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. [Research Report] RR-8168, INRIA. 2012. hal-00761171

HAL Id: hal-00761171
https://inria.hal.science/hal-00761171
Submitted on 5 Dec 2012

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,
Marc Glisse, Xavier Goaoc
Complexity analysis of random geometric structures made simpler

Olivier Devillers∗, Marc Glisse†, Xavier Goaoc‡

Project-Teams Geometrica and Vegas

Research Report n° 8168 — December 2012 — 17 pages

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 \( \tilde{\Theta}(n^{\frac{d}{d-1}}) \) and \( \tilde{\Theta}(n) \). We then prove that if we start with \( n \) points well-spread on a sphere, e.g., an \((\epsilon, \kappa)\)-sample of that sphere, and perturb that sample by moving each point randomly and uniformly within distance at most \( \delta \) of its initial position, then the expected complexity of the convex hull of the resulting point set is \( \tilde{\Theta}\left((\sqrt{n})^{1-\frac{d}{2}}\left(\frac{1}{\sqrt{n}}\right)^{d-\frac{d}{2}}\right) \).

Key-words: Convex hull, Voronoi diagram, point distribution, smooth analysis

* The work in this paper has been partially supported by ANR blanc PRESAGE (ANR-11-BS02-003)

∗ Projet Geometrica, INRIA Sophia Antipolis - Méditerranée
† Projet Geometrica, INRIA Saclay - Île de France
‡ Projet Vegas, INRIA Nancy - Grand Est
Analyse simplifiée de la complexité de structures géométriques aléatoires

Résumé : L’analyse en moyenne de structure de données et d’algorithmes géométriques est fréquemment utilisée en géométrie algorithmique, un domaine où l’analyse dans le cas le pire est souvent très pessimiste. La difficulté de ce type d’analyse fait que les modèles probabilistes utilisés restent simples et relativement éloignées de données réalistes. Nous présentons une nouvelle approche pour l’analyse des structures géométriques. Nos résultats sont seulement à des facteurs logarithmiques près, mais notre méthode est plus simple que les classiques du domaine et nous réussissons à analyser de nouveau type de distribution liée à la smooth analysis.

Nous illustrons notre méthode sur deux structures classiques: l’enveloppe convexe et la triangulation de Delaunay. Plus précisément, nous démontrons simplement le fait, classique, que $n$ points uniformément distribués dans une boule de $\mathbb{R}^d$ ont une enveloppe convexe et une triangulation de Delaunay dont l’espérance de la taille est respectivement $\tilde{\Theta}(n^{\frac{d-1}{d}})$ et $\Theta(n)$. Nous démontrons ensuite que si on se donne ensemble de $n$ points bien distribués sur une sphère, par exemple un $\epsilon, \kappa$-échantillon de la sphère, et qu’on le perturbe ensuite en déplaçant chaque point uniformément d’une distance $\delta$ à partir de sa position initiale, alors l’espérance de la taille de l’enveloppe convexe de ces points est $\tilde{\Theta} \left( (\sqrt{n})^{1-\frac{1}{d}} \left( \frac{1}{\sqrt{\pi}} \right)^{d-\frac{1}{2}} \right)$.

Mots-clés : Enveloppe convexe, diagramme de Voronoï, distribution de points, analyse lisse
1 Introduction

Understanding the complexity of geometric structures (Delaunay triangulations, convex hulls, arrangements...) is a 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 \( \tilde{\Theta}(n^{\frac{d+1}{d-1}}) \) bound\(^1\) on the expected number of \( k \)-dimensional faces of the convex hull of \( n \)

---

\(^1\) The notations \( O, \tilde{O}, \tilde{\Theta}, \tilde{\Omega} \) refer to asymptotic bounds up to a logarithmic factor, e.g. \( f(n) = \tilde{O}(n) \) means \( \exists a, b \in \mathbb{Z}, f(n) = O(n \log^a n) \) and \( f(n) = \tilde{\Omega}(n \log^b n) \).
points uniformly distributed in a smooth convex body \( K \subset \mathbb{R}^d \),

(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 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 \cite{12} and Dwyer \cite{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 \( \delta \) of a fixed point set \( P^* \) is the random point set obtained by choosing a point uniformly in each \((L^p)\) ball of radius \( \delta \) centered in a point of \( P^* \). In this direction, we obtain:

(iii) a \( \tilde{\Theta}\left(\sqrt{n}^{-\frac{1}{2}}\left(\frac{1}{5}\delta^{-\frac{d}{2}}\right)^{d-\frac{1}{2}}\right) \) bound on the expected number of \( k \)-dimensional faces of the convex hull of a \( L^2\)-perturbation of amplitude \( \delta \in \left[n^{-\frac{2}{d}}, 1\right] \) of an \( \left(n^{-\frac{1}{d}}, \kappa\right) \) sample of a smooth convex body in \( \mathbb{R}^d \); we also obtain a bound of \( \tilde{\Theta}\left(\sqrt{n}^{\frac{1}{3}}\left(\frac{1}{5}\delta^{-\frac{d}{2}}\right)\right) \) on the expected number of vertices of the convex hull of a \( L^\infty\)-perturbation of amplitude \( \delta \in \left[n^{-2}, 1\right] \) of an \( \left(n^{-1}, \kappa\right) \) sample of a smooth convex body in the plane.

(Recall that an \((\epsilon, \kappa)\)-sample of a set \( \Gamma \) is a point set \( P \subset \Gamma \) such that any ball of radius \( \epsilon \) centered on \( \Gamma \) contains between 1 and \( \kappa \) points from the sample \( P \).) The only previous bound in the direction of (iii) is due to Damerow and Sohler \cite{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 \( O\left(n^{\frac{1}{2}}\delta^{-\frac{d}{2}}\right) \) under \( L^\infty \) perturbation.

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 \( H_\kappa \) 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 Figure 1 form an edge in the hypergraph induced by halfplanes but are not an edge of the convex hull. Thus, \((\mathbb{R}^n)\) independently distributed, \(E\) eventually amounts to bounding from above the \(k\) contains more than \(W\) by inflating \(W\) lower tails implies that the probability that \(W\) becomes \(\mathbb{R}\) induced by \(R\) on \(P\). 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

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

since

\[
(1 - p) \mathbb{E} \left[ \left( \binom{|P \cap C'|}{k} \right) \left| W' \cap P > k \right| \right] = \mathbb{E} \left[ \left( \binom{|P \cap C'|}{k} \right) \right] - p \mathbb{E} \left[ \left( \binom{|P \cap C'|}{k} \right) \left| W' \cap P \leq k \right| \right] \leq \mathbb{E} \left[ \left( \binom{|P \cap C'|}{k} \right) \right].
\]

we get

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

To make the contribution \(p \left( \binom{|P|}{k} \right)\) negligible we want that \(p = O \left( \frac{n}{k} \right)\). 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 \(\left( \binom{|P \cap C'|}{k} \right)\).

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, \(\mathbb{E} \left[ (|P \cap C'|)^k \right] = O \left( \mathbb{E} \left[ |P \cap C'| \right]^k \right),\)

see Lemma 2, and Inequality (1) becomes

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

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

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 \(P\).
This will suffice for our purpose as we care only for the asymptotic orders of magnitude of
\(E [ |H_k|] \).

For this approach to work effectively, we break up the set \(R \) of ranges into \(m \) pieces \(R = R_1 \cup R_2 \cup \ldots \cup R_m \) 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 \(O(1) \), yielding
\(E [ |H_k|] = \tilde{O}(m) \). Since each witness contains some element of \(H_k \) with high probability,
if no more than \(O(1) \) witnesses have a common intersection then \(E [ |H_k|] = \Omega(m) \) 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 \(O(n^{-k}) \).

**Proof.** Let \(X_i \) be the indicator function of the event that the \(i^{th} \) point from \(P \) belongs to \(W \).
We write \(X = X_1 + \ldots + X_n \) and let \(\alpha = E [X] \). Chernoff’s bound for lower tails yields that for any \(\delta \in (0,1) \)
\[
P [ X < (1 - \delta)\alpha ] < \left( \frac{e^{-\delta}}{(1 - \delta)^{1+\delta}} \right)^\alpha .
\]
In particular, setting \(\delta = 1 - \frac{k+1}{k \log n} \) we get that \((1 - \delta)\alpha = k + 1 \) and:
\[
P [ X < k + 1 ] < \left( \frac{e^{-\delta}}{(1 - \delta)^{1+\delta}} \right) = e^{-\alpha - \frac{k+1}{k \log n} \alpha} \left( \frac{k + 1}{k \log n} \right)^{-\alpha - \frac{k+1}{k \log n}}.
\]
Since \(\alpha = k \log n \), the previous bound becomes
\[
P [ X < k + 1 ] < e^{k+1-k \log n -(k+1) \log \left( \frac{k+1}{k \log n} \right)} = \tilde{O} \left( e^{-k \log n} \right) = \tilde{O} \left( n^{-k} \right),
\]
which proves the statement.

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 \(E [X] \geq 1 \) then \(E [X^k] = O \left( E [X]^k \right) \).

**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
\[
E [X^k] = \sum_{1 \leq i_1, i_2, \ldots, i_k \leq n} E [X_{i_1} \cdot X_{i_2} \ldots X_{i_k}] = \sum_{\ell=1}^k \sum_{1 \leq i_1, i_2, \ldots, i_k \leq n} \sum_{|\{i_1, i_2, \ldots, i_k\}| = \ell} E [X_{i_1} \cdot X_{i_2} \ldots X_{i_k}] .
\]
Since the \(X_i \)’s are 0-1 variables, for any positive integers \(a_1, a_2, \ldots, a_t \)
\[
E [X_{i_1}^{a_1} \cdot X_{i_2}^{a_2} \ldots X_{i_t}^{a_t}] = E [X_{i_1} \cdot X_{i_2} \ldots X_{i_t}] .
\]
Letting \( p(\ell, k) \) denote the number of partition of \( \{1, 2, \ldots, k\} \) in subsets of size \( \ell \), we can thus write
\[
\mathbb{E} \left[ X^k \right] = \sum_{\ell=1}^{k} \sum_{1 \leq i_1 < i_2 < \ldots < i_\ell \leq n} p(\ell, k) \mathbb{E}[X_{i_1} \cdot X_{i_2} \ldots X_{i_\ell}] .
\]

Since \( X_i \) and \( X_j \) are independent if \( i \neq j \) the previous identity rewrites as
\[
\mathbb{E} \left[ X^k \right] = \sum_{\ell=1}^{k} \left( p(\ell, k) \sum_{1 \leq i_1 < i_2 < \ldots < i_\ell \leq n} \mathbb{E}[X_{i_1}] \cdot \mathbb{E}[X_{i_2}] \ldots \mathbb{E}[X_{i_\ell}] \right) .
\]

Thus,
\[
\mathbb{E} \left[ X^k \right] \leq \sum_{\ell=1}^{k} \left( p(\ell, k) \sum_{1 \leq i_1 < i_2 < \ldots < i_\ell \leq n} \mathbb{E}[X_{i_1}] \cdot \mathbb{E}[X_{i_2}] \ldots \mathbb{E}[X_{i_\ell}] \right)
\]
and since
\[
\sum_{1 \leq j_1, j_2, \ldots, j_\ell \leq n} \mathbb{E}[X_{j_1}] \cdot \mathbb{E}[X_{j_2}] \ldots \mathbb{E}[X_{j_\ell}] = \left( \sum_{i=1}^{n} \mathbb{E}[X_i] \right)^\ell = \mathbb{E}[X]^\ell
\]
we finally obtain that
\[
\mathbb{E} \left[ X^k \right] \leq \sum_{\ell=1}^{k} p(\ell, k) \mathbb{E}[X]^\ell \leq \left( \sum_{\ell=1}^{k} p(\ell, k) \right) \mathbb{E}[X]^k
\]
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 \( \mathbb{E} \left[ |H(k)| \right] = \widetilde{\Theta} \left( n^{\frac{d}{d+1}} \right) \).

Theorem 3 immediately implies an \( \widetilde{O} \left( n^{\frac{d}{d+1}} \right) \) 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 \( \frac{v}{d+1} \) \( i \)-dimensional faces. The expected number of \( i \)-faces of the convex hull of \( P \) is known to be \( \Theta \left( n^{\frac{i}{d+1}} \right) \) [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 \( S^{d-1} \) of \( \mathbb{R}^d \), envisioned as the unit sphere \( \partial K \), by spherical caps of radius \( r \) which we denote by \( Sc_1, Sc_2, \ldots, Sc_m \). 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 \( r = \widetilde{O} \left( n^{-\frac{1}{d+1}} \right) \) (Claim 3.1). We choose the
cover so that \( m = O(r^{1-d}) = \tilde{O} \left( n^{\frac{d}{d+1}} \right) \). We let \( R_i \) stand for the set of halfspaces in \( \mathbb{R}^d \) with inner normal in \( S_{c_i} \) (see Figure 2).

The witness \( W_i \) for \( R_i \) is defined as the halfspace cutting out \( S_{c_i} \) from \( \partial K \); 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 \( S_{c_i} \), \( W_i \) contains \( k \log n \) points on average and the probability that it contains less than \( k \) points is therefore \( O(n^{-k}) \) by Lemma 1. The volume of \( C_i \) 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 \( \mathbb{E} |H(k)| = \tilde{O} \left( m \right) = \tilde{O} \left( \frac{n^{d-1}}{\pi^{d+1}} \right) \). Each witness region \( W_i \) contains an element of \( H(k) \) with high probability. In a minimal cover of \( \partial K \) 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 \( \tilde{O}(m) \). Altogether, \( \mathbb{E} |H(k)| = \tilde{O}(m) = \tilde{O} \left( n^{\frac{d-1}{d+1}} \right) \).

The proof used two elementary computations that we now detail.

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

Proof. The halfspace \( H \) cuts out a spherical cap of radius \( r \) and height \( \Theta(r^2) \) from \( \partial K \), 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 \( k \log n \frac{n}{n} \). Thus \( r = \Theta \left( \left( k \log n \right)^{\frac{1}{d+1}} \right) = \tilde{O} \left( n^{-\frac{1}{d+1}} \right) \) (see Figure 3).
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, Vol $(C_i \cap K) \leq \text{Vol } (H' \cap K) = \Theta \left( (3r)^{d+1} \right) = \tilde{O} \left( \frac{1}{n} \right)$.  

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 $k \in \{1, 2, \ldots, d-1\}$ we have $E \left[ |H_{(k)}| \right] = \tilde{O} (n)$.

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 $h = O \left( \left( \frac{k \log n}{n} \right)^{1/d} \right) = \tilde{O} \left( n^{-\frac{1}{d}} \right)$. We cover $K$ by balls
$W_1, W_2, \ldots, W_m$ of radius $h$; we can choose these balls so that $m = O(h^{-d}) = \tilde{O}(n)$ (see Figure 5).

For $i \in \{1, 2, \ldots, m\}$ we let $C_i$ denote the ball with same center as $W_i$ and radius $3h$. 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 $O\left((nh^d)^k\right) = \tilde{O}(1)$. 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 $O\left(\mathbb{E}\left[|P \cap C_i|^k\right]\right)$ which is $O\left(\mathbb{E}\left[|P \cap C_i|^k\right]\right) = \tilde{O}(1)$ by Lemma 2.

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 $\partial K$ by $m' = O(h^{1-d}) = \tilde{O}\left((n^{-\frac{d}{2}})\right)$ spherical caps $S_{c_1}, S_{c_2}, \ldots, S_{c_{m'}}$ of radius $h$. We let $A_i$ denote the cone with apex $O$ supported by $S_{c_i}$. We let $R_{m+1}$ 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 $C_{m+1}$ as the union of the balls in $R_{m+1}$ (here we need no witness) (see Figure 6).
We group the balls in \( L \), the larger balls “almost tangent” to \( K \) of elements of \( O \). We finally obtain \( R \) denote the union of the balls of \( \alpha \) and \( \beta \) induced by a range of \( m \) balls centered in \( C \). The volume of \( C \) is, again, \( \tilde{O} \left( \frac{h^d}{n} \right) \). Since, for \( 1 \leq i \leq m \), all ranges of \( R \) are contained in \( C \), the expected number of elements of \( H(k) \) induced by a range of \( m \) balls is trivially \( O \left( \mathbb{E} \left[ |P \cap C|^{m} \right] \right) = \tilde{O}(1) \). 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 \( K \).

We group the balls in \( R \)\( \cup_{i=1}^{m'} R_{i} \) in packets according to their radius. We let \( M = \left[ \log \frac{1}{n} \right] = \tilde{O}(1) \), let \( \alpha_j = h \cdot 2^{j+2} \) for \( j \in \{0, 1 \ldots M\} \) and put \( \alpha_{M+1} = +\infty \). We cover \( K \) by spherical caps \( \cap_{j \in \{0, 1 \ldots M\}} \alpha_j \) of radius \( r_j \), here \( r_j \) is chosen so that the region \( W_{i,j} \) is the intersection of \( K \) with the ball of radius \( \alpha_j \), centered outside of \( K \) and intersecting \( \partial K \) in \( \cap_{j \in \{0, 1 \ldots M\}} \alpha_j \) contains on average \( k \log n \) points of \( P \) (see Figure 7). Remark that \( r_j = \tilde{O} \left( 2^{\frac{1}{M+n} n^{-\frac{1}{2}}} \right) \), \( \log n \) points of \( P \), the probability that \( |W_{i,j} \cap P| < k + 1 \) 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 \( O \left( \mathbb{E} \left[ |P \cap C|^{k} \right] \right) = \tilde{O}(1) \). Altogether, the number of elements of \( H(k) \) induced by a range of \( \bigcup_{i=1}^{M} \bigcup_{j=1}^{m} R_{i,j} \) therefore sums up to

\[
\sum_{j=1}^{M} \tilde{O}(m_j) = \tilde{O} \left( n^{1 \frac{1}{2}} \sum_{j=1}^{M} \left( 2^{\frac{1}{M+n} n^{0.5}} \right)^{k} \right) = \tilde{O} \left( M n^{1 \frac{1}{2}} \right) = \tilde{O} \left( n^{1 \frac{1}{2}} \right).
\]

Since the cases discussed above cover \( R \),

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

we finally obtain

\[
\mathbb{E} \left[ |H(k)| \right] = \tilde{O}(m) + \tilde{O}(m') + \tilde{O} \left( M n^{1 \frac{1}{2}} \right) = \tilde{O}(n),
\]

which concludes the proof.

\[\square\]

Figure 7: Proof of Theorem 4, larger balls “almost tangent” to \( K \).
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 $B(O, 1 - 4h)$.

Proof. Let $B \in R$. Since $B$ intersects $K$ and $K$ is contained in $\bigcup_{i=1}^{m} W_j$, the ball $B$ intersects $W_i$ with $i \in \{1, 2, \ldots, m\}$ (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 $i \in \{1, 2, \ldots, m\}$ that contains the point $x$. Since $W_i$ is completely contained both in $K$ and in $B$, we have that $B \in R_i$.

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

Proof. Let $B$ stand for the ball of radius $\alpha_j$ that intersects $\partial K$ in $S_{c_{i,j}}$. Let $H$ denote the halfspace that intersects $\partial K$ in $S_{c_{i,j}}$ (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 $\alpha_j - \Theta \left( \frac{r_j}{\alpha_j} \right)$.

Thus,

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

The condition $\text{Vol} (W_{i,j}) = \frac{k \log n}{n} \text{Vol} (K)$ therefore translates into $r_j = \Theta \left( \left( 2^{j+2} \frac{k \log n}{n} \right)^{\frac{1}{d+1}} \right)$.

Plugging $h = \Theta \left( n^{-\frac{1}{2}} \right)$ in that expression we get $r_j = \tilde{O} \left( 2^{\frac{d}{2}} n^{-\frac{1}{2}} \right)$.

Figure 8: Proof of Claim 4.1

Figure 9: Proof of Claim 4.2
Claim 4.3. The volume of $C_{i,j} \cap K$ is $\tilde{O} \left( \frac{1}{n} \right)$ for $j \in \{0, 1, \ldots, p\}$ and $i \in \{0, 1, \ldots, m_j\}$.

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 $\tilde{O} \left( \frac{1}{n} \right)$ 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). 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.

\[ \square \]

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 $\partial K$ of $\mathbb{R}^d$. 

RR n° 8168
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 \( P^* = \{p_i^*, 0 < i < n\} \) be an \((\epsilon, \kappa)\)-sample of \( \partial K \) for \( \epsilon = n^{-\epsilon/\delta} \) and some constant \( \kappa \), \( x_i, 0 \leq i < n \) be \( n \) independent and identically distributed points in \( K \), and \( \delta \in \left[ n^{-\epsilon/\delta}, \frac{1}{2} \right] \). 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 \( \mathbb{E}[|H_{(k)}|] = \Theta \left((\sqrt{n})^{1-\frac{d}{2}} \left(\frac{1}{\sqrt{\delta}}\right)^{d-\frac{d}{2}}\right) \).

**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 \( r = \tilde{O} \left(\delta^{\frac{d+1}{2\epsilon}} n^{-\frac{1}{2d}}\right) \) (Claim 5.1) and \( m = O(r^{1-d}) = \tilde{O} \left(\delta^{\frac{d+1}{2\epsilon}} n^{\frac{d+1}{2d}}\right) \). 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 \( \mathbb{E}[|H_{(k)}|] = \tilde{O}(m) = \tilde{O} \left(\delta^{\frac{d+1}{2\epsilon}} n^{\frac{d+1}{2d}}\right) \).

**Claim 5.1.** \( r = \Theta \left(\delta^{\frac{d+1}{2\epsilon}} n^{-\frac{1}{2d}}\right) \).

**Proof.** Let \( W \) be a witness that cuts out a spherical cap of radius \( r \geq \epsilon \) and height \( h = \Theta(r^2) \) from \( \partial K \). Since \( P^* \) is an \((\epsilon, \kappa)\) sample the number of points \( p^* \in P^* \) such that \( (1-\delta)p^* + \delta K \cap W \neq \emptyset \) is \( \Theta \left(\frac{d+1}{2\epsilon}\right) \) (see Figure 11). For such a point the probability that the perturbed point \( p \) falls in \( W \) is \( \frac{\text{Vol}(W \cap (p^* + \delta K))}{\text{Vol}(\partial K)} \leq \Theta \left(\frac{(\sqrt{2})^{d-1}r^{d+1}}{\delta}\right) = \Theta \left(\delta^{-\frac{d+1}{2\epsilon}} r^{d+1}\right) \). Notice that if the height of the cap \( (1-\delta)p^* + \delta K \cap W \) is at least \( \frac{1}{2} \) the probability is now lower bounded by \( \Theta \left(\delta^{-\frac{d+1}{2\epsilon}} r^{d+1}\right) \) (with a different constant) and the number of points \( p^* \in P^* \) satisfying that property is also \( \Theta \left(\frac{(d+1)}{\epsilon}\right) \). Summing these probabilities we get that the expected number of points of \( P \) in \( W \) is \( \Theta \left(\frac{(d+1)}{\epsilon} \cdot \delta^{-\frac{d+1}{2\epsilon}} r^{d+1}\right) = \Theta \left(\eta \delta^{-\frac{d+1}{2\epsilon}} r^{2d}\right) \), since we want this quantity to be \( k \log n \) we set \( r = \Theta \left(\delta^{\frac{d+1}{2\epsilon}} n^{-\frac{1}{2d}}\right) \). The condition \( r \geq \epsilon \) yields the lower bound \( \epsilon^2 \) on \( \delta \), the upper bound \( \frac{1}{2} \) on \( \delta \) 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 \( \delta \) 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 below (see [2] for details).

<table>
<thead>
<tr>
<th>Dimension</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation</td>
<td>$3\delta^{-2}n^{1/2}$</td>
<td>$4.2\delta^{-2}n^{1/2}$</td>
</tr>
<tr>
<td>Range of $\delta$</td>
<td>$[10/n^2, 0.1]$</td>
<td>$[100/n, 0.1]$</td>
</tr>
</tbody>
</table>

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 $\tilde{\Theta}(\delta^{-2}n^{1/2})$ for the expected size of the convex hull. An experimental confirmation of this bound is depicted on the 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 the Figure 12-bottom-right. For large $\delta$, 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 \( \delta \), starting with all points on the unit-radius regular \( n \)-gon gives an expected complexity of \( \tilde{\Theta}(n^{\frac{4}{3}}) \) which is less than the expected complexity of \( \Theta(n^{\frac{1}{3}}) \) 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 \( \Pi \) containing \( x \) and the normal to \( \partial K \) in \( x \), the curvature of \( \Pi \cap \partial K \), seen as a plane curve, at \( x \) is at least \( \rho_{\min} \) and at most \( \rho_{\max} \). Some constants appearing in the analysis change, e.g. certain ratios of volumes becomes \( O\left(\left(\frac{\rho_{\min}}{\rho_{\max}}\right)^{d+1}\right) \) instead of \( 3^{d+1} \), 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 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 [2].

References


