

# On spanners of geometric graphs

Joachim Gudmundsson\*      Michiel Smid†

April 21, 2006

## Abstract

Given a connected geometric graph  $G$ , we consider the problem of constructing a  $t$ -spanner of  $G$  having the minimum number of edges. We prove that for every  $t$  with  $1 < t < \frac{1}{4} \log n$ , there exists a connected geometric graph  $G$  with  $n$  vertices, such that every  $t$ -spanner of  $G$  contains  $\Omega(n^{1+1/t})$  edges. This bound almost matches the known upper bound, which states that every connected weighted graph with  $n$  vertices contains a  $t$ -spanner with  $O(tn^{1+2/(t+1)})$  edges. We also prove that the problem of deciding whether a given geometric graph contains a  $t$ -spanner with at most  $K$  edges is **NP**-hard. Previously, this **NP**-hardness result was only known for non-geometric graphs.

## 1 Introduction

Let  $G = (V, E)$  be a connected undirected graph in which every edge  $e$  has a positive weight  $\omega(e)$ . We define the weight of a path in  $G$  to be the sum of the weights of the edges on this path. For any two vertices  $u$  and  $v$  of  $G$ , we denote the weight of a shortest path in  $G$  between  $u$  and  $v$  by  $\delta_G(u, v)$ . For

---

\*National ICT Australia Ltd., Sydney, Australia. Funded by the Australian Government's Backing Australia's Ability initiative, in part through the Australian Research Council.

†School of Computer Science, Carleton University, Ottawa, Ontario, Canada K1S 5B6. E-mail: [michiel@scs.carleton.ca](mailto:michiel@scs.carleton.ca). This author was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Part of this work was done while visiting NICTA.

a given subgraph  $G' = (V, E')$  of  $G$  (hence,  $E' \subseteq E$ ), we define the *dilation of  $G'$  with respect to  $G$*  to be the value

$$\max \left\{ \frac{\delta_{G'}(u, v)}{\delta_G(u, v)} : u, v \in V, u \neq v \right\}.$$

For a given real number  $t > 1$ , we say that  $G'$  is a  $t$ -spanner of  $G$ , if the dilation of  $G'$  with respect to  $G$  is at most  $t$ .

The following problem has been studied extensively in the literature: Given a connected weighted graph  $G$ , and given a real number  $t > 1$ , does  $G$  contain a  $t$ -spanner having “few” edges?

Althöfer *et al.* [1] showed that for every connected weighted graph  $G$  with  $n$  vertices and for every real number  $t \geq 3$ , there exists a  $t$ -spanner of  $G$  that contains  $O(n^{1+2/(t-1)})$  edges. This result was improved by Baswana and Sen [2] and Roditty *et al.* [16], who showed that for every integer  $t \geq 3$ , any connected weighted graph with  $n$  vertices contains a  $t$ -spanner with  $O(tn^{1+2/(t+1)})$  edges.

The following lower bound was proved by Althöfer *et al.* [1]: For every real number  $t > 1$ , there exists a connected weighted graph  $G$  with  $n$  vertices, such that every  $t$ -spanner of  $G$  contains  $\Omega(n^{1+4/(3(t+2))})$  edges.

We remark that the corresponding problem for unweighted graphs has been considered before by Peleg and Schäffer [15]; see also the book by Peleg [14].

In this paper, we consider the above spanner problem for *geometric graphs*. A graph  $G = (S, E)$  is called a geometric graph, if the vertex set  $S$  of  $G$  is a set of points in  $\mathbb{R}^d$ , and the weight of every edge  $\{u, v\}$  in  $E$  is equal to the Euclidean distance  $|uv|$  between  $u$  and  $v$ .

Since the upper bounds in [1, 16, 2] mentioned above are valid for arbitrary connected weighted graphs, they also hold for geometric graphs. The graph constructed in the proof of the lower bound in [1], however, is not a geometric graph. The difficulty is in mapping the vertices to points in the plane, such that the weight of each edge  $\{u, v\}$  is exactly equal to the Euclidean distance  $|uv|$ . In Section 2, we prove the following theorem, which states that the lower bound of Althöfer *et al.* can almost be achieved by geometric graphs:

**Theorem 1** *For every sufficiently large integer  $n$ , and for every real number  $t$  with  $1 < t < \frac{1}{4} \log n$ , there exists a connected geometric graph  $G$  with  $2n$  vertices, such that every  $t$ -spanner of  $G$  contains  $\Omega(n^{1+1/t})$  edges.*

The proof of Theorem 1 uses an  $n \times n$  connected bipartite graph with  $\Omega(kn)$  edges and whose girth is  $\Omega(\log n / \log k)$ . The probabilistic method has been used to prove the existence of a dense (not necessarily bipartite) graph with high girth; see, for example, Mitzenmacher and Upfal [13]. This existence proof can easily be extended to bipartite graphs. Lazebnik and Ustimenko [12] used algebraic methods to give an explicit construction of a dense bipartite graph with high girth. Chandran [7] used a purely combinatorial approach to construct such a graph, which is, however, not bipartite. In Section 3, we modify Chandran's construction and obtain a simple deterministic algorithm that produces a bipartite graph that we can use to prove Theorem 1.

The spanner problem naturally leads to the following optimization problem: Given a connected weighted graph  $G$  with  $n$  vertices, and given a real number  $t > 1$ , compute a  $t$ -spanner of  $G$ , having the minimum number of edges.

Cai [4] proved that, for any fixed  $t \geq 2$ , this optimization problem is **NP**-hard for unweighted graphs (or, equivalently, for graphs in which all edges have weight one). Cai and Corneil [5] considered the problem for weighted graphs, and showed it to be **NP**-hard for any fixed  $t > 1$ . The problem has also been shown to be **NP**-hard for restricted classes of graphs, such as planar graphs (see Brandes and Handke [3]), chordal graphs, and bipartite graphs (see Venkatesan et al. [20]).

However, the complexity of the optimization problem has not been considered for geometric graphs. In Section 4, we prove this version of the problem to be **NP**-hard as well. Our proof of this result consists of modifying the approach of Cai [4]: We show that any Boolean formula  $\varphi$  in 3-conjunctive normal form can be transformed, in polynomial time, to a geometric graph  $G$  and an integer  $K$ , such that  $\varphi$  is satisfiable if and only if  $G$  contains a  $t$ -spanner with at most  $K$  edges. Again, the main difficulty is in defining  $G$  in such a way that its vertices are points in the plane and the weight of each edge  $\{u, v\}$  is exactly equal to the Euclidean distance  $|uv|$ . Recall that the transformation from  $\varphi$  to the pair  $(G, K)$  has to be done on a Turing machine. Since Turing machines can only deal with finite strings, we take care that the vertices of  $G$  are points in the plane having *rational* coordinates. Thus, the decision version of the optimization problem for geometric graphs is formally defined as follows, for any fixed rational number  $t > 1$ :

**Problem** GEOMMINSPANNER( $t$ ):

**Instance:** A connected geometric graph  $G = (S, E)$ , where  $S \subseteq \mathbb{Q}^2$ , and a positive integer  $K$ .

**Question:** Does  $G$  contain a  $t$ -spanner with at most  $K$  edges?

In Section 4, we prove the following result:

**Theorem 2** *For any rational number  $t > 1$ , problem  $\text{GEOMMINSPANNER}(t)$  is  $\text{NP-hard}$ .*

We do not know if  $\text{GEOMMINSPANNER}(t)$  is in  $\text{NP}$ , because it is not known how to decide, on a Turing machine and in polynomial time, if any given subgraph  $G'$  of a geometric graph  $G$  is a  $t$ -spanner of  $G$ . (The difficulty is in determining whether a rational number is less than a sum of square roots of rational numbers.)

## 1.1 Related work

The problem of constructing geometric spanners with few edges has been considered for point sets. A graph  $G'$ , whose vertex set is a set  $S$  of points in  $\mathbb{R}^d$ , is said to be a  $t$ -spanner for  $S$ , if  $G'$  is a  $t$ -spanner of the complete geometric graph on  $S$ . Salowe [17], Vaidya [19], and Callahan and Kosaraju [6] have shown that, for any set  $S$  of  $n$  points in  $\mathbb{R}^d$ , and for any real constant  $t > 1$ , a  $t$ -spanner for  $S$  with  $O(n)$  edges can be computed in  $O(n \log n)$  time. See also the survey papers by Eppstein [8], Gudmundsson and Knauer [9], and Smid [18].

Gudmundsson *et al.* [10] have shown that if  $S$  is a set of  $n$  points in  $\mathbb{R}^d$ ,  $t > 1$  is a real number, and  $G$  is a  $(1 + \epsilon)$ -spanner for  $S$ , then  $G$  contains a  $t$ -spanner with  $O(n)$  edges. (See also Gudmundsson *et al.* [11]).

Thus, the problem of constructing sparse spanners of geometric graphs  $G$  has been considered for the cases when  $G$  is the complete geometric graph or when  $G$  itself is a spanner of its vertex set. The problem has not been considered for arbitrary geometric graphs  $G$ .

## 2 A geometric graph that contains only dense spanners

In this section, we will prove Theorem 1. Consider a connected (not necessarily geometric) graph  $G$ , in which every edge  $e$  has a positive weight  $\omega(e)$ .

Recall that the *girth* of  $G$  is the minimum number of edges on any cycle in  $G$ . We denote by  $\omega(C)$  the weight of any cycle  $C$  in  $G$ . Thus,  $\omega(C)$  is equal to the sum of the weights of the edges on  $C$ . We define the *weighted girth* of  $G$  to be the quantity

$$\min \left\{ \frac{\omega(C)}{\omega(e)} : C \text{ is a cycle in } G, e \text{ is an edge of maximum weight on } C \right\}.$$

The following lemma relates the girth of  $G$  to its weighted girth.

**Lemma 1** *Let  $G$  be a connected graph, in which every edge  $e$  has a positive weight  $\omega(e)$ . Let  $g$  and  $g_\omega$  be the girth and weighted girth of  $G$ , respectively. Then  $g \geq g_\omega$ .*

**Proof.** Let  $C$  be an arbitrary cycle in  $G$ , let  $e$  be an edge of maximum weight on  $C$ , and let  $m$  be the number of edges on  $C$ . Then,  $\omega(C) \leq m \cdot \omega(e)$ . By the definition of weighted girth, we have  $\omega(C)/\omega(e) \geq g_\omega$ . It follows that  $m \geq g_\omega$ . Hence, we have shown that every cycle in  $G$  contains at least  $g_\omega$  edges. ■

The next lemma relates the dilation of every proper subgraph of  $G$  to the weighted girth of  $G$ .

**Lemma 2** *Let  $G$  be a connected graph in which every edge  $e$  has a positive weight  $\omega(e)$ . Let  $g_\omega$  be the weighted girth of  $G$ . Let  $f$  be an arbitrary edge of  $G$ , and let  $G'$  be the graph obtained by deleting  $f$  from  $G$ . Then the dilation of  $G'$  with respect to  $G$  is at least  $g_\omega - 1$ .*

**Proof.** Let  $u$  and  $v$  be the vertices of  $f$ , i.e.,  $f = \{u, v\}$ , and let  $t$  denote the dilation of  $G'$  with respect to  $G$ . If there is no path in  $G'$  between  $u$  and  $v$ , then  $t = \infty$  and the lemma holds. Otherwise, let  $P$  be a path of minimum weight in  $G'$  between  $u$  and  $v$ . Let  $C$  be the cycle in  $G$  obtained by adding  $f$  to  $P$ , and let  $e$  be an edge of maximum weight on  $C$ . Then  $\omega(f) \leq \omega(e)$  and

$$\frac{\delta_{G'}(u, v)}{\delta_G(u, v)} = \frac{\omega(P)}{\omega(f)} = \frac{\omega(C) - \omega(f)}{\omega(f)} = \frac{\omega(C)}{\omega(f)} - 1 \geq \frac{\omega(C)}{\omega(e)} - 1 \geq g_\omega - 1.$$

Since  $t \geq \delta_{G'}(u, v)/\delta_G(u, v)$ , the proof is complete. ■

The previous two lemmas are valid for arbitrary (i.e., not necessarily geometric) connected weighted graphs. The next lemma shows that any

connected bipartite graph with girth  $g$  can be transformed to a connected geometric graph whose weighted girth is  $\Omega(g)$ . We say that a graph  $G$  is an  $n \times n$  bipartite graph, if its vertex set can be partitioned into two sets  $L$  and  $R$ , each having size  $n$ , such that every edge of  $G$  is between a vertex in  $L$  and a vertex in  $R$ .

**Lemma 3** *Let  $G$  be a connected  $n \times n$  bipartite graph with  $m$  edges and girth  $g$ . Then for every real number  $\epsilon$  with  $0 < \epsilon < 1$ , there exists a set  $S$  of  $2n$  points in the plane and a connected geometric graph with vertex set  $S$  that consists of  $m$  edges and whose weighted girth is at least  $(1 - \epsilon)g$ .*

**Proof.** Let the vertex set of  $G$  be  $L \cup R$ , where  $L \cap R = \emptyset$ ,  $|L| = |R| = n$ , and every edge of  $G$  is between some vertex of  $L$  and some vertex of  $R$ . Let  $\ell_1$  be the vertical line segment with endpoints  $(0, 0)$  and  $(0, \epsilon/2)$ , and let  $\ell_2$  be the vertical line segment with endpoints  $(1 - \epsilon, 0)$  and  $(1 - \epsilon, \epsilon/2)$ . We embed the graph  $G$  in the plane, by mapping the vertices of  $L$  to a set  $S_L$  of  $n$  points on  $\ell_1$ , and mapping the vertices of  $R$  to a set  $S_R$  of  $n$  points on  $\ell_2$ . Let  $S$  be the union of  $S_L$  and  $S_R$ , and let  $G'$  denote the embedded geometric graph. Since  $0 < \epsilon < 1$ , a simple calculation shows that the length of each edge of  $G'$  is in the interval  $[1 - \epsilon, 1]$ . Consider an arbitrary cycle  $C$  in  $G'$ , and let  $e$  be a longest edge on  $C$ . Since  $C$  contains at least  $g$  edges, we have  $\omega(C) \geq (1 - \epsilon)g$ . Thus, since  $\omega(e) \leq 1$ , we have  $\omega(C)/\omega(e) \geq (1 - \epsilon)g$ . Since this lower bound holds for any cycle in  $G'$ , the lemma follows. ■

The previous lemmas imply that we can prove Theorem 1, by constructing a dense bipartite graph whose girth is large. The following lemma states that such a graph exists; the proof will be given in Section 3.

**Lemma 4** *Let  $n$  and  $k$  be positive integers with  $n \geq 3k + 4$  and  $k \geq 2$ . There exists a connected  $n \times n$  bipartite graph with  $kn$  edges, in which the degrees of all vertices are in  $\{k - 1, k, k + 1\}$ , and whose girth is at least*

$$\frac{\log(3n/8)}{\log(k + 1)} + 1 = \log_k n - O(1).$$

Consider the bipartite graph of Lemma 4, and denote its girth by  $g$ . By Lemma 3, we can transform this graph to a geometric graph  $G'$ , whose weighted girth is at least  $(1 - \epsilon)g$ . Then, Lemma 2 implies that every proper subgraph of  $G'$  has dilation at least  $(1 - \epsilon)g - 1$ . Thus, we obtain the following result.

**Lemma 5** *Let  $n$  and  $k$  be positive integers with  $n \geq 3k + 4$  and  $k \geq 2$ , and let  $\epsilon$  be a real number with  $0 < \epsilon < 1$ . There exists a connected geometric graph  $G$  with  $2n$  vertices and  $kn$  edges, such that for every proper subgraph  $G'$  of  $G$ , the dilation of  $G'$  with respect to  $G$  is at least*

$$(1 - \epsilon) \frac{\log(3n/8)}{\log(k+1)} - \epsilon = (1 - \epsilon) \log_k n - O(1).$$

We are now ready to prove Theorem 1. Let  $n$  be a sufficiently large integer, and let  $t$  be a real number with  $1 < t < \frac{1}{4} \log n$ . Define  $\epsilon = 2t / \log n$  and

$$k = (n/4)^{(1-\epsilon)/(t+\epsilon)} - 1.$$

Observe that, by our restriction on  $t$ , the exponent  $(1 - \epsilon)/(t + \epsilon)$  is in the interval  $(0, 1)$ . Therefore, since  $n$  is sufficiently large, we have  $k \geq 2$  and  $n \geq 3k + 4$ . Let  $G$  be the geometric graph in Lemma 5. We claim that this graph has the properties stated in Theorem 1. Indeed, let  $G'$  be an arbitrary  $t$ -spanner of  $G$ . If  $G'$  is a proper subgraph of  $G$ , then, by Lemma 5,

$$t \geq (1 - \epsilon) \frac{\log(3n/8)}{\log(k+1)} - \epsilon.$$

However, our choice of  $k$  implies that

$$t = (1 - \epsilon) \frac{\log(n/4)}{\log(k+1)} - \epsilon < (1 - \epsilon) \frac{\log(3n/8)}{\log(k+1)} - \epsilon.$$

Thus,  $G'$  is equal to  $G$  and, therefore, the number of edges of  $G'$  is equal to

$$kn = \Omega(n^{1+(1-\epsilon)/(t+\epsilon)}).$$

Since  $0 < \epsilon < 1/2$  and  $t > 1$ , we have

$$\frac{1 - \epsilon}{t + \epsilon} \geq \frac{1 - 2\epsilon}{t} = \frac{1}{t} - \frac{4}{\log n}.$$

It follows that the number of edges of  $G'$  is

$$\Omega(n^{1+1/t-4/\log n}) = \Omega(n^{1+1/t}).$$

This completes the proof of Theorem 1.

### 3 Constructing a dense bipartite graph with high girth

In this section, we prove Lemma 4. That is, we construct a connected  $n \times n$  bipartite graph with  $kn$  edges, in which the degrees of all vertices are in  $\{k-1, k, k+1\}$ , and whose girth is  $\Omega(\log_k n)$ . Our construction is a modification of a construction due to Chandran [7], who proved the same result for general, i.e., non-bipartite, graphs.

All graphs in this section are connected and unweighted. (Equivalently, all edge weights are equal to one.) Thus, for any two vertices  $u$  and  $v$  of a graph  $G$ , we denote by  $\delta_G(u, v)$  the minimum number of edges on any path in  $G$  between  $u$  and  $v$ .

The algorithm that constructs a dense bipartite graph with high girth is denoted by  $\text{BIPARTITEHIGHGIRTH}(n, k)$  and is given in Figure 1. This algorithm takes as input two integers  $n$  and  $k$  with  $n \geq 3k + 4$  and  $k \geq 2$ . As we will prove in Sections 3.1 and 3.2, the algorithm returns a connected  $n \times n$  bipartite graph  $G$  with  $kn$  edges and girth at least  $\log_k n - O(1)$ , such that each vertex has a degree in  $\{k-1, k, k+1\}$ .

The algorithm starts by initializing the graph  $G$  to be a Hamiltonian cycle in the complete bipartite graph on  $L \cup R$ . Then, it makes a sequence of  $(k-2)n$  iterations, which are numbered using a counter  $i$  which runs from  $2n+1$  to  $kn$ . In the  $i$ -th iteration, the algorithm takes an ordered pair  $(u, v)$  in  $(L \times R) \cup (R \times L)$ , such that, in the current graph  $G$ , (i)  $u$  has minimum degree, (ii)  $v$  has degree at most  $\lceil i/n \rceil$ , (iii) the edge  $\{u, v\}$  is not in  $G$ , and (iv) the distance between  $u$  and  $v$  is as large as possible. Then, it adds the edge  $\{u, v\}$  to  $G$ . Observe that it is not obvious that a pair  $(u, v)$  for which (i), (ii), and (iii) hold actually exists. We will show in Lemma 8 that such a pair  $(u, v)$  does exist. In particular, this will show that the set  $T$  is never empty and, therefore, it is possible to choose the pair  $(u, v)$  in  $T$  for which  $\delta_G(u, v)$  is maximum.

#### 3.1 Analyzing the size and the degree

We number the iterations of the for-loop according to the value of the variable  $i$ . In other words, the iterations are numbered  $2n+1, 2n+2, \dots, kn$ . Iteration  $j$  denotes the iteration in which the value of the variable  $i$  is equal to  $j$ . In this section, we will prove the following lemma.

**Algorithm** BIPARTITEHIGHGIRTH( $n, k$ )

**Input:** Integers  $n$  and  $k$ , such that  $n \geq 3k + 4$  and  $k \geq 2$ .

**Output:** A connected  $n \times n$  bipartite graph  $G$  with  $kn$  edges and girth at least  $\log_k n - O(1)$ , such that the degree of each vertex is in  $\{k - 1, k, k + 1\}$ .

let  $L$  and  $R$  be two disjoint sets, each having size  $n$ ;

let  $V = L \cup R$ ;

initialize  $G$  to be a Hamiltonian cycle in the complete bipartite graph on  $L \cup R$ ;

**for**  $i = 2n + 1$  **to**  $kn$

**do** let  $M$  be the set of all vertices in  $V$  having minimum degree in  $G$ ;

let  $P = ((M \cap L) \times R) \cup ((M \cap R) \times L)$ ;

let  $T$  be the set of all ordered pairs  $(u, v)$  in  $P$ , such that  $\{u, v\}$  is not an edge in  $G$  and  $\deg_G(v) \leq \lceil i/n \rceil$ ;

let  $(u, v)$  be any pair in  $T$ , such that  $\delta_G(u, v)$  is maximum;

add the edge  $\{u, v\}$  to  $G$

**endfor**;

return the graph  $G$

Figure 1: *The algorithm that constructs a dense bipartite graph with high girth.*

---

**Lemma 6** *Let  $d$  be an integer with  $2 \leq d \leq k$ . At the moment when iteration  $dn$  of the for-loop is completed, the following are true:*

1. *The graph  $G$  consists of  $dn$  edges.*
2. *The degree in  $G$  of every vertex of  $V$  is in  $\{d - 1, d, d + 1\}$ .*
3. *Let  $X$  and  $Z$  be the sets of vertices of  $V$ , whose degrees in  $G$  are equal to  $d - 1$  and  $d + 1$ , respectively. Then,  $|X| = |Z|$ .*

Thus, for  $d = k$ , this lemma implies the claims in Lemma 4 about the number of edges and the degrees of the vertices.

The proof of Lemma 6 is by induction on  $d$ . If  $d = 2$ , then we consider the situation just before the for-loop starts. At that moment,  $G$  is a Hamiltonian

cycle in the complete bipartite graph with vertex set  $L \cup R$ . Thus,  $G$  consists of  $2n$  edges, the degree of every vertex is equal to two, and the sets  $X$  and  $Z$  in the third claim are both empty. As a result, Lemma 6 holds for  $d = 2$ .

We choose an integer  $d$  such that  $2 \leq d < k$ , and assume that Lemma 6 holds for  $d$ . We will prove in Lemmas 7–10 below that the lemma then also holds for  $d+1$ . To prove this, we consider iterations  $dn+1, dn+2, \dots, (d+1)n$  of the for-loop. We will refer to this sequence of  $n$  iterations as the *current batch*. Observe that during the current batch, the value of  $\lceil i/n \rceil$  is equal to  $d+1$ .

**Lemma 7** *At the end of the current batch, the degree in  $G$  of every vertex of  $V$  is less than or equal to  $d+2$ .*

**Proof.** Let  $x$  be an arbitrary vertex in  $V$ . We have to prove that  $\deg_G(x) \leq d+2$  at the end of the current batch.

Consider any edge  $\{u, v\}$ , where  $v = x$ , that is added to  $G$  during the current batch, because the algorithm chooses the pair  $(u, v)$  in  $T$ . It follows from the algorithm that, prior to the moment this edge is added,  $\deg_G(v) \leq d+1$ . Therefore, the addition of edges of this type cannot lead to a degree of  $x$  that is larger than  $d+2$ .

Consider any edge  $\{u, v\}$ , where  $u = x$ , that is added to  $G$  during the current batch, because the algorithm chooses the pair  $(u, v)$  in  $T$ . Assume that this addition makes the degree of  $x$  to be at least  $d+3$ . It follows from the algorithm that, prior to the addition of  $\{u, v\}$ ,  $x$  has minimum degree in  $G$ . In other words, just before  $\{u, v\}$  is added to  $G$ , the degree of every vertex is at least  $d+2$ . In particular, the degree of  $v$  is at least  $d+2$  at that moment. But this implies that, during the iteration in which  $\{u, v\}$  is added to  $G$ , the ordered pair  $(u, v)$  is not in the set  $T$ . This is a contradiction. ■

**Lemma 8** *In each iteration of the current batch, exactly one edge is added to the graph  $G$ .*

**Proof.** By the induction hypothesis, the graph  $G$  consists of  $dn$  edges at the beginning of the current batch. During this batch, at most  $n$  edges are added to  $G$ . It follows that, at any moment during the current batch,

$$\sum_{v \in V} \deg_G(v) \leq 2(d+1)n. \quad (1)$$

Consider one iteration of the current batch, and let  $G'$  be the graph  $G$  at the start of this iteration. Let  $u$  be a vertex of  $V$ , whose degree in  $G'$  is minimum. We may assume without loss of generality that  $u \in L$ .

We claim that, at the start of this iteration, there exists a vertex  $v$  in  $R$ , such that  $\{u, v\}$  is not an edge in  $G'$  and  $\deg_{G'}(v) \leq d + 1$ . Assuming this claim is true, it follows from the algorithm that, during this iteration, the set  $T$  is non-empty and, therefore, an edge is added to  $G'$ . (This edge need not be  $\{u, v\}$  though.)

It remains to prove the claim. Let  $d'$  be the degree of  $u$  in  $G'$ , and let  $v_1, v_2, \dots, v_{d'}$  be all vertices of  $R$  that are connected to  $u$  by an edge of  $G'$ . It follows from the induction hypothesis that

$$\sum_{j=1}^{d'} \deg_{G'}(v_j) \geq d'(d - 1).$$

Moreover, by (1), we have

$$\sum_{v \in R} \deg_{G'}(v) = \frac{1}{2} \sum_{v \in V} \deg_{G'}(v) \leq (d + 1)n. \quad (2)$$

Assume that the claim does not hold. Then, we have  $\deg_{G'}(v) \geq d + 2$  for each  $v \in R \setminus \{v_1, v_2, \dots, v_{d'}\}$ . It follows that

$$\sum_{v \in R} \deg_{G'}(v) \geq d'(d - 1) + (n - d')(d + 2). \quad (3)$$

By combining (2) and (3), we obtain

$$d'(d - 1) + (n - d')(d + 2) \leq (d + 1)n,$$

which can be rewritten as  $n \leq 3d'$ . By Lemma 7, we have  $d' \leq d + 2 \leq k + 1$ , which implies that  $n \leq 3k + 3$ , contradicting our assumption that  $n \geq 3k + 4$ .

■

**Lemma 9** *At the end of the current batch, the degree in  $G$  of every vertex of  $V$  is greater than or equal to  $d$ .*

**Proof.** Consider the sets  $X$  and  $Z$  of vertices of  $V$ , whose degrees in  $G$ , at the beginning of the current batch, are equal to  $d - 1$  and  $d + 1$ , respectively. Since, by the induction hypothesis,  $|X| = |Z|$ , we have  $|X| \leq n$ .

It follows from the algorithm and Lemma 8 that in each iteration of the current batch, one edge  $\{u, v\}$ , where  $u$  has minimum degree in the current graph  $G$ , is added to  $G$ . The induction hypothesis implies that, after this edge has been added, the degree of  $u$  is at least  $d$ . Therefore, after the first  $|X|$  iterations of the current batch,  $G$  does not contain any vertex of degree at most  $d - 1$ . ■

**Lemma 10** *Let  $X'$ ,  $Y'$ , and  $Z'$  be the sets of vertices of  $V$ , whose degrees in  $G$  are equal to  $d$ ,  $d + 1$ , and  $d + 2$ , respectively, at the end of the current batch. Then,  $|X'| = |Z'|$ .*

**Proof.** We observe that, by Lemmas 7–9,

$$|X'| + |Y'| + |Z'| = 2n$$

and

$$d|X'| + (d + 1)|Y'| + (d + 2)|Z'| = 2(d + 1)n.$$

By multiplying the first equation by  $d + 1$ , and subtracting the result from the second equation, the lemma follows. ■

### 3.2 A lower bound on the girth

Let  $G$  be the graph that is returned by algorithm `BIPARTITEHIGHGIRTH`( $n, k$ ). In this section, we will prove the claim in Lemma 4 about the girth of the graph  $G$ .

Let  $g$  be the girth of  $G$ . Since  $G$  is a bipartite graph,  $g$  is even. We will prove that

$$g \geq \frac{\log(3n/8)}{\log(k + 1)} + 1. \quad (4)$$

Let  $C$  be a cycle in  $G$  consisting of  $g$  edges, and let  $\{u, v\}$  be the last edge of  $C$  that is added to  $G$ . Let  $j$  be the integer such that  $\{u, v\}$  is added to  $G$  during iteration  $j$  of the for-loop. We may assume that  $j \geq 2n + 1$ , because otherwise,  $C$  is a Hamiltonian cycle in the complete bipartite graph on  $L \cup R$

and, therefore,  $g = 2n$ , in which case (4) obviously holds. Let  $d = \lceil j/n \rceil$ , and let  $G_j$  be the graph  $G$  at the start of iteration  $j$ . Consider the ordered pair  $(u, v)$  in  $T$  that corresponds to the edge  $\{u, v\}$ . We observe that

$$\delta_{G_j}(u, v) \leq g - 1.$$

We may assume without loss of generality that  $u \in L$ . Define

$$B = \{x \in R : \delta_{G_j}(u, x) \geq g\}.$$

Let  $x$  be an arbitrary element in  $B$ . Then  $\{u, x\}$  is not an edge in  $G_j$ , because, otherwise,  $\delta_{G_j}(u, x) = 1 < g$ . Also, we have

$$\delta_{G_j}(u, x) \geq g > g - 1 \geq \delta_{G_j}(u, v).$$

Since the edge  $\{u, v\}$  is added to  $G_j$  in iteration  $j$ , it follows from the algorithm that  $(u, x) \notin T$ . Thus, the definition of  $T$  implies that  $\deg_{G_j}(x) \geq d+1$ . In fact, by Lemma 6, we have  $\deg_{G_j}(x) = d+1$ . Hence, we have

$$B \subseteq \{x \in R : \deg_{G_j}(x) = d+1\}.$$

Let  $G'$  be the graph  $G$  at the end of iteration  $dn$ , and define

$$Z_R = \{x \in R : \deg_{G'}(x) = d+1\}.$$

Since  $dn \geq j$ , and using Lemma 6, we obtain

$$B \subseteq Z_R.$$

Define

$$X_R = \{x \in R : \deg_{G'}(x) = d-1\}$$

and

$$Y_R = \{x \in R : \deg_{G'}(x) = d\}.$$

By Lemma 6, we have

$$|X_R| + |Y_R| + |Z_R| = n.$$

Also, the definitions of  $X_R$ ,  $Y_R$ , and  $Z_R$ , together with Lemma 6, imply that

$$(d-1)|X_R| + d|Y_R| + (d+1)|Z_R| = dn.$$

It follows that  $|X_R| = |Z_R|$ , implying that  $|Z_R| \leq n/2$ . Thus, since  $B \subseteq Z_R$ , we have  $|B| \leq n/2$  and, hence,

$$|R \setminus B| \geq n/2.$$

Since

$$R \setminus B = \{x \in R : \delta_{G_j}(u, x) \leq g - 1\},$$

and since, by Lemma 6, the degree of every vertex of  $G_j$  is at most  $d + 1$ , it follows that

$$\begin{aligned} |R \setminus B| &\leq (d + 1) + (d + 1)^3 + (d + 1)^5 + \dots + (d + 1)^{g-1} \\ &\leq (k + 1) + (k + 1)^3 + (k + 1)^5 + \dots + (k + 1)^{g-1} \\ &= (k + 1) \frac{(k + 1)^g - 1}{(k + 1)^2 - 1} \\ &\leq \frac{(k + 1)^{g+1}}{(k + 1)^2 - 1} \\ &\leq \frac{(k + 1)^{g+1}}{\frac{3}{4}(k + 1)^2} \\ &\leq \frac{4}{3}(k + 1)^{g-1}. \end{aligned}$$

By combining the lower and upper bounds on the size of  $R \setminus B$ , we obtain

$$n/2 \leq \frac{4}{3}(k + 1)^{g-1}.$$

The latter inequality is equivalent to (4). This completes the proof of Lemma 4.

## 4 The NP-hardness proof

We now prove Theorem 2, i.e., the decision problem  $\text{GEOMMINSpanner}(t)$  is **NP**-hard. Throughout this section, we fix a rational number  $t > 1$ . Recall that  $3SAT$  is the problem of deciding whether or not any given Boolean formula in 3-conjunctive normal form is satisfiable. It is well known that  $3SAT$  is **NP**-complete. To prove Theorem 2, it suffices to design a polynomial-time reduction from  $3SAT$  to  $\text{GEOMMINSpanner}(t)$ . Note that *time* refers to the number of steps made by, say, a Turing machine. Alternatively, time expresses the number of bit operations made in the reduction. In Section 4.2,

we present such a reduction, together with its correctness proof. Our approach is to modify Cai's reduction in [4], which shows that constructing a  $t$ -spanner with the minimum number of edges in any unweighted graph is **NP**-hard. First, in Section 4.1, we introduce so-called forced paths, which are paths in a geometric graph  $G$  that must be in any  $t$ -spanner of  $G$ .

## 4.1 Forced paths

Recall that we have fixed a rational number  $t > 1$ . We fix an even integer  $k$ , such that  $k \geq 4$  and  $k \geq t + 1$ .

Let  $\ell > 0$  be a rational number, and let  $x = (x_1, x_2)$  and  $y = (y_1, y_2)$  be two distinct points in  $\mathbb{Q}^2$ . Let  $\mu$  be a rational number, such that

$$1/|xy| \leq \mu \leq 1/|xy| + 1/\ell, \quad (5)$$

and define the rational number  $\lambda$  as  $\lambda = \ell\mu/k$ . Let  $v$  be the point in  $\mathbb{Q}^2$  defined as

$$v = (\lambda(y_2 - x_2), \lambda(x_1 - y_1)).$$

Observe that the vector from the origin to  $v$  is orthogonal to the line segment joining  $x$  and  $y$ . For  $i = 0, 1, \dots, k/2$ , we define the points  $a_i$  and  $b_i$  in  $\mathbb{Q}^2$  as

$$a_i = x + iv$$

and

$$b_i = y + iv.$$

Finally, we define  $P$  to be the path consisting of the edges

1.  $\{a_0, a_1\}, \{a_1, a_2\}, \dots, \{a_{k/2-1}, a_{k/2}\}$ ,
2.  $\{a_{k/2}, b_{k/2}\}$ , and
3.  $\{b_{k/2}, b_{k/2-1}\}, \dots, \{b_2, b_1\}, \{b_1, b_0\}$ .

See Figure 2(a) for an illustration. We will refer to the path  $P$  as the *forced path* of  $x$  and  $y$  (with respect to  $\ell$ ), and denote it by  $FP(x, y; \ell)$ . Lemma 12 explains this terminology. Before we state this lemma, we prove upper and lower bounds on the length of the path  $P$ :

**Lemma 11** *The length  $|P|$  of the forced path  $P = FP(x, y; \ell)$  satisfies*

$$\ell \leq |P| \leq \ell + 2|xy|.$$

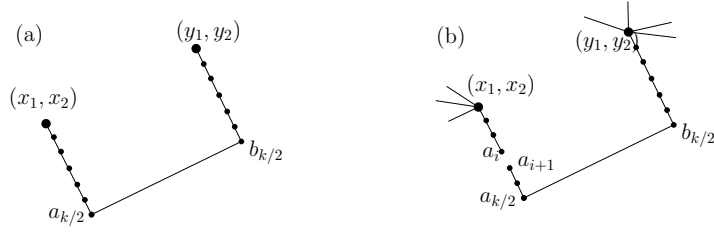


Figure 2: (a) The forced path  $FP(x, y; \ell)$  of  $x$  and  $y$ . (b) Illustrating the proof of Lemma 12.

**Proof.** We first observe that, for each  $i$  with  $0 \leq i < k/2$ ,

$$|a_i a_{i+1}| = |v| = \lambda |xy| = (\ell\mu/k) |xy| \geq \ell/k,$$

where the inequality follows from the left inequality in (5), and, similarly,

$$|b_i b_{i+1}| = |v| = \lambda |xy| = (\ell\mu/k) |xy| \geq \ell/k.$$

Since  $P$  consists of  $k$  edges, each having length at least  $\ell/k$ , plus one additional edge of length  $|a_{k/2} b_{k/2}| = |xy|$ , it follows that  $|P| \geq \ell$ . To prove the upper bound on the length of  $P$ , we first observe that  $|P| = (\ell\mu + 1) |xy|$ . It follows from the right inequality in (5) that  $\ell\mu \leq 1 + \ell/|xy|$ . Therefore, we have

$$|P| \leq (2 + \ell/|xy|) |xy| = \ell + 2|xy|.$$

This completes the proof of the lemma. ■

**Lemma 12** *Let  $G$  be a connected geometric graph, whose vertices are points in  $\mathbb{Q}^2$ , and let  $x$  and  $y$  be two distinct vertices of  $G$  that are not connected by an edge, such that  $|xy| \leq \ell/(t-1)$ . Assume that  $G$  contains the forced path  $P = FP(x, y; \ell)$ . Also, assume that each vertex of  $P \setminus \{x, y\}$  has degree two in  $G$ . Then, every  $t$ -spanner of  $G$  contains the path  $P$ .*

**Proof.** Let  $G'$  be an arbitrary  $t$ -spanner of  $G$ . Let  $i$  be any integer with  $0 \leq i < k/2$ , and assume that the edge  $\{a_i, a_{i+1}\}$  of  $P$  is not an edge in  $G'$ ; see Figure 2(b). Then,

$$\delta_{G'}(a_i, a_{i+1}) > |P| - |a_i a_{i+1}| > (k-1) |a_i a_{i+1}|.$$

Since  $k \geq t + 1$ , it follows that

$$\delta_{G'}(a_i, a_{i+1}) > t|a_i a_{i+1}|,$$

contradicting the fact that  $G'$  is a  $t$ -spanner of  $G$ . Thus, all edges  $\{a_i, a_{i+1}\}$ , with  $0 \leq i < k/2$ , are contained in  $G'$ . By a symmetric argument, all edges  $\{b_i, b_{i+1}\}$ , with  $0 \leq i < k/2$ , are contained in  $G'$ .

Assume that the edge  $\{a_{k/2}, b_{k/2}\}$  of  $P$  is not an edge in  $G'$ . Then,

$$\delta_{G'}(a_{k/2}, b_{k/2}) > |P| = (\ell\mu + 1)|xy| \geq (\ell/|xy| + 1)|xy|.$$

Since  $|xy| \leq \ell/(t - 1)$ , it follows that

$$\delta_{G'}(a_{k/2}, b_{k/2}) > t|xy| = t|a_{k/2}b_{k/2}|,$$

which is again a contradiction. Thus,  $G'$  contains the edge  $\{a_{k/2}, b_{k/2}\}$ . ■

**Lemma 13** *Assume that  $\ell > 0$  is a rational constant. Given the distinct points  $x$  and  $y$  in  $\mathbb{Q}^2$ , the path  $FP(x, y; \ell)$  can be constructed in time that is polynomial in  $L$ , where  $L$  is the total number of bits in the binary representations of the numerators and denominators of the coordinates of  $x$  and  $y$ .*

**Proof.** Given the points  $x = (x_1, x_2)$  and  $y = (y_1, y_2)$ , we have to compute a rational number  $\mu$ , such that

$$0 \leq \mu - \sqrt{\frac{1}{(x_1 - y_1)^2 + (x_2 - y_2)^2}} \leq 1/\ell. \quad (6)$$

That is, we have to approximate the square root in (6) within an absolute precision of  $1/\ell$ . Since  $\ell$  is a constant, we can compute, in time that is polynomial in  $L$ , a rational number  $\mu$  that satisfies (6) and for which the total number of bits in the binary representations of its numerator and denominator is polynomial in  $L$ . Given  $\mu$ , and using our assumption that  $\ell$  and  $k$  are constants, the rational number  $\lambda$ , the point  $v$ , and the points  $a_i$  and  $b_i$  ( $0 \leq i \leq k/2$ ) can all be computed in time that is polynomial in  $L$ . ■

## 4.2 The reduction

We are now ready to give the reduction from *3SAT* to  $\text{GEOMMINSPANNER}(t)$ . Recall that  $t > 1$  is a rational number, and  $k$  is an even integer, such that  $k \geq 4$  and  $k \geq t + 1$ . We define the rational number  $\ell$  as

$$\ell = 2(t - 1)/3.$$

We consider  $t$ ,  $k$ , and  $\ell$  to be constants.

We need the following lemma, which will be used to obtain points on the unit-circle that have rational coordinates and that are close together.

**Lemma 14** *Let  $\rho = \min(2/3, \ell/4)$ , let  $C$  be the circle of radius  $\rho/2$  centered at the point  $(1, 0)$ , let  $i$  be an integer, such that  $i \geq 4/\rho$ , and let  $Q(i)$  be the point*

$$Q(i) = \left( \frac{i^2 - 1}{i^2 + 1}, \frac{2i}{i^2 + 1} \right).$$

*Then,  $Q(i)$  has rational coordinates, is on the unit-circle, and is contained in the interior of the circle  $C$ .*

**Proof.** It is obvious that  $Q(i)$  has rational coordinates and that this point is on the unit-circle. A straightforward calculation shows that the distance between  $Q(i)$  and the center  $(1, 0)$  of  $C$  is less than  $\rho/2$ . This proves that  $Q(i)$  is in the interior of the circle  $C$ . ■

Let  $\varphi$  be a Boolean formula in 3-conjunctive normal form, with variables  $x_1, x_2, \dots, x_N$ , consisting of  $M$  clauses  $c_1, c_2, \dots, c_M$ . Thus, for each  $j$  with  $1 \leq j \leq M$ , the clause  $c_j$  is of the form  $c_j = y_1 \vee y_2 \vee y_3$ , where each of  $y_1$ ,  $y_2$ , and  $y_3$  is either a variable or the negation of a variable.

Our task is to map  $\varphi$  to an instance of  $\text{GEOMMINSPANNER}(t)$ , i.e., a connected geometric graph  $G$ , whose vertex set is a set of points in  $\mathbb{Q}^2$ , and an integer  $K$ , such that  $\varphi$  is satisfiable if and only if  $G$  contains a  $t$ -spanner having at most  $K$  edges.

Let  $z$  denote the origin in  $\mathbb{R}^2$ , and define

$$i^* = \lceil 4/\rho \rceil = \left\lceil \frac{4}{\min(2/3, \ell/4)} \right\rceil.$$

For each  $i$  with  $1 \leq i \leq N$ , we define the following geometric graph  $G_i$  (see Figure 3(a)):

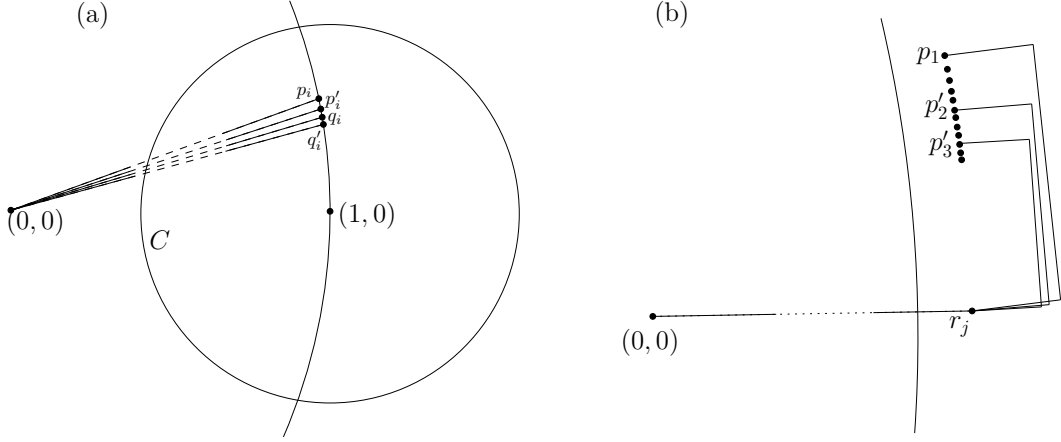


Figure 3: (a) The graph  $G_i$  (without the five forced paths), and (b) the graph  $G^j$ , where  $c_j = (x_1 \vee \overline{x_2} \vee \overline{x_3})$ .

1. Let  $p_i = Q(i^* + 4i)$ ,  $p'_i = Q(i^* + 4i + 1)$ ,  $q_i = Q(i^* + 4i + 2)$ , and  $q'_i = Q(i^* + 4i + 3)$ .
2. The graph  $G_i$  contains the four edges  $\{z, p_i\}$ ,  $\{z, p'_i\}$ ,  $\{z, q_i\}$ , and  $\{z, q'_i\}$ .
3. The graph  $G_i$  contains the five forced paths  $FP(p_i, p'_i; \ell)$ ,  $FP(p_i, q_i; \ell)$ ,  $FP(p_i, q'_i; \ell)$ ,  $FP(p'_i, q_i; \ell)$ , and  $FP(p'_i, q'_i; \ell)$ .

For each  $j$  with  $1 \leq j \leq M$ , we define the following geometric graph  $G^j$  (see Figure 3(b)): Write the clause  $c_j$  as  $c_j = y_1 \vee y_2 \vee y_3$ .

1. Let  $r_j = Q(i^* + 4N + 3 + j)$ .
2. The graph  $G^j$  contains the edge  $\{z, r_j\}$ .
3. For each  $m$  with  $1 \leq m \leq 3$ , if  $y_m$  is equal to the variable, say,  $x_i$ , then  $G^j$  contains the forced path  $FP(r_j, p_i; \ell)$ . On the other hand, if  $y_m$  is equal to the negation of the variable, say,  $x_i$ , then  $G^j$  contains the forced path  $FP(r_j, p'_i; \ell)$ .

We define  $G$  to be the union of the graphs  $G_i$  ( $1 \leq i \leq N$ ) and the graphs  $G^j$  ( $1 \leq j \leq M$ ). Observe that  $G$  is a connected geometric graph, whose vertices are points in  $\mathbb{Q}^2$ . Recall that each forced path consists of

$k + 1$  edges. The graph  $G$  consists of  $1 + (5k + 4)N + (3k + 1)M$  vertices and  $(5k + 9)N + (3k + 4)M$  edges. We define

$$K = (5k + 6)N + (3k + 3)M.$$

Let  $L$  be the number of bits in the representation of the Boolean formula  $\varphi$ . Then,  $L$  is proportional to  $(N + M) \log N$ . Since each vertex of  $G$  can be represented by  $O(\log N + \log M) = O(\log N)$  bits, it follows from Lemma 13 that the graph  $G$  can be constructed in time that is polynomial in  $L$ .

In the rest of this section, we will prove that the Boolean formula  $\varphi$  is satisfiable if and only if the graph  $G$  contains a  $t$ -spanner with at most  $K$  edges.

We first prove upper and lower bounds on the lengths of the forced paths in  $G$ :

**Lemma 15** *The length of each forced path in the graph  $G$  is in the interval  $[\ell, 3\ell/2]$ .*

**Proof.** By Lemma 14, the Euclidean distance between the two endpoints of any forced path is less than  $\rho$ , which is at most  $\ell/4$ . The claim then follows from Lemma 11. ■

The next lemma explains our choice for the integer  $K$ .

**Lemma 16** *Let  $G'$  be an arbitrary  $t$ -spanner of  $G$ . Then, the following two claims are true:*

1.  $G'$  contains at least  $K$  edges.
2. If  $G'$  consists of exactly  $K$  edges, then, for each  $i$  with  $1 \leq i \leq N$ , exactly one of the edges  $\{z, p_i\}$  and  $\{z, p'_i\}$  is in  $G'$ .

**Proof.** We first observe that, by Lemma 14, the Euclidean distance between the two endpoints of any forced path is less than  $\rho$ , which is at most  $2/3$ . Since  $\ell/(t - 1) = 2/3$ , it then follows from Lemma 12 that all forced paths in  $G$  are contained in  $G'$ . The total number of edges in these forced paths is equal to  $(5N + 3M)(k + 1) = K - N$ . We will prove below that, for each  $i$  with  $1 \leq i \leq N$ , the graph  $G'$  contains at least one of the four edges  $\{z, p_i\}$ ,  $\{z, p'_i\}$ ,  $\{z, q_i\}$ , and  $\{z, q'_i\}$ . This will imply that  $G'$  contains at least  $K$  edges and, thus, prove the first claim.

Let  $i$  be any integer with  $1 \leq i \leq N$ , and assume that none of the edges  $\{z, p_i\}$ ,  $\{z, p'_i\}$ ,  $\{z, q_i\}$ , and  $\{z, q'_i\}$  is contained in  $G'$ . Then, any path in  $G'$  between  $z$  and  $q_i$  contains at least one edge of length one and at least two forced paths. Since, by Lemma 15, the length of each forced path is at least  $\ell$ , it follows that

$$\delta_{G'}(z, q_i) \geq 1 + 2\ell = 1 + 2 \cdot 2(t-1)/3 > t = t \cdot \delta_G(z, q_i),$$

contradicting the fact that  $G'$  is a  $t$ -spanner of  $G$ .

To prove the second claim, assume that  $G'$  consists of exactly  $K$  edges. Let  $i$  be an integer with  $1 \leq i \leq N$ . It follows from the argument above that  $G'$  contains exactly one of the edges  $\{z, p_i\}$ ,  $\{z, p'_i\}$ ,  $\{z, q_i\}$ , and  $\{z, q'_i\}$ . If  $G'$  contains  $\{z, q'_i\}$ , then, by the same argument as above, we must have  $\delta_{G'}(z, q_i) > t \cdot \delta_G(z, q_i)$ , contradicting our assumption that  $G'$  is a  $t$ -spanner of  $G$ . Similarly, if  $G'$  contains  $\{z, q_i\}$ , then  $\delta_{G'}(z, q'_i) > t \cdot \delta_G(z, q'_i)$ , which is also a contradiction. Thus,  $G'$  contains exactly one of the edges  $\{z, p_i\}$  and  $\{z, p'_i\}$ . ■

In the next two lemmas, we prove the correctness of our reduction.

**Lemma 17** *If  $G$  contains a  $t$ -spanner with at most  $K$  edges, then the Boolean formula  $\varphi$  is satisfiable.*

**Proof.** Let  $G'$  be a  $t$ -spanner of  $G$  consisting of at most  $K$  edges. Then, by Lemma 16,  $G'$  contains exactly  $K$  edges and, for each  $i$  with  $1 \leq i \leq N$ ,  $G'$  contains exactly one of the edges  $\{z, p_i\}$  and  $\{z, p'_i\}$ .

For each  $i$  with  $1 \leq i \leq N$ , if  $\{z, p_i\}$  is an edge of  $G'$ , then we give the variable  $x_i$  the value *true*, otherwise, we give the variable  $x_i$  the value *false*. We claim that for this assignment of truth values, the Boolean formula  $\varphi$  evaluates to *true*. To prove this, let  $j$  be any integer with  $1 \leq j \leq M$ , and consider the clause  $c_j$  in  $\varphi$ . For ease of notation, let us assume that  $c_j = x_1 \vee \bar{x}_2 \vee \bar{x}_3$ . To prove that  $c_j$  evaluates to *true*, we have to show that at least one of the edges  $\{z, p_1\}$ ,  $\{z, p'_2\}$ , and  $\{z, p'_3\}$  is in  $G'$ . Assume that neither of these edges is in  $G'$ . Observe that  $\{z, r_j\}$  is not an edge in  $G'$ , because otherwise,  $G'$  contains more than  $K$  edges. Thus, every path in  $G'$  between  $z$  and  $r_j$  contains at least one edge of length one and at least two forced paths. Therefore, we have

$$\delta_{G'}(z, r_j) \geq 1 + 2\ell > t = t \cdot \delta_G(z, r_j).$$

This contradicts our assumption that  $G'$  is a  $t$ -spanner of  $G$ . ■

**Lemma 18** *If the Boolean formula  $\varphi$  is satisfiable, then  $G$  contains a  $t$ -spanner with at most  $K$  edges.*

**Proof.** Assume that  $\varphi$  is satisfiable. We fix an assignment of truth values for the variables  $x_1, x_2, \dots, x_N$  for which  $\varphi$  evaluates to *true*. Define the following subgraph  $G'$  of  $G$ :

1.  $G'$  contains all forced paths in  $G$ .
2. For each  $i$  with  $1 \leq i \leq N$ , if  $x_i = \text{true}$ , then  $G'$  contains the edge  $\{z, p_i\}$ , otherwise,  $G'$  contains the edge  $\{z, p'_i\}$ .

We first observe that  $G'$  contains exactly  $K$  edges. To show that  $G'$  is a  $t$ -spanner of  $G$ , it suffices to show the following claim: For each edge  $\{a, b\}$  of  $G$  that is not in  $G'$ , we have  $\delta_{G'}(a, b) \leq t|ab|$ .

Let  $i$  be any index with  $1 \leq i \leq N$ . We may assume without loss of generality that  $\{z, p'_i\}$  is an edge in  $G'$ . Consider the edge  $\{z, p_i\}$  of  $G$ , which is not an edge in  $G'$ . The edge  $\{z, p'_i\}$  and the forced path  $FP(p_i, p'_i; \ell)$  form a path in  $G'$  between  $z$  and  $p_i$ . Thus, using Lemma 15, we have

$$\delta_{G'}(z, p_i) \leq 1 + 3\ell/2 = t = t|zp_i|.$$

In a similar way, it can be shown that  $\delta_{G'}(z, q_i) \leq t = t|zq_i|$  and  $\delta_{G'}(z, q'_i) \leq t = t|zq'_i|$ .

Let  $j$  be any index with  $1 \leq j \leq M$ . Write the clause  $c_j$  as  $c_j = y_1 \vee y_2 \vee y_3$ , and consider the edge  $\{z, r_j\}$  of  $G$ , which is not an edge in  $G'$ . Since  $c_j$  evaluates to *true*, at least one of the literals in  $c_j$  is true. We may assume without loss of generality that  $y_1$  is *true*. If  $y_1 = x_i$ , for some  $i$ , then  $G'$  contains the edge  $\{z, p_i\}$  and the forced path  $FP(r_j, p_i; \ell)$ . It follows that

$$\delta_{G'}(z, r_j) \leq 1 + 3\ell/2 = t = t|zr_j|.$$

On the other hand, if  $y_1 = \bar{x}_i$ , for some  $i$ , then  $G'$  contains the edge  $\{z, p'_i\}$  and the forced path  $FP(r_j, p'_i; \ell)$ . Thus, in this case, we have

$$\delta_{G'}(z, r_j) \leq 1 + 3\ell/2 = t = t|zr_j|.$$

Hence, we have shown that  $G'$  is a  $t$ -spanner of  $G$ . ■

This concludes the proof of Theorem 2.

## 5 Concluding remarks

We have shown that there exist connected geometric graphs that do not contain sparse spanners. More specifically, we have constructed a connected geometric graph  $G$  with  $n$  vertices, such that every  $t$ -spanner of  $G$  contains  $\Omega(n^{1+1/t})$  edges. This bound comes close to the known upper bound of Baswana and Sen [2] and Roditty *et al.* [16], which states that every connected weighted graph with  $n$  vertices contains a  $t$ -spanner with  $O(tn^{1+2/(t+1)})$  edges. The main idea in our proof is to construct a geometric bipartite graph with  $kn$  edges and girth  $\Omega(\log_k n)$ . We leave as an open problem to close the gap between our lower bound and the upper bound in [2, 16].

A  $t$ -spanner of a geometric graph  $G$  is a subgraph  $G'$  that approximates  $G$ , in the sense that distances in  $G$  are approximated (within a multiplicative factor of  $t$ ) by distances in  $G'$ . Thus, if  $G$  is dense and  $G'$  is sparse, then  $G'$  can be regarded to be a “good” approximation of  $G$ . Our lower bound implies that there exist geometric graphs  $G$  that do not contain such a “good” approximation. We leave open the problem of finding classes of geometric graphs that contain sparse  $t$ -spanners. It is known that (i) the class of complete geometric graphs on sets of points in  $\mathbb{R}^d$  and (ii) the class of  $(1 + \epsilon)$ -spanners on sets of points in  $\mathbb{R}^d$ , have this property.

We also showed that computing a  $t$ -spanner with the minimum number of edges of a given geometric graph  $G$  is **NP**-hard. It would be interesting to prove the same result for the complete geometric graph  $G$  on any given set of points in  $\mathbb{R}^d$ .

## References

- [1] I. Althöfer, G. Das, D. P. Dobkin, D. Joseph, and J. Soares. On sparse spanners of weighted graphs. *Discrete & Computational Geometry*, 9:81–100, 1993.
- [2] S. Baswana and S. Sen. A simple linear time algorithm for computing a  $(2k - 1)$ -spanner of  $O(n^{1+1/k})$  size in weighted graphs. In *Proceedings of the 30th International Colloquium on Automata, Languages and Programming*, volume 2719 of *Lecture Notes in Computer Science*, pages 384–396, Berlin, 2003. Springer-Verlag.

- [3] U. Brandes and D. Handke. NP-completeness results for minimum planar spanners. *Discrete Mathematics and Theoretical Computer Science*, 3:1–10, 1998.
- [4] L. Cai. NP-completeness of minimum spanner problems. *Discrete Applied Mathematics*, 48:187–194, 1994.
- [5] L. Cai and D. Corneil. Tree spanners. *SIAM Journal on Discrete Mathematics*, 8:359–387, 1995.
- [6] P. B. Callahan and S. R. Kosaraju. Faster algorithms for some geometric graph problems in higher dimensions. In *Proceedings of the 4th ACM-SIAM Symposium on Discrete Algorithms*, pages 291–300, 1993.
- [7] L. S. Chandran. A high girth graph construction. *SIAM Journal on Discrete Mathematics*, 16:366–370, 2003.
- [8] D. Eppstein. Spanning trees and spanners. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 425–461. Elsevier Science, Amsterdam, 2000.
- [9] J. Gudmundsson and C. Knauer. Dilation and detours in geometric networks. In T. F. Gonzalez, editor, *Handbook on Approximation Algorithms and Metaheuristics*. Chapman & Hall/CRC, Boca Raton, 2006.
- [10] J. Gudmundsson, C. Levcopoulos, G. Narasimhan, and M. Smid. Approximate distance oracles for geometric graphs. In *Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms*, pages 828–837, 2002.
- [11] J. Gudmundsson, G. Narasimhan, and M. Smid. Fast pruning of geometric spanners. In *Proceedings of the 22nd Symposium on Theoretical Aspects of Computer Science*, volume 3404 of *Lecture Notes in Computer Science*, pages 508–520, Berlin, 2005. Springer-Verlag.
- [12] F. Lazebnik and V.A. Ustimenko. Explicit construction of graphs with an arbitrary large girth and of large size. *Discrete Applied Mathematics*, 60:275–284, 1995.
- [13] M. Mitzenmacher and E. Upfal. *Probability and Computing*. Cambridge University Press, Cambridge, UK, 2005.

- [14] D. Peleg. *Distributed Computing: A Locality-Sensitive Approach*. Monographs on Discrete Mathematics and Applications. Society for Industrial and Applied Mathematics, Philadelphia, 2000.
- [15] D. Peleg and A. A. Schäffer. Graph spanners. *Journal of Graph Theory*, 13:99–116, 1989.
- [16] L. Roditty, M. Thorup, and U. Zwick. Deterministic constructions of approximate distance oracles and spanners. In *Proceedings of the 32nd International Colloquium on Automata, Languages and Programming*, volume 3580 of *Lecture Notes in Computer Science*, pages 261–272, Berlin, 2005. Springer-Verlag.
- [17] J. S. Salowe. Constructing multidimensional spanner graphs. *International Journal of Computational Geometry & Applications*, 1:99–107, 1991.
- [18] M. Smid. Closest-point problems in computational geometry. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 877–935. Elsevier Science, Amsterdam, 2000.
- [19] P. M. Vaidya. A sparse graph almost as good as the complete graph on points in  $K$  dimensions. *Discrete & Computational Geometry*, 6:369–381, 1991.
- [20] G. Venkatesan, U. Rotics, M.S. Madanlal, J.A. Makowsky, and C. Pandu Rangan. Restrictions of minimum spanner problems. *Information and Computation*, 136:143–164, 1997.