# Path ideals of rooted trees and their graded Betti numbers

Rachelle R. Bouchat [a], Huy Tài Hà [b], Augustine O'Keefe [b]

[a] *Department of Mathematics, Slippery Rock University, 1 Morrow Way, Slippery Rock, PA 16057, United States*
[b] *Department of Mathematics, Tulane University, 6823 St. Charles Ave., New Orleans, LA 70118, United States*

**A B S T R A C T**

Let $\Gamma$ be a rooted (and directed) tree, and let $t$ be a positive integer. The path ideal $I_t(\Gamma)$ is generated by monomials that correspond to directed paths of length $(t-1)$ in $\Gamma$. In this paper, we study algebraic properties and invariants of $I_t(\Gamma)$. We give a recursive formula to compute the graded Betti numbers of $I_t(\Gamma)$ in terms of path ideals of subtrees. We also give a general bound for the regularity, explicitly compute the linear strand, and investigate when $I_t(\Gamma)$ has a linear resolution.

© 2011 Elsevier Inc. All rights reserved.

## 1. Introduction

The construction of edge ideals associated to (hyper)graphs (cf. [5,9]) provides a viewpoint complementary to the Stanley–Reisner correspondence in the study of monomial ideals. Edge ideals also provide a framework to study (hyper)graph theoretic questions from an algebraic perspective. Let $\Gamma = (V, E)$ be a finite, simple graph over the vertex set $V = \{x_1, \ldots, x_n\}$. Let $k$ be any field and identify the vertices in $V$ with the variables in the polynomial ring $S = k[x_1, \ldots, x_n]$. The edge ideal of $\Gamma$ is generated by monomials of the form $x_i x_j$, where $e = \{x_i, x_j\}$ is an edge in $\Gamma$. Note that an edge can be viewed as a path of length 1. Thus, for a given positive integer $t$, a more general construction is obtained by considering monomials corresponding to paths of length $(t-1)$ in $\Gamma$. This is the *path ideal* construction.

Path ideals were first introduced by Conca and De Negri in [3], and their algebraic properties have been investigated by various authors in the literature (cf. [2,3,6,7]). In this paper, we shall study path ideals of *rooted trees*. Recall that a *tree* is a graph in which there exists a unique path between every pair of distinct vertices; a *rooted tree* is a tree together with a fixed vertex called the *root*. In particular, in a rooted tree there exists a unique path from the root to any given vertex. We can also view a rooted tree as a *directed* graph by assigning to each edge the direction that goes "away" from

the root. *Throughout this paper, a rooted tree will always be viewed as a directed, rooted tree in this sense.* If $\{x_i, x_j\}$ is an edge in a rooted tree $\Gamma$, then we write $(x_i, x_j)$ for the "directed" edge whose direction is from $x_i$ to $x_j$. The path ideal of a rooted tree is defined in precise form as follows.

**1.1. Definition.** Let $t \geqslant 1$ be a given integer, and let $\Gamma$ be a rooted tree with vertex set $V = \{x_1, \ldots, x_n\}$.

(1) A directed *path* of length $(t-1)$ is a sequence of distinct vertices $x_{i_1}, \ldots, x_{i_t}$, in which $(x_{i_j}, x_{i_{j+1}})$ is the directed edge from $x_{i_j}$ to $x_{i_{j+1}}$ for any $j = 1, \ldots, t-1$.
(2) The *path ideal* of length $(t-1)$ associated to $\Gamma$ is the monomial ideal

$$I_t(\Gamma) := (x_{i_1} \cdots x_{i_t} \mid x_{i_1}, \ldots, x_{i_t} \text{ is a path in } \Gamma) \subset S = k[x_1, \ldots, x_n].$$

In particular, when $t = 1$, $I_1(\Gamma) = (x_1, \ldots, x_n)$ is the maximal homogeneous ideal, which is well understood. Hence, all our results in this paper will be for path ideals of length at least 1 (i.e., $t \geqslant 2$).

Due to the correspondence between paths and monomials we shall often abuse notation and use $x_{i_1} \cdots x_{i_t}$ to denote both the monomial $x_{i_1} \cdots x_{i_t}$ in $k[x_1, \ldots, x_n]$ and the path $x_{i_1}, \ldots, x_{i_t}$ in $\Gamma$.

A *rooted forest* is a disjoint union of rooted trees. For a rooted forest $\Delta$, we define the path ideal $I_t(\Delta)$ to be the sum of the path ideals of the connected components of $\Delta$.

A path ideal $I_t(\Gamma)$ is a squarefree monomial ideal, so it can also be realized as the edge ideal of a hypergraph or the Stanley–Reisner ideal of a simplicial complex. Given a path ideal, the corresponding hypergraph and simplicial complex are in general very complicated. The goal of this paper is then to investigate algebraic properties and invariants of a path ideal $I_t(\Gamma)$ via the combinatorial structures of the rooted tree $\Gamma$. We are interested in invariants associated to the minimal free resolution of $I_t(\Gamma)$, namely, the graded Betti numbers, the Castelnuovo–Mumford regularity, and the projective dimension.

We now provide an overview of the structure of the paper and our results. In Section 2, we recall some useful notation and terminology, and prove our first main result; here we give a recursive formula to compute the graded Betti numbers of path ideals (Theorem 2.7 and Remark 2.9). Section 3 is devoted to studying the Castelnuovo–Mumford regularity of path ideals. The main result of this section, Theorem 3.4, provides a general bound for the regularity of a path ideal in terms of the number of leaves and the number of pairwise disjoint paths of length $t$ in the tree. In Section 4, we study the linear strand of $I_t(\Gamma)$ and classify all rooted trees $\Gamma$ for which $I_t(\Gamma)$ has a linear resolution. Our first result of this section, Theorem 4.2, gives a precise formula for graded Betti numbers $\beta_{i,i+t}(I_t(\Gamma))$ on the linear strand of $I_t(\Gamma)$. Our next result in this section, Theorem 4.5, shows that $I_t(\Gamma)$ has a linear resolution if and only if it has linear first syzygies; this is the case if any only if $\Gamma$ belongs to a special class of rooted trees, which we will call broom graphs. In Section 5, we restrict our attention to rooted trees occurring as path graphs. For a path graph $\Gamma$, in Theorem 5.3, we characterize which graded Betti numbers of $I_t(\Gamma)$ are nonzero. As a consequence, we compute the regularity of $S/I_t(\Gamma)$ explicitly in Corollary 5.4. We also recover He and Van Tuyl's formula for the projective dimension of $S/I_t(\Gamma)$ in this case (Corollary 5.1).

## 2. Path ideals and graded Betti numbers

From this point forward, $\Gamma$ will denote a rooted tree (also viewed as a directed tree) with vertex set $V = \{x_1, \ldots, x_n\}$, $k$ will denote a field of arbitrary characteristic, and $t$ will denote a given positive integer. Then $S = k[x_1, \ldots, x_n]$ will denote the corresponding polynomial ring.

### 2.1. Induced subgraphs and examples

We will now introduce some combinatorial terminology and provide examples of path ideals.

**2.1. Definition.** Let $\Gamma$ be a rooted tree with root $x$. For a given vertex $y$ in $\Gamma$, the *level* of $y$, denoted level($y$), is defined to be the length of the unique path from $x$ to $y$. The *height* of $\Gamma$, denoted height($\Gamma$), is the maximal level of vertices in $\Gamma$.

Sometimes we will need to consider rooted forests. The level of a vertex $y$ in a rooted forest $\Delta$ is defined to be the level of $y$ inside the connected component of $\Delta$ containing $y$. The height of a rooted forest $\Delta$ is defined to be the largest height of its connected components.

**2.2. Example.** Consider the following rooted tree $\Gamma$.
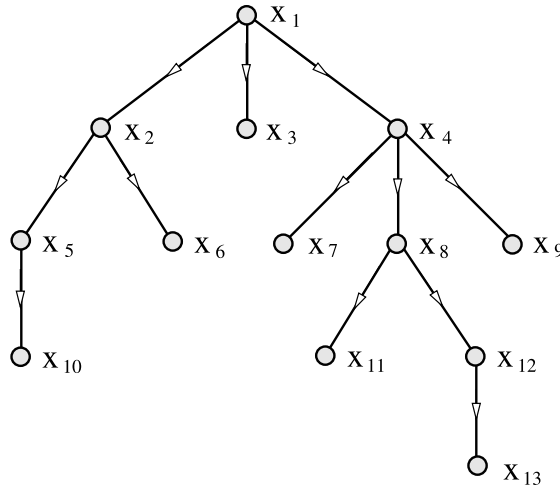


The edges in $\Gamma$ are given directions that go "away" from the root, making $\Gamma$ a directed tree. For instance, there is a unique path $x_1, x_2$ going from the root $x_1$ to the vertex $x_2$, and a unique path $x_1, x_2, x_6$ going from the root $x_1$ to the vertex $x_6$; and so, the direction of the edge $\{x_2, x_6\}$ is from $x_2$ to $x_6$. It can also be seen that the highest level in $\Gamma$ is 4 (level$(x_{13}) = 4$), so height$(\Gamma) = 4$.

For $t = 2, 3, 4$, and 5, we have the following path ideals associated to $\Gamma$:

$$I_2(\Gamma) = (x_1x_2, x_1x_3, x_1x_4, x_2x_5, x_2x_6, x_4x_7, x_4x_8, x_4x_9, x_5x_{10}, x_8x_{11}, x_8x_{12}, x_{12}x_{13}),$$

$$I_3(\Gamma) = (x_1x_2x_5, x_1x_2x_6, x_1x_4x_7, x_1x_4x_8, x_1x_4x_9, x_2x_5x_{10}, x_4x_8x_{11}, x_4x_8x_{12}, x_8x_{12}x_{13}),$$

$$I_4(\Gamma) = (x_1x_2x_5x_{10}, x_1x_4x_8x_{11}, x_1x_4x_8x_{12}, x_4x_8x_{12}x_{13}),$$

$$I_5(\Gamma) = (x_1x_4x_8x_{12}x_{13}).$$

Notice that the path ideals of $\Gamma$ depend on the choice of the root of $\Gamma$.

**2.3. Definition.** Let $\Gamma$ be a rooted tree, and let $x$ be a vertex in $\Gamma$.

(1) A vertex $z$ in $\Gamma$ is the *parent* of $x$ if and only if $(z, x)$ is a directed edge in $\Gamma$. A vertex $y$ is called a *child* of $x$ if $(x, y)$ is a directed edge in $\Gamma$.
(2) A vertex $z \neq x$ is an *ancestor* of $x$ if there is a path from $z$ to $x$. A vertex $y \neq x$ is a *descendant* of $x$ if there is a path from $x$ to $y$.
(3) The vertex $x$ is called a *leaf* of $\Gamma$ if $x$ has no descendants.
(4) The vertex $x$ is called the *root* of $\Gamma$ if $x$ has no ancestors.
(5) The *degree* of a vertex $x$ in $\Gamma$, denoted by $\deg_\Gamma(x)$, is the number of edges in $\Gamma$ incident to $x$.

**2.4. Definition.**

(1) Let $G$ be a finite simple graph. A subgraph $H$ of $G$ is called an *induced subgraph* if for every pair of vertices $x, y$ in $H$ the following condition holds: if $\{x, y\}$ is an edge in $G$, then it is also an edge in $H$.

(2) Let $\Gamma$ be a rooted tree. An *induced subtree* (or *forest*) of $\Gamma$ is a directed subtree (or forest) that is also an induced subgraph of $\Gamma$.
(3) Let $\Gamma$ be a rooted tree and let $x$ be a vertex in $\Gamma$. The *induced subtree rooted at $x$* of $\Gamma$ is the induced subtree of $\Gamma$ over the vertex set $\{x\} \cup \{y \mid y$ is a descendant of $x\}$ (with $x$ considered as its root).

**Notation.** Let $\Gamma$ be a rooted tree, and let $P$ be a collection of vertices in $\Gamma$. We shall denote by $\Gamma \backslash P$ the induced subforest of $\Gamma$ obtained by removing the vertices in $P$ and the edges incident to these vertices. If $P$ consists of a single element $x$, then we write $\Gamma \backslash x$ for $\Gamma \backslash \{x\}$.

*2.2. Minimal free resolutions*

Let $S = k[x_1, \ldots, x_n]$, and let $M$ be a finitely generated graded $S$-module. Associated to $M$ is a *minimal free resolution*, which is a finite complex of the form

$$0 \to \bigoplus_j S(-j)^{\beta_{p,j}(M)} \xrightarrow{\delta_p} \bigoplus_j S(-j)^{\beta_{p-1,j}(M)} \xrightarrow{\delta_{p-1}} \cdots \xrightarrow{\delta_1} \bigoplus_j S(-j)^{\beta_{0,j}(M)} \to M \to 0$$

where the maps $\delta_i$ are exact and where $S(-j)$ denotes the translation of $S$ obtained by shifting the degrees of elements of $S$ by $j$. The numbers $\beta_{i,j}(M)$ are called the *graded Betti numbers* of $M$, and they provide the number of minimal generators of degree $j$ occurring in the $i$th-syzygy module of $M$.

If $M$ is generated in degree $t$, then the *linear strand* of $M$ is given by the Betti numbers $\beta_{i,i+t}(M)$, for $i > 0$. In this case, $M$ is said to have linear first syzygies if $\beta_{1,j}(M) = 0$ for all $j \neq 1 + t$; and more generally, $M$ is said to have a *linear resolution* if $\beta_{i,j}(M) = 0$ for all $i > 0$ and $j \neq i + t$.

We are interested in the two following invariants that measure the "size" of the minimal free resolution.

**2.5. Definition.** Let $S$ and $M$ be as above.

(1) The *projective dimension* of $M$, denoted by $\mathrm{pd}(M)$, is the length of the minimal free resolution associated to $M$.
(2) The *Castelnuovo–Mumford regularity* (or simply, *regularity*), denoted by $\mathrm{reg}(M)$, is a measure of the width of the minimal free resolution of $M$ and is defined as

$$\mathrm{reg}(M) := \max\{ j - i \mid \beta_{i,j}(M) \neq 0 \}.$$

**2.6. Lemma.** *Let $S = k[x_1, \ldots, x_n]$, and let $M$ be a graded $S$-module. Let $y_1, \ldots, y_m$ be indeterminates, and denote by $R$ the polynomial ring $k[x_1, \ldots, x_n, y_1, \ldots, y_m]$. Then*

$$\mathrm{reg}(M) = \mathrm{reg}(M \otimes_S R),$$

*where the second regularity is computed for the $R$-module $M \otimes_S R$.*

**Proof.** It is clear that the ring extension $S \to R$ is flat. Thus, tensoring with $R$ maps a minimal free resolution to a minimal free resolution. The result now follows from the definition of regularity. □

Lemma 2.6 allows us to look at extensions of ideals in rings with more variables when discussing regularity. For instance, if $\Gamma$ is a rooted tree corresponding to the polynomial ring $R$ and $\Delta$ is an induced, rooted subtree of $\Gamma$ corresponding to the polynomial ring $S$ (i.e. the variables in $R$ correspond to the vertices in $\Gamma$, and the variables in $S$ correspond to the vertices in $\Delta$), then we can abuse notation and write $I_t(\Delta)$ for both the path ideal of $\Delta$ in $S$ and also for its extension in the bigger ring $R$ when discussing regularity.

*2.3. Mapping cone decomposition of path ideals*

To study the minimal free resolutions of the quotient rings $S/I_t(\Gamma)$, we provide an inductive construction of the path ideals via the mapping cone construction. This construction is a generalization of the method provided for edge ideals in [1]. This method will allow the decomposition of a given path ideal into a collection of simpler path ideals corresponding to smaller trees.

Given a short exact sequence

$$0 \longrightarrow M_1 \longrightarrow M_2 \longrightarrow M_3 \longrightarrow 0$$

where $M_1$, $M_2$, and $M_3$ are graded $S$-modules, the mapping cone is a method to construct a free resolution for $M_3$ knowing free resolutions of $M_1$ and $M_2$ (for more details on the mapping cone construction we refer the reader to [10]). In general, given *minimal* free resolutions for $M_1$ and $M_2$, the mapping cone construction does not necessarily give a *minimal* free resolution of $M_3$. However, in the case of path ideals, we shall show that the mapping cone construction does indeed provide a minimal free resolution for a particular short exact sequence.

**2.7. Theorem.** *Let $\Gamma$ be a rooted tree with vertex set $V = \{x_1, \ldots, x_n\}$ and height $h \geqslant t - 1$. Let $x_{i_t}$ denote a leaf of $\Gamma$ of level at least $(t-1)$. Then by letting $x_{i_1}, \ldots, x_{i_t}$ denote the path terminating at $x_{i_t}$, the mapping cone procedure applied to the sequence*

$$0 \to \left(S/I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t})\right)(-t) \xrightarrow{x_{i_1} \cdots x_{i_t}} S/I_t(\Gamma \backslash x_{i_t}) \longrightarrow S/I_t(\Gamma) \to 0$$

*provides a minimal free resolution of $S/I_t(\Gamma)$. In particular, for any $i$ and $j$, we have*

$$\beta_{i,j}\left(S/I_t(\Gamma)\right) = \beta_{i,j}\left(S/I_t(\Gamma \backslash x_{i_t})\right) + \beta_{i-1, j-t}\left(S/I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t})\right).$$

**Proof.** Since $x_{i_t}$ does not divide a minimal generator of $I_t(\Gamma \backslash x_{i_t})$,

$$I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t}) = I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_{t-1}}).$$

However, this implies that the exact sequence

$$0 \longrightarrow \left(S/I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t})\right)(-t) \xrightarrow{x_{i_1} \cdots x_{i_t}} S/I_t(\Gamma \backslash x_{i_t}) \longrightarrow S/I_t(\Gamma) \to 0 \qquad (2.1)$$

factors as

$$0 \longrightarrow (S/I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t}))(-t) \xrightarrow{x_{i_1} \cdots x_{i_t}} S/I_t(\Gamma \backslash x_{i_t}) \longrightarrow S/I_t(\Gamma) \longrightarrow 0. $$

$$\Big\downarrow {\scriptstyle x_{i_t}} \qquad \nearrow {\scriptstyle x_{i_1} \cdots x_{i_{t-1}}} \qquad\qquad\qquad\qquad\qquad (2.2)$$

$$(S/I_t(\Gamma \backslash x_{i_t}) : ((x_{i_1} \cdots x_{i_{t-1}})))(-t+1)$$

Let

$$0 \longrightarrow \cdots \xrightarrow{\phi_2} F_1 \xrightarrow{\phi_1} F_0 = S \xrightarrow{\phi_0} S/I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t}) \longrightarrow 0, \quad \text{and} \qquad (2.3)$$

$$0 \longrightarrow \cdots \xrightarrow{\psi_2} G_1 \xrightarrow{\psi_1} G_0 = S \xrightarrow{\psi_0} S/I_t(\Gamma \backslash x_{i_t}) \longrightarrow 0 \qquad (2.4)$$

be minimal free resolutions of $(S/I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t}))$ and $S/I_t(\Gamma \backslash x_{i_t})$ respectively. Then the mapping cone construction applied to the short exact sequence (2.1) provides a free resolution of $S/I_t(\Gamma)$ given by

$$0 \longrightarrow \cdots \xrightarrow{\sigma_3} G_2 \oplus F_1(-t) \xrightarrow{\sigma_2} G_1 \oplus S(-t) \xrightarrow{\sigma_1} S \xrightarrow{\sigma_0} S/I_t(\Gamma) \longrightarrow 0,$$

where the maps $\sigma_i$ are defined by $\sigma_1 = [\psi_1 - \delta_0]$ and

$$\sigma_i = \begin{bmatrix} \psi_i & (-1)^i \delta_{i-1} \\ 0 & \phi_{i-1} \end{bmatrix} \quad \text{for } i > 1 \qquad (2.5)$$

$(\delta_i : F_i(-t) \to G_i$ are resulted from the homomorphism $(S/I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t}))(-t) \xrightarrow{x_{i_1} \cdots x_{i_t}}$ $S/I_t(\Gamma \backslash x_{i_t}))$.

From the factorization given in (2.2), the entries of the matrix of $\delta_i$ are not units. Furthermore, since (2.3) and (2.4) are minimal free resolutions, the matrix representation of $\sigma_i$ in (2.5) cannot contain units. Therefore, the mapping cone construction applied to (2.1) provides a minimal free resolution of $S/I_t(\Gamma)$. In particular, this implies that

$$\beta_{i,j}\big(S/I_t(\Gamma)\big) = \beta_{i,j}\big(S/I_t(\Gamma \backslash x_{i_t})\big) + \beta_{i-1,j-t}\big(S/I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t})\big)$$

for all $i$ and $j$.  □

Theorem 2.7 provides an inductive method to study algebraic properties of $I_t(\Gamma)$ as the colon ideal $I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t})$ can be realized as a disjoint union of path ideals of varying lengths.

**2.8. Lemma.** *Let $\Gamma$ be a rooted tree of height $h \geqslant t - 1$, let $x_{i_t}$ be a leaf at the highest level in $\Gamma$, and let $x_{i_1}, \ldots, x_{i_t}$ be the unique path of length $(t-1)$ terminating at $x_{i_t}$. Let $x_{i_0}$ be the only parent of $x_{i_1}$, if it exists. For $j = 0, \ldots, t$, let $\Gamma_j$ be the induced subtree of $\Gamma$ rooted at $x_{i_j}$, and let $\Delta_j = \Gamma_j \backslash \Gamma_{j+1}$ for $j = 0, \ldots, t - 1$. Then*

$$I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t}) = I_t\big(\Gamma \backslash \{x_{i_0}, \ldots, x_{i_t}\}\big) + (x_{i_0}) + \sum_{j=0}^{t-1} I_{t-j}\big(\Delta_j \backslash \{x_{i_0}, \ldots, x_{i_t}\}\big).$$

**Proof.** Let $G$ be the set of minimal generators of $I_t(\Gamma \backslash x_{i_t})$, i.e. elements in $G$ corresponding to paths of length $(t-1)$ in $\Gamma \backslash x_{i_t}$. Clearly,

$$I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t}) = \sum_{Q \in G} (Q) : (x_{i_1} \cdots x_{i_t}).$$

Observe first that $Q_1 = x_{i_0} x_{i_1} \cdots x_{i_{t-1}}$ is a path of length $(t-1)$ in $\Gamma \backslash x_{i_t}$, and $(Q_1) : (x_{i_1} \cdots x_{i_{t-1}}) = (x_{i_0})$. Consider a path $Q$ of length $(t-1)$ in $\Gamma \backslash x_{i_t}$ that does not contain $x_{i_0}$. There are three possibilities for $Q$.

**Case 1:** $Q$ contains none of the vertices in $\{x_{i_0}, \ldots, x_{i_t}\}$, and $Q$ is not a path in the induced subtree rooted at $x_{i_0}$. This is the case if and only if $(Q) : (x_{i_1} \cdots x_{i_t}) = (Q) \subseteq I_t(\Gamma \backslash \{x_{i_0}, \ldots, x_{i_t}\})$.

**Case 2:** $Q$ contains none of the vertices in $\{x_{i_0}, \ldots, x_{i_t}\}$, and $Q$ is a path in the induced subtree rooted at $x_{i_0}$. This is the case if and only if $Q$ is a path of length $(t-1)$ in the rooted forest $\Gamma_0 \backslash \{x_{i_0}, \ldots, x_{i_t}\}$.

**Case 3:** $Q$ contains some but not all the vertices $\{x_{i_1}, \ldots, x_{i_{t-1}}\}$, and $Q$ does not contain $x_{i_0}$. Let $s$ be the largest index such that $x_{i_s}$ is in $Q$. Since $x_{i_t}$ is a leaf of highest level in $\Gamma$, $Q$ can contain at most $t - s$ descendants of $x_{i_s}$. This implies that $Q$ must contain all the vertices $x_{i_1}, \ldots, x_{i_s}$. This is the case only if $Q \backslash \{x_{i_1}, \ldots, x_{i_s}\}$ is a path of length $t - s - 1$ in $\Gamma_s \backslash \{x_{i_0}, \ldots, x_{i_t}\}$. Furthermore, because $x_{i_t}$ is of highest level in $\Gamma$, any path of length $(t-s-1)$ in $\Gamma_s$ must be from a child of $x_{i_s}$ (other than $x_{i_{s+1}}$) to a leaf in $\Gamma_s$. Thus, Case 3 appears if and only if $(Q) : (x_{i_1} \cdots x_{i_t}) \subseteq I_{t-s}(\Delta_s \backslash \{x_{i_0}, \ldots, x_{i_t}\})$.  □

**2.9. Remark.** Note that in Lemma 2.8, $I_t(\Gamma \backslash \{x_{i_0}, \ldots, x_{i_t}\}) = I_t(\Gamma \backslash \Gamma_0)$ (or $I_t(\Gamma \backslash \Gamma_1)$ if $x_{i_0}$ does not exist) since $x_{i_t}$ is a leaf at the highest level in $\Gamma$. Observe further that the minimal generators of the ideals $I_t(\Gamma \backslash \{x_{i_0}, \ldots, x_{i_t}\})$, $(x_{i_0})$, and $I_{t-j}(\Delta_j \backslash \{x_{i_0}, \ldots, x_{i_t}\})$ involve pairwise disjoint sets of vertices. Thus, the minimal free resolution of $S/[I_t(\Gamma \backslash \{x_{i_0}, \ldots, x_{i_t}\}) + (x_{i_0}) + \sum_{j=0}^{t-1} I_{t-j}(\Delta_j \backslash \{x_{i_0}, \ldots, x_{i_t}\})]$ is obtained by taking the tensor product of the minimal free resolutions of $S/I_t(\Gamma \backslash \{x_{i_0}, \ldots, x_{i_t}\})$, $S/(x_{i_0})$, and $S/I_{t-j}(\Delta_j \backslash \{x_{i_0}, \ldots, x_{i_t}\})$ for $j = 0, \ldots, t-1$. Together with Theorem 2.7, this gives a recursive formula to compute the graded Betti numbers of $I_t(\Gamma)$. In particular, the graded Betti numbers of $I_t(\Gamma)$ do not depend on the characteristic of the ground field $k$. This fact was proved in [6, Theorem 3.1]. It is also a corollary of a more general recursive formula for the graded Betti numbers of simplicial forests given in [4, Theorem 5.8].

## 3. Regularity of path ideals

In this section, we give a bound for the regularity of $I_t(\Gamma)$. From the Alexander duality (cf. [8, Theorem 5.59]), one obtains the following trivial bound $\operatorname{reg}(S/I_t(\Gamma)) = \operatorname{reg}(I_t(\Gamma)) - 1 = \operatorname{pd}(S/I_t(\Gamma)^\vee) - 1 \leqslant n - 1$, where $I_t(\Gamma)^\vee$ is the Alexander dual of $I_t(\Gamma)$. We are seeking a bound for $\operatorname{reg}(S/I_t(\Gamma))$ that is, in general, better than $n - 1$. Our bound will be based on the number of leaves and the number of pairwise disjoint paths of length $t$ in $\Gamma$.

**3.1. Definition.** Let $\Gamma$ be a rooted tree. We define $l_t(\Gamma)$ to be the number of leaves in $\Gamma$ whose level is at least $t - 1$ and $p_t(\Gamma)$ to be the maximal number of pairwise disjoint paths of length $t$ in $\Gamma$ (i.e., $p_t(\Gamma) = \max\{|D| \mid D \text{ is a set of disjoint paths of length } t \text{ in } \Gamma\}$). Note that, in general, $t l_t(\Gamma) \ll n$ and $t p_t(\Gamma) \ll n$.

In the next few corollaries, $\Gamma$ will denote a rooted tree of height $h \geqslant t - 1$, and $x_{i_t}$ will denote a leaf of highest level in $\Gamma$. Let $x_{i_1}, \ldots, x_{i_t}$ be the unique path of length $(t - 1)$ terminating at $x_{i_t}$, and let $x_{i_0}$ be the parent of $x_{i_1}$ (if it exists). Set $P = \{x_{i_0}, \ldots, x_{i_t}\}$ (or $\{x_{i_1}, \ldots, x_{i_t}\}$ if $x_{i_0}$ does not exist). Furthermore for $j = 0, \ldots, t - 1$, let $\Gamma_j$ be the induced subtree of $\Gamma$ rooted at $x_{i_j}$, and let $\Delta_j = \Gamma_j \backslash \Gamma_{j+1}$.

**3.2. Corollary.** *We have*

$$\operatorname{reg}\big(S/\big(I_t(\Gamma\backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t})\big)\big) = \operatorname{reg}\big(S/I_t(\Gamma\backslash P)\big) + \sum_{j=0}^{t-1} \operatorname{reg}\big(S/I_{t-j}(\Delta_j\backslash P)\big).$$

**Proof.** It is easy to see that for $j = 0, \ldots, t - 1$ the minimal generators of the ideals $I_t(\Gamma\backslash P)$, $(x_{i_0})$, and $I_{t-j}(\Delta_j\backslash P)$ involve pairwise disjoint sets of vertices. Thus, by Lemma 2.8, the minimal free resolution of $S/(I_t(\Gamma\backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t}))$ is given by the tensor product of the minimal free resolution of $S/I_t(\Gamma\backslash P)$, $S/(x_{i_0})$, and $S/I_{t-j}(\Delta_j\backslash P)$ for $j = 0, \ldots, t - 1$. This implies that

$$\operatorname{reg}\big(S/\big(I_t(\Gamma\backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t})\big)\big) = \operatorname{reg}\big(S/I_t(\Gamma\backslash P)\big) + \operatorname{reg}\big(S/(x_{i_0})\big) + \sum_{j=0}^{t-1} \operatorname{reg}\big(S/I_{t-j}(\Delta_j\backslash P)\big).$$

The conclusion now follows from the fact that $\operatorname{reg}(S/(x_{i_0})) = 0$.   □

**3.3. Corollary.** *We have*

$$\operatorname{reg}\big(S/I_t(\Gamma)\big) = \max\bigg\{\operatorname{reg}\big(S/I_t(\Gamma\backslash x_{i_t})\big), \operatorname{reg}\big(S/I_t(\Gamma\backslash P)\big)$$

$$+ \sum_{j=0}^{t-1} \operatorname{reg}\big(S/I_{t-j}(\Delta_j\backslash P)\big) + (t - 1)\bigg\}.$$

*In particular, by considering $\Gamma_0$ in place of $\Gamma$ we have*

$$\operatorname{reg}\big(S/I_t(\Gamma_0)\big) = \max\bigg\{\operatorname{reg}\big(S/I_t(\Gamma_0\backslash x_{i_t})\big), \sum_{j=0}^{t-1} \operatorname{reg}\big(S/I_{t-j}(\Delta_j\backslash P)\big) + (t - 1)\bigg\}.$$

*If $x_{i_0}$ does not exist, then*

$$\operatorname{reg}\big(S/I_t(\Gamma_1)\big) = \max\bigg\{\operatorname{reg}\big(S/I_t(\Gamma_1\backslash x_{i_t})\big), \sum_{j=1}^{t-1} \operatorname{reg}\big(S/I_{t-j}(\Delta_j\backslash P)\big) + (t - 1)\bigg\}.$$

**Proof.** It follows from Theorem 2.7 that

$$\operatorname{reg}\big(S/I_t(\Gamma)\big) = \max\big\{\operatorname{reg}\big(S/I_t(\Gamma\backslash x_{i_t})\big), \operatorname{reg}\big(S/I_t(\Gamma\backslash x_{i_t}) : (x_{i_1}\cdots x_{i_t})\big) + (t-1)\big\}.$$

The first conclusion follows by applying Corollary 3.2. The second conclusion follows by observing that since $x_{i_t}$ is a leaf at the highest level, we have $I_t(\Gamma_0\backslash P) = (0)$ (or $I_t(\Gamma_1\backslash P) = (0)$ if $x_{i_0}$ does not exist). $\square$

We are ready to prove our next theorem.

**3.4. Theorem.** *Let $\Gamma$ be a rooted tree over the vertex set $V = \{x_1, \ldots, x_n\}$. Then*

$$\operatorname{reg}\big(S/I_t(\Gamma)\big) \leqslant (t-1)\big[l_t(\Gamma) + p_t(\Gamma)\big].$$

**Proof.** We shall use induction on both $t$ and $n$. For $t = 1$, the ideal $I_t(\Gamma)$ is the maximal homogeneous ideal of $S = k[x_1, \ldots, x_n]$, and the assertion is clearly true. Assume that $t \geqslant 2$. The assertion is also true if $n \leqslant t$, so we may assume that $n > t$.

Let $h = \operatorname{height}(\Gamma)$. Observe that if $h < t - 1$ then $I_t(\Gamma) = (0)$, making the assertion vacuous. We shall assume that $h \geqslant t - 1$. Consider first the case when $h = t - 1$. In this case, any path of length $(t-1)$ in $\Gamma$ must be from the root to a leaf (at level $(t-1)$) of $\Gamma$ and so $p_t(\Gamma) = 0$. Without loss of generality, assume that $x_1$ is the root of $\Gamma$. Then $I_t(\Gamma) = x_1 I_{t-1}(\Gamma\backslash x_1)$ and $p_{t-1}(\Gamma\backslash x_1) = 0$. By the induction hypothesis, we have

$$\operatorname{reg}\big(S/I_{t-1}(\Gamma\backslash x_1)\big) \leqslant (t-2)\big[l_{t-1}(\Gamma\backslash x_1) + p_{t-1}(\Gamma\backslash x_1)\big] = (t-2)l_{t-1}(\Gamma\backslash x_1).$$

Observe further that $l_{t-1}(\Gamma\backslash x_1) = l_t(\Gamma) \geqslant 1$ (since $h = t - 1$, $\Gamma$ must have at least a leaf at level $(t-1)$). Therefore, we have

$$\operatorname{reg}\big(S/I_t(\Gamma)\big) = \operatorname{reg}\big(S/I_{t-1}(\Gamma\backslash x_1)\big) + 1 \leqslant (t-2)l_{t-1}(\Gamma\backslash x_1) + 1 \leqslant (t-1)l_t(\Gamma),$$

and the assertion is true.

Consider now the case when $h \geqslant t$. Let $x_t$ be a leaf at the highest level, and let $x_{i_0}, \ldots, x_{i_t}$ be the unique path of length $t$ terminating at $x_{i_t}$. Let $P = \{x_{i_0}, \ldots, x_{i_t}\}$, let $\Gamma_j$ be the induced subtree of $\Gamma$ rooted at $x_{i_j}$, and let $\Delta_j = \Gamma_j\backslash\Gamma_{j+1}$ for $j = 0, \ldots, t - 1$. It follows from Corollary 3.3 that

$$\operatorname{reg}\big(S/I_t(\Gamma)\big) = \max\Bigg\{ \operatorname{reg}\big(S/I_t(\Gamma\backslash x_{i_t})\big), \operatorname{reg}\big(S/I_t(\Gamma\backslash P)\big)$$

$$+ \sum_{j=0}^{t-1} \operatorname{reg}\big(S/I_{t-j}(\Delta_j\backslash P)\big) + (t-1) \Bigg\}. \tag{3.1}$$

Observe that $l_t(\Gamma\backslash x_{i_t}) \leqslant l_t(\Gamma)$ and $p_t(\Gamma\backslash x_{i_t}) \leqslant p_t(\Gamma)$. Thus, by the induction hypothesis, we have

$$\operatorname{reg}\big(S/I_t(\Gamma\backslash x_{i_t})\big) \leqslant (t-1)\big[l_t(\Gamma\backslash x_{i_t}) + p_t(\Gamma\backslash x_{i_t})\big] \leqslant (t-1)\big[l_t(\Gamma) + p_t(\Gamma)\big].$$

It can also be seen that $l_t(\Gamma\backslash P) = l_t(\Gamma\backslash\Gamma_0) \leqslant l_t(\Gamma) - l_t(\Gamma_0) + 1$ and $\sum_{j=0}^{t-1} l_{t-j}(\Delta_j\backslash P) = l_t(\Gamma_0) - 1$. Thus, by the induction hypothesis, we have

$$\operatorname{reg}\big(S/I_t(\Gamma\backslash P)\big) + \sum_{j=0}^{t-1} \operatorname{reg}\big(S/I_{t-j}(\Delta_j\backslash P)\big) + t - 1$$

$$\leqslant (t-1)\Bigg[ l_t(\Gamma\backslash\Gamma_0) + p_t(\Gamma\backslash\Gamma_0) + \sum_{j=0}^{t-1}\big(l_{t-j}(\Delta_j\backslash P) + p_{t-j}(\Delta_j\backslash P)\big) \Bigg] + (t-1)$$

$$= (t-1)l_t(\Gamma) + (t-1)\Bigg[ p_t(\Gamma\backslash\Gamma_0) + \sum_{j=0}^{t-1} p_{t-j}(\Delta_j\backslash P) \Bigg] + (t-1).$$

Moreover, $p_t(\Gamma\backslash\Gamma_0) \leqslant p_t(\Gamma) - p_t(\Gamma_0)$, $p_t(\Gamma_0) = 1$, and $p_{t-j}(\Delta_j\backslash P) = 0$ (because height$(\Delta_j\backslash P) \leqslant t - j - 1$ for any $j$). Hence,

$$\text{reg}\big(S/I_t(\Gamma\backslash P)\big) + \sum_{j=0}^{t-1} \text{reg}\big(S/I_{t-j}(\Delta\backslash P)\big) + t - 1$$

$$\leqslant (t-1)l_t(\Gamma) + (t-1)\big[p_t(\Gamma) - 1\big] + (t-1)$$

$$\leqslant (t-1)\big[l_t(\Gamma) + p_t(\Gamma)\big].$$

The theorem is proved by the use of (3.1). $\quad\square$

**3.5. Remark.** The bound in Theorem 3.4 is sharp when $\Gamma$ is a disjoint union of paths of length $(t-1)$. For instance, if $\Gamma$ is a directed path $x_1 \to x_2 \to \cdots \to x_t$ of length $(t-1)$, then $l_t(\Gamma) = 1$ and $p_t(\Gamma) = 0$. Hence, $\text{reg}(S/I_t(\Gamma)) = \text{reg}(S/(x_1\cdots x_t)) = t - 1 = (t-1)[l_t(\Gamma) + p_t(\Gamma)]$.

## 4. Linear strand and linear resolution

In this section, we compute the linear strand of $I_t(\Gamma)$ for a rooted tree $\Gamma$, and classify all rooted trees $\Gamma$ for which $I_t(\Gamma)$ has a linear resolution.

We start by investigating the linear strand of $I_t(\Gamma)$. Note that $\beta_{0,t}(I_t(\Gamma))$ is just the number of paths of length $(t-1)$ in $\Gamma$. Therefore we will be interested in $\beta_{i,i+t}(I_t(\Gamma))$ for $i \geqslant 1$. Note also that the path ideal $I_t(\Gamma)$ can be realized as the edge ideal of a *hyper-tree*. In [4], the second author and Van Tuyl gave a formula for the linear strand of the edge ideal of any hyper-tree. However, the structure of the hyper-tree corresponding to $I_t(\Gamma)$ is quite complicated. We shall use the combinatorial data of $\Gamma$ to provide an explicit formula for the linear strand of $I_t(\Gamma)$.

**4.1. Lemma.** *Let $\Gamma$ be a rooted tree of height $h \geqslant t - 1$, let $x_{i_t}$ be a leaf at the highest level in $\Gamma$, and let $x_{i_1}, \ldots, x_{i_t}$ be the unique path of length $(t-1)$ terminating at $x_{i_t}$. Then for $i > 0$,*

$$\beta_{i,i}\big(S/\big(I_t(\Gamma\backslash x_{i_t}) : (x_{i_1}\cdots x_{i_t})\big)\big) = \begin{cases} \binom{\deg_\Gamma(x_{i_{t-1}})-2}{i} & \text{if } h = t - 1 \text{ and } t \neq 2, \\ \binom{\deg_\Gamma(x_{i_{t-1}})-1}{i} & \text{if } h > t - 1 \text{ or } t = 2. \end{cases}$$

**Proof.** Let $x_{i_0}$ be the parent of $x_{i_1}$ if it exists; and let $P = \{x_{i_0}, \ldots, x_{i_t}\}$. Furthermore, for $j = 0, \ldots, t - 1$, let $\Gamma_j$ be the induced subtree of $\Gamma$ rooted at $x_{i_j}$, and let $\Delta_j = \Gamma_j\backslash\Gamma_{j+1}$. Observe, as before, that the minimal generators of the ideals $I_t(\Gamma\backslash P)$, $(x_{i_0})$, and $I_{t-j}(\Delta_j\backslash P)$ involve pairwise disjoint sets of vertices. Thus, by Lemma 2.8, the minimal free resolution of $S/(I_t(\Gamma\backslash x_{i_t}) : (x_{i_1}\cdots x_{i_t}))$ is the tensor product of the minimal free resolutions of $S/I_t(\Gamma\backslash P)$, $S/(x_{i_0})$, and $S/I_{t-j}(\Delta_j\backslash P)$ for $j = 0, \ldots, t - 1$. Therefore, the contribution to $\beta_{i,i}(S/(I_t(\Gamma\backslash x_{i_t}) : (x_{i_1}\cdots x_{i_t})))$ comes from $\beta_{i,i}(S/x_{i_0} \otimes S/I_1(\Delta_{t-1}\backslash P)) = \beta_{i,i}(S/((x_{i_0}) + I_1(\Delta_{t-1}\backslash P)))$.

Observe that the number of vertices of $\Delta_{t-1}\backslash P$ is $\deg(x_{i_{t-1}}) - 2$ if $t > 2$ or $t = 2$ and $h > t - 1$ (i.e., when $x_{i_{t-2}}$ exists). If $t = 2$ and $h = t - 1$ (i.e., when $x_{i_{t-2}}$ does not exist) then the number of vertices of $\Delta_{t-1}\backslash P$ is $\deg(x_{i_{t-1}}) - 1$. Observe further that $x_{i_0}$ exists if $h > t - 1$. Hence, the conclusion follows from the fact that the minimal free resolution of $S/(x_{i_0} + I_1(\Delta_{t-1}\backslash P))$ is the Koszul complex. $\quad\square$

**4.2. Theorem.** *Let $\Gamma$ be a rooted tree over the vertex set $V$. Then for $i \geqslant 1$,*

$$\beta_{i,i+t}\big(I_t(\Gamma)\big) = \begin{cases} \sum_{v\in V} \binom{\deg_\Gamma(v)}{i+1} & \text{if } t = 2, \\ \sum_{\text{level}(v)\geqslant t-1} \binom{\deg_\Gamma(v)}{i+1} + \sum_{\text{level}(v)=t-2} \binom{\deg_\Gamma(v)-1}{i+1} & \text{if } t > 2. \end{cases}$$

**Proof.** Let $h = \text{height}\,\Gamma$. The assertion is vacuously true if $h < t - 1$ so we may assume that $h \geqslant t - 1$. We shall use induction on $n$, the number of vertices in $\Gamma$. Again, the assertion is vacuously true if $n = t$. Assume that $n > t$. Let $x_{i_t}$ be a leaf at the highest level in $\Gamma$, and let $x_{i_1}, \ldots, x_{i_t}$ be the

unique path of length $(t-1)$ terminating at $x_{i_t}$. For simplicity, let $\Gamma' = \Gamma \backslash x_{i_t}$. As observed before, $\beta_{i,i+t}(I_t(\Gamma)) = \beta_{i+1,i+t}(S/I_t(\Gamma))$. By Theorem 2.7, we have

$$\beta_{i+1,i+t}\big(S/I_t(\Gamma)\big) = \beta_{i+1,i+t}\big(S/I_t(\Gamma')\big) + \beta_{i,i}\big(S/\big(I_t(\Gamma \backslash x_{i_t}) : (x_{i_1} \cdots x_{i_t})\big)\big).$$

If $t = 2$ then by the induction hypothesis and Lemma 4.1, we have

$$\beta_{i+1,i+2}\big(S/I_2(\Gamma)\big) = \sum_{v \in V \backslash x_{i_2}} \binom{\deg_{\Gamma'}(v)}{i+1} + \binom{\deg_\Gamma(x_{i_1})-1}{i}.$$

Since $\deg_{\Gamma'}(v) = \deg_\Gamma(v)$ for all $v \notin \{x_{i_1}, x_{i_2}\}$, $\deg_{\Gamma'}(x_{i_1}) = \deg_\Gamma(x_{i_1}) - 1$, and $\deg_\Gamma(x_{i_2}) = 1$, we have

$$\beta_{i+1,i+2}\big(S/I_2(\Gamma)\big) = \sum_{\substack{v \neq x_{i_1} \\ v \neq x_{i_2}}} \binom{\deg_{\Gamma'}(v)}{i+1} + \binom{\deg_{\Gamma'}(x_{i_1})}{i+1} + \binom{\deg_\Gamma(x_{i_1})-1}{i}$$

$$= \sum_{\substack{v \neq x_{i_1} \\ v \neq x_{i_2}}} \binom{\deg_\Gamma(v)}{i+1} + \binom{\deg_\Gamma(x_{i_1})-1}{i+1} + \binom{\deg_\Gamma(x_{i_1})-1}{i}$$

$$= \sum_{\substack{v \neq x_{i_1} \\ v \neq x_{i_2}}} \binom{\deg_\Gamma(v)}{i+1} + \binom{\deg_\Gamma(x_{i_1})}{i+1}$$

$$= \sum_{v \neq x_{i_1}} \binom{\deg_\Gamma(v)}{i+1} = \sum_{v \in V} \binom{\deg_\Gamma(v)}{i+1}.$$

Here, the last equality follows by adding $0 = \binom{1}{i+1}$.

Assume now that $t > 2$. Consider the case when $h = t - 1$. By the induction hypothesis and Lemma 4.1, we have

$$\beta_{i+1,i+t}\big(S/I_t(\Gamma)\big) = \beta_{i+1,i+t}\big(S/I_t(\Gamma')\big) + \binom{\deg_\Gamma(x_{i_{t-1}})-2}{i}$$

$$= \sum_{\substack{v \neq x_{i_t} \\ \text{level}(v) \geqslant t-1}} \binom{\deg_{\Gamma'}(v)}{i+1}$$

$$+ \sum_{\substack{v \neq x_{i_t} \\ \text{level}(v) = t-2}} \binom{\deg_{\Gamma'}(v)-1}{i+1} + \binom{\deg_\Gamma(x_{i_{t-1}})-2}{i}.$$

Since $h = t - 1$ and $x_{i_t}$ is at the highest level, we have $\text{level}(x_{i_{t-1}}) = t - 2$. Also, $\deg_{\Gamma'}(v) = \deg_\Gamma(v)$ for all $v \notin \{x_{i_{t-1}}, x_{i_t}\}$ and $\deg_{\Gamma'}(x_{i_{t-1}}) = \deg_\Gamma(x_{i_{t-1}}) - 1$. Thus,

$$\beta_{i+1,i+t}\big(S/I_t(\Gamma)\big) = \sum_{\substack{v \neq x_{i_t} \\ \text{level}(v) \geqslant t-1}} \binom{\deg_{\Gamma'}(v)}{i+1} + \sum_{\substack{v \neq x_{i_t}, \, v \neq x_{i_{t-1}} \\ \text{level}(v) = t-2}} \binom{\deg_{\Gamma'}(v)-1}{i+1}$$

$$+ \binom{\deg_{\Gamma'}(x_{i_{t-1}})-1}{i+1} + \binom{\deg_\Gamma(x_{i_{t-1}})-2}{i}$$

$$= \sum_{\substack{v \neq x_{i_t} \\ \text{level}(v) \geqslant t-1}} \binom{\deg_\Gamma(v)}{i+1} + \sum_{\substack{v \neq x_{i_t}, \, v \neq x_{i_{t-1}} \\ \text{level}(v) = t-2}} \binom{\deg_\Gamma(v)-1}{i+1}$$

$$+ \binom{\deg_\Gamma(x_{i_{t-1}}) - 2}{i+1} + \binom{\deg_\Gamma(x_{i_{t-1}}) - 2}{i}$$

$$= \sum_{\substack{v \neq x_{i_t} \\ \text{level}(v) \geqslant t-1}} \binom{\deg_\Gamma(v)}{i+1}$$

$$+ \sum_{\substack{v \neq x_{i_t},\, v \neq x_{i_{t-1}} \\ \text{level}(v) = t-2}} \binom{\deg_\Gamma(v) - 1}{i+1} + \binom{\deg_\Gamma(x_{i_{t-1}}) - 1}{i+1}.$$

As before, we can add $0 = \binom{1}{i+1}$ to the sum and the assertion follows.

Now we consider the case $h > t - 1$. By the induction hypothesis and Lemma 4.1, we have

$$\beta_{i+1,i+t}\big(S/I_t(\Gamma)\big) = \beta_{i+1,i+t}\big(S/I_t(\Gamma')\big) + \binom{\deg_\Gamma(x_{i_{t-1}}) - 1}{i}$$

$$= \sum_{\substack{v \neq x_{i_t} \\ \text{level}(v) \geqslant t-1}} \binom{\deg_{\Gamma'}(v)}{i+1}$$

$$+ \sum_{\substack{v \neq x_{i_t} \\ \text{level}(v) = t-2}} \binom{\deg_{\Gamma'}(v) - 1}{i+1} + \binom{\deg_\Gamma(x_{i_{t-1}}) - 1}{i}.$$

Since $h > t - 1$ and $\text{level}(x_{i_t}) = h$, we have $\text{level}(x_{i_{t-1}}) \geqslant t - 1$. As before, $\deg_{\Gamma'}(v) = \deg_\Gamma(v)$ for all $v \notin \{x_{i_{t-1}}, x_{i_t}\}$ and $\deg_{\Gamma'}(x_{i_{t-1}}) = \deg_\Gamma(x_{i_{t-1}}) - 1$. Thus, using a similar computation as above, we get

$$\beta_{i+1,i+t}\big(S/I_t(\Gamma)\big) = \sum_{\substack{v \neq x_{i_t},\, v \neq x_{i_{t-1}} \\ \text{level}(v) \geqslant t-1}} \binom{\deg_{\Gamma'}(v)}{i+1} + \sum_{\text{level}(v) = t-2} \binom{\deg_{\Gamma'}(v) - 1}{i+1}$$

$$+ \binom{\deg_{\Gamma'}(x_{i_{t-1}})}{i+1} + \binom{\deg_\Gamma(x_{i_{t-1}}) - 1}{i}$$

$$= \sum_{\text{level}(v) \geqslant t-1} \binom{\deg_\Gamma(v)}{i+1} + \sum_{\text{level}(v) = t-2} \binom{\deg_\Gamma(v) - 1}{i+1}. \qquad \square$$

The rest of the section is devoted to classifying all rooted trees $\Gamma$ for which $I_t(\Gamma)$ has a linear resolution. Recall that $I_t(\Gamma)$ has a linear resolution if and only if $\beta_{i,j}(I_t(\Gamma)) = 0$ for all $j \neq i + t$. Our characterization is based on the following special class of rooted trees.

**4.3. Definition.** A *broom graph* of type $t$ consists of a *handle*, which is a directed path $x_0, \ldots, x_s$, such that every edge in $\Gamma$ (not on the handle) has the form $(x_i, y)$ for some $i \geqslant s - t$.
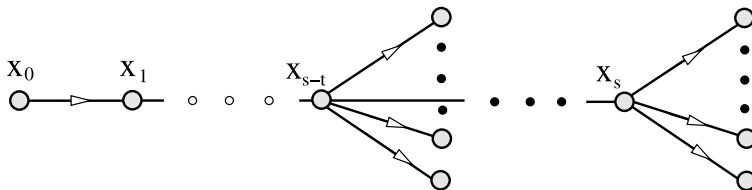


Since $t$ is usually fixed in our context, we often omit the phrase "of type $t$" and refer to a broom graph of type $t$ simply as a broom graph. Note also that in a broom graph as defined $s$ is related to the height of the graph.

**4.4. Remark.** If $x$ is a leaf of level strictly less than $(t-1)$, then $I_t(\Gamma \backslash x)$ and $I_t(\Gamma)$ have the same generators (in different polynomial rings). This implies that $\beta_{i,j}(I_t(\Gamma \backslash x)) = \beta_{i,j}(I_t(\Gamma))$ for all $i, j$. Thus we can successively remove leaves at level strictly less than $(t-1)$ from a rooted tree without changing the graded Betti numbers of its path ideal. We call this process the *cleaning* process of $\Gamma$. The rooted tree obtained after the cleaning process is called the *clean form* of $\Gamma$, denoted $C(\Gamma)$.

We are now ready to state our characterization.

**4.5. Theorem.** *Let $\Gamma$ be a rooted tree of height $h \geqslant t-1$. Then the following are equivalent:*

(1) $I_t(\Gamma)$ *has linear first syzygies.*
(2) $I_t(\Gamma)$ *has a linear resolution.*
(3) $C(\Gamma)$ *is a broom graph of height at most $(2t-1)$.*

**Proof.** Without loss of generality, we may assume that $\Gamma$ is already in its clean form, i.e., $C(\Gamma) = \Gamma$. In this case, all leaves of $\Gamma$ have level at least $(t-1)$.

We shall use induction on $n$, the number of vertices in $\Gamma$. The statement is clearly true for $n \leqslant t$ (since in this case $I_t(\Gamma)$ is either $(0)$ or $(x_1 \cdots x_t)$). Assume that $n > t$. As before, let $x_{i_t}$ be a leaf at the highest level in $\Gamma$, and let $x_{i_1}, \ldots, x_{i_t}$ be the unique path of length $(t-1)$ terminating at $x_{i_t}$. Let $x_{i_0}$ be the parent of $x_{i_1}$ if it exists; and let $P = \{x_{i_0}, \ldots, x_{i_t}\}$. Furthermore, for $j = 0, \ldots, t-1$, let $\Gamma_j$ be the induced subtree of $\Gamma$ rooted at $x_{i_j}$, let $\Delta_j = \Gamma_j \backslash \Gamma_{j+1}$, and let $\Gamma' = \Gamma \backslash x_{i_t}$. By Theorem 2.7, we have

$$\beta_{i+1,j}\big(S/I_t(\Gamma)\big) = \beta_{i+1,j}\big(S/I_t(\Gamma')\big) + \beta_{i,j-t}\big(S/\big(I_t(\Gamma') : (x_{i_1} \cdots x_{i_t})\big)\big).$$

This implies that $I_t(\Gamma)$ has linear first syzygies (respectively, has a linear resolution) if and only if $I_t(\Gamma')$ has linear first syzygies (respectively, has a linear resolution), and $I_t(\Gamma') : (x_{i_1} \cdots x_{i_t})$ is generated in degree one (respectively, is generated in degree one and has a linear resolution).

By Lemma 2.8 the minimal free resolution of $S/(I_t(\Gamma') : (x_{i_1} \cdots x_{i_t}))$ is the tensor product of the minimal free resolutions of $S/I_t(\Gamma \backslash P)$, $S/(x_{i_0})$, and $S/I_{t-j}(\Delta_j \backslash P)$. It follows that $I_t(\Gamma') : (x_{i_1} \cdots x_{i_t})$ is generated in degree one if and only if $I_t(\Gamma \backslash P) = (0)$ and $I_{t-j}(\Delta_j \backslash P) = (0)$ for all $j < t-1$. Note that if this is the case then $I_t(\Gamma') : (x_{i_1} \cdots x_{i_t})$ has a linear resolution. Thus, (1) and (2) are equivalent.

Now we show that $(2) \Rightarrow (3)$. Since $\Gamma$ is in its clean form, if $I_t(\Gamma \backslash P) = (0)$ then $\Gamma \backslash \Gamma_0$ must be a directed path from the root of $\Gamma$ to the parent of $x_{i_0}$ of length strictly less than $(t-1)$. To show that $\Gamma$ is a broom, it now suffices to show that for any $j$, $\Delta_j \backslash P$ consists of isolated vertices. Suppose for some $j < t-1$, $\Delta_j \backslash P$ contains a path of length at least 1. In this case, we can find a path of length at least 2 of the form $x_{i_j}, y_1, \ldots, y_s$ in $\Delta_j$ terminating at a leaf $y_s$ for some $2 \leqslant s \leqslant t-j$. Here, the second inequality is due to the fact that $x_{i_t}$ is of highest level. Recall that for $I_t(\Gamma)$ to have a linear resolution, $I_t(\Gamma') = I_t(\Gamma \backslash x_{i_t})$ must also have a linear resolution. Thus by induction and successively removing vertices, we will reduce $\Gamma$ to a rooted tree $\Gamma''$ in which $x_{i_{j+s}}$ is a leaf at the highest level and $I_t(\Gamma'')$ has a linear resolution. Let $\Delta''_{t-s}$ be the graph rooted at $x_{i_j}$ obtained from $\Gamma''$ in the same fashion as how $\Delta_j$ was obtained from $\Gamma$ (with $x_{i_{j+s}}$ replacing the role of $x_{i_t}$). Let $P''$ be the set of vertices on the unique path of length $(t-1)$ terminating at $x_{i_{j+s}}$ (this path exists since the level of $y_s$ is at least $(t-1)$). By a similar argument as with $\Delta_j$, for $I_t(\Gamma'')$ to have a linear resolution we must have $I_s(\Delta''_{t-s} \backslash P'') = (0)$. However, this is not true since $y_1 \cdots y_s \in I_s(\Delta''_{t-s} \backslash P'')$. We have now shown that for each $j < t-1$, $\Delta_j \backslash P$ consists of isolated vertices. Since $x_{i_t}$ is of highest level, $\Delta_{t-1} \backslash P$ also consists of isolated vertices. We can conclude that if $I_t(\Gamma)$ has a linear resolution, then $\Gamma$ is a broom of height at most $(2t-1)$.

Conversely, suppose that $\Gamma$ is a broom of height at most $(2t-1)$ in its clean form. By definition, it is easy to see that in this case, $\Gamma \backslash P$ consists of a path of length at most $(t-2)$ along with isolated vertices, and $\Delta_j \backslash P$ consists of isolated vertices for any $j$. Thus, $I_t(\Gamma \backslash P) = (0)$, $I_{t-j}(\Delta_j \backslash P) = (0)$ for all $j < t-1$, and $I_1(\Delta_{t-1} \backslash P)$ has a linear resolution. Moreover, since $\Gamma \backslash x_{i_t}$ is also a broom of height at most $(2t-1)$, by the induction hypothesis, $I_t(\Gamma \backslash x_{i_t})$ has a linear resolution. Thus, $I_t(\Gamma)$ has a linear resolution. Therefore, $(3) \Rightarrow (2)$.  $\square$

**4.6. Remark.** For any rooted tree $\Gamma$, a cellular complex supporting the linear strand of $I_t(\Gamma)$ is described by the minimal generators of $I_t(\Gamma)$; specifically, if $M_1, \ldots, M_u$ are minimal generators of $I_t(\Gamma)$, then $\{M_1, \ldots, M_u\}$ form a cell if and only if it is maximal with respect to the property that

$$\deg\big(\gcd(M_1, \ldots, M_u)\big) = t - 1.$$

When $\Gamma$ is a broom graph of height at most $(2t - 1)$, this cellular complex also supports the minimal free resolution of $I_t(\Gamma)$. In this case, set $D = \max\{\deg_\Gamma(x) \mid x \in \Gamma\}$ and let $v$ be a vertex of degree $D$ with highest level. Then it follows from Theorem 4.2 that this cellular complex supporting the minimal free resolution of $I_t(\Gamma)$ has dimension (equivalently, the projective dimension of $I_t(\Gamma)$) equal to

$$\begin{cases} D - 2 & \text{if } t > 2 \text{ and } \mathrm{level}(v) = (t - 2), \\ D - 1 & \text{otherwise.} \end{cases}$$

## 5. Specialization to path graphs

In this last section of the paper, we restrict our attention to a simple class of rooted trees, namely path graphs. The *path graph* over the vertex set $V = \{x_1, \ldots, x_n\}$ is the directed tree whose directed edges (after a possible re-indexing) are $e_i = (x_i, x_{i+1})$ for $i = 1, \ldots, n - 1$.



Let $L_n$ denote the path graph over $n$ vertices. Clearly,

$$I_t(L_n) = (x_1 x_2 \cdots x_t, \ x_2 x_3 \cdots x_{t+1}, \ \ldots, \ x_{n-t+1} x_{n-t+2} \cdots x_n).$$

It is also easy to see that

$$I_t(L_{n-1}) : (x_{n-t+1} x_{n-t+2} \cdots x_n) = \begin{cases} (x_{n-t}) + I_t(L_{n-(t+1)}) & \text{if } n > t, \\ 0 & \text{otherwise} \end{cases}$$

where we take $I_t(L_{n-(t+1)}) = (0)$ if $n - (t + 1) < t$.

Since the minimal resolution of $S/((x_{n-t}) + I_t(L_{n-(t+1)}))$ is the tensor product of the minimal free resolutions for $S/(x_{n-t})$ and $S/I_t(L_{n-(t+1)})$, it follows from Theorem 2.7 that

$$\beta_{i,j}\big(S/I_t(L_n)\big) = \beta_{i,j}\big(S/I_t(L_{n-1})\big) + \beta_{i-1,j-t}\big(S/I_t(L_{n-1}) : (x_{n-t+1} x_{n-t+2} \cdots x_n)\big)$$
$$= \beta_{i,j}\big(S/I_t(L_{n-1})\big) + \beta_{i-1,j-t}\big(S/I_t(L_{n-(t+1)})\big) + \beta_{i-2,j-t-1}\big(S/I_t(L_{n-(t+1)})\big).$$
$$\tag{5.1}$$

In [6] He and Van Tuyl computed the projective dimension of $S/I_t(L_n)$. Using (5.1), we can easily recover their formula.

**5.1. Corollary.** *Let $L_n$ be a path graph of over $n \geqslant t$ vertices. Then the projective dimension of $S/I_t(L_n)$ is given by*

$$\mathrm{pd}\big(S/I_t(L_n)\big) = \begin{cases} \frac{2(n-d)}{t+1} & \text{if } n \equiv d \bmod (t+1) \text{ for } 0 \leqslant d \leqslant t - 1, \\ \frac{2n-(t-1)}{t+1} & \text{if } n \equiv t \bmod (t+1). \end{cases}$$

**Proof.** The recursive formula (5.1) gives

$$\mathrm{pd}\big(S/I_t(L_n)\big) = \max\big\{\mathrm{pd}\big(S/I_t(L_{n-1})\big), \mathrm{pd}\big(S/I_t(L_{n-(t+1)})\big) + 2\big\}.$$

We can now proceed by using inducting on $n$. Using the same line of arguments as in [6, Theorem 4.1], the result follows. $\square$

**5.2. Remark.** Corollary 5.1, in fact, gives us that for $n \not\equiv 0, t \bmod (t+1)$,

$$\mathrm{pd}\big(S/I_t(L_n)\big) = \mathrm{pd}\big(S/I_t(L_{n-1})\big) = \mathrm{pd}\big(S/I_t(L_{n-(t+1)})\big) + 2,$$

and for $n \equiv 0, t \bmod (t+1)$,

$$\mathrm{pd}\big(S/I_t(L_n)\big) = \mathrm{pd}\big(S/I_t(L_{n-1})\big) + 1 = \mathrm{pd}\big(S/I_t(L_{n-(t+1)})\big) + 2.$$

Our last result characterizes which graded Betti numbers of $S/I_t(L_n)$ are nonzero.

**5.3. Theorem.** *Let $L_n$ be a path graph over $n \geqslant t$ vertices. Then the following are equivalent*:

(1) $\beta_{i,j}(S/I_t(L_n)) \neq 0$,
(2) $j - i = s(t-1)$ *for some integer $s$ satisfying $0 \leqslant s \leqslant \min\{i, \lceil \frac{n-t+1}{t+1} \rceil\}$ and $i \leqslant \min\{2s, \mathrm{pd}(S/I_t(L_n))\}$.*

**Proof.** The statement can be easily verified for $n = t$. We shall assume that $n > t$ and use induction on $n$.

We shall first show that (1) $\Rightarrow$ (2). To have $\beta_{i,j}(S/I_t(L_n)) \neq 0$, clearly $i \leqslant \mathrm{pd}(S/I_t(L_n))$. It suffices to show that $j - i = s(t-1)$ for some $s$ satisfying $0 \leqslant s \leqslant \min\{i, \lceil \frac{n-t+1}{t+1} \rceil\}$ and $i \leqslant 2s$. By the recursive formula (5.1), we have that at least one of the graded Betti numbers $\beta_{i,j}(S/I_t(L_{n-1}))$, $\beta_{i-1,j-t}(S/I_t(L_{n-(t+1)}))$ or $\beta_{i-2,j-t-1}(S/I_t(L_{n-(t+1)}))$ must be nonzero.

If $\beta_{i,j}(S/I_t(L_{n-1})) \neq 0$ then, by the induction hypothesis, we have $j - i = s(t-1)$ where $0 \leqslant s \leqslant \min\{i, \lceil \frac{(n-1)-t+1}{t+1} \rceil\} \leqslant \min\{i, \lceil \frac{n-t+1}{t+1} \rceil\}$ and $i \leqslant 2s$. And so (2) follows. If $\beta_{i-1,j-t}(S/I_t(L_{n-(t+1)})) \neq 0$, then by the induction hypothesis, we have $(j-i) - (t-1) = (j-t) - (i-1) = s'(t-1)$ for some $0 \leqslant s' \leqslant \min\{i-1, \lceil \frac{n-t+1}{t+1} \rceil - 1\}$ and $i - 1 \leqslant 2s'$. By taking $s = s' + 1$, (2) again follows. In the case where $\beta_{i-2,j-t-1}(S/I_t(L_{n-(t+1)})) \neq 0$, a similar argument again implies that once again (2) follows.

We proceed to show that (2) $\Rightarrow$ (1). Suppose that $j - i = s(t-1)$ where $0 \leqslant s \leqslant \min\{i, \lceil \frac{n-t+1}{t+1} \rceil\}$ and $i \leqslant \min\{2s, \mathrm{pd}(S/I_t(L_n))\}$. Clearly, (1) holds if $s = 0$, so we may assume that $s > 0$.

If $s \leqslant i - 1$ (i.e., $s - 1 \leqslant i - 2$), then set $s' = s - 1$. By Remark 5.2, we have $i - 2 \leqslant \mathrm{pd}(S/I_t(L_{n-(t+1)}))$. Since $(j-t-1) - (i-2) = (j-i) - (t-1) = s'(t-1)$ and $i - 2 \leqslant 2s - 2 = 2s'$, the induction hypothesis now implies that $\beta_{i-2,j-t-1}(S/I_t(L_{n-(t+1)})) \neq 0$, and (1) follows from (5.1).

It remains to consider the case where $s = i$. This, in particular, implies that $s = i \leqslant \lceil \frac{n-t+1}{t+1} \rceil$. Observe that if $n \equiv t \bmod (t+1)$, $n \geqslant t+1$ implies that $\lceil \frac{n-t+1}{t+1} \rceil + 1 = \lceil \frac{n+2}{t+1} \rceil \leqslant \lceil \frac{2n-t+1}{t+1} \rceil = \mathrm{pd}(S/I_t(L_n))$. On the other hand, if $n \equiv d \bmod (t+1)$ for some $d \leqslant t-1$, then $n \geqslant (t+1) + d$ and we have $\lceil \frac{n-t+1}{t+1} \rceil + 1 = \lceil \frac{n+2}{t+1} \rceil \leqslant \lceil \frac{2n-2d}{t+1} \rceil = \mathrm{pd}(S/I_t(L_n))$. Therefore we have $i \leqslant \lceil \frac{n-t+1}{t+1} \rceil \leqslant \mathrm{pd}(S/I_t(L_n)) - 1 = \mathrm{pd}(S/I_t(L_{n-(t+1)})) + 1$; that is, $i - 1 \leqslant \mathrm{pd}(S/I_t(L_{n-(t+1)}))$. Also, $i - 1 = s - 1 \leqslant 2(s-1)$. Now, set $s' = s - 1$, and observe that $(j-t) - (i-1) = (j-i) - (t-1) = s'(t-1)$ and $i - 1 \leqslant \min\{2s', \mathrm{pd}(S/I_t(L_{n-(t+1)}))\}$. The induction hypothesis implies that $\beta_{i-1,j-t}(S/I_t(L_{n-(n+1)})) \neq 0$, and (1) follows again from (5.1). $\square$

As a consequence of Theorem 5.3, we can compute the regularity of $S/I_t(L_n)$ and some graded Betti numbers of $I_t(L_n)$ explicitly.

**5.4. Corollary.** *Let $L_n$ be a path graph over $n \geqslant t$ vertices. Then the Castelnuovo–Mumford regularity of $S/I_t(L_n)$ is given by*

$$\mathrm{reg}\big(S/I_t(L_n)\big) = (t-1)\left\lceil \frac{n-t+1}{t+1} \right\rceil.$$

**Proof.** The conclusion follows from Theorem 5.3 noticing, by Corollary 5.1, that $\lceil \frac{n-t+1}{t+1} \rceil \leqslant \mathrm{pd}(S/I_t(L_n))$. $\square$

**5.5. Corollary.** *Let $L_n$ be a path graph over $n \geqslant t$ vertices. Then*

$$\beta_{i,it}\big(S/I_t(L_n)\big) = \binom{n - it + 1}{i}.$$

**Proof.** Observe first that Theorem 5.3 implies that $\beta_{i-2,it-t-1}(S/I_t(L_{n-(t+1)})) = 0$ since $(it - t - 1) - (i - 2) = (i - 1)(t - 1) > (i - 2)(t - 1)$. Now, by induction, assume that the result holds for all $L_m$ with $m < n$. It then follows from (5.1) that

$$\begin{aligned}
\beta_{i,it}\big(S/I_t(L_n)\big) &= \beta_{i,it}\big(S/I_t(L_{n-1})\big) + \beta_{i-1,it-t}\big(S/I_t(L_{n-(t+1)})\big) \\
&= \binom{(n-1) - it + 1}{i} + \binom{(n - (t+1)) - (i-1)t + 1}{i - 1} \\
&= \binom{n - it}{i} + \binom{n - it}{i - 1} \\
&= \binom{n - it + 1}{i}. \quad \square
\end{aligned}$$

**5.6. Remark.** As before, the linear strand of $I_t(L_n)$ is supported by a cellular complex. It is not hard to see that this complex is also a path graph, namely, the path graph $L_{n-t+1}$.

### Acknowledgments

### References

[1] Rachelle Bouchat, Free resolutions of some edge ideals of simple graphs, J. Commut. Algebra 2 (2010) 1–36.
[2] Paulo Brumatti, Aparecida Francisco da Silva, On the symmetric and Rees algebras of $(n, k)$-cyclic ideals, Mat. Contemp. 21 (2001) 27–42, 16th School of Algebra, Part II (Portuguese), Brasília, 2000.
[3] Aldo Conca, Emanuela De Negri, $M$-sequences, graph ideals, and ladder ideals of linear type, J. Algebra 211 (1999) 599–624.
[4] Huy Tài Hà, Adam Van Tuyl, Splittable ideals and the resolutions of monomial ideals, J. Algebra 309 (2007) 405–425.
[5] Huy Tài Hà, Adam Van Tuyl, Monomial ideals, edge ideals of hypergraphs, and their graded Betti numbers, J. Algebraic Combin. 27 (2008) 215–245.
[6] Jing (Jane) He, Adam Van Tuyl, Algebraic properties of the path ideal of a tree, arXiv:0902.0902.
[7] Gaetana Restuccia, Rafael H. Villarreal, On the normality of monomial ideals of mixed products, Comm. Algebra 29 (8) (2001) 3571–3580.
[8] Ezra Miller, Bernd Sturmfels, Combinatorial Commutative Algebra, Grad. Texts in Math., vol. 227, Springer, 2004.
[9] Rafael H. Villarreal, Rees algebras of edge ideals, Comm. Algebra 23 (1995) 3513–3524.
[10] Charles A. Weibel, An Introduction to Homological Algebra, Cambridge Stud. Adv. Math., vol. 38, Cambridge University Press, Cambridge, 1994.