SUBGRAPHS OF COPRIME GRAPHS ON SETS OF CONSECUTIVE INTEGERS

Ethan Berkove
Department of Mathematics, Lafayette College, Easton, Pennsylvania
berkovee@lafayette.edu

Michael Brilleslyper
Department of Applied Mathematics, Florida Polytechnic University, Lakeland, Florida
mbrilleslyper@floridapoly.edu

Received: 7/10/21, Accepted: 4/18/22, Published: 5/20/22

Abstract

Let $\mathcal{A}_n^k$ denote the set of k consecutive positive integers starting with n. The coprime graph associated to this set, $G(\mathcal{A}_n^k)$, is the graph whose vertices are elements of $\mathcal{A}_n^k$ and whose edges connect pairs of integers if and only if they are coprime. In this paper, we focus on complete subgraphs and complete bipartite subgraphs of $G(\mathcal{A}_n^k)$. We investigate how the sizes of these graphs change with n and k, and show that there are situations where there is no complete bipartite subgraph with k vertices. This happens precisely when $\mathcal{A}_n^k$ is a *stapled sequence*; we provide some numerical results on how frequently these occur. We also prove a result about the average size of the smaller bipartition of the most balanced bipartite subgraph.

1. Introduction

Many problems in graph theory involve assigning labels to the components of a graph subject to some set of conditions. Collectively, these are known as *graph labeling problems*, and they are well-studied. Gallian maintains a regularly updated survey of graph labeling results [12] which, as this paper was being written, contained a bibliography listing over 3000 references. In a *coprime labeling*, the vertices of a graph $G(V, E)$ are labeled with distinct elements of $\{1, 2, \ldots, m\}$, where $m \geq |V|$, in such a way that when v and w are adjacent, then the integer labels associated to v and w are relatively prime. (In this paper we will usually simply refer to a vertex by its label.) The minimum value of m for which a coprime labeling of G is possible is denoted $\mathfrak{pr}(G)$, the *minimum coprime number*. When $m = |V|$, the labeling is called a *prime labeling*. Prime labelings were first introduced by Entringer [21]; prime labelings are the subject of Section 7.2 in [12]. Coprime labelings are an emerging variation of graph vertex labeling problems; results of this type can be found, for example, in [1], [3], and [16].

Prime and coprime labeling can be interpreted in the context of the *coprime graph on the integers*, $G(\mathbb{Z})$, which is the graph that has elements of $\mathbb{Z}$ as its vertex set with an edge between m and n if and only if $(m, n) = 1$. Given a set of integers $\mathcal{A}$, one can also study $G(\mathcal{A})$, the subgraph induced by restricting the vertex set of $G(\mathbb{Z})$ to $\mathcal{A}$. Let $\mathcal{A}_n^k$ be the set of k consecutive integers starting at n. A finite graph $G(V, E)$ that admits a prime labeling is a (possibly non-induced) subgraph of $G(\mathcal{A}_1^{|V|})$. In a coprime labeling, one asks for a value of k so that the graph in question is a subgraph of $G(\mathcal{A}_1^k)$.

It is therefore an interesting question to study the coprime graph as a function of its label set. Erdős, in one of the earliest such references [8], investigated the size of the largest complete subgraph of $G(\mathcal{A}_1^k)$. Results have also been proved for cycles [9], tripartite subgraphs [20], and Hamiltonicity [2]. Erdős and Selfridge [10] proved results about the size of the largest complete subgraph in the more general setting where the labels are a set of consecutive integers starting at some integer value greater than 1. We remark that a labeling using the set of consecutive integers starting at n is called a *n-prime labeling* in [2] and in Section 3 of [22].

Our focus is on problems related to those listed in the previous paragraph, studying subgraphs of the coprime graph restricted to sets of consecutive integers, $\mathcal{A}_n^k$. Among the major results in this paper, we show in Theorem 1 that isolated vertices in the coprime graph can be characterized as satisfying one of two conditions. We show in Theorem 2 that a complete bipartite subgraph $K_{m,k-m}$ of the coprime graph is most balanced when the vertices in one bipartition are the so-called central vertices of the graph. In fact, it is possible for there to be no such bipartite subgraph; this happens when the label set is a *stapled sequence* (Corollary 4). We end with a result about the most balanced bipartite subgraph, showing in Theorem 5 that the average size of the smaller bipartition vertex set over all possible intervals of length k grows as $\frac{k}{\ln k} e^{-\gamma}$, where γ is the Euler-Mascheroni constant.

Many of our results follow from analysis of the common factor graph, which is the complement graph of the coprime graph. Integers that are isolated vertices in the common factor graph play an important role in a number of subgraphs of the coprime graph. For example, these integers are also vertices that are part of every maximum complete subgraph of the coprime graph. It is less obvious that the same integers have a part to play in bipartite subgraphs too. Our analysis of which integers are isolated vertices is aided by the *modular table*, a matrix of values that tracks the remainders of vertex labels modulo a collection of small primes. The Chinese Remainder Theorem implies that for a fixed number of consecutive vertex labels there are only finitely many distinct modular tables.

This paper is organized as follows. In Section 2, we formally define the coprime graph, the common factor graph, and the modular table, providing some basic properties of these objects. In Section 3, our focus shifts to maximum complete subgraphs and complete bipartite subgraphs of the coprime graph on sets of consecutive integers. The collection of isolated vertices in the common factor graph is an important component contributing to the maximum size of both subgraphs, and this section includes results on how the size of this collection changes as a function of the length and starting value of the label set. Results on stapled sequences are also recalled in this section. Section 4 is a methods section that outlines an algorithm based on the modular table which quickly finds stapled sequences of a particular size; results are in Table 2. A result on the asymptotic average size of the smaller bipartition of the most balanced complete bipartite subgraph of the coprime graph is presented in Section 5.

2. Basic Constructions

Let $\mathcal{A}$ be a set of k distinct positive integers. Usually $\mathcal{A}$ consists of consecutive positive integers, and we will let $\mathcal{A}_n^k$ denote the set

$$\{n, n+1, \ldots, n+k-1\}$$

of size k. When $n = 1$, we may drop the subscript and write $\mathcal{A}^k$ to denote the set $\{1, 2, \ldots, k\}$. As noted in the introduction, a prime labeling uses $\mathcal{A}^k$ as the label set, where k is equal to the number of vertices, $|V|$, whereas in a coprime labeling $k > |V|$ is possible.

Throughout this paper we will regularly consider two constructions: the coprime graph and its complement, the common factor graph. We recall these definitions.

Definition 1. The *coprime graph* of $\mathcal{A}$ is a graph whose k vertices are elements of $\mathcal{A}$, and where two vertices are connected by an edge if and only if the two vertices are coprime to each other.

Definition 2. The *common factor graph* of $\mathcal{A}$ is a graph whose k vertices are elements of $\mathcal{A}$, and where two vertices are connected by an edge if and only if the vertices have a common factor $d \geq 2$.

We denote the coprime graph of $\mathcal{A}$ by $G(\mathcal{A})$ and the common factor graph by $\overline{G}(\mathcal{A})$. When $\mathcal{A} = \mathcal{A}_n^k$ is a set of consecutive integers, the following result tells us more about the structure of the common factor graph.

Proposition 1. [7, Section 2] *The graph $\overline{G}(\mathcal{A}_n^k)$ is either connected or consists of a single connected subgraph and one or more isolated vertices.*

Sketch of proof. One notes that all even numbers must be in a component, and that if a prime $p \neq 2$ divides more than one element of $\mathcal{A}_n^k$, then at least one of those elements is even, which implies that the component with p-divisible elements is in the same component as the even numbers. (See also the proof of Theorem 2.) $\square$

We recall some basic definitions from graph theory. Given a graph G and vertices u, v in the graph, the *distance* between u and v, $d(u, v)$, is the length of the shortest path between u and v. We say $d(u, v) = \infty$ if no such path exists. We define the *eccentricity* of a vertex u to be $\max_{v \in V} d(u, v)$. Finally, the *center* of a graph is the subgraph induced by the vertices of minimum eccentricity.

We note that a vertex that is isolated in $\overline{G}(\mathcal{A}_n^k)$ is connected to every other vertex in $G(\mathcal{A}_n^k)$. Therefore, these vertices will be part of every maximum complete subgraph of $G(\mathcal{A}_n^k)$ and hence provide a lower bound on their size. Furthermore, whenever there are isolated vertices in $\overline{G}(\mathcal{A}_n^k)$, these vertices are also precisely the ones in the center of $G(\mathcal{A}_n^k)$. We will also see that isolated vertices in the common factor graph have an important role to play in complete bipartite subgraphs of $G(\mathcal{A}_n^k)$. The importance of these vertices motivates the following definition.

Definition 3. We call a vertex in the coprime graph $G(\mathcal{A}_n^k)$ a *central vertex* whenever that vertex is isolated in the common factor graph $\overline{G}(\mathcal{A}_n^k)$.

Using a direct construction of the common factor graph to count the number of connected components is computationally expensive. Instead, we describe a tabular approach based on modular arithmetic that stems from the observation that if p is a prime and $p \geq k$, then at most one integer in $\mathcal{A}_n^k$ can have p as a factor. Hence, we may restrict our attention to primes strictly less than the interval length to determine the edges in the common factor graph, rather than the entire prime factorization of the vertices.

Given an interval $\mathcal{A}_n^k$, we construct an $r \times k$ *modular table* of remainders as follows. The r rows are indexed by $2, 3, 5, \ldots p_r$, the set of primes less than k. The k *column numbers* are indexed by elements $m \in \mathcal{A}_n^k$. The entry in the prime p_i row and column with index m is $m \bmod p_i$.

Example 1. Table 1 shows the modular table (on the left) for the set $\mathcal{A}_{181}^9 = \{181, 182, \ldots, 189\}$. In the table on the right, we only mark entries where the column number is evenly divisible by the row prime.

	181	182	183	184	185	186	187	188	189
2	1	0	1	0	1	0	1	0	1
3	1	2	0	1	2	0	1	2	0
5	1	2	3	4	0	1	2	3	4
7	6	0	1	2	3	4	5	6	0

	181	182	183	184	185	186	187	188	189
2		0		0		0		0	
3			0			0			0
5					0				
7		0							0

Table 1: Modular tables for $\mathcal{A}_{181}^9$.

Theorem 1. *There are two conditions where a vertex associated to a column number is isolated in the common factor graph:*

1. *When the column contains no* 0*.*
2. *When each* 0 *in the column is the only* 0 *in its row.*

Proof. From the sketch of Proposition 1, any prime divisor of a vertex in $\overline{G}(\mathcal{A}_n^k)$ that is isolated does not evenly divide any other element of $\mathcal{A}_n^k$. The first condition corresponds to the situation where a vertex's prime divisors are all k or greater. The second condition corresponds to the situation where some or all of the prime divisors of a vertex are less than k. □

We note that our modular table is similar to a table described in [14, Section 2], and that the second condition in Theorem 1 is closely related to [14, Definition 2.5] (although used in a different context). Referring to Table 1, we see that the columns corresponding to the prime number 181 and $187 = 11 \times 17$ have no zeros, and the column corresponding to $185 = 5 \times 37$ has one 0, which is the only 0 in its row. Thus, the common factor graph of $\mathcal{A}_{181}^9$ consists of one large connected component consisting of $\{182, 183, 184, 186, 188, 189\}$ and three isolated vertices: $181, 185$ and 187. The common factor graph of $\mathcal{A}_{181}^9$ is shown in Figure 1.

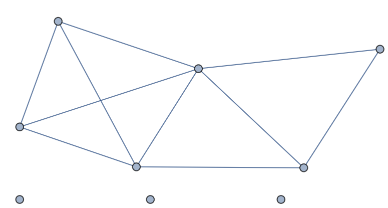

Figure 1: The common factor graph of $\mathcal{A}_{181}^9$.

We are interested in the properties of modular tables for various values of k. The following number theoretic function will appear regularly in our analysis of the enumeration of such tables.

Definition 4. Let m be a positive integer, p_i the i^{th} prime number, and $\pi(m)$ the number of primes less than or equal to m. Then the *primorial* of m, denoted by $m\#$, is defined by $m\# = \prod_{i=1}^{\pi(m)} p_i$.

The primorial function grows very quickly. It is known, for example [18], that

$$p_r\# = e^{(1+o(1))r \log r}, \tag{1}$$

where p_r is the rth prime.

We recall that the entries in Table 1 are the remainders upon division by the set of prime numbers less than 9, the size of $\mathcal{A}_{181}^9$. As a modular table is completely determined by the entries in its first column, by the Chinese Remainder Theorem there are $2 \cdot 3 \cdot 5 \cdot 7 = 210 = (9-1)\# = 8\#$ distinct 4×9 modular tables. Therefore, as n varies, the modular tables associated to $\mathcal{A}_n^9$ repeat with period $8\#$. More generally, for a fixed value of k, data pulled from the coprime and common factor graphs associated to $\mathcal{A}_n^k$ are periodic with period the product of the primes in the modular table, specifically $(k-1)\#$.

3. Subgraphs of the Coprime Graph

Although we will focus on bipartite subgraphs in this section, we start with some results on complete subgraphs of coprime graphs. Using an equivalent formulation, Erdős and Selfridge investigated the minimum and maximum sizes of the largest complete subgraphs of $G(\mathcal{A}_n^k)$ over all values of n [10]. For example, they proved the following result, where $F(n,k)$ is the size of a maximum complete subgraph of $G(\mathcal{A}_n^k)$:

Proposition 2. [10, Theorem 1] *The following inequality holds:*

$$\max_n F(n,k) > \pi(k) + \left(\log 2 - \frac{1}{2} - o(1)\right) \frac{k}{(\log k)^2}.$$

Here, $\log x$ denotes the natural log. Erdős and Selfridge also noted that

$$\max_n F(n,k) < (2 + o(1)) \frac{k}{\log k},$$

which follows from an application of the Selberg Sieve. We prove a couple of elementary results about $F(n,k)$ as a function of two variables.

Proposition 3. *For all n and k,*

$$|F(n,k) - F(n+1,k)| \leq 1.$$

In addition,

$$0 \leq F(n,k+1) - F(n,k) \leq 1.$$

Proof. First assume that $F(n,k) \geq F(n+1,k)$, and take a complete subgraph of maximum size in $G(\mathcal{A}_n^k)$. If the vertex n is in the subgraph, removal of that vertex implies that $F(n+1,k) \geq F(n,k) - 1$. Otherwise, by the inequality assumption, $F(n,k) = F(n+1,k)$. There is an analogous argument when $F(n,k) \leq F(n+1,k)$.

To prove the second assertion, we note that $F(n,k+1) \geq F(n,k)$ as $\mathcal{A}_n^k \subset \mathcal{A}_n^{k+1}$. If there were a complete subgraph of size $F(n,k) + a$ in $G(\mathcal{A}_n^{k+1})$ with $a > 1$, then the removal of the vertex $n+k$ from the complete subgraph would result in a complete subgraph of size greater than $F(n,k)$ in $G(\mathcal{A}_n^k)$, which is impossible. $\square$

We remark that, given a subset $\mathcal{A} \subseteq \mathcal{A}_n^k$, one can check if the subset's induced subgraph, $G(\mathcal{A}) \subseteq G(\mathcal{A}_n^k)$, is a complete subgraph using the modular table for $\mathcal{A}_n^k$. The condition is straightforward: consider the columns associated to $\mathcal{A}$–there can be at most one 0 per row in these columns. However, in practice, this condition does not seem to lead to a straightforward way to find a maximum complete subgraph in a coprime graph. We note that the general problem of finding a maximum complete subgraph of an arbitrary graph is an NP-hard problem [13].

It might appear from the proof of Proposition 3 that results about maximum complete subgraphs could be proven using induction. Unfortunately, relationships between such subgraphs are subtle, even in sets of consecutive integers with a lot of overlap. For example, $F(11,10) = F(11,11) = 6$, but the maximum complete subgraphs in $G(\mathcal{A}_{11}^{10})$ that have the greatest overlap with the maximum complete subgraph with vertex set $\{11,13,17,19,20,21\}$ in $G(\mathcal{A}_{11}^{11})$ have the vertex sets $\{11,13,15,16,17,19\}$ and $\{11,13,14,15,17,19\}$, each sharing only four elements.

For the rest of this section we will consider complete bipartite subgraphs of the coprime graph. This is a statement about $\mathcal{A}_n^k$, and whether it can be partitioned into two subsets so that every element in the first set is coprime with every element in the second set. We note that we allow integers within the same subset to be coprime with each other, so that when we pass to the graphical setting, the subgraphs corresponding to the subsets need not be induced subgraphs of $G(\mathcal{A}_n^k)$. As a follow-up question, we wish to determine the complete bipartite subgraph, $K_{m,k-m}$, should it exist, whose bipartitions are closest in size. We refer to this bipartite graph as the *most balanced*. We will approach this problem through an analysis of the common factor graph $\overline{G}(\mathcal{A}_n^k)$. Recall from Proposition 1 that $\overline{G}(\mathcal{A}_n^k)$ consists of a large connected component and some isolated vertices.

Theorem 2. *Given a set of consecutive positive integers $\mathcal{A}_n^k$, the most balanced complete bipartite subgraph, $K_{m,k-m}$, of $G(\mathcal{A}_n^k)$ has the integers in the connected component of the common factor graph $\overline{G}(\mathcal{A}_n^k)$ in one bipartition, and the isolated vertices in $\overline{G}(\mathcal{A}_n^k)$, i.e., the central vertices of $G(\mathcal{A}_n^k)$, in the other bipartition.*

Proof. The argument parallels the argument for Proposition 1. Even integers in $\mathcal{A}_n^k$ must be in the same partition, V_1, because they share a factor of 2. Similarly, if p is a prime factor of an even number in $\mathcal{A}_n^k$, vertices that are multiples of p must be in V_1 as well. We continue this process prime-by-prime, which terminates because the initial set, $\mathcal{A}_n^k$, is finite. The vertices which remain in the other bipartition, V_2, have prime factors that do not occur as factors of any other vertex in $\mathcal{A}_n^k$. This follows since when a prime p divides more than one element of $\mathcal{A}_n^k$, at least one of those elements is even. In particular, each of the vertices in V_2 is coprime with the other elements of $\mathcal{A}_n^k$. This implies that their eccentricity is 1, and hence they are central. Together, the central vertices form the largest set of vertices that can be in V_2. □

Notice that the size of the set of isolated vertices in $\overline{G}(\mathcal{A}_n^k)$ provides a lower bound for $F(n,k)$. The example for $F(11,10)$ above shows that this lower bound is not necessarily tight. Also, since the vertices in V_2 are pairwise coprime, vertices in V_2 can be moved to V_1 while maintaining coprimality between the vertices in the bipartitions.

Theorem 2 tells us that the problem of finding complete bipartite subgraphs of the coprime graph whose bipartitions include all vertices is the same problem as finding isolated vertices in the common factor graph. We will use the coprime graph formulation in what follows, because of its connection with the modular table.

Corollary 1. *Suppose $\overline{G}(\mathcal{A}_n^k)$ has $m > 0$ connected components. Then $G(\mathcal{A}_n^k)$ contains $m-1$ different-sized complete bipartite subgraphs with k vertices.*

Example 2. The common factor graph of $\mathcal{A}_{181}^9$ shown in Figure 1.

The graph $\overline{G}(\mathcal{A}_{181}^9)$ has three isolated vertices: 181, 185, and 187. Therefore, $K_{8,1}$, $K_{7,2}$, and $K_{6,3}$ are the only complete bipartite subgraphs of $G(\mathcal{A}_{181}^9)$ with 9 vertices, depending on whether the smaller bipartition contains one, two, or all three central vertices. The most balanced complete bipartite subgraph is $K_{6,3}$.

We note a couple of basic results on the number of isolated vertices, the first about prime labeling. Recall that $\mathcal{A}^k = \{1, 2, \ldots, k\}$.

Lemma 1. *The number of isolated vertices of $\overline{G}(\mathcal{A}^k)$ is given by $\pi(k) - \pi\left(\lfloor \frac{k}{2} \rfloor\right) + 1$*

Proof. From the proof of Proposition 1, we know that isolated vertices are integers whose prime factors do not occur as factors of any other integer in $\mathcal{A}^k$. Since $\mathcal{A}^k$ contains 1, and any prime less than or equal to $\lfloor \frac{k}{2} \rfloor$ will divide more than one element of $\mathcal{A}^k$, integer labels of isolated vertices are either 1 or a prime greater than $\lfloor \frac{k}{2} \rfloor$. □

Example 3. For the first 30 values of k, the number of isolated vertices in $\overline{G}(\mathcal{A}^k)$ as given by Lemma 1 are

$$1, 2, 3, 2, 3, 2, 3, 3, 3, 2, 3, 3, 4, 3, 3, 3, 4, 4, 5, 5, 5, 4, 5, 5, 5, 4, 4, 4, 5, 5.$$

The last value in the list means that using the label set $\mathcal{A}^{30}$, there are five distinct complete bipartite graphs with 30 vertices that admit prime labelings, namely $K_{29,1}$, $K_{28,2}$, $K_{27,3}$, $K_{26,4}$, and $K_{25,5}$ (with vertex set $\{1, 17, 19, 23, 29\}$). Lemma 1 and the Prime Number Theorem imply that the number of isolated vertices in $\overline{G}(\mathcal{A}^k)$ grows as $\frac{k}{2 \ln k}$. We remark that the sequence in Example 3 is OEIS A076225, which is one greater than OEIS A056171. There is a related sequence, OEIS A080359, where the mth term in the sequence is the first time m appears in OEIS A056171. It is also the smallest positive integer n where $n!$ has m distinct prime factors.

The next result describes how the number of isolated vertices in $\overline{G}(\mathcal{A}_n^k)$ changes as the length and starting point of the sequence of integers changes. From Proposition 3, one might expect these changes to be small. In contrast:

Proposition 4. *Given any positive integer M, there exist values of n and k so that*

1. *the difference between the number of isolated vertices in $\overline{G}(\mathcal{A}_n^k)$ and $\overline{G}(\mathcal{A}_{n+1}^k)$ is at least M.*

2. *the number of isolated vertices in $\overline{G}(\mathcal{A}_n^{k+1})$ is at least M less than in $\overline{G}(\mathcal{A}_n^k)$.*

On the other hand, the number of isolated vertices in $\overline{G}(\mathcal{A}_n^{k+1})$ can be no more than one greater than in $\overline{G}(\mathcal{A}_n^k)$.

Proof. The examples for both numbered parts of the proposition come from sets involving integers that satisfy the second condition in Theorem 1. Let $N = p\#$ for some prime p. We note that in the modular table for intervals of length $p+1$, column N consists of all 0's.

For the first claim, consider the intervals $\mathcal{A}_{N-p}^p$ and $\mathcal{A}_{N-p+1}^p$. We claim that the values $\{N-p_r\}$, where p_r is any prime strictly greater than $\frac{p}{2}$ and less than p, will be isolated vertices in the graph $\overline{G}(\mathcal{A}_{N-p}^p)$. This follows as $N \notin \mathcal{A}_{N-p}^p$, using the observation that the only prime less than p which divides $N-p_r$ is p_r. This implies that in the modular table for $\mathcal{A}_{N-p}^p$, the column $N-p_r$ contains only one 0, and that 0 is the only one in its row. However, since $N \in \mathcal{A}_{N-p+1}^p$, none of the column numbers $\{N-p_r\}$ correspond to isolated vertices in $\overline{G}(\mathcal{A}_{N-p+1}^p)$. From the discussion after Lemma 1, the number of primes between $\frac{p}{2}$ and p grows as $\frac{k}{2\ln(k)}$ and hence will eventually take on a value greater than or equal to any fixed M. Therefore, by picking a suitable prime p, the number of isolated vertices in $\mathcal{A}_{N-p+1}^p$ can be made to be at least M less than $\mathcal{A}_{N-p}^p$, for any M. An analogous argument with the sets $\mathcal{A}_N^p$ and $\mathcal{A}_{N+1}^p$ shows that the number of isolated vertices can increase as well.

We use similar examples for the second claim of the proposition. Consider the sets $\mathcal{A}_{N-p}^p$ and $\mathcal{A}_{N-p}^{p+1}$. Then the same reasoning as above shows that by choosing a suitably large value of p, $\overline{G}(\mathcal{A}_{N-p}^{p+1})$ has at least M fewer isolated vertices than $\overline{G}(\mathcal{A}_{N-p}^p)$, for any fixed value of M.

For the final claim, when $\overline{G}(\mathcal{A}_n^{k+1})$ has more isolated vertices than $\overline{G}(\mathcal{A}_n^k)$, consider those graphs' respective modular tables. By Theorem 1, the addition of the $(k+1)$st column will not convert a non-isolated vertex in $\overline{G}(\mathcal{A}_n^k)$ into an isolated one in $\overline{G}(\mathcal{A}_n^{k+1})$. Therefore, any increase happens as the $(k+1)$st column satisfies one of the conditions of Theorem 1. □

Although Proposition 4 suggests that it may be difficult to say much about the overall behavior of the number of isolated vertices in $\overline{G}(\mathcal{A}_n^k)$, there is a characterization for this number for small values of n relative to k. This next result can be thought of as a significant generalization of Lemma 1.

Theorem 3. *For $n > 3$ and $1 < n \leq k+1$, the number of isolated vertices in $\overline{G}(\mathcal{A}_n^k)$ is given by $\pi(n+k-1) - \pi\left(\frac{n+k-1}{2}\right)$.*

Proof. Since $\mathcal{A}_n^k = \{n, n+1, \ldots, n+k-1\}$, we are only considering column numbers less than or equal to $2k$. We can check directly that the claim holds for $3 < n \leq 9$, so assume that $n > 9$. We claim that any column number that both satisfies the hypotheses of this theorem and is an isolated vertex must be prime. To see this, take a composite column number $m \leq 2k$ and let p_1 be its smallest prime divisor. We note that if $p_1 \leq \frac{k}{2}$, then $\mathcal{A}_n^k$ will contain at least two integers evenly divisible by p_1, and so m is not isolated. However, if $p_1 > \frac{k}{2}$, then $m \geq p_1^2 > \left(\frac{k}{2}\right)^2 > 2k$ as we are assuming that $n > 9$, and hence $k \geq 9$. This implies that $m \notin \mathcal{A}_n^k$.

Therefore, we may restrict our attention to prime column numbers. We note that not all prime numbers in $\mathcal{A}_n^k$ correspond to isolated vertices, for example, those that are less than or equal to $\frac{k}{2}$. We can be more specific. Say that the prime column number p lies in the interval $[k, n+k]$. Then in order for the 0 in p's column to only appear once in the modular table, we must have $2p > n+k-1$, which implies that $p > \frac{n+k-1}{2}$. We conclude that the only possible column numbers which correspond to isolated vertices are prime numbers that are larger than $\frac{n+k-1}{2}$ and less than or equal to $n+k-1$. □

The reader may wish to compare Theorem 3 to results about prime and coprime labelings of complete bipartite graphs from a different point of view in [11, Proposition 2.3] and [16, Corollary 4.1].

The statements of Lemma 1, Proposition 4, and Theorem 3 show how the number of central vertices in the coprime graph changes with $\mathcal{A}_n^k$. The next two corollaries follow from Theorem 3.

Corollary 2. *The number of isolated vertices of $\overline{G}(\mathcal{A}_n^k)$—equivalently the number of central vertices of $G(\mathcal{A}_n^k)$—roughly increases as n increases from 1 to k.*

Proof. We analyze $\pi(n+k) - \pi\left(\frac{n+k}{2}\right)$, as n goes from 1 to k, using the Prime Number Theorem: $\pi(x) \sim \frac{x}{\log x}$. We use the fact that

$$f(x) = \pi(x) - \pi\left(\frac{x}{2}\right) \sim \frac{x}{\log(x)} - \frac{\frac{x}{2}}{\log\left(\frac{x}{2}\right)} = \frac{x}{2}\,\frac{\log\left(\frac{x}{4}\right)}{\log(x)\log\left(\frac{x}{2}\right)}.$$

One checks that

$$f'(x) = \frac{\left(\log\left(\frac{x}{8}\right) - 1\right)\log^2(x) + (2+\log(2))\log(4)\log(x) - \log(2)\log(4)}{2\log^2\left(\frac{x}{2}\right)\log^2(x)}.$$

The numerator is a cubic polynomial in $\log(x)$ with one real root: $x \sim 1.41078$. For the n and k values in Proposition 4, $f'(x) > 0$. Therefore, $\pi(n+k) - \pi\left(\frac{n+k}{2}\right)$ tends to increase as n goes from 1 to k. $\square$

Corollary 3. *For a fixed k, the maximum number of isolated vertices of $\overline{G}(\mathcal{A}_n^k)$—equivalently the number of central vertices of $G(\mathcal{A}_n^k)$—is at least* $\pi(2k) - \pi(k) \sim k\,\frac{\log\left(\frac{k}{2}\right)}{\log(2k)\log(k)}$.

We get a more specific result using the following inequalities from [6]:

$$\pi(x) \geq \frac{x}{\log x - 1} \text{ for } x \geq 5393 \text{ and } \pi(x) \leq \frac{x}{\log x - 1.1} \text{ for } x \geq 60184.$$

Therefore,

$$\pi(2x) - \pi(x) \geq \frac{2x}{\log 2x - 1} - \frac{x}{\log x - 1.1} \tag{2}$$

for $x \geq 60184$. One can show by basic calculus techniques that the ratio of the right-hand side of Inequality (2) and $\frac{x}{\log x}$ approaches 1. A *Mathematica* plot confirms that

$$\pi(2x) - \pi(x) \geq 0.9\left(\frac{x}{\log x}\right)$$

holds for $x \geq 1329$.

We remark that the formula in Theorem 3 does not hold for $n = 3$ (consider $\mathcal{A}_3^3$) and in general for $n > k+1$ (consider $\mathcal{A}_7^5$ and $\mathcal{A}_8^6$, for examples with undercounts and overcounts, respectively). Also, the number of isolated vertices in Corollary 3 gives a lower bound for $F(n,k)$ as defined at the start of this section, and matches the dominant term in Proposition 2. The graph in Figure 2 shows how the number of isolated vertices with $k = 600$ grows as n increases.

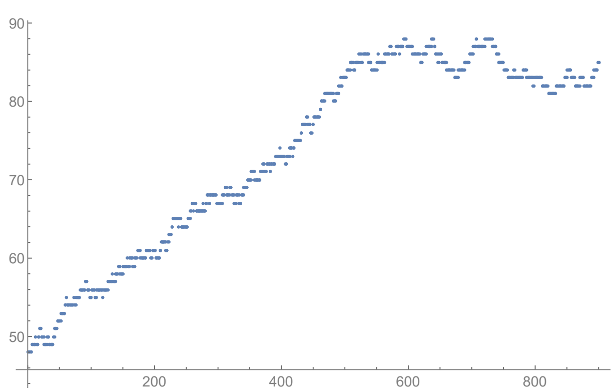

Figure 2: The number of isolated vertices in $\overline{G}(\mathcal{A}_n^{600})$ with $1 \leq n \leq 900$.

For the final question in this section, we investigate whether it is always possible to split the entries of $\mathcal{A}_n^k$ into two bipartitions of a complete bipartite subgraph. In particular, are there intervals $\mathcal{A}_n^k$ where there are no central vertices in $G(\mathcal{A}_n^k)$? In light of Theorem 2, an equivalent question is whether, for all intervals $\mathcal{A}_n^k$, the modular table either has a column with no 0's, or a column where each 0 is the only one in its row. If neither condition holds, then the large connected component described in Proposition 1 contains every integer in the sequence and there are no isolated vertices in $\overline{G}(\mathcal{A}_n^k)$. Although this is restrictive, intervals for which both situations fail can be realized for all $k \geq 17$.

Definition 5. Let $\mathcal{A}_n^k$ be an interval with the property that in its modular table, every column contains at least one zero and at least one zero in each column is in a row with another zero. Then $\mathcal{A}_n^k$ is a *stapled sequence*.

Stapled sequences have a long history. They were first studied by Pillai [19] in the 1930's and around the same time by Brauer [4]. Eggleton noted [7, Theorem 3] that there are intervals $\mathcal{A}_n^k$ for all $k \geq 17$ where $\overline{G}(\mathcal{A}_n^k)$ contained no isolated vertices. Gassko, who coined the term "stapled sequence," conducted further investigations into their properties [14, 15]. For additional background and history of these sequences, see [7, 14].

Corollary 4. *For all $k \geq 17$ there exist sequences $\mathcal{A}_n^k$ where there are no central vertices in $G(\mathcal{A}_n^k)$, and hence for which there are no complete bipartite subgraphs $K_{m,k-m}$ in $G(\mathcal{A}_n^k)$.*

The first stapled sequence is

$$\mathcal{A}_{2184}^{17} = \{2184, 2185, \ldots, 2200\}.$$

Since stapled sequences can be identified by a property of a modular table, the existence of one stapled sequence implies the existence of infinitely many, repeating with period $(k-1)\#$. A list of the smallest n so that $\mathcal{A}_n^k$ is a stapled sequence is OEIS A090318.

4. Stapled Sequences: Numerical Methods

The modular table framework provides an approach to readily determine the number of isolated vertices in $\overline{G}(\mathcal{A}_n^k)$. We briefly describe our algorithm, implemented in *Mathematica*, to find the number of isolated vertices in $\overline{G}(\mathcal{A}_n^k)$ and in particular stapled sequences (see Figure 2). Given the interval length k, we know that modular tables repeat with period $(k-1)\#$, which makes the search-space large, but finite. Moreover, stapled sequences cannot contain any prime number larger than k, so we can restrict our searches to prime gaps–consecutive sequences of composite integers.

We construct a modular table for every possible interval and then use the conditions on zeros described in Definition 5 to determine if the interval is a stapled sequence. We store only the interval's starting value, which is sufficient to reconstruct the entire interval given its length. Here are the basic steps of the algorithm.

Stapled Sequence Search Algorithm

1. Set the interval length and determine the modular table period.
2. Determine the list of row primes needed to build the modular table.
3. Sequentially find the prime gaps larger than the interval length.
4. Construct a large modular table for the full length of the prime gap using only the primes less than the interval length.
5. Select sub-tables of adjacent columns to create modular tables of the correct size.
6. Check the placement of zeros in each sub-table to see if they satisfy the requirements of Definition 5.
7. Store the starting values of each stapled sequence found.

The reader interested in the actual code should contact the authors. We remark that our current implementation searches through the $25\# = 23\# = 223{,}092{,}870$ total intervals of length 26 to find that 1750 are stapled sequences (see Table 2). The search takes about 12 minutes on a reasonably modern PC.

Using our code, we have found all stapled sequences of lengths 17 through 31 within a period. These data can be found in Table 2. The primorial growth rate given in Equation (1) implies the modular table period grows super-exponentially, which means that the search for all stapled sequences quickly becomes prohibitively difficult; we expect our algorithm to require about a week of computation to find all stapled sequences of length 32 within one period.

An interesting question to ask is how stapled sequences of different lengths may be related to one another. If a stapled sequence has even length $2m$, then it either starts or ends with an odd integer. In this case, the even integer adjacent to the odd endpoint can be included to make a stapled sequence of odd length $2m+1$. However, not every stapled sequence of odd length comes about in this manner, since this would imply that there are stapled sequences of length 16. Stapled sequences of odd length $2m+1$ have endpoints of equal parity. If the interval has odd endpoints, then two additional even integers can be included to make a new stapled sequence of length $2m+3$. Gassko also noted that stapled sequences of prime length p naturally give rise to stapled sequences of length $p+1$ [14, Lemma 3.7].

Interval length (n)	Stapled sequence count	Period ($(n-1)\#$)
17	2	30,030
18	4	510,510
19	10	510,510
20	128	9,699,690
21	250	9,699,690
22	192	9,699,690
23	226	9,699,690
24	1916	223,092,870
25	2568	223,092,870
26	1750	223,092,870
27	1666	223,092,870
28	568	223,092,870
29	534	223,092,870
30	3772	6,469,693,230
31	4472	6,469,693,230

Table 2: The number of stapled sequences within a period.

Given these observations, one might expect the number of stapled sequences to be monotonically increasing with the interval length k. A quick look at Table 2 shows this is not true. Gassko's observation about stapled sequences of lengths p and $p+1$ does appear, although we note that the increase also coincides with a large increase in the length of the period. It is not even true that there must be more stapled sequences of length $2m+1$ than of length $2m$. For example, the number of stapled sequences decreases from interval lengths of 25 to 29, so in particular from 26 to 27 as well from 28 to 29. To see what can happen, consider the two intervals $\{771320, \ldots, 771345\}$ and $\{771321, \ldots, 771346\}$ of length 26. Both of these sets are stapled sequences that are sub-intervals of same stapled sequence of length 27, namely $\{771320, \ldots, 771346\}$. This is a situation where one stapled sequence contains two sub-intervals that are both stapled sequences. Eggleton [7] refers to this situation as a *constellation.* For the reader interested in learning more, we recommend [7] and [15].

5. Average Size of the Smallest Bipartition

In Section 3, we showed that the number of isolated vertices of $\overline{G}(\mathcal{A}_n^k)$ with k fixed can take on a wide range of values (Corollaries 3 and 4, and Figure 2), and in general, can change significantly with n and k (Proposition 4). In other words, the shape of the most balanced maximal bipartite subgraph of $G(\mathcal{A}_n^k)$ varies widely with

n and k. In this section, we will show that, despite this variation, we can still say something about the average number of central vertices in coprime graphs. More specifically, we will fix a value of k and determine the average number of isolated vertices in $\overline{G}(\mathcal{A}_n^k)$ over an entire period, that is, as n ranges from 1 to $(k-1)\#$. We begin with some notation.

Definition 6. Denote the function $\mathtt{fr}(k)$ on integer values $k \geq 2$ by

$$\mathtt{fr}(k) = \prod_{p \leq k} \left(1 - \frac{1}{p}\right),$$

where the product is over primes less than or equal to k

The function $\mathtt{fr}(k)$, which is a decreasing function of k, gives the *fraction* of integers coprime to a set of small primes in certain intervals starting at 1. This follows from a sieve result, and is the subject of the next lemma.

Lemma 2. *Let p_r be a prime and $N = p_r\#$. Then for any $m \leq p_r$, there are $N\mathtt{fr}(m)$ integers in $[1, N]$ that are coprime to $\prod_{p_i \leq m} p_i$.*

Proof. Given $m \leq p_r$, let $p_1, p_2, \ldots, p_i$ be the primes less than or equal to m, and note that N is evenly divisible by $p_i\#$. The size of the set of integers we are counting can be readily determined using an inclusion-exclusion argument:

$$\begin{aligned} N - \sum_j \frac{N}{p_j} + \sum_{j,k} \frac{N}{p_j p_k} + \ldots + (-1)^i \frac{N}{p_1 p_2 \cdots p_i} \\ = N \left(1 - \sum_j \frac{1}{p_j} + \sum_{j,k} \frac{1}{p_j p_k} + \ldots + (-1)^i \frac{1}{p_1 p_2 \cdots p_i}\right) \\ = N\mathtt{fr}(m). \end{aligned}$$

□

When $m = p_r$, Lemma 2 is the well-known result that the number of positive integers less than or equal to N and coprime to N is $\varphi(N)$, where φ is the standard Euler totient function.

Proposition 5. *For a fixed interval length k, let S be the sum of all isolated vertices in $\overline{G}(\mathcal{A}_n^k)$, as n ranges over all values in $[1, N]$. Then*

$$kN\mathtt{fr}(k) \leq S \leq kN\mathtt{fr}(k/2).$$

Proof. Theorem 1 provides two conditions where a vertex associated to a column number in the modular table corresponds to an isolated vertex. The first condition, when the column contains no 0, implies that the column number is coprime to

$(k-1)\#$. By the note after Lemma 2, there are $\varphi((k-1)\#)$ such integers, each of which is a column number for a column with no 0 in k modular tables of length k. Therefore, such integers account for

$$S_1 = k\varphi\left((k-1)\#\right) = kN\texttt{fr}(k-1) \geq kN\texttt{fr}(k) \tag{3}$$

isolated vertices in $G(\mathcal{A}_n^k)$ over all intervals of length k over a full period, making S_1 a lower bound for S, the total sum of isolated vertices.

To construct the upper bound, we add to S_1 an estimate of the number of isolated vertices that arise from the second condition of Lemma 2, column numbers where each 0 in the column is the only one in its row. We denote this component of the sum by S_2. Let $p_1, p_2, \ldots, p_r$ be an ordered list of primes less than k. By the condition on 0's, it is sufficient in the modular table to only consider rows with primes greater than $\frac{k}{2}$. By Bertrand's postulate [17, Theorem 8.7], we know that there is a smallest prime strictly between $\frac{k}{2}$ and k, say p_j.

We stratify the count for S_2, adding one prime divisor at a time. By Lemma 2, the number of integers in $[1, N]$ that are divisible by p_j but none of $p_1, p_2, \ldots p_{j-1}$ is

$$N\texttt{fr}(p_{j-1}) - N\texttt{fr}(p_j).$$

In addition, when $p_j \geq \frac{k}{2}$ there are $2p_j - k$ modular tables of length k which have precisely one column number divisible by p_j. That is, there are as many as

$$N\left(\texttt{fr}(p_{j-1}) - \texttt{fr}(p_j)\right)(2p_j - k) \tag{4}$$

isolated vertices to be added to the S_2 count that come from column numbers divisible by p_j but not primes $p_1, \ldots, p_{j-1}$. This is a likely overcount, since not all $2p_j - k$ placements of p_j will necessarily result in isolated vertices; there could be another 0 in a row with prime label larger than p_j that is not the only 0 in its row.

We repeat this process for all primes from p_j to p_r, noting that the resulting sets of column numbers are disjoint and include all integers divisible by at least one of $p_j, p_{j+1}, \ldots, p_r$ but not any of $p_1, p_2, \ldots, p_{j-1}$. Therefore,

$$S_2 \leq N\left(\texttt{fr}(p_{j-1}) - \texttt{fr}(p_j)\right)(2p_j - k) + \cdots + N\left(\texttt{fr}(p_{r-1}) - \texttt{fr}(p_r)\right)(2p_r - k).$$

We can replace $2p_j - k$ with $2p_r - k$ in each term, increasing the right-hand side and turning it into a telescoping sum. Using the fact that $\texttt{fr}(n)$ is a decreasing function and $2p_r - k \leq k$, we have the following upper bound for S_2:

$$\begin{aligned} S_2 &\leq N\left(\texttt{fr}(p_{j-1}) - \texttt{fr}(p_j)\right)(2p_r - k) + \cdots + N\left(\texttt{fr}(p_{r-1}) - \texttt{fr}(p_r)\right)(2p_r - k) \\ &\leq N\left(\texttt{fr}(p_{j-1}) - \texttt{fr}(p_r)\right)(2p_r - k) \\ &\leq N\left(\texttt{fr}(k/2) - \texttt{fr}(k-1)\right)(2p_r - k) \\ &\leq kN\left(\texttt{fr}(k/2) - \texttt{fr}(k-1)\right). \end{aligned} \tag{5}$$

Therefore, the total sum of isolated vertices in coprime graphs, S, over an entire period that comes from both cases of Lemma 2 is bounded above by $S_1 + S_2$. By Equation (3) and Inequality (5),

$$S_1 + S_2 \leq kN\mathtt{fr}(k-1) + kN\Big(\mathtt{fr}(k/2) - \mathtt{fr}(k-1)\Big) = kN\mathtt{fr}(k/2).$$

□

To finish the argument, we use Merten's third theorem.

Theorem 4. [5, Theorem 5.2.1] *Let γ be the Euler-Mascheroni constant. Then*

$$\prod_{p \leq n} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log n}\left(1 + O\left(\frac{1}{\log n}\right)\right).$$

Theorem 5. *Given a fixed interval length k, the average number of isolated vertices in $\overline{G}(\mathcal{A}_n^k)$—equivalently the number of central vertices of $G(\mathcal{A}_n^k)$—over the the interval $[1, N]$ grows as $\frac{k}{\ln k}e^{-\gamma}$.*

Proof. We apply Merten's theorem to the upper and lower bounds for the sum S in Proposition 5. For the lower bound,

$$kN\mathtt{fr}(k) = kN\frac{e^{-\gamma}}{\log k}\left(1 + O\left(\frac{1}{\log k}\right)\right).$$

For the upper bound,

$$kN\mathtt{fr}(k/2) = kN\frac{e^{-\gamma}}{\log(k/2)}\left(1 + O\left(\frac{1}{\log(k/2)}\right)\right).$$

However,

$$\frac{1}{\log(k/2)} = \frac{1}{\log k}\left(1 + O\left(\frac{1}{\log k}\right)\right).$$

This implies that both bounds have the same growth function. We average over the entire period by dividing both bounds by N, and the result follows. □

In the context of Section 3, Theorem 5 also provides a statement about a lower bound on the average size of the largest complete subgraph of $G(\mathcal{A}_n^k)$ as well as the average size of the smaller bipartition in complete bipartite subgraphs of $G(\mathcal{A}_n^k)$.

Acknowledgement. It is our pleasure to acknowledge Wil Hubert ('18, United States Air Force Academy), whose senior capstone project and early investigations initiated many of the ideas and directions taken in this paper.

References

[1] J. Asplund and N. B Fox, Minimum coprime labelings for operations on graphs, *Integers* **52** (2019), #A24, 26 pp.

[2] M. H. Bani Mostafa A and E. Ghorbani, Hamiltonicity of a coprime graph, *Graphs Combin.* **37** (2021), no. 6, 2387-2395.

[3] A. Berliner, N. Dean, J. Hook, A. Marr, A. Mbirika, and C. McBee, Coprime and prime labelings of graphs, *J. Integer Seq.* **19** (2016), no. 5, Article 16.5.8, 14 pp.

[4] A. Brauer, On a property of k consecutive integers, *Bull. Amer. Math. Soc.* **47** (1941), 328-331.

[5] A. Cojocaru and M. Ram Murty, *An Introduction to Sieve Methods and their Applications*, London Mathematical Society Student Texts **66**, Cambridge University Press, Cambridge, 2006.

[6] P. Dusart, Estimates of some functions over primes without R.H., `arXiv:1002.0442` (2010).

[7] R. B. Eggleton, Common factors of integers: a graphic view, *Discrete Math.* **65** (1987), no. 2, 141-147.

[8] P. Erdős, Remarks on number theory. IV. Extremal problems in number theory. I, *Mat. Lapok* **13** (1962), 228-255.

[9] P. Erdős and G. Sárközy, On cycles in the coprime graph of integers, The Wilf Festschrift (Philadelphia, PA, 1996), *Electron. J. Combin.* **4** (1997), no. 2, Research Paper 8, 11 pp.

[10] P. Erdős and J. L. Selfridge, Complete prime subsets of consecutive integers, *Proc. Manitoba Conf. on Numerical Mathematics* (Winnipeg, MB, 1971), Dept. Comput. Sci., Univ. Manitoba, Winnipeg, MB, 1971, 1-14.

[11] H. Fu and K. Huang, On prime labellings. Graph theory and applications (Hakone, 1990), *Discrete Math.* **127** (1994), no. 1-3, 181-186.

[12] J. Gallian, A dynamic survey of graph labeling - version 24, *Electron. J. Combin.* (2021), #DS6, 576pp.

[13] M. Garey and D. Johnson, *Computers and Intractability: a Guide to the Theory of NP-Completeness*, A Series of Books in the Mathematical Sciences, W. H. Freeman and Co., San Francisco, CA, 1979.

[14] I. Gassko, Stapled sequences and stapling coverings of natural numbers, *Electron. J. Combin.* **3** (1996), no. 1, Research Paper 33, 20 pp.

[15] I. Gassko, Common factor graphs of stapled sequences, Proceedings of the Twenty-eighth Southeastern International Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 1997), *Congr. Numer.* **126** (1997), 163-173.

[16] C. Lee, Coprime and prime labelings of graphs, *J. Integer Seq.* **23** (2020), no. 11, Article 20.11.4, 15 pp.

[17] I. Niven, H. Zuckerman, and H. Montgomery, *An Introduction to the Theory of Numbers*, Fifth edition, John Wiley & Sons, Inc., New York, NY, 1991.

[18] OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences, `https://oeis.org/A002110` (2022).

[19] S. S. Pillai, On m consecutive integers. I, *Proc. Indian Acad. Sci., Sect. A.* **11** (1940), 6-12.

[20] G. Sárközy, Complete tripartite subgraphs in the coprime graph of integers, *Discrete Math.* **202** (1999), no. 1-3, 227-238.

[21] A. Tout, A. N. Dabboucy, and K. Howalla, Prime labeling of graphs, *Nat. Acad. Sci. Letters* **11** (1982), 365-368.

[22] S. Vaidya and U. Prajapati, Some results on prime and k-prime labeling, *J. of Mathematics Research* **3** (2011), no. 1, 66-75.