

Random Points on an Algebraic Manifold*

Paul Breiding[†] and Orlando Marigliano[‡]

Abstract. Consider the set of solutions to a system of polynomial equations in many variables. An algebraic manifold is an open submanifold of such a set. We introduce a new method for computing integrals and sampling from distributions on algebraic manifolds. This method is based on intersecting with random linear spaces. It produces independent and identically distributed samples, works in the presence of multiple connected components, and is simple to implement. We present applications to computational statistical physics and topological data analysis.

Key words. sampling, approximating integrals, geometrical probability, algebraic geometry, statistical physics, topological data analysis

AMS subject classifications. 60B05, 62D05, 51H30

DOI. 10.1137/19M1271178

1. Introduction. In statistics and applied mathematics, *manifolds* are useful models for continuous data. For example, in *computational statistical physics* the state space of a collection of particles is a manifold. Each point on this manifold records the positions of all particles in space. The field of *information geometry* interprets a statistical model as a manifold, each point corresponding to a probability distribution inside the model. *Topological data analysis* studies the geometric properties of a point cloud in some Euclidean space. Learning the manifold that best explains the position of the points is a research topic in this field.

Regardless of the context, there are two fundamental computational problems associated with a manifold \mathcal{M} .

- (1) Approximate the Lebesgue integral $\int_{\mathcal{M}} f(x)dx$ of a given function f on \mathcal{M} .
- (2) Sample from a probability distribution with a given density on \mathcal{M} .

These problems are closely related in theory. In applications however, they may occur separately. For instance, in section 2 we present an example from computational physics that involves only (1). On the other hand, the subsequent example from topological data analysis is about (2).

In general these problems are easy for manifolds which admit a differentiable surjection $\mathbb{R}^k \rightarrow \mathcal{M}$, also called *parametrized* manifolds. They are harder for *nonparametrized* manifolds, which are usually represented as the sets of nonsingular solutions to some system of

*Received by the editors June 28, 2019; accepted for publication (in revised form) May 29, 2020; published electronically August 24, 2020.

<https://doi.org/10.1137/19M1271178>

Funding: The first author has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 787840).

[†]Technische Universität Berlin, Berlin, Germany (breiding@math.tu-berlin.de).

[‡]Max-Planck Institute for Mathematics in the Sciences, Leipzig, Germany (orlando.marigliano@mis.mpg.de).

differentiable *implicit* equations

$$(1.1) \quad F_i(x) = 0 \quad (i = 1, \dots, r).$$

The standard techniques to solve (2) in the nonparametrized case involve moving randomly from one sample point on \mathcal{M} to the next nearby, and fall under the umbrella term of *Markov chain Monte Carlo (MCMC)* [12, 14, 15, 22, 26, 31, 33].

In this paper, we present a new method that solves (1) and (2) when the functions F_i are *polynomial*. In simple terms, the method can be described as follows. First, we choose a random linear subspace of complementary dimension and calculate its intersection with \mathcal{M} . Since the implicit equations are polynomial, the intersection can be efficiently determined using numerical polynomial equation solvers [4, 10, 17, 34, 40]. The number of intersection points is finite and bounded by the *degree* of the system (1.1). Next, if we want to solve (1) we evaluate a modified function \bar{f} at each intersection point, sum its values, and repeat the process to approximate the desired integral. Else if we want to solve (2), after a rejection step, we pick one of the intersection points at random to be our sample point. We then repeat the process to obtain more samples of the desired density.

Compared to MCMC sampling, our method has two main advantages. First, we have the option to generate points that are independent of each other. Second, the method is global in the sense that it also works when the manifold has multiple distinct connected components, and does not require picking a starting point $x_0 \in \mathcal{M}$.

The main theoretical result supporting the method is Theorem 1.1. It is in line with a series of classical results commonly known as *Crofton's formulas* or *kinematic formulas* [38] that relate the volume of a manifold to the expected number of its intersection points with random linear spaces. Previous uses of such formulas in applications can be found in [8, 28, 29, 35]. Traditionally, one starts by sampling from the set of linear spaces intersecting a ball containing \mathcal{M} . Our contribution is to suggest an alternative sampling method for linear spaces $\{x \in \mathbb{R}^N | Ax = b\}$, namely, by sampling (A, b) from the Gaussian distribution on $\mathbb{R}^{n \times N} \times \mathbb{R}^n$. We argue in section 7 that our method is more exact and converges at least as quickly as the above.

Until this point, we have assumed that \mathcal{M} is the set of nonsingular solutions to an implicit system of polynomial equations (1.1). In fact, our main theorem also holds for open submanifolds of such a set of solutions. We call them *algebraic manifolds*. Throughout this article we fix an n -dimensional algebraic manifold $\mathcal{M} \subset \mathbb{R}^N$.

To state our result, we fix a measurable function $f: \mathcal{M} \rightarrow \mathbb{R}_{\geq 0}$ with finite integral over \mathcal{M} . We define the auxiliary function $\bar{f}: \mathbb{R}^{n \times N} \times \mathbb{R}^n \rightarrow \mathbb{R}$ as follows:

$$\bar{f}(A, b) := \sum_{x \in \mathcal{M}: Ax=b} \frac{f(x)}{\alpha(x)}, \quad \text{where} \quad \alpha(x) := \frac{\sqrt{1 + \langle x, \Pi_{N_x \mathcal{M}} x \rangle}}{(1 + \|x\|^2)^{\frac{n+1}{2}}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}^{n+1}}$$

and where $\Pi_{N_x \mathcal{M}}: \mathbb{R}^N \rightarrow N_x \mathcal{M}$ denotes the orthogonal projection onto the normal space of \mathcal{M} at $x \in \mathcal{M}$. Note that this projection can be computed from the implicit equations for \mathcal{M} , because $N_x \mathcal{M}$ is the row-span of the Jacobian matrix $J(x) = [\frac{\partial F_i}{\partial x_j}(x)]_{1 \leq i \leq r, 1 \leq j \leq N}$. Therefore, if $Q \in \mathbb{R}^{N \times r}$ is the Q-factor from the QR-decomposition of $J(x)^T$, then we have $N_x \mathcal{M} = QQ^T$. This means that we can easily compute $\bar{f}(A, b)$ from $\mathcal{M} \cap \mathcal{L}_{A,b}$.

The operator $f \mapsto \bar{f}$ allow us to state the following main result, making our new method precise.

Theorem 1.1. *Let $\varphi(A, b)$ be the probability density for which the entries of $(A, b) \in \mathbb{R}^{n \times N} \times \mathbb{R}^n$ are independent and identically distributed (i.i.d.) standard Gaussian. In the notation introduced above*

(1) *the integral of f over \mathcal{M} is the expected value of \bar{f} :*

$$\int_{\mathcal{M}} f(x) \, dx = \mathbb{E}_{(A,b) \sim \varphi} \bar{f}(A, b);$$

(2) *assume that $f : \mathcal{M} \rightarrow \mathbb{R}$ is nonnegative and that $\int_{\mathcal{M}} f(x) \, dx$ is positive and finite. Let $X \in \mathcal{M}$ be the random variable obtained by choosing a pair $(A, b) \in \mathbb{R}^{n \times N} \times \mathbb{R}^n$ with probability*

$$\psi(A, b) := \frac{\varphi(A, b) \bar{f}(A, b)}{\mathbb{E}_{\varphi}(\bar{f})}$$

and choosing one of the finitely many points X of the intersection $\mathcal{M} \cap \mathcal{L}_{A,b}$ with probability $f(x)\alpha(x)^{-1}\bar{f}(A, b)^{-1}$. Then X is distributed according to the scaled density $f(x)/(\int_{\mathcal{M}} f(x) \, dx)$ associated with $f(x)$.

Using the formula for $\alpha(x)$ we can already evaluate $\bar{f}(A, b)$ for an integrable function f . Thus we can approximate the integral of f by computing the empirical mean

$$E(f, k) = \frac{1}{k}(\bar{f}_1(A, b) + \cdots + \bar{f}_k(A, b))$$

of a sample drawn from $(A, b) \sim \varphi$. The next lemma yields a bound for the rate of convergence of this approach. It is an application of Chebyshev's inequality and proved in section 5.

Lemma 1.2. *Assume that $|f(x)|$ and $\|x\|$ are bounded on \mathcal{M} . Then, the variance $\sigma^2(\bar{f})$ of $\bar{f}(A, b)$ is finite and for $\varepsilon > 0$ we have $\text{Prob}\{|E(f, k) - \int_{\mathcal{M}} f(x) \, dx| \geq \varepsilon\} \leq \frac{\sigma^2(\bar{f})}{\varepsilon^2 k}$.*

In (5.1) we provide a deterministic bound for $\sigma^2(\bar{f})$, which involves the degree of the ambient variety of \mathcal{M} and upper bounds for $\|x\|$ and $|f(x)|$ on \mathcal{M} . In our experiments we also use the empirical variance $s^2(\bar{f})$ of a sample for estimating $\sigma^2(\bar{f})$.

For sampling $(A, b) \sim \psi$ in the second part of Theorem 1.1 we could use MCMC sampling. Note that this would employ MCMC sampling for the flat space $\mathbb{R}^{n \times N} \times \mathbb{R}^n$, which is easier than MCMC for nonlinear spaces like \mathcal{M} . Nevertheless, in this paper we use the simplest method for sampling ψ , namely, *rejection sampling*. This is used in the experiment section and explained in section 3.3.

1.1. Outline. In section 2 of this paper, we show applications of this method to examples in topological data analysis and statistical physics. We discuss preliminaries for the proof of the main theorem in section 3 and give the full proof in section 4. In section 5 we prove Lemma 1.2. We prove a variant of our theorems for projective algebraic manifolds in section 6. In section 7 we review other methods that make use of a kinematic formula for sampling. Finally, in section 8 we briefly discuss the limitations of our method and possible future work.

1.2. Notation. The Euclidean inner product on \mathbb{R}^N is $\langle x, y \rangle := x^T y$ and the associated norm is $\|x\| := \sqrt{\langle x, x \rangle}$. The unit sphere in \mathbb{R}^N is $\mathbb{S}^{N-1} := \{x \in \mathbb{R}^N : \|x\| = 1\}$. For a

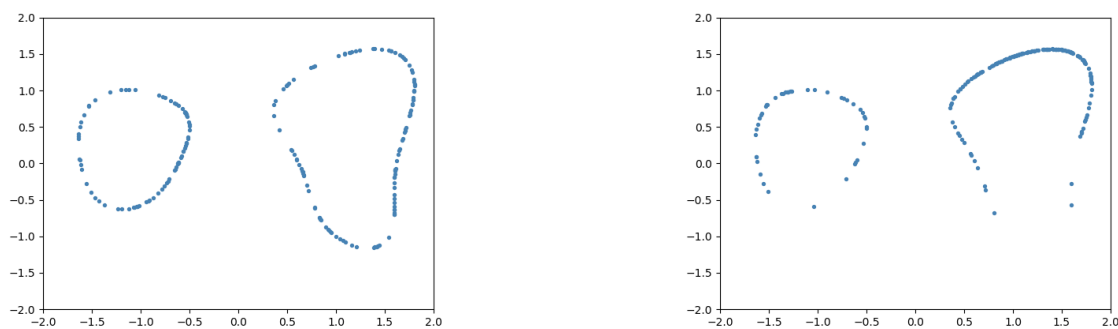


Figure 2.1. Left picture: a sample of 200 points from the uniform distribution on the curve (2.1). Right picture: a sample of 200 points from the uniform distribution scaled by e^{2y} .

function $f : \mathcal{M} \rightarrow \mathcal{N}$ between manifolds we denote by $D_x f$ the derivative of f at $x \in \mathcal{M}$. The tangent space of \mathcal{M} at x is denoted $T_x \mathcal{M}$ and the normal space is $N_x \mathcal{M}$.

2. Experiments. In this section we apply our main results to examples. All experiments have been performed on macOS 10.14.2 on a computer with Intel Core i5 2.3 GHz (two cores) and 8 GB RAM memory. For computing the intersections with linear spaces, we use the numerical polynomial equation solver `HomotopyContinuation.jl` [10]. For plotting we use `Matplotlib` [25]. For sampling from the distribution $\psi(A, b)$ we use rejection sampling as described in section 3.3.

As a first simple example, we consider the plane curve \mathcal{M} given by the equation

$$(2.1) \quad x^4 + y^4 - 3x^2 - xy^2 - y + 1 = 0.$$

We have $\text{vol}(\mathcal{M}) = \mathbb{E}_\varphi(\bar{1})$. We can therefore estimate the volume of the curve \mathcal{M} by taking a sample of i.i.d. pairs (A, b) and computing the empirical mean $E(1, k)$ of $\bar{1}$ of the sample. A sample of $k = 10^5$ yields $E(1, k) = 11.2$. In Lemma 1.2 we take $\varepsilon = 0.1$ and the variance of the sample s^2 , and get an upper bound of $\frac{s^2}{\varepsilon^2 k} = 0.008$. Therefore, we expect that 11.2 is a good approximation of the true volume. We can also take the deterministic upper bound from (5.1) for the variance σ^2 of $\bar{1}$. Here, we take $\sup_{x \in \mathcal{M}} \|x\| = \sqrt{8}$. To get an estimate with accuracy at least $\varepsilon = 0.1$ with probability at least 0.9 we need a sample of size $k \geq \frac{\sigma^2}{\varepsilon^2 \cdot 0.9} \geq 1421300$. Taking such a sample size we get an estimated volume of ≈ 11.217 . The code that produced this result is available at [2].

Next, we use the second part of Theorem 1.1 to generate random samples on \mathcal{M} ; code is available at [3]. We show in the left picture of Figure 2.1 a sample of 200 points drawn uniformly from the curve. The right picture shows 200 points drawn from the density given by the normalization of $f(x, y) = e^{2y}$. As can be seen from the pictures the points drawn from the second distribution concentrate in the upper half of \mathcal{M} , whereas points from the first distribution spread equally around the curve. This experiment also shows how our method generates global samples. The curve has more than one connected component, which is not an obstacle for our method.

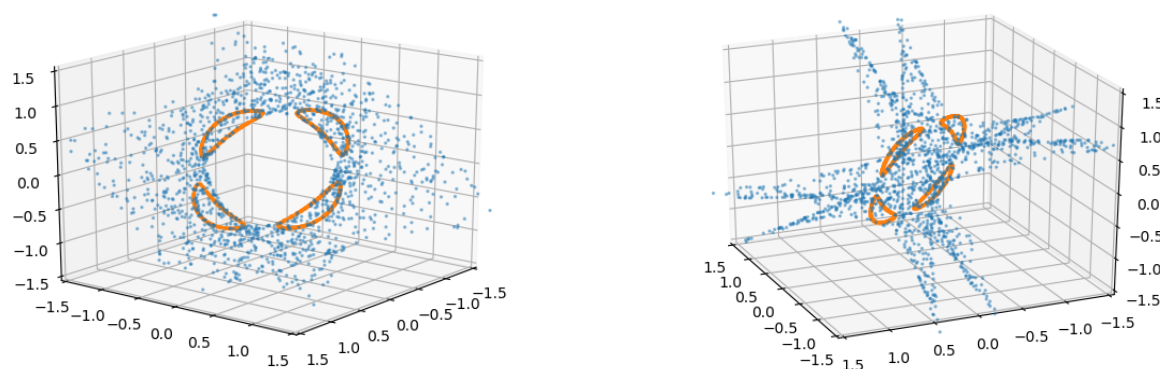


Figure 2.2. The blue points are a sample of 1569 points from the complex Trott curve (2.2) seen as a variety in \mathbb{R}^4 projected to \mathbb{R}^3 . The orange points are a sample of 1259 points from the real part of the Trott curve.

Our method is particularly appealing for hypersurfaces like (2.1) because intersecting a hypersurface with a linear space of dimension 1 reduces to solving a single univariate polynomial equation. This can be done very efficiently, for instance, using the algorithm from [39], and so for hypersurfaces we can easily generate large sample sets.

The pictures suggest we use sampling for visualization. For instance, we can visualize a semialgebraic piece of the complex and real part of the Trott curve T , defined by the equation

$$(2.2) \quad 144(x_1^4 + x_2^4) - 225(x_1^2 + x_2^2) + 350x_1^2x_2^2 + 81 = 0.$$

The associated complex variety in \mathbb{C}^2 can be seen as a real variety $T_{\mathbb{C}}$ in \mathbb{R}^4 . We sample from the real Trott curve T and the complex Trott curve $T_{\mathbb{C}}$ intersected with the box $-1.5 < \text{Real}(x_1), \text{Imag}(x_1), \text{Real}(x_2), \text{Imag}(x_2) < 1.5$. Then, we take a random projection $\mathbb{R}^4 \rightarrow \mathbb{R}^3$ to obtain a sample in \mathbb{R}^3 (the projected sample is *not* uniform on the projected semialgebraic variety). The outcome of this experiment is shown in Figure 2.2.

2.1. Application to statistical physics. In this section we want to apply Theorem 1.1 to study a physical system of N particles $q = (q_1, \dots, q_N) \in \mathcal{M}$, where $\mathcal{M} \subseteq (\mathbb{R}^3)^N$ is the manifold that models the spacial constraints of the q_i . In our example we have $N = 6$ and the q_i are the spacial positions of carbon atoms in a *cyclohexane molecule*. The constraints of this molecule are the following algebraic equations:

$$(2.3) \quad \mathcal{M} = \{q = (q_1, \dots, q_6) \in (\mathbb{R}^3)^6 \mid \|q_1 - q_2\|^2 = \dots = \|q_5 - q_6\|^2 = \|q_6 - q_1\|^2 = c^2\},$$

where c is the *bond length* between two neighboring atoms (the vectors $q_i - q_{i+1}$ are called *bonds*). In our example we take $c^2 = 1$ (unitless). Due to rotational and translational invariance of the equations we define q_1 to be the origin, $q_6 = (c, 0, 0)$, and q_5 to be rotated, such that its last entry is equal to zero. We thus have 11 variables.

Lelièvre, Rousset, and Stoltz [32] write “In the framework of statistical physics, macroscopic quantities of interest are written as averages over [...] probability measures on all the admissible microscopic configurations.” As the probability measure we take the canonical ensemble [32]. If $E(q)$ denotes the total energy of a configuration q , the density in the canonical

ensemble is proportional to $f(q) = e^{-E(q)}$. That is, a configuration is most likely to appear when its energy is minimal. We model the energy of a molecule using an interaction potential, namely, the *Lennard-Jones potential* $V(r) = \frac{1}{4}(\frac{c}{r})^{12} - \frac{1}{2}(\frac{c}{r})^6$; see, e.g., [32, equation (1.5)]. Then, the energy function of a system is

$$E(q) = \sum_{1 \leq i < j \leq N} V(\|q_i - q_j\|).$$

In this example we consider as the quantity the average angle between neighboring bonds $q_{i-1} - q_i$ and $q_{i+1} - q_i$:

$$\theta(q) = \frac{\angle(q_6 - q_1, q_2 - q_1) + \cdots + \angle(q_5 - q_6, q_1 - q_6)}{6},$$

where $\angle(b_1, b_2) := \arccos \frac{\langle b_1, b_2 \rangle}{\|b_1\| \|b_2\|}$. We compute the macroscopic state of $\theta(q)$ by determining its distribution $\text{Prob}\{\theta(q) = \theta_0\} = \frac{1}{H} \int_{\theta(q)=\theta_0} f(q) dq$, where $H = \int_V f(q) dq$ is the normalizing constant. For comparing the probabilities of different values for θ it suffices to compute

$$\rho(\theta_0) = \int_{\theta(q)=\theta_0} f(q) dq.$$

We approximate this integral as $\rho(\theta_0) \approx \frac{\mu_1(\theta_0)}{\mu_2(\theta_0)}$, where

$$\mu_1(\theta_0) = \int_{\substack{\theta(q) > \theta_0 - \Delta\theta \\ \theta(q) < \theta_0 + \Delta\theta}} f(q) dq \quad \text{and} \quad \mu_2(\theta_0) = \int_{\substack{\theta(q) > \theta_0 - \Delta\theta \\ \theta(q) < \theta_0 + \Delta\theta}} 1 dq$$

for some $\Delta\theta > 0$ (in our experiment we take $\Delta\theta = 3^\circ$), and we approximate both $\mu_1(\theta_0)$ and $\mu_2(\theta_0)$ for several values of θ by their empirical means $E(f, k)$ and $E(1, k)$, and using [Theorem 1.1](#). We take $k = 10^4$ samples, respectively. The code for this is available at [1].

Figure 2.3 shows both the values of the empirical means in the logarithmic scale and the ratio of $\mu_1(\theta_0)$ and $\mu_2(\theta_0)$.

The left picture shows the approximations of $\mu_1(\theta_0)$ and $\mu_2(\theta_0)$ by the empirical means $E(f, k)$ and $E(1, k)$. Both integrals were approximated independently, each by an empirical mean obtained from 10^4 intersections with linear spaces. The right picture shows the ratio of the empirical means, which approximate $\rho(\theta_0)$.

How good is our estimate? From the plot above we can deduce that $\varepsilon = 10^2$ is a good accuracy for both $\mu_1(\theta_0)$ and $\mu_2(\theta_0)$. Using the variances s_1^2 and s_2^2 of the samples, respectively, we get $\frac{s_1^2}{\varepsilon^2 k} \leq 0.002$ and $\frac{s_2^2}{\varepsilon^2 k} \leq 0.002$. Hence, by [Lemma 1.2](#) we expect that the probability that the empirical mean $E(f, k)$ deviates from $\mu_1(\theta_0)$ by more than ε and that the probability that $E(1, k)$ deviates from $\mu_2(\theta_0)$ by more than ε are both at most 0.2%. We conclude that our approximation of $\rho(\theta) = H \text{Prob}\{\theta(q) = \theta_0\}$ is a good approximation.

In fact, [Figure 2.3](#) shows a peak at around $\theta = 114^\circ$. It is known that the total energy of the cyclohexane system is minimized when all angles between consecutive bonds achieve 110.9° ; see [11, Chapter 2]. Therefore, our experiment truly gives a good approximation of the molecular geometry of cyclohexane. An example where all the angles between consecutive bonds are 110.9° is shown in [Figure 2.4](#).

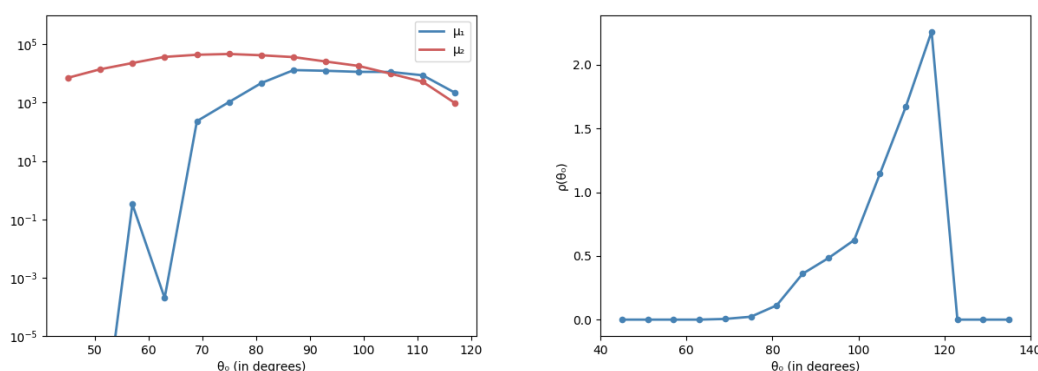


Figure 2.3. The left picture shows the approximations of $\mu_1(\theta_0)$ and $\mu_2(\theta_0)$ by the empirical means $E(f, k)$ and $E(1, k)$. Both integrals were approximated independently, each by an empirical mean obtained from 10^4 intersections with linear spaces. The right picture shows the ratio of the empirical means, which approximate $\rho(\theta_0)$.

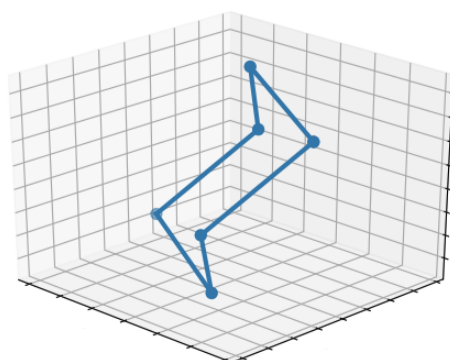


Figure 2.4. The picture shows a point from the variety (2.3), for which the angles between two consecutive bonds are all equal to 110.9° degrees. This configuration is also known as the “chair” [9].

2.2. Application to topological data analysis. We believe that [Theorem 1.1](#) will be useful for researchers working in topological data analysis using persistent homology (PH). PH is a tool to estimate the homology groups of a topological space from a finite point sample. The underlying idea is as follows: for varying t , put a ball of radius t around each point and compute the homology of the union of those balls. One then looks at topological features that persist for large intervals in t . It is intuitively clear that the point sample should be large enough to capture all of the topological information of its underlying space and, on the other hand, the sample should be small enough to remain feasible for computations. Dufresne et al. [18] comment “Both the theoretical framework for the PH pipeline and its computational costs drive the requirements of a suitable sampling algorithm.” (For an explanation of the PH pipeline, see [18, section 2] and the references therein.) They develop an algorithm that takes as input a denseness parameter ϵ and outputs a sample where each point has at most distance ϵ to its nearest neighbor. At the same time, their method is trying to keep the sample size as small as possible. In the context of topological data analysis we see our algorithm as an alternative to [18].

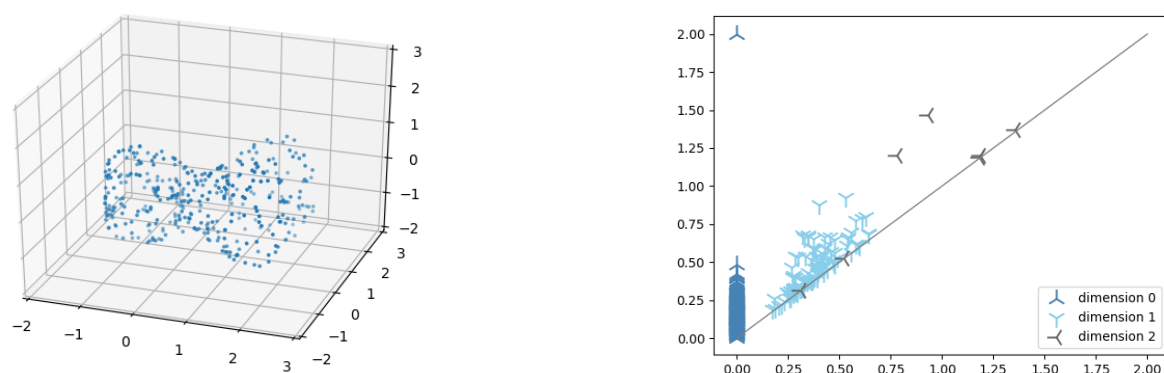


Figure 2.5. The left picture shows a sample of 386 points from the variety (2.4). The right picture shows the corresponding persistence diagram.

In the following we use [Theorem 1.1](#) for generating samples as input for the PH pipeline from [18]. The output of this pipeline is a *persistence diagram*. It shows the appearance and the vanishing of topological features in a 2-dimensional plot. Each point in the plot corresponds to an i -dimensional “hole,” where the x -coordinate represents the time t when the hole appears, and the y -coordinate is the time when it vanishes. Points that are far from the line $x = y$ should be interpreted as signals coming from the underlying space. The number of those points is used as an estimator for the Betti number β_i . For computing persistence diagrams we use [Ripser](#) [5].

First, we consider two toy examples from [18, section 5]: the surface \mathcal{S}_1 is given by

$$(2.4) \quad 4x_1^4 + 7x_2^4 + 3x_3^4 - 3 - 8x_1^3 + 2x_1^2x_2 - 4x_1^2 - 8x_1x_2^2 - 5x_1x_2 + 8x_1 - 6x_2^3 + 8x_2^2 + 4x_2 = 0.$$

[Figure 2.5](#) shows a sample of 386 points from the uniform distribution on \mathcal{S}_1 . The associated persistence diagram suggest one connected component, two 1-dimensional and two 2-dimensional holes. The latter two come from two sphere-like features of the variety. The outcome is similar to the diagram from [18, Figure 6]. Considering that the diagram in this reference was computed using 1500 points [20], we think that the quality of our diagram is good.

The second example is the surface \mathcal{S}_2 given by

$$144(x_1^4 + x_2^4) - 225(x_1^2 + x_2^2)x_3^2 + 350x_1^2x_2^2 + 81x_3^4 + x_1^3 + 7x_1^2x_2 + 3(x_1^2 + x_1x_2^2) - 4x_1 - 5(x_2^3 - x_2^2 - x_2) = 0.$$

[Figure 2.6](#) shows a sample of 651 points from the uniform distribution on \mathcal{S}_2 . The persistence diagram on the right suggest one or five connected components. The true answer is five connected components. The diagram from [18, Figure 6] captures the correct homology more clearly, but was generated from a sample of 10000 points [20].

The next example is from a specific application in kinematics. We quote [18, section 5.3] “Consider a regular pentagon in the plane consisting of links with unit length, and with one of the links fixed to lie along the x -axis with leftmost point at $(0, 0)$. The set of all possible configurations of this regular pentagon is a real algebraic variety.” The equations of the

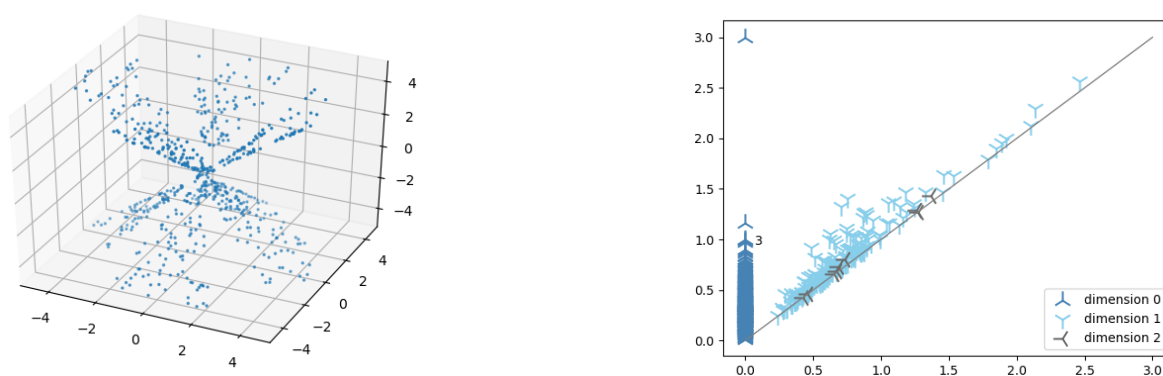


Figure 2.6. The left picture shows a sample of 651 points from the variety \mathcal{S}_2 . The right picture shows the corresponding persistence diagram.

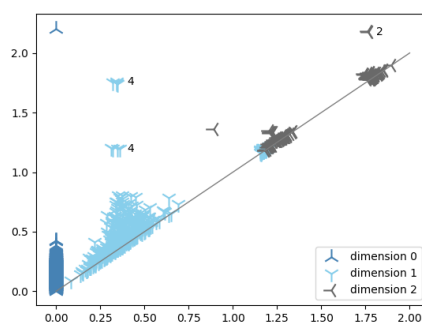


Figure 2.7. The picture shows the persistence diagram of a sample of 1400 points from the variety given by (2.5).

configuration space are

$$(2.5) \quad (x_1 + x_2 + x_3)^2 + (1 + x_4 + x_5 + x_6)^2 - 1 = 0, \quad x_1^2 + x_4^2 = 1, \quad x_2^2 + x_5^2 = 1, \quad x_3^2 + x_6^2 = 1.$$

Here, the zeroth homology is of particular importance because, if the variety is connected, “the mechanism has one assembly mode which can be continuously deformed to all possible configurations” [18]. Figure 2.7 shows the persistence diagram of a sample of 1400 points from the configuration space. It suggests that the variety indeed has only one connected component. We moreover observe eight holes of dimension 1 and one or three 2-dimensional holes. The correct Betti numbers are $\beta_0 = 1, \beta_1 = 8, \beta_2 = 1$; see [21].

3. Preliminaries. In this section we first define the degree of a real algebraic variety and explain why the number of intersection points of \mathcal{M} with a linear space of the right codimension does not exceed the degree of its ambient variety. Then, we recall the coarea formula of integration and discuss some consequences. Finally, we explain how to sample from $\psi(A, b)$ using rejection sampling and we prove an algorithm for sampling $\mathcal{L}_{A,b}$ in implicit form.

3.1. Real algebraic varieties. For the purpose of this paper, a (real, affine) *algebraic variety* is a subset \mathcal{V} of \mathbb{R}^N such that there exists a set of polynomials F_1, \dots, F_k in N variables such that \mathcal{V} is their set of common zeros. All varieties have a *dimension* and a *degree*. The dimension of \mathcal{V} is defined as the dimension of its subspace of nonsingular points \mathcal{V}_0 , which is a manifold. For the degree we give a definition in the following steps.

An algebraic variety \mathcal{V} in \mathbb{R}^N is *homogeneous* if for all $t \in \mathbb{R} \setminus \{0\}$ and $x \in \mathcal{V}$ we have $tx \in \mathcal{V}$. Homogeneous varieties are precisely the ones where we can choose the F_i above to be homogeneous polynomials. Naturally, homogeneous varieties live in the $(N-1)$ -dimensional *real projective space* \mathbb{P}^{N-1} . This space is defined as the set $(\mathbb{R}^N \setminus \{0\})/\sim$, where $x \sim y$ if x and y are collinear. It comes with a canonical projection map $p: (\mathbb{R}^N \setminus \{0\}) \rightarrow \mathbb{P}^{N-1}$. Then, a *projective variety* is defined as the image of a homogeneous variety \mathcal{V} under p . Its dimension is $\dim \mathcal{V} - 1$.

Similarly, we define *complex affine*, homogeneous, and projective varieties by replacing \mathbb{R} with \mathbb{C} in the previous definitions. We can pass from real to complex varieties as follows. Let $\mathcal{V} \subset \mathbb{R}^N$ be a real affine variety. Its *complexification* $\mathcal{V}_{\mathbb{C}}$ is defined as the complex affine variety

$$\mathcal{V}_{\mathbb{C}} := \{x \in \mathbb{C}^n : f(x) = 0 \text{ for all real polynomials } f \text{ vanishing on } \mathcal{V}\}.$$

The “all” is crucial here. Consider, for instance, the variety in \mathbb{R}^2 defined by $x_1^2 + x_2^2 = 0$. Obviously, this variety is a single point $\{(0,0)\}$, but $\{x \in \mathbb{C}^2 : x_1^2 + x_2^2 = 0\} = \{(t, \sqrt{-1}t) : t \in \mathbb{C}\}$ is one dimensional. Nevertheless, the polynomials $x_1 = 0, x_2 = 0$ also vanish on $\{(0,0)\}$ and so the complexification of $\mathcal{V} = \{(0,0)\}$ is $\mathcal{V}_{\mathbb{C}} = \{(0,0)\}$. The following lemma is important.

Lemma 3.1 (see [41, Lemma 8]). *The real dimension of \mathcal{V} and the complex dimension of its complexification $\mathcal{V}_{\mathbb{C}}$ agree.*

The *Grassmannian* is a smooth algebraic variety $G(k, \mathbb{C}^N)$ that parametrizes linear subspaces of \mathbb{C}^N of dimension k . Furthermore, k -dimensional affine-linear subspaces of \mathbb{C}^N can be seen as $(k+1)$ -dimensional linear subspaces of \mathbb{C}^{N+1} and are parametrized by the *affine Grassmannian* $G_{\text{Aff}}(k, \mathbb{C}^N)$. A *projective linear space* of dimension k is the image of a linear space $\mathcal{L} \in G(k+1, \mathbb{C}^N)$ under the projection p . This motivates us to define the projective Grassmannian as $G(k, \mathbb{P}^{N-1}) := \{p(\mathcal{L}) : \mathcal{L} \in G(k+1, \mathbb{C}^N)\}$.

Now, we have gathered all the material to give a precise definition of the degree: let $\mathcal{V} \subset \mathbb{P}_{\mathbb{C}}^{N-1}$ be a complex projective variety of dimension n . There exists a unique natural number d and a lower-dimensional subvariety \mathcal{W} of $G(N-n, \mathbb{P}^{N-1})$ with the property that for all linear spaces $\mathcal{L} \in G(N-n, \mathbb{P}^{N-1}) \setminus \mathcal{W}$ the intersection $\mathcal{V} \cap \mathcal{L}$ consists of d distinct points [23, section 18]. Furthermore, the number of such intersection points only decreases when $\mathcal{L} \in \mathcal{W}$. This number d is called the degree of the projective variety \mathcal{V} . The degree of a complex affine variety $\mathcal{V} \subset \mathbb{C}^N$ is defined as the degree of the smallest projective variety containing the image of \mathcal{V} under the embedding $\mathbb{C}^N \hookrightarrow \mathbb{P}_{\mathbb{C}}^N$ sending x to $p([1, x])$.

The definition of degree of complex varieties is standard in algebraic geometry. In this article, however, we are solely dealing with real varieties. We therefore make the following definition, which is not standard in the literature, but which fits in our setting.

Definition 3.2. *The degree of a real affine variety \mathcal{V} is the degree of its complexification. The degree of a real projective variety \mathcal{V} is the degree of the image of the complexification of $p^{-1}(\mathcal{V})$ under p .*

Using Lemma 3.1 we make the following conclusions, after passing from the Grassmannians $G_{\text{Aff}}(N - n, \mathbb{R}^N)$ and $G(N - n, \mathbb{R}^N)$ to the parameter spaces $\mathbb{R}^{n \times N} \times \mathbb{R}^n$ and $\mathbb{R}^{n \times N}$.

Lemma 3.3. *Let $\mathcal{V} \subset \mathbb{R}^N$ be an affine variety of dimension n and degree d . Except for a lower-dimensional subset of $\mathbb{R}^{n \times N} \times \mathbb{R}^n$, all affine linear $\mathcal{L}_{A,b} = \{x \in \mathbb{R}^N : Ax = b\} \subset \mathbb{R}^N$ intersect \mathcal{V} in at most d many points.*

Let $\mathcal{V} \subset \mathbb{P}^{N-1}$ be a projective variety of dimension n and degree d . Except for a lower-dimensional subset of $\mathbb{R}^{n \times N}$, all linear subspaces $\mathcal{L}_A = \{x \in \mathbb{P}^{N-1} : Ax = 0\} \subset \mathbb{P}^{N-1}$ intersect \mathcal{V} in at most d many points.

3.2. The coarea formula. The coarea formula of integration is a key ingredient in the proof of the main theorems. This formula says how integrals transform under smooth maps. A well-known special case is integration by substitution. The coarea formula generalizes this from integrals defined on the real line to integrals defined on differentiable manifolds.

Let \mathcal{M}, \mathcal{N} be Riemannian manifolds and dv, dw be the respective volume forms. Furthermore, let $h : \mathcal{M} \rightarrow \mathcal{N}$ be a smooth map. A point $v \in \mathcal{M}$ is called a *regular point* of h if $D_v h$ is surjective. Note that a necessary condition for regular points to exist is $\dim \mathcal{M} \geq \dim \mathcal{N}$.

For any $v \in \mathcal{M}$ the Riemannian metric on \mathcal{M} defines orthogonality on $T_v \mathcal{M}$. For a regular point v of h this implies that the restriction of $D_v h$ to the orthogonal complement of its kernel is a linear isomorphism. The absolute value of the determinant of that isomorphism is the *normal Jacobian* of h at v . Let us summarize this in a definition.

Definition 3.4. *Let $h : \mathcal{M} \rightarrow \mathcal{N}$ be a smooth map and $v \in \mathcal{M}$ be a regular point of h . Let $(\cdot)^\perp$ denote the orthogonal complement. The normal Jacobian of h at v is defined as*

$$\text{NJ}(h, v) := \left| \det \left(D_v h|_{(\ker D_v h)^\perp} \right) \right|.$$

We also need the following theorem (see, e.g., [13, Theorem A.9]).

Theorem 3.5. *Let \mathcal{M}, \mathcal{N} be smooth manifolds with $\dim \mathcal{M} \geq \dim \mathcal{N}$ and let $h : \mathcal{M} \rightarrow \mathcal{N}$ be a smooth map. Let $w \in \mathcal{N}$ be such that all $v \in h^{-1}(w)$ are regular points of h . Then, the fiber $h^{-1}(w)$ over w is a smooth submanifold of \mathcal{M} of dimension $\dim \mathcal{M} - \dim \mathcal{N}$ and the tangent space of $h^{-1}(w)$ at v is $T_v h^{-1}(w) = \ker D_v h$.*

A point $w \in \mathcal{N}$ satisfying the properties in the previous theorem is called a *regular value* of h . By Sard's lemma the set of all $w \in \mathcal{N}$ that are not a regular value of h is a set of measure zero. We are now equipped with all we need to state the coarea formula. See [24, (A-2)] for a proof.

Theorem 3.6 (the coarea formula of integration). *Suppose that \mathcal{M}, \mathcal{N} are Riemannian manifolds, and let $h : \mathcal{M} \rightarrow \mathcal{N}$ be a surjective smooth map. Then we have for any function $a : \mathcal{M} \rightarrow \mathbb{R}$ that is integrable with respect to the volume measure of \mathcal{M} that*

$$\int_{v \in \mathcal{M}} a(v) dv = \int_{w \in \mathcal{N}} \left(\int_{u \in h^{-1}(w)} \frac{a(u)}{\text{NJ}(h, u)} du \right) dw,$$

where du is the volume form on the submanifold $h^{-1}(w)$.

The following corollary from the coarea formula is important.

Corollary 3.7. Let $h: \mathcal{M} \rightarrow \mathcal{N}$ be a smooth surjective map of Riemannian manifolds.

(1) Let X be a random variable on \mathcal{M} with density β . Then $h(X)$ is a random variable on \mathcal{N} with density

$$\gamma(y) = \int_{x \in h^{-1}(y)} \frac{\beta(x)}{\text{NJ}(h, x)} dx.$$

(2) Let ψ be a density on \mathcal{N} and for all $y \in \mathcal{N}$, let ρ_y be a density on $h^{-1}(y)$. The random variable X on \mathcal{M} obtained by independently taking $Y \in \mathcal{N}$ with density ψ and $X \in h^{-1}(Y)$ with density ρ_Y has density

$$\beta(x) = \psi(h(x))\rho_{h(x)}(x)\text{NJ}(h, x).$$

Proof. The first part follows directly from the coarea formula. For the second part, it suffices to note that for measurable $\mathcal{U} \subset \mathcal{M}$ we have

$$\begin{aligned} \int_{y \in h(\mathcal{U})} \int_{x \in h^{-1}(y) \cap \mathcal{U}} \psi(h(x))\rho_{h(x)}(x) dx dy &= \int_{y \in h(\mathcal{U})} \int_{x \in h^{-1}(y) \cap \mathcal{U}} \frac{\beta(x)}{\text{NJ}(h, x)} dx dy \\ &= \int_{x \in \mathcal{U}} \beta(x) dx; \end{aligned}$$

see also [13, Remark 17.11]. ■

3.3. Sampling from the density on affine-linear subspaces. Our method for sampling from an algebraic manifold involves taking a distribution φ on the parameter space $\mathbb{R}^{n \times N} \times \mathbb{R}^n$ of hyperplanes of the right dimension which is easy to sample, and turning it into another density ψ . In this section, we explain how to sample from ψ with rejection sampling. In the following we denote elements of $\mathbb{R}^{n \times N} \times \mathbb{R}^n$ by (A, b) .

Proposition 3.8. Let κ be any number satisfying $0 < \kappa \cdot \sup_{(A,b)} \bar{f}(A, b) \leq 1$. Consider the binary random $Z \in \{0, 1\}$ with $\text{Prob}\{Z = 1 \mid (A, b)\} = \kappa \bar{f}(A, b)$. Then, ψ is the density of the conditional random variable $((A, b) \mid Z = 1)$.

Proof. We denote the density of the conditional random variable $((A, b) \mid Z = 1)$ by λ . Bayes' theorem implies $\lambda(A, b) \text{Prob}\{Z = 1\} = \text{Prob}\{Z = 1 \mid (A, b)\} \varphi(A, b)$, which, by assumption, is equivalent to

$$\lambda(A, b) = \frac{\kappa \bar{f}(A, b) \varphi(A, b)}{\text{Prob}\{Z = 1\}}.$$

By the definition of Z we have $\text{Prob}\{Z = 1\} = \kappa \mathbb{E}_{(A,b) \sim \varphi} \bar{f}(A, b)$. Hence,

$$\lambda(A, b) = \frac{\bar{f}(A, b) \varphi(A, b)}{\mathbb{E}_{(A,b) \sim \varphi} \bar{f}(A, b)} = \psi(A, b).$$

This finishes the proof. ■

Proposition 3.8 shows that ψ is the density of a conditional distribution. A way to sample from such distributions is by rejection sampling: for sampling $((A, b) \mid Z = 1)$ we may sample from the joint distribution (A, b, Z) and then keep only the points with $Z = 1$. The strong

law of large numbers implies the correctness of rejection sampling. Indeed, if (A_i, b_i, Z_i) is a sequence of i.i.d. copies of (A, b, Z) and \mathcal{U} is a measurable set with respect to the Lebesgue measure on $\mathbb{R}^{n \times N} \times \mathbb{R}^n$, then we have

$$\begin{aligned} \frac{\#\{i \mid (A_i, b_i) \in \mathcal{U}, Z_i = z, i \leq n\}}{\#\{i \mid Z_i = z, i \leq n\}} &= \frac{\frac{1}{n} \#\{i \mid (A_i, b_i) \in \mathcal{U}, Z_i = z, i \leq n\}}{\frac{1}{n} \#\{i \mid Z_i = z, i \leq n\}} \\ &\xrightarrow{\text{a.s.}} \frac{\text{Prob}\{(A, b) \in \mathcal{U}, Z = z\}}{\text{Prob}\{Z = z\}} \\ &= \text{Prob}_{(A, b) \mid Z=z}(\mathcal{U}). \end{aligned}$$

For sampling Z , however, we must compute a suitable κ . This can be done as follows. Let d be the degree of the ambient variety of \mathcal{M} . We assume we know upper bounds K for $f(x)$ and C for $\|x\|^2$, both as x ranges over \mathcal{M} , and set

$$\kappa = \frac{1}{dK} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}^{n+1}} \frac{1}{(1+C)^{\frac{n+1}{2}}}.$$

Then we have $0 < \kappa \bar{f}(A, b) \leq 1$ for all (A, b) as needed. With κ , we have everything we need to carry out the sampling method.

How to obtain the upper bounds K and C ? For K , we might just know the maximum of f . For example, if we want to sample from the uniform distribution, then we may use $f = 1$. In more complicated cases, we could approximate $\max f$ by repeatedly sampling $(A, b) \sim \varphi$ and recording the highest value f takes on the points in the intersection $\mathcal{M} \cap \mathcal{L}_{(A, b)}$. Casella and Robert [16] call this approach *stochastic exploration*.

We might know C a priori, for example, because we restrict the manifold \mathcal{M} to a box in \mathbb{R}^N . We can also restrict the manifold to a box after determining by sampling what the size of the box should be. We could also estimate $\max \|x\|^2$ by sampling as for $\max f$. Sometimes we can also use *semidefinite programming* [6] to bound polynomial functions like $\|x\|^2$ on a variety. Note that the probability for rejection increases as C increases. We thus seek to find a C which is as small as possible. If our given function f is invariant under translation, we may translate \mathcal{M} to decrease C . For instance, sampling from the uniform distribution on the circle $(x_1 - 100)^2 + (x_2 - 100)^2 = 1$ is the same as sampling on $x_1^2 + x_2^2 = 1$ and then translating by adding $(100, 100)$ to each sample point. The difference between the two is that for the first variety we need $C = 101$, whereas for the second we can use $C = 1$.

3.4. Sampling linear spaces in explicit form. Sometimes it is useful to sample the linear space $\mathcal{L}_{A, b}$ in explicit form, and not in implicit form $Ax = b$. For instance, if \mathcal{V} is a hypersurface given by an equation $F(x) = 0$, then intersecting \mathcal{V} with a line $u + tv$ can be done by solving the univariate equation $F(u + tv) = 0$. The next lemma shows how to pass from implicit to explicit representation in the Gaussian case.

Lemma 3.9. *Let $\mathcal{L}_{A, b} = \{x \in \mathbb{R}^N \mid Ax = b\}$ be a random affine linear space given by i.i.d. standard Gaussian entries for $A \in \mathbb{R}^{n \times N}$ and $b \in \mathbb{R}^n$. Consider another random linear space*

$$\mathcal{K}_{u, v_1, \dots, v_{N-n}} = \{u + t_1 v_1 + \dots + t_{N-n} v_{N-n} \mid t_1, \dots, t_{N-n} \in \mathbb{R}\},$$

where u, v_1, \dots, v_{N-m} are obtained as follows. Sample a matrix $U \in \mathbb{R}^{(N-n+1) \times (N+1)}$ with i.i.d. standard Gaussian entries, and let

$$\begin{pmatrix} u \\ 1 \end{pmatrix}, \begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} v_{N-n} \\ 0 \end{pmatrix} \in \text{rowspan}(U).$$

Then, we have $\mathcal{K}_{u,v_1,\dots,v_{N-m}} \sim \mathcal{L}_{A,b}$.

Proof. Consider the linear space $\tilde{\mathcal{L}}_{A,b} := \{z \in \mathbb{R}^{N+1} \mid [A, -b]z = 0\}$. This is a random linear space in the Grassmannian $G(N+1-n, \mathbb{R}^{N+1})$. The affine linear space is given as $\mathcal{L}_{A,b} := \{u + t_1 v_1 + \dots + t_{N-n} v_{N-n}\}$, where

$$\begin{pmatrix} u \\ 1 \end{pmatrix}, \begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} v_{N-n} \\ 0 \end{pmatrix} \in \ker([A, -b]).$$

Now the kernel of $[A, -b]$ is a random linear space in $G(n, \mathbb{R}^{N+1})$, which is invariant under orthogonal transformations. By [30] there is a unique orthogonally invariant probability distribution on the Grassmannian $G(n, \mathbb{R}^{N+1})$. Since $\text{rowspan}(U)$ is also orthogonally invariant, we find that $\text{rowspan}(U) \sim \ker([A, -b])$, which concludes the proof. ■

4. Proof of Theorem 1.1. We begin by giving an alternate description of the function $\alpha(x)$.

Lemma 4.1. $\alpha(x) = \int_{A \in \mathbb{R}^{n \times N}} \varphi(A, Ax) |\det(A|_{T_x \mathcal{M}})| \, dA$.

Proof. Let $\alpha'(x)$ be the right-hand side of the formula. Let $U \in O(N)$ be an orthogonal matrix such that $Ux = (0, \dots, 0, x_N)^T$ and consider the manifold $\mathcal{N} = U \cdot \mathcal{M}$. We have $T_{Ux} \mathcal{N} = UT_x \mathcal{M}$ and $\det(A|_{T_x \mathcal{M}}) = \det(AU^T|_{T_{Ux} \mathcal{N}})$. After the change of variables $A \mapsto AU^T$ we get

$$\alpha'(x) = \int_{A \in \mathbb{R}^{n \times N}} |\det(A|_{T_{Ux} \mathcal{N}})| \varphi(A, AUx) \, dA.$$

By definition of the Gaussian density, we have $\varphi(A, AUx) = \frac{1}{(\sqrt{2\pi})^n} \phi(AR)$, where ϕ is the Gaussian density on $\mathbb{R}^{n \times N}$ and $R = \text{diag}(1, \dots, 1, \sqrt{1+x_N^2}) \in \mathbb{R}^{N \times N}$. Let us write $B = AR$. A change of variables from A to B yields

$$\alpha'(x) = \frac{1}{(1+x_N^2)^{\frac{n}{2}} (\sqrt{2\pi})^n} \int_{B \in \mathbb{R}^{n \times N}} |\det(BR^{-1}|_{T_{Ux} \mathcal{N}})| \phi(B) \, dB.$$

Let $W \in \mathbb{R}^{N \times n}$ be a matrix whose columns form an orthonormal basis for $T_{Ux} \mathcal{N}$ and write $M := R^{-1}W$. Then we have $\det(BR^{-1}|_{T_{Ux} \mathcal{N}}) = \det(BM)$ and so

$$\alpha'(x) = \frac{1}{(1+x_N^2)^{\frac{n}{2}} \sqrt{2\pi}^n} \mathbb{E}_{B \sim \phi} |\det(BM)|.$$

We now write $\mathbb{E}_{B \sim \phi} |\det(BM)| = \mathbb{E}_{B \sim \phi} \det(M^T B^T B M)^{\frac{1}{2}}$. By [37, Theorem 3.2.5] the matrix $C := M^T B^T B M \in \mathbb{R}^{n \times n}$ is a *Wishart matrix* with covariance matrix $M^T M$. By [37, Theorem

3.2.15], we have $\mathbb{E} \det(C)^{\frac{1}{2}} = \det(M^T M)^{\frac{1}{2}} \frac{1}{\sqrt{\pi}} \sqrt{2}^n \Gamma(\frac{n+1}{2})$. Altogether, this shows that

$$\alpha'(x) = \frac{\det(M^T M)^{\frac{1}{2}}}{(1+x_N^2)^{\frac{n}{2}}} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}^{n+1}}.$$

Moreover, we have

$$\begin{aligned} M^T M &= W^T R^{-T} R^{-1} W \\ &= W^T \operatorname{diag}\left(1, \dots, 1, \frac{1}{1+x_N^2}\right) W \\ &= \mathbf{1} - \frac{1}{1+x_N^2} W \operatorname{diag}(0, \dots, 0, x_N^2) W. \end{aligned}$$

The second summand in the last expression is a rank-one matrix with the single nonzero eigenvalue $-\frac{\|W^T Ux\|^2}{1+\|Ux\|^2}$. Taking determinants we get

$$\det(M^T M) = 1 - \frac{\|W^T Ux\|^2}{1+\|Ux\|^2} = \frac{1+\|Ux\|^2 - \|W^T Ux\|^2}{1+\|Ux\|^2} = \frac{1+\|x\|^2 - \|\Pi_{T_x \mathcal{M}} x\|^2}{1+\|x\|^2},$$

where $\Pi_{T_x \mathcal{M}}$ denotes the orthogonal projection onto the tangent space. Since we have $\|x\|^2 - \|\Pi_{T_x \mathcal{M}} x\|^2 = \langle x, \Pi_{N_x \mathcal{M}} x \rangle$, this implies

$$\alpha'(x) = \frac{\sqrt{1+\langle x, \Pi_{N_x \mathcal{M}} x \rangle}}{(1+x_N^2)^{\frac{n+1}{2}}} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}^{n+1}} = \alpha(x).$$

This concludes the proof. ■

We are now prepared to prove Theorem 1.1.

Proof of Theorem 1.1. We first prove the first part. The support of $\bar{f}(A, b)$ is a full dimensional subset of $\mathbb{R}^{n \times N} \times \mathbb{R}^n$ and it is contained in the complement of the set of all (A, b) for which $\mathcal{M} \cap \mathcal{L}_{A, b} = \emptyset$. We let \mathcal{X} denote the interior of the support of $\bar{f}(A, b)$, so that $\mathbb{E}_{(A, b) \sim \varphi} \bar{f}(A, b) = \int_{\mathcal{X}} \bar{f}(A, b) \varphi(A, b) d(A, b)$.

Let $\pi: \mathbb{R}^{n \times N} \times \mathcal{M} \rightarrow \mathcal{X}$ be the map sending a pair (A, x) to (A, Ax) . We have $D_{(A, x)} \pi(\dot{A}, \dot{x}) = (\dot{A}, \dot{A}x + A\dot{x})$, so the derivative of π can be identified with the matrix $\begin{pmatrix} \mathbf{1} & 0 \\ * & A \end{pmatrix}$. This shows that $\operatorname{NJ}(\pi, (A, x)) = |\det(A|_{T_x \mathcal{M}})|$. Therefore, by Theorem 3.6,

$$\mathbb{E}_{(A, b) \sim \varphi} \bar{f}(A, b) = \int_{\mathbb{R}^{n \times N} \times \mathcal{M}} \frac{f(x)}{\alpha(x)} |\det(A|_{T_x \mathcal{M}})| \varphi(A, Ax) d(A, x).$$

The projection $\mathbb{R}^{n \times N} \times \mathcal{M} \rightarrow \mathcal{M}$ on the second factor has normal Jacobian one everywhere. Applying Theorem 3.6 again yields

$$\mathbb{E}_{(A, b) \sim \varphi} \bar{f}(A, b) = \int_{\mathcal{M}} \frac{f(x)}{\alpha(x)} \left(\int_{\mathbb{R}^{n \times N}} |\det(A|_{T_x \mathcal{M}})| \varphi(A, Ax) dA \right) dx = \int_{\mathcal{M}} f(x) dx,$$

the second inequality by Lemma 4.1. This proves the first part.

Now, we prove the second part, where we assume that $f: \mathcal{M} \rightarrow \mathbb{R}_{>0}$ is nonnegative. Recall that $\psi(A, b) = \frac{\varphi(A, b)\bar{f}(A, b)}{\mathbb{E}_\varphi(\bar{f})}$. Since $\mathbb{E}_\varphi(\bar{f}) = \int_{\mathcal{M}} f(x) dx$ is positive and finite by the first part of the theorem, we find that ψ is a well-defined probability density. The support of ψ is contained in the closure of \mathcal{X} and therefore $\mathcal{M} \cap \mathcal{L}_{A, b}$ is almost surely nonempty and finite.

Let $Y = (A, x) \in \mathbb{R}^{n \times N} \times \mathcal{M}$ be the random variable defined by first choosing $(A, b) \sim \psi$ and then taking $x \in \mathcal{M} \cap \mathcal{L}_{A, b}$ with probability $f(x)\alpha(x)^{-1}\bar{f}(A, b)^{-1}$. By construction, $\pi(Y) \sim \psi$. We use [Corollary 3.7\(2\)](#) and find that Y has density

$$\beta(A, x) = \frac{\psi(A, Ax)f(x)\text{NJ}(\pi, (A, x))}{\alpha(x)\bar{f}(A, Ax)}.$$

Recall that $\text{NJ}(\pi, (A, x)) = |\det(A|_{T_x\mathcal{M}})|$ and that the projection $\mathbb{R}^{n \times N} \times \mathcal{M} \rightarrow \mathcal{M}$ on the second factor has normal Jacobian one everywhere. Therefore, by [Corollary 3.7\(1\)](#), the random point $x \in \mathcal{M}$ has density γ with

$$\begin{aligned} \gamma(x) &= \int_{A \in \mathbb{R}^{n \times N}} \beta(A, x) dA \\ &= \frac{f(x)}{\alpha(x)} \int_{A \in \mathbb{R}^{n \times N}} \frac{\psi(A, Ax)|\det(A|_{T_x\mathcal{M}})|}{\bar{f}(A, Ax)} dA \\ &= \frac{f(x)}{\alpha(x)\mathbb{E}_\varphi(\bar{f})} \int_{A \in \mathbb{R}^{n \times N}} \varphi(A, Ax)|\det(A|_{T_x\mathcal{M}})| dA. \end{aligned}$$

Using [Lemma 4.1](#) yields $\gamma(x) = \frac{f(x)}{\mathbb{E}_\varphi(\bar{f})}$. This finishes the proof of [Theorem 1.1](#). ■

5. Proof of [Lemma 1.2](#). First, we can bound $\alpha(x) \geq \frac{1}{1 + \sup_{x \in \mathcal{M}} \|x\|^2} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}^{n+1}}$. Let d be the degree of the ambient variety of \mathcal{M} . With probability one $\mathcal{M} \cap \mathcal{L}_{A, b}$ consists of at most d points and so we have

$$\begin{aligned} \mathbb{E}_{(A, b) \sim \varphi} \bar{f}(A, b)^2 &= \mathbb{E}_{(A, b) \sim \varphi} \left(\sum_{x \in \mathcal{M} \cap \mathcal{L}_{A, b}} \frac{f(x)}{\alpha(x)} \right)^2 \\ &\leq \mathbb{E}_{(A, b) \sim \varphi} \left(\sum_{x \in \mathcal{M} \cap \mathcal{L}_{A, b}} \frac{|f(x)|}{\alpha(x)} \right)^2 \\ &\leq \mathbb{E}_{(A, b) \sim \varphi} \left(\sum_{x \in \mathcal{M} \cap \mathcal{L}_{A, b}} \frac{\sup_{x \in \mathcal{M}} |f(x)|}{\inf_{x \in \mathcal{M}} \alpha(x)} \right)^2 \\ &\leq d^2 \left(1 + \sup_{x \in \mathcal{M}} \|x\|^2 \right)^{n+1} \frac{\pi^{n+1}}{\Gamma(\frac{n+1}{2})^2} \sup_{x \in \mathcal{M}} f(x)^2. \end{aligned}$$

We also have $\sigma^2(\bar{f}) \leq \mathbb{E}_{(A, b) \sim \varphi} \bar{f}(A, b)^2$ and, therefore,

$$(5.1) \quad \sigma^2(\bar{f}) \leq d^2 \left(1 + \sup_{x \in \mathcal{M}} \|x\|^2 \right)^{n+1} \frac{\pi^{n+1}}{\Gamma(\frac{n+1}{2})^2} \sup_{x \in \mathcal{M}} f(x)^2$$

is finite. We may therefore use Chebyshev's inequality to deduce that

$$(5.2) \quad \text{Prob} \left\{ \left| \mathbb{E}(f, k) - \int_{\mathcal{M}} f(x) dx \right| \geq \varepsilon \right\} \leq \frac{\sigma^2(\bar{f})}{\varepsilon^2 k}.$$

This finishes the proof.

6. Sampling from projective manifolds. In this section we prove variations of Theorem 1.1 for projective algebraic manifolds.

Real projective space \mathbb{P}^{N-1} from section 3 is a compact Riemannian manifold with a canonical metric, the *Fubini–Study* metric. Namely, let $p: \mathbb{R}^N \setminus \{0\} \rightarrow \mathbb{P}^{N-1}$ be the canonical projection. Restricted to the unit sphere \mathbb{S}^{N-1} , the projection p identifies antipodal points. We define a subset $\mathcal{U} \subset \mathbb{P}^{N-1}$ to be open if and only if $p|_{\mathbb{S}^{N-1}}^{-1}(\mathcal{U})$ is open. This gives \mathbb{P}^{N-1} the structure of a differential manifold. The Riemannian structure on \mathbb{P}^{N-1} is defined as $\langle \dot{a}, \dot{b} \rangle := \langle D_x p^{-1} \dot{a}, D_x p^{-1} \dot{b} \rangle$ for $\dot{a}, \dot{b} \in T_x \mathbb{P}^{N-1}$. This metric is called the *Fubini–Study* metric, and it induces the *standard measure* on \mathbb{P}^{N-1} .

We say that \mathcal{M} is a *projective algebraic manifold* if it is an open submanifold of the smooth part of a real projective variety $\mathcal{V} \subset \mathbb{P}^{N-1}$. We assume \mathcal{M} to be n -dimensional, and consider a function $f: \mathcal{M} \rightarrow \mathbb{R}_{\geq 0}$ with a well-defined scaled probability density $f(x)/\int_{\mathcal{M}} f(x) dx$. For $A \in \mathbb{R}^{n \times N}$, define the linear space $\mathcal{L}_A = \{x \in \mathbb{P}^{N-1} \mid Ax = 0\}$ and write

$$\bar{f}(A) := \sum_{x \in \mathcal{M} \cap \mathcal{L}_A} f(x).$$

In this section, we denote by φ_ℓ the density of the multivariate standard normal distribution on \mathbb{R}^ℓ .

Theorem 6.1. *In the notation introduced above,*

(1) *let f be an integrable function on \mathcal{M} . We have*

$$\int_{\mathcal{M}} f(x) d(x) = \text{vol}(\mathbb{P}^n) \mathbb{E}_{A \sim \varphi_{n \times N}} \bar{f}(A);$$

(2) *let f be nonnegative and assume that the integral $\int_{\mathcal{M}} f(x) d(x)$ is finite and nonzero. Let $X \in \mathcal{M}$ be the random variable obtained by choosing $A \in \mathbb{R}^{n \times N}$ with probability $\psi(A) := \frac{\varphi(A) \bar{f}(A)}{\mathbb{E}_{A \sim \varphi_{n \times N}} \bar{f}(A)}$ and one of the finitely many points $X \in \mathcal{M} \cap \mathcal{L}_A$ with probability $f(x)/\bar{f}(A)$. Then X is distributed according to the density $f(x)/\int_{\mathcal{M}} f(x) dx$.*

Remark 6.2. In [27, section 2.4] Lairez proved a similar theorem for the uniform distribution on complex projective varieties.

Sampling \mathcal{L}_A with $A \sim \varphi_{n \times N}$ yields a special distribution on the Grassmannian $G(N - n - 1, \mathbb{P}^{N-1})$. By [30] there is unique orthogonally invariant probability measure ν on $G(N - n - 1, \mathbb{P}^{N-1})$. Since the distribution of the kernel of a Gaussian A is invariant under orthogonal transformations, the projective plane $\mathcal{L}_A = \{x \in \mathbb{P}^{N-1} : Ax = 0\}$ has distribution ν .

Furthermore, setting $f = 1$ in Theorem 6.1 gives the formula

$$\text{vol}(\mathcal{M}) = \text{vol}(\mathbb{P}^n) \mathbb{E}_{A \sim \varphi_{n \times N}} |\mathcal{M} \cap \mathcal{L}_A|.$$

This is the *kinematic formula for projective manifolds* from [24, Theorem 3.8] in disguise.

Before we can prove [Theorem 6.1](#), we have to prove an auxiliary lemma, similar to [Lemma 4.1](#).

Lemma 6.3. *For any $x \in \mathcal{M}$ we have*

$$\int_{A \in \mathbb{R}^{n \times N} : Ax=0} |\det(A|_{T_x \mathcal{M}})| \varphi_{n \times N}(A) \, dA = \frac{1}{\text{vol}(\mathbb{P}^n)}.$$

In particular, the integral is independent of x .

Proof. Let $\mathcal{H}(x) := \{A \in \mathbb{R}^{n \times N} \mid Ax = 0\}$. It is a linear subspace of $\mathbb{R}^{n \times N}$ of codimension n . Let $U \in \mathbb{R}^{N \times n}$ be a matrix whose columns form an orthonormal basis for $T_x \mathcal{M}$, so that $\det(A|_{T_x \mathcal{M}}) = \det(AU)$. Furthermore, let $O \in \mathbb{R}^{N \times N}$ be an orthogonal matrix with $Ox = e_1$, where $e_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^N$. Then, $\mathcal{H}(e_1)O = \mathcal{H}(x)$. Making a change of variables $A \mapsto AO$ we get

$$(6.1) \quad \int_{A \in \mathcal{H}(x)} |\det(A|_{T_x \mathcal{M}})| \varphi_{n \times N}(A) \, dA = \int_{A \in \mathcal{H}(e_1)} |\det(AOU)| \varphi_{n \times N}(AO) \, dA.$$

We have $\varphi_{n \times N}(AO) = \varphi_{n \times N}(A)$, because the Gaussian distribution is orthogonally invariant. Moreover, any $A \in \mathcal{H}(e_1)$ is of the form $A = [0, A']$ with $A' \in \mathbb{R}^{n \times (N-1)}$, and we have $\varphi_{n \times N}(A) = \frac{1}{\sqrt{2\pi}^n} \varphi_{n \times (N-1)}(A')$. Let us denote by O' the lower $(N-1) \times n$ part of OU , so that $AOU = A'O'$. It follows that [\(6.1\)](#) is equal to

$$\frac{1}{\sqrt{2\pi}^n} \int_{A' \in \mathbb{R}^{n \times (N-1)}} |\det(A'O')| \varphi_{n \times (N-1)}(A') \, dA'.$$

We show that O' has orthonormal columns: since $\mathcal{M} \subset \mathbb{S}^{N-1}$, the tangent space $T_x \mathcal{M}$ is orthogonal to x , which implies $U^T x = 0$. Furthermore, $e_1^T OU = (U^T O^T e_1)^T = (U^T x)^T$. It follows that the first row of OU contains only zeros and so the columns of O' must be pairwise orthogonal and of norm one. A standard Gaussian matrix multiplied with a matrix with orthonormal columns is also standard Gaussian, so we have

$$\int_{A' \in \mathbb{R}^{n \times (N-1)}} |\det(A'O')| \varphi_{n \times (N-1)}(A') \, dA' = \int_{M \in \mathbb{R}^{n \times n}} |\det(M)| \varphi_{n \times n}(M) \, dM.$$

This implies

$$\int_{A \in \mathcal{H}(x)} \varphi_{n \times N}(A) |\det(A|_{T_x \mathcal{M}})| \, dA = \frac{\mathbb{E}_{M \sim \varphi_{n \times n}} |\det(M)|}{\sqrt{2\pi}^n}.$$

Finally, we compute $\text{vol}(\mathbb{P}^n) = \frac{1}{2} \text{vol}(\mathbb{S}^n) = \frac{\sqrt{\pi}^{n+1}}{\Gamma(\frac{n+1}{2})}$, and by [\[37, Theorem 3.2.15\]](#), we have $\mathbb{E}_{M \sim \varphi_{n \times n}} \det(M^T M)^{\frac{1}{2}} = \frac{1}{\sqrt{\pi}} \sqrt{2}^n \Gamma(\frac{n+1}{2})$. This finishes the proof. ■

Proof of Theorem 6.1. We define $\mathcal{I} := \{(A, x) \in \mathbb{R}^{n \times N} \times \mathcal{M} \mid Ax = 0\}$. It is an algebraic subvariety of $\mathbb{R}^{n \times N} \times \mathcal{M}$. One can show that \mathcal{I} is smooth, but for our purposes it suffices to integrate over the dense subset of \mathcal{I} that is obtained by removing potential singularities from \mathcal{I} .

Let π_1 and π_2 be the projections from \mathcal{I} to $\mathbb{R}^{n \times N}$ and \mathcal{M} , respectively. Applying [Theorem 3.6](#) first to π_1 and then to π_2 yields

$$\mathbb{E}_{A \sim \varphi_{n \times N}} \bar{f}(A) = \int_{\mathcal{M}} f(x) \left(\int_{A \in \pi_1(\pi_2^{-1}(x))} \varphi_{n \times N}(A) \frac{\text{NJ}(\pi_2, (A, x))}{\text{NJ}(\pi_1, (A, x))} dA \right) dx.$$

By [\[7, section 13.2, Lemma 3\]](#), the ratio of normal Jacobians in the integrand equals $|\det(A|_{T_x \mathcal{M}})|$. We get

$$\begin{aligned} \mathbb{E}_{A \sim \varphi_{n \times N}} \bar{f}(A) &= \int_{\mathcal{M}} f(x) \left(\int_{A \in \pi_1(\pi_2^{-1}(x))} \varphi_{n \times N}(A) |\det(A|_{T_x \mathcal{M}})| dA \right) dx \\ &= \frac{1}{\text{vol}(\mathbb{P}^n)} \int_{\mathcal{M}} f(x) dx; \end{aligned}$$

the second equality by [Lemma 6.3](#). This proves the first part.

Let now $Y \in \mathcal{I}$ be the random variable obtained by choosing $A \in \mathbb{R}^{n \times N}$ with distribution $\psi(A) = \frac{\varphi(A) \bar{f}(A)}{\mathbb{E}_{A \sim \varphi_{n \times N}} \bar{f}(A)}$ and, independently of A , a point $x \in \mathcal{M} \cap \mathcal{L}_A$ with probability $f(x) \bar{f}(A)^{-1}$. Then, by construction, $X = \pi_2(Y)$. Let γ be the density of X . Applying the first part of [Corollary 3.7](#) to π_1 and then the second part to π_2 , we have

$$\begin{aligned} \gamma(x) &= \int_{(A, x) \in \pi_2^{-1}(x)} \frac{\psi(A) f(x) \text{NJ}(\pi_2, (A, x))}{\bar{f}(A) \text{NJ}(\pi_1, (A, x))} dA \\ &= \frac{f(x)}{\mathbb{E}_{\varphi}(\bar{f})} \int_{(A, x) \in \pi_2^{-1}(x)} \varphi(A) |\det(A|_{T_x \mathcal{M}})| dA \\ &= \frac{f(x)}{\mathbb{E}_{\varphi}(\bar{f})} \frac{1}{\text{vol}(\mathbb{P}^n)} \\ &= \frac{f(x)}{\int_{\mathcal{M}} f(x) dx}; \end{aligned}$$

the last penultimate equality again by [Lemma 6.3](#), and the last equality by the first part of the theorem. This finishes the proof. ■

7. Comparison with previous work. We now briefly review the use of kinematic formulas in applications and compare our method to them.

In [\[8\]](#), the authors use a Crofton-type formula for curves to establish a link between a discrete *cut metric* on a grid, which is an object in combinatorial optimization, and a Euclidean metric on \mathbb{R}^2 . This is then applied to a problem in image segmentation.

In [\[28\]](#), the authors use the Cauchy formula and Crofton formula to compute Minkowski measures (e.g., surface area, perimeter) of discrete binary 2-dimensional or 3-dimensional pictures given as a grid of white or black pixels. Since the picture is discrete, the set of lines is also appropriately discretized, as well as Crofton's formula itself. So the difference to our method here lies in the discretization. In [\[29\]](#), a more efficient way to evaluate the discretized Crofton's formula is proposed, using run-length encoding.

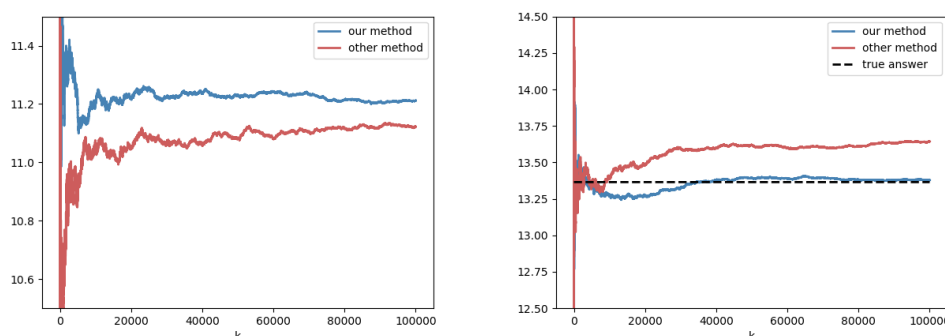


Figure 7.1. The plot shows estimates for the volumes of two curves \mathcal{M} obtained from empirical estimates for $1 \leq k \leq 10^5$ samples. On the left, \mathcal{M} is the curve from (2.1). On the right, \mathcal{M} is the ellipse $\{x \in \mathbb{R}^2 \mid (x/3)^2 + y^2 - 1 = 0\}$. Its volume is known and shown by the black line. The blue curve shows our method, and the red curve shows the method from [35], where we have used the circle of radius 3 for \mathcal{E} . Our method is guaranteed to converge to the true volume of \mathcal{M} for $k \rightarrow \infty$, while the other method is not, as exemplified by the plot. Our method seems to converge at least as quickly as the other method, if not slightly faster.

In [35], the authors use Crofton's formula to approximate the volume of a body \mathcal{M} . They use a different sampling method, which goes as follows: (1) Find a compact body \mathcal{E} containing \mathcal{M} , of known volume $\text{vol}(\mathcal{E})$, such that the space of lines intersecting \mathcal{E} is approximately the same as the space of lines intersecting \mathcal{M} . For example, \mathcal{E} could be a sphere containing \mathcal{M} . (2) Sample uniformly from the set of lines that intersect \mathcal{E} . (3) Compute the total number of intersection points of all the sampled lines with \mathcal{E} (call it g) and with \mathcal{M} (call it h). (4) Approximate the volume of \mathcal{M} as $\frac{h}{g} \text{vol}(\mathcal{E})$.

As the authors write in [35], this method can only give an approximation for the volume of \mathcal{M} . Its accuracy depends on the choice of \mathcal{E} , i.e., on how well the uniform distribution on the set of lines intersecting \mathcal{E} approximates the same with respect to \mathcal{M} . On the other hand, our method is guaranteed to converge to the true volume given enough samples. We tested both methods on the curve \mathcal{M} from (2.1) as well as on the ellipse $\mathcal{M}_1 = \{x \in \mathbb{R}^2 \mid (x/3)^2 + y^2 - 1 = 0\}$, choosing \mathcal{E} to be the centered circle of radius 3. We plotted the results in Figure 7.1.

Finally, we want to mention [19, 36]. In these works the authors derive MCMC methods for sampling \mathcal{M} by intersecting it with random subspaces moving according to the *kinematic measure* in \mathbb{R}^N . This is related to our discussion from the introduction, where we proposed sampling from $\psi(A, b)$ using MCMC methods. Taking this approach and comparing it to [19, 36] is left for future work as we discuss in the next section.

8. Discussion. We explained a new method to sample from a manifold \mathcal{M} described by polynomial implicit equations. The implementation we described generates independent samples from the density $\psi(A, b)$ by rejection sampling, hence, independent points $x \in \mathcal{M}$. As we made experiments, we observed some downsides of our method. Namely, our method becomes slow when the degree of the variety is large in which case it is not easy to find a good κ , and the rejection rate in the sampling process becomes infeasibly large.

But we could also sample from $\psi(A, B)$ using an MCMC method with the goal of improving the rejection rate, at the cost of introducing dependencies between samples. In contrast to the known MCMC methods for nonlinear manifolds, our method would employ MCMC on a flat space. We name using MCMC methods for sampling $\psi(A, B)$ as a possible direction for future research.

Acknowledgments. We would like to thank the following people for helping us with this article: Diego Cifuentes, Benjamin Gess, Christiane Gorgen, Tony Lelievre, Mateusz Michalek, Max von Renesse, Bernd Sturmfels, and Sascha Timme. We also would like to thank two anonymous referees for valuable comments on the paper.

REFERENCES

- [1] PAUL BREIDING, *Energy minimization*, <https://juliahomotopycontinuation.org/examples/energy-minimization/>.
- [2] PAUL BREIDING, *Monte Carlo integration*, <https://www.juliahomotopycontinuation.org/examples/monte-carlo-integration/>.
- [3] PAUL BREIDING, *Sampling*, <https://www.juliahomotopycontinuation.org/examples/sampling/>.
- [4] D. BATES, J. HAUENSTEIN, A. SOMMESE, AND C. WAMPLER, *Bertini: Software for Numerical Algebraic Geometry*, <https://doi.org/10.7274/R0H41PB5> (2013).
- [5] U. BAUER, *Ripser: A Lean C++ Code for Computation of Vietoris–Rips Persistence Barcodes*, <https://github.com/Ripser/ripser>.
- [6] G. BLEKHERMAN, P. PARRILO, AND R. THOMAS, *Semidefinite Optimization and Convex Algebraic Geometry*, MOS-SIAM Ser. Optim. 13, SIAM, Philadelphia, 2012.
- [7] L. BLUM, F. CUCKER, M. SHUB, AND S. SMALE, *Complexity and Real Computation*, Springer, New York, 1998, <https://doi.org/10.1007/978-1-4612-0701-6>.
- [8] Y. BOYKOV AND V. KOLMOGOROV, *Computing geodesics and minimal surfaces via graph cuts*, in Proceedings of the Ninth IEEE International Conference on Computer Vision, IEEE Computer Society, Los Alamitos, CA, 2003, pp. 23–33.
- [9] C. BRAMMER AND D. NELSON, *Toward consistent terminology for cyclohexane conformers in introductory organic chemistry*, J. Chem. Educ. 88 (2011), pp. 292–294.
- [10] P. BREIDING AND S. TIMME, *HomotopyContinuation.jl - a package for solving systems of polynomial equations in Julia*, Mathematical Software – ICMS 2018, Lecture Notes in Computer Science 10931, Springer, Cham, Switzerland, 2018.
- [11] W. H. BROWN AND L. S. BROWN, *Organic Chemistry, Enhanced Edition*, Vol. 5, Brooks Cole, Pacific Grove, CA, 2010.
- [12] M. BRUBAKER, M. SALZMANN, AND R. URTASUN, *A family of MCMC methods on implicitly defined manifolds*, Proc. Mach. Learn. Res. (PMLR), 22 (2012), pp. 161–172.
- [13] P. BURGISSER AND F. CUCKER, *Condition: The Geometry of Numerical Algorithms*, Grundlehren Math. Wiss. 349, Springer, Heidelberg, 2013.
- [14] S. BYRNE AND M. GIROLAMI, *Geodesic Monte Carlo on embedded manifolds*, Scand. J. Stat., 41 (2013), pp. 825–845.
- [15] B. CALDERHEAD AND M. GIROLAMI, *Riemann manifold Langevin and Hamiltonian Monte Carlo methods*, J. Roy. Statist. Soc. Ser. B, 73 (2011), pp. 123–214.
- [16] G. CASELLA AND C. ROBERT, *Monte Carlo Statistical Methods*, Springer Texts Statist., Springer, New York, 2004.
- [17] T. CHEN, T. LEE, T. LI, AND N. OVENHOUSE, *HOMAPS: A Software Package for Solving Polynomial Systems by the Polyhedral Homotopy Continuation Method*.
- [18] E. DUFRESNE, P. B. EDWARDS, H. A. HARRINGTON, AND J. D. HAUENSTEIN, *Sampling real algebraic varieties for topological data analysis*, in International Conference on Machine Learning and Applications, IEEE, Piscataway, NJ, 2018, pp. 1531–1536.

- [19] A. EDELMAN AND O. MANGOUBI, *Integral Geometry for Markov Chain Monte Carlo: Overcoming the Curse of Search-Subspace Dimensionality*, preprint, <https://arxiv.org/abs/1503.03626>, 2015.
- [20] P. EDWARDS, *private communication*.
- [21] M. FARBER AND F. SCHÜTZ, *Homology of planar polygon spaces*, *Geom. Dedicata*, 125 (2007), pp. 75–92.
- [22] J. GOODMAN, M. HOLME-CERFON, AND E. ZAPPA, *Monte Carlo on manifolds: Sampling densities and integrating functions*, *Comm. Pure Appl. Math.*, 71 (2018), pp. 2609–2647.
- [23] J. HARRIS, *Algebraic Geometry, A First Course*, *Grad. Text in Math.* 133, Springer, New York, 1992.
- [24] R. HOWARD, *The Kinematic Formula in Riemannian Homogeneous Spaces*, *Mem. Amer. Math. Soc.* 106, 1993, <https://doi.org/10.1090/memo/0509>.
- [25] J. D. HUNTER, *Matplotlib: A 2D graphics environment*, *Comput. Sci. Eng.*, 9 (2007), pp. 90–95, <https://doi.org/10.1109/MCSE.2007.55>.
- [26] M. H. KALOS AND P. A. WHITLOCK, *Monte Carlo Methods*, 2nd ed., Wiley-Blackwell, Weinheim, Germany, 2008, <https://doi.org/10.1002/9783527626212>.
- [27] P. LAIREZ, *Rigid continuation paths I. Quasilinear average complexity for solving polynomial systems*, *J. Amer. Math. Soc.*, 33 (2020), pp. 487–526.
- [28] D. LEGLAND, K. KIÉU, AND M.-F. DEVAUX, *Computation of Minkowski measures on 2D and 3D binary images*, *Image Anal. Stereol.*, 26 (2007), pp. 83–92.
- [29] G. LEHMANN AND D. LEGLAND, *Efficient n-dimensional surface estimation using Crofton formula and run-length encoding*, *Insight J.* (2012).
- [30] K. LEICHTWEISS, *Zur Riemannschen Geometrie in Grassmannschen Mannigfaltigkeiten*, *Math. Z.*, 76 (1961), pp. 334–366.
- [31] T. LELIÈVRE, M. ROUSSET AND G. STOLTZ, *Hybrid Monte Carlo methods for sampling probability measures on submanifolds*, *Numer. Math.*, 143 (2019), pp. 379–421.
- [32] T. LELIÈVRE, M. ROUSSET, AND G. STOLTZ, *Free Energy Computations: A Mathematical Perspective*, Imperial College Press, London, 2010.
- [33] T. LELIÈVRE, M. ROUSSET, AND G. STOLTZ, *Langevin dynamics with constraints and computation of free energy differences*, *Math. Comp.*, 81 (2012), pp. 2071–2125, <https://doi.org/10.1090/S0025-5718-2012-02594-4>.
- [34] A. LEYKIN, *Homotopy continuation in macaulay2*, *Mathematical Software – ICMS 2018. Lecture Notes in Comput. Sci.* 10931, Springer, Cham, Switzerland, 2018, pp. 328–334.
- [35] X. LI, W. WANG, R. R. MARTIN, AND A. BOWYER, *Using low-discrepancy sequences and the Crofton formula to compute surface areas of geometric models*, *Comput.-Aided Des.*, 35 (2003), pp. 771–782.
- [36] O. MANGOUBI, *Integral Geometry, Hamiltonian Dynamics, and Markov Chain Monte Carlo*, *Ph.D Thesis*, Massachusetts Institute of Technology, Cambridge, MA, 2016.
- [37] R. MUIRHEAD, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.
- [38] L. SANTALÓ, *Integral Geometry and Geometric Probability*, *Encyclopedia Math. Appl.*, Addison-Wesley, Reading, PA, 1976.
- [39] J. SKOWRON AND A. GOULD, *General Complex Polynomial Root Solver and Its Further Optimization for Binary Microlenses*, preprint, <https://arxiv.org/abs/1203.1034>, 2012.
- [40] J. VERSCHELDE, *PHCpack: A General-Purpose Solver for Polynomial Systems by Homotopy Continuation*.
- [41] H. WHITNEY, *Elementary structure of real algebraic varieties.*, *Ann. of Math.* (22), 66 (1957), pp. 545–556.