The Erdős-Sós Conjecture for Geometric Graphs

Luis F. Barba * Ruy Fabila-Monroy † Dolores Lara † Jesús Leaños § Cynthia Rodríguez † Gelasio Salazar ¶ Francisco Zaragoza ‡

Abstract

Let \( f(n, k) \) be the minimum number of edges that must be removed from some complete geometric graph \( G \) on \( n \) points, so that there exists a tree on \( k \) vertices that is no longer a planar subgraph of \( G \). In this paper we show that \( \left( \frac{1}{2} \right) \frac{n^2}{k-1} - \frac{n}{2} \leq f(n, k) \leq 2 \frac{n(n-2)}{k-2} \). For the case when \( k = n \), we show that \( 2 \leq f(n, n) \leq 3 \). For the case when \( k = n \) and \( G \) is a geometric graph on a set of points in convex position, we show that at least three edges must be removed.

1 Introduction

One of the most notorious problems in extremal graph theory is the Erdős-Sós Conjecture, which states that every simple graph with average degree greater than \( k - 2 \) contains every tree on \( k \) vertices as a subgraph. This conjecture was recently proved true for all sufficiently large \( k \) (unpublished work of Ajtai, Komlós, Simonovits, and Szemerédi).

In this paper we investigate a variation of this conjecture in the setting of geometric graphs. Recall that a geometric graph \( G \) consists of a set \( S \) of points in the plane (these are the vertices of \( G \)), plus a set of straight line segments, each of which joins two points in \( S \) (these are the edges of \( G \)). In particular, any set \( S \) of points in the plane in general position naturally induces a complete geometric graph. For brevity, we often refer to the edges of this graph simply as edges of \( S \). If \( S \) is in convex position then \( G \) is a convex geometric graph. A geometric graph is planar if no two of its edges cross each other. An embedding of an abstract graph \( H \) into a geometric graph \( G \) is an isomorphism from \( H \) to a planar geometric subgraph of \( G \). For \( r \geq 0 \), an \( r \)-edge is an edge of \( G \) such that in one of the two open semi-planes defined by the line containing it, there are exactly \( r \) points of \( G \).

In this paper all point sets are in general position and \( G \) is a complete geometric graph on \( n \) points. It is well known that \( G \) contains every tree on \( k \) vertices as a planar subgraph \([2]\), for every integer \( 1 \leq k \leq n \).

Moreover, it is possible to embed any such tree into \( G \), when the image of a given vertex is predefined \([4]\).

Let \( T \) be a subset of edges of \( G \), which we call forbidden edges. If \( T \) is a tree for which every embedding into \( G \) uses an edge of \( T \), then we say that \( T \) forbids \( T \). In this paper we study the question of what is the minimum size of \( T \) so that there is a tree on \( k \) vertices that is forbidden by \( T \). Let \( f(n, k) \) be the minimum of this number taken over all complete geometric graphs on \( n \) points. As \( f(2, 2) = 1 \), \( f(3, 3) = 2 \), \( f(4, 4) = 2 \) and \( f(n, 2) = \binom{n}{2} \), we assume throughout the paper that \( n \geq 5 \) and \( k \geq 3 \).

We show the following bounds on \( f(n, k) \).

Theorem 1

\[
\left( \frac{1}{2} \right) \frac{n^2}{k-1} - \frac{n}{2} \leq f(n, k) \leq 2 \frac{n(n-2)}{k-2}
\]

Theorem 2

\[ 2 \leq f(n, n) \leq 3 \]

In the case when \( G \) is a convex complete geometric graph, we show that the minimum number of edges needed to forbid a tree on \( n \) vertices is three. Some results shown in \([3]\) are closely related to this problem.

An equivalent formulation of the problem studied in this paper is to ask how many edges must be removed from \( G \) so that it no longer contains some planar subtree on \( k \) vertices. A different but related problem is to ask how many edges must be removed from \( G \), so that it no longer contains any planar subtree on \( k \) vertices. For the case of \( k = n \), in \([5]\), it is proved that if any \( n - 2 \) edges are removed from \( G \), it still contains a planar spanning subtree. Note that if the \( n - 1 \) edges incident to any vertex of \( G \) are removed, then \( G \) no longer contains a spanning subtree. In general, for \( 2 \leq k \leq n - 1 \), in \([1]\), it is proved that if any set of \( \frac{n(n-k+1)}{2} - 1 \) edges are removed from \( G \), it still contains a planar subtree on \( k \) vertices. In the same paper it is also shown that this bound is tight.

2 Spanning Trees

In this section we consider the case when \( k = n \). Let \( T \) be a tree on \( n \) vertices. Consider the following algorithm to embed \( T \) into \( G \). Choose a vertex \( v \) of \( T \); root \( T \) at \( v \). For every vertex of \( T \) choose an arbitrary
order of its children. Suppose that the neighbors of $v$ are $u_1, \ldots, u_m$, and let $n_1, \ldots, n_m$ be the number of nodes in their corresponding subtrees. Choose a convex hull point $p$ of $G$ and embed $v$ into $p$. Sort the remaining points of $G$ counter-clockwise by angle around $p$. Choose $m + 1$ rays centered at $p$ so that the wedge between two consecutive rays is convex and between the $i$-th ray and the $(i+1)$-th ray there are exactly $n_i$ points of $G$. Let $S_i$ be this set of points. For each $u_i$ choose a convex hull vertex of $S_i$ visible from $p$ and embed $u_i$ into this point. Recursively embed the subtrees rooted at each $u_i$ into $S_i$. Note that this algorithm provides an embedding of $T$ into $G$. We will use this embedding frequently throughout the paper. See Figure 1.

For every integer $n \geq 2$ we define a tree $T_n$ as follows: If $n = 2$, then $T_n$ consists of only one edge; if $n$ is odd, then $T_n$ is constructed by subdividing once every edge of a star on $\frac{n+1}{2}$ vertices; if $n$ is even and greater than 2, then $T_n$ is constructed by subdividing an edge of $T_{n-1}$. These trees are particular cases of spider trees. See Figure 2.

We prove the lower bound of $f(n, n) \geq 2$ of Theorem 2.

Theorem 3 If $G$ has only one forbidden edge, then any tree on $n$ vertices can be embedded into $G$, without using the forbidden edge.

Proof. Let $e$ be the forbidden edge of $G$. Let $T$ be a tree on $n$ vertices. Choose a root for $T$. Sort the children of each node of $T$, by increasing size of their corresponding subtree. Embed $T$ into $G$ with the embedding algorithm, choosing at all times the rightmost point as the root of the next subtree. Suppose that $e$ is used in this embedding. Let $e := (p, q)$ so that $u$ is embedded into $p$ and $v$ is embedded into $q$ (note that $u$ and $v$ are vertices of $T$).

Suppose that the subtree rooted at $v$ has at least two nodes. In the algorithm, we embedded this subtree rooted at $v$ into a set of at least two points. We chose a convex hull point $(q)$, of this set visible from $p$ to embed $v$. In this case we may choose another convex hull point visible from $p$ to embed $v$ and continue with the algorithm. Note that $(p, q)$ is no longer used in the final embedding.

Suppose that $v$ is a leaf, and that $v$ has a sibling $v'$ whose subtree has at least two nodes. Then we may change the order of the children of $u$ so that $e$ is no longer used in the embedding, or if it is, then $v'$ is embedded into $q$, but then we proceed as above.

Suppose that $v$ is a leaf, and that all its siblings are leaves. The subtree rooted at $u$ is a star. We choose a point distinct from $p$ and $q$ in the point set where this subtree is embedded, and embed $u$ into this point. Afterwards we join it to the remaining points. This produces an embedding that avoids $e$.

Assume then, that $v$ is a leaf and that it has no siblings. We distinguish the following cases:

1. $u$ has no siblings. In this case, the subtree rooted at the parent of $u$ is a path of length two. It is always possible to embed this subtree without using $e$. See Figure 3.

2. $u$ has a sibling $u'$ whose subtree is not an edge. We may change the order of the siblings of $u$, with respect to their parent, so that the subtree rooted at $u'$ will be embedded into the point set containing $p$ and $q$. In the initial order—increasing by size of the corresponding subtrees—$u'$ is after $u$. We may assume that in the new ordering, the order of the siblings of $u$ before it, stays the same. Therefore $p$ is the rightmost point of the set into which the subtree rooted at $u'$ will be embedded. Embed $u'$ into $p$. Either we find an embedding not using $e$, or this embedding fails into one of the cases considered before.

3. $u$ has at least one sibling, all whose corresponding subtrees are edges

Suppose that $u$ has no grandparent; then $T$ is equal to $T_n$ and $n$ is odd. Let $w$ be the parent of $u$. Embed $w$ into $p$. Let $p_1, \ldots, p_{n-1}$ be the points of $G$ different from $p$ sorted counter-clockwise by angle around $p$; choose $p_1$ so that the angle between two consecutive points is less than $\pi$. Let $u_1, \ldots, u_{(n-1)/2}$ be the neighbors of

Figure 1: An embedding of a tree using the algorithm.

Figure 2: $T_7$ and $T_8$. 
Lemma 4. Let $T$ be a tree on $n$ vertices. If $G$ is a convex geometric graph, then $T$ can be embedded into $G$ using at most two convex hull edges of $G$.

\textbf{Proof.} If $T$ is a star, then any embedding of $T$ into $G$ uses only two convex hull edges. If $T$ is a path then it can be embedded into $G$ using at most two convex hull edges. Therefore, we may assume that $T$ is neither a star nor a path.

Since $T$ is not a path, it has a vertex of degree at least three. Choose this vertex as the root. Since $T$ is not a star, the root has a child whose subtree has at least two nodes. Sort the children of $T$ so that this node is first. Embed $T$ into $G$ with the embedding algorithm.

Let $u$ and $v$ be vertices of $T$, so that $u$ is the parent of $v$. Suppose that the subtree rooted at $v$ has at least two nodes. Then in the embedding algorithm we have at least two choices to embed $v$ once the ordering of the children of $u$ has been chosen. At least one of which is such that $(u, v)$ is not embedded into a convex hull edge. Therefore, we may assume that the embedding is such that all the convex hull edges used are incident to a leaf.

Since the first child of the root is not a leaf, there is at most one convex hull edge incident to the root in the embedding. Note that any vertex of $T$, other than the root, is incident to at most one convex hull edge in the embedding. If $n/2$ or more convex hull edges are used, then there are at least $n/2$ non-leaf vertices, each adjacent to a leaf. These vertices must be all the vertices in $T$ and there are only $n/2$ such pairs ($n$ must also be even). Therefore every non-leaf vertex has at most one child which is a leaf. In particular the root has at most one child which is a leaf. Since the root was chosen of degree at least three it has a child which is not a leaf nor the first child; we place this vertex last in the ordering of the children of the root. The leaf adjacent to the root can no longer be a convex hull edge and the embedding uses less than $n/2$ convex hull edges.

Theorem 5. If $G$ is a convex geometric graph and has at most two forbidden edges, then any tree on $n$ vertices can be embedded into $G$, without using a forbidden edge.

\textbf{Proof.} Let $f_0$ be an embedding given by Lemma 4, of $T$ into $G$. For $0 \leq i \leq n$, let $f_i$ be the embedding produced by rotating $f_0$, $i$ places to the right. Assume that in each of these rotations at least one forbidden edge is used, as otherwise we are done. Let $e_1, \ldots, e_m$ be the edges of $T$ that are mapped to a forbidden edge. Some tree on $n$ vertices. Lemma 4 can be proved easily using a previous result (Theorem 2.1 of [3]). We provide a self-contained proof for completeness.
edge in some rotation. Assume that the two forbidden edges are an \( l \)-edge and an \( r \)-edge respectively.

Suppose that \( l \neq r \). Then, each edge of \( T \) can be embedded into a forbidden edge at most once in all of the \( n \) rotations. Thus \( m \geq n \). This is a contradiction, since \( T \) has \( n - 1 \) edges.

Suppose that \( l = r \). Then, each of the \( e_i \) is mapped twice to a forbidden edge. Thus \( m \geq n/2 \). By Lemma 4, \( f_0 \) uses less than \( n/2 \) convex hull edges. Therefore, \( l \) and \( r \) must be greater than 0. But a set of \( n/2 \) or more \( r \)-edges, with \( r > 0 \), must contain a pair of edges that cross. And we are done, since \( f_0 \) is an embedding.

\( \square \)

3 Bounds on \( f(n,k) \)

In this section we prove Theorem 1. First we show the upper bound which can also be seen as a consequence of Theorem 2.2 of [3]. However, we provide a self-contained proof for completeness.

**Lemma 6** If \( G \) is a convex geometric graph, then forbidding three consecutive convex hull edges of \( G \) forbids the embedding of \( T_n \).

**Proof.** Recall that \( T_n \) comes from subdividing a star, let \( v \) be the non leaf vertex of this star. Let \( (p_1, p_2), (p_2, p_3), (p_3, p_4) \) be the forbidden edges, in clockwise order around the convex hull of \( G \). Note that in any embedding of \( T_n \) into \( G \), an edge incident to a leaf of \( T_n \), must be embedded into a convex hull edge. Thus, the leaves of \( T_n \) nor its neighbors can be embedded into \( p_2 \) or \( p_3 \), without using a forbidden edge. Thus, \( v \) must be embedded into \( p_2 \) or \( p_3 \). Without loss of generality assume that \( v \) is embedded into \( p_2 \). But then, the embedding must use \( (p_2, p_3) \) or \( (p_3, p_4) \). \( \square \)

**Lemma 7** If \( G \) is a convex geometric graph, then forbidding any three pairs of consecutive convex hull edges of \( G \) forbids the embedding of \( T_n \).

**Proof.** Let \( p_1, p_2 \) and \( p_3 \) be the vertices in the middle of the three pairs of consecutive forbidden edges of \( G \). Note that a leaf of \( T_n \), nor its neighbor can be embedded into \( p_1, p_2 \) or \( p_3 \), without using a forbidden edge. But at most two points do not fall into this category. \( \square \)

**Lemma 8** \( f(n, k) \leq 2 \frac{n(n-2)}{k-2} \)

**Proof.** Let \( G \) be a complete convex geometric graph. We forbid every \( r \)-edge of \( G \) for \( r = 0, \ldots, \left[ \frac{n-2}{k-2} - 2 \right] \). Note that, in total we are forbidding at most \( n \left( \left[ \frac{n-2}{k-2} - 2 \right] + 1 \right) \leq 2 \frac{n(n-2)}{k-2} \) edges. As every subset of points of \( G \) is in convex position, it suffices to show that every induced subgraph \( H \) of \( G \) on \( k \) vertices is in one of the two configurations of Lemma 6 and 7.

Assume then, that \( H \) does not contain three consecutive forbidden edges in its convex hull nor three pairs of consecutive forbidden edges in its convex hull. \( H \) has at most two (non-adjacent) pairs of consecutive forbidden edges in its convex hull. Therefore every forbidden edge of \( H \) in its convex hull—with the exception of at most two—must be preceded by an \( \ell \)-edge (of \( G \)), with \( \ell > \left[ \frac{n-2}{k-2} - 2 \right] \). \( H \) contains at least \( \frac{k+2}{2} \) of these edges. The points separated by these edges amount to more than \( \frac{k-2}{2} \left[ \frac{n-2}{k-2} - 2 \right] \geq n - k \) points of \( G \). Together with the \( k \) points of \( H \) this is strictly more than \( n \)—a contradiction. \( \square \)

Now, we show the lower bound of Theorem 1.

**Lemma 9** \( f(n, k) \geq \left( \frac{k}{2} \right) \frac{n^2}{k-2} - \frac{n}{2} \)

**Proof.** Let \( F \) be a set of edges whose removal from \( G \) forbids some \( k \)-tree. Let \( H := G \setminus F \). Note that \( H \) contains no complete \( K_k \) as a subgraph, otherwise any \( k \)-tree can be embedded in this subgraph [2]. By Turán’s Theorem [6], \( H \) cannot contain more than \( \left( \frac{k-2}{2} \right) \frac{n^2}{k-2} \) edges. Thus \( F \) must have size at least \( \left( \frac{k}{2} \right) \frac{n^2}{k-2} - \frac{n}{2} \). \( \square \)

**Acknowledgments**

Part of this work was done at the “First Workshop in Combinatorial Optimization at Cineast”. It was continued during a visit of L.F. Barba, R. Fabila-Monroy, J. Leaños and G. Salazar to Abacus research center.

**References**


