On the Spanning Ratio of Gabriel Graphs and β-Skeletons

Prosenjit Bose[1], Luc Devroye[2], William Evans[3], and David Kirkpatrick[3]

[1] School of Computer Science, Carleton University, Ottawa, Ontario, Canada. E-mail: jit@cs.carleton.ca. Research supported by NSERC.

[2] School of Computer Science, McGill University, Montreal, Canada. E-mail: luc@cs.mcgill.ca. Research supported by NSERC and by FCAR.

[3] Department of Computer Science, University of British Columbia, Vancouver, Canada. E-mail: {will,kirk}@cs.ubc.ca. Research supported by NSERC.

Abstract. The spanning ratio of a graph defined on n points in the Euclidean plane is the maximal ratio over all pairs of data points (u, v), of the minimum graph distance between u and v, over the Euclidean distance between u and v. A connected graph is said to be a k-spanner if the spanning ratio does not exceed k. For example, for any k, there exists a point set whose minimum spanning tree is not a k-spanner. At the other end of the spectrum, a Delaunay triangulation is guaranteed to be a 2.42-spanner[11]. For proximity graphs *inbetween* these two extremes, such as Gabriel graphs[8], relative neighborhood graphs[16] and β-skeletons[12] with $\beta \in [0, 2]$ some interesting questions arise. We show that the spanning ratio for Gabriel graphs (which are β-skeletons with $\beta = 1$) is $\Theta(\sqrt{n})$ in the worst case. For all β-skeletons with $\beta \in [0, 1]$, we prove that the spanning ratio is at most $O(n^\gamma)$ where $\gamma = (1 - \log_2(1 + \sqrt{1 - \beta^2}))/2$. For all β-skeletons with $\beta \in [1, 2)$, we prove that there exist point sets whose spanning ratio is at least $\left(\frac{1}{2} - o(1)\right) \sqrt{n}$. For relative neighborhood graphs[16] (skeletons with $\beta = 2$), we show that there exist point sets where the spanning ratio is $\Omega(n)$. For points drawn independently from the uniform distribution on the unit square, we show that the spanning ratio of the (random) Gabriel graph and all β-skeletons with $\beta \in [1, 2]$ tends to ∞ in probability as $\sqrt{\log n / \log \log n}$.

1 Introduction.

Many problems in geometric network design, pattern recognition and classification, geographic variation analysis, geographic information systems, computational geometry, computational morphology, and computer vision use the underlying *structure* (also referred to as the *skeleton* or *internal shape*) of a set of data points revealed by means of a *proximity graph* (see for example [16], [13], [7], [9]). A proximity graph attempts to exhibit the relation between points in a point set. Two points are joined by an edge if they are deemed *close* by some proximity measure. It is the measure that determines the type of graph that results. Many different measures of proximity have been defined, giving rise to many different types of proximity graphs. An extensive survey on the current research in proximity graphs can be found in Jaromczyk and Toussaint [9].

We are concerned with the spanning ratio of proximity graphs. Consider n points in $\mathbb{R}^2$, and define a graph on these points, such as the Gabriel graph [8], or the relative neighborhood graph [16]. For a pair of data points (u, v), the length of the shortest path measured by Euclidean distance is denoted by $L(u, v)$, while the direct Euclidean distance is $D(u, v)$. The *spanning ratio* of the graph is defined by

$$S \stackrel{\text{def}}{=} \max_{(u,v)} \frac{L(u,v)}{D(u,v)} ,$$

where the maximum is over all $\binom{n}{2}$ pairs of data points. Note that if the graph is not connected, the spanning ratio is infinite. In this paper, we will concentrate on connected graphs.

Graphs with small spanning ratios are important in some applications (see [7] for a survey on spanners). The history for the Delaunay triangulation is interesting. First, Chew [2, 3] showed that in the worst case, $S \geq \pi/2$. Subsequently, Dobkin et al.[5] showed that the Delaunay triangulation was a $((1+\sqrt{5})/2)\pi \approx 5.08$ spanner. Finally, Keil and Gutwin [10, 11] improve this to $2\pi/(3\cos(\pi/6))$ which is about 2.42. It is conjectured that the spanning ratio of the Delaunay triangulation is $\pi/2$. The complete graph has $S = 1$, but is less interesting because the number of edges is not linear but quadratic in n. In this paper, we concentrate on the parameterized family of proximity graphs known as β-skeletons [12] with β in the interval $[0, 2]$. The family of β-graphs contains certain well-known proximity graphs such as the Gabriel graph [8] when $\beta = 1$ and the relative neighborhood graph [16] when $\beta = 2$. As graphs become sparser, their spanning ratios increase. For example, it is trivial to show that there are minimal spanning trees with n vertices for which $S \geq n - 1$, whereas the Delaunay triangulation has a constant spanning ratio.

In this note, we probe the expanse inbetween these two extremes. We show that for any n, in the plane, there exists a point set whose Gabriel graph satisfies $S \geq c\sqrt{n}$, where c is a universal constant. We also show that for any Gabriel graph in the plane, $S \leq c'\sqrt{n}$ for another constant c'. For all β-skeletons with $\beta \in [0, 1]$, we prove that the spanning ratio is at most $O(n^\gamma)$ where $\gamma = (1 - \log_2(1 + \sqrt{1 - \beta^2}))/2$. For all β-skeletons with $\beta \in [1, 2)$, we prove that there exist point sets whose spanning ratio is at least $\left(\frac{1}{2} - o(1)\right)\sqrt{n}$. For relative neighborhood graphs, we show that there exist point sets where the spanning ratio is $\Omega(n)$. The second part of the paper deals with point sets drawn independently from the uniform distribution on the unit square. We show that the spanning ratio of the (random) Gabriel graph and all β-skeletons with $\beta \in [1, 2]$ tends to ∞ in probability as $\sqrt{\log n / \log\log n}$.

2 Preliminaries

We begin by defining some of the graph theoretic and geometric terminology used in this paper. For more details see [1] and [15].

A *graph* $G = (V, E)$ consists of a finite non empty set $V(G)$ of *vertices*, and a set $E(G)$ of unordered pairs of vertices known as *edges*. An edge $e \in E(G)$ consisting of vertices u and v is denoted by $e = uv$; u and v are called the *endpoints* of e and are said to be *adjacent* vertices or *neighbors*. The *degree* of a vertex $v \in V(G)$, denoted by $deg_G(v)$ (or just $deg(v)$ when no confusion will result), is the number of edges of $E(G)$ which have v as an endpoint. A *path* in a graph G is a finite non-null sequence $P = v_1 v_2 \ldots v_k$ where the vertices $v_1, v_2 \ldots v_k$ are distinct and $v_i v_{i+1}$ is an edge for each $i = 1 \ldots k-1$. The vertices v_1 and v_k are known as the *endpoints* of the path. A *cycle* is a path whose endpoints are the same. A graph is *connected* if, for each pair of vertices $u, v \in V$, there is a path from u to v.

Intuitively speaking, a *proximity graph* on a finite set $P \subset \mathbb{R}^2$ is obtained by connecting pairs of points of P with line segments if the points are considered to be *close* in some sense. Different definitions of closeness give rise to different proximity graphs. One technique for defining a proximity graph on a set of points is to select a geometric region defined by two points of P—for example the smallest disk containing the two points—and then specifying that a segment is drawn between the two points if and only if this region contains no other points from P. Such a region will be referred to as a *region of influence* of the two points. Four such definitions follow.

Given a set P of points in $\mathbb{R}^2$, the *relative neighborhood graph of* P, denoted by $RNG(P)$, has a segment between points u and v in P if the intersection of the open disks of radius $D(u, v)$ centered at u and v is empty. This region of influence is referred to as the *lune* of u and v. Equivalently, $u, v \in S$ are adjacent if and only if

$$D(u, v) \leq \max[D(u, w), D(v, w)], \text{ for all } w \in S, w \neq u, v.$$

The *Gabriel graph of* P, denoted by $GG(P)$, has as its region of influence the closed disk having segment $\overline{uv}$ as diameter. That is, two vertices $u, v \in S$ are adjacent if and only if

$$D^2(u, v) < D^2(u, w) + D^2(v, w), \text{ for all } w \in S, w \neq u, v.$$

A *Delaunay triangulation* of a set P of points in the plane, denoted by $DT(P)$, is a triangulation of P such that for each interior face, the triangle which bounds that face has the property that the circle circumscribing the triangle contains no other points of the graph in its interior. A set P may admit more than one Delaunay triangulation, but only if P contains four or more co-circular points. A list of properties of the Delaunay triangulation can be found in [15].

We describe another graph, a *minimum spanning tree*, which is not defined in terms of a region of influence. Given a set P of points in the plane, consider a connected straight-line graph G on P, that is, a graph having as its edge set E a collection of line segments connecting pairs of vertices of P. Define the *weight* of G to be the sum of all of the edge lengths of G. Such a graph is called a *minimum spanning tree of* P, denoted by $MST(P)$, if its weight is no greater than the weight of any other connected straight-line graph on P. (It is easy to see that such a graph must be a tree.) In general, a set P may have many minimum spanning trees (for example, if P consists of the vertices of a regular polygon).

The following relationships among the different proximity graphs hold for any finite set P of points in the plane.

Lemma 1. *[15] $MST(P) \subseteq RNG(P) \subseteq GG(P) \subseteq DT(P)$*

Given a finite set P of distinct points in $\mathbb{R}^2$, we define the β-skeleton of P. β-skeletons are a family of graphs having vertex set P, parameterized by the value of β. For each pair x, y of points in P, we define the region of influence for a given value of β, and denote this region as $R(x, y, \beta)$.

1. *For $\beta = 0$, $R(x, y, \beta)$ is the line segment $\overline{xy}$.*
2. *For $0 < \beta < 1$, $R(x, y, \beta)$ is the intersection of the two disks of radius $D(x, y)/(2\beta)$ passing through both x and y.*
3. *For $1 \leq \beta < \infty$, $R(x, y, \beta)$ is the intersection of the two disks of radius $\beta D(x, y)/2$ and centered at the points $(1 - \beta/2)x + (\beta/2)y$ and $(\beta/2)x + (1 - \beta/2)y$, respectively.*
4. *For $\beta = \infty$, $R(x, y, \beta)$ is the infinite strip perpendicular to the line segment $\overline{xy}$*

Now consider the set of segments $\overline{xy}$ with $x, y \in P$ such that $R(x, y, \beta) \cap P \setminus \{x, y\} = \emptyset$ (i.e. the set of edges $\overline{xy}$ whose region of influence contains no points of $P \setminus \{x, y\}$). This set of distinct points and segments naturally defines a graph called the *β-skeleton* of P [12]. The β-skeleton for a fixed value of $\beta = k$ shall be referred to as the k-skeleton. Notice that different values of the parameter β give rise to different graphs. Note also that different graphs may result for the same value of β if the regions of influence are constructed with open rather than closed disks, however, these boundary effects do not alter our results. When necessary, we will explicitly state whether the region of influence is open or closed. These graphs will be referred to as open β-skeletons and closed β-skeletons, respectively. The closed 1-skeleton is the Gabriel graph and the open 2-skeleton is the relative neighborhood graph.

Also, as the value of β increases, the graphs become sparser since the region of influence increases in size. β-skeletons with $\beta \leq 2$ are connected. Therefore, we will concentrate on the interval $\beta \in [0, 2]$.

Observation 1 *If $k \leq k'$, then the k'-skeleton is a subset of the k-skeleton of a point set.*

3 A lower bound on the spanning ratio.

We begin with a deterministic lower bound on the spanning ratio of β-skeletons. The example developed in this section is essential for the understanding of the results on random Gabriel graphs.

Theorem 1. *For any $n \geq 2$, there exists a set of n points in the plane whose β-skeleton with $\beta \in [1,2]$ has spanning ratio*

$$S \geq \left(\frac{1}{2} - o(1)\right)\sqrt{n} \ .$$

Note: the closed 1-skeleton is the Gabriel graph and that all β-skeletons with $\beta > 1$ are subgraphs of the Gabriel graph. Therefore, it suffices to prove the theorem for the Gabriel graph. Also, the $1/2 - o(1)$ factor can be improved to $2/3$.

Proof. Let $m = \lfloor n/2 \rfloor$. Place points p_i and q_i at locations $(-r_i, y_i)$ and (r_i, y_i) respectively $(1 \leq i \leq m)$ where

$$r_i = 1 - (i-1)/n$$
$$y_i = (i-1)/\sqrt{n}$$

If n is odd place the remaining point at the same location as p_1.

We claim that for each pair p_i, q_i, the circle with diameter $p_i q_i$ contains the points p_{i+1} and q_{i+1} $(1 \leq i \leq m-1)$. Let d be the distance from the center of the circle with diameter $p_i q_i$ to the point p_{i+1}. For p_{i+1} to lie within this circle, d must be at most r_i. By construction,

$$d = \sqrt{(r_i - 1/n)^2 + 1/n} \ .$$

Thus we require $(r_i - 1/n)^2 + 1/n \leq r_i^2$ or, equivalently, $r_i \geq 1/2 + 1/(2n)$, which holds for $1 \leq i \leq m-1$.

It follows that when $i \leq j$, edge $p_i q_j$ does not belong to the Gabriel graph of these points (unless $i = j = m$), since p_{i+1} lies in or on the circle with diameter $p_i q_j$. Similarly, when $i > j$, edge $p_i q_j$ is precluded by point q_{j+1}.

The Euclidean distance between p_1 and q_1 is two. However, the shortest path from p_1 to q_1 using Gabriel graph edges is at least $2y_m$, which results in a spanning ratio of

$$S = y_m = (\lfloor n/2 \rfloor - 1)/\sqrt{n} = \left(\frac{1}{2} - o(1)\right)\sqrt{n} \ .$$

■

When β is in the interval $(0,1]$, Eppstein[6] presents an elegant fractal construction that provides a non-constant lower bound on the spanning ratio. His result is summarized in the following theorem.

Theorem 2. *For any $n = 5^k + 1$, there exists a set of n points in the plane whose β-skeleton with $\beta \in (0,1]$ has spanning ratio $\Omega(n^c)$, where $c = \log_5(5/(3 + 2sin(\theta)))$ and $\theta < (\pi - sin^{-1}(\beta))/2$.*

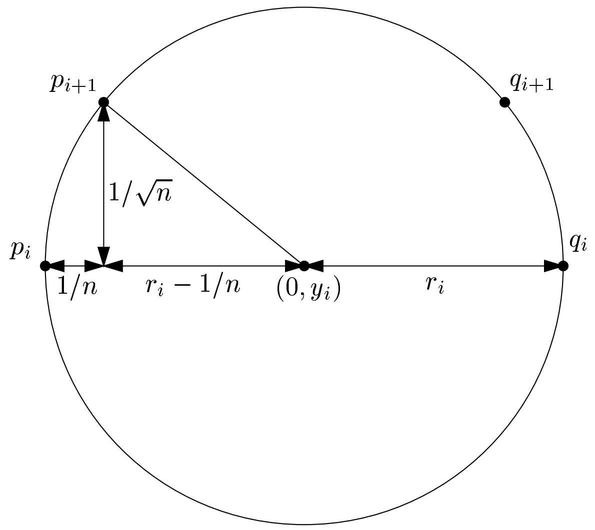

Fig. 1. Illustration for proof of Theorem 1

As noted before, the spanning ratio leaps to infinity for $\beta > 2$, since past this point, the graph may be disconnected. Therefore, it only makes sense to consider spanning ratios when $\beta \in [0, 2]$. When $\beta = 0$, the β-skeleton of a point set has spanning ratio 1. Note that for Gabriel graphs, the above result implies a ratio of $\Omega(n^c), 0.077 < c < 0.078$, thus for $\beta \geq 1$, Theorem 1 provides a much stronger bound of $\Omega(\sqrt{n})$.

4 Lower bound for relative neighborhood graphs

In this section, we show that there exist point sets where the spanning ratio for relative neighborhood graphs (open 2-skeletons) is $\Omega(n)$.

Lemma 2. *The spanning ratio for the relative neighborhood graph of a set of n points in the plane can be $\Omega(n)$.*

Proof. Refer to Figure 2. Let $\theta = 60 - \epsilon$ and $\alpha = 60 + 2\epsilon$. We will fix ϵ later. Since $\alpha + 2\theta = \pi$, the points $a_0, a_1, \ldots, a_n$ are colinear. Similarly, the points $b_0, b_1, \ldots, b_n$ are colinear. The point a_{i+1} blocks the edge $\overline{a_i, b_i}$. An edge $\overline{a_i, b_j}$ for $i < j$ is blocked by a_{i+1} and an edge $\overline{a_i, b_j}$ for $i > j$ is blocked by b_{i+1}. Thus, the only edges in the RNG of these points are $\overline{a_i, a_{i+1}}$, $\overline{b_i, b_{i+1}}$ and $\overline{a_n, b_n}$. Let $A_i = \|a_{i+1} - a_i\|$. Let $B_i = \|b_{i+1} - b_i\|$.

Triangle(a_0, a_1, b_0) and Triangle(a_1, b_1, b_0) are similar, therefore, $B_0 = A_0^2/A$. By the same argument, $A_1 = A_0^3/A^2$, and $B_1 = A_0^4/A^3$. In general, $A_i = A_0^{2i+1}/A^{2i}$ and $B_i = A_0^{2i+2}/A^{2i+1}$.

We choose an ϵ so that $A_0/A > (1/2)^{1/2n}$. Let L be the length of the path from a_0 to b_0. $L > \sum_{i=0}^{n-1} A_i + B_i = \sum_{i=0}^{2n-1} A_0(A_0/A)^i$. Since $A_0/A > (1/2)^{1/2n}$, we have that $\sum_{i=0}^{2n-1} A_0(A_0/A)^i > 1/2 \sum_{i=0}^{2n-1} A_0 = A_0 n$. Therefore, $L > A_0 n$. ■

Lemma 3. *For any $\beta \leq 2$, the ratio $L(x, y)/D(x, y) \leq n - 1$.*

Proof. Let G be the β-skeleton of a set of n points P. Note that the minimum spanning tree $MST(P)$ is contained in G. Let x, y be two points in P.

Let $L(x, y)$ be the length of the unique path from x to y in $MST(P)$. This path has at most $n - 1$ edges and each edge must have length at most $D(x, y)$, otherwise, $MST(P)$ can be made shorter. Therefore, $L(x, y) \leq (n - 1)D(x, y)$. ■

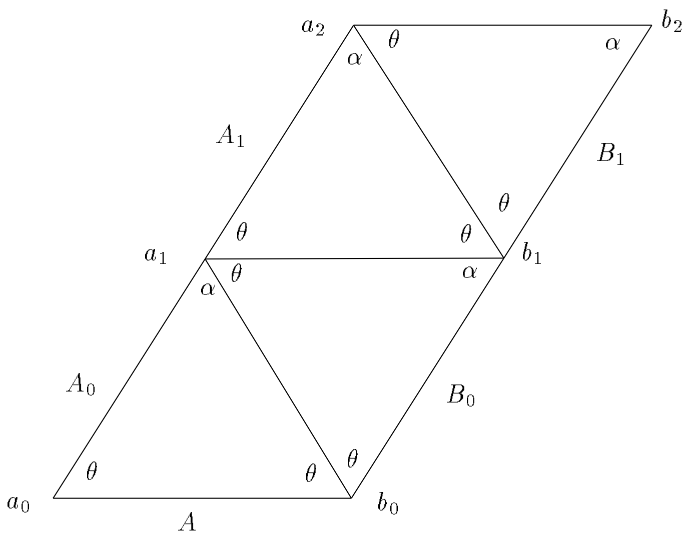

Fig. 2. RNG tower

5 Upper bound

The upper bound established in this section applies to β-skeletons for $\beta \in [0, 1]$. The β-skeleton of a point set P for $\beta \in [0, 1]$ is a graph in which points x and y in P are connected by an edge if and only if there is no other point $v \in P$ such that $\angle xvy > \pi - \arcsin\beta$. Recall that the Gabriel graph is the closed 1-skeleton of P and is a subgraph of the Delaunay triangulation of P.

To upper bound the spanning ratio of β-skeletons, we show that there exits a special walk $SW(x,y)$ in the β-skeleton between the endpoints of any Delaunay edge xy. We upper bound the length $|SW(x,y)|$ of $SW(x,y)$ as a multiple of $D(x,y)$. We then combine this with an upper bound on the spanning ratio of Delaunay triangulations [10, 11] to obtain our result.

Let $DT(P)$ be the Delaunay triangulation of a points set P. In order to describe the walk between the endpoints of a Delaunay edge, we define the *peak* of a Delaunay edge.

Lemma 4. *Let xy be an edge of $DT(P)$. Either xy is an edge of the β-skeleton of P or there exists a unique z (called the peak of xy) such that triangle(xyz) is in $DT(P)$ and z lies in the β-region of xy.*

Proof. Suppose $xy \in DT(P)$ is not an edge in the β-skeleton of P. Then there exists a point $v \in P$ such that $\angle xvy > \pi - \arcsin\beta$. Since xy is an edge of $DT(P)$, there exists a unique z on the same side of xy as v such that disc(xyz) is empty. This implies $\angle xzy \geq \angle xvy$ and thus z lies in the β-region of xy. Since $\beta \leq 1$, disc(xyz) contains that part of the β-region of xy which lies on the other side of xy from z. Since this circle is empty, z is unique.

■

We now define the walk $SW(x,y)$ between the endpoints of the Delaunay edge xy. (Note that in a walk edges may be repeated. See Bondy and Murty for details [1].)

$$SW(x,y) = \begin{cases} xy & \text{if } xy \in \beta\text{-skeleton of } P \\ SW(x,z) \cup SW(z,y) & \text{otherwise } (z \text{ is the peak of } xy) \end{cases}$$

Lemma 5. *Given a set P of n points in the plane. If $xy \in DT(P)$ then the number of edges in $SW(x,y)$ is at most $6n-12$.*

Proof. Since a Delaunay edge is adjacent to at most two Delaunay triangles, an edge can occur at most twice in the walk $SW(x,y)$. Since there are at most $3n-6$ edges in $DT(S)$ by Euler's Formula, $SW(x,y)$ can consist of at most $6n-12$ edges.

■

Lemma 6. *Let P be a set of n points in the plane. For all $x, y \in S$, if $xy \in DT(P)$ then*

$$|SW(x,y)| \leq m^{\gamma} D(x,y)$$

where $\gamma = (1 - \log_2(1 + \sqrt{1-\beta^2}))/2$ and m is the number of edges in $SW(x,y)$.

Proof. The proof is by induction on the number of edges m in $SW(x,y)$. When $m=1$, i.e. $SW(x,y)$ is simply the line segment from x to y, the lemma clearly holds.

If $m>1$, then $|SW(x,y)| = |SW(x,z)| + |SW(z,y)|$ for z the peak of xy. Let k be the number of edges in $SW(x,z)$. Thus, $m-k$ is the number of edges in $SW(z,y)$. Let $a=D(x,y)$, $b=D(x,z)$, and $c=D(y,z)$. Since xz and zy are Delaunay edges, by induction, $|SW(x,z)| \leq bk^\gamma$ and $|SW(z,y)| \leq c(m-k)^\gamma$. Thus it suffices to prove that

$$bk^\gamma + c(m-k)^\gamma \leq am^\gamma \ .$$

As a function of k the left-hand side of the equation is maximized when $k = mc^\phi/(b^\phi+c^\phi)$ where $\phi = 1/(\gamma-1)$. With this substitution for k, after factoring m^γ, it remains to show,

$$b\left(\frac{c^\phi}{b^\phi+c^\phi}\right)^\gamma + c\left(\frac{b^\phi}{b^\phi+c^\phi}\right)^\gamma \leq a \ .$$

By the law of cosines, $a^2 = b^2+c^2-2bc\cos A$ where A is the angle at the peak z. Thus we need only show,

$$b\left(\frac{c^\phi}{b^\phi+c^\phi}\right)^\gamma + c\left(\frac{b^\phi}{b^\phi+c^\phi}\right)^\gamma - \sqrt{b^2+c^2-2bc\cos A} \leq 0 \ .$$

This inequality holds for b,c if and only if it holds for αb,αc for all $\alpha > 0$. Thus we may assume that $b+c=1$. The left-hand side, as a function of b, is maximized at $b=1/2$, and the inequality holds as long as $\gamma \geq (1-\log_2(1-\cos A))/2$. The angle A is minimized (thus maximizing $(1-\log_2(1-\cos A))/2$) when z lies on the boundary of the β-region. For such z, $1-\cos A = 1+\sqrt{1-\beta^2}$.

■

Theorem 3. *The spanning ratio of the β-skeleton of a set P of n points in the plane is at most*

$$\frac{4\pi(6n-12)^\gamma}{3\sqrt{3}}$$

where $\gamma = (1-\log_2(1+\sqrt{1-\beta^2}))/2$.

Proof. Given two arbitrary points x,y in P, let $M = e_1, e_2, \ldots, e_j$ represent the shortest path between x and y in $DT(P)$. Keil and Gutwin [10, 11] have shown that the length of P is at most $2\pi/(3\cos(\pi/6))$ times $D(x,y)$.

For each edge e_i in M, by Lemma 5 and Lemma 6, we know there exists a path in the β-skeleton whose length is at most $(6n-12)^\gamma$ times the length of e_i. Therefore, the shortest path between x and y in the β-skeleton has length at most $2\pi(6n-12)^\gamma/(3\cos(\pi/6))$ times $D(x,y)$. The theorem follows.

■

Corollary 1. *The spanning ratio of the Gabriel graph of an n-point set is at most*

$$\frac{4\pi}{3}\sqrt{2n-4}\ .$$

When β lies strictly between 0 and 1, there is a gap between the upper bound and lower bound on the spanning ratio of β-skeletons. As noted in section 3, the spanning ratio is at least $\Omega(n^c)$ where $c = \log_5(5/(3+2\sin(\theta)))$ and $\theta < (\pi - \sin^{-1}(\beta))/2$. We have shown here that the spanning ratio is at most $O(n^\gamma)$ where $\gamma = (1-\log_2(1+\sqrt{1-\beta^2}))/2$. Refer to Figure 3 for a graph of the exponents of the upper and lower bound. The gap is closed for Gabriel graphs ($\beta = 1$). For Gabriel graphs, the lower bound construction given in section 3, together with the upper bound given here, show that the spanning ratio is indeed $\Theta(\sqrt{n})$.

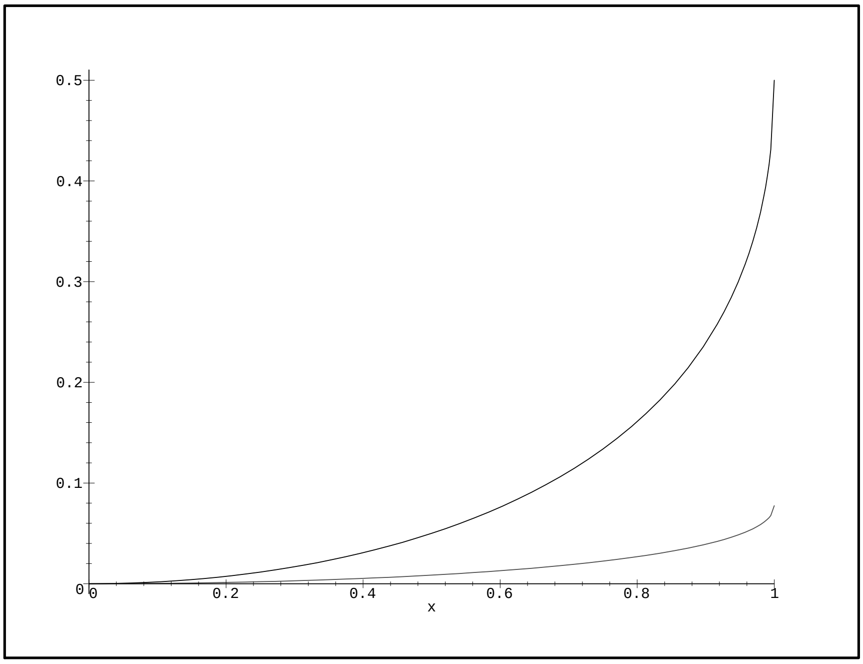

Fig. 3. Gap between upper and lower bound of spanning ratio

6 Random Gabriel graphs.

If n points are drawn uniformly and at random from the unit square $[0,1]^2$, the spanning ratio of the induced Gabriel graph grows unbounded in probability. In particular, we have the following.

Theorem 4. *If n points are drawn uniformly and at random from the unit square $[0,1]^2$, and S is the spanning ratio of the induced Gabriel graph then*

$$\mathbf{P}\left\{ S < c\sqrt{\frac{a\log n}{\log\log n}} \right\} \le 2e^{-2n^{1-12a-o(1)}}$$

for constants c and $a < 1/12$. Thus, for $a < 1/12$, with probability tending exponentially quickly to one,

$$S \ge c\sqrt{a\log n/\log\log n}\ .$$

Proof. The main idea is to show that a set of n points randomly distributed in the unit square contains many tower-like structures of size $c\log n/\log\log n$ each of which has spanning ratio approximately the square root of its size. We first define what a tower-like structure is and then show that the expected number of such structures is large.

A tower-like structure resembles the towers of section 3 but the points may be slightly perturbed. For $i = 1, \ldots, k$, let A_i and B_i be discs both of radius d/k (the constant d will be specified later) located at (r_i, y_i) and $(-r_i, y_i)$ respectively, where the sequences r_i and y_i are given below.

$$r_i = 1 - \frac{i-1}{2k}$$
$$y_i = (i-1)\sqrt{\frac{1/2-(1+\sqrt{2})d}{k}\left(1-\frac{1/2-(1+\sqrt{2})d}{k}\right)}\ .$$

The value of d is chosen so that y_i is positive ($d < 1/(2+2\sqrt{2})$).

Let C be the smallest square enclosing the A_i and B_i within a border of width y_k. (So C extends from $(-3y_k/2 - d/k, -y_k - d/k)$ to $(3y_k/2 + d/k, 2y_k + d/k)$.)

Assume that each of the A_i and B_i contain exactly one point and C contains no other data point beyond these $2k$ points. We claim that among the points in C, the only edges are those connecting A_1 with A_2, A_2 with A_3, and so forth, up to A_{k-1} and A_k. Then A_k connects with B_k, B_k with B_{k-1} and so forth down to B_1. The proof of this claim is rather technical and can be found in the full version of the paper. Note that the A_i's and B_i's are disjoint.

Let u and v be the points in A_1 and B_1 respectively. We have $D(u,v) \le 2 + 2d/k$. Also, any path from u to v entirely in C must be equal in length to the chain, which is longer than $2y_k$. If the path leaves C, then at least two edges leave C, and those edges have length at least $2y_k$, taken together. Thus, $L(u,v) \ge 2y_k$. Therefore,

$$S \ge \frac{L(u,v)}{D(u,v)} \ge \frac{y_k}{1+d/k} \ge c\sqrt{k}$$

for sufficiently large k where c is a constant that depends on d.

Let bC denote the scaled down set $\{bx : x \in C\}$.

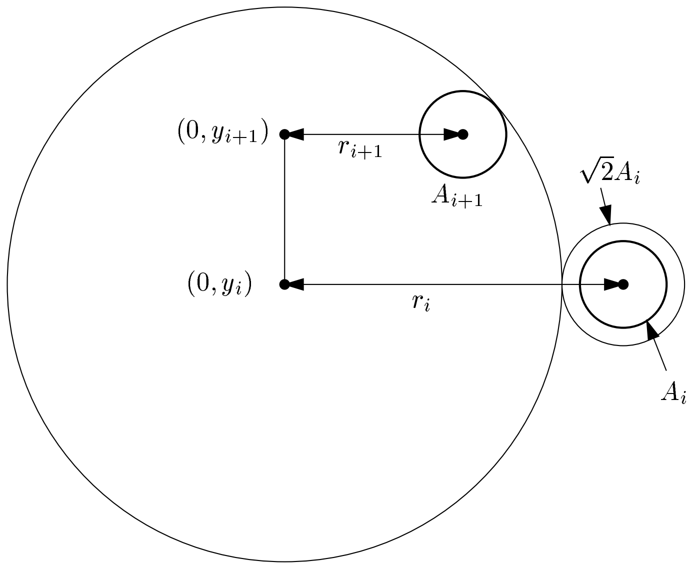

Fig. 4. The construction of A_i and A_{i+1}.

Divide $[0,1]^2$ into n non-overlapping *tiles* of size $1/\sqrt{n} \times 1/\sqrt{n}$. For $b = 1/(4\sqrt{kn})$, bC fits within one of these tiles. Thus we may place n non-overlapping copies of bC within the unit square. For a given data set, we call a tile *tower-like* if it contains exactly $2k$ data points, one each for bA_i and bB_i, $1 \le i \le k$ within it. Let N be the number of tiles that are tower-like.

Clearly, since the distribution is uniform,

$$\mathbf{E}N = n\mathbf{P}\{\text{a tile is tower-like}\} .$$

Pick one tile and partition the n data points over the following disjoint sets: the bA_i's, the bB_i's, $bC - \cup bA_i \cup bB_i$, and $[0,1]^2 - bC$. The cardinalities of these sets, taken together, form a multinomial random vector with probabilities given by the areas of the sets involved. For example, area $(bA_i) = b^2\pi d^2/k^2$. According to the formula for the multinomial distribution,

$$\begin{aligned}
\mathbf{P}\{\text{a tile is tower-like}\} &= \frac{n!}{(n-2k)!}\left(\frac{b^2\pi d^2}{k^2}\right)^{2k}(1-1/n)^{n-2k} \\
&\ge (n-2k+1)^{2k}\left(\frac{\pi d^2}{16nk^3}\right)^{2k}(1-1/n)^{n} \\
&\ge \frac{1}{4}\left(\frac{(n-2k+1)\pi d^2}{16nk^3}\right)^{2k} \\
&\ge \frac{1}{4}\left(\frac{\pi d^2}{32k^3}\right)^{2k}
\end{aligned}$$

provided that n is sufficiently large and $k < (n+2)/4$. We conclude that

$$\mathbf{E}N \geq \frac{n}{4}\left(\frac{\pi d^2}{32k^3}\right)^{2k} .$$

If $k = a \log n / \log\log n$ for a constant $a < 1/6$, then

$$\mathbf{E}N \geq n^{1-6a-o(1)} \to \infty .$$

For each one of these tower-like squares, there is a pair of data points for which the spanning ratio is at least

$$c\sqrt{k} \geq c\sqrt{\frac{a \log n}{\log\log n}} .$$

Change one of the n data points. That will change the number N by at most one. But then, by McDiarmid's inequality [14], we have

$$\mathbf{P}\{|N - \mathbf{E}N| \geq t\} \leq 2e^{-2t^2/n} .$$

In particular, for fixed $\epsilon > 0$,

$$\mathbf{P}\{|N - \mathbf{E}N| \geq \epsilon \mathbf{E}N\} \leq 2e^{-2\epsilon^2 n^{1-12a-o(1)}} \to 0$$

when $a < 1/12$. This shows that $N/\mathbf{E}N \to 1$ in probability for such a choice of a (and thus k), and thus that for every $\epsilon > 0$,

$$\mathbf{P}\{N < (1-\epsilon)\mathbf{E}N\} \to 0 .$$

As another application, we have

$$\begin{aligned}
\mathbf{P}\{S < c\sqrt{a \log n / \log\log n}\} &\leq \mathbf{P}\{N = 0\} \\
&= \mathbf{P}\{N - \mathbf{E}N \leq -\mathbf{E}N\} \\
&\leq 2e^{-2n^{1-12a-o(1)}} \\
&\to 0 .
\end{aligned}$$

Note that this probability decreases exponentially quickly with n.

■

Remark 1. We have implicitly shown several other properties of random Gabriel graphs. For example, a Gabriel graph partitions the plane into a finite number of polygonal regions. The outside polygon which extends to ∞ is excluded. Let D_n be the maximal number of vertices in these polygons. Then $D_n \to \infty$ in probability, because D_n is larger than the maximal size of any tower that occurs in the point set, and this was shown to diverge in probability. From what transpired above, this is bounded from below in probability by $\Omega(a \log n / \log\log n)$. □

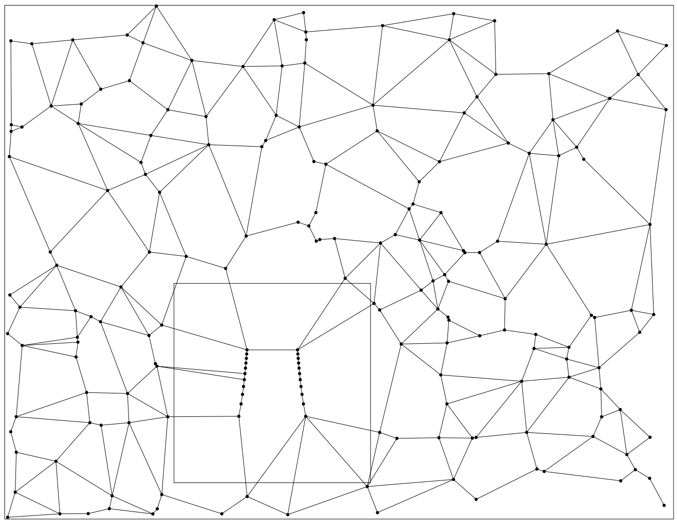

Fig. 5. Gabriel graph with tower-like square.

	$\beta = 0$	$0 < \beta < 1$	$\beta = 1$	$1 < \beta < 2$	$\beta = 2$	$\beta > 2$
Lower Bound	1	$\Omega(n^c)$[6]	$\Omega(\sqrt{n})$	$\Omega(\sqrt{n})$	$\Omega(n)$	∞
Upper Bound	1	$O(n^\gamma)$	$O(\sqrt{n})$	$O(n)$	$O(n)$	∞

$$c = \log_5(5/(3 + 2\sin(\theta))) \text{ and } \theta < (\pi - \sin^{-1}(\beta))/2.$$
$$\gamma = (1 - \log_2(1 + \sqrt{1 - \beta^2}))/2.$$

Table 1. Summary of Results on the Spanning Ratio of β-skeletons

7 Conclusion

We studied the spanning ratio of β-skeletons with β ranging from 0 to 2. This class of proximity graphs includes the Gabriel graph and the relative neighborhood graph. Table 1 summarizes our results. For $\beta > 2$, β-skeletons lose connectivity; thus, their spanning ratio leaps to infinity. For points drawn independently from the uniform distribution on the unit square, we showed that the spanning ratio of the (random) Gabriel graph (and all β-skeletons with $\beta \in [1, 2]$) tends to ∞ in probability as $\sqrt{\log n / \log\log n}$

Several open problems arise from this investigation. It would be interesting to close the gap between upper and lower bounds for β-skeletons in the ranges $0 < \beta < 1$ and $1 < \beta < 2$. Also, for random point sets, it would be interesting to try to find a matching upper bound for the spanning ratio.

References

1. J.A. Bondy and U.S.R. Murty. *Graph theory with applications.* North Holland, 1976.
2. L.P. Chew. There is a planar graph almost as good as the complete graph. In *Proceedings of the 2nd Annual ACM Symposium on Computional Geometry*, pages 169–177, 1986.
3. L.P. Chew. There are planar graphs almost as good as the complete graph. *Journal of Computers and Systems Sciences*, 39:205–219, 1989.
4. L.Devroye. The expected size of some graphs in computational geometry. *Computers and Mathematics with Applications*, 15:53–64, 1988.
5. D.P. Dobkin, S.J. Friedman, and K.J. Supowit. Delaunay graphs are almost as good as complete graphs. In *Proceedings of the 28th Annual Symposium on the Foundations of Computer Science*, pages 20–26, 1987. Also in *Discrete and Computational Geometry*, vol. 5, pp. 399–407, 1990.
6. D. Eppstein. Beta-skeletons have unbounded dilation, Tech. Report 96-15, Dept. of Comp. Sci, University of California, Irvine, 1996.
7. D. Eppstein. Spanning trees and spanners, Handbook of Computational Geometry (J. Sack and J. Urrutia eds.), North Holland, pp. 425–462, 2000.
8. K.R. Gabriel and R.R. Sokal. A new statistical approach to geographic variation analysis. *Systematic Zoology*, 18:259–278, 1969.
9. J. W. Jaromczyk and G. T. Toussaint. Relative neighborhood graphs and their relatives. *Proceedings of the IEEE,* 80(9), pp. 1502-1517, 1992.
10. J.M. Keil and C.A. Gutwin. The Delaunay triangulation closely approximates the complete Euclidean graph. In *Proc. 1st Workshop Algorithms Data Struct.*, volume 382 of *Lecture Notes in Computer Science*, pages 47–56. Springer-Verlag, 1989.
11. J.M. Keil and C.A. Gutwin. Classes of graphs which approximate the complete Euclidean graph. *Discrete and Computational Geometry*, 7:13–28, 1992.
12. D. G. Kirkpatrick and J. D. Radke. A framework for computational morphology. *Computational Geometry, G. T. Toussaint,* Elsevier, Amsterdam, 217-248, 1985.
13. 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*, 12:205–222, 1980.
14. C. McDiarmid. On the method of bounded differences. *Surveys in Combinatorics*, 141:148–188, 1989.
15. F. P. Preparata and M. I. Shamos, *Computational geometry – an introduction.* Springer-Verlag, New York, 1985.
16. G. T. Toussaint. The relative neighborhood graph of a finite planar set. *Pattern Recognition,* 12: 261-268, 1980.