ON CLOSED GRAPHS

DAVID A. COX AND ANDREW ERSKINE

Abstract. A graph is closed when its vertices have a labeling by \([n]\) with a certain property first discovered in the study of binomial edge ideals. In this article, we prove that a connected graph has a closed labeling if and only if it is chordal, claw-free, and has a property we call narrow, which holds when every vertex is distance at most one from all longest shortest paths of the graph. After proving our main result, we also explore other aspects of closed graphs, including the number of closed labelings and clustering coefficients.

1. Introduction

In this paper, \(G\) will always be a simple graph with finite vertex set \(V(G)\) and edge set \(E(G)\). We first recall the definition of closed graph.

Definition 1.1. A labeling of \(G\) is a bijection \(V(G) \simeq [n] = \{1, \ldots, n\}\), and given a labeling, we typically assume \(V(G) = [n]\). A labeling is closed if whenever we have distinct edges \(\{j, i\}, \{i, k\} \in E(G)\) with either \(j > i < k\) or \(j < i > k\), then \(\{j, k\} \in E(G)\). Finally, a graph is closed if it has a closed labeling.

A labeling of \(G\) gives a direction to each edge \(\{i, j\} \in E(G)\) where the arrow points from \(i\) to \(j\) when \(i < j\), i.e., the arrow points to the bigger label. The following picture illustrates what it means for a labeling to be closed:

\[
\begin{align*}
&j \quad \cdots \quad k \\
&\downarrow \quad \cdots \quad \downarrow \\
&i \quad \cdots \quad i
\end{align*}
\]

(1.1)

Whenever the arrows point away from \(i\) (as on the left) or towards \(i\) (as on the right), closed means that \(j\) and \(k\) are connected by an edge.

Closed graphs were first encountered in the study of binomial edge ideals. The binomial edge ideal of a labeled graph \(G\) is the ideal \(J_G\) in the polynomial ring \(k[x_1, \ldots, x_n, y_1, \ldots, y_n]\) (\(k\) a field) generated by the binomials

\[
f_{ij} = x_i y_j - x_j y_i
\]

for all \(i, j\) such that \(\{i, j\} \in E(G)\) and \(i < j\). A key result, discovered independently in \([3]\) and \([5]\), is that the above binomials form a Gröbner basis of \(J_G\) for lex order with \(x_1 > \cdots > x_n > y_1 > \cdots > y_n\) if and only if the labeling is closed. The name “closed” was introduced in \([3]\).

The algebraic properties of binomial edge ideals are explored in \([2]\) and \([7]\), and a generalization is studied in \([6]\). The paper \([1]\) gives a more general Gröbner basis

2010 Mathematics Subject Classification. 05C75 (primary), 05C25, 05C78 13P10 (secondary).

Key words and phrases. closed graph, binomial edge ideal, clustering coefficient.
criterion for a graph to be closed and, more relevant to our paper, gives a necessary and sufficient condition for a graph to have a closed labeling. The criterion given in [1] uses the simplicial complex formed by all cliques (complete subgraphs) of $G$.

The main goal of this paper is to characterize when a graph $G$ has a closed labeling in terms of properties that can be seen directly from the graph. Our starting point is the following result proved in [3].

**Proposition 1.2.** Every closed graph is chordal and claw-free.

“Chordal” is a standard term in graph theory (every cycle of length $\geq 4$ has a chord), and “claw-free” means that $G$ has no induced subgraph of the form

![Diagram of chordal and claw-free graph](image)

Definition 1.3. A connected graph $G$ is narrow if for every $v \in V(G)$ and every longest shortest path $P$ of $G$, either $v \in V(P)$ or there is $w \in V(P)$ with $\{v, w\} \in E(G)$.

In other words, a connected graph is narrow if every vertex is distance at most one from every longest shortest path. Here is an example of a graph that is chordal and claw-free but not narrow:

![Diagram of narrow graph](image)

Narrowness fails because the vertex $D$ is distance two from the longest shortest path $ACF$.

We can now state the main result of this paper.

**Theorem 1.4.** A connected graph is closed if and only if it is chordal, claw-free, and narrow.

Since a graph is closed if and only if its connected components are closed [5], we get the following immediate corollary of Theorem 1.4.
Corollary 1.5. A graph is closed if and only if it is chordal, claw-free, and its connected components are narrow.

The independence of the three conditions (chordal, claw-free, narrow) is easy to see. The graph (1.2) is chordal and narrow but not claw-free, and the graph (1.3) is chordal and claw-free but not narrow. Finally, the 4-cycle is claw-free and narrow but not chordal.

The paper is organized as follows. In Section 2 we recall some known properties of closed graphs and prove some new ones, and in Section 3 we introduce an algorithm for labeling connected graphs. Section 4 uses the algorithm to prove Theorem 1.4 and Section 5 presents some additional results about closed graphs.

2. Properties of Closed Labelings

2.1. Directed Paths. A path in $G$ is $P = v_0v_1 \cdots v_{\ell-1}v_\ell$ where $\{v_j, v_{j+1}\} \in E(G)$ for $j = 0, \ldots, \ell - 1$. A single vertex is regarded as a path of length zero. When $G$ is labeled, we assume as usual that $V(G) = [n]$. Then a path $P = i_0i_1 \cdots i_{\ell-1}i_\ell$ is directed if either $i_j < i_{j+1}$ for all $j = 0, \ldots, \ell - 1$ or $i_j > i_{j+1}$ for all $j = 0, \ldots, \ell - 1$.

Here is a useful result from [3].

Proposition 2.1. A labeling on a graph $G$ is closed if and only if for all vertices $i, j \in V(G) = [n]$, all shortest paths from $i$ to $j$ are directed.

Corollary 2.2. A connected closed graph $G$ has $\leq 2$ leaves (vertices of degree 1).

Proof. Let $u, v, w$ be leaves and fix a closed labeling of $G$. Without loss of generality we may assume $u < v < w$, and let $u', v', w'$ be the unique vertices adjacent to $u, v, w$ respectively. A shortest path from $u$ to $v$ is directed and must pass through $u'$ and $v'$. This implies $u < u' \leq v' < v$ since $u < v$. The same argument applied to $v$ and $w$ would imply $v < v' \leq w' < w$. Thus $v' < v$ and $v < v'$, which proves that three leaves cannot exist.

Since an induced subgraph of a closed graph is closed, Corollary 2.2 implies that every connected induced subgraph of a closed graph has $\leq 2$ leaves, a property we call three-leaf free. This generalizes the claw-free property of closed graphs stated in Proposition 1.2.

2.2. Neighborhoods and Intervals. Given a vertex $v \in V(G)$, the neighborhood of $v$ in $G$ is

$$N_G(v) = \{w \in V(G) \mid \{v, w\} \in E(G)\}.$$  

When $G$ is labeled and $i \in V(G) = [n]$, we have a disjoint union

$$N_G(i) = N^>_G(i) \cup N^<_G(i),$$

where

$$N^>_G(i) = \{j \in N_G(i) \mid j > i\}$$

$$N^<_G(i) = \{j \in N_G(i) \mid j < i\}.$$
This is the notation used in [1], where the following basic result is also proved.

**Proposition 2.3.** A labeling on a graph $G$ is closed if and only if for all vertices $i \in V(G) = [n]$, $N^+_G(i)$ and $N^-_G(i)$ are complete.

Given $i, j \in [n]$ with $i < j$, we get the interval $[i, j] = \{k \in [n] | i \leq k \leq j\}$. Then we have the following useful characterization of when a labeling of a connected graph is closed.

**Proposition 2.4.** A labeling on a connected graph $G$ is closed if and only if for all $i \in V(G) = [n]$, $N^+_G(i)$ is complete and equal to $[i+1, i+r]$, $r = |N^+_G(i)|$.

**Proof.** First suppose that the labeling is closed. Then $N^+_G(i)$ is complete by Proposition 2.3. It remains to show that $N^+_G(i)$ is an interval of the desired form.

Pick $j \in N^+_G(i)$ and $k \in [n]$ with $i < k < j$. Since $G$ is connected, there must be a path from $i$ to $k$, and therefore there must be a shortest path. Choose such a path $P = i_0i_1i_2\cdots i_m$ with $i = i_0$ and $i_m = k$. Then $P$ is directed by Proposition 2.1. Since $i < k$, we have $i = i_0 < i_1 < i_2 < \cdots < i_m = k$. Thus $i_1 \in N^+_G(i)$ and hence $\{i_1, j\} \in E(G)$ since $N^+_G(i)$ is complete. Since $i_1 < j$, we have $j \in N^+_G(i_1)$.

We now prove by induction that $j \in N^+_G(i_u)$ for all $u = 1, \ldots, m$. The base case is proved in the previous paragraph. Also, if $j \in N^+_G(i_u)$, then $\{j, i_{u+1}\} \in E(G)$ since $\{i_u, i_{u+1}\} \in E(G)$ and the labeling is closed. This completes the induction. Since $k = i_m$, it follows that $j \in N^+_G(k)$. Then we have $\{i, j\}, \{k, j\} \in E(G)$ with $i < j < k$. Thus $\{i, k\} \in E(G)$ since the labeling is closed, and then $i < k$ implies that $k \in N^+_G(i)$. It follows that $N^+_G(i)$ is an interval of the desired form.

Conversely, suppose that $N^+_G(i)$ is complete and $N^+_G(i) = [i+1, \ldots, i+r]$, $r = |N^+_G(i)|$, for all $i \in V(G)$. By Proposition 2.3, it suffices to show that $N^+_G(i)$ is complete. Given $j \not= k$ in $N^+_G(i)$, we can assume without loss of generality that $j < k$. Then we have $j < k < i$ with $i \in N^+_G(j)$. By hypothesis, $N^+_G(j)$ is an interval starting at $j + 1$. Since it contains $i$, it must contain $k$. But $k \in N^+_G(j)$ implies $\{j, k\} \in E(G)$, which completes the proof that $N^+_G(i)$ is complete. \qed

This proposition will be very useful in the proof of our main theorem.

2.3. **Layers.** The following subsets of $V(G)$ will play a key role in our study of closed graphs.

**Definition 2.5.** Let $G$ be a connected graph labeled so that $V(G) = [n]$. Then the $N^{th}$ layer of $G$ is the set

$$L_N = \{i \in [n] | d(i, 1) = N\}.$$ 

Thus $L_N$ consists of all vertices that are distance $N$ from the vertex 1. Note that $L_0 = \{1\}$ and $L_1 = N_G(1) = N^+_G(1)$. Furthermore, since $G$ is connected, we have a disjoint union

$$V(G) = L_0 \cup L_1 \cup \cdots \cup L_h,$$

where $h = \max\{d(i, 1) | i \in [n]\}$. Here are some simple properties of layers.

**Lemma 2.6.** Let $G$ be a connected graph labeled so that $V(G) = [n]$. Then:

1. If $i \in L_N$ and $\{i, j\} \in E(G)$, then $j \in L_{N-1}$, $L_N$, or $L_{N+1}$.
2. If $P$ is a path in $G$ connecting $i \in L_N$ to $j \in L_M$ with $N \leq M$, then for every integer $N \leq m \leq M$, there exists $k \in V(P)$ with $k \in L_m$. 

Proof. Statement (1) follows easily from the definition of layer. We prove (2) by induction on the length of $P$, the base case of length zero being obvious. Suppose that the result holds for paths of length $\ell$ and let $P = \ell_0 \ell_1 \cdots \ell_{\ell+1}$ be a path with $i_0 = i$ and $i_{\ell+1} = j$. Then $i_1 \in L_{\ell-1}$, $L_{\ell}$ or $L_{\ell+1}$ by the first part of the lemma. Applying the inductive hypothesis to the path $P' = \ell_1 \cdots \ell_{\ell+1}$ of length $\ell$, one sees without difficulty that the result holds for $P$. \hfill $\Box$

We next prove some properties of layers when the labeling is closed.

**Proposition 2.7.** Let $G$ be a connected graph with a closed labeling satisfying $V(G) = [n]$. Then:

1. Each layer $L_N$ is complete.
2. If $d = \max\{L_N\}$, then $L_{N+1} = N_G(d)$.

Proof. We first show that

\[(2.1) \quad r \in L_M, \ s \in L_{M+1}, \ \{r, s\} \in E(G) \implies r < s.\]

To see why, take a shortest path from $1$ to $r \in L_M$. This path has length $M$, so appending the edge $\{r, s\}$ gives a path of length $M + 1$ to $s$. Since $s \in L_{M+1}$, this is a shortest path and hence is directed by Proposition 2.1. Thus $r < s$.

Although (1) follows from (2) and Proposition 2.3, we prove (1) separately since it is needed in the proof of (2). We proceed by induction on $N \geq 0$. The base case is trivial since $L_0 = \{1\}$. Now assume $L_N$ is complete and take $i, j \in L_{N+1}$ with $i \neq j$. Take a shortest path $P_1$ from $1$ to $i$ and let $k \in L_N$ be the vertex of $P_1$ that is distance $N$ from $1$ in $P_1$. Thus $k \in L_N$. Similarly, a shortest path $P_2$ from $1$ to $j$ has a vertex $l$ with $l \in L_N$. Then $k < i$ and $l < j$ by (2.1).

If $k = l$, then $i > k < j$, which implies $\{i, j\} \in E(G)$ since the labeling is closed. If $k \neq l$, then $\{l, k\} \in E(G)$ since $L_N$ is complete. Without loss of generality, assume $l > k$. Then $l > k < i$ implies $\{l, i\} \in E(G)$ since the labeling is closed. Since $l \in L_N$ and $i \in L_{N+1}$, we conclude that $l < i$ by (2.1). Then $i > l < j$ and thus $\{i, j\} \in E(G)$ since the labeling is closed. We conclude that $L_{N+1}$ is complete.

We now turn to (2). To prove $L_{N+1} \subseteq N_G(d)$, let $d = \max\{L_N\}$, take $i \in L_{N+1}$. Then we have a shortest path from $1$ to $i$ of length $N + 1$. This path must have a vertex $k \in L_N$ such that $\{k, i\} \in E(G)$. Then $k < i$ by (2.1) and therefore $i \in N_G(d)$.

To prove the opposite inclusion, take $i \in N_G(d)$. Since $\{i, j\} \in E(G)$ and $d \in L_N$, we have $i \in L_M$ for $M = N - 1, N, N + 1$ by Lemma 2.1. If $i \in L_{N+1}$, then (2.1) would imply $i < d$, contradicting $i \in N_G(d)$. If $i \in L_N$, then $i \leq \max\{L_N\} = d$, which again contradicts $i \in N_G(d)$. Hence $i \in L_{N+1}$. \hfill $\Box$

2.4. **Longest Shortest Paths.** When the labeling of a connected graph is closed, the diameter of the graph determines the number of layers as follows.

**Proposition 2.8.** Let $G$ be a connected graph with a closed labeling. Then:

1. $\text{diam}(G)$ is the largest integer $h$ such that $L_h \neq \emptyset$.
2. If $P$ is a longest shortest path of $G$, then one endpoint of $P$ is in $L_0$ or $L_1$ and the other is in $L_h$, where $h = \text{diam}(G)$.
Proof. For (1), let \( h \) be the largest integer with \( L_h \neq \emptyset \). Thus \( h \) measures the height of \( G \). Since points in \( L_h \) have distance \( h \) from 1, we have \( h \leq \text{diam}(G) \).

For the opposite inequality, it suffices to show that \( d(i, j) \leq h \) for all \( i, j \in V(G) \) with \( i \neq j \). We can assume \( G \) has more than one vertex, so that \( h \geq 1 \). Suppose \( i \in L_N \) and \( j \in L_M \) with \( N \leq M \). If \( N = 0 \), then \( i = 1 \) and \( d(i, j) = d(1, j) = M \leq h \) since \( j \in L_M \). Also, if \( M = N \), then \( i, j \in L_N \), so that \( d(i, j) = 1 \leq h \) since \( L_N \) is complete by Proposition 2.7. Finally, if \( 0 < N < M \), we have distance \( \geq 1 \) from \( i \in L_M \) and a longest shortest path which forces \( N \) is complete by Proposition 2.7. Finally, if \( 0 < N < M \), let \( d_u = \max(L_u) \) for each integer \( u \). By Proposition 2.7, we know that \( j \in N_G^+(d_M-1) \). Hence, if \( i \neq d_N \), then \( P = id_Nd_{N+1} \cdots d_{M-2}d_{M-1}j \) is a path of length \( M - N + 1 \). If \( i = d_N \), then \( P = id_N \cdots d_{M-1}j \) is a path of length \( M - N \). Thus we have a path from \( i \) to \( j \) of length at most \( M - N + 1 \), so that \( d(i, j) \leq M - N + 1 \leq M \leq h \).

For (2), let \( i \) and \( j \) be the endpoints of the longest shortest path \( P \) with \( i \in L_N \) and \( j \in L_M \) and \( N \leq M \). If \( 0 < N < M \), then the previous paragraph implies

\[
\text{diam}(G) = d(i, j) \leq M - N + 1 \leq M \leq h = \text{diam}(G),
\]

which forces \( N = 1 \) (so \( i \in L_1 \)) and \( M = h \) (so \( j \in L_h \)). The remaining cases \( N = 0 \) and \( N = M \) are straightforward and are left to the reader. \( \square \)

Recall from Definition 1.3 that a connected graph \( G \) is narrow when every vertex is distance at most one from every longest shortest path. Narrowness is a key property of connected closed graphs.

**Theorem 2.9.** Every connected closed graph is narrow.

**Proof.** Let \( G \) be a connected graph with a closed labeling. Pick a vertex \( i \in V(G) \) and a longest shortest path \( P \). Since \( G \) is connected, \( i \in L_N \) for some integer \( N \). By Proposition 2.8, the endpoints of \( P \) lie in \( L_0 \) or \( L_1 \) and \( L_h \), \( h = \text{diam}(G) \). Then Lemma 2.6 implies that \( P \) has a vertex \( i_M \) in \( L_M \) for every \( 1 \leq M \leq h \).

If \( N \geq 1 \), then either \( i = i_N \in V(P) \) or \( i \neq i_N \), in which case \( \{i, i_N\} \in E(G) \) since \( L_N \) is complete by Proposition 2.7. On the other hand, if \( N = 0 \), then \( i \in L_0 \), hence \( i = 1 \). Then \( \{i, i_1\} = \{1, i_1\} \in E(G) \) since \( i_1 \in L_1 = N_G(1) \). In either case, \( i \) is distance at most one from \( P \). \( \square \)

### 3. A Labeling Algorithm

We introduce Algorithm 1 which labels the vertices of a connected graph. This algorithm will play a key role in the proof of Theorem 1.4

The algorithm works as follows. First a candidate for 1 is found by choosing an endpoint of a longest shortest path of minimal degree. We then go through the vertices in \( N_G(1) \) and label them 2, 3, \ldots, first labeling vertices with the fewest number of edges connected to unlabeled vertices. This process is repeated for the unlabeled vertices connected to vertex 2, and vertex 3, and so on until every vertex is labeled. Informally, we know that every vertex will be labeled because we first label everything in \( N_G^+(1) \), then label everything in \( N_G^+(2) \) not already labeled, and so on. Since the input graph is connected, this process must eventually reach all of the vertices.

In Lemma 3.1 below we prove carefully that Algorithm 1 produces a labeling of \( G \). Lemma 3.2 below explains the meaning of the function \( l \) that appears in the algorithm. Finally, the labeling allows us to define the layers \( L_N \) as in Definition 2.5.

Lemma 3.3 below describes how the layers interact with Algorithm 1.
Algorithm 1: Labeling Algorithm

Input: A connected graph $G$ with $n$ vertices

Output: A labeling $V(G) = [n]$ and a function $l : [n] \to \{0, \ldots, n - 1\}$

1. $i := 1$
2. $j := 0$
3. $v_0 :=$ endpoint of a longest shortest path with minimal degree;
4. label $v_0$ as $i$
5. $l(i) := j$
6. $i := i + 1$
7. $J := \{v_0\}$
8. $j := j + 1$
9. while $j \leq |V(G)|$ do
   10. $S := N_G(j) \setminus J$
   11. while $S \neq \emptyset$ do
       12. $v :=$ pick $v \in S$ such that $|N_G(v) \setminus J| = \min_{u \in S} |N_G(u) \setminus J|$
       13. label $v$ as $i$
       14. $l(i) := j$
       15. $i := i + 1$
       16. $J := J \cup \{v\}$
       17. $S := S \setminus \{v\}$
   end
   19. $j := j + 1$
end

Lemma 3.1. Let $G$ be a connected graph. Then:

1. At every stage of Algorithm 1, the variables $j$ and $i$ satisfy $j < i$.
2. The algorithm produces a labeling of $G$.
3. The function $l$ in the algorithm is defined on $[n]$ and maps to $\{0, \ldots, n - 1\}$.

Proof. First observe that every time a vertex is labeled, and only when a vertex is labeled, that vertex is added to the set $J$ in the algorithm. Therefore, throughout the algorithm, $J$ is always the set of vertices that are already labeled.

The first vertex, $v_0$, is labeled 1 on Line 4. At this point the variables $i$ and $j$ in the algorithm are both incremented from $i = 1, j = 0$ to $i = 2, j = 1$. Then every time through the loop that begins on Line 9, the set $S$ is reset to be the vertices in $N_G(j)$ that are not already labeled. The loop that begins on Line 11 picks a vertex $v \in S$, and labels it $i$, and then increments $i$. The set $S$ maintains the property of being the vertices of $N_G(j)$ that are not labeled since $v$ is removed from $S$.

We know that every time $S$ is set on Line 10, $S$ will not contain any labeled vertices and therefore no vertex is labeled more than once. We also know that since $i$ is always incremented by 1 every time a vertex is labeled, $J$ always is of the form $J = \{1, 2, \ldots, i\}$. Hence no vertex is labeled with a value greater than $n = |V(G)|$.

We now prove (1), which asserts that $j < i$ at all times. This is true when $1$ is labeled since $j = 0 < 1 = i$ at this stage of the algorithm. Now suppose that when $i = i_0$ is labeled, we have $j = j_0 < i_0$. There are two possibilities for the value of $j$ when $i_0 + 1$ is labeled.
The first occurs when $S \setminus \{i_0\}$ is not yet empty and therefore when the algorithm resets $S := S \setminus \{i_0\}$ on Line 17, we have $S \neq \emptyset$. Then the algorithm continues through the loop starting on Line 11 and $j$ has not yet incremented. Since the next vertex is labeled $i = i_0 + 1$ in this loop while $j = j_0$, we know by our assumption that $j = j_0 < i_0 < i_0 + 1 = i$.

The second possibility occurs when $S \setminus \{i_0\} = \emptyset$ immediately after $i_0$ is labeled and therefore $S = \emptyset$ on Line 17. At this point, the algorithm does not continue the loop on Line 11 but instead increments $j$ until $N_G(j) \setminus J \neq \emptyset$. If any unlabeled vertices remain, then since $G$ is connected, some labeled vertex must be adjacent to an unlabeled vertex. At this point in the algorithm, $J = \{1, \ldots, i_0\}$ remains unchanged as long as we continue to increment $j$. Since $S = N_G(j) \setminus J$, skipping the loop beginning on Line 11 means $S = \emptyset$, so that every vertex in $N_G(j)$ is labeled. It follows that the first labeled vertex $j_1$ adjacent to an unlabeled vertex is where we stop incrementing $j$ and instead assign the label $i_0 + 1$ to one of the unlabeled vertices of $N_G(j_1)$. Since $j_1$ is labeled, we have $j_1 \leq i_0$. Hence $j = j_1 = i_0 < i_0 + 1 = i$. This completes the induction, and (1) is proved.

For (2), pick an arbitrary $v \in V(G)$. Since $G$ is connected, there is a path $P = v_0 v_1 v_2 \cdots v_m$ from $1 = v_0$ to $v_m = v$. Note that $v_0 = 1$ is labeled.

Suppose $v_u$ is labeled at some point in the algorithm. We claim the same is true for $v_{u+1}$. To see why, assume $v_u$ is labeled $i_0$ when $j = j_0$. Thus $j_0 < i_0$ by (1). Each time the algorithm goes through the loop beginning on Line 9, $j$ is incremented by one. If $v_{u+1}$ is labeled by the time we get to $j = i_0 - 1$, the claim is true. If we get to $j = i_0$ without $v_{u+1}$ being labeled, then $v_{u+1} \in N_G(v_u)$ is unlabeled, so $v_{u+1}$ will be in the set $S = N_G(v_u) \setminus J$. Since the loop beginning on Line 11 eventually labels every vertex in $S$, it follows that $v_{u+1}$ will be labeled. By induction, we see that at some point, $v$ will be labeled. It follows that all $n$ vertices of $V(G)$ are labeled by a distinct integer $i \leq n$. This gives a labeling with $V(G) = [n]$, and (2) is proved.

Finally, (3) follows immediately from (1) and (2) since $l(i)$ is always set to the current value of $j$ on Line 5 or 14.

The next lemma explains what the function $l$ means in terms of the labeling.

**Lemma 3.2.** Let $G$ be a graph with the labeling given by Algorithm 1. Then $l(1) = 0$, and for every $i \in [n]$ with $i > 1$, $l(i) = \min(N_G(i))$.

**Proof.** The first time in the algorithm when $l(i)$ is given a value is when $i = 1$ on Line 5. Since $j = 0$ and $l(i) := j$, we get $l(1) = 0$. For a vertex $i > 1$, when that vertex is labeled on Line 13 immediately afterward on Line 14, $l(i) := j$. At this point, we have $S \subseteq N_G(j)$, and therefore $i \in N_G(j)$, which implies $j \in N_G(i)$.

Fix some $i_0 \in [n]$ with $i_0 > 1$ and fix the $j_0 \in [n]$ such that $j = j_0 = l(i_0)$ when $i_0$ is labeled. To prove that $l(i_0) = \min(N_G(i_0))$, suppose that there is $s \in N_G(i_0)$ such that $s < l(i_0)$. This means that in an earlier loop on Line 10 when $j = s$, we know that that $i_0$ is not yet labeled so $i_0 \notin J$, and therefore $i_0 \in N_G(a) \setminus J = S$. This means that at some iteration of the loop on Line 12 when $j = s$, $i_0$ is labeled. But we know that $i_0$ is labeled when $j = l(i_0) > s$, a contradiction. □

In Lemma 3.2, $l(i)$ plays two roles. First, in the context of Algorithm 1, $l(i)$ is the value of the variable $j$ when the label $i$ is assigned to a vertex. Second, in terms of the labeling of the graph, $l(i)$ is the smallest vertex that is adjacent to $i$. Together, these dual roles of $l(i)$ are also present in the following lemma.
Lemma 3.3. Let $G$ be a graph with the labeling given by Algorithm 1. Then:

1. If $l(t) < l(s)$, then $t < s$.
2. If $t \in L_N$, then $l(t) \in L_{N-1}$ if $N > 0$.
3. If $t \in L_N$ and $s \in L_M$ with $N < M$, then $t < s$.

Proof. For (1), suppose that $s, t \in [n]$ satisfy $l(t) < l(s)$. Recall that $l(t)$ (resp. $l(s)$) is the value of $j$ when the label $t$ (resp. $s$) was assigned in Algorithm 1. Then $l(t) < l(s)$ implies that the label $s$ was assigned later than $t$ in the algorithm. Since the labels are assigned in numerical order, we must have $t < s$.

We prove (2) and (3) simultaneously by induction on $N \geq 1$ (the case $N = 0$ of (3) is trivially true). Since $j$ is assigned the value 1 on Line 8 of the first time Algorithm 1 gets to Line 10 we have $S = N_G(1) \setminus J = N_G(1) = L_1$. Every vertex in $S = L_1$, is labeled during the loop starting on Line 9 so $l(t) = 1$ for all $t \in L_1$. Hence (2) holds when $N = 1$. Also, for $t \in L_1$ and $s \in L_M$ with $1 < M$, we know that $s \notin S = L_1$ the first time $S$ is set. Since labels are assigned in numerical order, we must have $t < s$. Hence (3) holds when $N = 1$.

Now assume that (2) and (3) hold for $M$ and every $N \leq N_0$. Given $t \in L_{N_0+1}$, a shortest path from 1 to $t$ gives $v \in L_{N_0}$ with $v \in N_G(t)$. Since $l(t) = \min(N_G(t))$ by Lemma 3.2 we have $l(t) \leq v$. We have $l(t) \in L_u$ for some $u$. If $u > N_0$, then the inductive hypothesis for (3) would imply $l(t) > v$, which contradicts $l(t) \leq v$. Hence $l(t) \in L_u$ for some $u \leq N_0$. But $t \in L_{N_0+1}$ and $(t, l(t)) \in E(G)$ imply $l(t) \in L_u$ for $u \leq N_0$ by Lemma 2.6(1). Hence $l(t) \in L_{N_0}$, proving (2) for $N_0 + 1$.

Turning to (3), pick $t \in L_{N_0+1}$ and $s \in L_M$ with $N_0 + 1 < M$. We just showed that $l(t) \in L_{N_0}$ and Lemma 2.6(1) implies that $l(s) \in L_u$, $u \geq M - 1$, since $s \in L_M$. Then $N_0 < M - 1 \leq u$, so our inductive hypothesis, applied to $l(t) \in L_{N_0}$ and $l(s) \in L_u$, implies $l(t) < l(s)$. Then $t < s$ by (1), proving (3) for $N_0 + 1$.

Note that the contrapositive of Lemma 3.3(1) states that $t \geq s$ implies $l(t) \geq l(s)$.

4. Proof of the Main Theorem

We now turn to the main result of the paper. Theorem 1.4 from the Introduction states that a connected graph is closed if and only if it is chordal, claw-free and narrow. One direction is now proved, since closed graphs are chordal and claw-free by Proposition 1.2 and connected closed graphs are narrow by Theorem 2.9.

The proof of converse is much harder. The key idea that the labeling constructed by Algorithm 1 is closed when the input graph is chordal, claw-free and narrow. Thus the proof of Theorem 1.4 will be complete once we prove the following result.

Theorem 4.1. Let $G$ be a connected, chordal, claw-free, narrow graph. Then the labeling produced by Algorithm 1 is closed.

Proof. By Proposition 2.4 it suffices to show that the labeling produced by Algorithm 1 has the property that for all $m \in V(G) = [n]$,

$$(4.1) \quad N_G^c(m) \text{ is complete and } N_G(m) = [i + m, i + r_m] \text{ for } r_m = |N_G^c(m)|.$$ 

We will prove this by induction on $m$.

In Lemma 4.2 below, we show that $(4.1)$ holds for $m = 1$, and in Lemma 4.3 below, we show that if $(4.1)$ holds for all $1 \leq u < m$, then it also holds for $m$. Thus, we will be done after proving Lemmas 4.2 and 4.3. □

4.1. The Base Case. After Algorithm 1 runs on a chordal, claw-free and narrow graph \( G \), the following lemma is the base case of the induction in the proof of Theorem 4.1.

**Lemma 4.2.** \( N_G^>(1) = [2,1 + r] \), \( r = |N_G^>(1)| \), and \( N_G^=(1) \) is complete.

**Proof.** We will first show that \( N_G^>(1) = [2,1 + r] \), \( r = |N_G^>(1)| \). The first time through the loop beginning on Line 9 in Algorithm 1, \( j = 1 \) and \( i = 2 \) and \( S = N_G(1) \). For each vertex in \( S \), the loop beginning on Line 11 labels that vertex \( i \), removes it from \( S \), and increments \( i \). This continues until \( S = \emptyset \), at which point every vertex in \( S \) has been labeled 2, 3, . . . , \( 1 + r \), where \( r \) is the initial size of \( S \). Hence \( N_G^=(1) = N_G(1) = [2,1 + r] \).

The proof that \( N_G^=(1) \) is complete is harder, with many cases to consider. Pick distinct vertices \( s,t \in N_G^=(1) \) and assume that \( \{s,t\} \notin E(G) \).

Note that \( s,t \in L_1 \) are distance 2 apart and therefore \( h = \text{diam}(G) \geq 2 \). Our choice of vertex 1 guarantees that there is a longest shortest path \( P \) with 1 as an endpoint. Let \( z \in V(G) \) be the other, so that \( P = v_0v_1 \cdots v_h, 1 = v_0 \) and \( v_h = z \). Both \( s,t \) are distance at least \( h - 1 \) from \( z \), since otherwise \( d(1,z) < h \). Since \( G \) has diameter \( h \), it follows that

\[
(4.2) \quad s \text{ and } t \text{ are distance } h - 1 \text{ or } h \text{ from } z.
\]

Since \( v_1 \in V(P) \) is the only vertex of \( P \) in \( L_1 \), \( s \) and \( t \) cannot both lie on \( P \). Therefore, either \( s \in V(P), t \in V(P) \), or \( s,t \notin V(P) \). We will show that each possibility leads to a contradiction, proving that \( \{s,t\} \in E(G) \).

**Case 1.** Both \( s,t \notin V(P) \). Since the subgraph induced on vertices 1, \( s,t \), \( v_1 \) cannot be a claw, either \( \{v_1,s\} \in E(G) \) or \( \{v_1,t\} \in E(G) \) or both. We consider each case separately.

**Case 1A.** Both \( \{v_1,s\}, \{v_1,t\} \in E(G) \), as shown in Figure 1. By (4.2), there are three possibilities: both \( s,t \) are distance \( h \) from \( z \), one of \( s,t \) is distance \( h \) from \( z \) while the other is distance \( h - 1 \) from \( z \), and both \( s,t \) are distance \( h - 1 \) from \( z \).

![Figure 1. The portion of the graph relevant to Case 1A.](image)

*First,* suppose that \( s,t \) are distance \( h \) from \( z \). Then neither \( s \) nor \( t \) is adjacent to \( v_2 \), or else there is a path shorter than length \( h \) from \( s \) or \( t \) to \( z \). Hence the subgraph induced on \( v_2, v_1, s, t \) is a claw, contradicting our assumption of claw-free.

*Second,* suppose that one of \( s,t \) is distance \( h - 1 \) from \( z \) and the other is distance \( h \) from \( z \). Without loss of generality, let \( t \) be the vertex of distance \( h \) from \( z \). Because the subgraph induced by \( s, t, v_2 \) cannot be a claw, either \( \{s,v_2\} \) or
{t, v₂} is an edge. But t has distance h from z, so {t, v₂} ∉ E(G), and hence {s, v₂} ∈ E(G). See Figure 2(a). The choice of 1 = v₀ on Line 3 of Algorithm 1 implies |N₂(1)| ≤ |N₀(t)|. Since {1, s} ∈ E(G) and {s, t} ∉ E(G), there must be t₂ ∈ N₀(t) such that {1, t₂} ∉ E(G), and hence t₂ ∈ L₂. We also know that {s, t₂} ∉ E(G) since otherwise there would be a 4-cycle 1s₁t₁ with no chords since {1, t₂}, {t, s} ∉ E(G), contradicting the assumption that G is chordal. Since G is narrow, this t₂ must either be in P or be adjacent to some vertex in P. We know that t₂ is distance 2 from 1 and therefore t₂ /∈ V(P) (if it were, then t₂ = v₂, which is impossible since {t, v₂} ∉ E(G) and {t, t₂} ∈ E(G)). Hence t₂ is adjacent to a vertex of V(P).

If {t₂, vₜ} ∈ E(G), then u ≤ 2 or else there is a path from t to z through t₂ of length less than h. Therefore either {t₂, v₂} ∈ E(G) or {t₂, v₁} ∈ E(G). We will show that both {t₂, v₂}, {t₂, v₁} ∈ E(G), as in Figure 2(b).

If {t₂, v₂} ∈ E(G), then we have a 4-cycle tt₂v₂v₁t, and since {t, v₂} ∉ E(G) and G is chordal, we must have {t₂, v₁} ∈ E(G). And if {t₂, v₁} ∈ E(G), then {t₂, v₂} ∈ E(G) since the subgraph induced on v₁, v₂, t₂, 1 must be claw-free. Hence {t₂, v₁}, {t₂, v₂} ∈ E(G).

It follows that we have Figure 2(b) with the 5-cycle 1sv₂t₂t₁. Of the 5 possible chords, {1, t₂}, {s, t}, {t, v₂}, {t₂, s} ∉ E(G) as shown above. Also {1, v₂} ∉ E(G) or else we have a contradiction since v₂ is distance 2 from 1. Therefore we have a 5-cycle with no chords, a contradiction since G is chordal.

Third, suppose that both s, t are distance h – 1 from z. Since the subgraph induced on v₂, s, t, v₁ is claw-free, without loss of generality, we may assume {s, v₂} ∈ E(G). Since {s, t}, {1, v₂} ∉ E(G), we also have {t, v₂} ∉ E(G), else there would be a 4-cycle 1sv₂t₁ with no chords, a contradiction. Hence we are in the situation of Figure 2(a). Then we must have h > 2, since if h = 2, then t would be distance h – 1 = 1 from v₂ = z, so t would be adjacent to v₂, contradicting {t, v₂} ∉ E(G). Hence vₚ ∈ P exists since P is a longest shortest path. However, since t is distance h – 1 from z, we can take a shortest path P₁ = tₜ₂ · · · tₜ from t to tₜ = z. We know that vₚ ≠ v₂ since {t, v₂} ∉ E(G). We also have {t₂, s} ∉ E(G), since otherwise we would have the 4-cycle 1st₂t₁ with no chords, a contradiction.
We claim that \( \{t_2, v_2\} \in E(G) \) or \( \{t_2, v_3\} \in E(G) \). To see why, we use the longest shortest path \( P_2 = 1sv_2v_3 \cdots v_h, \) \( v_h = z \). Since \( G \) is narrow, \( t_2 \) is either a vertex or adjacent to a vertex of \( P_2 \). The former cannot occur since \( t_2 \neq v_2 \) and \( t_2, v_2 \in L_2 \). So \( t_2 \) is adjacent to a vertex of \( P_2 \). The vertex cannot be 1 or \( s \) (shown above), and since \( t_2 \in L_2 \), the vertex cannot be \( v_u \) for \( u > 3 \) by Lemma 2. If \( \{t_2, v_2\} \in E(G) \), then we have the 4-cycle \( \{t, v_2\} \) with no chords since \( \{t, s\} \) is distance 1 from \( t \), \( v_2 \) is distance 2 from 1. The vertex cannot be 1 or \( s \) (shown above), and since \( t_2 \in L_2 \), the vertex cannot be \( v_u \) for \( u > 3 \) by Lemma 2. If \( \{t_2, v_3\} \in E(G) \), then we have the 6-cycle \( \{t, v_3\} \) with no chords since \( \{t, s\} \) is distance 1 from \( t \), \( v_3 \) is distance 2 from 1. The vertex cannot be 1 or \( s \) (shown above), and since \( t_2 \in L_2 \), the vertex cannot be \( v_u \) for \( u > 3 \) by Lemma 2.

If \( \{t_2, v_2\} \in E(G) \), then the 5-cycle \( t_2v_2s1tt_2 \) in Figure 3(a) has no chords since \( \{t_2, s\}, \{t_2, 1\}, \{v_2, 1\}, \{v_2, t\}, \{t, s\} \notin E(G) \), contradicting chordal. Similarly, if \( \{t_2, v_3\} \in E(G) \), we have the 6-cycle \( t_2v_3s1tt_2 \) in Figure 3(b), which has the same impossible chords in addition to \( \{1, v_3\}, \{s, v_3\}, \{t, v_3\}, \{v_2, t_2\} \notin E(G) \), again contradicting chordal.

**Figure 3.** Both (a) and (b) cannot have the dotted edge or the graph has a 5-cycle or 6-cycle with no chord.

**Case 1B.** We are still in Case 1 with \( s, t \notin V(P) \), but now without loss of generality, we consider the case where \( \{s, v_1\} \in E(G) \) and \( \{t, v_1\} \notin E(G) \), as shown in Figure 4(a). First note that \( \{t, v_2\} \notin E(G) \), since if \( \{t, v_2\} \in E(G) \), then we would have the 4-cycle \( tv_1v_2t_1 \) with no chords since \( \{v_1, t\}, \{v_2, 1\} \notin E(G) \).

We claim that there exists \( t_2 \in N_G(t) \) with \( t_2 \notin N_G(1) \). To prove the existence of \( t_2 \), recall from (2.4) that \( t \) has distance \( h \) or \( h - 1 \) from the endpoint \( z \) of \( P \).

- If \( t \) is distance \( h \) from \( z \), then there must exist \( t_2 \in N_G(t) \) with \( t_2 \notin N_G(1) \) since \( |N_G(1)| \leq |N_G(t)| \) by Line 3 of Algorithm 1 and \( s \in N_G(1) \) but \( s \notin N_G(t) \).
- If \( t \) is distance \( h - 1 \) from \( z \), then there is a shortest path \( P_3 = tt_2 \cdots t_h \)

from \( t \) to \( t_h = z \), where \( t_2 \) is distance \( h - 2 \) from \( z \) and distance \( 2 \) from 1. In either case, we get \( t_2 \in N_G(t) \) with \( t_2 \notin N_G(1) \).

For this \( t_2 \), it follows that \( t_2 \in L_2 \). We also have \( t_2 \neq v_2 \) since \( \{t, v_2\} \notin E(G) \). Furthermore, \( \{t_2, s\} \notin E(G) \), since otherwise we would have the 4-cycle \( t_2s1tt_2 \)

with no chords as \( \{t_2, 1\}, \{t, s\} \notin E(G) \). Similarly, \( \{t_2, v_1\} \notin E(G) \) or else we would have the 4-cycle \( t_2v_11tt_2 \) with no chords since \( \{t_2, 1\}, \{t, v_1\} \notin E(G) \). Note also that \( \{t_2, v_2\} \notin E(G) \), since otherwise we would have the 5-cycle \( t_2v_2v_1tt_2 \) with
The Inductive Step. After Algorithm 1 runs on a chordal, claw-free and narrow graph \( G \), we now prove that the resulting labeling satisfies the inductive step in the proof of Theorem 4.1.

4.2. The Inductive Step. After Algorithm 1 runs on a chordal, claw-free and narrow graph \( G \), we now prove that the resulting labeling satisfies the inductive step in the proof of Theorem 4.1.

Since \( G \) is narrow, either \( t_2 \in V(P) \) or \( t_2 \) is adjacent to a vertex of \( P \). However, \( t_2 \in V(P) \) would imply \( t_2 = v_2 \) since both lie in \( L_2 \), contradicting \( t_2 \neq v_2 \). Thus \( \{t_2, v_u\} \in E(G) \) for some \( u \geq 1 \). Since \( t_2 \in L_2 \) and \( v_u \in L_u \), we have \( u \leq 3 \) by Lemma 2.6(1). We just proved \( \{t_2, v_3\} \notin E(G) \), so we must have \( \{t_2, v_3\} \notin E(G) \).

This gives the 6-cycle \( t_2v_3v_2v_1s \). Since Figure 4(b) is an induced subgraph, the only possible chords are \( \{1, v_3\}, \{t, v_3\}, \{v_1, v_3\} \). However, this gives Figure 4(b) as an induced subgraph. Hence \( 2 \) gives Figure 4(b) as an induced subgraph.

FIGURE 4. The portion of the graph relevant to Case 1B

We claim that \( \{t, v_3\} \notin E(G) \), since otherwise the 4-cycle \( 1s_2t1 \) has no chords as \( \{t, s\}, \{1, v_3\} \notin E(G) \).

Since \( G \) is narrow, \( t_2 \) must either be in \( P \) or be adjacent to a vertex in \( P \). However, \( t_2 \in V(P) \) would imply \( t_2 = v_2 \) since \( v_2 \in L_2 \), and the latter would give \( \{t_2, v_2\} = \{t, t_2\} \notin E(G) \), which we just showed to be impossible. Hence \( t_2 \notin V(P) \), so that \( \{t_2, v_u\} \notin E(G) \) for some \( u \). Note that \( u < 4 \) by Lemma 2.6(1).

We claim that \( u = 3 \).

To see why, first note that \( \{t_2, s = v_1\} \notin E(G) \), since otherwise we would have the 4-cycle \( 1s_2t1 \) with no chords as \( \{t, s\}, \{t_2, 1\} \notin E(G) \). We also know that \( \{t_2, v_2\} \notin E(G) \), as otherwise we would have the 5-cycle \( t_2v_3s_2t_2 \) with no chords since \( \{t_2, s\}, \{t_2, 1\}, \{s, t\}, \{t_2, v_2\}, \{v_2, 1\} \notin E(G) \). Once again, see Figure 3(a).

Thus we must have \( \{t_2, v_3\} \notin E(G) \). However, this gives a 6-cycle \( t_2v_3v_2s_2t_2 \) with the same impossible chords as before along with \( \{t, v_3\}, \{1, v_3\}, \{s, v_3\}, \{v_2, t_2\} \notin E(G) \), as in Figure 3(b). This contradicts chordal and completes the proof of Lemma 1.2.

4.2. The Inductive Step. After Algorithm 1 runs on a chordal, claw-free and narrow graph \( G \), we now prove that the resulting labeling satisfies the inductive step in the proof of Theorem 4.1.
Lemma 4.3. If for every vertex $1 \leq u < m$, $N_G^r(u) = \{u+1, u+r_u\}$, $r_u = |N_G^r(u)|$, and $N_G^r(u)$ is complete, then $N_G^r(m) = \{m+1, m+r_m\}$, $r_m = |N_G^r(m)|$, and $N_G^r(m)$ is complete.

**Proof.** By assumption, $N_G^r(m-1) = \{m, m-1+r_{m-1}\}$ is complete. This implies that $m+1, \ldots, m-1+r_{m-1} \in N_G^r(m)$. Furthermore, when the loop beginning on Line 11 of Algorithm 1 for $j = m-1$ terminates, we have $i = (m-1)+r_{m-1}+1 = m+r_{m-1}$ and $j$ increments to $j = m$. Then, when $S$ is set on Line 10 every unlabeled vertex in $N_G(m)$ is added to $S$. During the loop starting on Line 11 every vertex in $S$ will be labeled with consecutive integers, starting at $i = m+r_{m-1}$ and continuing until the final vertex in $N_G(m)$ is labeled $i = m+r_{m-1}+r-1$, where $r$ is the original size of $S$. Thus everything in $N_G^r(m)$ is labeled either $m+1, \ldots, m-1+r_{m-1}$ if it was labeled prior when $j < m$, or $m+r_{m-1}, m+r_{m-1}+1, \ldots, m+r_{m-1}+r-1$, if it was labeled when $j = m$. It follows that $N_G^r(m) = \{m+1, m+2, \ldots, m+r_{m-1}+r-1\}$. Since $r_{m-1}+r-1 = |N_G^r(m)| = r_m$, we have $N_G^r(m) = \{m+1, m+r_m\}$, proving the first assertion of the lemma.

To show that $N_G^r(m)$ is complete, pick distinct vertices $s, t \in N_G^r(m)$. We will prove that $\{s, t\} \in E(G)$. The arguments will be similar to the proof of Lemma 4.2 with some differences since $m > 1$.

Let $P = v_0v_1 \cdots v_{q-1}v_q$ be a shortest path from $1 = v_0$ to $v_q = m$. Thus $v_u \in L_u$. Note that $q \geq 1$ since $m > 1$. Since $s, t \in N_G^r(m)$, it follows that $s, t$ are at most distance $q + 1$ from 1. Then $s, t \in L_q$ and Lemma 3.3(3) imply that $s \in L_q$ or $L_{q+1}$, and the same holds for $t$ since $t > m$. Hence, $s$ and $t$ are either both distance $q + 1$ from 1, both distance $q$ from 1, or one of $s$ and $t$ is distance $q$ from 1 and the other is distance $q + 1$ from 1. We will consider each of these cases separately.

**Case 1.** Both $s$ and $t$ are distance $q + 1$ from 1. Then $s, t \in L_{q+1}$ and it follows that $\{s, v_q-1\}, \{t, v_q-1\} \notin E(G)$ by Lemma 8.3(1). Since the subgraph induced on $s, t, m, v_{q-1}$ cannot be a claw, we must have $\{s, t\} \in E(G)$.

**Case 2.** Both $s$ and $t$ are distance $q$ from 1, so $s, t \in L_q$. Without loss of generality, assume $s < t$ and choose a shortest path $P_1 = w_0w_1 \cdots w_q = t$ with $w_u \in L_u$. By Lemma 3.3(3), $q_{q-1} < m$ since $m \in L_q$ and $w_{q-1} \in L_{q-1}$. Then since $w_{q-1} < m < s < t$ and $t \in N_G^r(w_{q-1})$. Since $N_G^r(w_{q-1})$ is an interval by hypothesis, we have $s \in N_G^r(w_{q-1})$. But then $\{s, t\} \in E(G)$ since we are also assuming that $N_G^r(w_{q-1})$ is complete.

**Case 3.** Without loss of generality, let $s$ be distance $q$ from 1 and $t$ distance $q + 1$ from 1, so $s \in L_q$ and $t \in L_{q+1}$, and hence $s < t$ by Lemma 3.3(4). We also have $l(m) \leq l(s)$ by Lemma 3.3(1) since $m < s$. We will consider separately the two possibilities that $l(m) < l(s)$ and $l(m) = l(s)$.

**Case 3A.** Suppose that $l(m) < l(s)$. Then $\{l(m), s\} \notin E(G)$ since $l(s) = \min(N_G(s))$ by Lemma 3.2. We also claim that $\{l(m), t\} \notin E(G)$. This follows because if $\{l(m), t\} \in E(G)$, then $l(t) \leq l(m)$ since $l(t) = \min(N_G(t))$. But then we would have $l(t) \leq l(m) < l(s)$, which by Lemma 3.3(1) would imply $t < s$, contradicting $s < t$. Hence $\{l(m), s\}, \{l(m), t\} \notin E(G)$. Since the subgraph induced on $l(m), m, s, t$ cannot be a claw, we must have $\{s, t\} \in E(G)$.

**Case 3B.** Suppose that $l(m) = l(s)$. We will assume $\{s, t\} \notin E(G)$ and derive a contradiction. The equality $l(m) = l(s)$ means that $m$ and $s$ were both labeled when $j = l(m) = l(s)$ in the loop starting on Line 9 of Algorithm 1. Consider the moment in the algorithm when the label $m$ is assigned. Since $m < s$ and $j = l(m) = l(s)$, this happens during an iteration of the loop on Line 11 for which
Figure 5. The subgraph induced on \( j = l(m) = l(s), m, s, t, s_2 \).

\( m, s \in S \). Line 12 guarantees that the vertices assigned the labels \( m \) and \( s \) satisfy 
\( |N_G(m) \setminus J| \leq |N_G(s) \setminus J| \). Since \( s \) is not yet labeled at this point and \( s < t, t \) is also not yet labeled and therefore \( t \notin J \). It follows that \( t \in N_G(m) \setminus J \) and \( t \notin N_G(s) \setminus J \). But, in order for \( |N_G(m) \setminus J| \leq |N_G(s) \setminus J| \) to hold, there must be \( s_2 \in N_G(s) \) with \( s_2 > m \) and \( s_2 \notin J \) and \( s_2 \notin N_G(m) \).

Let us study \( s_2 \). If \( \{s_2, l(m)\} \in E(G) \), then \( s_2 \in N_G(l(m)) \). But we also have \( m \in N_G^2(l(m)) \). Since \( l(m) < m \), \( N_G^2(l(m)) \) is complete by the hypothesis of the lemma, so we would have \( \{m, s_2\} \in E(G) \). This contradicts our choice of \( s_2 \). Hence \( \{s_2, l(m)\} \notin E(G) \). We also have \( \{s_2, t\} \notin E(G) \), since otherwise the 4-cycle \( s_2ts_2m \) would have no chords as \( \{s_2, m\}, \{s, t\} \notin E(G) \). Also, since \( m \in L_q \), Lemma 3.3(2) implies that \( j = l(m) = l(s) \in L_{q-1} \).

We also claim that \( s_2 \in L_{q+1} \). Lemma 2.6(1), \( s \in L_q \), and \( s_2 > m \in L_q \) imply that \( s_2 \in L_q \) or \( L_{q+1} \). If \( s_2 \in L_q \), then \( l(s_2) \in L_{q-1} \) by Lemma 3.3(2). From here, \( m \in L_q \) implies \( m > l(s_2) \) by Lemma 3.3(3). Hence we have \( l(s_2) < m < s_2 \).

The hypothesis of the lemma implies that \( N_G^2(l(s_2)) \) is complete and is an interval. Since \( s_2 \in N_G^2(l(s_2)) \), it follows that \( m \in N_G^2(l(s_2)) \), which contradicts our choice of \( s_2 \). Hence \( s_2 \in L_{q+1} \) and we have Figure 5.

Let \( z \) be a vertex of distance \( h = \text{diam}(G) \) from 1 and pick a longest shortest path \( P_2 = w_0w_1 \cdots w_h \) from 1 = \( w_0 \) to \( w_h = z \), so \( w_u \in L_u \). Since \( G \) is narrow, \( t \) and \( s_2 \) must each either be in \( P_2 \) or be adjacent to a vertex in \( P_2 \). We will consider each of these cases. However, we first note that \( h > 2 \), since \( h = 2 \) and \( G \) narrow imply that \( s_2 \) must be adjacent to \( t, m \), or \( l(m) \) (since \( l(m) = 1 \) in this situation), all of which we already showed cannot occur.

First, suppose that \( t \in V(P_2) \). Then \( t \in L_{q+1} \) implies that \( t = w_{q+1} \). Since \( l(m) \in L_{q-1} \), there is a path of length \( q - 1 \) connecting 1 to \( l(m) \). Using \( t = w_{q+1} \), it follows that \( P_3 = 1 \cdots l(m)mtw_{q+2} \cdots z \) is a path of length \( h = \text{diam}(G) \). Since \( G \) is narrow, \( s_2 \) must be adjacent to some vertex \( P_3 \). Then \( \{s_2, t\}, \{s_2, m\} \notin E(G) \) and Lemma 2.6(1) imply that \( \{s_2, w_{q+2}\} \in E(G) \). This gives the 5-cycle \( mss_2w_{q+2}tm \) with no chords since \( \{s_2, t\}, \{m, s_2\}, \{s, t\} \notin E(G) \) and \( \{w_{q+2}, m\}, \{w_{q+2}, s\} \notin E(G) \) since \( w_{q+2} \in L_{q+2} \) but \( s, m \in L_q \). See Figure 6. Hence we have a contradiction since \( G \) is chordal.

Second, suppose that \( s_2 \in V(P_2) \). Then \( s_2 = w_{q+1} \). Arguing as in the First, we arrive at Figure 6 with the same 5-cycle with no chords, again a contradiction.

Third, suppose that \( s_2, t \notin V(P_2) \). First note that \( P_2 \) was an arbitrary longest shortest path starting at 1. Thus the above First and Second give a contradiction.
whenever $s_2$ or $t$ are on any longest shortest path starting at 1. Hence we may assume that $s_2$ and $t$ are not on any shortest path of length $h$ starting at 1.

Since $G$ is narrow, $s_2 \in L_{q+1}$ is adjacent to a vertex of $P_2$, which must be $w_q$, $w_{q+1}$, or $w_{q+2}$ by Lemma 2.6. However, if $\{s_2, w_{q+2}\} \in E(G)$, then we would get a path of length $h$ from 1 to $z$ by taking any shortest path from 1 to $s_2$, followed by $\{s_2, w_{q+2}\}$, and then continuing along $P_2$ from $w_{q+2}$ to $z$. This longest shortest path starts at 1 and contains $s_2$, contradicting the previous paragraph. Hence $\{s_2, w_{q+2}\} \notin E(G)$ and $s_2$ must be adjacent to $w_q$ or $w_{q+1}$, and the same is true for $t$ by a similar argument.

In fact, we must have $\{s_2, w_q\} \in E(G)$, since otherwise $\{s_2, w_{q+1}\} \in E(G)$ and the subgraph induced on $w_q, w_{q+1}, w_{q+2}, s_2$ would be a claw. A similar argument shows that $\{t, w_q\} \in E(G)$. Since $w_{q-1} \in L_{q-1}$ and $s_2, t \in L_{q+1}$, this implies that the subgraph induced on $t, s_2, w_q, w_{q-1}$ is a claw, again contradicting claw-free. This final contradiction completes the proof of Lemma 4.3.

In Lemmas 4.2 and 4.3, the chordal hypothesis is applied only to cycles of length 4, 5, or 6. Hence, in Theorem 1.4 and Corollary 1.5, we can replace chordal with the weaker hypothesis that all cycles of length 4, 5, or 6 have a chord.

5. Further Study of Closed Graphs

5.1. Exchangeable Vertices. A closed graph with $\geq 2$ vertices has at least two closed labelings, since the reversal of a closed labeling is clearly closed. But there may be other closed labelings, as shown by the simple example

\begin{figure}[h]
\centering
\begin{tikzpicture}
\node (w) at (0,0) {$w_{q+2}$};
\node (s) at (1,1) {$s_2$};
\node (m) at (0,-1) {$m$};
\node (j) at (1,-2) {$j$};
\node (l) at (-1,-1) {$L_q$};
\node (p) at (1,2) {$L_{q+2}$};
\node (q) at (2,1) {$L_{q+1}$};
\node (r) at (3,0) {$L_q$};
\node (s) at (4,1) {$L_{q-1}$};
\draw (w) -- (s) -- (p);
\draw (m) -- (j) -- (l);
\draw (s) -- (q);
\draw (m) -- (r);
\end{tikzpicture}
\caption{A 5-cycle with no chords.}
\end{figure}
To explore what makes this work, we need some definitions.

**Definition 5.1.** Let $G$ be a graph.

1. The **full neighborhood** of a vertex $v \in V(G)$ is the set
   
   $$N^*_G(v) = \{v\} \cup N_G(v).$$

2. $v, w \in V(G)$ are **exchangeable**, written $v \sim w$, if $N^*_G(v) = N^*_G(w)$.

Note that in the left-hand graph of (5.1), vertices 1 and 2 are exchangeable, which is why when we switch them to obtain the right-hand graph of (5.1), we still have a closed graph. Here is the general result.

**Proposition 5.2.** Let $G$ be a graph with a closed labeling and $V(G) = [n]$. If $i, j \in V(G)$, $i \neq j$, are exchangeable, then the labeling that switches $i$ and $j$ is also closed.

**Proof.** Define $\phi : [n] \to [n]$ by $\phi(i) = j$, $\phi(j) = i$, and $\phi(k) = k$ for $k \in [n] \setminus \{i, j\}$. Pick $u, v, w \in V(G)$ with $\{u, v\}, \{v, w\} \in E(G)$, $u \neq w$, and $\phi(u) > \phi(v) < \phi(w)$ or $\phi(u) < \phi(v) > \phi(w)$. We need to prove that $\{u, w\} \in E(G)$.

If $\{i, j\} \cap \{u, v, w\} = \emptyset$, then $u \neq v < w$ or $u < v < w$. Hence $\{u, w\} \in E(G)$ since the original labeling is closed.

We next consider the case when $\{i, j\} \cap \{u, v, w\} \neq \emptyset$. For the time being, we will assume that $\phi(u) > \phi(v) < \phi(w)$. There are several cases to consider.

First suppose that $i = v$. If $j$ is one of $u, w$, then without loss of generality we may assume $j = u$. Then

$$w \in N^*_G(v) = N^*_G(i) = N^*_G(j) = N^*_G(u)$$

implies $\{u, w\} \in E(G)$. If $j$ is neither of $u, w$, then $\phi(u) > \phi(i) < \phi(w)$ means that $u > j < w$. Then $\{u, w\} \in E(G)$ since the original labeling is closed and $j \sim i = v$.

The proof when $j = v$ is similar and is omitted. Without loss of generality, two cases remain:

- $i = u$ and $j$ is neither of $v, w$. Thus $\phi(u) > \phi(v) < \phi(w)$ means that $j > v < w$. Then $\{j, w\} \in E(G)$ since the original labeling is closed and $j \sim i = u$. Using $j \sim i = u$ again, we conclude that $\{u, w\} \in E(G)$.
- $i = u$ and $j = w$. Then $\phi(u) > \phi(v) < \phi(w)$ means $j > v < i$. Then $\{u, w\} = \{i, j\} \in E(G)$ since the original labeling is closed.

The proof when $\phi(u) < \phi(v) > \phi(w)$ is similar and is omitted. \qed

Exchangeability $v \sim w$ is clearly an equivalence relation on $V(G)$, and the equivalence class of $v$ will be denoted

$$e(v) = \{w \in V(G) \mid w \sim v\} = \{w \in V(G) \mid N^*_G(w) = N^*_G(v)\}.$$

Equivalence classes are complete, since $v \sim w$ implies $v \in N^*_G(v) = N^*_G(w)$, so that if in addition $v \neq w$, then $\{v, w\} \in E(G)$.

Since permutations are generated by transpositions, Proposition 5.2 implies that when $G$ has a closed labeling, every permutation of $e(i)$ yields a new closed labeling. When $G$ is connected and closed, equivalence classes have the following structure.

**Proposition 5.3.** Let $G$ be a connected graph with a closed labeling and $V(G) = [n]$. Then for every $i \in [n]$, the equivalence class $e(i)$ is an interval.
Proof. It suffices to show that if $i < j$ are exchangeable and $i < k < j$, then $N_G^*(k) = N_G^*(i)$. First note that $\{i, k\} \in E(G)$ since $j \in N_G^*(i)$ and $N_G^*(i)$ is an interval by Proposition 2.7. Then $\{j, k\} \in E(G)$ since $i \sim j$.

Now take $m \in N_G^*(k)$. We need to show $m \in N_G^*(i)$. If $m = k$, this follows from the previous paragraph. If $\{m, k\} \in E(G)$, there are two possibilities:

- If $m < k$, then $m < k > i$, so $\{m, i\} \in E(G)$ since the labeling is closed.
- If $m > k$, then $m > k < j$, so either $m = j$ or $\{m, j\} \in E(G)$ by closed.

Since $N_G^*(i) = N_G^*(j)$, both possibilities imply $m \in N_G^*(i)$.

Conversely, take $m \in N_G^*(i)$. If $m = i$, then we are done since $\{i, k\} \in E(G)$ by the first paragraph of the proof. If $\{m, i\} \in E(G)$, then $\{m, j\} \in E(G)$ since $i \sim j$.

Again, there are two possibilities:

- If $m < i$, then $m < i < k < j$, so $\{m, k\} \in E(G)$ since $N_G^*(m)$ is an interval.
- If $m > i$, then $m > i < k$, so either $m = k$ or $\{m, k\} \in E(G)$ by closed.

Thus $m \in N_G^*(k)$ and the proof is complete. \hfill \Box

5.2. Collapsed Graphs. Some graphs have no nontrivial exchangeable vertices.

Definition 5.4. A graph $G$ is collapsed if all exchangeable vertices are equal, i.e., $N_G^*(v) = N_G^*(w)$ implies $v = w$.

Theorem 5.5. Let $G$ be a closed graph with $\geq 3$ vertices. Then the following are equivalent:

1. $G$ has exactly two closed labelings.
2. $G$ is connected and collapsed.

Proof. The proof of (1) $\Rightarrow$ (2) is easy. If $G$ is not connected, then $G$ is a disjoint union $G = G_1 \cup G_2$, where $G_i$ is closed. We may assume $G_1$ has at least two vertices, so $G_1$ has at least two labelings. Then we get at least four closed labelings of $G$: two where 1 is in $G_1$, and two where 1 is in $G_2$. Also, if $G$ is not collapsed, then some equivalence class $e(i)$ has at least two elements. If $|e(i)| \geq 3$, then switching labels within $e(i)$ gives at least 6 closed labelings, and if $|e(i)| = 2$, then $G$ has at least one more vertex, which makes it easy to see $G$ has at least four closed labelings.

The proof of (2) $\Rightarrow$ (1) will take more work. First note that $G$ has diameter $\text{diam}(G) = h \geq 2$. This follows because $h = 1$ would imply that $G$ is complete, which is impossible since $G$ is collapsed with $\geq 3$ vertices, and $h = 0$ is impossible since $G$ is connected with $\geq 3$ vertices.

Fix a closed labeling with $V(G) = [n]$. This gives layers $L_0 = \{1\}, L_1, \ldots, L_h$ associated with the labeling, and Proposition 2.8 implies that every longest shortest path has one endpoint in $L_0$ or $L_1$ and the other in $L_h$.

Let $\phi : [n] \to [n]$ be another closed labeling which we will call the $\phi$-labeling. Pick $1' \in [n]$ such that $l(1') = 1$. Then some longest shortest path of $G$ begins at $1'$. By the previous paragraph, $1' \in L_0 \cup L_1$ or $1' \in L_h$. Replacing $\phi$ with its reversal if necessary, we may assume that $1' \in L_0 \cup L_1$. We claim that $\phi$ is the identity function. This will prove the theorem.

We first show that $1' = 1$, i.e., $\phi(1) = 1$. Recall that $L_1 = N_G^*(1)$ and that $L_1$ is complete by Proposition 2.7. It follows that $N_G^*(1) = L_0 \cup L_1$ is also complete. The same argument implies that $N_G^*(1')$ is complete. Now suppose $1 \neq 1'$ and pick $m \in N_G^*(1')$ different from 1. Then $\{1, m\} \in E(G)$ since $1 \in N_G^*(1')$ and $N_G^*(1')$ is complete. This implies $m \in L_1 = N_G^*(1)$, and then the inclusion
$N_G^*(1') \subseteq N_G^*(1)$ follows easily. The opposite inclusion follows by interchanging the two labelings. Hence we have proved $N_G^*(1') = N_G^*(1)$. Since we are assuming $1 \neq 1'$, this contradicts Proposition 5.3, we can order the equivalence classes.

Now suppose that vertices $1, \ldots, u - 1 \in [n]$ have the same $\phi$-label as in the original labeling, i.e., $\phi(j) = j$ for $1 \leq j \leq u - 1$. Then pick $u' \in [n]$ such that $\phi(u') = u$. To prove that $u' = u$, i.e., $\phi(u) = u$, suppose that $u' \neq u$. Since $\phi$ is the identity on $1, \ldots, u - 1$ and $\phi(u') = u$, we have $u' > u$ and $\phi(u') < \phi(u)$.

We first show that $\{u, u'\} \in E(G)$. Since $G$ is connected, Proposition 2.4 implies that every vertex is connected by an edge to its successor in any closed labeling. For the original labeling, this gives $\{u - 1, u\} \in E(G)$, and for the $\phi$-labeling, this gives $\{u - 1, u'\} \in E(G)$ since $\phi(u - 1) = u - 1$ and $\phi(u') = u$. Proposition 2.4 implies that $N_G^*(u - 1)$ (in the original labeling) is complete, and $\{u, u'\} \in E(G)$ follows.

We next prove that $N_G^*(u) \subseteq N_G^*(u')$. Pick $m \in N_G^*(u)$. Then:

- If $m = u$, then $m \in N_G^*(u')$ since $\{u, u'\} \in E(G)$.
- If $m > u$, then either $m = u'$, in which case $m \in N_G^*(u')$ is obvious, or $m \neq u'$, in which case $m \in N_G^*(u')$ since $m > u < u'$ implies $\{m, u'\} \in E(G)$ as the original labeling is closed.
- If $m < u$, then $m \in N_G^*(u')$ since $\phi(m) = m < u < \phi(u) > \phi(u')$ implies $\{m, u'\} \in E(G)$ as the $\phi$-labeling is closed.

This proves $N_G^*(u) \subseteq N_G^*(u')$. By symmetry, we get $N_G^*(u') = N_G^*(u)$, which contradicts $u' \neq u$ since $G$ is collapsed. We conclude that $u' = u$, and then $\phi$ is the identity by induction on $u$. This completes the proof.

Now suppose that $G$ is a connected graph with a closed labeling. Since each equivalence class is an interval by Proposition 5.3, we can order the equivalence classes

\[(5.2) \quad E_1 < E_2 < \cdots < E_r\]

so that if $i \in E_a$ and $j \in E_b$, then $i < j$ if and only if $a < b$. This induces an ordering on $V(G)/\sim = \{E_1, \ldots, E_r\}$. Then define the graph $G/\sim$ with vertices

\[(5.3) \quad V(G/\sim) = V(G)/\sim = \{E_1, \ldots, E_r\}\]

and edges

\[(5.4) \quad E(G/\sim) = \{\{E_a, E_b\} \mid \{i, j\} \in E(G) \text{ for some } i \in E_a, j \in E_b\}.\]

Since $i \sim i'$ and $j \sim j'$ imply that $\{i, j\} \in E(G)$ if and only if $\{i', j'\} \in E(G)$, we can replace “for some” with “for all” in (5.4).

**Proposition 5.6.** Let $G$ be a connected graph with a closed labeling and exchangeable equivalence classes $E_1, \ldots, E_r$. Then:

1. The quotient graph $G/\sim$ defined in (5.3) and (5.4) is connected, collapsed, and closed with respect to the labeling (5.2).
2. If $r > 1$, then $G$ has precisely

\[2 \prod_{a=1}^{r} |E_a|!\]

closed labelings.
Proposition 5.8. Let \( G \) be a connected graph with a closed labeling and \( V(G) = [n] \). If \( u_s = m_N + s - 1 \in L_N \) is the \( s^{th} \) vertex of \( L_N \) and \( b_s > 0 \), then
\[
\{ v \in L_{N+1} \mid \{ u_s, v \} \in E(G) \} = [m_{N+1}, m_{N+1} + b_s - 1].
\]
Thus \( b_s \) determines how \( u_s \) links to \( L_{N+1} \).
Proposition 5.9. Let $A = \{v \in L_{N+1} \mid \{u_s, v\} \in E(G)\}$. Note that every $v \in A$ satisfies $v > u_s$ by (2.1). It follows easily that

$$A = N_G^>(u_s) \cap L_{N+1}.$$  

We know that $L_{N+1}$ is an interval, and the same is true for $N_G^>(u_s)$ by Proposition 2.3. Hence $A$ is an interval. However, if $v \in A$ and $v \neq m_{N+1}$, then $m_{N+1} < v > u_s$ and closed imply $\{u_s, m_{N+1}\} \in E(G)$ since $\{m_{N+1}, v\} \in E(G)$ by the completeness of $L_{N+1}$. Hence $m_{N+1} \in A$, and from here, the proposition follows without difficulty. □

Here is an important property of the sequence $S_N$.

Proposition 5.9. Let $G$ be a graph with a closed labeling and $V(G) = [n]$. If $N < \text{diam}(G)$, then the sequence $S_N = (b_1, b_2, \ldots, b_{N+1})$ of the layer $L_N$ has the following properties:

1. The last element of $S_N$ is $a_{N+1}$, i.e., $b_{a_{N+1}} = a_{N+1}$.
2. $S_N$ is increasing, i.e., $b_s \leq b_{s+1}$ for $s = 1, \ldots, a_N - 1$.

Proof. For (1), note that the last vertex of $L_N$ connects to every vertex of $L_{N+1}$ by Proposition 2.7(2). It follows that $b_{a_N} = |L_{N+1}| = a_{N+1}$.

For (2), let $u_s$ be the $s$th vertex of $L_N$, $1 \leq s \leq a_N - 1$. If $b_s = 0$, then $b_s \leq b_{s+1}$ clearly holds. If $b_s > 0$, then $u_s$ connects to $m_{N+1} + b_s - 1$ by Proposition 5.8 and it connects to $u_{s+1}$ since $L_N$ is complete. Then $m_{N+1} + b_s - 1 > u_s < u_{s+1}$ implies that $u_{s+1}$ connects to $m_{N+1} + b_s - 1$ since the labeling is closed. Using Proposition 5.8 again, we obtain

$$m_{N+1} + b_s - 1 \in [m_{N+1}, m_{N+1} + b_{s+1} - 1],$$

and $b_s \leq b_{s+1}$ follows. □

Theorem 5.10. Fix $n$ and an integer partition $n = a_0 + a_1 + \cdots + a_h$ with $a_0 = 1$ and $a_N \geq 1$ for $N = 1, \ldots, h$. Also set $\mathcal{L}_0 = \{1\}$ and

$$\mathcal{L}_N = [a_0 + \cdots + a_{N-1} + 1, a_0 + \cdots + a_N]$$

for $N = 1, \ldots, h$, so that $|\mathcal{L}_N| = a_N$. Then the number of graphs $G$ satisfying the conditions:

1. $V(G) = [n]$,
2. $G$ is connected and closed with respect to the labeling $V(G) = [n]$, and
3. The $N$th layer of $G$ is $\mathcal{L}_N$ for $N = 0, \ldots, h$,

is given by the product

$$\prod_{N=0}^{h-1} \left(\frac{a_{N+1} + a_N - 1}{a_N - 1}\right).$$

Proof. Let $G$ satisfy (1), (2) and (3). Each layer of $G$ is complete, and every edge of $G$ connects to the same layer or an adjacent layer by Lemma 2.6(1). Then Proposition 5.8 shows that the edges of $G$ are uniquely determined by $S_0, \ldots, S_{h-1}$.

By Proposition 5.9 each $S_N = (b_1, b_2, \ldots, b_{N+1})$ is an increasing sequence of non-negative integers of length $a_N$ that ends at $a_{N+1}$. It is well known that the number of such sequences equals the binomial coefficient $\binom{a_{N+1} + a_N - 1}{a_N - 1}$. This follows, for
example, since such sequences correspond bijectively to monomials in \( x_1, \ldots, x_a \) of degree \( a+1 \) by encoding \((b_1, b_2, \ldots, b_a)\) into the monomial
\[
x_1^{b_{a+1}-b_a} x_2^{b_a-b_{a-1}} \cdots x_{a+1}^{b_2-b_1} = x_{a+1}^{a+1-a} = x_{a+1}.
\]
of degree \( b_{a+1} = a+1 \).

It follows that the product in the statement of the proposition is an upper bound for the number of graphs satisfying (1), (2) and (3).

To complete the proof, we need to show that every sequence counted by the product corresponds to a graph \( G \) satisfying (1), (2) and (3). First note that the minimal element of \( \mathcal{L}_N \) is
\[
m_N = a_0 + \cdots + a_{N-1} + 1
\]
when \( N > 0 \). Now suppose we have sequences \( S_0, \ldots, S_{h-1} \), where each \( S_N = (b_1, b_2, \ldots, b_{a_N}) \) is an increasing sequence of nonnegative integers of length \( a_N \) that ends at \( a_{N+1} \). This determines a graph \( G \) with \( V(G) = [n] \) and the following edges:

(A) All possible edges connecting elements in the same level \( \mathcal{L}_N \).

(B) For each \( N = 0, \ldots, h-1 \), all edges \( \{u_s, v\} \), where \( u_s \) is the \( s \)-th vertex of \( \mathcal{L}_N \) and \( v \) is any vertex in the interval \([m_N+1, m_{N+1}+b_s-1] \subseteq \mathcal{L}_{N+1}\) from Proposition 5.8.

Once we prove that \( G \) is closed and connected with \( \mathcal{L}_N \) as its \( N \)-th layer, the theorem will be proved.

Since \( b_{a_N} = a_{N+1} \), we see that for \( N = 0, \ldots, h-1 \), the last element of \( \mathcal{L}_N \) connects to all elements of \( \mathcal{L}_{N+1} \). This allows us to construct a path from 1 to any \( u \) in \( \mathcal{L}_N \) for \( N = 1, \ldots, h \). It follows that \( G \) is connected and that all \( u \in \mathcal{L}_N \) satisfy \( d(1, u) \leq N \) when \( N > 0 \). Since every edge of \( G \) connects to the same layer or an adjacent layer, any path connecting 1 to \( u \) must have length at least \( N \). It follows that \( \mathcal{L}_N \) is indeed the \( N \)-th layer of \( G \).

It remains to show that \( G \) is closed with respect to the natural labeling given by \( V(G) = [n] \). A vertex of \( G \) is the \( s \)-th vertex \( u_s \) of \( \mathcal{L}_N \) for some \( s \) and \( N \). We will show that \( N_{G} (u_s) \) satisfies Proposition 2.4. The formula (5.6) for \( \mathcal{L}_N \) and the description of the edges of \( G \) given in (A) and (B) make it clear that
\[
N_{G} (u_s) = [u_{s+1}, a_0 + \cdots + a_N] \cup [m_N+1, m_{N+1}+b_s-1]
\]
where the second equality follows from \( m_{N+1} = a_0 + \cdots + a_N + 1 \). To show that \( N_{G} (u_s) \) is complete, take distinct vertices \( v, w \in N_{G} (u_s) \). If both lie in \( \mathcal{L}_N \) or \( \mathcal{L}_{N+1} \), then \( \{v, w\} \in V(G) \) by (A). Otherwise, we may assume without loss of generality that \( v = u_t \), \( t \geq s \), and \( w \in [m_N+1, m_{N+1}+b_s-1] \). Note that \( u_t \) links to every vertex in \([m_N+1, m_{N+1}+b_t-1]\) by (B). We also have \( b_s \leq b_t \) since \( S_N \) is increasing. It follows that \( \{v, w\} = \{u_t, w\} \in E(G) \). Hence \( N_{G} (u_s) \) is complete, so that \( G \) is closed by Proposition 2.4. \( \square \)

5.4. Local Clustering Coefficients. In a social network, one can ask how often a friend of a friend is also a friend. Translated into graph theory, this asks how often a path of length two has an edge connecting the endpoints of the path. The illustration [1] from the Introduction indicates that this should be a frequent occurrence in a closed graph.

There are several ways to quantify the “friend of a friend” phenomenon. For our purposes, the most convenient is the local clustering coefficient of vertex \( v \) of a
graph \( G \), which is defined by

\[
C_v = \begin{cases} 
\frac{\text{number of pairs of neighbors of } v \text{ connected by an edge}}{\text{number of neighbors of } v} & \text{deg}(v) \geq 2 \\
0 & \text{deg}(v) \leq 1.
\end{cases}
\]

Local clustering coefficients are discussed in [4] pp. 201–204.

**Proposition 5.11.** Let \( v \) be a vertex of a closed graph \( G \) of degree \( d = \text{deg}(v) \geq 2 \). Then the local clustering coefficient \( C_v \) satisfies the inequality

\[
C_v \geq \frac{1}{2} - \frac{1}{2(d-1)}.
\]

Furthermore, \( d \geq 3 \) implies that \( C_v \geq \frac{1}{4} \).

**Proof.** Pick a closed labeling of \( G \) and let \( a = |N_G^2(v)| \) and \( b = |N_G^3(v)| \). Then \( a + b = |N_G(v)| = \text{deg}(v) = d \). Since the labeling is closed, any pair of vertices in \( N_G^2(v) \) or in \( N_G^3(v) \) is connected by an edge. It follows that at least

\[
\frac{1}{2}a(a-1) + \frac{1}{2}b(b-1)
\]

pairs of neighbors of \( v \) are connected by an edge. Since the total number of such pairs is \( \frac{1}{2}d(d-1) \) and \( d = a+b \), we obtain

\[
C_v \geq \frac{a(a-1) + b(b-1)}{d(d-1)} = \frac{a^2 + b^2 - d}{d(d-1)} \geq \frac{\frac{1}{2}d^2 - d}{d(d-1)} = \frac{1}{2} - \frac{1}{2(d-1)},
\]

where we use \( a^2 + b^2 - \frac{1}{2}d^2 = \frac{1}{2}(a-b)^2 \geq 0 \). When \( d \geq 4 \), this inequality for \( C_v \) easily gives \( C_v \geq \frac{1}{4} \). When \( d = 3 \), then \( a + b = 3 \), \( a, b \in \mathbb{Z} \), implies that \( a^2 + b^2 \geq 5 \), in which case the left half of (5.7) gives \( C_v \geq \frac{5-3}{3(3-1)} = \frac{1}{3} \). \( \square \)

A global version of the clustering coefficient defined by Watts and Strogatz is

\[
C_{WS} = \frac{1}{n} \sum_{v \in V(G)} C_v, \quad n = |V(G)|.
\]

(See reference [323] of [4]. A different global clustering coefficient is discussed in [4] pp. 199–204.) To estimate \( C_{WS} \) for a closed graph, we need the the following lemma.

**Lemma 5.12.** Let \( G \) be a connected closed graph of diameter \( \text{diam}(G) = h \) and let \( c \) be the number of vertices \( v \in G \) with \( \text{deg}(v) = 2 \) and \( C_v = 0 \). Then \( c \leq h - 1 \).

**Proof.** Fix a closed labeling for \( G \) with \( V(G) = [n] \) and pick \( v \in V(G) \) with \( \text{deg}(v) = 2 \) and \( C_v = 0 \). We claim that \( v \) is in a layer of its own. To see why, let \( v \in L_N \) and suppose there is \( s \in L_N \) with \( s \neq v \). Then \( \{v,s\} \in E(G) \) since layers are complete by Proposition 2.7(1). Furthermore, \( |L_N| \geq 2 \), so \( N > 0 \). Then \( \{s,d\}, \{v,d\} \in E(G) \) for \( d = \max\{|L_{N-1}|\} \) by Proposition 2.7(2). Since \( \text{deg}(v) = 2 \), we must have \( N_G(v) = \{s,d\} \), and then \( \{s,d\} \in E(G) \) contradicts \( C_v = 0 \). Thus \( \{v\} \) is a layer when \( \text{deg}(v) = 2 \) and \( C_v = 0 \).

Note that if \( \{v\} = L_0 \), then the two vertices in \( N_G(v) = L_1 \) would be linked by an edge. The same holds if \( \{v\} = L_h \), for here the two vertices would be in \( L_{h-1} \) since \( L_h \) is the highest layer by Proposition 2.8. It follows that each of the \( c \) vertices with \( \text{deg}(v) = 2 \) and \( C_v = 0 \) requires a separate layer distinct from \( L_0 \) or \( L_h \). Since there are only \( h - 1 \) intermediate layers, we must have \( c \leq h - 1 \). \( \square \)
Proposition 5.13. Let $G$ be a connected closed graph $G$ with $n > 1$ vertices and diameter $h$. Then
\[ C_{WS} \geq \frac{1}{3} - \frac{h + 1}{3n}. \]

Proof. Since $n > 1$ and $G$ is connected, all vertices of $G$ have degree $\geq 1$. Thus we can write $V(G)$ as the disjoint union
\[ V(G) = A \cup B \cup C \cup D, \]
where $A$ consists of vertices of degree $\geq 3$, $B$ consists of vertices of degree 2 with $C_v = 1$, $C$ consists of vertices of degree 2 with $C_v = 0$, and $D$ consists of the leaves (which have $C_v = 0$). Since $C_v \geq \frac{1}{3}$ for $v \in A$ by Proposition 5.11 we have
\[ C_{WS} \geq \frac{1}{n} \left( \frac{1}{3} \cdot |A| + 1 \cdot |B| + 0 \cdot |C| + 0 \cdot |D| \right) \geq \frac{|A| + |B|}{3n} = \frac{n - (|C| + |D|)}{3n}. \]
However, Lemma 5.12 implies $|C| \leq h - 1$ and Corollary 2.2 implies $|D| \leq 2$. The proposition follows immediately. 

This proposition shows that the clustering coefficient $C_{WS}$ is reasonably large when the diameter is small compared to the number of vertices. At the other extreme, one easily checks that the inequality of Proposition 5.13 is an equality when $G$ is a path graph (both sides are zero).

Acknowledgements

Theorem 1.4 is based on the senior honors thesis of the second author, written under the direction of the first author. We are grateful to Amherst College for the Post-Baccalaureate Summer Research Fellowship that supported the writing of this paper. Thanks also to Amy Wagaman for suggesting that we look at clustering coefficients and to Michael Ching for the comment at the end of Section 4.

References

[1] M. Crupi and G. Rinaldo, Binomial edge ideals with quadratic Gröbner bases, Electron. J. Combin. 18 (2011), Paper 211, 13pp.
[2] V. Ene, J. Herzog and T. Hibi, Cohen-Macaulay binomial edge ideals, Nagoya Math. J. 204 (2011), 57–68.
[3] J. Herzog, T. Hibi, F. Hreinsdóttir, T. Kahle and J. Rauh, Binomial edge ideals and conditional independence statements, Adv. in Appl. Math. 45 (2010), 317–333.
[4] M. E. J. Newman, Networks: An Introduction, Oxford Univ. Press, Oxford, 2010.
[5] M. Ohtani, Graphs and ideals generated by some 2-minors, Commun. Algebra 39 (2011), 905–917.
[6] J. Rauh, Generalized binomial edge ideals, Adv. in Appl. Math. 50 (2013), 409–414.
[7] S. Saeedi Madani and D. Kiani, Binomial edge ideals of graphs, Electron. J. Combin. 19 (2012), Paper 44, 6pp.

Department of Mathematics, Amherst College, Amherst, MA 01002-5000, USA
E-mail address: dac@math.amherst.edu

Department of Mathematics, Amherst College, Amherst, MA 01002-5000, USA
E-mail address: aerskine13@amherst.edu