

Complexity Analysis of Random Geometric Structures Made Simpler

Olivier Devillers, Marc Glisse, Xavier Goaoc

► **To cite this version:**

Olivier Devillers, Marc Glisse, Xavier Goaoc. Complexity Analysis of Random Geometric Structures Made Simpler. 29th Annual Symposium on Computational Geometry, Jun 2013, Rio, Brazil. pp.167-175, 2013, <10.1145/2462356.2462362>. <hal-00833774>

HAL Id: hal-00833774

<https://hal.inria.fr/hal-00833774>

Submitted on 13 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Complexity Analysis of Random Geometric Structures Made Simpler*

Olivier Devillers
INRIA Sophia Antipolis – Méditerranée, †

Marc Glisse
INRIA Saclay – Île de France, ‡

Xavier Goaoc
INRIA Nancy - Grand Est, §

ABSTRACT

Average-case analysis of data-structures or algorithms is commonly used in computational geometry when the, more classical, worst-case analysis is deemed overly pessimistic. Since these analyses are often intricate, the models of random geometric data that can be handled are often simplistic and far from “realistic inputs”. We present a new simple scheme for the analysis of geometric structures. While this scheme only produces results up to a polylog factor, it is much simpler to apply than the classical techniques and therefore succeeds in analyzing new input distributions related to smoothed complexity analysis.

We illustrate our method on two classical structures: convex hulls and Delaunay triangulations. Specifically, we give short and elementary proofs of the classical results that n points uniformly distributed in a ball in \mathbb{R}^d have a convex hull and a Delaunay triangulation of respective expected complexities $\tilde{\Theta}(n^{\frac{d-1}{d+1}})$ and $\tilde{\Theta}(n)$. We then prove that if we start with n points well-spread on a sphere, e.g. an (ϵ, κ) -sample of that sphere, and perturb that sample by moving each point randomly and uniformly within distance at most δ of its initial position, then the expected complexity of the convex hull of the resulting point set is

$$\tilde{\Theta}\left(\left(\sqrt{n}\right)^{1-\frac{1}{d}}\left(\frac{1}{\sqrt[4]{\delta}}\right)^{d-\frac{1}{d}}\right).$$

1. INTRODUCTION

Understanding the complexity of geometric structures (Delaunay triangulations, convex hulls, arrangements...) is a

*Part of this work is supported by: ANR blanc PRESAGE (ANR-11-BS02-003)

†<http://www-sop.inria.fr/members/Olivier.Devillers/>

‡<http://geometrica.saclay.inria.fr/team/Marc.Glisse/>

§<http://www.loria.fr/~goaoc/>

key problem in computational geometry as this understanding guides the design of geometric algorithms. This paper proposes a new simple method to analyze the expected complexity of structures close to geometric hypergraphs, e.g. convex hulls and Delaunay triangulations, for distributions of random point sets.

Worst-case, average case, and smoothed complexity.

Complexity analyses of geometric structures traditionally consider the *worst-case* scenario to provide a guarantee on how bad things can get. The worst-case is usually achieved by overly constrained constructions that vanish under small perturbation; such analyses often give a pessimistic image of the structure that can be misleading when it comes to algorithmic choices. For instance, worst-case analysis alone suggests that Delaunay-based reconstruction methods in \mathbb{R}^3 should be inefficient, as the worst-case complexity of the Delaunay triangulation in \mathbb{R}^3 is quadratic, whereas these methods prove effective in practice [3]. The worst-case analysis can sometimes be refined by introducing additional parameters to restrict the input model: fatness [6], spread [9], “reasonable” samples [1]... There are, however, contexts in which realistic input models remain elusive; for example, in spite of much efforts there is still no satisfying model of “plausible” computer graphics scenes.

An alternative, probabilistic, analysis averages the complexity function on a suitable distribution on the space of inputs. This *average-case* analysis can usually demonstrate that the complexity function seldom achieves its maximum, and therefore that worst-case bounds are rather pessimistic. Yet, the input distributions considered are often so simplistic that average-case bounds are rarely seen as relevant to depict “typical” situations. About ten years ago, Teng and Spielman [13] proposed a complexity model that stands in-between the worst-case and the average-case. The *smoothed complexity* is defined, informally, 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 behavior on “real data”, which is usually given with bounded precision and subject to measurement noise. Teng and Spielman proved that the simplex algorithm has polynomial smoothed complexity, explaining the apparent difficulty to construct exponential lower bounds and the effectiveness of this algorithm on practical input. Since this seminal paper, smoothed analysis has been applied to a variety of problems [14], including, in computational geometry, the complexity of visibility maps

in a terrain [7] and the number of maximal points (bounding also the number of vertices of the convex hull) [5].

Problem and results.

Performing a smoothed complexity analysis of a geometric structure divides into two different sub-problems. First, one needs to estimate the average complexity of a random perturbation of a given, fixed, input. This turns out to be a challenge in itself already for convex hulls or Delaunay triangulations as many of the techniques devised for average-case analysis do not readily apply to this type of distributions. Next, one needs to identify the inputs whose perturbations give rise to the largest expected complexity. Perhaps surprisingly, already for convex hulls in the plane the answer does not seem straightforward (see below).

Our main result is a simple method to analyze random *geometric hypergraphs*, which we explain in Section 2 and use to derive simple proofs of two classical bounds:

- (i) a $\tilde{\Theta}(n^{\frac{d-1}{d+1}})$ bound¹ on the expected number of k -dimensional faces of the convex hull of n points uniformly distributed in a unit ball $K \subset \mathbb{R}^d$,
- (ii) a $\tilde{\Theta}(n)$ bound on the expected number of k -dimensional faces of the Delaunay triangulation of n points uniformly distributed in a unit ball $K \subset \mathbb{R}^d$.

Note that, as we discuss in the conclusion, both bounds extend to points uniformly distributed in a smooth convex body in \mathbb{R}^d . Bounds (i) and (ii) are not new and are, in fact, weaker than the earlier bounds proven by, respectively, Raynaud [12] and Dwyer [8]. Yet, the analyses of Raynaud and Dwyer only avoid the logarithmic factors at the cost of tedious and ad hoc analyses. In contrast, our method is simple and systematic and enables us to analyze certain *random perturbations*; here a random (L^p -)perturbation of amplitude δ of a fixed point set P^* is the random point set obtained by choosing a point uniformly in each (L^p) ball of radius δ centered in a point of P^* . In this direction, we obtain:

- (iii) a $\tilde{\Theta}\left((\sqrt{n})^{1-\frac{1}{d}}\left(\frac{1}{\sqrt[4]{\delta}}\right)^{d-\frac{1}{d}}\right)$ bound on the expected number of k -dimensional faces of the convex hull of a L^2 -perturbation of amplitude $\delta \in [n^{-\frac{2}{d-1}}, 1]$ of an $(n^{-\frac{1}{d-1}}, \kappa)$ sample of a smooth convex body in \mathbb{R}^d ; we also obtain a bound of $\tilde{\Theta}\left(n^{\frac{1}{5}}\left(\frac{1}{\delta}\right)^{\frac{2}{5}}\right)$ on the expected number of vertices of the convex hull of a L^∞ -perturbation of amplitude $\delta \in [n^{-2}, 1]$ of an (n^{-1}, κ) sample of a smooth convex body in the plane.

(Recall that an (ϵ, κ) -sample of a set Γ is a point set $P \subset \Gamma$ such that any ball of radius ϵ centered on Γ contains between 1 and κ points from the sample P .) The only previous bound in the direction of (iii) is due to Damerow and Sohler [5] who studied the expected number of dominating points in point sets subjected to independent random noise on each coordinate axis; in particular, they proved a bound of $\tilde{O}\left(n^{\frac{2}{3}}\delta^{-\frac{2}{3}}\right)$ under L^∞ perturbation.

¹ The notations \tilde{O} , $\tilde{\Theta}$, $\tilde{\Omega}$ refer to asymptotic bounds up to a logarithmic factor, e.g. $f(n) = \tilde{\Theta}(n)$ means $\exists a, b \in \mathbb{Z}$, $f(n) = O(n \log^a n)$ and $f(n) = \Omega(n \log^b n)$.

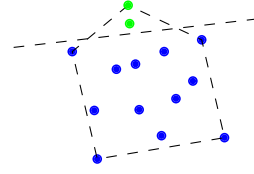


Figure 1: Geometric hypergraph induced by half-planes vs convex hull

One of the interesting consequences of Bound (iii) is that in expectation, the convex hull of a L^2 -perturbation of large enough amplitude of a regular n -gon is *smaller* than the convex hull of uniformly distributed points in a disk. This suggests that there is more to the problem of finding the extremal configurations for smoothed complexity of the convex hull than meets the eye, already in the plane.

2. OUTLINE OF THE METHOD

Range spaces and geometric hypergraphs.

A *range space* is a pair (X, R) where X is a set, the “ground” set, and R is a family of subsets of X , which we call the *ranges*. The *geometric hypergraph* H induced by a range space (X, R) on a finite set of points P is the family of intersections of P with the ranges of R :

$$H = \{P \cap r : r \in R\}.$$

(When the ground set is clear from the context, we also say that H is the geometric hypergraph induced by R on P ; in this paper, the ground set is always $X = \mathbb{R}^d$.) In other words, a subset $Q \subseteq P$ is a hyperedge if and only if there exists $r \in R$ such that $Q = r \cap P$; in that case we say that the hyperedge Q is *induced* by the range r . If H is a geometric hypergraph we denote by $H_{(k)}$ the set of hyperedges of H of cardinality k .

In this paper, we analyze the complexity of the convex hull and the Delaunay triangulation of a random point set P by considering the geometric hypergraphs induced on P by, respectively, the set of all halfspaces and the set of all balls. The convex hull and the Delaunay triangulations are typically proper subsets of the geometric hypergraphs; for instance the two green points on the Figure 1 form an edge in the hypergraph induced by halfplanes but are not an edge of the convex hull. This gap will not be an issue: the upper bounds obtained for geometric hypergraphs will be good enough for geometric structures, and the lower bounds can be handled by direct arguments.

The witness & collector method.

Let (\mathbb{R}^d, R) denote a range space and let P be a finite set of random points independently distributed in \mathbb{R}^d . Let H denote the geometric hypergraph induced by R on P .

Let $R' \subseteq R$ and let $H'_{(k)}$ denote the subset of $H_{(k)}$ induced by ranges in R' . Our goal is to estimate the expected size of $H'_{(k)}$. We first identify some region W' that contains with high probability at least $k+1$ points of P ; we call W' the *witness* of R' . We then define the *collector* C' of R' as the union of all ranges in R' not containing W' . Observe that

when W' contains at least $k+1$ points, the ranges of R' that contains W' cannot define any element of $H'_{(k)}$ and the size of $H'_{(k)}$ is therefore bounded from above by the number of k -element subsets of $P \cap C'$. Letting p stand for the probability that W' contains at most k points, we obtain the bound

$$\mathbb{E}[|H'_{(k)}|] \leq (1-p) \mathbb{E} \left[\binom{|P \cap C'|}{k} \middle| |W' \cap P| > k \right] + p \binom{|P|}{k},$$

since

$$\begin{aligned} & (1-p) \mathbb{E} \left[\binom{|P \cap C'|}{k} \middle| |W' \cap P| > k \right] \\ &= \mathbb{E} \left[\binom{|P \cap C'|}{k} \right] - p \mathbb{E} \left[\binom{|P \cap C'|}{k} \middle| |W' \cap P| \leq k \right] \\ &\leq \mathbb{E} \left[\binom{|P \cap C'|}{k} \right] \end{aligned}$$

we get

$$\mathbb{E}[|H'_{(k)}|] \leq \mathbb{E} \left[\binom{|P \cap C'|}{k} \right] + p \binom{|P|}{k}. \quad (1)$$

To make the contribution $p \binom{|P|}{k}$ negligible we want that $p = O(n^{-k})$. Chernoff's bound for lower tails implies that the probability that W' contains many points increases with the expected number of points in W' (see Lemma 1 below). Thus, p can be made arbitrarily small simply by inflating W' . There is a caveat: as C' grows with W' , increasing the probability that W' contains more than k points also increases the expectation of $\binom{|P \cap C'|}{k}$.

Bounding from above the expected number of k -element subsets of P contained in C' essentially amounts to bounding from above the k th moment of $|P \cap C'|$. Since P is a set of points independently distributed, $\mathbb{E}[(|P \cap C'|)^k] = O(\mathbb{E}[|P \cap C'|]^k)$, see Lemma 2, and Inequality (1) becomes

$$\mathbb{E}[|H'_{(k)}|] \leq O(\mathbb{E}[|P \cap C'|]^k) + \mathbb{P}[|W' \cap P| \leq k] \binom{|P|}{k}. \quad (2)$$

This will suffice for our purpose as we care only for the asymptotic orders of magnitude of $\mathbb{E}[|H'_{(k)}|]$.

For this approach to work effectively, we break up the set R of ranges into m pieces $R = R_1 \cup R_2 \cup \dots \cup R_m$ and define a witness and a collector for each piece. In our applications, we typically manage to bring down the expected number of points in a collector to $\tilde{O}(1)$, yielding that $\mathbb{E}[|H_{(k)}|] = \tilde{O}(m)$. Since each witness contains some element of $H_{(k)}$ with high probability, if no more than $O(1)$ witnesses have a common intersection then $\mathbb{E}[|H_{(k)}|] = \Omega(m)$ as well. We therefore often obtain bounds that are tight up to a polylogarithmic factor.

Technical preliminaries.

We conclude this high-level overview with the two technical lemmas mentioned previously.

LEMMA 1. *Let P be a set of n random points of \mathbb{R}^d independently distributed. If W is a subset that contains on*

average $k \log n$ points of P then the probability that W contains less than $k+1$ points of P is $\tilde{O}(n^{-k})$.

PROOF. Let X_i be the indicator function of the event that the i^{th} point from P belongs to W . We write $X = X_1 + \dots + X_n$ and let $\alpha = \mathbb{E}[X]$. Chernoff's bound for lower tails yields that for any $\delta \in (0, 1)$

$$\mathbb{P}[X < (1-\delta)\alpha] < \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}} \right)^\alpha.$$

In particular, setting $\delta = 1 - \frac{k+1}{k \log n}$ we get that $(1-\delta)\alpha = k+1$ and:

$$\begin{aligned} \mathbb{P}[X < k+1] &< \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}} \right)^\alpha \\ &= e^{-\alpha + \frac{k+1}{k \log n} \alpha} \left(\frac{k+1}{k \log n} \right)^{-\alpha \frac{k+1}{k \log n}} \end{aligned}$$

Since $\alpha = k \log n$, the previous bound becomes

$$\begin{aligned} \mathbb{P}[X < k+1] &< e^{k+1 - k \log n - (k+1) \log \left(\frac{k+1}{k \log n} \right)} \\ &= \tilde{O}(e^{-k \log n}) = \tilde{O}(n^{-k}), \end{aligned}$$

which proves the statement. \square

Our second lemma will reduce the estimation from the k th to the first moment of the number of point in the collector.

LEMMA 2. *If $X = \sum_{i=1}^n X_i$, where the X_i are independently distributed random variables with value in $\{0, 1\}$ and $\mathbb{E}[X] \geq 1$ then $\mathbb{E}[X^k] = O(\mathbb{E}[X]^k)$.*

PROOF. The statement is a special case of a classical inequality for sums of random variables [11, Th 2.12]; we give a simple, elementary, proof.

Expanding $X^k = (\sum_{i=1}^n X_i)^k$ we obtain

$$\begin{aligned} \mathbb{E}[X^k] &= \sum_{1 \leq i_1, i_2, \dots, i_k \leq n} \mathbb{E}[X_{i_1} \cdot X_{i_2} \dots X_{i_k}] \\ &= \sum_{\ell=1}^k \sum_{\substack{1 \leq i_1, i_2, \dots, i_\ell \leq n \\ |\{i_1, i_2, \dots, i_\ell\}| = \ell}} \mathbb{E}[X_{i_1} \cdot X_{i_2} \dots X_{i_\ell}]. \end{aligned}$$

Since the X_i 's are 0-1 variables, for any positive integers a_1, a_2, \dots, a_t

$$\mathbb{E}[X_1^{a_1} \cdot X_2^{a_2} \dots X_t^{a_t}] = \mathbb{E}[X_1 \cdot X_2 \dots X_t].$$

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

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

Since X_i and X_j are independent if $i \neq j$ the previous identity rewrites as

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

Thus,

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

and since

$$\sum_{1 \leq j_1, j_2, \dots, j_\ell \leq n} \mathbb{E}[X_{j_1}] \cdot \mathbb{E}[X_{j_2}] \dots \mathbb{E}[X_{j_\ell}] = \left(\sum_{i=1}^n \mathbb{E}[X_i] \right)^\ell = \mathbb{E}[X]^\ell$$

we finally obtain that

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

the last inequality following from the fact that $\mathbb{E}[X] \geq 1$. \square

3. CONVEX HULL OF A UNIFORM SAMPLE OF A BALL

As a warm-up we apply our technique to the classical problem of estimating the number of faces of a random polytope. Let $K = B(O, 1)$ be the unit ball in \mathbb{R}^d and let P denote a set of n points chosen independently and uniformly in K . We let H denote the geometric hypergraph induced on P by the set R of halfspaces in \mathbb{R}^d that intersect K .

THEOREM 3. *For any fixed $k \geq 1$ we have $\mathbb{E}[|H_{(k)}|] = \tilde{\Theta}\left(n^{\frac{d-1}{d+1}}\right)$.*

Theorem 3 immediately implies an $\tilde{O}\left(n^{\frac{d-1}{d+1}}\right)$ upper bound on the average number of faces of the convex hull of P . To obtain a matching lower bound, it suffices to remark that $H_{(1)}$ is exactly the set of vertices of that convex hull and that a polytope with v vertices has at least $\frac{v}{i+1}$ i -dimensional faces. The expected number of i -faces of the convex hull of P is known to be $\Theta\left(n^{\frac{d-1}{d+1}}\right)$ [12]; the known proof is long and computational (more than 7 pages long, still leaving substantial computations to the reader). In contrast, our proof is short and elementary.

PROOF. We break-up R in smaller range spaces by covering the space of directions \mathbb{S}^{d-1} of \mathbb{R}^d , envisioned as the unit sphere ∂K , by spherical caps of radius r which we denote by Sc_1, Sc_2, \dots, Sc_m . We choose r to be the radius of the spherical cap $H \cap \partial K$ where H is a halfspace such that $H \cap K$ contains on average $k \log n$ points of P . Specifically $r = \tilde{O}\left(n^{-\frac{1}{d+1}}\right)$ (Claim 3.1). We choose the cover so that $m = O(r^{1-d}) = \tilde{O}\left(n^{\frac{d-1}{d+1}}\right)$. We let R_i stand for the set of halfspaces in \mathbb{R}^d with inner normal in Sc_i (see Figure 2).

The witness W_i for R_i is defined as the halfspace cutting out Sc_i from ∂K ; the collector C_i is then defined as the union of the halfspaces in R_i that do not contain W_i . By construction of Sc_i , W_i contains $k \log n$ points of P on average and the probability that it contains less than k points is therefore $O(n^{-k})$ by Lemma 1. The volume of $C_i \cap K$ is $\tilde{O}\left(\frac{1}{n}\right)$ (Claim 3.2) so $\mathbb{E}[|C_i \cap P|] = \tilde{O}(1)$. Following Inequality (2), the expected number of elements of $H_{(k)}$ induced by a range of R_i is also $\tilde{O}(1)$. Summing over all subsets R_i 's of ranges, we get that $\mathbb{E}[|H_{(k)}|] = \tilde{O}(m) = \tilde{O}\left(n^{\frac{d-1}{d+1}}\right)$. Each witness

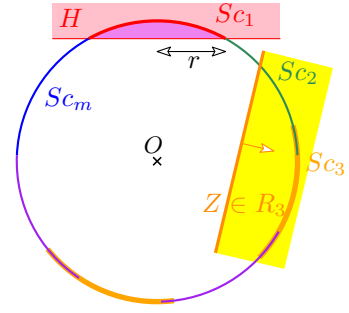


Figure 2: Proof of Theorem 3

region W_i contains an element of $H_{(k)}$ with high probability. In a minimal cover of ∂K by spherical caps of radius r the number of caps that cover a given point is, for any fixed dimension, $O(1)$. It follows that an element of $H_{(k)}$ can be contained in at most a constant number of witnesses and $\mathbb{E}[|H_{(k)}|]$ is also $\tilde{\Omega}(m)$. Altogether, $\mathbb{E}[|H_{(k)}|] = \tilde{\Theta}(m) = \tilde{\Theta}\left(n^{\frac{d-1}{d+1}}\right)$. \square

The proof used two elementary computations that we now detail.

CLAIM 3.1. $r = \tilde{O}\left(n^{-\frac{1}{d+1}}\right)$

PROOF. The halfspace H cuts out a spherical cap of radius r and height $\Theta(r^2)$ from ∂K , so

$$\text{Vol}(H \cap K) = \Theta\left(r^{d-1}\right) \cdot \Theta\left(r^2\right) = \Theta\left(r^{d+1}\right).$$

In order to have on average $k \log n$ points of P in $H \cap K$, the volume of this intersection should be $k \frac{\log n}{n}$. Thus $r = \Theta\left(\left(k \frac{\log n}{n}\right)^{\frac{1}{d+1}}\right) = \tilde{\Theta}\left(n^{-\frac{1}{d+1}}\right)$ (see Figure 3). \square

CLAIM 3.2. *The volume of $C_i \cap K$ is $\tilde{O}\left(\frac{1}{n}\right)$.*

PROOF. Let H denote a half-space with normal in Sc_i that does not contain $W_i \cap K$ (C_i is the union of all such half-spaces). Observe that $H \cap \partial K$ is a spherical cap with center in Sc_i . Moreover, for any half-space X the intersection $X \cap K$ is the convex hull of $X \cap \partial K$ (see Figure 4). The fact that H does not contain $W_i \cap K$ therefore implies that $H \cap \partial K$ does not contain $W_i \cap \partial K = Sc_i$. It follows that $H \cap \partial K$ has radius at most $2r$ and is thus contained in a spherical cap S' with same center as Sc_i and radius $3r$. Let H' denote the half-space such that $H' \cap \partial K = S'$ and observe that what precedes implies that $C_i \cap K$ is contained in $H' \cap K$. By the same argument as in the proof of Claim 3.1, $\text{Vol}(C_i \cap K) \leq \text{Vol}(H' \cap K) = \Theta\left((3r)^{d+1}\right) = \tilde{O}\left(\frac{1}{n}\right)$. \square

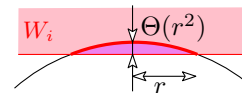


Figure 3: Proof of Claim 3.1

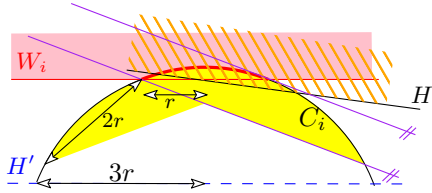


Figure 4: Proof of Claim 3.2

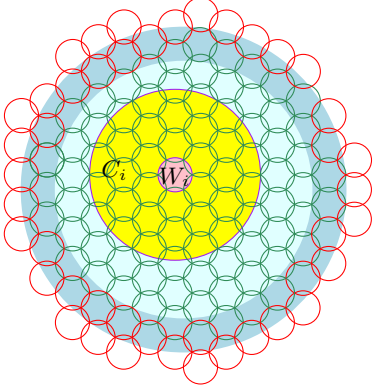


Figure 5: Proof of Theorem 4, balls intersecting K deeply inside.

4. DELAUNAY TRIANGULATION OF A UNIFORM SAMPLE OF A BALL IN \mathbb{R}^D

We continue with an application of our technique to another classical problem: estimating the number of faces of the Delaunay triangulation of a random set of points. As in the previous section, let K stand for the unit ball in \mathbb{R}^d , let P denote a set of n points distributed independently and uniformly in K . We let H denote the geometric hypergraph induced on P by the set R of balls in \mathbb{R}^d that intersect K .

THEOREM 4. For $k \in \{1, 2, \dots, d-1\}$ we have

$$\mathbb{E}[|H_{(k)}|] = \tilde{\Theta}(n).$$

Theorem 4 immediately implies an upper bound of $\tilde{O}(n)$ on the number of k -dimensional faces of the Delaunay triangulation of P , for $0 \leq k \leq d-1$, as any such face is in $H_{(k+1)}$. Since every vertex is contained in a k -dimensional face of the Delaunay triangulation, the number of such faces is at least $\frac{n}{k+1}$ and a matching lower bound is therefore obvious. A sharper bound of $\Theta(n)$ on the expected size of this Delaunay triangulation was previously established by Dwyer [8] but through considerably more tedious arguments.

PROOF. We proceed a bit differently from before and first fix the witness and the collector, and only then identify which subset of ranges they take care of (namely those ranges that contain the witness or are contained in the collector). The balls in R will be separated in three groups: the balls intersecting K “deeply inside”, the small balls close to the boundary of K , and the larger balls “almost tangent” to K .

Balls intersecting K deeply inside.

Let $h \in \mathbb{R}$ be such that the expected number of points of P falling in a ball of radius h contained in K is $k \log n$. It is clear that $h = O\left(\left(k \frac{\log n}{n}\right)^{1/d}\right) = \tilde{O}\left(n^{-\frac{1}{d}}\right)$. We cover K by balls W_1, W_2, \dots, W_m of radius h ; we can choose these balls so that $m = O(h^{-d}) = \tilde{O}(n)$ (see Figure 5). For $i \in \{1, 2, \dots, m\}$ we let C_i denote the ball with same center as W_i and radius $5h$. If $W_i \subseteq K$ then we let R_i stand for the set of balls that contain W_i or are contained in C_i ; else, if $W_i \not\subseteq K$ we simply let R_i be the set of balls contained in C_i .

If $W_i \subseteq K$ then W_i contains on average $k \log n$ points of P and, by Lemma 1, the probability that $|W_i \cap P| < k + 1$ is $O(n^{-k})$. Following Inequality (2), the expected number of elements of $H_{(k)}$ induced by a range of R_i is $O\left(\mathbb{E}[|P \cap C_i|^k]\right) = O\left((nh^d)^k\right) = \tilde{O}(1)$. When $W_i \not\subseteq K$, all ranges of R_i are contained in C_i and the expected number of elements of $H_{(k)}$ induced by a range of R_i is trivially $O\left(\mathbb{E}\left[\binom{|C_i \cap P|}{k}\right]\right)$ which is $O\left(\mathbb{E}[|P \cap C_i|^k]\right) = \tilde{O}(1)$ by Lemma 2. The number of elements of $H_{(k)}$ induced by a range of $\bigcup_{i=1}^m R_i$ therefore sums up to $\tilde{O}(m)$.

Small balls close to the boundary of K .

Any ball left in $R \setminus \bigcup_{i=1}^m R_i$ has radius at least $2h$ and does not intersect the ball $B(O, 1 - 4h)$ (Claim 4.1). We cover ∂K by $m' = O(h^{1-d}) = \tilde{O}\left(n^{1-\frac{1}{d}}\right)$ spherical caps $Sc_1, Sc_2, \dots, Sc_{m'}$ of radius h . We let A_i denote the cone with apex O supported by Sc_i . We let R_{m+i} denote the set of balls of radius between $2h$ and $4h$, centered in A_i that intersect K but do not intersect $B(O, 1 - 4h)$. We define C_{m+i} as the union of the balls in R_{m+i} (here we need no witness) (see Figure 6).

The volume of C_{m+i} is $O(h^d) = \tilde{O}\left(\frac{1}{n}\right)$. Since, for $1 \leq i \leq m'$, all ranges of R_{m+i} are contained in C_{m+i} the expected number of elements of $H_{(k)}$ induced by a range of R_{m+i} is trivially $O\left(\mathbb{E}\left[\binom{|C_{m+i} \cap P|}{k}\right]\right)$ which, by Lemma 2, is $O\left(\mathbb{E}[|P \cap C_{m+i}|^k]\right) = \tilde{O}(1)$. The number of elements of $H_{(k)}$ induced by a range of $\bigcup_{i=1}^{m'} R_{m+i}$ therefore sums up to $\tilde{O}(m')$.

Larger balls “almost tangent” to K .

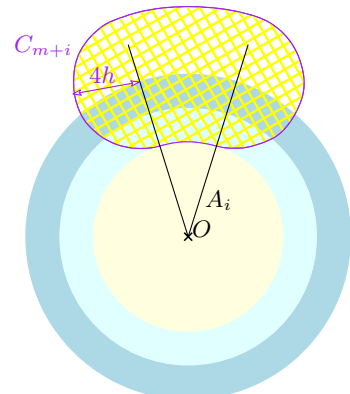


Figure 6: Proof of Theorem 4, small balls close to the boundary of K .

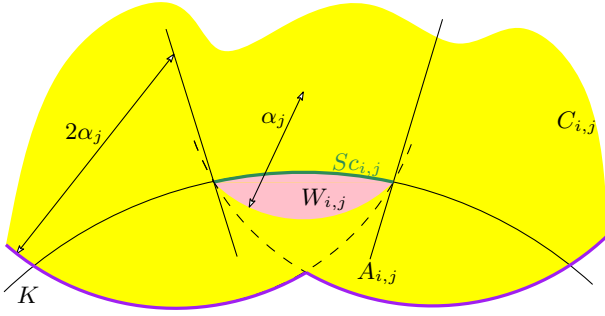


Figure 7: Proof of Theorem 4, larger balls “almost tangent” to K .

We group the balls in $R \setminus \bigcup_{i=1}^{m+m'} R_i$ in packets according to their radius. We let $M = \lceil \log_2 \frac{1}{4h} \rceil = \tilde{O}(1)$, let $\alpha_j = h \cdot 2^{j+2}$ for $j \in \{0, 1, \dots, M\}$ and put $\alpha_{M+1} = +\infty$. We cover ∂K by spherical caps $Sc_{1,j}, Sc_{2,j}, \dots, Sc_{m_j,j}$ of radius r_j ; here r_j is chosen so that the region $W_{i,j}$ defined as the intersection of K with the ball of radius α_j , centered outside of K and intersecting ∂K in $Sc_{i,j}$ contains on average $k \log n$ points of P (see Figure 7). Remark that $r_j = \tilde{O}\left(2^{\frac{j}{d+1}} n^{-\frac{1}{d}}\right)$ (Claim 4.2) and that $m_j = O(r_j^{1-d}) = \tilde{O}\left(\left(2^{-\frac{d-1}{d+1}}\right)^j n^{1-\frac{1}{d}}\right)$. We denote by $A_{i,j}$ the cone with apex O and supported by $Sc_{i,j}$. Finally, we define $R_{i,j}$ as the set of balls of radius between α_j and α_{j+1} that are centered in $A_{i,j}$, intersect K but do not intersect $B(O, 1 - 4h)$. We let $C_{i,j}$ denote the union of the balls of $R_{i,j}$ that do not contain $W_{i,j}$. The volume of $C_{i,j} \cap K$ is, again, $\tilde{O}\left(\frac{1}{n}\right)$ (Claim 4.3).

Since $W_{i,j}$ contains on average $k \log n$ points of P , the probability that $|W_i \cap P| < k + 1$ is $O(n^{-k})$ by Lemma 1. Following Inequality (2), the expected number of elements of $H_{(k)}$ induced by a range of $R_{i,j}$ is $O\left(\mathbb{E}[|P \cap C_i|]^k\right) = \tilde{O}(1)$. Altogether, the number of elements of $H_{(k)}$ induced by a range of $\bigcup_{j=1}^M \bigcup_{i=1}^{m_j} R_{i,j}$ therefore sums up to

$$\begin{aligned} \sum_{j=1}^M \tilde{O}(m_j) &= \tilde{O}\left(n^{1-\frac{1}{d}} \sum_{j=1}^M \left(2^{-\frac{d-1}{d+1}}\right)^j\right) \\ &= \tilde{O}\left(Mn^{1-\frac{1}{d}}\right) = \tilde{O}\left(n^{1-\frac{1}{d}}\right). \end{aligned}$$

Since the cases discussed above cover R ,

$$R = \left(\bigcup_{i=1}^m R_i\right) \cup \left(\bigcup_{i=1}^{m'} R_{m+i}\right) \cup \left(\bigcup_{j=1}^M \bigcup_{i=1}^{m_j} R_{i,j}\right),$$

we finally obtain

$$\mathbb{E}[|H_{(k)}|] = \tilde{O}(m) + \tilde{O}(m') + \tilde{O}\left(Mn^{1-\frac{1}{d}}\right) = \tilde{O}(n),$$

which concludes the proof. \square

The proof used three elementary computations that we now detail.

CLAIM 4.1. A ball in $R \setminus \bigcup_{i=1}^m R_i$ has radius at least $2h$ and does not intersect $B(O, 1 - 4h)$.

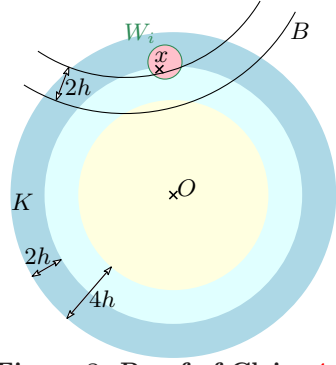


Figure 8: Proof of Claim 4.1

PROOF. Let $B \in R$. Since B intersects K and K is contained in $\bigcup_{j=1}^m W_j$, the ball B intersects W_i with $i \in \{1, 2, \dots, m\}$ (see Figure 8). If B has radius at most $2h$ then it is contained in C_i and it follows that $B \in R_i$. Now, assume that B has radius $r(B) > 2h$. If B intersects $B(O, 1 - 4h)$ then the ball with same center as B and radius $r(B) - 2h$ intersects $B(O, 1 - 2h)$ in some point x . There is some ball W_i with $i \in \{1, 2, \dots, m\}$ that contains the point x . Since W_i is completely contained both in K and in B , we have that $B \in R_i$. \square

CLAIM 4.2. $r_j = \tilde{O}\left(2^{\frac{j}{d+1}} n^{-\frac{1}{d}}\right)$.

PROOF. Let B stand for the ball of radius α_j that intersects ∂K in $Sc_{i,j}$. Let H denote the halfspace that intersects ∂K in $Sc_{i,j}$ (see Figure 9).

Since $\alpha_j \leq 1$ the volume of $W_{i,j}$ is at least the volume of $H \cap B$ and at most twice that volume. The distance between the center of B and the hyperplane bounding H is $\alpha_j - \Theta\left(\frac{r_j^2}{\alpha_j}\right)$. Thus,

$$\text{Vol}(W_{i,j}) = \Theta(\text{Vol}(H \cap B)) = \Theta\left(\frac{r_j^{d+1}}{\alpha_j}\right).$$

The condition $\text{Vol}(W_{i,j}) = k \frac{\log n}{n} \text{Vol}(K)$ hence translates into $r_j = \Theta\left(\left(2^{j+2} h k \frac{\log n}{n}\right)^{\frac{1}{d+1}}\right)$. Plugging $h = \tilde{\Theta}\left(n^{-\frac{1}{d}}\right)$ in that expression we get $r_j = \tilde{\Theta}\left(2^{\frac{j}{d+1}} n^{-\frac{1}{d}}\right)$ \square

CLAIM 4.3. The volume of $C_{i,j} \cap K$ is $\tilde{O}\left(\frac{1}{n}\right)$ for $j \in \{0, 1, \dots, p\}$ and $i \in \{0, 1, \dots, m_j\}$.

PROOF. We prove the claim by showing that the volume of $C_{i,j} \cap K$ is of the same order as that of $W_{i,j}$, which is $\tilde{O}\left(\frac{1}{n}\right)$

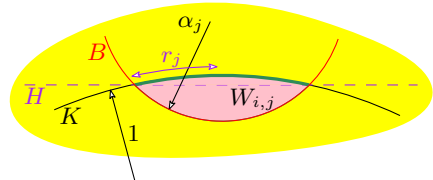


Figure 9: Proof of Claim 4.2

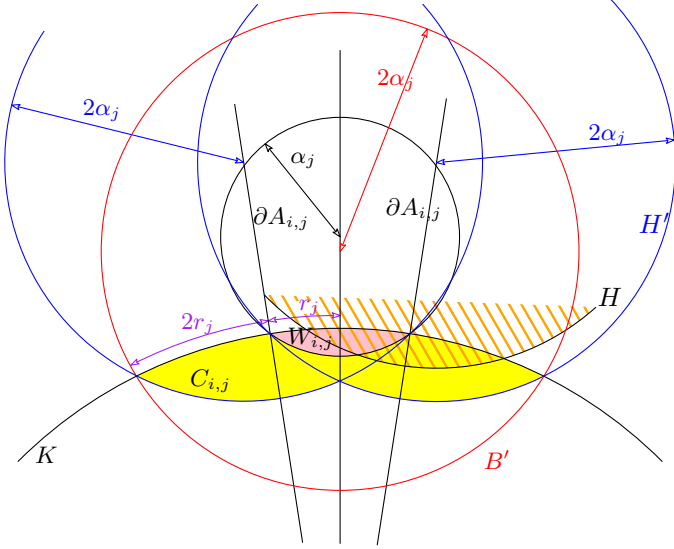


Figure 10: Proof of Claim 4.3

by construction. Similarly to the proof of Claim 3.2, let H denote a ball centered in $A_{i,j}$ and radius $2\alpha_j$ that does not contain $W_{i,j} \cap K$ ($C_{i,j}$ is the union of all such balls). Observe that $H \cap \partial K$ is a spherical cap with center in $S_{C_{i,j}}$ and not containing $S_{C_{i,j}}$, thus $H \cap \partial K$ is contained in a spherical cap S' with same center as $S_{C_{i,j}}$ and radius $3r_j$. Let B' denote the ball with radius $2\alpha_j$ such that $B' \cap \partial K = S'$. For any ball H of $R_{i,j}$, H verifies $H \cap K \subset H' \cap K$ where H' is a ball of radius $2\alpha_j$ of $R_{i,j}$ and whose center is as close as possible to the origin so that H' does not contain $W_{i,j}$ (see Figure 10). Observe that what precedes implies that $C_{i,j} \cap K$ is contained in $B' \cap K$. Then,

$$\text{Vol}(B' \cap K) \leq \Theta\left(\frac{(3r_j)^{d+1}}{2\alpha_j}\right) = \Theta(\text{Vol}(W_{i,j}))$$

by the same argument as in the proof of Claim 4.2. \square

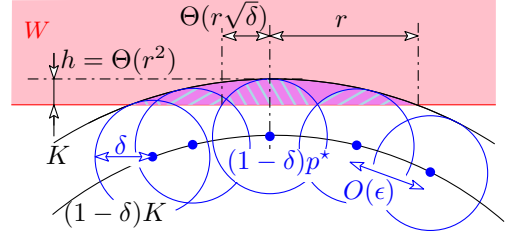


Figure 11: Proof of Claim 5.1

5. CONVEX HULL OF LOCAL PERTURBATIONS OF A POINT SET

We now apply our scheme to a new problem. In the spirit of smoothed analysis where the complexity of a problem is studied on average under some noise, we compute the size of the convex hull of perturbations of certain bad situations for the worst-case: points nicely sampled on the unit sphere ∂K of \mathbb{R}^d .

As in Section 3, H is the empty-region hypergraph induced by the set R of halfspaces in \mathbb{R}^d but with a different point set P . Let $P^* = \{p_i^*, 0 \leq i < n\}$ be an (ϵ, κ) -sample of ∂K for $\epsilon = n^{-\frac{1}{d}}$ and some constant κ , $x_i, 0 \leq i < n$ be n independent and identically distributed points in K , and $\delta \in \left[n^{-\frac{2}{d}}, \frac{1}{2}\right]$. The point set P is defined as $P = \{p_i = (1-\delta)p_i^* + \delta x_i\}$. Scaling the points in P^* before perturbing ensures that $P \subset K$.

THEOREM 5. *For any fixed $k \geq 1$ we have $\mathbb{E}[|H_{(k)}|] = \tilde{\Theta}\left((\sqrt{n})^{1-\frac{1}{d}} \left(\frac{1}{\sqrt{\delta}}\right)^{d-\frac{1}{d}}\right)$.*

PROOF. We break up R in m ranges and define witnesses and collectors as in Section 3, but using a different value for r . The value of r comes from the fact that $H \cap K$ has to contain on average $k \log n$ points, since the point distribution has changed this condition now yields $r = \tilde{O}\left(\delta^{\frac{1+d}{4d}} n^{-\frac{1}{2d}}\right)$

(Claim 5.1) and $m = O(r^{1-d}) = \tilde{O}\left(\delta^{\frac{1-d^2}{4d}} n^{\frac{d-1}{2d}}\right)$. Note that the expected number of points in a collector is $\tilde{O}(1)$, as in the witness, which can be proved with exactly the same argument as in Claim 3.2. We then conclude $\mathbb{E}[|H_{(k)}|] = \tilde{\Theta}(m) = \tilde{\Theta}\left(\delta^{\frac{1-d^2}{4d}} n^{\frac{d-1}{2d}}\right)$. \square

CLAIM 5.1. $r = \tilde{\Theta}\left(\delta^{\frac{1+d}{4d}} n^{-\frac{1}{2d}}\right)$.

PROOF. Let W be a witness that cuts out a spherical cap of radius $r \geq \epsilon$ and height $h = \Theta(r^2)$ from ∂K . Since P^* is an (ϵ, κ) sample the number of points $p^* \in P^*$ such that $((1-\delta)p^* + \delta K) \cap W \neq \emptyset$ is $\Theta\left(\frac{r^{d-1}}{\epsilon^{d-1}}\right)$ (see Figure 11). For such a point the probability that the perturbed point p falls in W is $\frac{\text{Vol}(W \cap ((1-\delta)p^* + \delta K))}{\text{Vol}(\delta K)} \leq \Theta\left(\frac{(r\sqrt{\delta})^{d-1} r^2}{\delta^d}\right) = \Theta\left(\delta^{-\frac{d+1}{2}} r^{d+1}\right)$. Notice that if the height of the cap $((1-\delta)p^* + \delta K) \cap W$ is at least $\frac{h}{2}$ the probability is now lower bounded by $\Theta\left(\delta^{-\frac{d+1}{2}} r^{d+1}\right)$ (with a different constant) and

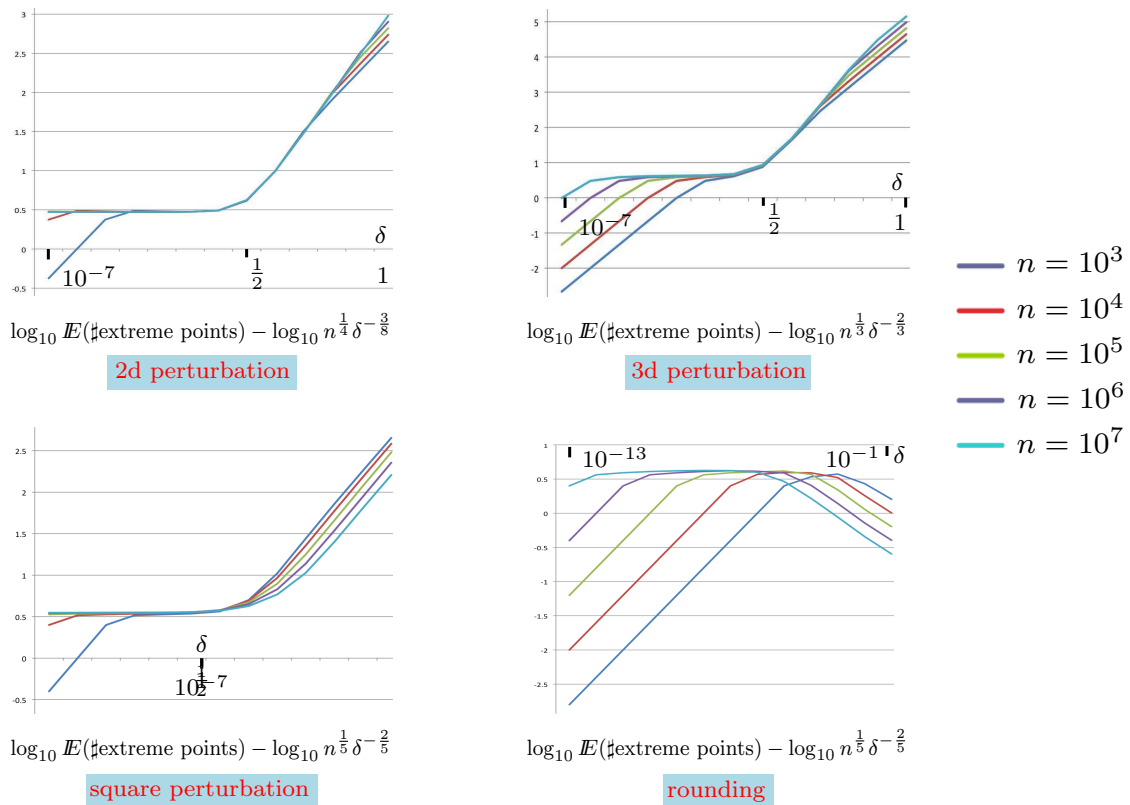


Figure 12: Experimental results

the number of points $p^* \in P^*$ satisfying that property is also $\Theta\left(\frac{r^{d-1}}{\epsilon^{d-1}}\right)$. Summing these probabilities we get that the expected number of points of P in W is $\Theta\left(\frac{r^{d-1}}{\epsilon^{d-1}} \delta^{-\frac{d+1}{2}} r^{d+1}\right) = \Theta\left(n \delta^{-\frac{d+1}{2}} r^{2d}\right)$, since we want this quantity to be $k \log n$ we set $r = \tilde{\Theta}\left(\delta^{\frac{1+d}{4d}} n^{-\frac{1}{4d}}\right)$. The condition $r \geq \epsilon$ yields the lower bound ϵ^2 on δ , the upper bound $\frac{1}{2}$ on δ ensures that $(1-\delta)p^*$ lies on a sphere of radius lower bounded by a constant. \square

Experimental remarks

We could confirm experimentally our theoretical estimates. More precisely, using CGAL [4], we simulated a uniform noise of amplitude δ on a set P^* of n points in 2 and 3 dimensions, and counted the number of extreme points of the perturbed set P .

We consider $n = 10^i$ for $i = 3$ to 7 and $\delta = 10^j$ for $j = -7$ to 5. Our count of the number of extreme points is averaged over 1000 trials for $n \leq 10^6$ and over 100 trials for $n = 10^7$. In the plane, we applied the perturbation to the vertices of a regular n -gon (see Figure 12-top-left); in 3D, we applied the perturbation to a random sample of the unit sphere (see Figure 12-top-right).

We observe on Figure 12 that all curves present a significant plateau, indicating that the behavior of the number of extreme points is correctly estimated by our formulas. More precisely, in the range $10^3 \leq n \leq 10^7$ these behaviors are modeled within a margin of error of 10% by the table (see [2] for details).

Dimension	2	3
Estimation	$3\delta^{-\frac{3}{8}} n^{\frac{1}{4}}$	$4.2\delta^{-\frac{2}{3}} n^{\frac{1}{3}}$
Range of δ	$[\frac{10}{n^2}, 0.1]$	$[\frac{100}{n}, 0.1]$

We can also consider perturbation schemes other than points uniformly sampled from a ball. In a preliminary report [2], we used the same witness collector technique to analyze perturbation with a square support in two dimensions and obtained a bound of $\tilde{\Theta}\left(\delta^{-\frac{2}{5}} n^{\frac{1}{5}}\right)$ for the expected size of the convex hull. An experimental confirmation of this bound is depicted on Figure 12-bottom-left.

It is also interesting to remark that the same formula predicts accurately the size of the convex hull of the vertices of the regular n -gon when a grid is “rounded down” to a coarser grid. The comparison between the experimental simulations and the complexity law under square perturbation is presented in Figure 12-bottom-right. For large δ , the two quantities cease to be related: when snap-rounding to a very coarse grid, the fact many points collide entails a significant reduction of the size of the convex hull (see [2] for a more detailed discussion).

6. CONCLUSION

We believe that the simple technique presented in this paper will find more applications in the near future; work in progress includes the analysis, by the authors of the present paper, of the expected complexity of the Delaunay triangulation of perturbations of certain point configurations in

3D, and the study, by Glisse et al. [10], of the average complexity of worst-case silhouette of a random polytope. Also, while the extra polylogarithmic factors do not seem a high price to pay given the simplicity of the method, the work of Glisse et al. [10] indicates that they can even be disposed of by considering *random* witnesses and collectors.

While it may seem natural that the smoothed complexity of planar convex hull would be achieved by perturbation of points initially in convex position, Theorem 5 suggests caution as for constant amplitude δ , starting with all points on the unit-radius regular n -gon gives an expected complexity of $\tilde{\Theta}\left(n^{\frac{1}{4}}\right)$ which is less than the expected complexity of $\Theta\left(n^{\frac{1}{3}}\right)$ obtained by starting with all points at the origin.

Let us finally remark that the proofs of Sections 3, 4 and 5 generalize mutatis mutandis to the case where K is a smooth convex body with “bounded curvature” in the following sense: there exist $\rho_{\min} > 0$ and $\rho_{\max} < \infty$ such that at every point $x \in \partial K$, for every 2-plane Π containing x and the normal to ∂K in x , the curvature of $\Pi \cap \partial K$, seen as a plane curve, at x is at least ρ_{\min} and at most ρ_{\max} . Some constants appearing in the analysis change, e.g. certain ratios of volumes become $O\left(\left(\frac{\rho_{\min}}{\rho_{\max}}\right)^{d+1}\right)$ instead of 3^{d+1} or 5^{d+1} , but this has no consequence on the estimates of orders of magnitude.

Acknowledgment

This work was initiated during the 8th McGill - INRIA Workshop on Computational Geometry at the Bellairs institute. The authors wish to thank all the participants for creating a pleasant and stimulating atmosphere, Nina Amenta and Sarel Har-Peled for useful discussions at the early stage of this work, and in particular Dominique Attali who was involved in the preliminary work [2].

7. REFERENCES

- [1] D. Attali, J. D. Boissonnat, and A. Lieutier. Complexity of the Delaunay triangulation of points on surfaces the smooth case. In *Proc. 19th Sympos. Comput. Geom.*, 201–210, 2003. doi:10.1145/777792.777823.
- [2] Dominique Attali, Olivier Devillers, and Xavier Goaoc. The effect of noise on the number of extreme points. Research Report 7134, INRIA, 2009. <http://hal.inria.fr/inria-00438409/>.
- [3] Frédéric Cazals and Joachim Giesen. Delaunay triangulation based surface reconstruction. In Jean-Daniel Boissonnat and Monique Teillaud, editors, *Effective Computational Geometry for Curves and Surfaces*, 231–276. Springer-Verlag, Mathematics and Visualization, 2006. doi:10.1007/978-3-540-33259-6_6.
- [4] CGAL, Computational Geometry Algorithms Library. <http://www.cgal.org>.
- [5] V. Damerow and C. Sohler. Extreme points under random noise. In *Proc. 12th European Sympos. Algorithms*, 264–274, 2004. doi:10.1007/978-3-540-30140-0_25.
- [6] Mark 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.
- [7] Mark de Berg, Herman Haverkort, and Constantinos P. Tsirogiannis. Visibility maps of realistic terrains have linear smoothed complexity. In *Proc. of the 25th Symposium on Computational geometry*, 163–168, 2009. doi:10.1145/1542362.1542397.
- [8] 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.
- [9] 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.
- [10] Marc Glisse, Sylvain Lazard, Julien Michel, and Marc Pouget. Silhouette of a random polytope, 2012. manuscript.
- [11] Valentin Petrov. *Limit Theorems of Probability Theory. Sequence of Independent Random Variables*. Number 4 in Oxford studies in probability. Clarendon Press, 1995. <http://www.citeulike.org/group/2854/article/1615442>.
- [12] H. Raynaud. Sur l’enveloppe convexe des nuages de points aleatoires dans R^n . *J. Appl. Probab.*, 7:35–48, 1970. <http://www.jstor.org/stable/10.2307/3212146>.
- [13] 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.
- [14] Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis: an attempt to explain the behavior of algorithms in practice. *Commun. ACM*, 52(10):76–84, 2009. doi:10.1145/1562764.1562785.