Smoothed complexity of convex hulls by witnesses and collectors

Olivier Devillers, Marc Glisse, Xavier Goaoc, Rémy Thomasse

Smoothed complexity of convex hulls by witnesses and collectors

Olivier Devillers, Marc Glisse, Xavier Goaoc, Rémy Thomasse

RESEARCH REPORT
N° 8787
October 2015
Project-Team Geometrica & Vegas

ISSN 0249-6399 ISRN INRIA/RR--8787--FR+ENG

Smoothed complexity of convex hulls by witnesses and collectors

Olivier Devillers[*†‡], Marc Glisse[§], Xavier Goaoc[¶], Rémy Thomasse[‖]

Project-Team Geometrica & Vegas

Research Report n° 8787 — October 2015 — 41 pages

Abstract: We present a simple technique for analyzing the size of geometric hypergraphs defined by random point sets. As an application we obtain upper and lower bounds on the smoothed number of faces of the convex hull under Euclidean and Gaussian noise and related results.

Key-words: Probabilistic analysis – Worst-case analysis – Gaussian noise

Part of this work is supported by: ANR blanc PRESAGE (ANR-11-BS02-003), ERC project GUDHI, Région PACA and Institut Universitaire de France.

[*] Inria, Centre de recherche Nancy - Grand Est, France.
[†] CNRS, Loria, France.
[‡] Université de Lorraine, France
[§] Inria, Centre de recherche Saclay - Île-de-France, France.
[¶] LIGM, Université Paris-Est Marne la Vallée, France.
[‖] Inria, Centre de recherche Sophia Antipolis - Méditerranée, France.

RESEARCH CENTRE NANCY – GRAND EST

615 rue du Jardin Botanique
CS20101
54603 Villers-lès-Nancy Cedex

Analyse lissée des enveloppes convexes par témoins et collecteurs

Résumé : Nous présentons une méthode simple pour l'analyse de la taille d'hypergraphes géométriques définis par des ensembles de points aléatoires. En appliquant cette technique nous obtenons des bornes inférieures et supérieures pour l'analyse lissée de du nombre de faces de l'enveloppe convexe de points soumis à un bruit euclidien ou gaussien.

Mots-clés : Analyse probabiliste – Analyse dans le cas le pire – bruit gaussien

1 Introduction

Let P^* be a finite set of points in $\mathbb{R}^d$ and consider a random perturbation $P = \{p^* + \eta(p^*) : p^* \in P^*\}$ where each point p^* is moved by some random vector $\eta(p^*)$, typically chosen independently. We are interested in the asymptotic behaviour of the expected number of faces (of all dimensions) of the convex hull of P, as a function of the number n of points and some parameter that describes the amplitude of the perturbations.

Formally, the *smoothed complexity of convex hulls* relative to a probability distribution μ on $\mathbb{R}^d$ is defined as

$$\mathcal{S}(n,\mu) = \max_{\substack{p_1^*, p_2^*, \ldots, p_n^* \in \mathbb{R}^d \\ \operatorname{diam}\{p_1^*, p_2^*, \ldots, p_n^*\} \leq 1}} \mathbb{E}\left[\operatorname{card} \mathrm{CH}\left(\{p_1^* + \eta_1, p_2^* + \eta_2, \ldots, p_n^* + \eta_n\}\right)\right]$$

where diam denotes the diameter, card S denotes the cardinality of a set S, $\mathrm{CH}(X)$ denotes the set of faces, of all dimensions, of the convex hull of X, and $\eta_1, \eta_2, \ldots, \eta_n$ are random variables chosen independently from the distribution μ. In this paper, we present upper and lower bounds on $\mathcal{S}(n, \mathcal{U}_{\delta\mathbb{B}})$, where $\mathcal{U}_{\delta\mathbb{B}}$ is the uniform distribution on the ball of radius δ centered in the origin in $\mathbb{R}^d$, and $\mathcal{S}(n, \mathcal{N}(0, \sigma^2 I_2))$, where $\mathcal{N}(0, \sigma^2 I_2)$ is the Gaussian distribution centered in the origin and with covariance matrix $\sigma^2 I_2$.

1.1 Context and Motivations

The interest for smoothed complexity arises from considerations in the analysis of algorithms in computational geometry and relates to the study of random polytopes in probabilistic geometry.

Analysis of Algorithms. To understand and predict the practical behaviour of an algorithm, a first step is to analyze how the amount of resources it requires grows with the size of the input. The basic building blocks of geometric algorithms are combinatorial structures induced by geometric data such as convex hulls or Voronoi diagrams of finite point sets, lattices of polytopes obtained as intersections of half-spaces, intersection graphs or nerves of families of balls... The size of these structures usually depends not only on the number of geometric primitives (points, half-spaces, balls...), but also on their relative position: for instance, the number of faces of the Voronoi diagram of n points in $\mathbb{R}^d$ is $\Theta(n)$ if these points form a regular grid but $\Theta\left(n^{\lceil d/2 \rceil}\right)$ if they lie on the moment curve. (We assume here a *Real RAM* model of computation, so the points have arbitrary real coordinates and the input size is simply the number n of points.)

There are two traditional approaches to account for how the complexity of a structure depends on the position of the points that induce it: the *worst-case complexity*, which measures the maximum of the complexity function over the input space, and the *average-case analysis*, which averages the complexity function against a suitable probability distribution on the space of inputs. Unfortunately, both approaches have shortcomings: the worst-case may be exceedingly pessimistic when the maximum is achieved only by constructions that are so brittle that it is unlikely they arise in practice,[1] whereas the input distributions considered for the average complexity are often unconvincing for lack of relevant and tractable statistical models to work with.

[1]For instance, while Delaunay triangulations in $\mathbb{R}^3$ have quadratic worst-case complexity, they appear to have near-linear size for the point sets arising in practice in the context of reconstruction [4]; one should thus not consider Delaunay-based reconstruction methods inefficient on the sole ground of worst-case analysis. The worst-case analysis can sometimes be refined by introducing additional parameters such as fatness [7] or spread [12], but *realistic input models* remain elusive in many contexts (*eg.* computer graphics scenes).

The *smoothed complexity* model, proposed by Spielman and Teng [19] in the early 2000's, interpolates between the worst-case and the average case model. Informally, it is defined as the maximum over the inputs of the expected complexity over small perturbations of that input. Intuitively, this "local averaging" mechanism disposes of configurations that vanish under small perturbation and models more accurately the behaviour on "real data", which is usually given with bounded precision and subject to measurement noise. In other words, the smoothed complexity quantifies the stability of bad configurations.

Stochastic Geometry. The study of random polytopes goes back to the celebrated *four point problem* of Sylvester [20] and is an important subject in probabilistic geometry. A well-established model of random polytopes consists in taking the convex hull of a family of random points distributed identically according to some measure. Our model of random polytope contains this model (in short: by taking all points in P^* in the origin) and naturally generalizes it. Starting with the seminal articles of Renyi and Sulanke [17, 18] in the 1960's, a series of works in stochastic geometry led to precise quantitative statements (*eg.* central limit theorems) for models such as convex hulls of points sampled i.i.d. from a Gaussian distribution or the uniform measure on a convex body; we refer the interested reader to the recent survey of Reitzner [16]. In the 1980's, Bárány and Larman [2] related random polytopes obtained from uniform distributions over convex bodies to the classical theory of *floating bodies* in convex geometry; we come back to this in Section 1.3 as several of the key ideas behind our results are already present in their work.

1.2 New Results

Our main results are a technique to analyze random geometric hypergraphs, which we call the *witness-collector technique*, as well as its application to the analysis of the smoothed complexity of convex hulls. Before we spell them out we need to clarify some terminology.

Random Geometric Hypergraphs. Let $\mathcal{X}$ be a set, $(\mathcal{X}, \mathcal{R})$ a range space (i.e. $\mathcal{R}$ is a family of subsets (ranges) of $\mathcal{X}$) and P a finite set of random elements of $\mathcal{X}$. The *random geometric hypergraph* induced by $(\mathcal{X}, \mathcal{R})$ on P is the set $\mathcal{H} = \{P \cap r : r \in \mathcal{R}\}$; that is, a subset $Q \subset P$ is a hyperedge of $\mathcal{H}$ if and only if there exists $r \in \mathcal{R}$ such that $r \cap P = Q$. Our analyses of random convex hulls proceed by analyzing random geometric hypergraphs where $\mathcal{X} = \mathbb{R}^d$, $\mathcal{R}$ is the set of all half-spaces of $\mathbb{R}^d$, and the elements of P are chosen independently (but not identically distributed!). Any face of the convex hull of P is a hyperedge of $\mathcal{H}$, but the converse is not true. It turns out, however, that the average size of $\mathcal{H}$ is close enough to that of $\mathrm{CH}(P)$ that our technique yields meaningful upper and lower bounds on the smoothed complexity of convex hulls (*cf.* Section 2.3).

Notations for Orders of Magnitude. Our goal is to understand how the order of magnitude of the smoothed complexity depends on the number n of points and the amplitude δ or σ of the perturbation. For the sake of the presentation, we do not keep track in our analyses of additive or multiplicative constants depending on fixed quantities such as the dimension of the space. Throughout the paper, we therefore write $a = O(b)$, $a = \Omega(b)$ and $a = \Theta(b)$ to mean that there exist positive reals c and c' such that, respectively, $a \leq cb$, $a \geq cb$ and $cb \leq a \leq c'b$; we also use $\Theta(b)$ (and similarly for $O()$ and $\Omega()$) as a shorthand for a quantity x for which $x = \Theta(b)$ holds. These notations do *not* carry any asymptotic meaning (since several variables may assume large and unrelated values); when used without stating any condition on n, σ or δ, these notations mean inequalities that hold for any $n \geq d + 1$, $\delta > 0$ and $\sigma > 0$.

The Witness-Collector Technique. Let $(\mathcal{X}, \mathcal{R})$ denote a range space. Our analyses are based on the following notion:

Definition 1. *A* system of witnesses and collectors *for a covering $R_1 \cup R_2 \cup \ldots \cup R_m$ of $\mathcal{R}$ is a family $\{(W_i^j, C_i^j)\}_{\substack{1 \le i \le m \\ 1 \le j \le \ell}}$ of pairs of subsets of $\mathcal{X}$ such that*

(a) for all i, j, any $r \in R_i$ contains W_i^j or is contained in C_i^j,
(b) for all i, $W_i^1 \subseteq W_i^2 \subseteq \ldots \subseteq W_i^\ell$,
(c) for all i, j, $W_i^j \subseteq C_i^j$.

We denote by $\mathcal{H}^{(k)}$ the set of hyperedges of cardinality k of a hypergraph $\mathcal{H}$. Our analyses are based on the following theorem, which we prove in Section 2:

Theorem 2. *Let $(\mathcal{X}, \mathcal{R})$ be a range space, let P be a set of n random elements of $\mathcal{X}$ chosen independently and let $\mathcal{H}$ denote the hypergraph induced by $\mathcal{R}$ on P.*

(i) If $\{(W_i^j, C_i^j)\}_{\substack{1 \le i \le m \\ 1 \le j \le \ln^2 n}}$ is a system of witnesses and collectors for a covering $R_1 \cup R_2 \cup \ldots \cup R_m$ of $\mathcal{R}$ such that $W_i^j \cap P$ and $C_i^j \cap P$ have average size $\Omega(j)$ and $O(j)$ respectively then $\mathbb{E}\left[\operatorname{card} \mathcal{H}^{(k)}\right] = O(m)$.

(ii) If every element of $\mathcal{H}^{(1)}$ is in at least one element of $\mathcal{H}^{(k)}$, and $\{W_i^1\}_{1 \le i \le m}$ is a family of disjoint subsets of $\mathcal{X}$ such that $\mathbb{E}\left[\operatorname{card}\left(W_i^1 \cap \mathcal{H}^{(1)}\right)\right] = \Omega(1)$ then $\mathbb{E}\left[\operatorname{card} \mathcal{H}^{(k)}\right] = \Omega(m)$.

In several of our applications we first construct a system $\{(W_i^j, C_i^j)\}$ of witnesses and collectors satisfying the assumptions of Theorem 2 (i), then use a subfamily of the W_i^1's that are disjoint to apply Theorem 2 (ii).

Applications. We present, in Sections 3 and 4, two designs of systems of witnesses and collectors suited to study the smoothed complexity of convex hulls relative to Euclidean and Gaussian perturbations with the following results (*cf.* Figures 1 and 2):

Smoothed Complexity. We obtain upper bounds on the smoothed complexity of convex hulls relative to Euclidean and Gaussian perturbations; in the Euclidean case we obtain sharper bounds for the smoothed number of vertices. We also analyze the convex hull of perturbations of points in convex position and delineate the main regimes in terms of the number of points and the amplitude of the perturbation; this provides lower-bounds on the Euclidean and Gaussian smoothed complexities of convex hulls.

Large Perturbations. We show that for $\delta = \Omega\left(n^{\frac{2}{d+1}}\right)$ the smoothed complexity of convex hulls relative to $\mathcal{U}_{\delta\mathbb{B}}$ is of the same order of magnitude as the expected complexity of the convex hull of random points chosen i.i.d. from $\mathcal{U}_{\delta\mathbb{B}}$, the classical model of random polytope. Our smoothed complexity upper bound also implies a similar result for Gaussian perturbation with $\sigma = \Omega(1)$.

Simple Analysis of Classical Random Polytopes. The classical model of random polytopes corresponds to the case where all points of p_i^* coincide. There, our systems of witnesses and collectors yield the order of magnitude of the expected number of faces with considerably less effort than earlier analyses.

A Surprising Phenomenon. We observed experimentally (Figure 2c) that the expected size of the convex hull of perturbations of points in convex position consistently decreases with the amplitude of the noise in the Gaussian model, whereas some non-monotonicity appears in the Euclidean model. Our analyses of perturbations of points in convex position provide a theoretical confirmation of this difference in behaviours (see Figures 1b and 2a).

As evidence that the witness-collector technique is relevant for the study of other geometric hypergraphs, we outline a design of witnesses and collectors that yields the order of magnitude of the number of faces in the Delaunay triangulations of a set of random points chosen uniformly and independently from the unit ball (Theorem 12); again, this is a well-known result but the proof (only sketched here) is considerably shorter than the original one.

1.3 Related Works

The results presented here appeared in preliminary form in research reports [1, 8] and proceedings of conferences [9, 10]. Note that the shift from *static* to *adaptative* witness-collectors in Section 2.2 is based on an idea which we learned from [13] and systematize here. We briefly position our results with respect to prominent related previous work.

Smoothed Number of Dominant Points. The only previous bound on the smoothed complexity of convex hulls is due to Damerow and Sohler [6]. They study the number of *dominant* points under Gaussian and ℓ^∞ perturbations (we included the results for the Gaussian case in Figures 1d and 2b). Their technique requires that the perturbation acts independently on each coordinate (thus restricting possible perturbations) so that the analysis of point dominance reduces to considerations on independent random permutations. The number of dominating points bounds from above the number of extreme points, but in probabilistic setting these two quantities typically have different orders of magnitude. As a consequence, the upper bounds are not sharp and there is no lower bound.

One may expect that when the magnitude of the perturbations is sufficiently large compared to the scale of the initial input, the initial position of the points does not matter and smoothed complexity is subsumed by some average-case analysis (up to constant multiplicative factors). The main insight of Damerow and Sohler [6] is a quantitative version of this claim. Specifically, they show that if n points from a region of diameter r are perturbed by a Gaussian noise of standard deviation $\Omega(r\sqrt{\ln n})$ or a ℓ^∞ noise of amplitude $\Omega(r\sqrt[3]{n/\ln n})$ then the expected number of dominating points is the same as in the average-case analysis. A smoothed complexity bound then follows by a simple rescaling argument: split the input domain into cells of size $r = O(\sigma/\sqrt{\ln n})$, assume that *each* cell contains all of the initial point set, and charge each of them with the average-case bound.

Our technique yields a similar subsuming of the smoothed complexity analysis by the average-case analysis for the number of faces (Lemma A.1) with the same threshold, thus extending Damerow and Sohler's main insight; we also obtain a similar statement in the Euclidean model (Lemma 3.8). The smoothed complexity bound we obtain by the rescaling argument (Corollary 13) is better than the one of Damerow and Sohler because the average number of extreme points is asymptotically smaller than the number of dominant points. It should be noted that the rescaling argument only applies to bound the number of vertices of the convex hull since faces of higher dimension may come from more than one cell; in any case, we further improve the bound obtained from the rescaling argument by a more direct analysis that accounts for faces of arbitrary dimension (Theorem 9).

Smoothed Complexity of a Simplex Algorithm. A substantial literature in the analysis of algorithms was devoted to explain the very good practical performance of the simplex algorithm,

any d	Range of δ	$\left[0, n^{\frac{2}{d+1}-\frac{1}{d-1}\lfloor\frac{d}{2}\rfloor}\right]$	$\left[n^{\frac{2}{d+1}-\frac{1}{d-1}\lfloor\frac{d}{2}\rfloor}, 1\right]$	$\left[1, 3n^{\frac{2}{d+1}}\right]$	$\left[3n^{\frac{2}{d+1}}, +\infty\right)$
	$\mathcal{S}(n, \mathcal{U}_{\delta\mathbb{B}})$	$O\left(n^{\lfloor\frac{d}{2}\rfloor}\right)$	$O\left(n^{2\frac{d-1}{d+1}}\delta^{-(d-1)}\right)$	$O\left(n^{2\frac{d-1}{d+1}}\right)$	$\Theta\left(n^{\frac{d-1}{d+1}}\right)$
any d	$\mathbb{E}\left[\text{card}\,\mathcal{H}^{(1)}\right] = O\left(n^{\frac{d-1}{d+1}} + \delta^{-\frac{2d}{d+1}} n^{1+2\frac{d-1}{(d+1)^2}}\right)$				

$d=2$	Range of δ	$[0, \frac{1}{\sqrt{n}}]$	$[\frac{1}{\sqrt{n}}, 1]$	$[1, n^{5/12}]$	$[n^{5/12}, n^{2/3}]$	$[n^{2/3}, +\infty]$
	$\mathcal{S}(n, \mathcal{U}_{\delta\mathbb{B}})$	$O(n)$	$O\left(\delta^{-\frac{2}{3}} n^{\frac{2}{3}}\right)$	$O\left(n^{2/3}\right)$	$O\left(\delta^{-\frac{4}{3}} n^{\frac{11}{9}}\right)$	$O\left(n^{1/3}\right)$

(a) Upper bounds on the smoothed complexity relative to Euclidean perturbations (Theorem 5 and Corollary 6).

Range of δ	$0 \leq \delta \leq n^{\frac{2}{1-d}}$	$n^{\frac{2}{1-d}} \leq \delta \leq 1$	$1 \leq \delta \leq n^{\frac{2}{d+1}}$	$n^{\frac{2}{d+1}} \leq \delta$
$\mathbb{E}\left[\text{card}\,\text{CH}(P)\right]$	$\Theta(n)$	$\Theta\left(n^{\frac{d-1}{2d}} \delta^{\frac{1-d^2}{4d}}\right)$	$\Theta\left(n^{\frac{d-1}{2d}} \delta^{\frac{(1-d)^2}{4d}}\right)$	$\Theta\left(n^{\frac{d-1}{d+1}}\right)$

(b) Expected complexity of a Euclidean perturbation P of a regular sample of the unit sphere in $\mathbb{R}^d$ (Theorem 7). This gives a lower bound on the smoothed complexity for Euclidean perturbation.

	any d	$d=2$
$\delta \geq \delta_0 \Rightarrow$ average-case behaviour	$\delta_0 = O\left(n^{\frac{2}{d+1}}\right)$	$\delta_0 = O\left(n^{2/3}\right)$

(c) Amplitude of an Euclidean perturbation for which the smoothed complexity behaves as the average-case complexity (Lemma 3.8).

	Our bounds ($d=2$)	Previous bound [6][a]
$\sigma \geq \sigma_0 \Rightarrow$ average-case behavior	$\sigma_0 = O(1)$	$\sigma_0 = O\left(\sqrt{\ln n}\right)$
$\mathcal{S}(n, \mathcal{N}(0, \sigma^2 I_2))$	$O\left(\sqrt{\ln n} + \sigma^{-1}\sqrt{\ln n}\right)$	$O\left(\ln n + \sigma^{-2}\ln^2 n\right)$

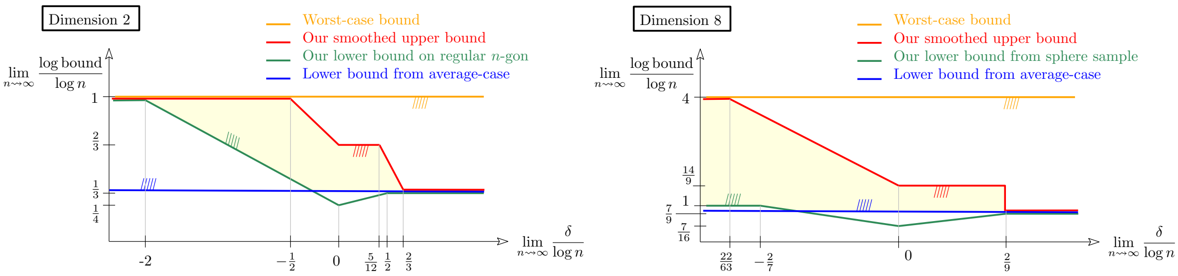
(d) Upper bounds for Gaussian perturbations (Theorem 9).

[a] This bound applies to *dominating point*, *cf.* the comparison to earlier work.

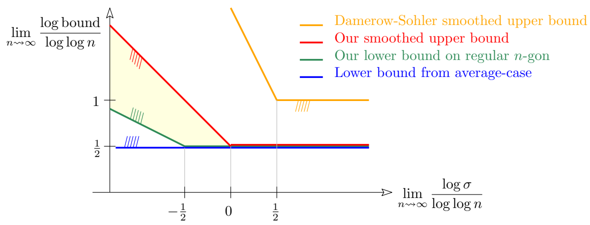
Range of σ	$0 \leq \sigma \leq \frac{1}{n^2}$	$\frac{1}{n^2} \leq \sigma \leq \frac{1}{\sqrt{\ln n}}$	$\frac{1}{\sqrt{\ln n}} \leq \sigma$
$\mathbb{E}\left[\text{card}\,\text{CH}(P)\right]$	$\Omega(n)$	$\Omega\left(\frac{\sqrt[4]{\ln(n\sqrt{\sigma})}}{\sqrt{\sigma}}\right)$	$\Omega\left(\sqrt{\ln n}\right)$

(e) Expected complexity of a Gaussian perturbation P of a regular n-gon in $\mathbb{R}^2$ (Theorem 10). This gives a lower bound on the smoothed complexity for Gaussian perturbation.

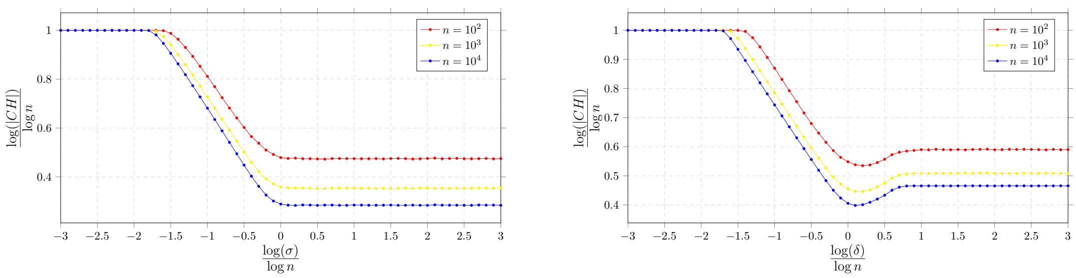
Figure 1: Summary of our bounds.



(a) A comparison of our smoothed complexity bound for Euclidean perturbation (Theorem 5 and Corollary 6) and two lower bounds, where the initial points are placed respectively at the vertices of a unit-size n-gon (Theorem 7) and in the origin. A data point with coordinates (x, y) means that for a perturbation with δ of magnitude n^x the expected size of the convex hull grows as n^y, subpolynomial terms being ignored. The worst-case bound is given as a reference. The constants in the $O()$ and $\Omega()$ have been ignored as their influence vanishes as $n \to \infty$ in this coordinate system.



(b) Comparison of the smoothed bounds for Gaussian perturbation in dimension 2 (Theorem 9 and [6]) and the lower bound perturbing the regular n-gon (Theorem 10). A data point with coordinates (x, y) means that for a perturbation of magnitude $\sigma = \ln^x n$ the expected size of the convex hull grows as $\ln^y n$.



(c) Experimental results for the complexity of the convex hull of a perturbation of the regular n-gon inscribed in the unit circle. Left: Gaussian perturbation of variance σ^2. Right: Euclidean perturbation of amplitude δ. Each data point corresponds to an average over 1000 experiments.

Figure 2: Plots summarizing the main results.

given that most of the pivoting rules had exponential worst-case complexity. This motivated the study of various models of random polytopes, and eventually the introduction of the smoothed complexity analysis model by Spielman and Teng [19]. We encourage the interested reader to consult their discussion of earlier literature, and simply compare our work to the smoothed complexity bound for convex hulls that is at the core of their analysis of the shadow-vertex pivot rule. They estimate the expected number of vertices of an arbitrary two-dimensional projection of a polytope given as an intersection of n halfspaces in d dimensions and perturbed by a Gaussian noise of standard deviation σ using techniques quite different from ours, see [19, Th 4.1]. Neither n nor d are fixed, so the number of vertices may be exponential in the input; their analysis shows that it is polynomial in n, d and $\frac{1}{\sigma}$. The question we consider is therefore, from the point of view of the model, of a rather different nature: we consider the dimension to be fixed rather than variable, specify the polytope as a convex hull of vertices rather than intersection of half-spaces, and estimate the number of faces rather than the two-dimensional silhouettes. More importantly, our intent is to understand a transition within the polynomial domain rather than identify a polynomial behaviour in place of an exponential worst-case bound.

Floating Bodies and Economic Cap Coverings. Bárány and Larman [2] established that the expected number of faces of the convex hull of n random points chosen uniformly from a convex body K is $\Theta\left(nK\left(\frac{1}{n}\right)\right)$, where $K(t)$ denotes the volume of the *wet part* of K with parameter t: the union of the intersections of K with a half-space that intersects it with volume at most t. This connection allowed them to transfer to the study of random polytopes various results from convex geometry, for which wet parts, or their complements the floating bodies, are classical objects.

When the ranges are half-spaces in $\mathbb{R}^d$, our systems of witnesses and collectors are essentially equivalent to the *economic cap covers* on which Bárány and Larman's proof is based (Bárány and Vu [3, § 5] also use the same idea in the proving of a central limit theorems for Gaussian polytopes). A first difference is that the analogue of our Condition (a) for economic cap covers is formulated in terms of wet parts, so the role of the range space is implicit. This has little effect as far as the ranges are half-spaces, but we note that the analogue of wet parts for other range spaces is not straightforward to define and study, whereas our presentation naturally extends to other range spaces (as the case of Delaunay triangulation sketched in Section 5 demonstrates). We also note that the constructions of systems of witnesses and collectors differ from the constructions of economic cap covers, but believe that this is a less essential distinction.

2 Witnesses and Collectors

In this section we first explain the idea behind Theorem 2 in a simpler setting in Section 2.1, then prove Theorem 2 in Section 2.2, then clarify its use for the analysis of convex hulls of random point sets in Section 2.3.

2.1 Principle: Static Witnesses and Collectors

Let $(\mathcal{X}, \mathcal{R})$ be a range space, P a random set of n elements of $\mathcal{X}$ chosen independently, $\mathcal{H}$ the hypergraph induced by $\mathcal{R}$ on P, and $k \in \mathbb{N}$. Let $R_1 \cup R_2 \cup \ldots \cup R_m$ be a covering of $\mathcal{R}$ and $\{(W_i^1, C_i^1)\}_{1 \leq i \leq m}$ a system of witnesses and collectors for that covering. Since $\ell = 1$, we shorten W_i^1 into W_i and C_i^1 into C_i and note that Condition (b) is trivial.

Conditioning on Loaded Witnesses. If $\mathrm{card}\,(W_i \cap P)$ is at least k then Condition (a) ensures that every hyperedge of size k in $\{r \cap P : r \in R_i\}$ is contained in C_i, so there are at most $\mathbb{E}\left[\mathrm{card}\,(C_i \cap P)^k\right]$ such hyperedges; otherwise we can use the trivial upper bound $\binom{n}{k}$. Conditioning on the event that $\mathrm{card}\,(W_i \cap P)$ is at least k for all i we therefore get

$$\begin{aligned}\mathbb{E}\left[\mathrm{card}\,\mathcal{H}^{(k)}\right] \quad &\leq \quad \mathbb{P}\left[\exists i, \mathrm{card}\,(W_i \cap P) < k\right]\binom{n}{k} \\ &\qquad + \mathbb{P}\left[\forall i, \mathrm{card}\,(W_i \cap P) \geq k\right] \cdot \sum_{i=1}^{m} \mathbb{E}\left[\mathrm{card}\,(C_i \cap P)^k\right]\end{aligned}$$

so if the witnesses are chosen so that $\mathrm{card}\,(W_i \cap P) \geq k$ with probability $1 - O(n^{-k})$,

$$\begin{aligned}\mathbb{E}\left[\mathrm{card}\,\mathcal{H}^{(k)}\right] \quad &= \quad O\left(\left(\sum_{i=1}^{m} \mathbb{P}\left[\mathrm{card}\,(W_i \cap P) < k\right]\right) \cdot \binom{n}{k}\right. \\ &\qquad \left. + \mathbb{P}\left[\forall i, \mathrm{card}\,(W_i \cap P) \geq k\right] \cdot \sum_{i=1}^{m} \mathbb{E}\left[\mathrm{card}\,(C_i \cap P)^k\right]\right) \\ &= \quad O\left(m + \sum_{i=1}^{m} \mathbb{E}\left[\mathrm{card}\,(C_i \cap P)^k\right]\right). \qquad (1)\end{aligned}$$

Role of $\mathbf{W_i \cap P}$ and $\mathbf{C_i \cap P}$. Chernoff's multiplicative bound implies that if $W_i \cap P$ has average size $\Omega(k \ln n)$ then indeed $\mathrm{card}\,(W_i \cap P) \geq k$ with probability $1 - O(n^{-k})$. More generally:

Lemma 2.1. *Let P be a set of random elements of $\mathcal{X}$ chosen independently and W a subset of $\mathcal{X}$.*

(a) $\mathbb{P}\left[W \cap P = \emptyset\right] \leq e^{-\mathbb{E}[\mathrm{card}(W \cap P)]}$.

(b) If $\mathbb{E}\left[\mathrm{card}\,(W \cap P)\right] \geq k+1$ *then* $\mathbb{P}\left[\mathrm{card}\,(W \cap P) < k\right] \leq e^{-\Omega(\mathbb{E}[\mathrm{card}(W \cap P)])}$.

(We defer the proof to Section 2.4.) The bound in Equation (1) is expressed in terms of the $\mathbb{E}\left[\mathrm{card}\,(C_i \cap P)^k\right]$ but can be controlled by $\mathbb{E}\left[\mathrm{card}\,(C_i \cap P)\right]$ since the elements of P are chosen independently:

Lemma 2.2. *If $V = \sum_{i=1}^{n} V_i$, where the V_i are independently distributed random variables with value in $\{0,1\}$ and $\mathbb{E}[V] \geq 1$ then $\mathbb{E}\left[V^k\right] = O\left(\mathbb{E}[V]^k\right)$.*

(Again, the proof is postponed to Section 2.4.) In the situations we consider, one can construct witnesses and collectors such that $W_i \cap P$ and $C_i \cap P$ both have expected size $\Theta(k \ln n)$; see [9] for several examples. Equation (1) and Lemma 2.2 then yield that $\mathbb{E}\left[\mathrm{card}\,\mathcal{H}^{(k)}\right]$ is of order m up to some logarithmic factors.

Shaving Log Factors. The use of a Chernoff bound to control the probability that witnesses contain fewer than k elements increases the expected size of the $W_i \cap P$ so that *all of them* are large for *most* realizations of P. By Condition (c), $W_i \subseteq C_i$, so this also overloads the collectors, resulting in the extra log factors. The idea that leads to the sharper bounds of Theorem 2, which we learned from [13], is to make W_i and C_i random variables depending on P. By adapting the witness-collector pairs used in the analysis to each realization of P, very few collectors will need to be large, and their contribution to the total will remain negligible.

It is perhaps worth pointing out that the above analysis holds for several of our constructions when only the first layer ($j = 1$) of witnesses and collectors is considered. Our proofs can therefore be further simplified should one not care about some extra logarithmic factors.

2.2 Proof of Theorem 2: Adaptative Witnesses and Collectors

We first prove the upper bound, in a format that will allow slightly more flexibility.

Lemma 2.3. *Let $(\mathcal{X}, \mathcal{R})$ be a range space, let P be a set of n random elements of $\mathcal{X}$ chosen independently and let $\mathcal{H}$ denote the hypergraph induced by $\mathcal{R}$ on P. If $R_1 \cup R_2 \cup \ldots \cup R_m$ is a covering of $\mathcal{R}$ and $\{(W_i^j, C_i^j)\}_{\substack{1\le i\le m\\ 1\le j\le \ln^2 n}}$ is a system of witnesses and collectors for that covering with*

$$\mathbb{P}\left[\operatorname{card}\left(W_i^j \cap P\right) < k\right] = O\left(e^{-\Omega(j)}\right) \quad \text{and} \quad \mathbb{E}\left[\operatorname{card}\left(C_i^j \cap P\right)\right] = O(j)$$

then $\mathbb{E}\left[\operatorname{card}\mathcal{H}^{(k)}\right]$ is $O(m)$.

Proof. Let $i \in \{1, 2, \ldots, m\}$. We let d_i denote the smallest j such that W_i^j contains at least k points and $C_i = C_i^{d_i}$, or, if no such W_i^j exists, $d_i = \infty$ and $C_i = \mathcal{X}$. (So d_i and C_i are random variables depending on P.) By Condition (a) and the definition of d_i, every hyperedge of $\mathcal{H}$ of size k induced by R_i is contained in C_i so:

$$\mathbb{E}\left[\operatorname{card}\mathcal{H}^{(k)}\right] \le \sum_{i=1}^{m} \mathbb{E}\left[\operatorname{card}\left(C_i \cap P\right)^k\right]. \tag{2}$$

Moreover, by Condition (b) we have $\mathbb{P}[d_i \ge j] = \mathbb{P}\left[\operatorname{card}\left(W_i^j \cap P\right) < k\right] = O\left(e^{-\Omega(j)}\right)$. We claim that

$$\mathbb{E}\left[\operatorname{card}\left(C_i^j \cap P\right) \mid d_i \ge j\right] \le \mathbb{E}\left[\operatorname{card}\left(C_i^j \cap P\right)\right] + k = O(j) \tag{3}$$

Indeed, working with the complement $\bar{C}_i^j$ of C_i^j, $\mathbb{E}\left[\operatorname{card}\left(\bar{C}_i^j \cap P\right)\right] = \sum_{p\in P} \mathbb{P}\left[p \notin C_i^j\right]$. For any $T \subset P$ we have

$$\mathbb{E}\left[\operatorname{card}\left(\bar{C}_i^j \cap P\right) \mid W_i^j \cap P = T\right] = \sum_{p\in P\setminus T} \mathbb{P}\left[p \notin C_i^j \mid p \notin W_i^j\right] \ge \sum_{p\in P\setminus T} \mathbb{P}\left[p \notin C_i^j\right],$$

the last inequality following from Condition (c). Thus,

$$\mathbb{E}\left[\operatorname{card}\left(\bar{C}_i^j \cap P\right)\right] \le \mathbb{E}\left[\operatorname{card}\left(\bar{C}_i^j \cap P\right) \mid W_i^j \cap P = T\right] + \operatorname{card} T.$$

By Condition (b), $d_i \ge j$ if and only if $\operatorname{card}\left(W_i^{j-1} \cap P\right) < k$. Total probabilities let us decompose this event:

$$\begin{aligned}
&\mathbb{E}[\operatorname{card}\left(\bar{C}_i^j \cap P\right) \mid d_i \ge j] \\
&= \sum_{T:\operatorname{card} T<k} \mathbb{E}\left[\operatorname{card}\left(\bar{C}_i^j \cap P\right) \mid W_i^{j-1} \cap P = T\right] \mathbb{P}\left[W_i^{j-1} \cap P = T \mid \operatorname{card}\left(W_i^{j-1} \cap P\right) < k\right] \\
&\ge \left(\mathbb{E}\left[\operatorname{card}\left(\bar{C}_i^j \cap P\right)\right] - k\right) \sum_{T:\operatorname{card} T<k} \mathbb{P}\left[W_i^{j-1} \cap P = T \mid \operatorname{card}\left(W_i^{j-1} \cap P\right) < k\right] \\
&= \mathbb{E}\left[\operatorname{card}\left(\bar{C}_i^j \cap P\right)\right] - k
\end{aligned}$$

Moving back to the complement yields Inequation (3). Now, since P has n points in total, conditioning on the value of d_i we obtain

$$\begin{aligned}
\mathbb{E}\left[\operatorname{card}\left(C_i \cap P\right)\right] &= \sum_{j=1}^{\ln^2 n} \mathbb{E}\left[\operatorname{card}\left(C_i^j \cap P\right) \cdot \mathbb{1}_{d_i=j}\right] + \mathbb{E}\left[n \cdot \mathbb{1}_{d_i=\infty}\right] \\
&\leq \sum_{j=1}^{\ln^2 n} \mathbb{E}\left[\operatorname{card}\left(C_i^j \cap P\right) \cdot \mathbb{1}_{d_i\geq j}\right] + \mathbb{E}\left[n \cdot \mathbb{1}_{d_i=\infty}\right] \\
&= \sum_{j=1}^{\ln^2 n} \mathbb{E}\left[\operatorname{card}\left(C_i^j \cap P\right) \ \middle|\ d_i \geq j\right] \mathbb{P}\left[d_i \geq j\right] + n \cdot \mathbb{P}\left[d_i = \infty\right] \\
&= \sum_{j=1}^{\ln^2 n} O\left(je^{-\Omega(j)}\right) + O\left(ne^{-\Omega(\ln^2 n)}\right)
\end{aligned}$$

so each collector C_i contains on average a constant number of elements of P. Lemma 2.2 and Equation (2) imply that $\mathbb{E}\left[\operatorname{card} \mathcal{H}^{(k)}\right] = O(m)$. $\square$

We now wrap-up the proof of our witness-collector theorem.

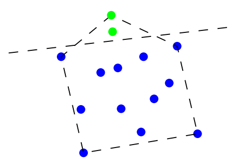
Proof of Theorem 2. Since $\mathbb{E}\left[\operatorname{card}\left(W_i^j \cap P\right)\right] = \Omega(j)$, there exists some constant $c > 0$ such that $\mathbb{E}\left[\operatorname{card}\left(W_i^j \cap P\right)\right] \geq cj$. For $j \geq \frac{k+1}{c}$, the Chernoff bound of Lemma 2.1 (b) thus ensures that $\mathbb{P}\left[\operatorname{card}\left(W_i^j \cap P\right) < k\right]$ is at most $e^{-\Omega(j)}$. Bounding that probability from above by 1 in the cases $j < \frac{k+1}{c}$ we get that $\mathbb{P}\left[\operatorname{card}\left(W_i^j \cap P\right) < k\right]$ is $O\left(e^{-\Omega(j)}\right)$. Statement (i) then follows readily from Lemma 2.3.

Now consider Statement (ii). We can charge each element of $\mathcal{H}^{(1)}$ to an element of $\mathcal{H}^{(k)}$ that contains it. Since each element of $\mathcal{H}^{(k)}$ is charged at most k times, we have $\operatorname{card} \mathcal{H}^{(k)} \geq \frac{1}{k} \operatorname{card} \mathcal{H}^{(1)}$. The assumptions ensure that each W_i^1 contains on average $\Omega(1)$ elements of $\mathcal{H}^{(1)}$ and that these elements are distinct. It follows that $\mathbb{E}\left[\operatorname{card} \mathcal{H}^{(1)}\right]$ and $\mathbb{E}\left[\operatorname{card} \mathcal{H}^{(k)}\right]$ are $\Omega(m)$. $\square$

2.3 The Special Case of Convex Hulls

Unless indicated otherwise, in the remainder of this paper the range space $(\mathcal{X}, \mathcal{R})$ considered is that of half-spaces in $\mathbb{R}^d$. Every element of $\mathcal{H}^{(1)}$ belongs to some element of $\mathcal{H}^{(k)}$, so the first condition of Theorem 2 (ii) holds for this range space.

In this setting, the elements of $\mathcal{H}^{(k)}$ are also called the *k-sets* of the point set P. The bounds that we establish are expressed with $O()$, $\Omega()$ and $\Theta()$ in which the multiplicative constants depend on k; they are therefore valid for any *fixed* k. For $k \leq d$, any $(k-1)$-dimensional face of $\mathrm{CH}(P)$ is a k-set, so the upper bound of Theorem 2 (i) applies to the size of the convex hull. The reverse is not true (*cf.* the figure on the right) but we remark that $\mathcal{H}^{(1)}$ is exactly the set of vertices of $\mathrm{CH}(P)$ and that every element of $\mathcal{H}^{(1)}$ belongs to an actual $(k-1)$-dimensional face of $\mathrm{CH}(P)$; the proof of Statement (ii) of Theorem 2 therefore provides, *mutatis mutandis*, a lower bound on the number of $(k-1)$-dimensional faces of $\mathrm{CH}(P)$. In the rest of the paper, we will navigate without further justification between the convex hull of a random point set P and the associated random geometric hypergraph.



2.4 Proofs of Lemmas 2.1 and 2.2

Proof of Lemma 2.1. Let V_i be the indicator function of the event that the i^{th} point from P belongs to W. We write $V = V_1 + \ldots + V_n$ and let $t = \mathbb{E}[V]$. Chernoff's bound for lower tails yields that for any $\delta \in (0,1)$

$$\mathbb{P}[V < (1-\delta)t] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^t = e^{-t(1-(1-\delta)(1-\ln(1-\delta)))}. \tag{4}$$

In particular,

$$\mathbb{P}[V = 0] \leq \lim_{\delta\to 1} \mathbb{P}[V < (1-\delta)t] = \lim_{\delta\to 1} e^{-t(1-(1-\delta)(1-\ln(1-\delta)))} = e^{-t}$$

which proves Statement (a). Moreover, for $1-\delta = \frac{k}{t}$, Equation (4) specializes into

$$\mathbb{P}[V < k] < e^{-t\left(1-\frac{k}{t}\left(1-\ln\frac{k}{t}\right)\right)}$$

Since $x \mapsto x(1-\ln x)$ is increasing on $(0,1)$, for $t \geq k+1$ we have

$$1 - \frac{k}{t}\left(1-\ln\frac{k}{t}\right) \geq 1 - \frac{k}{k+1}\left(1-\ln\frac{k}{k+1}\right) > 0$$

and Statement (b) follows. □

Proof of Lemma 2.2. The statement is a special case of a classical inequality for sums of random variables [14, Th 2.12]; we give a simple, elementary, proof.

Expanding $V^k = \left(\sum_{i=1}^n V_i\right)^k$ we obtain

$$\begin{aligned}\mathbb{E}\left[V^k\right] &= \sum_{1\leq i_1,i_2\ldots i_k\leq n} \mathbb{E}\left[V_{i_1}\cdot V_{i_2}\ldots V_{i_k}\right] \\ &= \sum_{\ell=1}^{k} \sum_{\substack{1\leq i_1,i_2\ldots i_k\leq n\\ |\{i_1,i_2\ldots i_k\}|=\ell}} \mathbb{E}\left[V_{i_1}\cdot V_{i_2}\ldots V_{i_k}\right].\end{aligned}$$

Since the V_i's have values in $\{0,1\}$, for any positive integers $a_1, a_2, \ldots a_t$ and $i_1, i_2, \ldots i_t$

$$\mathbb{E}\left[V_{i_1}^{a_1}\cdot V_{i_2}^{a_2}\ldots V_{i_t}^{a_t}\right] = \mathbb{E}\left[V_{i_1}\cdot V_{i_2}\ldots V_{i_t}\right].$$

Letting $p(\ell,k)$ denote the number of partition of $\{1,2,\ldots,k\}$ in ℓ subsets, we can thus write

$$\mathbb{E}\left[V^k\right] = \sum_{\ell=1}^{k} \sum_{\substack{1\leq i_1,i_2\ldots i_\ell\leq n\\ i_a\neq i_b \text{ if } a\neq b}} p(\ell,k)\,\mathbb{E}\left[V_{i_1}\cdot V_{i_2}\ldots V_{i_\ell}\right].$$

Since V_i and V_j are independent if $i \neq j$ the previous identity rewrites as

$$\mathbb{E}\left[V^k\right] = \sum_{\ell=1}^{k}\left(p(\ell,k)\sum_{\substack{1\leq i_1,i_2\ldots i_\ell\leq n\\ i_a\neq i_b \text{ if } a\neq b}} \mathbb{E}\left[V_{i_1}\right]\cdot\mathbb{E}\left[V_{i_2}\right]\ldots\mathbb{E}\left[V_{i_\ell}\right]\right).$$

Thus,

$$\mathbb{E}\left[V^k\right] \leq \sum_{\ell=1}^{k} \left(p(\ell,k) \sum_{1\leq i_1,i_2\ldots i_\ell \leq n} \mathbb{E}\left[V_{i_1}\right]\cdot\mathbb{E}\left[V_{i_2}\right]\ldots\mathbb{E}\left[V_{i_\ell}\right]\right)$$

and since

$$\sum_{1\leq j_1,j_2,\ldots,j_\ell\leq n} \mathbb{E}\left[V_{j_1}\right]\cdot\mathbb{E}\left[V_{j_2}\right]\ldots\mathbb{E}\left[V_{j_\ell}\right] = \left(\sum_{i=1}^{n}\mathbb{E}\left[V_i\right]\right)^{\ell} = \mathbb{E}\left[V\right]^{\ell}$$

we finally obtain that

$$\mathbb{E}\left[V^k\right] \leq \sum_{\ell=1}^{k} p(\ell,k)\,\mathbb{E}\left[V\right]^{\ell} \leq \left(\sum_{\ell=1}^{k} p(\ell,k)\right)\mathbb{E}\left[V\right]^{k}$$

the last inequality following from the fact that $\mathbb{E}\left[V\right] \geq 1$. □

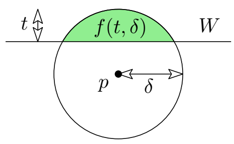
3 Euclidean Perturbations

We first consider the complexity of convex hulls of points perturbed under Euclidean perturbations.

Terminology and Notations. We denote by $\rho\mathbb{B}$ the ball of radius ρ centered at the origin of $\mathbb{R}^d$. Given $X \subset \mathbb{R}^d$ we denote by $\mathrm{vol}_k(X)$ its k-dimensional volume and by ∂X its boundary. We say that two half-spaces are *parallel* if they have the same inner normal. The *intersection depth* of a half-space W and a ball $p + \delta\mathbb{B}$ is $\delta - \bar{d}(p, W)$, where $\bar{d}(p, W)$ is the signed distance of p to ∂W (positive if and only if $p \notin W$).

3.1 Preliminaries: Ball/Half-space Intersection

We denote by $f(t, \delta)$ the volume of the intersection of $p + \delta\mathbb{B}$ with a half-space that intersects it with depth t. Note that $t \mapsto f(t, \delta)$ is increasing on $[0, 2\delta]$ for any fixed δ.



Claim 3.1. *For any $\lambda \geq 1$ and any $t \geq 0$, $f(\lambda t, \delta) \leq \lambda^{\frac{d+1}{2}} f(t, \delta)$.*

Proof. First assume that $\lambda t \leq 2\delta$. Let ν_{d-1} denote the volume of a $(d-1)$-dimensional ball of radius 1. By integrating along the direction of the inner normal to the half-space, we find

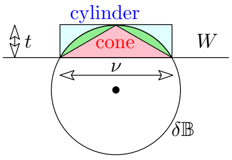
$$\begin{aligned} f(\lambda t,\delta) = \nu_{d-1}\int_0^{\lambda t}\left(2x\delta - x^2\right)^{\frac{d-1}{2}} dx &= \nu_{d-1}\int_0^{t}\lambda^{\frac{d-1}{2}}\left(2x\delta - \lambda x^2\right)^{\frac{d-1}{2}}\lambda dx \\ &\leq \nu_{d-1}\int_0^{t}\lambda^{\frac{d+1}{2}}\left(2x\delta - x^2\right)^{\frac{d-1}{2}} dx = \lambda^{\frac{d+1}{2}} f(t,\delta) \end{aligned}$$

which proves the claim. The case $\lambda t > 2\delta$ then follows easily:

$$f(\lambda t,\delta) = \mathrm{vol}_d\left(\delta\mathbb{B}\right) = f\left(\frac{2\delta}{t}t,\delta\right) \leq \left(\frac{2\delta}{t}\right)^{\frac{d+1}{2}} f(t,\delta) \leq \lambda^{\frac{d+1}{2}} f(t,\delta).$$

□

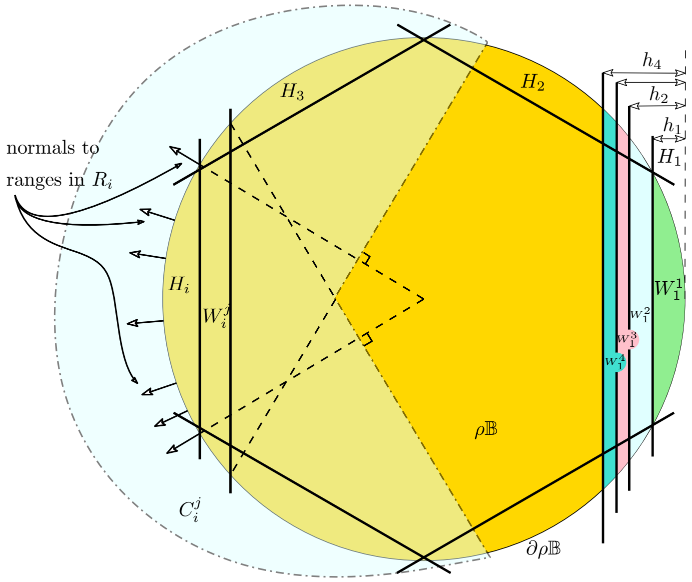
Claim 3.2. *For* $t \in [0, \delta]$, $f(t, \delta) = \Theta\left(t^{\frac{d+1}{2}} \delta^{\frac{d-1}{2}}\right)$.



Proof. Let W be a half-space that intersects $\delta\mathbb{B}$ with depth t and let $\nu = (\partial W) \cap \delta\mathbb{B}$. The region $W \cap \delta\mathbb{B}$ is sandwiched between a cone and a right cylinder with heights t and bases ν, with respective volumes $t\,\mathrm{vol}_{d-1}(\nu)/d$ and $t\,\mathrm{vol}_{d-1}(\nu)$. Since ν is a ball of radius r, with $r^2 = \delta^2 - (\delta - t)^2 = 2t\delta - t^2$, it has $(d-1)$-dimensional volume $\Theta\left((t\delta)^{\frac{d-1}{2}}\right)$ and the claim follows. $\square$

3.2 Witness-Collector Construction

Our systems of witnesses and collectors for Euclidean perturbations are based on the construction summarized in the following figure.



Definition. Our construction is parameterized by a radius ρ, usually chosen so that the *perturbed* point set remains inside $\rho\mathbb{B}$, and a sequence of positive reals $h_1 < h_2 < \ldots < h_\ell$. We let $H_1, H_2, \ldots H_m$ be an inclusion-minimal cover of $\partial(\rho\mathbb{B})$ by half-spaces of intersection depth h_1 with $\rho\mathbb{B}$.

Claim 3.3. $m = \Theta\left((\rho/h_1)^{\frac{d-1}{2}}\right)$.

Proof. If $h_1 \geq \rho$ then $m \leq 2$ so assume $h_1 < \rho$. The intersection $H_i \cap \partial(\rho\mathbb{B})$ is a spherical cap with radius $r = \Theta\left(\sqrt{h_1\rho}\right)$. Since the H_i form a minimal cover of $\partial(\rho\mathbb{B})$,

$$m = \Theta\left(\frac{\mathrm{vol}_{d-1}(\partial(\rho\mathbb{B}))}{\mathrm{vol}_{d-1}(\partial(\rho\mathbb{B}) \cap H_i)}\right)$$

The statement then follows from the fact that $\mathrm{vol}_{d-1}(\partial(\rho\mathbb{B}))$ and $\mathrm{vol}_{d-1}(\partial(\rho\mathbb{B}) \cap W_i^1)$ are, respectively, proportional to ρ^{d-1} and r^{d-1}. $\square$

We then define the range R_i as the set of half-spaces whose inner normal is parallel to a vector from the origin to a point of $H_i \cap \partial(\rho\mathbb{B})$. We define W_i^j as the intersection of $\rho\mathbb{B}$ with the half-space parallel to H_i and with intersection depth h_j with $\rho\mathbb{B}$. We define C_i^j as the union of the half-spaces of R_i that do not contain W_i^j.

Lemma 3.4. *$R_1 \cup R_2 \cup \ldots \cup R_m$ covers the set of half-spaces in $\mathbb{R}^d$ and $\{(W_i^j, C_i^j)\}_{\substack{1\le i\le m\\ 1\le j\le \ell}}$ is a system of witnesses and collectors for that covering. Moreover, a constant fraction of the W_i^1 are pairwise disjoint.*

Proof. The definition readily ensures that the union of the R_i is the set of all half-spaces and that Condition (a) holds. The monotonicity of the h_i implies that Condition (b) is also satisfied. Let $x \in W_i^j$.If $x \notin \partial H_i$, then let H denote the half-space parallel to H_i with x on its boundary. If $x \in \partial H_i$, we have to tilt the plane slightly: let H be a half space in R_i with x on its boundary but not parallel to H. In both cases H is in R_i and does not contain W_i^j and thus $x \in H \subset W_i^j$ and Condition (c) holds. As the cover of $\partial\rho\mathbb{B}$ by $\{H_i\}_{i=1,2,\ldots,m}$ is inclusion-minimal, we can extract a family $I \subseteq \{1, 2, \ldots, m\}$ of size $\Omega(m)$ such that the $\{W_i^1\}_{i\in I}$ are pairwise disjoint. □

In our analysis we will need some control over the intersection of C_i^j with $\rho\mathbb{B}$:

Claim 3.5. *$C_i^j \cap \rho\mathbb{B}$ is contained in a half-space parallel to H_i with intersection depth at most $9h_j$ with $\rho\mathbb{B}$.*

Proof. For any half-space H, the region $H \cap \rho\mathbb{B}$ is the convex hull of $H \cap \partial\rho\mathbb{B}$. It follows that $H \in R_i$ does not contain W_i^j if and only if $H \cap \partial\rho\mathbb{B}$ does not contain $W_i^j \cap \partial\rho\mathbb{B}$. This implies that for any $H \in R_i$ the spherical cap $H \cap \partial\rho\mathbb{B}$ is contained in a cap with same center as $W_i^j \cap \partial\rho\mathbb{B}$ and three times its radius. A half-space cutting out a cap of radius r_x in $\partial\rho\mathbb{B}$ intersects $\rho\mathbb{B}$ with depth $h_x = \Theta\left(\frac{r_x^2}{\rho}\right)$. Tripling the radius of a cap thus multiplies the depth of intersection by 9, and the statement follows. □

Claim 3.6. *If $\mathbb{E}\left[\text{card}\left(W_i^1 \cap P\right)\right] = \Omega(1)$ then $\mathbb{E}\left[\text{card}\left(W_i^1 \cap \mathcal{H}^{(1)}\right)\right] = \Omega(1)$*

Proof. If $W_i^1 \cap P$ is non-empty then W_i^1 contains the point of P extreme in direction $\vec{u}_i$ and $W_i^1 \cap \mathcal{H}^{(1)}$ is therefore non-empty. We thus have

$$\mathbb{E}\left[\text{card}\left(W_i^1 \cap \mathcal{H}^{(1)}\right)\right] \ge \mathbb{P}\left[W_i^j \cap \mathcal{H}^{(1)} \ne \emptyset\right] \ge \mathbb{P}\left[W_i^1 \cap P \ne \emptyset\right] \ge 1 - e^{-\mathbb{E}\left[\text{card}\left(W_i^1 \cap P\right)\right]} = \Omega(1),$$

the last inequality following from the Chernoff bound of Lemma 2.1 (a). □

3.3 Warm-up: Average-Case Analysis Made Easy

As a first example, let us use a system of witnesses and collectors to give a short[2] proof of a classical result of Raynaud.

Theorem 3 (Raynaud [15])**.** *Let $P = \{p_1, p_2, \ldots, p_n\}$ be a set of random points uniformly and independently distributed in a ball of $\mathbb{R}^d$. For any fixed k, the expected number of $k-$dimensional faces of the convex hull of P is $\Theta\left(n^{\frac{d-1}{d+1}}\right)$.*

[2] Raynaud's original argument was more than 7 pages long, still leaving substantial computations to the reader.

Proof. The problem is invariant under scaling, so we can choose the ball to be $\mathbb{B}$. We use our construction of Section 3.2 with $\rho = 1$. Using Claim 3.2, we find that setting $h_j = (j/n)^{\frac{2}{d+1}}$ yields

$$\mathbb{E}\left[\mathrm{card}\left(W_i^j \cap P\right)\right] = n\frac{f(h_j, 1)}{\mathrm{vol}\,(\mathbb{B})} = \Theta(j).$$

Claim 3.3 gives $m = \Theta\left((\rho/h_1)^{\frac{d-1}{2}}\right) = \Theta\left(n^{\frac{d-1}{d+1}}\right)$. With Claims 3.1 and 3.5 this implies

$$\mathbb{E}\left[\mathrm{card}\left(C_i^j \cap P\right)\right] \leq n\ O\left(\frac{f(h_j, 1)}{\mathrm{vol}\,(\mathbb{B})}\right) = O(j)$$

so $\mathbb{E}\left[\mathrm{card}\,\mathrm{CH}(P)\right] = O\left(n^{\frac{d-1}{d+1}}\right)$ by Theorem 2 (i). Moreover, a constant fraction of the W_i^1 are pairwise disjoint, and Claim 3.6 ensures that $\mathbb{E}\left[\mathrm{card}\left(W_i^1 \cap \mathcal{H}^{(1)}\right)\right] = \Omega(1)$; Theorem 2 (ii) thus implies that $\mathbb{E}\left[\mathrm{card}\,\mathrm{CH}(P)\right] = \Omega\left(n^{\frac{d-1}{d+1}}\right)$. □

3.4 Upper Bounds on the Smoothed Complexity

We now bound from above $\mathcal{S}\left(n, \mathcal{U}_{\delta\mathbb{B}}\right)$, using various arguments whose effectiveness vary with the value of δ.

Charging Argument. Our first smoothed complexity bound relies on a charging argument between the witness and the collector that form a pair. Let P^* be some point set of diameter at most 1 in $\mathbb{R}^d$. Without loss of generality we assume that P^* is contained in $\mathbb{B}$, and use a system of witnesses and collectors similar to the one presented in Section 3.2 with $\rho = 1 + \delta$.

We make an important change, though: the depth of intersection of each witness W_i^j depends on i, and is adapted to P^*. We start with an inclusion-minimal covering $H_1, H_2, \dots, H_m$ of $\partial(\rho\mathbb{B})$ by half-spaces whose intersection depth with $\rho\mathbb{B}$ is $\Theta\left(\left(\frac{r}{1+\delta}\right)^2\right)$. Each cuts out a spherical caps of radius $r = \delta n^{-\frac{2}{d+1}}$ on $\partial(\rho\mathbb{B})$, and $m = O\left(n^{2-\frac{4}{d+1}}\left(1 + \frac{1}{\delta}\right)^{d-1}\right)$ by Claim 3.3. For $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, \lceil \ln^2 n \rceil$ we define:

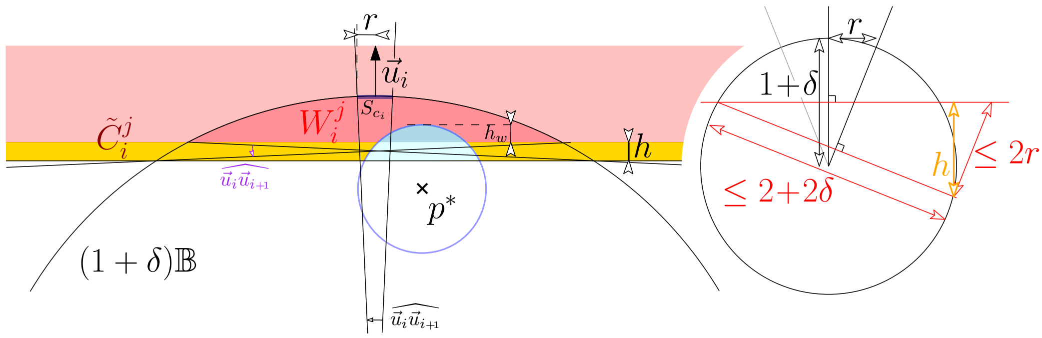
- R_i as the set of half-spaces whose inner normal is parallel to a vector from the origin to a point of $H_i \cap \partial(\rho\mathbb{B})$,
- W_i^j as the intersection of $\rho\mathbb{B}$ with a half-space parallel to H_i positioned so that $\mathbb{E}\left[W_i^j \cap P\right] = j$,
- C_i^j as the union of the half-spaces of R_i that do not contain W_i^j.

The proof of Lemma 3.4 readily implies that $\{(W_i^j, C_i^j)\}_{\substack{1\leq i\leq m\\ 1\leq j\leq \ell}}$ is a system of witnesses and collectors for the covering of the set of half-spaces in $\mathbb{R}^d$ by $R_1 \cup R_2 \cup \ldots \cup R_m$. To apply Theorem 2 (i) it remains to control the expected number of points of P in the collectors.

Claim 3.7. *For any perturbed point $p \in P$, with* $\mathrm{card}\, P \geq 2^{\frac{d+1}{2}}$,

$$\mathbb{P}\left[p \in C_i^j\right] = O\left(\frac{1}{n} + \mathbb{P}\left[p \in W_i^j\right]\right).$$

Proof. Let $p^* \in P^*$ and p its perturbed copy. We fix some indices $1 \leq i \leq m$ and $1 \leq j \leq \lceil \ln^2 n \rceil$ and write $w = \mathbb{P}\left[p \in W_i^j\right]$ and $c = \mathbb{P}\left[p \in C_i^j\right]$.



Refer to the figure above and let $\tilde{C}_i^j$ be the halfspace with normal $\vec{u}_i$ containing $C_i^j \cap (1+\delta)\mathbb{B}$ and with minimal intersection depth with $(1+\delta)\mathbb{B}$. Let h denote the difference of the intersection depth of the half space cutting out W_i^j and $\tilde{C}_i^j$ with $(1+\delta)\mathbb{B}$ and h_w denote the intersection depth at which W_i^j intersects $B(p^*, \delta)$. Observe that $\tilde{C}_i^j$ intersects $B(p^*, \delta)$ with depth at most $h_w + h$. Since the diameter of $\tilde{C}_i^j \cap P$ is at most $2 + 2\delta$, considerations on similar triangles show that $h \leq 2r$. If $h_w \leq 2r$ then we obtain the first part of the announced bound on c:

$$c \leq \frac{f(2r+h,\delta)}{f(2\delta,\delta)} \leq \frac{f(4\delta n^{-\frac{2}{d+1}},\delta)}{f(2\delta,\delta)} = \frac{f(4n^{-\frac{2}{d+1}},1)}{f(2,1)} = \frac{1}{f(2,1)} \int_0^{4n^{-\frac{2}{d+1}}} \left(2x - x^2\right)^{\frac{d-1}{2}} dx$$
$$\leq \frac{1}{f(2,1)} \int_0^{4n^{-\frac{2}{d+1}}} (2x)^{\frac{d-1}{2}} dx = O\left(\frac{1}{n}\right).$$

If $h_w > 2r$ then we can assume that $c > 2w$, as otherwise the claim holds trivially. In particular $h_w \leq \delta$. Since $h \leq 2r = 2n^{-\frac{2}{d+1}}$, the hypothesis $n \geq 2^{\frac{d+1}{2}}$ ensures $h < \delta$ and the depths of intersection of both W_i^j and $\tilde{C}_i^j$ are in the interval $[0, 2\delta]$. We then have

$$c \leq \frac{f(h_w+h,\delta)}{f(2\delta,\delta)} = \frac{f\left(\left(1+\frac{h}{h_w}\right)h_w, \delta\right)}{f(2\delta,\delta)} \leq \left(1+\frac{h}{h_w}\right)^{\frac{d+1}{2}} w \leq 2^{\frac{d+1}{2}} w,$$

the last inequality coming from $h_w > 2r \geq h$. □

Claim 3.7 implies that, for n bigger than the constant $2^{\frac{d+1}{2}}$,

$$\mathbb{E}\left[\mathrm{card}\left(C_i^j \cap P\right)\right] = O\left(1 + \mathbb{E}\left[\mathrm{card}\left(W_i^j \cap P\right)\right]\right) = O(j)$$

and Theorem 2 (i) provides the following bound:

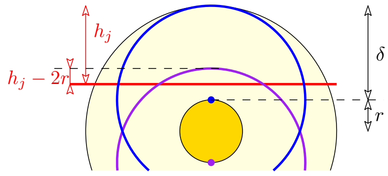
Proposition 4. $\mathcal{S}\left(n, \mathcal{U}_{\delta\mathbb{B}}\right) = O\left(n^{2\frac{d-1}{d+1}} + n^{2\frac{d-1}{d+1}}\delta^{-(d-1)}\right)$.

Large Perturbations. As $\delta \to \infty$ the bound of Proposition 4 does not tend to $\Theta\left(n^{\frac{d-1}{d+1}}\right)$, the average-case complexity bound. We thus complement it by a variation on the same system of witnesses and collectors better suited for the analysis of large perturbations.

Lemma 3.8. *For $\delta \geq 3n^{\frac{2}{d+1}}$ we have $\mathcal{S}(n, \mathcal{U}_{\delta\mathbb{B}}) = \Theta\left(n^{\frac{d-1}{d+1}}\right)$.*

Proof. We again assume, without loss of generality, that P^* is contained in $\mathbb{B}$ and use the construction of Section 3.2 with $\rho = 1+\delta$ and $h_j = (1+\delta)\left(\frac{j}{n}\right)^{\frac{2}{d+1}}$. By Claim 3.3 we have

$$m = \Theta\left((1+\delta)/h_1\right)^{\frac{d-1}{2}} = \Theta\left(n^{\frac{d-1}{d+1}}\right).$$



For any point p^* in $\mathbb{B}$, we have

$$\frac{f(h_j-2,\delta)}{\mathrm{vol}\,(\delta\mathbb{B})} \leq \mathbb{P}\left[p \in W_i^j\right] \leq \frac{f(h_j,\delta)}{\mathrm{vol}\,(\delta\mathbb{B})}$$

Since $h_j \geq 3$, Claims 3.1 and 3.2 imply that $\mathbb{P}\left[p \in W_i^j\right] = \Theta(\frac{j}{n})$. By Claims 3.1 and 3.5 we get $\mathbb{P}\left[p \in C_i^j\right] = \Theta(\frac{j}{n})$ as well, so Theorem 2 (i) applies. A constant fraction of the W_i^1 are pairwise disjoint, by Lemma 3.4, and $\mathbb{E}\left[\mathrm{card}\left(W_i^1 \cap P\right)\right] = \Omega(1)$. Using Claim 3.6, it follows that Theorem 2 (ii) also applies, and $\mathbb{E}\left[\mathrm{card}\,\mathcal{H}^{(k)}\right] = \Theta(m) = \Theta\left(n^{\frac{d-1}{d+1}}\right)$. □

Smoothed Number of Faces. Combining Proposition 4 and Lemma 3.8 we obtain the following upper bound on the smoothed number of faces of any dimension:

Theorem 5.

Range of δ	$\left[0, n^{\frac{2}{d+1}-\frac{1}{d-1}\lfloor\frac{d}{2}\rfloor}\right]$	$\left[n^{\frac{2}{d+1}-\frac{1}{d-1}\lfloor\frac{d}{2}\rfloor}, 1\right]$	$\left[1, 3n^{\frac{2}{d+1}}\right]$	$\left[3n^{\frac{2}{d+1}}, +\infty\right)$
$\mathcal{S}(n, \mathcal{U}_{\delta\mathbb{B}})$	$O\left(n^{\lfloor\frac{d}{2}\rfloor}\right)$	$O\left(n^{2\frac{d-1}{d+1}}\delta^{-(d-1)}\right)$	$O\left(n^{2\frac{d-1}{d+1}}\right)$	$\Theta\left(n^{\frac{d-1}{d+1}}\right)$

In dimension 2, a Euclidean noise of amplitude above $n^{-1/3}$ suffices to guarantee an expected sub-linear complexity. In dimension 3, the second bound is uninteresting as it exceeds the worst-case bound. In dimension d, a Euclidean noise of amplitude above $n^{-4/(d^2-1)}$ suffices to guarantee an expected sub-quadratic complexity.

Smoothed Number of Vertices. The bounds of Theorem 5 may be improved by a rescaling argument like the one used by Damerow and Sohler [6]: splitting the input into small cells and accounting separately for the contribution of each cell using a scaled version of Lemma 3.8. This only applies to the number of vertices, as a face of dimension 1 or more may involve perturbation of points coming from more than one cell.

Corollary 6. *For any d, $\mathbb{E}\left[\mathrm{card}\,\mathcal{H}^{(1)}\right] = O\left(n^{\frac{d-1}{d+1}} + \delta^{-\frac{2d}{d+1}}n^{1+2\frac{d-1}{(d+1)^2}}\right)$, and for $d=2$ we have:*

Range of δ	$[0, \frac{1}{\sqrt{n}}]$	$[\frac{1}{\sqrt{n}}, 1]$	$[1, n^{5/12}]$	$[n^{5/12}, n^{2/3}]$	$[n^{2/3}, +\infty]$
$\mathcal{S}(n, \mathcal{U}_{\delta\mathbb{B}})$	$O(n)$	$O\left(\delta^{-\frac{2}{3}}n^{\frac{2}{3}}\right)$	$O\left(n^{2/3}\right)$	$O\left(\delta^{-\frac{4}{3}}n^{\frac{11}{9}}\right)$	$O\left(n^{1/3}\right)$

Proof. We continue to assume that $P^* \subset \mathbb{B}$ and we cover $\mathbb{B}$ with $m' = \Theta(1 + r^{-d})$ disjoint cells of size $r = \frac{1}{3}\delta n^{-\frac{2}{d+1}}$. We partition P^* into $P_1^* \cup P_2^* \cup \ldots \cup P_{m'}^*$ by taking its intersection with each of the covering cells; we let P_i denote the perturbation of P_i^* and $n_i = \mathrm{card}\,P_i$. Every vertex of $\mathrm{CH}(P)$ is a vertex of some $\mathrm{CH}(P_i)$, and we can apply Lemma 3.8 to bound the number of vertices of $\mathrm{CH}(P_i)$ from above by $n_i^{\frac{d-1}{d+1}}$. If $m' > 1$, the sum is maximized with $\forall i$, $n_i = \frac{n}{m'}$ which bounds from above the number of vertices of $\mathrm{CH}(P)$ by

$$m'O\left(\left(\tfrac{n}{m'}\right)^{\frac{d-1}{d+1}}\right) = O\left(\left(\left(\delta n^{-\frac{2}{d+1}}\right)^{-d}\right)^{\frac{2}{d+1}} n^{\frac{d-1}{d+1}}\right) = O\left(\delta^{-\frac{2d}{d+1}} n^{\frac{4d}{(d+1)^2}+\frac{d-1}{d+1}}\right)$$
$$= O\left(\delta^{-\frac{2d}{d+1}} n^{1+2\frac{d-1}{(d+1)^2}}\right).$$

This proves the first statement. For the second statement, in two dimensions, we proceed differently in each regime:

$\delta \leq \frac{1}{\sqrt{n}}$. In this case, the worst-case bound is used.

$1 \leq \delta \leq n^{5/12}$. This case is solved using Proposition 4.

$n^{2/3} \leq \delta$. Here, Lemma 3.8 yields the result.

$n^{5/12} \leq \delta \leq n^{2/3}$. This case is handled through the first statement of the present corollary.

$\frac{1}{\sqrt{n}} \leq \delta \leq 1$. For the remaining case, we apply the same partitioning idea, but using Proposition 4 instead of Lemma 3.8 as an upper bound for one cell. Namely, considering a partitioning induced by covering cells of size δ, we get sets P_i^* whose convex hull has size $n_i^{\frac{2}{3}}$. Summing on the $\frac{1}{\delta^2}$ cells and using the concavity of $x \mapsto x^{\frac{2}{3}}$, we have

$$\sum_{i=1}^{O(\delta^{-2})} n_i^{\frac{2}{3}} = O\left(\delta^{-2}(\delta^2 n)^{\frac{2}{3}}\right) = O\left(\left(\tfrac{n}{\delta}\right)^{\frac{2}{3}}\right) \qquad \square$$

3.5 Lower Bound: Points in Convex Position

We finally analyze the expected complexity of Euclidean perturbations of some particular point configuration: points in convex position that are "nicely spread out"; more precisely, we take P^* to be an (ε, κ)-*sample* of a sphere with fixed radius, *ie.* a sample such that any ball of radius ε centered on the sphere contains between 1 and κ points of the sample.

Our motivation for studying this class of configurations is that they are natural candidates to realize the smoothed complexity of convex hulls in 2 and 3 dimensions and therefore provide an interesting lower bound. In light of Theorem 2 (ii), setting up the witnesses W_i^1 is enough to obtain a lower bound on the expected size of the convex hull; we give a complete analysis since at this stage it comes easily and makes it clear that the lower bound obtained by our choice of W_i^1 is sharp for these configurations.

Theorem 7. *Let $P^* = \{p_i^* : 1 \leq i \leq n\}$ be an $\left(\Theta\left(n^{\frac{1}{1-d}}\right), \Theta(1)\right)$-sample of the unit sphere in $\mathbb{R}^d$ and let $P = \{p_i = p_i^* + \eta_i\}$ where $\eta_1, \eta_2, \ldots, \eta_n$ are random variables chosen independently from $\mathcal{U}_{\delta\mathbb{B}}$. For any fixed k, $\mathbb{E}\left[\mathrm{card}\, \mathcal{H}^{(k)}\right]$ is*

Range of δ	$[0, n^{\frac{2}{1-d}}]$	$[n^{\frac{2}{1-d}}, 1]$	$[1, n^{\frac{2}{d+1}}]$	$[n^{\frac{2}{d+1}}, +\infty)$
$\mathbb{E}\left[\mathrm{card}\, \mathcal{H}^{(k)}\right]$	$\Theta(n)$	$\Theta\left(n^{\frac{d-1}{2d}} \delta^{\frac{1-d^2}{4d}}\right)$	$\Theta\left(n^{\frac{d-1}{2d}} \delta^{\frac{(1-d)^2}{4d}}\right)$	$\Theta\left(n^{\frac{d-1}{d+1}}\right)$

The last bound corresponds to the average-case behaviour which applies for δ sufficiently large, as follows from Lemma 3.8. We thus only have to analyze the range $\delta \leq n^{\frac{2}{d+1}}$. Note that the first bound merely reflects that a point remains extreme when the noise is small compared to the distance to the nearest hyperplane spanned by points in its vicinity, and that the bounds reveal that as the amplitude of the perturbation increases, the expected size of the convex hull does not vary monotonically (see Figures 2a and 2c): the lowest expected complexity is achieved by applying a noise of amplitude roughly the diameter of the initial configuration.

The following claim will be useful to position the witnesses and control the collectors.

Claim 3.9. *Under the assumptions of Theorem 7, let $j \leq \ln^2 n$, let H be a half-space such that $\mathbb{E}[\mathrm{card}(H \cap P)] = \Theta(j)$ and let h denote its depth of intersection with $(1+\delta)\mathbb{B}$.*

(i) If $\delta = O\left(\frac{j}{n}\right)^{\frac{2}{d-1}}$ then $h = \Theta\left(\left(\frac{j}{n}\right)^{\frac{2}{d-1}}\right)$

and if $\Omega\left(\frac{j}{n}\right)^{\frac{2}{d-1}} \leq \delta \leq O\left(n^{\frac{2}{d+1}}\right)$ then $h = \Theta\left(\left(\frac{j}{n}\right)^{\frac{1}{d}} \delta^{\frac{d+1}{2d}}\right)$.

(ii) If H' is a half-space that intersects $(1+\delta)\mathbb{B}$ with depth $9h$ then

$$\mathbb{E}[\mathrm{card}(H' \cap P)] = O(\mathbb{E}[\mathrm{card}(H \cap P)]).$$

Proof. The region $S \subseteq \partial\mathbb{B}$ in which we can center a ball of radius δ that intersects H is the intersection of $\partial\mathbb{B}$ with a half-space parallel to H and that intersects it with depth h; S is thus a spherical cap of $\mathbb{B}$ of radius $\sqrt{2h-h^2} = \Theta\left(\sqrt{h}\right)$ and $(d-1)$-dimensional area $\Theta\left(h^{\frac{d-1}{2}}\right)$. By the sampling condition in the definition of P^*, each ball of radius $n^{\frac{1}{1-d}}$ centered on $\partial\mathbb{B}$ contains $\Theta(1)$ points of P^*. In total there are thus $\Theta\left(nh^{\frac{d-1}{2}}\right)$ points $p^* \in P^*$ such that $(p^* + \delta\mathbb{B}) \cap H \neq \emptyset$. For the rest of this proof call these points *relevant*. How much a relevant point contributes to $\mathbb{E}[\mathrm{card}\, H \cap P]$ depends on how h compares to δ.

If $h \leq \delta$ then H intersects any ball $p^* + \delta\mathbb{B}$ with depth at most δ, and Claim 3.2 bounds the contribution of any relevant point p^* to $\mathbb{E}[\mathrm{card}\, H \cap P]$ by

$$\frac{\mathrm{vol}(H \cap (p^* + \delta\mathbb{B}))}{\mathrm{vol}(\delta\mathbb{B})} \leq \frac{f(h,\delta)}{f(2\delta,\delta)} = O\left(\frac{h^{\frac{d+1}{2}}\delta^{\frac{d-1}{2}}}{\delta^d}\right) = O\left(\left(\frac{h}{\delta}\right)^{\frac{d+1}{2}}\right).$$

Shrinking h by a factor two, we obtain that a constant fraction (depending only on d) of the relevant points contribute for at least $\frac{f(h/2,\delta)}{f(2\delta,\delta)} = \Omega\left(\left(\frac{h}{\delta}\right)^{\frac{d+1}{2}}\right)$ to $\mathbb{E}[\mathrm{card}(H \cap P)]$, hence

$$\Theta(j) = \mathbb{E}[\mathrm{card}(H \cap P)] = \Theta\left(nh^{\frac{d-1}{2}}\left(\frac{h}{\delta}\right)^{\frac{d+1}{2}}\right) = \Theta\left(n\delta^{-\frac{d+1}{2}}h^d\right)$$

and $h = \Theta\left(\left(\frac{j}{n}\right)^{\frac{1}{d}}\delta^{\frac{d+1}{2d}}\right)$. The condition $h \leq \delta$ thus amounts to $\delta = \Omega\left(\left(\frac{j}{n}\right)^{\frac{2}{d-1}}\right)$, giving the second regime.

If $h > \delta$ then a constant fraction of the relevant points p^* are such that H intersects $p^* + \delta\mathbb{B}$ with depth at least $\delta/2$, thus containing a constant fraction of each of these balls (and the rest of the relevant points contribute less). It follows that $\Theta(j) = \Theta\left(nh^{\frac{d-1}{2}}\right)$ and $h = \Theta\left(\left(\frac{j}{n}\right)^{\frac{2}{d-1}}\right)$. The condition $h > \delta$ amounts to $\delta = O\left(\left(\frac{j}{n}\right)^{\frac{2}{d-1}}\right)$, giving the first regime.

Observe that in either case, the number of points in $H \cap P$ depends polynomially on h. Thus, multiplying the depth by 9 multiplies the expected number of points by a constant (depending only on d) and statement (ii) follows. $\square$

Proof of Theorem 7. We use our construction of Section 3.2 with $\rho = 1+\delta$. We fix h_j such that each W_i^j contains $\Theta(j)$ points of P; the values of h_j are given by Claim 3.9(i). By Claim 3.5, C_i^j is contained in a half-space that intersects $(1+\delta)\mathbb{B}$ with depth at most $9h_j$. Claim 3.9(ii) thus ensures that

$$\mathbb{E}\left[\mathrm{card}\left(C_i^j \cap P\right)\right] = O\left(\mathbb{E}\left[\mathrm{card}\left(W_i^j \cap P\right)\right]\right) = O(j)$$

and we can apply Theorem 2 (i). Lemma 3.4 and Claim 3.6 further guarantee that we can apply Theorem 2 (ii). By Claim 3.3, $m = \Theta\left(\left(\frac{1+\delta}{h_1}\right)^{\frac{d-1}{2}}\right)$ and we have three regimes.
If $\delta = O\left(\frac{1}{n}\right)^{\frac{2}{d-1}}$ then Claim 3.9(i) yields $h_1 = \Theta\left(\left(\frac{1}{n}\right)^{\frac{2}{d-1}}\right)$ and

$$m = \Theta\left(\left(\frac{1+\delta}{h_1}\right)^{\frac{d-1}{2}}\right) = \Theta\left(\left(\frac{1}{\left(\frac{1}{n}\right)^{\frac{2}{d-1}}}\right)^{\frac{d-1}{2}}\right) = \Theta\left(n\right).$$

If $\Omega\left(\frac{j}{n}\right)^{\frac{2}{d-1}} \leq \delta \leq O\left(n^{\frac{2}{d+1}}\right)$ then Claim 3.9(i) yields $h_1 = \Theta\left(\left(\frac{1}{n}\right)^{\frac{1}{d}}\delta^{\frac{d+1}{2d}}\right)$. If $\delta \leq 1$ then

$$m = \Theta\left(\left(\frac{1+\delta}{h_1}\right)^{\frac{d-1}{2}}\right) = \Theta\left(\left(\frac{1}{\left(\frac{1}{n}\right)^{\frac{1}{d}}\delta^{\frac{d+1}{2d}}}\right)^{\frac{d-1}{2}}\right) = \Theta\left(n^{\frac{d-1}{2d}}\delta^{\frac{1-d^2}{4d}}\right)$$

and if $\delta \geq 1$ then

$$m = \Theta\left(\left(\frac{1+\delta}{h_1}\right)^{\frac{d-1}{2}}\right) = \Theta\left(\left(\frac{\delta}{\left(\frac{1}{n}\right)^{\frac{1}{d}}\delta^{\frac{d+1}{2d}}}\right)^{\frac{d-1}{2}}\right) = \Theta\left(n^{\frac{d-1}{2d}}\delta^{\frac{(1-d)^2}{4d}}\right)$$

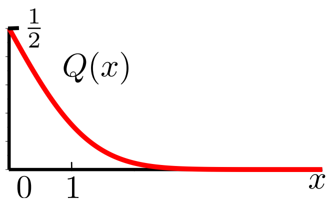
Up to multiplicative constants, the boundaries between the regimes can be set as in the statement of the Theorem. □

4 Gaussian Perturbation

The Gaussian model raises two difficulties compared to the Euclidean model: the computations are more technical and the fact that the perturbations have non-compact support requires to adapt the witness-collector construction. We expect some of the results to extend to arbitrary dimension *mutatis mutandis*, but for the sake of the presentation only spell out the analysis in the two-dimensional case.

4.1 Preliminaries

Recall that if $X \sim \mathcal{N}(\mu, \sigma^2)$ then for any $t \geq 0$ we have $\mathbb{P}\left[X \geq \mu + t\sigma\right] = Q(t)$, where Q is the tail probability of the standard Gaussian distribution:



$$\forall x \in \mathbb{R}, \quad Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt.$$

The solution to the functional equation $f(x)e^{f(x)} = x$ is called the Lambert function $\mathcal{W}_0$ [5, Equation (3.1)]. For $x \geq 0$ the definition of $\mathcal{W}_0(x)$ is non-ambiguous and satisfies

$$\forall x \geq 1.01, \qquad \mathcal{W}_0(x) = \Theta(\ln x). \tag{5}$$

This essentially follows from [5, Equations (4.6) and (4.9)]; note that the constant 1.01 is arbitrary and any constant strictly larger than 1 would do (the constants in the $\Theta()$ would change but we do not care). The following inequalities will be useful:

Lemma 4.1.

(i) For $x > 0$, $$\frac{e^{-\frac{x^2}{2}}}{x+\frac{1}{x}} < \sqrt{2\pi}Q(x) < \frac{e^{-\frac{x^2}{2}}}{x}.$$

(ii) For $x > 1/4$, $$Q\left(x + \frac{1}{x}\right) = \Theta(Q(x)).$$

(iii) *(a)* $$\sum_{i=0}^{n} e^{-i^2 x} = O\left(1 + \frac{1}{\sqrt{x}}\right)$$

(b) For any constant $\gamma > 0$, for $x > \frac{\gamma}{n^2}$, $$\sum_{i=0}^{n} e^{-i^2 x} = \Omega\left(\frac{1}{\sqrt{x}}\right)$$

(c) For any constant $\gamma > 0$, for $x < \frac{\gamma}{n^2}$, $$\sum_{i=0}^{n} e^{-i^2 x} = \Omega(n)$$

Proof. The upper bound of statement (i) comes from

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt < \int_x^\infty \frac{t}{\sqrt{2\pi}x} e^{-\frac{t^2}{2}} dt = \int_{\frac{x^2}{2}}^\infty \frac{e^{-t}}{x\sqrt{2\pi}} dt = \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}}$$

and the lower bound comes from the fact that

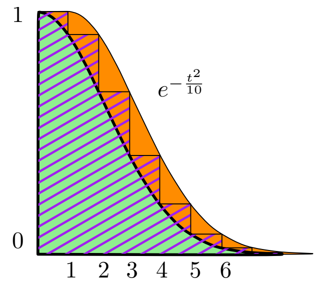
$$\left(1 + \frac{1}{x^2}\right) Q(x) = \int_x^\infty \left(1 + \frac{1}{x^2}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt > \int_x^\infty \left(1 + \frac{1}{t^2}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}}.$$

Now, for statement (ii), we have $Q(x) \geq Q(x+\frac{1}{x})$ since Q is a decreasing function. Moreover, from statement (i) we have

$$\begin{aligned}
\sqrt{2\pi}Q(x+\tfrac{1}{x}) &> \frac{x+\frac{1}{x}}{1+\left(x+\frac{1}{x}\right)^2} e^{-\frac{\left(x+\frac{1}{x}\right)^2}{2}} \\
&= \left(\frac{x^4+x^2}{x^4+3x^2+1} e^{-1-\frac{1}{2x^2}}\right) \left(\frac{e^{-\frac{x^2}{2}}}{x}\right) > \left(\frac{x^4+x^2}{x^4+3x^2+1} e^{-1-\frac{1}{2x^2}}\right) \sqrt{2\pi}Q(x)
\end{aligned}$$

Statement (ii) then follows from noting that the image of $[1/4, +\infty)$ under the function $x \mapsto \frac{x^4+x^2}{x^4+3x^2+1} e^{-1-\frac{1}{2x^2}}$ is contained in some closed interval of $(0, +\infty)$.

The proof of Statements (iii-a) and (iii-b) follows from a standard comparison between series and integrals:



if $x > \frac{\gamma}{n^2}$,

$$\sum_{i=0}^{n} e^{-i^2 x} \geq \int_0^{n+1} e^{-t^2 x} dt \geq \int_0^{n\sqrt{x}} e^{-u^2} \frac{du}{\sqrt{x}}$$
$$\geq \frac{\int_0^{\sqrt{\gamma}} e^{-u^2} du}{\sqrt{x}} \geq \Omega\left(\frac{1}{\sqrt{x}}\right)$$

and for any $x > 0$,

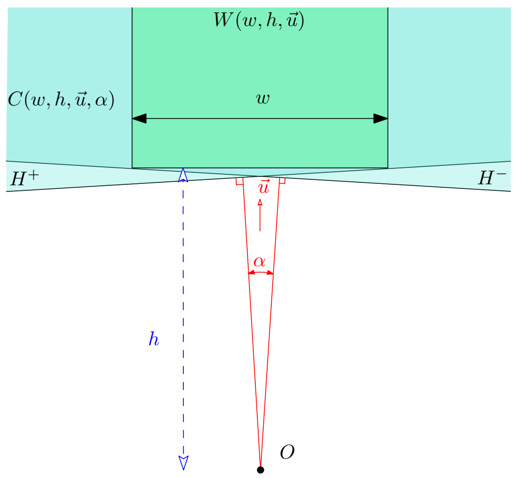
$$\sum_{i=0}^{n} e^{-i^2 x} \leq 1 + \int_0^{n} e^{-t^2 x} dt \quad \leq 1 + \int_0^{n\sqrt{x}} e^{-u^2} \frac{du}{\sqrt{x}} \leq 1 + \int_0^{\infty} e^{-u^2} \frac{du}{\sqrt{x}} \quad = O\left(1 + \frac{1}{\sqrt{x}}\right).$$

Statement (iii-c) is trivial since, when $x < \frac{\gamma}{n^2}$, $\sum_{i=0}^{n} e^{-i^2 x} \geq n \cdot e^{-\gamma} = \Omega(n)$. ☐

4.2 Witness-Collector Construction

One Witness-Collector Pair. The pairs witness-collectors that we use to analyze Gaussian perturbations are based on the following basic construction. Let w, h and α be positive reals and $\vec{u}$ some vector in the plane.

- We define $R(\vec{u}, \alpha)$ as the set of half-planes whose inner normal makes an angle at most $\frac{\alpha}{2}$ with $\vec{u}$.
- We define $W(w, h, \vec{u})$ as the semi-infinite half strip with axis of symmetry $O + \mathbb{R}\vec{u}$, with width w and distance h to the origin. To save breath we define the *height* of a semi-infinite half strip as its distance to the origin – so $W(w, h, \vec{u})$ has height h.
- We define $C(w, h, \vec{u}, \alpha)$ as the union of the half-planes in $R(\vec{u}, \alpha)$ that do not contain $W(w, h, \vec{u})$.



The following more explicit description of $C(w, h, \vec{u}, \alpha)$ will be convenient:

Claim 4.2. *$C(w, h, \vec{u}, \alpha) = H^- \cup H^+$ where H^- and H^+ are the half-planes whose inner normals make an angle of $\pm\frac{\alpha}{2}$ with $\vec{u}$, that contain $W(w, h, \vec{u})$ and have one of the corners of $W(w, h, \vec{u})$ on their boundary.*

Proof. This follows from observing that any halfplane through $\partial H^+ \cap \partial H^-$ and contained in $C(w, h, \vec{u}, \alpha)$ also contains $W(w, h, \vec{u})$. ☐

This construction has the following properties:

(a') Any halfplane whose inner normal makes an angle at most $\frac{\alpha}{2}$ with $\vec{u}$ contains $W(w, h, \vec{u})$ or is contained in $C(w, h, \vec{u}, \alpha)$.

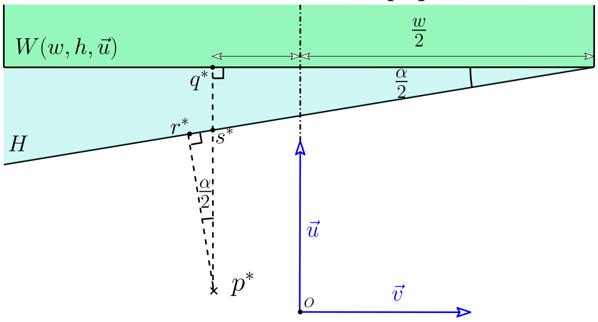
(b') If $h_j > h_{j+1}$ and $w_j < w_{j+1}$ then $W(w_j, h_j, \vec{u}) \subseteq W(w_{j+1}, h_{j+1}, \vec{u})$.

(c') $W(w,h,\vec{u}) \subseteq C(w,h,\vec{u},\alpha)$.
Families of pairs $(W(w,h,\vec{u}), C(w,h,\vec{u},\alpha))$ therefore combine easily into systems of witnesses and collectors. We will control the expected number of points in a witness by setting w and h adequately and tune α accordingly thanks to the next fact. We say that a point p^* is *in the slab of* $W(w,h,\vec{u})$ if the ray $p^* + \mathbb{R}^+\vec{u}$ intersects $W(w,h,\vec{u})$.

Claim 4.3. *Let $\vec{u}$ be arbitrary and let $\vec{v}$ denote a unit vector orthogonal to $\vec{u}$. If $p^* \in \mathbb{R}^2$ is in the slab of $W(w,h,\vec{u})$ and outside the interior of $C(w,h,\vec{u},\alpha)$ then*

$$d(p^*, C(w,h,\vec{u},\alpha)) = d(p^*, W(w,h,\vec{u}))\cos\frac{\alpha}{2} - \left(\frac{w}{2} + |\overrightarrow{Op^*}\cdot\vec{v}|\right)\sin\frac{\alpha}{2}.$$

Proof. Let H denote the half-plane contained in $C(w,h,\vec{u},\alpha)$ and whose distance to p^* is minimal. Let q^* and r^* denote respectively the orthogonal projections of p^* on $W(w,h,\vec{u})$ and $C(w,h,\vec{u},\alpha)$. Let s^* denote the intersection of p^*q^* with the boundary of H.



The assumptions ensure that

$$|p^*q^*| = d(p^*, W(w,h,\vec{u}))$$

and

$$|p^*r^*| = d(p^*, C(w,h,\vec{u},\alpha)).$$

With $\vec{v} \in \mathbb{S}^1, \vec{v} \perp \vec{u}$, we have

$$|q^*s^*| = \left(\frac{w}{2} + |\overrightarrow{Op^*}\cdot\vec{v}|\right)\tan\frac{\alpha}{2} \quad \text{and} \quad |p^*r^*| = |p^*s^*|\cos\frac{\alpha}{2} = (|p^*q^*| - |q^*s^*|)\cos\frac{\alpha}{2}$$

and the statement follows. □

System of Witnesses and Collectors. Our construction is parameterized by some positive real α and two sequences of positive reals $h_1 > h_2 > \cdots > h_\ell$ and $w_1 \leq w_2 \leq \cdots \leq w_\ell$. We choose an inclusion-minimal cover of $\partial\mathbb{B}$ by half-planes $H_1, H_2, \ldots, H_m$ each intersecting $\partial\mathbb{B}$ in a circular arc of angle α; we let $\vec{u_i}$ denote the center of $H_i \cap \partial\mathbb{B}$ and note that $m = \Theta\left(\frac{1}{\alpha}\right)$. We define R_i as the set of half-planes whose inner normal is parallel to a vector from the origin to a point of $H_i \cap \partial\mathbb{B}$ and let

$$W_i^j = W(w_j, h_j, \vec{u}_i) \quad \text{and} \quad C_i^j = C(w_j, h_j, \vec{u}_i, \alpha).$$

Lemma 4.4. *$R_1 \cup R_2 \cup \ldots \cup R_m$ covers the set of half-planes and $\{(W_i^j, C_i^j)\}_{\substack{1\leq i\leq m\\1\leq j\leq \ell}}$ is a system of witnesses and collectors for that covering. Moreover, some $\Omega\left(\frac{h_1}{w_1}\right)$ of the W_i^1 are pairwise disjoint.*

Proof. The definition readily ensures that the union of the R_i is the set of all half-planes and that Condition (a) holds. The monotonicity of the h_i and the w_i implies that Condition (b) is also satisfied. Claim 4.2 implies that each W_i^j is contained in the corresponding C_i^j, so Condition (c) holds. Each W_i^1 is contained in a wedge with apex the origin and opening angle $\Theta\left(\frac{w_1}{h_1}\right)$. Some $\Omega\left(\frac{h_1}{w_1}\right)$ of these wedges are disjoint (except in the origin), so the corresponding W_i^1's are pairwise disjoint. □

4.3 Warm-up: Gaussian Polygons Made Easy

To illustrate our construction, we revisit the classical problem of computing the expected number of faces of the convex hull from a Gaussian distribution:

Theorem 8 (Rényi and Sulanke [17])**.** *Let $P = \{p_1, p_2, \dots, p_n\}$ be a set of random points chosen independently from $\mathcal{N}(0, I_2)$. The expected number of vertices of the convex hull of P is $\Theta\left(\sqrt{\ln n}\right)$.*

$$\alpha = \frac{1}{h_1} = \Theta\left(\frac{1}{\sqrt{\ln n}}\right) \qquad w_j = 2 \qquad h_j = \sqrt{\mathcal{W}_0\left(\frac{n^2}{j^2}\right)}$$

Proof. We use the construction of Section 4.2 with $\ell = \ln^2 n$ and the values of α, w_j and h_j set as specified on the right. Lemma 4.4 ensures that we obtain a system of witnesses and collectors, so it only remains to analyze the expected number of points in W_i^j and C_i^j. We complete each vector $\vec{u}_i$ into a direct, orthonormal frame $(O, \vec{v}_i, \vec{u}_i)$; in that frame, the coordinates of any point $p \in P$ write (x_i, y_i) where x_i, y_i are independent random variables chosen from $\mathcal{N}(0,1)$. The probability for p to be in W_i^j therefore writes

$$\mathbb{P}\left[p \in W_i^j\right] = \mathbb{P}\left[y_i > h_j\right]\mathbb{P}\left[|x_i| < 1\right] = \Theta\left(Q(h_j)\right).$$

Lemma 4.1 (i) yields $Q(x) = \Theta\left(\frac{1}{x}e^{-\frac{x^2}{2}}\right)$ for $x > 1$ so, since $j \leq \ln^2 n$,

$$Q(h_j) = \Theta\left(\frac{e^{-\frac{1}{2}\mathcal{W}_0\left(\frac{n^2}{j^2}\right)}}{\sqrt{\mathcal{W}_0\left(\frac{n^2}{j^2}\right)}}\right) = \Theta\left(\frac{1}{\sqrt{\mathcal{W}_0\left(\frac{n^2}{j^2}\right)e^{\mathcal{W}_0\left(\frac{n^2}{j^2}\right)}}}\right) = \Theta\left(\frac{j}{n}\right) \tag{6}$$

and $\mathbb{E}\left[\operatorname{card}\left(W_i^j \cap P\right)\right] = n\Theta\left(Q(h_j)\right) = \Theta(j)$. Since for $n \geq 3$, $\alpha < \frac{1}{\sqrt{\mathcal{W}_0(3^2)}} < \frac{\pi}{4}$, $\tan\frac{\alpha}{2} < 0.5$ and $\frac{2h_j}{w_j} \geq \frac{2h_\ell}{w_j} = \sqrt{\mathcal{W}_0\left(\frac{n^2}{\ln^4 n}\right)} \geq 1$. This means that $\tan\frac{\alpha}{2} < \frac{2h_j}{w_j}$, so the origin is not in C_i^j. By Claims 4.2 and 4.3, C_i^j is contained in the union of two half-planes with height $\tilde{h_j} = h_j\cos\frac{\alpha}{2} - \sin\frac{\alpha}{2} = h_j - O(\alpha)$. Thus,

$$\mathbb{E}\left[\operatorname{card}\left(C_i^j \cap P\right)\right] \leq 2nQ\left(\tilde{h_j}\right) = 2\left(nQ\left(h_j\right)\right)\left(\frac{Q\left(\tilde{h_j}\right)}{Q\left(h_j\right)}\right)$$

We already observed that $nQ\left(h_j\right) = \Theta(j)$. By Lemma 4.1 (i) we have

$$\frac{Q\left(\tilde{h_j}\right)}{Q\left(h_j\right)} = \frac{e^{-\frac{1}{2}\left(\tilde{h_j}^2 - h_j^2\right)}h_j}{\tilde{h_j}} = \frac{h_j}{\tilde{h_j}}e^{O\left(h_j\alpha + \alpha^2\right)}$$

and with Equation (5) we finally obtain

$$\mathbb{E}\left[\operatorname{card}\left(C_i^j \cap P\right)\right] = O\left(je^{O(h_j\alpha)}\right) = O\left(je^{O\left(\sqrt{1-\frac{\ln j}{\ln n}}\right)}\right) = O(j),$$

and Property (b') holds. Theorem 2 (i) then yields that $\mathbb{E}\left[\operatorname{card}\mathcal{H}^{(1)}\right] = O(m) = O(\sqrt{\ln n})$.

Let H_i denote the halfplane with same height and inner normal as W_i^1 and $p_{\vec{u}_i}$ be the point of P extremal in direction $\vec{u}_i$. By construction $p_{\vec{u}_i}$ belongs to $\mathcal{H}^{(1)}$, thus

$$\mathbb{E}\left[\text{card}\left(W_i^1 \cap \mathcal{H}^{(1)}\right)\right] \geq \mathbb{P}\left[p_{\vec{u}_i} \in W_i^1\right] = \mathbb{P}\left[p_{\vec{u}_i} \in H_i\right] \mathbb{P}\left[p_{\vec{u}_i} \in W_i^1 \mid p_{\vec{u}_i} \in H_i\right]$$

We have

$$\mathbb{P}\left[p_{\vec{u}_i} \in H_i\right] = \mathbb{P}\left[P \cap H_i \neq \emptyset\right] \geq \mathbb{P}\left[P \cap W_i^1 \neq \emptyset\right] \geq 1 - \tfrac{1}{e} > \frac{1}{2}$$

by Lemma 2.1 (a). Gaussian noise perturbs points independently in directions x and y of frame $(O, \vec{v}_i, \vec{u}_i)$. The choice of $p_{\vec{u}_i}$ in P depends only on the y perturbation, thus knowing that $p_{\vec{u}_i} \in H_i$, deciding if it is in W_i^1 or in $H_i \setminus W_i^1$ depends only on the coordinate along direction v_i, thus

$$\begin{aligned}
\mathbb{P}\left[p_{\vec{u}_i} \in W_i^1 \mid p_{\vec{u}_i} \in H_i\right] &= \sum_{p \in H_i \cap P} \mathbb{P}\left[p \in W_i^1 \mid p_{\vec{u}_i} = p\right] \mathbb{P}\left[p_{\vec{u}_i} = p \mid p_{\vec{u}_i} \in H_i\right] \\
&= \sum_{p \in H_i \cap P} \mathbb{P}\left[|x_p| \leq 1\right] \mathbb{P}\left[p_{\vec{u}_i} = p \mid p_{\vec{u}_i} \in H_i\right] \\
&= \mathbb{P}\left[|x_{p_1}| \leq 1\right] \sum_{p \in H_i \cap P} \mathbb{P}\left[p_{\vec{u}_i} = p \mid p_{\vec{u}_i} \in H_i\right] \\
&= \mathbb{P}\left[|x_{p_1}| \leq 1\right] \quad = 1 - 2Q(1) > \frac{1}{2}
\end{aligned} \tag{7}$$

Together we get that Lemma 4.4 ensures that we can also apply Theorem 2 (ii) and get that $\mathbb{E}\left[\text{card}\,\mathcal{H}^{(1)}\right] = \Omega(\sqrt{\ln n})$ as well. □

4.4 Upper Bound on the Smoothed Complexity

As in the Euclidean case, for large Gaussian perturbation the smooth complexity is identical to the i.i.d. case. It is possible to obtain a Gaussian analogue of Lemma 3.8 and apply the rescaling argument to get a smooth complexity for any scale of perturbation (see Appendix A). This bound is, however, worse than what we can obtain by a charging argument in the spirit of Claim 3.7 and Proposition 4.

Theorem 9. $\mathcal{S}(n, \mathcal{N}(0, \sigma^2 I_2)) = O\left(\frac{\sqrt{\ln n}}{\sigma} + \sqrt{\ln n}\right)$.

Let P^* be some point set of diameter at most 1 in the plane and, without loss of generality, assume that P^* is contained in $\mathbb{B}$. We use a system of witnesses and collectors similar to the one presented in Section 4.2 with $\ell = \ln^2 n$. As in the Euclidean case, a key difference is that the depth of intersection of each witness W_i^j depends on i, and is adapted to P^*. Specifically, we set w and α to the values on the right, choose the $\vec{u}_i$ regularly spaced on $\mathbb{S}^1$ with $\widehat{\vec{u}_i\vec{u}_{i+1}} = \Theta(\alpha)$. We then define $R_i = R(\vec{u}_i, \alpha)$, $W_i^j = W(h_i^j, w, \vec{u}_i)$ and $C_i^j = C(h_i^j, w, \vec{u}_i, \alpha)$ where h_i^j depends on P^* and is tuned so that the expected number of points in the witnesses are what they should be.

$$w = 2(1+\sigma)$$

$$\alpha = \frac{\sqrt{(2+\sigma)^2 + \frac{2\sqrt{2}\sigma}{\sqrt{\ln n}} + 2\sqrt{2}\sigma^2} - (2+\sigma)}{(1+\sigma\sqrt{\ln n})}$$

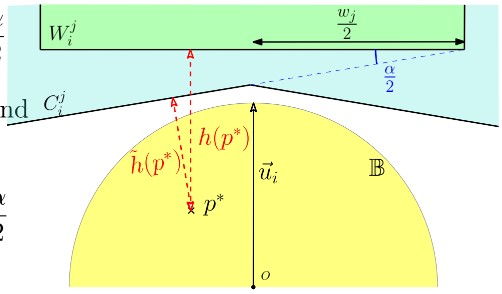
$$h_i^j \text{ s. t. } \mathbb{E}\left[\text{card}\left(P \cap W(h_i^j, w, \vec{u}_i)\right)\right] = j$$

We first relate the distances from a point p^* of P^* to a witness W_i^j and a collector C_i^j:

Claim 4.5. *If $p^* \in \mathbb{R}^2$ is in the slab of W_i^j and outside the interior of C_i^j, then*

$$d(p^*, W_i^j) - d(p^*, C_i^j) \le \frac{\sigma}{\sqrt{2 \ln n}}.$$

Proof. Let $h(p^*)$ and $\tilde{h}(p^*)$ denote, respectively, $d(p^*, W_i^j)$ and $d(p^*, C_i^j)$. By Claim 4.3,



$$\tilde{h}(p^*) = h(p^*) \cos \frac{\alpha}{2} - \left(\frac{w}{2} + |\overrightarrow{Op^*} \cdot \vec{v}_i| \right) \sin \frac{\alpha}{2}$$

and with $1 - \cos x \le \frac{x^2}{2}$, $\sin |x| < |x|$, $\frac{w}{2} = 1 + \sigma$, and $|\overrightarrow{Op^*} \cdot \vec{v}_i| \le 1$, this becomes

$$\begin{aligned} h(p^*) - \tilde{h}(p^*) &\le h(p^*) - h(p^*) \cos \frac{\alpha}{2} + (2+\sigma) \sin \frac{\alpha}{2} \\ &\le h(p^*) \frac{\alpha^2}{8} + (2+\sigma) \frac{\alpha}{2} \end{aligned}$$

The distance from p^* to $W_i^j = W(h_i^j, w, \vec{u}_i)$ is maximized when p^* is located at the point of $\partial\mathbb{B}$ with outer normal $-\vec{u}_i$ and all other points of P^* are at the symmetric position, at the point of $\partial\mathbb{B}$ with normal $\vec{u}_i$. The same argument as in Equation (6) and the observation that $\ln(x) > \mathcal{W}_0(x)$ for $x \ge 3$ yield the upper bound

$$h(p^*) \le 2 + \sigma \sqrt{\mathcal{W}_0(n^2)} \le 2 \left(1 + \sigma \sqrt{\ln n} \right).$$

Injecting this in the above inequality we get

$$h(p^*) - \tilde{h}(p^*) \le \left(1 + \sigma \sqrt{\ln n} \right) \frac{\alpha^2}{4} + (2+\sigma) \frac{\alpha}{2}$$

The polynomial

$$P(\alpha) = \left(\left(1 + \sigma \sqrt{\ln n} \right) \frac{\alpha^2}{4} + (2+\sigma) \frac{\alpha}{2} - \frac{\sigma}{\sqrt{2 \ln n}} \right)$$

can be checked to be negative for

$$0 \le \alpha \le \frac{\sqrt{(2+\sigma)^2 + \frac{2\sqrt{2}\sigma}{\sqrt{\ln n}} + 2\sqrt{2}\sigma^2} - (2+\sigma)}{(1 + \sigma\sqrt{\ln n})},$$

and that concludes the proof. □

The distance from a point p^* to W_i^j and C_i^j determines the probability that the perturbation of p^* belongs to either of these sets.

Claim 4.6. $\mathbb{P}\left[p \in W_i^j\right] = \Theta\left(Q\left(\frac{d(p^*, W_i^j)}{\sigma}\right)\right)$ *and* $\mathbb{P}\left[p \in C_i^j\right] = O\left(Q\left(\frac{d(p^*, C_i^j)}{\sigma}\right)\right)$.

Proof. A perturbed point p is in W_i^j if it satisfies two conditions: (α) its displacement from p^* along $\vec{u}_i$ should be greater than $d(p^*, W_i^j)$, and (β) its displacement in the orthogonal direction is in the slab of width w_j. The conditions are independent, (α) is true with probability $Q\left(\frac{d(p^*, W_i^j)}{\sigma}\right)$ and (β) is true with constant probability since $w = 2 + 2\sigma$ ensures that the allowed orthogonal displacement for p^* is larger than σ. The statement for W_i^j follows. As for the collectors, the probability that a perturbed point p is in C_i^j is bounded from above by the sum of the probabilities to be in H^+ and to be in H^-, which are both $Q\left(\frac{d(p^*, C_i^j)}{\sigma}\right)$. □

Combining the two previous claims we now get that witness and collector get, on average, essentially the same number of points.

Claim 4.7. *For any $p^* \in P^*$, we have* $\mathbb{P}\left[p \in C_i^j\right] = O\left(\frac{1}{n} + \mathbb{P}\left[p \in W_i^j\right]\right)$.

Proof. Let $h(p^*)$ and $\tilde{h}(p^*)$ denote, respectively, $d(p^*, W_i^j)$ and $d(p^*, C_i^j)$. Since $w \geq 2$ any point in P^* is in the slab of W_i^j.

First assume that p^* is not in C_i^j. Claim 4.5 then ensures that $\tilde{h}(p^*) \geq h(p^*) - \frac{\sigma}{\sqrt{2\ln n}}$. If $h(p^*) > \sigma\sqrt{2\ln n} + \frac{\sigma}{\sqrt{2\ln n}}$ then by Claim 4.6, Lemma 4.1 (i), and the fact that Q is decreasing, we have

$$\begin{aligned}\mathbb{P}\left[p \in C_i^j\right] = O\left(Q\left(\frac{\tilde{h}(p^*)}{\sigma}\right)\right) &= O\left(Q\left(\frac{h(p^*)}{\sigma} - \frac{1}{\sqrt{2\ln n}}\right)\right)\\ &= O\left(Q\left(\sqrt{2\ln n} + \frac{1}{\sqrt{2\ln n}} - \frac{1}{\sqrt{2\ln n}}\right)\right)\\ &= O\left(Q\left(\sqrt{2\ln n}\right)\right) = O\left(\frac{1}{n}\right)\end{aligned}$$

and the statement follows. If $h(p^*) \leq \sigma\sqrt{2\ln n} + \frac{\sigma}{\sqrt{2\ln n}}$ then we have

$$\mathbb{P}\left[p \in C_i^j\right] = O\left(Q\left(\frac{\tilde{h}(p^*)}{\sigma}\right)\right) = O\left(Q\left(\frac{h(p^*)}{\sigma} - \frac{1}{\sqrt{2\ln n}}\right)\right)$$

If $h(p^*) \leq \frac{\sigma}{4} + \frac{\sigma}{\sqrt{2\ln n}} \leq \sigma\left(\frac{1}{4} + \frac{1}{\sqrt{2\ln 3}}\right)$ then $h(p^*)$ is bounded from above by 2σ and

$$\mathbb{P}\left[p \in W_i^j\right] = \Omega\left(Q\left(2\right)\right) = \Omega(1).$$

Then, $\mathbb{P}\left[p \in C_i^j\right] \leq 1 = O\left(\mathbb{P}\left[p \in W_i^j\right]\right)$ and the statement also holds. Thus, we can suppose that $h(p^*) \geq \frac{\sigma}{4} + \frac{\sigma}{\sqrt{2\ln n}}$ and use Lemma 4.1 (ii) to get:

$$\mathbb{P}\left[p \in C_i^j\right] = O\left(Q\left(\frac{h(p^*)}{\sigma} - \frac{1}{\sqrt{2\ln n}} + \frac{1}{\frac{h(p^*)}{\sigma} - \frac{1}{\sqrt{2\ln n}}}\right)\right).$$

Since

$$\frac{1}{\frac{h(p^*)}{\sigma} - \frac{1}{\sqrt{2\ln n}}} \geq \frac{1}{\sqrt{2\ln n} + \frac{1}{\sqrt{2\ln n}} - \frac{1}{\sqrt{2\ln n}}} = \frac{1}{\sqrt{2\ln n}}$$

we get

$$\mathbb{P}\left[p \in C_i^j\right] = O\left(Q\left(\frac{h(p^*)}{\sigma}\right)\right) = O\left(\mathbb{P}\left[p \in W_i^j\right]\right)$$

and the statement also holds.

Finally assume that $p^* \in C_i^j$. In such a case Claims 4.3 and 4.5 do not apply directly, but we have $\frac{1}{2} \leq \mathbb{P}\left[p \in C_i^j\right] \leq 1$ so we have to argue that $\mathbb{P}\left[p \in W_i^j\right] = \Omega(1)$. Let us move from p^* in the direction $-\vec{u}_i$ until we reach some point $\bar{p}^*$ on the boundary of C_i^j; observe that $\mathbb{P}\left[\bar{p}^* + \eta \in C_i^j\right] \geq \frac{1}{2}$ where $\eta \sim \mathcal{N}(0, \sigma^2 I_2)$. Now $\bar{p}^*$ satisfies the hypotheses of Claim 4.5 and the above analysis implies that $\mathbb{P}\left[\bar{p}^* + \eta \in W_i^j\right] = \Omega\left(\mathbb{P}\left[\bar{p}^* + \eta \in C_i^j\right]\right) = \Omega(1)$. Moving from p^* to $\bar{p}^*$ only increased the distance to W_i^j, so we also have $\mathbb{P}\left[p \in W_i^j\right] \geq \mathbb{P}\left[\bar{p}^* + \eta \in W_i^j\right] = \Omega(1)$. □

We now have all the ingredient to prove our upper bound on the smoothed complexity under Gaussian noise.

Proof of Theorem 9. We set up our witnesses and collectors as described above. Since the parameter w is fixed and each sequence $\{h_i^j\}_j$ is decreasing, Lemma 4.4 yields that $\{(W_i^j, C_i^j)\}_{\substack{1\leq i\leq m\\ 1\leq j\leq \ell}}$ is a system of witnesses and collectors for the covering $R_1 \cup R_2 \cup \ldots \cup R_m$ of the set of half-planes. Each parameter h_i^j is set so that $\mathbb{E}\left[\text{card}\left(W_i^j \cap P\right)\right] = j$ and Claim 4.7 implies that $\mathbb{E}\left[\text{card}\left(C_i^j \cap P\right)\right] = O(j)$. Theorem 2 (i) thus implies that

$$\mathcal{S}(n, \mathcal{N}(0, \sigma^2 I_2)) = O\left(\frac{1}{\alpha}\right) = O\left(\frac{\left(1 + \sigma\sqrt{\ln n}\right)}{(2+\sigma)\left(\sqrt{1 + \frac{2\sqrt{2}}{(2+\sigma)^2}\left(\sigma^2 + \frac{\sigma}{\sqrt{\ln n}}\right)} - 1\right)}\right). \tag{8}$$

If $\sigma \leq \frac{1}{\sqrt{\ln n}}$ then Equation (8) simplifies into

$$\mathcal{S}(n, \mathcal{N}(0, \sigma^2 I_2)) = O\left(\frac{1}{\sqrt{1 + \frac{1}{(1+\sigma)^2}\left(\sigma^2 + \frac{\sigma}{\sqrt{\ln n}}\right)} - 1}\right).$$

Notice that in this case, $\frac{1}{(1+\sigma)^2}\left(\sigma^2 + \frac{\sigma}{\sqrt{\ln n}}\right)$ is bounded by some constant C and since for $0 < x < C$, $\sqrt{1+x} - 1 = \Theta(x)$,

$$\mathcal{S}(n, \mathcal{N}(0, \sigma^2 I_2)) = O\left(\frac{1}{\frac{1}{(1+\sigma)^2}\left(\sigma^2 + \frac{\sigma}{\sqrt{\ln n}}\right)}\right) = O\left(\frac{\sqrt{\ln n}}{\sigma}\right).$$

If $\frac{1}{\sqrt{\ln n}} \leq \sigma \leq 1$ then Equation (8) simplifies into

$$\mathcal{S}(n, \mathcal{N}(0, \sigma^2 I_2)) = O\left(\frac{\sigma\sqrt{\ln n}}{\sqrt{1 + \Theta(\sigma^2)} - 1}\right) = O\left(\frac{\sigma\sqrt{\ln n}}{\sigma^2}\right) = O\left(\frac{\sqrt{\ln n}}{\sigma}\right)$$

If $1 \leq \sigma$ then Equation (8) simplifies into

$$\mathcal{S}(n, \mathcal{N}(0, \sigma^2 I_2)) = O\left(\frac{\sigma\sqrt{\ln n}}{\sigma\Theta(1)}\right) = O(\sqrt{\ln n}).$$

In each case we therefore have $\mathcal{S}(n, \mathcal{N}(0, \sigma^2 I_2)) = O\left(\frac{\sqrt{\ln n}}{\sigma} + \sqrt{\ln n}\right)$. □

4.5 Lower Bound on Smoothed Complexity: Points in Convex Position

We finally investigate lower bounds on the smoothed complexity by analyzing the size of the convex hull of a Gaussian perturbation of points in convex position, as in Section 3.5.

Theorem 10. *Let $P^* = \{p_i^*, 1 \le i \le n\}$ be the set of vertices of a regular n-gon of radius 1 in $\mathbb{R}^2$. Let $P = \{p_i = p_i^* + \eta_i\}$ where $\eta_1, \eta_2, \ldots, \eta_n$ are random vectors in $\mathbb{R}^2$ chosen independently from $\mathcal{N}(0, \sigma^2 I_2)$. The expected number of vertices of the convex hull of P is*

Range of σ	$\left[0, \frac{1}{n^2}\right]$	$\left[\frac{1}{n^2}, \frac{1}{\sqrt{\ln n}}\right]$	$\left[\frac{1}{\sqrt{\ln n}}, +\infty\right)$
$\mathbb{E}\left[\operatorname{card} \mathcal{H}^{(1)}\right]$	$\Omega(n)$	$\Omega\left(\frac{\sqrt[4]{\ln(n\sqrt{\sigma})}}{\sqrt{\sigma}}\right)$	$\Omega\left(\sqrt{\ln n}\right)$

We use the witness-collector construction presented in Section 4.2. We only care about the lower-bound, so, shortening W_i^1 into W_i, we need only define one level of witnesses $\{W_i\}_{1 \le i \le m}$ to apply Theorem 2 (ii).

Parameters Setting. We set h_1 and w_1 depending on σ and n as summarized below. We let $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ denote a family of vectors in $\mathbb{S}^1$ such that $\vec{u}_i$ is aligned with $p^*_{\left\lfloor \frac{2\pi i}{m} \right\rfloor}$, so these vectors are more or less equally spaced on $\mathbb{S}^1$. The witnesses are defined as $W_i = W(w_1, h_1, \vec{u}_i)$. We choose m maximal so that the $\{W_i\}$ are pairwise disjoint; Lemma 4.4 ensures that we can set $m = \Omega\left(\min\left(n, \frac{h_1}{w_1}\right)\right)$.

	$0 \le \sigma < \frac{2}{n^2}$	$\frac{2}{n^2} \le \sigma < \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$	$\frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)} \le \sigma$
w_1	2σ	$2\sigma + 2\sqrt{\sigma}\left(\frac{3}{2}\mathcal{W}_0\left(\frac{2}{3}\left(n\sqrt{\sigma}\right)^{\frac{4}{3}}\right)\right)^{-1/4}$	$2\sigma + 2$
h_1	1	$1 + \sigma\sqrt{\frac{3}{2}\mathcal{W}_0\left(\frac{2}{3}\left(n\sqrt{\sigma}\right)^{\frac{4}{3}}\right)}$	$1 + \sigma\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$

Preparation. Let $i \in \{1, 2 \ldots, m\}$. As in Section 4.3, we let $(O, \vec{v}_i, \vec{u}_i)$ denote some orthonormal frame and let H_i be the halfplane supporting W_i with inner normal $\vec{u}_i$. We renumber the points of P^* with indices in $\{-\frac{n-1}{2}, \ldots, \frac{n-1}{2}\}$ so that p_0^* is the point in direction $\vec{u}_i$. For the sake of the presentation we assume that n is odd (the case of even n follows with trivial modifications). We write (x_i, y_i) for the coordinates of p_i in $(O, \vec{v}_i, \vec{u}_i)$ and denote by $p_{\vec{u}_i} \in \mathcal{H}^{(1)}$ the point of P extremal in direction $\vec{u}_i$. Our goal is to prove that $\mathbb{E}\left[\operatorname{card}\left(W_i \cap \mathcal{H}^{(1)}\right)\right]$ is $\Omega(1)$ in order to apply Theorem 2 (ii); in the light of

$$\mathbb{E}\left[\operatorname{card}\left(W_i \cap \mathcal{H}^{(1)}\right)\right] \ge \mathbb{P}\left[p_{\vec{u}_i} \in W_i\right]$$

we set out to bound from below the probability that $p_{\vec{u}_i}$ lies in W_i. We write z_j for the distance from p_j^* to H_i and note that

$$z_0 = h_1 - 1 \qquad \text{and} \qquad z_j = h_1 - 1 + 1 - \cos\frac{2\pi j}{n}$$

For $t \in [-\frac{1}{2}, \frac{1}{2}]$ we have $8t^2 \le 1 - \cos(2\pi t) \le 20t^2$, hence

$$\begin{aligned} h_1 - 1 + 8\frac{j^2}{n^2} \qquad &\le z_j \le \qquad h_1 - 1 + 20\frac{j^2}{n^2} \\ \frac{8j^2}{n^2} \quad &\le z_j - z_0 \le \quad \frac{20j^2}{n^2}. \end{aligned}$$

Analysis for Small σ. We start with the case $\sigma < \frac{2}{n^2}$, where the analysis is simpler but already uses the main ingredients of the general case. Since $h_1 = 1$, we have $z_0 = 0$ and therefore p_0^* lies on the boundary of H_i. We condition on the event $\{p_0 \in H_i \text{ and } p_{\vec{u}_i} = p_0\}$ and obtain:

$$\mathbb{P}\left[p_{\vec{u}_i} \in W_i\right] \geq \mathbb{P}\left[p_0 \in W_i \mid p_0 \in H_i \text{ and } p_{\vec{u}_i} = p_0\right] \mathbb{P}\left[p_0 \in H_i \text{ and } p_{\vec{u}_i} = p_0\right] \tag{9}$$

We bound each of these terms in turn.

Claim 4.8. *When $\sigma < \frac{2}{n^2}$, $\mathbb{P}\left[p_0 \in H_i \text{ and } p_{\vec{u}_i} = p_0\right] = \Omega(1)$.*

Proof. Using the independence of the random variables $\{y_j\}_j$ we write

$$\mathbb{P}\left[p_0 \in H_i \text{ and } p_{\vec{u}_i} = p_0\right] \geq \mathbb{P}\left[y_0 \geq h_1 \text{ and } \forall j \neq 0, y_j \leq h_1\right] = \mathbb{P}\left[y_0 \geq h_1\right] \prod_{j \neq 0} \mathbb{P}\left[y_j \leq h_1\right]$$

As $p_0^* \in H_i$, the point p_0 has probability at least $\frac{1}{2}$ of remaining in the half-plane H_i after a Gaussian perturbation, so $\mathbb{P}\left[y_0 \geq h_1\right] \geq \frac{1}{2}$. Moreover, $y_j \sim \mathcal{N}(h_1 - z_j, \sigma^2)$ so Lemma 4.1 (i) and the bounds on z_j and σ lead to:

$$\mathbb{P}\left[y_j \geq h_1\right] = \mathbb{P}\left[y_j - \mathbb{E}\left[y_j\right] \geq z_j\right] = Q\left(\tfrac{z_j}{\sigma}\right) \leq Q\left(\frac{8j^2}{n^2\sigma}\right) \leq Q\left(4j^2\right) \leq e^{-2j^2},$$

and $\mathbb{P}\left[y_j \leq h_1\right] \geq 1 - e^{-2j^2}$. Taking the logarithm we obtain

$$\ln \mathbb{P}\left[p_0 \in H_i \text{ and } p_{\vec{u}_i} = p_0\right] = \ln \mathbb{P}\left[y_0 \geq h_1\right] + 2\sum_{j=1}^{\frac{n-1}{2}} \ln \mathbb{P}\left[y_j \leq h_1\right] \geq \ln \tfrac{1}{2} + 2\sum_{j=1}^{\frac{n-1}{2}} \ln\left(1 - e^{-2j^2}\right)$$

Then, using that for $t \in (0, \frac{1}{2}]$ we have $\ln(1-t) > -2t$ and Lemma 4.1 (iii-a) we get

$$-\ln \mathbb{P}\left[p_0 \in H_i \text{ and } p_{\vec{u}_i} = p_0\right] \leq \ln 2 - 2\sum_{j=1}^{\frac{n-1}{2}} -2e^{-2j^2} \leq \ln 2 + 4\sum_{j=0}^{\frac{n-1}{2}} e^{-2j^2} = O(1),$$

and

$$\mathbb{P}\left[p_{\vec{u}_i} = p_0\right] = e^{-O(1)} = \Omega(1).$$

□

Equation (9) finally implies that $\mathbb{P}\left[p_{\vec{u}_i} \in W_i\right]$ is $\Omega(1)$, so $\mathbb{E}\left[\text{card}\left(W_i \cap \mathcal{H}^{(1)}\right)\right]$ is indeed $\Omega(1)$ for this range of σ.

Relevant Points. The contribution of the jth point to $\mathbb{E}\left[\text{card}\left(H_i \cap P\right)\right]$ is $Q\left(\frac{z_j}{\sigma}\right)$. The gist of our analysis for larger σ is to split the points into two parts, the relevant points where $Q\left(\frac{z_j}{\sigma}\right) = \Theta\left(Q\left(\frac{z_0}{\sigma}\right)\right)$ and the irrelevant ones. The expected number of points in H_i is (up to a constant multiplicative factor) at least the number of relevant points times $Q\left(\frac{z_0}{\sigma}\right)$; fine tuning z_0 so that this product is $\Omega(1)$ then amounts to solving some functional equation. Specifically, we call a point p_j *relevant* if $|j| \leq j_m = \min\left(\left\lfloor \frac{n\sigma}{\sqrt{z_0}} \right\rfloor, \frac{n-1}{2}\right)$. We denote by P_r the relevant points. The same conditioning as in Equation (9) yields

$$\mathbb{P}\left[p_{\vec{u}_i} \in W_i\right] \geq \mathbb{P}\left[p_{\vec{u}_i} \in W_i \mid p_{\vec{u}_i} \in H_i \cap P_r\right] \mathbb{P}\left[p_{\vec{u}_i} \in H_i \cap P_r\right]. \tag{10}$$

One of the terms can be bounded as easily as for small σ.

Claim 4.9. *When* $\sigma \geq \frac{2}{n^2}$, $\mathbb{P}\left[p_{\vec{u}_i} \in W_i \mid p_{\vec{u}_i} \in H_i \cap P_r\right] \geq \frac{1}{2}$.

Proof. First, note that the parameter w_1 is set so that in the orthogonal projection on the $\vec{v}_i$-axis, the image of the witness contains the image of the ball $B(p_j, \sigma)$ whenever p_j is relevant. This ensures that

$$\mathbb{P}\left[p_{\vec{u}_i} \in W_i \mid p_{\vec{u}_i} \in H_i \text{ and } p_{\vec{u}_i} \text{ is relevant}\right] \geq 1 - 2Q(1) \geq \frac{1}{2}.$$

□

Counting Relevant Points in $\mathbf{H_i}$. Bounding the remaining probability requires different quantitative analysis depending on the range of σ but are based on the same principle: counting the expected number of relevant points in H_i. Since H_i has inner normal $\vec{u}_i$, we have

$$\mathbb{P}\left[p_{\vec{u}_i} \in H_i \mid p_{\vec{u}_i} \in P_r\right] = \mathbb{P}\left[H_i \cap P_r \neq \emptyset\right].$$

Thus, by the Chernoff bound of Lemma 2.1 (a), to show that the right-hand term is $\Omega(1)$ it suffices to show that H_i contains on average $\Omega(1)$ relevant points. Notice that

$$\mathbb{P}\left[p_j \in H_i\right] = \mathbb{P}\left[y_j - \mathbb{E}\left[y_j\right] > z_j\right] = \Omega\left(Q\left(\frac{z_j}{\sigma}\right)\right),$$

so the expected number of relevant points in H_i writes

$$\Omega\left(\sum_{j=-j_m}^{j_m} Q\left(\frac{z_j}{\sigma}\right)\right) = \Omega\left(Q\left(\frac{z_0}{\sigma}\right)\sum_{j=0}^{j_m} \frac{1}{\frac{z_j}{\sigma} + \frac{\sigma}{z_j}} \frac{z_0}{\sigma} e^{-\frac{1}{2\sigma^2}(z_j^2 - z_0^2)}\right). \tag{11}$$

Recall that $z_j = z_0 + \Theta\left(\frac{j^2}{n^2}\right)$. How we evaluate Equation (11) depends on the range of σ.

Large σ. When $\sigma \geq \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$, every point is relevant, *ie.* $j_m = \frac{n-1}{2}$, since

$$z_0 = \sigma\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)} \quad \text{implies} \quad \frac{n\sigma}{\sqrt{z_0}} = n\sqrt{\frac{\sigma}{\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}}} \geq \frac{n}{2}.$$

Claim 4.10. *When* $\sigma \geq \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$, $\mathbb{P}\left[p_{\vec{u}_i} \in H_i \cap P_r\right] = \Omega(1)$.

Proof. Since every point is relevant, this probability equals the probability that $H_i \cap P$ is non-empty. Computations similar to that of Equation (6) yield that $Q\left(\frac{z_0}{\sigma}\right) = \Theta\left(\frac{1}{n}\right)$. Moreover, $z_j \geq \frac{\sigma}{2}$ so $\frac{1}{\frac{z_j}{\sigma} + \frac{\sigma}{z_j}} = \Theta\left(\frac{\sigma}{z_j}\right)$. Also, $z_j = \Theta(z_0)$ and $z_j^2 - z_0^2 = \Theta\left(\frac{j^2 z_0}{n^2}\right)$. Injecting these three relations in Equation (11) we obtain that the expected number of (relevant) points in H_i writes

$$\mathbb{E}\left[\text{card}\left(H_i \cap P\right)\right] = \Omega\left(\frac{1}{n}\sum_{j=0}^{\frac{n-1}{2}} \frac{z_0}{z_j} e^{-j^2\Theta\left(\frac{z_0}{n^2\sigma^2}\right)}\right) = \Omega\left(\frac{1}{n}\sum_{j=0}^{\frac{n-1}{2}} e^{-j^2\Theta\left(\frac{z_0}{n^2\sigma^2}\right)}\right)$$

Since $\frac{z_0}{n^2\sigma^2} < \frac{4}{n^2}$, Lemma 4.1 (iii-c) implies that

$$\sum_{j=0}^{\frac{n-1}{2}} e^{-j^2\Theta\left(\frac{z_0}{n^2\sigma^2}\right)} = \Omega(n),$$

so we finally get that H_i contains $\Omega(1)$ (relevant) points on average. The Chernoff bound of Lemma 2.1 (a) yields that $\mathbb{P}\left[H_i \cap P \neq \emptyset\right]$ is $\Omega(1)$, and so is $\mathbb{P}\left[p_{\vec{u}_i} \in H_i \cap P_r\right] = \Omega(1)$. □

Intermediate σ. When $\frac{2}{n^2} \leq \sigma < \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$ we have $z_0 = \sigma\sqrt{\frac{3}{2}\mathcal{W}_0\left(\frac{2}{3}(n\sqrt{\sigma})^{\frac{4}{3}}\right)}$. The function $x \mapsto \frac{x}{\sqrt{\frac{3}{2}\mathcal{W}_0\left(\frac{2}{3}(n\sqrt{x})^{\frac{4}{3}}\right)}}$ is increasing. Let us define σ_0 as the solution of

$$\sigma_0 = \frac{1}{4}\sqrt{\frac{3}{2}\mathcal{W}_0\left(\frac{2}{3}(n\sqrt{\sigma_0})^{\frac{4}{3}}\right)}$$

Using that $\mathcal{W}_0$ is the solution to $f(x)e^{f(x)} = x$, we obtain that

$$\sigma_0 = \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}.$$

Then, for $\sigma < \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$ we have

$$\frac{n\sigma}{\sqrt{z_0}} \leq n\sqrt{\frac{\sigma}{\sqrt{\frac{3}{2}\mathcal{W}_0\left(\frac{2}{3}(n\sqrt{\sigma})^{\frac{4}{3}}\right)}}} < n\sqrt{\frac{\sigma_0}{\sqrt{\frac{3}{2}\mathcal{W}_0\left(\frac{2}{3}(n\sqrt{\sigma_0})^{\frac{4}{3}}\right)}}} = \frac{n}{2}$$

and $j_m = \left\lfloor \frac{n\sigma}{\sqrt{z_0}} \right\rfloor$. Notice that

$$\mathbb{P}\left[p_{\vec{u}_i} \in H_i \cap P_r\right] \geq \mathbb{P}\left[H_i \cap P_r \neq \emptyset\right] \mathbb{P}\left[H_i \cap (P \setminus P_r) = \emptyset\right].$$

The two quantities on the right-hand side are independent.

Claim 4.11. *When $\frac{2}{n^2} \leq \sigma < \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$, $\mathbb{P}\left[P_r \cap H_i \neq \emptyset\right] = \Omega(1)$.*

Proof. Note that z_0 is set so that $Q\left(\frac{z_0}{\sigma}\right) = \Theta\left(\frac{1}{j_m}\right) = \Theta\left(\frac{\sqrt{z_0}}{n\sigma}\right)$. Indeed, using Lemma 4.1 (i) and the fact that $z_0 = \Omega(\sigma)$, $Q\left(\frac{z_0}{\sigma}\right) = \Theta\left(\frac{e^{-\frac{z_0^2}{2\sigma^2}}}{\frac{z_0}{\sigma}}\right)$. The choice for z_0 comes from the resolution of the equation $\frac{1}{x}e^{-\frac{x^2}{2}} = \frac{\sqrt{x}}{n\sqrt{\sigma}}$ using the definition of the function $\mathcal{W}_0$.

Moreover, for $|j| \leq j_m$ we have $z_j = \Theta(z_0)$ and $z_j^2 - z_0^2 = \Theta\left(\frac{j^2 z_0}{n^2}\right)$. Indeed, $\sigma \geq \frac{2}{n^2}$ implies that $z_0 = \Omega(\sigma)$ and $z_j < z_0 + O\left(\frac{j_m^2}{n^2}\right) = O\left(\frac{z_0^2+\sigma^2}{z_0}\right) = O(z_0)$ and $z_j = z_0 + \Theta\left(\frac{j^2}{n^2}\right) = \Omega(z_0)$. Also, $z_j = \Omega(\sigma)$ so $\frac{1}{\frac{z_j}{\sigma}+\frac{\sigma}{z_j}} = \Omega\left(\frac{\sigma}{z_j}\right)$. Injecting these relations into Equation (11) we obtain that the expected number of relevant points in H_i is

$$\Omega\left(\frac{1}{j_m}\sum_{j=-j_m}^{j_m}\frac{1}{\frac{z_j}{\sigma}+\frac{\sigma}{z_j}}\frac{z_0}{\sigma}e^{-\frac{1}{2\sigma^2}(z_j^2-z_0^2)}\right) = \Omega\left(\frac{1}{j_m}\sum_{j=0}^{j_m}e^{-j^2\Theta\left(\frac{z_0}{n^2\sigma^2}\right)}\right).$$

Again, Lemma 4.1 (iii-b) ensures that

$$\sum_{j=0}^{j_m}e^{-j^2\Theta\left(\frac{z_0}{n^2\sigma^2}\right)} = \Omega\left(\frac{n\sigma}{\sqrt{z_0}}\right)$$

and the expected number of relevant points in H_i is $\Omega(1)$. The Chernoff bound of Lemma 2.1 (a) yields that $\mathbb{P}\left[H_i \cap P \neq \emptyset\right]$ is $\Omega(1)$. □

It remains to bound the third quantity:

Claim 4.12. *When* $\frac{2}{n^2} \leq \sigma < \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$, $\mathbb{P}\left[H_i \cap (P \setminus P_r) = \emptyset\right] = \Omega(1)$.

Proof. Every irrelevant point p_j belongs to H_i with probability $Q\left(\frac{z_j}{\sigma}\right)$. The probability that H_i contains no irrelevant point is therefore at least

$$\left(\prod_{j=j_m+1}^{\frac{n-1}{2}} 1 - Q\left(\frac{z_j}{\sigma}\right)\right)^2$$

Lemma 4.1 (i) and the fact that $z_j = z_0 + \Theta\left(\frac{j^2}{n^2}\right)$ ensure that

$$1 - Q\left(\frac{z_j}{\sigma}\right) \geq 1 - Q\left(\frac{z_0}{\sigma}\right) e^{-\frac{1}{2\sigma^2}(z_j^2 - z_0^2)} = 1 - Q\left(\frac{z_0}{\sigma}\right) e^{-j^2\Theta\left(\frac{z_0}{n^2\sigma^2}\right)}$$

so the probability that H_i contains no irrelevant point is at least

$$\gamma = \left(\prod_{j=j_m+1}^{\frac{n-1}{2}} 1 - \frac{1}{j_m} e^{-j^2\Theta\left(\frac{1}{j_m^2}\right)}\right)^2.$$

Taking the logarithm, and using $\ln(1-t) \geq -2t$ for $t \in [0,1]$, we get

$$-\ln\gamma = -2\sum_{j=j_m+1}^{\frac{n-1}{2}} \ln\left(1 - \frac{1}{j_m} e^{-j^2\Theta\left(\frac{1}{j_m^2}\right)}\right) \leq \frac{4}{j_m}\sum_{j=j_m+1}^{\frac{n-1}{2}} e^{-j^2\Theta\left(\frac{1}{j_m^2}\right)} \leq \frac{4}{j_m}\sum_{j=0}^{\frac{n-1}{2}} e^{-j^2\Theta\left(\frac{1}{j_m^2}\right)}$$

and Lemma 4.1 (iii-a) yields $0 \leq -\ln\gamma \leq O(1)$. It follows that the probability that H_i contains no irrelevant point is at least $e^{-O(1)} = \Omega(1)$. □

Wrapping Up. We can now obtain our lower bound.

Proof of Theorem 10. Lemma 4.4 and the preceding analysis ensure that the assumptions of Theorem 2 (ii) are satisfied, and we thus have $\mathbb{E}\left[\operatorname{card}\operatorname{CH}(P)\right] = \Omega\left(\min\left(n, h_1/w_1\right)\right)$. We treat separately the three regimes.

If $\sigma < \frac{2}{n^2}$ then

$$\mathbb{E}\left[\operatorname{card}\operatorname{CH}(P)\right] = \Omega\left(\min\left(n, \left(\frac{1}{O\left(\frac{1}{n^2}\right)}\right)\right)\right) = \Omega(n)$$

which is the first regime announced in Theorem 10. (Note that the boundaries between the regimes can be set up to a multiplicative constant.)

If $\frac{2}{n^2} \leq \sigma < \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$ then

$$\mathbb{E}\left[\operatorname{card}\operatorname{CH}(P)\right] = \Omega\left(\frac{1 + \sigma\sqrt{\ln\left(n\sqrt{\sigma}\right)}}{\sqrt{\sigma}\left(\sqrt{\sigma} + \frac{1}{\sqrt[4]{\ln\left(n\sqrt{\sigma}\right)}}\right)}\right)$$

We simplify this expression by comparing σ and $\frac{1}{\sqrt{\ln\left(n\sqrt{\sigma}\right)}}$. Specifically, if $\sigma \leq \frac{1}{\sqrt{\ln n}}$ then

$$\sigma = O\left(\frac{1}{\sqrt{\ln\left(n\sqrt{\sigma}\right)}}\right) \quad \text{and} \quad \mathbb{E}\left[\operatorname{card}\operatorname{CH}(P)\right] = \Omega\left(\frac{\sqrt[4]{\ln\left(n\sqrt{\sigma}\right)}}{\sqrt{\sigma}}\right)$$

which is the second regime announced in Theorem 10.

If $\frac{1}{\sqrt{\ln n}} \leq \sigma < \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)} = O\left(\sqrt{\ln n}\right)$ then $\frac{1}{\sqrt{\ln\left(n\sqrt{\sigma}\right)}} = O(\sigma)$ and

$$\mathbb{E}\left[\operatorname{card}\operatorname{CH}(P)\right] = \Omega\left(\frac{\sigma\sqrt{\ln\left(n\sqrt{\sigma}\right)}}{\sigma}\right) = \Omega\left(\sqrt{\ln\left(n\sqrt{\sigma}\right)}\right) = \Omega\left(\sqrt{\ln n}\right)$$

If $\sigma \geq \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)} = \Omega\left(\sqrt{\ln n}\right)$ then

$$\mathbb{E}\left[\operatorname{card}\operatorname{CH}(P)\right] = \Omega\left(\frac{1+\sigma\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}}{\sigma+1}\right) = \Omega\left(\sqrt{\ln n}\right)$$

The lower bound is the same as in the case $\frac{1}{\sqrt{\ln n}} \leq \sigma < \frac{1}{4}\sqrt{\mathcal{W}_0\left(\frac{n^2}{4}\right)}$. Merging the two conditions we obtain that

$$\sigma \geq \frac{1}{\sqrt{\ln n}} \Rightarrow \mathbb{E}\left[\operatorname{card}\operatorname{CH}(P)\right] = \Omega\left(\sqrt{\ln n}\right)$$

which is the third regime announced in Theorem 10. □

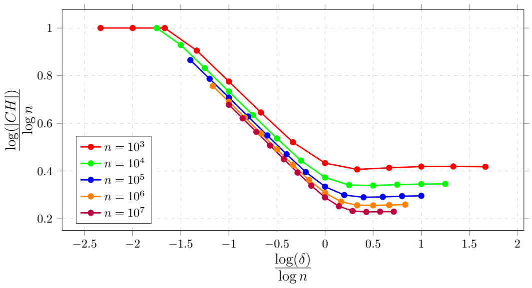
5 Concluding remarks

5.1 Poisson distribution

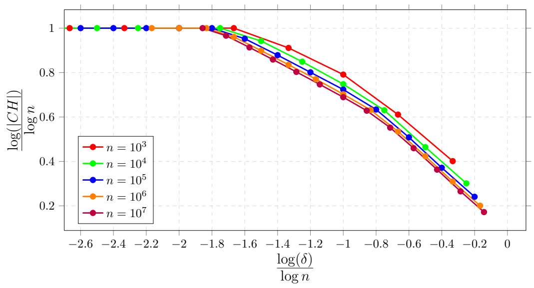
Theorem 2 is established for a set of n independent elements. Except for some technicalities in the presentation, nothing prevents making n a random variable to prove *eg.* analogs of Theorems 3 and 8 for Poisson distributions. (As this was not required for our application to smoothed complexity analysis, we opted for a simpler presentation where n is fixed.)

5.2 Silhouette of Polytopes

Glisse, Lazard, Michel and Pouget [13] used the witness and collector approach to study the expected size of the silhouette of a 3D random convex polytope defined as the convex hull of a Poisson point process of intensity n on the unit sphere. The silhouette of the polytope from a given viewpoint is the two dimensional convex hull of the projection of the points, thus the problem reduces to the size of the convex hull of i.i.d. points in a disk for the distribution corresponding to the projection of a Poisson point process. Glisse *et al.* analyzed the size of that convex hull using a system of witnesses and collectors adapted to that distribution and proved that the worst point of view yields a silhouette of expected size $\Theta\left(\sqrt{n}\right)$.



(a) Experimental results for the complexity of the convex hull of a ℓ^∞ perturbation of amplitude δ of the regular n-gon inscribed in the unit circle. Each data point corresponds to an average over 1000 experiments.



(b) Experimental results for the complexity of the convex hull of a rounding of the regular n-gon inscribed in the unit circle on a grid of pixel size δ.

Figure 3: Experimental results for the ℓ^∞ perturbation and rounding.

5.3 ℓ^∞ Perturbation and Snap-Rounding

Systems of witnesses and collectors can be designed for perturbations that are uniform in the ball for other metrics. In [1], denoting $\square$ the unit square in 2D, we prove the following theorem:

Theorem 11. *Let $P^* = \{p_i^* : 1 \leq i \leq n\}$ be an $(\Theta(n), \Theta(1))$-sample of the unit circle in $\mathbb{R}^2$ and let $P = \{p_i = p_i^* + \eta_i\}$ where $\eta_1, \eta_2, \ldots, \eta_n$ are random variables chosen independently from $\mathcal{U}_{\delta\square}$. For any fixed k, and $\delta \in [n^{-2}, 1]$*

$$\mathbb{E}\left[\operatorname{card} \mathcal{H}^{(k)}\right] = \Theta\left(n^{\frac{1}{5}} \delta^{-\frac{2}{5}}\right)$$

As in the Euclidean case, the witnesses and collectors are parallel half planes, but the partition of ranges must be adapted to cope with the lack of rotational symmetry. The angle α_i of the set of directions covered by R_i is no longer constant and is much smaller when the ranges are almost horizontal or vertical than when they are oblique. The bound of Theorem 11 is confirmed experimentally (*cf.* the slopes of $-\frac{2}{5}$ in the plots of Figure 3a). Theorem 11 implies that for $\delta \in [n^{-2}, 1]$, $\mathcal{S}(n, \mathcal{U}_{\delta\square})$ is $\Omega\left(n^{\frac{1}{5}} \delta^{-\frac{2}{5}}\right)$. It is also known to be $O\left(\left(\frac{n \ln n}{\delta}\right)^{\frac{2}{3}}\right)$, for all ranges of δ, by the upper bound obtained by Damerow and Sohler for dominant points under ℓ^∞ noise [6].

Snap Rounding. Given a grid whose pixel have size δ, rounding points with real coordinates at the center of their pixel have some similarity with ℓ^∞ noise. Actually, for a single point, and if the origin of the grid is random, the two processus are identical, but when several points are involved things are different: clearly rounding creates collisions while noising separates identical points. However for the regular n-gon, provided that $\delta < \frac{1}{n}$ the two processes gives convex hulls of similar size as confirmed by Figure 3b.

5.4 Delaunay Triangulation

Systems of witnesses and collectors can also be used to prove the following well known result of Dwyer [11]:

Theorem 12 (Dwyer [11])**.** *The expected complexity of the Delaunay triangulation of n random points uniformly distributed in the unit ball $\mathbb{B}$ of dimension d is $\Theta(n)$.*

In a preliminary version [9] we gave a proof, considerably simpler than Dwyer's, of this result up to logarithmic factors; these factors can be removed thanks to Theorem 2 using a system of witnesses and collectors that we now outline.

The faces of dimension k of the Delaunay triangulation are hyperedges of size $k+1$ in the hypergraph where the ranges are balls in $\mathbb{R}^d$. More precisely, given a set P of n points in general position, $k+1$ points define a face of the Delaunay triangulation $DT(P)$ iff there exists a ball with the $k+1$ points on its boundary and no other points inside. Thus the hypergraph define using the balls as ranges may be a strict superset of the Delaunay faces. Our proof splits the ranges in three subsets and builds a system of witnesses and collectors for each these subsets.

Balls Centered Deep Inside $\mathbb{B}$. Let $r_j = O\left(\left(\frac{j}{n}\right)^{\frac{1}{d}}\right)$ denote the radius of a ball completely contained in $\mathbb{B}$ and expected to contain j points. We use a minimal covering of $\mathbb{B}$ with balls of radius r_1 and keep the balls centered inside $(1 - r_1 \ln^2 n)\mathbb{B}$ to define our first level witnesses W_i^1. We define W_i^j as the ball concentric with W_i^1 with radius r_j, and C_i^j as the ball concentric with W_i^j with radius $r_j + 2r_1$. We finally let R_i be the set of balls centered in W_i^1. This system of witnesses and collectors verifies the hypotheses of Theorem 2 (i), and a constant fraction of the first layer $\{(W_i^1, C_i^1)\}_i$ verifies the hypotheses of Theorem 2 (ii). Altogether, they allow to conclude that the number of Delaunay balls centered in $(1 - r_1 \ln^2 n)\mathbb{B}$ is $\Theta(n)$.

Balls Centered Near $\partial\mathbb{B}$. The Delaunay balls centered in an annulus of width $2r_1 \ln^2 n$ around $\partial\mathbb{B}$ can be counted easily since their number is sublinear. To this aim we can cover the above annulus by collectors of diameter $O(r_1 \ln^2 n)$ and use associated empty witnesses.

Balls Centered outside $\mathbb{B}$. Balls centered outside $\mathbb{B}$ are a bit more delicate, since they can have a large radius but, possibly, a small probability to be empty. A first remark is that ball of infinite radius are half plane and are counted by Theorem 3. Actually, the construction of Theorem 3 can be adapted to count all balls of radii between α and 2α by using balls of radius α to define the witnesses and balls of radii 2α for the collectors. Then it is possible to sum on various values of α to cover all the possible radii. As a side result we get the expected size of the α-shape of points uniformly distributed in $\mathbb{B}$.

Acknowledgements

This work was initiated during the 8^{th} McGill - INRIA Workshop on Computational Geometry at the Bellairs institute in 2009. The authors wish to thank all the participants for creating a pleasant and stimulating atmosphere, Nina Amenta and Sariel Har-Peled for useful discussions at the early stage of this work, and in particular Dominique Attali who was involved in the preliminary work [1].

References

[1] D. Attali, O. Devillers, and X. Goaoc. The Effect of Noise on the Number of Extreme Points. Research Report RR-7134, Inria, 2009. URL `https://hal.inria.fr/inria-00438409`.

[2] I. Bárány and D. Larman. Convex bodies, economic cap coverings, random polytopes. *MATHEMATIKA*, 35:274–291, 1988. ISSN 2041-7942. doi:10.1112/S0025579300015266.

[3] I. Bárány and V. Vu. Central limit theorems for gaussian polytopes. *The Annals of Probability*, 35(4):1593–1621, 2007.

[4] F. Cazals and J. Giesen. Delaunay triangulation based surface reconstruction. In J.-D. Boissonnat and M. Teillaud, editors, *Effective Computational Geometry for Curves and Surfaces*, pages 231–276. Springer-Verlag, Mathematics and Visualization, 2006. doi:10.1007/978-3-540-33259-6_6. URL `http://www.springerlink.com/index/j1h1893177371m97.pdf`.

[5] R. M. Corless, G. H. Gonnet, D. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Advances in Computational Mathematics*, 5(1):329–359, 1996. doi:10.1007/BF02124750.

[6] V. Damerow and C. Sohler. Extreme points under random noise. In *Proc. 12th European Sympos. Algorithms*, pages 264–274, 2004. doi:10.1007/978-3-540-30140-0_25.

[7] M. de Berg. Improved bounds on the union complexity of fat objects. *Discrete & Computational Geometry*, 40:127–140, 2008. doi:10.1007/s00454-007-9029-7. URL `http://www.springerlink.com/index/Q3917G253G836885.pdf`.

[8] O. Devillers, M. Glisse, and X. Goaoc. Complexity analysis of random geometric structures made simpler. Research Report RR-8168, INRIA, 2012. URL `https://hal.inria.fr/hal-00761171`.

[9] O. Devillers, M. Glisse, and X. Goaoc. Complexity analysis of random geometric structures made simpler. In *Symposium on Computational Geometry*, pages 167–176, 2013. doi:10.1145/2462356.2462362.

[10] O. Devillers, M. Glisse, X. Goaoc, and R. Thomasse. On the smoothed complexity of convex hulls. In *Symposium on Computational Geometry*, pages 224–239, 2015. doi:10.4230/LIPIcs.SOCG.2015.224.

[11] R. Dwyer. The expected number of k-faces of a Voronoi diagram. *Internat. J. Comput. Math.*, 26(5):13–21, 1993. doi:10.1016/0898-1221(93)90068-7. URL `http://www.sciencedirect.com/science/article/pii/0898122193900687`.

[12] J. Erickson. Nice point sets can have nasty Delaunay triangulations. *Discrete and Computational Geometry*, 30(1):109–132, 2003. doi:10.1007/s00454-003-2927-4. URL `http://www.springerlink.com/index/91UKK7QDWVPX77D1.pdf`.

[13] M. Glisse, S. Lazard, J. Michel, and M. Pouget. Silhouette of a random polytope. Research Report 8327, INRIA, 2013. URL `https://hal.inria.fr/hal-00841374/`.

[14] V. Petrov. *Limit Theorems of Probability Theory. Sequence of Independent Random Variables.* Number 4 in Oxford studies in probability. Clarendon Press, 1995. URL `http://www.citeulike.org/group/2854/article/1615442`.

[15] H. Raynaud. Sur l'enveloppe convexe des nuages de points aleatoires dans R^n. *J. Appl. Probab.*, 7:35–48, 1970. URL `http://www.jstor.org/stable/10.2307/3212146`.

[16] M. Reitzner. Random polytopes. In *New perspectives in stochastic geometry*, pages 45–76. Oxford Univ. Press, Oxford, 2010.

[17] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gewählten Punkten I. *Z. Wahrsch. Verw. Gebiete*, 2:75–84, 1963. doi:10.1007/BF00535300.

[18] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gewählten Punkten II. *Z. Wahrsch. Verw. Gebiete*, 3:138–147, 1964. doi:10.1007/BF00535973.

[19] D. A. Spielman and S.-H. Teng. Smoothed analysis: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51:385 – 463, 2004. doi:10.1145/990308.990310.

[20] J. J. Sylvester. On a special class of questions on the theory of probabilities. *Birmingham British Association Report*, pages 8–9, 1865.

A Large Perturbation and Rescaling for Gaussian perturbation

Lemma A.1. *Let $P^* = \{p_1^*, p_2^* \dots, p_n^*\}$ be a set of points in $\mathbb{R}^2$ with diameter $2r$. Let P be the set of $p_i = p_i^* + \eta_i$, where $\eta_i \sim \mathcal{N}(0, \sigma^2 I_2)$ are independent, for $\sigma \geq 3r\sqrt{\ln n}$. Then, the expected number of vertices of the convex hull of P is $O\left(\sqrt{\ln n}\right)$.*

Proof. We can suppose that P^* is included in $r\mathbb{B}$. We set up the Gaussian witness-collector construction of Section 4.2, with $\ell = \ln^2 n$ and the values of α, w_j and h_j set as specified on the right.

$$\boxed{\begin{aligned} \alpha &= \tfrac{\sigma}{h_1 - r} = \Theta\left(\tfrac{1}{\sqrt{\ln n}}\right) \\ w_j &= w = \tfrac{7}{2}\sigma \\ h_j &= r + \sigma\sqrt{\mathcal{W}_0\left(\tfrac{n^2}{j^2}\right)} \end{aligned}}$$

Every point $p \in P$ writes $p = x_i\vec{v}_i + y_i\vec{u}_i$ with $x_i, y_i \sim \mathcal{N}(0, \sigma^2)$ and $\vec{u}_i, \vec{v}_i \in \mathbb{S}^1, \vec{v}_i \perp \vec{u}_i$. Thus, the probability for p to be in W_i^j is upper-bounded by $\mathbb{P}\left[y_i > h_j\right]$ and is lower-bounded by $\mathbb{P}\left[y_i > h_j\right]\mathbb{P}\left[|x_i| \leq \frac{w-r}{2}\right]$. It's easy to see that $\mathbb{P}\left[|x_i| \leq \frac{w-r}{2}\right]$ is lower-bounded from below by $\mathbb{P}\left[|x_i| \leq \frac{5w}{14}\right]$ (since $r \leq \sigma = \frac{2w}{7}$), which is a constant.
Now, $\mathbb{P}\left[y_i > h_j\right] \in \left[Q\left(\frac{h_j+r}{\sigma}\right), Q\left(\frac{h_j-r}{\sigma}\right)\right]$ so $\mathbb{P}\left[y_i > h_j\right] < Q\left(\sqrt{\mathcal{W}_0\left(\frac{n^2}{j^2}\right)}\right) = \Theta\left(\frac{j}{n}\right)$. As

$$\sigma \geq 3r\sqrt{\ln n} \geq 2r\sqrt{\ln(n^2)} \geq 2r\sqrt{\mathcal{W}_0\left(n^2\right)} \geq 2r\sqrt{\mathcal{W}_0\left(\frac{n^2}{j^2}\right)},$$

we get $\frac{2r}{\sigma} \leq \frac{1}{\sqrt{\mathcal{W}_0\left(\frac{n^2}{j^2}\right)}} = \frac{\sigma}{h_j - r}$. and $Q\left(\frac{h_j+r}{\sigma}\right) = Q\left(\frac{h_j-r}{\sigma} + \frac{2r}{\sigma}\right) \geq Q\left(\frac{h_j-r}{\sigma} + \frac{\sigma}{h_j-r}\right)$. Now, using Lemma 4.1 (ii) , as long as $\frac{h_j-r}{\sigma} > \frac{1}{4}$ (which is true for $n \geq 3$ and $j \leq \ln^2 n$) we get that $Q\left(\frac{h_j-r}{\sigma} + \frac{\sigma}{h_j-r}\right) = \Omega\left(Q\left(\frac{h_j-r}{\sigma}\right)\right) = \Omega\left(\frac{j}{n}\right)$ using Equation (6).

We obtain that for any $p^* \in P^*$, $\mathbb{P}\left[p \in W_i^j\right] = \Theta\left(\frac{j}{n}\right)$, and so $\mathbb{E}\left[\text{card}\left(P \cap W_i^j\right)\right] = \Theta\left(j\right)$. Condition (a') therefore holds.

Since $n \geq 3$, $\alpha < \frac{1}{\sqrt{\mathcal{W}_0(3^2)}} < \frac{\pi}{4}$, $\tan\frac{\alpha}{2} < 0.5$ and $\frac{2h_j}{w} \geq 2\frac{h_\ell - r}{w} = \frac{4}{7}\sqrt{\mathcal{W}_0\left(\frac{n^2}{\ln^4 n}\right)} \geq 0.5$. Thus, $\tan\frac{\alpha}{2} < \frac{2h_j}{w}$, so the origin is not in C_i^j. Using Claim 4.3 with p^* at the origin, the collector C_i^j is included in the union of two half-spaces with height $\tilde{h_j} = h_j\cos\frac{\alpha}{2} - \frac{w}{2}\sin\frac{\alpha}{2} \geq h_j\cos\frac{\alpha}{2} - \frac{7}{8}\alpha\sigma$, since $\sin(x) < x$ for $x > 0$.

Now,

$$\begin{aligned} h_j \cos\tfrac{\alpha}{2} &= h_j + (\cos\tfrac{\alpha}{2} - 1)h_j \\ &\geq h_j - \frac{1}{2}\left(\tfrac{\alpha}{2}\right)^2 h_j = h_j - \frac{1}{8}\alpha\sigma\left(\frac{h_j}{h_1 - r}\right) \\ &\geq h_j - \frac{1}{8}\alpha\sigma\left(\frac{h_j}{h_j - r}\right) \geq h_j - \frac{1}{8}\alpha\sigma \end{aligned}$$

since $\cos x - 1 \geq -\frac{1}{2}x^2$. Thus, we obtain $\tilde{h_j} \geq h_j - \alpha\sigma$.

Note that $Q\left(\frac{h_j - r}{\sigma}\right) \geq Q\left(\frac{\tilde{h_j} - r}{\sigma} + \alpha\right) = Q\left(\frac{\tilde{h_j} - r}{\sigma} + \frac{\sigma}{h_1 - r}\right)$. Since $\tilde{h_j} \leq h_1$, $\frac{\sigma}{h_1 - r} \leq \frac{\sigma}{\tilde{h_j} - r}$ and

$$\begin{aligned} Q\left(\frac{h_j - r}{\sigma}\right) &> Q\left(\frac{\tilde{h_j} - r}{\sigma} + \frac{\sigma}{h_1 - r}\right) \\ &> Q\left(\frac{\tilde{h_j} - r}{\sigma} + \frac{\sigma}{\tilde{h_j} - r}\right) \\ &= \Omega\left(Q\left(\frac{\tilde{h_j} - r}{\sigma}\right)\right) \end{aligned}$$

using Lemma 4.1 (ii) (since $\frac{\tilde{h_j} - r}{\sigma} > \frac{1}{4}$ for $n \geq 3$ and $j \leq \ln^2 n$), we get $Q\left(\frac{\tilde{h_j} - r}{\sigma}\right) = O\left(\frac{j}{n}\right)$ by Equation (6).

Thus, $\mathbb{E}\left[\operatorname{card}\left(P \cap C_i^j\right)\right] = O(j)$, and Theorem 2 (i) can be applied and we obtain $\mathbb{E}\left[\operatorname{card}\mathcal{H}^{(1)}\right] = \mathbb{E}\left[\operatorname{card}\mathrm{CH}(P)\right] = O\left(\frac{1}{\alpha}\right) = O\left(\sqrt{\ln n}\right)$.

□

Corollary 13.

$$\mathcal{S}\left(n, \mathcal{N}(0, \sigma^2)\right) = O\left(\frac{\ln(n)}{\sigma^2}\sqrt{\ln(n\sigma^2)} + \sqrt{\ln n}\right).$$

Proof. Let P^* be some point of diameter at most 1. Without loss of generality, we assume that $P^* \subset \mathbb{B}$. We cover $\mathbb{B}$ with $m'' = \Theta\left(1 + \frac{1}{r^2}\right)$ disjoint cells of size $r = \frac{\sigma}{3\sqrt{\ln n}}$ We break down P^* into $P_1^* \cup P_2^* \cup \cdots \cup P_{m''}^*$ by taking its intersection with each covering cells; we let P_i denote the perturbation of P_i^* and $n_i = \operatorname{card} P_i$. Every vertex of $\mathrm{CH}(P)$ is a vertex of some $\mathrm{CH}(P_i)$, and we can apply Lemma A.1 to bound the number of vertices of $\mathrm{CH}(P_i)$ from above by $\sqrt{\ln n_i}$. If $m'' > 1$, the sum is maximized with $\forall i, n_i = \frac{n}{m''}$, which bounds from above the number of vertices of $\mathrm{CH}(P)$ by

$$m'' O\left(\sqrt{\ln\left(\tfrac{n}{m''}\right)}\right) = O\left(\tfrac{1}{r^2}\sqrt{\ln(nr^2)}\right) = O\left(\frac{\ln(n)}{\sigma^2}\sqrt{\ln(n\sigma^2)}\right)$$

using Equality (5). □

RESEARCH CENTRE
NANCY – GRAND EST

615 rue du Jardin Botanique
CS20101
54603 Villers-lès-Nancy Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399