RESULTS AND PROBLEMS ON SATURATION NUMBERS FOR LINEAR FORESTS

Guantao Chen, Jill R. Faudree, Ralph J. Faudree, Ronald J. Gould, Michael S. Jacobson, and Colton Magnant

ABSTRACT. A graph G is called *H-saturated* if G contains no copy of H, but for any edge e in the complement of G, the graph $G + e$ contains some copy of H. The minimum size of an n-vertex H-saturated graph is denoted by $\mathbf{sat}(n, H)$ and is called the *saturation number of* H. In [**?**], Kászonyi and Tuza determined the values of $\mathbf{sat}(n, H)$ when H is a path or a disjoint union of edges. In this paper, we determine the values of $\mathbf{sat}(n, H)$ for the disjoint union of paths (a linear forest) within a constant depending only on H. Moreover, we obtain exact values for some special classes and include several conjectures.

1. INTRODUCTION

A graph G is called *H-saturated* if G contains no copy of H, but for any edge e in the complement of G, the graph $G+e$ contains some copy of H. The set of H-saturated graphs of order n is denoted by $\mathbf{SAT}(n, H)$, and the *saturation number*, denoted $\mathbf{sat}(n, H)$, is the minimum size of a graph in $\mathbf{SAT}(n, H)$. The maximum size of a graph in $\mathbf{SAT}(n, H)$ is the well known Turán extremal number (see [**?**]), and is usually denoted by $ex(n, H)$. The graphs in $\mathbf{SAT}(n, H)$ of minimum size will be denoted by $\underline{\mathbf{SAT}}(n, H)$, and those with a maximum number of edges will be denoted by $\overline{\mathbf{SAT}}(n, H)$. Thus, all graphs in $\underline{\mathbf{SAT}}(n, H)$ have $\mathbf{sat}(n, H)$ edges and graphs in $\overline{\mathbf{SAT}}(n, H)$ have $ex(n, H)$ edges.

The notion of the saturation number of a graph was introduced by Erdős, Hajnal, and Moon in [**?**] in which the authors proved
$\mathbf{sat}(n, K_t) = \binom{t-2}{2} + (n-t+2)(t-2)$ and $\underline{\mathbf{SAT}}(n, K_t) = \{K_{t-2} + \overline{K}_{n-t+2}\}$.

Since then, $\mathbf{sat}(n, H)$ and $\underline{\mathbf{SAT}}(n, H)$ have been investigated for a range of graphs H. Some examples include trees, cycles, bipartite graphs, matchings, friendship graphs, and books. However, the exact value of $\mathbf{sat}(n, H)$ and a complete characterization of $\underline{\mathbf{SAT}}(n, H)$ are known for very few special classes of graphs H. For a summary of known results see [**?**] and for results on trees see [**?**]. There has been extensive study of extremal numbers, $ex(n, H)$, and a well-developed theory of $\overline{\mathbf{SAT}}(n, H)$ has been established. This is not so for saturation numbers, $\mathbf{sat}(n, H)$, and the graphs in $\underline{\mathbf{SAT}}(n, H)$. For example, many of the hereditary and monotone properties that hold for $\overline{\mathbf{SAT}}(n, H)$ do not hold for $\underline{\mathbf{SAT}}(n, H)$.

For $t \geq 2$, let $F = P_{k_1} \cup P_{k_2} \cup \cdots \cup P_{k_t}$ be a linear forest where P_k denotes a path on k vertices and $k_1 \geq k_2 \geq \cdots \geq k_t \geq 2$. An example of what is known about $ex(n, F)$ for a disjoint union of paths F can be found in results by Bushaw and Kettle [**?**]. We will show that the value of $\mathbf{sat}(n, F)$ is determined by the size of the shortest path, P_{k_t}. We determine the value of $\mathbf{sat}(n, F)$ within a constant that is a function of the order of F. Moreover, we will improve the constant for linear forests composed of paths of equal length. Exact values will be determined in some small order cases.

Only finite graphs without loops or multiple edges will be considered. Notation will be standard, and generally follow the notation of [**?**]. We begin by defining a particular family of trees which will be used repeatedly in the remainder of the paper.

A *perfect degree three tree* is a tree such that every vertex has degree 3 or degree 1, and all vertices of degree 1 are the same distance from the center. Thus, all perfect degree three trees differ only by their diameter. For $k \geq 2$, we will let T_k denote the perfect degree three tree whose longest path contains k vertices. By this definition, $T_2 = K_2$ and $T_3 = K_{1,3}$. See Figure **??** for more examples.

(A) T_5 (B) T_6

FIGURE 1. Examples of perfect degree three trees.

In some instances it will be useful to view T_k as a rooted (or doubly rooted) tree. Specifically, if k is odd, let the root r be the unique vertex in the center of T_k. Viewed in this way, the tree has $\lceil \frac{k}{2} \rceil$ *levels*, the root has three children, all vertices in the middle levels have two children, and all vertices of degree 1 are in the bottom level. If k is even, the center consists of two adjacent vertices. In this case, all vertices have two children except for those of degree 1, all of which are in the bottom of the $k/2$ levels.

Observe that T_{k-1} is P_k-saturated for $k \geq 4$. In addition, any graph obtained from T_{k-1} by adding more pendant vertices to those already adjacent to vertices of degree 1 remains P_k-saturated. As Theorem **??** will show, for $k \geq 5$, the graphs T_{k-1} are building blocks for graphs in $\underline{\mathbf{SAT}}(n, P_k)$. For ease of reference, when $k \geq 3$, we let

$$a_k = |V(T_{k-1})| = \begin{cases} 3 \cdot 2^{m-1} - 2 & \text{if } k = 2m \\ 4 \cdot 2^{m-1} - 2 & \text{if } k = 2m+1, \end{cases}$$

and we define $a_2 = 1$.

Theorem 1.1. **[?]** *Let P_k be a path on $k \geq 5$ vertices and let T_{k-1} and a_k be the tree and the order of the tree defined above, respectively. Then, for $n \geq a_k$, every graph in $\underline{\mathbf{SAT}}(n, P_k)$ consists of a forest with $\lfloor n/a_k \rfloor$ components and $\mathbf{sat}(n, P_k) = n - \lfloor n/a_k \rfloor$. Furthermore, if T is a P_k-saturated tree, then $T_{k-1} \subseteq T$.*

Note that P_2-saturated graphs consist of the set of empty graphs. The set of P_3-saturated graphs consist of independent edges with at most one isolated vertex, depending upon the parity of n. While T_3 is P_4-saturated, the set of P_4-saturated graphs of *minimum size* consists of independent edges with possibly a single K_3 as a component, depending upon the parity of n.

In this paper we investigate the saturation number of graphs consisting of a disjoint union of paths. For convenience, we refer to such graphs as *linear forests*. The following theorem established the saturation number for a matching, tK_2, which can be viewed as a particular family of linear forests.

Theorem 1.2. **[?]**
(a) *For $n \geq \max\{2k, 3k-3\}$, $\mathbf{sat}(n, kP_2) = 3k - 3$.*
(b) *For $n \geq \max\{2k, 3k-3\}$, $\underline{\mathbf{SAT}}(n, kP_2) = \{(k-1)K_3 \cup \overline{K}_{n-3k+3}\}$ or $k = 2$, $n = 4$ and $\underline{\mathbf{SAT}}(n, kP_2) = \{K_{1,3}, K_3 \cup K_1\}$.*

Our first result establishes bounds on the saturation number for an arbitrary linear forest.

Theorem 1.3 (G. Chen, JF, R. Faudree, R. Gould, M. Jacobson, C. Magnant). *For $t \geq 2$, let $F = P_{k_1} \cup P_{k_2} \cup \cdots \cup P_{k_t}$ be a linear forest with $k_1 \geq k_2 \geq \cdots \geq k_t$ and let $k = k_t$ and $q = \left(\sum_{i=1}^{t} k_i\right) - 1$. Then, for n sufficiently large,*

$$\begin{aligned} n - \left\lfloor \frac{n}{a_k} \right\rfloor &\leq \mathbf{sat}(n, F) \leq \binom{q}{2} + n - q - \left\lfloor \frac{n-q}{a_k} \right\rfloor, \quad \text{if } k \neq 4 \\ n - \left\lfloor \frac{n}{2} \right\rfloor &\leq \mathbf{sat}(n, F) \leq \binom{q}{2} + n - q - \left\lfloor \frac{n-q}{2} \right\rfloor, \quad \text{if } k = 4. \end{aligned}$$

The proof of Theorem **??** will be given in Section 3. The above result shows that $\mathbf{sat}(n, F)$ is determined by the order of the smallest component of F. An immediate corollary of Theorem **??** is the following:

Corollary 1.1. *If $F = P_{k_1} \cup P_{k_2} \cup \cdots \cup P_{k_t}$ is a linear forest with $k_1 \geq k_2 \geq \cdots \geq k_t \geq 2$ and $k = k_t$, then*

$$\begin{aligned} \mathbf{sat}(n, F) &= n - \left\lfloor \frac{n}{a_k} \right\rfloor + c(n) \quad \text{if } k \neq 4, \text{ and} \\ \mathbf{sat}(n, F) &= n - \left\lfloor \frac{n}{2} \right\rfloor + c(n) \quad \text{if } k = 4, \end{aligned}$$

for some constant $c(n)$ with $0 \leq c(n) \leq \binom{q}{2} - q + \left\lceil \frac{q}{a_k} \right\rceil$.

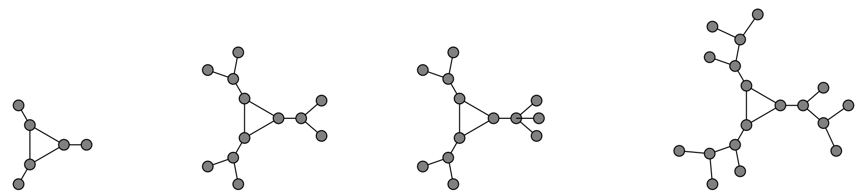
Since $a_{k+1} > a_k$ if $k > 2$, $\mathbf{sat}(n, F)$ increases as the smallest component in F increases provided n is sufficiently large. As a consequence of the the monotone property, we have the following result.

Corollary 1.2. *Let F and F^* be two linear forests such that the order of the smallest components in F and F^* are k and k^*. If $k > k^*$ and $(k, k^*) \neq (4, 3)$, then $\mathbf{sat}(n, F) > \mathbf{sat}(n, F^*)$ provided n is sufficiently large.*

We will now define two families of graphs that will be used throughout the remainder of the paper.

For $k \geq 3$, let N_k be obtained from a K_3 by attaching a perfect degree three tree T_{k-1} at each vertex of the K_3 beginning with a pendant vertex of the tree. Let N_4^* be the graph obtained from N_4 by adding another single pendant edge to one of the centers of the stars in the construction of N_4. Observe that $|V(N_k)| = 3a_k$. (See Figure **??**.)

For $k \geq 5$, we define $Z(n, k)$ to be the graph on n vertices consisting of $\lfloor n/a_k \rfloor$ vertex disjoint copies of a T_{k-1} such that the remaining $r = n - a_k \lfloor n/a_k \rfloor$



(A) N_3 (B) N_4 (C) N_4^* (D) N_5

FIGURE 2. Examples of the family N_k

vertices are attached as pendant vertices to the same vertex in the penultimate level of a single copy of T_{k-1}. (See Figure **??**.) The graph $Z(n,k)$ is shown to be one of the graphs in $\underline{\mathbf{SAT}}(n,P_k)$ in [**?**]. For $k=4$, we define $Z(n,k)=\lfloor n/2\rfloor K_2$ when n is even, and $Z(n,k)=K_3\cup\lfloor (n-3)/2\rfloor K_2$ when n is odd. For $k=3$, we define $Z(n,k)=\lfloor n/2\rfloor K_2$. For $k=2$, $Z(n,k)=\overline{K}_n$.

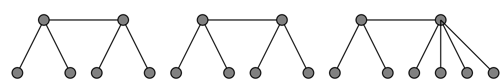


FIGURE 3. The graph $Z(20,5)$

When all components of F are paths with the same order, we make the following conjecture:

Conjecture 1.1. *Let $t\geq 2$ be an integer. For n sufficiently large,*

(1) $\mathbf{sat}(n,tP_3)=\left\lfloor\frac{n+6t-6}{2}\right\rfloor$ *and* $(t-1)N_3\bigcup\left\lfloor\frac{n-6t+6}{2}\right\rfloor P_2\in\underline{\mathbf{SAT}}(n,tP_3)$.

(2) $\mathbf{sat}(n,tP_4)=\begin{cases}(n+12t-12)/2 & \textit{if } n \textit{ is even}\\ (n+12t-11)/2 & \textit{if } n \textit{ is odd.}\end{cases}$
Moreover,

$$\begin{array}{rcll}(t-1)N_4\cup(1/2)(n-12t+12)P_2 & \in & \underline{\mathbf{SAT}}(n,tP_4) & \textit{if } n \textit{ is even and}\\ N_4^*\cup(t-2)N_4\cup(1/2)(n-12t+11)P_2 & \in & \underline{\mathbf{SAT}}(n,tP_4) & \textit{if } n \textit{ is odd.}\end{array}$$

(3) *For* $k\geq 5$, $\mathbf{sat}(n,tP_k)=n-\left\lfloor\frac{n}{a_k}\right\rfloor+3(t-1)$, *and*
$(t-1)N_k\cup Z(n-3(t-1)a_k,k)\in\underline{\mathbf{SAT}}(n,tP_k)$.

Conjecture **??** will be shown to be true for tP_3's for $t=2,3$ in Theorem **??** and for $2P_4$'s in Theorem **??**.

2. Lemmas

We next prove seven lemmas that will be used repeatedly in the proofs of the theorems.

Lemma 2.1. *Let F be a linear forest and G an F-saturated graph. If w is a vertex of G of degree 2 and u, v are the neighbors of w, then $uv \in E(G)$.*

PROOF. Let F, G, u, v, and w be as stated in the lemma. Assume that $uv \notin E(G)$. Then, $G + uv$ contains a copy of F. Clearly, one of these paths contains uv. Replacing uv with uw or vw or uwv, shows that there is a copy of F in G, contradicting the assumption that $F \not\subseteq G$. □

Lemma 2.2. *For every integer $k \geq 3$, N_k is $2P_k$ saturated, and $N_k \cup Z(n - 3a_k, k)$ is $2P_k$-saturated.*

PROOF. Recall that the graph N_k has a cycle C with 3 vertices and three attached trees each isomorphic to T_{k-1}. Furthermore, recall that the tree T_{k-1} contains a P_{k-1} and is P_k-saturated. Thus, the addition of an edge to N_k between two vertices of the same copy of T_{k-1} will result in a P_k using at most one vertex of the cycle. All other edges of the complement of N_k lie between vertices of distinct copies of T_{k-1}, say T and T'. If one of the two endpoints of the added edge e lies on or between the root (or closest root) and the vertex on C in its respective tree, say T, then there exists a P_k using the longest path in T, the added edge, and ending in $T' - V(C)$. If both endpoints lie such that the the root (or closest root) lies between the endpoint and C, then there is a P_k beginning in $T \cap C$, through T to the added edge and ending in $T' - C$. In all three cases, the addition of an edge to N_k produces a P_k that uses at most two of the T_k's and at most one vertex of C. Thus a second vertex disjoint P_k can be constructed in the remaining T_k and the remaining two vertices of C. Finally, if N_k is $2P_k$-saturated, then $N_k \cup Z(n - 3a_k, k)$ is $2P_k$-saturated. □

An immediate corollary of Lemma **??** is the following:

Corollary 2.1. *For every integer $k \geq 3$ and $t \geq 2$, $(t-1)N_k$ is tP_k saturated, and also $(t-1)N_k \cup Z(n - (t-1)3a_k, k)$ is tP_k-saturated.*

Lemma 2.3. *If $m \geq 2$ and $k \geq 3$ are integers, then no mP_k-saturated graph is a tree.*

PROOF. Suppose the contrary: *there is a tree T which is mP_k-saturated*. Pick a vertex $u \in T$ and treat T as a rooted tree with root u. For each vertex v, let T_v denote the subtree of G that consists of v and all its descendants. Clearly, $T_u = T$.

Since T is mP_k-saturated, T contains a copy of $m-1$ disjoint paths P_k. Let v be the root of a P_k with the lowest rank (that is, the vertex of P_k closest to u is most distant from the root u). Since T does not contain mP_k, $T - T_v$ does not contain a $(m-1)P_k$.

Select a vertex $w \in V(T_v)$ such that $vw \notin E(T)$. Then $T + vw$ contains a copy of mP_k. By the minimality of the rank of v, $T - T_v$ contains $(m-1)P_k$, a contradiction. Note that such a vertex w must exists unless $k = 3$, v is the middle vertex of this path, and v has precisely two children (say x_1 and x_2) in T both of which are pendant. In this case, the same argument applies to $T + x_1x_2$. This completes the proof of Lemma **??**. □

Lemma 2.4. *Let $t \geq 2$ and let $k_1 \geq k_2 \geq \cdots \geq k_t \geq 5$ be integers. If a tree T is $P_{k_1} \cup P_{k_2} \cup \cdots \cup P_{k_t}$ saturated, then $|T| \geq a_{k_t}$.*

PROOF. Since $t \geq 2$ and $k_t \leq k_i$ for all $i \leq t$, the graph T must contain a path P of length k_t. Let $k = k_t$. Let C be the center of P. Note that $|C| = 1$ if k is odd and $|C| = 2$ is k is even. Starting with C, we perform a Breadth-First-Search and partition $V(T)$ into $V_0 = C$, V_1, V_2, … such that all vertices in V_i have distance i from C. Clearly, $V_i \neq \emptyset$ for $i \leq k/2 - 1$.

For each vertex v, let T_v be the subtree consisting of v and all descendants of v in the tree T under the Breadth-First-Search. (See Figure **??**). Let $\ell(v)$ be the length of a longest path in T_v starting at v. By Lemma **??**, T does not have vertices of degree 2. Thus, every vertex $x \in V_i$ is either a pendant vertex or has at least two children in V_{i+1}.

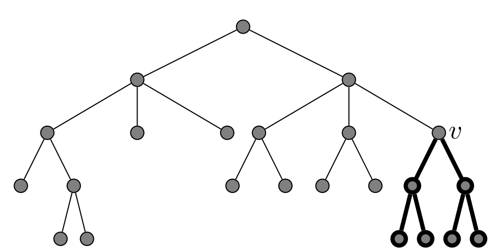


FIGURE 4. The subtree T_v is shown in bold.

Recall that a Breadth-First-Search rooted at the center of T_{k-1} would have exactly $\lceil (k-1)/2 \rceil$ levels such that every vertex has exactly two children except the last level in which all are pendant. Thus, by comparing T with tree T_{k-1}, if $|V(T)| < |V(T_{k-1})| = a_k$, then, intuitively, some vertex of T must have pendant children prematurely. More rigorously, if $|V(T)| < |V(T_{k-1})| = a_k$, there exists a vertex v who has a child w such that $\ell(w) > \ell(w^*)$ for all other children w^* of v and $\ell(w^*) \leq k/2 - 1$. Let u be the predecessor of v. (See Figure **??**.)

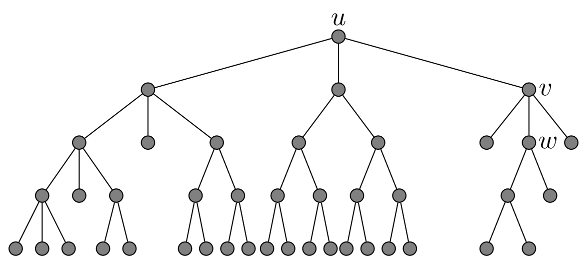


FIGURE 5. Vertices $v,\ w$ and u from the proof of Lemma **??**

Now $T + uw$ must contain a copy of $F := P_{k_1} \cup P_{k_2} \cup \cdots \cup P_{k_t}$ such that one of the paths uses edge uw. If the new path follows edge uw from u to w and remains in T_w, then there exists a copy of F in T by replacing edge uw with the path uvw. Thus, the new path must contain the path $uwvw^*$ and remain in T_{w^*}. But $\ell(w^*) < \ell(w)$ again implies there exists a copy of F in T. Thus, $|V(T)| \geq a_k$. □

Lemma 2.5. *If H is a connected graph with $8 \leq |H| \leq 11$, $|E(H)| = |H|$ and such that all vertices of degree 2 lie on a triangle, then either H contains a $2P_4$ or the unique cycle in H is a C_3.*

PROOF. Since H is connected and $|E(H)| = |H|$, H has a unique cycle C with no chords and $H - E(C)$ is a forest. If $C = K_3$, the result follows, thus assume $|V(C)| \geq 4$. Now, to avoid vertices of degree 2 on C, there must exist a set of $|V(C)|$ independent edges between $V(C)$ and $V(H) - V(C)$. But for cycles on 4 or more vertices, such a graph always contains a $2P_4$. Thus, under the hypotheses of the lemma, the unique cycle in H is a C_3 or H contains a $2P_4$. □

Lemma 2.6. *If H is a connected graph with $8 \leq |H| \leq 9$, $|E(H)| = |H|+1$, and such that all vertices of degree 2 lie on a triangle, then either H contains a $2P_4$ or the longest cycle in H is a C_5.*

PROOF. Assume that $|E(H)| = |H|+1$, and let C be the longest cycle in H. If $|C| \geq 7$, then clearly there is a $2P_4$ in H. If C has no chords and $|C| = 6$, then there will always be a vertex of degree 2 on C not on a triangle. Assume there is a chord in C. Then this is the unique chord in C. So there are two consecutive vertices on C not incident to the chord, say x and y. Now C along with the two independent edges that must be adjacent to x and y contain a P_8 and therefore a copy of $2P_4$. Thus, under the hypothesis of the lemma, either H contains a $2P_4$ or the longest cycle in H is a C_5. □

Lemma 2.7. *If H is a connected graph containing a triangle with $8 \leq |H| \leq 11$ and $|E(H)| = |H|$, then H is not $2P_4$-saturated.*

PROOF. Suppose H satisfies the hypotheses of the Lemma and is $2P_4$-saturated. By assumption, the only cycle in H is C_3 and $H - E(C_3)$ is a forest. Label the vertices of the cycle u_i for $i = 1, 2, 3$ and label the tree rooted at u_i as T_i. Observe that Lemma **??** implies that no vertex of T_i can have degree 2 other than possibly u_i. This fact will be used repeatedly in the following argument.

If all the vertices of $H - V(C_3)$ are pendant, H is not $2P_4$-saturated. Thus, there is at least one tree, T_i, containing a path on at least three vertices starting at u_i. Lemma **??** implies that such a tree T_i could not be simply a path, forcing T_i to have at least 4 vertices. Since $|V(H)| \leq 11$, at most two trees can have such paths. We simply consider the two cases.

If u_i has a path of length at least 2 in T_i, for $i = 1, 2$. Then, to avoid a $2P_4$, neither tree can have a path of length more than 2. Let u_iv_i be an edge in T_i. Then $H + v_1u_2$ illustrates that H cannot be $2P_4$-saturated.

Assume that only one tree, say T_1, has a path of length at least 2 beginning at u_1. If the longest path in T_1 starting at u_1 contains 5 or more vertices, the additional vertices implied by Lemma **??** forces H to contain $2P_4$. Thus, the longest path in T_1 from u_1 has either 3 or 4 vertices. If the path has 4 vertices, $T_2 \cong T_3 \cong K_1$ to avoid a $2P_4$ in H. Such an H is not $2P_4$-saturated. If the path has 3 vertices (say, u_1, v_1, v_2), adding edge u_1v_2 illustrates that H cannot be $2P_4$-saturated. □

3. Proof of Theorem ??

Let $F = P_{k_1} \cup P_{k_2} \cup \cdots \cup P_{k_t}$ be a linear forest with $k_1 \geq k_2 \geq \cdots \geq k_t \geq 2$ and let $k = k_t$ and $q = \left(\sum_{i=1}^{t} k_i\right) - 1$. To establish the upper bound, we construct a saturated graph G_F for F according to the value of k.

- If $k = 2$, let $G_F := K_q \cup \overline{K_{n-q}}$.
- If $k = 3$, let $G_F := K_q \cup \lfloor \frac{n-q}{2} \rfloor K_2 \cup (n - q - 2\lfloor \frac{n-q}{2} \rfloor)K_1$.
- If $k = 4$, let
$$G_F := \begin{cases} K_q \cup \frac{n-q}{2} K_2 & \text{provided } n-q \text{ is even} \\ K_q \cup \frac{n-q-3}{2} K_2 \cup K_3 & \text{provided } n-q \text{ is odd.} \end{cases}$$
- If $k \geq 5$, let $G_F := K_q \cup Z(n - q, k_t)$.

We claim that $G_F \in \mathbf{SAT}(n, F)$ for $k \geq 5$, and that all the other cases are similar. Certainly K_q is not big enough to contain a copy of F and $Z(n - q, k_t)$ does not contain any path of F so G_F contains no copy of F as a subgraph. Any edge added with one vertex in K_q trivially produces a copy of F as a subgraph. Also any edge added within $Z(n - q, k_t)$ produces a P_{k_t} and the remaining paths $P_{k_1}, \ldots, P_{k_{t-1}}$ in F can be embedded in K_q to produce a copy of F and complete the proof.

We now show the lower bound holds. Let $G \in \underline{\mathbf{SAT}}(n, F)$.

- If $k = 2$, then $a_k = 1$, so $|E(G)| \geq n - \lfloor \frac{n}{a_k} \rfloor = 0$. The result holds vacuously.
- If $k = 3$, since G is F-saturated, there can be at most one isolated vertex in G. Consequently, $|E(G)| \geq \frac{n}{2}$.
- If $k = 4$, then any isolated vertex in G would imply that all of the other components of G would have at least 3 vertices and $|E(G)| > 2(n-1)/3 > n/2$. Thus, clearly $\mathbf{sat}(n, F) \geq n/2$.
- Suppose $k \geq 5$. Consider any component G' of G that is a tree. Observe that G' must be F'-saturated for some sub-forest F' of F. Since $k \geq 5$, any edge added to H' must produce a path P_m for $m \geq k$. Therefore, by Lemma **??**, $|G'| \geq a_k$. Hence, if r is the number of vertices in the components of G that are not trees, then the number of edges in G is at least $r + (n - r) - \lfloor \frac{n-r}{a_k} \rfloor \geq n - \frac{n}{a_k}$.

This completes the proof of Theorem **??**.

4. Improvement of the constant bound

In this section we prove several theorems that improve the bounds of Theorem **??** in some special families of linear forests. Specifically, we improve the bounds in the cases when all the paths in the linear forest have the same length and when the forest has exactly two paths. Also, we show that Conjecture **??** holds in some small order cases.

We begin with linear forests such that all paths have the same length.

Theorem 4.1. *For n sufficiently large, $t \geq 2$ and $k \geq 5$,*

$$n - \left\lfloor \frac{n}{a_k} \right\rfloor \leq \mathbf{sat}(n, tP_k) \leq n - \left\lfloor \frac{n}{a_k} \right\rfloor + 3(t-1).$$

PROOF. By Corollary **??**, we know that $(t-1)N_k$ is tP_k-saturated, and also that $G = (t-1)N_k \cup Z(n-(t-1)3a_k, k)$ is tP_k-saturated. The number of edges in G is $(t-1)3a_k + (n-(t-1)3a_k) - \lfloor (n-(t-1)3a_k)/a_k \rfloor = n - \lfloor n/a_k \rfloor + 3(t-1)$. This verifies the upper bound for $\mathbf{sat}(n, tP_k)$. The lower bound is a direct consequence of Theorem **??**. □

In the case when $t = 2$, the following result gives a very close bound for two paths.

Theorem 4.2. *For n sufficiently large and $5 \leq k \leq \ell \leq \lceil (3k-2)/2 \rceil$,*

$$n - \left\lfloor \frac{n}{a_k} \right\rfloor \leq \mathbf{sat}(n, P_k \cup P_\ell) \leq n - \left\lfloor \frac{n}{a_k} \right\rfloor + 3.$$

Also, the graph $N_k \cup Z(n-3a_k, k) \in \mathbf{SAT}(n, P_k \cup P_\ell)$.

PROOF. Consider the graph $G = N_k \cup Z(n-3a_k, k)$. Since each of the trees T_{k-1} in N_k contains a P_{k-1}, the graph N_k has a P_{2k-2} containing two vertices of the triangle in N_k. Since any copy of P_k will contain at least 2 vertices of the triangle in N_k, there does not exist a $P_k \cup P_k$ in G. The addition of any edge in $Z(n-3a_k, k)$ will give a P_k disjoint from N_k. Also, each edge added to a $T_{k-1} \in N_k$ produces a P_k disjoint from a P_{2k-2} in N_k. An edge added between two different copies of T_{k-1} with one in N_k and one in $Z(n-3a_k, k)$ will produce a P_k disjoint from a P_{2k-2} in N_k. Finally, an edge added between two different copies of T_{k-1} in N_k will produce either a P_k in the two copies of T_{k-1}, and a disjoint P_l using vertices in the other copy of T_{k-1} along with some vertices in one of the initial T_{k-1}, or a P_l in the two copies of T_{k-1} disjoint from a P_k in the third T_{k-1} along with two vertices of the triangle. It is in this final case that the longest path P_l possible is $l = \lceil (3k-2)/2 \rceil$, since not all of the vertices of one of the T_{k-1} are available. Thus, $G \in \mathbf{SAT}(n, P_k \cup P_l)$. The lower bound is a direct consequence of Theorem **??**. □

The previous result provides support for the following conjecture.

Conjecture 4.1. *For n sufficiently large, $k \geq 4$, and $k \leq \ell \leq \lceil (3k-2)/2 \rceil$,*

$$\begin{aligned} \mathbf{sat}(n, P_k \cup P_\ell) &= n - \left\lfloor \frac{n}{a_k} \right\rfloor + 3, \text{ and} \\ N_k \cup Z(n - 3a_k, k) &\in \underline{\mathbf{SAT}}(n, P_k \cup P_\ell). \end{aligned}$$

The next two theorems support Conjecture **??**.

Theorem 4.3. *Let $1 \leq t \leq 3$ and $n \geq 6t$ be two positive integers. Then,*

$$\mathbf{sat}(n, tP_3) = \left\lfloor \frac{n+6t-6}{2} \right\rfloor \text{ and } (t-1)N_3 \cup \left\lfloor \frac{n-6t+6}{2} \right\rfloor P_2 \in \underline{\mathbf{SAT}}(n, tP_3).$$

Theorem 4.4. *(a) For an even integer $n \geq 12$ and $t = 1, 2$,*

$$\mathbf{sat}(n, tP_4) = \frac{n+12t-12}{2} \text{ and } (t-1)N_4 \cup \left(\frac{n-12t+12}{2} \right) P_2 \in \underline{\mathbf{SAT}}(n, tP_4).$$

(b) For an odd integer $n \geq 13$ and $t = 1, 2$,

$$\mathbf{sat}(n, P_4) = \frac{n+3}{2} \text{ and } K_3 \cup \left(\frac{n-3}{2} \right) P_2 \in \underline{\mathbf{SAT}}(n, P_4),$$

$$\mathbf{sat}(n, 2P_4) = \frac{n+13}{2} \text{ and } N_4^* \cup \left(\frac{n-13}{2} \right) P_2 \in \underline{\mathbf{SAT}}(n, 2P_4).$$

PROOF. (Theorem **??**) It is readily seen that graph $(t-1)N_3 \cup \lfloor \frac{n-6t+6}{2} \rfloor P_2$ is tP_3-saturated and it has $\lfloor \frac{n+6t-6}{2} \rfloor$ edges. Let G be a tP_3-saturated graph of order $n \geq 6t$. We will show that $|E(G)| \geq \lfloor \frac{n+6t-6}{2} \rfloor$ for $t = 1, 2, 3$.

In [**?**], the case when $t = 1$ was proven. Observe that for $t \geq 2$, if G contains N_3 as a connected component then the theorem follows by induction. That is, if G is tP_3-saturated and G contains a connected component isomorphic to N_3, then $G - N_3$ is a $(t-1)P_3$-saturated graph on $n - 6 > 6(t-1)$ vertices. Thus, $|E(G)| = 6 + |E(G - V(N_3))| \geq 6 + \lfloor \frac{n+6t-12}{2} \rfloor = \lfloor \frac{n+6t-6}{2} \rfloor$. We will first prove some general structural properties of components in $\underline{\mathbf{SAT}}(n, tP_3)$.

For ease of reference, let A represent a connected component of a tP_3-saturated graph G such that $|V(A)| \geq 3$. We will begin by proving several properties of connected components of tP_3-saturated graphs. By Lemma **??** no connected component of G of order 3 or more can be a tree. Thus, A must contain at least one cycle and $|E(A)| \geq |V(A)|$. By Lemma **??**, the neighbors of any vertex of degree 2 in G must be adjacent. Thus, if A contains a unique cycle, C, on 4 or more vertices, then $A - E(C)$ is a forest consisting of $|V(C)|$ nontrivial trees.

Furthermore, observe that no vertex of a tP_3-saturated graph can be adjacent to two vertices of degree 1 because the copy of P_3 obtained by adding the edge between these two pendant vertices can be replaced by one already in the graph. Consequently, if a component A contains a unique cycle, C, on 4 or more vertices, in fact $A - E(C)$ is a forest consisting of $|V(C)|$ copies of K_2. If a component A contains a single cycle on 3 vertices, the preceding argument implies that each vertex of the cycle is adjacent to at most one vertex of degree 1. But, unless $A = K_3$, A must have at least 6 vertices so $A - E(C)$ is again a forest of K_2's.

We now show that if A is a connected component of a tP_3-saturated graph G such that $|E(A)| = |V(A)|$, then $A = N_3$. By assumption, A contains a unique cycle. Clearly, $A \neq K_3$ since K_3 itself is not tP_3-saturated for any t and no edge added to K_3 can produce a new copy of P_3 that can't be replaced by an existing copy. Thus, A must take the form of a chordless cycle such that each vertex of the cycle is adjacent to a single pendant vertex. Since A is not complete, A itself must be $(r+1)P_3$-saturated for some $r \geq 1$. Thus, A contains exactly r copies of P_3. Since each copy of P_3 must use at least two vertices of the cycle, A must have either $4r$ or $4r+2$ vertices. (That is, the cycle is either C_{2r} or C_{2r+1}.) Label the vertices of C cyclically, $v_1, v_2, \cdots$. If $r \geq 2$, the edge v_1v_3 is not in C. But the graph $A + v_1v_3$ does not contain $(r+1)P_3$. Thus, $A = N_3$.

Consider the case when $t = 2$. The argument above implies that, if $G \in \underline{\mathbf{SAT}}(n, 2P_3)$ and G does not have a component isomorphic to N_3, then the component of G containing the unique copy of P_3 must have order less than 6 and would consequently be complete. But this forces $|E(G)| > (n+6)/2$, a contradiction. Thus Theorem **??** holds for $t = 2$.

Consider the case when $t = 3$. The argument above implies that, if $G \in \underline{\mathbf{SAT}}(n, 3P_3)$ and G does not have a component isomorphic to N_3, then the part of G containing copies of P_3 must lie in a single connected component, A, with between 9 and 11 vertices, such that $|E(A)| > |V(A)|$. But $|E(G)| < \lfloor \frac{n+6t-6}{2} \rfloor = \lfloor \frac{n+12}{2} \rfloor$ only if $|V(A)| = 9$ and $|E(A)| = 10$. A case analysis shows that no such graph is $3P_3$-saturated. Thus Theorem **??** holds for $t = 3$. □

The lemmas from Section 2 will now be used to prove Theorem **??**.

PROOF. (Theorem **??**) It is already known ([**?**]) and easily verified that $\frac{n}{2}P_2 \in \mathbf{SAT}(n, P_4)$ for n even and $K_3 \cup \frac{n-3}{2}P_2 \in \mathbf{SAT}(n, P_4)$ for n odd. If $G \in \mathbf{SAT}(n, P_4)$ has an isolated vertex, then all of the remaining components of G

would have to have at least 3 vertices, and so G would have at least $2(n-1)/3$ edges. This implies the previously defined graphs are minimal and this takes care of the case $t = 1$.

Consider the case $t = 2$. It is straightforward to verify that N_4, and also N_4^*, is $2P_4$-saturated. Also $N_4 \cup \frac{n-12}{2}P_2 \in \mathbf{SAT}(n, 2P_4)$ and $N_4^* + \frac{n-13}{2}P_2 \in \mathbf{SAT}(n, 2P_4)$ for n even and odd respectively. Assume $G \in \underline{\mathbf{SAT}}(n, 2P_4)$. Let H be the component of G that contains a P_4. Since $G - H$ must be P_4-saturated, all vertices except at most one must have degree at least 1. If H is complete, then $|H| \geq 6$, which would imply that $|E(G)| \geq 15 + (n-6)/2 = (n+24)/2$. Thus, we can assume that H is not complete, and so $|H| \geq 8$, since the addition of any edge in H will result in a $2P_4$. Thus, $G - H$ will be the disjoint union of independent edges if $|G - H|$ is even, and the disjoint union of a K_3 and independent edges if $|G - H|$ is odd. Lemma **??** implies that H is not a tree, and so $|E(H)| \geq |H|$. Also, Lemma **??** implies that any vertex of degree 2 must be on a triangle.

We will first consider the case when n is even. If $|H| = m \geq 12$, then $|E(G)| \geq m + \lfloor (n-m)/2 \rfloor \geq (n+12)/2$, thus we can assume that $8 \leq |H| < 12$. More specifically, if $|H| = 11$, then $|E(G)| \geq 11 + 3 + (n-14)/2 > (n+12)/2$. If $|H| = 10$ and $|E(H)| \geq 11$, then $|E(G)| \geq 11 + (n-10)/2 \geq (n+12)/2$. Therefore, if $|H| = 10$, then $|E(H)| = 10$. However, Lemma **??** and Lemma **??** imply that H is not $2P_4$-saturated. If $|H| = 9$ and $|E(H)| \geq 9$, then $|E(G)| \geq 9 + 3 + (n-12)/2 = (n+12)/2$. Therefore, we can assume that $|H| = 8$.

If $|H| = 8$ and $|E(H)| \geq 10$, then $|E(G)| \geq 10 + (n-8)/2 = (n+12)/2$. Therefore, we can assume that $|H| = 8$ and $|E(H)| = 8$ or 9. Lemmas **??**, **??** and **??** imply that $|E(H)| = 9$ and that the longest cycle C in H is a C_5. Assume that $|C| = 5$. If C has no chords, then H will contain a vertex of degree 2 not on a triangle, which contradicts Lemma **??**. If C contains a chord, then 2 of the remaining vertices in $H - C$ must be adjacent to the 2 vertices of degree 2 in C not on a triangle. Any possibility for the adjacency of the remaining vertex of $H - C$ will result in a $2P_4$, a vertex of degree 2 not on a triangle, or 2 vertices of degree 1 adjacent to the same vertex. Thus, H is not $2P_4$-saturated. If $|C| = 4$ and there is a chord e of C in H, then $H - \{E(C) \cup e\}$ is a forest. To avoid a vertex of degree 2 not on a triangle and 2 vertices of degree 1 adjacent to the same vertex, the remaining vertices of $H - C$ must be adjacent to distinct vertices of C, giving a $2P_4$. Thus, H is not $2P_4$-saturated. If $|C| = 4$, and there is no chord in C, then a vertex in $H - C$ is adjacent to 2 nonconsecutive vertices of C forming a $K_{2,3}$. Each of the 3 vertices of degree 2 in the $K_{2,3}$ will be incident to a pendant edge. It is easily checked, by adding a chord, that this graph H is not $2P_4$-saturated. If $|C| = 3$, then H contains either 2 vertex disjoint triangles T_1 and T_2 connected by an edge, or the triangles share a vertex. In the case of vertex disjoint triangles, any choice of the remaining 2 edges will result in a $2P_4$, a pair of vertices of degree 1 with a common adjacency, or 2 edges incident to distinct vertices of say T_1, with 1 of the these vertices incident to the edge between T_1 and T_2. It is easily checked that this last graph is not $2P_4$-saturated. In the case of triangles sharing a vertex, any choice of the remaining 3 edges will result in a $2P_4$, a pair of vertices of degree 1 with a common adjacency, or the 3 independent edges incident to the shared vertex of T_1 and T_2, and an additional vertex of each of the triangles. It is easily checked that this last graph is not $2P_4$-saturated.

We now consider the case when n is odd. If $|H| = m \geq 13$, then $|E(G)| \geq m + (n-m)/2 \geq (n+13)/2$. Thus, we can assume that $8 \leq |H| \leq 12$. If $|H| = 12$, then $|E(G)| \geq 12 + 3 + (n-15)/2 > (n+13)/2$. If $|H| = 11$ and $|E(H)| \geq 12$, then $|E(G)| \geq 12 + (n-11)/2 \geq (n+13)/2$. Therefore, if $|H| = 11$, then $|E(H)| = 11$. However, Lemma **??** and Lemma **??** imply that H is not $2P_4$-saturated. If $|H| = 10$ and $|E(H)| \geq 10$, then $|E(G)| \geq 10 + 3 + (n-13)/2 = (n+13)/2$. Next consider the case when $|H| = 8$. If $|E(H)| \geq 9$, then $|E(G)| \geq 9 + 3 + (n-11)/2 = (n+13)/2$. If $|E(H)| = 8$, then Lemma **??** and Lemma **??** imply that H is not $2P_4$-saturated. The only case remaining is $|H| = 9$. If $|E(H)| \geq 10$, then $|E(G)| \geq (n+13)/2$. If $|E(H)| = 9$, then Lemma **??** and Lemma **??** imply that H is not $2P_4$-saturated. Finally, if $|E(H)| = 10$, Lemmas **??** and **??** imply that the longest cycle C in H is a C_5. A similar case analysis by cycle length shows that no such $2P_4$-saturated graphs exist. This completes he proof of Theorem **??**. □

More can be said about saturation numbers involving P_3 if there are some copies of P_2 in the linear forest. Our next result illustrates this.

Theorem 4.5. *For n sufficiently large and $t \geq 0$,*

$\mathbf{sat}(n, tP_3 \cup 3P_2) = 3t + 6$ *and* $(t+2)K_3 \cup \overline{K}_{n-3(t+2)} \in \underline{\mathbf{SAT}}(n, tP_3 \cup 3P_2)$.

PROOF. Clearly for $t \geq 0$, $(t+2)K_3 \cup \overline{K}_{n-3(t+2)} \in \mathbf{SAT}(n, tP_3 \cup 3P_2)$, and so $\mathbf{sat}(n, tP_3 \cup 3P_2) \leq 3t + 6$. Let $G \in \underline{\mathbf{SAT}}(n, tP_3 \cup 3P_2)$. Let H be the subgraph of G of components with at least 3 vertices. So, $G = H \cup mK_2 \cup (n - 2m - |H|)K_1$, where $m \geq 0$.

We proceed by induction. If $t = 0$, then by Theorem **??**, $\mathbf{sat}(n, 3P_2) = 6$. If one of the components of H is a K_3, then $G - K_3 \in \mathbf{SAT}(n-3, (t-1)P_3 \cup 3P_2)$ and by induction $|E(G - K_3) \geq 3(t-1) + 6$, and the result follows. Thus it is sufficient assume no component of G is a K_3 and produce a contradiction.

We will begin by establishing some properties of H. By Lemma **??**, any vertex in H of degree 1 will be adjacent to a vertex of degree at least 3. Also, if two vertices u and v of degree 1 are adjacent to a vertex w, then the addition of any edge e between w and an isolated vertices of G will result in a $tP_3 \cup 3P_2$. Clearly, the edge e can be replaced by one of uw or vw, a contradiction. Thus, no two vertices of degree 1 in H are adjacent to the same vertex. Thus, no component of H is a tree and $|E(H)| \geq |V(H)|$.

Next we will show that any connected component of H, say A, such that $|E(A)| = |V(A)|$ must have a particular structure. Clearly A must have a unique cycle, C. The same argument that implies that no component of H is a tree also forces $A - E(C)$ to be a forest of K_2's if $|V(C)| \geq 4$. If $|V(C)| = 3$, then $A - E(C)$ is a forest of K_2's and K_1's.

However, if $|V(C)| \geq 4$ and we label the vertices on the cycle $v_1, v_2, \cdots$, then $A + v_1v_3$ is not $rP_3 \cup sP_2$-saturated for any r and s. Similiarly, if $|V(C)| = 3$ and $A - E(C)$ has only one or two K_2's, the graph A is not $(rP_3 \cup sP_2)$-saturated for any r and s. Thus, $A = N_3$.

Now we split the argument into cases according to the value of m, the number of K_2's in G.

If $m \geq 3$, then the addition of an edge e between vertices in $(n-2m-|H|)K_1$ should result in a $tP_3 \cup 3P_2$, but one of the edges in $3K_2$ could play that role, a contradiction. Hence, we can assume that $m < 3$.

If $m = 0$, then either H is a complete graph with exactly $3t + 5$ vertices or it is not complete with at least $3t + 6$ vertices. Clearly in each of these cases, $|E(H)| \geq 3t + 6$.

If $m = 1$, then $|H| \geq 3t + 4$. If H is complete, then $|E(G)| \geq \binom{3t+4}{2} + 1 \geq 3t + 6$. If H is not complete, then $|E(G)| \geq (3t + 5) + 1$, unless $|E(H)| = 4t + 4 = |V(H)|$. So H is the disjoint union of components isomorphic to N_3. For $t = 1$, we see that the graph $N_3 \cup K_2$ is not $(P_3 \cup 3P_2)$-saturated. If $t \geq 2$, then H would have to contain tN_3. Therefore, $|E(G)| \geq 6t + 1 \geq 3t + 6$.

If $m = 2$, then the addition of an edge between the two independent edges of $G - H$ will result in a P_3, and so H would have to contain at least $3(t-1) + 6 = 3t + 3$ vertices. Thus, if $|E(H)| > |H|$, then $|E(G)| \geq (3t + 4) + 2$, so we can assume that H would have to contain tN_3. For $t = 1$, the graph $N_3 \cup 2K_2$ already contains $P_3 \cup 3P_2$, and so is not saturated. If $t \geq 2$, then H would have to contain

tN_3, and so $|E(G)| \geq 6t + 2 \geq 3t + 6$. This completes the proof of Theorem **??**. $\square$

The previous theorem has recently been generalized in [**?**]

Theorem 4.6. *For n sufficiently large and $s \geq 3$,*

$$\mathbf{sat}(n, tP_3 \cup sP_2) = 3(s+t-1) \text{ and } (s+t-1)K_3 \cup \overline{K}_{n-3(s+t-1)} \in \underline{\mathbf{SAT}}(n, tP_3 \cup sP_2).$$

References

[BK11] N. Bushaw and N. Kettle, *Turán Numbers of Multiple Paths and Equibipartite Forests*, **Combinatorics, Probability, and Computing** 20 (2011) No. 6, 837-853.

[CLZ11] G. Chartrand, L. Lesniak, and P. Zhang **Graphs and Digraphs**, CRC Press, Boca Raton, FL, (2011).

[EHM64] P, Erdős, A. Hajnal, and J. W. Moon, *A Problem in Graph Theory*, **Amer. Math. Monthly** 71, (1964), 1107-1110.

[FFGJ09] J. R. Faudree, R. J. Faudree, R. J. Gould, and M. S. Jacobson, *Saturation Numbers for Trees*, **Electron. J. Combin.** 16 (2009), 19 pp.

[FFS] J. R. Faudree, R. J. Faudree, J. R. Schmitt, *A Survey of Minimum Saturated Graphs*, **Electron. J. Combin.** 18 (2011), 36 pp.

[JF] A. Jambulapati, R. J. Faudree, *A Collection of Results on Saturation Numbers*, preprint.

[KT86] L. Kászonyi, and Zs Tuza, *Saturated Graphs with Minimal Number of Edges*, **J. Graph Theory** 10, (1986), 203-210.

[P1891] J. Petersen, Die Theorie der Regularen Graphs, *Acta Math.* **15** (1891), 193-220.

[T41] Paul Turán, *Eine Extremalaufgabe aus der Graphentheorie*, **Mat. Fiz. Lapok, 48**, (1941) 436-452.

Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303

E-mail address: gchen@gsu.edu

Department of Mathematics Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-6660

E-mail address: jrfaudree@alaska.edu

Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152

E-mail address: rfaudree@memphis.edu

Department of Math and Computer Science, Emory University, Atlanta, GA 30322

E-mail address: rg@mathcs.emory.edu

Department of Mathematics, University of Colorado Denver, Denver, CO 80217

E-mail address: Michael.Jacobson@ucdenver.edu

Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460-8093

E-mail address: cmagnant@georgiasouthern.edu