New Results on Ramsey Multiplicity and Graph Commonality

Sarvasva Raghuvanshi

under the direction of
James Hirst
Massachusetts Institute of Technology

Research Science Institute
January 15, 2016

Abstract

If a graph G has v vertices, a *copy* of G inside a larger graph K is a subgraph H of K on v vertices such that G is a subgraph of H. Similarly, we define an *anticopy* of G to be a subgraph H of K on v vertices such that G is a subgraph of $\bar{H}$. A graph F is common if and only if the minimum density of copies and anticopies of F in any graph G is $2^{1-|E(F)|}$, where $|E(F)|$ denotes the number of edges in the graph F. Note that this minimum is attained when G is a random graph with edge density one-half. In this paper, we propose a modern proof that the graph formed from any number of disjoint copies of a common graph is itself common. This novel proof leads to innovative partial results and opens other questions about the commonality of disjoint graphs. We then prove that the graph obtained from a pentagon by adding a chord is common, resolving a central open problem in the field of graph commonality.

Summary

Social network-based companies like Twitter and Facebook rely on their ability to analyze large networks and exploit structures within these networks for a profit. We can study these large networks by analyzing what structures we are guaranteed to find inside any large network with a minimum frequency. These small structures are known as *common graphs.* This paper provides a modern proof of the existing result that the structure formed from any number of disjoint copies of a common graph is also common. The method of proof yields novel partial results and poses new questions about the commonality of disjoint graphs. This paper also resolves the question of the commonality of the pentagon with a chord, an outstanding open question in the field of graph commonality.

The work in this research paper has applications to analysis of any large network, like the Internet or a social network, by paving the road towards identifying substructures that are high-value research targets. This paper may even have applications in ecology and urban planning by identifying common habitat patterns (like a forest with two lakes, which can be modeled in a large graph by a pentagon with a chord). These specific habitats can be given a priority for development or preservation depending on their frequency and environmental impact.

1 Introduction

For social network-based companies like Facebook and LinkedIn, the primary business model is strongly dependent on the corporation's ability to create new connections between previously unacquainted people. Social networks in general can be described by graphs, where people are represented by vertices, with an edge between two vertices if and only if the two corresponding people are acquaintances or friends on the social network in question.

This analysis of graphs as social networks has been around for decades, as a well-known theorem in combinatorics known as the Party Theorem [1], which states that among any six people at a dinner party, either three of them all know each other or three of them are strangers. This problem can be formulated in terms of graph theory — any graph on six vertices must either contain three vertices that are connected to each other (a triangle) or three vertices all disconnected from each other (the complement of a triangle).

This problem can be further generalized; in 1930, Ramsey [2] proved that one can always find any complete graph (a graph with every edge drawn in) inside a sufficiently large graph K or its complement $\bar{K}$. The complement of a graph, denoted by the graph with a bar on top, is the graph obtained by replacing every edge in the graph with a non-edge, and vice-versa. Ramsey's Theorem naturally generalizes to any graph, as can be observed by noting that any graph is a subgraph of the complete graph on its vertices.

For a graph G, $V(G)$ is the set of vertices of G and $E(G)$ is the set of edges of G. For a set A, $|A|$ refers to the cardinality, or size, of the set. Let us define a *copy* of a graph G inside a larger graph K to be a subgraph $H \subset K$ on $|V(G)|$ vertices such that G is a subgraph of H. Similarly, we define an *anticopy* of G to be a subgraph H of K on $|V(G)|$ vertices such that G is a is a subgraph of $\bar{H}$. Note that an anticopy of G in K is a copy of G in $\bar{K}$. We let a *presence* of G in K be either a copy or an anticopy of G in K.

Combinatorialists wish to count the minimum number of presences of a particular graph

G that we are guaranteed to find inside a large K. In a random large graph K, we expect to see $2^{1-e} \cdot \binom{n}{v}$ presences of G, where e denotes the number of edges in G, v denotes the number of vertices in G, and n denotes the number of vertices in K. The intuition for this is that the probability that all e edges of G are present in K is 2^{-e}, which is also the probability that every edge of G is missing. It is known that for some graphs G, the frequency of presences of G in a random K, which is $2^{1-e} \cdot \binom{n}{v}$, ends up being minimal. Such graphs G are called *common* graphs.

Applications for determining commonality of graphs include analysis of large networks, such as social networks or the Internet. Another important application of determining which graphs are common is that of property testing, where a large graph K is analyzed and compared to a random graph on n vertices using metrics such as the number of presences of particular graphs.

Currently, the mathematical community knows the commonality of only a few graphs and classes of graphs. Goodman [1] showed that the triangle K_3 is common. Thomason [3] expanded the result to show that the complete graphs on n vertices are uncommon for $n > 3$. Jagger, et. al. [4] generalized this to show that all graphs containing K_4 as a subgraph are uncommon. Furthermore, removing two edges from a K_4 yields a triangle with a pendant edge, which was proven by Sidorenko [5] to be uncommon (see Figure 1).

This result was further generalized by Fox [6] to show that any odd cycle (triangle, pentagon, etc.) with a pendant edge attached is uncommon (see Figure 1). However, Sidorenko [5] proved that all cycles are common. A few other classes of graphs have also had their commonality verified or disproven, but the latest contribution to the field was by Hatami, et. al. [7], who showed that the 5-wheel is common (see Figure 1).

This paper's main result is a proof of the commonality of the pentagon with a chord[1] (see

[1] The commonality of the pentagon with a chord was proposed by Jagger, et. al. in [4] as one of the most important unsolved problems in the field of graph commonality.

Figure 1: (from left to right) K_4, the triangle with a pendant edge, and the 5-wheel, whose commonalities are determined. Lastly, the pentagon with a chord, whose commonality we prove in this paper.

Figure 1), a graph belonging to the category of odd cycles with a chord, a class of graphs whose commonality is completely undetermined until now. However, before that, we analyze disconnected graphs, including a modern proof that a graph formed from any number of disjoint clones of a common graph is common, originally proven by Jagger, et. al. [4].

2 An Analysis of Disconnected Graphs

Most commonality studies to date have focused on connected graphs. For instance, Jagger, et. al. [4] disregarded the problem of determining the commonality of disconnected graphs. However, interestingly, they immediately afterwards described the vast potential for research in the area of disjoint graphs.

In particular, it is unknown in generality when the disjoint union of two common graphs produces another common graph. It is known that the disjoint union of a triangle and a lone edge, both of which are common graphs, is uncommon [4]. However, the disjoint union of two triangles is common [8]. It is also undetermined whether the union of a common graph and an uncommon graph must always be uncommon. Furthermore, the commonality of the disjoint union of two uncommon graphs is undecided [4]. Disjoint graphs are so sparsely studied that even the commonality of the graph composed of a disjoint triangle and square has not yet definitively been established. However, my floating point calculations suggest that the disjoint triangle and square is common, up to an error of order 10^{-32}.

However, in 1996, Jagger, et. al. [4] proved that any number of disjoint clones of a common graph form a disconnected graph that is itself common. Here we propose a modern proof of this theorem using Hölder's Inequality.

2.1 A Modern Proof of Commonality of Disjoint Clones of Common Graphs

Let us define a *homomorphism* of a graph F inside another graph G as a map from the vertices of F to the vertices of G that preserves edges, but not necessarily non-edges, between vertices. We then define the *homomorphism density* $t(F, G)$ of a graph F in another graph G as the number of homomorphisms of F in G divided by $\binom{|V(G)|}{|V(F)|}$.

Note that, by definition, a graph F is common if and only if

$$t(F, G) + t(F, \bar{G}) \geq 2^{1-|E(F)|} \tag{1}$$

for all graphs G, where $\bar{G}$ denotes the complement of G. Commonness is a particularly interesting asymptotic property of graphs because it is a tight bound; it can be attained for any graph F by taking G to be the random graph with edge density one-half.

Theorem 2.1 (Jagger, et. al., 1996)**.** *Let F be a common graph. Then, for all positive integers n, the graph F_n composed of n disjoint clones of F is common.*

Proof. Because all the clones of F are disjoint, $t(F_n, G) = t(F, G)^n$. Similarly, $t(F_n, \bar{G}) = t(F, \bar{G})^n$. Thus, we have

$$t(F_n, G) + t(F_n, \bar{G}) = t(F, G)^n + t(F, \bar{G})^n.$$

To prove our theorem, we use Hölder's Inequality, which can be stated as

$$\sum_{i=1}^{k} a_i b_i \leq \left(\sum_{i=1}^{k} a_i^p\right)^{1/p} \left(\sum_{i=1}^{k} b_i^q\right)^{1/q},$$

where the a_i and b_i are non-negative reals, and p and q are positive reals such that $\frac{1}{p}+\frac{1}{q}=1$.

For our application of Hölder's Inequality, we let $k=2$ and plug in $a_1=a_2=1$, $b_1=t(F_n,G)$, $b_2=t(F_n,\bar{G})$, $p=\frac{n}{n-1}$, and $q=n$. We verify that all the a_i and b_i are non-negative because subgraph densities are non-negative, and that $\frac{1}{p}+\frac{1}{q}=\frac{n-1}{n}+\frac{1}{n}=1$. Thus, by Hölder's Inequality, we have

$$t(F,G)+t(F,\bar{G}) \leq 2^{(n-1)/n} \cdot \left(t(F,G)^n + t(F,\bar{G})^n\right)^{\frac{1}{n}}.$$

Because $t(F,G)+t(F,\bar{G}) \geq 2^{1-|E(F)|}$, we have $(t(F,G)+t(F,\bar{G}))^n \geq 2^{n-n\cdot|E(F)|}$, yielding,

$$2^{n-n\cdot|E(F)|} \leq (t(F,G)+t(F,\bar{G}))^n \leq 2^{n-1} \cdot (t(F,G)^n + t(F,\bar{G})^n).$$

Simplifying, we see that our expression above reduces to $2^{1-n|E(F)|} \leq t(F,G)^n + t(F,\bar{G})^n$. Because F_n has $n \cdot |E(F)|$ edges, inequality 1 implies that F_n is common. □

Our proof affords us the opportunity to investigate the equality case of Hölder's Inequality to glean some insight into graphs composed of disjoint clones of uncommon graphs, such as in the following proposition.

Theorem 2.2. *If there is a large graph G' that is a counterexample to the commonality of F and also has $t(F,G')=t(F,\bar{G}')$, then the graph formed from any number of disjoint clones of F is a graph that is also uncommon.*

Proof. The equality case of Hölder's Inequality is $b_i = ca_i^{p-1}$ for all i and some constant c. In our proof of Theorem 2.1, $a_1=a_2=1$, so the equality case occurs when $b_1=b_2$, or

$t(F_n, G) = t(F_n, \bar{G})$. Because $t(F_n, G) = t(F, G)^n$ and $t(F_n, \bar{G}) = t(F, \bar{G})^n$, the equality case is just

$$t(F, G) = t(F, \bar{G}). \tag{2}$$

Recall that by definition, a graph F is uncommon if and only if there exists some graph G' such that $t(F, G') + t(F, \bar{G'}) < 2^{1-|E(F)|}$. This is equivalent to

$$(t(F, G') + t(F, \bar{G'}))^n < 2^{n-n\cdot|E(F)|}. \tag{3}$$

In the equality case of Hölder's Inequality, we have

$$(t(F, G') + t(F, \bar{G'}))^n = 2^{n-1} \cdot (t(F, G')^n + t(F, \bar{G'})^n). \tag{4}$$

Combining inequality 3 and equation 4, we get

$$t(F, G')^n + t(F, \bar{G'})^n < 2^{1-n\cdot|E(F)|},$$

which is exactly the definition of uncommonality. □

This new theorem naturally yields the question of which uncommon graphs have counterexamples G' to commonality with $t(F, G') = t(F, \bar{G'})$.

Another class of poorly studied disconnected graphs is that of an uncommon graph with a disjoint edge. We propose the following partial result.

Theorem 2.3. *If F is an uncommon graph such that there is a graph G' with $t(F, G') + t(F, \bar{G'}) < 2^{1-|E(F)|}$ such that G' has edge density one-half, then the graph F' formed from adding a disjoint edge to F is also uncommon.*

Proof. We claim that G' is also a counterexample to the commonality of F'. We note, from the definition of commonality in inequality 1, that F' is uncommon if

$$t(F', G') + t(F', \bar{G}') < 2^{1-|E(F')|} = 2^{-|E(F)|}.$$

Also, because G' is disjoint, we know $t(F', G') = \frac{1}{2}t(F, G)$ because the probability that the additional edge is present in G' is $\frac{1}{2}$. Similarly, $t(F', \bar{G}') = \frac{1}{2}t(F, G)$. Hence,

$$t(F', G') + t(F', \bar{G}') = \frac{1}{2}(t(F, G') + t(F, \bar{G}')) < 2^{-|E(F)|},$$

proving uncommonality. □

The difference between the conditions of Theorems 2.2 and 2.3 is particularly insightful into the general commonality of graphs because the relationship between the existence of counterexamples with $t(F, G') = t(F, \bar{G}')$ and counterexamples with edge density $\frac{1}{2}$ is unknown. However, letting G' be a complete bipartite graph and F be a triangle shows that a counterexample to commonality can be edge symmetric (the density of edges and non-edges is equal) but not necessarily F-symmetric (the density of F-copies and F-anticopies is equal).

3 Algebraic Structure of Graphs

Sidorenko's original proof in [5] that the triangle is common is based on an asymptotic combinatorial argument that explicitly calculates the prevalence of homomorphisms of the triangle in a random graph and then uses a monovariant to establish that the number of homomorphisms of the triangle in a random graph is minimal. His proof of the uncommonality of the triangle with a pendant edge provides a coloring with fewer homomorphisms of the triangle with a pendant edge than in a random graph. The proof that K_4 is uncommon also is based on an explicit coloring, though it was discovered using a computer program [9].

Modern methods for disproving commonality include finding explicit colorings, which are very large and are discovered through a computer search, and a method of Fourier analysis developed in [4]. There is also a technique that allows for a graph to be proven to be common algebraically, in terms of smaller graphs. The technique is known as the method of *flag algebras*, which was developed by Razborov [10].

3.1 The Algebra of Homomorphism Densities

Flag algebras can be intuitively explained through adjacency matrices of simple graphs. For any simple graph, there exists a symmetric adjacency matrix to describe that graph. Each entry in the matrix is either a 1 or a 0, representing an edge or a non-edge, respectively, between two vertices.

Graphs naturally generalize to weighted graphs, which can be viewed as symmetric adjacency matrices with values in the interval $[0, 1]$, rather than only in the set $\{0, 1\}$. Let us also denote the set of all weighted graphs as W_0. We are interested in weighted graphs because they generalize our problem, as the set of adjacency matrices with values in $\{0, 1\}$ is a subset of W_0.

Recall that, by definition, a graph F is common if and only if $t(F, G)+t(F, \bar{G}) \geq 2^{1-|E(F)|}$ for all $G \in W_0$. Also, note that the homomorphism density of F in G can be expressed as the expected value of the product of all the values in G's adjacency matrix corresponding to the edges of F. In other words,

$$t(F, G) = \mathbb{E}_{\varphi}\left(\prod_{e \in F} A_G(\varphi(e))\right), \tag{5}$$

where φ is a homomorphism of F in G, and $A_G(\varphi(e))$ represents the value of the edge e in the adjacency matrix of G.

Taking the complement of a graph is almost equivalent[2] to subtracting each entry of its adjacency matrix from 1. Hence, we can also write

$$t(F, \bar{G}) = \mathbb{E}_{\varphi}\left(\prod_{e \in F}(1 - A_G(\varphi(e)))\right). \tag{6}$$

Thus, commonality is equivalent to

$$\mathbb{E}_{\varphi}\left(\prod_{e \in F} A_G(\varphi(e)) + \prod_{e \in F}(1 - A_G(\varphi(e)))\right) \geq 2^{1-|E(F)|}.$$

To prove commonality, we employ a method known as the *Cauchy-Schwarz method*, or the *sum of squares method*. However, before we can meaningfully define multiplication or squares of graphs, we must first define *labeled graphs*.

A labeled graph is a simple graph such that some number of vertices have been marked and every marked vertex has a unique label. The unlabeled vertices are all indistinguishable (ignoring their connectivity). We denote a k-labeled graph to be a graph with k labeled vertices. Thus, a 0-labeled graph has no labels, and a $|V(G)|$-labeled graph is completely labeled. A partially labeled graph is merely a labeled graph that is neither 0-labeled nor completely labeled.

The homomorphism density of a partially labeled graph is defined similarly to the homomorphism density of a simple graph. The labeled vertices are mapped to fixed vertices in G through an extra parameter x, and the homomorphism proceeds normally from the unlabeled vertices. Thus, the density of a labeled graph is denoted $t_x(F, G)$, where x is a parameter that describes the images of the labeled vertices under the homomorphic mapping.

We also define a form of multiplication of partially labeled graphs known as the gluing

[2]The only difference between $1 - G$ and $\bar{G}$ is along the diagonal of the adjacency matrix where $1 - G$ contains 1s (loops from edges to themselves), whereas the complement does not. However, as the size of a graph increases, the measure of the diagonal tends towards zero, and thus the difference between $1 - G$ and $\bar{G}$ is negligible.

product. We can multiply two k-labeled graphs by *identifying*, or overlapping, their labeled vertices. If there are two labeled vertices with an edge between them in both terms of the product, then we keep both edges in the result. Although this may create a multigraph, the definition of density remains the same and all aforementioned results still hold. The graph formed by multiplying two labeled graphs is known as the *product* of the two original graphs. Note that the homomorphism density of the product of two graphs is the product of their homomorphism densities.

Just like we can multiply k-labeled graphs, we can add them to form what is known as a quantum graph. A quantum graph is merely a linear combination of k-labeled graphs for some k. We define the homomorphism density of a quantum graph as the sum of the homomorphism densities of its terms.

Note that both addition and multiplication of k-labeled quantum graphs yield other k-labeled quantum graphs. However, our original problem of commonality is stated in terms of simple graphs. To convert from labeled graphs back to unlabeled graphs, we introduce an unlabeling operator $[\![Q]\!]$ that is applied to a graph or quantum graph Q and yields out the unlabeled graph or quantum graph with the same structure.

To convert from a labeled graph's density to an unlabeled graph's density, the unlabeling operator takes the expected homomorphism density over every way to fix the labeled vertices in the larger graph G. Thus, although the labeling operator does not preserve homomorphism densities exactly, it does preserve non-negativity of densities because expectation preserves non-negativity. In other words, $t([\![Q]\!], G) \geq 0$ if $t_x(Q, G) \geq 0$ for all x.

We have thus created a structure that expresses addition and multiplication of graph densities in terms of the underlying graphs themselves[3]. In fact, we have established an algebra

[3] We glossed over a relatively minor detail in this discussion — we are able to interchange any two graphs with the same density in every G. For instance, because adding a disconnected vertex to a graph does not introduce any new edges, we can always add or remove disconnected vertices from a graph to produce a new graph equivalent to the original graph.

over the field of the real numbers because we have some notion of addition, multiplication of a graph (and its density) by a scalar, and multiplication of two graphs. We can therefore drop the t density notation and simply write the graphs themselves to represent their densities.

Because our algebra is over the reals, we can define squares of graphs as the products of graphs by themselves. Squares of quantum graphs are identically defined. Also, because graph densities are all real numbers, we see that the square of a density is always non-negative. Thus, the square of a graph always has non-negative homomorphism density. Note that this is true even if we extend our edge weights to include negative numbers. In that case, homomorphism densities, which are defined as in equation 5, can be negative. However, because of the non-negativity of real number squares, the squares of graphs and quantum graphs still have non-negative homomorphism densities.

3.2 Reduction of the Commonality Definition

Because we have established the validity of our density algebra over edge weights that include negative numbers, we consider the transformation $A_{1,G}(\varphi(e)) := 2A_G(\varphi(e)) - 1$, which maps $[0, 1]$, the original codomain of A_G, to $[-1, 1]$, the codomain of this new $A_{1,G}$ function. If we replace each edge weight $A_G(\varphi(e))$, or entry in the adjacency matrix of G, with $A_{1,G}(\varphi(e)) = 2A_G(\varphi(e)) - 1$, then we can consider the resulting graph and adjacency matrix as a weighted graph on the interval $[-1, 1]$. We denote the set of all such weighted graphs as W_1. Note that we can define a new density

$$t_1(F, G) = \mathbb{E}_\varphi \left(\prod_{e \in F} A_{1,G}(\varphi(e)) \right)$$

similar to equation 5 above.

We now return to our original definition of commonality in terms of homomorphic densities, but using our new system of edge weights. The resulting inequality is

$$\mathbb{E}_{\varphi}\left(\prod_{e\in F}\left(\frac{1+A_{1,G}(\varphi(e))}{2}\right)+\prod_{e\in F}\left(\frac{1-A_{1,G}(\varphi(e))}{2}\right)\right)\geq 2^{1-|E(F)|}. \tag{7}$$

We see that expanding the inside of the expectation leads to some summation over all subgraphs F' of F. The intuition behind this is that each term of each product requires a choice between choosing $\frac{1}{2}$, which represents not including that edge in the subgraph, or $\frac{A_{1,G}(\varphi(e))}{2}$, which represents including that edge.

In the expansion of inequality 7, the empty graph yields a constant term of $2 \cdot 2^{-|E(F)|} = 2^{1-|E(F)|}$, which cancels with the constant term on the right hand side of inequality 7. Also, the opposite signs inside the two products lead to cancellation of all the terms with an odd number of edges.

Thus, inequality 7 above is equivalent to

$$\mathbb{E}_{\varphi}\left(\sum_{\substack{F'\subseteq F\\|E(F')|\neq 0\\|E(F')|\equiv 0 \pmod 2}}\prod_{e\in F'}A_{1,G}(\varphi(e))\right)\geq 0=\sum_{\substack{F'\subseteq F\\|E(F')|\neq 0\\|E(F')|\equiv 0 \pmod 2}}\mathbb{E}_{\varphi}\left(\prod_{e\in F'}A_{1,G}(\varphi(e))\right)\geq 0,$$

which can be rewritten as

$$\sum_{\substack{F'\subseteq F\\|E(F')|\neq 0\\|E(F')|\equiv 0 \pmod 2}} F'\geq 0, \tag{8}$$

dropping the t notation in favor of simply writing the graphs to represent their densities.

The main idea of the transformation from W_0 to W_1 is as follows: the flag algebras method differs from traditional methods of proof in that it writes the density of G in an arbitrary large graph K with edges weighted in the interval $[-1,1]$ as the sum of the densities of all of the subgraphs of G with an even number of edges. To prove that the original graph G is common, we must only show that this sum of subgraph densities is always non-negative.

The left-hand side of inequality 8 is the density of a quantum graph formed from the sum of all the subgraphs of F with an even number of edges. To prove non-negativity of this quantum graph's density, we sum together inequalities of the form $t(Q^2, G) = t(Q, G)^2 \geq 0$.

The beauty of the transformation to W_1 and the sum of squares method is the sheer simplicity of the proofs that result from it.

4 A Flag Algebras Proof of the Main Result

After defining the algebraic structure in the previous section, we can hand over the proof of commonality to a computer that searches the space of squares of quantum graphs for a set of squares that sum to the expression we are trying to prove is non-negative. We wrote a program to prove our main result, which follows.[4] We also used the NEOS online semi-definite solver [11] for numerical calculations.

Theorem 4.1. *The graph known as the pentagon with a chord is common.*

Proof. Our flag algebras program requires the input of some set of labeled quantum graphs $F_1, F_2, \ldots, F_m$, which serves as the basis the computer uses to search for a sum of squares proof of non-negativity. However, for the sake of minimizing computational complexity and ease of proof verification, it is convenient to use *induced density*, a slightly different algebra structure than defined in Section 3 [12]. When using induced density, the definition of graph multiplication involves removal of double edges. In other words, if two labeled vertices both contain an edge in the multiplicand graphs, then the product contains only one edge between those two vertices. This is contrary to the product defined in Section 3, where multiple edges were retained. However, squares of graphs are still non-negative in the induced basis [8].

[4]I would like to thank my mentor, James Hirst, for providing much of the code base necessary for the project.

Converting from the standard density of a graph F in W_0 to induced density involves applying an invertible linear transformation on F defined by,

$$\mathrm{Hom}(F) = \sum_{F' \supseteq F} F', \tag{9}$$

where $|V(F')| = |V(F)|$ and F' denotes the homomorphism density of F' in W_0.

$\mathrm{Hom}(F)$ is invertible because we can apply a Möbius Inversion to get the inverse of $\mathrm{Hom}(F)$, defined as

$$\mathrm{Ind}(F) = \sum_{F' \supseteq F} (-1)^{|E(F)|-|E(F')|} F', \tag{10}$$

where the same conditions hold as for the definition of Ind.

We recall that we are trying to prove that $t(F,G) + t(F,\bar{G}) - \frac{1}{32} \geq 0$ for all G, where F is the pentagon with a chord. The $\frac{1}{32}$ comes from the definition of commonality and the fact that F has six edges.

Also recall from Section 3 that we can approximate $\bar{G}$ with $1 - G$, which represents subtracting each entry of G's adjacency matrix from one. Then, by the definition of homomorphism density from equation 6,

$$t(F,\bar{G}) = \mathbb{E}_{\varphi}\left(\prod_{e \in F} (1 - A_G(\varphi(e)))\right).$$

Expanding this product like in section 3 gives us that

$$t(F,\bar{G}) = \sum_{F' \subseteq F} (-1)^{|E(F')|} t(F', G).$$

Thus, our proof of commonality reduces to proving that

$$\sum_{F' \subseteq F} (-1)^{|E(F')|} t(F', G) + t(F,G) - \frac{1}{32} \geq 0. \tag{11}$$

Because a disjoint vertex has no edges, it has homomorphic density of 1. Hence, the constant term $-\frac{1}{32}$ can actually be written as $-\frac{1}{32} \cdot \circ$, where $\circ$ represents a single disjoint vertex. Thus, the left hand side of inequality 11 is the homomorphism density of some quantum graph f.

Let $g = \mathrm{Hom}(f)$. Then, taking Hom of both sides of the inequality $f \geq 0$ yields $g \geq 0$, because the Hom of a zero quantum graph is another zero quantum graph.

In order to prove that $g \geq 0$, we are searching for a positive semidefinite matrix[5] Y of real numbers and a vector of quantum graphs z such that $z^T Y z = g$.

Our computer program indicates that the vector g (computed by plugging in the pentagon with a chord for F) is

$$\{31/32, 59/16, 33/16, -5/16, -5/8, 81/32, 1/8, -15/8, -15/32, -15/16, -5/32, -5/32,\\ 17/16, -15/8, -15/16, -7/8, -15/8, -15/8, -15/8, -5/16, -15/8,\\ -15/8, -5/8, -5/16, -5/16, -15/32, -3/8, -7/8, 1/8, 17/16, 33/16, 81/32, 59/16, 31/32\},$$

where the entries of the vector are coefficients of the 5-vertex partially labeled graphs (written with their adjacency matrices in lexicographical order). Consider the matrices:

$$Y_1 = \begin{pmatrix} \frac{31}{32} & -\frac{119}{416} & -\frac{383}{832} & -\frac{187}{832} & \frac{1}{416} \\ -\frac{119}{416} & \frac{431}{416} & \frac{89}{832} & \frac{17}{832} & -\frac{365}{416} \\ -\frac{383}{832} & \frac{89}{832} & \frac{83}{208} & -\frac{109}{832} & \frac{71}{832} \\ -\frac{187}{832} & \frac{17}{832} & -\frac{109}{832} & \frac{107}{208} & -\frac{149}{832} \\ \frac{1}{416} & -\frac{365}{416} & \frac{71}{832} & -\frac{149}{832} & \frac{31}{32} \end{pmatrix},$$

[5]A positive semidefinite matrix has all non-negative eigenvalues. The positive semidefiniteness of Y guarantees non-negativity of g because there exists a linear change of basis (like an eigenvalue decomposition) that can turn Y into a symmetric matrix Y' with all positive entries. Then, $z^T Y' z$ is guaranteed a sum of squares, which is non-negative. Since we performed an invertible linear transformation on Y to generate Y', $z^T Y z$ must also be expressible as a sum of squares.

$$Y_2 = \begin{pmatrix} \frac{5}{13} & -\frac{243}{1664} & -\frac{563}{832} & \frac{563}{832} & \frac{243}{1664} & -\frac{5}{13} \\ -\frac{243}{1664} & \frac{103}{832} & \frac{295}{1664} & -\frac{475}{1664} & -\frac{1}{64} & \frac{63}{1664} \\ -\frac{563}{832} & \frac{295}{1664} & \frac{665}{416} & -\frac{547}{416} & -\frac{59}{128} & \frac{799}{832} \\ \frac{563}{832} & -\frac{475}{1664} & -\frac{547}{416} & \frac{41}{32} & \frac{531}{1664} & -\frac{591}{832} \\ \frac{243}{1664} & -\frac{1}{64} & -\frac{59}{128} & \frac{531}{1664} & \frac{131}{832} & -\frac{479}{1664} \\ -\frac{5}{13} & \frac{63}{1664} & \frac{799}{832} & -\frac{591}{832} & -\frac{479}{1664} & \frac{33}{52} \end{pmatrix},$$

$$Y_3 = \begin{pmatrix} \frac{1}{2} \end{pmatrix}, Y_4 = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & -\frac{1}{8} & -\frac{1}{8} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{16} & \frac{1}{16} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{16} & \frac{1}{16} \end{pmatrix}, \text{and } Y_5 = \begin{pmatrix} \frac{9}{8} & \frac{137}{208} \\ \frac{137}{208} & \frac{9}{13} \end{pmatrix}.$$

Then, let Y be the 18 by 18 matrix formed by lining Y_1, Y_2, Y_3, Y_4, and Y_5 along the major diagonal and then padding the rest of the matrix with zeros. It can be verified that Y is positive semidefinite.

The program also tells us that the vector z is

$$\begin{aligned} \{&(1,0,0,0,0,0),(0,0,1,0,0,0),(0,0,0,1,0,0),(0,0,0,0,1,0),(0,0,0,0,0,1), \\ &(1,0,0,0,0,0,0,0),(0,1,1,0,0,0,0,0),(0,0,0,1,0,0,0,0), \\ &(0,0,0,0,1,0,0,0),(0,0,0,0,0,1,1,0),(0,0,0,0,0,0,0,1),(0,1,-1,0,0,0,0,0), \\ &(1,0,0,0,0,0,0,0),(0,1,0,0,0,0,0,0),(0,0,1,0,1,0,0,0), \\ &(0,0,0,1,0,1,0,0),(0,0,1,0,-1,0,0,0),(0,0,0,1,0,-1,0,0)\}, \end{aligned}$$

where each ordered n-tuple represents a quantum graph. The values of the elements of the n-tuples are the coefficients of each labeled graph that form the quantum graph. The labeled graphs are ordered lexicographically on their lowest adjacency matrix. The first five quantum graphs are linear combinations of the 1-labeled graphs on 3 vertices. The rest are linear combinations of the 3-labeled graphs on 4 vertices.

It can be verified that $z^T Y z = g$ holds. Because Y is positive semidefinite, this indicates that there exists a representation of g as a sum of squares, which means that it must be non-negative. □

Note that our switch to the induced basis works because Ind and Hom are linear transformations. Thus, we can apply them to quantum graphs as well as simple graphs. Furthermore, the invertibility of Ind and Hom guarantees us that the basis spanned by the quantum graphs $\text{Ind}(F_1)$, $\text{Ind}(F_2), \ldots, \text{Ind}(F_m)$ is the same as the basis of graphs spanned by $F_1, F_2, \ldots, F_m$. Hence, we can have our program search using the basis $\text{Ind}(F_1)$, $\text{Ind}(F_2), \ldots, \text{Ind}(F_m)$ and have it be theoretically equivalent to searching with the more computationally expensive basis $F_1, F_2, \ldots, F_m$.

The creation of this proof relied on a combination of my program and some experimentation regarding which vectors $F_1, F_2, \ldots, F_m$ (the elements of z) are necessary to create a search space large enough for the program to find a proof.

Also, we attempted many different strategies to try to come up with simple proofs of the commonality of the pentagon with a chord in W_1 by hand. However, we could find no simple proofs, and experimentation with the program's search space, defined by the z vector, hinted that the provided commonality proof is minimal or at least close to minimal.

5 Conclusion

In this paper, we provided a modern proof that the graph composed of any number of disjoint clones of a common graph is itself common. Our application of Hölder's Inequality led to new insights and partial results about the nature of graph commonality and uncommonality. This opened new doors into research possibilities regarding the commonality of disjoint graphs. More specifically, it is unknown in what cases the disjoint union of two common graphs is common. Similarly defined problems, such as the disjoint union of two uncommon graphs, are also undecided.

We then proved a completely new result, closing a problem originally proposed by Jagger, et. al. [4] central to the field of graph commonality. Namely, we proved that the pentagon with a chord is a common graph. We used the technique of flag algebras to produce a space for a computer search, and then wrote a program using the NEOS Server [11] to calculate the final matrices for the proof itself.

This research is important in the context of graph commonality in general because it provides the first step towards exactly enumerating the set of all non-bipartite common graphs.[6] The simplest classes of non-bipartite graphs are odd cycles and odd cycles with a pendant edge, with the former proven common in [5] and the latter proven uncommon in [6]. Thus, the simplest class of non-bipartite graphs to investigate is that of the odd cycles with a chord. In this paper, by establishing the commonality of the pentagon with a chord, we break the ground for research to continue regarding this category of graphs. I am currently pursuing further research in this direction using theoretical flag algebras.

The future applications of my research on graph commonality include analysis of large networks, including the Internet and social networks. Subnetworks of people that are represented by common graphs are guaranteed to occur relatively often within the databases of companies like Facebook and Google, making such social or web structures high-value research targets. By identifying common graph-based structures that can be efficiently monetized, companies can gain the competitive edge in the market and increase revenues.

My research to uncover knowledge about common graphs is even relevant to topics like ecology and city planning. For instance, Urban and Keitt [14] outlined how ecological habitats can be modeled using large networks. Common graphs represent frequently occurring patterns in the environment; for instance, a pentagon with a chord could represent a patch

[6] The inclination among mathematicians to investigate non-bipartite graphs arises from the much-hailed Sidorenko's Conjecture in [13], which (if true) implies that all bipartite graphs are common. Proving Sidorenko's Conjecture in any generality is esteemed to be an extremely difficult problem, so commonality research often focuses on non-bipartite graphs.

of forest with two lakes inside it. Such an analysis would allow urban designers to identify specific regions for preservation to mitigate human impact on the local ecosystem.

With such a diverse range of applications for research on common graphs and such a breadth of open questions in the field, I am continuing research for a better understanding of the commonality of odd cycles with a chord in general, as well as a deeper study of disconnected graphs.

6 Acknowledgments

I would like to thank my mentor, James Hirst from MIT, for his tireless dedication in helping me complete my research project. I would also like to thank my RSI tutor, the Rose-Hulman Institute of Technology's Dr. John Rickert, as well as MIT's Dr. Tanya Khovanova, who introduced me to the presentation and article-writing process. My regards also go out to Professors David Jerison and Ankur Moitra from MIT for overseeing my project. I am grateful to the Center for Excellence in Education (CEE), the Research Science Institute (RSI) and its sponsors, and the Massachusetts Institute of Technology (MIT) for facilitating my research. Finally, my thanks also extend to my sponsors, Mr. Jonathan A. Marcus, Dr. Radhakrishnan Jagadeesan and Dr. Lalita Jagadeesan from De-Paul University, and Professor Chris Kendrick and Professor Kasturi Haldar from Loyola University.

References

[1] A. W. Goodman. On sets of acquaintances and strangers at any party. *The American Mathematical Monthly*, 66(9):778–783, 1959.

[2] F. P. Ramsey. On a problem in formal logic. *Proc. London Math. Soc.*, 3(30):264–286, 1930.

[3] A. Thomason. A disproof of a conjecture of Erdõs in Ramsey theory. *J. London Math. Soc*, 39:246–255, 1989.

[4] C. Jagger, P. Šťovóček, and A. Thomason. Multiplicities of subgraphs. *Combinatorica*, 16(1):123–141, 1996.

[5] A. Sidorenko. Cycles in graphs and gunctional inequalities. *Mathematical Notes of the Academy of Sciences of the USSR*, 46(5):877–882, 1989.

[6] J. Fox. There exist graphs with super-exponential ramsey multiplicity constant. *Journal of Graph Theory*, 57(2):89–98, 2008.

[7] H. Hatami, J. Hladky, D. Kral, S. Norine, and A. Razborov. Non-three-colorable common graphs exist. arXiv:1105.0307, 2011.

[8] L. Lovász. *Large Networks and Graph Limits*. American Mathematical Society, 2012.

[9] F. Franek and V. Rödl. 2-colorings of complete graphs with a small number of monochromatic {K4} subgraphs. *Discrete Mathematics*, 114(1–3):199–203, 1993.

[10] A. A. Razborov. Flag algebras. *The Journal of Symbolic Logic*, 72(04):1239–1282, 2007.

[11] Neos sdpa online solver. `http://www.neos-server.org/neos/solvers/sdp:SDPA/SPARSE_SDPA.html`. Accessed: 2015-07-28.

[12] J. Hirst. The inducibility of graphs on four vertices. *Journal of Graph Theory*, 75(3):231–243, 2014.

[13] A. Sidorenko. A correlation inequality for bipartite graphs. *Graphs and Combinatorics*, 9(2-4):201–204, 1993.

[14] D. Urban and T. Keitt. Landscape connectivity: A graph-theoretic perspective. *Ecology*, 82(5):pp. 1205–1218, 2001.