INTERSECTING FAMILIES OF SPANNING TREES

PETER FRANKL, GLENN HURLBERT, FERDINAND IHRINGER, ANDREY KUPAVSKII*, NATHAN LINDZEY, KAREN MEAGHER**, AND VENKATA RAGHU TEJ PANTANGI

Abstract. A family $\mathcal{F}$ of spanning trees of the complete graph on n vertices K_n is t-*intersecting* if any two members have a forest on t edges in common. We prove an Erdős–Ko–Rado result for t-intersecting families of spanning trees of K_n. In particular, we show there exists a constant $C > 0$ such that for all $n \geq C(\log n)t$ the largest t-intersecting families are the families consisting of all trees that contain a fixed set of t disjoint edges (as well as the stars on n vertices for $t = 1$). The proof uses the spread approximation technique in conjunction with the Lopsided Lovász Local Lemma.

1. Introduction

A *spanning tree* T of a simple graph $G = (V, E)$ is a connected subgraph $T \subseteq G$ on $n-1$ edges that contains all of V. We recall a classic result of Nash-Williams that determines the number of edge-disjoint spanning trees of a graph.

Theorem 1.1 (Nash-Williams). *A graph $G = (V, E)$ has ℓ edge-disjoint spanning trees if and only if for every partition $V = V_1 \sqcup \cdots \sqcup V_k$ such that $V_i \neq \emptyset$, there are at least $\ell(k-1)$ cross edges of G, i.e., edges $uv \in E$ such that u and v do not belong to the same partition class.*

Let $\mathcal{T}(G)$ be the collection of all spanning trees of a graph G. We define *the spanning tree disjointness graph* $\Gamma(G)$ on $\mathcal{T}(G)$ such that two vertices $T, T' \in \mathcal{T}(G)$ are adjacent if they have no edges in common.

Let K_n be the complete graph on n vertices. Its spanning tree disjointness graph $\Gamma(K_n)$ has n^{n-2} vertices by Cayley's theorem, and Nash-Williams' theorem implies that there are $\ell = \lfloor n/2 \rfloor$ edge-disjoint spanning trees of K_n (which is best possible). In other words, its *clique*

Key words and phrases. Extremal Graph Theory, Erdős–Ko–Rado Theorems, Trees, Probabilistic Combinatorics, Lovász Local Lemma.

*Research was supported by the Foundation for the Advancement of Theoretical Physics and Mathematics "BASIS".

**Research supported in part by an NSERC Discovery Research Grant, Application No.: RGPIN-2018-03952.

number $\omega(\Gamma(G))$ equals $\lfloor n/2 \rfloor$. Moreover, using Nash-Williams' theorem, one can determine $\omega(\Gamma(G))$ for arbitrary G in polynomial time. Closed formulas for the clique numbers of $\Gamma(G)$ have been computed for special graph classes (see [14], for example).

In this work, we consider the natural complementary problem of determining the *independence number* $\alpha(\Gamma(G))$ of the spanning tree disjointness graph. The independent sets of the spanning tree disjointness graph of G are families of spanning trees of G such that any two trees *intersect*, i.e., have an edge in common. Independent sets of this graph are called *intersecting families*. We will focus our efforts on $G = K_n$, leaving more general classes of graph as an open problem.

For many structured classes of graphs, there are well-known relations between their clique and independence numbers, e.g., *the clique-coclique bound* (see [8], for example).

Theorem 1.2. *If* Γ *is a vertex-transitive graph on* n *vertices, then*

$$\omega(\Gamma)\alpha(\Gamma) \leq n.$$

For even n, the theorem above would give us

$$n/2 \cdot \alpha(\Gamma(K_n)) \leq n^{n-2}.$$

It is a simple exercise to show that the size of a family $\mathcal{F}$ of spanning trees of K_n that all share a fixed edge is $2n^{n-3}$, which would seem to imply that $\mathcal{F}$ is a maximum independent set of $\Gamma(K_n)$.

Of course, the graph $\Gamma(K_n)$ is *not* vertex-transitive. Indeed, it is highly irregular: the orbits of the natural action of S_n on its vertices are the isomorphism classes of n-vertex trees. Moreover, one can show that $\mathcal{F}$ unioned with the set of all *n-stars* of K_n, i.e., spanning trees with one vertex of degree $n-1$, in fact produces a larger independent set of size $2n^{n-3} + n - 2$ (which turns out to be maximum).

The irregularity of the spanning tree disjointness graph is the source of much difficulty. Indeed, many other natural families of disjointness graphs for other combinatorial domains enjoy much more combinatorial regularity, making them amenable to elegant algebraic and combinatorial methods for computing their independence numbers (e.g., see [8]). This collection of results and techniques is sometimes referred to as *Erdős–Ko–Rado combinatorics*, which we briefly describe below.

1.1. **Erdős–Ko–Rado Combinatorics.**

Given a positive integer m, we set $[m] := \{1, 2, \ldots, m\}$. Given $1 \leq k \leq m$, let $\binom{[m]}{k}$ denote the set of k-subsets of $[m]$. We say that a family of subsets of $[m]$ is *t-intersecting* if any two of its elements intersect in at least t elements.

The Erdős–Ko–Rado (EKR) Theorem [4] is one of the most celebrated results in extremal combinatorics and states that, for $m \geq 2k$, a 1-intersecting family $\mathcal{F}$ in $\binom{[m]}{k}$ has size at most $\binom{m-1}{k-1}$ with equality for $m > 2k$ if and only if $\mathcal{F}$ consists of all k-sets that contain a fixed element of $[m]$. This result determines the size and structure of the largest intersecting families of k-subsets of $[m]$. Many generalizations of the EKR theorem have been proved for other classes of objects possessing a natural notion of intersection (see [6, 7], for example, or [2, 8] and the references therein). Before we state our main results, we discuss some previous work that is somewhat related to our tree intersection problem.

In 1978 Simonovits and Sós introduced the following problem [15, 16]: for a given family of graphs L, what is $f(n, L)$, the largest number of graphs on n vertices such that the intersection of every pair is in L. They show the following results:

- if $L = \{k\text{-vertex graphs}\}$, then $f(n, L) = o(n^{k+2})$;
- if $L = \{\text{complete graphs}\}$, then $f(n, L) \geq 2^{n-2}$;
- if $L = \{\text{stars}\}$, then $f(n, L) = 2^{n-1}$;
- if $L = \{\text{connected graphs}\}$, then $f(n, L) \geq 2^{\binom{n-1}{2}}$;
- if $L = \{\text{non-empty paths}\}$, then $f(n, L) = o(n^4)$; and
- if $L = \{\text{non-empty cycles}\}$, then $f(n, L) = \binom{n}{2} - 2$.

In 2012, Ellis, Filmus, and Friedgut [3] determined the largest family of triangle-intersecting (K_3-intersecting) family of graphs, resolving a conjecture by Simonovits and Sós from 1976.

Theorem 1.3 (Ellis, Filmus, and Friedgut [3]). *Let $\mathcal{F}$ be a triangle-intersecting family of graphs on n vertices. Then $|\mathcal{F}| \leq \frac{1}{8}2^{\binom{n}{2}}$. Equality holds if and only if $\mathcal{F}$ consists of all graphs containing a fixed triangle.*

1.2. **Main Results.** The main result of this work is an EKR theorem for n-vertex spanning trees (*trees*, for short), which in some sense lies in between the intersection problems mentioned above. Let $\mathcal{T}_n := \mathcal{T}(K_n)$ denote the set of labelled spanning trees of K_n, which we sometimes refer to as the *ambient family*. We say that $A, B \in \mathcal{T}_n$ are *t-intersecting* if they share at least t edges, and we often simply use *intersecting* in place of 1-intersecting. Recall that an n-star is a tree that is isomorphic to the complete bipartite graph $K_{1,n}$. Every tree is connected, so any star must intersect with any other tree. A set of trees is t-intersecting (or intersecting) if the trees in the set are pairwise t-intersecting (or intersecting).

We answer the following question for sufficiently large n with respect to t: *what is the size and structure of the largest t-intersecting families of $\mathcal{T}_n$?* This amounts to characterizing the maximum independent sets of the *spanning tree t-disjointness graph* $\Gamma_t(K_n)$ defined such that $T, T' \in \mathcal{T}_n$ are not adjacent if they have t or more edges in common. It also modifies the problem by Simonovits and Sós discussed above by replacing "number of graphs" with "number of spanning trees", and taking L to be the set of all forests with at least t edges.

We use the *spread approximation technique* developed in [11] to prove the following theorem.[1]

Theorem 1.4. *Suppose that $n \geq 2^{19}$ and $1 < t \leq \frac{n}{4032 \log_2 n}$. Let $\mathcal{F}$ be a t-intersecting family of trees on n vertices. Then*

$$|\mathcal{F}| \leq 2^t n^{n-t-2}.$$

Moreover, equality holds if and only if $\mathcal{F}$ consists of all trees containing a fixed set of t pairwise disjoint edges.

The stars cause a slight difference between the $t = 1$ case and the $t \geq 2$ case, so we state the result for the $t = 1$ case separately.

Theorem 1.5. *For $n \geq 2^{19}$, let $\mathcal{F}$ be a 1-intersecting family of trees on n vertices. Then*

$$|\mathcal{F}| \leq 2n^{n-3} + (n-2).$$

Moreover, equality holds if and only if $\mathcal{F}$ consists of all stars plus all trees containing one fixed edge.

Apart from the spread approximation technique, the main ingredient of the proof is a bound on the number of trees that contain a certain number of fixed edges, but avoid a given tree (see Section 3). For this, we use probabilistic techniques, most importantly, the *Lopsided Lovász Local Lemma* (LLLL) introduced by Erdős and Spencer [5].

2. Notation and Background

Let $2^{[m]}$ denote the power set of $[m]$. We start by identifying elements of $\mathcal{T}_n$ as sets. We can represent a tree $T \in \mathcal{T}_n$ with its edge set, since an edge is an element of $\binom{[n]}{2}$. That is, each tree can be identified with an element of $2^{\binom{[n]}{2}}$. In fact, we can identify every graph on n vertices as a subset of $\binom{[n]}{2}$. If the graph is a tree, then this subset will have size exactly $n-1$; if it is a forest with m components, it will have size exactly

[1] No attempt was made to optimize the constant 4032 or the constant 2^{19} in the hypotheses of the theorem below.

$n-m$. Conversely, any element in $F \in 2^{\binom{[n]}{2}}$ can be viewed as the set of edges in a graph on n vertices. If this graph is a forest, we say F *spans a forest* and this forest is *spanned by* F. The size of a forest, denoted by $|F|$, is the number of edges in the forest, so this representation preserves the size of a forest. The definition of t-intersecting trees is also consistent with this identification, trees are t-intersecting if and only if their corresponding (edge) sets are t-intersecting.

Cayley's theorem is a famous result that proves there are exactly n^{n-2} labelled trees on n vertices. We will use a generalization of this from [13, Lemma 6] that counts the number of labelled trees that contain a specific subforest.

Lemma 2.1 (Lu, Mohr, Székely)**.** *Let F be a forest in the complete graph on n vertices that has m components of sizes $q_1, q_2, \dots, q_m$. Then the number of spanning trees in K_n, that contain F is*

$$q_1 q_2 \cdots q_m \; n^{n-2-\sum_{i=1}^{m}(q_i-1)}.$$

A subset of $\mathcal{T}_n$ in which all trees contain a common edge is clearly intersecting. Similarly, for any forest F with t edges the set of all trees in $\mathcal{T}_n$ that contain F is t-intersecting; we called such sets *trivially t-intersecting*. We will define notation that will be used to describe the set of all trees that contain a common forest.

Given $\mathcal{F} \subset 2^{[m]}$, $\mathcal{S} \subset 2^{[m]}$, and $X \subset [m]$, we define

$$\begin{aligned} \mathcal{F}(X) &:= \{A \setminus X : \; A \in \mathcal{F},\, X \subset A\}, \\ \mathcal{F}[X] &:= \{A : \; A \in \mathcal{F},\, X \subset A\}, \\ \mathcal{F}[\mathcal{S}] &:= \bigcup_{X \in \mathcal{S}} \mathcal{F}[X]. \end{aligned}$$

If $F \in 2^{\binom{[n]}{2}}$ spans a forest, then the set $\mathcal{T}_n[F]$ is the set of all trees in $\mathcal{T}_n$ whose edges include the elements in F; this set is trivially $|F|$-intersecting. The size of $\mathcal{T}_n[F]$ can be easily found using Lemma 2.1. We note one particular case that we consider later.

Lemma 2.2. *If F is forest with exactly ℓ disjoint edges then*

$$|\mathcal{T}_n[F]| = 2^{\ell} n^{n-2-\ell}.$$

Moreover, if F' is any other forest with ℓ edges then $|\mathcal{T}_n[F']| \leq |\mathcal{T}_n[F]|$.

Proof. The size of $\mathcal{T}_n[F]$ clearly follows from Lemma 2.1.

Let F' be any other forest with ℓ edges and components of sizes $p_1, \dots, p_k$. Set $p_{k+1} = 1, \dots, p_\ell = 1$, this simply adds isolated vertices to F' and does not change the number of trees that contain the forest.

Then $\sum_{i=1}^{\ell} p_i = 2\ell$, and using the AM/GM inequality, the maximum value of $p_1 \cdots p_\ell$ is achieved with ℓ of the values equal to 2. $\square$

We also give a simple lower bound on the number of trees that contain a fixed forest.

Corollary 2.3. *Let F be a forest with t edges in the complete graph on n vertices, then the number of spanning trees in K_n, that contain F is at least n^{n-t-2}.*

Proof. Let $q_1, q_2, \ldots, q_m$ be the number of vertices in each of the components in F. From Lemma 2.1, along with the fact that each $q_1 \geq 2$, and $\sum_{i=1}^{m} q_i = t + m$, it follows that

$$|\mathcal{T}_n[F]| = q_1 q_2 \cdots q_m \; n^{n-2-\sum_{i=1}^{m}(q_i-1)} \geq n^{n-2-t}. \qquad \square$$

If two trees do not intersect, then we say these trees *avoid* each other; this happens if the two trees have no edges in common. Recall that a *star* is a tree that is isomorphic to $K_{1,n}$. Let $\mathfrak{S} \subset \mathcal{T}_n$ be the set of all stars in $\mathcal{T}_n$. Note that $|\mathfrak{S}| = n$. Since any tree $T \in \mathcal{T}_n$ is connected, no T avoids any of the stars in $\mathfrak{S}$.

For a given tree $T_0 \notin \mathfrak{S}$, we will consider the set of trees that contain some fixed forest $F \not\subseteq T_0$ and avoid T_0, outside of F. We denote this set as

$$\mathcal{T}_n[T_0; F] := \{T \in \mathcal{T}_n[F] \;:\; (T \cap T_0) \setminus F = \emptyset\}.$$

We are interested in the following minimization over all forests F of size t:

$$\mathfrak{D}_t := \min_{\substack{F \text{ forest} \\ |F|=t}} \quad \min_{\substack{T_0 \in \mathcal{T}_n \setminus \mathfrak{S} \\ |T_0 \cap F| < t}} \quad |\mathcal{T}_n[T_0; F]|.$$

The value of $\mathfrak{D}_t$ is important for EKR theorems. If $\mathcal{F}$ is a t-intersecting set of trees such that $T_0 \in \mathcal{F}$, but T_0 does not contain the forest F, then at least $\mathfrak{D}_t$ of the elements from $\mathcal{T}_n[F]$ do not belong to $\mathcal{F}$, as they do not t-intersect with T_0.

3. A Lower Bound on $\mathfrak{D}_t$

The main result of this section is the following lower bound on $\mathfrak{D}_t$.

Proposition 3.1. *If $n \geq 55 + t$, then $\mathfrak{D}_t > n^{n-t-27}$.*

We will apply the Lopsided Lovász Local Lemma (Lemma 3.2) to prove Proposition 3.1, so first we need to set our notation.

Let $A_1, A_2, \ldots, A_N$ be events. A *negative dependency graph* is a simple graph $G = ([N], E)$ such that

$$\Pr\left[\bigwedge_{j\in S} \bar{A}_j\right] \neq 0 \Rightarrow \Pr\left[A_i \,\middle|\, \bigwedge_{j\in S} \bar{A}_j\right] \leq \Pr[A_i]$$

for all $i \in [N]$ and $S \subseteq \{j \in [N] : \{i,j\} \notin E\}$.

Lemma 3.2 (Lopsided Lovász Local Lemma [5] (LLLL))**.** *Let $A_1, A_2, \ldots, A_N$ be events with negative dependency graph $G = ([N], E)$. If $x_1, x_2, \ldots, x_N \in [0,1)$ and*

$$\Pr[A_i] \leq x_i \prod_{ij\in E} (1 - x_j) \quad \textit{for all } i \in [N],$$

then

$$\Pr\left[\bigwedge_{i=1}^{N} \bar{A}_i\right] \geq \prod_{i=1}^{N} (1 - x_i) > 0.$$

We will use the following result of Lu, Mohr, and Székely [13].

Theorem 3.3. [13, Theorem 4] *Let $\mathcal{F}$ be a family of forests of K_n. The graph of events $\{A_F : F \in \mathcal{F}\}$, where $A_F \sim A_{F'}$ if there exist connected components of C of F and C' of F' that are neither identical nor disjoint, is a negative dependency graph.*

Let F be a forest with t edges in K_n, and e be an edge of K_n with $e \notin F$. Define A_e to be the event that the edge e is in a spanning tree of $\mathcal{T}_n[F]$ drawn uniformly at random. Recall that the *line graph* $\mathcal{L}(G)$ of a graph $G = (V, E)$ is the graph with vertex set E where $e, e' \in E$ are adjacent if they share an endpoint. Theorem 3.3 implies the following.

Corollary 3.4. *Let $T \in \mathcal{T}_n$ and F be a forest in K_n. The line graph $\mathcal{L}(T \setminus F)$ is a negative dependency graph for the events $\{A_e\}_{e\in T\setminus F}$.*

To apply the LLLL, we need to first establish bounds on $\Pr[A_e]$.

Lemma 3.5. *Let F be a forest in K_n, e be an edge in $K_n \setminus F$, and A_e be the event that e is in a spanning tree of $\mathcal{T}_n[F]$ drawn uniformly at random. Then*

$$\Pr[A_e] \leq \frac{2}{n}.$$

Proof. If $F \cup \{e\}$ is not a forest, then $\Pr[A_e] = 0$, so we may assume that $F \cup \{e\}$ is a forest.

By Lemma 2.1, if F has components of sizes $q_1, q_2, \ldots, q_m$, then the number of spanning trees containing F equals $q_1 q_2, \ldots, q_m\, n^{n-2-t}$. Similarly, if $r_1, r_2, \ldots, r_{m'}$ are the sizes of the components of the forest

$F \cup \{e\}$, then the number of spanning trees containing the forest $F \cup \{e\}$ equals $r_1 r_2 \cdots r_{m'}\ n^{n-2-(t+1)}$. Thus the probability that e is an edge of a tree in $\mathcal{T}_n[F]$ drawn uniformly at random is

$$\Pr[A_e] = \frac{r_1 r_2, \ldots, r_{m'}\ n^{n-2-(t+1)}}{q_1 q_2, \ldots, q_m\ n^{n-2-t}}.$$

There are three cases to consider for the value of m', since e can meet 0, 1 or 2 components of F.

(1) If e meets no component of F, then $m' = m+1$. Without loss of generality $q_i = r_i$ for $i \in [m]$ and $r_{m+1} = 2$. Then

$$\Pr[A_e] = \frac{2}{n}.$$

(2) If e meets one component of F, then $m' = m$. Without loss of generality $q_i = r_i$ for $i \in [m-1]$ and $r_m = q_m + 1$. Then, as $q_m \geq 2$,

$$\Pr[A_e] = \frac{q_m + 1}{q_m n} \leq \frac{3}{2n}.$$

(3) If e meets two components of F, then $m' = m-1$. Without loss of generality $q_i = r_i$ for $i \in [m-2]$ and $r_{m-1} = q_{m-1} + q_m$. Then, as $q_m \geq 2$,

$$\Pr[A_e] \leq \frac{q_{m-1} + q_m}{q_{m-1} q_m\, n} \leq \frac{1}{n},$$

where the last inequality follows from the fact that $a + b \leq ab$ for any pair of integers $a, b \geq 2$. □

We say that a forest T is *d-star-like* if it has an edge that is incident to at least $(n-1)/d$ other edges of T. If a forest is d-star-like, then its line graph will have a vertex with degree at least $(n-1)/d$.

The following lemma shows that if T_0 is not *d-star-like*, then for any forest F such that $|T_0 \cap F| \leq t-1$, the value of $\mathcal{T}_n[T_0; F]$ must be large.

Lemma 3.6. *Let $n \geq 5$. Let F be a forest with t edges in K_n. Suppose that T_0 is also a forest of K_n that is not 6-star-like. Then at least $e^{-4} n^{n-t-2}$ of the trees $T \in \mathcal{T}_n$ include F and avoid $T_0 \setminus F$, so in particular*

$$|\mathcal{T}_n[T_0; F]| \geq e^{-4} n^{n-t-2}.$$

Proof. Let $\{e_1, e_2, \ldots, e_m\}$ be the edge set of $T_0 \setminus F$; this means $m \leq n-1$. Let A_{e_i} be the event that e_i is in a spanning tree from $\mathcal{T}_n[F]$ drawn uniformly at random. First, we show that the hypotheses of Lemma 3.2 hold for the events $A_{e_1}, \ldots, A_{e_m}$.

Define

$$x_{e_1} = x_{e_2} = \cdots = x_{e_m} = 4/n,$$

and recall from Lemma 3.5 that $\Pr[A_e] \le \frac{2}{n}$. From Corollary, 3.4, the line graph $\mathcal{L}(T_0 \setminus F)$ is the negative dependency graph for $A_{e_1}, \ldots, A_{e_m}$. If we set $E = E(\mathcal{L}(T_0 \setminus F))$, then we must show for each $e_i \in T_0 \setminus F$ that

$$\Pr[A_{e_i}] \le x_{e_i} \prod_{\{e_i, e_j\} \in E} (1 - x_{e_j}). \tag{2}$$

The assumption that T_0 is not 6-star-like guarantees that the maximum degree of $\mathcal{L}(T_0 \setminus F)$ is less than $n/6$, thus

$$\frac{4}{n}\left(1 - \frac{4}{n}\right)^{n/6} \le x_{e_i} \prod_{\{e_i, e_j\} \in E} (1 - x_{e_j}).$$

But, as $\Pr[A_{e_i}] \le \frac{2}{n}$, it is sufficient to show that

$$\frac{2}{n} \le \frac{4}{n}\left(1 - \frac{4}{n}\right)^{n/6}.$$

Multiplying by n gives

$$2 \le 4\left(1 - \frac{4}{n}\right)^{n/6}.$$

One can check that this inequality holds for $n = 5$. Moreover, $\left(1 - \frac{4}{x}\right)^x$ is an increasing function of x for $x \ge 4$, and so the displayed inequality holds, provided that $n \ge 5$. Thus, Equation (2) holds for all $n \ge 5$.

Applying LLLL, we deduce that the probability that a spanning tree in $\mathcal{T}_n[F]$, drawn uniformly at random, contains none of the edges $e_1, \ldots, e_m$ is

$$\Pr\left[\bigwedge_{i=1}^{m} \bar{A}_{e_i}\right] \ge \prod_{i=1}^{m} (1 - x_i) = \left(1 - \frac{4}{n}\right)^m \ge e^{-4}.$$

Corollary 2.3 now shows the number of trees that contain F and avoid T_0 outside of F is

$$|\mathcal{T}_n[T_0; F]| \ge e^{-4}\, \mathcal{T}_n[F] \ge e^{-4} n^{n-t-2}. \qquad \square$$

Before we begin the proof of Proposition 3.1, we first establish some basic graph-theoretical notation. Let T be a forest. If v is a vertex of T, then the operation $T \setminus \{v\}$ removes the vertex v from T and removes all edges of T that are incident to v. If E is a set of edges, then the operation $T \setminus E$ removes from T the edges in E that are contained in

T. If F is a forest, then the operation $T \setminus F$ removes the edge set of F from T.

Proof of Proposition 3.1. We show for any tree T_0 and forest F_0 with t edges, that

$$\mathcal{T}_n[T_0; F_0] \geq n^{n-t-27}.$$

If T_0 is not 6-star-like, then the proof is immediate from Lemma 3.6, so we may assume that T_0 is 6-star-like.

Denote the vertex in T_0 of largest degree by v_0. Since T_0 is 6-star-like, we have $\deg(v_0) \geq (n-1)/12$.

Consider the neighborhood of v_0 in F_0. If F_0 has at least one edge that includes v_0, then we set $F_0' = F_0$. If F_0 has no edge incident to v_0, then we select a u_0 so that $F_0' = F_0 \cup \{v_0, u_0\}$ is also a forest. Because v_0 is not in F_0, any vertex u_0 will do, and since T_0 is not a star, we pick u_0 so that the edge $\{v_0, u_0\}$ is not in T_0.

Set $E_0 = \{\{v_0, x_1\}, \{v_0, x_2\}, \ldots, \{v_0, x_m\}\}$ to be all the edges of F_0' that contain v_0. Note that $E_0 \neq \emptyset$.

We count the number of trees T that contain F_0', avoid $T_0 \setminus F_0'$, and have the additional property that the neighborhood of v_0 in T is exactly the neighborhood of v_0 in F_0'. By our choice of u_0, every such tree T must also avoid $T_0 \setminus F_0$ and contain F_0; therefore, this count gives a lower bound on $|\mathcal{T}_n[T_0; F_0]|$. Moreover, since in every such tree T the neighborhood of v_0 is fixed, we may replace T_0 with $T_0 \setminus \{v_0\}$ so that this set of trees is precisely those trees $T \in \mathcal{T}_n[T_0 \setminus \{v_0\}; F_0']$ in which every neighbor of v_0 in T_0 is also a neighbor of v_0 in F_0'. We denote this set by $\mathcal{T}_n^*[T_0 \setminus \{v_0\}; F_0']$. We will obtain a lower bound on $|\mathcal{T}_n^*[T_0 \setminus \{v_0\}; F_0']|$ by considering another related set of trees.

Recall that E_0 is the set of all edges in F_0' that are incident to v_0. We define the path

$$E_0' = \{\{x_1, x_2\}, \{x_2, x_3\}, \ldots, \{x_{m-1}, x_m\}\};$$

if $|E_0| = 1$, then E_0' is the empty path.

No edge of the form $\{x_i, x_j\}$ belongs to F_0'; otherwise, the vertices $\{v_0, x_i, x_j\}$ would form a cycle in F_0'. We define $F_1 = (F_0' \setminus \{v_0\}) \cup E_0'$ so that F_1 is a forest in $K_n \setminus \{v_0\}$, and define $T_1 = (T_0 \setminus \{v_0\}) \setminus E_0'$ so that T_1 is a forest. Note that $|F_1| \leq |F_0|$.

We claim that

$$|\mathcal{T}_{n-1}[T_1; F_1]| \leq |\mathcal{T}_n^*[T_0 \setminus \{v_0\}; F_0']|,$$

which in turn gives us a lower bound on $|\mathcal{T}_n[T_0; F_0]|$. To see this, observe that every tree $T' \in \mathcal{T}_{n-1}[T_1; F_1]$ can be assigned to a unique tree $(T' \setminus E_0') \cup E_0 = T \in \mathcal{T}_n^*[T_0 \setminus \{v_0\}; F_0']$.

Next, we need a lower bound on the size of $\mathcal{T}_{n-1}[T_1; F_1]$. If T_1 is not 6-star-like, then as before, we apply Lemma 3.6 with F_1. This gives

$$e^{-4}(n-1)^{(n-1)-t-2} \leq |\mathcal{T}_{n-1}[T_1; F_1]|,$$

which is a lower bound on $|\mathcal{T}_n[T_0; F_0]|$.

If T_1 is 6-star-like, then we repeat the procedure above starting with T_1 in place of T_0. Again, we pick a vertex v_1 in T_1 of largest degree, so that $\deg(v_1) \geq \frac{n-2}{12}$, and we repeat until T_i is not 6-star-like.

We claim that we repeat this process at most 12 times, i.e., $i \leq 12$. Indeed, after 12 iterations we obtain the forest

$$T_{12} = (T_0 \setminus \{v_0, v_1, \ldots, v_{11}\}) \setminus (E_0' \cup E_1' \cup \cdots \cup E_{11}'),$$

which cannot be 6-star-like since it would only have at most

$$(n-1) - \frac{1}{12}\left((n-1) + \cdots + (n-11)\right) < \frac{n-13}{6}$$

edges from T_0 left, where in the last inequality we use $n \geq 56$.

Applying Lemma 3.6 to K_{n-12} in place of K_n, F_{12} in place of F_0, and T_{12} in place of T_0 gives the worst case bound of

$$\mathcal{T}_{n-12}[T_{12}; F_{12}] \geq e^{-4}(n-12)^{n-12-t-2} \geq e^{-4}n^{n-t-26} \geq n^{n-t-27}$$

for $n \geq 55$. To see this, for the second inequality, it suffices to show that

$$(n-t-14)\log(n-12) \geq (n-t-26)\log n.$$

Since $t \geq 1$, it suffices to show that $\log(n-12) \geq (1 - \frac{12}{n-14})\log n$ for all $n \geq t+14$, which follows from basic calculus. The third inequality holds since we have $n \geq 55 > e^4$, completing the proof. $\square$

4. Trees have Spread Properties

In this section, we provide a bound on the spreadness of the family of all spanning trees. Spreadness is a very handy definition that captures the amount of quasi-randomness of families of sets and is the basis for constructing spread approximations of families.

Definition 4.1 (r-spread)**.** *Given $r > 1$, a family $\mathcal{F} \subset 2^{[m]}$ is* r-spread *if*

$$|\mathcal{F}(X)| \leq r^{-|X|}\,|\mathcal{F}|,$$

for all $X \subset [m]$.

Definition 4.2 ((r,t)-spread)**.** *A family $\mathcal{F} \subset 2^{[m]}$ is said to be* (r,t)-spread, *if for each set $T \subset [m]$, with size at most t, the family $\mathcal{F}(T)$ is r-spread. That is, for any $T \subset [m]$ with $|T| \leq t$ and any set U with $T \subseteq U \subseteq [m]$*

$$|\mathcal{F}(U)| \leq r^{-(|U|-|T|)}\,|\mathcal{F}(T)|.$$

Lemma 4.3. *The family* $\mathcal{T}_n \subset 2^{\binom{[n]}{2}}$ *is* $(n/2, n-1)$*-spread.*

Proof. Let T be any subset from $\binom{[n]}{2}$ of size no more than $n-1$. If the graph corresponding to T is not a forest, then $\mathcal{T}_n(T) = \emptyset$. So trivially $\mathcal{T}_n(T)$ is $n/2$-spread.

Assume the graph corresponding to T is a forest with components of sizes $q_1, q_2, \dots, q_\ell$. Let $T \subset U \subset \binom{[n]}{2}$ be such that the graph corresponding to U is a forest and $|U| = |T|+1$. The forest corresponding to U is a forest formed by adding a single edge to the forest corresponding to T. Just as in the proof of Lemma 3.5, the components in U will have either sizes equal to : $q_1+q_2, q_3, \dots, q_\ell$; $q_1+1, q_2, \dots, q_\ell$; or $q_1, \dots, q_\ell, 2$. In any case, if the sizes of the components in U are $p_1, \dots, p_k$, then ratio of $q_1 \dots q_\ell$ to $p_1 \dots p_k$ is at least $1/2$. Thus, by Lemma 2.1

$$\frac{\mathcal{T}_n(T)}{\mathcal{T}_n(U)} = \frac{q_1 \dots q_\ell n^{n-2-\sum_{i=1}^{\ell} q_i - 1}}{p_1 \dots p_k n^{n-2-\sum_{i=1}^{\ell} p_i - 1}} = \frac{n^{n-2-|T|}}{2n^{n-2-|U|}} \geq \frac{n}{2}. \qquad \square$$

5. Theorems for Spread Approximations

We will follow the method of spread approximations, introduced in [11] and developed in subsequent papers [9, 10]. The first result allows us to get a low-uniformity approximation for a bulk of the family. Let $\binom{[m]}{\leq k}$ denote the set of subsets $[m]$ of size at most k.

Theorem 5.1. [9, Theorem 12] *Let integers* $m, k, t, q, r, r_0 \geq 1$ *satisfy* $r \geq 2q$ *and* $r_0 > r > 2^{12}\log_2(2k)$. *Suppose that* $\mathcal{A} \subset 2^{[m]}$ *is* r_0*-spread and that* $\mathcal{F} \subset \mathcal{A} \cap \binom{[m]}{\leq k}$ *is* t*-intersecting. Then there exists a* t*-intersecting family* $\mathcal{S} \subset \binom{[m]}{\leq q}$ *and an* $\mathcal{F}' \subset \mathcal{F}$ *such that*

(1) $|\mathcal{F}'| \leq (r_0/r)^{-q-1}|\mathcal{A}|$,
(2) $\mathcal{F} \setminus \mathcal{F}' \subset \mathcal{A}[\mathcal{S}]$, *and*
(3) $\mathcal{F}(B)$ *is* r*-spread for every* $B \in \mathcal{S}$.

We say that $\mathcal{A}$ is the *ambient family*, $\mathcal{S}$ is the *spread approximation* to $\mathcal{F}$, and $\mathcal{F}'$ is the *remainder.*

Next, we apply another general result from [9]. This result is used after the spread approximation is found, and gives us the desired conclusion that the structure of $\mathcal{S}$ for large t-intersecting families is trivial.

Theorem 5.2. [9, Theorem 14] *Let* $\epsilon \in (0, 1]$, *and* $m, r_0, q, t \geq 1$ *be such that* $\epsilon r_0 \geq 24q$. *Let* $\mathcal{A} \subset 2^{[m]}$ *be an* (r_0, t)*-spread family and let* $\mathcal{S} \subset \binom{m}{\leq q}$ *be a non-trivial* t*-intersecting family. Then there exists a* t*-element set* T *such that* $|\mathcal{A}[\mathcal{S}]| \leq \epsilon|\mathcal{A}[T]|$.

Remark. In the paper [9], it is required that $\mathcal{A}$ is *weakly* (r, t)-spread, but, as the name suggests, this is implied by its (r, t)-spreadness.

6. Proof of Theorem 1.4

Recall from Lemma 2.2 that the largest trivial t-intersecting family of trees is the family of all trees containing a fixed set of t disjoint edges provided that $t \leq n/2$. The size of this family is

$$2^t n^{n-t-2}.$$

Proof of Theorem 1.4. Let $\mathcal{F}$ be a t-intersecting family of $\mathcal{T}_n$. As $\mathcal{T}_n$ is $n/2$-spread by Lemma 4.3, we can apply Theorem 5.1, with the following parameters:

(1) $\mathcal{A} = \mathcal{T}_n$, the set of all labelled trees on n vertices,
(2) $m = \binom{n}{2}$, corresponding to the set of all possible edges,
(3) $q = 42t \log_2 n$,
(4) $r_0 = \dfrac{n}{2}$, $r = \dfrac{n}{4}$, and
(5) $k = n - 1$.

With these values we have $r > 2^{12} \log_2(2k)$ for $n \geq 2^{19}$. We also need that $r \geq 2q$, so

$$\frac{n}{4} \geq 2(42t \log_2 n),$$

which is implied by the assumed bound $t \leq \frac{n}{4032 \log_2 n}$.

The ambient family $\mathcal{T}_n$ with these parameters satisfies the conditions of Theorem 5.1. Applying this theorem yields that there exists a t-intersecting family $\mathcal{S}$ of sets of size at most $q = 42t \log_2 n$ and a remainder $\mathcal{F}' \subset \mathcal{F}$. By Theorem 5.1, the bound on the remainder is

$$\begin{aligned}|\mathcal{F}'| \leq \left(\frac{r_0}{r}\right)^{-q-1} |\mathcal{T}_n| &\leq 2^{-42t \log_2 n - 1} |\mathcal{T}_n| \\ &\leq n^{-42t} |\mathcal{T}_n| = n^{-42t} n^{n-2} = n^{n-42t-2}.\end{aligned}$$

Next we show that if the family $\mathcal{F} \setminus \mathcal{F}' \subset \mathcal{T}_n[\mathcal{S}]$ is close to extremal, then $\mathcal{S}$ is trivial. This is done by applying Theorem 5.2 to the family $\mathcal{T}_n[\mathcal{S}]$. By Lemma 4.3 $\mathcal{F}$ is $(n/2, t)$-spread.

In order to apply Theorem 5.2 we need that $\epsilon r_0 \geq 24q$. We will use $\epsilon = \frac{1}{2}$, and $q := 42t \log_2 n$, so we need that

$$n \geq 4032t \log_2 n.$$

This is satisfied since we assume that $t \leq \frac{n}{4032 \log_2 n}$.

By Theorem 5.2, if $\mathcal{S}$ is non-trivially t-intersecting, then

$$|\mathcal{F} \setminus \mathcal{F}'| \leq \epsilon |\mathcal{T}_n[F]|$$

for some t-element F (i.e., a forest on t edges). Since $|\mathcal{T}_n[F]| \geq n^{n-2-t}$ and $|\mathcal{F}'| \leq n^{n-42t-2}$, we have

$$|\mathcal{F}| = |\mathcal{F} \setminus \mathcal{F}'| + |\mathcal{F}'| \leq \left(\frac{1}{2} + n^{-41t}\right) |\mathcal{T}_n[F]| < |\mathcal{T}_n[F]|$$

in the case where $\mathcal{S}$ is non-trivial,

Thus, we may assume $\mathcal{S}$ is a trivial t-intersecting set, i.e., $\mathcal{S} = \{F\}$ where F is a forest on t edges. Without loss of generality, we may assume that $\mathcal{F}' \cap \mathcal{T}_n[F] = \emptyset$. If $\mathcal{F}'$ is empty, then $\mathcal{F}$ is trivial t-intersecting and $|\mathcal{F}| = 2^t n^{n-t-2}$, so we assume that $\mathcal{F}'$ is non-empty and pick $T_0 \in \mathcal{F}'$.

First, we assume that we can pick a $T_0 \in \mathcal{F}'$ that is not a star. Proposition 3.1 states that n^{n-t-27} is a lower bound on the number of trees on n vertices that contain F and avoid T_0, outside F. This means that the number of sets in $\mathcal{T}_n[T_0; F]$, that is, the number of trees that contain F and do not intersect T_0 outside $T_0 \cap F$, for any forest F of t edges is much larger than $|\mathcal{F}'|$. So

$$\begin{aligned} |\mathcal{F}| &\leq |\mathcal{T}_n(F)| - |\mathcal{T}_n[T_0; F]| + |\mathcal{F}'| \\ &\leq 2^t n^{n-t-2} - n^{n-t-27} + n^{n-42t-2} < 2^t n^{n-t-2}. \end{aligned}$$

Now assume that every $T_0 \in \mathcal{F}'$ is a star. If $t = 1$, there is nothing to do, as we have noted, a star intersects any tree and our claim is $\mathcal{F} = \mathcal{F} \cup \mathfrak{S}$. Thus, we assume $t \geq 2$. No two stars are t-intersecting, and thus $|\mathcal{F}'| = 1$. Let T_0 be the unique star in $\mathcal{F}'$, and denote its only non-leaf vertex by v. If v is adjacent to an edge in F, then, since $|F| > 1$, there are at least 2 vertices in F that are not v. In this case, we can construct at least two trees from $\mathcal{T}_n[F]$ that do not intersect T_0 outside F, by simply taking F and making every vertex not in F incident to any of the vertices in F that are not v (and perhaps adding edges between the components of F, without using the vertex v). If v is not adjacent to an edge of F, then any tree with F in which v is a leaf will not be t-intersecting with T_0.

Thus $|\mathcal{F} \setminus \mathcal{F}'| \leq |\mathcal{T}_n[F]| - 2$, and we have

$$|\mathcal{F} \setminus \mathcal{F}'| + |\mathcal{F}'| \leq |\mathcal{T}_n[F]| - 1,$$

as desired. This completes the proof of the main result. $\square$

7. Open Problems

In the next section we discuss our main conjectures for a complete result for t-intersecting trees.

7.1. **Conjectures for Larger t.** In this section, we discuss, the form we expect the largest t-intersecting family of trees to take when t is not bounded away from n. For any $t \leq n/2$ we conjecture that a trivial t-intersecting family of trees is still the largest possible.

Conjecture 7.1. *If $t \leq n/2$, then the largest t-intersecting family of trees is the family of all trees that contain a fixed set of t disjoint edges.*

If $t > n/2$ then it is not possible to have a set of t disjoint edges. We will consider an example of a t-intersecting set that is larger than the trivial t-intersecting one.

Example 1. Assume t is even and $\frac{3(t+2)}{2} \leq n < 2t$, $n \geq 8$. Let F be a forest that consists of $t/2$ disjoint copies of a path on 3 vertices. Note that the number of vertices is sufficient to host such a forest. By Lemma 2.1 the number of trees that contains F is

$$3^{t/2} n^{n-2-t}.$$

Since $\frac{3(t+2)}{2} \leq n$, there is also a forest consisting of $t/2+1$ disjoint copies of a path on 3 vertices. If F is the set of edges in this forest, then $|F| = t+2$. The collection of all trees that contain at least $t+1$ of these edges is a t-intersecting set. Let us count the number of such trees using a simple inclusion-exclusion argument.

The number of trees that contain all the $t+2$ edges in F is

$$3^{t/2+1} n^{n-2-(t+2)} = 3^{t/2+1} n^{n-t-4}.$$

Next we count the number of trees that contain at least $t+1$ of the edges in F. By Lemma 2.1, the number of trees that contain a forest with $(t+2)/2-1 = t/2$ components of size 3 and one more component of size 2 is

$$2(3^{t/2}) n^{n-2-(t+1))} = 2(3^{t/2}) n^{n-t-3};$$

this number includes the trees that contain all of F. There are $t+2$ ways to pick the component of size 2 and, adjusting for over-counting the tree that contain all of F, the number of such trees that contain at least $t+1$ of the edges in F is

$$\begin{aligned} &(t+2)(2)3^{t/2} n^{n-t-3} - (t+1)3^{t/2+1} n^{n-t-4} \\ &= 3^{t/2} n^{n-t-4} \left((t+2)2n - (t+1)3\right) \\ &= 3^{t/2} n^{n-t-4} \left(2nt + 4n - 3t - 3\right). \end{aligned}$$

The second construction is larger if

$$3^{t/2} n^{n-t-4} \left(2nt + 4n - 3t - 3\right) > 3^{t/2} n^{n-2-t},$$

which happens if and only if

$$n^2 - (4+2t)n + 3t + 3 < 0.$$

Since this parabola is negative for $n = 2$ and for $n = 2t$, it is negative for all $n \in [3(t+2)/2, \dots, 2t]$. So the second construction is always larger.

We end with a "Complete Intersection Theorem"-type conjecture. Define $F_{n,\ell}$ to be any spanning forest on n vertices with ℓ edges in which each component has either k or $k+1$ vertices (so the size of the components are as equal as possible). Then let $\mathcal{F}_{n,t,j}$ be the family of all trees that contain at least $t+j$ of the $t+2j$ edges of $F_{n,t+2j}$.

Conjecture 7.2. *For any t and n, there exists a j so that $\mathcal{F}_{n,t,j}$ is the largest t-intersecting set of trees.*

7.2. **Other Directions.** We believe that other ambient graph classes, beyond the complete graphs, should admit similar results. Our expectation is that symmetric graphs and sufficiently dense graphs would be good candidates, so we expect that complete multipartite graphs and quasi-random graphs will work. Indeed, Lemma 2.1 and other enumerative properties of spanning trees of K_n have multipartite analogues, see [12, 1], for example.

We can generalize the spanning tree t-disjointness graph $\Gamma_t(K_n)$ to $\Gamma_t(G)$, for any graph G: the graph in which the vertices are all spanning trees of G and two trees are adjacent if and only if they have fewer than t edges in common. This opens up a world of different problems. For example, for a fixed $t \geq 2$, is it NP-hard to compute the clique (or independence) number of $\omega(\Gamma_t(G))$ for arbitrary G? Can we find bounds or exact values of either $\omega(\Gamma_t(G))$ or $\alpha(\Gamma_t(G))$ for interesting families of graphs G? In fact, the clique number of $\omega(\Gamma_t(K_n))$ is open.

References

[1] Miklos Bona and Stephan Wagner. On the intersection of pairs of trees. *arXiv preprint arXiv:2309.00097*, 2025.

[2] David Ellis. Intersection problems in extremal combinatorics: theorems, techniques and questions old and new. In *Surveys in combinatorics 2022*, volume 481 of *London Math. Soc. Lecture Note Ser.*, pages 115–173. Cambridge Univ. Press, Cambridge, 2022.

[3] David Ellis, Yuval Filmus, and Ehud Friedgut. Triangle-intersecting families of graphs. *Journal of the European Mathematical Society*, 14(3):841–885, 2012.

[4] Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.

[5] Paul Erdős and Joel Spencer. Lopsided Lovász local lemma and latin transversals. *Discrete Applied Mathematics*, 30(2):151–154, 1991.

[6] Péter Frankl and Mikhail Deza. On the maximum number of permutations with given maximal or minimal distance. *J. Combinatorial Theory Ser. A*, 22(3):352–360, 1977.

[7] Péter Frankl and Richard M Wilson. The Erdős–Ko–Rado theorem for vector spaces. *J. Combin. Theory Ser. A*, 43(2):228–236, 1986.

[8] Chris Godsil and Karen Meagher. *Erdős–Ko–Rado Theorems: Algebraic Approaches*, volume 149 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2016.

[9] Andrey Kupavskii. Erdős–Ko–Rado type results for partitions via spread approximations. *arXiv preprint arXiv:2309.00097*, 2023.

[10] Andrey Kupavskii. Intersection theorems for uniform subfamilies of hereditary families. *arXiv preprint, arxiv:2311.02246*, 2023.

[11] Andrey Kupavskii and Dmitrii Zakharov. Spread approximations for forbidden intersections problems. *Advances in Mathematics*, 445:109653, 2024.

[12] Danyi Li, Wuxian Chen, and Weigen Yan. Enumeration of spanning trees of complete multipartite graphs containing a fixed spanning forest. *Journal of Graph Theory*, 104(1):160–170, 2023.

[13] Linyuan Lu, Austin Mohr, and László Székely. Quest for negative dependency graphs. In *Recent advances in harmonic analysis and applications*, volume 25 of *Springer Proc. Math. Stat.*, pages 243–258. Springer, New York, 2013.

[14] Edgar M. Palmer. On the spanning tree packing number of a graph: a survey. *Discrete Mathematics*, 230(1):13–21, 2001. Catlin.

[15] Miklós Simonovits and Vera T. Sós. Intersection theorems for graphs. In *Problemes combinatoires et théorie des graphes*, volume 260, pages 389–391. Colloq. Internat. CNRS, Paris, 1978.

[16] Miklós Simonovits and Vera T. Sós. Intersection theorems for graphs. II. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, pages 1017–1030. Colloq. Math. Soc. János Bolyai, 18 North-Holland Publishing Co., Amsterdam-New York, 1978.

Rényi Institute, Budapest, Hungary
Email address: peter.frankl@gmail.com

Department of Mathematics and Applied Mathematics, Virginia Commonwealth University, Richmond, VA, 23220 USA
Email address: ghurlbert@vcu.edu

Department of Mathematics, Southern University of Science and Technology, No. 1088 Xueyuan Blvd, Nanshan District, Shenzhen, 518055, Guangdong Province, P.R. China.
Email address, Corresponding author: Ferdinand.Ihringer@gmail.com

CombGeo lab, Moscow Institute of Physics and Technology, Dolgoprudniy, Russia
Email address: kupavskii@ya.ru

Department of Mathematical Sciences, University of Memphis, Memphis, TN, 38111 USA
Email address: nathan.lindzey@memphis.edu

Department of Mathematics and Statistics, University of Regina, Regina, SK, S4S 0A2, Canada
Email address: karen.meagher@uregina.ca

Email address: pvrt1990@gmail.com