TURÁN NUMBERS FOR FORESTS OF PATHS IN HYPERGRAPHS*

NEAL BUSHAW† AND NATHAN KETTLE‡

Abstract. The *Turán number* of an r-uniform hypergraph H is the maximum number of edges in any r-graph on n vertices which does not contain H as a subgraph. Let $\mathcal{P}_\ell^{(r)}$ denote the family of r-uniform loose paths on ℓ edges, $\mathcal{F}(k,l)$ denote the family of hypergraphs consisting of k disjoint paths from $\mathcal{P}_\ell^{(r)}$, and $L_\ell^{(r)}$ denote an r-uniform linear path on ℓ edges. We determine precisely $\mathrm{ex}_r(n;\mathcal{F}(k,l))$ and $\mathrm{ex}_r(n;k \cdot L_\ell^{(r)})$, as well as the Turán numbers for forests of paths of differing lengths (whether these paths are loose or linear) when n is appropriately large dependent on k,l,r for $r \geq 3$. Our results build on recent results of Füredi, Jiang, and Seiver, who determined the extremal numbers for individual paths, and provide more hypergraphs whose Turán numbers are exactly determined.

Key words. Turán, extremal, hypergraphs, paths, hyperpaths

AMS subject classifications. 05C35, 05C65, 05D05

DOI. 10.1137/130913833

1. Introduction and background. Extremal graph theory is that area of combinatorics which is concerned with finding the largest, smallest, or otherwise optimal structures with a given property. Often, the area is concerned with finding the largest (hyper)graph avoiding some subgraph. We build on earlier work of Füredi, Jiang, and Seiver [18], who determined the extremal numbers when the forbidden hypergraph is a single linear path or a single loose path. In this paper, we determine precisely the exact Turán numbers when the forbidden hypergraph is a forest of loose paths or a forest of linear paths; our main results appear in section 2. This is one of only a few papers which gives exact Turán numbers for an infinite family of hypergraphs, in this case, several such families.

The *Turán number*, or *extremal number*, of an r-uniform hypergraph F is the maximum number of edges in any r-graph H on n vertices which does not contain F as a subgraph. This is a natural generalization of the classical Turán number for 2-graphs; we restrict ourselves to the case of r-uniform hypergraphs, as allowing the extremal number to count edges of different sizes obscures the true extremal structure.

Throughout, we use standard terminology and notation (see, e.g., [7]). A *hypergraph* is a pair $H = (V, E)$ consisting of a set V of vertices and a set $E \subseteq \mathcal{P}(V)$ of edges. If $E \subseteq \binom{V}{r}$, then H is an *r-uniform hypergraph*; in this paper, we will restrict ourselves to this setting. If $|V| = n$, we will assume without loss of generality that $V = [n] = \{1, 2, \dots, n\}$. For two hypergraphs G and H, we write $G \subseteq H$ if there is an injective homomorphism from G into H. By disjoint, we will always mean *vertex disjoint*; we use $G \cup H$ to denote the disjoint union of (hyper)graphs G and H. Similarly, for $k \in \mathbb{N}$, we use $k \cdot G$ to denote k (vertex-)disjoint copies of G. We also make use of the indicator function; $\mathbb{1}_E = \left\{ \begin{smallmatrix} 1 \text{ if } E \text{ holds,} \\ 0 \text{ else.} \end{smallmatrix} \right.$

†School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287 (bushaw@asu.edu).

‡Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge CB3 0WB, UK (n.kettle@dpmms.cam.ac.uk).

Developing an understanding of the Turán numbers and extremal graphs for a general r-graph F is a classical and longstanding problem in extremal graph theory. The field began to take off in the 1940s, in the case of 2-graphs, when Pál Turán determined the extremal numbers for complete graphs of all orders; it is through this result that Turán's name became synonymous with the field.

While this problem is well solved up to asymptotics when the forbidden graph has chromatic number at least three by the Erdős–Stone Theorem [10], things remain much murkier for bipartite graphs. We discuss this in section 1.1; first, we formally define the Turán number for r-graphs as follows.

DEFINITION 1. *The r-uniform hypergraph Turán number, or extremal number, of a family $\mathcal{F}$ of r-uniform hypergraphs is defined as the following:*

$$\mathrm{ex}_r(n;\mathcal{F}) = \max\{|E(H)| : |V(H)| = n \ \forall F \in \mathcal{F}, F \not\subseteq H\}.$$

For a single hypergraph F, we will often write $\mathrm{ex}_r(n;F)$ for $\mathrm{ex}_r(n;\{F\})$. We modify this definition slightly for lists of hypergraphs. As opposed to the above definition for a family of hypergraphs, where any member of the family is forbidden, here we are forbidding disjoint copies of all graphs in the list from appearing simultaneously.

DEFINITION 2. *The r-uniform hypergraph Turán number of a list of r-uniform hypergraphs $F_1, F_2, \ldots, F_k$ is defined as the following:*

$$\mathrm{ex}_r(n;F_1, F_2, \ldots, F_k) = \mathrm{ex}_r(n;F_1 \cup F_2 \cup \cdots \cup F_k).$$

A hypergraph H is called *extremal for F* if H is F-free and $|E(H)| = \mathrm{ex}_r(n;F)$; we denote by $\mathrm{Ex}_r(n;F)$ the family of n vertex graphs which are extremal for F; similarly, H is extremal for the list $F_1, F_2, \ldots, F_k$ if it does not contain disjoint copies of all graphs in the list and $|E(H)| = \mathrm{ex}_r(n;F_1, F_2, \ldots, F_k)$.

1.1. Background for graphs. Before discussing paths in hypergraphs, on which this paper focuses, we briefly discuss the related results for paths in graphs of which the results in section 2 are generalizations. We'll make use of the standard notation $\mathrm{ex}(n,F) = \mathrm{ex}_2(n,F)$.

In 1959, Erdős and Gallai proved the following result giving the extremal numbers for paths of a given length [12]. We note that the bound in Theorem 1.1 is attained by taking disjoint copies of $K_{\ell-1}$. This extremal construction is unique as long as n is divisible by $\ell - 1$, and gives a tight bound in this case.

THEOREM 1.1 (Erdős and Gallai, 1959). *For any $n, \ell \in \mathbb{N}$,*

$$\mathrm{ex}(n;P_\ell) \leq \left(\frac{\ell-2}{2}\right)n.$$

We note that a path can be viewed as an extreme kind of tree; it is the tree with the largest diameter for its number of vertices. The opposite extreme is the star; here the diameter is only two. Forbidding the star S_ℓ with ℓ leaves is, in fact, simply imposing a maximum degree condition, and so $\mathrm{ex}(n;S_\ell) \leq (\frac{\ell-2}{2})n$. This bound is tight, with the extremal graphs being all $(\ell-2)$-regular graphs. For general trees, this result is notoriously difficult and is known as the Erdős–Sós conjecture [11].

CONJECTURE 1 (Erdős and Sós, 1963). *For any tree T on ℓ vertices, $\mathrm{ex}(n;T) \leq (\frac{\ell-2}{2})n$.*

In 2008, a proof of this conjecture was announced for very large trees by Ajtai, Komlós, Simonovits, and Szemerédi; this will appear in a series of upcoming papers [1, 2, 3]. A sketch of this result can be found in a recent survey of Füredi and

Simonovits [20]. For small trees, however, the conjecture is largely open. For a survey of other Turán results for connected bipartite graphs, see, e.g., [4, 6, 9, 8].

In [8], the present authors determined precisely the extremal numbers for forests where each component has the same number of vertices in each bipartite class (such a forest is called *equibipartite*), assuming the Erdős–Sós conjecture holds for those trees in the forest, as well as the extremal numbers for forests of paths of the same odd length. Our main results in this paper give the hypergraph versions of these theorems for forests of (linear or loose) paths. For comparison, these graph results are included below.

THEOREM 1.2 (B.-K, 2011). *For $n \geq 7k$, the following holds:*

$$\mathrm{ex}(n; k \cdot P_3) = \binom{k-1}{2} + (n-k+1)(k-1) + \left\lfloor \frac{n-k+1}{2} \right\rfloor.$$

THEOREM 1.3 (B.-K., 2011). *For $k \geq 2$, $\ell \geq 4$, and $n \geq 2\ell + 2k\ell \left(\lceil \frac{\ell}{2} \rceil + 1\right) \left(\binom{\ell}{\lfloor \frac{\ell}{2} \rfloor}\right)$, the following holds:*

$$\mathrm{ex}(n; k \cdot P_\ell) = \binom{k\left\lfloor \frac{\ell}{2} \right\rfloor - 1}{2} + \left(k \left\lfloor \frac{\ell}{2} \right\rfloor - 1\right) \left(n - k \left\lfloor \frac{\ell}{2} \right\rfloor + 1\right) + \mathbb{1}_{\{\ell \text{ is odd}\}}.$$

THEOREM 1.4 (B.-K., 2011). *Let F be an equibipartite forest on 2ℓ vertices which is comprised of at least two trees. If the Erdős–Sós conjecture holds for each component tree in F, then for $n \geq 3\ell^2 + 32\ell^5\binom{2\ell}{\ell}$,*

$$\mathrm{ex}(n; F) = \begin{cases} \binom{\ell-1}{2} + (\ell-1)(n-\ell+1) & \text{if } H \text{ admits a perfect matching,} \\ (\ell-1)(n-\ell+1) & \text{otherwise.} \end{cases}$$

With a generalization of these theorems in mind, we now proceed to a discussion of the hypergraph Turán problem.

1.2. Background for hypergraphs. In general, Turán theory for r-uniform hypergraphs with $r \geq 3$ is much less developed than the theory for 2-graphs. In the same paper in which Turán proved his fundamental theorem on the extremal numbers for complete graphs [26], he posed the natural question of determining $\mathrm{ex}_r(n; K_t^{(r)})$, where $K_t^{(r)}$ denotes the complete r-uniform graph on t vertices. Surprisingly, this problem remains open in all cases for $r > 2$, even up to asymptotics.

Determining extremal numbers precisely for hypergraphs is difficult indeed. Those results that do exist tend to be asymptotics, and exact results, with a few exceptions discussed below, are virtually always for small graphs on a few vertices. Such exact results exist (for large n) for the Fano plane, 4-books with 2, 3, or 4 pages, and a few other similarly small objects (see, e.g., [23] for a survey of hypergraph Turán results). As a sample of a typical exact theorem in this area, we state the Erdős–Ko-Rado theorem below; this is perhaps the classical extremal result for hypergraphs [13].

THEOREM 1.5 (Erdős and Ko-Rado, 1961). *If H is an r-uniform hypergraph on $n \geq 2r$ vertices in which every pair of edges intersects, then $|E(H)| \leq \binom{n-1}{r-1}$. That is, if we let $M_2^{(r)}$ denote the r-graph consisting of 2 disjoint edges, then for $n \geq 2r$,*

$$\mathrm{ex}_r(n; M_2^{(r)}) = \binom{n-1}{r-1}.$$

We now continue to the main objects of focus in this paper: paths in hypergraphs. One can think of a matching in hypergraphs as being a forest of paths of length one.
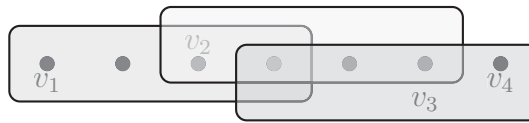
FIG. 1. *A 4-uniform Berge path on 3 edges.*

Thus perhaps a starting point for this history is the following conjecture of Erdős [14]. If one considers r-uniform hypergraphs forbidding an $s+1$-matching, there are two natural constructions. Either one can take $r(s+1) - 1$ vertices and all edges, or one can take a set of s vertices and all edges intersecting this set; thus the Erdős conjecture is as follows.

CONJECTURE 2 (Erdős, 1965). *Let $M_{s+1}^{(r)}$ denote $s+1$ disjoint r-edges, $A_r = \binom{[r(s+1)-1]}{r}$, and $B_r(n) = \{F \in \binom{[n]}{r} : F \cap [s] \neq \emptyset\}$. Then*

$$\mathrm{ex}_r(n; M_s^{(r)}) = \max\{|A_r|, |B_r(n)|\}.$$

In the same paper as his conjecture, Erdős proved this result for $n > n_0(r,s)$; it is this result to which the main theorems of this paper correspond. In fact, both Theorems 2.1 and 2.3, with $\ell = 1$, prove the above conjecture for large n. The bound on $n_0(r,s)$ has been gradually improved over the years: Bollobás, Daykin, and Erdős [5] proved $n_0(r,s) \leq 2r^3 s$; this was later improved by Huang, Loh, and Sudakov [21] to $n_0(r,s) \leq 3r^2 s$. In the case of $r = 3$, there has been significant recent progress: Frankl, Rödl, and Rucinski [16] showed $n_0(3,s) \leq 4s$, Łuczak and Mieczkowska [25] proved the conjecture for $r = 3$ and $s > s_0$, and finally Frankl [17] settled this case for all s and n.

We now move to paths of longer length in hypergraphs, noting first that there are several natural generalizations of paths in graphs to paths in hypergraphs. Thus we give three different definitions of paths in hypergraphs; we present these from most general to most specific.

DEFINITION 3. *A Berge path of length ℓ in a hypergraph H is a family of distinct edges $\{F_1, \ldots, F_\ell\} \subseteq E(H)$ along with a family of vertices $\{v_1, \ldots, v_{\ell+1}\} \subseteq V(H)$ such that for each $i \in \{1, \ldots, \ell\}$, $v_i, v_{i+1} \in F_i$.*

DEFINITION 4. *A loose path of length ℓ in a hypergraph H is a family of distinct edges $\{F_1, \ldots, F_\ell\} \subseteq E(H)$ such that $F_i \cap F_j \neq \emptyset$ iff $|i-j| = 1$. We use $\mathcal{P}_\ell^{(r)}$ to denote the family of r-uniform loose paths on ℓ edges.*

DEFINITION 5. *A linear path of length ℓ in a hypergraph H is a family of distinct edges $\{F_1, \ldots, F_\ell\} \subseteq E(H)$ such that $|F_i \cap F_j| = 1$ if $|i - j| = 1$, and $F_i \cap F_j = \emptyset$ otherwise. We use $L_\ell^{(r)}$ to denote an r-uniform linear path on ℓ edges.*

We provide examples of Berge, loose, and linear 4-paths in Figures 1, 2, and 3, respectively.

The following results were proven by Füredi, Jiang, and Seiver [18]. These results are quite significant, as mentioned in the introduction, as they provide exact Turán numbers for an infinite family of hypergraphs. In this sense, Theorems 1.6 and 1.7 are the first of their kind.

THEOREM 1.6 (Füredi, Jiang, and Seiver, 2011). *Let $r \geq 3$, $\ell \geq 1$. Letting $t = \lfloor \frac{\ell+1}{2} \rfloor - 1$, $c_\ell = \mathbb{1}_{\{\ell \text{ is even}\}}$, and n sufficiently large,*

$$\mathrm{ex}_r(n; \mathcal{P}_\ell^{(r)}) = \binom{n-1}{r-1} + \binom{n-2}{r-1} + \cdots + \binom{n-t}{r-1} + c_\ell.$$
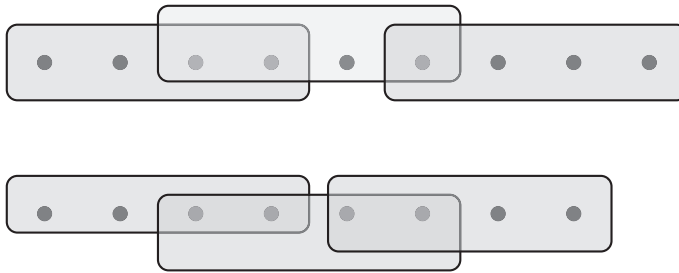
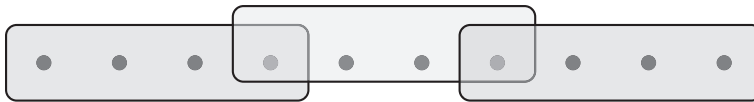FIG. 2. *Two 4-uniform loose paths, each on 3 edges.*

FIG. 3. *A 4-uniform linear path on 3 edges.*

The unique extremal family consists of all the r-sets in $[n]$ which meet some fixed set S of t vertices, plus one additional r-set disjoint from S when ℓ is even.

THEOREM 1.7 (Füredi, Jiang, and Seiver, 2011; Kostochka, Mubayi, and Verstraëte, 2013). *Let* $r \geq 3$, $\ell \geq 1$. *Letting* $t = \lfloor \frac{\ell+1}{2} \rfloor - 1$, $d_\ell = 0$ *if* ℓ *is odd and* $d_\ell = \binom{n-t-2}{r-2}$ *if* ℓ *is even, and* n *sufficiently large,*

$$\mathrm{ex}_r(n; L_\ell^{(r)}) = \binom{n-1}{r-1} + \binom{n-2}{r-1} + \cdots + \binom{n-t}{r-1} + d_\ell.$$

For ℓ odd, the unique extremal family consists of all the r-sets in $[n]$ which meet some fixed set S of t vertices. For ℓ even, we have these edges plus all the r-sets in $[n] \setminus S$ containing some two fixed elements not in S.

In Theorem 1.7, the case $r \geq 4$ was proved by Füredi, Jiang, and Seiver [18], and the case $r = 3$ by Kostochka, Mubayi, and Verstraëte [24].

In addition to these results for paths, we will need the following result due to Keevash, Mubayi, and Wilson (Theorem 1.3 in [22]).

THEOREM 1.8 (Keevash, Mubayi, and Wilson, 2006). *Let H be an r-uniform hypergraph on n vertices with no singleton intersection, where $r \geq 3$. Then*

$$|E(H)| \leq \binom{n}{r-2}.$$

This gives us an upper bound on the number of edges in a hypergraph where no two edges intersect at *exactly* one vertex. This will be useful to us in the case of linear paths, as the intersection of consecutive edges in these paths is precisely a single vertex.

2. Main results. As mentioned in the introduction, our main results give exact Turán numbers for forests of hyperpaths. We have several theorems here, dealing with forests of linear paths and loose paths, and with all paths having the same length or differing lengths. As every linear r-path on ℓ edges is isomorphic, the results for these types of paths avoid some notational difficulties associated with loose paths. Thus we state these results first, although our proofs will appear in the reverse order, as loose

paths avoid the difficulties associated with singleton intersections and thus avoid the use of Theorem 1.8 above.

THEOREM 2.1. *Let $r \geq 3$, $\ell \geq 1$, $k \geq 2$, and n sufficiently large. Then letting $t = k\lfloor \frac{\ell+1}{2} \rfloor - 1$, $d_\ell = 0$ if ℓ is odd, and $d_\ell = \binom{n-t-2}{r-2}$ if ℓ is even,*

$$\mathrm{ex}_r(n; k \cdot L_\ell^{(r)}) = \binom{n-1}{r-1} + \cdots + \binom{n-t}{r-1} + d_\ell.$$

THEOREM 2.2. *Let $r \geq 3$, $k \geq 2$, $\ell_1, \ldots, \ell_k \geq 1$, and n sufficiently large. Setting $t = \sum_{i \in [k]} \lfloor \frac{\ell_i+1}{2} \rfloor - 1$ and $d_l = 0$ when at least one of the ℓ_i is odd and $d_l = \binom{n-t-2}{r-2}$ otherwise, the following holds:*

$$\mathrm{ex}_r(n; L_{\ell_1}^{(r)}, \ldots, L_{\ell_k}^{(r)}) = \binom{n-1}{r-1} + \cdots + \binom{n-t}{r-1} + d_\ell.$$

We now continue to forests of loose paths. Here, we need some extra notation. Our results have the same flavor as the above; we give the maximum number of edges in a graph which does not have k vertex disjoint loose paths of given lengths. However, not all loose paths of a given length are isomorphic. Thus we define the following family of graphs, where each member consists of k disjoint paths:

$$\mathcal{F}(k,\ell) = \{P_1 \cup \cdots \cup P_k : P_i \in \mathcal{P}_\ell^{(r)} \text{ for each } i \in [k]\}.$$

We also define a similar but slightly more complicated family for paths of differing lengths:

$$\mathcal{F}'(\ell_1, \ell_2, \ldots, \ell_k) = \{P_1 \cup \cdots \cup P_k : P_i \in \mathcal{P}_{\ell_i}^{(r)} \text{ for each } i \in [k]\}.$$

With these definitions, we can state our remaining results as follows.

THEOREM 2.3. *Let $r \geq 3$, $\ell \geq 1$, $k \geq 1$, and n sufficiently large. Then letting $t = k\lfloor \frac{\ell+1}{2} \rfloor - 1$, $c_\ell = \mathbb{1}_{\{\ell \text{ is even}\}}$,*

$$\mathrm{ex}_r(n; \mathcal{F}(k,l)) = \binom{n-1}{r-1} + \cdots + \binom{n-t}{r-1} + c_\ell.$$

THEOREM 2.4. *Let $r \geq 3$, $k \geq 2$, $\ell_1, \ldots, \ell_k \geq 1$, and n sufficiently large. Letting $t = \sum_{i \in [k]} \lfloor \frac{\ell_i+1}{2} \rfloor - 1$ and $c_\ell = \mathbb{1}_{\{every \, \ell_i \, is \, even\}}$,*

$$\mathrm{ex}_r(n; \mathcal{F}'(\ell_1, \ldots \ell_k)) = \binom{n-1}{r-1} + \cdots + \binom{n-t}{r-1} + c_\ell.$$

The proofs for each of these results follows a standard form. Given a hypergraph with many edges (i.e., more than the claimed extremal number), we assert that any individual path of the appropriate type found in our hypergraph must have many edges incident to its vertex set; otherwise, the hypergraph induced by the vertices outside this path contains the rest of the specified forest by induction. We then examine sets of vertices which are incident to many edges. From here, we determine that there is a set of vertices which is large enough that we can construct our entire forest using these vertices as building blocks, in the case that a particular path is found in the rest of the graph.

This is enough to give us the structural result we need, and we can then simply count potential edges. This is complicated somewhat in the case of linear paths, as

we need to find edges which intersect in only a single vertex; nevertheless, our overall scheme is the same. In order to avoid redundancy, we shall prove only Theorems 2.1 and 2.4. The differing lengths in Theorem 2.2 can be dealt with precisely as in the proof of Theorem 2.4, and Theorems 1.6 and 2.4 together give Theorem 2.3.

3. Forests of loose paths. We note some ambiguity in the base case: instead of forbidding a particular loose path of length ℓ, the family of loose paths of length ℓ is forbidden. Thus, as in Theorem 1.6, we are not able to guarantee the existence of a *particular* loose path of a given length, only that *some* ℓ-edge loose path exists.

For ease of notation, we let $t = \sum_{i \in [k]} \lfloor \frac{\ell_i + 1}{2} \rfloor - 1$, $c_\ell = \mathbb{1}_{\{\text{every } \ell_i \text{ is even}\}}$, and define

$$h(n, r, \{\ell_1, \ldots, \ell_k\}) = \binom{n-1}{r-1} + \cdots + \binom{n-t}{r-1} + c_\ell.$$

We note that the hypergraph on n vertices that has every edge incident to a specified set S of t vertices, along with a single edge disjoint from S when all of the paths are even, has exactly $h(n, r, \{\ell_1, \ldots, \ell_k\})$ edges and does not contain a copy of any member of $\mathcal{F}'(\ell_1, \ell_2, \ldots, \ell_k)$.

We start with a lemma that will be of use in the proofs of each of our main theorems.

LEMMA 3.1. *Let c and t be positive constants and n large enough. Then each r-hypergraph H on n vertices with at least cn^r edges and a specified set of vertices $T \subset V(H)$ of size $|T| \leq t$ contains a pair of 1-intersecting edges that is vertex disjoint from T.*

Proof. The number of edges that meet T is at most $t\binom{n-1}{r-1}$, so there are at least $cn^r - t\binom{n-1}{r-1}$ edges in the hypergraph H restricted to vertex set $V(H) \setminus T$. As this is more than $\binom{n-t}{r-2}$ for n sufficiently large, we can find a pair of 1-intersecting edges by Theorem 1.8. □

Proof of Theorem 2.4. The case $k = 1$ is provided by Theorem 1.6. We proceed by induction on k; thus assume that $k \geq 2$ and that H is a hypergraph on n vertices with $|E(H)| = m > h(n, r, \{\ell_1, \ldots, \ell_k\})$. If any of the ℓ_i is even, we rearrange the list so that ℓ_1 is even for convenience.

Since $h(n, r, \{\ell_1, \ldots, \ell_k\}) > h(n, r, \ell_1)$, for n large enough, we can find at least one loose path on ℓ_1 vertices inside H. Consider one of these ℓ_1-paths, say, on vertex set P. Certainly $|E(V(H) \setminus P)| \leq h(n - |P|, r, \{\ell_2, \ldots, \ell_k\})$, or else by induction, the graph on $V(\mathcal{H}) \setminus P$ contains a member of $\mathcal{F}(\ell_2, \ldots, \ell_k)$; these alongside the loose path on P form a member of $\mathcal{F}(\ell_1, \ldots, \ell_k)$.

Letting n_P denote the number of edges of H incident to vertices in P, we have that

$$\begin{aligned} n_P &\geq m - h(n - |P|, r, \{\ell_2, \ldots, \ell_k\}) \\ &\geq h(n, r, \{\ell_2, \ldots, \ell_k\}) - h(n - (\ell + 1), r, \{\ell_2, \ldots, \ell_k\}) \\ &= \frac{\lfloor \frac{\ell_1 + 1}{2} \rfloor n^{r-1}}{(r-1)!} + O(n^{r-2}). \end{aligned}$$

(3.1)

We now focus on counting sets of vertices which can be used to easily "finish" edges started by vertices in P. With this in mind, for every set R of $r - 1$ vertices from $V(H) \setminus P$, we define

$$A_R = \{E' \in E(H) : R \subseteq E' \text{ and } E' \setminus R \in P\}.$$

We now break the $(r-1)$ subsets of $V(H) \setminus P$ into two sets, dependent on the size of their respective A_R:

$$A = \left\{ R \in (V(H) \setminus P)^{(r-1)} : |A_R| \leq \left\lfloor \frac{\ell_1 + 1}{2} \right\rfloor - 1 \right\},$$

$$B = \left\{ R \in (V(H) \setminus P)^{(r-1)} : |A_R| \geq \left\lfloor \frac{\ell_1 + 1}{2} \right\rfloor \right\}.$$

Counting edges entirely contained in P and the edges incident to the sets A, B defined above, we have that

$$n_P \leq \binom{|V(P)|}{2}\binom{n}{r-2} + \left(\left\lfloor \frac{\ell_1 + 1}{2} \right\rfloor - 1 \right)|A| + |V(P)||B|$$

$$(3.2) \qquad \leq \binom{|V(P)|}{2}\binom{n}{r-2} + \left(\left\lfloor \frac{\ell_1 + 1}{2} \right\rfloor - 1 \right)\binom{n}{r-1} + r\ell_1 |B|.$$

By comparison of the upper and lower bounds on n_P, (3.1), and (3.2), we have that

$$(3.3) \qquad\qquad |B| \geq \frac{n^{r-1}}{r\ell_1 (r-1)!} + O(n^{r-2}).$$

To each set $R \in B$ we associate a set of $\lfloor \frac{\ell_1+1}{2} \rfloor$ vertices from A_R arbitrarily. From (3.3), we see that some set of $\lfloor \frac{\ell_1+1}{2} \rfloor$ vertices is chosen many times; here "many" is at least

$$(3.4) \qquad \binom{|V(P)|}{\lfloor \frac{\ell_1+1}{2} \rfloor}^{-1} \frac{n^{r-1}}{(r-1)!r\ell_1} + O(n^{r-2}) \geq \binom{r\ell_1}{\lfloor \frac{\ell_1+1}{2} \rfloor}^{-1} \frac{n^{r-1}}{(r-1)!r\ell_1} + O(n^{r-2}).$$

Thus each loose path on ℓ_1 vertices in H contains a subset of $\lfloor \frac{\ell_1+1}{2} \rfloor$ vertices that has many common edge-finishing $(r-1)$-sets in the rest of the graph.

Let U be such a set of $\lfloor \frac{\ell_1+1}{2} \rfloor = \ell'$ vertices and X_U be the set of common edge-finishing $(r-1)$-sets. Since

$$|E\left(V\left(H\right) \setminus U\right)| > h(n - \ell', r, \{\ell_2, \ldots, \ell_k\}),$$

we can find $k-1$ vertex disjoint loose paths of appropriate lengths on vertices inside $V(H) \setminus U$, say, on vertex set W with $|W| < (\ell_2 + \cdots + \ell_k)r$.

We shall now find a loose path with ℓ_1 edges, all of which are of the form $X' \cup \{u\}$, where $X' \in X_U$ and $u \in U$. Applying Lemma 3.1 to the $(r-1)$-hypergraph on vertex set $V(\mathcal{H}) \setminus U$ and edge set X_U with $c < \frac{1}{(r-1)!r\ell_1 \binom{r\ell_1}{\ell'}}$ and $t = (\ell_2 + \cdots + \ell_k)r + \ell'(2r-3)$, we see from (3.4) that we can find a pair of 1-intersecting edges Y_1 and Z_1 disjoint from W. We can repeat this argument ℓ' more times to find vertex disjoint 1-intersecting edges $Y_2, Z_2, \ldots, Y_{\ell'+1}, Z_{\ell'+1}$, all also disjoint from W. Setting $U = \{u_1, u_2, \ldots, u_{\ell'}\}$, we can make a loose path of length ℓ_1, with edges

$$Z_1 \cup \{u_1\}, Y_2 \cup \{u_1\}, Z_2 \cup \{u_2\}, \ldots, Z_{\ell'} \cup \{u_{\ell'}\}, Y_{\ell'+1} \cup \{u_{\ell'}\},$$

where the last edge is only required if ℓ_1 is odd. In fact this path is not just loose but also linear. Thus we have constructed k disjoint loose paths and so our initial graph cannot have more than $h(n, r, \{\ell_1, \ldots, \ell_k\})$ edges. $\square$

Remark 1. As mentioned in the proof sketch, a major step is examining sets of vertices which are incident to many edges. From here, we determine that there is a set

of vertices which is large enough that we can construct our entire forest using these vertices as building blocks in the case that a particular path is found in the rest of the graph. When proving Theorem 2.2, this step could also be deduced from Theorem 6.2 of [18] for $r \geq 4$; embedding the necessary linear forests can then be carried out as above.

4. Multiple linear paths. We now proceed to forests of linear paths. The techniques are similar to the proofs for loose paths, but as the intersection of edges in linear paths have a particular shape (i.e., just a single vertex), we require some extra tools. The difference arises in the last steps, where we are building a linear path out of common neighborhoods. Instead of simply taking any two intersecting edges, as in the case for loose paths, we need to find edges which intersect appropriately for building linear paths.

Proof of Theorem 2.1 For ease of notation, we define $a(n, r, k, \ell) = \binom{n-1}{r-1} + \cdots + \binom{n-t}{r-1} + d_\ell$.

The case $k = 1$ is provided by Theorem 1.7. We proceed by induction on k. Let $k \geq 2$, and let H be a hypergraph on n vertices and with $|E(H)| = m > a(n, r, k, \ell)$. Since $a(n, r, k, \ell) > a(n, r, 1, \ell)$, for n large enough, we can find at least one linear path inside H.

As in section 3, consider one of these linear paths, say, on vertex set P. Certainly $|E(V(H) \setminus P)| \leq a(n - |P|, r, k - 1, \ell)$, or else by induction, the graph on $V(H) \setminus P$ contains $(k - 1) \cdot L_\ell^{(r)}$; these along with the linear path on P form $k \cdot L_\ell^{(r)}$. As before, we let n_P denote the number of edges of H incident to vertices in P; by identical calculations, we have that

$$(4.1) \qquad n_P \geq \frac{\lfloor \frac{\ell+1}{2} \rfloor n^{r-1}}{(r-1)!} + O(n^{r-2}).$$

Again we focus on counting sets of vertices which can be used to build edges started by vertices in P, defining A_R, A, and B identically to the proof for loose paths; by the same counting arguments, we get that

$$(4.2) \qquad |B| \geq \frac{\frac{n^{r-1}}{(r-1)!} + O(n^{r-2})}{r\ell}.$$

To each set $R \in B$, we now associate a set of $\lfloor \frac{\ell+1}{2} \rfloor$ vertices from A_R arbitrarily. From (4.2), we see that some set U of $\lfloor \frac{\ell+1}{2} \rfloor$ vertices is chosen many times; here "many" is again at least

$$(4.3) \qquad \binom{r\ell}{\lfloor \frac{\ell+1}{2} \rfloor}^{-1} \frac{n^{r-1}}{(r-1)!r\ell} + O(n^{r-2}).$$

Thus, as before, each linear path found in H has a set of $\lfloor \frac{\ell+1}{2} \rfloor$ vertices which have many common edge-finishing $(r - 1)$-sets in the rest of the graph, and we can again find $(k - 1) \cdot L_\ell^{(r)}$ on vertices inside $V(H) \setminus U$, say, on vertex set W.

We are now in the same position we were at the end of the proof of Theorem 2.4, and so we can construct a linear path with ℓ edges and thus we have constructed $k \cdot L_\ell^{(r)}$.

We note that the hypergraph on n vertices in which each edge is incident to a specified set S of t vertices, along with all edges disjoint from S containing some two fixed vertices not in S when k is even, gives a graph with exactly $\binom{n-1}{r-1} + \cdots + \binom{n-t}{r-1} + d_\ell$ edges and without k vertex disjoint linear paths of length ℓ; thus our result gives the exact value of the extremal function. $\square$

Modifying the above proof in the same manner as was used to deal with differing path lengths in the proof of Theorem 2.4, we obtain Theorem 2.2 for multiple linear paths of varying lengths; the lower bound is again given by the hypergraph on n vertices in which each edge is incident to at least one of a specified set S of t vertices, along with all edges disjoint from S containing some two fixed elements not in S when each of the ℓ_i is even.

5. Open problems. Using different methods, Fredi and Jiang [19] (for $r \geq 5$) and subsequently Kostochka, Mubayi, and Verstraëte [24] (for all $r \geq 3$) found the extremal number of the linear cycle $C_\ell^{(r)}$, the hypergraph formed from $L_{\ell-1}^{(r)}$ and one other edge that shares exactly one vertex with each of the two end edges of $L_{\ell-1}^{(r)}$ for all r and l. In particular for $r \geq 3$, $l \geq 4$, and $(r, l) \neq (3, 4)$, the extremal number satisfies

$$\mathrm{ex}_r(n; C_\ell^{(r)}) = \mathrm{ex}_r(n; L_\ell^{(r)}),$$

and the extremal hypergraphs are the same as well.

Using similar methods to those in this paper, the "eventual extremal number" (that is, the extremal number for large n) for hypergraphs consisting of linear cycles and linear paths can be determined. The main reason this is possible is because of the common structure shared by the extremal hypergraph for linear paths and linear cycles.

PROBLEM 1. *What hypergraphs have an eventual extremal hypergraph structure similar to the linear path, that is, consist of every edge adjacent to a set of t vertices, and $o(n^{r-1})$ other edges?*

As well as paths, there are several notions of trees in hypergraphs as well. We define the tight r-tree inductively as follows. Every r-graph consisting of a single edge is an r-tree. Suppose that $\mathcal{T}$ is an r-tree, and $E' \in E(\mathcal{T})$. Then for any $S \in \binom{E'}{r-1}$ and $v \notin V(\mathcal{T})$, the tree defined with edge set $E(\mathcal{T}) \cup \{S \cup \{v\}\}$ is a tight r-tree.

Using this definition, Kalai (see, e.g., [15]) proposed the following generalization of the Erdős–Sós conjecture.

CONJECTURE 3 (Kalai, 1984). *Let $r \geq 2$, and let $\mathcal{T}$ be a tight r-tree on ℓ vertices. Then for n sufficiently large,*

$$\mathrm{ex}_r(n; \mathcal{T}) \leq \frac{\ell - r}{r} \binom{n}{k-1}.$$

This conjecture remains open in virtually all cases. For general r, it is proven only in the case of trees containing an edge intersecting every other edge in $k-1$ vertices; thus the tree is essentially a star [15].

PROBLEM 2. *What are the extremal numbers for hyperforests containing nonpath components?*

This seems to be a quite difficult problem. There are extraordinarily few examples of even individual nonpath hypertrees for which the extremal numbers are known (see, e.g., [9], [23] for a survey).

REFERENCES

[1] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi, *On the Approximative Solution of the Erdős-Sós Conjecture on Trees*, manuscript.

[2] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi, *The Solution of the Erdős-Sós Conjecture for Large Trees*, manuscript.

[3] M. Ajtai, J. Komlós, M. Simonovits, and E. Szemerédi, *Some Elementary Lemmas on the Erdős-Sós Conjecture for Trees*, manuscript.

[4] N. Alon, M. Krivelevich, and B. Sudakov, *Turán numbers of bipartite graphs and related Ramsey-type questions*, Combin. Probab. Comput., 12 (2003), pp. 477–494.

[5] B. Bollobás, D. E. Daykin, and P. Erdős-Sós, *Sets of independent edges of a hypergraph*, Quart. J. Math. Oxford Ser. (2), 27 (1976), pp. 25–32.

[6] B. Bollobás, *Extremal Graph Theory*, London Math. Soc. Monogr. 11, Academic Press, London, 1978.

[7] B. Bollobás, *Modern graph theory*, Grad. Texts in Math. 184, Springer-Verlag, New York, 1998.

[8] N. Bushaw and N. Kettle, *Turán numbers of multiple paths and equibipartite forests*, Combin. Probab. Comput., 20 (2011), pp. 837–853.

[9] N. Bushaw, *Problems in Extremal Combinatorics*, Ph.D. thesis, University of Memphis, Memphis, TN, 2012.

[10] P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc., 52 (1946), pp. 1087–1091.

[11] P. Erdős, *Extremal problems in graph theory*, Theory of Graphs and its Applications, Czech Republic Academy of Science, Prague, 1964, pp. 29–36.

[12] P. Erdős and T. Gallai, *On maximal paths and circuits of graphs*, Acta Math. Acad. Sci. Hungar., 10 (1959), pp. 337–356.

[13] P. Erdős, C. Ko, and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math. Oxford Ser. (2), 12 (1961), pp. 313–320.

[14] P. Erdős, *A problem on independent r-tuples*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math., 8 (1965), pp. 93–95.

[15] P. Frankl and Z. Füredi, *Exact solution of some Turán-type problems*, J. Combin. Theory Ser. A, 45 (1987), pp. 226–262.

[16] P. Frankl, V. Rödl, and A. Ruciński, *On the maximum number of edges in a triple system not containing a disjoint family of a given size*, Combin. Probab. Comput., 21 (2012), pp. 141–148.

[17] P. Frankl, *On the maximum number of edges in a hypergraph with given matching number*, arXiv e-prints, 2012.

[18] Z. Füredi, T. Jiang, and R. Seiver, *Exact solution of the hypergraph Turán problem for k-uniform linear paths*, Combinatorica, to appear.

[19] Z. Füredi and T. Jiang, *Hypergraph Turán numbers of linear cycles*, J. Combin. Theory Ser. A, 123 (2014), pp. 252–270.

[20] Z. Füredi and M. Simonovits, *The history of degenerate (bipartite) extremal graph problems*, in Erdös Centennial, L. Lovasz, I. Ruzsa, V. T. Sos, and D. Palvolgyi eds., Bolyai Soc. Math. Stud. 25, Springer-Verlag, Heidelberg. 2013, pp. 169–264.

[21] H. Huang, P. Loh, and B. Sudakov, *The size of a hypergraph and its matching number*, Combin. Probab. Comput., 21 (2012), pp. 442–450.

[22] P. Keevash, D. Mubayi, and R. Wilson, *Set systems with no singleton intersection*, SIAM J. Discrete Math., 20 (2006), pp. 1031–1041.

[23] P. Keevash, *Hypergraph Turán problems*, in Surveys in Combinatorics 2011, London Math. Soc. Lecture Note Ser. 392, Cambridge University Press, Cambridge, 2011, pp. 83–139.

[24] A. Kostochka, D. Mubayi, and J. Verstraëte, *Turan problems and shadows* I: *Paths and cycles*, arXiv e-prints, 2013.

[25] T. Łuczak and B. Mieczkowska, *On Erdős' extremal problem on matchings in hypergraphs*, J. Combin. Theory Ser. A, 124 (2014), pp. 178–194.

[26] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok, 48 (1941), pp. 436–452.