

Asymptotic Bounds on the Combinatorial Diameter of Random Polytopes

Gilles Bonnet^{*†‡} Daniel Dadush^{§¶} Uri Grupel^{||}
 University of Groningen Centrum Wiskunde & Informatica University of Innsbruck

Sophie Huiberts^{**} Galyna Livshyts^{††}
 Centrum Wiskunde & Informatica Georgia Institute of Technology

December 28, 2021

Abstract

The combinatorial diameter $\text{diam}(P)$ of a polytope P is the maximum shortest path distance between any pair of vertices. In this paper, we provide upper and lower bounds on the combinatorial diameter of a random “spherical” polytope, which is tight to within one factor of dimension when the number of inequalities is large compared to the dimension. More precisely, for an n -dimensional polytope P defined by the intersection of m i.i.d. half-spaces whose normals are chosen uniformly from the sphere, we show that $\text{diam}(P)$ is $\Omega(nm^{\frac{1}{n-1}})$ and $O(n^2m^{\frac{1}{n-1}} + n^{54^n})$ with high probability when $m \geq 2^{\Omega(n)}$.

For the upper bound, we first prove that the number of vertices in any fixed two dimensional projection sharply concentrates around its expectation when m is large, where we rely on the $\Theta(n^2m^{\frac{1}{n-1}})$ bound on the expectation due to Borgwardt [Math. Oper. Res., 1999]. To obtain the diameter upper bound, we stitch these “shadow paths” together over a suitable net using worst-case diameter bounds to connect vertices to the nearest shadow. For the lower bound, we first reduce to lower bounding the diameter of the dual polytope P° , corresponding to a random convex hull, by showing the relation $\text{diam}(P) \geq (n-1)(\text{diam}(P^\circ) - 2)$. We then prove that the shortest path between any “nearly” antipodal pair vertices of P° has length $\Omega(m^{\frac{1}{n-1}})$.

*g.f.y.bonnet@rug.nl; Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, The Netherlands; CogniGron (Groningen Cognitive Systems and Materials Center), University of Groningen, The Netherlands.

†Partially funded by the CogniGron research center and the Ubbo Emmius Funds (Univ. of Groningen).

‡Partially funded by the DFG Priority Program (SPP) 2265 Random Geometric Systems, project P23.

§dadush@cwi.nl; Centrum Wiskunde & Informatica, The Netherlands

¶Supported by the ERC Starting grant QIP-805241.

||uri.grupel@uibk.ac.at; University of Innsbruck, Austria

**s.huiberts@cwi.nl; Centrum Wiskunde & Informatica, The Netherlands

††glivshyts6@math.gatech.edu; Georgia Institute of Technology, United States

Contents

1	Introduction	2
1.1	Diameter of Random Polytopes	3
1.2	Prior work	4
1.3	Proof Overview	5
1.3.1	The Upper Bound	5
1.3.2	The Lower Bound	7
1.4	Organization	7
2	Preliminaries	7
2.1	Density Estimates	8
2.2	Cap Volumes	9
2.3	Poisson Processes	11
2.4	Concentration for Nearly-Independent Random Variables	11
3	Shadow size and upper bounding the diameter	11
3.1	Only ‘nearby’ constraints are relevant	13
3.2	Locality, independence, and concentration	16
3.3	Concentration of the shadow size around its mean	18
3.4	Upper bound on the diameter	19
4	Lower Bounding the Diameter of $P(A)$	21

1 Introduction

When does a polyhedron have small (combinatorial) diameter? This question has fascinated mathematicians, operation researchers and computer scientists for more than half a century. In a letter to Dantzig in 1957, motivated by the study of the simplex method for linear programming, Hirsch conjectured that any n -dimensional polytope with m facets has diameter at most $m - n$. While recently disproved by Santos [32] (for unbounded polyhedra, counter-examples were already given by Klee and Walkup [23]), the question of whether the diameter is bounded from above by a polynomial in n and m , known as the *polynomial Hirsch conjecture*, remains wide open. In fact, the current counter-examples violate the conjectured $m - n$ bound by at most 25 percent.

The best known general upper bounds on the combinatorial diameter of polyhedra are the $2^{n-3}m$ bound by Barnette and Larman [3, 26, 4], which is exponential in n and linear in m , and the *quasi-polynomial* $m^{\log_2 n + 1}$ bound by Kalai and Kleitman [22]. The Kalai-Kleitman bound was recently improved to $(m - n)^{\log_2 n}$ by Todd [36] and $(m - n)^{\log_2 O(n/\log n)}$ by Sukegawa [35]. Similar diameter bounds have been established for graphs induced by certain classes of simplicial complexes, which vastly generalize 1-skeleta of polyhedra. In particular, Eisenbrand et al [18] proved both Barnette-Larman and Kalai-Kleitman bounds for so-called connected-layer families (see Theorem 26), and Labbé et al [25] extended the Barnette-Larman bound to pure, normal, pseudo-manifolds without boundary.

Moving beyond the worst-case bounds, one may ask for which families of polyhedra does the Hirsch conjecture hold, or more optimistically, are there families for which we can significantly beat the Hirsch conjecture? In the first line, many interesting classes induced by combinatorial optimization problems are known, including the class of polytopes with vertices in $\{0, 1\}^n$ [28],

Leontief substitution systems [20], transportation polyhedra and their duals [1, 11, 9], as well as the fractional stable-set and perfect matching polytopes [27, 31].

Related to the second vein, there has been progress on obtaining diameter bounds for classes of “well-conditioned” polyhedra. If P is a polytope defined by an integral constraint matrix $A \in \mathbb{Z}^{m \times n}$ with all square submatrices having determinant of absolute value at most Δ , then diameter bounds polynomial in m, n and Δ have been obtained [17, 5, 13, 29]. The best current bound is $O(n^3 \Delta^2 \log(\Delta))$, due to [13]. Extending on the result of Naddef [28], strong diameter bounds have been proved for polytopes with vertices in $\{0, 1, \dots, k\}^n$ [24, 15, 16]. In particular, [24] proved that the diameter is at most nk , which was improved to $nk - \lceil n/2 \rceil$ for $k \geq 2$ [15] and to $nk - \lceil 2n/3 \rceil - (k - 2)$ for $k \geq 4$ [16].

1.1 Diameter of Random Polytopes

With a view of beating the Hirsch bound, the main focus on this paper will be to analyze the diameter of random polytopes, which one may think of as well-conditioned on “average”. Coming both from the average case and smoothed analysis literature [6, 7, 34, 37, 14], there is tantalizing evidence that important classes of random polytopes may have very small diameters.

In the average-case context, Borgwardt [6, 7] proved that for $P := Ax \leq 1$, $A \in \mathbb{R}^{m \times n}$ where the rows of A are drawn from any rotational symmetric distribution (RSD), that the expected number of edges in any fixed 2 dimensional projection of P – the so-called *shadow bound* – is $O(n^2 m^{\frac{1}{n-1}})$. Borgwardt also showed that this bound is tight up to constant factors when the rows of A are drawn uniformly from the sphere, that is, the expected shadow size is $\Theta(n^2 m^{\frac{1}{n-1}})$. In the smoothed analysis context, A has the form $\bar{A} + \sigma G$, where \bar{A} is a fixed matrix with rows of ℓ_2 norm at most 1 and G has i.i.d. $\mathcal{N}(0, 1)$ entries and $\sigma > 0$. Bounds on the expected size of the shadow in this context were first studied by Spielman and Teng [34], later improved by [37, 14], where the best current bound is $O(n^2 \sqrt{\log m} / \sigma^2)$ due to [14] when $\sigma \leq \frac{1}{\sqrt{n \log m}}$.

From the perspective of short paths, these results imply that if one samples objectives v, w uniformly from the sphere, then there is a path between the vertices maximizing v and w in P of expected length $O(n^2 m^{\frac{1}{n-1}})$ in the RSD model, and expected length $O(n^2 \sqrt{\log m} / \sigma^2)$ in the smoothed model. That is, “most pairs” of vertices (with respect to the distribution in the last sentence), are linked by short expected length path. Note that both of these bounds scale either sublinearly or logarithmically in m , which is far better than $m - n$. While these bounds provide evidence, they do not directly upper bound the diameter, since this would need to work for all pairs of vertices rather than most pairs.

A natural question is thus whether the shadow bound is close to the true diameter. In this paper, we show that this is indeed the case, in the setting where the rows of A are drawn uniformly from the sphere and when m is (exponentially) large compared to n . More formally, our main result is as follows:

Theorem 1. *Suppose that $n, m \in \mathbb{N}$ satisfy $n \geq 2$ and $m \geq 2^{\Omega(n)}$. Let $A^\top := (a_1, \dots, a_M) \in \mathbb{R}^{n \times M}$, where M is Poisson distributed with $\mathbb{E}[M] = m$, and a_1, \dots, a_M are sampled independently and uniformly from \mathbb{S}^{n-1} . Then, letting $P(A) := \{x \in \mathbb{R}^n : Ax \leq 1\}$, with probability at least $1 - m^{-n}$, we have that*

$$\Omega(nm^{\frac{1}{n-1}}) \leq \text{diam}(P(A)) \leq O(n^2 m^{\frac{1}{n-1}} + n^5 4^n).$$

In the above, we note that the number of constraints M is chosen according to a Poisson distribution with expectation m . This is only for technical convenience (it ensures useful independence properties, see Proposition 9), and with small modifications, our arguments also work in the case where $M := m$ deterministically. Also, since the constraints are chosen from

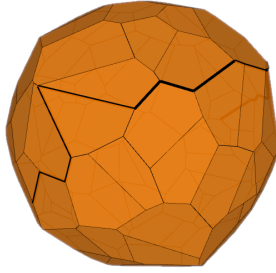


Figure 1: Diameter Achieving Path for Random Spherical Polytope with 100 Constraints

the sphere, M is almost surely equal to the number of facets of $P(A)$ above (i.e., there are no redundant inequalities).

From the bounds, we see that $\text{diam}(P(A)) \leq O(n^2 m^{\frac{1}{n-1}})$ with high probability as long as $m \geq 2^{\Omega(n^2)}$. This shows that the shadow bound indeed upper bounds the expected diameter when $m \rightarrow \infty$. Furthermore, the shadow bound is tight to within one factor of dimension in this regime. We note that in the upper bound is already non-trivial when $m \geq \Omega(n^5 4^n)$, since then $O(n^2 m^{\frac{1}{n-1}} + n^5 4^n) \leq m - n$.

While our bounds are only interesting when m is exponential, the bounds are nearly tight asymptotically, and as far as we are aware, they represent the first non-trivial improvements over worst-case upper bounds for a natural class of polytopes defined by random halfspaces.

Our work naturally leaves two interesting open problems. The first is whether the shadow bound upper bounds the diameter when m is polynomial in n . The second is to close the factor n gap between upper and lower bound in the large m regime.

1.2 Prior work

Lower bounds on the diameter of $P(A)$, $A^\top = (a_1, \dots, a_m) \in \mathbb{R}^{n \times m}$, were studied by Borgwardt and Huhn [8]. They examined the case where each row is sampled from a RSD with radial distribution $\Pr_a[\|a\|_2 \leq r] = \frac{\int_0^r (1-t^2)^\beta t^{n-1} dt}{\int_0^1 (1-t^2)^\beta t^{n-1} dt}$, for $r \in [0, 1]$, $\beta \in (-1, \infty)$. Restricting their results to the case $\beta \rightarrow -1$, corresponding to the uniform distribution on the sphere (where the bound is easier to state), they show that $\mathbb{E}[\text{diam}(P(A))] \geq \Omega(m^{\frac{1}{n} + \frac{1}{n(n-1)^2}})$. We improve their lower bound to $\Omega(nm^{1/(n-1)})$ when $m \geq 2^{\Omega(n)}$, noting that $m^{1/(n-1)} = O(1)$ for $m = 2^{O(n)}$.

In terms of upper bounds, the diameter of a *random convex hull of points*, instead of a random intersection of halfspaces, has been implicitly studied. Given $A^\top = (a_1, \dots, a_m) \in \mathbb{R}^{n \times m}$, let us define

$$Q(A) := \text{conv}(\{a_1, \dots, a_m\}) \tag{1}$$

to be the convex hull of the rows of A . When the rows of A are sampled uniformly from \mathbb{B}_2^n , the question of when the diameter of $Q(A)$ is exactly 1 (i.e., every pair of distinct vertices is connected by an edge) was studied by Bárány and Füredi[2]. They proved that with probability $1 - o(1)$, $\text{diam}(Q(A)) = 1$ if $m \leq 1.125^n$ and $\text{diam}(Q(A)) > 1$ if $m \geq 1.4^n$.

In dimension 3, letting $a_1, \dots, a_M \in \mathbb{S}^2$ be chosen independently and uniformly from the 2-sphere, where M is Poisson distributed with $\mathbb{E}[M] = m$, Glisse, Lazard, Michel and Pouget [19] proved that with high probability the maximum number of edges in any 2-dimensional projection

of $Q(A)$ is $\Theta(\sqrt{m})$. This in particular proves that the combinatorial diameter is at most $O(\sqrt{m})$ with high probability.

It is important to note that the geometry of $P(A)$ and $Q(A)$ are strongly related. Indeed, as long as $m = \Omega(n)$ and the rows of A are drawn from a symmetric distribution, $P(A)$ and $Q(A)$ are polars of each other. That is, $Q(A)^\circ = P(A)$ and $P(A)^\circ := \{x \in \mathbb{R}^n : \langle x, y \rangle \leq 1, \forall y \in P(A)\} = Q(A)$ ¹.

As we will see, our proof of Theorem 1 will in fact imply similarly tight diameter bounds for $\text{diam}(Q(A))$ as for $\text{diam}(P(A))$, yielding analogues and generalizations of the above results, when $A^\top = (a_1, \dots, a_M) \in \mathbb{R}^{n \times M}$ and M is Poisson with $\mathbb{E}[M] = m$. More precisely, we will show that for $m \geq 2^{\Omega(n)}$, with high probability

$$\Omega(m^{\frac{1}{n-1}}) \leq \text{diam}(Q(A)) \leq O(nm^{\frac{1}{n-1}} + n^5 4^n).$$

In essence, for m large enough, our bounds for $\text{diam}(Q(A))$ are a factor $\Theta(n)$ smaller than our bounds for $\text{diam}(P(A))$. This relation will be explained in Section 4.

1.3 Proof Overview

In this section, we give the high level overview of our approach for both the upper and lower bound in Theorem 1.

1.3.1 The Upper Bound

In this overview, we will say that an event holds with high probability if it holds with probability $1 - m^{-\Omega(n)}$. To prove the upper bound on the diameter of $P(A)$, we proceed as follows. For simplicity, we will only describe the level high strategy for achieving a $O(n^2 m^{\frac{1}{n-1}} + 2^{O(n)})$ bound. To begin, we first show that the vertices of $P(A)$ maximizing objectives in a suitable net N of the sphere \mathbb{S}^{n-1} , are all connected to the vertex maximizing e_1 , with a path of length $O(n^2 m^{\frac{1}{n-1}} + 2^{O(n)})$ with high probability. Second, we will show that with high probability, for all $v \in \mathbb{S}^{n-1}$, there is a path between the vertex of $P(A)$ maximizing v and the corresponding maximizer of closest objective $v' \in N$ of length at most $2^{O(n)} \log m$. Since every vertex of $P(A)$ maximizes some objective in \mathbb{S}^{n-1} , by stitching at most 4 paths together, we get that the diameter of $P(A)$ is at most $O(n^2 m^{\frac{1}{n-1}} + 2^{O(n)} \log m) = O(n^2 m^{\frac{1}{n-1}} + 2^{O(n)})$ with high probability.

We only explain the strategy for the first part, as the second part follows easily from the same techniques. The key estimate here is the sharp $\Theta(n^2 m^{\frac{1}{n-1}})$ bound on the expected number of vertices in a fixed two dimensional projection due to Borgwardt [6, 7], the so-called *shadow bound*, which allows one to bound the expected length of paths between vertices maximizing any two fixed objectives (see Section 3 for a more detailed discussion). We first strengthen this result by proving that the size of the shadow sharply concentrates around its expectation when m is large (Theorem 15), allowing us to apply a union bound on a suitable net of shadows, each corresponding a two dimensional plane spanned by e_1 and some element of N above. To obtain such concentration, we show that the shadow decomposes into a sum of nearly independent “local shadows”, corresponding to the vertices maximizing a small slice of the objectives in the plane, allowing us to apply concentration results on nearly-independent sums.

¹Precision: $P(A) = Q(A)^\circ$ always holds and $P(A)^\circ = Q(A)$ requires that $0 \in Q(A)$ which, as a direct consequence of Wendel’s theorem [33, Theorem 8.2.1], happens with probability $1 - o(1)$ when $m \geq cn$ for some $c > 2$. In general $P(A)^\circ = \text{conv}(A \cup \{0\})$ holds.

Independence via Density We now explain the local independence structure in more detail. For this purpose, we examine the smallest $\epsilon > 0$ such that rows of A are ϵ -dense on \mathbb{S}^{n-1} , that is, such that every point in \mathbb{S}^{n-1} is at distance at most ϵ from some row of A . Using standard estimates on the measure of spherical caps and the union bound, one can show with high probability that $\epsilon := \Theta((\log m/m)^{1/m})$ and that any spherical cap of radius $t\epsilon$ contains at most $O(t^{n-1} \log m)$ rows of A for any fixed $t \geq 1$ (see Lemma 7 and Corollary 5).

We derive local independence from the fact that the vertex v of $P(A)$ maximizing a unit norm objective w is defined by constraints $a \in A$ which are distance at most 2ϵ from w (see Lemma 21 for a more general statement). This locality implies that the number of vertices in a projection of $P(A)$ onto a two dimensional subspace $W \ni w$ maximizing objectives at distance ϵ from w (i.e., the slice of objectives) depends only on the constraints in A at distance at most $O(\epsilon)$ from w . In particular, the number of relevant constraints for all objectives at distance ϵ from w is at most $2^{O(n)} \log m$ by the estimate in the last paragraph. By the independence properties of Poisson processes (see Proposition 9), one can in fact conclude that this local part of the shadow on W is independent of the constraints in A at distance more than $O(\epsilon)$ from w .

Given the above, we decompose the shadow onto W into $k = O(1/\epsilon)$ pieces, by placing k equally spaced objectives $w_0, \dots, w_{k-1}, w_k = w_0$ on $\mathbb{S}^{n-1} \cap W$, so that $\|w_i - w_{i+1}\|_2 \leq \epsilon$, $0 \leq i \leq k-1$, and defining $K_i \geq 0$, $0 \leq i \leq k-1$, to be the number of vertices maximizing objectives in $[w_i, w_{i+1}]$. This subdivision partitions the set of shadow vertices, so Borgwardt's bound applies to the expected sum: $\mathbb{E}[\sum_{i=0}^{k-1} K_i] = O(n^2 m^{1/(n-1)})$. Furthermore, as argued above, each K_i is (essentially) independent of all K_j 's with $|i-j \bmod k| = \Omega(1)$. This allows us to apply a Bernstein-type concentration bound for sums of nearly-independent bounded random variables to $\sum_{i=0}^{k-1} K_i$ (see Lemma 10).

Unfortunately, the worst-case upper bounds we have for each K_i , $0 \leq i \leq k-1$, are rather weak. Namely, we only know that in the worst-case, K_i is bounded by the total number of vertices induced by constraints relevant to the interval $[w_i, w_{i+1}]$, where $\|w_i - w_{i+1}\| \leq \epsilon$. As mentioned above, the number of relevant constraints is $2^{O(n)} \log m$ and hence the number of vertices is at most $(2^{O(n)} \log m)^n$. With these estimates, we can show high probability concentration of the shadow size around its mean when $m \geq 2^{\Omega(n^3)}$. One important technical aspect ignored above is that both the independence properties and the worst-case upper bounds on each K_i crucially relies only on conditioning A to be "locally" ϵ -dense around $[w_i, w_{i+1}]$ (see Definition 22 and Lemma 25 for more details).

Abstract Diameter Bounds to the Rescue To allow tight concentration to occur for $m = 2^{\Omega(n^2)}$, we adapt the above strategy by successively following shortest paths instead of the shadow path on W . More precisely, between the maximizer v_i of w_i and v_{i+1} of w_{i+1} , $0 \leq i \leq k-1$, we follow the shortest path from v_i to v_{i+1} in the subgraph induced by the vertices v of $P(A)$ satisfying $\langle v, w_{i+1} \rangle \geq \langle v_i, w_{i+1} \rangle$. We now let K_i , $0 \leq i \leq k-1$, denote the length of the corresponding shortest path. For such local paths, one can apply the abstract Barnette–Larman style bound of [18] to obtain much better worst-case bounds. Namely, we can show $K_i \leq 2^{O(n)} \log m$, $0 \leq i \leq k-1$, instead of $(2^{O(n)} \log m)^n$ (see Lemma 27). Crucially, the exact same independence and locality properties hold for these paths as for the shadow paths, due to the generality of our main locality lemma (Lemma 21). Furthermore, as these paths are only shorter than the corresponding shadow paths, their expected sum is again upper bounded by Borgwardt's bound. With the improved worst-case bounds, our concentration estimates are sufficient to show that all paths indexed by planes in the net N have length $O(n^2 m^{\frac{1}{n-1}} + 2^{O(n)})$ with high probability.

1.3.2 The Lower Bound

For the lower bound, we first reduce to lower bounding the diameter of the polar polytope $P(A)^\circ = Q(A)$, where we show that $\text{diam}(P(A)) \geq (n-1)(\text{diam}(Q(A)) - 2)$ (see Lemma 29). This relation holds as long as $P(A)$ is a simple polytope containing the origin in its interior (which holds with probability $1 - 2^{-\Omega(m)}$). To prove it, we show that given any path between vertices v_1, v_2 of $P(A)$ of length D , respectively incident to distinct facets F_1, F_2 of $P(A)$, one can extract a facet path, where adjacent facets share an $n-2$ -dimensional intersection (i.e., a ridge), of length at most $D/(n-1) + 2$. Such facet paths exactly correspond to paths between vertices in $Q(A)$, yielding the desired lower bound.

For $m \geq 2^{\Omega(n)}$, proving that $\text{diam}(P(A)) \geq \Omega(nm^{1/(n-1)})$ reduces to showing that $\text{diam}(Q(A)) \geq m^{1/(n-1)}$ with high probability. For the $Q(A)$ lower bound, we examine the length of paths between vertices of $Q(A)$ maximizing antipodal objectives, e.g., $-e_1$ and e_1 . From here, one can easily derive an $\Omega((m/\log m)^{\frac{1}{n-1}})$ lower bound on the length of such a path, by showing that every edge of $Q(A)$ has length $\epsilon := \Theta((\log m/m)^{\frac{1}{n-1}})$ and that the vertices in consideration are at distance $\Omega(1)$. This is a straightforward consequence of $Q(A)$ being tightly sandwiched by a Euclidean ball, namely $(1 - \epsilon^2/2)B_2^n \subseteq Q(A) \subseteq B_2^n$ (Lemma 17) with high probability. This sandwiching property is itself a consequence of the rows of A being ϵ -dense on \mathbb{S}^{n-1} , as mentioned in the previous section.

Removing the extraneous logarithmic factor (which makes the multiplicative gap between our lower and upper bound go to infinity as $m \rightarrow \infty$), requires a much more involved argument as we cannot rely on a worst-case upper bound on the length of edges. Instead, we first associate any antipodal path above to a continuous curve on the sphere from $-e_1$ to e_1 (Lemma 33), corresponding to objectives maximized by vertices along the path. From here, we decompose any such curve into $\Omega(m^{\frac{1}{n-1}})$ segments whose endpoints are at distance $\Theta(m^{-1/(n-1)})$ on the sphere. Finally, by appropriately bucketing the breakpoints (Lemma 34) and applying a careful union bound, we show that for any such curve, an $\Omega(1)$ fraction of the segments induce at least 1 edge on the corresponding path with overwhelming probability (Theorem 35). For further details on the lower bound, including how we discretize the set of curves, we refer the reader to Section 4.

1.4 Organization

In Section 2, we introduce some basic notation as well as background materials on Poisson processes (Section 2.3), the measure of spherical caps (Section 2.2), and concentration inequalities for independent random variables (Section 2.4). In Section 3, we prove the upper bound. Halfway into that section, we also prove Theorem 15, a tail bound on the shadow size that is of independent interest. We prove the lower bound in Section 4.

2 Preliminaries

For notational simplicity in the sequel, it will be convenient to treat A as a subset of \mathbb{S}^{n-1} instead of a matrix. For $A \subseteq \mathbb{S}^{n-1}$, we will slightly abuse notation and let $P(A) := \{x \in \mathbb{R}^n : \langle x, a \rangle \leq 1, \forall a \in A\}$ and $Q(A) := \text{conv}(A)$. We denote the indicator of a random event X by $1[X]$, i.e., $1[X] = 1$ if X and $1[X] = 0$ otherwise.

For completeness sake, we first define paths and diameters.

Definition 2. For any polyhedron $P \subseteq \mathbb{R}^n$, a path is a sequence $v_1, v_2, \dots, v_k \in P$ of vertices, such that each line segment $[v_i, v_{i+1}]$, $i \in [k-1]$, is an edge of P . A path is monotone with respect

to an inner product $\langle w, \cdot \rangle$ if $\langle w, v_{i+1} \rangle \geq \langle w, v_i \rangle$ for every $i \in [k-1]$.

The distance between vertices $v_1, v_2 \in P$ is the minimum number k such that there exists a path $v'_1, v'_2, \dots, v'_{k+1}$ with $v_1 = v'_1$ and $v'_k = v_2$. The diameter of P is the maximal distance between any two of its vertices.

2.1 Density Estimates

In this section, we give bounds on the fineness of the net induced by a Poisson distributed subset of \mathbb{S}^{n-1} . Roughly speaking, if A is $\text{Pois}(\mathbb{S}^{n-1}, m)$ distributed then A will be $\Theta((\log m/m)^{1/(n-1)})$ -dense, see Definition 3. While this estimate is standard in the stochastic geometry, it is not so easy to find a reference giving quantitative probabilistic bounds, as more attention has been given to establishing exact asymptotics as $m \rightarrow \infty$ (see [30]). We provide a simple proof of this fact here, together with the probabilistic estimates that we will need.

Definition 3. For $w \in \mathbb{S}^{n-1}$ and $r \geq 0$, we denote by $C(w, r) = \{x \in \mathbb{S}^{n-1} : \|w - x\| \leq r\}$ the spherical cap of radius r centered at w .

We say $A \subseteq \mathbb{S}^{n-1}$ is ε -dense in the sphere for $\varepsilon > 0$ if for every $w \in \mathbb{S}^{n-1}$ there exists $a \in A$ such that $a \in C(w, \varepsilon)$.

Lemma 4. For $m \geq n \geq 2$ and $0 < p < m^{-n}$, have $\varepsilon = \varepsilon(m, n, p) > 0$ satisfy $\sigma(C(v, \varepsilon)) = 3e \log(1/p)/m < 1/12$. Then, for $A \sim \text{Pois}(\mathbb{S}^{n-1}, m)$,

$$\Pr[\exists v \in \mathbb{S}^{n-1} : C(v, \varepsilon) \cap A = \emptyset] \leq p$$

and for every $t \geq 1$,

$$\Pr[\exists v \in \mathbb{S}^{n-1} : |C(v, t\varepsilon) \cap A| \geq 45 \log(1/p)t^{n-1}] \leq p.$$

Proof. Let $N \subseteq \mathbb{S}^{n-1}$ denote the centers of a maximal packing of spherical caps of radius $\varepsilon/(2n)$. By maximality, N is ε/n -dense, i.e., an ε/n net. Comparing volumes, by Lemma 6, we see that

$$1 \geq |N|\sigma(C(v, \varepsilon/(2n))) \geq |N|(2n)^{-(n-1)}\sigma(C(v, \varepsilon)),$$

so $|N| \leq (2n)^{n-1}/\sigma(C(v, \varepsilon)) \leq (2n)^{n-1}m$. By way of a net argument, using that $|C(v, (1-1/n)\varepsilon) \cap A| \sim \text{Pois}(m\sigma(C(v, (1-1/n)\varepsilon)))$, $\forall v \in \mathbb{S}^{n-1}$, we analyze our first probability

$$\begin{aligned} \Pr[\exists v \in \mathbb{S}^{n-1} : C(v, \varepsilon) \cap A = \emptyset] &\leq \Pr[\exists v \in N : C(v, (1-1/n)\varepsilon) \cap A = \emptyset] \\ &\leq |N| \max_{v \in N} \Pr[C(v, (1-1/n)\varepsilon) \cap A = \emptyset] \\ &\leq (2n)^{n-1} m e^{-m\sigma(C(v, (1-1/n)\varepsilon))} \\ &\leq (2n)^{n-1} m e^{-(1-1/n)^{n-1} m\sigma(C(v, \varepsilon))} \\ &\leq (2n)^{n-1} m e^{-3 \log(1/p)} \leq p. \end{aligned}$$

We now prove the second estimate. By Lemma 6, we have that $m\sigma(C(v, (1+1/n)t\varepsilon)) \leq (1+1/n)^{n-1}t^{n-1}m\sigma(C(v, \varepsilon)) \leq 3e^2t^{n-1}\log(1/p)$. Write $\lambda := 3e^2t^{n-1}\log(1/p)$. By a similar net argument as above, we see that

$$\begin{aligned}
\Pr[\exists v \in \mathbb{S}^{n-1} : |C(v, t\varepsilon) \cap A| \geq 2\lambda] &\leq \Pr[\exists v \in N : |C(v, (1+1/n)t\varepsilon) \cap A| \geq 2\lambda] \\
&\leq |N| \max_{v \in N} \Pr[|C(v, (1+1/n)t\varepsilon) \cap A| \geq 2\lambda] \\
&\leq |N| \Pr_{X \sim \text{Pois}(\lambda)}[X \geq 2\lambda] \leq |N| e^{-\frac{(2\lambda - m\sigma(C(v, (1+1/n)t\varepsilon)))^2}{4\lambda}} \\
&\quad (\text{by the Poisson tailbound, Lemma 8}) \\
&\leq |N| e^{-\frac{\lambda}{4}} \leq (2n)^{n-1} m e^{-3 \log(1/p)} \leq p.
\end{aligned}$$

The proof is complete when we observe that $2\lambda \leq 45t^{n-1} \log(1/p)$. \square

We now give effective bounds on the density estimate ε above. Note that taking the $(n-1)^{\text{th}}$ root of the bounds for ε^{n-1} below yields $\varepsilon = \Theta((\log m/m)^{1/(n-1)})$ for $m = n^{\Omega(1)}$ and $p = 1/m^{-n}$. The stated bounds follow directly from the cap measure estimates in Lemma 7.

Corollary 5. *Let $\varepsilon > 0$ be as in Lemma 4, i.e., satisfying $\sigma(C(v, \varepsilon)) = 3\varepsilon \log(1/p)/m \leq 1/12$. Then $\varepsilon \in [0, \sqrt{2(1 - \frac{2}{\sqrt{n}})}]$ and*

$$12\varepsilon \log(1/p)/m \leq \varepsilon^{n-1} \leq (\sqrt{2})^{n-1} \cdot 18\sqrt{n} \log(1/p)/m.$$

Proof. The claim $\varepsilon \in [0, \sqrt{2(1 - \frac{2}{\sqrt{n}})}]$ follows by Lemma 7 part 1 and our assumption that $\sigma(C(v, \varepsilon)) \leq 1/12$. The lower bound on ε^{n-1} follows from the upper bound from Lemma 7 part 2

$$3\varepsilon \log(1/p)/m = \sigma(C(v, \varepsilon)) \leq \frac{1}{2(1 - \varepsilon^2/2)\sqrt{n}} (\varepsilon \sqrt{1 - \varepsilon^2/4})^{n-1} \leq \frac{\varepsilon^{n-1}}{4},$$

where the last inequality follows since $\varepsilon \in [0, \sqrt{2(1 - \frac{2}{\sqrt{n}})}]$. For the upper bound on ε , we rely on the corresponding estimate in Lemma 7 part 2:

$$3\varepsilon \log(1/p)/m = \sigma(C(v, \varepsilon)) \geq \frac{(\varepsilon \sqrt{1 - \varepsilon^2/4})^{n-1}}{6(1 - \varepsilon^2/2)\sqrt{n}} \geq \frac{(\varepsilon \sqrt{1 - \varepsilon^2/4})^{n-1}}{6\sqrt{n}} \geq \frac{(\varepsilon/\sqrt{2})^{n-1}}{6\sqrt{n}},$$

where the last inequality follows from $\varepsilon \in [0, \sqrt{2}]$. The desired inequalities now follow by rearranging. \square

2.2 Cap Volumes

For a subset $C \subseteq \mathbb{S}^{n-1}$, we write $\sigma(C) := \sigma_{n-1}(C)$ to denote the measure of C with respect to the uniform measure on \mathbb{S}^{n-1} . In particular, $\sigma(\mathbb{S}^{n-1}) = 1$. For $v \in \mathbb{S}^{n-1}, \varepsilon \geq 0$, let $C(v, \varepsilon) := \{x \in \mathbb{S}^{n-1} : \|x - v\| \leq \varepsilon\}$ denote the spherical cap of radius ε around v . Throughout the paper, $\|\cdot\|$ will denote the Euclidean norm.

We will need relatively tight estimates on the measure of spherical caps. The following lemma gives useful upper and lower bounds on the ratio of cap volumes.

Lemma 6. *For any $s, \varepsilon > 0$ and $v \in \mathbb{S}^{n-1}$ we have*

$$\frac{\sigma(C(v, (1+s)\varepsilon))}{(1+s)^{n-1}} \leq \sigma(C(v, \varepsilon)) \leq \frac{\sigma(C(v, (1-s)\varepsilon))}{(1-s)^{n-1}},$$

assuming for the first inequality that $(1+s)\varepsilon \leq 2$ and for the second that $s < 1$ and $\varepsilon \leq 2$.

Proof of Lemma 6. First we write the area of the cap as the following integral, for any $r \in [0, 2]$

$$\sigma(C(v, r)) = c_{n-1} \int_0^{r^2/2} \sqrt{2t - t^2}^{n-3} dt,$$

where $c_{n-1} := \text{vol}_{n-2}(\mathbb{S}^{n-2})/\text{vol}_{n-1}(\mathbb{S}^{n-1})$. Note that $\sqrt{2t - t^2}$ is the radius of the slice $\mathbb{S}^{n-1} \cap \{x \in \mathbb{S}^{n-1} : \langle x, v \rangle = 1 - t\} = (1 - t)v + \sqrt{2t - t^2}(\mathbb{S}^{n-1} \cap v^\perp)$. The scaling of the volume of the central slice by $\sqrt{2t - t^2}^{n-3}$ instead of $\sqrt{2t - t^2}^{n-2}$ is to account for the curvature of the sphere. With this integral in our toolbox, we can prove our desired inequalities. We start with the first one, assuming that $(1 + s)^2 r^2 / 2 \leq 2$ so that we only take square roots of positive numbers.

$$\begin{aligned} \sigma(C(v, (1 + s)\varepsilon)) &= c_{n-1} \int_0^{(1+s)^2 r^2 / 2} \sqrt{2t - t^2}^{n-3} dt \\ &= c_{n-1} (1 + s)^2 \int_0^{r^2/2} \sqrt{2(1 + s)^2 u - (1 + s)^4 u^2}^{n-3} du \\ &\leq c_{n-1} (1 + s)^2 \int_0^{r^2/2} \sqrt{2(1 + s)^2 u - (1 + s)^2 u^2}^{n-3} du \\ &= (1 + s)^{n-1} c_{n-1} \int_0^{r^2/2} \sqrt{2u - u^2}^{n-3} du \\ &= (1 + s)^{n-1} \sigma(C(v, \varepsilon)). \end{aligned}$$

The second inequality is proven in a similar fashion, assuming that $1 - s > 0$:

$$\begin{aligned} \sigma(C(v, (1 - s)\varepsilon)) &= c_{n-1} \int_0^{(1-s)^2 r^2 / 2} \sqrt{2t - t^2}^{n-3} dt \\ &= c_{n-1} (1 - s)^2 \int_0^{r^2/2} \sqrt{2(1 - s)^2 t - (1 - s)^4 t^2}^{n-3} dt \\ &\geq c_{n-1} (1 - s)^2 \int_0^{r^2/2} \sqrt{2(1 - s)^2 t - (1 - s)^2 t^2}^{n-3} dt \\ &= (1 - s)^{n-1} c_{n-1} \int_0^{r^2/2} \sqrt{2t - t^2}^{n-3} dt \\ &= (1 - s)^{n-1} \sigma(C(v, \varepsilon)). \end{aligned}$$

□

We now give absolute estimates on cap volume measure due to [10]. We note that [10] parametrize spherical caps with respect to the distance of their defining halfspace to the origin. The following lemma is derived using the fact that the cap $C(v, \varepsilon)$, $\varepsilon \in [0, \sqrt{2}]$, $v \in \mathbb{S}^{n-1}$, is induced by intersecting \mathbb{S}^{n-1} with the halfspace $\langle v, x \rangle \geq 1 - \varepsilon^2/2$, whose distance to the origin is exactly $1 - \varepsilon^2/2$.

Lemma 7. [10, Lemma 2.1] For $n \geq 2$, $\varepsilon \in [0, \sqrt{2}]$, $v \in \mathbb{S}^{n-1}$, the following estimates holds:

- If $\varepsilon \in [\sqrt{2(1 - \frac{2}{\sqrt{n}})}, \sqrt{2}]$, then $\sigma(C(v, \varepsilon)) \in [1/12, 1/2]$.
- If $\varepsilon \in [0, \sqrt{2(1 - \frac{2}{\sqrt{n}})}]$, then

$$\frac{1}{6(1 - \varepsilon^2/2)\sqrt{n}} (\varepsilon \sqrt{1 - \varepsilon^2/4})^{n-1} \leq \sigma(C(v, \varepsilon)) \leq \frac{1}{2(1 - \varepsilon^2/2)\sqrt{n}} (\varepsilon \sqrt{1 - \varepsilon^2/4})^{n-1}.$$

2.3 Poisson Processes

The Poisson distribution $\text{Pois}(\lambda)$ with parameter $\lambda \geq 0$ has probability mass function $f(x, \lambda) := e^{-\lambda} \frac{\lambda^x}{x!}$, $x \in \mathbb{Z}_+$. We note that $\text{Pois}(0)$ is the random variable taking value 0 with probability 1. Recall that $\mathbb{E}[\text{Pois}(\lambda)] = \lambda$. We will rely on the following standard tail-estimate (see [12, Theorem 1]):

Lemma 8. *Let $X \sim \text{Pois}(\lambda)$. Then for $x \geq 0$, we have that*

$$\max\{\Pr[X \geq \lambda + x], \Pr[X \leq \lambda - x]\} \leq e^{-\frac{x^2}{2(\lambda+x)}}. \quad (2)$$

We define a random subset A to be distributed as $\text{Pois}(\mathbb{S}^{n-1}, \lambda)$, $\lambda \geq 0$, if $A = \{a_1, \dots, a_M\}$, where $|A| = M \sim \text{Pois}(\lambda)$ and a_1, \dots, a_M are uniformly and independently distributed on \mathbb{S}^{n-1} . Note that $\mathbb{E}[|A|] = \lambda$. In standard terminology, A is called a homogeneous Poisson point process on \mathbb{S}^{n-1} with intensity $\lambda > 0$.

A basic fact about such a Poisson process is that the number of samples landing in disjoint subsets are independent Poisson random variables.

Proposition 9. *Let $A \sim \text{Pois}(\mathbb{S}^{n-1}, \lambda)$. Let $C_1, \dots, C_k \subseteq \mathbb{S}^{n-1}$ be pairwise disjoint measurable sets. Then, the random variables $|A \cap C_i|$, $i \in [k]$, are independent and $|A \cap C_i| \sim \text{Pois}(\lambda \sigma(C_i))$, $i \in [k]$.*

2.4 Concentration for Nearly-Independent Random Variables

For a random variable $X \in \mathbb{R}$, let $\text{Var}[X] := \mathbb{E}[X^2] - \mathbb{E}[X]^2$ denote its variance.

We will use the following variant on Bernstein's inequality that is a direct consequence of [21, Theorem 2.3], which proves a more general result using the fractional chromatic number of the dependency graph.

Lemma 10. *Suppose that Y_1, \dots, Y_k are random variables taking values in $[0, M]$ and $\text{Var}(Y_i) \leq \sigma^2$ for each $i \in [k]$. Assume furthermore that there exists a partition $I_1 \cup I_2 \cup \dots \cup I_q = \{Y_1, \dots, Y_k\}$ such that the random variables in any one set I_j are mutually independent. Then for any $t \geq 0$ we get*

$$\Pr \left[\left| \sum_{i=1}^k Y_i - \mathbb{E} \left[\sum_{i=1}^k Y_i \right] \right| \geq t \right] \leq 2 \exp \left(\frac{-8t^2}{25q(k\sigma^2 + Mt/3)} \right)$$

When we use the above lemma, we will bound the variance of the random variables using the following inequality:

Lemma 11. *Let $Y \in [0, M]$ be a random variable and $\mathbb{E}[Y] = \mu$. Then $\text{Var}(Y) \leq \mu(M - \mu)$.*

Proof. The inequality follows from $\text{Var}(Y) = \mathbb{E}[Y^2] - \mu^2 \leq M\mathbb{E}[Y] - \mu^2 = \mu(M - \mu)$, where we have used that $Y^2 \leq MY$ for $Y \in [0, M]$. \square

3 Shadow size and upper bounding the diameter

In the first part of this section, we prove a concentration result on the number of *shadow vertices* of $P(A)$. This addresses an open problem from [6]. In the second part, we use the resulting tools to prove Theorem 16, our high-probability upper bound on the diameter of $P(A)$. We start by defining a useful set of paths for which we know their expected lengths.

Definition 12. Let $P \subseteq \mathbb{R}^n$ be a polyhedron and $W \subseteq \mathbb{R}^n$ be a two-dimensional linear subspace. We denote by $\mathcal{S}(P, W)$ the set of shadow vertices: the vertices of P that maximize a non-zero objective function $\langle w, \cdot \rangle$ with $w \in W$.

From standard polyhedral theory, we get a characterization of shadow vertices:

Lemma 13. Let $P(A)$ be a polyhedron given by $A \subseteq \mathbb{R}^n$ and $w \in \mathbb{R}^n \setminus \{0\}$. A vertex $v \in P(A)$ maximizes $\langle w, \cdot \rangle$ if and only if $w\mathbb{R}_+ \cap \text{conv}\{a \in A : \langle a, v \rangle = 1\} \neq \emptyset$.

Hence for $W \subseteq \mathbb{R}^n$ a two-dimensional linear subspace, a vertex $v \in P(A)$ is a shadow vertex $v \in \mathcal{S}(P(A), W)$ if and only if $\text{conv}\{a \in A : \langle a, v \rangle = 1\} \cap W \setminus \{0\} \neq \emptyset$.

The set of shadow vertices for a fixed plane W induces a connected subgraph in the graph consisting of vertices and edges of P , and so any two shadow vertices are connected by a path of length at most $|\mathcal{S}(P, W)|$. As such, for nonzero $w_1, w_2 \in W$, we might speak of a *shadow path* from w_1 to w_2 to denote a path from a maximizer of $\langle w_1, \cdot \rangle$ to a maximizer of $\langle w_2, \cdot \rangle$ that stays inside $\mathcal{S}(P, W)$ and is monotonous with respect to $\langle w_2, \cdot \rangle$. The shadow path was studied by Borgwardt:

Theorem 14 ([6, 7]). Let $m \geq n$ and fix a two-dimensional linear subspace $W \subseteq \mathbb{R}^n$. Pick any probability distribution on \mathbb{R}^n that is invariant under rotations and let the entries of $A \subseteq \mathbb{R}^n$, $|A| = m$, be independently sampled from this distribution. Then, almost surely, for any linearly independent $w_1, w_2 \in W$ there is a unique shadow path from w_1 to w_2 . Moreover, the vertices in $\mathcal{S}(P(A), W)$ are in one-to-one correspondence to the vertices of $\pi_W(P(A))$, the orthogonal projection of $P(A)$ onto W . The expected length of the shadow path from w_1 to w_2 is at most

$$\mathbb{E}[|\mathcal{S}(P(A), W)|] = O(n^2 m^{\frac{1}{n-1}}).$$

This upper bound is tight up to constant factors for the uniform distribution on \mathbb{S}^{n-1} .

We prove a tail bound for the shadow size when $A \sim \text{Pois}(\mathbb{S}^{n-1}, m)$. This result answers a question of Borgwardt in the asymptotic regime, regarding whether bounds on higher moments of the shadow size can be given. To obtain such concentration, we show that the shadow decomposes into a sum of nearly independent ‘‘local shadows’’, using that A will be ε -dense per Lemma 4, allowing us to apply standard concentration results for sums of nearly independent random variables.

Theorem 15 (Shadow Size Concentration). Let $e^{\frac{-m}{18\sqrt{n}(76\sqrt{2})^{n-1}}} < p < m^{-2n}$ and let

$$t_p := \max\left(\sqrt{O(Un^2 m^{\frac{1}{n-1}} \log(1/p))}, O(U \log(1/p))\right)$$

for $U := O(n2^{n^2}(\log(1/p))^n)$. If $A \sim \text{Pois}(\mathbb{S}^{n-1}, m)$ then the shadow size satisfies

$$\Pr\left[|\mathcal{S}(P(A), W)| - \mathbb{E}[|\mathcal{S}(P(A), W)|] > t_p\right] \leq 4p.$$

After that, we extend the resulting tools to obtain our upper bound on the diameter.

Theorem 16 (Diameter Upper Bound). Let $e^{\frac{-m}{18\sqrt{n}(76\sqrt{2})^{n-1}}} < p < m^{-2n}$. If $A = \{a_1, \dots, a_M\} \in \mathbb{S}^{n-1}$, where M is Poisson with $\mathbb{E}[M] = m$, and a_1, \dots, a_M are uniformly and independently distributed in \mathbb{S}^{n-1} . Then, we have that

$$\Pr[\text{diam}(P(A)) > O(n^2 m^{\frac{1}{n-1}} + n4^n \log^2(1/p))] \leq O(\sqrt{p}).$$

Proof. From Corollary 5, for $\varepsilon := \varepsilon(m, n, p)$, we know that $\varepsilon^{n-1} \leq \frac{1}{76^{n-1}}$ given the lower bound on p . In particular, $\varepsilon < 1/76$.

Let $N \subseteq \mathbb{S}^{n-1}$ be a fixed minimal ε -net. Consider the following statements:

- For every $n \in N$, any two vertices in $\mathcal{S}(P(A), \text{span}(e_1, n))$ are connected by a path of length at most $O(n^2 m^{\frac{1}{n-1}}) + t$, where t is defined in Theorem 28.
- A is ε -dense.
- For any $x \in \mathbb{S}^{n-1}$ we have $|A \cap C(x, (2 + 2/n)\varepsilon)| \leq 45e2^n \log(1/p)$.

For given $n \in N$, the first event holds with probability at least $1 - 4p$ by Theorem 28. The net N has $|N| \leq (4/\varepsilon)^n$ points, which is at most $4^n \cdot m$ by Corollary 5. By the union bound the first statement holds for all $n \in N$ simultaneously with probability at least $1 - \sqrt{p}$. From Lemma 4 we know that the second statement holds with probability at least $1 - p$ and the third statement holds with probability at least $1 - p$. We conclude that all three statements hold simultaneously with probability at least $1 - O(\sqrt{p})$.

We will show that the above conditions imply the bound on the combinatorial diameter of $P(A)$.

First, observe that we only need to show an upper bound for all $w \in \mathbb{S}^{n-1}$ on the length of a path connecting any vertex maximizing $\langle w, \cdot \rangle$ to a vertex maximizing $\langle e_1, \cdot \rangle$. The combinatorial diameter of $P(A)$ is at most twice that upper bound.

Let $w \in \mathbb{S}^{n-1}$ and pick $n \in N$ such that $\|w - n\| \leq \varepsilon$. By the first statement, there is a path from the vertex maximizing $\langle n, \cdot \rangle$ to the vertex maximizing $\langle e_1, \cdot \rangle$ of length $O(n^2 m^{\frac{1}{n-1}}) + t$.

By the second two statements, E_{w_1, w_2} is satisfied for every $w_1, w_2 \in \mathbb{S}^{n-1}$. We conclude from Lemma 27 that there is a path from any vertex maximizing $\langle w, \cdot \rangle$ to the vertex maximizing $\langle n, \cdot \rangle$ of length $45en4^n \log(1/p)$.

Therefore, when all three statements hold the combinatorial diameter of $P(A)$ is at most $O(n^2 m^{\frac{1}{n-1}}) + t_p + 45en4^n \log(1/p)$. Now we fill in t_p and obtain an upper bound of

$$O(n^2 m^{\frac{1}{n-1}} + n4^n \log^2 p).$$

□

3.1 Only ‘nearby’ constraints are relevant

We will start by showing that, with very high probability, constraints that are ‘far away’ from a given point on the sphere will not have any impact on the local shape of paths. That will result in a degree of independence between different parts of the sphere, which will be essential in getting concentration bounds on key quantities.

Lemma 17. *If $A \subseteq \mathbb{S}^{n-1}$ is ε -dense for $\varepsilon \in [0, \sqrt{2})$ then $\mathbb{B}_2^n \subseteq P(A) \subseteq \left(1 - \frac{\varepsilon^2}{2}\right)^{-1} \mathbb{B}_2^n$.*

Proof. The first inclusion follows immediately from the construction of $P(A)$. We now show the second inclusion. Taking $x \in P(A) \setminus \{0\}$, we must show that $\|x\| \leq (1 - \varepsilon^2/2)^{-1}$. For this purpose, choose $a \in A$ such that $\|a - x/\|x\|\| \leq \varepsilon$, which exists by our assumption that A is ε -dense. Since $\varepsilon^2 \geq \|a - x/\|x\|\|^2 = 2(1 - \langle a, x/\|x\| \rangle)$, we have that $\langle a, x/\|x\| \rangle \geq 1 - \varepsilon^2/2$. Since $x \in P(A)$, we have $1 \geq \langle a, x \rangle \geq (1 - \varepsilon^2/2)\|x\|$, and the bound follows by rearranging. □

Lemma 18. *If $w \in \mathbb{S}^{n-1}$, $\alpha < 1$, $\|v\| \leq (1 - \alpha)^{-1}$ and $\langle v, w \rangle \geq 1$ then $\|v/\|v\| - w\|^2 \leq 2\alpha$.*

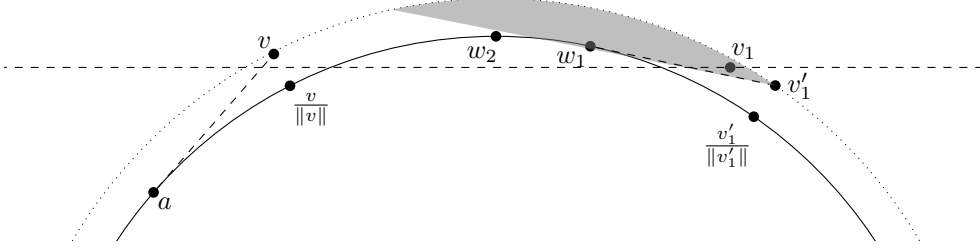


Figure 2: Illustration of the proof of Lemma 19. The inner (resp. outer dotted) curve represents part of the sphere \mathbb{S}^{n-1} (resp. $(1 - \varepsilon^2/2)^{-1}\mathbb{S}^{n-1}$). The horizontal dashed line represents the hyperplane $\{x \in \mathbb{R}^d : \langle x, w_2 \rangle = \langle v_1, w_2 \rangle\}$. The two oblique dashed line segments represent parts of the hyperplanes tangent to the unit sphere at the points a and w_1 . The grey area represents the set B .

Proof. We have $1 \leq \langle v, w \rangle = \|v\| \cdot \langle v/\|v\|, w \rangle \leq (1 - \alpha)^{-1} \langle v/\|v\|, w \rangle$. Hence $1 - \|v/\|v\| - w\|^2/2 = \langle v/\|v\|, w \rangle \geq 1 - \alpha$, which exactly implies that $\|v/\|v\| - w\|^2 \leq 2\alpha$ as required. \square

We will use the above lemmas to prove the main technical estimate of this subsection: if $A \subseteq \mathbb{S}^{n-1}$ is ε -dense and $w_1, w_2 \in \mathbb{S}^{n-1}$ satisfy $\|w_1 - w_2\| \leq 2\varepsilon/n$ then any vertex on any path on $P(A)$ starting at a maximizer of $\langle w_1, \cdot \rangle$ that is non-decreasing with respect to $\langle w_2, \cdot \rangle$ can only be tight at constraints $\langle a, x \rangle = 1$ induced by $a \in A \cap C(w_2, (2 + 2/n)\varepsilon)$. All other constraints are strictly satisfied by every vertex on such a monotone path.

Lemma 19. *Let $\varepsilon \in [0, 1]$ and assume that $w_1, w_2 \in \mathbb{S}^{n-1}$ satisfy $\|w_1 - w_2\| \leq (1 - \varepsilon^2/2)$. Let $v_1, v \in \mathbb{R}^n$ satisfy $\langle w_1, v_1 \rangle \geq 1$ and $\langle w_2, v \rangle \geq \langle v_1, w_2 \rangle$, and assume $\|v_1\|, \|v\| \leq (1 - \varepsilon^2/2)^{-1}$. Last, let $a \in \mathbb{S}^{n-1}$ satisfy $\langle a, v \rangle \geq 1$. Then we have $\|w_2 - a\| \leq 2\varepsilon + \|w_1 - w_2\|$.*

Proof. By Lemma 18, since $\langle w_1, v_1 \rangle, \langle a, v \rangle \geq 1$ and $\|w_1\| = \|a\| = 1$, we get that $\|w_1 - v_1/\|v_1\|\|, \|a - v/\|v\|\| \leq \varepsilon$.

If $w_1 = w_2$, then by assumption $\langle v, w_2 \rangle \geq \langle v_1, w_2 \rangle = \langle v_1, w_1 \rangle \geq 1$. Thus, Lemma 18 implies that $\|w_2 - v/\|v\|\| \leq \varepsilon$. By the triangle inequality, we conclude that $\|w_2 - a\| \leq \|w_2 - v/\|v\|\| + \|v/\|v\| - a\| \leq 2\varepsilon$, as needed.

Now assume that $w_1 \neq w_2$. To prove the lemma, we show that it suffices to v'_1 such that the following two inequalities hold:

$$\left\| \frac{v}{\|v\|} - w_2 \right\| \leq \left\| \frac{v'_1}{\|v'_1\|} - w_2 \right\|, \quad \left\| \frac{v'_1}{\|v'_1\|} - w_1 \right\| \leq \varepsilon. \quad (3)$$

Indeed, given v'_1 as above, the triangle inequality and the first inequality of (3) imply that

$$\begin{aligned} \|w_2 - a\| &\leq \left\| w_2 - \frac{v}{\|v\|} \right\| + \left\| \frac{v}{\|v\|} - a \right\| \\ &\leq \left\| w_2 - \frac{v'_1}{\|v'_1\|} \right\| + \left\| \frac{v}{\|v\|} - a \right\| \\ &\leq \|w_2 - w_1\| + \left\| w_1 - \frac{v'_1}{\|v'_1\|} \right\| + \left\| \frac{v}{\|v\|} - a \right\|. \end{aligned}$$

From here, by the second inequality of (3) and $\|a - v/\|v\|\| \leq \varepsilon$, we get that

$$\|w_2 - a\| \leq \|w_2 - w_1\| + \varepsilon + \varepsilon,$$

which is the claim of the lemma. To construct v'_1 , let

$$B := \left\{ x \in \mathbb{R}^d : \langle w_1, x \rangle \geq 1, \|x\| \leq \left(1 - \frac{\varepsilon^2}{2}\right)^{-1} \right\}.$$

and define v'_1 to be the minimizer of $\langle w_2, \cdot \rangle$ in B . Since $w_1 \neq w_2$, it is direct to verify the v'_1 is unique and satisfies $\|v'_1\| = (1 - \frac{\varepsilon^2}{2})^{-1}$.

From Lemma 18 we have that any point $x \in B$ satisfies $\|x/\|x\| - w_1\| \leq \varepsilon$, and in particular this is true for v'_1 , making the second inequality of (3) hold. Note that $v_1 \in B$ as well. It remains to show the first inequality of (3). For this, we claim that

$$\langle w_2, v \rangle \geq \max\{0, \langle w_2, v'_1 \rangle\},$$

By assumption, recall that $\langle w_2, v \rangle \geq \langle w_2, v_1 \rangle$. The first inequality now follows since $\langle w_2, v_1 \rangle \geq \langle w_1, v_1 \rangle - \|w_1 - w_2\| \|v_1\| \geq 1 - \|w_1 - w_2\| (1 - \varepsilon^2/2)^{-1} \geq 0$, by our assumption on $\|w_1 - w_2\|$. The second inequality now follows from $\langle w_2, v_1 \rangle \geq \langle w_2, v'_1 \rangle$, which holds since $v_1 \in B$ and v'_1 minimizes w_2 over B .

Using that $\|v\| \leq (1 - \varepsilon^2/2)^{-1} = \|v'_1\|$, we conclude that

$$\langle w_2, \frac{v}{\|v\|} \rangle \geq \langle w_2, \frac{v}{\|v'\|} \rangle \geq \langle w_2, \frac{v'_1}{\|v'_1\|} \rangle,$$

where the first inequality uses $\langle w_2, v \rangle \geq 0$. The first inequality of (3) now follows from the fact that $u \in \mathbb{S}^{n-1} \mapsto \|u - w_2\|$ is a decreasing function of $\langle u, w_2 \rangle$, and thus the proof is complete. \square

To round out this subsection, we prove that the conclusion of Lemma 19 holds whenever $v, v_1 \in P(A)$ and A is ε -dense in a neighbourhood around w_2 .

Definition 20. Given sets $A, C \subseteq \mathbb{S}^{n-1}$ and $\varepsilon > 0$, we say that A is ε -dense for C if for every $c \in C$ there exists $a \in A$ such that $\|a - c\| \leq \varepsilon$.

Lemma 21. Let $A \subseteq \mathbb{S}^{n-1}$ be compact and ε -dense for $C(w_2, 4\varepsilon)$, $\varepsilon > 0$. Let $v_1, v \in P(A)$ and $w_1, w_2 \in \mathbb{S}^{n-1}$ satisfying $\langle w_1, v_1 \rangle \geq 1$, $\langle w_2, v \rangle \geq \langle w_2, v_1 \rangle$ and $\|w_1 - w_2\| \leq \varepsilon$. Now let $a \in \mathbb{S}^{n-1}$ satisfy $\langle a, v \rangle \geq 1$. Then we have $\|v_1\|, \|v\| \leq (1 - \varepsilon^2/2)^{-1}$ and $\|w_2 - a\| \leq 2\varepsilon + \|w_1 - w_2\|$.

Note also the contrapositive of the above statement: for w_1, w_2, v_1, v, A satisfying the conditions above, we have for $a \in \mathbb{S}^{n-1}$ that $\|w_2 - a\| > 2\varepsilon + \|w_1 - w_2\|$ implies $\langle a, v \rangle < 1$.

Proof of Lemma 21. First, observe that if $\varepsilon \geq 1$ then the conclusion is trivially satisfied since $\|w_2 - a\| \leq 2 \leq 2\varepsilon + \|w_1 - w_2\|$. From now on, assume $\varepsilon < 1$.

Let $A \subseteq A' \subseteq \mathbb{S}^{n-1}$ be ε -dense, such that $A' \cap C(w_2, 3\varepsilon) \subseteq A$. One valid choice is to take any ε -net $N \subseteq \mathbb{S}^{n-1}$ and set $A' = A \cup (N \setminus C(w_2, 3\varepsilon))$. Then any $x \in C(w_2, 4\varepsilon)$ has an $a \in A \subseteq A'$ with $\|a - x\| \leq \varepsilon$ and any $y \notin C(w_2, 4\varepsilon)$ has some $b \in N$ with $\|y - b\| \leq \varepsilon$ and $b \notin C(w_2, 3\varepsilon)$. Moreover we have $(N \setminus C(w_2, 3\varepsilon)) \cap C(w_2, 3\varepsilon) = \emptyset$ so this choice of A' satisfies our requirements.

If $v, v_1 \in P(A')$, then $\|v\|, \|v_1\| \leq (1 - \varepsilon^2/2)^{-1}$ by Lemma 17 and we can apply Lemma 19 to the set A' and vectors w_1, w_2, v, v_1 and a to conclude $\|w_2 - a\| \leq 2\varepsilon + \|w_1 - w_2\|$ as required.

We now prove that both the case $v_1 \notin P(A')$ and the case $v_1 \in P(A'), v \notin P(A')$ lead to contradiction. First, observe that given w_1 and w_2 , the set of pairs (v_1, v) that satisfy $\langle w_1, v_1 \rangle \geq 1, \langle w_2, v \rangle \geq \langle w_2, v_1 \rangle$ and $\|v_1\|, \|v\| \leq (1 - \varepsilon^2/2)^{-1}$ is a closed convex set and contains (w_1, w_1) .

If $v_1 \notin P(A')$, let (x, y) be the convex combination of (v_1, v_1) and (w_1, w_1) such that $x = y \in P(A')$ and there exists $a' \in A' \setminus A$ such that $\langle a', x \rangle = 1$. Such a' will exist because A' is compact.

Otherwise we have $v \notin P(A')$ and let (x, y) be a convex combination of (v_1, v) and (w_1, w_1) such that $x, y \in P(A')$ and there exists $a' \in A' \setminus A$ such that $\langle a', x \rangle = 1$. Such a' will exist because A' is compact.

Either way, apply Lemma 19 to A', w_1, w_2, x, y and a' to find that $\|w_2 - a'\| \leq 2\varepsilon + \|w_1 - w_2\|$. This contradicts the earlier claim that $a' \in A' \setminus A$. From this contradiction we conclude that $v, v_1 \in P(A')$, which finishes the proof. \square

3.2 Locality, independence, and concentration

With an eye to Lemma 21, this subsection is concerned with proving concentration for sums of random variables that behave nicely when A is dense in given neighbourhoods. The specific random variables that we will use this for are the paths between the maximizers of nearby objective vectors $w_1, w_2 \in \mathbb{S}^{n-1}$.

Definition 22. Given m, n, p , let $\varepsilon = \varepsilon(m, n, p) > 0$ be as in Lemma 4 and $A \subseteq \mathbb{R}^n$ be a random set. For $x, y \in \mathbb{S}^{n-1}$ define the event $E_{x,y}$ as:

- A is ε -dense for $C(x, \|x - y\| + 4\varepsilon)$, and
- for every $z \in [x, y]$ we have

$$\left| A \cap C\left(\frac{z}{\|z\|}, (2 + 2/n)\varepsilon\right) \right| \leq 45e2^n \log(1/p)$$

A random variable K is called (x, y) -local if $E_{x,y}$ implies that K is a function of $A \cap C(x, 5\varepsilon + \|x - y\|)$.

In particular, we will use that if K is (x, y) -local then $K1[E_{x,y}]$ is a function of $A \cap C(x, 5\varepsilon + \|x - y\|)$.

To help prove that certain path are local random variables, we will use the following helper lemma.

Lemma 23. Let $w_1, w_2 \in \mathbb{S}^{n-1}$, and have $w_1 = v_1, v_2, \dots, v_{k+1} = w_2$ be equally spaced on a shortest geodesic segment on \mathbb{S}^{n-1} connecting w_1 and w_2 . Then for every $i \in [k]$ we have $\|w_1 - w_2\|/4k \leq \|v_i - v_{i+1}\| \leq \|w_1 - w_2\|/k$.

Proof. By the triangle inequality, we have $\|w_1 - w_2\| \leq \sum_{i=1}^k \|v_i - v_{i+1}\|$. Since each of the line segments $[v_i, v_{i+1}]$ has identical length, this gives us the second inequality $\|v_i - v_{i+1}\| \leq \|w_1 - w_2\|/k$.

Furthermore, we know that the geodesic segment connecting w_1 and w_2 has length at most $\pi\|w_1 - w_2\|$. From this we get $\sum_{i=1}^k \|v_i - v_{i+1}\| \leq \pi\|w_1 - w_2\|$ and hence $\|w_1 - w_2\|/4k \leq \|w_1 - w_2\|/\pi k \leq \|v_i - v_{i+1}\|$. \square

Many paths on $P(A)$ turn out to be such local random variables. One example are short segments of the shadow paths from Theorem 14.

Lemma 24. Let $w_1, w_2 \in \mathbb{S}^{n-1}$ satisfy $\|w_1 - w_2\| \leq \varepsilon$. Then the length of the shadow path on $P(A)$ from w_1 to w_2 is a (w_1, w_2) -local random variable. If $\|w_1 - w_2\| \leq \varepsilon$ then E_{w_1, w_2} implies that this path has length at most $2n(45e2^n \log(1/p))^n$.

Proof. Let us first assume that $\|w_1 - w_2\| \leq 2\varepsilon/n$. Consider the points $v_1, v \in P(A \cap C(w_2, 5\varepsilon))$ such that $\langle w_1, v_1 \rangle \geq 1$ and $\langle w_2, v \rangle \geq \langle w_2, v_1 \rangle$. By Lemma 21, assuming E_{w_1, w_2} , any such

points have bounded norm. Hence, we can take v_1 to be a vertex maximizing $\langle w_1, \cdot \rangle$ and $v \in P(A \cap C(w_2, 5\varepsilon))$ be any vertex on the shadow path from w_1 to w_2 .

Again by Lemma 21, assuming E_{w_1, w_2} , every $a \in A$ such that $\langle a, v \rangle = 1$ satisfies $a \in C(w_2, (2 + 2/n)\varepsilon)$, meaning that $v, v_1 \in P(A)$ as well.

Now Lemma 13 implies that if E_{w_1, w_2} then any vertex of $P(A \cap C(w_2, 5\varepsilon))$ on its shadow path from w_1 to w_2 is a shadow vertex of $P(A)$ on the shadow path from w_1 to w_2 . Hence the shadow path on $P(A)$ from w_1 to w_2 is a (w_1, w_2) -local random variable.

The upper bound follows because every vertex on the shadow path is visited at most once and, assuming E_{w_1, w_2} , almost surely every vertex on the shadow path is induced by n constraints out of $A \cap C(w_2, (2 + 2/n)\varepsilon)$. The total number of subsets of size n of $A \cap C(w_2, (2 + 2/n)\varepsilon)$ is at most $|A \cap C(w_2, (2 + 2/n)\varepsilon)|^n \leq (45e2^n \log(1/p))^n$ by E_{w_1, w_2} .

To extend the conclusion to the case when $2\varepsilon/n < \|w_1 - w_2\| \leq \varepsilon$, pick $w_1 = v_1, v_2, \dots, v_{2n+1} = w_2$ evenly spaced on the shortest geodesic segment connecting w_1 and w_2 . For every $k \in [2n]$, by Lemma 23 the shadow path from v_k to v_{k+1} satisfies $\|v_k - v_{k+1}\| \leq 2\varepsilon/n$ and is thus a (v_k, v_{k+1}) -local random variable and $E_{v_k, v_{k+1}}$ implies that this shadow path has length at most $(45e2^n \log(1/p))^n$ when $E_{v_k, v_{k+1}}$.

Now observe that the shadow path from w_1 to w_2 is obtained by concatenating the shadow paths from v_k to v_{k+1} for $k \in [2n]$. Since E_{w_1, w_2} implies $E_{v_k, v_{k+1}}$ for every $k \in [2n]$, each of the shadow paths from v_k to v_{k+1} is a (w_1, w_2) -local random variable. Hence the shadow path from w_1 to w_2 is a (w_1, w_2) -local random variable and has length at most $2n(45e2^n \log(1/p))^n$. \square

Lemma 25. *Let $0 < p < m^{-2n}$ and let $\varepsilon = \varepsilon(m, n, p) < 1/76$ be as in Lemma 4 and let $k \geq 2\pi/\varepsilon$ be the smallest number divisible by 76. Let $W \subseteq \mathbb{R}^n$ be a fixed $2D$ linear subspace and let $w_1, \dots, w_k, w_{k+1} = w_1 \in W \cap \mathbb{S}^{n-1}$ be equally spaced around the circle. Assume for every $i \in [k]$ that $K_i \geq 0$ is a (w_i, w_{i+1}) -local random variable and there exists $U \leq m^n$ such that $K_i \leq U$ whenever $E_{w_i, w_{i+1}}$. Furthermore assume that $\mathbb{E}[\sum_{i=1}^k K_i] \leq O(n^2 m^{\frac{1}{n-1}})$. Then*

$$\Pr \left[\left| \sum_{i \in [k]} K_i - \mathbb{E} \left[\sum_{i \in [k]} K_i \right] \right| \geq t_p \right] \leq 4p$$

for $t_p = \max \left(\sqrt{O(Un^2 m^{\frac{1}{n-1}} \log(1/p))}, O(U \log(1/p)) \right)$.

Proof. Let F denote the event that E_{v_1, v_2} holds for every $v_1, v_2 \in \mathbb{S}^{n-1}$. By Lemma 4 we have $\Pr[F] \geq 1 - 2p$.

Define $E_i := E_{w_i, w_{i+1}}$, $i \in [k]$. Our first observation is that $\Pr[\sum_{i=1}^k K_i = \sum_{i=1}^k K_i 1[E_i]] \geq \Pr[F] \geq 1 - p$. Since both sums only take values in the interval $[0, km^n]$, it follows that

$$\left| \mathbb{E} \left[\sum_{i=1}^k K_i \right] - \mathbb{E} \left[\sum_{i=1}^k K_i 1[E_i] \right] \right| \leq 2km^n p \leq 1.$$

From the above statements we deduce that

$$\begin{aligned}
\Pr \left[\left| \sum_{i=1}^k K_i - \mathbb{E} \left[\sum_{i=1}^k K_i \right] \right| > t_p \right] &\leq \Pr \left[\left| \sum_{i=1}^k K_i - \mathbb{E} \left[\sum_{i=1}^k K_i 1[E_i] \right] \right| > t_p - 1 \right] \\
&\leq \Pr[\neg F] + \Pr \left[F \wedge \left| \sum_{i=1}^k K_i - \mathbb{E} \left[\sum_{i=1}^k K_i 1[E_i] \right] \right| > t_p - 1 \right] \\
&\leq 2p + \Pr \left[F \wedge \left| \sum_{i=1}^k K_i 1[E_i] - \mathbb{E} \left[\sum_{i=1}^k K_i 1[E_i] \right] \right| > t_p - 1 \right] \\
&\leq 2p + \Pr \left[\left| \sum_{i=1}^k K_i 1[E_i] - \mathbb{E} \left[\sum_{i=1}^k K_i 1[E_i] \right] \right| > t_p - 1 \right]
\end{aligned}$$

We will now upper bound the last term.

For $j \in [76]$ define $I_j = \{i \in [k] \mid i \equiv j \pmod{76}\}$, forming a partition $I_1 \cup \dots \cup I_{76} = [k]$. Observe that w_1, \dots, w_k are placed on a unit circle and every $[w_i, w_{i+1}]$ is an edge of $\text{conv}(w_1, \dots, w_k)$. As such we know that $\sum_{i \in [k]} \|w_i - w_{i+1}\| \leq 2\pi$. Since $k \geq 2\pi/\varepsilon$ that gives us $\|w_i - w_{i+1}\| \leq \varepsilon$ for every $i \in [k]$. Next, from $\varepsilon \leq 1/76$ we know that $k \leq 2\pi/\varepsilon + 76 \leq 8/\varepsilon$. Since $k \geq 4$ we have $\sum_{i \in [k]} \|w_i - w_{i+1}\| \geq 4$ and hence $\|w_i - w_{i+1}\| \geq 4/k \geq \varepsilon/2$ for every $i \in [k]$. Last, we use that $\|w_i - w_{i+76}\| \leq \sum_{j=i}^{i+75} \|w_j - w_{j+1}\| \leq 76\varepsilon \leq 1$ to deduce

$$\|w_i - w_{i+76}\| \geq \frac{1}{\pi} \sum_{k=i}^{i+75} \|w_k - w_{k+1}\| \geq \frac{76}{\pi} \cdot \varepsilon/2 > 12\varepsilon.$$

This lets us conclude that if $i, i' \in I_j$ are distinct then $\|w_i - w_{i'}\| > 12\varepsilon$. In particular, for any $j \in [76]$ the random variables $K_i 1[E_i]$ for $i \in I_j$ are mutually independent since they are functions of A intersected with disjoint subsets of \mathbb{S}^{n-1} due to being local random variables.

For any $i \in [k]$, the random variable $K_i 1[E_i] \in [0, U]$ has variance at most

$$\mathbb{E}[K_i 1[E_i]] \cdot U \leq \frac{O(n^2 m^{\frac{1}{n-1}})}{k} \cdot U$$

by Lemma 11.

We apply Lemma 10 to the random variables $K_i 1[E_i]$ for $i \in [k]$ and obtain

$$\Pr \left[\left| \sum_{i=1}^k K_i 1[E_i] - \mathbb{E} \left[\sum_{i=1}^k K_i 1[E_i] \right] \right| > t_p - 1 \right] \leq 2 \exp \left(\frac{-8(t_p - 1)^2}{1900(UO(n^2 m^{\frac{1}{n-1}}) + (t_p - 1)U)} \right).$$

By filling in t_p , we find that the right-hand side of the above inequality is at most $2p$.

Putting the bounds together we get our desired inequality

$$\Pr \left[\sum_{i \in [k]} K_i \geq \mathbb{E} \left[\sum_{i \in [k]} K_i \right] + t \right] \leq 4p.$$

□

3.3 Concentration of the shadow size around its mean

To illustrate the use of the above technical result, we show in this subsection that $|\mathcal{S}(P(A), W)|$ is concentrated around its mean when $m > 2^{O(n^3)}$.

Recall that by Theorem 14 we have $\mathbb{E}[|\mathcal{S}(P(A), W)|] = \Theta(n^2 m^{\frac{1}{n-1}})$.

Theorem 15 (Shadow Size Concentration). *Let $e^{\frac{-m}{18\sqrt{n}(76\sqrt{2})^{n-1}}} < p < m^{-2n}$ and let*

$$t_p := \max \left(\sqrt{O(Un^2 m^{\frac{1}{n-1}} \log(1/p))}, O(U \log(1/p)) \right)$$

for $U := O(n2^{n^2} (\log(1/p))^n)$. If $A \sim \text{Pois}(\mathbb{S}^{n-1}, m)$ then the shadow size satisfies

$$\Pr \left[\left| |\mathcal{S}(P(A), W)| - \mathbb{E}[|\mathcal{S}(P(A), W)|] \right| > t_p \right] \leq 4p.$$

Proof. From Corollary 5, we know that $\varepsilon^{n-1} \leq \frac{1}{76^{n-1}}$. As such, the lower bound on p implies that $\varepsilon(m, n, p) < 1/76$.

Let w_1, \dots, w_k be as in Lemma 25 and let K_i denote the number of edges on the shadow path from w_i to w_{i+1} . By Lemma 24, each K_i is a (w_i, w_{i+1}) -local random variable which satisfies $K_i \leq 2n(45e2^n \log(1/p))^n$ when $E_{w_i, w_{i+1}}$.

By Theorem 14 we get $\sum_{i \in [k]} K_i = |\mathcal{S}(P(A), W)|$ almost surely, hence $\mathbb{E}[\sum_{i \in [k]} K_i] \leq O(n^2 m^{\frac{1}{n-1}})$. We apply Lemma 25 to $\sum_{i \in [k]} K_i$:

$$\Pr \left[\left| |\mathcal{S}(P(A), W)| - \mathbb{E}[|\mathcal{S}(P(A), W)|] \right| > t \right] = \Pr \left[\left| \sum_{i \in [k]} K_i - \mathbb{E} \left[\sum_{i \in [k]} K_i \right] \right| > t \right] \leq 4p.$$

□

3.4 Upper bound on the diameter

In this section we prove our high probability upper bound on $\text{diam}(P(A))$. We start by proving that for fixed W the vertices in $\mathcal{S}(P(A), W)$ are connected by short paths, where we aim for an error term smaller than that of Theorem 15. We require the following abstract diameter bound from [18]. We will only need the Barnette–Larman style bound.

Theorem 26. *Let $G = (V, E)$ be a connected graph, where the vertices V of G are subsets of $\{1, \dots, k\}$ of cardinality n and the edges E of G are such that for each $u, v \in V$ there exists a path connecting u and v whose intermediate vertices all contain $u \cap v$.*

Then the following upper bounds on the diameter of G hold:

$$2^{n-1} \cdot k - 1 \text{ (Barnette–Larman)}, \quad k^{1+\log n} - 1 \text{ (Kalai–Kleitman)}.$$

To confirm that the above theorem indeed gives variants of the Barnette–Larman and Kalai–Kleitman bounds, let $A = \{a_1, \dots, a_m\} \subseteq \mathbb{S}^{n-1}$ be in general position. For a vertex $x \in P(A)$, we denote $A_x = \{a \in A : \langle a, x \rangle = 1\}$. Consider the following sets

$$\begin{aligned} V &= \{A_x : x \text{ is a vertex of } P(A)\}, \\ E &= \{\{A_x, A_y\} : [x, y] \text{ is an edge of } P(A)\}. \end{aligned}$$

One can check that $G = (V, E)$ satisfies almost surely the assumptions of theorem 26 which therefore shows that the combinatorial diameter of $P(A)$ is less than $\min(2^{n-1} \cdot m - 1, m^{1+\log n} - 1)$. Up to a constant factor difference, these bounds correspond to the same bounds described in the introduction.

Now we use the Barnette–Larman style bound to bound the length of the local paths.

Lemma 27. *Let $w_1, w_2 \in \mathbb{S}^{n-1}$ satisfy $\|w_1 - w_2\| \leq \varepsilon$, where $\varepsilon = \varepsilon(m, n, p)$ is as in Lemma 4. Furthermore, let K denote the maximum over all $w \in [w_1, w_2]$ of the length of the shortest path from a maximizer $v_w \in P(A)$ of $\langle w, \cdot \rangle$ to the maximizer of $\langle w_2, \cdot \rangle$ of which every vertex $v \in P(A)$ on the path satisfies $\langle w_2, v \rangle \geq \langle w_2, v_w \rangle$. Then K is a (w_1, w_2) -local random variable and E_{w_1, w_2} implies that K_i is at most $45en4^n \log(1/p)$.*

Proof. We start by assuming $\|w_1 - w_2\| \leq 2\varepsilon/n$. Let $w \in [w_1, w_2]$ and let $v_w \in P(A)$ be a vertex maximizing $\langle w, \cdot \rangle$. By Lemma 21, assuming E_{w_1, w_2} , for every vertex $v \in P(A)$ satisfying $\langle w_2, v \rangle \geq \langle w_2, v_w \rangle$ and every $a \in A$ such that $\langle a, v \rangle \geq 1$ we have $a \in A \cap C(w_2, (2 + 2/n)\varepsilon)$.

First, this implies that if E_{w_1, w_2} and if $v \in \mathbb{R}^n$ is satisfies $\langle w_2, v \rangle \geq \langle w_2, v_w \rangle$ then we need only inspect $A \cap C(w_2, (2 + 2/n)\varepsilon)$ to decide if v is a vertex of $P(A)$. From this we conclude that if E_{w_1, w_2} then the shortest path described in the lemma statement can be computed knowing only $A \cap C(w_2, (2 + 2/n)\varepsilon)$. This implies that the path length is a (w_1, w_2) -local random variable.

Second, assuming E_{w_1, w_2} we consider the sets

$$\begin{aligned} \widehat{V} &= \{v \in P(A) : v \text{ is a vertex and } \langle w_2, v \rangle \geq \langle w_2, v_1 \rangle\}, \\ \widehat{A} &= \{a \in A : \text{there exist a vertex } v \in \widehat{V} \text{ such that } \langle a, v \rangle = 1\} \subseteq A \cap C\left(w_2, \left(2 + \frac{2}{n}\right)\varepsilon\right). \end{aligned}$$

The last inclusion follows directly from Lemma 21.

Recall the notation $A_v = \{a \in A : \langle a, v \rangle = 1\}$ for vertices $v \in P(A)$. We will apply Theorem 26 to the graph

$$\begin{aligned} V &= \{A_v : v \in \widehat{V}\} \simeq \widehat{V}, \\ E &= \{\{A_{v_1}, A_{v_2}\} : v_1, v_2 \in \widehat{V}, [v_1, v_2] \text{ is an edge of } P(A)\}. \end{aligned}$$

We need to check that the assumptions of Theorem 26 are met. First we note that almost surely $P(A)$ is a simple polytope and thus the vertices of the graph (V, E) are subsets of A of cardinality n . Consider two vertices $A_v = \{a_{i_1}, \dots, a_{i_n}\}, A_{v'} = \{a_{i'_1}, \dots, a_{i'_n}\} \in V$. Observe that the set

$$F = \{x \in P(A) : \langle x, a \rangle = 1 \ \forall a \in A_v \cap A_{v'}\}$$

is the minimum face of $P(A)$ containing both v and v' . We build paths $v_0 = v, v_1, \dots, v_k$ and $v'_0 = v', v'_1, \dots, v'_{k'}$ satisfying the following monotonicity properties

$$\begin{aligned} \langle w_2, v \rangle = \langle w_2, v_0 \rangle &\leq \langle w_2, v_1 \rangle \leq \dots \leq \langle w_2, v_k \rangle = \operatorname{argmax}\{\langle w_2, x \rangle : x \in F\}, \\ \langle w_2, v' \rangle = \langle w_2, v'_0 \rangle &\leq \langle w_2, v'_1 \rangle \leq \dots \leq \langle w_2, v'_{k'} \rangle = \operatorname{argmax}\{\langle w_2, x \rangle : x \in F\}. \end{aligned}$$

Moreover one can assume that $v_k = v'_{k'}$ by potentially completing the paths moving along the edges of $\operatorname{argmax}\{\langle w_2, x \rangle : x \in F\}$ (in the case this face contains more than one vertex). By construction all vertices v_i and v'_i belong to \widehat{V} . Stitching the two paths and adopting the dual point of view we found a path $A_v = A_{v_0}, \dots, A_{v_k} = A_{v'_{k'}}, \dots, A_{v'_0} = A_{v'}$ whose vertices contain the intersection $A_v \cap A_{v'}$.

We can thus apply Theorem 26 and conclude that there is a path in the graph (V, E) from A_{v_1} to A_{v_2} of length at most $2^{n-1} \cdot |A \cap C(w_2, (2 + 2/n)\varepsilon)|$. It follows that $K \leq 2^{n-1} \cdot |A \cap C(w_2, (2 + 2/n)\varepsilon)|$.

To extend the conclusion to the case when $2\varepsilon/n < \|w_1 - w_2\| \leq \varepsilon$, we do the same as in the proof of Lemma 24. \square

Theorem 28. Let $0 < p < m^{-2n}$ and let

$$t_p = \max \left(\sqrt{O(Un^2 m^{\frac{1}{n-1}} \log(1/p))}, O(U \log(1/p)) \right)$$

for $U = O(n4^n \log(1/p))$. If $W \subseteq \mathbb{R}^n$ is a fixed $2D$ linear subspace and $A \sim \text{Pois}(\mathbb{S}^{n-1}, m)$, the largest distance T between any two shadow vertices satisfies

$$\Pr[T \geq O(n^2 m^{\frac{1}{n-1}}) + t_p] \leq 4p$$

Proof. Let w_1, \dots, w_k be as in Lemma 25 and let K_i denote the maximum over all $w \in [w_i, w_{i+1}]$ of the length of the shortest path from a shadow vertex v_w maximizing $\langle w, \cdot \rangle$ to a vertex maximizing $\langle w_{i+1}, \cdot \rangle$ such that every vertex v on this path satisfies $\langle w_{i+1}, v \rangle \geq \langle w_{i+1}, v_w \rangle$. From Lemma 27 we know that K_i is a (w_i, w_{i+1}) -local random variable and $K_i \leq 45en4^n \log(1/p)$ whenever $E_{w_i, w_{i+1}}$. Now recall Theorem 14. Observe that $T \leq \sum_{i \in [k]} K_i$ almost surely by concatenating the above-mentioned paths, and note that that $\sum_{i \in [k]} K_i \leq \mathcal{S}(P(A), W)$ holds almost surely, which implies $\mathbb{E}[\sum_{i \in [k]} K_i] = O(n^2 m^{\frac{1}{n-1}})$. We apply Lemma 25 to $\sum_{i \in [k]} K_i$ and get the desired result. \square

Theorem 16 (Diameter Upper Bound). Let $e^{\frac{-m}{18\sqrt{n}(76\sqrt{2})^{n-1}}} < p < m^{-2n}$. If $A = \{a_1, \dots, a_M\} \subseteq \mathbb{S}^{n-1}$, where M is Poisson with $\mathbb{E}[M] = m$, and a_1, \dots, a_M are uniformly and independently distributed in \mathbb{S}^{n-1} . Then, we have that

$$\Pr[\text{diam}(P(A)) > O(n^2 m^{\frac{1}{n-1}} + n4^n \log^2(1/p))] \leq O(\sqrt{p}).$$

4 Lower Bounding the Diameter of $P(A)$

To begin, we first reduce to lower bounding the diameter of the polar polytope P° , corresponding to a convex hull of m uniform points on \mathbb{S}^{n-1} , via the following simple lemma.

Lemma 29 (Diameter Relation). For $n \geq 2$, let $P \subseteq \mathbb{R}^n$ be a simple bounded polytope containing the origin in its interior and let $Q = P^\circ := \{x \in \mathbb{R}^n : \langle x, y \rangle \leq 1, \forall y \in P\}$ denote the polar of P . Then, $\text{diam}(P) \geq (n-1)(\text{diam}(Q) - 2)$.

Proof. If $\text{diam}(Q) \leq 1$, the statement is trivial, so we may assume that $\text{diam}(Q) \geq 2$. Let $a_1, a_2 \in Q$ be vertices of Q at distance $\text{diam}(Q) \geq 2$. Since P is bounded, note that 0 is in the interior of Q and hence $a_1, a_2 \neq 0$. We must show that there exists a path from a_1 to a_2 of length $L \geq 2$ such $\text{diam}(P) \geq (n-1)(L-2)$.

Let $F_i := \{x \in P : \langle a_i, x \rangle = 1\}$, $i \in [2]$, the corresponding facets of P . Pick the two vertices $v_1 \in F_1$, $v_2 \in F_2$ whose distance in P is minimized. Let $v_1 := w_0, \dots, w_D := v_2$ be a shortest path from v_1 to v_2 in P . Here w_0, \dots, w_D are all vertices of P , and $[w_i, w_{i+1}]$, $0 \leq i \leq D-1$, are edges of P . By definition, $D \leq \text{diam}(P)$.

To complete the proof, we will extract a walk from a_1 to a_2 in Q from the path w_0, \dots, w_D of length at most $D/(n-1) + 2$. For this purpose, let $Q_i := Q \cap \{x \in \mathbb{R}^n : \langle x, w_i \rangle = 1\}$, $0 \leq i \leq D$, denote the facet of Q induced by w_i . By our assumption that P is simple, each Q_i , $i \in [D]$, is a $(n-1)$ -dimensional simplex, and hence there exists $S_i \subseteq \text{vertices}(Q)$, $|S_i| = n$, such that $Q_i := \text{conv}(a : a \in S_i)$. In particular, the combinatorial diameter of each Q_i , $0 \leq i \leq D$, is 1. That is, every distinct pair of vertices of Q_i induces an edge of Q_i , and hence an edge of Q .

By the above discussion, note that if $a_1, a_2 \in S_0$, then a_1, a_2 are adjacent in Q . Since we assume that the distance between a_1, a_2 is at least 2, we conclude that $a_1, a_2 \notin S_0$, and hence

that $D \geq 1$. Furthermore, since we assume that v_1, v_2 are at minimum distance in P subject to $v_1 \in F_1, v_2 \in F_2$, we conclude that $a_1 \in S_0 \setminus \cup_{j=1}^D S_j$ and $a_2 \in S_L \setminus \cup_{j=0}^{D-1} S_j$, since otherwise we could shortcut the path.

We now define a walk $a_1 = u_0, \dots, u_L = a_2$, for some $L \geq 2$, from a_1 to a_2 in Q as follows. Letting $l_0 = 0$ and $S_{D+1} := \emptyset$, for $i \geq 1$ inductively define $l_i := \max\{j \geq l_{i-1} : \cap_{r=l_{i-1}}^j S_r \neq \emptyset\}$ and let $L = \min\{i \geq 1 : l_i = D\} + 1$. For $1 \leq i \leq L-1$, choose u_i from $\cap_{r=l_{i-1}}^{l_i} S_r$ arbitrarily. To relate the length of the walk to D , we will need the following claim.

Claim 30. *For any interval $I \subseteq \{0, \dots, D\}$, $|\cap_{i \in I} S_i| \geq n - |I| + 1$.*

Proof. First note that $|S_j \cap S_{j+1}| = n-1 = |S_j| - 1$, $0 \leq j \leq D-1$, since P is simple and $S_j \cap S_{j+1}$ indexes the tight constraints of an edge of P . In particular, $|S_j \setminus S_{j+1}| = 1$, $0 \leq j \leq D-1$. Thus, for an interval $I = \{c, c+1, \dots, d\} \subseteq \{0, \dots, D\}$, we see that $|\cap_{i=c}^d S_i| \geq |\cap_{i=c}^{d-1} S_i| - |S_{d-1} \setminus S_d| = |\cap_{i=c}^{d-1} S_i| - 1 \geq |S_c| - (d-c) = n+1 - |I|$. \square

Applying the claim to the interval $I = \{l_{i-1}, \dots, l_i+1\}$, $1 \leq i \leq L-1$, we see that $\cap_{r=l_{i-1}}^{l_i+1} S_r = \emptyset$ implies that either $l_i = D$ or that $|I| \geq n+1 \Leftrightarrow l_i - l_{i-1} \geq n-1$. In particular, $l_i - l_{i-1} \geq n-1$ for $0 \leq i \leq L-2$ and $l_{L-1} - l_{L-2} \geq 1$ (since $l_{L-1} = D$ and $l_{L-2} < D$).

Let us now verify that $a_1 = u_0, u_1, \dots, u_L = a_2$ induces a walk in Q . Here, we must check that $[u_i, u_{i+1}]$, $0 \leq i \leq L-1$, is an edge of Q . By construction u_i, u_{i+1} are both vertices of the simplex Q_{l_i} . Furthermore, $u_i \neq u_{i+1}$, since either $u_i = a_1 \neq u_{i+1}$ or $u_{i+1} = a_2 \neq u_i$ or $u_{i+1} \in S_{l_i+1}$ and $u_i \notin S_{l_i+1}$. Thus, $[u_i, u_{i+1}]$ is indeed an edge of Q_i and thus of Q , as explained previously. Note by our assumption that a_1 and a_2 , we indeed have $2 \leq \text{diam}(Q) \leq L$.

We can now compare the diameters of P and Q as follows:

$$\text{diam}(P) \geq D = l_{L-1} - l_0 = \sum_{i=1}^{L-1} (l_i - l_{i-1}) \geq \sum_{i=1}^{L-2} (n-1) = (n-1)(L-2) \geq (n-1)(\text{diam}(Q) - 2),$$

as needed. \square

We then associate any ‘‘antipodal’’ path to a continuous curve on the sphere corresponding to objectives maximized by vertices along the path. From here, we decompose any such curve into $\Omega(m^{\frac{1}{n-1}})$ segments whose endpoints are at distance $\Theta(m^{-1/(n-1)})$ on the sphere. Finally, we apply a suitable union bound, to show that for any such curve, an $\Omega(1)$ fraction of the segments induce at least 1 edge on the corresponding path.

Building on Lemma 29, we turn to proving the lower bound for $Q(A)$.

For a discrete set $N \subseteq S^{n-1}$, a point $x_0 \in N$ and a positive number $\varepsilon > 0$ we denote by

$$X_k := X_k(N, x_0, \varepsilon) = \{\mathbf{x} \in N^k : x_i \neq x_j \text{ and } 6\varepsilon \leq \|x_i - x_{i+1}\| \leq 8\varepsilon \text{ for any } 0 \leq i < j \leq k\}$$

the set of all sequences of k distinct points in N with jumps of length between 6ε and 8ε (including an extra initial jump between x_0 and x_1).

Lemma 31. *Let $\varepsilon > 0$. If $N \subseteq S^{n-1}$ is a maximal ε -separated set, then*

$$|X_k| \leq (17^{n-1})^k.$$

Note that a maximal ε -separated set is also an ε -net.

Proof of Lemma 31. For any $x \in N$ we find an upper bound for the number of points $y \in N$ such that $6\varepsilon \leq \|x - y\| \leq 8\varepsilon$. Recall that $C(x, r)$ denotes the closed spherical cap centered at x with radius $r > 0$. Since N is ε -separated, for any different points $y_1, y_2 \in N$ we have

$$\text{int}(C(y_1, \varepsilon/2)) \cap C(y_2, \varepsilon/2) = \emptyset.$$

Taking a union of spherical caps centered at all points inside the annulus, we obtain a subset of the inflated annulus

$$C(x, 17\varepsilon/2) \setminus \text{int}(C(x, 11\varepsilon/2)).$$

Since the caps $C(y, \varepsilon/2)$, $y \in N$, have pairwise disjoint interiors, the volume of their union is the sum of the volumes. Hence, the maximal number of points in the annulus is bounded by

$$\frac{\sigma(C(x, 17\varepsilon/2)) - \sigma(C(x, 11\varepsilon/2))}{\sigma(C(x, \varepsilon/2))} \leq \frac{\sigma(C(x, 17\varepsilon/2))}{\sigma(C(x, \varepsilon/2))}.$$

Using Lemma 6 we have

$$|\{y \in N : 6\varepsilon \leq \|x - y\| \leq 8\varepsilon\}| \leq \frac{(17/2)^{n-1}}{(1/2)^{n-1}} = 17^{n-1}.$$

Thus, the overall number of paths in X_k is bounded by

$$|X_k| \leq 17^{k(n-1)}.$$

□

Lemma 32. *Let $f: [0, 1] \rightarrow S^{n-1}$ be a continuous function. Let $\varepsilon > 0$ and $N \subseteq S^{n-1}$ be an ε -net, such that $f(0) \in N$. There exist $k \in \mathbb{N}_0$, $0 \leq t_0 < t_1 < \dots < t_k \leq 1$ and $x_0, \dots, x_k \in N$ such that*

1. $\|f(t_i) - x_i\| \leq \varepsilon$ for any $i \in \{0, \dots, k\}$,
2. $\|f(t) - x_i\| \geq \varepsilon$ for any $i \in \{0, \dots, k\}$ and $t > t_i$,
3. $(x_1, \dots, x_k) \in X_k(N, x_0, \varepsilon)$,
4. $\|x_k - f(1)\| < 7\varepsilon$.

Proof. We build the desired couple of sequences (x_i) and (t_i) by induction. We start by taking $x_0 = f(0)$ and

$$t_0 = \sup\{t \geq 0 : \|f(t) - x_0\| \leq \varepsilon\}.$$

Note that with these choices, we have a couple of (very short) sequences for which 1-3 are fulfilled.

Assume that x_0, \dots, x_ℓ and $0 \leq t_0 < \dots < t_\ell \leq 1$ are sequences for which 1-3 hold true.

If $\|x_\ell - f(1)\| < 7\varepsilon$ then we may take $k = \ell$, and we are done.

Assume otherwise, and define

$$t' = \min\{t \in [t_\ell, 1] : \exists x_{\ell+1} \in N \text{ with } \|f(t) - x_{\ell+1}\| \leq \varepsilon \text{ and } \|x_{\ell+1} - x_\ell\| \geq 6\varepsilon\},$$

Since 4 is not fulfilled, the set is non-empty (it contains 1) and t' is well defined. We take $x_{\ell+1}$ as it appears in the definition of t' . Set

$$t_{\ell+1} = \sup\{t \in [0, 1] : \|f(t) - x_{\ell+1}\| \leq \varepsilon\}.$$

By 2 for any $i \leq \ell$

$$\|f(t_{\ell+1}) - x_i\| > \varepsilon,$$

hence $x_i \neq x_{\ell+1}$. Combining this with the definition of $t_{\ell+1}$ and $x_{\ell+1}$ we only need to show that $\|x_\ell - x_{\ell+1}\| \leq 8\varepsilon$ in order to get that $0 \leq t_0 < \dots < t_\ell < t_{\ell+1} \leq 1$ and $x_0, \dots, x_{\ell+1}$ fulfill 1-3.

By the minimality of t' , for any $s \in (t_\ell, t')$ we have $\|x_\ell - f(s)\| \leq 7\varepsilon$, otherwise there would be $x' \in N$ such that $\|x' - f(s)\| \leq \varepsilon$ but $\|x_\ell - x'\| \geq 6\varepsilon$, hence $t' \leq s$ in contradiction to the definition of s . Hence

$$\|x_\ell - x_{\ell+1}\| \leq \|x_\ell - f(s)\| + \|f(s) - f(t')\| + \|f(t') - x_{\ell+1}\| \leq 7\varepsilon + \|f(s) - f(t')\| + \varepsilon.$$

This holds for all $s \in (t_\ell, t')$. By continuity of f we may take $s \nearrow t'$ and have $\|f(s) - f(t')\| \rightarrow 0$. Thus $\|x_\ell - x_{\ell+1}\| \leq 8\varepsilon$.

Since N is finite and the points x_0, \dots, x_ℓ are distinct the process must end at most after $|N|$ steps. \square

Lemma 33. *Let $A \subseteq S^{n-1}$ be a finite subset of the sphere. Let $[a_0, a_1], [a_1, a_2], \dots, [a_{\ell-1}, a_\ell]$ be a path along the edges of $Q(A)$. There exists a continuous function $f: [0, 1] \rightarrow S^{n-1}$ and $0 = s_0 < s_1 < \dots < s_{\ell+1} = 1$ such that $f(0) = a_0$, $f(1) = a_\ell$, and for any $i \in \{0, 1, \dots, \ell\}$ and any $t \in [s_i, s_{i+1}]$,*

$$a_i \in \operatorname{argmin}_{a \in A} (\|f(t) - a\|).$$

Proof. First we consider the case where the path consist of a single edge, i.e. $\ell = 1$. Consider a point $x \in S^{n-1}$ and a real $r > 0$ such that the cap $C(x, r)$ contains a_0 and a_1 on its boundary and no point of A in its interior. A possible choice is given by the circumscribed cap of any facet of $Q(A)$ which contains $[a_0, a_1]$ as an edge. Now we set f such that it interpolates a_0 , x and a_1 by two geodesic segments,

$$f(t) = \frac{\tilde{f}(t)}{\|\tilde{f}(t)\|}, \quad \tilde{f}(t) = \begin{cases} (1-2t)a_0 + 2tx, & t \in [0, \frac{1}{2}], \\ (2-2t)x + (2t-1)a_1, & t \in [\frac{1}{2}, 1]. \end{cases}$$

By construction we get that for any $t \in [0, \frac{1}{2}]$ (resp. $t \in [\frac{1}{2}, 1]$), the cap $C(f(t), \|f(t) - a_0\|)$ (resp. $C(f(t), \|f(t) - a_1\|)$) is a subset of $C(x, r)$. Thus it contains a_0 (resp. a_1) on its boundary and no point of A in its interior. This implies that $f(0) = a_0$, $f(1) = a_1$, and

$$a_0 \in \operatorname{argmin}_{a \in A} (\|f(t) - a\|), \quad t \in [0, \frac{1}{2}],$$

$$a_1 \in \operatorname{argmin}_{a \in A} (\|f(t) - a\|), \quad t \in [\frac{1}{2}, 1].$$

This yields the proof in the case $\ell = 1$ (with $s_0 = 0 < s_1 = \frac{1}{2} < s_{1+1} = 1$). The general case follows by concatenating and renormalizing the functions corresponding to each edge. \square

Lemma 34. *Let $A \subseteq S^{n-1}$ be a finite subset of the sphere, containing two points $a_+, a_- \in A$ such that $\|a_+ - a_-\| \geq 1$. Let $\varepsilon > 0$ and $N \subseteq S^{n-1}$ be a maximal ε -separated set, such that $a_+ \in N$. Set $x_0 = a_+$ and $k_0 = \lceil 1/8\varepsilon \rceil - 1$. It holds that*

$$\operatorname{diam}(Q(A)) \geq \min_{k \geq k_0} \min_{\mathbf{x} \in X_k(N, x_0, \varepsilon)} \sum_{0 \leq i \leq k-1} 1[C(x_i, \varepsilon/2) \cap A \neq \emptyset] 1[C(x_{i+1}, \varepsilon/2) \cap A \neq \emptyset].$$

Proof. The diameter of $Q(A)$ is at least the combinatorial distance between a_+ and a_- , i.e., the minimal number of edges required to form a path between these two vertices. Note that this minimum is realized for a path without loops. Let $[a_0, a_1], [a_1, a_2], \dots, [a_{\ell-1}, a_\ell]$ be such a path. Here we denote $a_0 = a_+ = x_0$ and $a_\ell = a_-$.

Consider a function f and a sequence $0 = s_0 < s_1 < \dots < s_{\ell+1} = 1$ as in Lemma 33, and consider $k \in \mathbb{N}_0$, $0 \leq t_0 < t_1 < \dots < t_k \leq 1$ and $x_0, \dots, x_k \in N$ as in Lemma 32. We set $j(0) \leq j(1) \leq \dots \leq j(k)$ such that $t_i \in [s_{j(i)}, s_{j(i)+1}]$. In particular, with this notation set up we have

$$\|x_i - x_{i+1}\| \geq 6\varepsilon, \quad i \in \{0, \dots, k-1\}, \quad (4)$$

$$\|a_{j(i)} - f(t_i)\| = \min_{a \in A} \|a - f(t_i)\|, \quad i \in \{0, \dots, k\}, \quad (5)$$

and

$$\|x_i - f(t_i)\| \leq \varepsilon, \quad i \in \{0, \dots, k\}. \quad (6)$$

From (6) we get $C(f(t_i), 3\varepsilon/2) \supset C(x_i, \varepsilon/2)$. Hence, if $C(x_i, \varepsilon/2) \cap A \neq \emptyset$, we have that $\|a_{j(i)} - f(t_i)\| \leq 3\varepsilon/2$ because of (5). Therefore if, for some $i \in \{0, \dots, k-1\}$, both caps $C(x_i, \varepsilon/2)$ and $C(x_{i+1}, \varepsilon/2)$ contain points of A , then

$$\begin{aligned} & \|a_{j(i)} - a_{j(i+1)}\| \\ & \geq \|x_i - x_{i+1}\| - \|x_i - f(t_i)\| - \|f(t_i) - a_{j(i)}\| - \|a_{j(i+1)} - f(t_{i+1})\| - \|f(t_{i+1}) - x_{i+1}\| \\ & \geq 6\varepsilon - \varepsilon - 3\varepsilon/2 - 3\varepsilon/2 - \varepsilon = \varepsilon > 0 \end{aligned}$$

and we get $a_{j(i+1)} \neq a_{j(i)}$ which implies that $j(i) < j(i')$ for any $i' > i$. This shows that if

$$i, i' \in I = \{i : C(x_i, \varepsilon/2) \cap A \neq \emptyset \text{ and } C(x_{i+1}, \varepsilon/2) \cap A \neq \emptyset\} \subseteq \{0, 1, \dots, k-1\},$$

with $i \neq i'$, then $a_{j(i)}$ and $a_{j(i')}$ are distinct vertices of the path. Therefore

$$l \geq |I| = \sum_{0 \leq i \leq k-1} 1[C(x_i, \varepsilon/2) \cap A \neq \emptyset] 1[C(x_{i+1}, \varepsilon/2) \cap A \neq \emptyset].$$

Also, we note that from

$$\begin{aligned} \|a_+ - a_-\| & \leq \|a_+ - x_0\| + \sum_{1 \leq i \leq k} \|x_i - x_{i-1}\| + \|x_k - a_-\| \\ & < \varepsilon + k \times 8\varepsilon + 7\varepsilon = 8(k+1)\varepsilon \end{aligned}$$

we have $k \geq k_0$, and therefore

$$(x_0, \dots, x_k) \in \cup_{k \geq k_0} X_k(N, x_0, \varepsilon).$$

□

Theorem 35 (Lower Bound for $Q(A)$). *There exist positive constants $c_2 < 1$ and $c_3 > 1$ independent of $n \geq 3$ and m such that the following holds. Let $A = \{a_1, \dots, a_M\} \in \mathbb{S}^{n-1}$, where M is Poisson with $\mathbb{E}[M] = m$, and a_1, \dots, a_M are uniformly and independently distributed in \mathbb{S}^{n-1} . Then, with probability at least $1 - e^{-c_3^{n-1} m^{1/(n-1)}}$, the combinatorial diameter of $Q(A)$ is at least $c_2 m^{1/(n-1)}$.*

Proof. Without loss of generality $m \geq (1/c_2)^{n-1}$ since otherwise the statement of the theorem is trivial.

In this proof the constants $1 < c_3 < c_4 < c_5 < c_6 < c_2^{-1}$ are large enough constants, independent from n and m .

We set $\varepsilon = c_6 m^{-1/(n-1)}$, and want to apply Lemma 34. Let N be an ε -net, obtained from a maximal ε -separated set, such that it contains a point a_+ from the set A . For independence properties needed later we take a_+ randomly and uniformly from the set A . With probability $1 - e^{-m/2}$ we have that A intersects the halfsphere $\{u \in \mathbb{S}^{n-1} : \langle a_+, u \rangle \leq 0\}$. In which case there exists a point $a_- \in A$ such that $\|a_+ - a_-\| \geq \sqrt{2} \geq 1$. Therefore we can apply Lemma 34 with $x_0 = a_+$. Combined with the union bound, we get

$$\Pr\left(\text{diam } Q(A) \leq c_2 m^{1/(n-1)}\right) \leq e^{-m/2} + \sum_{k \geq k_0} \sum_{\mathbf{x} \in X_k(N, x_0, \varepsilon)} \Pr\left(\sum_{0 \leq i \leq k-1} B_i \leq c_2 m^{1/(n-1)}\right),$$

where

$$k_0 = \lceil 1/8\varepsilon \rceil + 1 \geq 1/8\varepsilon = m^{1/(n-1)}/8c_6,$$

and the summands in the probability are Bernoulli random variables

$$B_i = 1[C(x_i, \varepsilon/2) \cap A \neq \emptyset]1[C(x_{i+1}, \varepsilon/2) \cap A \neq \emptyset].$$

For $1 \leq i \leq k-1$, they are identically distributed, with failure probability

$$\begin{aligned} \Pr(B_i = 0) &\leq 2\Pr(C(x_i, \varepsilon/2) \cap A = \emptyset) = 2\exp(-m\sigma(C(x_i, \varepsilon/2))) \\ &\leq 2\exp(-m(\varepsilon/4)^{n-1}) = 2\exp\left(-\left(\frac{c_6}{4}\right)^{n-1}\right) =: 1-p. \end{aligned}$$

Note that we used Lemma 6 to lower bound the cap's volume $\sigma(C(x_i, \varepsilon/2)) \geq (\varepsilon/4)^{n-1}\sigma(C(x_i, 2))$. Since N forms a maximal ε -separated set and the x_i are distinct, the caps $C(x_i, \varepsilon/2)$ are disjoint and therefore the random variables B_1, B_3, B_5, \dots are independent. Next we exploit this independence. Let $k \geq k_0$, and set $K = \lfloor k/2 \rfloor$. Note that $K \geq 1/16\varepsilon = m^{1/(n-1)}/16c_6$. Assuming that $c_2 \leq 1/32c_6$, we have

$$\Pr\left(\sum_{0 \leq i \leq k-1} B_i \leq c_2 m^{1/(n-1)}\right) \leq \Pr\left(\sum_{1 \leq i \leq K} B_{2i-1} \leq \frac{K}{2}\right) = \sum_{1 \leq i \leq \lfloor K/2 \rfloor} \binom{K}{i} p^i (1-p)^{K-i}.$$

Now we bound p by 1, $(1-p)^{K-i}$ by $(1-p)^{K/2}$ and $\sum \binom{K}{i}$ by 2^K , which provides us the bound

$$\Pr\left(\sum_{0 \leq i \leq k-1} B_i \leq c_2 m^{1/(n-1)}\right) \leq (2(1-p)^{1/2})^K = \left(e^{\left(-\frac{1}{2}\left(\frac{c_6}{4}\right)^{n-1} + \frac{3}{2} \ln 2\right)}\right)^K \leq \left(e^{(-c_5^{n-1})}\right)^K.$$

Thus, with the bound $|X_k| \leq (17^{n-1})^k$ from lemma 31, and the fact that $K \geq k/2$, we get

$$\begin{aligned}
\Pr\left(\text{diam } Q(A) \leq c_2 m^{-1/(n-1)}\right) &\leq e^{-m/2} + \sum_{k \geq k_0} \left(e^{(-\frac{1}{2}(c_5)^{n-1} + (n-1) \ln 17)}\right)^k \\
&\leq e^{-m/2} + \sum_{k \geq k_0} (e^{-(c_4)^{n-1}})^k \\
&= e^{-m/2} + \frac{e^{-k_0 c_4^{n-1}}}{1 - e^{-(c_4)^{n-1}}} \\
&\leq e^{-m/2} + \frac{e^{-\frac{m^{1/(n-1)}}{8c_6} c_4^{n-1}}}{1 - e^{-c_4^{n-1}}} \\
&\leq e^{-c_3^{n-1} m^{1/(n-1)}}. \quad \square
\end{aligned}$$

Acknowledgment This work was done in part while the authors were participating in the following programs:

- the *Probability, Geometry and Computation in High Dimensions* semester at the Simons Institute for the Theory of Computing,
- the *Interplay between High-Dimensional Geometry and Probability* trimester at the Hausdorff Institute for Mathematics,
- and the *Discrete Optimization* trimester at the Hausdorff Institute for Mathematics.

References

- [1] Michel L Balinski. The Hirsch conjecture for dual transportation polyhedra. *Mathematics of Operations Research*, 9(4):629–633, 1984.
- [2] Imre Bárány and Zoltán Füredi. On the shape of the convex hull of random points. *Probability Theory and Related Fields*, 77(2):231–240, February 1988. doi:10.1007/bf00334039.
- [3] David Barnette. Wv paths on 3-polytopes. *Journal of Combinatorial Theory*, 7(1):62–70, July 1969. doi:10.1016/s0021-9800(69)80007-4.
- [4] David Barnette. An upper bound for the diameter of a polytope. *Discrete Mathematics*, 10(1):9–13, 1974. doi:10.1016/0012-365x(74)90016-8.
- [5] Nicolas Bonifas, Marco Di Summa, Friedrich Eisenbrand, Nicolai Hähnle, and Martin Niemeier. On sub-determinants and the diameter of polyhedra. *Discrete & Computational Geometry*, 52(1):102–115, 2014.
- [6] Karl Heinz Borgwardt. *The simplex method: a probabilistic analysis*, volume 1 of *Algorithms and Combinatorics: Study and Research Texts*. Springer-Verlag, Berlin, 1987. doi:10.1007/978-3-642-61578-8.
- [7] Karl Heinz Borgwardt. Erratum: A sharp upper bound for the expected number of shadow vertices in lp-polyhedra under orthogonal projection on two-dimensional planes. *Mathematics of Operations Research*, 24(4):925–984, 1999. URL: <http://www.jstor.org/stable/3690611>.

- [8] Karl Heinz Borgwardt and Petra Huhn. A lower bound on the average number of pivot-steps for solving linear programs valid for all variants of the simplex-algorithm. *Mathematical Methods of Operations Research*, 49(2):175–210, April 1999. doi:10.1007/s186-1999-8373-5.
- [9] Steffen Borgwardt, Jesús A De Loera, and Elisabeth Finhold. The diameters of network-flow polytopes satisfy the Hirsch conjecture. *Mathematical Programming*, 171(1):283–309, 2018.
- [10] Andreas Brieden, Peter Gritzmann, Ravindran Kannan, Victor Klee, László Lovász, and Miklós Simonovits. Deterministic and randomized polynomial-time approximation of radii. *Mathematika*, 48(1-2):63–105, 2001.
- [11] Graham Brightwell, Jan Van den Heuvel, and Leen Stougie. A linear bound on the diameter of the transportation polytope. *Combinatorica*, 26(2):133–139, 2006.
- [12] Clément Canonne. A short note on poisson tail-bounds. Github repository link: <https://github.com/ccanonne/probabilitydistributiontoolbox/blob/master/poissonconcentration.pdf>, 2019.
- [13] Daniel Dadush and Nicolai Hähnle. On the shadow simplex method for curved polyhedra. *Discrete Computational Geometry*, 56(4):882–909, June 2016. doi:10.1007/s00454-016-9793-3.
- [14] Daniel Dadush and Sophie Huiberts. A friendly smoothed analysis of the simplex method. *SIAM Journal on Computing*, 49(5):STOC18–449, 2019.
- [15] Alberto Del Pia and Carla Michini. On the diameter of lattice polytopes. *Discrete & Computational Geometry*, 55(3):681–687, 2016.
- [16] Antoine Deza and Lionel Pournin. Improved bounds on the diameter of lattice polytopes. *Acta Mathematica Hungarica*, 154(2):457–469, 2018.
- [17] Martin Dyer and Alan Frieze. Random walks, totally unimodular matrices, and a randomised dual simplex algorithm. *Mathematical Programming*, 64(1):1–16, 1994.
- [18] Friedrich Eisenbrand, Nicolai Hähnle, Alexander Razborov, and Thomas Rothvoss. Diameter of polyhedra: Limits of abstraction. *Mathematics of Operations Research*, 35(4):786–794, 2010.
- [19] Marc Glisse, Sylvain Lazard, Julien Michel, and Marc Pouget. Silhouette of a random polytope. *Journal of Computational Geometry*, 7(1):14, 2016.
- [20] Richard C Grinold. The Hirsch conjecture in Leontief substitution systems. *SIAM Journal on Applied Mathematics*, 21(3):483–485, 1971.
- [21] Svante Janson. Large deviations for sums of partly dependent random variables. *Random Structures & Algorithms*, 24(3):234–248, 2004. doi:<https://doi.org/10.1002/rsa.20008>.
- [22] Gil Kalai and Daniel J. Kleitman. A quasi-polynomial bound for the diameter of graphs of polyhedra. *Bull. Amer. Math. Soc.*, 26(2):315–317, July 1992. doi:10.1090/s0273-0979-1992-00285-9.
- [23] Victor Klee, David W Walkup, et al. The d -step conjecture for polyhedra of dimension $d < 6$. *Acta Mathematica*, 117:53–78, 1967.

- [24] Peter Kleinschmidt and Shmuel Onn. On the diameter of convex polytopes. *Discrete mathematics*, 102(1):75–77, 1992.
- [25] Jean-Philippe Labbé, Thibault Manneville, and Francisco Santos. Hirsch polytopes with exponentially long combinatorial segments. *Mathematical Programming*, 165(2):663–688, 2017.
- [26] D.G. Larman. Paths on polytopes. *Proc. London Math. Soc. (3)*, s3-20(1):161–178, January 1970. doi:10.1112/plms/s3-20.1.161.
- [27] Carla Michini and Antonio Sassano. The Hirsch Conjecture for the fractional stable set polytope. *Mathematical Programming*, 147(1):309–330, 2014.
- [28] Denis Naddef. The Hirsch conjecture is true for $(0, 1)$ -polytopes. *Mathematical Programming: Series A and B*, 45(1-3):109–110, 1989.
- [29] Hariharan Narayanan, Rikhav Shah, and Nikhil Srivastava. A spectral approach to polytope diameter, 2021. arXiv:2101.12198.
- [30] A Reznikov and EB Saff. The covering radius of randomly distributed points on a manifold. *International Mathematics Research Notices*, 2016(19):6065–6094, 2016.
- [31] Laura Sanità. The diameter of the fractional matching polytope and its hardness implications. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 910–921. IEEE, 2018.
- [32] Francisco Santos. A counterexample to the Hirsch Conjecture. *Annals of Mathematics*, 176(1):383–412, July 2012. doi:10.4007/annals.2012.176.1.7.
- [33] Rolf Schneider and Wolfgang Weil. *Stochastic and integral geometry*. Probability and its Applications (New York). Springer-Verlag, Berlin, 2008. doi:10.1007/978-3-540-78859-1.
- [34] Daniel A Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM (JACM)*, 51(3):385–463, 2004.
- [35] Noriyoshi Sukegawa. An asymptotically improved upper bound on the diameter of polyhedra. *Discrete & Computational Geometry*, 62(3):690–699, 2019.
- [36] Michael J Todd. An improved Kalai–Kleitman bound for the diameter of a polyhedron. *SIAM Journal on Discrete Mathematics*, 28(4):1944–1947, 2014.
- [37] Roman Vershynin. Beyond Hirsch conjecture: walks on random polytopes and smoothed complexity of the simplex method. *SIAM J. Comput.*, 39(2):646–678, 2009. Preliminary version in FOCS ‘06. URL: <http://dx.doi.org/10.1137/070683386>, doi:10.1137/070683386.