INTERSECTIONS WITH RANDOM GEOMETRIC OBJECTS

Prosenjit Bose
School of Computer Science
Carleton University
Ottawa, Ontario, Canada, K1S 5B6

and

Luc Devroye
School of Computer Science
McGill University
Montreal, Québec, Canada H3A 2A7

ABSTRACT. We present a systematic study of the expected complexity of the intersection of geometric objects. We first study the expected size of the intersection between a random Voronoi diagram and a *generic* geometric object that consists of a finite collection of line segments in the plane. Using this result, we explore the intersection complexity of a random Voronoi diagram with the following target objects which may or may not be random: a line segment, a Voronoi diagram, a minimum spanning tree, a Gabriel graph, a relative neighborhood graph, a Hamiltonian circuit, a furthest point Voronoi diagram, a convex hull, a k-d tree, and a rectangular grid.

KEYWORDS AND PHRASES. Voronoi diagram, range searching, intersection complexity, probabilistic analysis, computational geometry.

CR CATEGORIES: 3.74, 5.25, 5.5.

The first author's work was supported by NSERC Grant OGP0183877 and FIR. The second author's work was supported by NSERC Grant A4456 and by FCAR Grant 90-ER-0291.

1. Introduction.

The complexity of the intersection of geometric objects is a fundamental component in determining the efficiency of many geometric algorithms. Numerous different applications fall into this category. We mention a few of the basic ones.

In computer graphics, ray tracing is an important technique used in rendering a scene (see Foley, van Dam, Feiner and Hughes (1990)). Given a scene consisting of geometric objects and a ray (usually denoting a ray from the viewpoint through a pixel of the viewing plane), the ray tracing or ray shooting problem consists of determining which objects are intersected by the ray. Naturally, the intersection complexity of the ray with the scene is a dominating factor in all algorithms designed to perform this task.

Another structure gaining prominence in graphics is the binary space partition tree (BSP tree) (see Foley, van Dam, Feiner and Hughes (1990) or Samet (1990)). A BSP tree is a structure that allows one to efficiently render a scene when the viewpoint is moving. A BSP tree is built by recursively dividing the objects in a scene with a line in the 2-dimensional case and a plane in the 3-dimensional case. Since each division may split objects into two parts, the process can lead to the proliferation of objects. As such, the size of the BSP tree is directly related to the intersection complexity of the divider (line or plane) with the scene.

In computational metrology—a field that deals with the science of measuring manufactured parts—the *out-of-roundness* problem falls into this category (see Srinivasan or Yap (1994) for a brief overview of computational metrology). The out-of-roundness problem refers to the following: Given that an object is designed to be circular, how can one verify whether the manufactured part is indeed circular? Ebara, Fukuyama, Nakano and Nakanishi (1989) and Roy and Zhang (1991) propose a method for deciding this. Given the manufactured object M, use a coordinate measuring machine to compute a set P of n points on the planar cross-section of M. In order to determine how close to a circle M is, compute the *annular width* of the set, i.e., the thickness of the thinnest annulus that contains the points. Ebara, Fukuyama, Nakano and Nakanishi (1989) and Roy and Zhang (1991) show that this can be computed in $O(n \log n + k)$ time where k is the number of intersections between the closest and furthest point Voronoi diagram of the set P.

As noted above, the efficiency of many algorithms depends on the complexity of the intersection of two geometric objects. In many situations, the size of the intersection dominates the time complexity of proposed solutions. Often the worst-case complexity of the intersection is much larger than the complexity of the algorithm. For example, given a set of n red line segments R and n blue line segments B, computing the set of proper intersections between R and B takes $O(n \log n + k)$ time where k is the number of intersections (Chan, 1994). In the worst case, k can be $\Omega(n^2)$. For this reason, we study the expected complexity of the intersection of geometric objects when one of the objects is random. The formal definition of a random geometric object is given in the following section. Intuition might lead one to believe that the expected size of the intersection of random geometric objects might be smaller than worst case intersection complexity.

An encyclopedic treatment of the topic is impossible. Instead, we study the intersection complexity of a few fundamental geometric objects. We begin by studying the expected size of the intersection between a random Voronoi diagram and a *generic* geometric object, which is defined as a finite collection of line segments in the plane. Using this result, we explore the intersection complexity of a random Voronoi diagram with the following target objects which may or may not be random: a line segment, a Voronoi diagram, a β-skeleton, a minimum spanning tree, a Gabriel graph, a relative neighborhood graph, a Hamiltonian circuit, a furthest point Voronoi diagram, a convex hull, a k-d tree, and a rectangular grid. In all cases the intersection complexity is much smaller than the worst-case complexity. If the base Voronoi diagram has size n and each of the other objects has size m, the worst-case intersection complexity is usually $\Omega(mn)$. However, all of our results indicate that the expected complexity is much smaller. Having developed the tools when the base object is a Voronoi diagram, we study the intersection complexity of other base objects such as the minimum spanning tree with the same set of target objects.

2. Basic results for line segment intersections.

Voronoi diagrams are partitions of the plane into polygonal cells in which each cell is the collection of points that has a given data point as its nearest neighbor (Okabe, Boots and Sugihara, 1992). If two cells share a boundary segment, we say that they are neighboring cells. The graph obtained by connecting data points in neighboring cells is called the Delaunay graph, after the Russian physicist Delone. In this paper, we assume that n data points $X_1, \ldots, X_n$ are drawn independently from a density f on a compact convex set C of the plane that contains at least a circle of positive radius (to avoid trivialities). We are interested in the (random) number of intersections between a random Voronoi diagram and certain other objects, such as line segments, spheres, grids, minimal spanning trees, traveling salesman tours, k-d tree partitions, relative neighborhood graphs, Gabriel graphs, other Voronoi diagrams, and the farthest neighbor graph. We begin by studying the intersection complexity of a Voronoi diagram with a line segment.

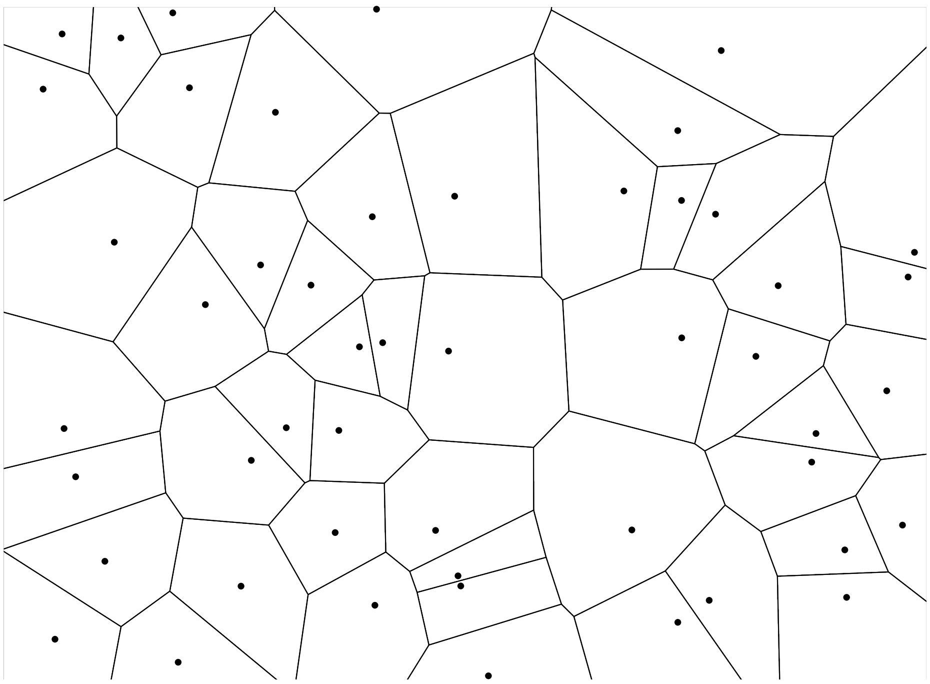

Figure 1. A random Voronoi diagram.

THEOREM 1. *Let $0 < \alpha \le f(x) \le \beta < \infty$ on C for fixed constants α, β. Let L be a fixed line segment contained in C, of length l. Then there exist constants d and n_0 depending upon α, β and C only such that*

$$\mathbf{E}N \le dl\sqrt{n}\ ,\ n \ge n_0\ ,$$

where N is the number of intersections between the random Voronoi diagram and L.

DISCUSSION. If L is of constant length, then $\mathbf{E}N = O(\sqrt{n})$, regardless of where L is located in C. However, when $l = O(1/\sqrt{n})$, we obtain $\mathbf{E}N = O(1)$. This, of course, is easily proved, as the expected area of a cell in a Voronoi diagram is about $\Theta(1/n)$, and its dimensions must be $\Theta(1/\sqrt{n})$. If L too is a random variable possibly dependent upon the Voronoi diagram, Theorem 1 does not apply. It does apply however if L is random and independent of the Voronoi diagram. $\square$

In what follows, $S_{x,r}$ denotes the closed circle of radius r centered at $x \in \mathbb{R}^2$.

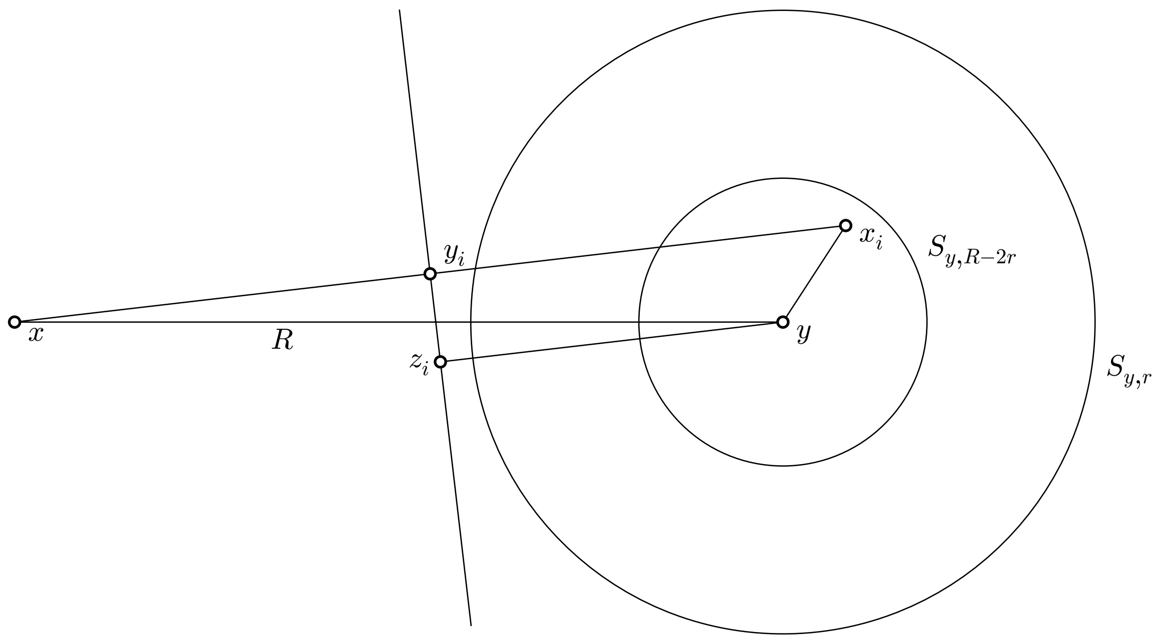

Figure 2. Illustration for Lemma 1.

LEMMA 1. *Let y, x be points of $\mathbb{R}^2$ at distance R from each other. Let V be a Voronoi diagram based upon points $x, x_1, \ldots, x_n$. For $r < R/2$, no boundary segment of the Voronoi cell of x intersects $S_{y,r}$ when one of the x_i's falls in $S_{y,R-2r}$.*

PROOF. Let x_i be a point in $S_{y,R-2r}$. Let $\delta = \|x_i - y\|$. Let $y_i = (x + x_i)/2$, and let z_i be the projection of y to the line through y_i which forms the bisector of point x and x_i. We need to show that $\|z_i - y\| \geq r$. Let $\Delta = \|z_i - y\|$. From the figure above, let γ be the angle $\angle x_i x y$ in triangle $\triangle(x_i x y)$, and θ be angle $\angle y x_i x$ in $\triangle(x_i x y)$. Assume that $0 \leq \theta \leq \pi/2$, and note the following two properties.

$$2\Delta + \delta \cos(\theta) = R\cos(\gamma)$$

$$\delta \sin(\theta) = R \sin(\gamma)$$

From this, we have the following after squaring the above and summing:

$$4\Delta^2 + 4\Delta\delta\cos(\theta) - R^2 + \delta^2 = 0$$

This expression is quadratic in Δ. At $\Delta = (R - \delta)/2$, we have the value:

$$\begin{aligned}(R-\delta)^2 + &2(R-\delta)\delta\cos(\theta) - R^2 + \delta^2 \\ &= \delta(2R - 2\delta)(\cos(\theta) - 1) \\ &\leq 0\end{aligned}$$

when $0 \leq \theta \leq \pi/2$. The polynomial is minimal when $\Delta = -\delta\cos(\theta)/2$, and takes the value $\delta^2 \sin^2(\theta) - R^2$ which is less than zero. Therefore, $\Delta \geq (R-\delta)/2 \geq (R - (R-2r))/2 = r$ when $0 \leq \theta \leq \pi/2$. Now consider the case when $\pi/2 < \theta \leq \pi$. Let c_i be the intersection of the line through x and x_i with $S_{y,R-2r}$. This intersection must be a chord of $S_{y,R-2r}$ since $\theta > \pi/2$. Let x_j be an arbitrary point on c_i with $\angle y x_j x \leq \pi/2$. Such a point must exist since c_i is a chord. But $\|z_i - y\| \geq \|z_j - y\| \geq r$, as required. □

LEMMA 2. *Let $x \in C$, and let X_1 be a random variable with density f as in Theorem 1. Then there exist constants $r_0 > 0$ and $\gamma > 0$ such that*

$$\inf_{x \in C} \mathbf{P}\{\|X_1 - x\| \leq r\} \geq \gamma r^2 \ , 0 < r \leq r_0 \ .$$

PROOF. Let $S_{z,b}$ be the nonempty circle contained in C, and let θ be the viewing angle at which $S_{z,b}$ can be seen from the point in C that is furthest from z. Note that for any

$x \in C$, $S_{x,r} \cap C$ must contain a slice of $S_{x,r}$ of angle at least $\theta > 0$ (uniformly over $x \in C$). Thus, if $r \leq \|x - z\|$ or if $r \leq b$,

$$\mathbf{P}\{\|X_1 - x\| \leq r\} \geq \int_{S_{x,r} \cap C} f(y)\, dy \geq \alpha \int_{S_{x,r} \cap C} dy \geq \frac{\alpha\theta}{2\pi} \int_{S_{x,r}} dy = \frac{\alpha\theta r^2}{2} \ .$$

This concludes the proof of Lemma 2. □

Lemma 3. *Assume that the conditions of Theorem 1 hold. Let N^* denote the number of X_i's with the property that one of its cell boundaries intersects $S_{x,r}$ where $x \in C$ and $r > 0$. Then there are positive constants c_i only depending upon C and f such that for $n \geq c_0$,*

$$\mathbf{E}N^* \leq c_1 n r^2 + c_2 \sqrt{n r^2} \ .$$

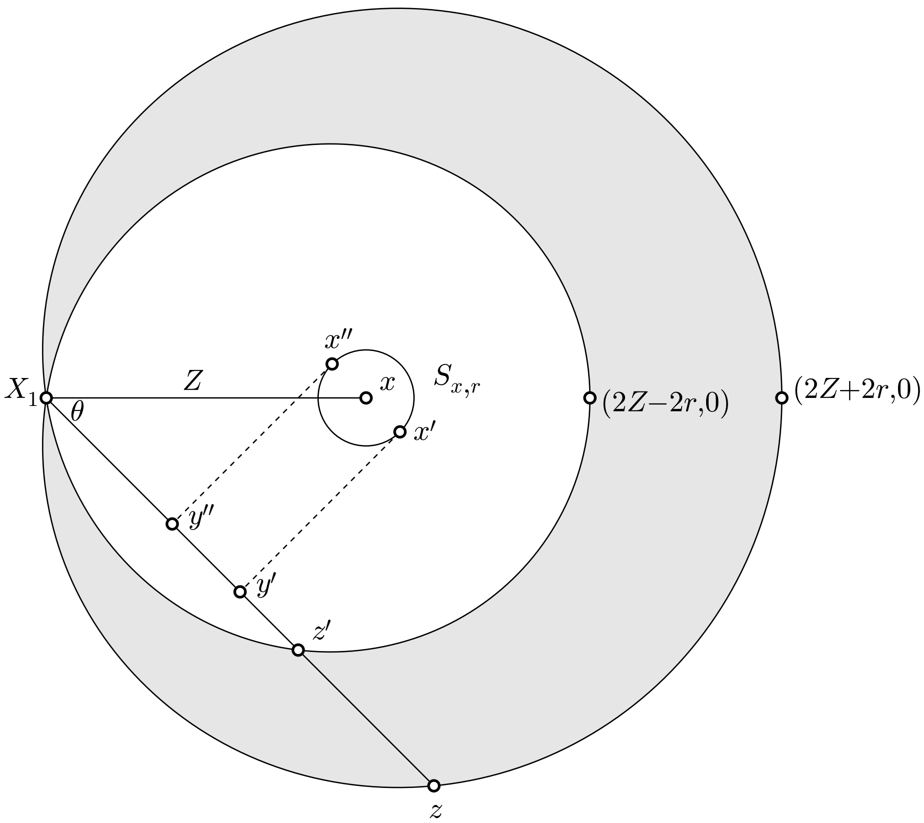

Figure 3. Illustration for part of the proof of Lemma 3.

PROOF. Let A denote the event that X_1 has a cell boundary that intersects $S_{x,r}$. Define $Z = \|X_1 - x\|$. We refer to figure 3 above, and observe that if $Z > r$, then

$$A \subseteq \cup_{i=2}^{n} [X_i \in V]$$

where V is the shaded region of figure 3 and the boundary of the gray set V has a parametric desciption given by the vector

$$(2(r + Z\cos\theta)\cos\theta, -2(r + Z\cos\theta)\sin\theta)\ , 0 \le \theta < 2\pi.$$

To see this, consider a line of angle θ through X_1. We claim that on this line, there is only one segment $[z', z]$ with the property that if $X_2 \in [z', z]$, then the perpendicular bisector of $[X_1, X_2]$ hits $S_{x,r}$. In the figure, y' and y'' denote the points midway between z and X_1 and z' and X_1 respectively. In this representation, $z'(\theta) = z(\theta + \pi)$. Thus, the distance from z to z' is at most $4r$. Denoting the dependence of z and z' on θ in an explicit manner, we see that the area of V is easily computed as

$$\begin{aligned}
\int_0^\pi \frac{\|z(\theta)\|^2 - \|z'(\theta)\|^2}{2}\, d\theta &\le 4r \int_0^\pi \frac{\|z(\theta)\| + \|z'(\theta)\|}{2}\, d\theta \\
&\le 4r \int_0^\pi 2Z d\theta \\
&= 8\pi r Z.
\end{aligned}$$

Thus, area $(V) \le 8\pi r\Delta$, if $X_1, x \in C$, where Δ is the diameter of C.

Let $W = S_{x,Z-2r}$. Let γ and r_0 be as in Lemma 2. Clearly $\mathbf{E}N^* = np$, where $p = \mathbf{P}\{A\}$. By Lemma 1,

$$\begin{aligned}
\mathbf{P}\{A|X_1\} &= \mathbf{P}\{A, \cap_{i=2}^n [X_i \notin W]|X_1\} + \mathbf{P}\{A, \cup_{i=2}^n [X_i \in W]|X_1\} \\
&= \mathbf{P}\{A, \cap_{i=2}^n [X_i \notin W]|X_1\} \\
&\le I_{Z>3r}\mathbf{P}\{A, \cap_{i=2}^n [X_i \notin W]|X_1\} + I_{Z\le 3r}\ .
\end{aligned}$$

Let c_1, c_2, c_3 denote positive constants. If X_1 is such that $Z > 3r$, then

$$\begin{aligned}
&\mathbf{P}\{A, \cap_{i=2}^n [X_i \notin W]|X_1\} \\
&\quad \le \mathbf{P}\{\cup_{j=2}^n [X_j \in V], \cap_{i=2}^n [X_i \notin W]|X_1\} \\
&\quad \le \sum_{j=2}^n \mathbf{P}\{[X_j \in V], \cap_{i=2, i\ne j}^n [X_i \notin W]|X_1\} \\
&\quad = \sum_{j=2}^n \mathbf{P}\{X_j \in V|X_1\}\mathbf{P}\{\cap_{i=2,i\ne j}^n [X_i \notin W]|X_1\} \\
&\quad = (n-1)\mathbf{P}\{X_2 \in V|X_1\}\mathbf{P}^{n-2}\{X_2 \notin W|X_1\} \\
&\quad \le 8\beta\pi Zrn\mathbf{P}^{n-2}\{X_2 \notin W|X_1\} \\
&\quad \le 8\beta\pi Zrn(1 - \gamma\text{min}^2(Z - 2r, r_0))^{n-2}
\end{aligned}$$

$$
\begin{aligned}
&\leq 8\beta\pi\Delta rn(1-\gamma r_0^2)^{n-2} + 8\beta\pi Zrn(1-\gamma(Z-2r)^2)^{n-2} I_{3r<Z\leq 3r+r_0} \\
&\leq 8\beta\pi\Delta rne^{-\gamma r_0^2(n-2)} + 16\beta\pi r^2 ne^{-\gamma(n-2)(Z-2r)^2} I_{3r<Z} + 8\beta\pi(Z-2r)rne^{-\gamma(n-2)(Z-2r)^2} I_{3r<Z} \\
&\stackrel{\text{def}}{=} I + II + III \ .
\end{aligned}
$$

The following fact will be used repeatedly. Let ϕ be a positive decreasing log-concave function. Then

$$
\int_r^\infty e^{\phi(t)} dt \leq \frac{e^{\phi(r)}}{-\phi'(r)} \ .
$$

We uncondition and note that

$$
\mathbf{P}\{A\} \leq \mathbf{P}\{Z \leq 3r\} + I + \mathbf{E}II + \mathbf{E}III \leq \beta\pi(3r)^2 + r\psi(n) + \mathbf{E}II + \mathbf{E}III \ ,
$$

where $\psi(n) = 8\beta\pi\Delta ne^{-\gamma r_0^2(n-2)} < 1/\sqrt{n}$ for all n large enough. Next, for $n \geq 4$, and by our assumption on f, and the inequality

$$
\int_r^\infty ue^{-cu^2} du = (1/2c)e^{-cr^2} \ ,
$$

$$
\begin{aligned}
\mathbf{E}II &= \mathbf{E}\left\{16\beta\pi r^2 ne^{-\gamma(n-2)(Z-2r)^2} I_{3r<Z}\right\} \\
&\leq \int_{3r}^\infty 16\beta\pi r^2 n 2\pi\beta z e^{-\gamma(n-2)(z-2r)^2} \, dz \\
&\leq \int_0^\infty 16\beta\pi r^2 n 2\pi\beta 3(r+u) e^{-\gamma(n-2)(r+u)^2} \, du \\
&\leq \int_r^\infty 96\beta^2\pi^2 r^2 nue^{-\gamma(n-2)u^2} \, du \\
&= \frac{96\beta^2\pi^2 r^2 ne^{-\gamma(n-2)r^2}}{2\gamma(n-2)} \\
&\leq \frac{96\beta^2\pi^2 r^2}{\gamma} \ .
\end{aligned}
$$

Finally, also for $n \geq 4$ and for $\gamma(n-2)2r > 4/r$,

$$
\begin{aligned}
\mathbf{E}III &= \mathbf{E}\left\{8\beta\pi(Z-2r)rne^{-\gamma(n-2)(Z-2r)^2} I_{3r<Z}\right\} \\
&\leq \int_0^\infty 8\beta\pi 2\pi\beta(r+u)^2 rne^{-\gamma(n-2)(r+u)^2} \, du \\
&\leq 16\beta^2\pi^2 rn \int_r^\infty u^2 e^{-\gamma(n-2)u^2} \, du \\
&\leq 16\beta^2\pi^2 rn \frac{r^2 e^{-\gamma(n-2)r^2}}{\gamma(n-2)2r - 2/r} \\
&\leq 32\beta^2\pi^2 r^2 e^{-\gamma(n-2)r^2}/\gamma \\
&\leq 32\beta^2\pi^2 r^2/\gamma \ .
\end{aligned}
$$

On the other hand, for $\gamma(n-2)2r \leq 4/r$,

$$\begin{aligned}
\mathbf{E}III &\leq 16\beta^2\pi^2 rn \int_r^\infty u^2 e^{-\gamma(n-2)u^2}\, du \\
&\leq 16\beta^2\pi^2 rn(\gamma(n-2))^{-3/2} \int_0^\infty \gamma(n-2)u^2 e^{-\gamma(n-2)u^2}\, d(\sqrt{\gamma(n-2)}u) \\
&\leq 4\beta^2\pi^2 r(\gamma(n-2))^{-3/2}\sqrt{2\pi} \\
&\leq 8\beta^2\pi^2 r\gamma^{-3/2}\sqrt{2\pi/(n-2)} \\
&\leq 16\beta^2\pi^{5/2} r\gamma^{-3/2}/\sqrt{n}\ .
\end{aligned}$$

In conclusion, there are positive constants c_i such that for $n \geq c_0$,

$$np \leq c_1 nr^2 + c_2\sqrt{nr^2}\ .$$

This completes the proof. □

LEMMA 4. *Under the assumptions of Lemma 3, for* $n \geq c_0$,

$$\mathbf{E}N' \leq 3c_1 nr^2 + 3c_2\sqrt{nr^2}\ ,$$

where N' *is the number of Voronoi cell boundaries that intersect* $S_{x,r}$, $x \in C$.

PROOF. All the cell boundaries counted in N' must be spawned by X_i's that are counted in N^* in Lemma 3. Call the latter collection of X_i's $\mathcal{X}$ (so that $|\mathcal{X}| = N^*$). The Voronoi diagram formed by $\mathcal{X}$ alone has at most $3|\mathcal{X}|-6$ boundary segments. Hence, $N' \leq 3N^*-6$. □

COROLLARY 1. If $r \leq \xi/\sqrt{n}$ for a constant ξ, then

$$\sup_{x\in C} \mathbf{E}N' \leq 3c_1\xi^2 + 3c_2\xi\ .$$

The uniform nature of the last inequality will be very useful in the sequel. □

PROOF OF THEOREM 1. If $l\sqrt{n} \geq 1$, L may be covered by $1 + \lceil l\sqrt{n}\rceil \leq 3l\sqrt{n}$ circles of radius $1/(2\sqrt{n})$ each. N is bounded by the sum of the number of intersections with these circles. By Corollary 1, the expected number of intersections with each circle is uniformly bounded by a constant c'. Hence, $\mathbf{E}N \le 3c' l\sqrt{n}$. If however $l\sqrt{n} < 1$, cover L by one circle of radius $l/2$. By Lemma 4,

$$\mathbf{E}N \le 3c_1 n l^2/4 + 3c_2\sqrt{n}l/2 \le (3c_1/4 + 3c_2/2)l\sqrt{n}\ . \ \square$$

3. Intersections with geometric objects.

A geometric object in this section is a finite collection of line segments in the plane. The number of line segments is denoted by M and the sum of the lengths of the segments is L. Possible geometric objects include a grid of lines, a convex hull, a polygon, or a minimal spanning tree.

THEOREM 2. *Let V be a random Voronoi diagram, and let G be a geometric object that is independent of V (note: it may be deterministic). The (random) length of G, after clipping to C, is L. Under the conditions of Theorem 1, there exist constants d and n_0 depending upon α, β and C only, such that*

$$\mathbf{E}N \le d\mathbf{E}L\sqrt{n}\ ,\ \ n \ge n_0\ ,$$

where N is the number of intersections between V and G.

PROOF. Immediate from Theorem 1 and the linearity of expectation. $\square$

Theorem 2 tells us that we need only find $\mathbf{E}L$ to obtain a bound on $\mathbf{E}N$. While we won't go into it here, the bounds of Theorems 1 and 2 cannot be improved by more than a constant multiplicative factor. The following eight sections describe various applications. The only new technical difficulties that may arise are related to the possible dependence between V and G. These are dealt with as we go along.

4. G is another Voronoi diagram.

Let the geometric object be another Voronoi diagram with n centers distributed as, but independent of, V. As $\mathbf{E}L = O(\sqrt{n})$ under the condition of Theorem 1 (the straightforward proof is omitted), we have $\mathbf{E}N = O(n)$: the expected number of intersections between two independent identically distributed Voronoi diagrams is linear in n. Strictly speaking, this result is only for intersections that occur within the convex hull of V (or within C) by the clipping present in Theorem 1. With some work, this restriction can be removed.

5. G is the Gabriel graph.

Given is a collection of n' points denoted by $Y_1, Y_2, \ldots, Y_{n'}$. The Gabriel graph is obtained by joining all pairs of points Y_i, Y_j for which the circle with Y_i and Y_j at opposite poles contains no Y_k, $1 \leq k \leq n'$ (Gabriel and Sokal, 1969). If the Y_i's are independent of the X_i's that define V, and if they have common density g as in Theorem 1, then $\mathbf{E}L = O(\sqrt{n'})$, which is easy to verify by elementary computations. Thus, $\mathbf{E}N = O\left(\sqrt{nn'}\right)$.

In particular, if $n' \in O(n)$ then $\mathbf{E}N = O(n)$. The relative neighborhood graph (Toussaint, 1980), the minimal spanning tree, and all β-skeletons with $\beta \geq 1$ (Kirkpatrick and Radke, 1985) are subgraphs of the Gabriel graph. Therefore, the result holds for these graphs as well.

6. G is a traveling salesman path.

The traveling salesman path (TSP) through n' points of $[0,1]^2$ has length not exceeding $\sqrt{8n'}$ as its length is at most twice that of the minimal spanning tree (see Moran (1988) for sharper bounds). For n' points in a convex compact set C, we have $L \leq \sqrt{8n'}\Delta$, where Δ is the side of the smallest square that covers C. As the TSP is entirely contained in C, we conclude that

$$\mathbf{E}N = O\left(\sqrt{nn'}\right) .$$

7. G is a random 2-d tree.

k-d Trees were invented by Bentley (1975) for use as dictionaries for multidimensional data. A 2-d tree in the plane partitions it by alternating vertical and horizontal cuts through data points. The total length L of the line segments determined by the 2-d tree has no nontrivial upper bound. However, for n' points uniformly distributed on the unit square, Chanzy and Devroye (1994) showed that $\mathbf{E}L = \Theta(n'^{(\sqrt{17}-3)/2})$ (this result is related to range searching in 2-d trees; see, e.g., Flajolet and Puech, 1986). We conclude that if the 2-d tree is independent of V, then

$$\mathbf{E}N = O\left(\sqrt{n}n'^{(\sqrt{17}-3)/2}\right)$$

in general and

$$\mathbf{E}N = O\left(n^{(\sqrt{17}-2)/2}\right) = O\left(n^{1.06155\ldots}\right)$$

in particular, when $n' = n$. We recall that this has only been rigorously proved for the uniform distribution on the unit square, but it should certainly also be true under the conditions of Theorem 1. It is interesting to note that the number of intersections between two similar-sized Voronoi diagrams is less than between a Voronoi diagram and a 2-d tree with an equal number of nodes. The reason is that the cells in the 2-d trees are in fact elongated rectangles.

8. G is a random convex hull or a convex polygon.

Consider a random convex polygon G that is independent of V and completely contained in C. As $L = O(1)$, it is clear that $\mathbf{E}N = O(\sqrt{n})$.

9. G is a rectilinear grid.

Let G be a rectilinear grid drawn through n' points in the plane that are picked independently from V. All lines are clipped to C as shown in the figure. As $L = O(n')$, we conclude that

$$\mathbf{E}N = O(n'\sqrt{n}) \ .$$

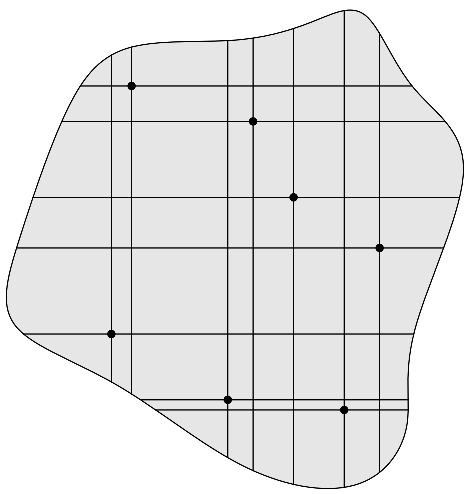

Figure 4. A rectilinear grid clipped to C.

10. G is the furthest-point Voronoi diagram.

The partition of the plane obtained by assigning locations to the furthest member of the point set is called the furthest-point Voronoi diagram. It remains unchanged if we remove all points that are not on the convex hull. Consider n points that gives rise to both the Voronoi diagram V and the furthest-point Voronoi diagram G. Under the assumptions of theorem 1, we would like to obtain a good bound for $\mathbf{E}N$ but can't use Theorem 1 directly because G and V are dependent.

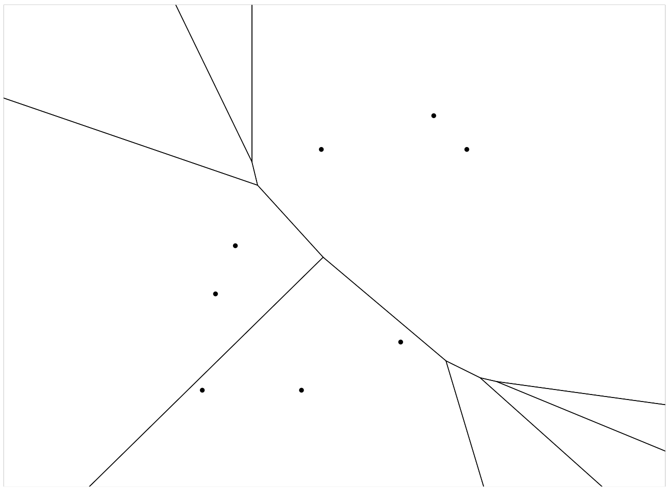

Figure 5. The furthest-point Voronoi diagram for a set of points in convex position. Adding non-convex hull points does not alter the furthest-point Voronoi diagram.

We will only focus on N', the number of intersections that fall entirely within the convex hull (CH) of the n points. The non-convex hull points are collected in a set $\mathcal{X}$. Given CH, the points in $\mathcal{X}$ are independent and have common density

$$\frac{fI_A}{\int_A f} \ ,$$

where A is the interior of the convex hull. The geometric object G, clipped to A has length not exceeding $(3N''-6)\Delta$, where N'' is the (random) cardinality of CH and Δ is the diameter of CH. Conditioning on CH, we have

$$\mathbf{E}\{N'|\mathrm{CH}\} \leq \frac{1}{\int_A f} \times \left(c_1 N'' + c_2(3N''-6)\Delta\sqrt{n}\right) \ ,$$

where N' is the number of intersections between the clipped G and the Voronoi diagram for $\mathcal{X}$, also clipped to A. Here we used the fact that the proportionality constant in Theorem

2 is proportional to β. Since Δ does not exceed a finite constant and $\mathbf{E}N'' = O(n^{1/3})$ under our assumption on f (Lemma A2), and because

$$\mathbf{P}\left\{\int_A f < \frac{\alpha\lambda(C)}{4\pi}\right\} \leq 8e^{-n\alpha\lambda(C)/64}$$

where $\lambda(C)$ is the area of C (Lemma A1, Appendix A), we see that

$$\mathbf{E}N' = O(n^{5/6}) \ .$$

The number N we want is slightly different, as we want to deal with the Voronoi diagram for all n points, not just $\mathcal{X}$. As always, intersections that fall outside the support set C are not counted. Observe that

$$N \leq N' + 3N''S \ ,$$

where N'' is the number of convex hull points (and thus $3N''$ is a bound on the number of edges of the furthest neighbor Voronoi diagram, by planarity), and S is the sum of the degrees in the Delaunay triangulation of these convex hull points. Call a point an inner point if it is not a convex hull point but if it has a Delaunay edge with a convex hull point. Let N''' be the total number of inner points. By planarity of the Delaunay triangulation, $S \leq 6(N'' + N''')$. Thus, the number of added intersections is not more than $18N''(N'' + N''')$. To conclude the claim made above about the total number of intersections being $O(n^{5/6})$ on average, it suffices to show that $\mathbf{E}N''N''' = O(n^{5/6})$, and that $\mathbf{E}N''^2 = O(n^{5/6})$. The latter result follows from an inequality of Devroye (1983) which states that $\mathbf{E}N''^2 \leq c(\mathbf{E}N'')^2$ for some universal constant c, and the fact that under the assumption of Theorem 1, $\mathbf{E}N'' = O(n^{1/3})$ (Lemma A2). By the Cauchy-Schwarz inequality,

$$\mathbf{E}\{N''N'''\} \leq \sqrt{\mathbf{E}N''^2\,\mathbf{E}N'''^2} = \sqrt{O(n^{2/3})\,O(n)} = O(n^{5/6}) \ ,$$

by the estimate above for $\mathbf{E}N''^2$, and Lemma A3 of Appendix B.

11. Minimum spanning tree as base.

In this section, we change our focus. Our base object is no longer the Voronoi diagram but the minimum spanning tree. Given a set X of n points $X_1, \ldots, X_n$ drawn independently from a density f on a compact convex set C of the plane that contains at least a circle of positive radius, let $\mathrm{MST}(X)$ be the minimum spanning tree of X.

Corollary 2. *Let $0 < \alpha \le f(x) \le \beta < \infty$ on C for fixed constants α, β. Let L be a fixed line segment contained in C, of length l. Then there exist constants d and n_0 depending upon α, β and C only such that*

$$\mathbf{E}N \le dl\sqrt{n}\ ,\ n \ge n_0\ ,$$

where N is the number of intersections between $\textsc{mst}(X)$ *and L.*

Proof. The maximum degree of a node in $\textsc{mst}(X)$ is 6. If L is contained in one cell of the Voronoi diagram of X, then $N \le 6$. If L intersects K cells of the Voronoi diagram of X then $N \le 6(K+1)$. Therefore, $\mathbf{E}N \le dl\sqrt{n}$ for some constant $d > 0$ and all $n \ge n_0$ by Theorem 1. □

All of the results obtained in the previous sections for the intersection of a Voronoi diagram with another geometric object hold for minimum spanning trees by Corollary 2.

12. Applications

Geometric range searching. Let P be a set of n points in the plane. Consider the problem of pre-processing P such that given a query half-plane, the set of points or the cardinality of the set of points contained in the query half-plane is quickly reported. A query time close to $\sqrt{n}$ was first achieved by Welzl (1988) using a *spanning tree with low crossing number*, where given a spanning tree T of P, the *crossing number of T with respect to a given half-plane H* is the number of edges of T crossed by H, and the *crossing number of T* is the maximum of crossing numbers of T with respect to all half-planes.

Let T be a spanning tree of P with crossing number K. For any given half-plane H, there are at most $K+1$ components of T once the crossing edges have been removed. Each such component is completely to one side of H. Once the at most $K+1$ components have been identified, the query can be easily answered. Note that the efficiency of this approach is based on the size of the crossing number. In the plane, there exist point sets such that the crossing number of every spanning tree on the set is $\Omega(\sqrt{n})$. However, given a point set P, several algorithms exist to construct a spanning tree with crossing number $O(\sqrt{n})$, which is referred to as a *spanning tree with low crossing number* (see Welzl (1992) for a survey, and Matoušek (1994) and Welzl (1992) for other applications). Moreover, all the known algorithms for computing spanning trees with low crossing number are more complicated than standard minimum spanning tree algorithms.

Corollary 3. *Under the conditions of Theorem 1, the minimum spanning tree of a point set is a spanning tree of low expected crossing number ($O(\sqrt{n})$).*

The out-of-roundness problem. Ebara, Fukuyama, Nakano and Nakanish (1989) and Roy and Zhang (1991) show that the out-of-roundness problem can be solved $O(n \log n + k)$ time where k is the number of intersections between the closest and furthest point Voronoi diagram of the point set P. In the worst case, k can be $\Omega(n^2)$. However, for random point sets satisfying the conditions of Theorem 1, $\mathbf{E}\{k\} = O(n^{5/6})$ so that the expected time for the out-of-roundness algorithm mentioned above is $O(n \log n)$.

13. References.

J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Communications of the ACM*, vol. 18, pp. 509–517, 1975.

T. Chan, "A simple trapezoid sweep algorithm for reporting red/blue segment intersections," in: *Proceedings of the 6th Canadian Conference on Computational Geometry*, pp. 263–268, 1994.

P. Chanzy and L. Devroye, "On range searching and nearest neighbor queries with 2-d trees," Manuscript, School of Computer science, McGill University, Montreal, 1994.

L. Devroye, "Moment inequalities for random variables in computational geometry," *Computing*, vol. 30, pp. 111–119, 1983.

L. Devroye, *Lecture Notes on Bucket Algorithms*, Birkhäuser Verlag, Boston, 1986.

L. Devroye, "The expected size of some graphs in computational geometry," *Computers and Mathematics with Applications*, vol. 15, pp. 53–64, 1988.

L. Devroye, "Coupled samples in simulation," *Operations Research*, vol. 38, pp. 115–126, 1990.

L. Devroye and B. Zhu, "Intersections of random line segments," *International Journal of Computational Geometry and Applications*, vol. 4(3), pp. 261–274, 1994.

H. Ebara, N. Fukuyama, H. Nakano and Y. Nakanishi, "Roundness algorithms using the Voronoi diagrams," in: *Proceedings of the 1st Canadian Conference on Comp. Geom.*, pp. 41, 1989.

L. Few, "The shortest path and the shortest road through n points in a region," *Mathematika*, vol. 2, pp. 141–144, 1955.

P. Flajolet and C. Puech, "Partial match retrieval of multidimensional data," *Journal of the ACM*, vol. 33, pp. 371–407, 1986.

J.D. Foley, A. van Dam, S.K. Feiner and J.F. Hughes, *Computer Graphics: Principles and Practice*, Addison-Wesley, Reading, MA, 1990.

K. R. Gabriel and R. R. Sokal, "A new statistical approach to geographic variation analysis," *Systematic Zoology*, vol. 18, pp. 259–278, 1969.

F. John, "Extremum problems with inequalities as subsidiary conditions," in: *Studies and Essays Presented to R.Courant*, pp. 187–204, Interscience, New York, 1948.

D. G. Kirkpatrick and J. D. Radke "A framework for computational morphology," *Computational Geometry, G. T. Toussaint ed.,* Elsevier, pp. 217–248, 1985.

D. W. Matula and R. R. Sokal, "Properties of Gabriel graphs relevant to geographic variation research and the clustering of points in the plane," *Geographical Analysis*, vol. 12, pp. 205–222, 1980.

J. Matoušek, "Geometric range searching," *ACM Computing Surveys*, vol. 26(4), pp. 421–461, 1994.

J. Møller, "Random tessellations in $\mathbb{R}^d$," *Advances in Applied Probability*, vol. 21, pp. 37–73, 1989.

S. Moran, "On the length of the optimal TSP circuits in sets of bounded diameter," *Journal of Combinatirial Theory*, vol. B37, pp. 113–141, 1984.

A. Okabe, B. Boots and K. Sugihara, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, John Wiley, Chichester, England, 1992.

A. Rényi and R. Sulanke, "Uber die konvexe Hülle von n zufällig gewählten Punkten," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 2, pp. 75–84, 1963.

A. Rényi and R. Sulanke, "Uber die konvexe Hülle von n zufällig gewählten Punkten," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 3, pp. 138–1247, 1964.

U. Roy and X. Zhang "Establishment of a pair of concentric circles with the minium radial separation for assessing roundness error," *Computer Aided Design*, vol. 24(3), pp. 161–168, 1992.

H. Samet, *Applications of Spatial Data Structures: Computer Graphics, Image Processing, and GIS*, Addison-Wesley, 1990.

V. Srinivasan, "How tall is the pyramid of Cheops? and other problems in computational metrology," *SIAM News*, to appear.

G. T. Toussaint, "Pattern recognition and geometrical complexity," in: *Fifth International Conference on Pattern Recognition*, pp. 1324–1347, 1980.

G. T. Toussaint, "The relative neighborhood graph of a finite planar set," *Pattern Recognition*, vol. 12, pp. 261–268, 1980.

E. Welzl, "Partition trees for triangle counting and other range searching problems," in: *Proceedings of the 4th ACM Symposium on Computational Geometry*, pp. 23–33, 1988.

E. Welzl, "On spanning-trees with low crossing number," in: *Lecture Notes in Computer Science*, vol. 594, Springer-Verlag, Berlin, pp. 233–249, 1992.

C. Yap, "Exact computational geometry and tolerancing metrology," *Snapshots of Discrete and Computational Geometry, Vol. 3*, eds. D. Avis and P. Bose, Technical Report #94.50, School of Computer Science, McGill University, pp. 34–48, 1994.

14. Appendix A: Some properties of random convex hulls.

LEMMA A1. *If A is the convex hull for a random cloud of n points drawn from a density f on a compact set $C \subseteq \mathbb{R}^2$ with $f \geq \alpha > 0$ on C, then*

$$\mathbf{P}\left\{\int_A f < \frac{\alpha\lambda(C)}{4\pi}\right\} \leq 8e^{-n\alpha\lambda(C)/64}\ .$$

Proof. By John's theorem (1948), there exists an ellipse $C'' \subseteq C$ such that $\lambda(C') \geq \lambda(C)/4$, where $\lambda(.)$ denotes Lebesgue measure or area. By a linear transformation, we may assume without loss of generality that C' is a circle. Let C'' be the circle concentric with C' of radius $\sqrt{2}$ times smaller than that of C'. Define a square T that is inscribed in C' but that contains C'' (see figure). There are four caps defined by $C' - T$, called A_1, A_2, A_3, A_4. Split each cap in half by a line that extends through the center of C'.

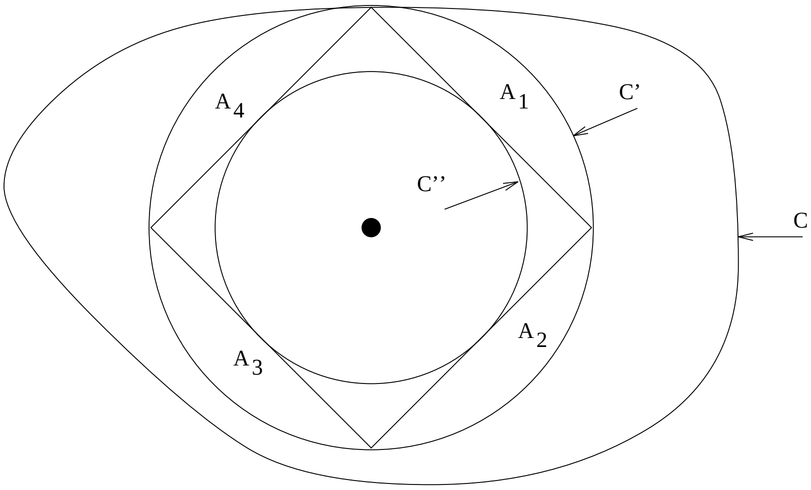

Figure 6. A convex region C is shown together with a large ellipse C' of area at least one fourth.

Let E be the event that one of the eight halves of the A_i's does not receive a data point. Clearly, by the assumption on f,

$$\mathbf{P}\{E\} \leq 8\,(1 - \alpha\lambda(C')/16)^n \leq 8e^{-n\alpha\lambda(C')/16} \leq 8e^{-n\alpha\lambda(C)/64}$$

where we used the fact that $4\lambda(A_1) = \lambda(C') - \lambda(C'') = \lambda(C')/2$. If E does not happen and all eight cap regions are occupied, then the convex hull must contain a small square Q with side equal to the radius of C', and thus,

$$\int_A f \geq \alpha\lambda(Q) I_{E^c} \geq \frac{\alpha\lambda(C')}{\pi} I_{E^c} \geq \frac{\alpha\lambda(C)}{4\pi} I_{E^c} \; .$$

Thus, $\int_A f < \frac{\alpha\lambda(C)}{4\pi}$ implies E, and the Lemma is proved. $\square$

Lemma A2. *Let N be the number of points on the convex hull of a cloud of n random points in the plane drawn independently from a density f on a compact convex set C, where $\inf_{x \in C} f(x) > 0$ and $\sup_{x \in C} f(x) < \infty$. Then $\mathbf{E}N = O(n^{1/3})$.*

PROOF. For the uniform density on a convex set with a smooth perimeter, this result is due to Rényi and Sulanke (1963, 1964). The proof for the slightly more general case is omitted. □

15. Appendix B: Inner points.

LEMMA A3. *Let C be a compact convex set of $\mathbb{R}^2$ and let f be a density supported on C such that (as in Theorem 1) $0 < \alpha \le f(x) \le \beta < \infty$ for $x \in C$. Draw n i.i.d. points from f and let N''' be the number of inner points. Then:*

(A). $\mathbf{E}N''' = O(\sqrt{n})$.

(B). $\mathbf{E}\{N'''^2\} = O(n)$.

PROOF. We prove Lemma A3 for the unit circle C, leaving the more general case as an exercise.

PROOF OF PART (A). Let $r > 0$ be a constant depending upon n. Define an outer doughnut O as the difference of two circles centered at the origin of radii 1 and $1-r$ respectively. Let A be the event that at least one convex hull point does not belong to the outer doughnut. Take any point u outside the outer doughnut. Draw the diagonal through u and the perpendicular to that diagonal at u. Both lines together define four slices of the circle, two congruent smaller pieces, and two congruent larger pieces. If u is on the convex hull, one of these four pieces must be empty. The area of the smallest small piece must be at least $(r/2) \times \sqrt{1-(1-r)^2} \ge r^{3/2}/2$. The probability content of this piece is at least $\alpha r^{3/2}/2$. Thus,

$$\mathbf{P}\{A\} \le 4n \left(1 - \alpha r^{3/2}/2\right)^{n-1} \le 4n e^{-(n-1)\alpha r^{3/2}/2} .$$

Clearly,

$$\mathbf{E}N''' \le n\mathbf{P}\{X_1 \in C - O; X_1 \text{ has a Delaunay edge with a point in O}; A^c\} + n\mathbf{P}\{A\} .$$

The last term is $o(1)$ if we choose $r = c(\log n/n)^{2/3}$ for a suitably chosen constant c. Denote the distance from X_1 to the outer doughnut by R. Clearly, R has a triangular density on $[0, 1-r]$, decreasing monotonically from $2/(1-r)$ at zero to 0. Consider the circle B of radius $R/2$ centered at X_1. Partition this circle into 16 equal sectors of angle $\pi/8$ each. If X_1 has a Delaunay edge with a member of the outer doughnut, then one of these 16 sectors must be empty (contain no data point). Thus, the probability that X_1 has such a Delaunay edge, given R, is not more than

$$16(1-p)^{n-1} \le 16 e^{-(n-1)p} ,$$

where $p = \alpha\pi R^2/64$ is a lower bound for the probability of any such sector. Therefore,

$$
\begin{aligned}
& n\mathbf{P}\{X_1 \in C - O; X_1 \text{ has a Delaunay edge with a point in O}\} \\
& \quad \le n\mathbf{E}\left\{16e^{-(n-1)\alpha\pi R^2/64} I_{X_1 \in C-O}\right\} \\
& \quad \le 32n \int_0^{1-r} e^{-(n-1)\alpha\pi t^2/64}\, dt \\
& \quad \le 32n \int_0^{\infty} e^{-(n-1)\alpha\pi t^2/64}\, dt \\
& \quad = 16n \sqrt{\frac{64\pi}{(n-1)\alpha\pi}} \\
& \quad = O(\sqrt{n})\ .
\end{aligned}
$$

This concludes the proof of part (A).

PROOF OF PART (B). We use once again the moment inequality of Devroye (1983), now applied to the random variable N_1+N_2, where N_1 is the number of points in the outer doughnut defined in the proof of part (A) above, and N_2 is the number of points not in the outer doughnut but with a Delaunay edge connected to a point in the outer doughnut. If A fails (notation of part (A)), then N_1 is a bound for the number of convex hull points, and N_2 is a bound for the number of inner points. The random variable $N_1 + N_2$ is a sum of indicator functions satisfying the conditions of the inequality (permutation invariance for the data, and an inclusion inequality). Thus,

$$\mathbf{E}(N_1 + N_2)^2 \le c\mathbf{E}^2(N_1 + N_2)$$

for some universal constant c. But we know that $\mathbf{E}N_1 \le \beta 2\pi rn = O((\log n)^{2/3} n^{1/3})$ if r is picked as in part (A). Also, $\mathbf{E}N_2 = O(\sqrt{n})$ as that is what is essentially shown in part (A). Thus, if N'' denotes the number of convex hull points,

$$
\begin{aligned}
\mathbf{E}N'''^2 &\le \mathbf{E}(N'' + N''')^2 \\
&\le \mathbf{E}\left\{I_{A^c}(N_1 + N_2)^2\right\} + \mathbf{E}\left\{n^2 I_A\right\} \\
&\le \mathbf{E}\left\{(N_1 + N_2)^2\right\} + n^2\mathbf{P}\{A\} \\
&\le O(n) + 4n^3 e^{-(n-1)\alpha r^{3/2}/2} \\
&\qquad \text{(by an inequality from part (A))} \\
&= O(n) \\
&\qquad \text{(if } r = (c\log n/n)^{2/3} \text{ and } c = 4/\alpha)\ .
\end{aligned}
$$

This is what we set out to show. □