Forum of Mathematics, Sigma (2019), Vol. 7, e10, 54 pages
doi:10.1017/fms.2019.3

VARIETIES OF SIGNATURE TENSORS

CARLOS AMÉNDOLA[1], PETER FRIZ[2] and BERND STURMFELS[3]

[1] Technische Universität München, Germany;
email: carlos.amendola@tum.de

[2] Technische Universität Berlin and WIAS Berlin, Germany;
email: friz@math.tu-berlin.de

[3] MPI for Mathematics in the Sciences, Leipzig and UC Berkeley, Germany;
email: bernd@mis.mpg.de

Received 23 April 2018; accepted 22 December 2018

Abstract

The signature of a parametric curve is a sequence of tensors whose entries are iterated integrals. This construction is central to the theory of rough paths in stochastic analysis. It is examined here through the lens of algebraic geometry. We introduce varieties of signature tensors for both deterministic paths and random paths. For the former, we focus on piecewise linear paths, on polynomial paths, and on varieties derived from free nilpotent Lie groups. For the latter, we focus on Brownian motion and its mixtures.

2010 Mathematics Subject Classification: 14Q15, 60H99

1. Introduction

A *path* is a continuous map $X : [0, 1] \to \mathbb{R}^d$. Unless otherwise stated, we assume that the coordinate functions $X_1, X_2, \ldots, X_d$ are (piecewise) continuously differentiable functions. In particular, their differentials $dX_i(t) = X_i'(t)\,dt$ obey the usual rules of calculus. The *kth signature tensor* of X has order k and format $d \times d \times \cdots \times d$. This tensor is denoted by $\sigma^{(k)}(X)$. Its d^k entries $\sigma_{i_1 i_2 \cdots i_k}$ are real numbers. These are computed as iterated integrals:

$$\sigma_{i_1 i_2 \cdots i_k} = \int_0^1 \int_0^{t_k} \cdots \int_0^{t_3} \int_0^{t_2} dX_{i_1}(t_1)\, dX_{i_2}(t_2) \cdots dX_{i_{k-1}}(t_{k-1})\, dX_{i_k}(t_k). \quad (1)$$

We also write $dX(t) = (dX_1(t), dX_2(t), \ldots, dX_d(t))$ for the differential of the path X. The signature tensor can be written as an integral over tensor products of

such differentials:

$$\sigma^{(k)}(X) = \int_{\Delta} dX(t_1) \otimes dX(t_2) \otimes \cdots \otimes dX(t_k). \tag{2}$$

This integral is over the simplex $\Delta = \{(t_1, \ldots, t_k) \in \mathbb{R}^k : 0 \leqslant t_1 \leqslant \cdots \leqslant t_k \leqslant 1\}$.

Signature tensors were introduced in the 1950s by Chen [7–9]. They are central to the theory of rough paths [29], a revolutionary view on stochastic analysis, with important contributions by Terry Lyons, Martin Hairer, and many others (cf. [18, 20, 31, 32]).

The *step-n signature* of a path X is the sequence of its first n signature tensors:

$$\sigma^{\leqslant n}(X) = \big(1, \sigma^{(1)}(X), \sigma^{(2)}(X), \sigma^{(3)}(X), \ldots, \sigma^{(n)}(X)\big). \tag{3}$$

Here $\sigma^{(1)}(X) = X(1) - X(0)$ is a vector, $\sigma^{(2)}(X)$ is a matrix, and so forth. As n goes to infinity, we get a tensor series $\sigma(X)$, referred as *the signature* of X. Chen [9] proved that, up to a natural equivalence relation, each path X is determined by its signature $\sigma(X)$. Hambly and Lyons [24] refined Chen's theory and obtained quantitative results for paths of bounded variation. This was extended in [5] to arbitrary geometric rough paths. Algorithms for recovering paths X from their signature $\sigma(X)$ are of considerable interest for applications [10, 30, 33–35].

This paper concerns paths X that depend on finitely many parameters. We focus on two families: polynomial paths and piecewise linear paths. In the first case, each coordinate X_i is a polynomial in t of degree m. In the second case, X_i is piecewise linear with m pieces. Each family has dm parameters. The integral (1) is a polynomial of degree k in these parameters. The signature tensor $\sigma^{(k)}(X)$ is a polynomial function of the dm parameters. These tensors trace out an algebraic variety in the space of all tensors. Our aim is to study such varieties.

Stochastic analysis is the natural home for paths and their signatures. The present article is different. It examines paths and their signatures through the lens of algebraic geometry [13], commutative algebra [17], and representation theory [28]. Our inspiration is drawn from algebraic statistics [16] and tensor decomposition [25]. While the tensors we defined have their entries in $\mathbb{R}$, the associated varieties comprise tensors over the complex numbers $\mathbb{C}$. We work with homogeneous polynomials, so they are complex projective varieties.

This paper is organized as follows. In Section 2, we fix notation, we define our objects of study, and we compute first examples. We explain the passage from the real tensors $\sigma(X)$, where X ranges over paths of interest, to homogeneous ideals and complex projective varieties. The resulting distinction between *image* and *variety* is ubiquitous in applied algebraic geometry.

Section 3 concerns signature matrices. The case $k = 2$ is surprisingly rich and interesting. Theorem 3.3 shows that the signature images for polynomial paths of degree m agree with those for piecewise linear paths with m segments. This result does not generalize to higher order tensors. Theorem 3.4 characterizes the signature ideals in terms of rank constraints.

In Section 4, we investigate the algebraic underpinnings of the theory of rough paths [**20**, Section 7]. We study the free Lie algebra and the associated nilpotent Lie groups [**39**]. Theorem 4.10 characterizes their vanishing ideals by identifying explicit Gröbner bases. The order k component of the nilpotent Lie group defines a variety that contains the signature tensors of all deterministic paths. It serves as the ambient space for the varieties in the next section.

Section 5 studies signature varieties for piecewise linear paths and for polynomial paths. Chen's theorem [**7**] gives a parametric representation for the former, via tensor products of exponentials. The two cases are similar, but different (Theorem 5.7). Both families interpolate between the Veronese variety, which encodes line segments, and the universal variety (Theorem 5.6). Numerous instances are computed explicitly. As an application, we answer a question due to Lyons and Xu [**34**] concerning axis-parallel paths. We conclude the section by introducing the *rough Veronese*, a variety comprising signatures of rough paths.

In Section 6, we examine the dimensions and identifiability of our signature varieties. Theorem 6.1 shows that the universal variety $\mathcal{U}_{d,k}$ is identifiable. We conjecture the same for signature varieties of piecewise linear and polynomial paths and we offer supporting evidence.

In Section 7, we shift gears, by turning to random paths and their expected signatures. We focus on Brownian motion and mixtures of Brownian motion. This leads to a non-Abelian refinement of the moment varieties in [**2**, **3**]. The varieties of expected signatures are computed in several cases. These are not contained in the universal variety from Section 4. However, all expected signatures are contained in the convex hull of the universal variety.

The main contribution of this paper is a new bridge between applied algebraic geometry and stochastic analysis, where rough paths are encoded in signature tensors. Varieties of such tensors offer a concise representation of geometric data seen in numerous applications.

2. From integrals to projective varieties

In this section we develop the foundations for signature varieties of paths.

2.1. Computing iterated integrals. A first observation is that the integrals (1) do not change if the path X undergoes translation. We therefore always assume $X(0) = 0$. For such a path X in $\mathbb{R}^d$, consider the entries of the signature matrix

$\sigma^{(2)}(X) = (\sigma_{ij})$. They can be expressed as one-dimensional integrals:

$$\sigma_{ij} = \int_0^1 \int_0^t dX_i(s)\, dX_j(t) = \int_0^1 X_i(t) X_j'(t)\, dt.$$

The symmetrization of this $d \times d$ matrix is the matrix $X(1) \cdot X(1)^T$. This symmetric matrix has rank $\leqslant 1$. Conversely, every $d \times d$ matrix S whose symmetrization $S + S^T$ has rank $\leqslant 1$ and is positive semidefinite arises from some path X in $\mathbb{R}^d$. The set of such matrices is the *signature image*

$$\mathcal{U}_{d,2}^{\text{im}} = \left\{ \tfrac{1}{2} v \otimes v + Q : v \in \mathbb{R}^d, Q \in \mathbb{R}^{d \times d} \text{ skew-symmetric} \right\} \subset \mathbb{R}^{d \times d}. \quad (4)$$

This converse is not obvious. It follows, for instance, from Theorem 3.3 below.

We are interested in polynomial equations that hold on the signature image. Consider the case $d = 2$. The symmetric part of our 2×2 matrix has rank 1, so its determinant vanishes. This means that the following quadratic equation holds on the signature image:

$$(\sigma_{12} + \sigma_{21})^2 - 4\sigma_{11}\sigma_{22} = 0. \quad (5)$$

Let $\mathcal{U}_{2,2}^{\mathbb{R}}$ denote the set of solutions $S = (\sigma_{ij})$ to this equation in the space of matrices $\mathbb{R}^{2\times 2}$. This is the smallest variety containing the signature image. The latter is the *semialgebraic set*

$$\mathcal{U}_{2,2}^{\text{im}} = \left\{ S \in \mathcal{U}_{2,2}^{\mathbb{R}} : \sigma_{11} \geqslant 0 \text{ and } \sigma_{22} \geqslant 0 \right\}.$$

For an algebraic geometer, it is natural to view the σ_{ij} as homogeneous coordinates for the projective space $\mathbb{P}^3$. In that ambient space, the quadratic equation (5) defines the following surface, which will be our *signature variety* in this case:

$$\mathcal{U}_{2,2} = \left\{ [\sigma_{11} : \sigma_{12} : \sigma_{21} : \sigma_{22}] \in \mathbb{P}^3 \text{ such that (5) holds} \right\}.$$

See the next subsection for the relevant philosophy.

We next present explicit formulas for the entries of the signature tensors for two special paths in $\mathbb{R}^d$. It is instructive to check the computation of these integrals by hand for $d \leqslant 3$ and $k \leqslant 3$.

EXAMPLE 2.1 (The canonical axis path). Let C_{axis} be the path from $(0, 0, \ldots, 0)$ to $(1, 1, \ldots, 1)$ given by d linear steps in the unit directions $e_1, e_2, \ldots, e_d$ (in that order). One finds its signature tensor $\sigma^{(k)}(C_{\text{axis}})$ either directly by integration, or via Chen's Formula (38) for piecewise linear paths. The entry $\sigma_{i_1 i_2 \cdots i_k}$ is zero unless $i_1 \leqslant i_2 \leqslant \cdots \leqslant i_k$. In that case, it equals $1/k!$ times the number of distinct permutations of the string $i_1 i_2 \cdots i_k$. For example, if $k = 4$ then $\sigma_{1111} = \frac{1}{24}$, $\sigma_{1112} = \frac{1}{6}$, $\sigma_{1122} = \frac{1}{4}$, $\sigma_{1123} = \frac{1}{2}$, $\sigma_{1234} = 1$ and $\sigma_{1243} = 0$. ◇

We next specify a blueprint for paths whose coordinates are polynomials.

EXAMPLE 2.2 (The canonical monomial path). Let C_{mono} denote the monomial path $t \mapsto (t, t^2, t^3, \ldots, t^d)$. It also travels from $(0, 0, \ldots, 0)$ to $(1, 1, \ldots, 1)$, but now along the rational normal curve. The entries $\sigma_{i_1 i_2 \cdots i_k}$ of the signature tensor $\sigma^{(k)}(C_{\text{mono}})$ are given in (1), by integrating $i_1 i_2 \cdots i_k$ times $t_1^{i_1-1} t_2^{i_2-1} \cdots t_k^{i_k-1}$ over the simplex Δ. This computation reveals, for $1 \leqslant i_1, i_2, \ldots, i_k \leqslant d$,

$$\sigma_{i_1 i_2 \cdots i_k} = \frac{i_1}{i_1} \cdot \frac{i_2}{i_1 + i_2} \cdot \frac{i_3}{i_1 + i_2 + i_3} \cdots \frac{i_k}{i_1 + i_2 + \cdots + i_k}.$$

The general linear group $G = \mathrm{GL}(d, \mathbb{R})$ acts on paths in $\mathbb{R}^d$. It also acts on signatures, by the kth tensor power representation. This action extends to noninvertible $d \times d$ matrices. The map $X \mapsto \sigma^{(k)}(X)$ is equivariant for this action. This allows us to derive formulas for the following two special classes of paths.

Let $\mathbf{X} = (x_{ij})$ be a $d \times d$ matrix. Its jth column represents either a linear step or a vector of coefficients associated to the monomial t^j. Thus, $\mathbf{X}C_{\text{axis}}$ is a piecewise linear path with d pieces, and $\mathbf{X}C_{\text{mono}}$ is a polynomial path of degree d. Conversely, all such paths arise from some choice of matrix $\mathbf{X}$. The signature tensor $\sigma^{(k)}(\mathbf{X}C_\bullet)$ is the image of $\sigma^{(k)}(C_\bullet)$ under action by $\mathbf{X}$. For $k = 2$, this action is multiplying the matrix $\sigma^{(2)}(C_\bullet)$ by $\mathbf{X}$ on the left and by $\mathbf{X}^T$ on the right.

EXAMPLE 2.3 ($d = 3, k = 2$). The signature matrices of the canonical paths are

$$\sigma^{(2)}(C_{\text{axis}}) = \begin{pmatrix} \frac{1}{2} & 1 & 1 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad \sigma^{(2)}(C_{\text{mono}}) = \begin{pmatrix} \frac{1}{2} & \frac{2}{3} & \frac{3}{4} \\ \frac{1}{3} & \frac{2}{4} & \frac{3}{5} \\ \frac{1}{4} & \frac{2}{5} & \frac{3}{6} \end{pmatrix}. \tag{6}$$

The symmetric part of each signature matrix is the same constant rank 1 matrix:

$$\sigma^{(2)}(C_{\text{axis}}) + \sigma^{(2)}(C_{\text{axis}})^T = \sigma^{(2)}(C_{\text{mono}}) + \sigma^{(2)}(C_{\text{mono}})^T = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

We encode cubic paths $X = \mathbf{X}C_{\text{mono}}$ and 3-step paths $X = \mathbf{X}C_{\text{axis}}$ by 3×3 matrices

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}.$$

In each case, the map $X \mapsto \sigma^{(2)}(X)$ from paths to signature matrices is encoded by the quadratic map $\mathbf{X} \mapsto \mathbf{X}\sigma^{(2)}(C_\bullet)\mathbf{X}^{\mathbf{T}}$. This takes the space of 3×3 matrices

to itself. The image consists of certain 3×3 matrices S whose symmetric part $\frac{1}{2}(S + S^T)$ has rank 1. ◇

EXAMPLE 2.4 ($d = 2, k = 3$). Consider a general quadratic path in the plane $\mathbb{R}^2$:

$$X(t) = (x_{11}t + x_{12}t^2, x_{21}t + x_{22}t^2)^T = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \cdot \begin{pmatrix} t \\ t^2 \end{pmatrix}. \tag{7}$$

Its third signature ($k = 3$) is a $2 \times 2 \times 2$ tensor. The eight entries of $\sigma^{(3)}(X)$ are

$$\begin{aligned}
\sigma_{111} &= \tfrac{1}{6}(x_{11} + x_{12})^3 \\
\sigma_{112} &= \tfrac{1}{6}(x_{11} + x_{12})^2(x_{21} + x_{22}) + \tfrac{1}{60}(5x_{11} + 4x_{12})(x_{11}x_{22} - x_{21}x_{12}) \\
\sigma_{121} &= \tfrac{1}{6}(x_{11} + x_{12})^2(x_{21} + x_{22}) + \tfrac{1}{60}(2x_{12})(x_{11}x_{22} - x_{21}x_{12}) \\
\sigma_{211} &= \tfrac{1}{6}(x_{11} + x_{12})^2(x_{21} + x_{22}) - \tfrac{1}{60}(5x_{11} + 6x_{12})(x_{11}x_{22} - x_{21}x_{12}) \\
\sigma_{122} &= \tfrac{1}{6}(x_{11} + x_{12})(x_{21} + x_{22})^2 + \tfrac{1}{60}(5x_{21} + 6x_{22})(x_{11}x_{22} - x_{21}x_{12}) \\
\sigma_{212} &= \tfrac{1}{6}(x_{11} + x_{12})(x_{21} + x_{22})^2 - \tfrac{1}{60}(2x_{22})(x_{11}x_{22} - x_{21}x_{12}) \\
\sigma_{221} &= \tfrac{1}{6}(x_{11} + x_{12})(x_{21} + x_{22})^2 - \tfrac{1}{60}(5x_{21} + 4x_{22})(x_{11}x_{22} - x_{21}x_{12}) \\
\sigma_{222} &= \tfrac{1}{6}(x_{21} + x_{22})^3.
\end{aligned} \tag{8}$$

This is the transformation of the special tensor $\sigma^{(3)}(C_{\text{mono}})$ in Example 2.2 under the action of the 2×2 matrix $\mathbf{X} = (x_{ij})$. The symmetrization of $\sigma^{(3)}(X)$ is a tensor of rank one, because

$$\begin{aligned}
\sigma_{112} + \sigma_{121} + \sigma_{211} &= \sigma_{11} \sqcup 2 = \sigma_{11}\sigma_2 = \tfrac{1}{2}\sigma_1^2\sigma_2 = \tfrac{1}{2}(x_{11} + x_{12})^2(x_{21} + x_{22}) \\
\sigma_{122} + \sigma_{212} + \sigma_{221} &= \sigma_1 \sqcup 22 = \sigma_1\sigma_{22} = \tfrac{1}{2}\sigma_1\sigma_2^2 = \tfrac{1}{2}(x_{11} + x_{12})(x_{21} + x_{22})^2.
\end{aligned} \tag{9}$$

The *shuffle relations* in the middle will be defined and play an important role in Section 4. Of course, the equality of left- and right-hand sides in (9) can be checked directly from (8).

We wish to find relations among the eight entries of $\sigma^{(3)}(X)$ that hold for all quadratic paths. Hence, we want to eliminate the parameters x_{ij} from (8). This task, known as *implicitization*, arises frequently in computer algebra. The standard approach using *Gröbner bases* is explained in the undergraduate textbook by Cox *et al.* [**13**, Section 3.3]. The output of the implicitization is the prime ideal $P_{2,3,2}$ of polynomial relations among the σ_{ijk}.

The ideal $P_{2,3,2}$ is generated by nine quadratic relations. Three of them are

degree	ideal generator
(4, 2)	$3\sigma_{121}^2 - \sigma_{111}\sigma_{122} - 10\sigma_{112}\sigma_{211} + 2\sigma_{121}\sigma_{211} + 2\sigma_{211}^2 + 11\sigma_{111}\sigma_{212} - 7\sigma_{111}\sigma_{221}$
(3, 3)	$10\sigma_{122}\sigma_{211} - 4\sigma_{112}\sigma_{212} - 7\sigma_{121}\sigma_{212} + 2\sigma_{112}\sigma_{221} - 4\sigma_{121}\sigma_{221} + 3\sigma_{111}\sigma_{222}$
(2, 4)	$3\sigma_{212}^2 - 10\sigma_{122}\sigma_{221} + 2\sigma_{212}\sigma_{221} + 2\sigma_{221}^2 - \sigma_{112}\sigma_{222} + 11\sigma_{121}\sigma_{222} - 7\sigma_{211}\sigma_{222}$.

These three constraints on $2 \times 2 \times 2$ signature tensors are specific to quadratic paths in $\mathbb{R}^2$. The other six generators are valid for all paths in $\mathbb{R}^2$. Those are shown in Example 4.13. All nine quadrics can be found with any computer algebra system that offers Gröbner bases. The computations with ideals reported in this paper were carried out with the software `Macaulay2` [21]. ◇

2.2. The many lives of a variety. We now step back to contemplate the title of this article. *Signature tensors* were explained in the previous subsection. Our next goal is to explain what is meant by their *varieties*. This is important because our presentation rests on conventions from algebraic geometry that may be unfamiliar to readers from stochastic analysis. The key point is that the geometric objects we study are complex projective varieties along with their homogeneous ideals.

Consider an arbitrary map $\sigma : \mathbb{R}^p \to \mathbb{R}^q$ whose coordinates σ_i are homogeneous polynomials in p variables of degree k. We are interested in the image of σ. Algebraically, this is represented by an ideal F in the polynomial ring in q variables, namely the kernel of the associated ring homomorphism from that polynomial ring to the polynomial ring in p variables. The ideal F is *homogeneous* and *prime*. It is generated by homogeneous polynomials. Being prime means that the quotient modulo F is an integral domain [13, Section 5.1, Proposition 4].

The image of σ is a *semialgebraic set* in $\mathbb{R}^q$, denoted by $\mathcal{F}^{\text{im}} = \text{image}(\sigma)$. This set is often very complicated. One works with outer approximations by algebraic varieties, namely the real and complex zeros of the ideal F. The *complex variety* $\mathcal{F}^{\mathbb{C}} = \mathcal{V}^{\mathbb{C}}(F) \subset \mathbb{C}^q$ contains the *real variety* $\mathcal{F}^{\mathbb{R}} = \mathcal{V}^{\mathbb{R}}(F) \subset \mathbb{R}^q$. Both of the following two inclusions are usually strict:

$$\mathcal{F}^{\text{im}} \subset \mathcal{F}^{\mathbb{R}} \subset \mathcal{F}^{\mathbb{C}}. \tag{10}$$

EXAMPLE 2.5. Let $p = 2, q = 4$ and $\sigma : u \mapsto u \otimes u$ the map that takes a vector $u = (u_1, u_2)$ to the associated rank one matrix $u \otimes u = \begin{pmatrix} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{pmatrix}$. The ideal F is generated by the linear form $\sigma_{12} - \sigma_{21}$ and the quadratic form $\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21}$. The variety $\mathcal{F}^{\mathbb{R}}$ is a real surface in $\mathbb{R}^4$, and $\mathcal{F}^{\mathbb{C}}$ is a complex surface in $\mathbb{C}^4$. These are the symmetric 2×2 matrices of rank at most 1. The image $\mathcal{F}^{\text{im}}$ is the subset of matrices in $\mathcal{F}^{\mathbb{R}}$ that are positive semidefinite. ◇

From a geometric point of view, it is natural to work modulo scaling. This means we work in a projective space: two nonzero vectors are identified when they are parallel. Our degree k map σ induces a rational map of complex projective spaces $\sigma' : \mathbb{P}^{p-1}_{\mathbb{C}} \dashrightarrow \mathbb{P}^{q-1}_{\mathbb{C}}$, and this restricts to a rational map of real projective spaces $\sigma'' : \mathbb{P}^{p-1}_{\mathbb{R}} \dashrightarrow \mathbb{P}^{q-1}_{\mathbb{R}}$. Being *rational* means that σ' may not be

everywhere defined. If $u \in \mathbb{C}^p \backslash \{0\}$ satisfies $\sigma(u) = 0$ then σ' is not defined at the corresponding point $[u] \in \mathbb{P}^{p-1}_{\mathbb{C}}$. Such points are known as *base points*.

Every subset $\mathcal{U}$ of $\mathbb{R}^q$ or $\mathbb{C}^q$ gives rise to a subset in $\mathbb{P}^{q-1}_{\mathbb{R}}$ or $\mathbb{P}^{q-1}_{\mathbb{C}}$, namely the set $[\mathcal{U}]$ of lines that are spanned by vectors in $\mathcal{U}$. We write this in symbols as

$$[\mathcal{U}] = \{[u] : u \in \mathcal{U} \backslash \{0\}\}.$$

If $\mathcal{U}\backslash\{0\}$ is an affine variety then $[\mathcal{U}]$ is a projective variety [**13**, Section 8.2].

With this, the chain of inclusions in (10) is now taken into projective space:

$$\text{image}(\sigma'') = [\mathcal{F}^{\text{im}}] \subset [\mathcal{F}^{\mathbb{R}}] \subset [\mathcal{F}^{\mathbb{C}}]. \tag{11}$$

We similarly have the inclusion $\text{image}(\sigma') \subset [\mathcal{F}^{\mathbb{C}}]$ in the complex projective space $\mathbb{P}^{q-1}_{\mathbb{C}}$. If we recast Example 2.5 in the projective setting then the left inclusion in (11) becomes an equality. Indeed, for this example, $[\mathcal{F}^{\text{im}}] = [\mathcal{F}^{\mathbb{R}}]$ in $\mathbb{P}^3_{\mathbb{R}}$ because every 2×2 matrix of rank 1 is either positive or negative semidefinite. Similarly, $\text{image}(\sigma') = [\mathcal{F}^{\mathbb{C}}]$ holds in $\mathbb{P}^3_{\mathbb{C}}$.

The key word *Zariski closure* refers to the passage in (10) or (11) from images on the left to varieties on the right. The Zariski closure is the smallest variety containing a given set.

We have now introduced quite a few decorations for the symbol $\mathcal{F}$, and these decorations are important when making precise geometric statements pertaining to the three dichotomies

image versus *closure of image*, *real* versus *complex*, *affine* versus *projective*.

In the literature in applied algebraic geometry, these distinctions are often swept under the rug. This is a matter of convenience and simplicity. For the most part, our article will follow that convention. The symbol $\mathcal{F}$ will stand for the complex projective variety $[\mathcal{F}^{\mathbb{C}}]$. Some of our results will pertain to the images $\mathcal{F}^{\text{im}}$ and $[\mathcal{F}^{\text{im}}]$ over the real numbers, and in those cases the decorations for $\mathcal{F}$ will be used and highlighted. Thus $\mathbb{P}^{q-1}$ is a projective space over $\mathbb{C}$, and $\mathcal{F}$ is a subvariety in that space. It is tacitly understood that $\mathcal{F}$ arose from some specific polynomial map σ and that $\mathcal{F}$ is a proxy for $\mathcal{F}^{\text{im}} = \text{image}(\sigma)$. For instance, the map σ in Example 2.5 defines a quadratic curve $\mathcal{F}$ in a plane $\mathbb{P}^2$ that lives in $\mathbb{P}^3$.

Both of the fields $\mathbb{R}$ and $\mathbb{C}$ are peripheral when it comes to computations with software. In most applications, including ours, the coordinates of σ have coefficients in the field $\mathbb{Q}$ of rational numbers. All polynomials seen in this article have rational coefficients. In particular, the ideal F lives in a polynomial ring with coefficients in $\mathbb{Q}$. The variety $\mathcal{F}$ is represented through its ideal F.

We conclude that the title of this paper refers to the algebro-geometric study of a certain *signature map* σ. Its coordinates are homogeneous polynomials of

degree k. We seek to find the homogeneous prime ideal F of relations among these polynomials. This is the *signature ideal* F. Its *signature variety* $\mathcal{F}$ is the complex projective variety, unless explicitly stated otherwise. The associated *signature image* $\mathcal{F}^{\text{im}}$ is a semialgebraic set in real affine space.

Parameter identifiability is an important issue, and we shall address this in Section 6. In Example 2.5, the map $\sigma : \mathbb{R}^2 \to \mathcal{F}^{\text{im}}$ is two-to-one. Each nonzero matrix in $\mathcal{F}^{\text{im}}$ has two preimages (u_1, u_2) and $(-u_1, -u_2)$. On the other hand, the map $\sigma' : \mathbb{P}^1 \to \mathcal{F}$ is one-to-one. Every point on the quadratic curve $\mathcal{F}$ is uniquely represented by a point $[u_1 : u_2]$ in the parameter line $\mathbb{P}^1$. We express this by saying that the variety $\mathcal{F}$ is *rationally identifiable*.

Throughout this paper we use calligraphic letters for varieties and roman letters for ideals. So, if $\mathcal{F}$ denotes a variety then we write F for its radical ideal. If $\mathcal{F}$ is irreducible then F is prime, and if $\mathcal{F}$ is projective then F is homogeneous. For any field $\mathbb{K}$, such as $\mathbb{C}$ or $\mathbb{R}$ or $\mathbb{Q}$, we also use the notation $\mathcal{F}(\mathbb{K})$ instead of $\mathcal{F}^{\mathbb{K}}$ for the set of points with coordinates in $\mathbb{K}$.

We illustrate this for some objects that will appear in Section 4. The (n-truncated) free Lie algebra over $\mathbb{R}$ with d generators is denoted by $\text{Lie}^n(\mathbb{R}^d)$. This is a linear subspace in an affine space of dimension $d + d^2 + \cdots + d^n$, namely the truncated tensor algebra over $\mathbb{R}$. Its ideal lives in a polynomial ring over $\mathbb{Q}$ in $d+d^2+\cdots+d^n$ variables. It is generated by the linear forms $\sigma_{I \sqcup\!\sqcup J}$ in Lemma 4.1. The corresponding Lie group $\mathcal{G}^{\text{im}}_{d,n} = \exp(\text{Lie}^n(\mathbb{R}^d))$ is also an affine variety, with ideal $G_{d,n}$ generated by the quadratic polynomials $\sigma_I \sigma_J - \sigma_{I \sqcup\!\sqcup J}$. The intersection of this ideal with the polynomial subring for a fixed level k, where the variables σ_I satisfy $|I| = k$, is a homogeneous prime ideal $U_{d,k}$. The associated projective variety $\mathcal{U}_{d,k}$ is our universal variety. Its subset of real points, denoted $\mathcal{U}_{d,k}(\mathbb{R})$, contains the signature tensor $\sigma^{(k)}(X)$ for any deterministic path X in $\mathbb{R}^d$. This is the Chen–Chow Theorem 4.4.

3. Varieties of signature matrices

We here study varieties of $d \times d$ signature matrices $S = \sigma^{(2)}(X)$ of paths X in $\mathbb{R}^d$. Any $d \times d$ matrix S can be written uniquely as the sum of its symmetric part and skew-symmetric part:

$$S = P + Q, \quad \text{where } P = \tfrac{1}{2}(S + S^T) \text{ and } Q = \tfrac{1}{2}(S - S^T).$$

The $\binom{d+1}{2}$ entries p_{ij} of P and the $\binom{d}{2}$ entries q_{ij} of Q serve as coordinates on the space of matrices, either $\mathbb{R}^{d \times d}$ or $\mathbb{P}^{d^2-1}$. Thus, in this section we transform our coordinates as follows:

$$\sigma_{ii} = p_{ii}, \quad \sigma_{ij} = p_{ij} + q_{ij} \quad \text{and} \quad \sigma_{ji} = p_{ij} - q_{ij} \quad \text{for } 1 \leqslant i < j \leqslant d.$$

The signature matrix of a path X in $\mathbb{R}^d$ has real entries. We here restrict to certain families of paths that depend on a parameter m. The resulting signature matrices $\sigma^{(2)}(X)$ form the signature image. This is a semialgebraic subset in the real affine space $\mathbb{R}^{d\times d}$, as discussed in Section 2.2. We shall represent it as an orbit of the group $\mathrm{GL}(d,\mathbb{R})$ with respect to its natural action on the tensor product $\mathbb{R}^{d\times d} = \mathbb{R}^d \otimes \mathbb{R}^d$. The signature variety is the Zariski closure of this orbit in the projective space $\mathbb{P}^{d^2-1}$ of complex $d \times d$ matrices.

3.1. Real signature matrices. We fix positive integers $m \leqslant d$. Let $S^{[d,m]}_{\mathrm{axis}}$ denote the $d \times d$ matrix whose upper left $m \times m$ block is the upper-triangular matrix $\sigma^{(2)}(C_{\mathrm{axis}})$ in Example 2.1, with zeros in all other entries. Similarly, we write $S^{[d,m]}_{\mathrm{mono}}$ for the $d \times d$ matrix whose upper left $m \times m$ block is the matrix $\sigma^{(2)}(C_{\mathrm{mono}})$ in Example 2.2. See equation (23) in Example 3.5 for the case $d = 3$, $m = 2$. The decompositions of our matrices into symmetric and skew-symmetric parts are

$$S^{[d,m]}_{\mathrm{axis}} = P^{[d,m]}_{\mathrm{axis}} + Q^{[d,m]}_{\mathrm{axis}} \quad \text{and} \quad S^{[d,m]}_{\mathrm{mono}} = P^{[d,m]}_{\mathrm{mono}} + Q^{[d,m]}_{\mathrm{mono}}. \tag{12}$$

A piecewise linear path X with m segments in $\mathbb{R}^d$ is represented by any $d \times d$ matrix $\mathbf{X} = (x_{ij})$ whose first m columns are the steps. The signature matrix of this path equals

$$\sigma^{(2)}(X) = \mathbf{X} \cdot S^{[d,m]}_{\mathrm{axis}} \cdot \mathbf{X}^T. \tag{13}$$

The signature image for piecewise linear paths with m segments in $\mathbb{R}^d$ is the orbit

$$\mathcal{L}^{\mathrm{im}}_{d,2,m} := \left\{ \mathbf{X} \cdot S^{[d,m]}_{\mathrm{axis}} \cdot \mathbf{X}^T : \mathbf{X} \in \mathbb{R}^{d\times d} \right\} \subset \mathbb{R}^{d\times d}. \tag{14}$$

A polynomial path of degree m in $\mathbb{R}^d$ is represented by any $d \times d$ matrix $\mathbf{X} = (x_{ij})$ whose first m columns give the coefficients of the polynomials. As in (13), it has signature matrix

$$\sigma^{(2)}(X) = \mathbf{X} \cdot S^{[d,m]}_{\mathrm{mono}} \cdot \mathbf{X}^T. \tag{15}$$

The signature image for polynomial paths of degree m in $\mathbb{R}^d$ is the resulting orbit

$$\mathcal{P}^{\mathrm{im}}_{d,2,m} := \left\{ \mathbf{X} \cdot S^{[d,m]}_{\mathrm{mono}} \cdot \mathbf{X}^T : \mathbf{X} \in \mathbb{R}^{d\times d} \right\} \subset \mathbb{R}^{d\times d}. \tag{16}$$

EXAMPLE 3.1 ($m = 1$ and Veronese). The signature image for linear paths equals

$$\mathcal{L}^{\mathrm{im}}_{d,2,1} = \mathcal{P}^{\mathrm{im}}_{d,2,1} = \left\{ \tfrac{1}{2} X X^T : X \in \mathbb{R}^d \right\}.$$

These are the symmetric positive definite matrices of rank $\leqslant 1$. The Zariski closure of the set $[\mathcal{L}^{\mathrm{im}}_{d,2,1}]$ in $\mathbb{P}^{d^2-1}$ is the *Veronese variety*, which consists of all complex symmetric rank 1 matrices, regarded up to scaling. ◇

EXAMPLE 3.2 ($m = 2$). The signature image for two-step paths equals

$$\mathcal{L}^{\mathrm{im}}_{d,2,2} = \left\{ \tfrac{1}{2}(X_1 + X_2)(X_1 + X_2)^T + \tfrac{1}{2}(X_1 X_2^T - X_2 X_1^T) : X_1, X_2 \in \mathbb{R}^d \right\}. \quad (17)$$

For a matrix $S = P + Q \in \mathcal{L}^{\mathrm{im}}_{d,2,2}$, the symmetric part P lies in $\mathcal{L}^{\mathrm{im}}_{d,2,1}$ and the skew-symmetric part Q has rank $\leqslant 2$. The orthogonal complement of $\mathrm{span}(X_1, X_2)$ lies in the kernel of P and of Q. Hence, the $d \times 2d$ matrix $[P\ Q]$, built by concatenating P and Q, has rank $\leqslant 2$.

The signature image $\mathcal{L}^{\mathrm{im}}_{d,2,2}$ is not closed. To see this, consider any matrix S with $P = 0$ and Q skew-symmetric of rank 2. Then S satisfies the constraints above but it does not lie in $\mathcal{L}^{\mathrm{im}}_{d,2,2}$. Otherwise, $P = 0$ would imply $X_1 = -X_2$ and hence $Q = 0$. However, S lies in the closure of $\mathcal{L}^{\mathrm{im}}_{d,2,2}$. One can construct a sequence of nearly antipodal pairs $(X_1^\varepsilon, X_2^\varepsilon) \in (\mathbb{R}^d)^2$ such that $Q^\varepsilon = Q$ for all $\varepsilon > 0$ and $P^\varepsilon \to 0$ for $\varepsilon \to 0$. The same phenomenon arises for polynomial paths of the form (7) of degree $m = 2$. A quadratic path with $P = 0$ satisfies $Q = 0$. Such paths are *tree-like* [24]. A typical case is the path $t \mapsto \big(t(1-t), t(1-t)\big)$. ◇

The following theorem elucidates the relationship between piecewise linear paths and polynomial paths in $\mathbb{R}^d$. They have the same $d \times d$ signature matrices.

THEOREM 3.3. *If $k = 2$ then the signature image for piecewise linear paths equals the signature image for polynomial paths. In symbols, $\mathcal{L}^{\mathrm{im}}_{d,2,m} = \mathcal{P}^{\mathrm{im}}_{d,2,m} \subset \mathbb{R}^{d\times d}$.*

Proof. We abbreviate the matrices in (12) by $A := S^{[d,m]}_{\mathrm{axis}}$ and $M := S^{[d,m]}_{\mathrm{mono}}$. It is enough to treat the case $m = d$, since this proves the result for all $m \leqslant d$. We shall build an invertible matrix H such that $HMH^T = A$. This will prove that A and M have the same orbit under the action by $\mathrm{GL}(d, \mathbb{R})$ which clearly yields the desired set identity in $\mathbb{R}^{d\times d}$. The proof that follows proceeds by induction on d; to this end, introduce the 'refined' notation $M^{[d]} = S^{[d,d]}_{\mathrm{mono}}$. For $d = 1$ we have $A = M = (\frac{1}{2})$ and $H = (1)$. To go from dimension d to $d+1$ we write

$$M^{[d+1]} = \left(\begin{array}{c|c} M & \begin{matrix}\frac{d+1}{d+2}\\ \vdots\end{matrix} \\ \hline \begin{matrix}\frac{1}{d+1} & \cdots\end{matrix} & \frac{d+1}{2(d+1)} \end{array}\right) = \left(\begin{array}{c|c} M & \vdots \\ \hline \cdots & \frac{1}{2} \end{array}\right).$$

Here $M = M^{[d]}$. Our goal is to find a row vector $x \in \mathbb{R}^d$ and a scalar $y \in \mathbb{R}$ with

$$\left(\begin{array}{c|c} x^T & y \\ \hline H & 0^T \end{array}\right)\left(\begin{array}{c|c} M & \vdots \\ \hline \cdots & \frac{1}{2} \end{array}\right)\left(\begin{array}{c|c} x & H^T \\ \hline y & 0 \end{array}\right) = \left(\begin{array}{c|c} * & ** \\ \hline o & HMH^T \end{array}\right) = \left(\begin{array}{c|c} \frac{1}{2} & 1\cdots 1 \\ \hline 0 & \\ \vdots & A \\ 0 & \end{array}\right) = S^{[d+1,m]}_{\text{axis}}. \quad (18)$$

The identity in the middle requires that $(*) = \frac{1}{2}$, $(**) = (1 \cdots 1)$ and $o = 0^T = (0, \ldots, 0)^T$. We carry out the (block) matrix multiplication, starting with

$$\left(\begin{array}{c|c} M & \vdots \\ \hline \cdots & \frac{1}{2} \end{array}\right)\left(\begin{array}{c|c} x & H^T \\ \hline y & 0 \end{array}\right) = \left(\begin{array}{c|c} Mx + y\begin{pmatrix}\vdots\end{pmatrix} & MH^T \\ \hline (\cdots)x + \frac{1}{2}y & (\cdots)H^T \end{array}\right).$$

The requirement $o = 0^T$ in (18) implies $H\left(Mx + y\begin{pmatrix}\vdots\end{pmatrix}\right) = 0$ and hence $Mx + y\begin{pmatrix}\vdots\end{pmatrix} = 0$. With this, the condition $(*) = \frac{1}{2}$ translates into one quadratic equation in $d + 1$ unknowns (x, y):

$$y\left((\cdots)x + \frac{1}{2}y\right) = y\left(\sum_{i=1}^{d} \frac{i}{d+1+i}x_i + \frac{1}{2}y\right) = 1/2. \quad (19)$$

The final condition $(**) = (1, \ldots, 1)$ gives d linear equations for $(x, y) \in \mathbb{R}^{d+1}$:

$$x^T MH^T + y(\cdots)H^T = (1, \ldots, 1) \quad \text{in } \mathbb{R}^d. \quad (20)$$

Here $(\cdots) = (1/(d+2), \ldots, d/(2d+1))$. The solution set to (20) is a line in $\mathbb{R}^{d+1}$. For any given $y \in \mathbb{R}$, there is a unique solution $x = x(y)$. (It is an a fortiori consequence of (21) below that M is invertible, as is MH^T using the induction hypothesis.) Substituting this into (19), we obtain a quadratic equation in one variable y. Let y be one of its complex solutions. We claim that y is real.

Let $\tilde{H}$ be the desired $(d+1) \times (d+1)$ transformation matrix on the left of (18). By Laplace expansion with respect to the last column, its determinant equals $\det \tilde{H} = \pm y \det H$. Taking determinants on both sides of (18), and using $\det H = \det H^T \neq 0$, we conclude

$$y^2 \cdot \det(H)^2 \cdot \det\left(M^{[d+1]}\right) = \frac{1}{2}\det(A) = \frac{1}{2^{d+1}}.$$

At this point it suffices to know that $\det(M^{[d]}) > 0$ for all $d \in \mathbb{N}$. In fact, we claim

$$\det(M^{[d]}) = d! \cdot \frac{\prod_{1 \leqslant i < j \leqslant d}(j-i)^2}{\prod_{i=1}^{d}\prod_{j=1}^{d}(i+j)}. \quad (21)$$

This identity holds because $M^{[d]}$ becomes a *Cauchy matrix* after we divide its jth column by j. This positivity of $\det(M^{[d]})$ implies that y^2 is positive and hence y is real. From (20), we obtain a (real) vector $x = x(y)$. The resulting matrix $\tilde{H}$ completes the induction step. □

3.2. Determinantal varieties. The precise characterization of signature images is a subtle matter. We already saw this in Example 3.2. In this article, we address the easier problem of characterizing the polynomial equations that vanish on these images. In other words, we study the signature varieties and their ideals. The inequalities that hold on the signature images are left to future research.

Let $M_{d,m}$ denote the prime ideal of homogeneous polynomials that vanishes on $\mathcal{L}^{\mathrm{im}}_{d,2,m} = \mathcal{P}^{\mathrm{im}}_{d,2,m}$. This lives in the polynomial ring over $\mathbb{Q}$ with d^2 variables, namely the $\binom{d+1}{2}$ variables p_{ij} and the $\binom{d}{2}$ variables q_{ij}. The corresponding variety $\mathcal{M}_{d,m}$ lives in the projective space $\mathbb{P}^{d^2-1}$ of $d \times d$ matrices. This signature variety is the Zariski closure of the signature image. Our main result in this section characterizes this variety in terms of the $d \times 2d$ matrix $[P\ Q]$.

THEOREM 3.4. *For each d and m, the following varieties in $\mathbb{P}^{d^2-1}$ coincide:*

(1) *The variety of signature matrices of piecewise linear paths with m segments.*

(2) *The variety of signature matrices of polynomial paths of degree m.*

(3) *The variety of matrices $S = P+Q$, with P symmetric and Q skew-symmetric, such that* $\mathrm{rank}(P) \leqslant 1$ *and* $\mathrm{rank}\big([P\ Q]\big) \leqslant m$.

For each fixed value of d, we have the following chain of varieties in $\mathbb{P}^{d^2-1}$:

$$\mathcal{M}_{d,1} \subset \mathcal{M}_{d,2} \subset \mathcal{M}_{d,3} \subset \cdots \subset \mathcal{M}_{d,d} = \mathcal{M}_{d,d+1} = \mathcal{M}_{d,d+2} = \cdots \tag{22}$$

Fix $m \leqslant d$. Then $\mathcal{M}_{d,m}$ is irreducible of dimension $md - \binom{m}{2} - 1$. If m is odd then its ideal $M_{d,m}$ is generated by the 2-minors of P and the $(m+1)$-pfaffians of Q. If m is even then $M_{d,m}$ is generated by the 2-minors of P, the $(m+2)$-pfaffians of Q, and the entries in $P \cdot C_m(Q)$ where $C_m(Q)$ is the circuit matrix of m-pfaffians.

The matrix Q is skew-symmetric. Its principal minors of odd size are zero. Each principal minor of even size m is the square of a polynomial of degree $m/2$ in the q_{ij}. These polynomials are the *m-pfaffians* of Q. The *circuit matrix* $C_m(Q)$ has format $d \times \binom{d}{m+1}$, with columns indexed by the $(m+1)$-element subsets I of $\{1, 2, \ldots, d\}$. The entry in row i and column I is 0 unless $i \in I$. In that case it equals the m-pfaffian of Q that is indexed by $I \backslash \{i\}$. Here the pfaffian must be

taken with the correct sign. For instance, if $d = 6$ and $m = 4$ then

$$C_4(Q) = \begin{bmatrix} 0 & q_{34}q_{56} - q_{35}q_{46} + q_{36}q_{45} & q_{24}q_{56} - q_{25}q_{46} + q_{26}q_{45} & \cdots & q_{23}q_{45} - q_{24}q_{35} + q_{25}q_{34} \\ -q_{34}q_{56} + q_{35}q_{46} - q_{36}q_{45} & 0 & q_{14}q_{56} - q_{15}q_{46} + q_{16}q_{45} & \cdots & q_{13}q_{45} - q_{14}q_{35} + q_{15}q_{34} \\ -q_{24}q_{56} + q_{25}q_{46} - q_{26}q_{45} & -q_{14}q_{56} + q_{15}q_{46} - q_{16}q_{45} & 0 & \cdots & q_{12}q_{45} - q_{14}q_{25} + q_{15}q_{24} \\ -q_{23}q_{56} + q_{25}q_{36} - q_{26}q_{35} & -q_{13}q_{56} + q_{15}q_{36} - q_{16}q_{35} & -q_{12}q_{56} + q_{15}q_{26} - q_{16}q_{25} & \cdots & q_{12}q_{35} - q_{13}q_{25} + q_{15}q_{23} \\ -q_{23}q_{46} + q_{24}q_{36} - q_{26}q_{34} & -q_{13}q_{46} + q_{14}q_{36} - q_{16}q_{34} & -q_{12}q_{46} + q_{14}q_{26} - q_{16}q_{24} & \cdots & q_{12}q_{34} - q_{13}q_{24} + q_{14}q_{23} \\ -q_{23}q_{45} + q_{24}q_{35} - q_{25}q_{34} & -q_{13}q_{45} + q_{14}q_{35} - q_{15}q_{34} & -q_{12}q_{45} + q_{14}q_{25} - q_{15}q_{24} & \cdots & 0 \end{bmatrix}.$$

The entries of this matrix are the 2-pfaffians of the 6×6 matrix Q. The 4-pfaffians are homogeneous quartics with 15 terms. For any m and d, the columns of the circuit matrix $C_d(Q)$ are canonical generators for the kernel of the matrix Q, provided Q has rank m. Think of Cramer's rule for $\ker(Q)$. These generators correspond to *circuits* in matroid theory [**37**].

EXAMPLE 3.5 ($d = 3, m = 2$). The variety $\mathcal{M}_{3,2}$ has dimension 4 and degree 6 in $\mathbb{P}^8$. It is the Zariski closure of the common orbit of the following two matrices under the $\mathrm{GL}(3, \mathbb{R})$-action:

$$S^{[3,2]}_{\text{axis}} = \begin{pmatrix} \frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{or} \quad S^{[3,2]}_{\text{mono}} = \begin{pmatrix} \frac{1}{2} & \frac{2}{3} & 0 \\ \frac{1}{3} & \frac{2}{4} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{23}$$

The variety $\mathcal{M}_{3,2}$ is cut out by the 2×2 minors of $P = (p_{ij})$ and the 3×3 minors of

$$[P\ Q] = \begin{bmatrix} p_{11} & p_{12} & p_{13} & 0 & q_{12} & q_{13} \\ p_{12} & p_{22} & p_{23} & -q_{12} & 0 & q_{23} \\ p_{13} & p_{23} & p_{33} & -q_{13} & -q_{23} & 0 \end{bmatrix}.$$

However, these do not generate the prime ideal $M_{3,2}$. For that we need nine quadrics, namely the six 2-minors of P and the three entries of $P \cdot C_2(Q)$ where $C_2(Q) = \big(q_{23}, -q_{13}, q_{12}\big)^T$. The fourfold $\mathcal{M}_{3,2}$ contains the Veronese surface $\mathcal{M}_{3,1} \simeq \mathbb{P}^2$, which lives in a $\mathbb{P}^5$ inside $\mathbb{P}^8$. The ideal $M_{3,1}$ of that surface is generated by $\{q_{12}, q_{13}, q_{23}\}$ and the 2-minors of P. ◇

Proof of Theorem 3.4. We consider pencils spanned by a symmetric matrix P and a skew-symmetric matrix Q. These are called *T-even matrix pencils* in the linear algebra literature. Their normal forms are found, for example, in [**27**, **40**]. They correspond to normal forms of $S = P + Q$ under the $\mathrm{GL}(d, \mathbb{C})$-action by *congruence* $S \mapsto \mathbf{X}S\mathbf{X}^T$. By Theorem 3.3, our two canonical pencils in (12) are in the same orbit. It consists of rank m pencils (P, Q) where P has rank 1. The closure of this orbit also contains orbits where the rank drops from m to $\leqslant m - 1$.

If m is odd then Q has rank $m - 1$. There is no constraint imposed on the rank 1 matrix P. The normal form is the direct sum of one pencil $\big((1), (0)\big)$ and $d - m$

zero pencils and $(m-1)/2$ pencils $\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}, \begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$. If m is even then image$(P) \subseteq$ image(Q) holds on the orbit in question. The normal form is the direct sum of one pencil $\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}, \begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$ and $(m-2)/2$ pencils $\left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}, \begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$. The rank constraints in item 3 characterize the closure of this orbit. We conclude that the variety in item 3 is indeed irreducible and equals those in items 1 and 2.

To prove our formula for the dimension, we start with two binomial identities:

$$md - \binom{m}{2} = \left[\binom{d}{2} - \binom{d-m}{2}\right] + m = \left[\binom{d}{2} - \binom{d-(m-1)}{2}\right] + d.$$

We claim that this is the dimension of the affine cone over $\mathcal{M}_{d,m}$. The parenthesized expressions give the dimension of the variety of skew-symmetric $d \times d$ matrices Q of rank $\leqslant m$. When m is odd then the formula on the right is used. In that case we can set $P = vv^T$ where v is any column vector in $\mathbb{R}^d$. When m is even then the formula in the middle is used. Given Q, we now choose v in the column span of Q, so there are m degrees of freedom for P.

It remains to establish the assertions about the prime ideal $M_{d,m}$ associated with $\mathcal{M}_{d,m}$. The polynomials we list are in that ideal, and they cut out the variety $\mathcal{M}_{d,m}$ set-theoretically. We must show that they generate a prime ideal. If m is odd then this follows from the well-known fact that the 2-minors of P and the $(m+1)$-pfaffians of Q both generate prime ideals. Since they share no variables, the sum of these two ideals is also prime. If m is even then we must argue that incorporating the entries of $P \cdot C_m(Q)$, which are homogeneous of bidegree $(1, m/2)$, does not destroy primality. This can be shown by computing the reduced Gröbner basis with respect to the lexicographic term order, where the matrix entries are ordered row-wise with the p_{ij} coming before the q_{ij}. The leading terms in that Gröbner basis are square-free. The initial ideal is radical, and hence so is the ideal in question. Since it cuts out the correct irreducible variety, Hilbert's Nullstellensatz ensures that the ideal is prime. □

Table 1 lists invariants of the prime ideals $M_{d,m}$. These can be computed in `Macaulay2` [**21**]. The same format is used in later sections when we pass from matrices to tensors. The column 'a' denotes the *ambient dimension*. This refers to the smallest linear subspace of $\mathbb{P}^{d^2-1}$ containing $\mathcal{M}_{d,m}$. For instance, $M_{3,1}$ is an ideal in a polynomial ring with 9 unknowns and it contains 3 linear forms. Hence, the ambient dimension is $5 = 9 - 3 - 1$. Indeed, the quadratic Veronese surface $\mathcal{M}_{3,1} \simeq \mathbb{P}^2$ lives in $\mathbb{P}^5$. The column 'dim' displays $\dim(\mathcal{M}_{d,m}) = md - \binom{m}{2} - 1$. The column 'deg' displays the degree of $\mathcal{M}_{d,m}$. For m odd, this is 2^{d-1} times the degree of the variety of skew-symmetric matrices of rank $m-1$. The latter is a Catalan number when $m = 3$, which explains $\deg(\mathcal{M}_{d,3}) = (2^{d-1})/(d-1)\binom{2d-4}{d-2}$. Note also $\deg(\mathcal{M}_{d,2}) = \binom{2d-2}{d-1}$. The last column 'gens' lists $\mu_2, \mu_3, \ldots$ where

Table 1. Invariants of the ideal $M_{d,m}$ that defines the variety of signature matrices $\mathcal{M}_{d,m}$.

d	k	m	a	dim	deg	gens	d	k	m	a	dim	deg	gens
2	2	1	2	1	2	1	5	2	2	24	8	70	100
2	2	2	3	2	2	1	5	2	3	24	11	80	55
3	2	1	5	2	4	6	5	2	4	24	13	40	50, 5
3	2	2	8	4	6	9	5	2	5	24	14	16	50
3	2	3	8	5	4	6	6	2	1	20	5	32	105
4	2	1	9	3	8	20	6	2	2	35	10	252	225
4	2	2	15	6	20	36	6	2	3	35	14	448	120
4	2	3	15	8	16	21	6	2	4	35	17	280	105, 36
4	2	4	15	9	8	20	6	2	5	35	19	96	105, 1
5	2	1	14	4	16	50	6	2	6	35	20	32	105

μ_i is the number of minimal generators of $M_{d,m}$ in degree i. For instance, $M_{6,3}$ is generated by 120 quadrics, while $M_{6,4}$ is generated by 105 quadrics and 36 cubics. The former express $\operatorname{rank}(P) = 1$. The latter come from $P \cdot C_4(Q)$.

4. Universal varieties from free Lie algebras

The kth signature tensor of a path X in $\mathbb{R}^d$ is an element $\sigma^{(k)}(X)$ of the tensor space $(\mathbb{R}^d)^{\otimes k}$. We denote the coordinates on this space by $\sigma_{i_1 i_2 \cdots i_k}$ for $1 \leqslant i_1, i_2, \ldots, i_k \leqslant d$. We write $\mathbb{R}[\sigma^{(k)}]$ for the polynomial ring generated by these d^k coordinates. This is the homogeneous coordinate ring of the projective space $\mathbb{P}^{d^k-1}$, taken over either $\mathbb{R}$ or $\mathbb{C}$. The general linear group $G = \mathrm{GL}(d, \mathbb{R})$ acts naturally on the spaces $(\mathbb{R}^d)^{\otimes k}$ and $\mathbb{P}^{d^k-1}$, and on the ring $\mathbb{R}[\sigma^{(k)}]$.

Signature matrices $\sigma^{(2)}(X)$ were characterized in Section 3. In this section, we deal with signature tensors of order $k \geqslant 3$. We shall derive their signature varieties in $\mathbb{P}^{d^k-1}$. This rests on the theory of free Lie algebras [39].

4.1. Tensor algebra, Lie polynomials and shuffles. For any positive integer n, we consider the *truncated tensor algebra*

$$T^n(\mathbb{R}^d) = \bigoplus_{k=0}^{n} (\mathbb{R}^d)^{\otimes k}. \tag{24}$$

This is a noncommutative algebra whose multiplication is the tensor (or *concatenation*) product, where tensors of order $\geqslant n+1$ are set to zero. We write $T^n_v(\mathbb{R}^d)$ for the hyperplane in $T^n(\mathbb{R}^d)$ that consists of tensor polynomials with constant term v. The standard basis of $\mathbb{R}^d$ is denoted by $e_1, e_2, \ldots, e_d$.

The induced standard basis of $T^n(\mathbb{R}^d)$ is abbreviated

$$e_{i_1 i_2 \cdots i_k} := e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k} \quad \text{for } 1 \leqslant i_1, i_2, \ldots, i_k \leqslant d \text{ and } 0 \leqslant k \leqslant n.$$

We refer to basis vectors with $k \geqslant 1$ as *words* and we often write them simply as $i_1 i_2 \cdots i_k$. The empty word spans the one-dimensional space $(\mathbb{R}^d)^{\otimes 0} \cong \mathbb{R}$ of constant terms in $T^n(\mathbb{R}^d)$.

Following [23, Section 4] and [39], the truncated tensor algebra $T^n(\mathbb{R}^d)$ also carries several additional algebraic structures. First, it is a *Lie algebra*, where the Lie bracket is the commutator

$$[P, Q] = P \otimes Q - Q \otimes P \quad \text{for } P, Q \in T^n(\mathbb{R}^d).$$

Next, $T^n(\mathbb{R}^d)$ is a commutative algebra with respect to the *shuffle product* ш, where words of length $\geqslant n+1$ are set to zero. The shuffle product of two words of lengths r and s (with $r + s \leqslant n$) is the sum over the $\binom{r+s}{s}$ ways of interleaving the two words. For a more formal definition see Reutenauer's book [39, Section 1.4]. Here are some examples for the shuffle product:

$$\begin{gathered} e_{12} \sqcup\!\sqcup e_{34} = e_{12} \sqcup\!\sqcup\!\sqcup e_{34} = e_{1234} + e_{1324} + e_{1342} + e_{3124} + e_{3142} + e_{3412}, \\ e_{3} \sqcup\!\sqcup e_{134} = e_{3} \sqcup\!\sqcup\!\sqcup e_{134} = e_{3134} + 2e_{1334} + e_{1343}, \quad e_{21} \sqcup\!\sqcup e_{21} = 2e_{2121} + 4e_{2211}, \\ e_{1} \sqcup\!\sqcup e_{111} = 4e_{1111}, \quad e_{11} \sqcup\!\sqcup e_{11} = 6e_{1111}, \quad e_{12} \sqcup\!\sqcup e_{21} = 2e_{1221} + e_{1212} + e_{2121} + 2e_{2112}. \end{gathered}$$

The tensor algebra also carries two *coproducts*: the *deconcatenation* coproduct $\Delta_{\otimes}$ and the *deshuffle coproduct* $\Delta_{ш}$, given by dualizing the corresponding products. One so obtains two bialgebra (in fact, Hopf algebra) structures that are in natural duality, see for example, [39, Proposition 1.9]. The truncated tensor algebra naturally inherits most of these structures.

The ring of polynomial functions on $T^n(\mathbb{R}^d)$ is denoted by $\mathbb{R}[\sigma^{(\leqslant n)}]$. The coordinate functions are σ_I, where I runs over all words of length $\leqslant n$. For two words I and J of length $\geqslant 1$, let $\sigma_{I \sqcup\!\sqcup J}$ be the linear form in $\mathbb{R}[\sigma^{(\leqslant n)}]$ that corresponds to the shuffle product $e_{I \sqcup\!\sqcup J}$. We refer to $\sigma_{I \sqcup\!\sqcup J}$ as a *shuffle linear form*. For instance, here are some examples of shuffle linear forms:

$$\begin{gathered} \sigma_{12 \sqcup\!\sqcup 34} = \sigma_{1234} + \sigma_{1324} + \sigma_{1342} + \sigma_{3124} + \sigma_{3142} + \sigma_{3412}, \\ \sigma_{3 \sqcup\!\sqcup 134} = \sigma_{3134} + 2\sigma_{1334} + \sigma_{1343}, \quad \sigma_{21 \sqcup\!\sqcup 21} = 2\sigma_{2121} + 4\sigma_{2211}, \\ \sigma_{1 \sqcup\!\sqcup 111} = 4\sigma_{1111}, \quad \sigma_{11 \sqcup\!\sqcup 11} = 6\sigma_{1111}, \\ \sigma_{12 \sqcup\!\sqcup 21} = 2\sigma_{1221} + \sigma_{1212} + \sigma_{2121} + 2\sigma_{2112}. \end{gathered} \tag{25}$$

An element of $T^n(\mathbb{R}^d)$ is a *Lie polynomial* if it can be obtained from the standard basis vectors $e_1, e_2, \ldots, e_d$ by iterating the operations of taking Lie brackets and linear combinations. The resulting set of all Lie polynomials of

degree $\leqslant n$ is a vector space, denoted by $\mathrm{Lie}^n(\mathbb{R}^d)$. Equivalently, $\mathrm{Lie}^n(\mathbb{R}^d)$ is the smallest Lie subalgebra of $T^n(\mathbb{R}^d)$ containing $\mathbb{R}^d$. By construction, all Lie polynomials are elements of $T_0^n(\mathbb{R}^d) = \{0\} \oplus \mathbb{R}^d \oplus \cdots \oplus (\mathbb{R}^d)^{\otimes n}$.

LEMMA 4.1. *Lie polynomials are characterized by the vanishing of all shuffle linear forms:*

$$\mathrm{Lie}^n(\mathbb{R}^d) = \left\{ P \in T_0^n(\mathbb{R}^d) : \sigma_{I \sqcup\!\sqcup J}(P) = 0 \quad \textit{for all (nonempty) words } I, J \right\}.$$

Proof. This is the equivalence between (i) and (iv) in [39, Theorem 3.1]. □

We next recall (for example, from [39, Section 1.1]) that the familiar series for the exponential function and the logarithm function determine polynomial maps from $T^n(\mathbb{R}^d)$ to itself:

$$\exp(P) = \sum_{r \geqslant 0} \frac{1}{r!} P^{\otimes r} \quad \text{and} \quad \log(1+P) = \sum_{r \geqslant 1} \frac{(-1)^{r-1}}{r} P^{\otimes r}.$$

The image of $T_0^n(\mathbb{R}^d)$ under the exponential function is the set $T_1^n(\mathbb{R}^d)$ of tensor polynomials with constant term 1. The logarithm function inverts the exponential function on its image:

$$\log(\exp(P)) = P \quad \text{for all } P \in T_0^n(\mathbb{R}^d).$$

The *step-n free nilpotent Lie group* is the image of the free Lie algebra under the exponential map:

$$\mathcal{G}^n(\mathbb{R}^d) := \exp(\mathrm{Lie}^n(\mathbb{R}^d)) \subset T_1^n(\mathbb{R}^d). \tag{26}$$

It is known (for example, [20, Theorem 7.30]) that $\mathcal{G}^n(\mathbb{R}^d)$ is a Lie group, with polynomial group law. It plays a central role as state space of geometric rough paths (cf. Section 5.4). Elements in $\mathcal{G}^n(\mathbb{R}^d)$ are also known as *group-like elements* [23, 39].

LEMMA 4.2. *Group-like elements are characterized by multiplicativity:*

$$\mathcal{G}^n(\mathbb{R}^d) = \left\{ P \in T_1^n(\mathbb{R}^d) : \sigma_{I \sqcup\!\sqcup J}(P) = \sigma_I(P)\sigma_J(P) \textit{ for all } I, J \textit{ with } |I|+|J| \leqslant n \right\}.$$

Proof. This is a reformulation of [39, Theorem 3.2(ii)]. See also [23, (4.2)]. □

REMARK 4.3 (From $\mathbb{R}$ to $\mathbb{K}$). The definitions above carry over verbatim from the real numbers $\mathbb{R}$ to an arbitrary field $\mathbb{K}$ of characteristic zero. Lemmas 4.1 and 4.2

remain unchanged. They furnish equational characterizations of the Lie algebra $\mathrm{Lie}^n(\mathbb{K}^d)$ and the Lie group $\mathcal{G}^n(\mathbb{K}^d)$. The former is a linear space and the latter is an affine algebraic variety.

We now come to the connection with paths. It is established by the following fundamental result due to Chen [8, 9]. This can also be viewed as a consequence of Chow's work in [12].

THEOREM 4.4 (Chen–Chow). *The step-n free nilpotent Lie group $\mathcal{G}^n(\mathbb{R}^d)$ is precisely the image of the step n signature map in (3) when that map is applied to all smooth paths in $\mathbb{R}^d$:*

$$\mathcal{G}^n(\mathbb{R}^d) = \bigl\{\sigma^{\leqslant n}(X) : X : [0,1] \to \mathbb{R}^d \textit{any smooth path}\bigr\}. \tag{27}$$

We revisit the $n = 2$ case in Section 3 from the perspective of this theorem.

EXAMPLE 4.5 ($n = 2$). The truncated tensor algebra $T^2(\mathbb{R}^d)$ consists of elements

$$P = \sum_{i=1}^d \sum_{j=1}^d \alpha_{ij} e_{ij} + \sum_{i=1}^d \beta_i e_i + \gamma.$$

By Lemma 4.1, P is in $\mathrm{Lie}^2(\mathbb{R}^d)$ if and only if $\gamma = 0$ and $\alpha_{i \sqcup\!\sqcup j} = \alpha_{ij} + \alpha_{ji} = 0$ for all i, j. Lemma 4.2 says that the exponentials of these Lie polynomials are precisely the expressions

$$\exp(P) = \sum_{i=1}^d \sum_{j=1}^d \sigma_{ij} e_{ij} + \sum_{i=1}^d \sigma_i e_i + 1.$$

where $\sigma_i \sigma_j$ equals the shuffle linear form $\sigma_{i \sqcup\!\sqcup j} = \sigma_{ij} + \sigma_{ji}$ for $1 \leqslant i, j \leqslant d$. These inhomogeneous quadratic equations cut out the step-2 free nilpotent Lie group $\mathcal{G}^2(\mathbb{R}^d)$. By eliminating $\sigma_1, \ldots, \sigma_d$ from these equations, we obtain the homogeneous ideal $M_{d,d}$ in Theorem 3.4. ◇

4.2. Gröbner basis for the free Lie group. In what follows we work over an algebraically closed field $\mathbb{K}$ of characteristic zero. Our varieties are defined over $\mathbb{Q}$. Computations refer to polynomials with rational coefficients. We consider the following ideal in the ring $\mathbb{K}[\sigma^{(\leqslant n)}]$ of polynomial functions on $T^n(\mathbb{K}^d)$:

$$G_{d,n} = \bigl\langle \sigma_I \sigma_J - \sigma_{I \sqcup\!\sqcup J} : \text{for all words } I \text{ and } J \text{ with } |I| + |J| \leqslant n \bigr\rangle. \tag{28}$$

This ideal is not homogeneous. Its affine variety $\mathcal{G}_{d,n}$ is precisely the free Lie group $\mathcal{G}^n(\mathbb{K}^d)$. The exponential map is a polynomial parametrization of this variety. The following example illustrates the associated implicitization problem, whose solution is given by Lemma 4.2.

EXAMPLE 4.6 ($d = 2, n = 3$). The linear space $\mathrm{Lie}^3(\mathbb{K}^2)$ has dimension 5. Its elements are

$$\sigma = re_1 + se_2 + t[e_1, e_2] + u[e_1, [e_1, e_2]] + v[[e_1, e_2], e_2], \quad \text{where } r, s, t, v, u \in \mathbb{K}.$$

$\mathrm{Lie}^3(\mathbb{K}^2)$ is the subspace of $T_0^3(\mathbb{K}^2) \simeq \mathbb{K}^{14}$ defined by nine shuffle relations like

$$\sigma_{1\sqcup 2} = \sigma_{12} + \sigma_{21} = 0, \quad \sigma_{1\sqcup 11} = 3\sigma_{111} = 0, \quad \sigma_{1\sqcup 12} = 2\sigma_{112} + \sigma_{121} = 0.$$

The exponential $\exp(\sigma)$ of the Lie polynomial σ is the following expression:

$$1 + re_1 + se_2 + \frac{r^2}{2}e_{11} + \left(\frac{rs}{2} + t\right)e_{12} + \left(\frac{rs}{2} - t\right)e_{21} + \cdots$$
$$+\left(\frac{rs^2}{6} - 2v\right)e_{212} + \left(\frac{rs^2}{6} - \frac{st}{2} + v\right)e_{221} + \frac{s^3}{6}e_{222}.$$

The coefficients of $\exp(\sigma)$ define the exponential map from $\mathrm{Lie}^3(\mathbb{K}^2) \simeq \mathbb{K}^5$ into $T_1^3(\mathbb{K}^2) \simeq \mathbb{K}^{14}$. Its image is the 5-dimensional variety $\mathcal{G}_{2,3}$. We compute its ideal $G_{2,3}$ using the computer algebra package `Macaulay2` as follows:

```
R = QQ[s11,s21,s22,s111,s121,s211,s212,s221,s222,s1,s2,s12,s112,s122,
        MonomialOrder=>Lex];
ExponentialMap = map( QQ[r,s,t,u,v] , R ,
{r^2/2,r*s/2-t,s^2/2,r^3/6,r^2*s/6-2*u,r^2*s/6-r*t/2+u,r*s^2/6-2*v,
r*s^2/6-s*t/2+v, s^3/6, r,s,r*s/2+t, r^2*s/6+r*t/2+u, r*s^2/6+s*t/2+v});
G23 = kernel ExponentialMap;
gens gb G23
G23 == ideal( s1^2-2*s11, s1*s2-s12-s21, s2^2-2*s22, s1*s11-3*s111,
s1*s12-2*s112-s121, s1*s21-s121-2*s211, s1*s22-s122-s212-s221,
s2*s11-s121-s211-s112,  s2*s12-2*s122-s212, s2*s21-2*s221-s212,
s2*s22-3*s222)
```

This computes the lexicographic Gröbner basis with nine elements to be seen in Theorem 4.10. The last command uses == to verify that $G_{2,3}$ is generated by the quadratic relations $\sigma_I\sigma_J - \sigma_{I\sqcup J}$. ◇

We are interested in the structure of the free Lie group $\mathcal{G}_{d,n}$ as an affine algebraic variety. To this end, we need to first record some combinatorial facts about free Lie algebras.

A word I on the alphabet $\{1, 2, \ldots, d\}$ is a *Lyndon word* if it is strictly smaller in lexicographic order than all of its rotations. Lyndon words are the Hall words for a particular Hall set (cf. [**39**, Ch. 4]). Since they are easy to define, and seen widely in the combinatorics literature, we use Lyndon words in what follows. With any Lyndon word I one associates an iterated Lie bracketing $b(I)$ by induction on $k = \mathrm{length}(I)$. If $k = 1$ and $I = i$ then $b(I) = e_i$. If $k \geqslant 2$ then it is

$b(I) = [b(I_1), b(I_2)]$ where $I = I_1 I_2$ and I_2 is the longest Lyndon word appearing as a proper right factor of I. Some authors refer to the bracketing image $b(I)$ of a Lyndon word I as *Lyndon bracket*. We point to [26]. That article also features connections to Gröbner bases. We use the following about Lyndon words.

PROPOSITION 4.7. *The bracketings $b(I)$ of Lyndon words I of length $\leqslant n$ form a basis for* $\mathrm{Lie}^n(\mathbb{K}^d)$. *The dimension of this Lie algebra, which is the number of Lyndon words, equals*

$$\lambda_{d,n} = \sum_{k=1}^{n} \sum_{\ell \text{ divides } k} \frac{\mu(\ell)}{k} d^{k/\ell}, \quad \textit{where } \mu \textit{ is the Möbius function.}$$

Proof. The inner sum is the number of Lyndon words of length k. See [39, Corollary 4.14]. □

EXAMPLE 4.8. There are $\lambda_{2,2} = 3$ Lyndon words of length $n \leqslant 2$ and $d = 2$. Their bracketings are $e_1 = b(1)$, $e_2 = b(2)$, $e_{12} - e_{21} = b(12)$. These three elements form a basis of $\mathrm{Lie}^2(\mathbb{K}^2)$. ◇

EXAMPLE 4.9 ($d = 2, k = 4$). The three Lyndon words have the bracketings

$$\begin{aligned} b(1112) &= [1,[1,[1,2]]] = e_{1112} - 3e_{1121} + 3e_{1211} - e_{2111}, \\ b(1122) &= [1,[[1,2],2]] = e_{1122} - 2e_{1212} + 2e_{2121} - e_{2211}, \\ b(1222) &= [[[1,2],2],2] = e_{1222} - 3e_{2122} + 3e_{2212} - e_{2221}. \end{aligned} \tag{29}$$

The shuffle forms $\sigma_{1 \sqcup\!\sqcup 111}$, $\sigma_{11 \sqcup\!\sqcup 11}$ and $\sigma_{12 \sqcup\!\sqcup 21}$ in (25) vanish on the span of (29), as seen in Lemma 4.1. Hence the subspace of shuffle forms in $\mathbb{K}[\sigma^{(4)}]$ has dimension $13 = 16 - 3$. ◇

We now present our main result in this subsection. It shows that the ideal $G_{d,n}$ is prime. We fix a *lexicographic term order* in the polynomial ring $\mathbb{K}[\sigma^{(\leqslant n)}]$. The underlying variable ordering is assumed to satisfy the following requirement: we have $\sigma_I > \sigma_J$ if $\mathrm{length}(I) > \mathrm{length}(J)$, or if J is obtained from I by rotation and J is lexicographically smaller than I. Consider any word I that is not a Lyndon word. In the following theorem, the polynomial ϕ_I is the unique expression of σ_I on $\mathcal{G}_{d,n}$ in terms of the unknowns σ_J that are indexed by the Lyndon words J.

THEOREM 4.10. *The ideal $G_{d,n}$ in (28) is prime. Its irreducible variety $\mathcal{G}_{d,n}$ has dimension $\lambda_{d,n}$ in $T^n(\mathbb{K}^d)$. The reduced Gröbner basis of $G_{d,n}$ consists of the polynomials $\sigma_I - \phi_I(\sigma_{\mathrm{lyndon}})$ where I runs over non-Lyndon words of length $\leqslant n$.*

Example 4.6 illustrates Theorem 4.10 for $d = 2, n = 3$. The vector of Lyndon variables equals $\sigma_{\text{lyndon}} = (\sigma_1, \sigma_2, \sigma_{12}, \sigma_{112}, \sigma_{122})$. Its length is $\lambda_{2,3} = 5$. The last line of the `Macaulay2` code is the right-hand side of (28). The output of the command `gens gb G23` is the reduced Gröbner basis in Theorem 4.10.

Proof. By Lemma 4.2, the Lie group $\mathcal{G}_{d,n}$ is the zero set of the polynomials $\sigma_I\sigma_J - \sigma_{I \sqcup\!\sqcup J}$. This affine variety is irreducible because it is the image of the linear space $\text{Lie}^n(\mathbb{K}^d)$ under the polynomial map exp. This map has a polynomial inverse, namely log. Hence the dimension of $\mathcal{G}_{d,n}$ agrees with that of $\text{Lie}^n(\mathbb{K}^d)$. The latter dimension is $\lambda_{d,n}$, by Proposition 4.7.

Let σ_{lyndon} denote the vector of all variables σ_J that are indexed by Lyndon words J. We claim that, for every non-Lyndon word I, there is a polynomial $\phi_I = \phi_I(\sigma_{\text{lyndon}})$ in the Lyndon variables σ_J such that the difference $\sigma_I - \phi_I(\sigma_{\text{lyndon}})$ lies in the ideal $G_{d,n}$. This follows from a theorem of Radford [**38**]. We shall prove it by induction on lexicographic order.

We construct the polynomial ϕ_I with the technique used by Melançon and Reutenauer in [**36**, Section 4]. The non-Lyndon word I has a unique factorization into Lyndon words $I_1 I_2 \cdots I_\kappa$. Recall that shuffle multiplication is associative and commutative. This ensures that the expression $\sigma_{I_1 \sqcup\!\sqcup I_2 \sqcup\!\sqcup \cdots \sqcup\!\sqcup I_\kappa} - \sigma_{I_1}\sigma_{I_2}\cdots\sigma_{I_\kappa}$ lies in the ideal $G_{d,n}$. Its highest term equals σ_I times a positive integer. We divide the expression by that integer and then subtract it from σ_I. Each of the non-Lyndon variables $\sigma_{I'}$ seen in the resulting polynomial has either $|I'| < |I|$ or $|I'| = |I|$ and I' comes before I. Iterating this process, we are done by induction.

The prime ideal of $\mathcal{G}_{d,n}$ contains $G_{d,n}$ which in turn contains the ideal generated by the expressions $\sigma_I - \phi_I(\sigma_{\text{lyndon}})$ for I non-Lyndon. The first ideal is prime by definition. The third ideal is prime because its initial monomial ideal is prime. This holds as it is generated by the variables σ_I for I non-Lyndon. Moreover, the two prime ideals have the same dimension, namely $\lambda_{d,n}$. This implies that all three ideals in our chain of inclusions are equal.

The argument also proves Radford's result that the Lyndon variables σ_J are algebraically independent modulo $G_{d,n}$, so the polynomials ϕ_I seen in the Gröbner basis are unique. □

4.3. The universal variety. In this paper we study signature tensors of a fixed order k. The elements in the Lie group $\mathcal{G}_{d,n}$ record these tensors simultaneously for all values of k between 1 and n. We must extract the homogeneous component of degree k, by projecting to the kth summand in (24). The algebraic counterpart to the geometric operation of projection is the elimination of variables.

The *universal ideal* for order k tensors of format $d \times d \times \cdots \times d$ is defined as

$$U_{d,k} := G_{d,k} \cap \mathbb{K}[\sigma^{(k)}]. \tag{30}$$

This is a homogeneous prime ideal in a polynomial ring in d^k variables. We define the *universal variety* $\mathcal{U}_{d,k}$ to be the zero set of the ideal $U_{d,k}$ in the projective space $\mathbb{P}^{d^k-1}$. As in earlier sections, the field $\mathbb{K}$ is algebraically closed of characteristic zero. In applications to paths in $\mathbb{R}^d$, we would take $\mathbb{K} = \mathbb{C}$, but we could also consider paths in $\mathbb{K}^d$.

COROLLARY 4.11. *If $\mathbb{R} \subset \mathbb{K}$ then the universal variety $\mathcal{U}_{d,k}$ is the smallest projective variety in $\mathbb{P}^{d^k-1}$ that contains the kth signature tensors $\sigma^{(k)}(X)$ of all smooth paths X in $\mathbb{R}^d$.*

Proof. This follows immediately from the Chen–Chow Theorem 4.4. □

The universal variety $\mathcal{U}_{d,k}$ is the ambient space for the varieties of signature tensors in Section 5 below. We saw this already in the matrix case in Section 3.

EXAMPLE 4.12 ($k = 2$). The universal variety $\mathcal{U}_{d,2}$ is the variety $\mathcal{M}_{d,d}$ in (22). Its points are the $d \times d$ matrices whose symmetrization has rank 1. It is universal in the sense that, for all $m \in \mathbb{N}$, it contains the signature images $\mathcal{L}^{\text{im}}_{d,2,m}$ and $\mathcal{P}^{\text{im}}_{d,2,m}$ and their varieties $\mathcal{M}_{d,m}$. ◇

The varieties $\mathcal{U}_{d,k}$ and its subvarieties in the next section are invariant under the action of $G = \text{GL}(d, \mathbb{K})$ on tensors. In describing their ideals, it is sometimes convenient to perform a change of basis that realizes the decomposition of $(\mathbb{K}^d)^{\otimes k}$ into irreducible G-modules.

EXAMPLE 4.13 ($d = 2, k = 3$). The universal variety $\mathcal{U}_{2,3}$ lives in the space $\mathbb{P}^7$ of $2 \times 2 \times 2$ tensors. We compute its ideal $U_{2,3}$ with the following two commands after Example 4.6. The output consists of six quadrics:

```
U23 = eliminate({s1,s2,s11,s12,s21,s22},G23)
codim U23, degree U23, betti mingens U23
```

To understand the output, we perform the change of coordinates

$$\begin{matrix} \sigma_{111} = 6\alpha_1 & \sigma_{112} = 2\alpha_2 - \beta_1 & \sigma_{121} = 2\alpha_2 - \gamma_1 & \sigma_{211} = 2\alpha_2 + \beta_1 + \gamma_1 \\ \sigma_{222} = 6\alpha_4 & \sigma_{221} = 2\alpha_3 - \beta_2 & \sigma_{212} = 2\alpha_3 - \gamma_2 & \sigma_{122} = 2\alpha_3 + \beta_2 + \gamma_2. \end{matrix} \tag{31}$$

Here $\alpha_i, \beta_j, \gamma_k$ are coordinates on three irreducible G-modules $S_3(\mathbb{K}^2)$, $S_{21}(\mathbb{K}^2)$, $S_{21}(\mathbb{K}^2)$ in $(\mathbb{K}^2)^{\otimes 3}$. These have dimensions $4, 2, 2$, by the Hook Length Formula [**28**, Section 2.8]. The new coordinates reveal that the ideal $U_{2,3}$ is generated by the 2×2 minors of the 2×4 matrix

$$\begin{pmatrix} 3\alpha_1 & \alpha_2 & \alpha_3 & 2\beta_1 + \gamma_1 \\ \alpha_2 & \alpha_3 & 3\alpha_4 & -2\beta_2 - \gamma_2 \end{pmatrix}. \tag{32}$$

The universal variety $\mathcal{U}_{2,3}$ has dimension 4 and degree 4 in $\mathbb{P}^7$. ◇

EXAMPLE 4.14 ($d = k = 3$). The universal variety $\mathcal{U}_{3,3}$ of $3 \times 3 \times 3$ signature tensors has dimension 13 and degree 24 in $\mathbb{P}^{26}$. Using `Macaulay2`, we find that its homogeneous prime ideal $U_{3,3}$ is minimally generated by 81 quadrics. To exhibit these quadrics, we first decompose the ambient tensor space into four irreducible G-modules:

$$(\mathbb{K}^3)^{\otimes 3} \simeq \mathrm{S}_3(\mathbb{K}^3) \oplus \mathrm{S}_{21}(\mathbb{K}^3) \oplus \mathrm{S}_{21}(\mathbb{K}^3) \oplus \mathrm{S}_{111}(\mathbb{K}^3) \quad \text{or} \quad \mathbb{K}^{27}_s \simeq \mathbb{K}^{10}_a \oplus \mathbb{K}^8_b \oplus \mathbb{K}^8_c \oplus \mathbb{K}^1_d.$$

The first and last component represent symmetric and skew-symmetric tensors respectively. The following linear change of coordinates makes the isomorphisms above explicit:

$$\begin{gathered}
a_{111} = s_{111}, \quad a_{112} = s_{112} + s_{121} + s_{211}, \quad a_{113} = s_{113} + s_{131} + s_{311}, \quad a_{122} = s_{122} + s_{212} + s_{221}, \\
a_{123} = s_{123} + s_{132} + s_{213} + s_{231} + s_{312} + s_{321}, \quad a_{133} = s_{133} + s_{313} + s_{331}, \quad a_{222} = s_{222}, \\
a_{223} = s_{223} + s_{232} + s_{322}, \quad a_{233} = s_{233} + s_{323} + s_{332}, \quad a_{333} = s_{333}, \quad b_{121} = -4s_{112} + 2s_{121} + 2s_{211}, \\
b_{122} = 2s_{122} + 2s_{212} - 4s_{221}, \quad b_{123} = 2s_{123} + 2s_{213} - 2s_{231} - 2s_{321}, \quad b_{131} = -4s_{113} + 2s_{131} + 2s_{311}, \\
b_{132} = 2s_{132} - 2s_{231} + 2s_{312} - 2s_{321}, \quad b_{133} = 2s_{133} + 2s_{313} - 4s_{331}, \quad b_{232} = -4s_{223} + 2s_{232} + 2s_{322}, \\
b_{233} = 2s_{233} + 2s_{323} - 4s_{332}, \quad c_{112} = 2s_{112} - 4s_{121} + 2s_{211}, \quad c_{113} = 2s_{113} - 4s_{131} + 2s_{311}, \\
c_{122} = 2s_{122} - 4s_{212} + 2s_{221}, \quad c_{123} = 2s_{123} - 2s_{213} - 2s_{312} + 2s_{321}, \quad c_{132} = 2s_{132} - 2s_{213} + 2s_{231} - 2s_{312}, \\
c_{133} = 2s_{133} - 4s_{313} + 2s_{331}, \quad c_{223} = 2s_{223} - 4s_{232} + 2s_{322}, \quad c_{233} = 2s_{233} - 4s_{323} + 2s_{332}, \\
d_{123} = s_{123} - s_{132} - s_{213} + s_{231} + s_{312} - s_{321}.
\end{gathered}$$

The ideal $U_{3,3}$ of the universal variety is generated by a space of $6 \cdot 2 + 3 \cdot 2 + 3 \cdot 3 + 6 \cdot 7 + 12 = 81$ quadrics. Its $\mathbb{Z}^3$-grading has 19 components that come in five symmetry classes:

six	(4, 2, 0)	$3a_{111}a_{122} - a_{112}^2$,	$6a_{111}b_{122} + 3a_{111}c_{122} + 2a_{112}b_{121} + a_{112}c_{112}$,
three	(3, 3, 0)	$9a_{111}a_{222} - a_{112}a_{122}$,	$2a_{112}b_{122} + a_{112}c_{122} + 2a_{122}b_{121} + a_{122}c_{112}$,
three	(4, 1, 1)	$3a_{111}a_{123} - 2a_{112}a_{113}$,	$6a_{111}b_{123} + 3a_{111}c_{132} - 3a_{111}d_{123} + 2a_{112}b_{131} + a_{112}c_{113}$,
			$6a_{111}b_{132} + 3a_{111}c_{123} + 3a_{111}d_{123} + 2a_{113}b_{121} + a_{113}c_{112}$.

six (3, 2, 1) $6a_{111}a_{223} - a_{112}a_{123}, 2a_{112}b_{132} + a_{112}c_{123} + a_{112}d_{123} - 2a_{113}b_{122} - a_{113}c_{122}, \ldots$ etc.

There are seven quadrics in degree (3, 2, 1). In addition, we have twelve quadrics in the central degree (2, 2, 2), like $4a_{112}b_{233} + 2a_{112}c_{233} - 4a_{122}b_{133} - 2a_{122}c_{133} - 3a_{123}d_{123} - 4a_{233}b_{121} - 2a_{233}c_{112}$. ◇

We computed the ideals $U_{d,k}$ for all values of d and k with $d + k \leqslant 7$. Since $\mathcal{U}_{d,2}$ equals the matrix signature variety $\mathcal{M}_{d,d}$ (cf. Table 1), we only consider $k \geqslant 3$. The results are listed in Table 2. In each case we computed, we found that $U_{d,2}$ is generated by quadrics. The last column gives the number of generators. The second-to-last column reports the degree of the variety $\mathcal{U}_{d,k} \subset \mathbb{P}^{d^k-1}$.

In the first version of this paper we asked whether $U_{d,k}$ is always generated by quadrics. This question has since been answered, to the negative, by Francesco Galuppi. In Section 6 we shall prove that the universal variety $\mathcal{U}_{d,k}$ has the expected dimension $\lambda_{d,k} - 1$.

Table 2. Invariants of the ideal $U_{d,k}$ that defines the universal variety $\mathcal{U}_{d,k}$.

d	k	a	dim	deg	gens
2	3	7	4	4	6
2	4	15	7	12	33
2	5	31	13	40	150
3	3	26	13	24	81
3	4	80	31	672	954
4	3	63	29	200	486

5. Piecewise linear paths and polynomial paths

This section is the heart of this paper. We introduce, study and relate the signature varieties of two natural families of paths. These all live in the universal varieties seen in Section 4.

5.1. Polynomial maps into tensor space. We now study paths $X : [0, 1] \to \mathbb{R}^d$ whose coordinates are polynomials of degree m or piecewise linear with m pieces. Each of these is represented by a real $d \times m$ matrix, also denoted by $X = (x_{ij})$. With this convention, a polynomial path has coordinate functions

$$X_i(t) = x_{i1}t + x_{i2}t^2 + x_{i3}t^3 + \cdots + x_{im}t^m. \tag{33}$$

The differential 1-forms seen in the iterated integrals (1) are

$$dX_i(t) = X_i'(t)\,dt = \bigl(x_{i1} + 2x_{i2}t + 3x_{i3}t^2 + \cdots + mx_{im}t^{m-1}\bigr)\,dt. \tag{34}$$

Each coordinate $\sigma_{i_1 i_2 \cdots i_k}$ of the tensor $\sigma^{(k)}(X)$ is a homogeneous polynomial of degree k in the dm unknowns x_{ij} with coefficients in $\mathbb{Q}$. Formulas for $d = 2$, $k = 3$ are shown in Example 2.4. The $d \times d \times \cdots \times d$ tensor $\sigma^{(k)}(X)$ can be computed from the $m \times m \times \cdots \times m$ tensor $\sigma^k(C_{\text{mono}})$ in Example 2.2 by multiplying each of its k sides with the $d \times m$ matrix X. This is the tensor analogue to the congruence action on matrix space seen in (15).

The x_{ij} are homogeneous coordinates on the projective space $\mathbb{P}^{dm-1}$ over an algebraically closed field $\mathbb{K}$ that contains $\mathbb{R}$. The matrix-tensor multiplication described above defines a rational map of degree k:

$$\sigma^{(k)} : \mathbb{P}^{dm-1} \dashrightarrow \mathbb{P}^{d^k-1}, \quad X \mapsto \sigma^{(k)}(X). \tag{35}$$

The Zariski closure of the image of this map is the *polynomial signature variety* $\mathcal{P}_{d,k,m}$. The homogeneous prime ideal $P_{d,k,m}$ of this variety in $\mathbb{K}[\sigma^{(k)}]$ is the *polynomial signature ideal*.

EXAMPLE 5.1 ($d = k = 3, m = 2$). The third signature variety $\mathcal{P}_{3,3,2}$ for quadratic paths in 3-space lies in the space of $3 \times 3 \times 3$ tensors. Its linear span is the hyperplane $\mathbb{P}^{25}$ defined by

$$\sigma_{123} - \sigma_{132} - \sigma_{213} + \sigma_{231} + \sigma_{312} - \sigma_{321} = 0. \tag{36}$$

It is best to write the 162 quadrics in its ideal $P_{3,3,2}$ as in Example 4.14. ◇

Piecewise linear paths are also represented by $d \times m$ matrices X. Their steps are the column vectors $X_1, \ldots, X_m \in \mathbb{R}^d$. To be explicit, our path has the following parametrization:

$$t \mapsto X_1 + \cdots + X_{i-1} + (mt - i + 1) \cdot X_i$$
$$\text{where } \frac{i-1}{m} \leqslant t \leqslant \frac{i}{m} \text{ and } i = 1, 2, \ldots, m.$$

The tensor $\sigma^{(k)}(X)$ is obtained from the 'upper-triangular' $m \times m \times \cdots \times m$ tensor $\sigma^k(C_{\text{axis}})$ in Example 2.1 by multiplying each of its k sides with the matrix X. This defines a rational map (35) of degree k. The closure of its image is the *piecewise linear signature variety* $\mathcal{L}_{d,k,m}$. Its homogeneous prime ideal $L_{d,k,m}$ in $\mathbb{K}[\sigma^{(k)}]$ is called the *piecewise linear signature ideal*.

Let us reconcile these definitions with those for $k = 2$ in Section 3.1, by viewing them through the lens of Section 2.2. We are ultimately interested in

$$\begin{aligned} \mathcal{P}^{\text{im}}_{d,k,m} &:= \{\sigma^{(k)}(X) : X : [0,1] \to \mathbb{R}^d \text{ polynomial path of degree} \leqslant m\}, \\ \mathcal{L}^{\text{im}}_{d,k,m} &:= \{\sigma^{(k)}(X) : X : [0,1] \to \mathbb{R}^d \text{ piecewise linear with } m \text{ segments}\}. \end{aligned} \tag{37}$$

These signature images are semialgebraic subsets of $(\mathbb{R}^d)^{\otimes k}$. In this section we study the polynomials that vanish on these sets. They are recorded in the ideals $P_{d,k,m}$ and $L_{d,k,m}$. Equivalently, we examine the tightest outer approximations of (37) by algebraic varieties. Finding inequalities for $\mathcal{P}^{\text{im}}_{d,k,m}$ inside $\mathcal{P}^{\mathbb{R}}_{d,k,m}$, and for $\mathcal{L}^{\text{im}}_{d,k,m}$ inside $\mathcal{L}^{\mathbb{R}}_{d,k,m}$, is left to future research.

REMARK 5.2. If $m \leqslant d$ then $\mathcal{P}_{d,k,m}$ and $\mathcal{L}_{d,k,m}$ are closures of $\text{GL}(d, \mathbb{K})$-orbits.

(a) If $m = 1$ then X is a linear path and $\mathcal{L}_{d,k,1}$ is the *Veronese variety*, whose points are symmetric tensors of rank 1. In general, m plays a role similar to that of *tensor rank* in multilinear algebra [**28**].

(b) The varieties $\mathcal{L}_{d,2,m} = \mathcal{P}_{d,2,m} = \mathcal{M}_{d,m}$ were determined in Theorem 3.4.

Let X be the piecewise linear path with steps $X_1, X_2, \ldots, X_m$ in $\mathbb{R}^d$. Chen [**7**, **8**] showed that the n-step signature of the path X is given by the tensor product of

tensor exponentials:

$$\sigma^{\leqslant n}(X) = \exp(X_1) \otimes \exp(X_2) \otimes \cdots \otimes \exp(X_m) \quad \in T^n(\mathbb{R}^d). \tag{38}$$

Hence the kth signature tensor of X is the following element of $(\mathbb{R}^d)^{\otimes k}$ or $\mathbb{P}^{d^k-1}$:

$$\sigma^{(k)}(X) = \text{the sorted expansion of } \frac{1}{k!}(X_1 + X_2 + X_3 + \cdots + X_m)^{\otimes k}. \tag{39}$$

Here, by 'sorted expansion' we mean that every rank one summand $X_{i_1} \otimes X_{i_2} \otimes \cdots \otimes X_{i_k}$ is to be replaced by the corresponding rank one summand where the k indices are sorted. This replacement is done after the expansion of the m^k terms and prior to summing them.

COROLLARY 5.3. *The kth signature tensor of a piecewise polynomial path equals*

$$\sigma^{(k)}(X) = \sum_{\tau} \prod_{\ell=1}^{m} \frac{1}{|\tau^{-1}(\ell)|\,!} \cdot X_{\tau(1)} \otimes X_{\tau(2)} \otimes X_{\tau(3)} \otimes \cdots \otimes X_{\tau(k)}. \tag{40}$$

The sum is over all weakly increasing functions $\tau : \{1, 2, \ldots, k\} \to \{1, 2, \ldots, m\}$.

EXAMPLE 5.4. The third signature ($k = 3$) of a piecewise linear path X equals

$$\begin{aligned} \sigma^{(3)}(X) \;=\; & \frac{1}{6} \cdot \sum_{i=1}^{m} X_i^{\otimes 3} + \frac{1}{2} \cdot \sum_{1 \leqslant i < j \leqslant m} \left(X_i^{\otimes 2} \otimes X_j + X_i \otimes X_j^{\otimes 2} \right) \\ & + \sum_{1 \leqslant i < j < l \leqslant m} X_i \otimes X_j \otimes X_l. \end{aligned}$$

The fourth signature ($k = 4$) of a two-step path ($m = 2$) is the $d \times d \times d \times d$ tensor

$$\sigma^{(4)}(X) = \frac{1}{24} \cdot \left[X_1^{\otimes 4} + 4X_1^{\otimes 3} \otimes X_2 + 6X_1^{\otimes 2} \otimes X_2^{\otimes 2} + 4X_1 \otimes X_2^{\otimes 3} + X_2^{\otimes 4} \right]. \tag{41}$$

The projective variety $\mathcal{L}_{d,4,2}$ parametrizes tensors in $\mathbb{P}^{d^4-1}$ of this special form. ◇

EXAMPLE 5.5 ($d = m = 2, k = 4$). Consider paths consisting of two segments $X_1 = (\mathtt{a}, \mathtt{b})$ and $X_2 = (\mathtt{A}, \mathtt{B})$. The following `Macaulay2` code realizes the equation (41) and it computes the ideal $L_{2,4,2}$:

```
R = QQ[s1111,s1112,s1121,s1122,s1211,s1212,s1221,s1222,
       s2111,s2112,s2121,s2122,s2211,s2212,s2221,s2222];
S = QQ[ a,b, A,B ];
f = map(S,R,{
   a*a*a*a + 4*a*a*a*A + 6*a*a*A*A + 4*a*A*A*A + A*A*A*A,
   a*a*a*b + 4*a*a*a*B + 6*a*a*A*B + 4*a*A*A*B + A*A*A*B,
   a*a*b*a + 4*a*a*b*A + 6*a*a*B*A + 4*a*A*B*A + A*A*B*A,
   a*a*b*b + 4*a*a*b*B + 6*a*a*B*B + 4*a*A*B*B + A*A*B*B,
   a*b*a*a + 4*a*b*a*A + 6*a*b*A*A + 4*a*B*A*A + A*B*A*A,
   a*b*a*b + 4*a*b*a*B + 6*a*b*A*B + 4*a*B*A*B + A*B*A*B,
   a*b*b*a + 4*a*b*b*A + 6*a*b*B*A + 4*a*B*B*A + A*B*B*A,
   a*b*b*b + 4*a*b*b*B + 6*a*b*B*B + 4*a*B*B*B + A*B*B*B,
   b*a*a*a + 4*b*a*a*A + 6*b*a*A*A + 4*b*A*A*A + B*A*A*A,
   b*a*a*b + 4*b*a*a*B + 6*b*a*A*B + 4*b*A*A*B + B*A*A*B,
   b*a*b*a + 4*b*a*b*A + 6*b*a*B*A + 4*b*A*B*A + B*A*B*A,
   b*a*b*b + 4*b*a*b*B + 6*b*a*B*B + 4*b*A*B*B + B*A*B*B,
   b*b*a*a + 4*b*b*a*A + 6*b*b*A*A + 4*b*B*A*A + B*B*A*A,
   b*b*a*b + 4*b*b*a*B + 6*b*b*A*B + 4*b*B*A*B + B*B*A*B,
   b*b*b*a + 4*b*b*b*A + 6*b*b*B*A + 4*b*B*B*A + B*B*B*A,
   b*b*b*b + 4*b*b*b*B + 6*b*b*B*B + 4*b*B*B*B + B*B*B*B});
P = kernel f;
toString mingens P
dim P, degree P, betti mingens P
```

The output produced by this code reveals that the variety $\mathcal{L}_{2,4,2}$ is a threefold of degree 24 in a hyperplane $\mathbb{P}^{14}$ inside the space $\mathbb{P}^{15}$ of $2 \times 2 \times 2 \times 2$ tensors. The ideal $L_{2,4,2}$ has 55 quadratic minimal generators. ◇

5.2. Inclusions and separating invariants. Our aim in this subsection is to compare polynomial paths and piecewise linear paths. Some intuition for our choice of these two families is offered in Remark 5.9. Of course, numerous other families would also be interesting, including piecewise-quadratic paths, trigonometric paths, and so forth. Polynomial paths and piecewise linear paths have the same signature tensors for all cases in Remark 5.2. Our next result states that this also holds when m is large:

THEOREM 5.6. *We have the following chains of inclusions between the kth Veronese variety and the kth universal variety. Here M and M' are constants that depend only on d and k:*

$$\nu_k(\mathbb{P}^{d-1}) = \mathcal{L}_{d,k,1} \subset \mathcal{L}_{d,k,2} \subset \mathcal{L}_{d,k,3} \subset \cdots \subset \mathcal{L}_{d,k,M-1} \subset \mathcal{L}_{d,k,M} = \mathcal{U}_{d,k} \subset \mathbb{P}^{d^k-1},$$
$$\nu_k(\mathbb{P}^{d-1}) = \mathcal{P}_{d,k,1} \subset \mathcal{P}_{d,k,2} \subset \mathcal{P}_{d,k,3} \subset \cdots \subset \mathcal{P}_{d,k,M'-1} \subset \mathcal{P}_{d,k,M'} = \mathcal{U}_{d,k} \subset \mathbb{P}^{d^k-1}.$$

Proof. Chow's theorem [**20**, Theorem 7.28] says that the elements in the step-n free Lie group $\mathcal{G}^n(\mathbb{R}^d)$ are precisely the signatures (38) of piecewise linear paths. The projection of $\mathcal{G}^n(\mathbb{R}^d)$ into the kth factor of (24) is the universal signature image $\mathcal{U}^{\mathrm{im}}_{d,k}$. We conclude

$$\mathcal{U}^{\mathrm{im}}_{d,k} = \bigcup_{m=1}^{\infty} \mathcal{L}^{\mathrm{im}}_{d,k,m}.$$

By passing to Zariski closures, we obtain the same result for the projective varieties, namely $\mathcal{U}_{d,k} = \bigcup_{m=1}^{\infty} \mathcal{L}_{d,k,m}$. The right-hand side is a nested family of irreducible varieties, all contained in $\mathcal{U}_{d,k} \subset \mathbb{P}^{d^k-1}$. The number of distinct varieties in such a nested family is certainly bounded above by d^k. Hence there exists an integer M such that $\mathcal{L}_{d,k,M} = \mathcal{U}_{d,k}$.

Similarly, among the inclusions $\mathcal{P}_{d,k,i} \subset \mathcal{P}_{d,k,i+1}$, only finitely many can be strict, since each set is an irreducible variety in $\mathbb{P}^{d^k-1}$. Hence there exists a positive integer M' such that $\mathcal{P}_{d,k,M'} = \mathcal{P}_{d,k,M'+j}$ for all $j > 0$. Suppose that this terminal signature variety $\mathcal{P}_{d,k,M'}$ is strictly contained in $\mathcal{U}_{d,k}$. Then there exists a piecewise linear path X such that $S = \sigma^{(k)}(X)$ is not in $\mathcal{P}_{d,k,M'}$. In particular, there exists a polynomial $f \in P_{d,k,M'}$ such that $f(S) = 1$.

By the Weierstrass Approximation Theorem, the piecewise linear path X can be approximated arbitrarily closely by a sequence of polynomial paths X_ϵ with $\epsilon \to 0$. Here we employ a reparametrization with the property that X slows down (to velocity zero) before each kink and then speeds up again. This allows us to use the C^1 version of Weierstrass Approximation, which is what is needed here. The signature tensors $S_\epsilon = \sigma^{(k)}(X_\epsilon)$ of the nearby paths depend continuously on ϵ, and they satisfy $f(S_\epsilon) = 0$ for all $\epsilon > 0$. This implies

$$0 = \lim_{\epsilon \to 0} f(S_\epsilon) = f\big(\lim_{\epsilon \to 0} S_\epsilon\big) = f(S) = 1.$$

From this contradiction we now conclude that $\mathcal{P}_{d,k,M'} = \mathcal{U}_{d,k}$. □

It was shown in Section 3 that the signature matrices of piecewise linear and polynomial paths are the same. This result does not extend to signature tensors:

THEOREM 5.7. *The two-segment paths and the quadratic paths in the plane $\mathbb{R}^2$ have different signature threefolds. More precisely, for $k \geqslant 3$ we have $\mathcal{L}_{2,k,2} \neq \mathcal{P}_{2,k,2}$ in $\mathbb{P}^{2^k-1}$.*

Proof and discussion. For the sake of exposition, we show different proof techniques for $k = 3$ and $k = 4$. The case $k \geqslant 5$ is obtained by embedding small tensors into bigger ones.

We begin with $k = 3$. A computation shows that $M = M' = 3$ in Theorem 5.6. The ideal $U_{2,3} = L_{2,3,3} = P_{2,3,3}$ is generated by six quadrics, displayed in

Example 4.13. Both $P_{2,3,2}$ and $L_{2,3,2}$ are generated by three quadrics modulo $U_{2,3}$. For $P_{2,3,2}$ these three generators are

$$\begin{array}{c}(2\beta_1+\gamma_1)^2-\mathbf{10}(\alpha_2\gamma_1+3\alpha_1\gamma_2),\\ (2\beta_1+\gamma_1)(2\beta_2+\gamma_2)+\mathbf{10}(\alpha_3\gamma_1+\alpha_2\gamma_2),\\ (2\beta_2+\gamma_2)^2-\mathbf{10}(\alpha_3\gamma_2+3\alpha_4\gamma_1).\end{array} \tag{42}$$

The generators of $L_{2,3,2}$ are obtained by replacing the coefficient **10** by **9**.

Now let $k=4$. The varieties $\mathcal{P}_{2,4,2}$ and $\mathcal{L}_{2,4,2}$ are orbit closures for the $\mathrm{GL}(d,\mathbb{K})$-action on $\mathbb{P}^{15}$. We can use invariant theory to show that these orbits are different. According to Diehl and Reizenstein [**15**, Remark 14], the space of $\mathrm{SL}(d,\mathbb{K})$-invariants linear forms on $(\mathbb{K}^2)^{\otimes 4}$ has dimension 2 and is spanned by

$$\begin{array}{rcl}\ell_1 & = & \sigma_{1212}-\sigma_{1221}-\sigma_{2112}+\sigma_{2121}\\ \text{and}\quad \ell_2 & = & \sigma_{1122}-\sigma_{1221}-\sigma_{2112}+\sigma_{2211}.\end{array} \tag{43}$$

Their ratio ℓ_1/ℓ_2 is an absolute invariant, that is a rational function on $\mathbb{P}^{15}$ that is constant on orbits. It takes value 0 on C_{axis} and value $1/5$ on C_{mono}. Hence the orbit closures $\mathcal{L}_{2,4,2}$ and $\mathcal{P}_{2,4,2}$ are different. Indeed, `Macaulay2` confirms that ℓ_1 and $5\ell_1-\ell_2$ are the unique linear forms in the ideals $L_{2,4,2}$ and $P_{2,4,2}$. This explains the hyperplane $\mathbb{P}^{14}$ in Example 5.5. □

REMARK 5.8. It is instructive to explore the geometric meaning of linear invariants such as (43). These specify hyperplanes that contain our signature varieties. For instance, the invariant (36) is the volume of the convex hull of the path in $\mathbb{R}^3$, provided the path contains no four coplanar points [**15**, Proposition 22], or it is a limit of such paths. This volume is zero for paths in $\mathbb{R}^3$ that lie in a plane. Therefore, the linear form (36) is contained in both of the ideals $P_{3,3,2}$ and $L_{3,3,2}$.

The analogous statement holds in all dimensions d. Consider the alternating sum over all permutations, $\sum_{I\in S_d}\mathrm{sign}(I)\sigma_I$. This invariant measures the volume of the convex hull of a path in $\mathbb{R}^d$, assuming no d points on the path lie in a hyperplane. If d is even then such a path can be closed. Example 5.1 can thus be generalized as follows. The projective variety $\mathcal{P}_{d,d,d-1}$ has dimension d^2-d-1, and it lies on one hyperplane, given by the invariant above.

REMARK 5.9. One might wonder why polynomial and piecewise linear paths are so similar. The ideals $P_{d,k,m}$ and $L_{d,k,m}$ have the same numbers of minimal generators in all cases discussed so far. However, this is not always the case. Our next example will show this: the ideal $L_{d,k,m}$ can have more minimal generators than $P_{d,k,m}$. Both the similarities and the differences of our two models can perhaps be understood via degenerations of polynomial maps $\mathbb{P}^1\to\mathbb{P}^{d-1}$ to trees

Table 3. Invariants of the ideals $P_{d,k,m}$, $L_{d,k,m}$ that define the varieties $\mathcal{P}_{d,k,m}$, $\mathcal{L}_{d,k,m}$.

d	k	m	a	dim	deg	gens
2	3	2	7	3	6	9
2	3	$\geqslant 3$	7	4	4	6
2	4	2	14	3	24	55
2	4	3	15	5	$192^{\mathcal{P}}, 64^{\mathcal{L}}$	$(33^{\mathcal{P}}, 34^{\mathcal{L}}), (0^{\mathcal{P}}, 3^{\mathcal{L}}), ?$
2	4	$\geqslant 4$	15	7	12	33
2	5	2	25	3	60	220
2	5	3	31	5	$1266^{\mathcal{P}}, 492^{\mathcal{L}}$	$(160^{\mathcal{P}}, 185^{\mathcal{L}}), ?$
2	6	2	41	3	120	670
2	6	3	62	5	$4352^{\mathcal{P}}, 1920^{\mathcal{L}}$	$(945^{\mathcal{P}}, 1056^{\mathcal{L}}), ?$
3	3	2	25	5	90	162
3	3	3	26	8	$756^{\mathcal{P}}, 396^{\mathcal{L}}$	$(83^{\mathcal{P}}, 91^{\mathcal{L}}) , ?$
3	4	2	65	5	600	1536
3	4	3	80	8	?	$(1242^{\mathcal{P}}, 1374^{\mathcal{L}}) , ?$

of lines in $\mathbb{P}^{d-1}$. Indeed, an algebraic geometer might speculate that our signature varieties are related to *Kontsevich's space of stable maps* from $\mathbb{P}^1$ to $\mathbb{P}^{d-1}$.

Table 3 summarizes the computational results we found for $\mathcal{P}_{d,k,m}$ and $\mathcal{L}_{d,k,m}$. Since $\mathcal{P}_{d,2,m} = \mathcal{L}_{d,2,m} = \mathcal{M}_{d,m}$, we only consider $k \geqslant 3$. The columns have the same meanings as in Tables 1 and 2. We use upper indices $\mathcal{P}$ and $\mathcal{L}$ to mark distinctions between the polynomial case and the piecewise linear case. All our computations were done with `Macaulay2` [**21**] and `Bertini` [**4**]. To compute the degree we imposed linear constraints on the variety in question. We then pulled the resulting equations back to the parameter space, where we counted their complex zeros using either Gröbner bases or numerical homotopy methods.

EXAMPLE 5.10 ($d = 2, k = 4$). We have $M = M' = 4$ in Theorem 5.6. The varieties $\mathcal{P}_{2,4,m}$ and $\mathcal{L}_{2,4,m}$ have dimensions 1, 3, 5, 7 for $m = 1, 2, 3, 4$. The case $m = 2$ was discussed above. The case $m = 3$ reveals the distinction. The varieties $\mathcal{P}_{2,4,3}$ and $\mathcal{L}_{2,4,3}$ live in $\mathbb{P}^{15}$. Both are 5-dimensional, but their degrees differ. The former has degree 192; the latter has degree 64. The universal ideal $U_{2,4} = P_{2,4,4} = L_{2,4,4}$ has 33 quadrics, namely 4, 7, 11, 7, 4 in bidegrees (2, 6), (3, 5), (4, 4), (5, 3), (6, 2). These are also the quadrics in $P_{2,4,3}$. But, $L_{2,4,3}$ has 34 quadrics and 3 cubics. The extra quadric in $L_{2,4,3} \setminus P_{2,4,3}$ has bidegree (4, 4). ◇

When computing the degrees of our signature varieties by pulling back linear equations to the parameter spaces, it is useful to know that the varieties are rationally identifiable. In particular, their dimensions should be $dm - 1$, as expected. This is the case for all instances in Table 3. We conjecture that it holds in general. We study this topic in Section 6.

5.3. A question of Lyons and Xu. This subsection concerns *axis-parallel paths*. Each step X_i in such a path is a multiple $a_i \cdot e_{v_i}$ of a standard basis vector e_{v_i}. Combinatorially, these are characterized by the sequence $v = (v_1, v_2, \ldots, v_m) \in \{1, 2, \ldots, d\}^m$ of steps that are chosen for the path. The signature tensors of order k of axis-parallel paths form an irreducible subvariety $\mathcal{A}_{v,k}$ of $\mathcal{L}_{d,k,m}$. This variety is parametrized by varying the step lengths $a_1, a_2, \ldots, a_m$.

Unlike all the other varieties seen so far in this paper, the *axis-parallel signature variety* $\mathcal{A}_{v,k}$ is not invariant under $\mathrm{GL}(d, \mathbb{K})$. It would be interesting to study the prime ideal $A_{v,k}$ that defines $\mathcal{A}_{v,k}$, as well as geometric invariants and the fibers of its parametrization.

We illustrate the study of fibers by providing an affirmative answer to the following question due to Lyons and Xu [**34**, Question 2.5]: *Can one find a nontrivial lattice path with length shorter than 2^{n+1} such that the first n levels in its signature are all zero?* Here *length* is taken to mean discrete standard lattice length: $l = |a_1| + |a_2| + \cdots + |a_m|$.

PROPOSITION 5.11. *There exists an axis-parallel path in the plane $\mathbb{R}^2$ with $m = 8$ steps in alternating axis directions and length $l = 14 < 16 = 2^{n+1}$ whose first $n = 3$ signature tensors are all zero.*

Derivation and proof. Consider the parametrization of the varieties $\mathcal{A}_{v,k}$ where $k = 1, 2, 3$ and $v = (1, 2, 1, 2, 1, 2, 1, 2)$. These represent axis paths with steps $(a_1, 0), (0, a_2), (a_3, 0), (0, a_4), (a_5, 0), (0, a_6), (a_7, 0), (0, a_0)$. For $k = 1$ we have

$$\sigma^{(1)}(X) = (\sigma_1, \sigma_2) \quad \text{where} \quad \begin{matrix} \sigma_1 &=& a_1 + a_3 + a_5 + a_7, \\ \sigma_2 &=& a_0 + a_2 + a_4 + a_6. \end{matrix}$$

The 2×2 signature matrix $\sigma^{(2)}(X)$ has entries

$$\begin{matrix} \sigma_{11} &=& \frac{1}{2}(a_1 + a_3 + a_5 + a_7)^2 \\ \sigma_{12} &=& a_0a_1 + a_0a_3 + a_0a_5 + a_0a_7 + a_1a_2 + a_1a_4 + a_1a_6 + a_3a_4 + a_3a_6 + a_5a_6 \\ \sigma_{21} &=& a_2a_3 + a_2a_5 + a_2a_7 + a_4a_5 + a_4a_7 + a_6a_7 \\ \sigma_{22} &=& \frac{1}{2}(a_0 + a_2 + a_4 + a_6)^2. \end{matrix}$$

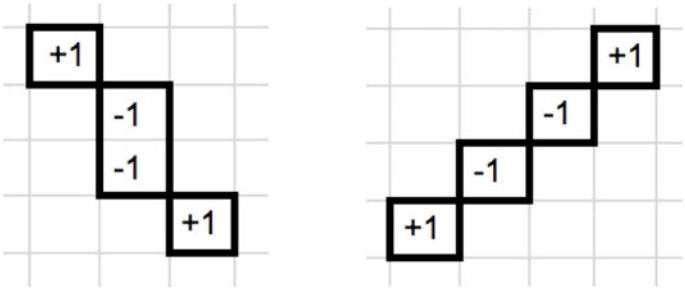

Figure 1. *Left:* the path with 8 steps and length 14 from Proposition 5.11; *Right:* path with 12 steps and length 16 from Lyons and Xu [34]. Both paths have the first 3 signature levels equal to zero. The Lévy areas are marked with the corresponding sign.

The signature tensor $\sigma^{(3)}(X)$ has format $2 \times 2 \times 2$. Its entries are the cubic polynomials

$$\begin{aligned}
\sigma_{111} &= \tfrac{1}{6}(a_1+a_3+a_5+a_7)^3\\
\sigma_{112} &= \tfrac{1}{2}a_0a_1^2+\tfrac{1}{2}a_0a_3^2+\tfrac{1}{2}a_0a_5^2+\tfrac{1}{2}a_0a_7^2+\tfrac{1}{2}a_1^2a_2+\tfrac{1}{2}a_1^2a_4+\tfrac{1}{2}a_1^2a_6+\tfrac{1}{2}a_3^2a_4+\tfrac{1}{2}a_3^2a_6+\tfrac{1}{2}a_5^2a_6+\\
&\quad a_1a_3a_4+a_1a_3a_6+a_1a_5a_6+a_3a_5a_6+a_0a_1a_3+a_0a_1a_5+a_0a_1a_7+a_0a_3a_7+a_0a_3a_5+a_0a_5a_7\\
\sigma_{121} &= a_1a_2a_3+a_1a_2a_5+a_1a_2a_7+a_1a_4a_5+a_1a_4a_7+a_1a_6a_7+a_3a_4a_5+a_3a_4a_7+a_3a_6a_7+a_5a_6a_7\\
&\dots \qquad \dots \quad \dots \quad \dots \quad \dots \quad \dots\\
\sigma_{222} &= \tfrac{1}{6}(a_0+a_2+a_4+a_6)^3
\end{aligned}$$

We seek a common zero of these $2+4+8$ expressions with no zero coordinate. The radical of the ideal they generate in $\mathbb{R}[a_1, \dots, a_7, a_0]$ is the intersection of 15 prime ideals. The first 14 primes all contain at least one of the parameters a_i as an element. The only exception is

$$\langle a_3+a_7,\ a_2+a_6,\ a_1+a_5,\ a_0+a_4,\ a_4a_5-a_5a_6+a_4a_7+a_6a_7\rangle.$$

Each zero of this ideal gives a path X with $\sigma^{(k)}(X)=0$ for $k=1,2,3$. We choose

$$a_0=-3,\quad a_1=1,\quad a_2=1,\quad a_3=-2,\quad a_4=3,\quad a_5=-1,\quad a_6=-1,\quad a_7=2.$$

This axis-parallel path with eight steps in $\mathbb{R}^2$ (see Figure 1) proves the claim. □

REMARK 5.12. The case $m=8$ is the smallest counterexample. If $m \leqslant 7$ then each zero of the corresponding ideal in $\mathbb{R}[a_1, \dots, a_m]$ represents a tree-like path. For instance, if $m=5$ then $\langle a_3, a_1+a_5, a_2+a_4\rangle$ is a prime ideal each of whose zeros gives a tree-like path.

5.4. Rough paths and the rough Veronese. We now connect our algebro-geometric study to the theory of rough paths which underlies much recent progress in stochastic analysis. This leads us to a rough version of the classical Veronese variety that represents linear paths ($m = 1$), as in Remark 5.2(a).

Every smooth path X in $\mathbb{R}^d$ lifts to a path in the Lie group $\mathcal{G}^m(\mathbb{R}^d)$, where $t \in [0, 1]$ is now mapped to the step-m signature of $X|_{[0,t]}$. In other words, we may replace the definite integrals in (1) and (2) with indefinite integrals that end at t instead of 1. We can also consider the same integrals starting at s and ending at t. The signature tensors of the resulting paths $X|_{[s,t]}$ are functions of s and t that satisfy the following *Hölder condition*:

$$\left|\sigma^{(k)}(X|_{[s,t]})\right| \lesssim |t-s|^k \quad \text{for } 0 \leqslant s < t \leqslant 1 \text{ and } k = 1, \ldots, m. \tag{44}$$

In Lyons' *rough path analysis* (see for example, [**18**, **20**, **30**] and references therein) this inequality now turns into a definition. A *rough path of order m* is any path in the Lie group $\mathcal{G}^m(\mathbb{R}^d)$ that satisfies (44) but with the exponent k replaced with k/m. Given any rough path of order m, *Lyons' Extension Theorem* provides a unique lift to a rough path of order n, for all $n > m$, such that the analytic estimate (44) remains valid for that extension. Every such rough path has a well-defined unique signature; that is, kth signature tensors for any positive integer k.

Theorem 5.6 furnishes two interpolations $\mathcal{L}_{d,k,\bullet}$ and $\mathcal{P}_{d,k,\bullet}$ between the Veronese $\nu_k(\mathbb{P}^{d-1})$ and the universal variety $\mathcal{U}_{d,k}$ for fixed d, k. In what follows we introduce a third such chain:

$$\nu_k(\mathbb{P}^{d-1}) = \mathcal{R}_{d,k,1} \subset \mathcal{R}_{d,k,2} \subset \cdots \subset \mathcal{R}_{d,k,k} = \mathcal{U}_{d,k}. \tag{45}$$

The rough Veronese variety $\mathcal{R}_{d,k,m}$ parametrizes the kth signature tensors of the following set of order m rough paths, indexed by Lie polynomials of degree m:

$$\mathbb{X}^L : t \mapsto \mathbb{X}_t = \exp(tL), \quad L \in \mathrm{Lie}^m(\mathbb{R}^d).$$

This is a path in the Lie group $\mathcal{G}^m(\mathbb{R}^d)$. It satisfies the Hölder condition $|\mathbb{X}^L_{s,t}| \lesssim |t-s|^{k/m}$ for $k \leqslant m$, with increments $\mathbb{X}^L_{s,t} = (\mathbb{X}^L_s)^{-1} \otimes \mathbb{X}^L_t$ taken with respect to the group structure in $\mathcal{G}^m(\mathbb{R}^d)$.

REMARK 5.13. We think of $\mathbb{X}^L$ as a $\mathcal{G}^m(\mathbb{R}^d)$-valued path that has constant speed L when viewed through its log-chart. In that sense, $\mathbb{X}^L$ is a log-linear path. For $m = 1$ we recover the linear paths in $\mathbb{R}^d$, whose kth signature tensors form the Veronese variety $\nu_k(\mathbb{P}^{d-1}) = \mathcal{L}_{d,k,1}$.

The signature of the geometric rough path $\mathbb{X}^L$ is simply the exponential $\sigma(\mathbb{X}^L) = \exp(L)$. The homogeneous component of degree k in this tensor series

is the signature tensor $\sigma^{(k)}(\mathbb{X}^L)$. Formally, we define the *rough Veronese variety* $\mathcal{R}_{d,k,m}$ to be the subvariety of $\mathbb{P}^{d^k-1}$ given by the degree k components of the exponential series $\exp(L)$ where L runs over $\mathrm{Lie}^m(\mathbb{R}^d)$. By construction, the inclusions (45) hold, and we have $\mathcal{R}_{d,k,m} = \mathcal{U}_{d,k}$ whenever $m \geqslant k$.

EXAMPLE 5.14 ($d = 2, m = 2$). The Lie polynomials in $\mathrm{Lie}^2(\mathbb{R}^2)$ are $\sigma = re_1 + se_2 + t[e_1, e_2]$, where r, s are parameters of degree 1 and t is a parameter of degree 2. The rough Veronese $\mathcal{R}_{2,k,2}$ is a surface in the space $\mathbb{P}^{2^k-1}$ of $2 \times 2 \times \cdots \times 2$ tensors. It parametrizes the degree k components of the exponential series $\exp(\sigma)$. For a concrete computation set $k = 3$ and revisit Examples 4.6 and 4.13. We modify our earlier `Macaulay2` code by setting $u = v = 0$. The output of the following fragment shows that $\mathcal{R}_{2,3,2}$ is a quartic surface in a $\mathbb{P}^5$ inside $\mathbb{P}^7$:

```
ExpMap = map( QQ[r,s,t] , R ,
{r^2/2, r*s/2-t, s^2/2, r^3/6, r^2*s/6,
  r^2*s/6-r*t/2, r*s^2/6, r*s^2/6-s*t/2, s^3/6,
  r, s, r*s/2+t, r^2*s/6+r*t/2, r*s^2/6+s*t/2 });
R232 = eliminate({s1,s2,s11,s12,s21,s22},kernel ExpMap)
codim R232, degree R232, betti mingens R232
```

The ideal $R_{2,3,2}$ is generated by the two linear forms $\sigma_{112} - 2\sigma_{121} + \sigma_{211}$ and $\sigma_{122} - 2\sigma_{212} + \sigma_{221}$, together with the six 2×2 minors of the 2×4 matrix $\left(\begin{smallmatrix} \sigma_{111} & \sigma_{121} & \sigma_{212} & \sigma_{112} \\ \sigma_{121} & \sigma_{212} & \sigma_{222} & \sigma_{122} \end{smallmatrix}\right)$. ◇

EXAMPLE 5.15 ($m = 2, k = 3$). Elements of $\mathrm{Lie}^2(\mathbb{R}^d)$ can be written as $X + Q$, where X is in $\mathbb{R}^d$ and Q is a skew-symmetric $d \times d$ matrix. The cubic component of $\exp(X + Q)$ equals

$$\tfrac{1}{6} X^{\otimes 3} + \tfrac{1}{2} (X \otimes Q + Q \otimes X) . \tag{46}$$

The rough Veronese $\mathcal{R}_{d,3,2}$ lives in $\mathbb{P}^{d^3-1}$. It consists of all $d \times d \times d$ tensors (46). For $d = 3$ we find that $\mathcal{R}_{3,3,2}$ has dimension 5, degree 24, and is cut out by 81 quadrics in a $\mathbb{P}^{18}$ in $\mathbb{P}^{26}$. ◇

REMARK 5.16. If we replace the skew-symmetric Q in (46) by a symmetric matrix, then $\mathcal{R}_{d,3,2}$ turns into the expected Brownian signature variety $\mathcal{B}_{d,3}$ in Section 7. This is no coincidence: $\mathbb{X}^L$ is a (deterministic) example of a Lie group valued (more precisely: $\mathcal{G}^m(\mathbb{R}^d)$-valued) *Lévy process*. These are stochastic processes with independent, stationary increments. Another example is Brownian motion $B = B(t, \omega)$ lifted to a rough path $\mathbb{B} = (1, B, \int B \otimes \circ\, dB)$ via iterated Stratonovich stochastic integration. The Lévy–Khintchine formula for such *Lévy rough paths* [**19**] provides a unified formalism to study the expected signatures.

This explains the resemblance of the rough Veronese and the expected Brownian signature variety.

A detailed study of the rough Veronese variety $\mathcal{R}_{d,k,m}$ was undertaken by Francesco Galuppi in response to this article. We might also consider a common generalization of $\mathcal{R}_{d,k,m}$ and the piecewise linear signature variety $\mathcal{L}_{d,k,m}$, namely the variety of signature tensors of piecewise log-linear paths in $\mathcal{G}^m(\mathbb{R}^d)$.

6. Dimension and identifiability

One of the motivations for this article is to develop a framework for recovering paths X from their signature tensors $\sigma^{(k)}(X)$. The material in this section serves as a step in that direction, outlining an approach based on algebraic geometry to [24, Problems 12 and 13].

Solutions for this reconstruction problem when the path X is piecewise linear were proposed by Lyons and Xu in [34, 35]. This problem amounts to solving a system of polynomial equations in the parameters x_{ij}, but this was not made explicit in the articles [34, 35]. In earlier work, Lyons and Sidorova [33] used signatures of paths for an application to sound compression. They phrase the question of how to solve for X given its signature as follows:

'. . . the problem of finding such a solution is equivalent to the problem of solving a system of polynomial equations. Provided that there exists a solution, this can be done for example by using Gröbner bases. In some sense, the remaining problem is to relate the level of truncation and the number of linear pieces in such a way that a solution exists. A disadvantage of this approach is a high complexity of the problem of solving polynomial systems' [33, Section 3].

State-of-the-art methods from applied algebraic geometry combine Gröbner bases with other tools, for example, tensor methods and numerics. We are confident that these can be developed to the point where the polynomial systems for path reconstruction will be solved in practice. Particularly promising would be an approach based on the theory of *tensor decomposition* [25]. The proof of Proposition 6.15 suggests a Tucker decomposition with a fixed core tensor.

6.1. Identifiability of the universal variety. The universal variety $\mathcal{U}_{d,k}$ comprises all kth order signature tensors of deterministic paths in $\mathbb{R}^d$. Its subvarieties $\mathcal{L}_{d,k,m}$ and $\mathcal{P}_{d,k,m}$ represent piecewise linear paths with m segments and polynomial paths of degree m, respectively. In what follows we examine whether the parameters of their natural parametrizations can be recovered from the tensors.

Table 4. The dimension of the universal variety $\mathcal{U}_{d,k}$ equals $\lambda_{d,k} - 1$.

$d \setminus k$	2	3	4	5	6	7	8	9
2	2	4	7	13	22	40	70	126
3	5	13	31	79	195	507	1317	3501
4	9	29	89	293	963	3303	11463	40583
5	14	54	204	828	3408	14568	63318	280318
6	20	90	405	1959	9694	49684	259474	1379194

Theorem 4.10 implies that the dimension of $\mathcal{U}_{d,k}$ is bounded above by $\lambda_{d,k} - 1$, where $\lambda_{d,k}$ is the number of Lyndon words of length $\leqslant k$ over the alphabet $\{1, 2, \ldots, d\}$. We compute these numbers using Proposition 4.7, and we display some values in Table 4. Note that each entry is considerably smaller than $d^k - 1$.

We now prove that the upper bound for the dimension of the universal variety is attained. We work over an algebraically closed field $\mathbb{K}$ of characteristic 0.

THEOREM 6.1. *The projection from the Lie group $\mathcal{G}_{d,n}$ onto the affine cone over the universal variety $\mathcal{U}_{d,n}$ is generically n-to-1. Hence, the projective variety $\mathcal{U}_{d,n}$ has the expected dimension $\lambda_{d,n} - 1$.*

Proof. Consider a general point S in the irreducible variety $\mathcal{G}_{d,n}$. Its coordinates σ_I are all nonzero. The point S is a sequence of tensors $S^{(1)}, S^{(2)}, S^{(3)}, \ldots, S^{(n)}$ whose formats are $d, d^2, d^3, \ldots, d^n$. The image of S under the map from $\mathcal{G}_{d,n}$ onto the affine cone over $\mathcal{U}_{d,n}$ is the nth order tensor $T = S^{(n)}$. We prove that T has precisely n preimages with coordinates in $\mathbb{K}$ under this map.

The symmetrization of T is a symmetric tensor of rank 1. It is the image of the vector $S^{(1)} = (\sigma_1, \sigma_2, \ldots, \sigma_d)$ under the nth Veronese map. This follows from the shuffle relation which states that $\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_n}$ is equal to $\sigma_{i_1 \sqcup\!\sqcup i_2 \sqcup\!\sqcup \cdots \sqcup\!\sqcup i_n}$ The latter is a linear form in the entries of the nth order tensor T, so it can be read off from T. By taking $(i_1, i_2, \ldots, i_n) = (1, 1, \ldots, 1)$, we see that there are n choices for σ_1, namely the distinct nth roots of $\sigma_{11\cdots1} = (1/n!)\sigma_1^n$. The other $d - 1$ entries of $S^{(1)}$ are found by setting $\sigma_i = \sigma_{i \sqcup\!\sqcup 1 \sqcup\!\sqcup \cdots \sqcup\!\sqcup 1}/\sigma_1^{n-1}$ for $i = 2, 3, \ldots, d$.

We can now recover the entries of $S^{(n-1)}$ uniquely. This is done by setting $\sigma_I = \sigma_{I \sqcup\!\sqcup i}/\sigma_i$ for every word I of length $n - 1$. This works because each σ_i is nonzero by our assumption on S. At this point, we proceed by downward induction on the order of the tensor, recovering next $S^{(n-2)}$ from $S^{(n-1)}$, then $S^{(n-3)}$ from $S^{(n-2)}$, and so forth. Finally, we recover $S^{(2)}$ from $S^{(3)}$.

This argument shows that T has precisely n preimages in $\mathcal{G}_{d,n}$. For any nth root of unity η, we can multiply the degree k part of the original point S by η^k and obtain another preimage S'. This proves that our map has degree n.

In particular, it preserves the dimension. The second assertion follows because we know from Theorem 4.10 that $\dim(\mathcal{G}_{d,n}) = \lambda_{d,n}$. Hence the projective variety $\mathcal{U}_{d,n}$ has dimension $\lambda_{d,n} - 1$ in $\mathbb{P}^{d^n-1}$. □

A key step in the proof was to extract the nth roots of a scalar in $\mathbb{K}$. If $\mathbb{K} = \mathbb{R}$ then the number of such roots is 1 if n is odd and 2 if n is even. This implies:

COROLLARY 6.2. *The map from the step-n free nilpotent Lie group $\mathcal{G}^n(\mathbb{R}^d)$ to the cone over the real universal variety $\mathcal{U}^{\mathbb{R}}_{d,n}$ is generically 1-to-1 if n is odd, and it is 2-to-1 if n is even.*

Recall from Theorem 4.10 that the signatures σ_I, where I runs over all Lyndon words of length n, can be used as coordinates for $\mathcal{G}^n(\mathbb{R}^d)$. In light of Theorem 6.1, we might use these Lyndon coordinates to also express the equations defining subvarieties of $\mathcal{U}_{d,n}$.

PROPOSITION 6.3. *Fix a sufficiently smooth path $X : [0, 1] \to \mathbb{R}^d$ that is not a loop, that is $X(1) \neq X(0)$. For any word I and odd integer $n > |I|$, the coordinate σ_I of the signature of X can be recovered uniquely from the nth signature tensor $\sigma^{(n)}(X)$. The same statement holds true for any rough path (cf. Section 5.4), provided its first signature tensor is nonzero.*

Proof. The hypothesis ensures that $\sigma_i \neq 0$ for $i = 1, 2, \ldots, d$, possibly after a linear change of coordinates in $\mathbb{R}^d$. The algorithm in the proof of Theorem 6.1 shows that there is a unique element in the Lie group $\mathcal{G}^n(\mathbb{R}^d)$ which maps to the signature tensor $\sigma^{(n)}(X)$. □

The hypothesis in Proposition 6.3 cannot be removed. It fails for loops.

EXAMPLE 6.4. Fix a parameter $c \in \mathbb{R}$. Using Chen's formula (38), we can construct piecewise linear paths X in the plane $\mathbb{R}^2$ whose level-3 signature is

$$\exp\big(c[e_1, e_2] + [e_1, [e_1, e_2]]\big) = \exp(L) \in \mathcal{G}^3(\mathbb{R}^2).$$

This is the signature of the order-3 rough path $\mathbb{X}^L$ with constant speed L. The third signature is $\sigma^{(3)}(X) = [e_1, [e_1, e_2]]$. For the Lyndon word $I = 12$, we have $\sigma_I = c$. This cannot be recovered from $\sigma^{(3)}(X)$. Note that $\sigma^{(i)}(X) = 0$ for $i = 1, 2$. ◇

REMARK 6.5. The rough Veronese variety $\mathcal{R}_{d,k,m}$ in Section 5.4 is identifiable in the same k-to-1 sense as the universal variety. Given a generic point $\sigma^{(k)}(\mathbb{X}^L)$ in $\mathcal{R}_{d,k,m}$, we recover $\sigma(\mathbb{X}^L) = \exp(L)$ via Theorem 6.1, and from this we compute L by taking the logarithm.

6.2. Recovering paths from signature tensors. Our main reconstruction problem is as follows. Given a real $d \times d \times \cdots \times d$ tensor S, we seek to find a piecewise linear or polynomial path X of small complexity m such that $S = \sigma^{(k)}(X)$. Of course, a necessary condition is that S lies in $\mathcal{L}^{\mathrm{im}}_{d,k,m}$ respectively $\mathcal{P}^{\mathrm{im}}_{d,k,m}$. Assuming this to be the case, we hope to express each parameter x_{ij} as a function of the entries σ_I of the tensor S.

In this subsection we discuss a variant of that reconstruction problem, namely recovery up to scaling. We replace the signature images above with the signature varieties $\mathcal{L}_{d,k,m}$ and $\mathcal{P}_{d,k,m}$. These are the closed images of our rational maps $\mathbb{P}^{md-1} \dashrightarrow \mathbb{P}^{d^k-1}$, $X \mapsto \sigma^{(k)}(X)$. We seek rational formulas that invert these maps on the signature varieties. As before, the points in the projective spaces have their coordinates in a field $\mathbb{K}$ that is algebraically closed.

The corresponding polynomial map $\mathbb{K}^{md} \to \mathbb{K}^{d^k}$ does not admit a rational inverse on the affine cones of our signature varieties. The reason is that the signature tensor $S = \sigma^{(k)}(X)$ remains unchanged if we multiply the parameter matrix $X = (x_{ij})$ by a kth root of unity.

REMARK 6.6 ($d = m = 2, k = 3$). Consider the eight equations seen in Example 2.4, where the $S = (\sigma_{ijk})$ obey the quadratic constraints in $P_{2,3,2}$. We seek to solve the eight equations for $x_{11}, x_{12}, x_{21}, x_{22}$. To do this, we must extract a third root, for example, as in $x_{11} + x_{21} = (6\sigma_{111})^{1/3}$. These roots of unity disappear when we work in projective space. In other words, if we regard both X and S up to scale, then we can compute formulas that express X as a rational function in the points S of the signature variety $\mathcal{P}_{2,3,2}$. Gröbner bases will find such formulas when they exist. We demonstrate this for the simplest piecewise linear paths.

EXAMPLE 6.7 ($d = m = 2, k = 3$). The reconstruction of two-segment paths in the plane from their $2 \times 2 \times 2$ signature tensors is solved by running the following `Macaulay2` code:

```
R = QQ[x11,x12,x21,x22,s111,s112,s121,s122,s211,s212,s221,s222,
     MonomialOrder => Lex];
M = matrix {
 { s111, 1/6*(x11+x21)^3                                                  },
 { s112, 1/6*x11^2*x12+1/2*x11^2*x22+1/2*x11*x21*x22+1/6*x21^2*x22        },
 { s121, 1/6*x11^2*x12+1/2*x11*x12*x21+1/2*x11*x21*x22+1/6*x21^2*x22},
 { s122, 1/6*x11*x12^2+1/2*x11*x12*x22+1/2*x11*x22^2+1/6*x21*x22^2        },
 { s211, 1/6*x11^2*x12+1/2*x11*x12*x21+1/2*x12*x21^2+1/6*x21^2*x22        },
 { s212, 1/6*x11*x12^2+1/2*x11*x12*x22+1/2*x12*x21*x22+1/6*x21*x22^2},
 { s221, 1/6*x11*x12^2+1/2*x12^2*x21+1/2*x12*x21*x22+1/6*x21*x22^2        },
 { s222, 1/6*(x12+x22)^3                                                 }};
I = saturate( minors(2,M) , ideal(x11+x21,x12+x22) );
toString gens gb I
```

The ideal I represents the graph of the map $\sigma^{(3)} : \mathbb{P}^3 \dashrightarrow \mathbb{P}^7$ whose image is the threefold $\mathcal{L}_{2,3,2}$. The output of the Macaulay2 code is the reduced Gröbner basis for I. This includes nine quadrics that generate $L_{2,3,2}$. In the Gröbner basis we also see the following polynomials:

$$\begin{aligned}(\sigma_{122}-\sigma_{212})\cdot \underline{x_{21}} - (\sigma_{121}-\sigma_{211})\cdot x_{22},\\ (\sigma_{121}-\sigma_{211})\cdot \underline{x_{12}} - (\sigma_{212}-\sigma_{221})\cdot x_{21},\\ 3\sigma_{211}\cdot \underline{x_{11}} - 3\sigma_{111}\cdot x_{12} + (\sigma_{211}-\sigma_{121})\cdot x_{21}.\end{aligned} \tag{47}$$

This is a rational formula for $X \in \mathbb{P}^3$ in terms of the point $\sigma^{(3)}(X)$ in $\mathcal{L}_{2,3,2} \subset \mathbb{P}^7$.

The same method works for quadratic paths in the plane, namely we can invert the map $\sigma^{(3)} : \mathbb{P}^3 \dashrightarrow \mathcal{P}_{2,3,2} \subset \mathbb{P}^7$ by running Macaulay2. In the lexicographic Gröbner basis we find

$$\begin{aligned}3(\sigma_{112}-\sigma_{211}+\sigma_{212})\cdot \underline{x_{21}} + (\sigma_{112}+\sigma_{121}-5\sigma_{211})\cdot x_{22},\\ (\sigma_{122}-2\sigma_{212}+\sigma_{221})\cdot \underline{x_{12}} + (\sigma_{112}-2\sigma_{121}+\sigma_{211})\cdot x_{22},\\ 5\sigma_{121}\cdot \underline{x_{11}} - 5\sigma_{111}\cdot x_{21} - (\sigma_{112}-5\sigma_{121}-\sigma_{211})\cdot x_{12} - 5\sigma_{111}\cdot x_{22}.\end{aligned} \tag{48}$$

This is a formula for $X \in \mathbb{P}^3$ in terms of the point $\sigma^{(3)}(X)$ in $\mathcal{P}_{2,3,2} \subset \mathbb{P}^7$. ◇

Counting parameters gives an upper bound on the dimension of our varieties:

$$\dim(\mathcal{L}_{d,k,m}) \leqslant \min\{\lambda_{d,k}-1, dm-1\} \quad \text{and} \quad \dim(\mathcal{P}_{d,k,m}) \leqslant \min\{\lambda_{d,k}-1, dm-1\}. \tag{49}$$

If the dimension equals $dm-1$ then the variety is *algebraically identifiable*. This means that, for some positive integer r, the map from $d \times m$ matrices that represent paths X to their signature tensors $\sigma^{(k)}(X)$ is r-to-1 even up to scaling. If $r = 1$ then the map is *birational*, and the variety is *rationally identifiable*. This best-case scenario happened in Example 6.7.

In general, our signature varieties can be *nonidentifiable*, that is the inequalities (49) are strict. This happens for the matrix case ($k = 2$) studied in Section 3. We saw in Theorem 3.4 that $\mathcal{L}_{d,2,m} = \mathcal{P}_{d,2,m}$ has dimension $dm - \binom{m}{2} - 1$. For $2 \leqslant m < d$, this is strictly less than the minimum of $dm-1$ and $\lambda_{2,d}-1 = (d+2)(d-1)/2$ on the right-hand side of (49). In the nonidentifiable case, the fibers of the map $X \mapsto \sigma^{(k)}(X)$ are positive-dimensional. It is interesting to study the geometry of such families of paths with fixed signature tensor.

EXAMPLE 6.8 ($k = m = 2$). We fix the $d \times d$ signature matrix S given by a two-segment path X or a quadratic path X. Suppose the path runs from $\mathbf{0} = (0, 0, \ldots, 0)$ to $\mathbf{1} = (1, 1, \ldots, 1)$. For a two-segment path, we write $X_1 = (a_1, a_2, \ldots, a_d)$ for the first step and hence $X_2 = (1-a_1, 1-a_2, \ldots, 1-a_d)$

for the second step. Then $\sigma^{(1)}(X) = \mathbf{1}$ and $S = \sigma^{(2)}(X) = \frac{1}{2}(\mathbf{1} \cdot \mathbf{1}^T + Q)$, where Q is the skew-symmetric matrix whose entries q_{ij} record the Lévy areas

$$q_{ij} = a_i - a_j \quad \text{for } 1 \leqslant i, j \leqslant d. \tag{50}$$

From this we see that $\mathcal{L}_{d,2,2}$ is not identifiable. The fiber over S is a curve, representing all paths obtained from X by adding a multiple of $\mathbf{1} = (1, 1, \ldots, 1)$ to $X_1 = (a_1, a_2, \ldots, a_d)$. For a quadratic path, we write the ith coordinate as $3a_i t + (1 - 3a_i)t^2$. Then we get the same signature matrix as above. In particular, (50) holds and the fiber is the same curve. ◇

EXAMPLE 6.9 ($k = 2, m = 3$). Let $d \geqslant 5$ and consider three-segment paths X that go from $\mathbf{0}$ to $(a_1, a_2, \ldots, a_d)$, then to $(b_1, b_2, \ldots, b_d)$, and finally to $\mathbf{1}$. We fix the signature matrix $S = \sigma^{(2)}(X) = \frac{1}{2}(\mathbf{1} \cdot \mathbf{1}^T + Q)$ as above. The Lévy areas of the coordinate projections are

$$q_{ij} = (a_i - 1)b_j - (a_j - 1)b_i \quad \text{for } 1 \leqslant i, j \leqslant d. \tag{51}$$

The q_{ij} are now fixed. The fiber over S, consisting of all three-segment paths with Lévy areas q_{ij}, is a threefold. Inside the 4-dimensional affine subspace of $\mathbb{R}^{2d}$ that is cut out by

$$\begin{aligned} &q_{ij}b_k - q_{ik}b_j + q_{jk}b_i \\ &\quad = q_{ij}(a_k - 1) - q_{ik}(a_j - 1) + q_{jk}(a_i - 1) = 0 \quad \text{for } 1 \leqslant i < j < k \leqslant d, \end{aligned}$$

our threefold is the hypersurface defined by any one of the equations in (51). ◇

We now come to the general case $k \geqslant 3$. We present two conjectures on identifiability.

CONJECTURE 6.10. *Let $k \geqslant 3$. Then equality holds in (49) for all integers d, $m \geqslant 2$.*

If this is true then we know the thresholds for stabilization in Theorem 5.6.

COROLLARY 6.11. *If Conjecture 6.10 holds then the constants in Theorem 5.6 are*

$$M = M' = \left\lceil \frac{\lambda_{d,k}}{d} \right\rceil. \tag{52}$$

Furthermore, for all $m \geqslant M$, we have $\mathcal{P}_{d,k,m} = \mathcal{L}_{d,k,m}$, and this variety equals the universal variety $\mathcal{U}_{d,k}$.

Table 5. The formula in (52) suggests the value M at which the signature varieties stabilize.

$d \setminus k$	3	4	5	6	7	8	9
2	3	4	7	12	21	36	64
3	5	11	27	66	170	440	1168
4	8	23	74	241	826	2866	10 146
5	11	41	166	682	2914	12 664	56 064
6	16	68	327	1616	8281	43 246	229 866

Proof. All three varieties $\mathcal{P}_{d,k,m}, \mathcal{L}_{d,k,m}$ and $\mathcal{U}_{d,k}$ are irreducible of the same dimension. The first two are contained in the third. Hence all three coincide. □

The following stronger conjecture generalizes the conclusion of Example 6.7.

CONJECTURE 6.12. *If $k \geqslant 3$ and m is strictly less than the constant M in Corollary 6.11 then both signature varieties $\mathcal{P}_{d,k,m}$ and $\mathcal{L}_{d,k,m}$ are rationally identifiable.*

The hypothesis $m < M$ is essential. Conjecture 6.12 fails when $m = M = \lambda_{d,k}/d$.

REMARK 6.13. The question of identifiability is delicate in the borderline case $\lambda_{d,k} = md$, when the signature variety exactly fills the universal variety. Here we expect algebraic identifiability (by Conjecture 6.10), but rational identifiability generally fails. This can be seen for $d = 2, k = m = 4$. The 7-dimensional variety $\mathcal{P}_{2,4,4} = \mathcal{L}_{2,4,4} = \mathcal{U}_{2,4}$ has degree 12 in $\mathbb{P}^{15}$. We have two parametrizations from the $\mathbb{P}^7$ of 2×4 matrices. The parametrization using quartic paths in $\mathbb{R}^2$ is 48-to-1. The parametrization using four-segment paths in $\mathbb{R}^2$ is only 4-to-1.

EXAMPLE 6.14 ($d = 2, m = k = 4$). Consider the four-step path in $\mathbb{R}^2$ given by

$$X = \begin{bmatrix} 29 & 15 & 13 & 2 \\ 23 & 26 & 6 & 27 \end{bmatrix}.$$

There are three other paths that have the same $2 \times 2 \times 2 \times 2$ signature tensor:

$$\begin{bmatrix} 36.74838 & -17.80169 & 37.75532 & 2.29799 \\ 27.39596 & -9.82926 & 40.23084 & 24.20246 \end{bmatrix},$$

$$\begin{bmatrix} 102.16286 & -131.13298 & 85.92484 & 2.04528 \\ 104.55786 & -136.84738 & 86.56467 & 27.72484 \end{bmatrix},$$

$$\begin{bmatrix} 38.53237 & 38.8057 & -79.20533 & 60.86735 \\ 28.69523 & 82.7734 & -147.7839 & 118.3152 \end{bmatrix}.$$

Each general point in $\mathcal{L}^{\text{im}}_{2,4,4}$ has four preimages. At least two of them are real. ◇

In the next result we reduce Conjectures 6.10 and 6.12 to 3-way tensors and paths that span their ambient space. This uses concepts from the theory of tensor decomposition [**25**].

PROPOSITION 6.15. *Fix integers d, k, m that satisfy $d \geqslant m \geqslant 1$ and $k \geqslant 3$.*

(a) *If $\mathcal{L}_{m,3,m}$ is rationally (respectively algebraically) identifiable then so is $\mathcal{L}_{d,k,m}$.*

(b) *If $\mathcal{P}_{m,3,m}$ is rationally (respectively algebraically) identifiable then so is $\mathcal{P}_{d,k,m}$.*

Proof. We claim that identifiability, for both piecewise linear and polynomial paths, can be reduced to the case $k = 3$ and $d = m$. In either case, the reduction to $k = 3$ follows from Theorem 6.1. If we are given the signature tensor $\sigma^{(k)}(X)$ of a path X for $k \geqslant 3$ then we can recover each tensor $\sigma^{(l)}(X)$ for $l \leqslant k$ up to a multiplicative constant. In particular, for $l = 3$, we obtain a unique point $\sigma^{(3)}(X)$ in the projective variety $\mathcal{L}_{d,3,m}$ or $\mathcal{P}_{d,3,m}$ respectively.

We next show that the case $d > m$ reduces to the case $d = m$. Our path is encoded in a $d \times m$ matrix X. Its third signature tensor has the format $d \times d \times d$:

$$\sigma^{(3)}(X) = [[C_\bullet; X, X, X]] = C_\bullet \times_1 X \times_2 X \times_3 X. \tag{53}$$

Here $C_\bullet$ is a tensor of format $m \times m \times m$ that is independent of X. It represents either C_{axis} in Example 2.1 or C_{mono} in Example 2.2. The notation used in (53) is standard in the literature on *tensor decomposition*, for example, [**14**] or [**25**, Section 4, equation (9)]. On the right in (53) we take the product of a small 3-way tensor with the same matrix on all three sides to create a larger 3-way tensor.

We claim that the *core tensor* $C_\bullet$ has rank $\geqslant m$. Consider the $m \times m$ matrix $\tilde{C}_\bullet$ seen in the first slice ($i = 1$) of that tensor. It suffices to show that this matrix is invertible. For the axis paths, the matrix $\tilde{C}_{\text{axis}}$ is invertible because it is upper-triangular with nonzero diagonal entries. For the monomial path, the $m \times m$ matrix $\tilde{C}_{\text{mono}}$ has the entries $(jk)/((j+1)(j+k+1))$. After scaling rows and columns, this is a Cauchy matrix. Similarly to (21), we compute

$$\det(\tilde{C}_{\text{mono}}) = \frac{d!}{d+1} \cdot \frac{\prod_{1 \leqslant i < j \leqslant d} (j-i)^2}{\prod_{i=1}^{d} \prod_{j=1}^{d} (i+j+1)} \neq 0.$$

We can assume that also the matrix X has rank m since the definitions of rational and algebraic identifiability refer to a path that is sufficiently general. Under these hypotheses, the *Tucker decomposition* of a 3-way tensor is unique in the following sense. All solutions (D, Y) to the system of d^3 polynomial equations given by $[[D; Y, Y, Y]] = \sigma^3(X)$ have the form $D = [[C_\bullet; g, g, g]]$ and $Y = \eta g^{-1} X$ for some $g \in \mathrm{GL}(m, \mathbb{R})$ and some $\eta \in \mathbb{C}$ with $\eta^3 = 1$.

Fix any such solution (D, Y). Consider the system of m^3 polynomial equations given by $[[C_\bullet; g, g, g]] = D$. By identifiability for $d = m$, this has a unique solution $g \in \mathrm{GL}(m, \mathbb{R})$. Here uniqueness is up to scale. We recover the matrix encoding our path by setting $X = gY$. □

We established the following partial result using computational methods.

LEMMA 6.16. *For both signature varieties $\mathcal{L}_{m,3,m}$ and $\mathcal{P}_{m,3,m}$ in $\mathbb{P}^{m^3-1}$, rational identifiability holds up to $m = 5$, and algebraic identifiability holds up to $m = 15$.*

Computational proof. We prove algebraic identifiability by examining the Jacobian matrix of the parametrization $X \mapsto \sigma^{(3)}(X)$. This is a matrix of format $m^2 \times m^3$ whose entries are homogeneous quadrics in the parameters x_{ij}. We show that this matrix has rank m^2. We do this by computing the rank for one random choice of integer parameters. This has rank m^2, and we conclude that the resulting variety in $\mathbb{P}^{m^3-1}$ has the expected dimension $m^2 - 1$.

We prove rational identifiability by checking it for one random matrix $X_0 \in \mathbb{Z}^{m \times m}$. The identity $\sigma^{(3)}(X) = \sigma^{(3)}(X_0)$ translates into a system of m^3 inhomogeneous cubic equations in m^2 unknowns x_{ij}, namely the entries of X. We compute a Gröbner basis for this system. The output shows that it has precisely three solutions, namely $X = \eta X_0$ where $\eta^3 = 1$. □

COROLLARY 6.17. *Conjecture 6.10 holds when $m \leqslant 15$. Conjecture 6.12 holds when $m \leqslant 5$.*

Proof. If $d \geqslant m$ then this follows immediately from Proposition 6.15 and Lemma 6.16. For $d < m$ this requires separate computations, like those in the derivation of Lemma 6.16. □

REMARK 6.18. Chen's formula (38) offers an alternative approach to algebraic identifiability of $\mathcal{L}_{d,k,m}$. The product rule of calculus is valid in the tensor algebra, and can be used to write the Jacobian of the map that takes $X = (X_1, \ldots, X_m)$ to (38) and then further to $\sigma^{(3)}(X)$. This leads to a tensor algebra formula for the Jacobian that is inductive in m. Using this, for small values of m, we also derived Proposition 6.15 and checked the rank of the Jacobian.

7. Expected signatures

We now shift gears, in that we replace our deterministic paths in $\mathbb{R}^d$ with random paths $X = X(\omega)$. The signature $\sigma^{(k)}(X)$ is a well-defined tensor-valued random variable, which takes values in the real universal variety $\mathcal{U}_{d,k}^{\mathbb{R}}$ in affine space $(\mathbb{R}^d)^{\otimes k}$, whenever sample paths are sufficiently regular, or a stochastic integration theory which respects the chain-rule is used. Thus, any such process, with randomly generated paths, induces a probability distribution on $\mathcal{U}_{d,k}^{\mathbb{R}}$. In this section we study projective varieties in $\mathbb{P}^{d^k-1}$ that arise from the expectation of such a probability distribution on paths.

7.1. Brownian motion. Following Lyons [11, 30], we are interested in the *expected signature tensor* $\mathbb{E}(\sigma^{(k)}(X))$. This is a convex combination in $(\mathbb{R}^d)^{\otimes k}$ of points in $\mathcal{U}_{d,k}^{\mathbb{R}}$, weighted by the probability distribution. Since the universal variety $\mathcal{U}_{d,k}^{\mathbb{R}}$ is not a linear space, the tensor $\mathbb{E}(\sigma^{(k)}(X))$ is usually not a point in $\mathcal{U}_{d,k}^{\mathbb{R}}$. The varieties of expected signature tensors that follow all live in $\mathbb{P}^{d^k-1}$. However, unlike the signature varieties in Section 5, they are now not subvarieties of $\mathcal{U}_{d,k}$.

The prototype of a stochastic process in $\mathbb{R}^d$ is the simple random walk $Z = Z(\omega, d)$ that starts at the origin $Z_0 = 0$ and takes equally likely axis-parallel steps $Z_{i+1} - Z_i$. These steps are independent and identically distributed (i.i.d.) in $\{e_1, -e_1, \ldots, e_d, -e_d\}$, with covariance given by $d^{-1}I$ where $I = \sum_{i=1}^{d} e_i^{\otimes 2}$ is the $d \times d$ identity matrix. A realization of this process, on $\{1, \ldots, m\}$ say, is viewed as random axis-parallel path in $\mathbb{R}^d$, moving linearly at unit speed from Z_i to Z_{i+1}.

The Central Limit Theorem suggests that we rescale the path as follows. Set

$$Z_t^{(m)} := m^{-1/2} Z_{mt}$$

where now $Z^{(m)}$ is a path on $[0, 1]$. This is the construction of Section 5.3, but with random steps $X_i := m^{-1/2}(Z_{i+1} - Z_i), i = 1, \ldots, m$. Under the given assumptions, one computes

$$\begin{aligned} \mathbb{E}\big(\sigma(Z^{(m)})\big) &= (\mathbb{E}(\exp(X_i)))^{\otimes m} \\ &= (2d)^{-m} \left(\sum_{i=1}^{d} \exp(m^{-1/2} e_i) + \sum_{i=1}^{d} \exp(-m^{-1/2} e_i) \right)^{\otimes m}. \end{aligned}$$

This expression converges, as $m \to \infty$, to the tensor exponential $\exp(1/(2d)I)$; equivalently the expected signature of $d^{1/2} Z^{(m)}$ converges to $\exp(\frac{1}{2} I)$. On the other hand, by the central limit theorem, have convergence in law $d^{1/2} Z_1^{(m)} \to N(0, I)$. More generally, by *Donsker's invariance principle* (a.k.a. functional central

limit theorem), the random path $d^{1/2}Z^{(m)}$ converges in law to a d-dimensional 'standard' Brownian motion B (with zero drift and unit covariance). Hence, the tensor series $\exp(\frac{1}{2}I)$ has the natural interpretation as *expected signature of Brownian motion*. This is known as *Fawcett's formula*, which is a starting point for stochastic integration and rough paths. For details we refer to the textbook [**18**].

We now fix $\mu \in \mathbb{R}^d$ and $\Sigma \in (\mathbb{R}^d)^{\otimes 2}$ symmetric and positive definite. By adjusting the discrete walk approximations, or by a direct transformation of the form $X_t := \mu t + \sqrt{\Sigma}B_t$, one obtains *Brownian motion* X on $[0, 1]$ with drift μ and covariance Σ. Note that $X_1 \sim N(\mu, \Sigma)$.

Both μ and Σ are elements in $T^n(\mathbb{R}^d)$, provided $n \geqslant 2$. Hence so is the sum $\mu + \frac{1}{2}\Sigma$ and its exponential. If we define $\sigma(X)$ via iterated *Stratonovich stochastic integration* then, by [**18**, Example 3.22], the following identity holds in the truncated tensor algebra $T^n(\mathbb{R}^d)$:

$$\mathbb{E}(\sigma(X)) = \exp\big(\mu + \tfrac{1}{2}\Sigma\big). \tag{54}$$

The coefficients $\sigma_{i_1 i_2 \cdots i_k}$ of this tensor series refine classical Gaussian moments. We think of them as 'noncommutative moments'. Indeed, $\sigma_{i_1 i_2 \cdots i_k}$ plays the role of the moment associated with the monomial $e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k}$ in $T^n(\mathbb{R}^d)$. We note that all moments of multivariate Gaussians are linear expressions in the expected signatures of Brownian motion.

PROPOSITION 7.1. *Fix a d-dimensional Gaussian $Z \sim N(\mu, \Sigma)$ and corresponding Brownian motion X with drift μ and covariance Σ. For $u \in \mathbb{N}^d$ consider the shuffle product $w(u) = 1^{\sqcup\!\sqcup u_1} \sqcup\!\sqcup \cdots \sqcup\!\sqcup d^{\sqcup\!\sqcup u_d}$. Then the Gaussian moment $m_u = \mathbb{E}\left(Z_1^{u_1} \cdots Z_d^{u_d}\right)$ is equal to the shuffle linear form $\sigma_{w(u)}$ evaluated at the coefficients of the expected signature $\mathbb{E}(\sigma(X))$ in (54).*

Here is an example for $d = 3$. The moment $m_{(1,1,1)} = \mathbb{E}(Z_1 Z_2 Z_3)$ equals

$$\sigma_{w(1,1,1)} = \sigma_{1 \sqcup\!\sqcup 2 \sqcup\!\sqcup 3} = \sigma_{123} + \sigma_{132} + \sigma_{213} + \sigma_{231} + \sigma_{312} + \sigma_{321}. \tag{55}$$

If we write $\Sigma = (s_{ij})$ then the six signatures in (55) are $\sigma_{ijk} = \frac{1}{6}\mu_i\mu_j\mu_k + \frac{1}{4}\mu_i s_{jk} + \frac{1}{4}s_{ij}\mu_k$. Their sum is the familiar formula for Gaussian moments: $m_{(1,1,1)} = \mu_1\mu_2\mu_3 + \mu_1 s_{23} + \mu_2 s_{13} + \mu_3 s_{12}$.

Proof. If we replace each noncommutative monomial $e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k}$ in the series (54) by the corresponding commutative monomial $t_{i_1} t_{i_2} \cdots t_{i_k}$ then we obtain the familiar moment generating function for Gaussians, displayed in, for example, [**2**, equation (1)]. This symmetrization map from tensor series to commutative series aggregates the coefficients as sums $m_u = \sigma_{w(u)}$. □

If the covariance matrix Σ is set to 0 then the path is linear, and all signature tensors are symmetric of rank 1. In this case, the moment m_u equals σ_I times the multinomial coefficient $\binom{|u|}{u}$ where $I = 1^{u_1}2^{u_2}\cdots d^{u_d}$. Geometrically: moment variety equals signature variety equals Veronese. This remains true if Σ is set to a fixed constant matrix (see Theorem 7.9).

In what follows, we consider Brownian motion X and its associated Gaussian Z parametrized by (μ, Σ) as above. We focus on the degree k component in the tensor series (54). This $d \times d \times \cdots \times d$ tensor is denoted by $\mathbb{E}(\sigma^{(k)}(X))$ and called the *expected Brownian signature tensor* of step k. Each of its entries $\sigma_{i_1 i_2 \cdots i_k}$ is a homogeneous polynomial of degree k in the $d + \binom{d+1}{2}$ model parameters. Here, the d parameters coming from μ have degree one, and the $\binom{d+1}{2}$ parameters coming from Σ have degree two. One can derive explicit formulas for $\sigma^{(k)}(X)$ in terms of μ and Σ, similar to those in Example 5.4. See Example 5.15 and Remark 7.7 for the analogous situation where Σ is skew-symmetric instead of symmetric.

We define the *expected Brownian signature variety* to be the subvariety of $\mathbb{P}^{d^k-1}$ that is parametrized by the tensors $\sigma^{(k)}(X)$. This projective variety is denoted by $\mathcal{B}_{d,k}$. Similarly, we can also define a projective variety $\mathcal{B}_{d,\leqslant n}$ that lives in the projective space associated with $T^n(\mathbb{R}^d)$. The premise of this section is the study of these varieties and their secant varieties.

EXAMPLE 7.2. Let $k = 2$. The expected Brownian signature matrix is given by $\frac{1}{2}(\mu\mu^T + \Sigma)$, and as such it is symmetric. It equals half of the Gaussian second order moment matrix. The variety $\mathcal{B}_{d,2}$ lives in $\mathbb{P}^{d^2-1}$ but it is actually a linear subspace of dimension $\binom{d+1}{2} - 1$, defined by the equations that reflect the symmetry: $b_{ij} - b_{ji} = 0$ for $1 \leqslant i < j \leqslant d$. ◇

The symmetry observed in Example 7.2 generalizes to expected signatures of order $k \geqslant 3$.

REMARK 7.3. The projective variety $\mathcal{B}_{d,k}$ lives in $\mathbb{P}^{d^k-1}$ and its dimension is bounded above by $d + \binom{d+1}{2} - 1$. It lives in a linear subspace of dimension $(d^k + d^{\lceil k/2 \rceil})/2 - 1$, namely the set of tensors that satisfy the reflectional symmetries $\sigma_{i_1 i_2 \cdots i_{k-1} i_k} = \sigma_{i_k i_{k-1} \cdots i_2 i_1}$. This holds because every term in the expansion of $\mathbb{E}(\sigma^{(k)}(X)) = \left(\mu + \frac{1}{2}\Sigma\right)^{\otimes k}$ has a matching term with respect to reflectional symmetry. For instance, for $k = 5$, the term $\mu \otimes \Sigma \otimes \mu \otimes \mu$ matches $\mu \otimes \mu \otimes \Sigma \otimes \mu$.

EXAMPLE 7.4. The expected Brownian signature tensor variety $\mathcal{B}_{2,3}$ is a cubic threefold in $\mathbb{P}^7$. Its ambient dimension is 5, thanks to the two relations $b_{122} = b_{221}$

Table 6. Invariants of the ideals $B_{d,k}$ and $B_{d,\leqslant n}$, that define the varieties $\mathcal{B}_{d,k}$ and $\mathcal{B}_{d,\leqslant n}$.

d	k	a	dim	deg	gens
d	2	$\binom{d+1}{2}-1$	$\binom{d+1}{2}-1$	1	0
d	$\leqslant 2$	$\binom{d+2}{2}-1$	$\binom{d+2}{2}-1$	1	0
2	3	5	3	3	3
2	$\leqslant 3$	11	5	16	8, 19
3	3	16	7	21	45
3	$\leqslant 3$	27	9	130	77, 175
2	4	9	4	12	10
2	$\leqslant 4$	22	5	102	63, 40

and $b_{112} = b_{211}$ from Remark 7.3. In that $\mathbb{P}^5$, our variety has codimension 2 and is cut out by the three 2×2 minors of

$$\begin{pmatrix} b_{111} & b_{121} & 2b_{122} - b_{212} \\ 2b_{112} - b_{121} & b_{212} & b_{222} \end{pmatrix}. \tag{56}$$

If we also consider the first and second order signatures, then we obtain the variety $\mathcal{B}_{2,\leqslant 3}$ sitting in $\mathbb{P}^{14}$ but with ambient space $\mathbb{P}^{11}$. The linear relation $b_{12} = b_{21}$ gets added to the two above. This resulting variety is 5-dimensional of degree 16. Its ideal $B_{2,\leqslant 3}$ is generated by 8 quadrics and 19 cubics. The five extra quadrics, beyond those in (56), are

$$\begin{matrix} b_1 b_{212} - b_2 b_{121}, \quad b_1 b_{222} - b_2(2b_{122} - b_{212}), \quad b_1(2b_{112} - b_{121}) - b_2 b_{111}, \\ b_1 b_{22} - b_2 b_{12} + 2b(b_{212} - b_{122}), \quad b_1 b_{12} - b_2 b_{11} + 2b(b_{112} - b_{121}). \end{matrix} \tag{57}$$

Here b is the homogenizing variable from the constant zero order signature. ◇

Table 6 becomes even more interesting once we compare with the Gaussian moment varieties in [2, 3]. We denote these now by $\mathcal{N}_{d,\leqslant n}$. They are subvarieties of $\mathbb{P}^{\binom{d+n}{d}-1}$ consisting of the vectors of all moments of order $\leqslant n$ of a Gaussian distribution on $\mathbb{R}^d$. These moments are polynomial expressions parametrized by the entries of the Gaussian's mean $\mu \in \mathbb{R}^d$ and covariance matrix Σ. Propositions 5, 7 and 9 in [2] describe the dimension, degree and minimal generators for $\mathcal{N}_{2,\leqslant 3}$, $\mathcal{N}_{2,\leqslant 4}$ and $\mathcal{N}_{3,\leqslant 3}$ respectively. The first two items match the ones obtained for $\mathcal{B}_{d,\leqslant n}$ with the same values of d, n. This highlights the similarity between expected signatures of Brownian motion and moments of Gaussians.

While the degrees of the generators of $\mathcal{N}_{d,\leqslant n}$ increase with d, n (for example, there are octic generators in $\mathcal{N}_{3,\leqslant 3}$), it appears that $\mathcal{B}_{d,\leqslant n}$ only has

generators up to degree 3. We record the following basic fact concerning the map between the two varieties given in Proposition 7.1.

PROPOSITION 7.5. *The linear projection from the expected signature variety $\mathcal{B}_{d,\leqslant n}$ onto the Gaussian moment variety $\mathcal{N}_{d,\leqslant n}$ is birational. Both varieties have dimension $\binom{d+1}{2} + d - 1$.*

Proof. The Gaussian moment variety is rationally identifiable. We can recover μ and Σ rationally from a general point in $\mathcal{N}_{d,\leqslant n}$. The preimage of that point in $\mathcal{B}_{d,\leqslant n}$ equals (54). □

We believe that the degree is also preserved. The next conjecture implies this.

CONJECTURE 7.6. *The birational map from $\mathcal{B}_{d,\leqslant n}$ to $\mathcal{N}_{d,\leqslant n}$ is defined everywhere on $\mathcal{B}_{d,\leqslant n}$. Equivalently, in the language of projective geometry, this birational map has no base points.*

REMARK 7.7 (Brownian motion in magnetic field and renormalization). We presented a fairly pedestrian derivation of the expected signature formula (54) for Brownian motion $X_t = \mu t + \sqrt{\Sigma} B_t$. A more sophisticated construction starts with the Brownian rough path

$$\mathbb{X}_t = \left(1, X_t, \int_0^t X \otimes \circ\, dX\right),$$

defined by Stratonovich stochastic integration. In a second step, one uses Lyons' extension theorem to get all signature tensors associated to $\mathbb{X}$. The third and last step is to take (component-wise) the expected value to arrive at (54). Let Q be a skew-symmetric $d \times d$ matrix and consider the 'perturbed' (translated) Brownian rough path $T_Q\mathbb{X} := \tilde{\mathbb{X}}$ given by

$$\tilde{\mathbb{X}}_t := \left(1, X_t, \int_0^t X \otimes \circ\, dX + tQ\right).$$

This construction is far from ad hoc: a systematic study of such perturbations is closely related to Hairer's theory of renormalized SPDEs [22]. Brownian rough paths of this form have arisen in a number of concrete situations with nonreversible noise, including the prominent example of Brownian motion in a magnetic field (cf. [18, Ch. 3]) and subsequent works on homogenization (by Melbourne and coworkers). On the level of expected signatures,

$$\mathbb{E}(\sigma(\tilde{\mathbb{X}})) = \exp\left(\mu + Q + \tfrac{1}{2}\Sigma\right) \tag{58}$$

follows from the Lévy–Khintchine formula in [19], by viewing $\tilde{\mathbb{X}}$ as Lévy rough paths with Lévy triplet $(\mu + Q, \Sigma, 0)$. In absence of a diffusive component, that is when $\Sigma = 0$, we are precisely in the $(m = 2)$ rough Veronese setting of Section 5.4, cf. Example 5.15.

The formula (58) generalizes (54). It replaces the symmetric matrix Σ with $\tilde{\Sigma} = \Sigma + 2Q$, where Q is skew-symmetric. The degree k part in the tensor polynomial $\exp\big(\mu + \frac{1}{2}\tilde{\Sigma}\big)$ is denoted $\mathbb{E}(\sigma^{(k)}(\tilde{X}))$ and is called the *expected magnetic Brownian signature tensor* of step k. Each entry $\sigma_{i_1 i_2 \cdots i_k}$ of this tensor is a homogeneous polynomial of degree k in the $d + d^2$ parameters. The d parameters from μ have degree 1, and the d^2 parameters, from Σ and Q, have degree 2. One can derive explicit formulas for $\mathbb{E}(\sigma^{(k)}(\mathbb{X}))$ in terms of μ and Σ and Q.

Finally, we define the *expected magnetic Brownian signature variety* to be the subvariety of $\mathbb{P}^{d^k-1}$ that is parametrized by the tensors $\sigma^{(k)}(X)$. This projective variety is denoted $\mathcal{B}^{\mathrm{mag}}_{d,k}$. Similarly, we can also define a projective variety $\mathcal{B}^{\mathrm{mag}}_{d,\leqslant n}$ that lives in the projective space associated with $T^n(\mathbb{R}^d)$. These varieties, whose detailed study is deferred to future work, are 'stochastic' generalizations of the rough Veronese, in the same way that Brownian and Gaussian varieties (as studied in [2, 3]) are natural generalizations of the classical Veronese.

7.2. Mixtures and secant varieties. The signature varieties of Brownian motion live above the moment varieties of Gaussians. There is a natural map from the latter onto the former, induced by tensor symmetrization $\mathbb{P}^{d^k-1} \dashrightarrow \mathbb{P}^{\binom{d+k-1}{k}}$. We can pass to secant varieties on either side of this linear projection.

Secant varieties are the algebraic representation of the statistical notion of *mixtures* of probability distributions. The moment varieties for mixtures of Gaussians were studied in detail in [3]. Mixtures of Brownian motion lift these varieties to the tensor space $\mathbb{P}^{d^k-1}$. In other words, the varieties that follow are non-Abelian extensions of those studied in [3].

A mixture of r Gaussians is parametrized by r pairs $(\mu_1, \Sigma_1), \ldots, (\mu_r, \Sigma_r)$ as above, along with nonnegative mixture probabilities $\alpha_1, \ldots, \alpha_r$ satisfying $\sum_{j=1}^r \alpha_j = 1$. Let $X^{(j)}$ denote Brownian motion with drift and covariance given by the jth pair (μ_j, Σ_j). We can then consider the mixed process $\bar{X}$ that is defined by selecting Brownian motion $X^{(j)}$ with probability α_j. As an immediate consequence from (54), the mixed expected signature is given by

$$\mathbb{E}\left(\sigma(\bar{X})\right) = \sum_{j=1}^{r} \alpha_j \cdot \exp\left(\mu^{(j)} + \frac{1}{2}\Sigma^{(j)}\right).$$

By focusing on the terms of degree k, or on all terms of degree $\leqslant n$, we obtain the signature varieties for Brownian mixtures, denoted $\mathcal{B}_{d,k,r}$ and $\mathcal{B}_{d,\leqslant n,r}$. These

live in projective spaces, and they map onto the moment varieties of Gaussian mixtures studied in [3].

As with our other signature varieties, we are interested in their equations, as well as in formulas for recovering parameters, that is identifiability as defined in the last section. We boldly conjecture that, unlike their commutative shadows in [3], the expected signature varieties $\mathcal{B}_{d,\leq n,r}$ are always algebraically identifiable. This was verified computationally in the first case where a secant of a Gaussian moment variety fails to be identifiable: $d = 3, n = 3, r = 2$. On the other hand, while algebraic identifiability works for generic parameters $(\mu_1, \Sigma_1), \ldots, (\mu_r, \Sigma_r)$, it may fail over specific submodels. We now illustrate this.

EXAMPLE 7.8 ($d = r = 2$). Consider mixtures of two Brownian paths in $\mathbb{R}^2$, each with identity covariance: $\Sigma_1 = \Sigma_2 = I$. This model has five parameters, namely the mixture parameter $\alpha \in (0, 1)$ and the entries of the two drift vectors: $\mu_1 = (a_1, a_2)$, $\mu_2 = (b_1, b_2)$. The mixed expected signature of this model is the bivariate tensor series

$$\mathbb{E}\big(\sigma(\bar{X})\big) = \alpha \cdot \exp\left(\mu_1 + \tfrac{1}{2}I\right) + (1-\alpha) \cdot \exp\left(\mu_2 + \tfrac{1}{2}I\right).$$

The first and second signature tensors have five distinct entries $\sigma_1, \sigma_2, \sigma_{11}, \sigma_{12} = \sigma_{21}$ and σ_{22}. We write their expressions in terms of the five model parameters in `Macaulay2` format:

```
R=QQ[a1,a2,b1,b2,p,s11,s12,s22,s1,s2,MonomialOrder=>Lex];
I = ideal(  p*a1 + (1-p)*b1 - s1, p*a2 + (1-p)*b2 - s2,
          (p/2)*(a1^2 +1) + ((1-p)/2)*(b1^2 +1) - s11,
          (p/2)*(a1*a2+0) + ((1-p)/2)*(b1*b2+0) - s12,
          (p/2)*(a2^2 +1) + ((1-p)/2)*(b2^2+1) - s22);
toString gens gb I
```

The command `gens gb I` computes a Gröbner basis. Prepending the command `toString` displays that Gröbner basis. This output consists of 14 polynomials, and it reveals that the model is *not* identifiable. The associated variety is precisely the Veronese surface in 5-space, after the linear change of coordinates coming from fixing identity covariances. The mixture model of two unit-covariance Brownian paths in $\mathbb{R}^2$ is cut out by the 3×3 determinant

$$\begin{pmatrix} 1 & s_1 & s_2 \\ s_1 & 2s_{11}-1 & 2s_{12} \\ s_2 & 2s_{12} & 2s_{22}-1 \end{pmatrix}.$$

This determinant is the first relation in s_{ij} seen in the Gröbner basis `gb I`. ◇

The last example generalizes to mixtures with common covariance matrix Σ.

THEOREM 7.9. *Consider mixtures of Brownian motion in $\mathbb{R}^d$ with $r > 1$ components having the same covariance Σ. Fix n such that $\binom{d+n}{n} \geqslant (d+1)r$. The r mean vectors are algebraically identifiable from the expected signature tensors up to order n, except in the following cases:*

$$n = 2 \quad \text{or } n = 3, d = 4, r = 7$$
$$\text{or} \quad n = 4, d = 2, r = 5 \quad \text{or } n = 4, d = 4, r = 14.$$

The same statement holds true for Brownian rough paths in the presence of a magnetic field, as in Remark 7.7, given by a fixed skew-symmetric matrix Q, and Σ replaced by $\tilde{\Sigma} = \Sigma + 2Q$.

Proof. This result is a slight generalization of [1, Theorem 6.2.3], where the analogue for homoscedastic Gaussian mixtures appears. The proof relies on the same fact: the subvariety $\mathcal{B}^{\tilde{\Sigma}}_{d,\leqslant n}$ of $\mathcal{B}_{d,\leqslant n}$ in question is isomorphic to the Veronese variety after a linear change of coordinates. Hence, the rth secant variety $\mathcal{B}^{\tilde{\Sigma}}_{d,\leqslant n,r}$ that corresponds to the mixture of processes is isomorphic to the rth secant of the Veronese variety. Indeed, consider the expansion of

$$\exp\left(\mu + \tfrac{1}{2}\tilde{\Sigma}\right),$$

where $\tilde{\Sigma}$ is a fixed matrix. While we cannot distribute the exponential due to the tensor noncommutativity, for each particular expected signature $\sigma_{i_1\cdots i_k}$, we can factor out the entries s_{ij} of the matrix $\tilde{\Sigma}$ that appear in each summand. This implies that $\sigma_{i_1\cdots i_k}$ equals $\mu_{i_1}\cdots\mu_{i_k}$ plus a linear combination of such μ-monomials of lower order with products of s_{ij} as coefficients.

It remains to see when the secant variety fails to have the expected dimension. The celebrated *Alexander–Hirschowitz Theorem* [6] from algebraic geometry provides a classification of all the cases where secant varieties of a Veronese variety have that special property. The listed values are the exceptions that are relevant for the above parameter conditions. □

Acknowledgements

We thank Joscha Diehl, Francesco Galuppi, and Anna Seigal for valuable comments on drafts of this paper. Fernando De Terán Vergara helped us with some references. Carlos Améndola was supported by the Einstein Foundation Berlin and the Deutsche Forschungsgemeinschaft (DFG) in the context of the

Emmy Noether junior research group KR 4512/1-1. Peter Friz was partially supported by the European Research Council through Consolidator Grant 683164, the Einstein Foundation Berlin and DFG research unit FOR2402. Bernd Sturmfels was supported by the US National Science Foundation (DMS-1419018).

References

[1] C. Améndola, 'Algebraic statistics of Gaussian mixtures', PhD Dissertation, TU Berlin, 2017.
[2] C. Améndola, J.-C. Faugère and B. Sturmfels, 'Moment varieties of Gaussian mixtures', *J. Algebr. Stat.* **7** (2016), 14–28.
[3] C. Améndola, K. Ranestad and B. Sturmfels, 'Algebraic identifiability of Gaussian mixtures', *Int. Math. Res. Not. IMRN* **21** (2018), 6556–6580.
[4] D. Bates, J. Hauenstein, A. Sommese and C. Wampler, *Numerically Solving Polynomial Systems with Bertini*, Software, Environments and Tools, 25 (SIAM, Philadelphia, 2013).
[5] H. Boedihardjo, X. Geng, T. Lyons and D. Yang, 'The signature of a rough path: uniqueness', *Adv. Math.* **293** (2016), 720–737.
[6] M. Brambilla and G. Ottaviani, 'On the Alexander–Hirschowitz theorem', *J. Pure Appl. Algebra* **212** (2008), 1229–1251.
[7] K.-T. Chen, 'Iterated integrals and exponential homomorphisms', *Proc. Lond. Math. Soc. (3)* **4** (1954), 502–512.
[8] K.-T. Chen, 'Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula', *Ann. of Math.* **65** (1957), 163–178.
[9] K.-T. Chen, 'Integration of paths – a faithful representation of paths by noncommutative formal power series', *Trans. Amer. Math. Soc.* **89** (1958), 395–407.
[10] I. Chevyrev and A. Kormilitzin, 'A primer on the signature method in machine learning'. Preprint, 2016, arXiv:1603.03788.
[11] I. Chevyrev and T. Lyons, 'Characteristic functions of measures on geometric rough paths', *Ann. Probab.* **44** (2016), 4049–4082.
[12] W.-L. Chow, 'Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung', *Math. Ann.* **117** (1940/1941), 98–105.
[13] D. Cox, J. Little and D. O'Shea, *Ideals, Varieties, and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra*, 2nd edn, Undergraduate Texts in Mathematics (Springer, New York, 1997).
[14] L. De Lathauwer, B. De Moor and J. Vandewalle, 'A multilinear singular value decomposition', *SIAM J. Matrix Anal. Appl.* **21** (2000), 1253–1278.
[15] J. Diehl and J. Reizenstein, "Invariants of multidimensional time series based on their iterated-integral signature', *J. Acta. Appl. Math.* (2018), https://doi.org/10.1007/s10440-018-00227-z.
[16] M. Drton, B. Sturmfels and S. Sullivant, *Lectures on Algebraic Statistics*, Oberwolfach Seminars, 39 (Birkhäuser, Basel, 2009).
[17] D. Eisenbud, *Commutative Algebra. With a View Toward Algebraic Geometry*, Graduate Texts in Mathematics, 150 (Springer, New York, 1995).
[18] P. Friz and M. Hairer, *A Course on Rough Paths*, Universitext (Springer, Cham, 2014), With an introduction to regularity structures.
[19] P. Friz and A. Shekhar, 'General rough integration, Lévy rough paths and a Lévy–Kintchine type formula', *Ann. Probab.* **45** (2017), 2707–2765.
[20] P. Friz and N. Victoir, *Multidimensional Stochastic Processes as Rough Paths*, Theory and Applications (Cambridge University Press, Cambridge, UK, 2010).

[21] D. Grayson and M. Stillman, 'Macaulay2, a software system for research in algebraic geometry', available at www.math.uiuc.edu/Macaulay2/.
[22] M. Hairer, 'A theory of regularity structures', *Invent. Math.* **198** (2014), 269–504.
[23] M. Hairer and D. Kelly, 'Geometric versus non-geometric rough paths', *Ann. Inst. Henri Poincaré Probab. Stat.* **51** (2015), 207–251.
[24] B. Hambly and T. Lyons, 'Uniqueness for the signature of a path of bounded variation and the reduced path group', *Ann. of Math.* **171** (2010), 109–167.
[25] T. Kolda and B. Bader, 'Tensor decompositions and applications', *SIAM Rev.* **51** (2009), 455–500.
[26] P. Lalonde and A. Ram, 'Standard Lyndon bases of Lie algebras and enveloping algebras', *Trans. Amer. Math. Soc.* **347** (1993), 1821–1831.
[27] P. Lancaster and L. Rodman, 'Canonical forms for symmetric/skew-symmetric real matrix pairs under strict equivalence and congruence', *Linear Algebra Appl.* **406** (2005), 1–76.
[28] J. M. Landsberg, *Tensors: Geometry and Applications*, Graduate Studies in Mathematics, 128 (American Mathematical Society, Providence, RI, 2012).
[29] T. Lyons, 'Differential equations driven by rough signals', *Rev. Mat. Iberoam.* **14** (1998), 215–310.
[30] T. Lyons, 'Rough paths, signatures and the modelling of functions on streams', in *Proc. International Congress of Mathematicians* (Kyung Moon Publishers, Seoul, 2014), 163–184.
[31] T. Lyons, M. Caruana and T. Lévy, *Differential Equations Driven by Rough Paths*, Lecture Notes in Mathematics, 1908 (Springer, Berlin, 2007).
[32] T. Lyons and Z. Qian, *System Control and Rough Paths* (Oxford University Press, Oxford, UK, 2002).
[33] T. Lyons and N. Sidorova, 'Sound compression – a rough path approach', in *Proceedings 4th Intern. Symp. Information and Communication Technologies* (Cape Town, 2005), 223–229.
[34] T. Lyons and W. Xu, 'Hyperbolic development and inversion of signature', *J. Funct. Anal.* **272** (2017), 2933–2955.
[35] T. Lyons and W. Xu, 'Inverting the signature of a path', *J. Eur. Math. Soc. (JEMS)* **20** (2018), 1655–1687.
[36] G. Melançon and C. Reutenauer, 'Lyndon words, free algebras and shuffles', *Canadian J. Math.* **16** (1989), 577–591.
[37] J. G. Oxley, *Matroid Theory* (Oxford University Press, PSA, 2006).
[38] D. E. Radford, 'A natural ring basis for the shuffle algebra and an application to group schemes', *J. Algebra* **58** (1979), 432–453.
[39] C. Reutenauer, *Free Lie Algebras*, London Mathematical Society Monographs, New Series, 7 (Oxford University Press, New York, 1993).
[40] R. C. Thompson, 'Pencils of complex and real symmetric and skew matrices', *Linear Algebra Appl.* **147** (1991), 323–371.