Ordered and convex geometric trees with linear extremal function

Zoltán Füredi* Alexandr Kostochka† Dhruv Mubayi‡ Jacques Verstraëte§

Revised September 28, 2019

Abstract

The extremal functions $\mathrm{ex}_{\rightarrow}(n,F)$ and $\mathrm{ex}_{\circlearrowright}(n,F)$ for ordered and convex geometric acyclic graphs F have been extensively investigated by a number of researchers. Basic questions are to determine when $\mathrm{ex}_{\rightarrow}(n,F)$ and $\mathrm{ex}_{\circlearrowright}(n,F)$ are linear in n, the latter posed by Brass-Károlyi-Valtr in 2003. In this paper, we answer both these questions for every tree F.

We give a forbidden subgraph characterization for a family $\mathcal{T}$ of ordered trees with k edges, and show that $\mathrm{ex}_{\rightarrow}(n,T) = (k-1)n - \binom{k}{2}$ for all $n \geq k+1$ when $T \in \mathcal{T}$ and $\mathrm{ex}_{\rightarrow}(n,T) = \Omega(n \log n)$ for $T \notin \mathcal{T}$. We also describe the family $\mathcal{T}'$ of the convex geometric trees with linear Turán number and show that for every convex geometric tree $F \notin \mathcal{T}'$, $\mathrm{ex}_{\circlearrowright}(n,F) = \Omega(n \log\log n)$.

Dedicated to the memory of B. Grünbaum

1 Introduction

An *ordered graph* refers to a graph whose vertex set is linearly ordered and a *convex geometric* or *cg* graph refers to a graph whose vertex set is cyclically ordered. Throughout this paper, an n-vertex ordered or cg graph will be assumed to have vertex set $[n] := \{1,2,\dots,n\}$ with the natural ordering $<$; in the cg setting we use $v < w < x$ to denote that w lies between v and x in the clockwise orientation. For n a positive integer and F an ordered (respectively, cg) graph, let the extremal function $\mathrm{ex}_{\rightarrow}(n,F)$ (respectively, $\mathrm{ex}_{\circlearrowright}(n,F)$) denote the maximum number of edges in an n-vertex ordered (respectively, cg) graph that does not contain F. Both $\mathrm{ex}_{\rightarrow}(n,F)$ and $\mathrm{ex}_{\circlearrowright}(n,F)$ have been extensively studied in the literature, in particular in the case where F is a forest. To describe the known results, we require some terminology.

Given subsets A, B of a linearly ordered set, write $A < B$ to denote that $a < b$ for every $a \in A$ and $b \in B$. The *interval chromatic number* $\chi_i(F)$ of an ordered graph F is the minimum k such that the vertex set of F can be partitioned into sets $A_1 < A_2 < \cdots < A_k$ such that no edge has both endpoints in any A_i. We call these sets *intervals* or *segments*. It is straightforward to see that if $\chi_i(F) > 2$, then $\mathrm{ex}_{\rightarrow}(n,F) = \Theta(n^2)$, since an ordered complete balanced bipartite graph with interval chromatic number two does not contain F.

*Research supported by grant KH 130371 from the National Research, Development and Innovation Office NKFIH and by the Simons Foundation Collaboration grant #317487.

†Research supported in part by NSF grant DMS-1600592 and by grants 18-01-00353A and 16-01-00499 of the Russian Foundation for Basic Research.

‡Research partially supported by NSF awards DMS-1300138 and DMS-1763317.

§Research supported by NSF award DMS-1556524.

1.1 Ordered graphs with interval chromatic number two

Since for every $k \geq 3$, the extremal function of cycle C_k grows at least as $n^{1+1/2k}$, if $\mathrm{ex}(n, F) = n^{1+o(1)}$, then F is acyclic. This motivates the following central conjecture in the area, due to Pach and Tardos [9]:

Conjecture 1.1. *For every forest F with $\chi_i(F) = 2$ we have $\mathrm{ex}_{\to}(n, F) = n(\log n)^{O(1)}$ as $n \to \infty$.*

Conjecture 1.1 remains open in general, though it was verified for all F with at most four edges by Füredi and Hajnal [3]. Based on their work, Tardos [11] determined the order of magnitude of $\mathrm{ex}_{\to}(n, F)$ for every ordered graph F with at most four edges (see Corollary 4.3 in [11]); in particular this verifies Conjecture 1.1 for acyclic graphs with at most four edges and interval chromatic number two. Klazar [4] showed that the partial case of the Füredi-Hajnal conjecture that $\mathrm{ex}_{\to}(n, F) = O(n)$ for every matching F, implies the Stanley-Wilf conjecture, which was proved by Marcus and Tardos [8]. We also point out that extremal problems for ordered forests have applications to theoretical computer science, to search trees and path-compression based data structures (see Bienstock and Györi [1], and Pettie [10] for a survey).

A particularly interesting phenomenon, discovered by Füredi and Hajnal [3], is that the order of magnitude of the extremal function for the ordered forest $\{13, 15, 24, 26\}$ consisting of two interlacing paths of length two is determined by the extremal theory for Davenport-Schinzel sequences, and in particular the extremal function has order of magnitude $\Theta(n\alpha(n))$, where $\alpha(n)$ is the inverse Ackermann function. Further progress towards the conjecture was made by Korándi, Tardos, Tomon and Weidert [5], in the equivalent reformulation of the problem in terms of forbidden 0-1 submatrices of 0-1 matrices, giving a wide class of graphs F for which $\mathrm{ex}_{\to}(n, F) = n^{1+o(1)}$ as $n \to \infty$. The following basic question closely related to Conjecture 1.1 also has information theoretic applications (see for instance [1]).

Problem 1.2. *Determine which ordered forests have linear extremal functions.*

The problem is not even solved for some forests with five edges. And the above example by Füredi and Hajnal of a 4-edge forest with extremal function involving the inverse Ackermann function indicates that the problem is likely to be hard. However, it turns out that for trees the situation is simpler.

In this paper, we resolve Problem 1.2 for ordered trees, and also determine the exact extremal function for ordered trees when the extremal function is linear. This exact result is perhaps surprising, since the situation in the unordered case is complicated, as represented by the Erdős-Sós conjecture. But the ordered situation has the benefit that most trees cannot have linear extremal function. On the other hand, the $\log n$ jump in complexity for trees with nonlinear extremal function is perhaps also interesting.

The description of the ordered trees with linear extremal functions is based on forbidden subtrees. These are ordered paths P, Q and R (shown below, Figure 1) and their mirror copies $\overline{Q}$ and $\overline{R}$ of Q and R, respectively.

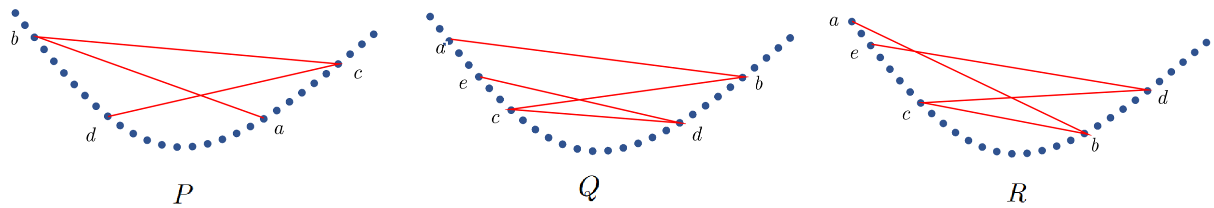

Figure 1 : Forbidden paths P, Q and R.

We are now ready to state our first main result.

Theorem 1.3. *Let T be an ordered tree with k edges and $\chi_i(T) = 2$. If T contains at least one of $P, Q, R, \overline{Q}, \overline{R}$, then $\mathrm{ex}_{\to}(n,T) = \Omega(n \log n)$ as $n \to \infty$, otherwise*

$$\mathrm{ex}_{\to}(n,T) = (k-1)n - \binom{k}{2}$$

for all $n \geq k+1$.

In particular, there is only one linear function that is the extremal function a k-edge ordered tree. As a corollary to Theorem 1.3, if T is any ordered forest containing a path of length four with two or more crossing edges, then $\mathrm{ex}_{\to}(n,T) = \Omega(n \log n)$.

1.2 Convex geometric graphs

Problem 1.2 was posed by Brass-Károlyi-Valtr [2] in the context of convex geometric graphs, and remains open. Using our methods for ordered graphs and some modifications of constructions due to Tardos, we are able to determine all cg trees with linear extremal function. For convenience, we assume that the vertex set V of any cgg we consider lies on a convex arc γ in the plane and the edges are strait line segments with end points on γ. We say that a cgg G is *crossing* or *has a crossing* if some pair of its edges intersect geometrically at an interior point of γ.

Definition 1. *Let $\mathcal{P} = \{P^0, P^1, P^2\}$ denote the family of three cg forests (depicted in Figure 2), where each P^i contains two copies $P = abcd, P' = a'b'c'd'$ of a three-edge path with the following properties:*

- *the center edges $bc \in P$ and $b'c' \in P'$ do not cross each other*
- *edge ab crosses edge cd at p, and $a'b'$ crosses $c'd'$ at p'*
- *the point p and the pair $\{b', c'\}$ lie on the opposite sides of $\{b, c\}$, similarly the point p' and the pair $\{b, c\}$ lie on the opposite sides of $\{b', c'\}$.*

We allow bc and $b'c'$ to share i endpoints where $0 \leq i \leq 2$ and we denote the corresponding member of $\mathcal{P}$ by P^i; hence $|V(P^i)| = 8 - i$. Note that P^1 and P^2 are connected while P^0 is not.

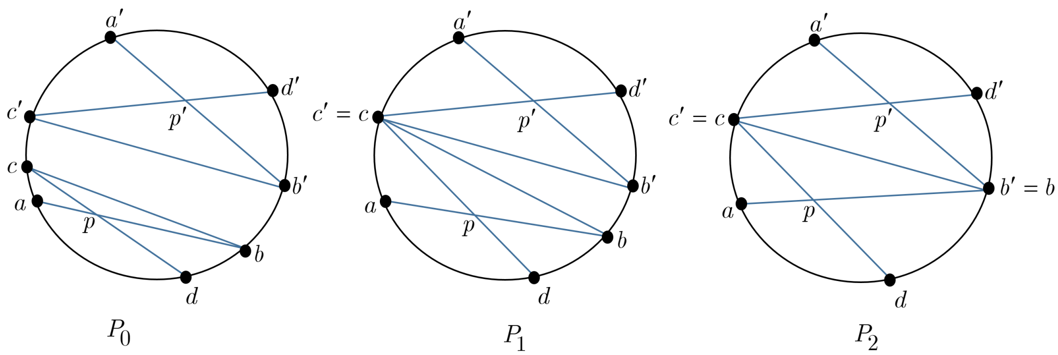

Figure 2 : The family $\mathcal{P}$.

Given vertices $a_1, \ldots, a_t$ in a cyclically ordered set $V \subset \gamma$ we write $a_1 < a_2 < \cdots < a_t$ to mean that the vertices are encountered in the order $a_1, a_2, \ldots, a_t, a_1$ when traversing γ in the clockwise direction. Given subsets A, B of V, write $A < B$ to denote that there are no elements $a, a' \in A$ and $b, b' \in B$ such that $a < b < a' < b'$. In other words, the intervals A and B appear as disjoint arcs/intervals of γ. The definition extends naturally to more than two intervals.

In analogy with the definition of interval chromatic number for ordered graphs, the *cyclic chromatic number* $\chi_c(G)$ of a cg graph G is the minimum k such that the vertex set of G can be partitioned into (nonoverlapping) intervals $A_1 < A_2 < \cdots < A_k$ and no edge has both endpoints in any A_i. It is again straightforward to see that if $\chi_c(G) > 2$, then $\mathrm{ex}_{\circlearrowright}(n, G) = \Theta(n^2)$. Consequently, as we are aiming for a characterization of those G for which $\mathrm{ex}_{\circlearrowright}(n, G) = O(n)$ we may restrict to G with $\chi_c(G) = 2$.

Theorem 1.4. *Fix $k > 2$ and let T be a cg tree with k edges and $\chi_c(T) = 2$. Then either $\mathrm{ex}_{\circlearrowright}(n, T) < (k-1)n$ or $\mathrm{ex}_{\circlearrowright} = \Omega(n \log\log n)$ where the former holds iff T contains no crossing four-edge path and no member of $\mathcal{P}$.*

It is interesting to contrast Theorem 1.4 with Theorem 1.3. As a general rule, determining $\mathrm{ex}_{\circlearrowright}(n, F)$ seems more difficult than determining $\mathrm{ex}_{\rightarrow}(n, F)$. Our experience suggests that both problems exhibit similar but different behavior. For example, it is not true that all k-edge cg trees with linear extremal function have the same extremal function. By definition and known bounds for Turán numbers, for each k-edge cg tree T, $\mathrm{ex}_{\circlearrowright}(n, T) \geq \mathrm{ex}(n, T) \geq \frac{k-1}{2}(n - k/2)$. Perles (see [6, 7]) proved that the k-edge cg non-crossing path P_k with cyclic chromatic number 2 satisfies $\mathrm{ex}_{\circlearrowright}(n, P_k) \sim \frac{k-1}{2}n$ which is almost twice less than the ordered Turán number of each ordered k-edge tree. We were not able to determine $\mathrm{ex}_{\circlearrowright}(n, D_k)$, where D_k is the cg double star with k edges and maximum number of crossings, namely with $\lfloor (k-1)/2 \rfloor \lceil (k-1)/2 \rfloor$ crossings. Perhaps this is the main impediment to obtaining an exact result in Theorem 1.4. Also, we do not know whether every nonlinear extremal function for a cg tree grows at least as $n \log n$.

2 Ordered trees

In this section we prove Theorem 1.3. In Section 2.1, we give the constructions which show that each of the ordered paths $P, Q, R, \overline{Q}, \overline{R}$ has extremal function of order at least $n \log n$. In Section 2.2, we describe the structure of the ordered trees of interval chromatic number two which do not contain $P, Q, \overline{Q}$ or $\overline{R}$. Then in Section 2.3, we determine the extremal function for all those trees.

2.1 Ordered trees with nonlinear extremal function

The paths P, Q and R are displayed in Figure 1. In this section, we present for each of $P, Q, R, \overline{Q}$ and $\overline{R}$ a construction of an n-vertex ordered graphs with $\Theta(n \log n)$ edges that does not contain $P, Q, R, \overline{Q}$ and $\overline{R}$, respectively. These results are not new, and if fact Tardos [11] showed that the extremal functions for $P, Q, R, \overline{Q}$ and $\overline{R}$ are all actually of order $n \log n$.

Construction avoiding P. We start with the simple construction that does not contain P: form an ordered graph on $[n]$ with edges ij such that $|i - j| = 2^h$ for some h. This graph has $\Omega(n \log n)$ edges. It does not contain P, since if $V(P) = \{ad, ac, bd\}$ where $a < b < c < d$, then for some h, i, j, $2^h = |a - d| < |a - c| + |b - d| = 2^i + 2^j$ whereas $\max\{i, j\} < h$ implies $2^i + 2^j \leq 2^h$, a contradiction. Another way to achieve this construction is to take the graph of the k-dimensional cube, where we construct the graph in the usual recursive manner and $n = 2^k$.

Construction avoiding Q and $\overline{Q}$. Bienstock and Györi [1] gave a construction showing $\mathrm{ex}_{\rightarrow}(n, Q) = \Omega(n \log n / \log \log n)$, and a simple construction giving $\mathrm{ex}_{\rightarrow}(n, Q) = \Omega(n \log n)$ was given by Füredi and Hajnal [3]. For $k \geq 0$ and $n = 2^k$, the construction F_n on $[2n]$ has edges only between two independent intervals I_n and J_n of size n with $I_n < J_n$. F_1 has a single edge. Having F_n, with intervals $I_n < J_n$ of size n, the vertex set of F_{2n} consists of four intervals I_n, I'_n and J_n, J'_n of length n, in that order. We put the edges of F_n between I_n and J_n and between I'_n and J'_n, and then add a matching M consisting of an edge from the ith vertex of I' to the ith vertex of J for each $i \in [n]$. For $n = 2^k$ and $f(n) = |F_n|$, we get F_{2n} has $2f(n) + n$ edges. We conclude $f(2n) = 2f(n) + n$ which implies $f(n) = \frac{1}{2} n \log_2 n + n$.

It is shown in Claim 3.3 in [3] that F_n does not contain Q, but to make the paper more self-contained, we also present a short proof of it for each $n = 2^k$ where $k \geq 0$ by induction on k. For $k = 0$, F_n has only one edge. Suppose now that F_n does not contain Q but F_{2n} contains a copy of Q, say a path $T = abcde$ as in the middle of Figure 1. Since $F_{2n} - M$ consists of two components isomorphic to F_n and neither of these components contains Q, some edge of M belongs to T. Since the length of each edge in M is n and the length of every other edge in F_{2n} is larger, the only edge of M in T is cd. In particular, $c \in I'_n$ and $d \in J_n$. Then by the definition of F_{2n}, $b \in J'_n$ and $e \in I_n$. Since $a < e$, also $a \in I_n$. But F_{2n} has no edges between I_n and J'_n, and thus $ab \notin F_{2n}$, a contradiction. This yields $\mathrm{ex}_{\rightarrow}(2n, Q) \geq f(n) = \frac{1}{2} n \log_2 n + n$ for all $n = 2^k$ where $k \geq 0$. The construction does not change if we reverse the ordering, and hence it also avoids $\overline{Q}$ as claimed.

Construction avoiding R. A similar type of construction avoids R. Again for $k \geq 0$ and $n = 2^k$, the construction H_n on $[2n]$ has edges only between two independent intervals I_n and J_n of size n with $I_n < J_n$. Again, H_1 has a single edge, and having H_n, with independent intervals $I_n < J_n$

of size n, the vertex set of H_{2n} consists of four intervals I_n, I'_n and J_n, J'_n of length n, in that order. But now we put the edges of H_n between I_n and J'_n and between I'_n and J_n, and then add a matching M consisting of an edge from the ith vertex of I' to the ith vertex of J' for each $i \in [n]$. Again $|H_n| = f(n) = \frac{1}{2}n \log_2 n + n$ for all $n = 2^k$ where $k \geq 0$.

We show that H_n does not contain R for each $n = 2^k$ where $k \geq 0$ by induction on k. For $k = 0$, H_n has only one edge. Suppose now that H_n does not contain R but H_{2n} contains a copy of R, say a path $T = abcde$ as on the right of Figure 1. As in F_{2n}, some edge of M belongs to T. Suppose this is an edge xy where $x \in \{a, c, e\}$ and $y \in \{b, d\}$. By the definition of H_{2n}, $x \in I'_n$ and $y \in J'_n$. Hence the second edge zy of T incident to y connects J'_n with I_n. It follows that $z < x$, and hence $x \notin \{a, e\}$, i.e., $x = c$. If $y = b$, then since $d > b$, d is also in J'_n, but only one edge in H_{2n} connects $c \in I'_n$ with J'_n. The last possibility is that $y = d$. Then $e \in I_n$ and hence $a \in I_n$. On the other hand, $b \in J_n$. But H_{2n} has no edges between I_n and J_n, and thus $ab \notin H_{2n}$, a contradiction.

Construction avoiding $\overline{R}$. Simply take the reverse to the construction avoiding R.

2.2 Structure of trees not containing $P, Q, R, \overline{Q}$ or $\overline{R}$

In this section, we consider trees which do not contain $P, Q, R, \overline{Q}$ or $\overline{R}$, and describe their structure. The *length* of an edge ij with $i, j \in [n]$ is $|i - j|$. Edges ij and $i'j'$ with $i < j$ and $i' < j'$ *cross* if $i < i' < j < j'$ or $i' < i < j' < j$.

Increasing trees. An *increasing tree* is an ordered tree of interval chromatic number two, with parts equal to intervals $I, J \subseteq [n]$, described as follows. A single edge is an increasing tree. Given an increasing tree, with longest edge ij where $i \in I$ and $j \in J$, we create an increasing tree with one more edge $i'j'$ with $i' \in I$ and $j' \in J$ by requiring $i' = i$ and $j' > j$ or $j' = j$ and $i' < i$. Note that an increasing tree has no crossing edges, and the edges have a unique ordering by the increasing order of their lengths. Also, increasing trees do not contain $P, Q, R, \overline{Q}$ or $\overline{R}$ since they have no crossing edges.

z-trees. A *z-tree* is an ordered tree Z with interval chromatic number two, say with parts equal to intervals I and J, consisting of a union of an increasing tree T with longest edge ij where $i \in I$ and $j \in J$, together with a set S_j of edges of the form hj with $h \in I$ and $h < i$ and a set S_i of edges of the form ik with $k \in J$ and $k > j$. These sets S_i and S_j are allowed to be empty. Note that any two of the edges hj and ik cross.

An example of a z-tree is below, where the tree increasing tree T is shown in solid edges whereas S_i and S_j are in dashed edges.

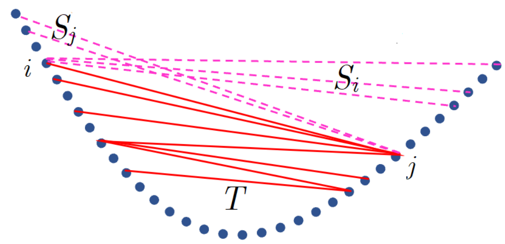

Figure 3 : A z-tree.

Note that the partition $E(T) \cup S_i \cup S_j$ of the edge-set of a z-tree and the edge ij is not uniquely determined by the z-tree. To make the partition unique, take a longest path P^* in a z-tree Z whose edges are strictly increasing in length, and let ij be defined to be the second-to-last edge of the path (see Figure 3). Then S_i is the set of edges ik with $k > j$ and S_j is the set of edges hj with $h < i$. The edge ij and the sets S_i, S_j are uniquely determined by Z, as is the increasing tree $T = Z - S_i - S_j$. By inspection, a z-tree does not contain $P, Q, R, \overline{Q}$ or $\overline{R}$, and we now show the converse:

Theorem 2.1. *If an ordered tree of interval chromatic number two does not contain* $P, Q, R, \overline{Q}$ *or* $\overline{R}$*, then it is a* z*-tree.*

Proof. We begin with the following observation:

> *If* T *is an ordered tree with* $\chi_i(T) = 2$*, not containing* $P, Q, R, \overline{Q}$ *or* $\overline{R}$*, then* T *contains no path of length four with at least one pair of crossing edges.*

The following claim is helpful:

> *If* T *is an ordered tree with* $\chi_i(T) = 2$ *containing crossing edges* $i'j$ *and* ij' *where* $i' < i < j < j'$ *and* $ij \notin E(T)$*, then* T *contains* P *or* Q *or* R *or* $\overline{Q}$ *or* $\overline{R}$*.*

We first prove the claim. Since T is a tree, there exists a path in T whose first and last edges are $i'j$ and ij'. If this path has length at least four, then we can find a subpath of length four with a crossing, a contradiction. Therefore the path must have length three. Since $\chi_i(T) = 2$, and $i' < i < j < j'$, i' and i are not adjacent and j and j' are not adjacent. Therefore the only possibility is that $i'j' \in E(T)$, but then the edges $i'j', i'j, ij'$ form a copy of P in T, a contradiction. This proves the claim.

Now we prove the theorem. Let Z be an ordered tree with $\chi_i(Z) = 2$ not containing $P, Q, R, \overline{Q}$ or $\overline{R}$, with intervals $I < J$. Let xy be an edge of Z such that $x \in I$ and $y \in J$ and y has degree 1 in Z (the case $x \in J$ and $y \in I$ is similar). Then $Z' = Z - \{y\}$ does not contain $P, Q, R, \overline{Q}$ or $\overline{R}$, so Z' is a z-tree. We may write $E(Z') = E(T) \cup S_i \cup S_j$, where T is an increasing tree with the longest edge ij with $i \in I$ and $j \in J$, and $S_i = \{ik : k > j\}$ and $S_j = \{hj : h < i\}$. The edge ij is the second-to-last edge of a longest path P^* in Z' whose edge lengths are increasing. Let $ik \in S_i$ be the last edge of the path where $k \in J$ and $k > j$ (the case that the last edge is hj with $h < i$ is similar).

Case 1. *The edge xy crosses an edge $ab \neq ik$ in P^* where $a \in I$ and $b \in J$.* We may assume that ab is the closest to ik edge on P^* with this property. In this case, either $x < a < b < y$ or $a < x < b < y$. By the claim, if $x < a < b < y$, then $ay \in E(Z)$, contradicting that y has degree 1 in Z. So $a < x < b < y$, and the claim gives $xb \in E(Z')$. Now the path $Q^* \subset P^*$ starting with the edges yx, xb and ba and ending with the edge ik has length at least four in Z and has a crossing, which is a contradiction. This completes Case 1.

Case 2. *The edge xy crosses no edge of $P^* - ik$.* If $x > i$ then $T \cup \{xy\}$ is an increasing tree and Z is therefore a z-tree. We conclude $x \leq i$. If $x < i$, then there is an edge $xj \in S_j$. But then the path $yxjik$ is a path of length four in Z with a crossing, a contradiction. We conclude $x = i$, and $y > j$, and now $E(Z) = E(T) \cup S_i' \cup S_j$ where $S_i' = S_i \cup \{iy\}$, so Z is a z-tree. □

2.3 Ordered trees with linear extremal function

This section is devoted to determining the extremal function for z-trees, thereby completing the proof of Theorem 1.3. We determine first the extremal function of increasing trees.

Lemma 2.2. *Let T be an increasing tree with k edges. Then $\mathrm{ex}_{\rightarrow}(n, T) = (k-1)n - \binom{k}{2}$ for $n \geq k+1$.*

Proof. Observe that the longest edge in T has length at least k. Therefore the ordered n-vertex graph G^* consisting of all edges ij with $i, j \in [n]$ such that $1 \leq |i-j| < k$ cannot contain T, and so

$$\mathrm{ex}_{\rightarrow}(n, T) \geq e(G^*) = \sum_{i=1}^{k-1}(n-i) = (k-1)n - \binom{k}{2}.$$

Now we establish equality. Suppose G is an n-vertex ordered graph that does not contain T. We prove by induction on k that $e(G) \leq (k-1)n - \binom{k}{2}$ for $n \geq k+1$. For $k = 1$, this is clear since any single edge is an increasing tree, so G in that case is empty. Suppose T is an increasing tree with $k+1$ edges. Let uv be the longest edge of T, where $u < v$, and suppose v is a leaf of T. Let $T' = T - uv$. Assuming $V(G) = [n]$, remove for every $i \leq n-k$ in $V(G)$ the longest edge ij with $j > i$. Then the total number of edges removed from G is at most $n-k$. We therefore obtain an ordered graph G' with at least $e(G') \geq e(G) - (n-k)$ edges. If G' contains the ordered tree T', say u is mapped to $i \in [n]$, then $i \leq n-k$, so there exists an edge $ij \in E(G) \backslash E(G')$ such that j is larger than any vertex in the embedding of T' in G. Then adding ij we get an embedding of $T' + uv = T$ in G, a contradiction. We conclude G' does not contain T', so by induction $e(G') \leq (k-1)n - \binom{k}{2}$. Therefore

$$e(G) \leq (k-1)n - \binom{k}{2} + (n-k) = kn - \binom{k+1}{2}.$$

This completes the proof. □

This proof extends to give the extremal function for z-trees in a fairly simple way:

Lemma 2.3. *Let Z be a z-tree with k edges. Then for $n \geq k+1$, $\mathrm{ex}_{\rightarrow}(n, Z) = (k-1)n - \binom{k}{2}$.*

Proof. We may write $Z = T \cup S_i \cup S_j$ where T is an increasing tree with the longest edge ij with $i < j$, S_i consists of the edges ig with $g > j$ and S_j consists of the edges hj with $h < i$. Suppose $|E(T)| = a$, $|S_j| = b$ and $|S_i| = c$. By definition, $a + b + c = k$. We construct an n-vertex ordered graph G^* with no copy of Z as follows: the vertex set of G^* is $[n]$, whereas the edge set consists of $E_a = \{xy : x < y, y - x < a\}$, $E_b = \{xy : x < y, x \leq b\}$ and $E_c = \{xy : x, y, y > n - c\}$. Let

$$f(a, b, c) = |E(G^*)| = |E_a \cup E_b \cup E_c| \textit{ and } E_0 = \{xy : 1 \leq x < y \leq n \text{ and } xy \notin E_a \cup E_b \cup E_c\}.$$

By the definitions of E_a, E_b and E_c, using $y' = y - a + 1$,

$$E_0 = \{xy : 1 + b \leq x \leq y - a \leq (n - c) - a\} = \{xy' : 1 + b \leq x < y' \leq n - c - a + 1\}.$$

Since $a + b + c = k$, we conclude $|E_0| = \binom{n-k+1}{2}$. Hence

$$f(a, b, c) = |E(G^*)| = \binom{n}{2} - |E_0| = \binom{n}{2} - \binom{n-k+1}{2} = (k - 1)n - \binom{k}{2}.$$

Furthermore, $E_a \cup E_b \cup E_c$ does not contain a copy of Z: since T has a edges, ij has length at least a, so $ij \notin E_a$. If $ij \in E_b$, then $i \leq b$. However, Z has b vertices preceding i, namely the vertices in S_j, so this is not possible. Similarly, if $ij \in E_c$, then $j > n - c$, but since Z has the c vertices in S_i after j, this is impossible, as well. Therefore G^* does not contain Z, and we have $\mathrm{ex}_{\rightarrow}(n, Z) \geq (k - 1)n - \binom{k}{2}$.

We now prove $\mathrm{ex}_{\rightarrow}(n, Z) = f(a, b, c)$ by induction on $|S_i| = c$. If $c = 0$, then Lemma 2.2 proves the required equality. If $c \geq 1$, then we observe $f(a, b, c) - f(a, b, c - 1) = n - k + 1$. Let G be an n-vertex ordered graph not containing Z. Following the notation above, with ij the longest edge of $T \subset Z$, for each vertex $g : b < g \leq n - a - c + 1$, delete the longest edge $gh \in E(G)$ with $h > g$. The number of deleted edges is $n - a - b - c + 1 = n - k + 1$. If this new graph G' contains $Z' = Z - ij'$ where j' is the last vertex of Z, then G contains Z: we observe $b < i \leq n - a - c + 1$, and so there is a longest edge $ij' \in E(G) \backslash E(G')$ which can be added to Z' to get Z. Therefore G' does not contain Z', and by induction, $e(G') \leq f(a, b, c - 1)$ which implies $e(G) \leq e(G') + n - k + 1 \leq f(a, b, c - 1) + n - k + 1 = f(a, b, c)$. This completes the proof. □

3 Convex geometric trees

In this section we prove Theorem 1.4. We denote a crossing four-edge path by the shorter notation crossing P_4. In Section 3.1, we give the constructions which show that each P^i and each crossing P_4 has extremal function of order at least $n \log \log n$. In Section 3.2, we describe the structure of the cg trees T with $\chi_c(T) = 2$ which contain neither a crossing P_4 nor any P^i and then show that these trees have linear extremal function.

3.1 Convex geometric trees with nonlinear extremal function

We begin by noting that Brass-Károlyi-Valtr [2] proved that $\mathrm{ex}_{\circlearrowright}(n, P^i) = \Theta(n \log n)$. Actually, they proved this only for P^0 but exactly the same proof (both upper and lower bounds) works for

P^1 and P^2 as well.

In order to present our constructions that avoid a crossing P_4, we need a theorem of Tardos [12]. The setup of his theorem is as follows. We are given a bipartite graph $G = (A, B, E)$ with a proper edge coloring c with d colors in which the colors are linearly ordered.

A walk $e_1e_2e_3e_4$ is called *fast* if $c(e_2) < c(e_3) < c(e_4) \le c(e_1)$. A walk $e_1e_2e_3e_4$ is called *slow* if it starts in B, $c(e_2) < c(e_3) < c(e_4)$ and $c(e_2) < c(e_1) \le c(e_4)$.

Theorem 3.1 (Tardos [12] p. 549)**.** *Let $G = (A, B, E)$ be a bipartite graph with a proper edge coloring with d colors. There exists a subgraph $G' = (A, B, E')$ of G without slow walks and with $|E'| > \frac{\log d}{480d}|E|$. Similarly, there exists a subgraph $G'' = (A, B, E'')$ of G without fast walks and with $|E''| > \frac{\log d}{480d}|E|$.*

We are now ready to present our constructions for Theorem 1.4

Construction. Let $v_1, \dots, v_n$ be in clockwise order on γ and form the vertex set V of our construction F_n, where $n = 2^k$. The edge set of F_n consists of $k-1$ matchings $M_1, \dots, M_{k-1}$, each of size $n/4$. For each $1 \le j \le k-1$,

$$M_j = \{v_{2i-1}v_{2i-2+2^j} \ : \ i \in [n/4]\}.$$

Let $V_1 = \{v_1, v_3, \dots, v_{n-1}\}$ and $V_2 = \{v_2, v_4, \dots, v_n\}$. For every edge $e = v_iv_j$ in F_n,v_j is the *left end* if $j < i$ and *right end* otherwise. Note that

(i) $|E(F_n)| = (k-1)n/4 = (\log_2 n - 1)n/4$;
(ii) the left ends of all edges are in V_1 and all right ends are in V_2;
(iii) F_n does not contain a path $v_{i_1}v_{i_2}v_{i_3}v_{i_4}$ such that $i_2 < i_4 < i_1 < i_3$.

Case (iii) is referred to by Tardos [12] as a *heavy path*. We consider $M_1, \dots, M_{k-1}$ as color classes of an edge coloring c of F_n in which the colors are ordered according to their indices.

By Theorem 3.1, F_n contains subgraphs $F_{n,1}$, $F_{n,2}$ and $F_{n,3}$ such that

(P1) $|F_{n,j}| \ge \frac{\log(k-1)}{480(k-1)}|E(F_n)| \ge \frac{\log(k-1)}{1920}n$ for each $1 \le j \le 3$;
(P2) $F_{n,1}$ does not contain fast walks;
(P3) $F_{n,2}$ does not contain slow walks starting in V_1;
(P4) $F_{n,3}$ does not contain slow walks starting in V_2.

We also will use the cg graph $F_{n,0}$ with the same vertex set V and $E(F_{n,0}) = \{v_iv_j \ : \ 1 \le i \le n/2, n/2+1 \le j \le n\}$. □

Definition 2. *We denote by L the cg three-edge path with interval chromatic number greater than two. In other words, L is the 3-edge cg path $xyzu$ such that $x < y < z < u$ (in the cyclic ordering).*

Remark. The cg graph $F_{n,0}$ does not contain L.

Definition 3. *A path $P = x_1x_2\dots x_s$ in a cgg is a* zigzag *if it has no crossing and for every $2 \le j \le s-2$, the sets $\{x_1,\dots,x_{j-1}\}$ and $\{x_{j+2},\dots,x_s\}$ are on different sides of the chord x_jx_{j+1}. Alternatively, P has no crossing and $\chi_c(P) = 2$.*

Perles (see [7] p. 292) proved that $\mathrm{ex}_{\circlearrowright}(n,P) = O(n)$ for any zigzag path P. Our main result is that the construction presented above contains no copy of a crossing P_4.

Theorem 3.2. *For every four-edge path P in a cgg apart from the zigzag path, $\mathrm{ex}_{\circlearrowright}(n,P) = \Omega(n\log\log n)$. Moreover, P is not contained in one of $F_{n,j}$ for $0 \le j \le 3$.*

Proof. Since each $F_{n,j}$ has at least $\Omega(n\log k) = \Omega(n\log\log n)$ edges, it suffices to show that for each crossing four-edge path P there is a j for which $P \not\subset F_{n,j}$. Every type of a cgg four-edge path corresponds to a cyclic permutation of [5]. So we need to consider $4! = 24$ types of them. By the remark above, it is enough to consider the types with cyclic chromatic number 2. Suppose such a path (i.e. with cyclic chromatic number 2 and distinct from a zigzag path) $P = abcdf$ can be embedded into $F_{n,1}$. Suppose that for $x \in \{a,b,c,d,f\}$, x is mapped onto v_{i_x}.

Assume i_c is odd (the proof for even i_c will be symmetric). By the structure of F_n, $i_b > i_c$ and $i_d > i_c$. So by symmetry we may suppose

$$i_c < i_d < i_b. \tag{1}$$

Again by the structure of F_n, $i_b > i_a$. If $i_d < i_a < i_b$, then the cyclic chromatic number of the path a,b,c,d is 3, a contradiction. The situation $i_c < i_a < i_d$ impossible because the lengths of the edges are of the form $2^j - 1$. So, the only possibility is $i_a < i_c < i_d < i_b$.

Once again by the structure of F_n, $i_d > i_f$. If $i_c < i_f < i_d$, then P is a zigzag path, contradicting our choice. If $i_f < i_a$, then $i_d - i_f \ge i_b - i_a$, since otherwise $2(i_d - i_f) < i_b - i_a$ and $2(i_b - i_c) < i_b - i_a$, a contradiction. But in this case, $F_{n,1}$ contains the fast walk $fdcba$ contradicting (P2). Thus, $i_a < i_f < i_c$. This means we need to consider only cyclical structure $(1,5,3,4,2)$ and (when we switch from the case of odd i_c to the even) $(5,1,3,2,4)$.

Case 1: $(1,5,3,4,2)$. We claim that $F_{n,2}$ does not contain it. Indeed, suppose it does. If i_c is odd, then repeating the above argument we come to $i_a < i_f < i_c$. But this means $F_{n,2}$ contains the slow walk $fdcba$ contradicting (P3). Thus assume i_c is even. Then by the structure of F_n, $i_d < i_c$ and $i_b < i_c$, say $i_d < i_b < i_c$. Then by the cyclic structure of our path, $i_b < i_f < i_c$. This is impossible, since $2(i_f - i_d) < i_c - i_d$ and $2(i_c - i_b) < i_c - i_d$.

Case 2: $(5,1,3,2,4)$. We claim that $F_{n,3}$ does not contain this path. The proof is symmetric to Case 1. ☐

3.2 Structure of trees avoiding $\mathcal{P}$ and a crossing P_4

We need the definitions of increasing trees and z-trees in the cg setting.

Convex geometric increasing trees. A *cg increasing tree* is obtained as follows. Start with an ordered increasing tree with vertex set $[n]$ (see the definition in Section 2.2) and view the linear ordering of the vertices as a cyclic ordering $n < n-1 < \cdots < 2 < 1$. Note that a cg increasing tree has no crossing edges, and the edges have a unique ordering by the increasing order of their lengths (when viewed in terms of the natural linear order on $[n]$).

Convex geometric z-trees. A *cg z-tree* is a cg tree Z with cyclic chromatic number two, obtained from a z-tree with ordered vertex set $[n]$ and intervals $I, J \subset [n]$ by viewing the linear ordering of the vertices as a cyclic ordering $n < n-1 < \cdots < 2 < 1$. Note that Z is a union of a cg increasing tree T with longest edge ij where $i \in I$ and $j \in J$, together with a set S_j of edges of the form hj with $h \in I$ and $i < h$ and a set S_i of edges of the form ik with $k \in J$ and $k < j$. These sets S_i and S_j are allowed to be empty. Note that any two of the edges hj and ik cross (See Figure 3 viewed as a cyclic ordering).

Our main structural result is the following.

Theorem 3.3. *Let T be a cg tree with $\chi_c(T) = 2$ that contains no crossing P_4 and no member of $\mathcal{P}$. Then T is a cg z-tree.*

Proof. Suppose we have an embedding of T in a circle γ. We will simultaneously refer to T as well as to the geometric properties of its embedding. Say that an edge e is *heavy* if both its endpoints have a neighbor on the same side of e. Since we have no L (see Definition 2), this means that every heavy edge e gives rise to a crossing three-edge path with central edge e.

Case 1. There is a heavy edge $e = i'j$. Suppose j' is a neighbor of i' and i is a neighbor of j such that both i and j' are on the same side of e, assume by symmetry that j' and i lie on the arc (j, i') of γ taken clockwise. Then $i < i' < j < j'$ otherwise we obtain L. This shows that all such neighbors of j lie clockwise of all such neighbors of i' in (j, i') and we obtain a double star as shown below. Moreover, each of these neighbors j', i has degree one otherwise we obtain a crossing P_4.

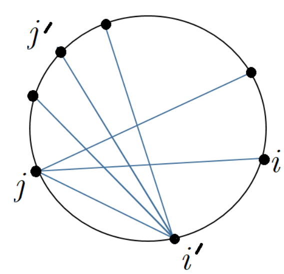

Figure 4.

Hence, to grow T further, we must consider neighbors of i' or j on the arc (i', j) of γ which omits j' and i (see Figure 4). We now claim that on the arc (i', j), the tree T is an increasing tree and moreover, there is no edge that crosses $e = i'j$. This will complete the proof in this case. First observe that i' and j cannot both have neighbors in (i', j) otherwise we get L as before or P^2 as shown below.

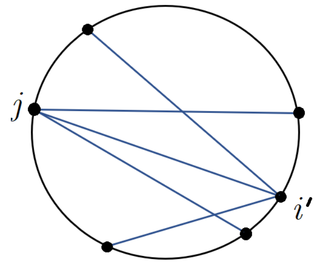

Figure 5.

So we may assume by symmetry that only j has neighbors in (i', j). If $j_1, j_2, \ldots, j_r$ are the neighbors of j in (i', j) in increasing clockwise order, then only j_r can have degree at least two, otherwise we get a copy of L or P^1; furthermore, all neighbors of j_r are in the segment (j_r, j); otherwise we have L or a crossing P_4. We now continue in this way: if j_r has neighbors $k_1, k_2, \ldots, k_m, j$ in increasing clockwise order in (j_r, j), then only k_1 can have degree at least two, otherwise we have a copy of L, or P^0, or a crossing P_4. Furthermore, all neighbors of k_1 are in the segment (j_r, k_1) otherwise we obtain L or a crossing P_4. We continue this process till we exhaust all of T.

Case 2. There is no heavy edge. In this case we claim the stronger statement that T is cg increasing tree. Start by choosing any edge $e = i'j$ and consider the neighbors of i' or of j in the arc (i', j) (traversed clockwise as usual). Since e is not heavy, at most one of i', j has neighbors in (i', j), say j. Then, using the fact that there is no heavy edge, we proceed as in the previous paragraph until we have exhausted all vertices of T in (i', j). Then we repeat this argument in the arc (j, i') to show that T is a cg increasing tree. □

We are now in a position to complete the proof of Theorem 1.4

Proof of Theorem 1.4. By Theorems 3.2 and 3.3, it suffices to show that every cg z-tree Z with $k \geq 2$ edges satisfies $\mathrm{ex}_{\circlearrowright}(n, Z) < (k-1)n$. Fix such a z-tree Z. By the definition of a cg z-tree, there is a linear order on the vertices of Z compatible with the cyclic ordering of the vertices of Z such that Z is an ordered z-tree $\vec{Z}$ in this order. Let G be an n-vertex cgg on $V(G) \subset \gamma$ not containing Z with $|G| = \mathrm{ex}_{\circlearrowright}(n, Z)$. View $V(G)$ as a linearly ordered set and $\vec{G}$ as the corresponding ordered graph (e.g., if the clockwise ordering of $V(G)$ is $n > n-1 > \cdots > 1$, then the linear ordering of $V(\vec{G})$ is $1 < 2 < \cdots < n$). Then $\vec{G}$ does not contain $\vec{Z}$, since G does not contain Z. So, by Theorem 1.3, $|G| = |\vec{G}| \leq (k-1)n - \binom{k}{2}$. □

Acknowledgement. This research was partly conducted during AIM SQuaRes (Structured Quartet Research Ensembles) workshops, and we gratefully acknowledge the support of AIM. We also thank the referees for their comments.

References

[1] D. Bienstock, E. Győri, An extremal problem on sparse 0-1 matrices. SIAM J. Discrete Math. 4 (1991), 17–27.

[2] P. Brass, Gy. Károlyi, P. Valtr, A Turán-type extremal theory of convex geometric graphs, in: B. Aronov et al. (Eds.), Discrete and Computational Geometry – The Goodman–Pollack Festscrift, Springer, Berlin, 2003, 277–302.

[3] Z. Füredi, P. Hajnal, Davenport-Schinzel theory of matrices. Discrete Math. 103 (1992), 233–251.

[4] M. Klazar, The Füredi-Hajnal conjecture implies the Stanley–Wilf conjecture, in: D.Krob, A.A.Mikhalev, A.V. Mikhalev (Eds.), Formal Power Series and Algebraic Combinatorics, Springer, Berlin, 2000, 250–255.

[5] D. Korándi, G. Tardos, I. Tomon, C. Weidert, On the Turán number of ordered forests https://arxiv.org/abs/1711.07723.

[6] Y.S. Kupitz and M.A. Perles, Extremal theory for convex matchings in convex geometric graphs, Discr. Comput. Geom. 15 (1996), 195220.

[7] W. Moser and J. Pach, Recent developments in combinatorial geometry, in: New Trends in Discrete and Computational Geometry (J. Pach, ed.), Springer, New York, 1993, 281–302.

[8] A. Marcus, G. Tardos, Excluded permutation matrices and the Stanley–Wilf conjecture, J. Combin. Theory Ser. A 107 (2004) 153–160.

[9] J. Pach, G. Tardos, Forbidden patterns and unit distances. Computational geometry (SCG'05), 1–9, ACM, New York, 2005.

[10] S. Pettie, Degrees of Nonlinearity in Forbidden 0-1 Matrix Problems Discrete Mathematics 311 (2011), 2396–2410.

[11] G. Tardos, On 0-1 matrices and small excluded submatrices, Journal of Combinatorial Theory, Series A Volume 111 (2005), 266–288.

[12] G. Tardos, Construction of locally plane graphs with many edges, in: Thirty essays on geometric graph theory (J. Pach, ed.), Springer, New York, 2013, 541–562.

Zoltán Füredi
Alfréd Rényi Institute of Mathematics
Hungarian Academy of Sciences
Reáltanoda utca 13-15.
H-1053, Budapest, Hungary.
E-mail: zfuredi@gmail.com.

Alexandr Kostochka
University of Illinois at Urbana–Champaign
Urbana, IL 61801
and Sobolev Institute of Mathematics
Novosibirsk 630090, Russia.
E-mail: kostochk@math.uiuc.edu.

Dhruv Mubayi
Department of Mathematics, Statistics and Computer Science
University of Illinois at Chicago
Chicago, IL 60607.
E-mail: mubayi@uic.edu.

Jacques Verstraëte
Department of Mathematics
University of California at San Diego
9500 Gilman Drive, La Jolla, California 92093-0112, USA.
E-mail: jverstra@math.ucsd.edu.