Open and closed random walks with fixed edgelengths in $\mathbb{R}^d$

Jason Cantarella,[*] Kyle Chapman,[*] Philipp Reiter,[*] and Clayton Shonkwiler[†]

(Dated: June 4, 2018)

In this paper, we consider fixed edgelength n-step random walks in $\mathbb{R}^d$. We give an explicit construction for the closest closed equilateral random walk to almost any open equilateral random walk based on the geometric median, providing a natural map from open polygons to closed polygons of the same edgelength. Using this, we first prove that a natural reconfiguration distance to closure converges in distribution to a Nakagami($d/2$, $d/d-1$) random variable as $n \to \infty$. We then strengthen this to an explicit probabilistic bound on the distance to closure for a random n-gon in any dimension with any collection of fixed edgelengths w_i. Numerical evidence supports the conjecture that our closure map pushes forward the natural probability measure on open polygons to something very close to the natural probability measure on closed polygons; if this is so, we can draw some conclusions about the frequency of local knots in closed polygons of fixed edgelength.

1. INTRODUCTION

Random walks in space with fixed edgelengths have been of interest to statistical physicists and chemists since Lord Rayleigh's day. These walks model polymers in solution (at least under θ-solvent conditions) [9, 14, 20] and are similarly interesting in computational geometry and mathematics as a space of "linkages" [3, 16]. While 2- and 3-dimensional walks are the most relevant to this case, high-dimensional random walks often shed light on the lower dimensional situation [21].

In this paper, we will consider the relationship between open and closed random walks of fixed edgelengths (in polymeric language, the shapes of linear and ring polymers). We will provide an explicit algorithm for finding the nearest closed polygon with given edgelengths to almost any collection of edge directions, and use our construction to provide tail bounds on the fraction of polygon space within a fixed distance of the closed polygons in any dimension. Our results will be strongest for equilateral polygons, but provide explicit bounds for any collection of edgelengths.

To establish notation, we describe random walks in $\mathbb{R}^d$ with (fixed) positive edgelengths w_i by their *edge clouds* $(w_1, \hat{x}_1), \ldots, (w_n, \hat{x}_n)$ where $\hat{x}_i \in S^{d-1}$ is the direction of the ith edge. The space of polygonal arms $\mathrm{Arm}(n, d, w)$ is topologically equivalent to $(S^{d-1})^n$. If we let $\omega_i = \frac{w_i}{\sum w_i}$ be the relative edgelengths, then we can define the submanifold $\{\boldsymbol{x} : \sum \omega_i \hat{x}_i = \vec{0}\}$ of closed polygons $\mathrm{Pol}(n, d, w)$.

Using Bernstein's inequality (e.g. [8]), there is an easy concentration inequality which suggests

———
[*]Mathematics Department, University of Georgia, Athens GA
[†]Department of Mathematics, Colorado State University, Fort Collins CO

that the endpoints of random arms are close together. For equilateral polygons in $\mathbb{R}^3$, this takes the simple form

Theorem 1. *If $\boldsymbol{x}$ is chosen randomly in $\mathrm{Arm}(n, 3, 1)$ with edges $\hat{x}_1, \ldots, \hat{x}_n$,*

$$\mathcal{P}\left(\frac{1}{n}\left\|\sum \hat{x}_i\right\| < t\right) \geq 1 - 3\,e^{-nt^2 \cdot \frac{3}{6 + 2\sqrt{3}t}}.$$

That is, the center of mass of a random edge cloud is very likely to be close to the origin. We can clearly close a random polygon in $\mathrm{Arm}(n, 3, 1)$ by subtracting the (small) $\frac{1}{n}\sum \hat{x}_i$ from each edge. That closed polygon is clearly near the original arm, but it is no longer equilateral. This raises the question of whether we can generally close a polygon in $\mathrm{Arm}(n, 3, 1)$ (or $\mathrm{Arm}(n, d, w)$) while preserving edgelengths and changing the polygon only a small amount. This question is the focus of this paper.

Given $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathrm{Arm}(n, d, w)$, we view both as vectors in $\mathbb{R}^{dn}$ and measure the distance between them accordingly. We call this the *chordal* distance because it does not measure the arc on the spheres of radius w_i for each pair of edges, but rather measures the straight line distance between edge vectors.

Our first main result is Proposition 10, which shows that the chordal distance between a random $\boldsymbol{x} \in \mathrm{Arm}(n, d, 1)$ and the nearest $\boldsymbol{y} \in \mathrm{Pol}(n, d, 1)$ converges in distribution to a Nakagami-$(d/2, d/d-1)$ random variable with PDF proportional to $x^{d-1}e^{-\frac{d-1}{2}x^2}$ as $n \to \infty$.

Our second main result is a general probabilistic bound on the chordal distance to closure for random polygons in any $\mathrm{Arm}(n, d, w)$. For equilateral polygons in $\mathbb{R}^3$, our main theorem (Corollary 19) takes the very simple form

$$\mathcal{P}\left(d_{\mathrm{chordal}}(\boldsymbol{x}, \mathrm{Pol}(n, 3, 1)) < t\right) \geq 1 - 6\exp\left(-t^2/4\right)$$

for $t < \frac{\sqrt{n}}{200\sqrt{2}}$.

Here is a broad overview of our arguments. Given a polygon $\boldsymbol{x}$ in $\mathrm{Arm}(n, d, w)$, we will provide an explicit construction for a nearby closed polygon in $\mathrm{Pol}(n, d, w)$, which we call the *geometric median closure* of $\boldsymbol{x}$ (denoted $\mathrm{gmc}(\boldsymbol{x})$). It will be clear how to construct the geodesic in $\mathrm{Arm}(n, d, w)$ from $\boldsymbol{x}$ to $\mathrm{gmc}(\boldsymbol{x})$. For equilateral polygons, we show $\mathrm{gmc}(\boldsymbol{x})$ is the closest closed polygon to $\boldsymbol{x}$ in chordal distance (Theorem 8).

The distance between $\boldsymbol{x}$ and $\mathrm{gmc}(\boldsymbol{x})$ depends on the norm $\|\vec{\mu}\|$ of the geometric median (or Fermat-Weber point) of the edge cloud (Proposition 12). For equilateral polygons, we will be able to leverage existing results of Niemiro [19] to find the asymptotic distribution of the geometric median of a random point cloud (Proposition 9). Combining this with the matrix Chernoff inequalities proves our first main result (Proposition 10).

The second main result follows from a concentration inequality for a random polygon in any $\mathrm{Arm}(n, d, w)$, which bounds the probability of a large $\|\vec{\mu}\|$ in terms of n, d, and w. This concentration result (Theorem 18) follows from parallel uses of the scalar and matrix Bernstein inequalities

to control the expected properties of a random edge cloud, together with the definition of the geometric median as the minimum of a convex function.

Last, we will observe that the pushforward measure on closed polygons obtained by closing random open polygons appears to converge rapidly to the uniform distribution on closed polygons (Conjecture 22). Since these closures involve only very small motions of any part of the polygon, local features (such as small knots) should be preserved – it would follow (Conjecture 23) that the rate of production of local knots in open and closed arcs should be almost the same.

2. CONSTRUCTING A NEARBY CLOSED POLYGON

As mentioned above, we view n-edge polygons (up to translation) in $\mathbb{R}^d$ as collections of edge vectors $\vec{x}_i \in \mathbb{R}^d$.[1] The vertices are obtained by summing the $\vec{x}_i$ from an arbitrary basepoint. In this section of the paper, we will assume only that the lengths of the edges of the polygon are fixed to some arbitrary $w_i = \|\vec{x}_i\|$. We will think of these fixed edgelength polygons in two ways:

- as a weighted point cloud on the unit sphere $S^{d-1} \subset \mathbb{R}^d$ where the points are denoted $\hat{x}_i = \vec{x}_i/\|\vec{x}_i\|$ and the weights are the w_i. We will call $(w_i, \hat{x}_i)$ the *edge cloud* of the polygon.

- as a point $\boldsymbol{x} \in \prod S^{d-1}(w_i) \subset (\mathbb{R}^d)^n = \mathbb{R}^{dn}$ (where $S^{d-1}(r)$ is the sphere of radius r). We will call $\boldsymbol{x}$ the *vector of edges* of the polygon.

The space of these polygons will be denoted $\mathrm{Arm}(n, d, w) = \prod S^{d-1}(w_i)$. Within this space, there is a submanifold $\mathrm{Pol}(n, d, w)$ of closed polygons defined by the condition $\sum w_i \hat{x}_i = \vec{0}$. (Equivalently, $\boldsymbol{x}$ is closed if it lies in the codimension d subspace of $\mathbb{R}^{dn}$ normal to the $\boldsymbol{n}^j = (\hat{e}_j, \ldots, \hat{e}_j)$, where $\hat{e}_1, \ldots, \hat{e}_d$ is the standard basis in $\mathbb{R}^d$.) Both $\mathrm{Arm}(n, d, w)$ and $\mathrm{Pol}(n, d, w)$ are Riemannian manifolds with standard metrics, but it will be useful to use two additional metrics as well:

Definition 2. The *chordal* metric on $\mathrm{Arm}(n, d, w)$ is given by

$$d_{\text{chordal}}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|_{\mathbb{R}^{dn}} = \left(\sum \|w_i \hat{x}_i - w_i \hat{y}_i\|_{\mathbb{R}^d}^2 \right)^{1/2}.$$

The *max-angular* metric on $\mathrm{Arm}(n, d, w)$ is given by

$$d_{\text{max-angular}}(\boldsymbol{x}, \boldsymbol{y}) = \max_i \angle(\vec{x}_i, \vec{y}_i).$$

---

[1] Throughout this paper, we use boldface to indicate elements of $\mathbb{R}^{dn}$, which we usually think of as vectors of edge vectors. We use a superscript arrow – as in $\vec{x}_i$ – to denote an arbitrary element of $\mathbb{R}^d$, though any such vector which is definitionally a unit vector we mark with a hat rather than an arrow.

We now make an important definition:

Definition 3. A *geometric median* (also known as a *Fermat-Weber point*) of an edge cloud $(w_i, \hat{x}_i)$ is any point $\vec{\mu}$ which minimizes the *weighted average distance function* $\mathrm{Ad}_{\boldsymbol{x}}(\vec{y})$ given by

$$\mathrm{Ad}_{\boldsymbol{x}}(\vec{y}) = \sum_i \omega_i \|\hat{x}_i - \vec{y}\|.$$

where $\omega_i = w_i / \sum w_i$. To clarify notation, we will only use $\vec{\mu}$ for points which are a geometric median of a weighted point cloud; the point cloud will be clear from the context.

This is a very old construction with a beautiful theory around it; see the nice review in [7]. We note that the geometric median differs from the center of mass (or geometric *mean*) of the points, which minimizes the weighted average of the *squared* distances between $\vec{y}$ and the $\hat{x}_i$ and that the geometric median is unique unless the points are all colinear and the geometric median is not one of the points.

This section is devoted to analyzing the following construction:

Definition 4. Suppose $\boldsymbol{x}$ is a polygon and $\vec{\mu}$ is a geometric median of its edge cloud $(w_i, \hat{x}_i)$ which is not one of the $\hat{x}_i$. The *geometric median closure* $\mathrm{gmc}(\boldsymbol{x})$ of $\boldsymbol{x}$ is the polygon whose edge cloud has the same weights and edge directions obtained by recentering the $\hat{x}_i$ on $\vec{\mu}$ and renormalizing: $\mathrm{gmc}(\boldsymbol{x})$ has edge cloud $\left(w_i, \frac{\hat{x}_i - \vec{\mu}}{|\hat{x}_i - \vec{\mu}|}\right)$.

If every geometric median of $(w_i, \hat{x}_i)$ is one of the $\hat{x}_i$, $\mathrm{gmc}(\boldsymbol{x})$ is not defined. If $\mathrm{gmc}(\boldsymbol{x})$ is defined, we say that $\boldsymbol{x}$ is *median-closeable*.

Of course, we need to justify our choice of name by proving that $\mathrm{gmc}(\boldsymbol{x})$ is closed. The key observation is the following Lemma, which follows by direct computation:

Lemma 5. *The function $\mathrm{Ad}_{\boldsymbol{x}}(\vec{y})$ is a convex function of $\vec{y}$. The gradient is given by*

$$\nabla \mathrm{Ad}_{\boldsymbol{x}}(\vec{y}) = \sum \omega_i \frac{\vec{y} - \hat{x}_i}{\|\vec{y} - \hat{x}_i\|}.$$

The Hessian of $\mathrm{Ad}_{\boldsymbol{x}}(\vec{y})$ is given by

$$\mathcal{H}\,\mathrm{Ad}_{\boldsymbol{x}}(\vec{y}) = \left(\sum_i \frac{\omega_i}{\|\hat{x}_i - \vec{y}\|}\right) I_d - \left(\sum_i \frac{\omega_i}{\|\hat{x}_i - \vec{y}\|^3}(\vec{y} - \hat{x}_i)(\vec{y} - \hat{x}_i)^T\right).$$

Proposition 6. *If $\boldsymbol{x}$ is median-closeable, $\mathrm{gmc}(\boldsymbol{x})$ is a unique closed polygon with edgelengths w_i.*

Proof. The proof follows from assembling several standard facts about the geometric median. These are in [18], but are easily checked by hand.

As it is a sum of convex functions, the average distance function $\mathrm{Ad}_{\boldsymbol{x}}$ is convex. Away from the $\hat{x}_i$, it is differentiable. If the points $\hat{x}_i$ are not colinear, $\mathrm{Ad}_{\boldsymbol{x}}$ is strictly convex and $\vec{\mu}$ is unique. If the points $\hat{x}_i$ are colinear, either the geometric median is one of the $\hat{x}_i$ or the set of geometric medians consists of the interval between two $\hat{x}_i$.

Any geometric median which is not one of the $\hat{x}_i$ must be a critical point of the average distance function. For any such $\vec{\mu}$, using Lemma 5,

$$\sum_i w_i \frac{\hat{x}_i - \vec{\mu}}{|\hat{x}_i - \vec{\mu}|} = -\left(\sum w_i\right) \nabla \mathrm{Ad}_{\boldsymbol{x}}(\vec{\mu}) = 0. \tag{1}$$

This implies that $\mathrm{gmc}(\boldsymbol{x})$ is closed.

If $\vec{\mu}$ is unique, then $\mathrm{gmc}(\boldsymbol{x})$ is obviously unique. If $\vec{\mu}$ is not unique, the $\hat{x}_i$ are colinear, and $\vec{\mu}$ is on the line segment between two of the $\hat{x}_i$. In this case, it is not hard to see that (1) implies that the edges of $\mathrm{gmc}(\boldsymbol{x})$ are two antipodal groups of points on S^{k-1}, each containing $n/2$ points, regardless of where we take $\vec{\mu}$ on the segment. $\qquad\square$

Our next goal is to prove an optimality property for the geometric median closure. We will start by proving a more general fact about recentering and renormalizing:

Proposition 7. *Suppose that $\hat{x}_i$ is any point cloud in $(S^{d-1})^n$, and $\vec{p} \in \mathbb{R}^d$ is not one of the $\hat{x}_i$. Given any set of weights w_i, we let $r(\boldsymbol{x}; \vec{p}, w)$ denote the renormalized, recentered, and reweighted point cloud, and $\vec{s}$ denote its weighted sum:*

$$r(\boldsymbol{x}; \vec{p}, w) := \left(w_i, \frac{\hat{x}_i - \vec{p}}{|\hat{x}_i - \vec{p}|}\right) \quad and \quad \vec{s} := \sum_i w_i \frac{\hat{x}_i - \vec{p}}{|\hat{x}_i - \vec{p}|}.$$

If $\boldsymbol{x} = (\hat{x}_1, \ldots, \hat{x}_n)$ and $\boldsymbol{r}$ is the vector of edges corresponding to the edge cloud $(w_i, \hat{r}_i)$, then

$$\boldsymbol{r} = \mathrm{argmin}_{\sum w_i \hat{y}_i = \vec{s}} \|\boldsymbol{y} - \boldsymbol{x}\|,$$

that is, $\boldsymbol{r}$ is the closest vector of edges to $\boldsymbol{x}$ (in $\mathbb{R}^{dn}$) with edge weights w_i and vector sum $\vec{s}$.

Proof. Suppose that $(w_i, \hat{y}_i)$ is a point cloud with the same weights which also has $\sum_i w_i \hat{y}_i = \vec{s}$ and $\boldsymbol{y}$ is the corresponding vector of edges in $\mathbb{R}^{dn}$. Let $\boldsymbol{v} = \boldsymbol{y} - \boldsymbol{r}$. Since $\sum_i \vec{y}_i = \sum_i \vec{r}_i$, we know $\sum \vec{v}_i = \vec{0}$.

Remembering that $\|\vec{y}_i\| = w_i = \|\vec{r}_i\|$, we compute

$$\langle \vec{v}_i, \vec{r}_i \rangle = \langle \vec{y}_i - \vec{r}_i, \vec{r}_i \rangle = \langle \vec{y}_i, \vec{r}_i \rangle - w_i^2 \leq 0.$$

Since $\vec{r}_i$ is a positive scalar multiple of $\hat{x}_i - \vec{p}$, this implies that $\langle \vec{v}_i, \hat{x}_i - \vec{p} \rangle \leq 0$, and so we have $\langle \vec{v}_i, \hat{x}_i \rangle \leq \langle \vec{v}_i, \vec{p} \rangle$. Since $\sum \vec{v}_i = \vec{0}$, we see

$$\langle \boldsymbol{v}, \boldsymbol{x} \rangle = \sum_i \langle \vec{v}_i, \hat{x}_i \rangle \leq \sum_i \langle \vec{v}_i, \vec{p} \rangle = \left\langle \sum_i \vec{v}_i, \vec{p} \right\rangle = 0,$$

or that $-\langle \boldsymbol{x}, \boldsymbol{v} \rangle \geq 0$. Using the facts $\langle \boldsymbol{y}, \boldsymbol{y} \rangle = \sum w_i^2 = \langle \boldsymbol{r}, \boldsymbol{r} \rangle$ and $\boldsymbol{y} = \boldsymbol{r} + \boldsymbol{v}$,

$$\begin{aligned} \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle &= \langle \boldsymbol{x}, \boldsymbol{x} \rangle + \langle \boldsymbol{y}, \boldsymbol{y} \rangle - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle \\ &= \langle \boldsymbol{x}, \boldsymbol{x} \rangle + \langle \boldsymbol{r}, \boldsymbol{r} \rangle - 2\langle \boldsymbol{x}, \boldsymbol{r} \rangle - 2\langle \boldsymbol{x}, \boldsymbol{v} \rangle \\ &= \langle \boldsymbol{x} - \boldsymbol{r}, \boldsymbol{x} - \boldsymbol{r} \rangle - 2\langle \boldsymbol{x}, \boldsymbol{v} \rangle \end{aligned}$$

so $\|\boldsymbol{x} - \boldsymbol{y}\| \geq \|\boldsymbol{x} - \boldsymbol{r}\|$, as claimed. $\square$

Combining Propositions 6 and 7 with Definition 4, we have

Theorem 8. *If $\boldsymbol{x}$ is a median-closeable equilateral polygon, $\mathrm{gmc}(\boldsymbol{x})$ is the closed equilateral polygon closest to $\boldsymbol{x}$ in the chordal metric.*

Remarks. This construction may seem unexpected, but it has deep roots. In [15], Kapovich and Millson provide an analogous closure construction which associates a unique closed equilateral polygon to any equilateral polygon where no more than half the edge vectors coincide by viewing the unit ball as the Poincaré ball model of hyperbolic space and (essentially) recentering and renormalizing in hyperbolic geometry around a point called the "conformal median" (see [6]) which is in many ways parallel to the geometric median. This is an example of a "Geometric Invariant Theory" (or GIT) quotient: see [13]. These ideas inspired our work above: we did not adopt them entirely only because working in hyperbolic geometry makes the whole endeavor seem much more abstract and because we have not managed to prove an optimality property for their construction analogous to Theorem 8.

3. ASYMPTOTICS OF THE GEOMETRIC MEDIAN AND THE DISTANCE TO CLOSURE

Now that we've established the connection between the geometric median and closure, we will establish some facts about the large-n behavior of the geometric median. Since the geometric median is a symmetric estimator of a large number of i.i.d. random variables, it seems natural to expect that the distribution of μ should converge to a multivariate normal, even though the classical central limit theorem doesn't apply. In fact, this is true:

Proposition 9. *Let n points $\hat{x}_i$ be sampled independently and uniformly on S^{d-1}, with geometric median $\vec{\mu}$. The random variable $\sqrt{n}\,\vec{\mu}$ converges in distribution to $\mathcal{N}\left(\vec{0}, \frac{d}{(d-1)^2} I_d\right)$ as $n \to \infty$.*

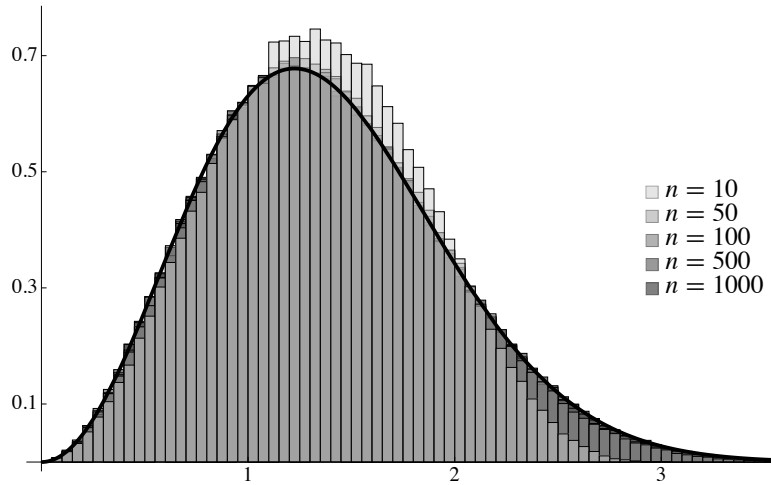FIG. 1: For various n, we generated 250,000 random elements of $\mathrm{Arm}(n, 3, 1)$ and computed $\sqrt{n}\,\|\vec{\mu}\|$, where $\vec{\mu}$ is the geometric median of the edge cloud. By Proposition 9, $\sqrt{n}\,\|\vec{\mu}\|$ converges to a Nakagami$\left(\frac{3}{2}, \frac{9}{4}\right)$ distribution, the pdf of which is the solid curve. Though we don't show it, the behavior in other dimensions is quite similar: by $n = 50$ the density of the limiting Nakagami$\left(\frac{d}{2}, \left(\frac{d}{d-1}\right)^2\right)$ distribution matches the histogram rather well.

This implies that $\|\sqrt{n}\,\vec{\mu}\|$ converges in distribution to a Nakagami$\left(\frac{d}{2}, \left(\frac{d}{d-1}\right)^2\right)$ random variable.

Proof. We start by defining $\mathrm{Ed}(\vec{y})$ to be the expected distance from $\vec{y} \in \mathbb{R}^d$ to the unit sphere; formally,

$$\mathrm{Ed}(\vec{y}) := \frac{1}{\mathrm{Vol}\, S^{d-1}} \int_{\hat{x} \in S^{d-1}} \|\hat{x} - \vec{y}\|\; \mathrm{dVol}_{S^{d-1}}\,.$$

Observe that, by symmetry, the minimizer of $\mathrm{Ed}(\vec{y})$ is the origin. Now the geometric median of a finite collection of points $\hat{x}_i$ uniformly sampled from the sphere is the minimizer of the average distance Ad to the $\hat{x}_i$ (Definition 3). For a large number of points, we expect Ad to be close to Ed as a function, and hence that the minimizers of the functions should be nearby as well.

In fact, Niemiro studied exactly this situation, showing[2] ([19, p. 1517], cf. Haberman [11]) that

$$\sqrt{n}\,\vec{\mu} \xrightarrow{d} \mathcal{N}(\vec{0}, \mathcal{H}^{-1}V\mathcal{H}^{-1})$$

---

[2] Under some technical hypotheses which are obviously satisfied in our case.

where V is the covariance matrix of a random point $\hat{x}$ on S^{d-1} and $\mathcal{H}$ is the Hessian of Ed, evaluated at the origin.

The off-diagonal elements of V are zero by symmetry. Using cylindrical coordinates on S^{d-1} with axis $\hat{e}_i$, the ith diagonal entry in the covariance matrix is computed by the integral

$$\sigma_i^2 = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} \int_{-1}^{1} x^2 (1-x^2)^{\frac{d-3}{2}} \, \mathrm{d}x = \frac{1}{d}.$$

We prove in the Appendix (Proposition 24) that the expected distance function $\mathrm{Ed}(\vec{y})$ is given as a function of $r = \|\vec{y}\|$ by

$$\mathrm{Ed}(r) = {}_2F_1\left(-\frac{1}{2}, \frac{1-d}{2}; \frac{d}{2}; r^2\right).$$

When d is odd, the standard Taylor series representation of the hypergeometric function truncates, and $\mathrm{Ed}(r)$ is a polynomial in r. For example, when $d = 3$ we have $\mathrm{Ed}(r) = 1 + r^2/3$. In turn, a straightforward computation shows that the Hessian of Ed evaluated at the origin is simply

$$\mathcal{H} = \mathcal{H}\,\mathrm{Ed}(\vec{0}) = \frac{d-1}{d} I_d,$$

where I_d is the $d \times d$ identity matrix. This completes the proof of the first statement. To get the second, we note that the norm of a Gaussian $\mathcal{N}(\vec{0}, \sigma^2 I_d)$ random variate is Nakagami$\left(\frac{d}{2}, d\sigma^2\right)$-distributed. $\square$

We now see that the geometric median is becoming asymptotically normal, and concentrating around the origin. We can use this to prove an asymptotic result for the distance to closure for equilateral polygons.

Proposition 10. *For a random equilateral n-gon $\boldsymbol{x}$ with edges $\hat{x}_i$ sampled independently and uniformly from S^{d-1}, the random variable $d_{chordal}(\boldsymbol{x}, \mathrm{Pol}(n, d, 1))$ converges in distribution to a Nakagami$\left(\frac{d}{2}, \frac{d}{d-1}\right)$ as $n \to \infty$.*

Proof. We know from Theorem 8 that $d_{\text{chordal}}(\boldsymbol{x}, \mathrm{Pol}(n, d, 1))$ is actually the chordal distance from $\boldsymbol{x}$ to $\mathrm{gmc}(\boldsymbol{x})$. To estimate this distance, we will make use of the recentering and renormalizing map $r(\boldsymbol{x}, \vec{p}, 1)$ from Proposition 7.

When $\|\mu\|$ is small, we can estimate

$$\|\mathrm{gmc}(\boldsymbol{x}) - \boldsymbol{x}\| = \left\|r(\boldsymbol{x}; \vec{\mu}, 1) - r(\boldsymbol{x}; \vec{0}, 1)\right\| \sim \|\vec{\mu}\| \left\|D_{\hat{\mu}} r(\boldsymbol{x}; \vec{0}, 1)\right\|$$

where $D_{\hat{\mu}} r(\boldsymbol{x}; \vec{0}, 1)$ is the derivative of $r(\boldsymbol{x}; \vec{v}, 1)$ with respect to the vector $\vec{v}$ in the direction of the unit vector $\hat{\mu} = \vec{\mu}/\|\vec{\mu}\|$ (while leaving the $\boldsymbol{x}$ variables constant).
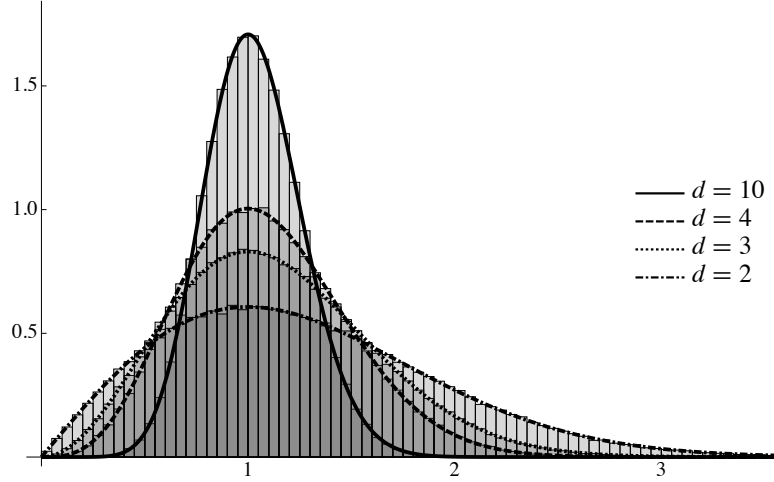
FIG. 2: For $d = 2, 3, 4, 10$, we generated 250,000 random elements of $\mathrm{Arm}(1000, d, 1)$ and computed their chordal distance to $\mathrm{Pol}(1000, d, 1)$ using Theorem 8. This plot shows the histograms of chordal distance to closure together with the densities of $\mathrm{Nakagami}\left(\frac{d}{2}, \frac{d}{d-1}\right)$ distributions.

Using the definition of $r(\boldsymbol{x}; \vec{p}, 1)$, a direct computation reveals that

$$\left\| D_{\hat{\mu}} r(\boldsymbol{x}; \vec{0}, 1) \right\| = \left(n - \sum_i \langle \hat{x}_i, \hat{\mu} \rangle^2 \right)^{\frac{1}{2}} = \sqrt{n} \left(1 - \frac{1}{n} \sum_i \langle \hat{x}_i, \hat{\mu} \rangle^2 \right)^{\frac{1}{2}}.$$

Since $\hat{\mu}$ is a unit vector, the sum is the Rayleigh quotient for the matrix $X = \frac{1}{n} \sum_i \hat{x}_i \hat{x}_i^T$, and so obeys the estimates

$$\lambda_{\min}(X) \leq \frac{1}{n} \sum_i \langle \hat{x}_i, \hat{\mu} \rangle^2 \leq \lambda_{\max}(X)$$

Now $\lambda_{\min}(X)$ and $\lambda_{\max}(X)$ are also random variables depending on the $\hat{x}_i$, but we can use the matrix Chernoff inequalities [22, Remark 5.3] to bound the probability that they are far from $\frac{1}{d}$.

It's quite standard to prove that $\mathcal{E}(\hat{x}_i \hat{x}_i^T) = \frac{1}{d} I_d$, so $\mathcal{E}(X) = \frac{1}{d} I_d$. The matrix Chernoff inequalities then reduce to

$$\mathcal{P} \left\{ \lambda_{\min}(X) < (1 - \delta) \frac{1}{d} \right\} \leq d \left(\frac{e^{-\delta}}{(1 - \delta)^{1-\delta}} \right)^{\frac{n}{d}} \tag{2}$$

and

$$\mathcal{P} \left\{ \lambda_{\max}(X) > (1 + \delta) \frac{1}{d} \right\} \leq d \left(\frac{e^{\delta}}{(1 + \delta)^{1+\delta}} \right)^{\frac{n}{d}} \tag{3}$$

For any $\delta > 0$, the quantities raised to $\frac{n}{d}$ are < 1, and so as $n \to \infty$ the probability that the bounds in (2) and (3) both hold $\to 1$. In turn, this means that for any fixed $\delta > 0$,

$$\mathcal{P}\left\{\left|\frac{1}{d} - \frac{1}{n}\sum_i \langle \hat{x}_i, \hat{\mu}\rangle^2\right| > \frac{\delta}{d}\right\} \to 0$$

and so the random variable $\frac{1}{n}\sum_i \langle \hat{x}_i, \hat{\mu}\rangle^2$ converges in probability to $\frac{1}{d}$. By the continuous mapping theorem, this means that $\left(1 - \frac{1}{n}\sum_i \langle \hat{x}_i, \hat{\mu}\rangle^2\right)^{1/2} \xrightarrow{p} \sqrt{\frac{d-1}{d}}$.

We can now rewrite the random variable $\|\vec{\mu}\| \left\|D_{\hat{\mu}} r(\boldsymbol{x}; \vec{0}, 1)\right\|$ as the product of $\|\sqrt{n}\,\vec{\mu}\|$, which by Proposition 9 converges in distribution to a Nakagami$\left(\frac{d}{2}, \left(\frac{d}{d-1}\right)^2\right)$ random variable, and $\left(1 - \frac{1}{n}\sum_i \langle \hat{x}_i, \hat{\mu}\rangle^2\right)^{1/2}$, which we have just proved converges in probability to the constant random variable $\sqrt{\frac{d-1}{d}}$.

Using Slutsky's theorem and a little algebra, this implies that the product converges in distribution to a Nakagami$\left(\frac{d}{2}, \frac{d}{d-1}\right)$ random variable, as claimed. $\qquad\square$

We have now learned something interesting: the distribution of chordal distances to closure should be converging to a distribution which doesn't depend on the number of edges! This is surprising because the diameter of $\mathrm{Arm}(n, d, 1)$ is clearly $\Theta(\sqrt{n}) \to \infty$. This means that some arms might indeed be very far from closure – but they are very rare. We will look for this feature in the more specific probability inequalities to come.

We can also see how fast the tail of the distribution of d_{chordal} can be expected to decay. The survival function of the Nakagami distribution is an incomplete Gamma function. Using [5, 8.10.1], we can show that there is a constant $C(d) > 0$ so that if x is Nakagami$\left(\frac{d}{2}, \frac{d}{d-1}\right)$, then

$$\mathcal{P}\left\{x < t\right\} \geq 1 - C(d)\, t^{d-2} e^{-\frac{d-1}{2}t^2}. \tag{4}$$

4. CONCENTRATION INEQUALITIES FOR $\|\vec{\mu}\|$ AND $d_{\mathbf{chordal}}$

We now know what to expect in the large-n limit, at least for equilateral polygons. For most applications in physics (4) is sufficient; even for small n, experimental evidence shows that the bound is quite close to the truth.

However, for mathematical applications, it's helpful to have a hard bound on the chordal distance to closure even if we sacrifice some accuracy. Learning from (4), we see that we should aim for a tail bound for d_{chordal} which does not depend on n and is proportional to $e^{-\alpha t^2}$ for some

$\alpha < \frac{d-1}{2}$. We will get exactly such a bound in Corollary 19 at the end of the section. Our bounds will apply for finite n, and also apply to the non-equilateral case, where it is not even clear what the large-n limit should mean.

4.1. A bound connecting $\|\vec{\mu}\|$ and d_{chordal}.

To start with, we prove a hard bound on the relationship between the geometric median and our two measures of distance in polygon space. First, we note that our procedure of recentering and renormalizing changes each $\hat{x}_i$ by a controlled amount.

Lemma 11. *If $\hat{x}_i \in S^{d-1}$ and $\vec{p} \in \mathbb{R}^d$ is any vector with $\|\vec{p}\| < 1$, then $\left\| \hat{x}_i - \frac{\hat{x}_i - \vec{p}}{\|\hat{x}_i - \vec{p}\|} \right\| \leq \sqrt{2}\|\vec{p}\|$ and $\angle(\hat{x}_i, \frac{\hat{x}_i - \vec{p}}{\|\hat{x}_i - \vec{p}\|}) \leq \arcsin \|\vec{p}\| < \frac{\pi}{2}\|\vec{p}\|$.*

Proof. This is a calculus exercise; it is straightforward to establish the (sharp) bound

$$\left\| \hat{x}_i - \frac{\hat{x}_i - \vec{p}}{\|\hat{x}_i - \vec{p}\|} \right\| \leq \sqrt{2 - 2\sqrt{1 - \|\vec{p}\|^2}}.$$

Further, it is easy to check that the right-hand side is a convex function of $\|\vec{p}\|$ which is equal to 0 when $\|\vec{p}\| = 0$, and $\sqrt{2}$ when $\|\vec{p}\| = 1$, so it is bounded above by the line $\sqrt{2}\|\vec{p}\|$. The angle bound is also straightforward. $\qquad\square$

We now can give a bound on the distance between a given $\boldsymbol{x} \in \text{Arm}(n, d, w)$ and $\text{Pol}(n, d, w)$ in terms of the norm of the geometric median $\vec{\mu}$ of the edge cloud $(\hat{x}_i, w_i)$.

Proposition 12. *If the edge cloud $(\hat{x}_i, w_i)$ has geometric median $\vec{\mu}$ with $\|\vec{\mu}\| < 1$,*

$$d_{\text{chordal}}(\boldsymbol{x}, \text{Pol}(n, d, w)) < \sqrt{2 \sum \omega_i^2} \, \|\vec{\mu}\| \quad \text{and} \quad d_{\text{max-angular}}(\boldsymbol{x}, \text{Pol}(n, d, w)) < \arcsin \|\vec{\mu}\|.$$

Proof. Since $\|\vec{\mu}\| < 1$, our polygon is median-closeable and $\text{gmc}(\boldsymbol{x})$ is a closed polygon with edge cloud $\left(w_i, \frac{\hat{x}_i - \vec{\mu}}{\|\hat{x}_i - \vec{\mu}\|} \right)$. Lemma 11 immediately yields the bound on $d_{\text{max-angular}}$; to get the chordal distance bound, we write

$$\sqrt{\sum \left\| w_i \hat{x}_i - w_i \frac{\hat{x}_i - \vec{\mu}}{\|\hat{x}_i - \vec{\mu}\|} \right\|^2} \leq \sqrt{\sum w_i^2 \cdot 2\|\vec{\mu}\|^2}.$$

$\qquad\square$

4.2. Strategy for the tail bound

To derive our explicit tail bound on the norm of the geometric median, our strategy is as follows. First, we will prove two probabilistic bounds: an upper bound on $\|\nabla \mathrm{Ad}_{\boldsymbol{x}}(\vec{0})\|$ and a positive lower bound on $\lambda_{\min}(\mathcal{H}\,\mathrm{Ad}_{\boldsymbol{x}}(\vec{0}))$. These will come from scalar and matrix versions of Bernstein's inequality.

If we restrict $\mathrm{Ad}_{\boldsymbol{x}}$ to a scalar function $\mathrm{Ad}_{\boldsymbol{x}}(z)$ on a ray from the origin, these bounds yield an upper bound on $|\mathrm{Ad}'_{\boldsymbol{x}}(0)|$ and a lower bound on $\mathrm{Ad}''_{\boldsymbol{x}}(0)$. We will get a uniform lower bound on $\mathrm{Ad}''_{\boldsymbol{x}}(z)$ for $z \in [0, 1/50]$ by showing that, $\mathrm{Ad}''_{\boldsymbol{x}}(z) \geq \mathrm{Ad}''_{\boldsymbol{x}}(0) - 7z$ on this interval. We prove this using the special structure of $\mathcal{H}\,\mathrm{Ad}_{\boldsymbol{x}}$.

By Taylor's theorem, there is some z_* in $[0, z]$ so that

$$\mathrm{Ad}'_{\boldsymbol{x}}(z) = \mathrm{Ad}'_{\boldsymbol{x}}(0) + z\,\mathrm{Ad}''_{\boldsymbol{x}}(z_*) \geq -|\mathrm{Ad}'_{\boldsymbol{x}}(0)| + \lambda z.$$

This means that for $z > |\mathrm{Ad}'_{\boldsymbol{x}}(0)|/\lambda$, this directional derivative must be positive: in particular, since the geometric median $\vec{\mu}$ is by definition a point where $\nabla \mathrm{Ad}(\vec{\mu}) = \vec{0}$, $\vec{\mu}$ can lie no farther than $|\mathrm{Ad}'_{\boldsymbol{x}}(0)|/\lambda$ from the origin.

4.3. A probabilistic bound on $\|\nabla \mathrm{Ad}_{\boldsymbol{x}}(\vec{0})\| = \|\sum \omega_i \hat{x}_i\|$

We want to bound the norm of the gradient $\nabla \mathrm{Ad}_{\boldsymbol{x}}(\vec{0})$, which we recall from Lemma 5 is equal to $\sum \omega_i \hat{x}_i$. We will start with a lemma which helps us understand the effect of variable weights ω_i.

Lemma 13. *For any collection of n non-negative real numbers w_i, if we define $\omega_i = w_i/\sum w_i$,*

$$n \geq 1 + n^2\,\mathrm{Var}(\omega_i) = n \sum \omega_i^2 \geq 1,$$

where $\mathrm{Var}(\omega_i)$ is the variance of $\{\omega_1, \ldots, \omega_n\}$. We have equality on the left precisely when all but one of the w_i equal zero and equality on the right precisely when all the w_i are equal.

Proof of Lemma. Starting with the definition of variance, and remembering that $\sum \omega_i = 1$,

$$\mathrm{Var}(\omega_i) = \frac{1}{n}\sum \omega_i^2 - \left(\frac{1}{n}\sum \omega_i\right)^2 = \frac{1}{n}\sum \omega_i^2 - \frac{1}{n^2}$$

Solving for $\sum \omega_i^2$,

$$\sum \omega_i^2 = \frac{1 + n^2\,\mathrm{Var}(\omega_i)}{n}$$

which proves the central equality. Since $\text{Var}(\omega_i) \geq 0$ with equality precisely when all the ω_i are equal, the inequality on the right follows easily.

To prove the inequality on the left, we invoke the Bhatia-Davis inequality [1], which says that since the $0 \leq \omega_i \leq 1$ and the mean of the ω_i is $^1/n$, we have $\text{Var}(\omega_i) \leq (1 - \frac{1}{n})(\frac{1}{n} - 0)$ with equality precisely when one $\omega_i = 1$ and the remainder are zero. $\square$

Now we can give our first result:

Proposition 14. *If we have n points $\hat{x}_i$ sampled independently and uniformly from S^{d-1}, and n weights $\omega_i \geq 0$ with $\sum_i \omega_i = 1$ and $\Omega = \max_i \omega_i$, then for any $t > 0$*

$$\mathcal{P}\left(\left\|\sum \omega_i \hat{x}_i\right\| > t\right) \leq d \exp\left(-\frac{3nt^2}{2nt\Omega\sqrt{d} + 6(1 + n^2 \text{Var}\,\omega_i)}\right).$$

If the ω_i are all equal (the polygon is equilateral), this simplifies to

$$\mathcal{P}\left(\left\|\frac{1}{n}\sum \hat{x}_i\right\| > t\right) \leq d \exp\left(-\frac{3nt^2}{6 + 2t\sqrt{d}}\right).$$

Proof. We will use Bernstein's inequality [8, Theorem 1.2]: Suppose $X_1, \ldots, X_n$ are independent random variables with $X_i - \mathcal{E}(X_i) \leq b$ for each i, the variance of each X_i is given by σ_i^2, and $X = \sum X_i$ (with variance $\sigma^2 = \sum \sigma_i^2$). Then for any $t > 0$,

$$\mathcal{P}\left(X > \mathcal{E}(X) + t\right) \leq \exp\left(-\frac{t^2}{2\sigma^2\left(1 + \frac{bt}{3\sigma^2}\right)}\right).$$

For any unit vector $\vec{v}$, we can set $X_i = \langle \omega_i \hat{x}_i, \vec{v}\rangle$. These random variables clearly have expectation 0 and $X_i - \mathcal{E}(X_i) \leq \omega_i \leq \Omega$. Using cylindrical coordinates on S^{d-1} with axis $\vec{v}$, the variance is computed by the integral

$$\sigma_i^2 = \frac{1}{\sqrt{\pi}}\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})}\int_{-1}^{1}(\omega_i x)^2(1 - x^2)^{\frac{d-3}{2}}\,\mathrm{d}x = \frac{\omega_i^2}{d}.$$

Using Lemma 13, this implies

$$\sigma^2 = \frac{1}{d}\sum \omega_i^2 = \frac{1 + n^2 \text{Var}\,\omega_i}{dn}.$$

This proves that for any $\vec{v}$,

$$\mathcal{P}\left(\left\langle \sum \omega_i \hat{x}_i, \vec{v}\right\rangle > t\right) \leq \exp\left(-ndt^2 \frac{3}{6 + 2\,dnt\,\Omega + 6n^2 \text{Var}\,\omega_i}\right).$$

Applying this inequality d times for $\vec{v} = \hat{e}_1, \ldots, \hat{e}_d$, and using the union bound, we can bound the L^∞ norm of $\sum \omega_i \hat{x}_i$:

$$\mathcal{P}\left(\left\|\sum \omega_i \hat{x}_i\right\|_\infty = \max_j \left(\sum \omega_i \hat{x}_i\right)_j > t\right) \leq d \exp\left(-ndt^2 \frac{3}{6 + 2\,dnt\,\Omega + 6n^2 \operatorname{Var}\omega_i}\right).$$

But we know that for any $\vec{u} \in \mathbb{R}^d$ we have $\frac{1}{\sqrt{d}}\|\vec{u}\| \leq \|\vec{u}\|_\infty$, so

$$\mathcal{P}\left(\frac{1}{\sqrt{d}}\left\|\sum \omega_i \hat{x}_i\right\| > t\right) \leq \mathcal{P}\left(\left\|\sum \omega_i \hat{x}_i\right\|_\infty > t\right) \leq d \exp\left(-ndt^2 \frac{3}{6 + 2\,dnt\,\Omega + 6n^2 \operatorname{Var}\omega_i}\right).$$

Replacing t by $\frac{t}{\sqrt{d}}$ yields the statement of the Proposition. $\qquad\square$

The terms Ω and $\operatorname{Var}\omega_i$ in the statement of Proposition 14 at first seem mysterious. However, if you read them in light of Lemma 13, they become clearer.

At one extreme, if one ω_i is close to 1 and the remaining ω_j are small, the sum $\|\sum_i \omega_i \hat{x}_i\| \sim 1$ regardless of n, and $\|\sum_i \omega_i \hat{x}_i\|$ cannot concentrate on 0 as $n \to \infty$. To see this in the statement of the Proposition, observe that in this case Ω and $\operatorname{Var}\omega_i$ approach their maximum values of $\Omega \sim 1$ and $1 + n^2 \operatorname{Var}\omega_i \sim n$, the n's in numerator and denominator cancel, and the exponent no longer depends on n at all.

At the other extreme, if the ω_i are all equal, Ω and $\operatorname{Var}\omega_i$ are minimized: $\Omega = 1/n$ and $\operatorname{Var}\omega_i = 0$. In this case, the denominator in the exponent does not depend on n and $\|\sum \omega_i \hat{x}_i\|$ concentrates on 0 as fast as possible. We can compare this result to that of Khoi [17], who showed in a different sense that the equilateral polygons are the "most flexible" of all the fixed edgelength polygons.

In the middle, if the ω_i are variable, but the number of comparably large ω_i increases, Ω and $\operatorname{Var}\omega_i$ act to slow the rate of concentration, but they do not stop it: $\|\sum_i \omega_i \hat{x}_i\|$ still concentrates on 0 as $n \to \infty$.

4.4. A probabilistic bound on $\lambda_{\min}(\mathcal{H}\operatorname{Ad}_{\boldsymbol{x}}(\vec{0})) = \lambda_{\min}\left(I - \sum \omega_i \hat{x}_i \hat{x}_i^T\right)$

We now want to bound the lowest eigenvalue of the Hessian of $\operatorname{Ad}_{\boldsymbol{x}}$ at the origin. Again using Lemma 5, we see that $\mathcal{H}\operatorname{Ad}_{\boldsymbol{x}}(\vec{0}) = I - \sum \omega_i \hat{x}_i \hat{x}_i^T$, where the quantities being summed are outer products of the vectors $\hat{x}_i$. That is, they are the symmetric, positive semidefinite projection matrices which project to the lines spanned by the $\hat{x}_i$. We now show

Proposition 15. *If we have n points $\hat{x}_i$ sampled independently and uniformly from S^{d-1}, and n weights $0 \leq \omega_i$ with $\sum_i \omega_i = 1$ and $\Omega = \max_i \omega_i$, then for any $t > 0$*

$$\mathcal{P}\left(\lambda_{\min}\left(I - \sum \omega_i \hat{x}_i \hat{x}_i^T\right) > \frac{d-1}{d} - t\right) \leq d \exp\left(-\frac{d}{d-1} \cdot \frac{3dnt^2}{2nt\Omega d + 6(1 + n^2 \operatorname{Var}\omega_i)}\right).$$

If the ω_i are all equal (the polygon is equilateral), this simplifies to

$$\mathcal{P}\left(\lambda_{\min}\left(I - \frac{1}{n}\sum \hat{x}_i \hat{x}_i^T\right) > \frac{d-1}{d} - t\right) \leq d\exp\left(-\frac{d}{d-1} \cdot \frac{3dnt^2}{2td+6}\right).$$

Proof. The statement is similar to the statement of Proposition 14, so it should not be surprising that this also follows from a Bernstein inequality, this time for matrices [22, Theorem 1.4]: suppose $X_1, \ldots, X_n$ are independent random symmetric $d \times d$ matrices, $\mathcal{E}(X_i) = 0$, $\lambda_{\max}(X_i) \leq b$, the "matrix variance" of each X_i is given by $\sigma_i^2 = \mathcal{E}(X_i^2)$, and $X = \sum X_i$ (with "scalar variance" $\sigma^2 = \|\sum \sigma_i^2\|$). Then for any $t > 0$,

$$\mathcal{P}\left(\lambda_{\max}(X) \geq t\right) \leq d\exp\left(-\frac{t^2}{2\sigma^2\left(1 + \frac{bt}{3\sigma^2}\right)}\right) \tag{5}$$

We will set $X_i = \omega_i\left(\hat{x}_i \hat{x}_i^T - \frac{1}{d}I_d\right)$. These are clearly symmetric $d \times d$ matrices.

We now prove $\mathcal{E}(X_i) = 0$. Since the $\hat{x}_i$ are uniformly sampled on S^{d-1}, their distribution is $O(d)$-invariant. This means we can first average $\hat{x}_i$ over any subgroup of $O(d)$ without changing $\mathcal{E}\left(\hat{x}_i \hat{x}_i^T\right)$. We'll choose the orthotope group of all 2^d possible diagonal matrices D with $D_{ii} = \pm 1$. For any vector $\vec{v} \in \mathbb{R}^d$:

$$\frac{1}{2^d}\left(\sum_D (D\vec{v})(D\vec{v})^T\right)_{ij} = \frac{1}{2^d}\sum_D D_{ii}D_{jj}v_i v_j$$

Now for each of the 4 possible combinations of signs $D_{ii} = \pm 1$ and $D_{jj} = \pm 1$, there are the same number 2^{d-2} of elements of the orthotope group with these signs. If $i \neq j$, two products are $+1$ and two are -1 and the terms cancel. If $i = j$ all the products are the same. Thus the average matrix $\frac{1}{2^d}\sum_D (D\vec{v})(D\vec{v})^T$ is a diagonal matrix with entries v_i^2.

Since the expectation of the square of a coordinate of a randomly distributed unit vector on S^{d-1} was computed in the proof of Proposition 14 to be $1/d$, we have $\mathcal{E}\left(\hat{x}_i \hat{x}_i^T\right) = \frac{1}{d}I_d$, proving that $\mathcal{E}(X_i) = 0$.

We now prove that $\lambda_{\max}(X_i) \leq \Omega\frac{d-1}{d}$. For any matrix A, the eigenvalues of $A + kI_d$ are simply k added to the eigenvalues of A (cf. [12, Theorem 2.4.8.1]). So

$$\lambda_{\max}(X_i) = \omega_i\left(\lambda_{\max}(\hat{x}_i \hat{x}_i^T) - \frac{1}{d}\right) = \omega_i\frac{d-1}{d} \leq \Omega\frac{d-1}{d}$$

since the largest eigenvalue of a projection matrix like $\hat{x}_i \hat{x}_i^T$ is 1.

Next, we want to show that $\sigma_i^2 = \mathcal{E}(X_i^2) = \omega_i^2\frac{d-1}{d^2}I_d$. A direct computation reveals

$$X_i^2 = \omega_i^2\left(\left(1 - \frac{2}{d}\right)\hat{x}_i \hat{x}_i^T + \frac{1}{d^2}I_d\right)$$

and the result follows from our previous computation that $\mathcal{E}\left(\hat{x}_i\hat{x}_i^T\right) = \frac{1}{d}I_d$. Summing the σ_i^2 and taking the operator norm, we get

$$\sigma^2 = \frac{d-1}{d^2}\sum\omega_i^2.$$

Plugging b and σ into (5) yields a bound on the probability that $\lambda_{\max}(X) > t$ or, since $\lambda_{\max}(X) = \lambda_{\max}(\sum\omega_i\hat{x}_i\hat{x}_i^T) - \frac{1}{d}$, that $\lambda_{\max}(\sum\omega_i\hat{x}_i\hat{x}_i^T) > \frac{1}{d} + t$. This completes the proof. $\square$

We note that this concentration inequality is better than Proposition 14: there is an extra factor of d in the numerator which means that the concentration gets faster as d increases. The effect of variable edgelengths is to slow (or stop) the concentration, just as in Proposition 14; the same comments on the role of Ω and $\mathrm{Var}\,\omega_i$ apply here.

4.5. A bound on the change in the radial second derivative

For any point $\vec{s} \in \mathbb{R}^d$, the second derivative of $\mathrm{Ad}_{\boldsymbol{x}}$ along the ray through $\vec{s}$ is given by evaluating the Hessian as a quadratic form on the vector $\vec{s}$ itself. Our last proposition gave us a lower bound on the result at the origin; we now show that this can't change too fast as we move away from the origin.

Proposition 16. *For $\|\vec{s}\| < 1$ we have*

$$\frac{\langle\mathcal{H}\,\mathrm{Ad}_{\boldsymbol{x}}(\vec{s})\vec{s},\vec{s}\rangle}{\langle\vec{s},\vec{s}\rangle} - \frac{\langle\mathcal{H}\,\mathrm{Ad}_{\boldsymbol{x}}(0)\vec{s},\vec{s}\rangle}{\langle\vec{s},\vec{s}\rangle} \geq -\|\vec{s}\|\frac{6 + \|\vec{s}\| + \|\vec{s}\|^2}{(1 - \|\vec{s}\|)^3}.$$

Since the fraction at right is increasing in $\|\vec{s}\|$, we can easily simplify the statement given a better upper bound on $\|\vec{s}\|$. In particular, for $\|\vec{s}\| < 1/50$, the right-hand side $\geq -7\|\vec{s}\|$.

Proof. Using Lemma 5, we see that

$$\langle\mathcal{H}\,\mathrm{Ad}_{\boldsymbol{x}}(\vec{s})\vec{s},\vec{s}\rangle - \langle\mathcal{H}\,\mathrm{Ad}_{\boldsymbol{x}}(0)\vec{s},\vec{s}\rangle =$$

$$\left(\left(\sum\frac{\omega_i}{\|\hat{x}_i - \vec{s}\|}\right) - 1\right)\langle\vec{s},\vec{s}\rangle - \sum\omega_i\left(\frac{\langle\hat{x}_i - \vec{s},\vec{s}\rangle^2}{\|\hat{x}_i - \vec{s}\|^3} - \langle\hat{x}_i,\vec{s}\rangle^2\right).$$

Using the estimates $1 - \|\vec{s}\| \leq \|\hat{x}_i - \vec{s}\| \leq 1 + \|\vec{s}\|$ and recalling that $\sum\omega_i = 1$, we can underestimate the right hand side by

$$-\frac{\|\vec{s}\|^3}{1 + \|\vec{s}\|} - \sum\omega_i\left(\frac{\langle\hat{x}_i - \vec{s},\vec{s}\rangle^2}{(1 - \|\vec{s}\|)^3} - \langle\hat{x}_i,\vec{s}\rangle^2\right) \geq -\frac{\|\vec{s}\|^3}{1 + \|\vec{s}\|} - \frac{5\|\vec{s}\|^3 + \|\vec{s}\|^4 + \|\vec{s}\|^5}{(1 - \|\vec{s}\|)^3}$$

where the second part follows from finding a common denominator, expanding, and cancelling, using Cauchy–Schwartz carefully to underestimate the inner product terms as needed. Observing that $1+\|\vec{s}\| > 1 > (1-\|\vec{s}\|)^3$ allows us to underestimate $-\|\vec{s}\|^3/{1+\|\vec{s}\|} \geq -\|\vec{s}\|^3/(1-\|\vec{s}\|)^3$, completing the proof. $\qquad\square$

4.6. Bounding the norm of the geometric median

We are now in a position to bound the norm of the geometric median! This will proceed in two stages: first, we'll use the Poincaré–Hopf index theorem to show that $\|\vec{\mu}\| < 1/50$ under certain hypotheses. Then we can immediately bootstrap to get a sharper bound.

Proposition 17. *If* $\|\sum \omega_i \hat{x}_i\| = \|\nabla \mathrm{Ad}_{\boldsymbol{x}}(\vec{0})\| < 5/1000$, $\lambda_{\min}(\mathcal{H}\,\mathrm{Ad}_{\boldsymbol{x}}(\vec{0})) > \frac{d-1}{d} - \frac{1}{100}$, *and* $d \geq 2$, *then* $\|\vec{\mu}\| \leq 1/50$.

Proof. Given our hypothesis on $\lambda_{\min}$ of the Hessian, we know that the $\hat{x}_i$ are not all colinear. This means that $\vec{\mu}$ is the unique point inside S^{d-1} where the vector field $\nabla \mathrm{Ad}_{\boldsymbol{x}}$ vanishes. We will now show that $\nabla \mathrm{Ad}_{\boldsymbol{x}}(\vec{y})$ has a zero inside the sphere of radius $1/50$; by uniqueness, this point must be the geometric median.

Along any ray from the origin, we may restrict $\mathrm{Ad}_{\boldsymbol{x}}$ to a scalar function $\mathrm{Ad}_{\boldsymbol{x}}(z)$. Using Proposition 16, on the interval $[0, 1/50]$ our hypotheses imply

$$\mathrm{Ad}'_{\boldsymbol{x}}(0) \geq -\frac{5}{1000} \quad \text{and} \quad \mathrm{Ad}''_{\boldsymbol{x}}(z) \geq \frac{1}{2} - \frac{1}{100} - \frac{7}{50} = \frac{7}{20}.$$

By Taylor's theorem, there is some $z_* \in [0, 1/50]$ so that

$$\mathrm{Ad}'_{\boldsymbol{x}}(1/50) = \mathrm{Ad}'_{\boldsymbol{x}}(0) + \frac{1}{50}\mathrm{Ad}''_{\boldsymbol{x}}(z_*) \geq -\frac{5}{1000} + \frac{1}{50} \cdot \frac{7}{20} = \frac{2}{1000} > 0.$$

This means that the directional derivative of $\mathrm{Ad}_{\boldsymbol{x}}$ in the outward direction is positive on the boundary of the sphere of radius $1/50$, or that $\nabla \mathrm{Ad}_{\boldsymbol{x}}(\vec{y})$ points outward on this sphere. In particular, this implies that the vector field has index 1 on the sphere, and so by the Poincaré–Hopf index theorem must vanish at some point inside the sphere. $\qquad\square$

We can now prove our main theorem.

Theorem 18. *If we have n points $\hat{x}_i$ sampled uniformly on S^{d-1} ($d \geq 2$), n weights $\omega_i > 0$ so that $\sum \omega_i = 1$, and $\max \omega_i = \Omega$, then for any $t < 5/1000$ we have*

$$\mathcal{P}\left(\|\vec{\mu}\| < \frac{t}{\frac{d-1}{d} - \frac{3}{20}}\right) \geq 1 - 2d \exp\left(-\frac{3nt^2}{2nt\Omega\sqrt{d} + 6(1 + n^2 \operatorname{Var}\omega_i)}\right). \tag{6}$$

If all the ω_i are equal (the polygon is equilateral)

$$\mathcal{P}\left(\|\vec{\mu}\| < \frac{t}{\frac{d-1}{d} - \frac{3}{20}}\right) \geq 1 - 2d \exp\left(-\frac{3nt^2}{2\sqrt{d}t + 6}\right).$$

For $d = 3$, we have the further simplification

$$\mathcal{P}\left(\|\vec{\mu}\| < t\right) \geq 1 - 6 \exp\left(-nt^2/9\right).$$

Proof. We first define two random events: $\lambda_{\min}(\mathcal{H} \operatorname{Ad}_{\boldsymbol{x}}(\vec{0})) > \frac{d-1}{d} - \frac{1}{100}$ (event A) and $\left\|\nabla \operatorname{Ad}_{\boldsymbol{x}}(\vec{0})\right\| < t < 5/1000$ (event B), which will happen for some choices of $\hat{x}_i$. Suppose both events occur.

As in Proposition 17, we restrict $\operatorname{Ad}_{\boldsymbol{x}}$ to a scalar function $\operatorname{Ad}_{\boldsymbol{x}}(z)$ on a ray; this time, the ray is assumed to pass through $\vec{\mu}$. By Taylor's theorem, if we evaluate at $z = \|\vec{\mu}\|$, there is some $0 \leq z_* \leq \|\vec{\mu}\|$ so that

$$0 = \operatorname{Ad}'_{\boldsymbol{x}}(\|\vec{\mu}\|) = \operatorname{Ad}'_{\boldsymbol{x}}(0) + \|\vec{\mu}\| \operatorname{Ad}''_{\boldsymbol{x}}(z_*). \tag{7}$$

Since we are assuming $A \wedge B$, the hypotheses of Proposition 17 are satisfied and $\|\vec{\mu}\| < 1/50$. In turn, this means that Proposition 16 holds at z_*, and

$$\operatorname{Ad}''_{\boldsymbol{x}}(z_*) \geq \operatorname{Ad}''_{\boldsymbol{x}}(0) - 7/50 \geq \lambda_{\min}(\mathcal{H} \operatorname{Ad}_{\boldsymbol{x}}(\vec{0})) - 7/50.$$

Since A, we have $\operatorname{Ad}''_{\boldsymbol{x}}(z_*) > \frac{d-1}{d} - \frac{3}{20}$. As before, since $\operatorname{Ad}'_{\boldsymbol{x}}(\vec{0})$ is a directional derivative, it satisfies $\operatorname{Ad}'_{\boldsymbol{x}}(0) \geq -\|\nabla \operatorname{Ad}_{\boldsymbol{x}}(\vec{0})\| > -t$. We can plug both estimates into (7) and solve for $\|\vec{\mu}\|$, obtaining

$$\|\vec{\mu}\| < \frac{t}{\frac{d-1}{d} - \frac{3}{20}}.$$

If we call this event C, we have shown that $A \wedge B \implies C$, and hence that $\mathcal{P}(C) \geq \mathcal{P}(A \wedge B)$. This means that

$$\mathcal{P}(\neg C) \leq \mathcal{P}(\neg(A \wedge B)) = \mathcal{P}(\neg A \vee \neg B) \leq \mathcal{P}(\neg A) + \mathcal{P}(\neg B). \tag{8}$$

Now $\mathcal{P}(\neg A)$ was bounded above in Proposition 15, while $\mathcal{P}(\neg B)$ was bounded above in Proposition 14. We now compare these upper bounds, noting that we have chosen $t_* = \frac{1}{100}$ in the statement of Proposition 15 while the t in Proposition 14 is smaller – less than $\frac{5}{1000} = \frac{1}{200}$. The bounds are

$$d \exp\left(-\frac{d}{d-1} \cdot \frac{3dnt_*^2}{2nt_*\Omega d + 6(1 + n^2 \operatorname{Var} \omega_i)}\right) \quad \text{and} \quad d \exp\left(-\frac{3nt^2}{2nt\Omega\sqrt{d} + 6(1 + n^2 \operatorname{Var} \omega_i)}\right).$$

Of course, it suffices to compare the absolute values of the fractions inside the exponential functions (since both are negative). We can simplify the comparison by rewriting these as

$$\frac{d}{d-1} \cdot \frac{3nt_*}{2n\Omega + \frac{6(1+n^2 \operatorname{Var}\omega_i)}{dt_*}} \quad \text{and} \quad \frac{3nt}{2n\Omega\sqrt{d} + \frac{6(1+n^2 \operatorname{Var}\omega_i)}{t}}.$$

It is now evident that if we compare the right fraction with the second fraction on the left, the numerator on the right is smaller and each term in the denominator is larger (recall $t < t_*$). Multiplying by $\frac{d}{d-1} > 1$ makes the left hand side even larger. Restoring the minus sign reverses this conclusion, and we see that our bound on $\mathcal{P}(\neg B)$ is larger than our bound on $\mathcal{P}(\neg A)$, as claimed. Returning this conclusion to (8), we see $\mathcal{P}(\neg C) \le 2\,\mathcal{P}(\neg B)$, which is the first statement of the Theorem.

The simplification when all the $\omega_i = 1/n$ is an immediate consequence. To simplify to $d = 3$, we observe that $\frac{t}{2/3 - 3/20} = \frac{60}{31}t$; substituting $t \to \frac{31}{60}t$ on the right hand side yields an expression in the form $1 - 6\exp(-f(t)nt^2)$, where $f(t)$ is a rational function bounded below by $1/9$ for $t \in [0, 5/1000]$. $\qquad\square$

We now make a few remarks. First, if you carefully examine Proposition 16, the lower bound on $\operatorname{Ad}''_{\boldsymbol{x}}(z)$ improves as $z \to 0$. One can wring some extra information out of this, but the improvement in the final bound is minimal. Similarly, it is clear that one could set $t_* < \frac{1}{100}$ in our bound on $\mathcal{P}(\neg A)$ without losing the conclusion, as long as $t_* < t$. Again, this does not significantly improve things.

5. DISTANCES AND ANGLES

We now want to restate our main Theorem 18 in terms of the chordal and max-angular distance from a random arm to the nearest closed polygon using Proposition 12.

Corollary 19. *If we have n points $\hat{x}_i$ sampled uniformly on S^{d-1} $(d \ge 2)$, n weights $\omega_i > 0$ so that $\sum \omega_i = 1$, and $\max \omega_i = \Omega$, then for any $t < 5/1000 \cdot 1/\sqrt{2\sum\omega_i^2}$ we have*

$$\mathcal{P}\left(d_{chordal}(\boldsymbol{x}, \operatorname{Pol}(n, d, w)) < \frac{t}{\frac{d-1}{d} - \frac{3}{20}}\right) \ge 1 - 2d\exp\left(\frac{-3t^2}{3 + t\Omega\sqrt{\frac{2dn}{1+n^2\operatorname{Var}\omega_i}}}\right).$$

If all the ω_i are equal (the polygon is equilateral), for $t < 5/1000 \cdot \sqrt{n/2}$ we have

$$\mathcal{P}\left(d_{chordal}(\boldsymbol{x}, \operatorname{Pol}(n, d, 1)) < \frac{t}{\frac{d-1}{d} - \frac{3}{20}}\right) \ge 1 - 2d\exp\left(\frac{-t^2}{1 + \frac{\sqrt{d}}{600}}\right).$$
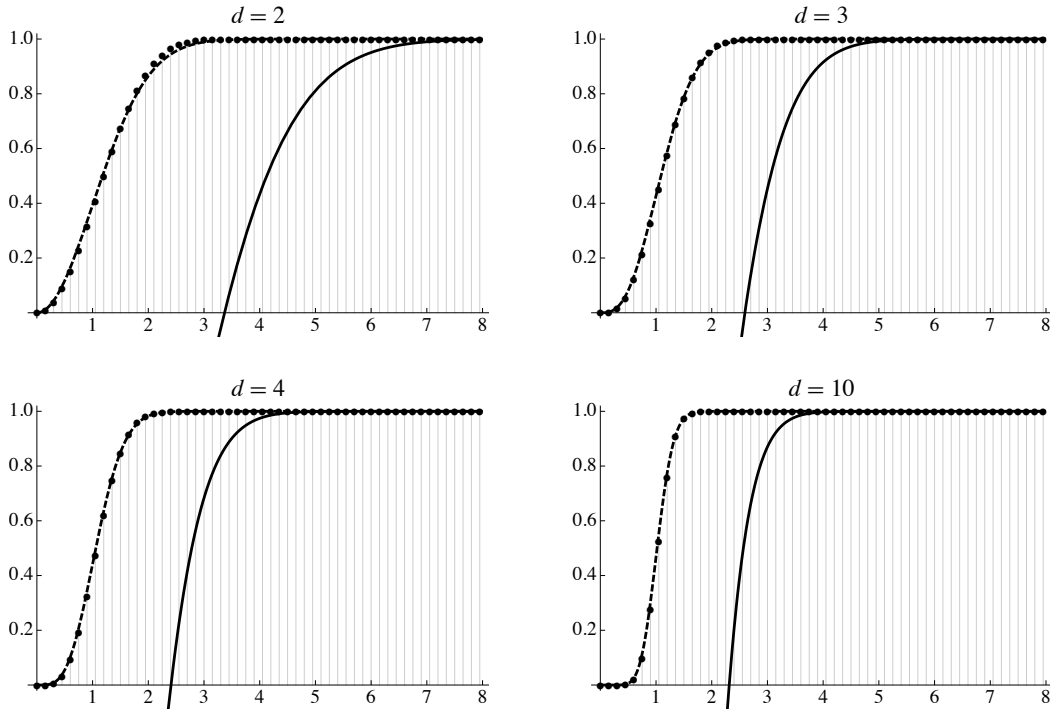
FIG. 3: For $d = 2, 3, 4, 10$, we generated 250,000 random elements of $\mathrm{Arm}(10, d, 1)$. The plots show the implied bound from Corollary 19 (solid), the empirical CDF of chordal distance to closure for those samples which were median-closeable (dots), and the CDF of the Nakagami $\left(\frac{d}{2}, \frac{d}{d-1}\right)$ distribution (dashed) given by Proposition 10 for the large-n limit (which is only slightly different, even though $n = 10$ is quite small). Though the hypotheses of Corollary 19 are only satisfied when $t < \frac{5}{1000}\sqrt{5} \approx 0.01118$, the data strongly suggests that the bound is valid on a much larger range. We see from the plots that the bound cannot be dramatically improved.

In dimension 3, this simplifies (again, for $t < {}^5/_{1000} \cdot \sqrt{n/_2}$), as

$$\mathcal{P}\left(d_{chordal}(\boldsymbol{x}, \mathrm{Pol}(n, 3, 1)) < t\right) \geq 1 - 6\exp\left(-t^2/4\right).$$

The problem with Corollary 19 is that the hypotheses (on t) are disappointingly restrictive: for $\mathrm{Arm}(n, 3, 1)$, we need $n > 538,519$ to extend the domain of t to the point where the right-hand side becomes positive! On the other hand, numerical experiments (Figure 3) comparing our bounds to experimental data and to the large-n Nakagami distribution proved in Proposition 10 show that the conclusions of Corollary 19 cannot be made much stronger. Further, these experiments suggest that, at least in the equilateral case, one should be able to entirely remove the upper bound on t– we leave this as

Conjecture 20. The conclusions of Corollary 19 hold for any $t > 0$.

We now proceed to prove Corollary 19.

Proof of Corollary 19. Proposition 12 tells us $d_{\text{chordal}}(\boldsymbol{x}, \text{Pol}(n, d, w)) < \sqrt{2 \sum \omega_i^2}\, \|\vec{\mu}\|$, so to get a bound on the probability that $d_{\text{chordal}}(\boldsymbol{x}, \text{Pol}(n, d, w)) < \frac{t}{\frac{d-1}{d} - \frac{3}{20}}$ we need to make the substitution $t \to t\sqrt{2 \sum \omega_i^2}$ on the right hand side of (6). Recalling that Lemma 13 shows $\sum \omega_i^2 = \frac{1 + n^2 \operatorname{Var} \omega_i}{n}$ and carefully simplifying yields the first result.

For the second result, it follows immediately from the assumption that $\omega_i = \frac{1}{n}$ that the first result simplifies to

$$\mathcal{P}\left(d_{\text{chordal}}(\boldsymbol{x}, \text{Pol}(n, d, 1)) < \frac{t}{\frac{d-1}{d} - \frac{3}{20}}\right) \geq 1 - 2d \exp\left(\frac{-3t^2}{3 + t\sqrt{\frac{2d}{n}}}\right).$$

Using our upper bound on t, we see that the right hand side obeys

$$1 - 2d \exp\left(\frac{-3t^2}{3 + t\sqrt{\frac{2d}{n}}}\right) > 1 - 2d \exp\left(\frac{-3t^2}{3 + \frac{\sqrt{d}}{200}}\right)$$

which immediately implies the second result.

For the third result, we simplify the fraction on the left hand side and substitute $t \to \frac{31}{60} t$ as we did above in the simplification of Theorem 18; the complicated constant that results as the coefficient of t^2 in the exponent is slightly less that $-1/4$. $\qquad\square$

The statements for the maximum angular change in edge direction are similar, but somewhat easier to prove because the relationship between $\|\vec{\mu}\|$ and the max-angular distance is simpler.

Corollary 21. *If we have n points $\hat{x}_i$ sampled uniformly on S^{d-1} ($d \geq 2$), n weights $\omega_i > 0$ so that $\sum \omega_i = 1$, and $\max \omega_i = \Omega$, then for any $t < 5/1000$ we have*

$$\mathcal{P}\left(d_{\text{max-angular}}(\boldsymbol{x}, \text{Pol}(n, d, w)) < \frac{t}{\frac{d-1}{d} - \frac{3}{20}}\right) \geq 1 - 2d \exp\left(-\frac{13nt^2}{9nt\Omega\sqrt{d} + 30(1 + n^2 \operatorname{Var} \omega_i)}\right).$$

If all the ω_i are equal (the polygon is equilateral), for $t < 5/1000$ we have

$$\mathcal{P}\left(d_{\text{max-angular}}(\boldsymbol{x}, \text{Pol}(n, d, 1)) < \frac{t}{\frac{d-1}{d} - \frac{3}{20}}\right) \geq 1 - 2d \exp\left(\frac{-26nt^2}{60 + \frac{9\sqrt{d}}{100}}\right).$$

In dimension 3, this simplifies (again, for $t < 5/1000$), as

$$\mathcal{P}\left(d_{\text{max-angular}}(\boldsymbol{x}, \text{Pol}(n, 3, 1)) < t\right) \geq 1 - 6 \exp\left(-nt^2/9\right).$$

Proof. We know from Proposition 12 that $d_{\text{max-angular}}(\boldsymbol{x}, \text{Pol}(n, d, w)) < \arcsin \|\vec{\mu}\|$. Since we're only going to apply this bound when $\|\vec{\mu}\| < \frac{t}{\frac{d-1}{d} - \frac{3}{20}} < \frac{1}{70}$ (since $d \geq 2$ and $t \leq 5/1000$), we can safely make the overestimate $\arcsin \|\vec{\mu}\| \leq \frac{14}{13} \|\vec{\mu}\|$.

Substituting $t \to \frac{13}{14} t$ in (6) leads us to replace 3 by $3(\frac{13}{14})^2 < 2.6$ in the coefficient of nt^2 in the numerator and 2 by $2(13/14) > 1.8$ in the coefficient of t in the denominator. Simplifying gives us the first statement.

To reach the second statement, we first observe that $\omega_i = \frac{1}{n}$ means $\text{Var}\,\omega_i = 0$ and $\Omega = \frac{1}{n}$. Substituting these into the first statement (and overestimating the t in the denominator by $5/1000$) yields the result.

Finally, the third statement (as in the proof of Corollary 19) requires us to substitute $t \to \frac{31}{60} t$ to simplify the left-hand side. The resulting complicated coefficient of nt^2 on the right-hand side is about $-0.115521 < -1/9$. $\qquad\square$

6. DISCUSSION

From Corollary 21 we see that closing a random arm is unlikely to change any edge very much. In particular, we should expect local features to be preserved by closure, as in the case of the local trefoil knot shown in Figure 4. This suggests that closing up an arm is unlikely to destroy any local knots: in other words, the probability of local knotting in the standard measure on $\text{Arm}(n, 3, w)$ should be essentially the same as the probability of local knotting in the pushforward measure on $\text{Pol}(n, 3, w)$ via the map $\boldsymbol{x} \mapsto \text{gmc}(\boldsymbol{x})$.

Of course, this map is not defined on all of $\text{Arm}(n, d, w)$, but we know from Theorem 18 that it is defined on all but an exponentially small fraction of $\text{Arm}(n, d, w) = \prod S^{d-1}(w_i)$; pushing forward the restriction of the product measure to the domain of gmc produces what we'll call the *pushforward measure* on $\text{Pol}(n, d, w)$. On the other hand, the standard probability measure on $\text{Pol}(n, d, w)$ is simply the volume measure induced by the Riemannian metric it inherits from $\text{Arm}(n, d, w)$. Since we've seen in Corollaries 19 and 21 that almost all of $\text{Arm}(n, d, w)$ is within a fixed distance of $\text{Pol}(n, d, w)$, it is reasonable to expect that this pushforward measure is close to the standard measure.

Indeed, this seems to be true. Rayleigh [20] showed that the distribution of end-to-end distances in a random element of $\text{Arm}(n, 3, 1)$ is

$$\Phi_n(\ell) = \frac{1}{2\pi^2\ell} \int_0^\infty x \sin \ell x \operatorname{sinc}^n x \, dx.$$

We note that a closed form for Φ_n is classical (see [14, 2.181]). Since a random closed polygon is formed from two random arms, conditioned on the hypothesis that their end-to-end distances are the same, the pdf of the length of the chord connecting vertices 0 and k in an polygon of n edges
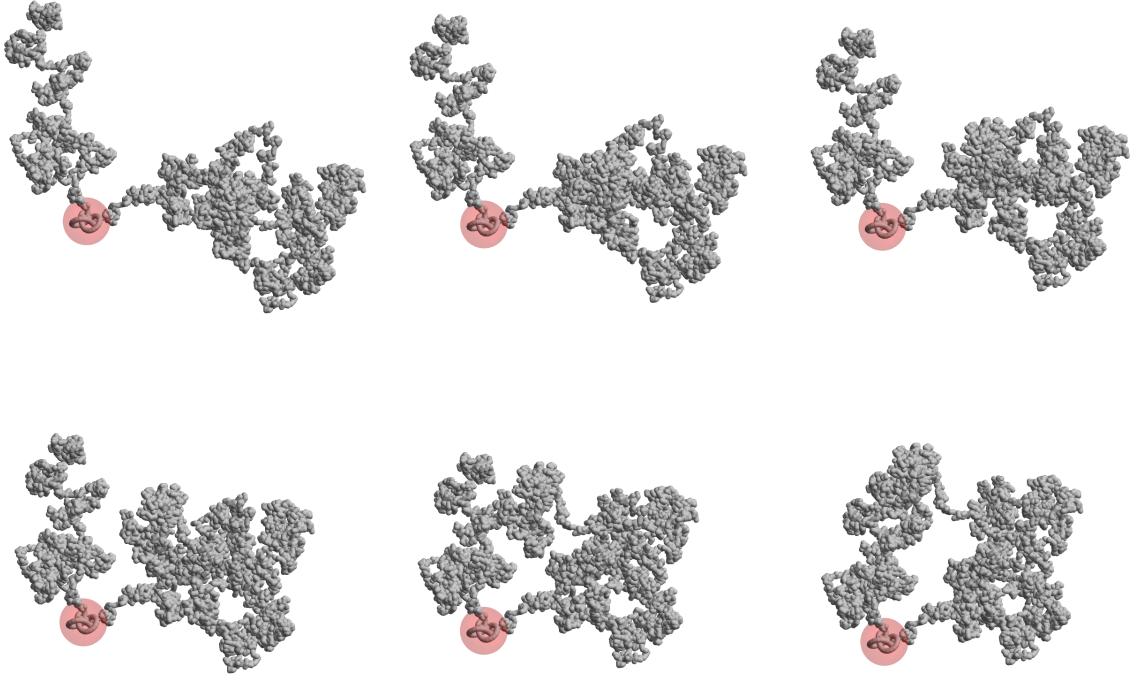
FIG. 4: A 10,000 step equilateral arm in $\mathbb{R}^3$ containing a small trefoil (top left) and its geometric median closure (bottom right). The intermediate images show equally-spaced points along the geodesic between the arm and its closure in $\mathrm{Arm}(10,000, 3, 1)$. The failure to close of the arm is ≈ 101.118 and the geometric median has norm $\|\vec{\mu}\| \approx 0.0151318$. The chordal distance between the arm and its closure is ≈ 1.23696 and $d_{\text{max-angular}} \approx 0.0151324$, which agrees with the bound $\arcsin \|\vec{\mu}\|$ to eleven decimal places.

turns out to be given by

$$\mathrm{Chord}_{n,k}(\ell) = \frac{1}{C(n)} 4\pi\ell^2 \Phi_k(\ell)\Phi_{n-k}(\ell).$$

where the factor of $4\pi\ell^2$ comes from the fact that vertex k lies on a sphere of radius ℓ and $C(n)$ is the volume of polygon space (which is known; see [2] for an identification between polygon space and a certain polytope which yields an explicit, though complicated, formula for $C(n)$).

Therefore, the extent to which the distributions of the chordlengths match $\mathrm{Chord}_{n,k}$ gives a sense of how close a given distribution on $\mathrm{Pol}(n, 3, w)$ is to the standard one. For $n = 4$ and 5, we can see in Figure 5 that the pushforward measure from $\mathrm{Arm}(n, 3, 1)$ is not particularly close to the standard measure. However, as n increases these statistics cannot distinguish between the pushforward measure and the standard measure; see Figure 6.
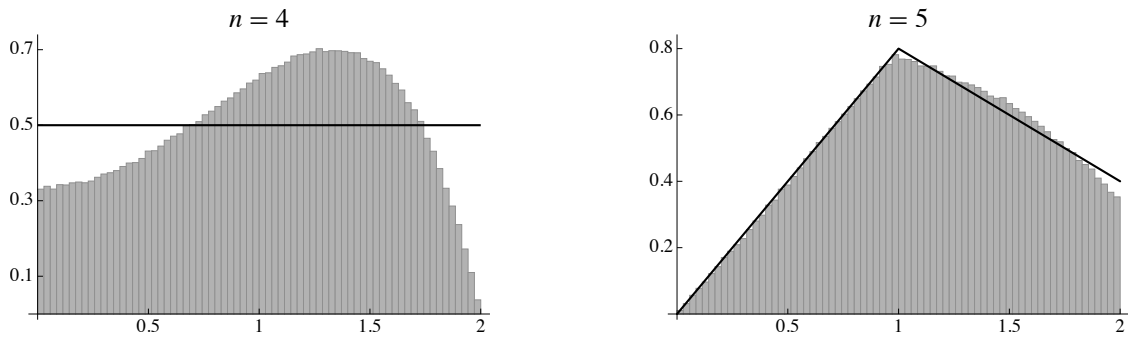
FIG. 5: For $n = 4, 5$, we generated 1,000,000 random equilateral n-edge arms in $\mathbb{R}^3$, computed their geometric median closures (when they existed), and then computed the distance from the first to the third vertex in the resulting closed n-gon. The histograms show the resulting distributions of chordlengths as well as the density of the chordlength for the standard distribution on $\mathrm{Pol}(n, 3, 1)$. Closure failed for 2474 quadrilaterals and for 117 pentagons.
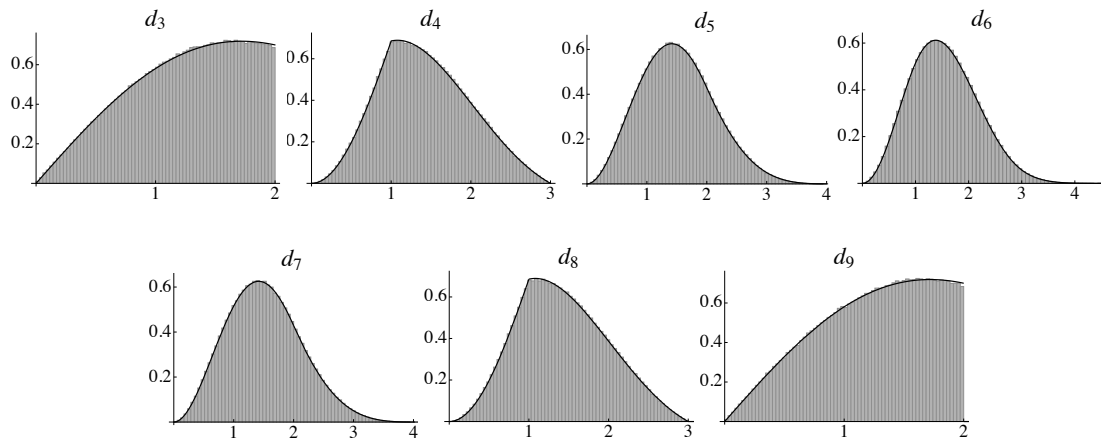

FIG. 6: We generated 1,000,000 random equilateral 10-edge arms in $\mathbb{R}^3$. All 1,000,000 had geometric median closures, and these are the histograms of distances from the first vertex to the ith vertex in the resulting closed 10-gons, along with the chordlength densities for the standard measure on equilateral 10-gons.

Conjecture 22. As $n \to \infty$, the pushforward measure from $\mathrm{Arm}(n, d, w)$ to $\mathrm{Pol}(n, d, w)$ converges to the standard measure.

Assuming the truth of this conjecture implies that, at least for large n, random elements of $\mathrm{Pol}(n, d, w)$ look essentially like geometric median closures of random elements of $\mathrm{Arm}(n, d, w)$. Since Corollary 21 implies that individual edges are practically unchanged by closure, this would mean that *all* local phenomena happen at essentially the same rate in $\mathrm{Arm}(n, d, w)$ and

$\mathrm{Pol}(n, d, w)$.

When $d = 3$, a particularly important local phenomenon is that of local knotting. Say that a subsegment ς of $\boldsymbol{x} \in \mathrm{Arm}(n, 3, w)$ is an *r-local knot* if it only intersects the boundary of a ball B of radius r at its endpoints and (B, ς) forms a knotted ball-arc pair. Let $K^{\mathrm{Arm}}(n, w, k, r)$ be the probability that a length-k arc of a random element of $\mathrm{Arm}(n, 3, w)$ is an r-local knot, and similarly for $K^{\mathrm{Pol}}(n, w, k, r)$.

Conjecture 23. For small r and large n and for $k \ll n$, $K^{\mathrm{Arm}}(n, w, k, r) \simeq K^{\mathrm{Pol}}(n, w, k, r)$.

Acknowledgements

Appendix: Proof of the hypergeometric formula for the expected distance to the sphere

Recall that $\mathrm{Ed}(\vec{y})$ is the expected distance from the point $\vec{y} \in \mathbb{R}^d$ to the unit sphere. Since it is spherically symmetric, $\mathrm{Ed}(\vec{y})$ depends only on $\|\vec{y}\|$.

Proposition 24. $\mathrm{Ed}(\vec{y})$ *is given as a function of* $r = \|\vec{y}\|$ *by*

$$\mathrm{Ed}(r) = {}_2F_1\left(-\frac{1}{2}, \frac{1-d}{2}; \frac{d}{2}; r^2\right).$$

We can compute $\mathrm{Ed}(r)$ by evaluating at $\vec{y} = (0, \ldots, 0, r)$:

$$\mathrm{Ed}(r) = \mathrm{Ed}((0, \ldots, 0, r)) = \frac{1}{\mathrm{Vol}\, S^{d-1}} \int_{\hat{x} \in S^{d-1}} \|\hat{x} - (0, \ldots, 0, r)\| \; \mathrm{dVol}_{S^{d-1}}$$

$$= \frac{1}{\mathrm{Vol}\, S^{d-1}} \int_{\hat{x} \in S^{d-1}} \sqrt{1 + r^2 - 2x_d r} \; \mathrm{dVol}_{S^{d-1}}, \quad \text{(A.1)}$$

where the integrand only depends on the last coordinate x_d of the point on S^{d-1}. Then the formula for Ed will follow from a more general formula for functions on the sphere which only depend on a single coordinate:

Lemma 25. *Suppose $\phi : S^{d-1} \to \mathbb{R}$ depends only on x_i; i.e. $\phi(\hat{x}) = \phi(x_i)$. Then*

$$\int_{\hat{x} \in S^{d-1}} \phi(\hat{x}) \; \mathrm{dVol}_{S^{d-1}} = \mathrm{Vol}\, S^{d-2} \int_{-1}^{1} \phi(x_i)(1 - x_i^2)^{\frac{d-3}{2}} \; dx_i.$$

Proof. Let $\pi_i : S^{d-1} \to \mathbb{R}$ be projection to the ith coordinate. Then the projection of $\nabla \pi_i$ to the tangent space of S^{d-1} has norm $\sqrt{1 - x_i^2}$, and hence the smooth coarea formula implies that

$$\int_{\hat{x} \in S^{d-1}} \phi(\hat{x}) \; \mathrm{dVol}_{S^{d-1}} = \int_{x_i \in [-1,1]} \int_{\vec{y} \in \pi_i^{-1}(x_i)} \frac{\phi(\vec{y})}{\sqrt{1 - x_i^2}} \; \mathrm{dArea}_{\pi_i^{-1}(x_i)} \; dx_i.$$

Since $\pi_i^{-1}(x_i)$ is a $(d - 2)$-dimensional sphere of radius $\sqrt{1 - x_i^2}$, which has area form $\left(1 - x_i^2\right)^{\frac{d-2}{2}} \mathrm{dArea}_{S^{d-2}}$, and since ϕ is constant on each level set, the above reduces to

$$\mathrm{Vol}\, S^{d-2} \int_{-1}^{1} \phi(x_i)(1 - x_i^2)^{\frac{d-3}{2}} \; dx_i$$

as desired. $\qquad\square$

Combining this result with (A.1) shows that

$$\mathrm{Ed}(r) = \frac{\mathrm{Vol}\, S^{d-2}}{\mathrm{Vol}\, S^{d-1}} \int_{-1}^{1} \sqrt{1 + r^2 - 2x_d r}(1 - x_d^2)^{\frac{d-3}{2}} \; dx_d, \quad \text{(A.2)}$$

which is the starting point of our derivation of the hypergeometric formula.

Proof of Proposition 24. Using $\mathrm{Vol}\, S^k = \frac{2\pi^{(k+1)/2}}{\Gamma((k+1)/2)}$ and the gamma function duplication formula $\Gamma(\zeta)\Gamma(\zeta + \frac{1}{2}) = 2^{1-2\zeta}\sqrt{\pi}\Gamma(2\zeta)$, we can write the ratio of sphere volumes as

$$\frac{\mathrm{Vol}\, S^{d-2}}{\mathrm{Vol}\, S^{d-1}} = \frac{2^{2-d}\Gamma(d-1)}{\Gamma(\frac{d-1}{2})\Gamma(\frac{d-1}{2})}.$$

Substituting this into (A.2) and completing the square inside the square root yields

$$(1+r)\frac{2^{2-d}\Gamma(d-1)}{\Gamma(\frac{d-1}{2})\Gamma(\frac{d-1}{2})}\int_{-1}^{1}\left(1-\frac{2r}{(1+r)^2}(1+x_d)\right)^{\frac{1}{2}}(1-x_d)^{\frac{d-3}{2}}(1+x_d)^{\frac{d-3}{2}}\,\mathrm{d}x_d.$$

Making the substitution $u=\frac{1+x_d}{2}$ produces

$$\mathrm{Ed}(r)=(1+r)\frac{2^{2-d}\Gamma(d-1)}{\Gamma(\frac{d-1}{2})\Gamma(\frac{d-1}{2})}\int_{-1}^{1}u^{\frac{d-3}{2}}(1-u)^{\frac{d-3}{2}}\left(1-\frac{4r}{(1+r)^2}u\right)^{\frac{1}{2}}\,\mathrm{d}x_d,$$

which is the standard integral representation of $(1+r)\,{}_2F_1\left(-\frac{1}{2},\frac{d-1}{2};d-1;\frac{4r}{(1+r)^2}\right)$. In turn, applying Kummer's quadratic transformation [10, 9.134.3] yields the desired formula

$$\mathrm{Ed}(r)={}_2F_1\left(-\frac{1}{2},\frac{1-d}{2};\frac{d}{2};r^2\right).$$

$\square$

[1] Rajendra Bhatia and Chandler Davis. A better bound on the variance. *The American Mathematical Monthly*, 107(4):353–357, 2000.

[2] Jason Cantarella and Clayton Shonkwiler. The symplectic geometry of closed equilateral random walks in 3-space. *The Annals of Applied Probability*, 26(1):549–596, 2016.

[3] Erik D Demaine and Joseph O'Rourke. *Geometric Folding Algorithms: Linkages, Origami, Polyhedra*. Cambridge University Press, Cambridge, 2007.

[4] Yuanan Diao. The knotting of equilateral polygons in $\mathbf{R}^3$. *Journal of Knot Theory and its Ramifications*, 4(2):189–196, 1995.

[5] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, release 1.0.18 of 2018–03–27. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[6] Adrien Douady and Clifford J Earle. Conformally natural extension of homeomorphisms of the circle. *Acta Mathematica*, 157(1-2):23–48, 1986.

[7] Zvi Drezner, Kathrin Klamroth, Anita Schöbel, and George O Wesolowsky. The Weber Problem. In Horst W Hamacher and Zvi Drezner, editors, *Facility Location*, pages 1–36. Springer-Verlag, Berlin, 2002.

[8] Devdatt P Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, Cambridge, 2009.

[9] Paul J Flory. *Statistical Mechanics of Chain Molecules*. Interscience Publishers, New York, 1969.

[10] Izrail S Gradshteyn and Iosif M Ryzhik. *Table of Integrals, Series, and Products*. Elsevier, Amsterdam, eighth edition, 2015.

[11] Shelby J Haberman. Concavity and estimation. *The Annals of Statistics*, 17(4):1631–1661, 1989.

[12] Roger A Horn and Charles R Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, second edition, 2013.

[13] Benjamin Howard, Christopher Manon, and John J Millson. The toric geometry of triangulated polygons in Euclidean space. *Canadian Journal of Mathematics*, 63(4):878–937, 2011.

[14] Barry D Hughes. *Random Walks and Random Environments: Volume 1: Random Walks*. Clarendon Press, Oxford, 1995.

[15] Michael Kapovich and John J Millson. The symplectic geometry of polygons in Euclidean space. *Journal of Differential Geometry*, 44(3):479–513, 1996.

[16] Michael Kapovich and John J Millson. Universality theorems for configuration spaces of planar linkages. *Topology*, 41(6):1051–1107, 2002.

[17] Vu The Khoi. On the symplectic volume of the moduli space of spherical and Euclidean polygons. *Kodai Mathematical Journal*, 28(1):199–208, 2005.

[18] Harold W Kuhn. A note on Fermat's problem. *Mathematical Programming*, 4(1):98–107, 1973.

[19] Wojciech Niemiro. Asymptotics for M-estimators defined by convex minimization. *The Annals of Statistics*, 20(3):1514–1533, 1992.

[20] Lord Rayleigh. On the problem of random vibrations, and of random flights in one, two, or three dimensions. *Philosophical Magazine Series 5*, 37(220):321–347, 1919.

[21] Joseph Rudnick and George Gaspari. The shapes of random walks. *Science*, 237(4813):384–389, 1987.

[22] Joel A Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of Computational Mathematics*, 12(4):389–434, 2012.