Upper and Lower Bounds on the Smoothed Complexity of the Simplex Method

Sophie Huiberts*[1], Yin Tat Lee†[2,3], and Xinzhi Zhang‡[2]

[1]Columbia University
[2]University of Washington
[2]Microsoft Research

November 23, 2022

Abstract

The simplex method for linear programming is known to be highly efficient in practice, and understanding its performance from a theoretical perspective is an active research topic. The framework of smoothed analysis, first introduced by Spielman and Teng (JACM '04) for this purpose, defines the smoothed complexity of solving a linear program with d variables and n constraints as the expected running time when Gaussian noise of variance σ^2 is added to the LP data. We prove that the smoothed complexity of the simplex method is $O(\sigma^{-3/2} d^{13/4} \log^{7/4} n)$, improving the dependence on $1/\sigma$ compared to the previous bound of $O(\sigma^{-2} d^2 \sqrt{\log n})$. We accomplish this through a new analysis of the *shadow bound*, key to earlier analyses as well. Illustrating the power of our new method, we use our method to prove a nearly tight upper bound on the smoothed complexity of two-dimensional polygons.

We also establish the first non-trivial lower bound on the smoothed complexity of the simplex method, proving that the *shadow vertex simplex method* requires at least $\Omega\Big(\min\big(\sigma^{-1/2} d^{-1/2} \log^{-1/4} d, 2^d\big)\Big)$ pivot steps with high probability. A key part of our analysis is a new variation on the extended formulation for the regular 2^k-gon. We end with a numerical experiment that suggests this analysis could be further improved.

1 Introduction

Introduced by Dantzig [Dan47], the simplex method is one of the primary methods for solving linear programs (LP's) in practice and is an essential component in many software packages for combinatorial optimization. It is a family of local search algorithms which begin by finding a vertex of the set of feasible solutions and iteratively move to a better neighboring vertex along the edges of the feasible polyhedron until an optimal solution is reached. These moves are known as *pivot steps*. Variants of the simplex method can be differentiated by the choice of *pivot rule*, which determines which neighbouring vertex is chosen in each iteration, as well as by the method for obtaining the initial vertex. Some well-known pivot rules are the most negative reduced cost rule, the steepest edge rule, and an approximate steepest edge rule known as the devex rule. In theoretical work, the parametric objective rule, also known as the shadow vertex rule, plays an important role.

Empirical evidence suggests that the simplex algorithm typically takes $O(d+n)$ pivot steps, see [Sha87; And04; Gol94] and the references therein. However, obtaining a rigorous explanation for this excellent performance has proven challenging. In contrast to the practical success of the simplex method, all studied variants are known to have super-polynomial or even exponential worst-case running times. For deterministic

*sophie@huiberts.me.
†yintat@uw.edu.
‡xinzhi20@cs.washington.edu. Research supported by NSF grant CCF-1813135, NSF CAREER Award and Packard Fellowships.

variants, many published bad inputs are based on *deformed cubes*, see [KM72; Jer73; AC78; GS79; Mur80; Gol76] and a unified construction in [AZ98]. For randomized and history-dependent variants, bad inputs have been constructed based on *Markov Decision Processes* [Kal92; MSW96; HZ15; FHZ11; Fri11; DFH22]. The fastest provable (randomized) simplex algorithm takes $O(2^{\sqrt{d \log n}})$ pivot steps in expectation [Kal92; MSW96; HZ15].

Average-case analyses of the simplex method have been performed for a variety of random distributions over linear programs [Bor82; Bor87; Bor99; Sma83; Hai83; Meg86; AKS87; Tod86; AM85]. While insightful, the results from average-case analyses might not be fully realistic due to the fact that "random" linear programs tend to have certain properties that "typical" linear programs do not.

To better explain why simplex algorithm performs well in practice, while avoiding some of the pitfalls of average-case analysis, Spielman and Teng [ST04] introduced the smoothed complexity framework. For any base LP data $\bar{A} \in \mathbb{R}^{n \times d}, \bar{b} \in \mathbb{R}^n, c \in \mathbb{R}^d \backslash \{0\}$ where the rows of $(\bar{A}, \bar{b})$ are normalized to have ℓ_2 norm at most 1, they consider the *smoothed LP* by adding independent Gaussian perturbations to the constraints:

$$\max_{x \in \mathbb{R}^d} c^\top x \quad \text{subject to} \quad (\bar{A} + \hat{A})x \le (\bar{b} + \hat{b}).$$

The entries of $\hat{A}$ and $\hat{b}$ are i.i.d. Gaussian random variables with mean 0 and variance σ^2. The *smoothed complexity* of a simplex algorithm $\mathcal{A}$ is defined to be the maximum (over $\bar{A}, \bar{b}, c$) expected number of, i.e.,

$$\mathcal{SC}_{\mathcal{A},n,d,\sigma} := \max_{\substack{\bar{A} \in \mathbb{R}^{n \times d}, \bar{b} \in \mathbb{R}^n, c \in \mathbb{R}^d \\ \|[\bar{A}, \bar{b}]\|_{\infty,2} \le 1}} \left(\mathbb{E}_{\hat{A},\hat{b}} \left[T_{\mathcal{A}}(\bar{A} + \hat{A}, \bar{b} + \hat{b}, c) \right] \right).$$

Here $T_{\mathcal{A}}(A, b, c)$ is the number of pivot steps that the algorithm $\mathcal{A}$ takes to solve the linear program $\max_{x \in \mathbb{R}^d} \{c^\top x : Ax \le b\}$. We may note that if $\sigma \to \infty$ then $\mathcal{SC}_{\mathcal{A},n,d,\sigma}$ approaches the average-case complexity of $\mathcal{A}$ on independent Gaussian distributed input data. In constrast, if $\sigma \to 0$ then $\mathcal{SC}_{\mathcal{A},n,d,\sigma}$ will approach the worst-case complexity of $\mathcal{A}$. As a result, most interest has been directed at understanding the dependence on σ in the regime where $\sigma \ge 2^{-\Omega(d)}$ but $\sigma \le 1/\operatorname{poly}(d)$.

The motivation for smoothed analysis lies in the observation that the above-mentioned worst-case instances are very "brittle" to perturbations, and computer implementations require great care in handling numerical inaccuracies to obtain the theorized running times even on problems with a small number of variables. When implemented with a larger number of variables, the limited accuracy of floating-point numbers make it impossible to reach the theorized running times.

An algorithm is said to have polynomial smoothed complexity if under the perturbation of constraints, it has expected running time $\operatorname{poly}(n, d, \sigma^{-1})$, and [ST04] proved that the smoothed complexity of the shadow simplex algorithm (which we will describe next) is at most $O(d^{55} n^{86} \sigma^{-30} + d^{70} n^{86})$. The best bound available in the literature is $O(\sigma^{-2} d^2 \sqrt{\log n})$ pivot steps due to [DH18], assuming $\sigma \le 1/\sqrt{d \log n}$. We note that assuming an upper bound on σ can be done without loss of generality; its influence can be captured as an additive term in the upper bound that does not depend on σ.

This work improves the dependence on σ of the smoothed complexity, obtaining an upper bound of $O(\sigma^{-3/2} d^{13/4} \log^{7/4} n)$ for $\sigma \le 1/d\sqrt{\log n}$. As a second contribution, we prove the first non-trivial lower bound on the smoothed complexity of a simplex method, finding that the shadow vertex simplex method requires $\Omega(\min(\frac{1}{\sqrt{\sigma d \sqrt{\log n}}}, 2^d))$ pivot steps.

Shadow Vertex Simplex Algorithm One of the most extensively studied simplex algorithms in theory is the shadow vertex simplex algorithm [GS55; Bor82]. Given an LP

$$\max_{x \in \mathbb{R}^d} c^\top x, \quad Ax \le b,$$

for $A \in \mathbb{R}^{n \times d}, b \in \mathbb{R}^n, c \in \mathbb{R}^d$, let $P = \{x \in \mathbb{R}^d : Ax \le b\}$ denote the feasible polyhedron of the linear program. The algorithm starts from an initial vertex $x_0 \in P$ that optimizes an initial objective c_0[1]. During the execution, it maintains an intermediate objective $c_\lambda = \lambda c + (1 - \lambda) c_0$ and a vertex that optimizes c_λ.

[1]There are many standard methods of finding such initialization with at most $O(d)$ overhead in running time, so we can assume that both x_0 and c_0 are already given. See the discussions in [DH18]

Thus by slowly increasing λ from 0 to 1 during different pivot steps, the temporary objective gradually changes from c_0 to c, revealing the desired solution at the end. Since each pivot step requires $\mathrm{poly}(d, n)$ computational work, theoretical analysis has focused on analyzing the number of pivot steps.

The algorithm is called shadow vertex simplex method because, after taking orthogonal projection of the feasible set onto the two-dimensional linear subspace $W = \mathrm{span}(c_0, c)$, the vertices visited by the algorithm project onto the boundary of the projection ("shadow") $\pi_W(P)$. Assuming certain non-degeneracy conditions, which will hold with probability 1 for the distributions we consider, this projection gives an injective relation between iterations of the method and vertices of the shadow is injective, meaning that we can upper bound the number of pivot steps in the algorithm by the number of vertices of the shadow polygon. This characterization makes the shadow vertex simplex method ideally suited for probabilistic analysis.

To analyze the "shadow size", the number of vertices of the shadow polygon, we follow earlier work and reduce to the case that $b = 1$ by [Ver09]. In this case, well-established principles of polyhedral duality show that

$$\mathrm{vertices}(\pi_W(P)) \leq \mathrm{edges}(W \cap \mathrm{conv}(0, a_1, \ldots, a_n)) \leq \mathrm{edges}(W \cap \mathrm{conv}(a_1, \ldots, a_n)) + 1.$$

Here, $P = \{x \in \mathbb{R}^d : Ax \leq \mathbf{1}_n\}$ for any matrix A with rows $a_1, \ldots, a_n$.

The smoothed complexity of shadow vertex simplex algorithm can thus be reduced to the smoothed complexity of a two-dimensional slice of a convex hull. For this purpose, let us define the maximum smoothed shadow size as

$$\mathcal{S}(n, d, \sigma) = \max_{\substack{\bar{a}_1, \cdots, \bar{a}_n \in \mathbb{R}^d \\ \max_{i \in [n]} \|\bar{a}_i\|_2 \leq 1 \\ W \subset \mathbb{R}^d}} \mathbb{E}_{\tilde{a}_1, \cdots, \tilde{a}_n \sim \mathcal{N}(0, \sigma^2)} \left[\mathrm{edges}\left(\mathrm{conv}(\bar{a}_1 + \tilde{a}_1, \cdots, \bar{a}_n + \tilde{a}_n) \cap W\right)\right] \tag{1}$$

The following upper bound we take from [DH18], who state that the analysis of [Ver09] can be strengthened to obtain the claimed bound. This upper bound should be understood as proving that there exists a shadow vertex rule based simplex algorithm which satisfies that smoothed complexity bound. The lower bound is due to [Bor87] and shows that the shadow vertex simplex rule can be made to follow paths of this length.

Theorem 1 (Smoothed Complexity of Shadow Vertex Simplex Algorithm)**.** *Given any $n \geq d \geq 2, \sigma > 0$, the smoothed complexity of the shadow vertex simplex algorithm satisfies*

$$\mathcal{S}(n, d, \sigma)/4 \leq \mathcal{SC}_{\textsc{ShadowSimplex}, n, d, \sigma} \leq 2 \cdot \mathcal{S}\left(n + d, d, \min(\sigma, \frac{1}{\sqrt{d}\log d}, \frac{1}{\sqrt{d \log n}})\right) + 4.$$

With this reduction, analyzing the smoothed complexity of the simplex method comes down to bounding the smoothed shadow size $\mathcal{S}(n, d, \sigma)$. As such, that will be the focus of the remainder of this paper.

1.1 Our Results

The previous best shadow bound is due to [DH18], who prove that $\mathcal{S}(n, d, \sigma) \leq O(d^2\sqrt{\log n}\sigma^{-2})$. We strengthen this result for small values of σ.

Theorem 2. *For $n \geq d \geq 3$ and $\sigma \leq \frac{1}{8d\sqrt{\log n}}$, the smoothed shadow size satisfies*

$$\mathcal{S}(n, d, \sigma) = O\left(\sigma^{-3/2} d^{13/4} \log^{5/4} n\right).$$

A full overview of bounds on the smoothed shadow size, including previous results in the literature, can be found in Table 1.

Second, we prove the first non-trival lower bound on the smoothed shadow size, establishing that $\mathcal{S}(4d - 15, d, \sigma) \geq \Omega(\min(\frac{1}{\sqrt{\sigma d \sqrt{\log d}}}, 2^d))$ for $d \geq 5$. This is lower bound is proven by constructing a polyhedron $P = \{x \in \mathbb{R}^d : Ax \leq \mathbf{1}_n\}$ and a two-dimensional subspace W such that any small perturbation of P, projected onto W, will have many vertices. The construction is based on an extended formulation similar to those first constructed by [BN01; Gli00].

Theorem 3. *For $d \geq 5$ and $\sigma \leq \frac{1}{480d\sqrt{\log(4d)}}$, the smoothed shadow size satisfies*

$$\mathcal{S}(4d-15, d, \sigma) = \Omega\left(\min\left(\frac{1}{\sqrt{d\sigma\sqrt{\log d}}}, 2^d\right)\right).$$

We remark that [DGGT16] showed a lower bound of $\Omega(\min(\sqrt{\log n} + \frac{\sqrt[4]{\log(n\sqrt{\sigma})}}{\sqrt{\sigma}}, n))$ expected vertices for the two-dimensional polygons, obtained by placing $\bar{a}_1, \ldots, \bar{a}_n$ equally spaced on the unit circle. We thus nearly match their dependence on σ for the analogous two-dimensional question. However, in two-dimensional case, one can only have a lower bound of $\Omega(n)$ even if there isn't any perturbation. It is only when we try to work in higher dimensions that we can show an exponential lower bound of $\Omega(2^d)$ any value of σ.

Also, it is possible that the exponents of σ in our bound can be further optimized. In Section 6.7, we describe numerical experiments which suggest that the actual shadow size for perturbations of our constructed polytope might be as high as $\Omega(\min(\sigma^{-3/4}, 2^d))$.

Reference	**Smoothed shadow size**	**Model**
[Bor87]	$\Theta(d^{3/2}\sqrt{\log n})$	Average-case, Gaussian distribution
[ST04]	$O(\sigma^{-6}d^3 n + d^6 n \log^3 n)$	Smooth
[DS05]	$O(\sigma^{-2}dn^2 \log n + d^2n^2 \log^2 n)$	Smooth
[Ver09]	$O(\sigma^{-4}d^3 + d^5 \log^2 n)$	Smooth
[DH18]	$O(\sigma^{-2}d^2\sqrt{\log n} + d^3 \log^{1.5} n)$	Smooth
This paper	$O(\sigma^{-3/2}d^{13/4} \log^{7/4} n + d^{19/4} \log^{13/4} n)$	Smooth
This paper	$\Omega(\min(\frac{1}{\sqrt{\sigma d\sqrt{\log d}}}, 2^d))$	Smooth

Table 1: Bounds of expected number of pivots in previous literature, assuming $d \geq 3$. Logarithmic factors are simplified. The lower bound of [Bor87] holds in the smoothed model as well.

Two-dimensional polygons To better understand the smoothed complexity of the intersection polygon $\mathrm{conv}(a_1, \ldots, a_n) \cap W$, we also analyze its two-dimensional analogue. Taking $\bar{a}_1, \ldots, \bar{a}_n \in \mathbb{R}^2$, each satisfying $\|a_i\|_2 \leq 1$, we are interested in the number of edges of the smoothed polygon $\mathrm{conv}(\bar{a}_1 + \hat{a}_1, \ldots, \bar{a}_n + \hat{a}_n)$, where $\hat{a}_1, \ldots, \hat{a}_n \sim N(0, \sigma^2)$ are independent. The previous best upper bound on the smoothed complexity of this polygon $O(\sigma^{-1} + \sqrt{\log n})$, due to [DH21]. Their analysis is based on an adaptation of the shadow bound by [DH18]. In Section 4 we improve this upper bound, obtaining the following theorem.

Theorem 4 (Two-Dimensional Upper Bound)**.** *Let $\bar{a}_1, \cdots, \bar{a}_n \in \mathbb{R}^2$ be $n > 2$ vectors with norm at most 1. For each $i \in [n]$, let a_i be independently distributed as $\mathcal{N}(\bar{a}_i, \sigma^2 I_{2\times 2})$. Then*

$$\mathbb{E}\left[\mathrm{edges}\left(\mathrm{conv}(a_1, \cdots, a_n)\right)\right] \leq O\left(\frac{\sqrt[4]{\log n}}{\sqrt{\sigma}} + \sqrt{\log n}\right).$$

Combined with the trivial upper bound of n vertices, this bound nearly matches the lower bound of $\Omega(\min(\sqrt{\log n} + \frac{\sqrt[4]{\log(n\sqrt{\sigma})}}{\sqrt{\sigma}}, n))$ in [DGGT16]. A full overview of previous results on the smoothed complexity of the two-dimensional convex hull can be found in Table 2.

Reference	Smoothed polygon complexity
[DS04]	$O(\log(n)^2 + \sigma^{-2}\log n)$
[Sch14]	$O(\log n + \sigma^{-2})$
[DGGT16]	$O(\sqrt{\log n} + \sigma^{-1}\sqrt{\log n})$
[DH21]	$O(\sqrt{\log n} + \sigma^{-1})$
This paper	$O(\sqrt{\log n} + \frac{\sqrt[4]{\log(n)}}{\sqrt{\sigma}})$
[DGGT16]	$\Omega(\min(\sqrt{\log n} + \frac{\sqrt[4]{\log(n\sqrt{\sigma})}}{\sqrt{\sigma}}, n))$

Table 2: Bounds on the smoothed complexity of a two-dimensional polygon.

1.2 Related work

Shadow Vertex Simplex Method The shadow vertex simplex algorithm has played a key role in many analyses of simplex and simplex-like algorithms. On well-conditioned polytopes, such as those of the form $\{x \in \mathbb{R}^d : Ax \leq b\}$ where A is integral with subdeterminants bounded by Δ, the shadow vertex method has been studied by in [DH16; BR13]. The shadow vertex method on polytopes all whose vertices are integral was studied in [BDKS21; Bla22]

On random polytopes of the form $\{x \in \mathbb{R}^d : Ax \leq \mathbf{1}_n\}$, assuming the constraint vectors are independently drawn from any rotationally symmetric distribution, the expected iteration complexity of the shadow vertex simplex method was studied by [Bor82; Bor87; Bor99]. In the case when the rows of A arise from a Poisson distribution on the unit sphere, concentration results and diameter bounds were proven in [BDGHL22]. The diameter of smoothed polyhedra was studied by [NSS22], who used the shadow bound of [DH18] to show that most vertices, according to some measure, are connected by short paths.

A randomized algorithm for solving linear programs in weakly polynomial time, inspired by the shadow vertex simplex method, was proposed in [KS06]. The shadow vertex algorithm was recently used as part of the analysis of an interior point method by [ADLNV22].

Extended Formulations For a polyhedron $P \subset \mathbb{R}^d$, an *extended formulation* is any polyhedron $Q \subset \mathbb{R}^{d'}$, $d' \geq d$, such that P can be obtained as an orthogonal projection of Q to some d-dimensional subspace. Importantly, Q can have much fewer facets than P. While there is a wider literature on extended formulations, here we describe only what is most relevant to the construction in Section 6.

The construction in our lower bound is based on an adaptation of the extended formulation by [BN01] of the regular 2^k-gon. They used this construction to obtain a polyhedral approximiation of the second order cone $\{x \in \mathbb{R}^{n+1} : \sum_{i=1}^n x_i^2 \leq x_{n+1}^2\}$. A variant on their construction using fewer variables and inequalities was given by [Gli00]. A more general construction based on *reflection relations* is used to construct extended formulation for the regular 2^k-gon, as well as other polyhedra, in [KP13]. Extended formulations for regular n-gons, when n is not a power of 2, can be found in [VGG17].

Approximations of the second order cone based on the work of [BN01; Gli00] have been used to solve second order conic programs, see, e.g., [BHMPT15]. These approximations were included in the solver `SCIP` until version 7.0 [Gam+20a; Gam+20b].

1.3 Proof Overview

1.3.1 Smoothed Complexity Upper Bound

We write the random polytope as $Q = \text{conv}(a_1, \cdots, a_d)$ where each a_i is sampled independently from $\mathcal{N}_d(\bar{a}_i, \sigma^2 I)$ such that $\|\bar{a}_i\| \leq 1$. Our goal is to upper bound the expected number of edges of the polygon $Q \cap W$ for any two-dimensional plane $W \subset \mathbb{R}^d$ and $\bar{a}_1, \cdots, \bar{a}_n$. This will immediately give us an upper bound of $\mathcal{S}(n, d, \sigma)$.

A classical result from smoothed analysis [ST04] states that the intersection polygon $Q \cap W$ is *non-degenerate* almost surely: Every edge on $Q \cap W$ is uniquely given by the intersection between W and a facet of Q spanned by exactly d vertices. For any index set $I \in \binom{[n]}{d}$, write E_I as the event that $\mathrm{conv}(a_i : i \in I) \cap W$ is an edge of $Q \cap W$. Non-degeneracy implies that every edge of $Q \cap W$ uniquely corresponds to an index set $I \in \binom{[n]}{d}$ such that E_I holds.

Before sketching our proof, we briefly review the approach of [DH18], and then discuss the main technical obstacles of obtaining a upper bound with better dependence on σ.

As a first step, they replace the Gaussian distribution with what they dub a Laplace-Gaussian distribution. The latter distribution approximates the probability density of the former, in particular having nearly-equal smoothed shadow size, while being $O(\sigma^{-1}\sqrt{d \log n})$-log-Lipschitz for any point on its domain. A probability distribution with probability density function μ is L-log-Lipschitz for some $L > 0$, if for any $x, y \in \mathbb{R}^d$, we have $|\log(\mu(x)) - \log(\mu(y))| \le L\|x - y\|$.

Next, conditional on E_I, write ℓ_I as the length of the edge on $Q \cap W$ that corresponds to I. [DH18] first showed that, for any family $S \subset \binom{[n]}{d}$, the expected number of edges of $Q \cap W$ coming from S is at most

$$\mathbb{E}\left[\sum_{I \in S} E_I\right] \le \frac{\mathbb{E}[\mathrm{perimeter}(Q \cap W)]}{\min_{I \in S} \mathbb{E}[\ell_I \mid E_I]}. \tag{2}$$

Taking $S = \{I \subset \binom{[n]}{d} : \Pr[E_I] \ge \binom{n}{d}^{-1}\}$, they find that the expected number of edges of $Q \cap W$ is at most

$$\mathbb{E}[\mathrm{edges}(Q \cap W)] \le 1 + \mathbb{E}\left[\sum_{I \in S} E_I\right] \le 1 + \frac{\mathbb{E}[\mathrm{perimeter}(Q \cap W)]}{\min_{I \in S} \mathbb{E}[\ell_I \mid E_I]}. \tag{3}$$

To upper bound the numerator of (3), they notice that $Q \cap W$ is a convex polygon contained in the two-dimensional disk centered at $\mathbf{0}$ with radius $\max_{i \in [n]} \|\pi_W(a_i)\|$. It then follows from convexity that the perimeter of $Q \cap W$ is at most the perimeter of such disk, i.e.

$$\mathbb{E}[\mathrm{perimeter}(Q \cap W)] \le 2\pi \cdot \mathbb{E}[\max_{i \in [n]} \|\pi_W(a_i)\|] \le 2\pi \cdot (1 + 4\sigma\sqrt{\log n}) \tag{4}$$

where the last step comes from Gaussian tail bound. For the denominator of (3), [DH18] showed for any $I \in \binom{[n]}{d}$ with $\Pr[E_I] \ge \binom{n}{d}^{-1}$ that, conditional on E_I, the expected edge length is at least

$$\mathbb{E}[\ell_I \mid E_I] \ge \Omega\left(\frac{\sigma^2}{d^2\sqrt{\log n}} \cdot \frac{1}{1 + \sigma\sqrt{d \log n}}\right) \tag{5}$$

Combining the two parts together, we get an upper bound of $O(\sigma^{-2} d^2 \sqrt{\log n} + d^3 \log^{1.5} n)$.

New Strategy of Counting Edges While [DH18] made the best analysis based of their edge-counting strategy (3), the strategy itself is sub-optimal in a fairly obvious way. A main drawback is that using the minimum expected length of edge $\min_{I \in \binom{[n]}{d}} \mathbb{E}[\ell_I \mid E_I]$ at the denominator of (3) is too pessimistic when the edges of $Q \cap W$ are long. Consider the case where an edge on $Q \cap W$ have length $\Omega(1)$ at the beginning. After the perturbation, is very likely that the length of this edge is still $\Omega(1)$, but [DH18] uses a lower-bound of $\Omega(\frac{\sigma^2}{d^2\sqrt{\log n}})$.

To improve this, we use a new edge-counting strategy that can handle the long and short edges separately with two different ways of counting the edges. Take any index set $I \in \binom{[n]}{d}$; conditional on E_I, we write e_I for its edge $\mathrm{conv}(a_i : i \in I) \cap W$. The next edge in clockwise direction we call e_{I+} and say it has length ℓ_{I+}. We say e_{I+} is likely to be long, if $\Pr[\ell_{I+} \ge t \mid E_I] \ge 0.05$ for some parameter $t > 0$. In this case, we will upper bound the number of such edges following a similar strategy as (3), which yields

$$\mathbb{E}[\text{number of } e_I \text{ s.t. } e_{I+} \text{ is likely to be long}] \le \frac{\mathbb{E}[\mathrm{perimeter}(P \cap W)]}{t} \le \frac{(2\pi + O(\sigma\sqrt{\log n}))}{t}.$$

where the second step uses the exact same upper bound of $\mathbb{E}[\mathrm{perimeter}(P \cap W)]$ as in (4).

In the other case, e_{I+} is unlikely to be long, i.e. $\Pr[\ell_{I+} \geq t \mid E_I] < 0.05$. Now we will upper bound the number of such edges by claiming that their exterior angles each are large in expectation. Let θ_I to be the exterior angle at the endpoint of $\mathrm{conv}(a_i : i \in I) \cap W$ that comes last in clockwise order. Our key observation is that $\sin(\theta_I) \cdot \ell_{I+}$ equals to the distance from the second vertex of e_{I+} to the extension line of e_I. So when the edge e_{I+} is likely to be short, a lower bound on this distance will imply a lower bound on $\mathbb{E}[\theta_I \mid E_I]$. See Figure 1 for an illustration. More formally, let p_{I+} denote the next vertex of $Q \cap W$ after e_I in clockwise order. suppose that for any $I \in \binom{n}{d}$,

$$\Pr[\mathrm{dist}(p_{I+}, \mathrm{affhull}(e_I)) \geq \gamma \mid E_I] \geq 0.1. \tag{6}$$

Then conditional on E_I, we have $\sin(\theta_I) \geq \frac{\gamma}{t}$ with probability at least 0.05, and the expectation of the exterior angle at the shared endpoint of e_I and e_{I+} is at least $\frac{\gamma}{20t}$.

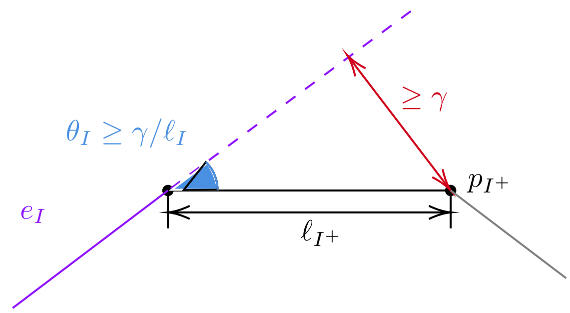

Figure 1: Illustration of the case when e_{I+} is short. In purple is the edge e_I, and the next edge in clockwise direction has length ℓ_{I+}. In red is the edge-to-vertex distance $\mathrm{dist}(p_{I+}, \mathrm{affhull}(e_I))$, and in blue is the angle θ_I. Suppose $\mathrm{dist}(p_{I+}, \mathrm{affhull}(e_I)) \geq \gamma$, then $\theta_I \geq \gamma/\ell_{I+}$.

On the other hand, the sum of exterior angles of a polygon equals to 2π. Therefore we can upper bound the number of short edges in expectation by at most

$$\mathbb{E}[\text{number of } e_I \text{ s.t. } e_{I+} \text{ is unlikely to be long}] \leq O(\frac{t}{\gamma}).$$

Summing up the number of edges in the two cases, we get an upper bound of the expected edge-count of $Q \cap W$ by at most

$$\mathbb{E}[\mathrm{edges}(Q \cap W)] \leq \frac{2\pi + O\sigma\sqrt{\log n})}{t} + O\left(\frac{t}{\gamma}\right) = O\left(\sqrt{\frac{1 + \sigma\sqrt{\log n}}{\gamma}}\right) \tag{7}$$

where the final step follows from optimizing $t > 0$. We summarize our result in Theorem 27. For details of the edge-counting strategy, see Section 3.

Two-dimensional Upper Bound In the second part of our proof, we need to show a lower bound of the expected distance from the affine hull of an edge of $Q \cap W$ to the next vertex in clockwise order, which is the quantity γ mentioned in (6).

As a warm-up, we first introduce our proof in $\mathbb{R}^2$, which will be explained in Section 4 in detail. In this case, W will become the entire two-dimensional space and will disappear. Therefore, we can focus on lower-bounding the distance from any edge e of the polygon $Q = \mathrm{conv}(a_1, \cdots, a_n)$ to any other vertices, i.e. it suffices to show that for any $I \in \binom{[n]}{2}$,

$$\Pr[\mathrm{dist}(\mathrm{conv}(a_j : j \notin I), \mathrm{affhull}(e_I)) \geq \gamma \mid E_I] \geq 0.1.$$

We can obtain a lower bound on this quantity for any L-log-Lipschitz distribution. Through an appropriate coordinate transformation we prove that, irrespective of the values of $a_j, j \notin I$, the distance $\mathrm{dist}(\mathrm{conv}(a_j, j \notin$

$I), \text{affhull}(e_I))$, conditional on being non-zero, follows a $2L$-log-Lipschitz distribution. We can directly calculate that we may choose $\gamma = \Omega(1/L)$. This result can be applied immediately to our Gaussian random variables $a_1, \cdots, a_n$, since each $a_i \sim \mathcal{N}(\bar{a}_i, \sigma I_{2\times 2})$ is $L = O(\sigma\sqrt{\log n})$-log-Lipschitz with overwhelming probability. Plugging into (7), we get that in the two-dimensional case, $\mathbb{E}[\text{edges}(Q)] \le O(\sqrt[4]{\log n}/\sqrt{\sigma} + \sqrt{\log n})$ as in Theorem 31.

Multi-Dimensional Upper Bound As in the two-dimensional case, it remains to lower-bound of the edge-to-vertex distance $\text{dist}(p_{I^+}, \text{affhull}(e_I))$ (see (6)) of $Q \cap W$. Analyzing this, however, becomes more challenging. In two-dimensional case, each edge is the convex hull of two vertices among $a_1, \cdots, a_n$ and is independent to the other vertices on $Q \cap W$. In contrast, if $d \ge 3$ then each edge on $Q \cap W$ will be the intersection between W and a $(d-1)$-dimensional facet of Q (which is the convex hull of d vertices), and each vertex will be the intersection between W and a $(d-2)$-dimensional ridge of Q (which is the convex hull of $d-1$ vertices). So the distribution of e_I and p_{I^+} are correlated.

To overcome these difficulties, we proposed a technique that first factors $\text{dist}(p_{I^+}, \text{affhull}(e_I))$ into the product of separate parts which are easier to analyze, and then use log-Lipschitzness of $a_1, \cdots, a_n$ to lower-bound each part with good probability. Fix without loss of generality $e = e_{[d]} = \text{conv}(a_1, \cdots, a_d) \cap W$, as the potential edge of interest. Consider the second endpoint p on e in clockwise direction and let $J \in \binom{[d]}{d-1}$ be the index set such that $\{p\} = \text{conv}(a_j : j \in J) \cap W$. Let $p' = \text{conv}(a_i : i \in J') \cap W$ (with $J' = \binom{[n]}{d-1}$) be the node next to the edge e in clockwise direction. From the non-degeneracy conditions, we know that J' only differs to J with two vertices almost surely, so we can assume without loss of generality that $J = \{a_2, \cdots, a_d\}$ and $J' = \{a_3, \cdots, a_d\} \cap \{a_k\}$ for some $k \in \{d+1, \cdots, n\}$.

The main idea of our analysis is the observation that if the radius of Q is bounded above by $O(1)$ (which happens with overwhelming probability due to Gaussian tail bound), then we can lower bound the two-dimensional edge-to-vertex distance $\text{dist}(p', \text{affhull}(e))$ by the product of two distances $\Omega(\delta \cdot r)$, where

- δ is the d-dimensional distance from the facet $\text{affhull}(a_1, \cdots, a_d)$ containing e, to the vertices that are not in the facet, i.e.
$$\delta = \text{dist}(\text{conv}(a_{d+1}, \cdots, a_n), \text{affhull}(a_1, \cdots, a_d)));$$
- r is the distance from the boundary of the ridge $\partial\, \text{conv}(a_2, \cdots, a_d)$ to the one-dimensional line $\text{affhull}(e)$, i.e. $r = \text{dist}(\text{affhull}(e), \partial\, \text{conv}(a_2, \cdots, a_d))$.

We will give the formal statement of the distance splitting lemma in Lemma 36.

We also remark that after conditioning on a fixed $(d-1)$-dimensional plane $\text{affhull}(a_1, \cdots, a_d)$ such that plane is at the exterior of Q. Then δ will solely depend on the position of a_k, and r will solely depends on the location of $a_1, \cdots, a_d$ on their affine plane. Therefore we can treat δ and r as independent variables after specifying $\text{affhull}(a_1, \cdots, a_d)$.

It then remains to show that r and δ are both unlikely to be too small. Similar to the two-dimensional case, we will also use log-Lipschitzness of $a_1, \cdots, a_d$ as our main tool.

- After specifying $\text{affhull}(a_1, \ldots, a_d)$, the lower bound on δ is derived from the remaining randomness in $a_{d+1}, \ldots, a_n$. Here we use both the L-log-Lipschitzness of the distributions of $a_{d+1}, \ldots, a_n$, as well as the knowledge that we only need to consider hyperplanes $\text{affhull}(a_1, \ldots, a_d)$ which are likely to have all points $a_{d+1}, \ldots, a_n$ on one side. This is made precise in Section 5.3.

- The lower bound of r resembles the proof of the "distance lemma" of [ST04]. First we show that each vertex of the ridge $\text{conv}(a_2, \cdots, a_d)$ is $\Omega(1/d^2 L)$-far away from the plane spanned by its other vertices, after projected onto $\text{affhull}(e)^\perp$. In the second step, we show that suppose we write $p = \sum_{i\in[d]} \lambda_i a_i$ as the convex combination, then with constant probability $\min_{i\in[d]} \lambda_i \ge \Omega(1/d^2 L)$. Combining the two steps, we get that $r \ge \Omega(1/d^4 L^2)$ with good probability. See Section 5.4 for more details.

We conclude our main result of the edge-to-vertex distance lower bound in Lemma 35. Readers are referred to Section 5 for detailed discussions.

1.3.2 Smoothed Complexity Lower Bound

Our smoothed complexity lower bound (Theorem 47) is based on two geometric observations using the inner and outer radius of the perturbed polytope. For a polytope P and a unit norm ball $\mathbb{B}$, its outer radius with center x is the smallest R such that there exists $P \subset R \cdot \mathbb{B} + x$. Its inner radius with center x is the largest r such that $r \cdot \mathbb{B} + x \subset P$.

The first observation is that, if a two-dimensional polygon T has inner ℓ_2-radius of r and outer ℓ_2-radius of $(1+\varepsilon) \cdot r$ with respect to the same center, then T has at least $\Omega(\varepsilon^{-1/2})$ edges (Lemma 56). This comes from the fact that every edge of T has length at most $O(r\sqrt{\varepsilon})$, whereas the perimeter of T is at least $2\pi r$.

Second, if two polytopes $Q, \tilde{Q} \subseteq \mathbb{R}^d$, each with inner radius t, have Hausdorff distance $\varepsilon < t/2$ to each other, then Q will approximate $\tilde{Q}$ in the way that (Lemma 55)

$$(1 - 2\varepsilon/t) \cdot Q \subseteq \tilde{Q} \subseteq (1 + \varepsilon/t) \cdot Q.$$

In particular, for any two-dimensional linear subspace W we have

$$(1 - O(\varepsilon)) \cdot Q \cap W \subseteq \tilde{Q} \cap W \subseteq (1 + O(\varepsilon)) \cdot Q \cap W. \tag{8}$$

To prove our lower bound, we construct a polytope $Q = \mathrm{conv}(\bar{a}_1, \cdots, \bar{a}_n) \subset \mathbb{R}^d$ and a two-dimensional linear subspace W such that $\Omega(1)B_1^d \subset Q \subset B_1^d$, and where $Q \cap W$ has inner radius $\frac{r}{(1+4^{-d})}$ and outer radius $r > 0$. Perturbing the vertices of Q, we obtain $\tilde{Q} = \mathrm{conv}(a_1, \cdots, a_n)$, where $a_i \sim N(\bar{a}_i, \sigma^2 I_{d\times d})$ for each $i \in [n]$. Note that $Q \subset \mathbb{B}_1^d$ implies that $\bar{a}_1, \ldots, \bar{a}_n$ satisfy the normalization requirement in (1). With high probability the Hausdorff distance in ℓ_1 between Q and $\tilde{Q}$ is bounded by $\max_{i\in[n]} \|a_i - \tilde{a}_i\|_1 \leq O(\sigma d\sqrt{\log n})$. Using (8), we bound the inner and outer radius of $\tilde{Q} \cap W$. A lower bound on the number of edges of $\tilde{Q} \cap W$ follows from Lemma 56 as described above.

We remark that the polytopes $Q = \mathrm{conv}(a_1, \ldots, a_n) \subset \mathbb{R}^d$ with $n = O(d)$ and two-dimensional subspaces W such that $Q \cap W$ has inner ℓ_2-radius $\frac{r}{1+4^{-d}}$ and outer ℓ_2-radius $r > 0$ were first obtained by [BN01] as an *extended formulation* for a regular 2^k-gon with $O(k)$ variables and $O(k)$ inequalities. Their polytope, however, has an outer and inner radius that differ by a factor $2^{\Omega(k)}$, meaning that we cannot apply Lemma 55 for $\sigma > 2^{-k}$. We construct an alternative such extended formulation where the ratio between inner and outer ℓ_1-radius is only $O(1)$. With an appropriate scaling to get $Q \subset B_1^d$, we find that the perturbed polytope $\tilde{Q}$ will have intersection $\tilde{Q} \cap W$ with inner radius $\frac{r}{1+4^{-d}}(1-\varepsilon)$ and outer radius $(1+\varepsilon)r$, where $\varepsilon = O(\sigma d\sqrt{\log n})$, and thus has $\Omega(\min(\frac{1}{\sqrt{\varepsilon}}, 2^d))$ edges, with high probability.

2 Preliminaries

We write $\mathbf{1}_n$ for the all-ones vector in $\mathbb{R}^n$, $\mathbf{0}_n$ for the all-zeroes vector in $\mathbb{R}^n$, and $I_{n\times n}$ for the n by n identity matrix. The standard basis vectors are denoted by $e_1, \ldots, e_n \in \mathbb{R}^n$. For a linear subspace $W \subset \mathbb{R}^n$ we denote the orthogonal projection onto W by π_W. The subspace of vectors orthogonal to a given vector $\omega \in \mathbb{R}^n$ is denoted $\omega^\perp$.

For a vector $x \in \mathbb{R}^n$, the ℓ_1 norm is $\|x\|_1 = \sum_{i\in[n]} |x_i|$, the ℓ_2-norm is $\|x\|_2 = \sqrt{\sum_{i\in[n]} x_i^2}$ and the ℓ_∞-norm is $\|x\|_\infty = \max_{i\in[n]} |x_i|$. A norm without a subscript is always the ℓ_2-norm. Given $p > 0, d \in \mathbb{Z}_+$, define $\mathbb{B}_p^d = \{x \in \mathbb{R}^d : \|x\|_p \leq 1\}$ as the d-dimensional unit ball of ℓ_p norm.

We write $[n] := \{1, \ldots, n\}$. The convex hull of vectors $a_1, \ldots, a_n$ is denoted $\mathrm{conv}(a_1, \ldots, a_n) = \mathrm{conv}(a_i : i \in [n])$, and similarly the affine hull as $\mathrm{affhull}(a_i : i \in [n])$.

2.1 Polytopes

Definition 5 (Polytope)**.** *A convex body $P \subset \mathbb{R}^d$ is called a polytope, if it can be written as $P = \{x \in \mathbb{R}^d : Ax \leq b\}$, for some $n \in \mathbb{Z}_+$ $A \in \mathbb{R}^{n\times d}, b \in \mathbb{R}^n$.*

Definition 6 (Valid Condition and Facet)**.** *Given a polytope $P \subset \mathbb{R}^d$, vector $c \in \mathbb{R}^d$ and $d \in \mathbb{R}$, we say the linear condition $x^\top c \leq d$ is valid for P if the condition holds for all $x \in P$.*

A subset $F \subset P$ is called a face of P, if $F = P \cap \{x \in \mathbb{R}^d : x^\top c = d\}$, for some valid condition $x^\top c \leq d$. A facet is a $d-1$-dimensional face, a ridge is a $d-2$-dimensional face, an edge is a 1-dimensional face and a vertex is a 0-dimensional face.

Definition 7 (Polar dual of a convex body)**.** *Let $P \subset \mathbb{R}^d$ be a convex sets. Define the polar dual of P as*

$$P^\circ = \{y \in \mathbb{R}^d : y^\top x \leq 1, \forall x \in P\}.$$

We state some basic facts from duality theory:

Fact 8 (Polar dual of polytope)**.** *Let $P \subset \mathbb{R}^d$ be a polytope given by the linear system $P = \{x \in \mathbb{R}^d, Ax \leq \mathbf{1}_n\} \subset \mathbb{R}^d$ for some $A \in \mathbb{R}^{n\times d}$. Then the polar dual of P equals to*

$$P^\circ := \mathrm{conv}(\mathbf{0}_d, a_1, a_2, \cdots, a_n).$$

where $a_1, \cdots, a_n \in \mathbb{R}^d$ are the row vectors of A. Moreover, P° is bounded iff $\mathbf{0}_d \in \mathrm{int}(\mathrm{conv}(a_1, \cdots, a_n))$.

Fact 9. *Let $P, Q \subset \mathbb{R}^d$ be two convex sets such that $P \subseteq Q$. Then $Q^\circ \subseteq P^\circ$.*

Fact 10. *Let $P \subset \mathbb{R}^d$ be a polytope, and let $W \subset \mathbb{R}^d$ be any $k \leq d$-dimensional linear subspace. Then the polar dual of $\pi_W(P)$ within W is equal to $P^\circ \cap W$.*

2.2 Probability Distributions

All probability distributions considered in this paper will admit a probability density function with respect to the Lebesgue measure.

Definition 11 (Gaussian distribution)**.** *The d-dimensional Gaussian distribution $\mathcal{N}_d(\bar{a}, \sigma^2 I)$ with support on $\mathbb{R}^d$, mean $\bar{a} \in \mathbb{R}^d$, and standard deviation σ, is defined by probability density*

$$(2\pi)^{-d/2} \cdot \exp(-\|s - \bar{a}\|^2/2\sigma^2).$$

at every $s \in \mathbb{R}^d$.

A basic property of Gaussian distribution is the strong tail bound:

Lemma 12 (Gaussian tail bound)**.** *Let $x \in \mathbb{R}^d$ be a random vector sampled from $\mathcal{N}_d(\mathbf{0}, \sigma^2 I)$. For any $t \geq 1$ and any $\theta \in \mathbb{S}^{d-1}$, we have*

$$\Pr[\|x\| \geq t\sigma\sqrt{d}] \leq \exp(-(d/2)(t-1)^2).$$

From this, one can upper-bound the maximum norm over n independent Gaussian random vectors with mean $\mathbf{0}_d$ and variance σ^2 by $O(\sigma\sqrt{d\log n})$ with dominating probability.

Corollary 13 (Global diameter of Gaussian random variables)**.** *For any $n \geq 2$, let $x_1, \cdots, x_n \in \mathbb{R}^d$ be i.i.d. random variables where each $x_i \sim \mathcal{N}_d(\mathbf{0}_d, \sigma^2 I)$. Then with probability at least $1 - \binom{n}{d}^{-1}$, $\max_{i\in[n]} \|x_i\| \leq 4\sigma\sqrt{d\log n}$.*

Proof. From Lemma 12, we have for each $i \in [n]$ that

$$\Pr[\|x_i\| > 4\sigma\sqrt{d\log n}] \leq \exp(-\frac{d(4\sqrt{\log n}-1)^2}{2}) \leq \exp(-2d\log n) \leq \binom{n}{d}^{-1}.$$

Then the statement follows from the union bound. □

A helpful technical substitute for the Gaussian distribution was introduced by [DH18]:

Definition 14 ((σ, r)-Laplace-Gaussian distribution)**.** *For any $\sigma, r > 0, \bar{a} \in \mathbb{R}^d$, define the d-dimensional (σ, r)-Laplace-Gaussian distribution with mean $\bar{a}$, or $LG_d(\bar{a}, \sigma, r)$, if its density function is proportional to*

$$f(x) = \begin{cases} \exp\left(-\frac{\|x-\bar{a}\|^2}{2\sigma^2}\right) & \text{, if } \|x-\bar{a}\| \leq r\sigma \\ \exp\left(-\frac{\|x-\bar{a}\|r}{\sigma} + \frac{r^2}{2}\right) & \text{, if } \|x-\bar{a}\| > r\sigma. \end{cases} \tag{9}$$

The Laplace-Gaussian random variables satisfies many desirable properties: Like Gaussian distribution, the distance to its mean is bounded above with high probability. Moreover, its probability density is log-Lipschitz throughout its domain (as a contrast, the probability density of Gaussian distribution is only log-Lipschitz close to the expectation). The definition of L-log-Lipschitz is as follows:

Definition 15 (L-log-Lipschitz random variable)**.** *Given $L > 0$, we say a random variable $x \in \mathbb{R}^d$ with probability density μ is L-log-Lipschitz (or μ is L-log-Lipschitz), if for all $x, y \in \mathbb{R}^d$, we have*

$$|\log(\mu(x)) - \log(\mu(y))| \leq L\|x - y\|,$$

or equivalently, $\mu(x)/\mu(y) \leq \exp(L\|x - y\|)$.

Lemma 16 (Properties of Laplace-Gaussian random variables, Lemma 45 of [DH18])**.** *Given any $n \geq d$, $\sigma > 0$. Let $a_1, \cdots, a_n \in \mathbb{R}^d$ be independent random variables each sampled form $LG_d(\bar{a}, \sigma, 4\sigma\sqrt{d\log n})$ (see Definition 14). Then $a_1, \cdots, a_n$ satisfy the follows:*

1. *(Log-Lipschitzness) For each $i \in [n]$, the probability density of a_i is $(4\sigma^{-1}\sqrt{d\log n})$-log-Lipschitz.*
2. *(Bounded maximum norm) With probability at least $1 - \binom{n}{d}^{-1}$, $\max_{i\in n} \|a_i\| \leq 4\sigma\sqrt{d\log n}$.*
3. *(Bounded expected radius of projection) For any $k \leq d$, any fixed k-dimensional linear subspace $H \subset \mathbb{R}^d$, we have $\mathbb{E}[\max_{i\in[n]} \|\pi_H(a_i)\|] \leq 4\sigma\sqrt{k\log n}$.*

2.3 Change of Variables

We will use the following change of variables, which is a standard tool in stochastic and integral geometry.

Definition 17 (Change of variables)**.** *Let $a_1, \cdots, a_d$ be d affine independent vectors in $\mathbb{R}^d$. Let $\theta \in \mathbb{S}^{d-1}, t \in \mathbb{R}$, be such that $\forall i \in [d], \theta^\top a_i = t$ and $\theta^\top e_1 > 1$.*

Let h be a fixed isometric embedding from $\mathbb{R}^{d-1} \to e_1^\perp$. Let $\tilde{R}_\theta : \mathbb{R}^d \to \mathbb{R}^d$ denote the rotation that rotates e_1 to θ in the two-dimensional subspace $\operatorname{span}(e_1, \theta)$, and is the identity transformation on $\operatorname{span}(e_1, \theta)^\perp$. Define $R_\theta = \tilde{R}_\theta \circ h$ to be the resulting isometric embedding from $\mathbb{R}^{d-1}$, identified with $e_1^\perp$, to $\theta^\perp$. Now define the transformation ϕ from $\theta \in \mathbb{S}^{d-1}, t \in \mathbb{R}, b_1, \cdots, b_d \in \mathbb{R}^{d-1}$ to $a_1, \cdots, a_d \in \mathbb{R}^d$ as follows:

$$\phi(\theta, t, b_1, \cdots, b_d) = (R_\theta(b_1) + t\theta, \cdots, R_\theta(b_d) + t\theta) = (a_1, \cdots, a_d). \tag{10}$$

Lemma 18 (Jacobian of the inverse isometric transformation, see Theorem 7.2.7 in [SW08])**.** *Let* $\phi : \mathbb{S}^{d-1} \times \mathbb{R} \times \mathbb{R}^{(d-1)\times d} \to \mathbb{R}^{d\times d}$ *be the transformation defined in Definition 17. The inverse transformation of* ϕ *is defined almost everywhere and has Jacobian equal to*

$$\left|\det\left(\frac{\partial\phi(a)}{\phi(a)}\right)\right| = C_d(d-1)! \cdot \mathrm{vol}_{d-1}(\mathrm{conv}(a_1, \cdots, a_d))$$

for some constant C_d *depending only on the dimension. As a consequence, if* $a_1, \dots, a_d$ *are points with probability density* $\mu(a_1, \dots, a_d)$ *and if* $\theta \in \mathbb{S}^{d-1}, t \in \mathbb{R}, b_1, \dots, b_d \in \mathbb{R}^{d-1}$ *have probability density proportional to*

$$\mathrm{vol}_{d-1}(\mathrm{conv}(b_1, \dots, b_d)) \cdot \mu(t\theta + R_\theta(b_1), \dots, t\theta + R_\theta(b_d))$$

then $\mathbb{E}[f(a_1, \dots, a_d)] = \mathbb{E}[f(\phi(\theta, t, b_1, \dots, b_d))]$ *for any measurable function* f*.*

In particular, we will use this transformation to condition on the value of θ and consider events in the variables $t, b_1, \dots, b_d$. For this purpose, we have the following fact.

Fact 19 (Log-Lipschitzness of the Position of Affine Hull)**.** *Let* $a_1, \cdots, a_d \in \mathbb{R}^d$ *be* d *independent* L*-log-Lipschitz random variables, and let* $(\theta, t, b_1, \dots, b_d) = \phi^{-1}(a_1, \dots, a_d)$*. Then conditional on the values of* $\theta, b_1, \dots, b_d$*, the random variable* t *is* (dL)*-log-Lipchitz.*

Proof. By Lemma 18, the joint probability density of $(a_1, \cdots, a_d)$ is proportional to

$$\mathrm{vol}_{d-1}(\mathrm{conv}(b_1, \cdots, b_d)) \cdot \prod_{i=1}^{d} \mu_i(R_\theta(b_i) + t\theta)$$

where μ_i is the probability density of a_i. Conditioning on $b_1, \dots, b_d$, the volume $\mathrm{vol}_{d-1}(\mathrm{conv}(b_1, \cdots, b_d))$ is fixed. The statement then follows from the fact that $\mu_i(R_\theta(b_i) + t\theta)$ is L-log-Lipschitz in t for any b_i and for each $i \in [d]$. □

2.4 Non-Degenerate Conditions

Definition 20 (Non-degenerate polytope)**.** *A polytope* $Q = \mathrm{conv}(a_1, \dots, a_n) \subset \mathbb{R}^d$ *is called non-degenerate, if it is simplicial (every facet is a simplex) and if, for* $i \in [n]$*,* $a_i \in \partial Q$ *implies that* a_i *is a vertex of* Q*.*

Definition 21 (Non-degenerate intersection with a 2D-plane)**.** *Let* $Q \subset \mathbb{R}^d$ *be a non-degenerate polytope and let* $W \subset \mathbb{R}^d$ *be a two-dimensional linear subspace. We say* Q *has non-degenerate intersection with* W*, if*

1. *The edges of the two-dimensional polygon* $Q \cap W$ *have one-to-one correspondence to the facets of* Q *that have non-empty intersection with* W*; and*

2. *The vertices of* $Q \cap W$ *have one-to-one correspondence to the* $(d-2)$*-dimensional faces (ridges) of* Q *that have non-empty intersection with* W

Lemma 22 (Non-degenerate conditions of random polytope)**.** *Given any* $n \ge d \ge 2$ *and any fixed two-dimensional plane* $W \subset \mathbb{R}^d$*. For* $a_1, \cdots, a_n \in \mathbb{R}^d$*, the polytope* $Q = \mathrm{conv}(a_1, \cdots, a_n)$ *satisfies the following properties everywhere except for a set of measure* 0*:*

1. Q *is non-degenerate;*

2. Q *has non-degenerate intersection with* W*;*

3. *For every normal vector* v *to any facet of* Q*,* $e_1^\top v \neq 0$*.*

Assume the polytope $Q = \mathrm{conv}(a_1, \dots, a_n)$ and the two-dimensional linear subspace $W \subset \mathbb{R}^d$ satisfy the conditions in Lemma 22. Then every edge of the two-dimensional polygon $W \cap Q$ correspond to a set of d vertices, and every vertex of $W \cap Q$ correspond to a set of $(d-1)$ vertices. The following lemma quantifies the relationship of such sets for adjacent vertices and edge:

Fact 23 (Properties of neighboring vertices on non-degenerate intersection polygon)**.** *Let* $W \subset \mathbb{R}^d$ *be a two-dimensional linear subspace,* $Q = \text{conv}(a_1, \ldots, a_n) \subset \mathbb{R}^d$ *is simplicial and has non-degeneracy intersection with* W*. Given* $J_1, J_2 \in \binom{[n]}{d-1}$*,* $I \in \binom{[n]}{d}$*, suppose (1)* $V_{J_1} = \text{conv}(a_j : j \in J_1) \cap W$ *and* $V_{J_2} = \text{conv}(a_j : j \in J_2) \cap W$ *are two adjacent vertices of* $Q \cap W$*, and (2)* $\text{conv}(a_i : i \in I) \cap W$ *is an edge of* $Q \cap W$ *that contains* V_{J_1} *but not contains* V_{J_2}*. Then we have* $|J_1 \backslash J_2| = |J_2 \backslash J_1| = 1$ *and* $|I \backslash J_2| = 2$*.*

Proof. Let $I' = J_1 \cap J_2$. Then $\text{conv}(V_{J_1}, V_{J_2}) = \text{conv}(a_i : i \in I) \cap W$ is an edge of the polygon $Q \cap W$. Since Q has non-degenerate intersection with W, we have that $|I'| = d$. Combining with $|J_1| = |J_2| = d - 1$ gives us that $|J_1 \backslash J_2| = |J_2 \backslash J_1| = 1$.

Next we consider $|I \backslash J_2|$. Since $J_1 \subset I$ and $|J_1 \backslash J_2| = 1$, it could only be the case that $|I \backslash J_2| \in \{1, 2\}$. If $|I \backslash J_2| = 1$, then by $|I| = |J_2| + 1$ we must have $J_2 \subset I$, but this contradicts to the fact that $J_2 \not\subset I$. Therefore we could only have $|I \backslash J_2| = 2$. □

3 Smoothed Complexity Upper Bound

In this section, we establish our key theorem for upper bounding the number of edges of a random polygon $\mathrm{conv}(a_1, \dots, a_n) \cap W$ for W a fixed 2-dimensional linear subspace and $a_1, \dots, a_n \in \mathbb{R}^d$. Specifically we show that if for any edge on the shadow polygon $\mathrm{conv}(a_1, \cdots, a_n) \cap W$, the expected distance between the affine hull of the edge and the next vertex of the shadow is not too small in expectation, then the expected number of edges of $\mathrm{conv}(a_1, \cdots, a_n) \cap W$ will be bounded from above.

Definition 24 (Edge event)**.** *For $I \subset [n]$, we write $F_I = \mathrm{conv}(a_i : i \in I)$. Define E_I to be the event that both F_I is a facet of $\mathrm{conv}(a_1, \dots, a_n)$ and $F_I \cap W \neq \emptyset$.*

Note that when $d = 2$ then $W = \mathbb{R}^2$ and the condition $F_I \cap W \neq \emptyset$ is guaranteed to hold.

Remark 25. *Any edge e of $\mathrm{conv}(a_1, \dots, a_n) \cap W$ can be written as $e = F_I \cap W$ for some $I \subset [n]$ for which E_I holds. Assuming non-degeneracy, this relation between edges and index sets is a one-to-one correspondence, and moreover every $I \subset [n]$ for which E_I holds satisfies $|I| = d$.*

To state the key theorem's assumption, we require one more definition:

Definition 26. *For any given two-dimensional linear subspace $W \subset \mathbb{R}^d$ we denote an arbitrary but fixed rotation as "clockwise". For the polygon $\mathrm{conv}(a_1, \cdots, a_n) \cap W$ of our interest, let $p_1, \dots, p_k$ denote its vertices in clockwise order and write $p_{k+1} = p_1, p_{k+2} = p_2$. Then for any edge $e = [p_{i-1}, p_i]$, we call p_i its second vertex in clockwise order and we call p_{i+1} the next vertex after e in clockwise order. The edge $[p_i, p_{i+1}]$ is the next edge after e in clockwise order.*

Note that the above terms are well-defined in the sense that they depend only on the polygon and the orientation of the subspace, not on the labels. With this definition in place, we can now state the theorem itself:

Theorem 27 (Smoothed complexity upper bound for continuous perturbations)**.** *Fix any $n, d \geq 2$, $\sigma \geq 0$, and any two-dimensional linear subspace $W \subseteq \mathbb{R}^d$. Let $a_1, \cdots, a_n \in \mathbb{R}^d$ be independently distributed each according to a continuous probability distribution.*

For any $I \in \binom{n}{d}$, conditional on E_I, define $y_I \in W$ as the outer unit normal of the edge $F_I \cap W$. Suppose for each $I \in \binom{n}{d}$ such that $\Pr[E_I] \geq 10\binom{n}{d}^{-1}$, we have

$$\Pr[y_I^\top p_2 - y_I^\top p_3 \geq \gamma \mid E_I] \geq 0.1,$$

where we write $[p_1, p_2] = F_I \cap W$ and $p_3 \in \mathrm{conv}(a_1, \cdots, a_n) \cap W$ as the next vertex after $F_I \cap W$ in clockwise order. Then we have

$$\begin{aligned}\mathbb{E}\left[\mathrm{edges}\left(\mathrm{conv}(a_1, \cdots, a_n)\right)\right] &\leq 10 + 80\pi\sqrt{\frac{\mathbb{E}[\max_{i\in[n]} \|\pi_W(a_i)\|]}{\gamma}} \\ &= O\left(\sqrt{\frac{\mathbb{E}[\max_{i\in[n]} \|\pi_W(a_i)\|]}{\gamma}}\right).\end{aligned}$$

Assuming non-degeneracy, y_I is well-defined if and only if E_I happens. In this case, we are guaranteed that $y_I^\top p_2 - y_I^\top q_3 > 0$.

To prove the above theorem, we show that any $I \in \binom{[n]}{d}$ with $\Pr[E_I] \geq \binom{n}{d}^{-1}$ can be charged to either a portion of the perimeter of the polygon $\mathrm{conv}(a_1, \dots, a_n) \cap W$ or to a portion of its sum 2π of exterior angles at its vertices.

Definition 28 (Exterior angle and length of the next edge)**.** *Given any $I \in \binom{[n]}{d}$, we define two random variables $\theta_I, \ell_{I+} \geq 0$. If E_I happens, write $v \in F_I \cap W$ for the second endpoint of $F_I \cap W$ in clockwise order. Let θ_I to be the (two-dimensional) exterior angle of $\mathrm{conv}(a_1, \cdots, a_n) \cap W$ at v; If E_I doesn't happen then let $\theta_I = 0$.*

Let ℓ_{I+} denote the following random variable: If E_I happens, then ℓ_{I+} equals to the length of the next edge after $F_I \cap W$ in clockwise order, i.e., the other edge of $\mathrm{conv}(a_1, \cdots, a_n) \cap W$ containing v. If E_I doesn't happen then let $\ell_{I+} = 0$.

Proof of Theorem 27. Since we have non-degeneracy with probability 1, by Lemma 22 and linearity of expectation we find

$$\mathbb{E}\left[\text{edges}\left(\text{conv}(a_1, \cdots, a_n) \cap W\right)\right] = \sum_{I \in \binom{[n]}{d}} \Pr[E_I].$$

We can give an upper bound of the expected number of edges of $\text{conv}(a_1, \cdots, a_n) \cap W$ by upper-bounding each $\Pr[E_I]$. Fix any $I \in \binom{[n]}{d}$ and let $t > 0$ be a parameter to be determined later. We consider three different possible upper bounds on $\Pr[E_I]$, at least one of which will always hold:

Case 1: $\Pr[E_I] \leq 10\binom{n}{d}^{-1}$**.**

Since $\sum_{I \in \binom{[n]}{d}} 10\binom{n}{d}^{-1} = 10$, one can immediately see that the total contribution of edges counted in this case is at most 10.

Case 2: $\Pr[E_I] > 10\binom{n}{d}^{-1}$ **and** $\Pr[\ell_{I^+} \geq t \mid E_I] \geq \frac{1}{20}$**.**

In this case, $\mathbb{E}[\ell_{I^+} \mid E_I] \geq \frac{t}{20}$, therefore we have

$$\Pr[E_I] = \frac{\mathbb{E}[\ell_{I^+}\mathbb{I}(E_I)]}{\mathbb{E}[\ell_{I^+} \mid E_I]} \leq \frac{20}{t} \cdot \mathbb{E}[\ell_{I^+}\mathbb{I}(E_I)].$$

Case 3: $\Pr[E_I] > 10\binom{n}{d}^{-1}$ **and** $\Pr[\ell_{I^+} \leq t \mid E_I] \geq \frac{19}{20}$**.** Conditional on E_I, without loss of generality we write $[p_1, p_2] = F_I \cap W$ and let p_3 to denote the next vertex after $F_I \cap W$ in clockwise direction. From the theorem's assumption we have $\Pr[\text{dist}(\text{affhull}(p_1, p_2), p_3) \geq \gamma \mid E_I] \geq \frac{1}{10}$. Then from the union bound,

$$\begin{aligned}
&\Pr[(\ell_{I^+} \leq t) \wedge (\text{dist}(\text{affhull}(p_1, p_2), p_3) \geq \gamma) \mid E_I] \\
\geq\ & 1 - \Pr[\ell_{I^+} > t \mid E_I] - \Pr[\text{dist}(\text{affhull}(p_1, p_2), p_3) < \gamma \mid E_I] \geq \frac{1}{20}.
\end{aligned}$$

Since $\theta_I \geq 0$ and

$$\theta_I \geq \sin(\theta_I) = \frac{\text{dist}(\text{affhull}(p_1, p_2), p_3)}{\ell_{I^+}}$$

we have $\mathbb{E}[\theta_I \mid E_I] \geq \frac{1}{20} \cdot \frac{\gamma}{t}$, and therefore we can upper bound $\Pr[E_I]$ by

$$\Pr[E_I] = \frac{\mathbb{E}[\theta_I\mathbb{I}(E_I)]}{\mathbb{E}[\theta_I \mid E_I]} \leq \frac{20t}{\gamma} \cdot \mathbb{E}[\theta_I\mathbb{I}(E_I)].$$

Readers are referred to Figure 1 for more illustration of the proof.

Combining the upper bounds for each $\Pr[E_I]$ for the above three cases, we get that

$$\begin{aligned}
\mathbb{E}\left[\text{vertices}\left(\text{conv}(a_1, \cdots, a_n) \cap W\right)\right] &= \sum_{I \in \binom{[n]}{d}} \Pr[E_I] \\
&\leq \sum_{I \in \binom{[n]}{d}} \left(10\binom{n}{d}^{-1} + \frac{20}{t} \cdot \mathbb{E}[\ell_{I^+}\mathbb{I}(E_I)] + \frac{20t}{\gamma} \cdot \mathbb{E}[\theta_I\mathbb{I}(E_I)]\right) \\
&\leq 10 + \frac{20}{t} \cdot \mathbb{E}[\sum_{I \in \binom{[n]}{d}} \ell_{I^+}\mathbb{I}(E_I)] + \frac{20t}{\gamma} \cdot \mathbb{E}[\sum_{I \in \binom{[n]}{d}} \theta_I\mathbb{I}(E_I)] \qquad (11)
\end{aligned}$$

To upper bound the second term of (11), we notice that $\sum_{I \in \binom{[n]}{d}} \ell_{I^+}\mathbb{I}(E_I)$ exactly equals the perimeter of $\text{conv}(a_1, \ldots, a_n) \cap W$. Since the shadow polygon $\text{conv}(a_1, \cdots, a_n) \cap W$ is contained in the two-dimensional disk of radius $\max_{i \in [n]} \|\pi_W(a_i)\|$, by the monotonicity of surface area for convex sets we have

$$\mathbb{E}[\sum_{I \in \binom{[n]}{d}} \ell_{I^+}\mathbb{I}(E_I)] \leq 2\pi \cdot \mathbb{E}[\max_{i \in [n]} \|\pi_W(a_i)\|].$$

To upper bound the third term of (11), we notice that the sum of exterior angles for any polygon always equals to 2π. Thus

$$\mathbb{E}[\sum_{I\in\binom{[n]}{d}} \theta_I \mathbb{I}(E_I)] = 2\pi \tag{12}$$

Finally, we combine (11 - 12) and minimize over all $t \geq 0$:

$$\begin{aligned}\mathbb{E}\left[\text{vertices}\left(\text{conv}(a_1, \cdots, a_n) \cap W\right)\right] &\leq \min_{t>0} \left(10 + \frac{40\pi \mathbb{E}[\max_{i\in[n]} \|\pi_W(a_i)\|]}{t} + \frac{40\pi t}{\gamma}\right) \\ &\leq 10 + 80\pi \sqrt{\frac{\mathbb{E}[\max_{i\in[n]} \|\pi_W(a_i)\|]}{\gamma}}.\end{aligned}$$

where in the final step, we set $t = \sqrt{\gamma \mathbb{E}[\max_{i\in[n]} \|\pi_W(a_i)\|]}$. □

We will not apply Theorem 27 directly to Gaussian distributed points $a_1, \ldots, a_n$. Instead, we will follow an approach introduced by [DH18]. First, we relate the shadow size for Gaussian distributed vectors to the shadow size for *Laplace-Gaussian* distributed vectors. We will then show how to use Theorem 27 to any log-Lipschitz probability distribution.

Lemma 29 (Lemma 46 of [DH18])**.** *Given any $n \geq d \geq 2, \sigma > 0$, any two-dimensional linear subspace $W \subset \mathbb{R}^d$, and any $\bar{a}_1, \cdots, \bar{a}_n \in \mathbb{R}^d$ with $\max_{i\in[n]} \|\bar{a}_i\| \leq 1$. For every $i \in [n]$, let $a_i \sim \mathcal{N}_d(\bar{a}_i, \sigma)$ and $\widehat{a}_i \sim LG_d(\bar{a}_i, \sigma, 4\sigma\sqrt{d \log n})$ be independently sampled. Then the following holds*

$$\mathbb{E}\left[\text{edges}(\text{conv}(a_1, \cdots, a_n) \cap W)\right] \leq 1 + \mathbb{E}\left[\text{edges}(\text{conv}(\widehat{a}_1, \cdots, \widehat{a}_n) \cap W)\right].$$

Although [DH18] state this lemma only for $d \geq 3$, their proof applies without change to the case $d = 2$.

4 Upper Bound for Two Dimension

In this section, we prove the smoothed complexity upper bound for $d = 2$ using the key lemma. In this case, the shadow plane W is the two-dimensional Euclidean space. From Theorem 27, it remains to lower bound the distance from the affine hull of an edge to its neighboring vertex in clockwise order (the quantity γ in Theorem 27), where our polygon $\mathrm{conv}(a_1, \cdots, a_n)$ is under Laplace-Gaussian perturbation. In fact, we can show a more general result for any L-log-Lipschitz distribution:

Lemma 30 (Lower Bound of γ in Two Dimension)**.** *Let $a_1, \cdots, a_n \in \mathbb{R}^2$ be n independent L-log-Lipschitz random variables. Then for any $I \in \binom{[n]}{2}$, conditional on E_I happens, the outer unit normal $y \in W$ of the edge $\mathrm{conv}(a_i : i \in I)$ satisfies*

$$\Pr[y^\top a_i - \max_{j \notin I} y^\top a_j \geq \frac{1}{L} \mid E_I] \geq 0.1,$$

for any $i \in I$.

Together with Theorem 27 and Lemma 29 from the previous section, Lemma 30 immediately tells us the upper bound for two-dimensional polygons under Gaussian perturbation:

Theorem 31 (Two-Dimensional Upper Bound)**.** *Let $\bar{a}_1, \cdots, \bar{a}_n \in \mathbb{R}^2$ be $n > 2$ vectors with norm at most 1. For each $i \in [n]$, let a_i be independently distributed as $\mathcal{N}_2(\bar{a}_i, \sigma^2 I)$. Then*

$$\mathbb{E}\left[\mathrm{edges}\left(\mathrm{conv}(a_1, \cdots, a_n)\right)\right] \leq O\left(\frac{\sqrt[4]{\log n}}{\sqrt{\sigma}} + \sqrt{\log n}\right).$$

Proof. For each $i \in [n]$, let $\hat{a}_i$ be independently sampled form the 2-dimensional Laplace-Gaussian distribution $LG_2(\bar{a}_i, \sigma, 4\sigma\sqrt{2\log n})$. It follows from Lemma 16 that $\hat{a}_i$ is $(4\sigma^{-1}\sqrt{2\log n})$-log-Lipschitz and $\mathbb{E}[\max_{i\in[n]} \|\hat{a}_i\|] \leq 1 + 4\sigma\sqrt{2\log n}$. We use Lemma 30 by setting $L = 4\sigma^{-1}\sqrt{2\log n}$, and Theorem 27 by setting $\gamma = \frac{1}{L} = \frac{\sigma}{4\sqrt{2\log n}}$, to find

$$\mathbb{E}\left[\mathrm{edges}\left(\mathrm{conv}(\hat{a}_1, \cdots, \hat{a}_n)\right)\right] \leq O\left(\frac{\sqrt[4]{\log n}}{\sqrt{\sigma}} + \sqrt{\log n}\right).$$

And from Lemma 29, we conclude that $\mathbb{E}\left[\mathrm{edges}\left(\mathrm{conv}(a_1, \cdots, a_n)\right)\right] \leq 1 + O(\frac{\sqrt[4]{\log n}}{\sqrt{\sigma}} + \sqrt{\log n})$. ☐

Proof of Lemma 30. Fix any set $I = \{i, i'\} \subset [n]$. Define $z \in \mathbb{S}^1$ and t to satisfy $z^\top a_i = z^\top a_{i'} = t$ and $z^\top e_1 > 0$. Both are well-defined with probability 1.

Note that E_I is now equivalent to either having $z^\top a_j < t$ for all $j \notin I$ or $z^\top a_j > t$ for all $j \notin I$. Write E_I^+ for the former case and E_I^- for the latter. The vector z is always defined, assuming non-degeneracy, and is equal to the outer normal unit vector y if E_I^+ and equal to $-y$ if E_I^-.

Using Fubini's theorem, we condition on the values of $a_j, j \notin I$. and z using Lemma 18. Let $\mu : \mathbb{R} \to \mathbb{R}_{\geq 0}$ denote the induced density of $t = y^\top a_i = y^\top a_{i'}$. Then from Fact 19, μ is $(2L)$-log-Lipschitz.

In the first case, for E_I^+, we have, still only considering the randomness over t,

$$\begin{aligned}
\Pr[(t - \max_{j\notin I} z^\top a_j \geq \frac{1}{L}) \wedge E_I^+] &= \int_{\max_{j\notin I} z^\top a_j + 1/L}^{\infty} \mu(x)\mathrm{d}x \\
&= \int_{\max_{j\notin I} z^\top a_j}^{\infty} \mu(z + 1/L)\mathrm{d}x \\
&\geq \int_{\max_{j\notin I} z^\top a_j}^{\infty} e^{-2}\mu(z)\mathrm{d}z \qquad \text{(By $(2L)$-log-Lipschitzness of μ)} \\
&= e^{-2}\Pr[E_I^+].
\end{aligned}$$

Similarly for the other case, E_I^-, we find

$$\Pr[(\min_{j\notin I} z^\top a_j - t \geq \frac{1}{L}) \wedge E_I^-] \geq e^{-2}\Pr[E_I^-].$$

Now observe that, for $i \in I$,

$$\begin{aligned}\Pr[y^\top a_i - \max_{j \notin I} y^\top a_j \geq \frac{1}{L} \wedge E_I] &= \Pr[(t - \max_{j \notin I} z^\top a_j \geq \frac{1}{L}) \wedge E_I^+] + \Pr[(\min_{j \notin I} z^\top a_j - t \geq \frac{1}{L}) \wedge E_I^-] \\ &\geq e^{-2} \Pr[E_I^+] + e^{-2} \Pr[E_I^-] = e^{-2} \Pr[E_I],\end{aligned}$$

finishing the proof since

$$\Pr[y^\top a_i - \max_{j \notin I} y^\top a_j \geq \frac{1}{L} \mid E_I] = \Pr[y^\top a_i - \max_{j \notin I} y^\top a_j \geq \frac{1}{L} \wedge E_I] / \Pr[E_I] \geq e^{-2} \geq 0.1.$$

□

5 Multi-Dimensional Upper Bound

5.1 Notations and roadmap

For the higher-dimensional case, we will establish the upper bound:

Theorem 32. *Given any $d > 2, n \geq d$, and $\sigma \leq \frac{1}{8d\sqrt{\log n}}$. Let $\bar{a}_1, \cdots, \bar{a}_n$ be n vectors with $\max_{i\in[n]} \|\bar{a}_i\| \leq 1$. For each $i \in [n]$, let a_i be independently distributed as $\mathcal{N}_d(\bar{a}_i, \sigma^2 I)$. Then*

$$\mathbb{E}[\text{edges}(\text{conv}(a_1, \ldots, a_n) \cap W)] = O\left(\sigma^{-3/2} d^{13/4} \log^{5/4} n\right). \tag{13}$$

Since we assume that $a_1, \ldots, a_n$ each have a continuous probability density function, we know that with probability 1 the polytope $\text{conv}(a_1, \ldots, a_n)$ is non-degenerate and has a non-degenerate intersection with W. In this case, the edges of the polygon $\text{conv}(a_1, \cdots, a_n) \cap W$ are given by the sets $F_I \cap W$ for which $I \in \binom{[n]}{d}$ and E_I holds (where E_I is defined in Definition 24). In addition, the vertices of the polygon $\text{conv}(a_1, \cdots, a_n) \cap W$ are given by the intersection between W and $(d-2)$-dimensional ridges of $\text{conv}(a_1, \cdots, a_n)$, which are convex hulls of $(d-1)$ vertices of $\text{conv}(a_1, \cdots, a_n)$. We define the following notations of ridge and the corresponding vertex:

Definition 33 (Ridge and vertex event)**.** *For any $J \subset [n]$, write $R_J = \text{conv}(a_j : j \in J)$. Define A_J to be the event that R_J is a ridge of $\text{conv}(a_1, \cdots, a_n)$ and $R_J \cap W \neq \emptyset$.*

Remark 34. *Any vertex v of $\text{conv}(a_1, \ldots, a_n)$ can be written as $v = R_J \cap W$ for some $J \subset [n]$ for which A_J holds. Assuming non-degeneracy, each J for which A_J holds satisfies $|J| = d - 1$ and the relation between vertices and index sets $J \in \binom{[n]}{d-1}$ with A_J is a one-to-one correspondence.*

Our proof follows from a similar structure as the two-dimensional upper bound (see section 4). Our main technical result is the following lower-bound of the edge-to-vertex distance on the shadow polygon:

Lemma 35 (Edge-to-vertex distance of shadow polygon in multi-dimension)**.** *Let $a_1, \cdots, a_n \in \mathbb{R}^d$ be independent L-log-Lipschitz random variables. For any $I \in \binom{[n]}{d}$ that satisfies $\Pr[E_I] \geq 10\binom{n}{d}^{-1}$, (See Definition 24 the definition of E_I) we have*

$$\Pr[y^\top p - y^\top p' \geq \Omega(\frac{1}{L^3 d^5 \log n}) \mid E_I] \geq 0.1,$$

where p is any point in $F_I \cap W$, and $p' \in \text{conv}(a_1, \ldots, a_n) \cap W$ is the next vertex after $F_I \cap W$ in clockwise direction. Here $y \in W$ is the outer unit normal to the edge $F_I \cap W$ on $\text{conv}(a_1, \cdots, a_n) \cap W$.

Then we can prove Theorem 32 directly by Lemma 35, and Theorem 27, Lemma 29 from Section 3.

Proof of Theorem 32. For each $i \in [n]$, let $\hat{a}_i$ be independently sampled form the Laplace-Gaussian distribution $LG_d(\bar{a}_i, \sigma, 4\sigma\sqrt{2\log n})$. From Lemma 16, we know that

1. Each $\hat{a}_i$ is $L = (4\sigma^{-1}\sqrt{d\log n})$-log-Lipschitz;
2. $\mathbb{E}[\max_{i\in[n]}]\|\pi_W(\hat{a}_i)\| \leq 1 + 4\sigma\sqrt{2\log n} \leq 1.5$.

Also from Lemma 35, we get that for any $p \in F_I \cap W$, if p' is the next vertex after the edge $F_I \cap W$ in clockwise order, then

$$\Pr[y_I^\top p \geq y_I^\top p' + \frac{1}{L^3 d^5 \log^2 n} \mid E_I] \geq 0.1.$$

here $y_I \in W$ is the outer unit normal vector of the polygon $\text{conv}(\hat{a}_1, \cdots, \hat{a}_n) \cap W$ on the edge $F_I \cap W$. Then we can use Theorem 27 by setting $L = 4\sigma^{-1}\sqrt{d\log n}$, $\gamma = \Omega(\frac{1}{L^3 d^5 \log^2 n})$ and $\mathbb{E}[\max_{i\in[n]} \|\pi_W(a_i)\|] = 1.5$, to find

$$\mathbb{E}\left[\text{edges}\left(\text{conv}(\hat{a}_1, \cdots, \hat{a}_n)\right)\right] \leq 10 + O(\sqrt{\sigma^{-3} d^{13/2} \log^{5/2} n}).$$

Finally, from Lemma 29, we conclude that

$$\begin{aligned}\mathbb{E}\left[\mathrm{edges}\left(\mathrm{conv}(a_1,\cdots,a_n)\right)\right] &\leq 11 + O(\sqrt{\sigma^{-3}d^{13/2}\log^{5/2} n}) \\ &= O\left(\sigma^{-3/2}d^{13/4}\log^{5/4} n\right).\end{aligned}$$

□

The rest of this section will be structured as follows. In section 5.2 we use a deterministic argument to establish sufficient criteria for the conclusion of Lemma 35 to hold. In section 5.3 and section 5.4, we prove that these conditions hold with good probability conditional on E_I. The proof of Lemma 35 is then finished in section 5.5.

5.2 Deterministic conditions for a good separator

In this subsection, we prove the following lemma:

Lemma 36. *Let $W \subset \mathbb{R}^d$ be a two-dimensional linear subspace, $Q = \mathrm{conv}(a_1,\cdots,a_n) \subset \mathbb{R}^d$ be a non-degenerate polytope with a non-degenerate intersection with W such that $\max_{i,j\in[n]}\|a_i - a_j\| \leq 3$ and $W \cap Q \neq \emptyset$. Fix any facet F of Q such that $F \cap W \neq \emptyset$ and any ridge $R \subset F$ of F such that $W \cap R$ is a singleton set $\{p\}$. Let $\delta, r \geq 0$ be such that*

1. *(distance between F and other vertices) $\forall a_k \notin F, \mathrm{dist}(\mathrm{affhull}(F), a_k) \geq \delta$;*
2. *(Inner radius of R) $\mathrm{dist}(F \cap W, \partial R) \geq r$.*

Then the outer unit normal vector $\bar{\theta} \in W$ to the edge $F \cap W$ satisfies

$$\bar{\theta}^\top p - \bar{\theta}^\top p' \geq \delta r/3$$

for any $p \in F \cap W$, here $p' \in Q \cap W$ is the next vertex after $F \cap W$ in clockwise order.

Proof. Write R' for the ridge of Q such that $\{p'\} = R' \cap W$. Since $p' \in Q \cap W$ is adjacent to vertex p and the edge $F \cap W$, by Fact 23 we may relabel the a_i such that $R' = \mathrm{conv}(a_1,\ldots,a_{d-1})$, $R = \mathrm{conv}(a_2,\ldots,a_d)$, and $F = \mathrm{conv}(a_2,\ldots,a_{d+1})$ without loss of generality. Let $\theta \in \mathbb{S}^{d-1}$ denote the outward unit normal to F. This normal vector satisfies

$$\delta \leq \min_{\substack{i\in[n]\\ a_i\notin F}} \theta^T(p - a_i) \leq \theta^\top(p - a_1).$$

Let $s \in \mathbb{S}^{d-1}$ be the unit vector indicating the direction of the (one-dimensional) line $F \cap W$, i.e., for which $F \cap W = s + F \cap W$. This vector is unique up to sign. Also, let $\bar{\theta} = \pi_W(\theta)/\|\pi_W(\theta)\|$ be the outward unit normal to $F \cap W$ in the two-dimensional plane W. Notice that θ and s form an orthonormal basis of W. Therefore we get

$$\bar{\theta}^\top(p - p') = \bar{\theta}^\top\pi_{s^\perp}(p - p') = \|\pi_{s^\perp}(p - p')\| \tag{14}$$

Here the last equality comes from $(p - p') \in W = \mathrm{span}(\bar{\theta}, s)$, thus $\pi_{s^\perp}(p - p') = \pi_{\bar{\theta}}(p - p')$.

Now we focus on the $(d-1)$-dimensional space $s^\perp$, and consider the projections $\pi_{s^\perp}(a_1),\cdots,\pi_{s^\perp}(a_d)$. Since the diameter of $\mathrm{conv}(a_1,\cdots,a_d)$ is at most 3, we have $\max_{i,j\in[d]}\|\pi_{s^\perp}(a_i) - \pi_{s^\perp}(a_j)\| \leq 3$. Because $\theta \in s^\perp$ and θ is a unit normal vector of R, we know that θ is also a unit normal of $\pi_{s^\perp}(R) = \pi_{s^\perp}(\mathrm{conv}(a_2,\cdots,a_d))$. This gives

$$\mathrm{dist}(\pi_{s^\perp}(a_1), \mathrm{affhull}(\pi_{s^\perp}(R))) = \theta^\top(p - a_1) \geq \delta.$$

Also, since $\mathrm{dist}(F \cap W, \partial R) \geq r$ where $F \cap W = \{p + st : t \in R\}$ is one-dimensional. After the projection to $s^\perp$ we have

$$\mathrm{dist}(\pi_{s^\perp}(p), \partial\pi_{s^\perp}(R)) = \mathrm{dist}(F \cap W, \partial R) \geq r.$$

Therefore by Lemma 37, we have

$$\|\pi_{s^\perp}(p) - \pi_{s^\perp}(p')\| \geq \mathrm{dist}(\pi_{s^\perp}(p), \mathrm{affhull}(\pi_{s^\perp}(R'))) \geq r\delta/3,$$

where the first step comes from $\pi_{s^\perp}(p') \in \mathrm{affhull}(\pi_{s^\perp}(R'))$. The lemma then follows from (14). □

Lemma 37. *Given $b_1, \cdots, b_d \in \mathbb{R}^{d-1}$ such that $\mathrm{conv}(b_1, \cdots, b_d)$ is non-degenerate. Suppose*

1. $\forall i, j \in [d]$, $\|b_i - b_j\| \le 3$;
2. $\mathrm{dist}(b_1, \mathrm{affhull}(b_2, \cdots, b_d)) \ge \delta$;
3. *There exists $q \in \mathrm{conv}(b_2, \cdots, b_d)$ such that $\mathrm{dist}(q, \partial(\mathrm{conv}(b_2, \cdots, b_d))) \ge r$.*

Then we have $\mathrm{dist}(q, \mathrm{affhull}(b_1, \cdots, b_{d-1})) \ge r\delta/3$.

Proof. For simplicity, write $B = \mathrm{conv}(b_2, \cdots, b_d)$ and $B' = \mathrm{conv}(b_1, \cdots, b_{d-1})$. Let $q' = \pi_{B'}(q)$ be the point closest to q on $\mathrm{affhull}(B')$.

Let $x = (B \cap B') \cap \mathrm{affhull}(b_1, q, q')$ be its intersection between the two-dimensional plane $\mathrm{affhull}(b_1, q, q')$ and the $(d-3)$-dimensional ridge $B \cap B'$ (which gives a unique point). (See Figure 2 for an illustration). Consider the triangle $\mathrm{conv}(b_1, q, x)$ and calculate its area in two different ways. On one hand, it has base $\mathrm{conv}(b_1, x)$ of length $\|b_1 - x\| \le 3$ with height $\mathrm{dist}(q, \mathrm{affhull}(b_1, x)) = \|q - q'\|$, which gives that the area of the triangle is at most $\frac{3\|q-q\|}{2}$. On the other hand, this triangle has base $\mathrm{conv}(x, q)$ of length $\|x - q\| \ge \mathrm{dist}(q, \partial(B)) \ge r$ with height $\mathrm{dist}(b_1, \mathrm{affhull}(x, q)) \ge \mathrm{dist}(b_1, B) \ge \delta$, which gives that the area of the triangle is at least $\frac{r\delta}{2}$.

Therefore we have $\mathrm{dist}(q, \mathrm{affhull}(B')) = \|q - q'\| \ge \frac{r\delta}{3}$ as desired. ∎

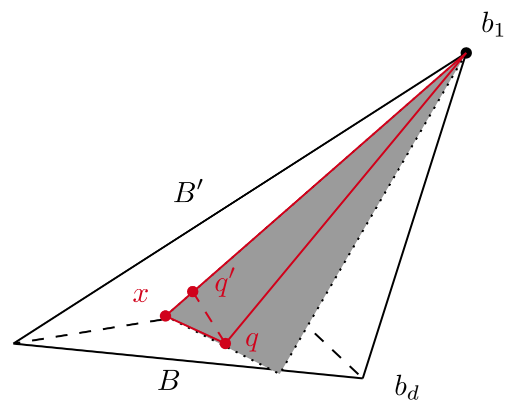

Figure 2: Illustration of Lemma 37 when $d - 1 = 3$. In gray is the intersection between the two-dimensional plane $\mathrm{affhull}(b_1, q, q')$ and $\mathrm{conv}(b_1, \cdots, b_d)$. The red triangle is $\mathrm{conv}(b_1, x, q)$. The bottom face is B and the back face is B'.

5.3 Randomized Lower-Bound for δ: Distance between vertices and facets

In this section, we show that the affine hull of a given facet F of the polytope $\mathrm{conv}(a_1, \cdots, a_n)$ is $\Omega(\frac{1}{Ld\log n})$-far away to other vertices with good probability, or in other words, the distance δ in Lemma 36 is at least $\Omega(\frac{1}{Ld\log n})$ with good probability. Our main result of this section is as follows:

Lemma 38 (Randomized lower-bound for δ)**.** *Let $a_1, \ldots, a_n \in \mathbb{R}^d$ be independent L-log-Lipshchitz random vectors. For any $I \in \binom{[n]}{d}$ such that $\Pr[E_I] \ge 10\binom{n}{d}^{-1}$, we have*

$$\Pr[\min_{k \in [n] \setminus I} \mathrm{dist}(\mathrm{affhull}(F_I), a_k) \ge \frac{1}{10e^3 dL \log n}) \mid E_I] \ge 0.72.$$

To show Lemma 38, we fix any $I \in \binom{[n]}{d}$ of consideration. Without loss of generality, assume $I = [d]$ and write $E = E_{[d]}$. To show Lemma 38, we define the following event B_ε indicating that the distance from $F_{[d]}$ to other vertices is at least ε.

Definition 39 (Separation by the margin of a facet)**.** *Let $\theta \in \mathbb{S}^{d-1}, t \in \mathbb{R}$ be as in Definition 17. For any $\varepsilon > 0$, let B_ε^+ denote the event that $\theta^\top a_i < t - \varepsilon$ for all $i \in [n] \setminus [d]$ and B_ε^- denote the event that $\theta^\top a_i > t + \varepsilon$ for all $i \in [n] \setminus [d]$. We write $B_\varepsilon = B_\varepsilon^+ \vee B_\varepsilon^-$.*

In the following lemma, we show that for sufficiently small ε, $\Pr[E \wedge B_\varepsilon]$ is a constant fraction of $\Pr[E]$.

Lemma 40. *For any $\varepsilon \leq \frac{1}{10e^3 Ld \log n}$ it holds that*

$$\Pr[E] \leq \binom{n}{d}^{-1} + \frac{5}{4} \cdot \Pr[E \wedge B_\varepsilon].$$

Proof. Writing random variables as subscripts to denote which expectation is over which variables, we start by using Fubini's theorem to write

$$\Pr_{a_1,\dots,a_n}[E] = \mathbb{E}_{a_1,\dots,a_d}[\Pr_{a_{d+1},\dots,a_n}[E]].$$

Fix any choice of $a_1,\dots,a_d \in \mathbb{R}^n$ subject to the non-degeneracy assumptions in Lemma 22 and $\mathrm{conv}(a_1,\dots,a_d) \cap W \neq \emptyset$. Define $\theta \in \mathbb{S}^{d-1}, t > 0$ as described in Definition 17, i.e. $\theta^\top a_i = t$ for each $i \in [d]$. Write $s_i = \theta^\top a_i$ for each $i \in [n] \setminus [d]$. We note that s_i is an L-log-Lipschitz random variable for all $i \in [n] \setminus [d]$. Moreover, over the remaining randomness in $a_{d+1},\dots,a_n$, we have $\Pr[E] = \Pr[B_0^+] + \Pr[B_0^-]$ and $\Pr[B_\varepsilon] = \Pr[B_\varepsilon^+] + \Pr[B_\varepsilon^-]$. We will prove that $\Pr[B_\varepsilon^+] \leq \frac{1}{2\binom{n}{d}} + \frac{5}{4}\Pr[B_0^+]$, and the appropriate statement will follow for B_ε^- analogously. Putting together this will prove the lemma.

If $\Pr[B_0^+] \leq \frac{1}{2}\binom{n}{d}^{-1}$ then the desired inequality holds directly. Otherwise, fix any $i \in [n] \setminus [d]$ and let μ_i denote the induced probability density function of s_i. We then have

$$\begin{aligned}
\Pr[s_i \geq t - \varepsilon \mid s_i \leq t] &= \frac{\int_{-\varepsilon}^0 \mu_i(t+s)\mathrm{d}s}{\int_{-\infty}^0 \mu_i(t+s)\mathrm{d}s} \\
&= \frac{\varepsilon L \int_{-1/L}^0 \mu_i(t+\varepsilon L s)\mathrm{d}s}{\int_{-\infty}^0 \mu_i(t+s)\mathrm{d}s} \\
&\leq e\frac{\varepsilon L \int_{-1/L}^0 \mu_i(t+s)\mathrm{d}s}{\int_{-\infty}^0 \mu_i(t+s)\mathrm{d}s} \\
&= e\frac{\varepsilon L \int_0^{1/L} \mu_i(t+s-1/L)\mathrm{d}s}{\int_{-\infty}^{1/L} \mu_i(t+s-1/L)\mathrm{d}s} \\
&\leq e^3\frac{\varepsilon L \int_0^{1/L} \mu_i(t+s)\mathrm{d}s}{\int_{-\infty}^{1/L} \mu_i(t+s)\mathrm{d}s} \\
&= e^3 \varepsilon L \Pr[s_i \geq t \mid s_i \leq t + 1/L] \\
&\leq e^3 \varepsilon L \Pr[s_i \geq t].
\end{aligned} \tag{15}$$

The first two inequalities above follow from L-log-Lipschitzness of μ_i. The third follows from the fact that $s_i \geq t + 1/L$ implies $s_i \geq t$. As such we can, for fixed t, θ, upper-bound the probability over $s_1,\dots,s_d$ that, conditional on B_0^+, there exists a vertex being ε-close to $\mathrm{affhull}(F_I)$:

$$\begin{aligned}
\Pr[\neg B_\varepsilon^+ \mid B_0^+] &= \Pr[\exists i \in [n] \setminus [d] : s_i \geq t - \varepsilon \mid B_0^+] && \text{(By union bound)} \\
&\leq \sum_{i \in [n] \setminus [d]} \Pr[s_i \geq t - \varepsilon \mid B_0^+] \\
&\leq \sum_{i \in [n] \setminus [d]} e^3 \varepsilon L \Pr[s_i \geq t \mid B_0^+] && \text{(By (15))} \\
&= e^3 \varepsilon L\, \mathbb{E}[\#\{i \in [n] \setminus [d] : s_i \geq t\} \mid B_0^+] \\
&= e^3 \varepsilon L\, \mathbb{E}[\#\{i \in [n] \setminus [d] : s_i \geq t\}]. && (16)
\end{aligned}$$

To interpret the last equality above, we observe that $\#\{i \in [n] \setminus [d] : s_i \geq t\} = 0$ if and only if B_0^+ happens. The upper bound $\Pr[B_0^+] \leq \frac{5}{4}\Pr[B_\varepsilon^+]$ will follow from (16) together with Claim 41 and our choice of ε. $\square$

Claim 41. *Conditional on θ, t, if $\Pr[B_0^+] \geq n^{-d}$ then $\mathbb{E}[\#\{i \in [n] \setminus [d] : s_i \geq t\}] \leq 2d \log n$. If $\Pr[B_0^- \mid \theta, t] \geq n^d$ then $\mathbb{E}[\#\{i \in [n] \setminus [d] : s_i \leq t\}] \leq 2d \log n$.*

Proof. We prove the first implication, and the second follows analogously. For each $i \in [n] \setminus [d]$, let $X_i \in \{0, 1\}$ have value 1 if and only if $s_i \geq t$. Since θ, t are fixed and depend only on $a_1, \ldots, a_d$, the random variables $X_{d+1}, \ldots, X_n$ are independent. Write $X = \sum_{i=d+1}^{n} X_i$. The Chernoff bound gives

$$\Pr[X = 0] \leq \exp\left(-\frac{\mathbb{E}[X]}{2}\right).$$

As such, $\mathbb{E}[X] > 2d \log n$ would imply $\Pr[X = 0] < n^{-d}$, contradicting the original assumption that $\Pr[X = 0] \geq n^{-d}$. It follows that $\mathbb{E}[X] \leq 2d \log n$. □

Now we can prove Lemma 38 using Lemma 40.

Proof of Lemma 38. Fix any $I \in \binom{[n]}{d}$. By Lemma 40, we have that $\Pr[E_I] \leq \binom{n}{d}^{-1} + \frac{5}{4} \cdot \Pr[E_I \wedge (\delta \geq \varepsilon)]$ for $\varepsilon = \frac{1}{10e^3 L d \log n}$. This gives that

$$\frac{\Pr[E_I \wedge (\delta \geq \varepsilon)]}{\Pr[E_I]} \geq \frac{4}{5} - \binom{n}{d}^{-1} \cdot \frac{4}{5 \Pr[E_I]}$$

Moreover, since $\Pr[E_I] \geq 10 \binom{n}{d}^{-1}$, we have

$$\begin{aligned} \Pr[(\delta \geq \varepsilon) \mid E_I] &= \frac{\Pr[E_I \wedge (\delta \geq \varepsilon)]}{\Pr[E_I]} \\ &\geq \frac{4}{5} - \binom{n}{d}^{-1} \cdot \frac{4}{5 \Pr[E_I]} \geq 0.72, \end{aligned}$$

as desired. □

5.4 Randomized Lower-Bound for r: Inner Radius of a Ridge Projected onto $(d-1)$-Dimensional Space

The remaining lemma, which has no analogue when $d = 2$, will require more technical effort. Its proof is similar to Lemma 4.1.1 (Distance bound) in [ST04].

Lemma 42 (Randomized Lower-bound for r)**.** *Let $a_1, \cdots, a_n \in \mathbb{R}^d$ be independent L-log-Lipschitz random vectors. Let D denote the event that $\forall i, j \in [n], \|a_i - a_j\| \leq 3$. Fix any $I \in \binom{[n]}{d}$ and any $J \in \binom{I}{d-1}$, we have*

$$\Pr[\operatorname{dist}(W \cap \operatorname{affhull}(a_i : i \in I), \partial \operatorname{conv}(a_j : j \in J)) \leq \frac{1}{19200 d^4 L^2} \mid E_I \wedge A_J] \leq 0.1 + \Pr[\neg D \mid E_I \wedge A_J].$$

Proof. We may assume without loss of generality that $I = [d]$ and $J = [d-1]$. Apply the change of variables ϕ as in Definition 17 to $\{a_i : i \in [d]\}$ and obtain

$$\phi(\theta, t, b_1, \cdots, b_d) = (a_1, \cdots, a_d).$$

where $\theta \in \mathbb{S}^{d-1}, t \in \mathbb{R}, b_1, \cdots, b_d \in \mathbb{R}^{d-1}$. For any $i \in [n]$, let μ_i denote the probability density function of a_i. Writing the conditioning as part of the pdf, we find that the joint probability density of $t, \theta, b_1, \cdots, b_d, a_{d+1}, \ldots, a_n$ is proportional to

$$\operatorname{vol}_{d-1}(b_1, \cdots, b_d) \cdot \prod_{i=1}^{d} \bar{\mu}_i(t, \theta, b_i) \cdot \prod_{i=d+1}^{n} \mu_i(a_i) \cdot 1[E_{[d]} \wedge A_{[d-1]}], \tag{17}$$

where $\text{vol}_{d-1}(\cdot)$ is the volume function of the $(d-1)$-dimensional simplex, and $\bar{\mu}_i(t, \theta, b_i) = \mu_i(t\theta + R_\theta(b_i))$ is the induced probability density of b_i, which is L-log-Lipschitz, and $1[\cdot]$ denotes the indicator function. Write S for the event that

$$\text{dist}(W \cap \text{affhull}(a_i : i \in I), \partial \,\text{conv}(a_j : j \in J)) \leq \frac{1}{19200 d^4 L^2}.$$

In this language, our goal is to prove that $\Pr[S] \leq 0.1 + \Pr[\neg D]$.

Let D' denote the event that $\|b_i - b_j\| \leq 3$ for all $i, j \in [d]$. Each of the events E_I, A_J, S, D', D are functions of the random variables $\theta, t, b_1, \ldots, b_d, a_{d+1}, \ldots, a_n$. We then use Fubini's theorem to write

$$\Pr_{\theta, t, b_1, \ldots, b_d, a_{d+1}, \ldots, a_n}[S] = \mathbb{E}_{\theta, t, a_{d+1}, \ldots, a_n}[\Pr_{b_1, \ldots, b_d}[S]]$$

With probability 1 over the choice of $\theta, t, a_{d+1}, \ldots, a_n$, the inner term satisfies all the conditions of Lemma 43. Specifically, since the value of $1[E_{[d]}]$ is already fixed, the intersection $\ell = (t\theta + \theta^\perp) \cap W$ is a line, the event $A_{[d-1]}$ is equivalent to $\ell \cap \text{conv}(b_1, \ldots, b_{d-1}) \neq \emptyset$. From Lemma 18, the joint probability distribution of $b_1, \ldots, b_d$ is thus proportional to

$$\text{vol}_{d-1}(b_1, \cdots, b_d) \cdot \prod_{i=1}^{d} \bar{\mu}_i(b_i) \cdot 1[\ell \cap \text{conv}(b_1, \ldots, b_{d-1} \neq \emptyset]$$

Applying Lemma 43 to the term $\Pr_{b_1, \ldots, b_d}[S]$ we find

$$\begin{aligned}\mathbb{E}_{\theta, t, a_{d+1}, \ldots, a_n}[\Pr_{b_1, \ldots, b_d}[S]] &\leq \mathbb{E}_{\theta, t, a_{d+1}, \ldots, a_n}[0.1 + \Pr_{b_1, \ldots, b_d}[\neg D']] \\ &= 0.1 + \Pr[\neg D'] \leq 0.1 + \Pr[\neg D],\end{aligned}$$

using Fubini's theorem for the equality and the fact that $\neg D'$ implies $\neg D$ for the final inequality. □

Lemma 43 (Randomized lower bound for r after change of variables)**.** *Let $b_1, \ldots, b_d \in \mathbb{R}^{d-1}$ be random vectors with joint probability density proportional to*

$$\text{vol}_{d-1}(b_1, \cdots, b_d) \cdot \prod_{i=1}^{d} \bar{\mu}_i(b_i)$$

where $\bar{\mu}_i$ is L-log-Lipschitz for each $i \in [d]$. Let D' denote the event that the set $\{b_1, \ldots, b_d\}$ has Euclidean diameter of at most 3. Given any fixed one-dimensional line $\ell \subset \mathbb{R}^{d-1}$, we have that

$$\begin{aligned}&\Pr\left[\left(\text{dist}\left(\ell, \partial \,\text{conv}(b_1, \ldots, b_{d-1})\right) < \frac{1}{19200 d^4 L^2}\right) \mid \ell \cap \text{conv}(b_1, \ldots, b_{d-1}) \neq \emptyset\right] \\ &\quad \leq 0.1 + \Pr[D' \mid \ell \cap \text{conv}(b_1, \ldots, b_{d-1}) \neq \emptyset].\end{aligned}$$

Proof. We can write the distance from ℓ to $\partial \,\text{conv}(b_1, \ldots, b_{d-1})$ as

$$\begin{aligned}&\text{dist}\left(\ell, \partial \,\text{conv}(b_1, \ldots, b_{d-1})\right) \\ &= \min_{i \in [d-1]} \lambda_i \cdot \text{dist}(\pi_{w^\perp}(b_i), \text{affhull}(\pi_{w^\perp}(b_j) : j \in [d-1], j \neq i)) \\ &\geq \min_{i \in [d-1]} \lambda_i \cdot \min_{k \in [d-1]} \text{dist}(\pi_{w^\perp}(b_k), \text{affhull}(\pi_{w^\perp}(b_j) : j \in [d-1], j \neq k))\end{aligned}$$

Abbreviate, for $k \in [d-1]$,

$$r_k = \text{dist}(\pi_{w^\perp}(b_k), \text{affhull}(\pi_{w^\perp}(b_j) : j \in [d-1], j \neq k)).$$

Let T denote the event that $\ell \cap \text{conv}(b_1, \ldots, b_{d-1}) \neq \emptyset$. We now find using the union bound, for $\alpha, \beta > 0$,

$$\begin{aligned}\Pr[\text{dist}(\ell, \partial \,\text{conv}(b_1, \ldots, b_{d-1})) < \alpha\beta \mid T] &\leq \Pr[\min_{i \in [d-1]} \lambda_i < \alpha \mid T] + \Pr[\min_{k \in [d-1]} r_k < \beta \mid T] \\ &\leq \Pr[\min_{i \in [d-1]} \lambda_i < \alpha \mid D' \wedge T] + \Pr[D' \mid T] + \Pr[\min_{k \in [d-1]} r_k < \beta \mid T]\end{aligned}$$

By Lemma 44 we know that $\Pr[\min_{i \in [d-1]} \lambda_i < \alpha \mid D' \wedge T] \leq 0.05$ for $\alpha = \frac{1}{120 d^2 L}$. By Lemma 45, we know that $\Pr[\min_{k \in [d-1]} r_k < \beta \mid T] \leq \sum_{k \in [d-1]} \Pr[r_k < \beta \mid T] \leq 0.05$ for $\beta = \frac{1}{160 d^2 L}$. This proves the lemma. □

The following two Lemmas are the key to prove Lemma 43. Let $\lambda \in \mathbb{R}^{d-1}_{\geq 0}$ be the unique solution to $\sum_{i=1}^{d-1} \lambda_i b_i = \ell \cap \mathrm{conv}(b_1, \cdots, b_{d-1})$ and $\sum_{i=1}^{d-1} \lambda_i = 1$. First, we use Lemma 44 to show that the every convex parameter λ_i is at least $\Omega(1/d^2 L)$ with constant probability.

Lemma 44 (Lower-bound for Convex Parameters of Vertices on the Ridge)**.** *Let $b_1, \cdots, b_d \in \mathbb{R}^{d-1}$ be random vectors with joint probability density proportional to*

$$\mathrm{vol}_{d-1}(b_1, \cdots, b_d) \cdot \prod_{i=1}^{d} \bar{\mu}_i(b_i)$$

where each $\bar{\mu}_i : \mathbb{R}^{d-1} \to \mathbb{R}_+$ is L-log-Lipschitz. Given any one-dimensional line $\ell \subset \mathbb{R}^{d-1}$ and conditional on $\ell \cap \mathrm{conv}(b_i : i \in [d-1]) \neq \emptyset$. Let $\lambda \in \mathbb{R}^{d-1}_{+}$ be the unique solution to $\sum_{i=1}^{d-1} \lambda_i b_i \in \ell \cap \mathrm{conv}(b_i : i \in [d-1])$. Let D' denote the event that $\forall i, j \in [d]$, $\|b_i - b_j\| \leq 3$. Then we have

$$\Pr\left[\forall i \in [d-1] : \lambda_i \geq \frac{1}{120 d^2 L} \middle| D' \wedge \ell \cap \mathrm{conv}(b_i : i \in [d-1]) \neq \emptyset\right] \geq 0.95.$$

Proof. By using the union bound, it suffices to prove for each $i \in [d-1]$ that

$$\Pr[\lambda_i < \frac{1}{120 d^2 L} \mid D' \wedge \ell \cap \mathrm{conv}(b_j : j \in [d-1]) \neq \emptyset] \leq \frac{1}{20d}.$$

Fix any $i \in [d-1]$, without loss of generality $i = 1$. We can assume $\ell = w\mathbb{R}$ for a non-zero $w \in \mathbb{R}^{d-1}$. Thus λ is defined to satisfy $\sum_{j=1}^{d-1} \lambda_j \pi_{w^\perp}(b_j) = 0$.

For any given values of $b_1 - b_j$, $j \in [d]$, which determine the shape of the simplex $\mathrm{conv}(b_j : j \in [d])$, we prove the result using the randomness in $\pi_{w^\perp}(b_1)$, the position of the simplex in the space $w^\perp$. For the remainder of this proof, we can consider all $b_j, j \in [d]$ to be functions of b_1. If we furthermore fix any value for $w^\top b_1$ then $\mathrm{vol}(\mathrm{conv}(b_j : j \in [d])$ is fixed, hence $\pi_{w^\perp}(b_1)$ has probability density $\mu'(\pi_{w^\perp}(b_1)) \propto \prod_{j=1}^{d} \mu_j(b_j)$, which is dL-log-Lipschitz in $\pi_{w^\perp}(b_1)$ with respect to the $d-2$-dimensional Lebesgue measure on $w^\perp$.

Write $M = \mathrm{conv}(\pi_{w^\perp}(b_1 - b_j) : j \in [d-1]) \subset w^\perp$, for which we can see that $\pi_{w^\perp}(b_1) \in M$ if and only if $\lambda \geq 0$. It then remains to show that

$$\Pr[\lambda_1 < \frac{1}{120 d^2 L} \mid D' \wedge \pi_{w^\perp}(b_1) \in M] < \frac{1}{20d}. \tag{18}$$

For any $j \in [d-1]$, let $l_j : M \to [0,1]$ be the function sending any point to its j'th convex coefficient, i.e., the functions satisfy $\sum_{j=1}^{d-1} l_j(x) = 1$ and $\sum_{j=1}^{d-1} l_j(x) \cdot \pi_{w^\perp}(b_1 - b_j) = x$ for every $x \in M$. For any $1 \geq \alpha \geq 0$, observe that l_1 takes values in the interval $[\alpha, 1]$ on the set $(1-\alpha)M$. Hence we get

$$\begin{aligned}
\Pr[\lambda_1 \geq \alpha \mid \pi_{w^\perp}(b_1) \in M] &= \frac{\int_M \mu'(x) \mathbf{1}[l_1(x) \geq \alpha] \mathrm{d}x}{\int_M \mu'(x) \mathrm{d}x} \\
&\geq \frac{\int_{(1-\alpha)M} \mu'(x) \mathrm{d}x}{\int_M \mu'(x) \mathrm{d}x} && (\forall x \in (1-\alpha)M,\ l_1(x) \geq \alpha) \\
&= \frac{(1-\alpha)^{d-2} \int_M \mu'((1-\alpha)x) \mathrm{d}x}{\int_M \mu'(x) \mathrm{d}x} \\
&\geq (1-\alpha)^{d-2} \max_{x \in M} e^{-dL\|\alpha x\|}. && (\text{By } d\text{-log-Lipschitzness of } \mu')
\end{aligned}$$

By definition of D', we know that M has Euclidean diameter at most 3. Thus we can bound $\|\alpha x\| \leq 3\alpha$ for any $x \in M$. Now take $\alpha = \frac{1}{120 d^2 L}$, we find

$$\begin{aligned}
\Pr[(\lambda_i < \frac{1}{120 d^2 L}) \mid D' \wedge \pi_{w^\perp}(b_i) \in M] &\leq 1 - \Pr[\lambda_i \geq \frac{1}{120 d^2 L} \mid D' \wedge \pi_{w^\perp}(b_i) \in M] \\
&\leq 1 - (1 - \frac{1}{120 d^2 L})^{d-2} e^{-1/40d} \leq 1/20d,
\end{aligned}$$

where the last line comes from $L \geq 1$. Thus (18) holds as desired. □

In the second part, we lower bound the distance between each vertex b_j (where $j \in [d-1]$) and the $(d-3)$-dimensional hyperplane spanned by the other vertices $\mathrm{affhull}(b_j : j \in [d-1], j \neq i)$. We show the following lemma:

Lemma 45. *Let $b_1, \cdots, b_d \in \mathbb{R}^{d-1}$ be random vectors with joint probability density proportional to*

$$\mathrm{vol}_{d-1}(b_1, \cdots, b_d) \cdot \prod_{i=1}^{d} \bar{\mu}_i(b_i)$$

where each $\bar{\mu}_i : \mathbb{R}^{d-1} \to [0,1]$ is L-log-Lipschitz. Given any one-dimensional line $\ell \subset \mathbb{R}^{d-1}$, and let $w \in \mathbb{S}^{d-2}$ be any unit direction of ℓ. For any $i \in [d-1]$ we have

$$\Pr\Big[\mathrm{dist}\left(\pi_{w^\perp}(b_i), \mathrm{affhull}(\pi_{w^\perp}(b_j) : j \in [d-1], j \neq i)\right) \geq \frac{1}{160d^2L} \;\Big|\; \ell \cap \mathrm{conv}(b_i : i \in [d-1]) \neq \emptyset\Big] \geq 1 - \frac{1}{20d}.$$

Proof. In the following arguments, we condition on $\ell \cap \mathrm{conv}(b_i : i \in [d-1]) \neq \emptyset$. Without loss of generality, set $i = d-1$ and assume that ℓ is a linear subspace, i.e., $\pi(\ell) = 0$.

We start with a coordinate transformation. Let $\phi \in w^\perp \cap \mathbb{S}^{d-2}$ denote the unit vector satisfying $\phi^\top b_1 = \phi^\top b_j > 0$ for all $j = 1, \ldots, d-2$. Note that ϕ is uniquely defined almost surely: $w^\perp$ is a $(d-2)$-dimensional linear space and we impose $(d-3)$ linear constraints $\{\phi^\top a_1 = \phi^\top a_j, \forall j \in [d-2]\}$. Almost surely, these give an one-dimensional linear subspace which, after adding the unit norm and $b_1^\top \phi > 0$ constraint, leaves a unique choice of ϕ.

Now define $h \in \mathbb{R}$ by $h = \phi^\top b_1$ and define $\alpha \in \mathbb{R}$ by $\alpha h = -\phi^\top b_{d-1}$. Since $0 \in \mathrm{conv}(\pi(b_i) : i \in [d-1])$ but $\phi^\top b_i > 0$ for all $i \in [d-2]$, we must have $\alpha \geq 0$ for otherwise ϕ would separate $\mathrm{conv}(\pi(b_i) : i \in [d-1])$ from 0. Again from almost-sure non-degeneracy we get $\alpha > 0$ and $h \neq 0$. We define the following coordinate transformation:

$$\begin{aligned} b_j &= h\phi + c_j, \quad \forall j \in [d-2] \\ b_{d-1} &= -\alpha h\phi + c_{d-1} \end{aligned}$$

where for each $j \in [d-1]$, $c_j \in \phi^\perp \cap H$ has $(d-3)$ degrees of freedom. From here on out, we consider the vertices $(b_1, \cdots, b_{d-1})$ to be a function of $(h, \alpha, \phi, c_1, \ldots, c_{d-1})$. Again by Lemma 18, the induced joint probability density on the random variables $(h, \alpha, \phi, c_1, \ldots, c_{d-1}, b_d)$, is proportional to

$$\begin{aligned} &\mathrm{vol}_{d-1}(\mathrm{conv}(b_1, \ldots, b_d)) \cdot \mathrm{vol}_{d-3}(\mathrm{conv}(c_1, \cdots, c_{d-2})) \cdot \prod_{j=1}^{d} \bar{\mu}_j(b_j) \\ \propto\ &\mathrm{vol}_{d-1}(\mathrm{conv}(b_1, \ldots, b_d)) \cdot \mathrm{vol}_{d-3}(\mathrm{conv}(\pi_{w^\perp}(c_1), \ldots, \pi_{w^\perp}(c_{d-2}))) \cdot \prod_{j=1}^{d} \bar{\mu}_j(b_j). \end{aligned}$$

Condition on the exact values of $(\alpha, \phi, c_1, \ldots, c_{d-1}, b_d)$. Note that the event $0 \in \mathrm{conv}(\pi(b_i) : i \in [d-1])$ depends only on these variables and not on h, and the same is true for $\mathrm{vol}_{d-3}(\mathrm{conv}(\pi_{w^\perp}(c_1), \ldots, \pi_{w^\perp}(c_{d-2})))$. By Lemma 18, the induced probability density on h is now proportional to

$$\mathrm{vol}_{d-1}(\mathrm{conv}(b_1, \ldots, b_d)) \cdot \prod_{j=1}^{d-1} \tilde{\mu}_j(h),$$

where $\tilde{\mu}_j(h) := \bar{\mu}_j(h\phi + c_j), j \in [d-2]$ and $\tilde{\mu}_{d-1}(h) := \bar{\mu}_{d-1}(\alpha h\phi + c_{d-1})$. Since each $\bar{\mu}_j$ is L-log-Lipschitz, it follows that the product $\prod_{j=1}^{d-1} \tilde{\mu}_j(h)$ is $(d-2+\alpha)L \leq d(1+\alpha)L$-log-Lipschitz in h.

Next, consider the volume term. We can write $\mathrm{vol}_{d-1}(\mathrm{conv}(b_1, \ldots, b_d))$ as a constant depending on d times the absolute value of the determinant of matrix

$$\begin{bmatrix} (b_1 - b_d)^\top \\ \vdots \\ (b_{d-1} - b_d)^\top \end{bmatrix} = \begin{bmatrix} (t\phi + c_1 - b_d)^\top \\ \vdots \\ (t\phi + c_{d-2} - b_d)^\top \\ (-\alpha t\phi + c_{d-1} - b_d)^\top \end{bmatrix} = \begin{bmatrix} (c_1 - b_d)^\top \\ \vdots \\ (c_{d-2} - b_d)^\top \\ (c_{d-1} - b_d)^\top \end{bmatrix} + h \cdot \begin{bmatrix} 1 \\ \vdots \\ 1 \\ -\alpha \end{bmatrix} \phi^\top.$$

where $\theta \in \mathbb{S}^{d-1}$ denotes a normal vector to H. Define

$$B := \begin{bmatrix} (c_1 - b_d)^\top \\ \vdots \\ (c_{d-2} - b_d)^\top \\ (c_{d-1} - b_d)^\top \\ \theta^\top \end{bmatrix}, v := \begin{bmatrix} 1 \\ \vdots \\ 1 \\ -\alpha \\ 0 \end{bmatrix}.$$

Then by the matrix determinant lemma, we can write the volume as the absolute value of an affine function of t (which is a convex function):

$$\begin{aligned} \mathrm{vol}_{d-1}(\mathrm{conv}(b_1, \cdots, b_d)) &= \left|\det(B) + hv\phi^\top\right| \\ &= \left|\det(B)(1 + \phi^\top B^{-1} v \cdot h)\right| \end{aligned}$$

Hence, we have found a convex function $k : \mathbb{R} \to \mathbb{R}_{\geq 0}$ and a $d(1+\alpha)L$-log-Lipschitz function $\nu : \mathbb{R} \to \mathbb{R}_{\geq 0}$ such that h has probability density proportional to $k(h) \cdot \nu(h)$.

To finalize the argument, we write $\mathrm{dist}(\pi(b_i), \mathrm{affhull}(\pi(b_j) : j \in [d-1], j \neq i)) = |(1+\alpha)h|$. It follows that the signed distance $(1+\alpha)h$ has a probability density function proportional to the product of a dL-log-Lipschitz function and a convex function. The result follows from Lemma 46 by plugging in the signed distance $(1+\alpha)h$ and $K = dL, \varepsilon = \frac{1}{160d^2L}$. □

Lemma 46. *Assume that $h : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a K-log-Lipschitz function and $g : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a convex function such that $\int_{-\infty}^{\infty} g(x) \cdot h(x)\mathrm{d}x = 1$. Suppose that $X \in \mathbb{R}$ is distributed with probability density $g(X) \cdot h(X)$. For any $\varepsilon > 0$ we have* $\Pr[X \in [-\varepsilon, \varepsilon]] \leq 8\varepsilon K$.

Proof. We can assume that $\varepsilon < 1/(8K)$, for otherwise the bound is trivial. First, we use the rudimentary upper bound

$$\Pr\left[X \in [-\varepsilon, \varepsilon]\right] \leq \Pr\left[X \in [-\varepsilon, \varepsilon] \mid X \in [-1/K, 1/K]\right] = \frac{\int_{-\varepsilon}^{\varepsilon} g(x) \cdot h(x)\mathrm{d}x}{\int_{-1/K}^{1/K} g(x) \cdot h(x)\mathrm{d}x}.$$

Log-Lipschitzness implies that for any $\gamma > 0$ we have

$$e^{-\gamma K} h(0) \int_{-\gamma}^{\gamma} g(x)\mathrm{d}x \leq \int_{-\gamma}^{\gamma} g(x) \cdot h(x)\mathrm{d}x \leq e^{\gamma K} h(0) \int_{-\gamma}^{\gamma} g(x)\mathrm{d}x,$$

and hence we get

$$\Pr[X \in [-\varepsilon, \varepsilon]] \leq e^{1+\varepsilon K} \frac{\int_{-\varepsilon}^{\varepsilon} g(x)\mathrm{d}x}{\int_{-1/K}^{1/K} g(x)\mathrm{d}x} \leq e^{1+\varepsilon K} \frac{2\varepsilon \max_{x \in [-\varepsilon, \varepsilon]} g(x)}{\int_{-1/K}^{1/K} g(x)\mathrm{d}x}$$

Since $g(x)$ is convex, at least one of

$$\max_{x \in [-\varepsilon, \varepsilon]} g(x) \leq \min_{x \in [-1/K, -\varepsilon]} g(x) \quad \text{or} \quad \max_{x \in [-\varepsilon, \varepsilon]} g(x) \leq \min_{x \in [\varepsilon, 1/K]} g(x)$$

holds. Without loss of generality, assume the second case holds. Then we bound

$$\frac{\max_{x \in [-\varepsilon, \varepsilon]} g(x)}{\int_{-1/K}^{1/K} g(x)\mathrm{d}x} \leq \frac{\max_{x \in [-\varepsilon, \varepsilon]} g(x)}{\int_{\varepsilon}^{1/K} g(x)\mathrm{d}x} \leq \frac{1}{1/K - \varepsilon}.$$

To summarize, we find $\Pr[X \in [-\varepsilon, \varepsilon]] \leq e^{1+\varepsilon K} \cdot \frac{2\varepsilon}{1/K - \varepsilon}$. Since $\varepsilon < 1/(8K)$ this implies

$$\Pr[X \in [-\varepsilon, \varepsilon]] \leq 2e^{8/7} \cdot \frac{7}{6} \cdot \varepsilon K \leq 8\varepsilon K.$$

□

5.5 Combining Together and Proof of Lemma 35

In this section, we combine the deterministic argument in Lemma 36, and the randomized arguments in Lemma 38 and Lemma 42. We can finally show the main technical lemma (Lemma 35).

Proof of Lemma 35. Without loss of generality, let $I = [d]$ and write $E = E_I$. Suppose $p' = A_J = \mathrm{conv}(a_j : j \in J) \cap W$ is the next vertex after the edge $F_I \cap W$. Here $J \in \binom{[n]}{d-1}$ and R_J is the $(d-2)$-dimensional ridge. With probability 1, the polytope $\mathrm{conv}(a_1, \ldots, a_n)$ is non-degenerate and $W \cap R'$ is a single point for any ridge R' of $\mathrm{conv}(a_1, \ldots, a_n)$ that intersects with W. We will show that conditional on E, each of the following conditions in the deterministic argument (Lemma 36) is satisfied with good probability:

1. (Bounded diameter) $\forall i, j \in [n]$, $\|a_i - a_j\| \le 3$;
2. (Lower bound of δ) $\min_{k \in [n] \setminus I} \mathrm{dist}(\mathrm{affhull}(F_I), a_k) \ge \Omega(\frac{1}{Ld \log n})$;
3. (Lower bound of r) $\forall J \in \binom{I}{d-1}$ for which the ridge $R_J = \mathrm{conv}(a_j : j \in J)$ has nonempty intersection with W, we have $\mathrm{dist}(F_I \cap W, \partial R_J) \ge \Omega(\frac{1}{d^4 L^2})$.

Note for the last point that Lemma 36 only requires this for the set J which indexes the second vertex of $F_I \cap W$ in clockwise direction, but we prove it for both of the sets J for which $R_J \cap W \neq \emptyset$.

First, we write D as the event that $\forall i, j \in [n]$ for which $\|a_i - a_j\| \le 3$. From Lemma 16, for any $\sigma \le \frac{1}{8\sqrt{d \log n}}$, with probability at least $1 - \binom{n}{d}^{-1}$, we have $\max_{i \in [n]} \|\hat{a}_i\| \le 1 + 4\sigma\sqrt{d \log n} \le \frac{3}{2}$, i.e. $\Pr[D] \ge 1 - \binom{n}{d}^{-1}$. Then we have

$$\Pr[D \mid E] = \Pr[D \wedge E] / \Pr[E] \le \frac{\Pr[D]}{\Pr[E]} \ge 0.9,$$

using the assumption that $\Pr[E_I] \ge 10\binom{n}{d}^{-1}$.

Next, we consider $\delta := \mathrm{dist}(\mathrm{affhull}(a_1, \ldots, a_d), \{a_{d+1}, \ldots, a_n\})$. Using Lemma 38, we have $\Pr[\delta \ge \frac{1}{10e^3 Ld \log n} \mid E] \ge 0.72$.

Finally, we consider $r := \max_J \mathrm{dist}(\mathrm{affhull}(a_1, \cdots, a_d) \cap W, \partial R_J)$ subject to all $J \in \binom{I}{d-1}$ such that A_J happens (in other words, $R_J = \mathrm{conv}(a_j : j \in J)$ is a ridge of F_I such that $R_J \cap W \neq \emptyset$). By union bound,

$$\begin{aligned}
&\Pr\left[\exists (J \in \binom{I}{d-1}, A_J), \mathrm{dist}(\mathrm{affhull}(F \cap W), \partial R_J \ge \frac{1}{19200 d^4 L^2} \mid E\right] \\
\ge& 1 - \sum_{J \in \binom{I}{d-1}} \Pr[A_J \wedge \mathrm{dist}(\mathrm{affhull}(F \cap W), \partial R_J < \frac{1}{19200 d^4 L^2} \mid E] \\
=& 1 - \sum_{J \in \binom{I}{d-1}} \Pr[\mathrm{dist}(\mathrm{affhull}(F \cap W), \partial R_J < \frac{1}{19200 d^4 L^2} \mid E \wedge A_J] \Pr[A_J \mid E]. \qquad (19)
\end{aligned}$$

From Lemma 42, for each $J \in \binom{I}{d-1}$, we know that

$$\Pr[\mathrm{dist}(\mathrm{affhull}(F \cap W), \partial R_J < \frac{1}{19200 d^4 L^2} \mid E \wedge A_J] \le 0.1 + \Pr[\neg D \mid E \wedge A_J],$$

Notice that when E happens, there are exactly two distinct ridges $R_J, R_{J'}$ that has nonempty intersection with W (or A_J happens), thus $\sum_{J \in \binom{I}{d-1}} \Pr[A_J \mid E] = 2$. Then (19) becomes

$$\begin{aligned}
&\Pr\left[\exists (J \in \binom{I}{d-1}, A_J), \mathrm{dist}(\mathrm{affhull}(F \cap W), \partial R_J \ge \frac{1}{19200 d^4 L^2} \mid E\right] \\
\ge& 1 - \sum_{J \in \binom{I}{d-1}} (0.1 + \Pr[\neg D \mid E \wedge A_J]) \Pr[A_J \mid E] \\
\ge& 1 - 0.1 \cdot 2 - 2 \cdot \Pr[\neg D \mid E] \ge 0.6
\end{aligned}$$

Therefore by union bound, the three conditions hold with probability at least $1 - (1 - 0.9) - (1 - 0.72) - (1 - 0.6) \ge 0.1$, and the lemma directly follows from Lemma 36. $\square$

6 Smoothed Complexity Lower Bound

In this section, we present the lower bound of the smoothed complexity by studying the intersection between the smoothed dual polytope $Q = \mathrm{conv}(a_1, \cdots, a_n) \subset \mathbb{R}^d$ (where each a_i is under Gaussian perturbation), and the two-dimensional shadow plane $W \subset \mathbb{R}^d$. Our main result is as follows:

Theorem 47. *For any $d \geq 5, n = 4d - 15, \sigma \leq \frac{1}{360d\sqrt{\log n}}$, there exists a two-dimensional linear subspace $W \subset \mathbb{R}^d$ and vectors $\bar{a}_1, \cdots, \bar{a}_n \in \mathbb{R}^d$, $\max_{i \in n} \|\bar{a}_i\|_1 \leq 1$ such that the following holds. Let $a_1, \cdots, a_n$ be independent Gaussian random variables where each $a_i \sim \mathcal{N}_d(\bar{a}_i, \sigma^2 I)$, then with probability at least $1 - \binom{n}{d}^{-1}$, we have*

$$\mathrm{edges}(\mathrm{conv}(a_1, \cdots, a_n) \cap W) \geq \Omega\left(\min\left(\frac{1}{\sqrt{d\sigma\sqrt{\log n}}}, 2^d\right)\right).$$

Theorem 47 is the direct consequence of the next theorem, which is a lower bound with the more general adversarial perturbations of bounded magnitude:

Theorem 48. *For any $d \geq 5, n = 4d - 15$, there exists a two-dimensional linear subspace $W \subset \mathbb{R}^d$ and vectors $\bar{a}_1, \cdots, \bar{a}_n \in \mathbb{R}^d$, $\max_{i \in n} \|\bar{a}_i\|_1 \leq 1$ such that the following holds. Assume $a_1, \cdots, a_n \in \mathbb{R}^d$ satisfy $\|a_i - \bar{a}_i\|_1 \leq \varepsilon$ for all $i \in [n]$ with $\varepsilon < \frac{1}{90}$, then*

$$\mathrm{edges}(\mathrm{conv}(a_1, \cdots, a_n) \cap W) \geq \Omega\left(\min\left(\frac{1}{\sqrt{\varepsilon}}, 2^d\right)\right).$$

The rest of this section is organized as follows: In Section 6.1, we construct a polytope P represented by a system of inequalities, and a two-dimensional shadow plane W. An informal intuition behind these inequalities is described in Section 6.2 In Section 6.3, we show that $\pi_W(P)$ approximates the unit disk $\mathbb{B}_2^2$. In Section 6.4, we analyze the largest ℓ_∞ ball contained in P and the smallest ℓ_∞ ball containing P. Section 6.5 investigates the polar polytope $Q = (P - x)^\circ$ of a shift of P and derives bounds on the radius of its largest contained ℓ_1 ball and smallest containing ℓ_1 ball. Finally, Section 6.6 shows that the small ratio between these radii imply that any perturbation $\tilde{Q}$ still has $\tilde{Q} \cap W$ approximates the unit disk $\mathbb{B}_2^2$ well and uses this fact to prove Theorem 47.

6.1 Construction of the Primal Polytope

In this subsection, we first construct the primal polytope and the two-dimensional plane W. For $k \in \mathbb{N}$, we construct a $(k+5)$-dimensional polytope. We will use the following vectors in the definition:

- Define $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \mathbb{R}^2$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
- For every $i \in [k]$, define the pair of orthogonal unit vectors $w_i = \begin{bmatrix} \cos(\pi/2^{i+2}) \\ \sin(\pi/2^{i+2}) \end{bmatrix} \in \mathbb{R}^2$ and $v_i = \begin{bmatrix} \sin(\pi/2^{i+2}) \\ -\cos(\pi/2^{i+2}) \end{bmatrix} \in \mathbb{R}^2$.

With these definitions in mind, let $P' \subset \mathbb{R}^{3k+5}$ denote the set of points $(x, y, p_0, \ldots, p_k, t, s)$, where $x, y, p_0, p_1, \cdots, p_k \in \mathbb{R}^2, t \in \mathbb{R}^k, s \in \mathbb{R}$, satisfying the following system of linear inequalities:

$$e_1^\top p_0 \geq |x|, e_2^\top p_0 \geq |y| \tag{20}$$
$$w_i^\top p_i = w_i^\top p_{i-1},\ \forall i \in [k] \tag{21}$$
$$t_i + is = v_i^\top p_i \geq |v_i^\top p_{i-1}|,\ \forall i \in [k] \tag{22}$$
$$e_1^\top p_k \leq 1 \tag{23}$$
$$\mathbf{0}_k \preceq t \preceq \mathbf{1}_k \tag{24}$$

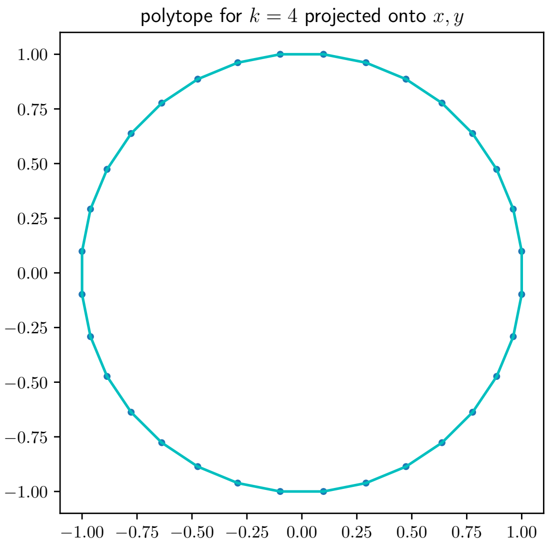

Figure 3: Vertices of the projected primal polytope $\pi_W(P)$ (see (26)) without perturbation for $k = 4$.

$$0 \le s \le 1. \tag{25}$$

We remark that p_0, t, s uniquely define the values of $p_1, p_2, \cdots, p_k$ via (21) and (22). As such, define the polytope $P \subset \mathbb{R}^{k+5}$ as:

$$P = \{(x, y, p_0, t, s) : \exists\ p_1, \cdots, p_k, \text{ s.t. } (x, y, p_0, \cdots, p_k, t, s) \in P'\}. \tag{26}$$

The plane W of interest is that spanned by the unit vectors in the x and y directions.

An plot of the vertices of the projected polytope $\pi_W(P)$ can be found in Figure 3 for $k = 4$. Note that the figure appears to depict a regular polygon with 2^{k+1} vertices.

6.2 Intuition of the Construction

To explain the intuition behind the equations (20 - 25), let us consider the simpler system of inequalities in variables $r_0, \ldots, r_k \in \mathbb{R}^2$:

$$e_1^\top r_0 \ge |x|, e_2^\top r_0 \ge |y| \tag{27}$$

$$w_i^\top r_i \ge w_i^\top r_{i-1},\ \forall i \in [k] \tag{28}$$

$$v_i^\top r_i \ge |v_i^\top r_{i-1}|,\ \forall i \in [k] \tag{29}$$

$$e_1^\top r_k \le 1 \tag{30}$$

$$e_2^\top r_i \ge 0,\ \forall i \in [k] \tag{31}$$

Let $R \subset \mathbb{R}^{2k+2}$ denote the set of vectors $(r_0, \ldots, r_k)$ satisfying the above inequalities. For each $i = 0, \ldots, k$, write $R_i = \{r_i : \exists r_0, \ldots, r_{i-1}, r_{i+1}, \ldots, r_k \text{ s.t. } (r_0, \ldots, r_k) \in R\} \subset \mathbb{R}^2$ for the projections of R onto the two-dimensional coordinate subspace of r_i. Also, let W be the two-dimensional plane spanned by the x and y directions, so that $\pi_W(R) = \{(x, y) : \exists r_0 \in R_0, r_0 \ge (|x|, |y|)^\top\}$. The vertices of these sets are depicted in Figure 4.

For these sets, we have the following observations. We will skip the proof since they only give illustrations of our analysis and will not be functional to the proof of Theorem 47.

1. The set R_k can be described by the inequalities $e_2^\top r_k \ge 0, e_1^\top r_k \le 1, v_i^\top r_i \ge 0$. These inequalities describe a small slice of the regular 2^{k+1}-gon.

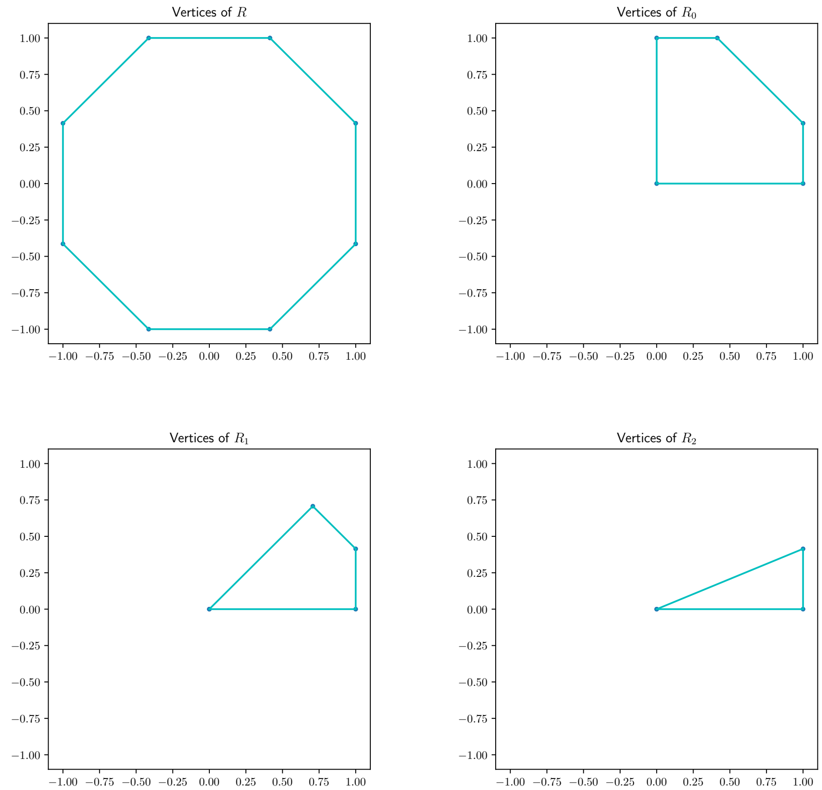

Figure 4: Vertices of $\pi_H(R), R_0, R_1, R_2$ for $k = 2$.

2. For each $i = 0, \ldots, k-1$, the set R_i is obtained by taking the union of R_{i+1} and its mirror image in the line spanned by w_i.
3. For each $i = 1, \ldots, k$, the set R_i can be described as the restriction of a regular 2^{k+1}-gon intersected with the set $\{r_i : e_2^\top r_i \geq 0, v_i^\top r_i \geq 0\}$.
4. The set R_0 can be described as the restriction of a regular 2^{k+1}-gon intersected with the non-negative orthant.
5. The set $R = \{(x, y) : \exists r_0 \in R_0, r_0 \geq (|x|, |y|)^\top\}$ is a regular 2^{k+1}-gon.

More careful inspection allows for the following additional observations, which we also state without proof:

1. Turning (28) from an inequality to an equality constraint does not change any of the sets $R_0, \ldots, R_k$.
2. Removing the constraint (31) does not change R_0.
3. Adding upper bounds $v_i^\top r_i \leq 2$ does not change R_0.

Each of the above-mentioned changes either serves to increase the size of the largest ball (in the affine hull of R) contained in the relative interior of R or to decrease the size of the smallest ball containing R. The addition of the variable s in the construction of P serves to further increase the size of the largest ball contained in P without increasing the size of the smallest ball containing P.

6.3 Projected Primal Polytope Approximates Two-Dimensional Unit Disk

In this subsection, we will show that the polytope P we constructed in (26) has a projection $\pi_W(P)$ which approximates the two-dimensional unit disk $\mathbb{B}_2^2 = \{x, y \in \mathbb{R} : x^2 + y^2 \leq 1\}$ within exponentially small error:

Lemma 49 (Projected primal polytope approximates the two-dimensional disk)**.** *For any $k \in \mathbb{N}$, let $P \subset \mathbb{R}^{k+5}$ be the polytope defined by the linear system (26) with variables $x, y, s \in \mathbb{R}, p_0 \in \mathbb{R}^2, t \in \mathbb{R}^k$. Let W be the two-dimensional subspace spanned by the directions of x and y. Then we have*

$$\mathbb{B}_2^2 \subset \pi_W(P) \subset (1 + 4^{-k-2})\mathbb{B}_2^2.$$

Lemma 49 directly follows from the next two lemmas. First, we show that the two-dimensional unit disk is contained in $\pi_W(P)$:

Lemma 50 (Inner radius of the projected primal polytope)**.** *For every $x, y \in \mathbb{R}$ with $x^2 + y^2 \leq 1$ there exist $p_0 \in \mathbb{R}^2$, $t \in \mathbb{R}^k$ and $s \in \mathbb{R}$ such that $(x, y, p_0, t, s) \in P$.*

Proof. Given such x, y, set $s = 0$ and set p_0, t such that (20) and (22) are satisfied with equality. This will result in $\sqrt{x^2 + y^2} = \|p_0\| = \|p_i\|$ for every $i \in [k]$. Since $t = |v_i^\top p_{i-1}| \leq \|v_i\| \cdot \|p_{i-1}\| \leq 1$, we know that (24) is satisfied. Furthermore, we have $e_1^\top p_k \leq \|e_1\| \cdot \|p_k\| = \|e_1\| \cdot \sqrt{x^2 + y^2} \leq 1$ which ensures that (23) is satisfied. □

In the next lemma, we will show that $\pi_W(P)$ is contained in scaled the two-dimensional disk $(1+4^{-k-2})\mathbb{B}_2^2$:

Lemma 51 (Outer radius of the projected primal polytope)**.** *For every $x, y \in \mathbb{R}$ with $\sqrt{x^2 + y^2} \geq 1+4^{-k-2}$ there exist no $p_0 \in \mathbb{R}^2, t \in \mathbb{R}^k$ and $s \in \mathbb{R}$ such that $(x, y, p_0, t, s) \in P$.*

Proof. Fix any $(x, y) \in \mathbb{R}^2$ and $p_0 \in \mathbb{R}^2$, such that $x^2 + y^2 > 1 + 4^{-k-2}$ and $p_0 \succeq \left[|x|, |y|\right]^\top$. Also fix any $p_1, \cdots, p_k \in \mathbb{R}^2$ satisfying (21) and (22). We will show that such $p_1, \cdots, p_k$ would violate (23), i.e. $e_1^\top p_k > 1$. To simplify our notation, for all $i \in \{0, 1, \cdots, k\}$, let $(p_i)_v = v_i^\top p_i \in \mathbb{R}$ and $(p_i)_w = w_i^\top p_i \in \mathbb{R}$. Then $p_i = (p_i)_v v_i + (p_i)_w w_i$.

Notice that for all $i \in [k]$, the increment of the first coordinate from p_{i-1} to p_i is

$$\begin{aligned} e_1^\top p_i - e_1^\top p_{i-1} &= e_1^\top \left((p_i)_w w_i + (p_i)_v v_i - (w_i^\top p_{i-1}) w_i - (v_i^\top p_{i-1}) v_i\right) \\ &= e_1^\top \left((p_i)_v v_i - (v_i^\top p_{i-1}) v_i\right) && \text{(By (21))} \\ &\geq e_1^\top v_i \left(|v_i^\top p_{i-1}| - v_i^\top p_{i-1}\right) && \text{(By } e_1^\top v_i > 0 \text{ and (22))} \\ &\geq 0 \end{aligned} \tag{32}$$

where the inequality in (32) is tight when $v_i^\top p_i = |v_i^\top p_{i-1}|$. Let $p_0^*, p_1^*, \cdots, p_k^* \in \mathbb{R}^2$ be the (unique) sequence defined by

$$\begin{aligned} p_0^* &= p_0 \\ w_i^\top p_i^* &= w_i^\top p_{i-1}^*, && \forall i \in [k] && \text{(Tight for (21))} \\ v_i^\top p_i^* &= |v_i^\top p_{i-1}^*|, && \forall i \in [k] && \text{(Tight for (22))} \end{aligned}$$

Then $e_1^\top p_i - e_1^\top p_{i-1} \geq e_1^\top p_i^* - e_1^\top p_{i-1}^*$ for each $i \in [k]$. Also, notice that $e_1^\top p_0 = e_1^\top p_0^*$, therefore for each $i \in [k]$, $e_1^\top p_i \geq e_1^\top p_i^*$.

It remains to show that $e_1^\top p_k^* \geq 1$. For all $i \in \{0, 1, \cdots, k\}$, let $\theta_i \in [-\pi, \pi]$ denote the angle between $p_i^* \in \mathbb{R}^2$ and e_1. Then since $e_1^\top p_0 \geq 0$ and $e_2^\top p_0 \geq 0$, we have $0 \leq \theta_0 \leq \frac{\pi}{2}$. For any $i \in [k]$, notice that p_i^* equals to p_{i-1}^* (if $\theta_{i-1} \leq \frac{\pi}{2^{i+2}}$), or equals to the mirror of p_{i-1}^* with respect to the line spanned by w_i (if $\theta_{i-1} \geq \frac{\pi}{2^{i+2}}$). By induction, this gives

$$\|p_i^*\|_2 = \|p_{i-1}^*\|_2 = \cdots = \|p_0^*\|_2 = \|p_0\|_2,$$

and

$$\theta_i = \frac{\pi}{2^{i+2}} - |\theta_{i-1} - \frac{\pi}{2^{i+2}}| \leq \frac{\pi}{2^{i+2}}.$$

Therefore, we get

$$e_1^\top p_k^* = \|p_k^*\| \cdot \cos(\theta_k) \geq \|p_0\| \cdot \cos\left(\frac{\pi}{2^{k+2}}\right) \geq (1 + 4^{-k-2}) \cdot (1 - \frac{1}{2 \cdot 4^{k+2}}) > 1.$$

Thus we have shown $x^2 + y^2 > 1 + 4^{-k-2}$ implies that $e_1^\top p_k \geq e_1^\top p_k^* > 1$ as desired. □

6.4 Inner and Outer Radius of the Primal Polytope

In this subsection, we will show that the primal polytope P has large inner radius and small outer radius.

Lemma 52 (Inner radius of the primal polytope). *For $k \in \mathbb{N}$, let $P \subset \mathbb{R}^{k+5}$ be the polytope defined by the linear system (26). Then there exists a point $(\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s})$ such that*

$$\frac{1}{30} \cdot \mathbb{B}_\infty^{k+5} \subset P - (\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s}) \subset \frac{3}{2}\mathbb{B}_\infty^{k+5}.$$

Proof. In Lemma 53 we find a point $(\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s})$ such that $\frac{1}{30}\mathbb{B}_\infty^{k+5} \subset P - (\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s})$.

On the other hand, we claim that $P \subset 2 \cdot \mathbb{B}_\infty^{k+5}$. Suppose $(x, y, p_0, t, s) \in P$. From Lemma 51 we know that $\|(x, y)\|_\infty \leq \sqrt{x^2 + y^2} \leq \|p_0\|_2 \leq 1 + 4^{-k-2}$. Since $\mathbf{0}_k \leq t \leq \mathbf{1}_k$ we get $\|t\|_\infty \leq 1$, and lastly we have $0 \leq s \leq 1$. Put together, we find that $\|(x, y, p_0, t, s)\|_\infty \leq 1 + 4^{-k-2}$. By the triangle inequality we find that $P - (\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s}) \subset (1 + 4^{-k-2} + \|(\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s})\|_\infty) \cdot \mathbb{B}_\infty^{k+5}$ and we know $\|(\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s})\|_\infty = 1/3$. □

Lemma 53 (Inner Radius of the Polytope). *For $\bar{x} = \bar{y} = 0, \bar{p}_0 = (1/6, 1/6)^\top, \bar{t} = \mathbf{1}_k/30, \bar{s} = \frac{1}{3}$, we have $(\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s}) + r \cdot \mathbb{B}_\infty^{k+5} \subseteq P$ for $r = \frac{1}{30}$.*

Proof. Fix any $(x, y, p_0, t, s) \in \mathbb{R}^{k+5}$ such that $\|(x - \bar{x}, y - \bar{y}, p_0 - \bar{p}_0, t - \bar{t}, s - \bar{s})\|_\infty \leq r$. Let $p_1, \cdots, p_k \in \mathbb{R}^2$ be uniquely defined by $w_i^\top p_i = w_i^\top p_{i-1}$ and $t_i + is = v_i^\top p_i$ with our fixed p_0, s and t. We will show that $(x, y, p_0, t) \in P$ by verifying (20 - 25). To simplify our notation, we let $(p_i)_v = v_i^\top p_i \in \mathbb{R}$ and $(p_i)_w = w_i^\top p_i \in \mathbb{R}$ for all $i \in \{0, 1, \cdots, k\}$. Then $p_i = (p_i)_v v_i + (p_i)_w w_i$.

First, observe that $e_1^\top p_0 \geq \frac{1}{6} - r \geq r \geq |x|$ and $e_2^\top p_0 \geq \frac{1}{6} - r \geq r \geq |y|$, confirming that (20) holds. Also, notice that $t_i \in [\bar{t}_i - r, \bar{t}_i + r] \subset [0, 1]$, and $s \in [\bar{s} - r, \bar{s} + r] \subset [0, 1]$. thus (24) and (25) hold. The equality constraint (21) holds by definition of $p_1, \cdots, p_k$.

To aid in the remaining steps of the proof, we show $w_i^\top p_i \geq 0$ for all $i \in \{0, 1, \cdots, k\}$. Notice that $w_0^\top p_0 \geq w_0^\top \bar{p}_0 - \|w_0\| \cdot \|p_0 - \bar{p}_0\| \geq \frac{\sqrt{2}}{6} - \sqrt{2}r \geq 0$. Also, for all $i \in [k]$,

$$\begin{aligned}
w_i^\top p_i &= w_i^\top p_{i-1} && \text{(By (21))} \\
&= w_i^\top \left((p_{i-1})_w w_{i-1} + (p_{i-1})_v v_{i-1}\right) \\
&= (p_{i-1})_w \cdot \cos(\frac{\pi}{2^{i+2}}) + (p_{i-1})_v \cdot \sin(\frac{\pi}{2^{i+2}}) && (33) \\
&\geq (p_{i-1})_w \cdot \cos(\frac{\pi}{2^{i+2}}). && \text{(By } (p_{i-1})_v = t_{i-1} + (i-1)s \geq 0\text{)}
\end{aligned}$$

It then follows by induction that $(p_i)_w \geq 0$ for all $i \in \{0, 1, \cdots, k\}$.

Next, we verify the inequality listed in (22) i.e. $v_i^\top p_i \geq |v_i^\top p_{i-1}|$ for all $i \in [k]$. Notice that for all $i \in [k]$,

$$\begin{aligned}
|v_i^\top p_{i-1}| &= \left|v_i^\top ((p_{i-1})_v v_{i-1} + (p_{i-1})_w w_{i-1})\right| \\
&\leq |(p_{i-1})_v| v_i^\top v_{i-1} + |(p_{i-1})_w| \cdot |v_i^\top w_{i-1}| && \text{(Triangle inequality)} \\
&= (t_{i-1} + (i-1)s) \cdot \cos(\frac{\pi}{2^{i+2}}) + (p_{i-1})_w \cdot \sin(\frac{\pi}{2^{i+2}}) && \text{(By (22) and } (p_{i-1})_w \geq 0\text{)} \quad (34)
\end{aligned}$$

We require an upper bound on $(p_{i-1})_w$. For all $i \in [k]$, from (33)

$$\begin{aligned}
(p_i)_w &= (p_{i-1})_w \cdot \cos(\frac{\pi}{2^{i+2}}) + (p_{i-1})_v \cdot \sin(\frac{\pi}{2^{i+2}}) \\
&\leq (p_{i-1})_w + (t_{i-1} + (i-1)s) \cdot \sin(\frac{\pi}{2^{i+2}}) && \text{(By } (p_{i-1})_w \geq 0 \text{ and (22))}
\end{aligned}$$

Let $t_0 = v_0^\top p_0$ and $\bar{t}_0 = v_0^\top \bar{p}_0 = 0$. We have that for all $i \in [k]$,

$$\begin{aligned}
(p_i)_w &\leq w_0^\top p_0 + \sum_{j=0}^{i-1} (t_j + js) \cdot \sin(\frac{\pi}{2^{j+3}}) \\
&\leq w_0^\top p_0 + \sum_{j=0}^{i-1} (t_j + js) \cdot \sin(\frac{\pi}{8}) \cdot 1.9^{-j} && \text{(By } \sin(x)/\sin(x/2) \geq 1.9 \text{ for all } x \leq \tfrac{\pi}{8}\text{)}
\end{aligned}$$

$$
\begin{aligned}
&\leq w_0^\top p_0 + \sum_{j=0}^{i-1} (\bar{t}_j + r + js) \cdot \sin(\frac{\pi}{8}) \cdot 1.9^{-j} \\
&\leq (\frac{\sqrt{2}}{6} + \sqrt{2}r) + \sin(\frac{\pi}{8}) \sum_{j=0}^{i-1} (\frac{1}{30} + r + js) \cdot 1.9^{-j} && (\text{By } w_0^\top \bar{p}_0 = \tfrac{\sqrt{2}}{6}, \bar{t}_j = \tfrac{1}{30}) \\
&\leq 0.263 + 2.226r + 0.898s. && (35)
\end{aligned}
$$

Plugging (35) back into (34), we have for all $i \in [k]$,

$$
\begin{aligned}
|v_i^\top p_{i-1}| &\leq (t_{i-1} + (i-1)s) \cdot \cos(\frac{\pi}{2^{i+2}}) + (0.263 + 2.226r + 0.898s) \cdot \sin(\frac{\pi}{2^{i+2}}) \\
&\leq (\frac{1}{30} + r + (i-1)s) + (0.263 + 2.226r + 0.898s) \cdot \sin(\frac{\pi}{8}) && (\text{By } \bar{t}_{i-1} = \tfrac{1}{30} \text{ and } i \geq 1) \\
&\leq 0.134 + 1.852r + (i - 0.656)s \\
&\leq 0.134 + 1.852r + is - 0.656(\bar{s} - r) \\
&\leq is && (\text{By } \bar{s} = \tfrac{1}{3}, r \leq \tfrac{1}{30}) \\
&\leq (p_i)_v.
\end{aligned}
$$

Therefore, $v_i^\top p_i \geq |v_i^\top p_{i-1}|$ for all $i \in [k]$ and (22) holds.

Finally, we verify (23). The increment of the first coordinate from p_{i-1} to p_i is

$$
\begin{aligned}
e_1^\top p_i - e_1^\top p_{i-1} &= e_1^\top \left((p_i)_v v_i - (v_i^\top p_{i-1}) v_i\right) && (\text{By } (21)) \\
&= \sin(\frac{\pi}{2^{i+2}}) \cdot (t_i + is - v_i^\top p_{i-1}) && (\text{By } e_1^\top v_i = \sin(\tfrac{\pi}{2^{i+2}}) \text{ and } (22)) \\
&= \sin(\frac{\pi}{2^{i+2}}) \cdot \left(t_i + is - v_i^\top ((p_{i-1})_w w_{i-1} + (p_{i-1})_v v_{i-1})\right) \\
&= \sin(\frac{\pi}{2^{i+2}}) \cdot \Big(t_i + is + (p_{i-1})_w \cdot \sin(\frac{\pi}{2^{i+2}}) - (t_{i-1} + (i-1)s) \cdot \cos(\frac{\pi}{2^{i+2}})\Big) && (36)
\end{aligned}
$$

where the last step comes from $v_i^\top v_{i-1} = \cos(\frac{\pi}{2^{i+2}})$ and $v_i^\top w_{i-1} = -\sin(\frac{\pi}{2^{i+2}})$. For all $i \geq 2$, we can show that in (36), the third term in the brackets is at most the fourth term:

$$
\begin{aligned}
(p_{i-1})_w \cdot \sin(\frac{\pi}{2^{i+2}}) &\leq (0.263 + 2.226r + 0.898s) \cdot \sin(\frac{\pi}{2^{i+2}}) && (\text{By } (35)) \\
&\leq (0.263 + 2.226r + 0.898s) \cdot \tan(\frac{\pi}{8}) \cdot \cos(\frac{\pi}{2^{i+2}}) && (\text{By } \tfrac{\pi}{2^{i+2}} \leq \tfrac{\pi}{8}) \\
&\leq (0.338 + 0.898 \cdot \frac{11}{30}) \cdot \tan(\frac{\pi}{8}) \cdot \cos(\frac{\pi}{2^{i+2}}) && (\text{By } r \leq \tfrac{1}{30} \text{ and } s \leq \bar{s} + r = \tfrac{11}{30}) \\
&\leq 0.277 \cdot \cos(\frac{\pi}{2^{i+2}}) \\
&\leq (\bar{t}_{i-1} - r + (i-1)s) \cdot \cos(\frac{\pi}{2^{i+2}}) && (\text{By } \bar{t}_{i-1} - r = 0 \text{ and } s \geq \bar{s} - r = 0.3) \\
&\leq (t_{i-1} + (i-1)s) \cdot \cos(\frac{\pi}{2^{i+2}}).
\end{aligned}
$$

Plugging back into (36), we have

$$
\begin{aligned}
e_1^\top p_i - e_1^\top p_{i-1} &\leq \sin(\frac{\pi}{2^{i+2}}) \cdot (t_i + is) \\
&\leq \sin(\frac{\pi}{2^{i+2}}) \cdot (\bar{t}_i + i\bar{s} + (i+1)r) \\
&\leq \sin(\frac{\pi}{8}) \cdot 1.9^{-(i-1)} \cdot (\frac{1}{30} + \frac{i}{3} + (i+1)r) && (\text{By } \sin(\tfrac{\pi}{2^{i+2}}) \leq \sin(\tfrac{\pi}{8}) \cdot 1.9^{-(i-1)})
\end{aligned}
$$

Therefore,

$$
e_1^\top p_k \leq e_1^\top p_0 + \sum_{i=1}^{k} (e_1^\top p_i - e_1^\top p_{i-1})
$$

$$\leq (\frac{1}{6} + r) + \sin(\frac{\pi}{8}) \cdot \sum_{i=1}^{k} 1.9^{-(i-1)} \cdot (\frac{1}{30} + \frac{i}{3} + (i+1)r)$$
$$\leq 0.562 + 3.514r \leq 1.$$

where the last inequality holds for any $r \leq \frac{1}{30}$. Therefore (23) holds, and $(x, y, p_0, t) \in P$. $\square$

6.5 Properties of the Dual Polytope

In this section, we will analyze the scaled polar dual polytope $Q = \frac{1}{30}(P - (\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s}))^{\circ}$. From well-known duality properties, we will find that Q satisfies the following desirable properties:

1. $Q \cap W$ approximates a two-dimensional disk;
2. The inner radius of Q is at least $\frac{1}{45}$ when centered at $\mathbf{0}$;
3. The outer radius of Q is at most 1 when centered at $\mathbf{0}$.

Lemma 54. *For any $k \in \mathbb{N}$, there exists a two-dimensional linear subspace $W \subset \mathbb{R}^{k+5}$ and $n = 4k+5$ points $a_1, \cdots, a_n \in \mathbb{B}_1^{k+5}$ such that $Q := \mathrm{conv}(a_1, \cdots, a_n)$ satisfies*

$$\frac{1}{30(1+4^{-k})} \cdot \mathbb{B}_2^{k+5} \cap W \subset Q \cap W \subset \frac{1}{30} \cdot \mathbb{B}_2^{k+5} \cap W$$

and

$$\frac{1}{45} \cdot \mathbb{B}_1^{k+5} \subset Q \subset \mathbb{B}_1^{k+5}$$

Proof. Let $P \subset \mathbb{R}^{k+5}$ be the polytope defined by the linear system in(26), and let

$$\tilde{P} = P - (\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s})$$

denote the polytope obtained from shifting its center $(\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s})$ to $\mathbf{0}_d$. Here

$$\bar{x} = \bar{y} = 0, \bar{p}_0 = (1/6, 1/6)^{\top}, \bar{t} = \mathbf{1}_k/30, \bar{s} = \frac{1}{3}.$$

Applying basic operations from linear algebra, we transform the constraints (20 - 25) into a matrix $A \in \mathbb{R}^{(4k+5)\times(k+5)}$ such that

$$\tilde{P} = \{z \in \mathbb{R}^{k+5} : Az \leq 1\}.$$

Let $\tilde{Q} = (\tilde{P})^{\circ} \subset \mathbb{R}^d$ denote the polar body of $\tilde{P}$. Since $\tilde{P}$ is bounded, $\tilde{Q}$ is the convex hull of the rows of the matrix A, i.e.

$$\tilde{Q} = \{A^{\top}\lambda : \lambda \in [0,1]^{4k+5} \text{ s.t. } \sum_{i=1}^{4k+5} \lambda_i = 1\}.$$

Then by Lemma 52 and Fact 9, the inner and outer ball of $\tilde{Q}$ satisfy

$$\frac{2}{3} \cdot \mathbb{B}_1^{k+5} \subset \tilde{Q} \subset 30 \cdot \mathbb{B}_1^{k+5}.$$

Also, by Lemma 49, Fact 10 and Fact 9, the inner and outer ball of $\tilde{Q} \cap W$ satisfy

$$\frac{1}{1+4^{-k-2}} \cdot \mathbb{B}_2^{k+5} \cap W \subset \tilde{Q} \cap W \subset \mathbb{B}_2^{k+5} \cap W.$$

The lemma then follows from taking $Q = \frac{1}{30}\tilde{Q}$. $\square$

6.6 Perturbation Analysis and Proof of the Lower Bound

In this subsection, we study the number of edges of the intersection polygon $Q \cap W$ after perturbation and prove our main theorem (Theorem 47). To show that our construction has many edges even after perturbation, we require the following two statements:

Lemma 55. *Let $\bar{a}_1, \ldots, \bar{a}_n \in \mathbb{R}^d$ be any points with $r\mathbb{B}_1^d \subset \operatorname{conv}(\bar{a}_1, \ldots, \bar{a}_n)$ for some $r > 0$. If $\varepsilon \leq r/2$ and points $a_1, \ldots, a_n \in \mathbb{R}^d$ satisfy $\|a_i - \bar{a}_i\|_1 \leq \varepsilon$ for all $i \in [n]$ then it follows that*

$$(1 - \frac{2\varepsilon}{r}) \operatorname{conv}(\bar{a}_1, \ldots, \bar{a}_n) \subset \operatorname{conv}(a_1, \ldots, a_n) \subset (1 + \frac{\varepsilon}{r}) \operatorname{conv}(\bar{a}_1, \ldots, \bar{a}_n).$$

Proof. Write $Q = \operatorname{conv}(a_1, \ldots, a_n)$ and $\bar{Q} = \operatorname{conv}(\bar{a}_1, \ldots, \bar{a}_n)$. The second inclusion follows by $Q \subset \bar{Q} + \varepsilon\mathbb{B}_1^d \subset \bar{Q} + \frac{\varepsilon}{r}\bar{Q}$. For the first inequality, we observe that $r\mathbb{B}_1^d \subset \bar{Q} \subset Q + \varepsilon\mathbb{B}_1^d \subset Q + \frac{r}{2}\mathbb{B}_1^d$. This implies that $\frac{r}{2}\mathbb{B}_1^d \subset Q$, for if there were to exist $x \in \frac{r}{2}\mathbb{B}_1^d$ such that $x \notin Q$ then, since Q is closed and convex, we could find $y \in \mathbb{R}^d$ such that $y^\top x > y^\top z$ for all $z \in Q$. Writing $f(S) = \max_{z \in S} y^\top z$ for $S \subset \mathbb{R}^d$, this would give

$$f(r\mathbb{B}_1^d) \geq f(x + \frac{r}{2}\mathbb{B}_1^d) = y^\top x + f(\frac{r}{2}\mathbb{B}_1^d) > f(Q) + f(\frac{r}{2}\mathbb{B}_1^d) \geq f(Q + \frac{r}{2}\mathbb{B}_1^d) \geq f(r\mathbb{B}_1^d).$$

By contradiction it follows that $\frac{r}{2}\mathbb{B}_1^d \subseteq Q$.

Now the desired result follows by $\bar{Q} \subset Q + \varepsilon\mathbb{B}_1^d \subset Q + \frac{\varepsilon}{r/2}Q$ and the fact that $(1+x)^{-1} > 1 - x$ for every $x > -1$. □

Lemma 56. *If a polygon $T \subset \mathbb{R}^2$ satisfies $\alpha \cdot \mathbb{B}_2^2 \subset T \subset \beta \cdot \mathbb{B}_2^2$ for some $\alpha, \beta > 0$ then T has at least $\Omega(\sqrt{\alpha/(\beta - \alpha)})$ edges.*

Proof. Without loss of generality, re-scale T so that $\mathbb{B}_2^2 \subset T \subset (1+\varepsilon) \cdot \mathbb{B}_2^2$, where $\varepsilon = \beta/\alpha - 1 > 0$.

Consider any edge $[q_1, q_2] \subset T$ and let $p \in [q_1, q_2]$ denote the minimum-norm point in this edge. Then we have $\|q_1 - p\|^2 = \|q_1\|^2 + \|p\|^2 - 2\langle q_1, p\rangle$. Since p is the minimum-norm point, we have $\langle p, q_1\rangle \geq \|p\|^2$, and hence $\|q_1 - p\|^2 \leq \|q_1\|^2 - \|p\|^2 \leq (1+\varepsilon)^2 - \|p\|^2$. Since p lies on the boundary of T we have $\|p\| \geq 1$, which implies that $\|q_1 - p\|^2 \leq (1+\varepsilon)^2 - 1 = 2\varepsilon + \varepsilon^2$. The analogous argument for $\|q_2 - p\|$ and the triangle inequality tell us that $\|q_1 - q_2\| \leq 2\sqrt{2\varepsilon + \varepsilon^2} \leq 4\sqrt{\varepsilon}$. The choice of the edge $[q_1, q_2]$ was arbitrary, hence every edge of T has length as most $2\sqrt{2\varepsilon + \varepsilon^2}$.

But T has perimeter at least 2π. Since the perimeter of a polygon is equal to the sum of the lengths of its edges, this implies that T has at least $\frac{2\pi}{4\sqrt{\varepsilon}}$ edges. □

Now, we can prove our generic lower bound Theorem 48 on the shadow size under adversarial ℓ_1-perturbations.

Proof of Theorem 48. Fix any $d \geq 5$, let $k = d - 5$ and observe that $n = 4k + 5$. Let $\bar{a}_1, \cdots, \bar{a}_n$ be as constructed in Lemma 54. Then we have

$$\frac{1}{30(1 + 4^{-k})} \cdot \mathbb{B}_2^{k+5} \cap W \subset \operatorname{conv}(\bar{a}_1, \cdots, \bar{a}_n) \cap W \subset \frac{1}{30} \cdot \mathbb{B}_2^{k+5} \cap W$$

and

$$\frac{1}{45} \cdot \mathbb{B}_1^{k+5} \subset \operatorname{conv}(\bar{a}_1, \cdots, \bar{a}_n) \subset \mathbb{B}_1^{k+5}$$

For any set of points $a_1, \cdots, a_n$ such that $\|a_i - \bar{a}_i\|_1 \leq \varepsilon$ for each $i \in [n]$, by Lemma 55 we have

$$\frac{1}{30(1 + 4^{-k})} \cdot \mathbb{B}_2^{k+5} \cap W \subset \operatorname{conv}(\bar{a}_1, \cdots, \bar{a}_n) \cap W \subset (1 + 90\varepsilon) \operatorname{conv}(a_1, \cdots, a_n) \cap W$$

and

$$(1 - 90\varepsilon) \operatorname{conv}(a_1, \cdots, a_n) \cap W \subset \operatorname{conv}(\bar{a}_1, \cdots, \bar{a}_n) \cap W \subset \frac{1}{30} \cdot \mathbb{B}_2^{k+5} \cap W.$$

Therefore, we can bound the inner and outer radius of $\mathrm{conv}(a_1, \cdots, a_n) \cap W$ by

$$\frac{1}{30 \cdot (1 + 4^{-k}) \cdot (1 + 90\varepsilon)} \cdot \mathbb{B}_2^2 \subseteq \mathrm{conv}(a_1, \cdots, a_n) \cap W \subseteq \cdot \frac{1}{30 \cdot (1 - 90\varepsilon)} \cdot \mathbb{B}_2^2$$

It follows from Lemma 56 that the polygon $\mathrm{conv}(a_1, \cdots, a_n) \cap W$ has at least $\Omega(\frac{1}{\sqrt{\varepsilon + 4^{-k}}})$ edges. □

Finally, we can prove our main result using Gaussian tail bound:

Proof of Theorem 47. Using concentration of Gaussian distribution in Corollary 13, we find that if $\sigma \leq \frac{1}{360d\sqrt{\log n}}$, then with probability at least $1 - \binom{n}{d}^{-1}$, we have $\max_{i \in [n]} \|a_i - \bar{a}_i\|_2 \leq 4\sigma\sqrt{d \log n} \leq \frac{1}{90\sqrt{d}}$. The result follows from Theorem 48 and the fact that $\|x\|_1 \leq \sqrt{d}\|x\|_2$ for every $x \in \mathbb{R}^d$. □

6.7 Experimental Results

To measure whether analysis in Theorem 47 is tight or not, we ran numerical experiments. Using `Python` and `Gurobi 9.5.2`, we constructed a matrix A such that

$$P - (\bar{x}, \bar{y}, \bar{p}_0, \bar{t}, \bar{s}) = \{x \in \mathbb{R}^{k+5} : Ax \leq \mathbf{1}_{4k+5}\},$$

as described earlier in this section. Writing R as the maximum Euclidean norm among the row vectors of A, we sampled $\hat{A}$ with independent Gaussian distributed entries with standard deviation σR and $\mathbb{E}[\hat{A}] = A$. To approximate the shadow size, we optimized the objective vectors $\cos(\frac{(i+0.3)\pi}{2^{k+4}})x + \sin(\frac{(i+0.3)\pi}{2^{k+4}})y$, with $i = 0, \ldots, 2^{k+5} - 1$, over the polyhedron $\{x \in \mathbb{R}^{k+5} : \hat{A}x \leq 1\}$ and counted the number of distinct values (x, y) found among the solutions. When $\sigma = 0$, our code found 2^{k+1} such points. For $\sigma > 0$, Theorem 47 shows that we expect to find at least $\Omega\Big(\min\Big(\frac{1}{\sqrt{d\sigma\sqrt{\log d}}}, 2^k\Big)\Big)$ distinct pairs (x, y).

For $k = 10, 15, 20$, we measured the shadow size for 20 different values of σ ranging from 0.01 to $0.0001/2^k$, The resulting data is depicted in Figure 5 along with a graph of the function $\sigma \mapsto \sigma^{-3/4}$. We observe that for each k, the measured shadow size appears to follow the graphed function up to a point, plateauing slightly above 2^{k+1} when σ is small. The fact that some measurements come out higher than 2^{k+1}, the shadow size for $\sigma = 0$, is not unexpected: the polytope P is highly degenerate, whereas the perturbed polytope is simple and can thus have many more vertices.

The measured shadow sizes appear to grow much faster than $1/\sqrt{\sigma}$ as σ gets small. These results suggest that the behaviour of the shadow size is substantially different in $d = 2$, where we have an upper bound of $O\left(\frac{\sqrt[4]{\log(n)}}{\sqrt{\sigma}} + \sqrt{\log n}\right)$, and $d > 3$, where one might expect a lower bound with a higher dependence on σ.

References

[AC78] David Avis and Vasek Chvátal. "Notes on Bland's pivoting rule". In: *Polyhedral Combinatorics*. Springer, 1978, pp. 24–34.

[ADLNV22] Xavier Allamigeon, Daniel Dadush, Georg Loho, Bento Natura, and László A. Végh. "Interior point methods are not worse than Simplex". In: *Proceedings of the 63th Annual IEEE Symposium on Foundations of Computer Science*. arXiv, 2022. DOI: 10.48550/ARXIV.2206.08810.

[AKS87] Ilan Adler, Richard M. Karp, and Ron Shamir. "A simplex variant solving an $m \times d$ linear program in $O(\min(m^2, d^2))$ expected number of pivot steps". In: *J. Complexity* 3.4 (1987), pp. 372–387. ISSN: 0885-064X. DOI: 10.1016/0885-064X(87)90007-0. URL: http://dx.doi.org/10.1016/0885-064X(87)90007-0.

[AM85] Ilan Adler and Nimrod Megiddo. "A simplex algorithm whose average number of steps is bounded between two quadratic functions of the smaller dimension". In: *Journal of the ACM (JACM)* 32.4 (1985), pp. 871–895.

[And04] Neculai Andrei. "On the complexity of MINOS package for linear programming". In: *Studies in Informatics and Control* 13.1 (2004), pp. 35–46.

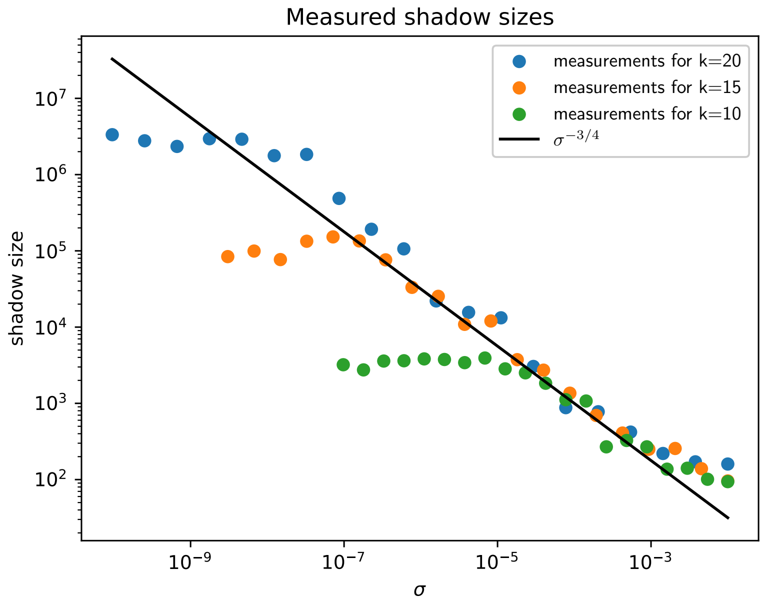

Figure 5: Measured shadow sizes for sampled perturbations of our construction, for different values of k and σ.

[AZ98] Nina Amenta and Günter M. Ziegler. "Deformed Products and Maximal Shadows". In: *Contemporary Math.* 223 (1998), pp. 57–90.

[BDGHL22] Gilles Bonnet, Daniel Dadush, Uri Grupel, Sophie Huiberts, and Galyna Livshyts. "Asymptotic Bounds on the Combinatorial Diameter of Random Polytopes". In: *38th International Symposium on Computational Geometry, SoCG 2022, June 7-10, 2022, Berlin, Germany*. Ed. by Xavier Goaoc and Michael Kerber. Vol. 224. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022, 18:1–18:15. DOI: 10.4230/LIPIcs.SoCG.2022.18. URL: https://doi.org/10.4230/LIPIcs.SoCG.2022.18.

[BDKS21] Alexander Black, Jesús De Loera, Sean Kafer, and Laura Sanità. *On the Simplex method for 0/1 polytopes*. 2021. DOI: 10.48550/ARXIV.2111.14050.

[BHMPT15] Andreas Bärmann, Andreas Heidt, Alexander Martin, Sebastian Pokutta, and Christoph Thurner. "Polyhedral approximation of ellipsoidal uncertainty sets via extended formulations: a computational case study". In: *Computational Management Science* 13.2 (Oct. 2015), pp. 151–193. DOI: 10.1007/s10287-015-0243-0. URL: https://doi.org/10.1007/s10287-015-0243-0.

[Bla22] Alexander E. Black. *Small Shadows of Lattice Polytopes*. 2022. DOI: 10.48550/ARXIV.2204.09129.

[BN01] Aharon Ben-Tal and Arkadi Nemirovski. "On polyhedral approximations of the second-order cone". In: *Mathematics of Operations Research* 26.2 (2001), pp. 193–205.

[Bor82] Karl-Heinz Borgwardt. "The Average number of pivot steps required by the Simplex-Method is polynomial". In: *Z. Oper. Research* 26.1 (1982), pp. 157–177.

[Bor87] Karl Heinz Borgwardt. *The Simplex Method: A Probabilistic Analysis*. Vol. 1. Springer-Verlag, Berlin, 1987.

[Bor99] Karl Heinz Borgwardt. "A Sharp Upper Bound for the Expected Number of Shadow Vertices in Lp-Polyhedra Under Orthogonal Projection on Two-Dimensional Planes". In: *Mathematics of Operations Research* 24.3 (1999), pp. 544–603.

[BR13] Tobias Brunsch and Heiko Röglin. "Finding Short Paths on Polytopes by the Shadow Vertex Algorithm". In: *Automata, Languages, and Programming*. ICALP '13. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 279–290. ISBN: 978-3-642-39206-1. DOI: 10.1007/978-3-642-39206-1_24.

[Dan47] George B Dantzig. "Maximization of a linear function of variables subject to linear inequalities". In: *Activity analysis of production and allocation* 13 (1947), pp. 339–347.

[DFH22] Yann Disser, Oliver Friedmann, and Alexander V. Hopp. "An exponential lower bound for Zadeh's pivot rule". In: *Mathematical Programming* (July 2022). DOI: 10.1007/s10107-022-01848-x.

[DGGT16] Olivier Devillers, Marc Glisse, Xavier Goaoc, and Rémy Thomasse. "Smoothed complexity of convex hulls by witnesses and collectors". In: *Journal of Computational Geometry* 7.2 (2016), pp. 101–144. DOI: 10.20382/jocg.v7i2a6.

[DH16] Daniel Dadush and Nicolai Hähnle. "On the shadow simplex method for curved polyhedra". In: *Discrete Comput. Geom.* 56.4 (2016), pp. 882–909. ISSN: 0179-5376. DOI: 10.1007/s00454-016-9793-3. URL: http://dx.doi.org/10.1007/s00454-016-9793-3.

[DH18] Daniel Dadush and Sophie Huiberts. "A friendly smoothed analysis of the simplex method". In: *SIAM Journal on Computing* 49.5 (2018), STOC18–449.

[DH21] Daniel Dadush and Sophie Huiberts. "Smoothed Analysis of the Simplex Method". In: *Beyond the Worst-Case Analysis of Algorithms*. Ed. by TimEditor Roughgarden. Cambridge University Press, 2021, pp. 309–333. DOI: 10.1017/9781108637435.019.

[DS04] Valentina Damerow and Christian Sohler. "Extreme points under random noise". In: *European Symposium on Algorithms*. Springer. 2004, pp. 264–274. DOI: 10.1007/978-3-540-30140-0_25.

[DS05] Amit Deshpande and Daniel A Spielman. "Improved smoothed analysis of the shadow vertex simplex method". In: *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*. 2005, pp. 349–356.

[FHZ11] Oliver Friedmann, Thomas Dueholm Hansen, and Uri Zwick. "Subexponential Lower Bounds for Randomized Pivoting Rules for the Simplex Algorithm". In: *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing*. 2011, pp. 283–292. ISBN: 978-1-4503-0691-1. DOI: 10.1145/1993636.1993675.

[Fri11] Oliver Friedmann. "A Subexponential Lower Bound for Zadeh's Pivoting Rule for Solving Linear Programs and Games". In: *Proceedings of 15th International Conference on Integer Programming and Combinatorial Optimization*. Ed. by Oktay Günlük and Gerhard J. Woeginger. Springer, 2011, pp. 192–206. DOI: 10.1007/978-3-642-20807-2_16.

[Gam+20a] Gerald Gamrath, Daniel Anderson, Ksenia Bestuzheva, Wei-Kun Chen, Leon Eifler, Maxime Gasse, Patrick Gemander, Ambros Gleixner, Leona Gottwald, Katrin Halbig, Gregor Hendel, Christopher Hojny, Thorsten Koch, Pierre Le Bodic, Stephen J. Maher, Frederic Matter, Matthias Miltenberger, Erik Mühmer, Benjamin Müller, Marc E. Pfetsch, Franziska Schlösser, Felipe Serrano, Yuji Shinano, Christine Tawfik, Stefan Vigerske, Fabian Wegscheider, Dieter Weninger, and Jakob Witzig. *The SCIP Optimization Suite 7.0*. Technical Report. Optimization Online, Mar. 2020. URL: http://www.optimization-online.org/DB_HTML/2020/03/7705.html.

[Gam+20b] Gerald Gamrath, Daniel Anderson, Ksenia Bestuzheva, Wei-Kun Chen, Leon Eifler, Maxime Gasse, Patrick Gemander, Ambros Gleixner, Leona Gottwald, Katrin Halbig, Gregor Hendel, Christopher Hojny, Thorsten Koch, Pierre Le Bodic, Stephen J. Maher, Frederic Matter, Matthias Miltenberger, Erik Mühmer, Benjamin Müller, Marc E. Pfetsch, Franziska Schlösser, Felipe Serrano, Yuji Shinano, Christine Tawfik, Stefan Vigerske, Fabian Wegscheider, Dieter Weninger, and Jakob Witzig. *The SCIP Optimization Suite 7.0*. ZIB-Report 20-10. Zuse Institute Berlin, Mar. 2020. URL: http://nbn-resolving.de/urn:nbn:de:0297-zib-78023.

[Gli00] François Glineur. *Computational Experiments With a Linear Approximation of Second-Order Cone Optimization*. Tech. rep. 2000. URL: https://web.archive.org/web/20220505114903/https://perso.uclouvain.be/francois.glineur/oldwww/Papers/Image0001.ps.gz.

[Gol76] Donald Goldfarb. "Using the steepest-edge simplex algorithm to solve sparse linear programs". In: *Sparse matrix computations*. Elsevier, 1976, pp. 227–240.

[Gol94] Donald Goldfarb. "On the Complexity of the Simplex Method". In: *Advances in Optimization and Numerical Analysis*. Springer, 1994, pp. 25–38. ISBN: 978-94-015-8330-5. DOI: 10.1007/978-94-015-8330-5_2.

[GS55] Saul Gass and Thomas Saaty. "The computational algorithm for the parametric objective function". In: *Naval research logistics quarterly* 2.1-2 (1955), pp. 39–45.

[GS79] Donald Goldfarb and William Y Sit. "Worst case behavior of the steepest edge simplex method". In: *Discrete Applied Mathematics* 1.4 (1979), pp. 277–285.

[Hai83] M. Haimovich. *The simplex method is very good! On the expected number of pivot steps and related properties of random linear programs*. Tech. rep. Columbia University, 1983.

[HZ15] Thomas Dueholm Hansen and Uri Zwick. "An improved version of the random-facet pivoting rule for the simplex algorithm". In: *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*. 2015, pp. 209–218.

[Jer73] Robert G Jeroslow. "The simplex algorithm with the pivot rule of maximizing criterion improvement". In: *Discrete Mathematics* 4.4 (1973), pp. 367–377.

[Kal92] Gil Kalai. "A subexponential randomized simplex algorithm". In: *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*. 1992, pp. 475–482.

[KM72] Victor Klee and George J Minty. "How good is the simplex algorithm". In: *Inequalities* 3.3 (1972), pp. 159–175.

[KP13] Volker Kaibel and Kanstantsin Pashkovich. "Constructing Extended Formulations from Reflection Relations". In: *Facets of Combinatorial Optimization: Festschrift for Martin Grötschel*. Ed. by Michael Jünger and Gerhard Reinelt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 77–100. ISBN: 978-3-642-38189-8. DOI: 10.1007/978-3-642-38189-8_4. URL: https://doi.org/10.1007/978-3-642-38189-8_4.

[KS06] Jonathan A. Kelner and Daniel A. Spielman. "A randomized polynomial-time simplex algorithm for linear programming". In: *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*. 2006, pp. 51–60. DOI: 10.1145/1132516.1132524.

[Meg86] Nimrod Megiddo. "Improved asymptotic analysis of the average number of steps performed by the self-dual simplex algorithm". In: *Math. Programming* 35.2 (1986), pp. 140–172. ISSN: 0025-5610. URL: https://doi.org/10.1007/BF01580645.

[MSW96] Jiří Matoušek, Micha Sharir, and Emo Welzl. "A subexponential bound for linear programming". In: *Algorithmica* 16.4-5 (1996), pp. 498–516.

[Mur80] Katta G Murty. "Computational complexity of parametric linear programming". In: *Mathematical programming* 19.1 (1980), pp. 213–219.

[NSS22] Hariharan Narayanan, Rikhav Shah, and Nikhil Srivastava. "A Spectral Approach to Polytope Diameter". In: *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*. Vol. 215. Leibniz International Proceedings in Informatics (LIPIcs). 2022, 108:1–108:22. DOI: 10.4230/LIPIcs.ITCS.2022.108.

[Sch14] Emanuel Schnalzger. "Lineare Optimierung mit dem Schatteneckenalgorithmus im Kontext probabilistischer Analysen". English translation by K.H. Borgwardt. PhD thesis. Universität Augsburg, 2014.

[Sha87] Ron Shamir. "The efficiency of the simplex method: a survey". In: *Management science* 33.3 (1987), pp. 301–334. DOI: 10.1287/mnsc.33.3.301.

[Sma83] Steve Smale. "On the average number of steps of the simplex method of linear programming". In: *Mathematical programming* 27.3 (1983), pp. 241–262.

[ST04] Daniel A Spielman and Shang-Hua Teng. "Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time". In: *Journal of the ACM (JACM)* 51.3 (2004), pp. 385–463.

[SW08] Rolf Schneider and Wolfgang Weil. *Stochastic and Integral Geometry*. Springer Berlin Heidelberg, 2008. DOI: 10.1007/978-3-540-78859-1. URL: https://doi.org/10.1007/978-3-540-78859-1.

[Tod86] Michael J. Todd. "Polynomial expected behavior of a pivoting algorithm for linear complementarity and linear programming problems". In: *Mathematical Programming* 35.2 (1986), pp. 173–192.

[Ver09] Roman Vershynin. "Beyond Hirsch conjecture: walks on random polytopes and smoothed complexity of the simplex method". In: *SIAM Journal on Computing* 39.2 (2009), pp. 646–678.

[VGG17] Arnaud Vandaele, Nicolas Gillis, and François Glineur. "On the linear extension complexity of regular n-gons". In: *Linear Algebra and its Applications* 521 (2017), pp. 217–239. ISSN: 0024-3795. DOI: https://doi.org/10.1016/j.laa.2016.12.023. URL: https://www.sciencedirect.com/science/article/pii/S0024379516306127.