoplus_{s<i\leq t}^{\mathbb{S}} f_i \Big) \odot_{\mathbb{S}} \Big(\bigoplus_{s<j\leq t}^{\mathbb{S}} g_j \Big) = \Big(\bigoplus_{s<i\leq j\leq t}^{\mathbb{S}} f_i \odot_{\mathbb{S}} g_j \Big) \oplus_{\mathbb{S}} \Big(\bigoplus_{s<j\leq i\leq t}^{\mathbb{S}} f_i \odot_{\mathbb{S}} g_j \Big).$$

This leads to the last term in (22) not being present and hence to a shuffle product instead of a quasi-shuffle product. ◻

We note that, in the tropical semiring, $\mathrm{ISS}^{\mathbb{R}_{\min+},\mathrm{idem}}$ is very degenerate, in the sense that, in the one-dimensional case,

$$\big\langle \mathrm{ISS}^{\mathbb{R}_{\min+},\mathrm{idem}}(z), [1^{a_1}]\cdots[1^{a_n}] \big\rangle = \big\langle \mathrm{ISS}^{\mathbb{R}_{\min+},\mathrm{idem}}(z), [1^{a_1+\cdots+a_n}] \big\rangle.$$

To get a more interesting object we can allow powers in $\mathbb{Z} \setminus \{0\}$ (instead of just $\mathbb{N}_{\geq 1}$), e.g.

$$\begin{aligned} \big\langle \mathrm{ISS}^{\mathbb{R}_{\min+},\mathrm{idem}}(z), [1^{-3}][1^5] \big\rangle &:= \bigoplus_{0<j_1\leq j_2}^{\mathbb{R}_{\min+}} z_{j_1}^{\odot_{\mathbb{R}_{\min+}} -3} \odot_{\mathbb{R}_{\min+}} z_{j_2}^{\odot_{\mathbb{R}_{\min+}} 5} \\ &= \min_{0<j_1\leq j_2} \{-3z_{j_1} + 5z_{j_2}\}. \end{aligned}$$

Without proof we state.

Proposition 4.4. *Let $\mathbb{R}_{\min+}$ be the tropical semiring. Define for $w \in A^*$, where $A = \mathbb{Z} \setminus \{0\}$,*

$$\left\langle \mathrm{ISS}^{\mathbb{R}_{\min+},(\!(\mathrm{idem})\!)}(z), w \right\rangle := \bigoplus_{\mathbb{R}_{\min+}} {}_{s<j_1\le j_2\le\cdots\le j_k\le t} \; z_{j_1}^{\odot_{\mathbb{R}_{\min+}} w_1} \odot_{\mathbb{R}_{\min+}} \cdots \odot_{\mathbb{R}_{\min+}} z_{j_k}^{\odot_{\mathbb{R}_{\min+}} w_k},$$

Then:

1 $\mathrm{ISS}^{\mathbb{R}_{\min+},(\!(\mathrm{idem})\!)}$ *is a shuffle character.*
2 $\mathrm{ISS}^{\mathbb{R}_{\min+},(\!(\mathrm{idem})\!)}$ *satifies Chen's identity.*
3 $\mathrm{ISS}^{\mathbb{R}_{\min+},(\!(\mathrm{idem})\!)}$ *is time warping invariant.*

Remark 4.5. *The iterated-sums signature over a field of characteristic* 0 *is, via the Hoffman exponential, in bijection to a certain* iterated-integrals *signature,* [16, Theorem 5.3]. *The iterated-sums signature satisfies a quasi-shuffle identity, whereas the iterated-integrals signature is a shuffle character. In fact, there is a whole family of signature-like maps, indexed by $\theta \in (-1, 1)$ obtained by composing the iterated sums signature with some linear transformation $A_{\theta\to 1}$ which generalize Hoffman's exponential (it being the case $\theta = 0$), see* [15, Remark 2.3].

When working over an idempotent semiring, however, only the cases $\theta = -1$ and $\theta = 0$ are well defined, and both maps coincide.

5. Conclusion

In (19) we introduced the iterated-sums signature, $\mathrm{ISS}^{\mathbb{S}}(z)$, over a commutative semiring $\mathbb{S}$. It stores all iterated sums (taken in the semiring) of a multidimensional time series $z = (z_1, z_2 \ldots)$ with entries $z_i \in \mathbb{S}^d$. As in the case over commutative rings, this object satisfies Chen's identity (Lemma 2.8) which here as well allows for an efficient computation. It also satisfies the quasi-shuffle identity (Lemma 2.11) that is, it behaves like a group-like element. Unlike for the usual ISS there is no proper Hopf algebra structure available here and it is in general *not* possible to take the logarithm of the signature.

In the one-dimensional case over commutative rings, the entries of the iterated-sums signature correspond to the evaluation of certain formal power series, namely quasisymmetric functions. Here, this is also true (Section 3), though it is more appropriate to speak of quasisymmetric *expressions*, since polynomial expressions over a semiring are *not* in one-to-one correspondence with polynomial functions.

In order to explicitly cover the expression (2) from the Introduction, we looked at the special case of idempotent semirings in Section 4.

Open questions

- The iterated-sums signature over the reals has a close connection to discrete control theory [33]. In the setting of the max-plus semiring:

  Is there a relation to discrete control theory in that semiring [12]*?*

- The ISS over a commutative ring contains, owing to the quasi-shuffle identity, many redundant entries. Working with the log-signature removes these redundancies. Over a general commutative semiring we cannot take the logarithm of the signature, so an open question is

  How to extract the "minimal" information contained in the signature?

- As seen in Remark 2.9, owing to the lack of additive inverses, Chen's identity only works in "one direction".

  Is there a way (with maybe larger object) of obtaining a general Chen's identity?

- Multidimensional time series are explicitly covered by the present work. Just as over the reals, this amounts to projecting the time series to coordinates before calculating the iterated-sums.

  In the semiring setting a more interesting approach seems possible. Indeed, one can turn a multidimensional real-valued time series into a one-dimensional semiring-valued time series. One example is via the map

$$\begin{aligned} \mathbb{R}^d &\longrightarrow \text{ bounded convex polytopes} \\ x &\mapsto \{x\}. \end{aligned}$$

  The resulting time series can then be considered in the semiring of polytopes, point 9. in Example 2.2. One can hope for tractable calculation, using the relation to the algebra of polynomials, [52, Theorem 2.25].
- Chen's identity, Lemma 2.8, applied to time points $0, t, t+1$ reads as

$$\begin{aligned} \left\langle \mathrm{ISS}^{\mathbb{S}}_{0,t+1}(z), w \right\rangle_{\mathbb{S}} &= \left\langle \mathrm{ISS}^{\mathbb{S}}_{0,t}(z), w \right\rangle_{\mathbb{S}} \oplus_{\mathbb{S}} \left(\left\langle \mathrm{ISS}^{\mathbb{S}}_{0,t}(z), w_1 \ldots w_{n-1} \right\rangle_{\mathbb{S}} \odot_{\mathbb{S}} \left\langle \mathrm{ISS}^{\mathbb{S}}_{t,t+1}(z), w_n \right\rangle_{\mathbb{S}} \right) \\ &= \left\langle \mathrm{ISS}^{\mathbb{S}}_{0,t}(z), w \right\rangle_{\mathbb{S}} \oplus_{\mathbb{S}} \left(\left\langle \mathrm{ISS}^{\mathbb{S}}_{0,t}(z), w_1 \ldots w_{n-1} \right\rangle_{\mathbb{S}} \odot_{\mathbb{S}} z_t^{\odot_{\mathbb{S}} w_n} \right), \end{aligned}$$

  where we use the notation of (19). This allows an iterative calculation of this value, with total cost of order $O(n \cdot t)$. This can be seen as a special case of dynamic programming.

  Is there a deeper connection to the dynamic programming literature?
- The iterated-integrals signature has been investigated from the perspective of algebraic geometry in [3].

  Is there interesting tropical algebraic geometry, that can be done on the objects introduced in this work?

References

[1] M. Akian, R. Bapat, and S. Gaubert, “Max-plus algebra,” *Handbook of linear algebra* **39** (2006) .

[2] M. Akian, S. Gaubert, and A. Guterman, “Linear independence over tropical semirings and beyond,” in *Tropical and idempotent mathematics*, vol. 495 of *Contemp. Math.*, pp. 1–38. Amer. Math. Soc., Providence, RI, 2009.

[3] C. Amendola, P. Friz, and B. Sturmfels, “Varieties of signature tensors,” in *Forum of Mathematics, Sigma*, vol. 7, Cambridge University Press. 2019.

[4] R. Antoine, F. Perera, and H. Thiel, “Tensor products and regularity properties of Cuntz semigroups,” *Mem. Amer. Math. Soc.* **251** no. 1199, (2018) viii+191.

[5] F. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat, “Synchronization and linearity: an algebra for discrete event systems,”.

[6] E. Baldwin and P. Klemperer, “Tropical geometry to analyse demand,” *Unpublished paper.[281]* (2013) .

[7] S. L. Bloom and Z. Ésik, “Iteration theories,” in *Iteration Theories*, pp. 159–213. Springer, 1993.

[8] M. Brandenburg, “Tensor categorical foundations of algebraic geometry,” `arXiv:1410.1716 [math.AG]`.

[9] M. Brandenburg, *Einführung in die Kategorientheorie: Mit ausführlichen Erklärungen und zahlreichen Beispielen*. Springer-Verlag, 2016.

[10] P. Butkovič, *Max-linear systems: theory and algorithms*. Springer Science & Business Media, 2010.

[11] G. Carlsson and S. Kališnik Verovšek, “Symmetric and r-symmetric tropical polynomials and rational functions,” *J. Pure Appl. Algebra* **220** no. 11, (2016) 3610–3627.

[12] C. G. Cassandras and S. Lafortune, *Introduction to discrete event systems*. Springer, New York, second ed., 2008.

[13] K.-T. Chen, “Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula,” *Ann. of Math. (2)* **65** (1957) 163–178.

[14] M. Cygan, F. V. Fomin, Ł. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh, *Parameterized algorithms*, vol. 4. Springer, 2015.

[15] J. Diehl, K. Ebrahimi-Fard, and N. Tapia, “Iterated-sums signature, quasisymmetric functions and time series analysis,” *Sem. Lothar. Combin.* **84B.86** (2020) .

[16] J. Diehl, K. Ebrahimi-Fard, and N. Tapia, “Time-Warping Invariants of Multidimensional Time Series,” *Acta Appl. Math.* (2020) .

[17] M. Droste and W. Kuich, “Chapter 1: Semirings and formal power series,” in *Handbook of weighted automata*, Monogr. Theoret. Comput. Sci. EATCS Ser., pp. 3–28. Springer, Berlin, 2009.

[18] L. A. Duffaut Espinosa, W. S. Gray, and K. Ebrahimi-Fard, “A combinatorial Hopf algebra for nonlinear output feedback control systems,” *J. Algebra* **453** (2016) 609–643.

[19] L. A. Duffaut Espinosa, W. S. Gray, and K. Ebrahimi-Fard, “Discrete-time approximations of Fliess operators,” *Numer. Math.* **137** (2017) 35–62.

[20] K. Ebrahimi-Fard, F. Fauvet, and D. Manchon, “A comodule-bialgebra structure for word-series substitution and mould composition,” *J. Algebra* **489** (2017) 552–581.

[21] K. Ebrahimi-Fard, S. J. Malham, F. Patras, and A. Wiese, “Flows and stochastic Taylor series in Itô calculus,” *Journal of Physics A: Mathematical and Theoretical* **48** no. 49, (2015) 495202.

[22] J. Eisner, “Parameter Estimation for Probabilistic Finite-State Transducers,” in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 1–8. Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, July, 2002.

[23] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, “Deep learning for time series classification: a review,” *Data Min. Knowl. Discov.* **33** no. 4, (2019) 917–963.

[24] J. G. Fletcher, “A more general algorithm for computing closed semiring costs between vertices of a directed graph,” *Communications of the ACM* **23** no. 6, (Jun, 1980) 350–351.

[25] M. Fliess, “Fonctionnelles causales non linéaires et indéterminées non commutatives,” *Bulletin de la société mathématique de France* **109** (1981) 3–40.

[26] G. Flint, B. Hambly, and T. J. Lyons, “Discretely sampled signals and the rough Hoff process,” *Stochastic Process. Appl.* **126** no. 9, (2016) 2593–2614.

[27] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*. No. 120 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010.

[28] J. G. Gaines, “The algebra of iterated stochastic integrals,” *Stochastics Stochastics Rep.* **49** no. 3-4, (1994) 169–179.

[29] I. M. Gessel, “Multipartite P-partitions and inner products of skew Schur functions,” in *Combinatorics and algebra (Boulder, Colo., 1983)*, vol. 34 of *Contemp. Math.*, pp. 289–317. Amer. Math. Soc., Providence, RI, 1984.

[30] K. Gimpel and N. A. Smith, “Cube Summing, Approximate Inference with Non-Local Features, and Dynamic Programming without Semirings,” in *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, EACL ’09, pp. 318–326. Association for Computational Linguistics, USA, 2009.

[31] J. S. Golan, *Semirings and affine equations over them: theory and applications*, vol. 556. Springer Science & Business Media, 2013.

[32] J. Goodman, “Semiring parsing,” *Comput. Linguist.* **25** no. 4, (1999) 573–605.

[33] W. S. Gray, G. Venkatesh, and L. A. D. Espinosa, “Nonlinear system identification for multivariable control via discrete-time Chen–Fliess series,” *Automatica* **119** (2020) 109085.

[34] B. Hambly and T. J. Lyons, “Uniqueness for the signature of a path of bounded variation and the reduced path group,” *Ann. Math.* **171** no. 1, (2010) 109–167.

[35] S. Kališnik and D. Lešnik, “Symmetric polynomials in upper-bound semirings,” *Journal of Symbolic Computation* (2020) .

[36] S. Kališnik and D. Lešnik, “Symmetric polynomials in tropical algebra semirings,” *J. Symbolic Comput.* **93** (2019) 100–119.

[37] P. Kidger, P. Bonnier, I. Perez Arribas, C. Salvi, and T. Lyons, “Deep Signature Transforms,” in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d’ Alché-Buc, E. Fox, and R. Garnett, eds., pp. 3105–3115. Curran Associates, Inc., 2019.

[38] F. J. Király and H. Oberhauser, “Kernels for sequentially ordered data,” *J. Mach. Learn. Res.* **20** (2019) Paper No. 31, 45.

[39] J. Kohlas and N. Wilson, “Semiring induced valuation algebras: exact and approximate local computation algorithms,” *Artificial Intelligence* **172** no. 11, (2008) 1360–1399.

[40] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, pp. 1097–1105. 2012.

[41] W. Kuich and A. Salomaa, *Semirings, automata, languages*, vol. 5 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, Berlin, 1986.

[42] S. Liao, T. Lyons, W. Yang, and H. Ni, “Learning stochastic differential equations using RNN with log signature features,” `arXiv:1908.08286 [cs.LG]`.

[43] K. W. Luoto, *Quasisymmetric functions and their applications*. ProQuest LLC, Ann Arbor, MI, 2009. Thesis (Ph.D.)–University of Washington.

[44] T. J. Lyons, “Differential equations driven by rough signals.,” *Rev. Mat. Iberoam.* **14** no. 2, (1998) 215–310.

[45] D. Maclagan and B. Strumfels, *Introduction to tropical geometry*, vol. 161 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2015.

[46] S. MacLane, *Categories for the working mathematician*. Springer-Verlag, New York-Berlin, 1971. Graduate Texts in Mathematics, Vol. 5.

[47] C. Malvenuto and C. Reutenauer, “Duality between quasi-symmetric functions and the Solomon descent algebra,” *J. Algebra* **177** no. 3, (1995) 967–982.

[48] F. Marty, *Des Immersions Ouvertes et Des Morphismes Lisses en Géométrie Relative.* PhD thesis, Université Toulouse III - Laboratoire Emile Picard, 2009.

[49] M. Mohri, “Semiring frameworks and algorithms for shortest-distance problems,” *J. Autom. Lang. Comb.* **7** no. 3, (2002) 321–350.

[50] J. Norton and S. Spiroff, “The tropical semiring in higher dimensions,” *Involve* **11** no. 3, (2018) 477–488.

[51] L. Pachter and B. Sturmfels, “Tropical geometry of statistical models,” *Proceedings of the National Academy of Sciences* **101** no. 46, (2004) 16132–16137.

[52] L. Pachter and B. Sturmfels, *Algebraic statistics for computational biology*, vol. 13. Cambridge university press, 2005.

[53] R. Ree, “Lie elements and an algebra associated with shuffles,” *Ann. of Math. (2)* **68** (1958) 210–220.

[54] C. Reutenauer, “Free Lie algebras,” in *Handbook of algebra, Vol. 3*, vol. 3 of *Handb. Algebr.*, pp. 887–903. Elsevier/North-Holland, Amsterdam, 2003.

[55] E. Riehl, *Category theory in context.* Mineola, NY: Dover Publications, 2016.

[56] M. Saers and D. Wu, “Handling ties correctly and efficiently in Viterbi training using the Viterbi semiring,” in *Language and automata theory and applications*, vol. 10792 of *Lecture Notes in Comput. Sci.*, pp. 284–295. Springer, Cham, 2018.

[57] J. Sakarovitch, *Elements of automata theory*. Cambridge University Press, 2009.

[58] C. Toth, P. Bonnier, and H. Oberhauser, “Seq2Tens: An Efficient Representation of Sequences by Low-Rank Tensor Projections,” *arXiv preprint arXiv:2006.07027* (2020) .

[59] J. Worthington, “A bialgebraic approach to automata and formal language theory,” in *Logical foundations of computer science*, vol. 5407 of *Lecture Notes in Comput. Sci.*, pp. 451–467. Springer, Berlin, 2009.

Appendix A. Categorial view on semirings

The aim of this section is to give a brief overview of the categorical setting for semirings, semimodules etc. Good references on category theory are [46, 55]. (See also [9] (in German)). For the particularities of monoidal categories, we refer to [48], [8, Section 4.1].

Recall that a **monoidal category** is a category C with a bifunctor $\otimes\colon \mathrm{C}\times\mathrm{C} \to \mathrm{C}$, and an object $1 \in \mathrm{C}$ called the *unit* such that there exist natural isomorphisms

$$((-) \otimes (-)) \otimes (-) \cong (-) \otimes ((-) \otimes (-)), \quad 1 \otimes (-) \cong (-), \quad (-) \otimes 1 \cong (-)$$

and satisfy some consistency relations. Essentially, this means that there is a notion of "tensor product" internal to the category. A monoidal category is **symmetric** if furthermore it is endowed with a "braiding" or "twisting" natural isomorphism $\tau_{X,Y}\colon X \otimes Y \to Y \otimes X$ such that $\tau_{Y,X} \circ \tau_{X,Y} = \mathrm{id}_{X\otimes Y}$. Examples of symmetric monoidal categories include Vect_k for any field k and Mod_R for any commutative ring R. In both cases $\otimes$ corresponds to the internal tensor product.

In any monoidal category, the notion of monoid makes sense. A **monoid** on a monoidal category C is an object M in C together with two arrows $\mu\colon M\otimes M \to M$ and $u\colon 1 \to M$ satisfying an associativity and unitality condition ([46, Section VII.3]) Additionally, in a symmetric monoidal category, one can also impose a commutativity constraint and obtain commutative monoids. As an example, monoids in Vect_k correspond to algebras over vector spaces. Dually, a comonoid in C is a monoid in C^{op}.[7] As monoids on the category of vector spaces correspond to algebras, comonoids in Vect_k correspond to coalgebras. The category of monoids in C is denoted by $\mathrm{Mon}(\mathrm{C})$.[8] Likewise, the category of commutative monoids in C is denoted by $\mathrm{CMon}(\mathrm{C})$.

Proposition A.1 ([46, Theorem VII.3.2]). *Let* C *be a monoidal category with countable coproducts and assume that for each* $A \in \mathrm{C}$ *the functors* $- \otimes A, A \otimes -$ *preserve countable coproducts. Then the forgetful functor* $U\colon \mathrm{Mon}(\mathrm{C}) \to \mathrm{C}$ *has a left adjoint* $F\colon \mathrm{C} \to \mathrm{Mon}(\mathrm{C})$*. On an object* X *in* C*, the underlying object of* $F(X)$ *is*

$$U(F(X)) = \coprod_{n=0}^{\infty} X^{\otimes n}$$

in C*, with the monoidal structure given by the tensor product.*

Example A.2. *In the category* Vect_k*,*

$$F(X) = T(X) = \bigoplus_{n\geq 0} X^{\otimes n},$$

is the tensor algebra over X*.*

Suppose that C is a symmetric monoidal category and let R be a commutative monoid object in C. A **left R-module** is an object M in C with an arrow $\mu_M\colon R \otimes M \to M$ defining an action of R on M. One can also define right R-modules in the obvious way, but since R is commutative both notions coincide and we just call them **R-modules**. The category of R-modules is denoted by Mod_R.[9]

A nice example of symmetric monoidal category is Ab, the category of abelian groups. The tensor product on Ab is obtained form the cartesian (or direct) product of abelian groups by modding out the relations $(a_1, b) + (a_2, b) - (a_1 + a_2, b)$ and $(a, b_1) + (a, b_2) - (a, b_1 + b_2)$. We denote it by $\otimes$ as usual. The unit for this tensor product is the group of integers with addition. Monoid objects over this category correspond to rings, that is, $\mathrm{Mon}(\mathrm{Ab}) = \mathrm{Ring}$. Given an object R in $\mathrm{CMon}(\mathrm{Ab})$, that is, a commutative ring, the notion of R-module corresponds to the usual notion of module over a ring.

[7] One may also consider bimonoids and Hopf monoids.

[8] The arrows are given by arrows in C that respect the monoid structure, [48, Definition 1.2.9].

[9] With obvious definition of morphisms, compare [48, Definition 1.2.11].

For any given monoid object R in a symmetric monoidal category $(\mathsf{C}, \otimes, 1)$, the category of modules Mod_R is also a symmetric monoidal category when endowed with the tensor product $M \otimes_R N$ defined as the **coequalizer** of the two maps $M \otimes R \otimes N \rightrightarrows M \otimes N$ given by the action of R on M and N. The unit for this tensor product is R seen as a module over itself. Hence, the complete data is $(\mathsf{Mod}_R, \otimes_R, R)$.

Now, we consider the category of monoids $\mathsf{Mon}(\mathsf{Mod}_R)$. We call objects in this category **algebras over** R. Likewise, comonoids in Mod_R are called **coalgebras**.

Theorem A.3 ([48, Proposition 1.2.14])**.** *Let* C *be a symmetric monoidal category. If* C *is either complete or cocomplete, then so are* $\mathsf{CMon}(\mathsf{C})$ *and* Mod_R *for any commutative monoid* R *in* C*.*

A.1. **The category of semirings.** We apply the above construction to the category $\mathsf{CMon} := \mathsf{CMon}(\mathsf{Set})$ of commutative monoids. This is a symmetric monoidal category, with product given by the tensor product of commutative monoids ([4, Appendix B]).

Proposition A.4. *1. The category* CMon *is complete and cocomplete.*

2. For R *an object in* $\mathsf{CMon}(\mathsf{CMon})$*, the category* Mod_R *is complete and cocomplete.*

Proof. 1. A commutative monoid is nothing else than a commutative monoid object in the symmetric monoidal category $(\mathsf{Set}, \times, \{*\})$ which is known to be complete and cocomplete ([55, Theorem 3.2.6, Proposition 3.5.1]). Therefore, CMon is also complete and cocomplete by Theorem A.3.

2. This follows from point 1. and Theorem A.3. □

Proposition A.5. *1. In* CMon*, for each* $A \in \mathsf{CMon}$ *the functors* $- \otimes A$ *and* $A \otimes -$ *preserve countable coproducts.*

2. For R *an object in* $\mathsf{CMon}(\mathsf{CMon})$*, the tensor product in* Mod_R *preserves countable coproducts.*

Proof. 1. By [4, Appendix B], CMon is closed. In particular, this means that $- \otimes A$ is left adjoint to $\mathsf{Hom}(A, -)$. Therefore, $- \otimes A$ preserves colimits [55, Theorem 4.5.3], and in particular coporducts. By symmetry of the tensor product, the same is true for $A \otimes -$.

2. By proposition A.4, CMon is complete. Therefore, by [8, Theorem 4.1.10], for any commutative monoid R in CMon, Mod_R is a cocomplete symmetric monoidal category with unit R and tensor product $\otimes_R$. □

The category of commutative monoids $\mathsf{CMon}(\mathsf{CMon})$ corresponds to **commutative semirings**, i.e., commutative rings without negative elements.[10] For a fixed commutative semiring R, objects in the category Mod_R are known as **semimodules**. The notions of **semialgebra**, **semi-coalgebra** and **Hopf semialgebra** follow (as monoid, comonoid and Hopf monoid in Mod_R).

Theorem A.6. *Let* R *be a commutative semiring.*

1. There exists a left adjoint $F : \mathsf{Set} \longrightarrow \mathsf{Mod}_R$ *to the forgetful functor.* $F(D)$ *is the* ***free*** R***-module over*** D*.*

2. There exists a left adjoint $F' : \mathsf{Mod}_R \longrightarrow \mathsf{Mon}(\mathsf{Mod}_R)$ *to the forgetful functor.* $F'(X)$ *is the* ***free semialgebra over*** X*.*

Proof. Using Proposition A.4 and Proposition A.5 we can apply Proposition A.1. □

[10] Also called *rigs*, rings without **n**egative elements.

$F(D)$ corresponds to the free $\mathbb{S}$-semimodule over a set D indicated in Section 2. It is **free** in the following sense. For every map ϕ from D into an $\mathbb{S}$-semimodule M there exists a unique $\mathbb{S}$-semimodule morphism $\Phi : \mathbb{F} \to M$ such that the following diagram commutes.

$$\begin{array}{ccc} D & \hookrightarrow & \mathbb{F} \\ & \overset{\phi}{\searrow} & \downarrow \Phi \\ & & M \end{array}$$

Appendix B. Quasi-shuffle via surjections

We remark that one can express the inductively defined quasi-shuffle product (16) explicitly via certain surjections [21, 20]. Let k, k_1, k_2 be positive integers such that $\max(k_1, k_2) \leq k \leq k_1 + k_2$. We introduce the notion of (k_1, k_2)-quasi-shuffle of type k. These are surjections

$$f : \{1, \ldots, k_1 + k_2\} \twoheadrightarrow \{1, \ldots, k\},$$

such that $f(1) < \cdots < f(k_1)$ and $f(k_1 + 1) < \cdots < f(k_1 + k_2)$. Note that for $k = k_1 + k_2$ one recovers the usual (k_1, k_2)-shuffle bijections. The set of (k_1, k_2)-quasi-shuffles of type k is denoted by $\mathrm{qSh(k_1, k_2; k)}$. The latter permit to express the quasi-shuffle product (16) of two words in closed form

$$\Big(a_{j_1} \cdots a_{j_{k_1}}\Big) * \Big(a_{j_{k_1+1}} \cdots a_{j_{k_1+k_2}}\Big) = \sum_{\max(k_1,k_2) \leq k \leq k_1+k_2} \; \sum_{f \in \mathrm{qSh(k_1,k_2;k)}} a^f_{i_1} \cdots a^f_{i_k}, \tag{25}$$

with $a^f_{i_l} := \prod_{m \in f^{-1}(\{l\})} a_m$. Note that for $f \in \mathrm{qSh(k_1, k_2; k)}$ the set $f^{-1}(\{l\})$ contains either one or two elements. In the case of a trivial bracket operation on A, the right-hand side of (25) reduces to the well-known formula expressing shuffle products of words in terms of shuffle permutations. Concretely, returning to Lemma 2.11, we see that (25) implies for two words over the alphabet (18) and $z \in \mathbb{S}^{d,\mathbb{N}_{\geq 1}}_{\mathbf{0}_\mathbb{S}}$

$$\begin{aligned} &\Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,t}(z), \Big(a_{j_1} \cdots a_{j_{k_1}}\Big) * \Big(a_{j_{k_1+1}} \cdots a_{j_{k_1+k_2}}\Big)\Big\rangle \\ &= \sum_{\max(k_1,k_2) \leq k \leq k_1+k_2} \; \sum_{f \in \mathrm{qSh(k_1,k_2;k)}} \Big\langle \mathrm{ISS}^{\mathbb{S}}_{s,t}(z), a^f_{i_1} \cdots a^f_{i_k} \Big\rangle \\ &= \sum_{\max(k_1,k_2) \leq k \leq k_1+k_2} \; \sum_{f \in \mathrm{qSh(k_1,k_2;k)}} \; \bigoplus_{\mathbb{S}\, s < j_1 < j_2 < \cdots < j_k \leq t} z_{j_1}^{\odot_\mathbb{S} a^f_{i_1}} \odot_\mathbb{S} \cdots \odot_\mathbb{S} z_{j_k}^{\odot_\mathbb{S} a^f_{i_k}} \end{aligned}$$

where $a^f_{i_l} := \big[\prod_{m \in f^{-1}(\{l\})} a_m\big]$.