Independent Sets and Eigenspaces

by

Michael William Newman

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of

Doctor of Philosophy

in

Combinatorics and Optimization

Waterloo, Ontario, Canada

Author's declaration for electronic submission of a thesis

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The problems we study in this thesis arise in computer science, extremal set theory and quantum computing. The first common feature of these problems is that each can be reduced to characterizing the independent sets of maximum size in a suitable graph. A second common feature is that the size of these independent sets meets an eigenvalue bound due to Delsarte and Hoffman. Thirdly, the graphs that arise belong to association schemes that have already been studied in other contexts.

Our first problem involves covering arrays on graphs, which arises in computer science. The goal is to find a smallest covering array on a given graph G. It is known that this is equivalent to determining whether G has a homomorphism into a *covering array graph*, $CAG(n, g)$. Thus our question: Are covering array graphs cores? A covering array graph has as vertex set the partitions of $\{1, \ldots, n\}$ into g cells each of size at least g, with two vertices being adjacent if their meet has size g^2. We determine that $CAG(9, 3)$ is a core. We also determine some partial results on the family of graphs $CAG(g^2, g)$. The key to our method is characterizing the independent sets that meet the Delsarte-Hoffman bound—we call these sets *ratio-tight*. It turns out that $CAG(9, 3)$ sits inside an association scheme, which will be useful but apparently not essential.

We then turn our attention to our next problem: the Erdős-Ko-Rado theorem and its q-analogue. We are motivated by a desire to find a unifying proof that will cover both versions. The EKR theorem gives the maximum number of pairwise disjoint k-sets of a fixed v-set, and characterizes the extremal cases. Its q-analogue does the same for k-dimensional subspaces of a fixed v-dimensional space over $GF(q)$. We find that the methods we developed for covering array graphs apply to the EKR theorem. Moreover, unlike most other proofs of EKR, our argument applies equally well to the q-analogue. We provide a proof of the characterization of the extremal cases for the q-analogue when $v = 2k$; no such proof has appeared before. Again, the graphs we consider sit inside of well-known association schemes; this time the schemes play a more central role.

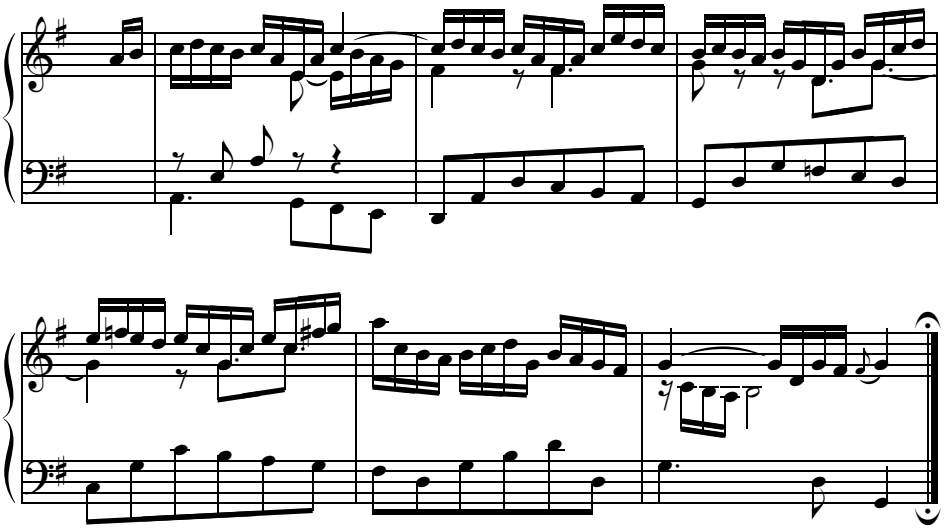
Finally, we deal with the problem in quantum computing. There are tasks that can be performed using quantum entanglement yet apparently are beyond the reach of methods using classical physics only. One particular task can be solved classically if and only if the graph $\Omega(n)$ has chromatic number n. The graph $\Omega(n)$ has as vertex set the set of all ± 1 vectors of length n, with two vertices adjacent if they are orthogonal. We find that n is a trivial upper bound on the chromatic number, and that this bound holds with equality if and only if the Delsarte-Hoffman bound on independent sets does too. We are thus led to characterize the ratio-tight independent sets. We are then able to leverage our result using a recursive argument to show that $\chi(\Omega(n)) > n$ for all $n > 8$. It is notable that the reduction to independent sets, the characterization of ratio-tight sets, and the recursive argument all follow from different proofs of the Delsarte-Hoffman bound. Furthermore, $\Omega(n)$ also sits inside a well-known association scheme, which again plays a central role in our approach.

Acknowledgements

I owe a big thank you to my supervisor Chris Godsil. I have benefited from his knowledge, support, encouragement and generosity in ways I didn't even imagine when I first contemplated this endeavour.

Thank you to my frabjous siblings and parents, Jesse, Jenna, Karin, Mom, Dad. You've always been there for me, and it shows.

Lastly, the principal ideas and directions of this thesis were mostly the result of William's stimulating conversation, and it was Timothy's clear thinking that worked through the hard parts.



Contents

Chapter 1

Introduction

The purpose of this thesis is to explore the theory and applications of a particular bound on independent sets in regular graphs, which we call the ratio bound.

Our interest in the ratio bound on independent sets is driven by the problems that we can solve as a consequence of it. These problems have their origin in computer science, extremal graph theory, and quantum computing; they seem to have little in common. In fact, it is not the problems that have something in common so much as their solutions. Each of the problems that we look at can be phrased in terms of a graph in an association scheme, and our solution in each case depends critically on a characterization of the independent sets that meet the ratio bound. We call these sets *ratio-tight*.

The characteristic vector of a ratio-tight set is a linear combination of $\mathbf{1}$ and an eigenvector for the least eigenvalue, which in turn implies that the partition induced by a ratio-tight set is equitable. This tells us where to look for them: in the eigenspace corresponding to the least eigenvector, suitably shifted by $\mathbf{1}$. It also tells us how they connect to the rest of the graph.

Association schemes have been used extensively in coding theory. Delsarte showed how to use the algebraic structure of an association scheme to obtain bounds on codes in the scheme. While a code can be thought of as an independent set, our use of association schemes is not motivated by coding theory. We use the algebraic structure of association schemes to obtain information on eigenspaces and how they relate to the graph.

We do not wish the reader to think that our purpose is simply to find eigenspaces and big independent sets. We use these tools to solve other problems. We will show that a partition graph is a core: this has consequences in the construction of covering arrays on graphs. We will determine when a particular colouring bound in a family of orthogonality graphs holds with equality: this has consequences in quantum computing. We deal with the Erdős-Ko-Rado theorem and its q-analogue, and show how all of these problems can be treated in a unified way.

1.1 Main Results

We use the ratio bound to solve specific open problems. In doing so, we develop a technique for characterizing ratio-tight maximum independent sets. We also give new generalizations of the ratio bound to non-regular graphs. Our problems come from three sources: covering arrays on graphs, the well-known Erdős-Ko-Rado theorem and its q-analogue, and quantum entanglement.

Covering Arrays

The partition graph $\mathcal{P}(3^3)$ has as its vertex set all partitions of $\{1, \ldots, 9\}$ into three cells of equal size, with two vertices being adjacent if each cell of one partition intersects every cell of the other. This graph is significant in studying covering arrays, which in turn find applications in software testing. The existence of a covering array of size 9 on an alphabet of size 3 for a graph G (see Section 4.2 for details) is equivalent to the existence of a homomorphism from G to $\mathcal{P}(3^3)$. It is both natural and of practical significance to ask whether $\mathcal{P}(3^3)$ is minimal with this property. Minimality would be equivalent to there being no homomorphism from $\mathcal{P}(3^3)$ to a proper subgraph: in other words, is $\mathcal{P}(3^3)$ a core?

We prove that $\mathcal{P}(3^3)$ is a core, and derive some partial results on more general partition graphs. We actually provide two proofs. The first is based on an analogy with a proof that the Petersen graph is a core. In doing so, we develop some new tools that may be useful for dealing with other questions about graph cores. The second proof uses a characterization of the maximum independent sets in $\mathcal{P}(3^3)$. There is some possibility of extending the full result to more general partition graphs.

Erdős-Ko-Rado

The Erdős-Ko-Rado theorem gives the size of the largest intersecting family of k-subsets of a v-set, and characterizes the extremal cases. In full generality, it asks for the largest family of k-subsets where the intersection of any two members has size at least t, for some $t \geq 1$. This was first posed by Erdős, Ko and Rado [22], who answered the question exactly for $t = 1$, and gave an answer that is valid for sufficiently large v in the case where $t > 1$. Later proofs appeared by Katona [38] and Daykin [17]. The bound was notably improved by Frankl [24] and the exact bound in all cases was finally determined by Wilson [51].

Such largest intersecting systems are exactly maximum independent sets in Kneser graphs. Thus this problem is a natural candidate for our technique. In fact, the bound of the EKR theorem *is* the ratio bound, so from our point of view, the only thing that needs to be proved is the characterization. This turns out to be a consequence of the equality conditions of the ratio bound. So the bound and characterization of the EKR theorem, is just one instance of the bound and equality condition of the ratio bound.

The result is not new, but our approach is novel. Furthermore, we can generalize EKR. What we call the qEKR theorem asks for the largest intersecting family of k-dimensional subspaces of a v-dimensional vector space over $GF(q)$. This was first answered (in most cases) by Hsieh [37], and later in its entirety by Frankl and Wilson [26]. Our approach works here as well, virtually unchanged. This is not true of most proofs of EKR. This generalizability is an important feature of our approach. We take it as a sign that this is not only a proof, but a proof that captures the essence of the problem. Furthermore, we give a proof of the characterization in the qEKR theorem for $v = 2k$, $t = 1$; this case has not appeared before in the literature.

Quantum Entanglement

The final problem we study comes from quantum computing. The question arises as to whether quantum entanglement can be simulated by classical physics. In practical terms, are there operations that can be carried out using quantum entanglement that are infeasible using classical physics? This can be seen as an argument in favour of the existence of quantum entanglement, or at least against the completeness of classical physics.

One approach to this problem is in the format of a question-answer challenge: Alice and Bob are both given a question consisting of a 2^k bit vector, and must answer with a k bit vector. They are not allowed to communicate, and remain ignorant of each other's questions. Their answers must be equal when the questions are, and distinct when the questions are at Hamming distance 2^{k-1}. This was first proposed by Buhrman, Cleve, and Wigderson [10], and also studied by Brassard, Cleve and Tapp [7]. The setup may seem somewhat odd; it was chosen because if Alice and Bob use quantum entanglement, they can trivially answer the challenge. So the question is then how well they can do without entanglement.

In the absence of entanglement, the challenge can be rephrased as a question on the chromatic number of a graph. Specifically, let $\Omega(n)$ be the graph with vertex set the set of all ± 1 vectors of length n, with two vertices adjacent if they are orthogonal. The challenge is then to determine for what values of n is $\chi(\Omega(n)) = n$, for n a power of two.

We prove that $\chi(\Omega(2^k)) = 2^k$ if and only if $k = 1, 2, 3$. The case $k = 3$ was first established by Gordon Royle. Equality was shown not to hold for $k = 4$ by Galliard, Tapp, and Wolf [28]. A deep result of Frankl and Rödl [25] has as a straightforward corollary that the size of a maximum independent set is exponentially less than the ratio bound for large enough k. Our result is self-contained, in that we obtain results for $k = 3, 4$ in a manner different from these other authors. More importantly, it is the first to exactly determine for which values the bound holds with equality.

We will prove the ratio bound in three different ways for these graphs, each time obtaining new information in the case of equality. Furthermore, we are able to apply our method for characterizing ratio-tight sets to prove that there are none: the equality conditions can be used to obtain information when there are

no ratio-tight sets. We view this as strong evidence in favour of the structural importance of the ratio bound: it is not simply a statement about the size of an independent set.

Generalizations

The ratio bound applies only to regular graphs. We find that the proof can be generalized to non-regular graphs. We do this, and in fact obtain a family of bounds. It is likely the case that there is no universally best bound among this family. We consider two natural special cases, show that either of them can be better, and use them to improve known bounds on the Erdős-Rényi graphs.

1.2 Significance

We have provided answers to the motivating problems of this thesis. In some cases, we were able to say more; in all cases, more questions were raised. In the process we were able to make a contribution to the theory of the ratio bound as well.

An extension of qEKR to the general situation of $t > 1$ seems within reach, using a finer analysis of the eigenvectors that we consider here. In the present work, our characterization of the ratio-tight sets rests on a partition of the rows of the eigenvectors into three classes: a single vertex, its neighbourhood, and the rest of the graph. A further partitioning of the neighbourhood based on the size of the intersection shows some promise of yielding a proof for $t > 1$.

Recent computations of De Klerk and Pasechnik have improved the bound on independent sets in $\Omega(16)$, using a new technique due to Schrijver [47]. This technique has important connections to association schemes and to the linear programming bound of Delsarte that is the historical origin of the ratio bound. Their computations give an upper bound of $\alpha(\Omega(16)) \leq 2304$, which we know by construction is tight. This provides evidence for a conjecture on the maximum independent sets in the graphs $\Omega(n)$.

Cameron and Ku [12] have recently characterized maximal intersecting systems in the symmetric group. Their result says that the maximum independent sets are the cosets of the stabilizer of a point. This is in close analogy with the Erdős-Ko-Rado theorem, and it is very likely that our method also applies here.

These are some of the further research directions that have emerged from this thesis.

The fundamental purpose of this thesis is twofold: to establish that the ratio bound is a powerful tool for analyzing structure in graphs, and that association schemes are a natural source of problems and techniques.

Of course, no one tool will solve all problems (else we would fear for our long term employment prospects). But as we will show in this thesis, the ratio bound can be used in different situations, and not just to answer a question about the largest independent set in a graph. In the end, the most important

feature of this bound is not about independent sets. It is about the relationships between a combinatorial structure (independent sets) and an algebraic structure (eigenspaces). It is this interaction that is the main source of richness.

Along the way we discovered that association schemes played a role at almost every step.

Partly this is because the problems we considered involved graphs in association schemes, but this comment can be misleading: We did not pick problems that were phrased in association schemes, the problems presented themselves from other areas of research. Partly this is because the ratio bound has a natural home in an association scheme. Historically, it was first established by Delsarte in an association scheme, and it can be derived in a natural way using the algebraic structure of the scheme. Mostly however, we feel that association schemes played a role because of what they are. They are a structure that has one foot in combinatorics, a rich source of problems and applications, and the other in algebra, a rich source of tools and techniques. We expect to find interesting and important mathematics there.

Chapter 2

Association Schemes

2.1 Introduction

This chapter is intended for background and reference. We do not assume that the reader has prior experience with association schemes, although if they do they will not likely find anything here that they did not already know. The purpose of this chapter is to introduce association schemes and survey some of the main results that will be useful to us. We will offer almost no proofs, although we will sketch many of the ideas behind the proofs. Our presentation has been strongly influenced by Godsil [29, 30]. Proofs of all unreferenced results we state can be found in [29], usually Chapter 12, and occasionally Chapters 10,11. Another useful reference on the subject is the book of Brouwer, Cohen and Neumaier [8]. Bannai and Ito [5] deal with association schemes more from the perspective of group theory. The association schemes we deal with are referred to by some authors (e.g., [5, 18]) as *symmetric association schemes*. There is a slightly more general definition (and most results generalize quite naturally), but we will not need it.

The partition graph of Chapter 4 is a graph in an association scheme. The theorem of Erdős, Ko and Rado that we consider in Chapter 5 can be rephrased in terms of Kneser graphs, which sit inside the Johnson scheme. The orthogonality graph of Chapter 6 sits inside the Hamming scheme. We are of course free to ignore these facts. These problems were posed by researchers who were not aware of the connection to schemes. On the other hand, we are free to use them, and use them to our advantage.

Our motivation for dealing with association schemes is that they are relevant to both the problems we are interested in and the techniques we are using to solve them. Accordingly, our survey will be brief and selective as we collect the particular tools relevant to our work. In an immediate sense, there are two things we need to know how to find: eigenspaces of graphs in association schemes, and eigenvalues of matrices in the algebra generated by the scheme. We will discover in both cases the required information can be read off the

matrix of eigenvalues of the scheme.

The methods we use to solve the problems posed in this thesis are based on an eigenvalue bound on independent sets due to Delsarte and Hoffman; we discuss this bound in some detail in Chapter 3. Significantly, it made its first appearance in association schemes, as Delsarte's linear programming bound. While it is true that Hoffman's extension of this bound to regular graphs is certainly sufficient for our purposes, this is another example of how association schemes arise naturally in the problems we consider.

Association schemes can be regarded as a generalization of distance-regular graphs, which in turn are a generalization of strongly regular graphs. This will partly motivate our approach. Some of our applications of association schemes could be described in terms of strongly regular graphs: Lemma 2.2.3 will be enough for our purposes, instead of the more general Lemma 2.5.4. Some of the association schemes we consider arise from distance-regular graphs: the graph $\Omega(n)$ is the distance-$\frac{n}{2}$ graph of the distance-regular Hamming graph. However, we feel that it is best to phrase our work in the more general context of association schemes.

2.2 Strongly Regular Graphs

We give a brief introduction to strongly regular graphs. Our aim is not to be comprehensive, but rather to develop the main ideas that will generalize to association schemes. The devotee of strongly regular graphs will find this section woefully incomplete; the devotee of association schemes will find it unnecessarily specific. The reader who is less familiar with these subjects will hopefully find it a concrete introduction to the theory of association schemes.

2.2.1 Definition. A *strongly regular graph* on n vertices is a graph that is neither complete nor edgeless, such that the following properties hold.

(a) Every vertex has k neighbours.

(b) Every pair of adjacent vertices has exactly a common neighbours.

(c) Every pair of non-adjacent vertices has exactly c common neighbours.

We refer to $(n, k; a, c)$ as the parameters of the graph. □

The first condition could have been stated more clumsily as "every pair of identical vertices has exactly k common neighbours". This is not an exercise in obfuscation; it leads us to a useful rewording of the definition: the number of common neighbours of any pair of vertices depends only on whether they are equal, adjacent, or distinct and non-adjacent.

A strongly regular graph with $c = 0$ is a disjoint union of cliques; a strongly regular graph with $c = k$ is the complement of a disjoint union of cliques. Moreover, the condition $0 < c < k$ is equivalent to both the graph and its complement being connected [29, p.178]. A *trivial* strongly regular graph has

$c = 0$ or $c = k$. Although much of what we will have to say applies to trivial strongly regular graphs, it will simplify things somewhat to deal only with non-trivial ones.

Some examples are perhaps in order. The Petersen graph is a $(10, 3; 0, 1)$ strongly regular graph. The cycle on 5 vertices is a $(5, 2; 0, 1)$ strongly regular graph; it is the smallest non-trivial strongly regular graph. The line graph of K_n is a $(\binom{n}{2}, 2(n-2); n-2, 4)$ strongly regular graph.

This last example is important for us. When we compute the rank of the matrices M in Chapter 6, what we are doing is computing the eigenvalues of a graph that sits in the algebra generated by the adjacency matrix of the line graph of the complete graph. We will in fact need to compute the eigenvalues of a matrix that lies in this exact algebra in Lemma 6.11.3: this will be an important step in our analysis of the orthogonality graphs of Chapter 6.

If X is a non-trivial strongly regular graph, then it follows that for distinct vertices, being at distance two is the same as being non-adjacent (and thus the diameter of any non-trivial strongly regular graph is two). Let A be the adjacency matrix of a strongly regular graph and $\overline{A}$ the adjacency matrix of the complement. Then the matrices I, A, and $\overline{A}$ partition the set of pairs $\{u, v\}$, where u and v are vertices of X according to whether the distance between x and y is 0, 1, or 2. We can express this as:

$$J = I + A + \overline{A}. \tag{2.1}$$

We regard this as a combinatorial statement: we have partitioned the edge set of the complete graph according to distance (0, 1 or 2) in X. It is not hard to show that the complement of a strongly regular graph is also strongly regular, and hence the partition obtained would be the same had we started with $\overline{X}$. We think of the partition as "having" two graphs, and there is no particular reason to favour one over the other. We could also refer to these two graphs as being the two classes of the partition.

Now we develop an algebraic viewpoint.

The matrices I, A, and $\overline{A}$ are all 01-matrices, and together with (2.1), this implies that they are linearly independent. The number of walks of length two from u to v depends only whether u and v are equal, adjacent, or distinct and non-adjacent. Furthermore, this number is the (u, v)-entry of A^2. Putting these observations together, we get:

$$A^2 = kI + aA + c\overline{A} = kI + aA + c(J - I - A). \tag{2.2}$$

This is not a trivial observation. The fact that A^2 is a linear combination of A, I and J is equivalent to the definition of a strongly regular graph. By a straightforward calculation we have:

$$A\overline{A} = (k - a - a)A + (k - c)\overline{A} = \overline{A}A. \tag{2.3}$$

The immediate consequence of this is that the matrices $\mathcal{A} = \{I, A, \overline{A}\}$ form a basis for the algebra they generate. Another way of saying this is that the algebra generated by $\mathcal{A}$ is exactly the real vector space spanned by $\mathcal{A}$. Furthermore,

this algebra is commutative. We will use the notation $\mathbb{R}[\mathcal{A}]$ to mean the real vector space spanned by $\mathcal{A}$; we warn the reader that

There is another consequence of (2.3).

2.2.2 Lemma. *A non-trivial strongly regular graph has exactly three distinct eigenvalues.*

Proof. We know the eigenvector $\mathbf{1}$ corresponds to the greatest eigenvalue, k. Using (2.2), it follows that any eigenvector orthogonal to $\mathbf{1}$ has an eigenvalue λ that satisfies

$$\lambda^2 = k\lambda + a\lambda + c(1-\lambda).$$

Thus there are at most three distinct eigenvalues. It is not hard to show that if there are less than three eigenvalues, the strongly regular graph is trivial. □

In fact, the converse is also true: any regular graph with exactly three distinct eigenvalues is strongly regular. We can say more. From (2.2), any eigenspace of $\overline{A}$ is a sum of eigenspaces of A. As $\overline{A}$ is also a non-trivial strongly regular graph, it has exactly three distinct eigenvalues and hence every eigenspace of $\overline{A}$ is an eigenspace of A. One way to express this is to say that I, A, and $\overline{A}$ are simultaneously diagonalizable.

For each eigenvalue λ of A, let U_j be a matrix whose columns form an orthogonal basis for the j-th eigenspace of A. Then the columns of U_j form an orthogonal basis for an eigenspace of $\overline{A}$ too. Let $E_j = \frac{1}{n} U_j U_j^T$. Let $E = E_0 + E_1 + E_2$ and let U be the matrix whose columns are the columns of U_1, U_2 and U_3. Then U has full rank, and furthermore, $EU = U$. Multiplying by U^{-1}, we see that $E = I$. It follows that

$$\begin{aligned} I &= E_0 + E_1 + E_2, \\ A &= kE_0 + \theta E_1 + \tau E_2, \\ \overline{A} &= (n-k-1)E_0 + (-1-\theta)E_1 + (-1-\tau)E_2. \end{aligned} \tag{2.4}$$

The second and third equations can be obtained by multiplying the first by A and $\overline{A} = J - I - A$, respectively. One of the consequences of this is that the matrices $\{E_0, E_1, E_2\}$ form a second basis for $\mathbb{R}[\mathcal{A}]$.

Denote the multiplicities of θ and τ as eigenvalues of A by m_θ and m_τ; these are also multiplicities of eigenspaces of $\overline{A}$. We can determine m_θ and m_τ by solving the following pair of equations (the first says that, counting multiplicities, A has n eigenvalues, and the second says that the trace of A is zero).

$$\begin{aligned} m_\theta + m_\tau + 1 &= n \\ \theta m_\theta + \tau m_\tau + k &= 0 \end{aligned} \tag{2.5}$$

We can summarize this information with a table of the eigenvalues. Each row corresponds to an eigenspace of the algebra.

mult.	I	A	$\overline{A}$
1	1	k	$n-k-1$
$\frac{(n-1)\tau+k}{\tau-\theta}$	1	θ	$-1-\theta$
$\frac{(n-1)\theta+k}{\theta-\tau}$	1	τ	$-1-\tau$

We point out one use for this table. If we have some matrix B which is a linear combination of I, A and $\overline{A}$, then we can trivially compute its eigenvalues. In fact, we know its eigenspaces, they are exactly the three eigenspaces we found above. So we can just multiply B by each of these three eigenspaces (one eigenvector each will suffice) to determine its eigenvalues. But it is even simpler than that.

2.2.3 Lemma. *Let A be a strongly regular graph on n vertices, of valency k, and remaining eigenvalues θ and τ. Let*

$$B = b_0 I + b_1 A + b_2 \overline{A}.$$

Then the eigenvalues of B are the entries of

$$\begin{pmatrix} 1 & k & n-k-1 \\ 1 & \theta & -1-\theta \\ 1 & \tau & -1-\tau \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}. \qquad \square$$

This is an extremely important observation. We will see that this is a special case of Lemma 2.5.4, but in fact for our purposes, we will only ever need to apply it for the algebra generated by a strongly regular graph, so Lemma 2.2.3 is actually sufficient for our purposes. To whet the reader's appetite, we will say that the computations required to prove the characterization of the Erdős-Ko-Rado theorem are essentially Lemma 2.2.3.

2.3 Distance-Regular Graphs

Distance-regular graphs are a natural generalization of strongly regular graphs to diameter greater than two. They are also metric association schemes. Motivated by the comments after Definition 2.2.1, we can define distance-regular graphs as follows.

2.3.1 Definition. A *distance-regular graph* is a graph such that the number of vertices at distance i from u and distance j from v depends only on the distance between u and v.

Note that if we take u and v at distance 0, i.e., $u = v$, then we see that distance-regular graphs are regular. If we add the condition that the diameter

of the graph is equal to 2, then we recover the definition of a strongly regular graph, in perhaps a more compact but less intuitive form.

Given a distance-regular graph X of diameter d, define A_i to be the matrix with rows and columns indexed by vertices, with (u, v)-entry equal to 1 when u and v are at distance i and 0 otherwise. Clearly A_i is non-zero if an only if $0 \leq i \leq d$, $A_0 = I$ and A_1 is the adjacency matrix of X. Let $\mathcal{A} = \{A_0, \ldots, A_d\}$ (these are sometimes called the *distance matrices* of X). Again, since the matrices of $\mathcal{A}$ are 01 and

$$J = A_0 + A_1 + \cdots + A_d, \tag{2.6}$$

we see that $A_0, \ldots, A_d$ are linearly independent. The (u, v)-entry of $A_i A_j$ is exactly the number of vertices that are at distance i from u and distance j from v. It follows directly from the definition of distance-regular that $A_i A_j$ is symmetric. More precisely, it follows that for each A_k, $(A_i A_j)$ is constant on the non-zero entries of A_k. Thus $A_i A_j$ is a linear combination of $A_0, \ldots, A_d$. The immediate consequence of this is that the matrices $A_0, \ldots, A_d$ form a basis for the algebra generated by $\mathcal{A}$.

We can summarize our observations thus far as

2.3.2 Lemma. *Let X be a distance-regular graph and $\mathcal{A} = \{A_0, \ldots, A_d\}$ the distance matrices of X. Then*

(a) $A_0 = I$

(b) $\sum_i A_i = J$

(c) A_i *is symmetric for each* i

(d) $A_i A_j = A_j A_i \in \mathbb{R}[\mathcal{A}]$ *for all* i, j

We will see that this is the definition of an association scheme. Distance-regular graphs are a little more specialized, as they also have the property that A_i is a polynomial of degree i in A_1 (see [29, p.197-198]). We can use this observation to show that the adjacency matrix A of a distance regular graph has at most most $d + 1$ eigenvalues, and that every eigenspace of every matrix in $\mathbb{R}[\mathcal{A}]$ is a sum of eigenspaces of A. It is also true [29, p.29] that A has at least $d + 1$ distinct eigenvalues, and this leads us straight to the matrix of eigenvalues of a distance-regular graph. But rather than developing this further, it is perhaps more appropriate at this point to formally introduce association schemes.

2.4 Definitions and Example

We saw that a strongly regular graph defined an algebra, and explored some of the relationships between the combinatorics and the algebra. These relationships generalized quite naturally to distance-regular graphs.

It is natural to ask whether the algebraic structure we have observed in the previous two sections occurs in more general contexts. The answer is yes: that is exactly what an association scheme is. We can think of it as a structure

having the algebraic properties we have met so far, but perhaps not because of any properties relating to distance in a graph.

Given the double nature of association schemes, we will have two definitions: one combinatorial and one algebraic. We start with the combinatorics.

2.4.1 Definition. An *association scheme* with d classes on n vertices is a set of graphs $X_1, \ldots, X_d$ on a fixed set of n vertices such that the following properties hold.

(a) The edge sets of the graphs partition the edge set of the complete graph on the same set of vertices.

(b) Given any two vertices x, y, the number of vertices z such that $xz \in E(X_i)$ and $yz \in E(X_j)$ depends only on i, j and k, where $xy \in E(X_k)$. □

We introduce some terminology. We will use the term "graph in an association scheme" to mean one of the X_i. Vertices x and y are i-related (or i-adjacent) if xy is an edge of X_i. We will also say that xy is an i-edge. Note that every pair of vertices is an i-edge for exactly one i.

It is not too hard to see that strongly regular graphs satisfy this definition with $d = 2$, and that more generally, so does the set of distance matrices of a distance-regular graph of diameter d. It may not be obvious that there are other examples, so we will provide one now.

2.4.2 Example. Construct a graph $\mathcal{P}(3^3)$ as follows. The vertices are the partitions of $V = \{1, \ldots, 9\}$ into three cells of size three. The meet of two partitions x and y, denoted by $x \wedge y$, is the set of non-empty intersections of one cell from each partition. So the meet of two partitions is another partition that is a common refinement of them. Two vertices are adjacent if their meet is the discrete partition.

In general, the meet of two partitions may have size $3, 9, 7, 6, 5$. If $|x \wedge y| = 3$ then $x = y$. If $|x \wedge y| = 9$, then xy is an edge of $\mathcal{P}(3^3)$. Let us define four graphs $X_1, \ldots, X_4$ on the same vertex set as $\mathcal{P}(3^3)$ such that $xy \in E(X_i)$ when $|x \wedge y| = 9, 7, 6, 5$ for $i = 1, 2, 3, 4$, respectively. Thus $X_1 = \mathcal{P}(3^3)$. Then these graphs form an association scheme with four classes on 280 vertices. There are several ways of proving this (including citing [41] or waiting for Lemma 2.4.4). For the moment, we will content ourselves with an intuitive approach that assumes no knowledge of association schemes.

2.4.3 Lemma. *The graphs $X_1 = \mathcal{P}(3^3), X_2, X_3, X_4$ form an association scheme with four classes.*

Proof. Let x and y be two vertices, and let $z = x \wedge y$ (so z is a partition but not a vertex). Let x' and y' be two other vertices with $z' = x' \wedge y'$, such that $|z| = |z'|$. Thus xy and $x'y'$ are both edges in the same graph. An automorphism of an association scheme is a permutation of the vertex set that

acts as an automorphism on each graph in the scheme. It will suffice to show that there is an automorphism of the scheme that maps x to x' and y to y'. This will force the condition of Definition 2.4.1 to be satisfied. Any permutation of σ of V induces a permutation of the vertices that is an automorphism of each graph. It is not immediately apparent that there are no other automorphisms of the scheme, but we can ignore that for the moment.

We will choose a particular labelling of the cells of x, y, x', y, with $1, 2, 3$. This will induce a labelling of the cells of z and z', by setting $z_{ij} = x_i \cap y_j$ and $z'_{ij} = x'_i \cap y'_j$. Then any permutation of V which maps the set z_{ij} to the set z'_{ij} will induce a permutation of the vertices mapping x to y and x' to y'. For this to work it is both necessary and sufficient to have $|z_{ij}| = |z'_{ij}|$ for all i, j.

So in fact it is sufficient to show how to label the cells of x and y so that $|z_{ij}|$ depends only on (i, j) (and on the type of edge between x and y of course). Then we can apply this "canonical labelling" to x' and y' as well. We now describe the canonical labellings for each edge-type.

If $|z| = 9$, then z and z' are both discrete. Label the cells of x, y in any order. It is easy to see that $|z_{ij}| = 1$ for all i, j.

If $|z| = 7$, then there is exactly one cell of x that intersects each cell of y: label this cell 1. Likewise, there is one cell of y that intersects each cell of x: label it 1. Label one of the remaining cells of x as 2. There is exactly one cell of y that intersects x_2 in two elements: label it 2. Label the remaining cells 3. Then

$$\begin{aligned} |z_{1j}| = |z_{i1}| &= 1 \qquad i, j = 1, 2, 3 \\ |z_{22}| = |z_{33}| &= 2 \\ |z_{23}| = |z_{32}| &= 0 \end{aligned}$$

If $|z| = 6$, then label any cell of x as 1. There is one cell of y that intersects x_1 in two elements: label it 1. There is one cell of x that intersects y_1 in one element: label it two. There is one cell of y that intersects x_2 in two elements: label it 2. Label the remaining cells of x and y as 3. Then

$$\begin{aligned} |z_{11}| = |z_{22}| = |z_{33}| &= 2 \\ |z_{21}| = |z_{32}| = |z_{13}| &= 1 \\ |z_{12}| = |z_{23}| = |z_{31}| &= 0 \end{aligned}$$

If $|z| = 5$, then there is one cell that is common to x and y: label it 1 in x and in y. Label one of the remaining cells of x as 2. There is exactly one cell of y that intersects x_2 in two elements: label it 2. Label the remaining cells 3. Then

$$\begin{aligned} |z_{11}| &= 3 \\ |z_{22}| = |z_{33}| &= 2 \\ |z_{23}| = |z_{32}| &= 1 \\ |z_{1j}| = |z_{i1}| &= 1 \qquad i, j = 2, 3 \end{aligned}$$

In each case, apply the same labelling to x' and y' and let σ be any permutation of V that maps z_{ij} to z'_{ij}. The result then follows. □

The above proof relies on the fact that for each value of $|x \wedge y|$ there is a unique "shape" of $x \wedge y$. We can always reorder the cells of each vertex so that the structure of $x \wedge y$ depends only on $|x \wedge y|$.

It may help to consider the following representation of an edge xy (we are thinking of an edge as an ordered pair of vertices for the moment, and every pair of vertices is an edge of some kind). Label the cells of x and y with $1, 2, 3$ in some order. Now label each element of V with two integers: the cells of x and y that contain it. Thus an edge may be represented by an ordered list of nine ordered pairs: the i-th ordered pair gives the cells of x and y, respectively, in which i is to be found. Call this a grid-labelling.

We may think of this as a deck of nine cards, where each card has an ordered pair on the face side, and the corresponding element of $\{1, \ldots, 9\}$ written on the back. Note that $|x \wedge y|$ is exactly the number of distinct face sides. Arrange the cards face up in some convenient way; for example, so that the first coordinates of each ordered pair correspond to columns in a 3×3 grid. The particular shape is fixed but arbitrary and there may be several ways of accomplishing it for a given edge. Call such an arrangement (with the back sides hidden) a grid-layout. Now if two edges have the same grid-layout, there is a permutation of V that maps one edge to the other: simply flip over the cards to give the mapping.

If $|x \wedge y| = 7$, then we can always choose a grid-layout that looks like the following

$$\begin{array}{ccc} 21 & 11 & 31 \\ 22 & 12 & 33 \\ 22 & 13 & 33 \end{array}$$

This means that all edges with $|x \wedge y| = 7$ are equivalent under the automorphism group of the scheme.

If $|x \wedge y| = 6$, we can always choose the following grid-layout.

$$\begin{array}{ccc} 11 & 21 & 33 \\ 11 & 22 & 33 \\ 13 & 22 & 32 \end{array}$$

If $|x \wedge y| = 5$, we can choose the following grid-layout.

$$\begin{array}{ccc} 11 & 23 & 33 \\ 11 & 22 & 33 \\ 11 & 22 & 32 \end{array}$$

There is an obvious generalization of $\mathcal{P}(3^3)$ to $\mathcal{P}(g^g)$, where $V = \{1, \ldots, g^2\}$. We could again define a family of graphs according to the size of the meet. Note that even for $g = 4$ the above proof fails to show that this family forms an association scheme. There are vertices x, y, x', y' such that $x \wedge y$ and $x' \wedge y'$ both consist of eight cells of size two, yet there is no permutation of V that maps x

to x' and y to y'. Using the above grid representation of an edge, an example would be two edges with grid-layouts as follows:

11	21	33	43		11	21	33	43
11	21	33	43		11	21	33	43
12	22	34	44		14	22	32	44
12	22	34	44		14	22	32	44

To make the example more concrete, label the 16 positions with $\{1, \dots 16\}$ in any way for the left-hand grid, and write down the two vertices so determined. There is no permutation of $\{1, \dots 16\}$ that will result in two vertices with the grid-layout on the right.

Note that it does not follow immediately that for $g = 4$ we do not have an association scheme. It is possible, in principle, that there are automorphisms of the scheme that do not arise from permutations of the V. More pragmatically, it seems highly unlikely that $\mathcal{P}(g^g)$ gives rise to an association scheme for $g > 3$, but we have not yet checked this.

We make one further comment on the above proof. It is perhaps slightly misleading, in that the reader may have the impression that the condition of Definition 2.4.1 holds *because* there are automorphisms that force it to. This is not the case. What we have shown is something stronger than being an association scheme. There are association schemes that do not have automorphisms that map a given i-edge to any other i-edge. The distinction is analogous to the difference between a distance-transitive graph and a distance-regular graph.

Returning to association schemes in general, we mention another way to show that a set of graphs is an association scheme. A group Γ of automorphisms acting on X is *generously transitive* if given any distinct vertices of X, there exists an automorphism in Γ that swaps them. The following result is well-known; a proof appears in [8, p.63].

2.4.4 Lemma. *Let $X_1, \dots, X_d$ be a set of graphs on a common vertex set V, whose edge set partitions the complete graph on V. If there is a group of automorphisms that is generously transitive on all X_i, then $X_1, \dots, X_d$ form an association scheme.* □

We could have used this to show that $\mathcal{P}(3^3)$ is an association scheme, but it is perhaps less intuitively obvious why this result is true.

We now introduce an equivalent algebraic definition of an association scheme.

2.4.5 Definition. An *association scheme* with d classes on n vertices is a set $\mathcal{A} = A_0, \dots, A_d$ of 01-matrices of size $n \times n$ such that the following properties hold.

(a) $A_0 = I$

(b) $\sum_i A_i = J$

(c) A_i is symmetric for each i

(d) $A_iA_j = A_jA_i \in \mathbb{R}[\mathcal{A}]$ for all i, j □

The matrices A_i are the adjacency matrices of the graphs X_i of Definition 2.4.1. Note that an association scheme consists of d graphs and $d+1$ matrices. So "graph" and "class" mean the same thing.

All 01-matrices in $\mathbb{R}[\mathcal{A}]$ are of the form $\sum_{i\in R} A_i$ for some $R \subseteq \{0,\dots,d\}$. Thus the only graphs whose adjacency matrices appear in $\mathbb{R}[\mathcal{A}]$ are unions (on the same set of vertices) of some of the X_i. By "graph in an association scheme" we continue to mean only the X_i. We will sometimes refer to $\sum_{i\in R} A_i$ as a union of classes.

2.5 Matrix of Eigenvalues

One of the main reasons for considering association schemes at all is because of the tight interplay between their combinatorial structure and their algebraic structure. As we saw with strongly regular graphs, there is more algebra that follows as a consequence of Definition 2.4.5.

The *Schur product* of two matrices is defined to be

$$(A \circ B)_{ij} = A_{ij}B_{ij}.$$

In writing down (2.6), we have tacitly observed that

$$A_i \circ A_j = \begin{cases} A_i & i = j, \\ 0 & i \neq j. \end{cases}$$

So the matrices $\{A_0,\dots,A_d\}$ are idempotent and orthogonal with respect to Schur multiplication; thus the A_i are referred to as the *principal Schur idempotents* of the scheme. We used these properties in Section 2.3 to show that $A_0,\dots,A_d$ are linearly independent. It also follows that $\mathbb{R}[\mathcal{A}]$ is closed under Schur multiplication (as well as matrix multiplication). This means that the vector space $\mathbb{R}[\mathcal{A}]$ is an algebra with respect to matrix multiplication and an algebra with respect to Schur multiplication.

We can regard Definition 2.4.5 as a statement of properties of a particular basis of $\mathbb{R}[\mathcal{A}]$. There is another important basis consisting of matrices that are idempotent and orthogonal with respect to matrix multiplication.

2.5.1 Theorem. *Let $\mathcal{A} = \{A_0,\dots,A_d\}$ be an association scheme. Then there exist matrices $\mathcal{E} = \{E_0,\dots,E_d\} \subseteq \mathbb{R}[\mathcal{A}]$ that are idempotent and pairwise orthogonal with respect to matrix multiplication such that the following properties hold.*

(a) $E_0 = \frac{1}{n}J$

(b) $\sum_j E_j = I$

(c) *$\mathcal{E}$ forms a basis for $\mathbb{R}[\mathcal{A}]$*

(d) *There exist numbers $p_i(j)$ such that $A_iE_j = p_i(j)E_j$* □

We refer to the matrices E_j as the *principal idempotents* of the scheme.

It is useful to compare this result with Definition 2.4.5. We could have said that the matrices A_i are pairwise orthogonal and idempotent with respect to Schur multiplication, but it was easier and more intuitive to give the equivalent condition that they are 01-matrices.

So Theorem 2.5.1 is dual to Definition 2.4.5, aside from the final condition. But this last condition is straightforward given the rest of the theorem: we are just expressing one basis in terms of another. Note that we cannot prove this theorem using duality, we are merely observing a correspondence: the proof of Theorem 2.5.1 is not trivial.

Moreover, since the matrices $\mathcal{E}$ form a basis, we can express A_i in terms of them as:

$$A_i = \sum_r p_i(r)E_r. \tag{2.7}$$

Matrix-multiplying both sides by E_j gives the result. There is of course a dual condition. The matrices $\mathcal{A}$ are a basis too, so we can express E_j in terms of them as:

$$E_j = \frac{1}{n}\sum_r q_j(r)A_r. \tag{2.8}$$

Schur-multiplying both sides by A_i gives that

$$E_j \circ A_i = \frac{1}{n}q_j(i)A_i,$$

which we can regard as an addendum to Definition 2.4.5.

The columns of each E_j span a common eigenspace for each A_i, so the $p_i(j)$ give the eigenvalues of each graph in the scheme. Accordingly we call these the *eigenspaces of the scheme*. Any eigenspace of any matrix in $\mathbb{R}[\mathcal{A}]$ is a sum of eigenspaces of the scheme. Recall that in Section 2.2 we derived a table of eigenvalues of a strongly regular graph. These eigenvalues are exactly the $p_i(j)$ of the association scheme of the strongly regular graph.

The matrix P, where $(P)_{ji} = p_i(j)$ is called the *matrix of eigenvalues* of the scheme. By analogy, the matrix Q, where $(Q)_{ji} = q_i(j)$ is called the *matrix of dual eigenvalues* of the scheme. The initial row of P, that is the values $p_i(0)$, are known as the valencies of the scheme. We denote them by v_i. They are in fact the valencies of the graphs in the scheme. The $q_i(0)$ are the multiplicities, denoted my m_j, and they are exactly the multiplicities of the eigenspaces of the scheme.

There are two ways of looking at the matrix of eigenvalues. It gives the eigenvalues of the matrices in the basis $\mathcal{A}$ (and hence every matrix in $\mathbb{R}[\mathcal{A}]$), but it is also a change of basis matrix. Thus given the matrices $A_0, \ldots, A_d$ and P, we can compute the matrices $E_0, \ldots, E_d$. In other words, given the eigenvalues, we can write down the eigenspaces. This is an important observation, and it is worth emphasizing.

2.5.2 Lemma. *Let $\mathcal{A} = \{A_0, \ldots, A_d\}$ be an association scheme. For some fixed but arbitrary i, let $\lambda_1, \ldots, \lambda_l$ be the distinct values of $p_i(j)$. Then the t-th eigenspace of A_i is spanned by the columns of*

$$\sum_{j\,:\,p_i(j)=\lambda_t} \sum_{r=0}^{d} q_j(r) A_r.$$

Proof. The actual proof of this statement is contained in a complete proof of Theorem 2.5.1 (we omit this: see e.g. [29] for details). We only show how this follows trivially given Theorem 2.5.1.

Using Definition 2.4.5, or indeed the final condition of Theorem 2.5.1, we see that the column space of each E_j is an invariant subspace of each A_i. Using (2.8) we see how to compute the E_j's from the A_i's. Each eigenspace is a sum of invariant subspaces. The only thing left to show is that there are no other eigenspaces unaccounted for. In other words, we need to show that the sum of the ranks of the E_j's is n.

If we let m_j be the rank of E_j, then since E_j is symmetric is can be written as $U_j U_j^T$, where U_j is an $n \times m_j$ full column rank matrix. Since the E_j's sum to I, the m_j's sum to at least n. Since the E_j's are pairwise orthogonal, the column spaces of distinct U_j's are orthogonal, and the result follows. □

This is not always the best way to find the eigenspaces of a scheme, if only because the E_j are "big"matrices. Note that in the above proof, the columns of the U_j form a basis for the j-th eigenspace. Nonetheless, the reader should notice that in order to find a complete set of eigenvectors for a $v \times v$ matrix (actually, for a whole algebra of them ...) we only need to compute a $(d+1) \times (d+1)$ matrix of eigenvalues. This is already much simpler than what we would expect from an arbitrary $v \times v$ matrix.

The matrices P and Q are related; in fact, they determine each other. To see this, substitute (2.8) into (2.7) to give

$$A_i = \frac{1}{n} \sum_j \left(p_i(j) \sum_k q_j(k) A_k \right).$$

Unrolling this, we find that

$$PQ = nI. \tag{2.9}$$

Thus the dual eigenvalues can be computed from the eigenvalues.

There is another standard relationship between P and Q. A proof appears in Godsil [8, Section 2.2].

2.5.3 Lemma. *Let $\Delta_m = \mathrm{diag}(e_0^T Q)$ be the diagonal matrix of multiplicities, and let $\Delta_v = \mathrm{diag}(e_0^T P)$ be the diagonal matrix of valencies. Then*

$$P^T \Delta_m = \Delta_v Q.$$ □

Returning to the graph $\mathcal{P}(3^3)$ of Example 2.4.2, we give the matrix of eigenvalues, computed by Mathon and Rosa in [41]. The columns (i.e., the classes of the scheme) are in order of the meet of two vertices being $3, 9, 7, 6, 5$.

$$P = \begin{pmatrix} 1 & 36 & 162 & 54 & 27 \\ 1 & -12 & -6 & 6 & 11 \\ 1 & 8 & -6 & -9 & 6 \\ 1 & 2 & -6 & 6 & -3 \\ 1 & -4 & 12 & -6 & -3 \end{pmatrix}$$

The rows of P are indexed by $A_0, \ldots, A_d$ and the columns by $E_0, \ldots, E_d$, so it is natural to number them starting from 0 and not 1. Thus our graph, $\mathcal{P}(3^3) = X_1$, corresponds to the column numbered 1, and its least eigenvalue occurs on the eigenspace numbered 1.

We shall defer for the moment an explanation of how they were able to compute this. It is possible, though highly unlikely, that they simply wrote down the four 280×280 matrices and directly evaluated and then factored their characteristic polynomials.

One thing we might notice is that the third column contains only three distinct values. This means that the graph X_2, where vertices are adjacent if their meet has size 7, is a strongly regular graph. This is completely obvious given P, and is in fact the immediate intent of Mathon and Rosa's note: this strongly regular graph was not known at the time. (Their note [41] is a tangent to a larger project).

For the moment we do not know the eigenspaces, that is to say, we do not know the matrices E_j. We can compute them using P. In practice, we first need Q, which we obtain easily using (2.9):

$$Q = \begin{pmatrix} 1 & 27 & 48 & 120 & 84 \\ 1 & -9 & \frac{32}{3} & \frac{20}{3} & -\frac{28}{3} \\ 1 & -1 & -\frac{16}{9} & -\frac{40}{9} & -\frac{56}{9} \\ 1 & 3 & -8 & -\frac{40}{3} & -\frac{28}{3} \\ 1 & 11 & \frac{32}{3} & -\frac{40}{3} & -\frac{28}{3} \end{pmatrix}$$

Now we can compute the matrices E_j using (2.7). In particular, using Lemma 2.5.2 we can compute E_1, whose columns span the eigenspace corresponding to the least eigenvalue of $\mathcal{P}(3^3)$:

$$E_1 = \frac{1}{280}\left(27A_0 - 9A_1 - 1A_2 + 3A_3 + 11A_4\right)$$

We will discover a nicer way to obtain this eigenspace in Lemma 4.5.1, but in principle, we already know all of the eigenspaces of $\mathcal{P}(3^3)$.

We return to association schemes in general. We turn to the multiplicities of the eigenspaces.

Denote by n_i the row-sum of A_i; this is just the valency of the graph it represents. Accordingly, n_i is the greatest eigenvalue of A_i, and the values n_i occupy the first row of P. Denote by m_j the rank of E_j, which is equal to the trace of E_j. For a strongly regular graph, we were able to compute the multiplicities as a function of the eigenvalues. What we were doing in (2.5) is, for each of I, A, and $\overline{A}$, writing down the sum of the eigenvalues (multiplicities included) and setting this equal to the trace. This may have been obscured somewhat by the fact that we already knew that m_0, the multiplicity of the eigenspace belonging to the constant vectors, is equal to 1. We take this approach now in full generality.

Let m_i be the multiplicity of the i-th eigenspace. Then the sum of the eigenvalues (multiplicities included) of A_i is $\sum_j m_j p_i(j)$. The trace of A_i is either n or 0, according to whether $i = 0$ or $i > 0$. Putting these observations together, we have

$$\begin{pmatrix} m_0 & m_1 & \cdots & m_d \end{pmatrix} P = \begin{pmatrix} n & 0 & \cdots & 0 \end{pmatrix}.$$

Since P is invertible, we can compute the multiplicities. More precisely, using (2.9), the multiplicities are exactly the zero-th row of Q. Of course, one can rewrite these in terms of P, since Q is determined by P.

In a practical sense, the multiplicities can function as constraints. Given a square matrix P that is alleged to be the matrix of eigenvalues of some association scheme, we can compute the alleged multiplicities. If these are not positive integers, then our P is fraudulent. This is not a trivial observation. We have in fact outlined the proof that a Moore graph of diameter two has valency 2, 3, 7 or 57.

Returning briefly to the graph $\mathcal{P}(3^3)$ of Example 2.4.2, we find that we already know the multiplicities, since we computed Q above. So for instance, $q_2(0) = Q_{02} = 27$ tells us that the eigenspace of $\mathcal{P}(3^3)$ corresponding to the least eigenvalue has dimension 27. Recall that we suggested that the columns of E_1 were perhaps not the best spanning set for this eigenspace. This is partly the reason why: it is, to say the least, computationally inefficient to carry around a spanning set of 280 vectors for a space of dimension 27. In Section 4.5, we will identify a better spanning set, one that is not only much smaller, but more importantly gives us other structural information as well. This is not an argument against association schemes at all: the matrix of eigenvalues will play a key role in deriving that spanning set as well.

Back once again to our general program. We have already met one important application of the matrix of eigenvalues, as Lemma 2.2.3. We now generalize this to association classes. Is is a straightforward consequence of our observations up to now; we record it as a lemma for future reference.

2.5.4 Lemma. *Let $A_0, \ldots, A_d$ be an association scheme and P its matrix of eigenvalues. Let $B = \sum_i b_i A_i$, and let $b = (b_1, \ldots, b_n)^T$.*

Then the vector of eigenvalues of B is given by Pb, and the corresponding eigenspaces of B are the column spaces of the matrices E_j. □

The reader should interpret this as meaning that the eigenvalues of any matrix in $\mathbb{R}[\mathcal{A}]$ are known quantities. Of course, we can multiply B by an eigenvector for each eigenspace of the scheme, but there is no need to.

We have already mentioned that we will use this on the association scheme of the line graph of a complete graph in Lemma 6.11.3. We will even use Lemma 2.5.4 to find the eigenvalues of a matrix in the association scheme with one class consisting of $A_0 = I, A_1 = J - I$. This will not be a trivial usage: we will use it to establish the characterization in the Erdős-Ko-Rado theorem.

2.6 Intersection Numbers

Definition 2.4.1 and Definition 2.4.5 are equivalent, but they have a different emphasis. We have seen that the matrix of eigenvalues follows naturally from the latter, but the former suggests another set of parameters, the *intersection numbers* $p_{ij}(k)$. We start with an algebraic approach to these.

Consider the matrix product A_iA_j. According to Definition 2.4.5, it is a linear combination of the matrices A_k. So there exist scalars $p_{ij}(k)$ such that

$$A_iA_j = p_{ij}(0)A_0 + p_{ij}(1)A_1 + \cdots + p_{ij}(d)A_d. \tag{2.10}$$

Thus the (x, y)-entry only depends on the value of k such that A_k has its (x, y)-entry equal to 1. There is such a value of k, and it is unique. If we notice that any symmetric 01-matrix with zero diagonal is the adjacency matrix of a graph, then we see that the numbers $p_{ij}(k)$ are exactly the quantities given in Definition 2.4.1. That is, given any two vertices x and y that are adjacent in the i-th graph, there are exactly $p_{ij}(k)$ vertices z such that x and z are adjacent in the j-th graph and y and z are adjacent in the k-th graph.

This observation is more important than it might seem. Equation (2.10) gives not only that $p_{ij}(k)$ exist, but that they must be non-negative integers. This says much more than just that $\mathbb{R}[\mathcal{A}]$ is closed under multiplication, it identifies another way in which the basis $A_0, \ldots, A_d$ is special. It is possible to express the numbers $p_{ij}(k)$ in terms of the entries of P. This gives another set of constraints on the eigenvalues of a scheme. It is also possible to express the entries of P in terms of the intersection numbers (see [8, p.45–46]).

Recalling the graph $\mathcal{P}(3^3)$, we see that computing the intersection numbers is straightforward. We may compute $p_{ij}(k)$ by picking any particular k-edge xy, and counting the number of vertices that are i-adjacent to x and j-adjacent to y. This is an exercise in counting. There are of course some redundancies in the set of all $p_{ij}(k)$ which we may either use to reduce our workload or double-check our results. The values are given in [41].

This is doubtless the way in which Mathon and Rosa computed the matrix of eigenvalues, i.e., as a function of the intersection numbers. Moreover, the fact

that the value of $p_{ij}(k)$ is independent of the particular choice of k-edge gives another proof that the association scheme is, in fact, an association scheme. Note that we do not in principle need any automorphisms to do this, although showing that any k-edge can be mapped to any other k-edge as we did following Example 2.4.2 is possibly the easiest approach.

Back to association schemes in general, whatever we do with the A_i's, we can do with the E_j's. So it should come as no surprise that there exist scalars $q_{ij}(k)$ such that

$$E_i \circ E_j = q_{ij}(0)E_0 + q_{ij}(1)E_1 + \cdots + q_{ij}(d)E_d. \tag{2.11}$$

These are known as the *Krein parameters* of the scheme. It is possible to show that these must be non-negative.

The intersection numbers can be quite useful, even if one is interested only in one of the graphs in an association scheme. Let $\omega(X)$ be the clique number of X, and $\alpha(X)$ be the independence number of X. If X_r is the r-th graph in an association scheme (so that A_r is the adjacency matrix of X_r), then we have

$$\omega(X_r) \leq 2 + p_{rr}(r) \tag{2.12}$$

$$\alpha(X_r) \leq \max_k \left\{ \sum_{i \neq r} \sum_{j \neq r} p_{ij}(k) : k \neq r \right\} \tag{2.13}$$

The first says that any two vertices in a clique of X_r must be r-adjacent and both be r-adjacent to all of the other vertices in the clique. The second says that any two vertices in an independent set of X_r must not be r-adjacent and must both not be r-adjacent to all of the other vertices in the independent set.

We can apply these ideas to our graph $\mathcal{P}(3^3)$. We obtain first of all:

$$\omega(\mathcal{P}(3^3)) = \omega(X_1) \leq 2 + p_{11}(1) = 4.$$

As we will see in Chapter 4, this bound is tight! It is perhaps surprising that such a naive approach might actually work.

For an independent set S, every pair of vertices in S is either 2-adjacent, 3-adjacent, or 4-adjacent. There are three possibilities, and using the intersection numbers in [41], we get:

$$\alpha(\mathcal{P}(3^3)) = \alpha(X_1) \leq \max\{210, 210, 218\} = 218.$$

As we will see in Chapter 4, this bound is ludicrous! It is perhaps reassuring that such a naive approach might actually fail.

There are some improvements we could make to this bound on independent sets, but we will refrain for now (but see Lemma 4.5.2).

2.7 Metric and Cometric Schemes

We have avoided any mention of particular orderings of the matrices A_i, or the matrices E_j. While it is true that most results are independent of the ordering, there are some cases where there is a natural ordering. These are the metric and cometric schemes.

An association scheme is metric (with respect to A_1) if each A_r can be written as a polynomial of degree r in A_1. This is equivalent to the graph A_1 being distance regular, with each A_r being the r-distance matrix.

The Johnson scheme, the Grassmann scheme, and the Hamming scheme are three classical schemes that we shall encounter in this thesis, and they are metric. We will introduce them more formally shortly.

We can define cometric association schemes in analogy with metric. A scheme is cometric (with respect to E_1) if E_r is a Schur-polynomial of degree r in E_1. By a *Schur polynomial* we mean a polynomial where instead of matrix multiplication we use Schur multiplication. Another way of describing this is for each r, there exists a polynomial w_r of degree r such that $w_r\left((E_1)_{ij}\right) = (E_r)_{ij}$.

The property of being cometric will have very practical consequences for us. It provides us with "nice" descriptions of the eigenspaces. Recall that the matrices $E_0, \ldots, E_d$ give the eigenspaces of the scheme. More precisely, the column space of each E_j corresponds to an eigenspace of the scheme. Assume we are dealing with a cometric scheme and we have a set of vectors that span the eigenspace corresponding to E_1 (this set of vectors may be the columns of E_1, or it may be some other convenient spanning set), then we can obtain all other eigenspaces by taking linear combinations of Schur products of these vectors.

It turns out that the Hamming, Johnson, and Grassmann schemes mentioned above are also cometric. In practice, this means that among other things, we can expect to have nice representations of the eigenspaces for these schemes.

We have the following characterizations of metric and cometric schemes, due to Delsarte [18, Thm. 5.6, 5.16]; see also [8, p. 58].

2.7.1 Lemma. *Let $A_0, \ldots, A_d$ be the Schur idempotents and $E_0, \ldots, E_d$ be the idempotents of an association scheme $\mathcal{A}$.*

Then $\mathcal{A}$ is metric with respect to A_1 if and only if the matrices $A_2, \ldots, A_d$ can be ordered such that $p_{1i}(i+1) \neq 0$ and $p_{1i}(j) = 0$ for all $j > i+1$.

Dually, $\mathcal{A}$ is cometric with respect to E_1 if and only if the matrices $E_2, \ldots, E_d$ can be ordered such that $q_{1i}(i+1) \neq 0$ and $q_{1i}(j) = 0$ for all $j > i+1$. □

Recall the graph $\mathcal{P}(3^3)$ from Example 2.4.2; we can easily compute the $p_{ij}(k)$ and $q_{ij}(k)$ from the matrix of eigenvalues (in fact, the intersection numbers are

explicitly given in [41]). We obtain

$$(p_{1,i}(j)) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 36 & 2 & 4 & 4 & 12 \\ 0 & 18 & 20 & 24 & 24 \\ 0 & 6 & 8 & 8 & 0 \\ 0 & 9 & 4 & 0 & 0 \end{pmatrix}$$

$$(q_{1,i}(j)) = \frac{1}{105} \begin{pmatrix} 0 & 105 & 0 & 0 & 0 \\ 2835 & 426 & 756 & 216 & 0 \\ 0 & 1344 & 539 & 224 & 560 \\ 0 & 960 & 560 & 1415 & 1400 \\ 0 & 0 & 980 & 980 & 875 \end{pmatrix}$$

We see that the association scheme it sits in is neither metric nor cometric with respect to any of the idempotents.

In fact, the graph $\mathcal{P}(3^3)$ has diameter two, meaning that were it to be distance regular it would be strongly regular, which it isn't. Recall that the graphs in this scheme correspond to the size of the meet of two partitions being 9, 7, 6 or 5, the first of these being our graph $\mathcal{P}(3^3)$. The second graph in this ordering is strongly regular, as a brief inspection of the matrix of eigenvalues will show, using Lemma 2.2.2.

Strongly Regular Graphs

Let X be a strongly regular graph. It defines an association scheme with two classes. This scheme is obviously metric with respect to the adjacency matrix of X, and it is almost as obviously metric with respect to the adjacency matrix of the complement of X. We can say a little more, and in fact, we only need that $\mathbb{R}[\mathcal{A}]$ is closed under matrix multiplication and Schur multiplication.

2.7.2 Lemma. *A strongly regular graph determines an association scheme that is metric with respect to each of its principal Schur idempotents and cometric with respect to each of its principal idempotents.*

Proof. Let the principal Schur idempotents be $I, A, \overline{A}$ and the principal idempotents be $\frac{1}{n}J, E, \overline{E}$. Since these both form bases, there are scalars a, b, c, a', b', c' such that

$$A^2 = aI + bA + c\overline{A},$$
$$E^{\circ 2} = aI + bE + c\overline{E}.$$

Rearranging, we have that $\overline{A}$ is a polynomial of degree 2 in A, and that $\overline{E}$ is a Schur polynomial of degree 2 in E. We did not actually specify "which" of the principal Schur idempotents was A, nor "which" of the principal idempotents was E, so we are done. □

Johnson Scheme

The Johnson scheme $J(v,k)$ is defined in terms of parameters v and k, with $v > k > 0$. The vertex set is the set of all k-sets of a fixed v-set. Vertices a and b are i-related if $|a \cap b| = k - i$. If $v \geq 2k$, then there are k classes.

The graph corresponding to being 1-related is better known as the Johnson graph with parameters v and k; sometimes it is denoted by $J(v,k)$ but we will not do this. The graph corresponding to being k-related is better known as the Kneser graph with parameters v and k, denoted $K_{v:k}$. It is not hard to show that in the Johnson graph $\mathrm{dist}(a,b) = k - |a \cap b|$. Thus being i-related and being at distance i in the Johnson graph are the same thing: the association scheme is the set of distance matrices of a Johnson graph. Thus the Johnson scheme is metric.

Let V be a fixed v-set. Let $W_{i,j}$ be the matrix whose rows are indexed by the i-subsets of V and whose columns are indexed by the j-subsets of V, and whose (a,b)-entry is 1 if the j-set b is contained in the i-set a, and 0 otherwise. Clearly, $W_{i,j}$ is identically zero if $i < j$. Consider the matrix $W_{k,2}$, and particularly, the column corresponding to the 2-set $\{1,2\}$. An entry in this column is 1 if and only if the corresponding k-set contains both 1 and 2. Another way of saying this is that this column is the Schur product of the columns corresponding to $\{1\}$ and $\{2\}$ in the matrix $W_{k,1}$. More generally, the column of $W_{k,i}$ corresponding to the i-set b is exactly the Schur product of all columns of $W_{k,1}$ corresponding to 1-subsets of b.

It turns out that the columns of $W_{k,i}$ span the sum of the first $i+1$ eigenspaces (recall that we are numbering them from 0). The columns of $W_{k,i}$ are obtained by taking all possible Schur products of k distinct columns of $W_{k,1}$. The j-th eigenspace is the orthogonal complement of $W_{k,j-1}$ in $W_{k,j}$. It follows that E_j, the j-th idempotent, is a Schur polynomial of degree j in E_1. Thus the Johnson scheme is cometric.

As a quick check on this, note that $W_{k,0}$ is the vector $\mathbf{1}$, which certainly spans the eigenspace corresponding to E_0. Also, $W_{k,k}$ is the identity matrix of order $n = \binom{v}{k}$, which certainly does span the sum of first $k+1$ eigenspaces (i.e., $\mathbb{R}^n$).

Grassmann Scheme

The Grassmann scheme $J_q(v,k)$ is defined in terms of parameters v and k, with $v > k > 0$, and a prime power q. The vertices are the k-dimensional subspaces of a fixed v-dimensional subspace over $GF(q)$. Two vertices are i-related if $\dim(a \cap b) = k - i$. If $v \geq 2k$, then there are k classes.

This scheme is the q-analogue of the Johnson scheme. We shall also meet this scheme in Chapter 5 when we deal with the q-analogue of the Erdős-Ko-Rado theorem.

The arguments we gave for the Johnson scheme apply almost verbatim to the Grassmann scheme. We now have a v-dimensional vector space V over $GF(q)$. We (re)define the matrices $W_{i,j}$ to be the matrix whose rows are indexed by the

i-subspaces of V, whose columns are indexed by the j-subspaces of V, and whose (a, b)-entry is 1 if the j-space b is contained in the i-space a, and 0 otherwise.

Again, the columns of $W_{k,i}$ span the first $i + 1$ eigenspaces of the scheme, and it follows that the Grassmann scheme is cometric.

Hamming Scheme

The Hamming scheme $H(n, q)$ is defined in terms of n and q, with $n, q > 0$. The vertex set is the set of all vectors of length n with entries from $0, \ldots, q-1$. Vertices a and b are i-related if they differ in i positions. There are n classes.

The graph corresponding to being 1-related is better known as the Hamming graph with parameters n and q. Again, it is straightforward to show that being i-related and being at distance i in the Hamming graph are the same thing. In Chapter 6 we will deal with the $\frac{n}{2}$-distance graph of the Hamming graph. Again, the Hamming scheme will provide us with a way to obtain the eigenspaces of the graph.

For the Hamming scheme, we have a convenient representation in terms of a Cayley graph. It will follow that we can use the theory of group characters to obtain a complete set of eigenvectors for the Hamming scheme. We will do this in Section 6.7. Our presentation there will focus on a particular graph in the scheme, but in fact the eigenspaces we find will be the eigenspaces of the scheme.

The vertices of the Hamming scheme $H(n, 2)$ are all 01-vectors of length n, or equivalently, all subsets of $\{1, \ldots, n\}$. Define the matrix W_j to be the matrix whose rows are indexed by the subsets of $\{1, \ldots, n\}$, whose columns are indexed by the j-subsets, and whose (S, a)-entry is

$$(-1)^{|S \cap a|}.$$

It is not hard to show that each column of W_j is a Schur polynomial in the columns of W_1. In fact, all we really need to observe is that if $a = \{a_1, \ldots, a_j\}$, then

$$(-1)^{|S \cap a|} = \prod_{t=1}^{j} (-1)^{|S \cap a_t|}.$$

It will turn out that the columns of W_j form a basis for the eigenspace corresponding to E_j. Thus the Hamming scheme is cometric.

It is interesting, though not directly relevant to our purposes, to note that the restriction of the Hamming scheme $H(n, 2)$ to the vertices of weight k is exactly $J(n, k)$. For comparison, we mention that the Grassmann scheme is a restriction of the bilinear forms scheme, which is itself the q-analogue of the Hamming scheme. We will not need this, but it does suggest strongly that theorems (and proofs!) dealing with the Johnson scheme should "q-generalize" to the Grassmann scheme in a natural way. This is the motivating idea behind Chapter 5.

Chapter 3

The Ratio Bound

3.1 Introduction

The central tool we use in this thesis is an eigenvalue bound on the size of an independent set. It was first established by Delsarte for the special case of a graph in an association scheme, and then generalized by Hoffman to regular graphs. It states that for any k-regular graph on n vertices with least eigenvalue τ, the size of an independent set is bounded above by

$$|S| \leq \frac{n}{1 - \frac{k}{\tau}}.$$

Furthermore, if equality holds then the characteristic vector of S is a linear combination of $\mathbf{1}$ and a τ-eigenvector.

The ratio bound for graphs in association schemes is a special case of Delsarte's linear programming bound. This in turn can be regarded as a specialization of the ϑ'-bound to association schemes. The ϑ' bound is Schrijver's variant of the ϑ-bound of Lovász, which is an upper bound on the Shannon capacity of a graph. These are all upper bounds on the size of an independent set: from this point of view, the ratio bound is the weakest. But it is the easiest to compute, and in many cases of interest it is tight.

We refer to independent sets that meet the ratio bound as *ratio-tight*. These are of course maximum independent sets, but as we will see throughout this thesis, they give us additional information about the graph. They tend to sit inside the graph in special ways, and they often have a "nice" combinatorial characterization: this is the case for the Kneser graphs, and the partition graphs of Chapter 4. We will find that the ratio-tight sets in the partition graph $\mathcal{P}(3^3)$ can be used to show that this graph is a core, that is, that there is no homomorphism from $\mathcal{P}(3^3)$ to a proper subgraph. In the graphs $\Omega(n)$ in Chapter 6, we will see that in addition to the eigenvector condition mentioned above, the ratio-tight sets play a crucial role in colouring, and a natural role in a recursive construction as well. In fact, we could say that our main result of Chapter 6 follows as a consequence of three different proofs of the ratio bound.

It is clear that more is involved here than a simple bound on the size of an independent set, so we are motivated to investigate the bound itself. We will derive a generalization of the ratio bound to non-regular graphs. More specifically, we will derive a template for building generalized ratio bounds. This can be applied to give bounds on independent sets in polarity graphs, and in particular the Erdős-Rényi graphs (see Section 3.11).

We consider graph products. A recent paper of Alon, Dinur, Friedgut and Sudakov deals with ratio-tight independent sets in products (though they do not use that language). We will show that this fits in very naturally with our approach, and offer an alternative proof of their result. This proof is interesting for two reasons. First, we argue that it is sufficient to prove the result for cliques. Secondly, our proof of this result is almost identical to a key step of our proof of the qEKR theorem in Chapter 5. We view this as evidence that our approach is a natural one to take when dealing with ratio-tight independent sets.

3.2 Cliques in Association Schemes

Cliques in association schemes were considered by Delsarte in his groundbreaking thesis [18]. Let $\mathcal{A} = \{A_0, \ldots, A_d\}$ be an association scheme on the vertex set V, and let $M \subseteq \{1, \ldots, d\}$. A set $Y \subseteq V$ is an *M-clique* if every pair of distinct vertices in Y are i-related for some $i \in M$. In other words, Y forms a clique in the graph consisting of the union of the classes indexed by M. Clearly if $|M| = 1$ this is just an ordinary clique.

Delsarte's motivation was from coding theory. Consider a code whose codewords are n-vectors in $\mathbb{Z}_q$: such a code is a subset of the vertices of the Hamming scheme. Two codewords are at distance i if and only if they are, as vertices in the scheme, i-related. A code with minimum distance δ is exactly an M-clique, where $M = \{\delta, \delta + 1, \ldots, d\}$. So in order to find a bound on the size of a code of given minimum distance, we need to bound an M-clique in the Hamming scheme. A similar argument applies to the Johnson scheme, or for that matter, any metric scheme.

We give a brief outline of Delsarte's arguments; for more details, the reader is referred to Delsarte [18, Chapter 3]. Let Y be a subset of the vertices of the scheme, and z_Y its characteristic vector. We define its *inner distribution* to be the row vector a_Y, where

$$(a_Y)_i = \frac{1}{|Y|} z_Y^T A_i z_Y.$$

Thus $(a_Y)_i$ is the average number of vertices of Y that are i-related to a fixed vertex of Y. Delsarte showed that for any inner distribution, the entries of aQ are all non-negative, where Q is the matrix of dual eigenvalues of the scheme. Clearly the entries of a sum to $|Y|$. If $|Y|$ is an M-clique then $(a_Y)_i = 0$ for $i \notin M$. It follows that we can bound $|Y|$ with a linear program.

Let x be a nonnegative vector such that $x_0 = 1$ and $x_i = 0$ for $i \notin M$. Then

for such x,

$$|Y| \leq \max\left\{x^T\mathbf{1} : (x^TQ)_i \geq 0\right\}$$

Using linear programming duality and the orthogonality relationship Lemma 2.5.3, we find that this is equivalent to the following, where y is a nonnegative vector such that $y_0 = 1$.

$$|Y| \leq \min\left\{y^T\mathbf{1} : (y^TP)_i \leq 0, i \in M\right\}$$

Thus we have the following result, which is a central tool in Delsarte's work.

3.2.1 Lemma. *Let $\mathcal{A} = \{A_0, \ldots, A_d\}$ be an association scheme and let M be a subset of $\{1, \ldots, d\}$. Then for any M-clique Y, we have:*

$$\begin{aligned}|Y| &\leq \max\left\{x^T\mathbf{1} : (x^TQ)_i \geq 0\,;\, x_i \geq 0\,;\, x_0 = 1\,;\, x_i = 0, i \notin M\right\}\\ &= \min\left\{y^T\mathbf{1} : (y^TP)_i \leq 0, i \in M\,;\, y_i \geq 0\,;\, y_0 = 1\right\}\end{aligned}$$ □

Consider the case where $M = \{i\}$, so that we are looking for a clique in the i-th graph. The eigenvalues of this graph are to be found in the i-th column of P, the matrix of eigenvalues of the scheme. Denote the greatest eigenvalue by k (thus $k = p_i(0)$, the entry in the 0-th row of the i-th column of P). Denote the least eigenvalue by τ; say it occurs in the j-th row. Now set $y_0 = 1, y_j = -\frac{k}{\tau}$ and $y_i = 0$ for all other entries. It is not too hard to see that this is a feasible solution to the dual linear program of Lemma 3.2.1, and thus we have established the following bound, from [18, p. 31].

3.2.2 Lemma. *Let $\mathcal{A} = \{A_0, \ldots, A_d\}$ be an association scheme, and Y a clique in A_i for some $1 \leq i \leq d$. Let k be the greatest eigenvalue of A_i and τ the least. Then*

$$|Y| \leq 1 + \frac{k}{-\tau}.$$ □

Let $\overline{M} = \{1, \ldots, d\} \setminus M$, and let Z be an $\overline{M}$-clique. Clearly Z is an independent set in the graph $\cup_{i\in M} A_i$. Thus Lemma 3.2.1 may be thought of as a bound on independent sets. It is not in general equivalent to the ratio bound. It follows from the work of Delsarte [18, Section 3.3] that the LP bound on an $\{i\}$-clique implies the ratio bound for the i-th graph of the scheme. Thus for instance, the ratio bound holds for strongly regular graphs. In fact, using the machinery of [18], one can show that the ratio bound holds for a union of classes in an association scheme.

We have another application of Lemma 3.2.2, based on an unpublished argument of Chris Godsil.

3.2.3 Lemma. *Let X be a graph in an association scheme, with valency k and eigenvalue λ. If we have a clique C and an independent set S such that*

$$|C| = 1 + \frac{k}{-\lambda}$$
$$|S| = \frac{|V(X)|}{1 + \frac{k}{-\lambda}}$$

then λ is the least eigenvalue of X, and C and S are both maximum.

Proof. Let τ be the least eigenvalue of X. Then by Lemma 3.2.2, we see that

$$1 + \frac{k}{-\lambda} = |C| \leq 1 + \frac{k}{-\tau}.$$

Unrolling the inequality, we find that $\lambda \leq \tau$. But since τ is the least eigenvalue equality must hold.

Knowing that λ is actually the least eigenvalue means that S is a ratio-tight independent set, and is hence maximum. □

Note that it is in principle possible to have an independent set of size

$$\frac{|V(X)|}{1 + \frac{k}{-\lambda}},$$

where λ is a non-least eigenvalue. It seems the independent set could be maximal by inclusion, or even maximum. We do not know of any examples where this occurs.

As an application of this, recall the partition graph of Example 2.4.2. It is not to hard to find a clique of size 4: we may regard a vertex as a parallel class in some affine plane of order three, and any affine plane of order three contains four parallel classes, which form a clique in $\mathcal{P}(3^3)$. It is also not hard to find an independent set of size 70: the set of all vertices that have 1 and 2 in the same cell. Taking the characteristic vector of such an independent set to be z, we find that

$$A\left(z - \frac{70}{280}\mathbf{1}\right) = -12\left(z - \frac{70}{280}\mathbf{1}\right).$$

Thus Lemma 3.2.3 implies that these cliques and independent sets are maximum. There are other ways to see this. In particular, since $\mathcal{P}(3^3)$ is vertex transitive we have

$$\alpha(\mathcal{P}(3^3))\,\omega(\mathcal{P}(3^3)) \leq |V(\mathcal{P}(3^3))|.$$

We see that merely the fact that $\mathcal{P}(3^3)$ lies in an association scheme is enough to give us a set of maximum independent sets. We did not even need to compute the eigenvalues or any other parameter of the scheme.

3.3 The Lovász ϑ-Bound

The results of this section constitute a summary of a paper of Schrijver [48], in which he explores the relationship between Delsarte's linear programming bound and Lovász' ϑ-bound. We include this to show how the linear programming bound of Delsarte and the ratio bound fit in with other graph bounds.

The *strong product* $G * H$ of two graphs G and H is the graph with vertex set $V(X) \times V(Y)$; vertices (x_1, y_1) and (x_2, y_2) are adjacent if x_1 and x_2 are equal or adjacent, and y_1 and y_2 are equal or adjacent.

Recall that the *Shannon capacity* of a graph G is

$$\Theta(G) = \sup_k \sqrt[k]{\alpha(G^{*k})} = \lim_k \sqrt[k]{\alpha(G^{*k})}.$$

We can think of the vertices of G as being letters in an alphabet, and edges representing pairs of letters that are not distinguishable. Then an independent set in G^{*k} represents a set of messages of length k such that any two of them can be distinguished in at least one position. Thus, for instance, if G is a vertex-disjoint union of c cliques then $\sqrt[k]{\alpha(G^{*k})} = c$ for all k (this would correspond to a rather silly alphabet to use). If S is an independent set in G, then the Cartesian product of k copies of S is an independent set in G^{*k}. Thus $\alpha(G)^k \leq \alpha(G^{*k})$ and it follows that

$$\alpha(G) \leq \Theta(G).$$

Equality need not hold. A well known result of Lovász [40] is that

$$\Theta(C_5) = \sqrt{\alpha(C_5^{*2})} = \sqrt{5} > \alpha(C_5).$$

Lovász introduced the following function, where we write $\rho(A)$ for the largest eigenvalue of A.

$$\begin{aligned}\vartheta(G) &= \min\left\{\rho(A) : A^T = A\,;\, A_{ij} = 1, i \not\sim j\right\},\\ &= \max\left\{\mathbf{1}^T B \mathbf{1} : B \succeq 0\,;\, tr(B) = 1\,;\, B_{ij} = 0, i \sim j\right\}.\end{aligned}$$

He showed that this is an upper bound on $\Theta(G)$. Equality need not hold, as shown by Haemers [35]. Lovász also showed that for regular graphs, $\vartheta(G)$ is at most the ratio bound on G. Thus if G is a k regular graph with least eigenvalue τ, then

$$\alpha(G) \leq \Theta(G) \leq \vartheta(G) \leq \frac{n}{1 - \frac{k}{\tau}}.$$

So we know the Shannon capacity of any graph that meets the ratio bound.

Schrijver introduced the function

$$\begin{aligned}\vartheta'(G) &= \min\left\{\rho(A) : A^T = A\,;\, A_{ij} \geq 1, i \not\sim j\right\},\\ &= \max\left\{\mathbf{1}^T B \mathbf{1} : B_{ij} \geq 0\,;\, B \succeq 0\,;\, tr(B) = 1\,;\, B_{ij} = 0, i \sim j\right\}.\end{aligned}$$

Clearly $\vartheta'(G)$ is a lower bound on $\vartheta(G)$. Schrijver further shows that

$$\alpha(G) \leq \vartheta'(G) \leq \vartheta(G).$$

Let G be a graph consisting of a union of classes in an association scheme, and let M be those classes not in the union. Schrijver showed that $\vartheta'(G)$ is exactly the linear programming bound of Delsarte for M-cliques, Lemma 3.2.1. Furthermore, in this case

$$\begin{aligned}\vartheta &\leq \max\left\{x^T\mathbf{1} : (x^TQ)_i \geq 0\,;\, x_0 = 1\,;\, x_i = 0, i \notin M\right\},\\ &= \min\left\{x^T\mathbf{1} : (y^TP)_i = 0, i \in M\,;\, y_i \geq 0\,;\, y_0 = 1\right\}.\end{aligned}$$

This is almost the same linear program as Lemma 3.2.1. There are examples in the Hamming scheme where $\vartheta'(G) < \vartheta(G)$; see [48] for details.

3.4 Proving the Ratio Bound

Delsarte first established the ratio bound in association schemes. He proved it using a linear programming argument based on the eigenvalues of the scheme, as we outlined in Section 3.2. It can also be proved using a graph quotient and interlacing. This approach is originally due to Haemers, and this is the argument given in Godsil and Royle [33, Section 9.6].

The proof we give is based on positive-semi definite matrices, and is an approach due to Godsil. We take this approach because it will allow to generalize the bound.

3.4.1 Theorem. *Let X be a k-regular graph on n vertices, S an independent set in X, and z the characteristic vector of S. Let A be the adjacency matrix of X and τ its least eigenvalue. Then*

$$|S| \leq \frac{n}{1 - \frac{k}{\tau}}.$$

Moreover, if equality holds then

$$A\left(z - \frac{s}{n}\mathbf{1}\right) = \tau\left(z - \frac{s}{n}\mathbf{1}\right).$$

Proof. Let A be the adjacency matrix of the k-regular graph X and τ its least eigenvalue. Then the matrix $A - \tau I$ is positive semi-definite. In particular, we have

$$y^T(A - \tau I)y \geq 0,$$

with equality holding if and only if $(A - \tau I)y = 0$.

Now let z be the characteristic vector of any independent set of X of size s and set

$$y = z - \frac{s}{n}\mathbf{1}.$$

Thus we have that

$$\left(z - \frac{s}{n}\mathbf{1}\right)^T (A - \tau I)\left(z - \frac{s}{n}\mathbf{1}\right) \geq 0.$$

Note that $z^T A z = 0$ and $z^T A \mathbf{1} = ks$. Upon expanding and rearranging terms, one discovers that:

$$s \leq \frac{n}{1 - \frac{k}{\tau}}.$$

Thus we have established the bound.

If equality holds, then

$$(A - \tau I)\left(z - \frac{s}{n}\mathbf{1}\right) = \mathbf{0},$$

and we conclude that $z - \frac{s}{n}\mathbf{1}$ is a τ-eigenvector, as desired. □

Note that if z is the characteristic vector of some independent set of size r and $z - \frac{r}{n}\mathbf{1}$ is a τ-eigenvector, then it is straightforward to show that

$$r = \frac{n}{1 - \frac{k}{\tau}}.$$

Thus every independent set whose characteristic vector lies in the sum of the τ-eigenspace and the k-eigenspace is a maximum independent set. It is not the case that every 01-vector in the sum of these two eigenspaces is a maximum independent set. For instance, the union of any two disjoint ratio-tight sets is a set whose characteristic vector lies in the sum of these two eigenspaces. In particular, for K_n, every 01-vector of length n is a linear combination of $\mathbf{1}$ and a τ-eigenvector. This is not really a surprise: the τ-eigenvectors are exactly $\mathbf{1}^\perp$, so the sum of $\mathbf{1}$ and the τ-eigenspace is $\mathbb{R}^n$.

3.5 Generalizing the Ratio Bound

We are motivated to generalize the ratio bound. Firstly, it occupies a central position in this thesis and so if only for the sake of completeness we feel bound to investigate it further. Secondly, we have a specific application in mind: the Erdős-Rényi graphs. These graphs are not regular, and hence the ratio bound described above does not apply. The graphs can be made regular by adding loops in a certain natural way, but then again the ratio bound does not directly apply.

Our proof in the previous section depends critically on two things. First, that the matrix $A - \tau I$ is positive semi-definite. Secondly, that the matrix $A - \tau I$ has a zero in position ij if $i \not\sim j$. This leads to the following very general approach:

(a) Choose a matrix B such that $B \succeq 0$ and $B_{ij} = 0$ whenever i and j are not joined by an edge.

(b) Let z be the characteristic vector of a maximum independent set S of size s. Compute $\left(z - \frac{s}{n}\mathbf{1}\right)^T B \left(z - \frac{s}{n}\mathbf{1}\right)$ and note that it is non-negative.

(c) Hope that upon simplifying the inequality something nice will happen involving $|S|$.

It will be useful to allow our graphs to contain loops. This forces us to reconsider our definition of an independent set. We will take the point of view that an independent set is a set of vertices with no two distinct elements adjacent. It will turn out to be useful when we consider the Erdős-Rényi graphs in Section 3.11.

We will defer the general strategy outlined above, and consider for the moment a restriction of it.

3.5.1 Lemma. *Let X be a graph on the vertex set V, with $|V| = n$ and vertex degrees $d_1, \ldots, d_n$. Let A be the adjacency matrix of X and $T = \mathrm{diag}(t_1, \ldots, t_n)$ be a diagonal matrix such that $A + T \succeq 0$. If S is a set of s vertices such that no two distinct elements of S are adjacent and S contains s_1 loops, then:*

$$\left(z - \frac{s}{n}\mathbf{1}\right)^T (A + T) \left(z - \frac{s}{n}\mathbf{1}\right) \geq 0.$$

In expanded form, we have

$$\frac{s^2}{n^2} \sum_{i \in V} (t_i + d_i) - 2\frac{s}{n} \sum_{i \in S} (t_i + d_i) + \sum_{i \in S} t_i \geq -s_1. \qquad \square$$

This gives a bound on s. However, it is difficult to apply in general, partly because the sums depend not only on s but on S. Furthermore, the bound obtained will depend on the choice of T. It is not known how to choose T optimally (or even if there is a single optimal choice for all graphs). We regard this as a template for producing bounds, and so we consider specific choices for T.

It will be useful to define the following parameters of a set S of size s:

$$\overline{d}_S = \frac{1}{s} \sum_{i \in S} d_i,$$

$$k_S = \frac{2}{s} \sum_{i \in S} d_i - \frac{1}{n} \sum_{i \in V} d_i.$$

Note that for k-regular graphs, $\overline{d}_S = k_S = k$. It will be seen that these two parameters behave, in some circumstances, as analogues to the degree of a regular graph.

3.6 Specific Bounds

We consider two particular choices for T. The bounds that arise can be regarded as coming from the adjacency matrix and the Laplacian matrix of the graph. We will in the course of this section rediscover the bound of Theorem 3.4.1.

Adjacency Matrix

Let $T = -\tau I$, where τ is the least eigenvalue of A. If the graph is regular and loopless, then we simply have a restatement of the ratio bound, which we reproduce here for convenience.

3.6.1 Corollary. *Let X be a connected k-regular graph with no loops and τ the least eigenvalue of its adjacency matrix. For any independent set S of size s we have*

$$s \leq n\frac{-\tau}{k-\tau}. \qquad \square$$

But Lemma 3.5.1 does not depend on the graph being regular. True, the terms involving the average degree are easier to deal with in a regular graph, but we do not need this condition. It turns out that the expression still simplifies nicely and we obtain the following bound.

3.6.2 Corollary. *Let X be a graph with no loops, and τ the least eigenvalue of its adjacency matrix. For any independent set S of size s, we have:*

$$s \leq n\frac{-\tau}{k_S-\tau}. \qquad \square$$

To be precise, one should say that Corollary 3.6.2 does not bound the size of an independent set, but provides a family of bounds, one for each value of k_S (or equivalently, one bound for each value of $\overline{d}_S$). We see that k_S plays a role analogous to that of the degree in Corollary 3.6.1. Looking at the ratio bound in hindsight, we see that the parameter k is both the greatest eigenvalue and the degree of the graph. Of course these two values are equal for regular graphs, but they are not equal in general. For non-regular graphs it appears that the the more general parameter k_S is more closely related to vertex degrees than to eigenvalues. However k_S is not an average degree: it can be zero or even negative.

Now we consider the case where the graph may have loops as well as not being regular. The bound of Lemma 3.5.1 is in fact a quadratic in s. This was perhaps not obvious in Corollary 3.6.1 and Corollary 3.6.2: there the constant term in the quadratic was zero, meaning we could cancel a factor of s. Now that we can no longer assume $s_1 = 0$, we get bounds that appear messier. The reader may verify that setting $s_1 = 0$ yields the previous bounds.

3.6.3 Corollary. *Let X be a k-regular graph and τ the least eigenvalue of its adjacency matrix. For any independent set S of size s containing s_1 loops, we have:*

$$s \leq n\frac{-\tau + \sqrt{\tau^2 + 4s_1\frac{k-\tau}{n}}}{2\,(k-\tau)}. \qquad \square$$

3.6.4 Corollary. *Let X be a graph and τ the least eigenvalue of its adjacency matrix. For any independent set S of size s containing s_1 loops, we have:*

$$s \leq n\frac{-\tau + \sqrt{\tau^2 + 4s_1\frac{k_S-\tau}{n}}}{2\left(k_S - \tau\right)}.$$

□

Laplacian Matrix

For a graph X with adjacency matrix A and diagonal matrix of degrees D, recall that the *Laplacian matrix* of X is $L = D - A$. We always have $L \succeq 0$, and in fact 0 is an eigenvalue of multiplicity equal to the number of components of X. The greatest eigenvalue of L is at most twice the maximum degree; it is also bounded by the number of vertices. See, for instance [4] for more details. If X is k-regular then $L = kI - A$, and the eigenvalues of L and A contain the same information. Accordingly we expect to recover previous bounds for regular graphs and hope to obtain new ones in the non-regular case. Note that graphs that differ only by the presence or absence of loops have the same Laplacian matrix. Thus, without loss of generality, we can delete all loop edges from the graph and set $s_1 = 0$ (this will not affect our results in any way, but it will simplify computations using this bound).

If we let μ be the greatest eigenvalue of L, then we may set $T = \mu I - D$ giving $A + T = \mu I - L \succeq 0$. If the graph is regular, then we recover Corollary 3.6.1, as expected. If it is not regular, then we obtain the following bound.

3.6.5 Corollary. *Let X be any graph, and μ the greatest eigenvalue of its Laplacian matrix. For any independent set S of size s, we have:*

$$s \leq n\frac{\mu - \overline{d}_S}{\mu}.$$

□

Note that $\overline{d}_S$ plays an analogous role to that of degree in Corollary 3.6.1. It is in some ways better behaved than k_S. For instance, $\overline{d}_S$ is only 0 in trivial cases; otherwise it is positive. Also, since $\overline{d}_S \geq \delta$, we can make the bound more usable at the cost of weakening the inequality slightly.

3.6.6 Corollary. *Let X be any graph with minimum degree δ, and μ the greatest eigenvalue of its Laplacian matrix. For any independent set S of size s, we have:*

$$s \leq n\frac{\mu - \delta}{\mu}.$$

□

Although this is weaker than Corollary 3.6.5, it truly is a bound, in that the right-hand side is a function of the graph alone. In Corollary 3.6.5, the right-hand side depends on the structure of the independent set as well.

3.7 Equality Conditions

We now consider what happens when Lemma 3.5.1 holds with equality.

3.7.1 Lemma. *Let X be a graph on the vertex set V, with $|V| = n$ and vertex degrees $d_1, \ldots, d_n$. Let A be the adjacency matrix of X and $T = \mathrm{diag}(t_1, \ldots, t_n)$ be a diagonal matrix such that $A + T \succeq 0$. Let S be a set of s vertices such that no two distinct elements of S are adjacent, and let S contain s_1 loops. If*

$$\left(z - \frac{s}{n}\mathbf{1}\right)^T (A+T)\left(z - \frac{s}{n}\mathbf{1}\right) = 0.$$

then for $i \in S$ we have

$$d_i = |\{j \notin S : j \sim i\}| = t_i\left(\frac{n}{s} - 1\right)$$

and for $i \notin S$, we have

$$|\{j \in S : j \sim i\}| = (d_i + t_i)\,\frac{s}{n}.$$

Proof. Recall that if B is a positive semidefinite matrix and $x^T B x = 0$ then $Bx = 0$. Applying this gives an eigenvector for $A + T$, namely

$$(A+T)(z - \frac{s}{n}\mathbf{1}) = 0.$$

The result follows. □

The first equation is just giving the degree of each vertex in S. We observe that if a particular choice of T is to lead to a bound that holds with equality, T must be chosen as a function of the vertex degrees. We will see that we recover the equality conditions of Theorem 3.4.1, with perhaps a little more context.

Adjacency Matrix

First consider a connected k-regular graph with no loops. If the bound of Corollary 3.6.1 holds with equality, then from Lemma 3.7.1 we have an eigenvector for A:

$$A\left(z - \frac{s}{n}\mathbf{1}\right) = \tau\left(z - \frac{s}{n}\mathbf{1}\right).$$

Since $\mathbf{1}$ is a k-eigenvector, we see that z lies in the sum of the eigenspaces corresponding to the greatest and least eigenvalues, k and τ. Using Lemma 3.7.1 with $t_i = -\tau$ and $d_i = k$, we see that each vertex in S has

$$-\tau\left(\frac{n}{s} - 1\right)$$

neighbours not in S, and each vertex not in S has

$$(k - \tau)\,\frac{s}{n}$$

neighbours in S.

Since all the neighbours of a vertex in S are outside S, we have an equation relating τ, s, and k:

$$k = -\tau\left(\frac{n}{s} - 1\right).$$

A partition $\Pi = \{\pi_1, \ldots, \pi_m\}$ of the vertex set of a graph is an *equitable partition* if the number of edges between a vertex $u \in \pi_i$ and the set of vertices of π_j depends only on i and j. The fact that vertices not in S have a constant number of neighbours in S implies that the partition $\{S, V\backslash S\}$ is equitable. We can summarize our findings for an adjacency matrix bound for a regular graph with no loops; this is just a restatement of Theorem 3.4.1 in greater detail. It will be useful for comparison.

3.7.2 Theorem. *Let X be a k-regular graph with no loops, and τ the least eigenvalue of its adjacency matrix. For any independent set S of size s and characteristic vector z, we have:*

$$s \leq n\frac{-\tau}{k-\tau}.$$

Furthermore, the following are equivalent:

(a) Equality holds.
(b) z is a linear combination of $\mathbf{1}$ and a τ-eigenvector.
(c) The bipartite subgraph induced by the partition $\{S, V(X) \setminus S\}$ is semi-regular.
(d) The partition $\{S, V(X) \setminus S\}$ is equitable. □

For general non-regular graphs, it does not seem to be the case that we obtain as much useful structural information in the case of equality. Part of the reason for this is that $\mathbf{1}$ is no longer an eigenvector of $A - \tau I$.

Laplacian Matrix

Recall that the form of the Laplacian bound was the same whether or not the graph was regular or had loops. If equality holds in Corollary 3.6.5, then we have an eigenvector for L:

$$L\left(z - \frac{s}{n}\mathbf{1}\right) = \mu\left(z - \frac{s}{n}\mathbf{1}\right).$$

Since $\mathbf{1}$ is a 0-eigenvector, we see that z lies in the sum of the eigenspaces corresponding to the greatest and least eigenvalues. In fact, we could have written

$$Lz = \mu\left(z - \frac{s}{n}\mathbf{1}\right),$$

although it is not as obviously an eigenvector condition. Using Lemma 3.7.1 with $t_i = \mu - d_i$, we see that each vertex in S has

$$\mu\left(1 - \frac{s}{n}\right) \tag{3.1}$$

neighbours not in S, and each vertex not in S has

$$\mu\frac{s}{n} \tag{3.2}$$

neighbours in S.

Note that this implies that the vertices of S all have the same degree, and gives a equation relating μ, s and the degree of a vertex in S. This does not quite say that the partition $\{S, V \setminus S\}$ is equitable. The only thing missing is that we cannot guarantee that every vertex not in S has the same number of neighbours not in S. However, if this condition were true, then the graph would necessarily be regular. So we can say informally that if the bound of Corollary 3.6.5 holds with equality, then the partition induced by S is as equitable as it could be without the graph being regular. We summarize our findings on the Laplacian bound for independent sets.

3.7.3 Theorem. *Let X be a graph with no loops, and μ the greatest eigenvalue of its Laplacian matrix. For any independent set S of size s and characteristic vector z, we have:*

$$s \leq n\frac{\mu - \overline{d}_S}{\mu}.$$

Furthermore, the following are equivalent:

(a) Equality holds.
(b) z is a linear combination of $\mathbf{1}$ and a μ-eigenvector.
(c) The bipartite subgraph induced by the partition $\{S, V(X) \setminus S\}$ is semi-regular. □

Recall that $\mathbf{1}$ is a 0-eigenvector of the Laplacian, and if the graph is k-regular, $\mathbf{1}$ is a k-eigenvector of the adjacency matrix. Furthermore, for connected graphs, $\mathbf{1}$ spans the 0-eigenspace of the Laplacian and for connected k-regular graphs, $\mathbf{1}$ spans the k-eigenspace of the adjacency matrix. So when equality holds in Theorem 3.7.2 or Theorem 3.7.3, z is a linear combination of eigenvectors belonging to the greatest and least eigenvalues. Based partly on the analogy between Theorem 3.7.2 and Theorem 3.7.3, and the fact that Theorem 3.7.2 is actually a special case of Theorem 3.7.3, it seems that the Laplacian matrix formulation is the natural generalization to non-regular graphs.

The equalities in (3.1) and (3.2) are actually equalities of integers. Unpacking this, we get that μ must be an integer, and $n \mid \mu s$.

Let us assume for the moment that equality holds in Theorem 3.7.3 and $\gcd(s, n) = 1$. Then since $0 < \mu \leq n$, not only does $n \mid \mu$ but $n = \mu$. This gives $s = n - \delta$. Thus we see that the graph would have to consist of an independent set of size s, every vertex of which is adjacent to all the remaining $\delta = \overline{d}_S$ vertices. So the bipartite subgraph induced by the partition $\{S, V(X) \setminus S\}$ must be complete bipartite.

3.8 Relating Bounds

We are motivated to compare the "adjacency matrix bound" and the "Laplacian matrix bound". The former seems more complicated, and from our results in Section 3.7, it would seem that the Laplacian version is preferred.

However, if we have a regular graph with loops, then applying Corollary 3.6.3 yields the same bound as removing the loops and applying Corollary 3.6.5. To see this, consider a k-regular graph where some of the vertices have loops. If we delete the loops, then $\overline{d}_S$ is not equal to k: for an independent set of size s that contains s_1 formerly-looped vertices, we would have

$$\overline{d}_S = k - \frac{s_1}{s}.$$

Substituting this value into Corollary 3.6.5 gives a quadratic whose solution is exactly the bound of Corollary 3.6.3. In fact we can see this directly. Let X be some graph, and A its adjacency matrix. Let E be a diagonal matrix of nonnegative integers such that the row sums of $A + E$ are constant, equal to k say. In other words, $A + E$ is the adjacency matrix of a graph X' derived from X by the addition of loops so as to make X' regular. Let τ be the least eigenvalue of $A + E$. Applying Corollary 3.6.3 to X' is equivalent to setting

$$T = -\tau I$$

in Lemma 3.5.1. Now observe that

$$(A + E) - \tau I = A - (kI - E) + (k - \tau)I.$$

We recognize $kI - E$ as being the diagonal matrix of degrees of the graph X. Also, $k - \tau$ is the greatest eigenvalue of the Laplacian matrix of X', which is the same thing as the greatest eigenvalue of the Laplacian matrix of X (note that τ is not the least eigenvalue of A).

In other words, applying Corollary 3.6.5 to a graph X is exactly equivalent to adding sufficient loops so as to make the new graph X' regular, and applying Corollary 3.6.3 to X' (this amounts to an appropriate choice of T in Lemma 3.5.1). Note that this does not mean that applying Corollary 3.6.5 to X is the same thing as applying Corollary 3.6.2 to X. The key observation is that Laplacian eigenvalues are independent of the number of loops on each vertex, whereas the eigenvalues of the adjacency matrix are not. If we want to apply Corollary 3.6.3 to X', then we must compute the least eigenvalue of X', which amounts to computing the greatest Laplacian eigenvalue of X. So this is really just a statement about relationships between Laplacian and adjacency matrix eigenvalues.

3.9 Comparing Bounds

Given that they are related but not equivalent, it is natural to ask which of Corollary 3.6.3 and Corollary 3.6.5 is stronger, if either. To answer this question, we will consider some examples.

3.9.1 Example. The complete bipartite graphs $K_{a,b}$, where $a < b$.

Clearly the only maximum independent set is the set of vertices of degree a. So in this case we may calculate:

$$\overline{d}_S = a, \qquad k_S = \frac{2a^2}{a+b}.$$

Also, the least eigenvalue of the adjacency matrix is $-\sqrt{ab}$ and the greatest eigenvalue of the Laplacian matrix is $a + b$. Applying Corollary 3.6.2 we get

$$s \leq \frac{(a+b)^2\sqrt{ab}}{(a+b)\sqrt{ab} + 2a^2}. \tag{3.3}$$

However applying Corollary 3.6.5 we get exactly the size of the maximum independent set.

$$s \leq b.$$

Thus we conclude that the bounds based on the adjacency and Laplacian matrices are not equal. Note that we may round down the bound in (3.3), and for $a < b \leq 22$ this will give the same result as the Laplacian. To give some idea of what "small" means, if $a < b \leq 22$, then (3.3) gives $s \leq b$; but for $K_{4,23}$, the bound of (3.3) gives $s \leq 24$. For fixed a, as $b \to \infty$, the floor of the bound of (3.3) tends towards $a + b - 1 = n - 1$. So asymptotically, the bound based on the adjacency matrix says only that it is not the case that the graph is edgeless, while the bound based on the Laplacian matrix gives the exact answer.

It is notable here that not only is Corollary 3.6.5 tight, but so is Corollary 3.6.6. The latter bound is in terms of the graph only, whereas the former depends on $\overline{d}_S$ and so retains an implicit dependence on the structure of S.

3.9.2 Example. Let the graphs X_m, $m > 1$ be constructed as follows. Let G_m be a copy of $\overline{K_m}$, and H_m be a copy of C_{2m+1}. Then X_m consists of the disjoint union of G_m and H_m, together with edges from every vertex of G_m to every vertex of H_m.

Owing to the structure, we can write down the adjacency and Laplacian matrices in a nice way.

$$A(X_m) = \begin{pmatrix} 0 & J \\ J & A(C_{m+1}) \end{pmatrix} \qquad L(X_m) = \begin{pmatrix} 0 & -J \\ -J & 2I - A(C_{m+1}) \end{pmatrix}$$

Using this structure we can determine the eigenvectors. We find that the least eigenvalue of the adjacency matrix is $\tau = 1 - \sqrt{2m^2 + m + 1}$, and the greatest eigenvalue of the Laplacian matrix is $\mu = 3m + 1$.

The maximum independent sets are of size m. There are two types: the set of vertices of G_m and an independent set of m vertices in H_m. For convenience, let α_A and α_L be the values of the bounds in Corollary 3.6.2 and Corollary 3.6.5, respectively.

If S is the vertex set of G_m, then we find that $\overline{d}_S = 2m+1$. It follows that the Laplacian bound of Corollary 3.6.5 gives m, whereas the adjacency matrix gives a somewhat messy expression that is greater than m. In other words,

$$|S| = \alpha_L < \alpha_A.$$

Now let S be a maximum independent set in H_m; we find that $\overline{d}_S = m+2$. Again, the Laplacian bound gives a much simpler expression, but it is not always better. For $2 \leq m \leq 24$ we find that

$$|S| = m < \alpha_A < \alpha_L = \frac{2m-1}{3m+1},$$

whereas for $m \geq 25$ we have

$$|S| = m < \alpha_L = \frac{2m-1}{3m+1} < \alpha_A.$$

A similar approach using different graphs in place of cycles for H_m can be used to construct more examples with either bound better. Thus the answer is that neither of these two bounds is uniformly better than the other. It does seem that the equality conditions relating to Corollary 3.6.5 are stronger than those for Corollary 3.6.3.

3.10 A Bound of Sarnak

We turn our attention to another eigenvalue bound on independent sets, due to Sarnak. We show that it is strictly weaker than the ratio bound.

Let λ be the maximum of the second largest eigenvalue and the absolute value of the least eigenvalue of the adjacency matrix. Sarnak [46] has shown the following bound for an independent set S in a k-regular graph.

3.10.1 Lemma.

$$|S| \leq n\frac{\lambda}{k}$$

We give an outline of the proof in [46]. *Proof.* If $x \perp \mathbf{1}$ then $\|Ax\|^2 \leq \|\lambda x\|^2$. We choose

$$x_i = \begin{cases} n-s, & i \in S \\ -s, & i \notin S \end{cases}$$

and compute the norms as follows:

$$\|\lambda x\|^2 = \lambda^2 \left(s(n-s)^2 + (n-s)s^2\right) = \lambda^2 ns(n-s),$$

$$\begin{aligned} \|Ax\|^2 &= \sum_{i \in S} \left((Ax)_i\right)^2 + \sum_{i \notin S} \left((Ax)_i\right)^2 \\ &\geq \sum_{i \in S} \left((Ax)_i\right)^2 \\ &= s^3k^2. \end{aligned} \tag{3.4}$$

Substituting into $\|Ax\|^2 \leq \|\lambda x\|^2$, we find that

$$s \leq \sqrt{n(n-s)}\frac{\lambda}{k}$$

which implies

$$s \leq n\frac{\lambda}{k}. \tag{3.5}$$

□

Note that at (3.4) we are neglecting some positive terms, and at (3.5), we are using $n - s \leq n$. Assuming that $0 < s < n$, either of these is sufficient to guarantee that the inequality in Lemma 3.10.1 is in fact strict. Based on these observations, we can improve this proof. If we write the adjacency matrix in the form

$$A = \begin{pmatrix} 0 & B \\ B^T & C \end{pmatrix},$$

then we can compute the missing contributions as follows:

$$\begin{aligned} \sum_{i \notin S} ((Ax)_i)^2 &= \sum_i \left((n-s)(B^T\mathbf{1})_i + (-s)(C\mathbf{1})_i\right)^2 \\ &\geq \frac{1}{n-s}\left(\sum_i \left((n-s)(B^T\mathbf{1})_i + (-s)(C\mathbf{1})_i\right)\right)^2 \qquad (3.6) \\ &= \frac{1}{n-s}\left((n-s)(sk) + (-s)((n-2s)k)\right)^2 \\ &= \frac{s^4k^2}{n-s}. \end{aligned}$$

This gives that

$$\|Ax\| \geq \frac{s^3k^2n}{n-s},$$

and then using $\|Ax\|^2 \leq \|\lambda x\|^2$ and rearranging gives exactly

$$s \leq \frac{n}{1+\frac{k}{\lambda}}. \tag{3.7}$$

This bound is strictly better than Lemma 3.10.1, but, as $\lambda \geq -\tau$, it is no stronger than the ratio bound. In the context of the present paper, this strengthening is not a surprise: the vector x in the proof of Lemma 3.10.1 is just a multiple of $z - \frac{s}{n}\mathbf{1}$. Note also that if(3.7) holds with equality, then (3.6) holds with equality, so in the extremal case we haven't neglected any terms.

3.11 Erdős-Rényi Graphs

The graphs known as the Erdős-Rényi graphs were first described as a family of graphs containing no four-cycles, with an asymptotically maximum number

of edges. We will denote them by $ER(q)$. The vertex set of $ER(q)$ is the set of points of $PG(2,q)$. Two distinct vertices x and y are adjacent if $x^T y = 0$.

These graphs are defined for any prime power q, but in the present work we will only deal with q odd.

We start with some general facts about these graphs which are well-known. The graphs $ER(q)$ are not regular. Most vertices have degree $q+1$, but there are $q+1$ vertices that have degree q; these are the *absolute vertices*. They are exactly the vertices x such that $x^T x = 0$. So were we to add a loop to each of the absolute vertices, the graph would be $(q+1)$-regular. The non absolute vertices can be partitioned into two classes. There are

$$\frac{q^2+q}{2}$$

external vertices, each of which is adjacent to exactly two absolute vertices; each absolute vertex is adjacent to q external vertices. There are

$$\frac{q^2-q}{2}$$

internal vertices, none of which are joined to any absolute vertex. For more details and background on these graphs, the reader is referred to [23, 9, 44, 50].

We are concerned with the size of a maximum independent set. One approach would be to add loops to each absolute vertex and apply the ratio bound. The least eigenvalue of this graph is $-\sqrt{q}$. So the value of the bound is

$$\frac{q^2+q+1}{1+\frac{q+1}{\sqrt{q}}} = \frac{q^2+q+1}{q+\sqrt{q}+1}\sqrt{q}.$$

However, the ratio bound deals with sets that contain no loops. Accordingly, this approach will give a bound on an independent set that contains no absolute vertices. So we must compensate by adding the absolute vertices back in:

$$\alpha(ER(q)) \leq \frac{(q^2+q+1)\sqrt{q}}{q+\sqrt{q}+1} + q + 1. \tag{3.8}$$

A similar approach, based on the bound of Sarnak (Lemma 3.10.1) was used by Williford [50] to provide an upper bound. Our goal is to improve this using the generalizations of the ratio bound we derived in Section 3.6 (In fact, our derivation of those generalizations were partly motivated by a desire to improve this bound).

We can apply Corollary 3.6.3 to the graph with loops added. Let S be an independent set of size s, containing s_1 absolute vertices. This gives

$$s \leq \frac{\sqrt{q} + \sqrt{q + 4s_1\frac{q+\sqrt{q}+1}{q^2+q+1}}}{2\frac{q+\sqrt{q}+1}{q^2+q+1}}.$$

This is not quite a bound yet, since the right hand side depends on the independent set, namely, in the parameter s_1. However, we know that $0 \leq s_1 \leq q+1$, so we obtain the following bound on an independent set:

$$\alpha(ER(q)) \leq \frac{\sqrt{q} + \sqrt{q + 4(q+1)\frac{q+\sqrt{q}+1}{q^2+q+1}}}{2\frac{q+\sqrt{q}+1}{q^2+q+1}}. \tag{3.9}$$

This bound is strictly better than (3.8).

As we noted above, applying Corollary 3.6.5 to the graphs $ER(q)$ gives exactly the same result as applying Corollary 3.6.3 to the graph with loops added.

Another approach would be to delete the vertices that have loops, apply Corollary 3.6.2 to this graph, and translate the result back to $ER(q)$. Let $ER_0(q)$ be the graph obtained by deleting the absolute vertices from $ER(q)$. In unpublished notes, Chris Godsil and Gordon Royle have computed the characteristic polynomial of $ER_0(q)$ to be

$$(\lambda - q)\lambda(\lambda+1)^q(\lambda^2 - q)^{(q^2-q-2)/2}.$$

This has least eigenvalue $-\sqrt{q}$. Let S be an independent set in $ER(q)$ of size s containing s_1 absolute vertices. Let S_0 be the set of non-absolute vertices of S. Note that even if S is maximal in $ER(q)$ this does not mean that S_0 is maximal in $ER_0(q)$. To apply Corollary 3.6.2 to $ER_0(q)$, we will need the average degree of S_0 in $ER_0(q)$. We have two different bounds on this parameter.

Recall that in $ER(q)$, each vertex of S_0 is adjacent to at most 2 absolute vertices: this means that the average degree of vertices of S_0 in the graph $ER_0(q)$ is at least $q-1$. So we can compute a bound on k_{S_0} as:

$$k_{s_0} \geq q - 2 + \frac{1}{q}.$$

Applying Corollary 3.6.2 to $ER_0(q)$ and adding s_1, we get the bound

$$s \leq \frac{q^2\sqrt{q}}{q - 2 + \frac{1}{q} + \sqrt{q}} + s_1.$$

To obtain a bound independent of s_1, we set $s_1 = q+1$ to obtain the following.

$$\alpha(ER(q)) \leq \frac{q^2\sqrt{q}}{q - 2 + \frac{1}{q} + \sqrt{q}} + q + 1 \tag{3.10}$$

On the other hand, each absolute vertex is adjacent to at most q vertices of S_0. So there are at most $q(q+1-s_1)$ edges between S_0 and the set of absolute vertices. So we have another bound on k_{S_0}:

$$k_{s_0} \geq q + 2 + \frac{1}{q} - \frac{2q(q+1-s_1)}{s - s_1}.$$

Applying Corollary 3.6.2 to $ER_0(q)$ and adding s_1, we get the bound

$$s \leq \frac{q^2\sqrt{q} + 2q(q+1)}{q+2+\frac{1}{q}+\sqrt{q}} + s_1 \frac{-q+2+\frac{1}{q}+\sqrt{q}}{q+2+\frac{1}{q}+\sqrt{q}}.$$

The term involving s_1 is negative for $q > 3$. So for $q > 3$, to obtain a bound independent of s_1 we set $s_1 = 0$.

$$\alpha(ER(q)) \leq \frac{q^2\sqrt{q} + 2q(q+1)}{q+2+\frac{1}{q}+\sqrt{q}} \tag{3.11}$$

For $5 \leq q \leq 23$, (3.11) is better than (3.10), but the reverse is true for $q \geq 25$. Neither bound is as good as (3.9).

Yet another approach would be to apply Corollary 3.6.2 directly to $ER(q)$ (recall that this graph has no loops). In unpublished notes, Chris Godsil and Gordon Royle have computed the characteristic polynomial of $ER(q)$ as:

$$(\lambda^3 - q\lambda^2 - 2q\lambda + q^2 + q)(\lambda^2 + \lambda + 1 - q)^q(\lambda^2 - q)^{(q^2-q-2)/2}$$

The least eigenvalue is a root of the cubic factor. We can approximate this eigenvalue using Newton's method. This gives us a lower bound on the least eigenvalue, which in turns provides us with an upper bound on the value of an upper bound on the size of an independent set. Let w be a lower bound on the least eigenvalue. As usual, S will be an independent set of size s, containing s_1 absolute vertices. It is straightforward to compute k_S:

$$\begin{aligned} k_S &= 2\left(q+1-\frac{s_1}{s}\right) - \left(q+1-\frac{q+1}{q^2+q+1}\right) \\ &= q+1+\frac{q+1}{q^2+q+1} - \frac{2s_1}{s}. \end{aligned}$$

Using Corollary 3.6.2, we obtain

$$s \leq \frac{(q^2+q+1)(-w)+2s_1}{q+1-w+\frac{q+1}{q^2+q+1}}.$$

Again, in order to obtain a bound independent of s_1, we set $s_1 = q+1$. The least eigenvalue is less than $-\sqrt{q}$, and it is the only eigenvalue less than $\sqrt{q}$; also the cubic factor of the characteristic polynomial in concave down for $\lambda \leq -\sqrt{q}$. So we know that iterating Newton's method starting with $\sqrt{q}$ will always give a lower bound on the least eigenvalue. Furthermore, it seems that two iterations are sufficient, since we will round the bound down to an integer value anyway. So we have the following bound.

$$s \leq \frac{(q^2+q+1)(-w)+q+1}{q+1-w+\frac{q+1}{q^2+q+1}}. \tag{3.12}$$

We close this section with a brief table summarizing the numerical values of the bounds we have derived. We also include exact values for the size of a maximum independent set; these are from Williford [50, Section 4.3]. The bounds are sorted according to their asymptotic order. The best bound we know of is (3.9).

q	$\alpha(ER(q))$	(3.9)	(3.12)	(3.8)	(3.10)	(3.11)
3	5	5.56	5.63	7.92	9.09	5.60
5	10	10.56	10.82	14.42	16.28	12.28
7	15	16.73	17.27	22.16	24.65	20.50
9	22	23.93	24.87	31	34.03	29.98
11	29	32.05	33.40	40.79	44.34	40.55
13	38	41.03	42.88	51.48	55.49	52.08

Recall that each of our generalizations was based on a particular choice of the matrix T in Lemma 3.5.1, making $A+T \succeq 0$. The bound of (3.9) is derived from a choice of T that makes $\mathbf{1}$ an eigenvector for $A+T$. This seems to be significant. We have not exhausted the potential of Lemma 3.5.1 for the Erdős-Rényi graphs. One avenue that seems promising is to try weighting the adjacency matrix as a function of the vertices being absolute, internal or external. Aside from this being the obvious partition of the vertices, the available data of [50] suggests that for q odd, the absolute vertices are all contained in a maximum independent set.

3.12 Polarities and Quotient Graphs

The Erdős-Rényi graphs are one example of a class of graphs that can be obtained using polarities. Let Y be a k-regular bipartite graph, and let σ be an automorphism of order two that swaps the colour classes of Y. The orbits of σ partition the vertices of Y. It is straightforward to show that this partition is equitable. Given two orbits of σ, C_i and C_j, define w_{ij} to be the number of edges from a vertex in C_i to the set of vertices of C_j. We define the *quotient graph* $X = Y/\sigma$ to have vertex set equal to the set of orbits of σ with w_{ij} arcs from C_i to C_j. We say that σ is a *polarity* of Y, and that X is a *polarity graph*. A vertex of X with a loop is said to be an *absolute vertex*; this corresponds to σ swapping two adjacent vertices of Y.

In a more general context, a quotient graph will be a weighted digraph. Since Y is bipartite, it follows that w_{ij} is either 0 or 1; also, $w_{ij} = w_{ji}$, because all cells of the partition have size 2. Moreover, since Y is k-regular, so is X.

We can use the quotient structure to obtain information about the eigenvectors. The eigenvectors of X are obtained from the eigenvectors of Y that are constant on each cell. In particular, the eigenvalues of X are exactly the eigenvalues of Y that correspond to eigenvectors that are constant on each cell (see [29, Section 5.1-2] for more details on quotient graphs).

We note one small result in this context, which is likely well-known.

3.12.1 Lemma. *Let Y be a bipartite graph, σ be an automorphism of Y of order two that swaps the colour classes, and $X = Y/\sigma$. If X contains a cycle of length $2m$ then so does Y.*

Proof. Let $C_1, \ldots, C_{2m}$ be the vertices of X that form a cycle of length $2m$, in that order. Each of these corresponds to a pair of vertices in Y. Let C_i correspond to the vertices x_i, y_i in Y, where all of the x_i are in one colour class. If $C_i \sim C_j$ in X, then $x_i \sim y_j$ and $x_j \sim y_i$ in X. Thus

$$x_1 \sim y_2 \sim x_3 \sim \ldots \sim y_{2m} \sim x_1$$

and

$$y_1 \sim x_2 \sim y_3 \sim \ldots \sim x_{2m} \sim y_1,$$

giving two cycles of length $2m$ in X. □

Notice that the cycles we obtain in Y are not necessarily induced, even if the cycle in X was: this can arise if one of the vertices in X has a loop on it, i.e., σ swaps two adjacent vertices of Y. If we try to apply this proof to odd cycles, then we discover that a cycle of length $2m+1$ in X gives a cycle of length $4m + 2$ in Y. The converse of this lemma is not true. A cycle in Y may give rise to a cycle of varying lengths in X: this is because a cycle in Y may contain exactly one vertex of an orbit of σ.

The Erdős-Rényi graphs can be obtained as quotient graphs in a straightforward manner. Let Y be the incidence matrix of $PG(2, q)$. This has as vertex set the set of all points and all lines of $PG(2, q)$, with two vertices being adjacent if the corresponding objects are incident in the geometry. Clearly this is a bipartite graph. Each point p can be represented by a vector $(p_1, p_2, p_3)^T$ in $GF(q)^3$, and each line l can be represented by a vector (l_1, l_2, l_3) in $GF(q)^3$. The point p and the line l are incident is $lp = 0$. Let σ be the map such that

$$\sigma(x_1, x_2, x_3)^T \leftrightarrow (x_1, x_2, x_3).$$

Then X/σ is exactly $ER(q)$ with a loop on each absolute vertex.

Polarities of $PG(2, q)$

We review some known facts about polarities in order to show that our bounds on the Erdős-Rényi graphs apply to a larger family of graphs. The derivation of the eigenvalues we offer is part of the folklore.

Let X be the incidence graph of $PG(2, q)$, and let σ be a permutation of order two of Y that swaps points and lines. Let N be the incidence matrix of $PG(2, q)$, with rows indexed by the points and columns indexed by the lines. Furthermore, order the columns of N so that the line corresponding to the i-th column is the image under σ of the point corresponding to the i-th row. Thus σ is an automorphism, and hence a polarity, if and only if the matrix N is symmetric. An absolute vertex corresponds to a 1 on the diagonal of N.

Let σ be a polarity of Y, and let $X = Y/\sigma$. Then N is the adjacency matrix of X and

$$\begin{pmatrix} 0 & N \\ N & 0 \end{pmatrix}$$

is the adjacency matrix of Y. Let a be the number of absolute vertices of X. Any point lies on exactly $q+1$ lines, and every two points lie on exactly one common line. It follows that

$$N^2 = (q+1)I + (J - I) = qI + J \tag{3.13}$$

The eigenvalues of N^2 are $q + q^2 + q + 1 = (q+1)^2$ and q, with multiplicities 1 and $q^2 + q$. It is not hard to guess the eigenvectors required to justify these claims, but one can also regard this a statement about the matrix of eigenvalues of an association scheme with one class.

It follows that the eigenvalues of N are $q+1$, $\sqrt{q}$ and $-\sqrt{q}$, with multiplicities 1, m_1, m_2. The sum of the multiplicities is the number of vertices of $PG(2,q)$, and the sum of eigenvalues according to their multiplicities is the trace of N, which is the number of absolute vertices: denote this by a. It is well-known that a unitary polarity has $q\sqrt{q}+1$ absolute vertices.

So we have the following two equations:

$$\begin{aligned} 1 + m_1 + m_2 &= q^2 + q + 1 \\ (q+1) + m_1\sqrt{q} - m_2\sqrt{q} &= a. \end{aligned}$$

We can solve these to get the multiplicities, although we don't actually need the multiplicities. What we really want is the least eigenvalue, which we now know is $-\sqrt{q}$, as long as $m_2 \neq 0$. Note that $m_2 = 0$ if and only if $a = q^2\sqrt{q} + q\sqrt{q} + q + 1$. But this is greater than the total number of vertices, thus $m_2 \neq 0$. It is in fact well known that a unitary polarity of $PG(2,q)$ has $q\sqrt{q}+1$ absolute vertices, but we don't actually need to know this. The point is that we can now apply Corollary 3.6.3 to the graph X. This gives a bound on the size of an independent set in the non-regular graph obtained by deleting the loops from X. It is in fact exactly (3.9). Note that if we simply apply the ordinary ratio bound directly and then add the number of absolute vertices, we get a weaker bound, exactly as we did in Section 3.11.

The reader should note that we still have not actually specified the graphs X or Y: the bound of (3.9) applies to *any* polarity graph of $PG(2,q)$. This includes the Erdős-Rényi graphs, but it also includes the graphs we obtain from a unitary polarity.

Generalized Quadrangles

We review some known properties of generalized quadrangles to show that our bounds on independent sets can be usefully applied to another family of graphs. The reader is directed to Payne and Thas [45] for more details. A *generalized quadrangle* with parameters (s,t) consists of a set of points, a set of lines, and an incidence relation between points and lines such that:

(a) Each point is incident with $t+1$ lines and two distinct points are incident with at most one common line.

(b) Each line is incident with $s+1$ points and two distinct lines are incident with at most one common point.

(c) Given a point p and a line l not incident with p, there exists a unique point q and a unique line m such that m is incident to p and q, and q is incident to m and l.

Consider the incidence graph of a generalized quadrangle. This is the graph whose vertex set is the union of the points and lines of the geometry, with two vertices adjacent exactly when their corresponding structures are incident. This is clearly a bipartite semi-regular graph. The existence part of the third condition implies that the diameter is four; the uniqueness part implies that the girth is eight. In fact, an incidence structure is a generalized quadrangle if and only if its incidence graph is bipartite semi-regular, and has diameter four and girth eight.

We will be interested in the generalized quadrangles $W(q)$. These have $s = t = q$, and they have $(q+1)(q^2+1)$ points and the same number of lines (see [45, p.37]). Let σ be a polarity of $W(q)$, and let $X = W(q)/\sigma$.

We will determine the eigenvalues of X. Our argument is essentially that of [45, 1.8.2]; alternatively, it can be thought of as a more general form of the derivation in the previous subsection. As before, if we order the points and lines according to the polarity σ, then the adjacency matrix of Y has the form

$$\begin{pmatrix} 0 & N \\ N & 0 \end{pmatrix},$$

where N is the (symmetric) incidence matrix of the generalized quadrangle, and hence N is the adjacency matrix of X. Let A be the matrix with rows and columns indexed by the points of the generalized quadrangle, and whose (i,j) entry is 1 if the distinct points i and j lie on a common line and 0 otherwise (we obtain the same matrix if we interchange the words "point" and "line"). We have

$$N^2 = (q+1)I + A. \tag{3.14}$$

The matrix A is the adjacency matrix of what is sometimes called the *point graph* of the geometry. Recalling (3.13), we see that $J - I$ is adjacency matrix for the point graph of $PG(2,q)$. In our present case, we will have to do slightly more work to obtain the eigenvalues of A.

The matrix A^2 is determined by the walks of length two; accordingly, we have

$$A^2 = (q+1)qI - (q-1)A + (q+1)(J-I-A) = (q^2-1)I - 2A + (q+1)J. \tag{3.15}$$

The vector $\mathbf{1}$ is an eigenvector of A corresponding to the eigenvalue $q(q+1)$, of multiplicity 1. The remaining eigenvectors are orthogonal to $\mathbf{1}$, and it follows

that their eigenvalues are roots of the equation

$$\lambda^2 + 2\lambda - (q^2 - 1) = 0.$$

Thus the remaining eigenvalues of A are $-1 \pm q$. Since we know the trace of A, we can compute the multiplicities. Again, we don't need them, except to notice that they are non-zero. The reader might notice that (3.15) implies that A is a strongly regular graph, and so what we have just done amounts to computing the matrix of eigenvalues for an association scheme with two classes.

Knowing the eigenvalues for A, we compute the eigenvalues for N^2 using (3.14): they are $(q+1)^2$, $2q$ and 0; their multiplicities are also known. The possible eigenvalues for N are then $q+1$, $\pm\sqrt{2q}$ and 0. Again, we can determine their multiplicities from the previous multiplicities and the trace of N. The trace of N is equal to q^2+1 (see [45]), which is the number of absolute points of the polarity: this is true for any polarity of $W(q)$. The multiplicity of $-\sqrt{2q}$ is not zero, hence it is the least eigenvalue. Now we can apply our generalized ratio bound.

3.12.2 Lemma. *Let X be the graph obtained by deleting the loops from a polarity graph of $W(q)$, and let S be an independent set of size s in X. Then*

$$s \leq \frac{\sqrt{2q} + \sqrt{2q + 4(q^2+1)\frac{q^2+q+\sqrt{2q}}{q^3+q^2+q+1}}}{2\frac{q^2+q+\sqrt{2q}}{q^3+q^2+q+1}}.$$

Proof. We apply Corollary 3.6.3 to the polarity graph of $W(q)$ (i.e., the graph X with loops added back in). Its least eigenvalue is $-\sqrt{2q}$, its valency is q^2+q and it has q^3+q^2+q+1 vertices. In this context we are bounding an independent set that contains s_1 loops. We set $s_1 = q^2+1$, the number of absolute vertices to obtain a bound independent of s_1. □

3.13 Graph Products

The results of this section are due to Alon, Dinur, Friedgut and Sudakov [2]. They deal with ratio-tight independent sets in graph products. They also derive an important result on independent sets that come close to meeting the ratio bound; we will not treat that result here. We show how their result fits in with our approach to ratio-tight maximum independent sets, and give an alternative proof.

By the *product* $G_1 \times G_2$ of two graphs G_1 and G_2, we mean the graph whose vertex set is the Cartesian product of the two vertex sets, with (x_1, x_2) and (y_1, y_2) being adjacent exactly when $x_1y_1 \in E(G_1)$ and $x_2y_2 \in E(G_2)$. We will denote this graph as $G_1 \times G_2$. We will write G^r for the product of r copies of G.

The following result is proved in [2].

3.13.1 Theorem. *Let G be a k-regular graph and S a ratio-tight independent set in G. Let $G^{\times r}$ be the product of r copies of G. Then $G^{\times r}$ has an independent set that is ratio-tight, and furthermore for any ratio-tight independent set R in $G^{\times r}$ there exists a ratio-tight independent set S in G such that R is the Cartesian product of S with $r-1$ copies of $V(G)$, in some order.* □

The eigenvectors of $G^{\times r}$ are exactly tensor products of r eigenvectors of G. It follows that the ratio of the greatest to the least eigenvalue is the same for G and for $G^{\times r}$. Furthermore, given any independent set S in G, the Cartesian product of S with $r-1$ copies of $V(G)$ is certainly an independent set in $G^{\times r}$, and this set is ratio-tight in $G^{\times r}$ if S is ratio-tight in G.

Thus the significance of this theorem lies in the "only if": *all* the ratio-tight independent sets of $G^{\times r}$, and hence all the maximum independent sets, are of this form. This fits in quite naturally with the main results of this thesis: Chapter 4, Chapter 5 and Chapter 6 all deal with characterizing the ratio-tight independent sets in specific families of graphs.

The approach used in [2] is based on the discrete Fourier transform. We offer a proof of Theorem 3.13.1 cast in our language. Furthermore, we will show that the general result follows from the specific case of complete graphs. Accordingly, we first prove the theorem for complete graphs.

3.13.2 Lemma. *Let $(K_n)^{\times r}$ be the product of r copies of the complete graph on n vertices. Then the ratio bound holds with equality for $(K_n)^{\times r}$, and every ratio-tight independent set in $(K_n)^{\times r}$ is the Cartesian product of a single vertex of K_n with $r-1$ copies of $V(K_n)$, in some order.*

Proof. First we note that the eigenvalues of K_n are $n-1$ and -1, with multiplicities 1 and $n-1$, respectively. The ratio bound is then 1, and so it holds with equality. It follows that the ratio bound holds with equality for $(K_n)^{\times r}$ as well.

Furthermore, for K_n, the sum of the eigenspaces corresponding to $n-1$ and -1 is $\mathbb{R}^n$. Thus a matrix whose columns span the sum of these two eigenspaces is the identity matrix. Eigenvectors of $(K_n)^{\times r}$ are tensor products of r eigenvectors of K_n. Eigenvectors for the greatest eigenvalue are obtained by taking r copies of $\mathbf{1}$; eigenvectors for the least eigenvalue are obtained by taking $r-1$ copies of $\mathbf{1}$ and one vector orthogonal to $\mathbf{1}$.

Let H_i be the tensor product of I with $r-1$ copies of $\mathbf{1}$, such that I appears in position i. Let H be the matrix whose columns are the set of all columns of all H_i, $1 \leq i \leq r$. Then the columns of H span the sum of the n^r-eigenspace and the $-n^{r-1}$-eigenspace. Note that the columns of H are not independent: the vector $\mathbf{1}^{\times r}$ lies in the column space of each H_i.

By the equality condition of the Theorem 3.4.1, every ratio-tight independent set in $(K_n)^{\times r}$ must lie in the column space of H. We now show that the only 01-vectors in the column space of H are the columns of H. Let $z = Hy$ be a 01 vector in the column space of H.

Given any indices $j_1, \ldots, j_r$ with $1 \leq j_l \leq n$, there is a row of H that contains a 1 in exactly the columns $j_1, \ldots, j_r$. Let $j_2, \ldots, j_r$ be fixed and vary j_1. The product of the corresponding row of H with y must always be zero or one, and thus the components of y corresponding to the columns of H_1 take on at most two distinct values. Now assume that y is not constant on the columns of H_1 and not constant on the columns of H_2. Then fixing $j_3, \ldots, j_r$ and letting j_1 and j_2 vary, we see that Hy takes on at least three distinct values. Since Hy is a 01 vector, this is impossible. This argument applies equally well to any of the H_i.

Thus we can partition y as

$$y = \left({y_1}^T {y_2}^T \ldots {y_r}^T\right)^T,$$

where for one index m, the vector y_m takes on at most two values and all other vectors y_i are multiples of $\mathbf{1}$. Now $H_i\mathbf{1}$ is a multiple of $\mathbf{1}$ for any i, so it follows that y_m is not constant. Thus $z = Hy$ is a linear combination of the columns of H_m and $\mathbf{1}^{\otimes r}$. In fact, since $H_i\mathbf{1}$ is independent of i, we have

$$z = H_m \sum_{l=i}^{r} y_i$$

Note that the rows of H_m are exactly r copies of the rows of I, so in fact z can be partitioned as

$$z = \left({z_1}^T {z_2}^T \ldots {z_r}^T\right)^T,$$

where the z_i are all equal.

Now further assume that z is the characteristic vector of a ratio-tight independent set in $(K_n)^{\times r}$. Then z_m is the characteristic vector of a ratio-tight set in K_n, and the theorem follows. □

Somewhat surprisingly, we will use this argument again in our proof of the q-analogue of Erdős-Ko-Rado in Section 5.5.

We now show how Lemma 3.13.2 implies Theorem 3.13.1.

Let X be some graph for which the ratio bound holds with equality, and $r > 1$. Let M be a matrix whose columns span the eigenspaces corresponding to the greatest and least eigenvectors of X. As in the proof above, define the matrices N_i to be the product of one copy of M with $r-1$ copies of $\mathbf{1}$, with the M in the i-th position. Let N be the matrix whose columns are the set of all columns of all N_i, $1 \leq i \leq r$. Then as in the above proof we have that the characteristic vector of a ratio tight set of $X^{\times r}$ lies in the column space of N. But the column space of N_i lies in the column space of H_i. Thus any 01 vector in the column space of N must lie in the column space of H. Since the intersection of the column spaces of any two distinct H_i or any two distinct N_i is spanned by $\mathbf{1}$, the result follows.

We can make a small extension to Theorem 3.13.1. Notice that at no point in either proof did we use any information about the graph other than the number of vertices and the ratio between its valency and its least eigenvalue. This is also

true of the proof in [2]. We now observe a consequence of this. Let $X_1, \ldots, X_r$ be graphs with common valency k and common least eigenvalue τ. If the ratio bound holds with equality in each of them, then we may apply a trivial variation of Theorem 3.13.1 to the graph

$$X_1 \times \cdots \times X_r.$$

The proof is completely analogous.

3.14 A Bound of Haemers

Haemers [34] had previously obtained a generalization of the ratio bound to non-regular graphs.

The following bound appears without proof in [11] and [15], apparently due to Haemers [34]. The original version with proof is apparently in Haemers' Thesis.

3.14.1 Lemma. *Let G a k-regular graph on v vertices with eigenvalues $k = \lambda_1 \geq \ldots \geq \lambda_v$. Let H some subgraph of G on v_1 vertices of average degree d_1. Then*

$$\frac{v_1(k-\lambda_v)}{v} + \lambda_v \leq d_1 \leq \frac{v_1(k-\lambda_2)}{v} + \lambda_2. \qquad \square$$

If we choose H to be an independent set, then $d_1 = 0$ and we recover the bound of the ratio bound. This proof is based on interlacing; a more publicly available version of Haemers' result appears with proof in Godsil and Royle [33, Section 9.6].

The following result is also obtained using a similar interlacing argument [34].

3.14.2 Lemma. *Let G be any graph on v vertices with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_v$ and minimum degree δ. Let S be an independent set of size s in G. Then*

$$s \leq v\frac{-\lambda_1\lambda_v}{\delta^2 - \lambda_1\lambda_v}. \qquad \square$$

If G is a regular graph, we again recover the ratio bound. We are motivated to compare this bound with the bounds for non-regular graphs we derived earlier.

It is not in general equal to any of the bounds we determine. Recall that in Section 3.9 we compared some of our bounds. For the complete bipartite graph of Example 3.9.1, Haemers' bound is equal to the bound of Corollary 3.6.5, which is optimal, and better than the bound of Corollary 3.6.2. For the graphs

$$\overline{K_m \cup \overline{C_{2m+1}}},$$

of Example 3.9.2, we showed that our bounds took on two different values depending on which of two types of independent sets one considered. Haemers' bound lies between them, and is asymptotically equal to the larger of these bounds.

We also computed the values of Haemers' bound for the Erdős-Rényi graphs. We obtain values that lie between (3.12) and (3.8).

3.15 Recent Developments

In the above analysis, we mostly deferred the general approach outlined at the beginning of Section 3.5. We were able to obtain results using a restricted approach, but it was clear that there was more to be said. We close this chapter with some comments on further generalizations of the ratio bound; this is the subject of ongoing research.

Let us start by applying directly the steps outlined at the beginning of Section 3.5.

3.15.1 Lemma. *Let S be an independent set of size s in a graph on n vertices. Let B be some positive semi-definite matrix such that $B_{ij} = 0$ when $i \not\sim j$. Then*

$$\sum_{i \in S} B_{ii} - 2\frac{s}{n} \sum_{i \in S, j \in V} B_{ij} + \frac{s^2}{n^2} \sum_{i,j \in V} B_{ij} \geq 0.$$

Proof. The matrix B is positive semi-definite, so we have:

$$\begin{aligned} 0 &\leq \left(z - \frac{s}{n}\mathbf{1}\right)^T B \left(z - \frac{s}{n}\mathbf{1}\right) \\ &= z^T z - 2\frac{s}{n} z^T B\mathbf{1} + \frac{s^2}{n^2}\mathbf{1}^T B\mathbf{1} \\ &= \sum_{i \in S} B_{ii} - 2\frac{s}{n} \sum_{i \in S, j \in V} B_{ij} + \frac{s^2}{n^2} \sum_{i,j \in V} B_{ij}. \end{aligned}$$

□

This seems very difficult to handle. It would be a quadratic inequality in s, were it not for the fact the the coefficients of the "linear" and "constant" terms are non-trivial functions of S.

In Section 3.7, we saw how the Laplacian bound seemed to capture more of the structure than the adjacency bound. One reason for this is that $\mathbf{1}$ is always an eigenvector of the Laplacian matrix, whereas it is only an eigenvector of A for regular graphs. This suggests that we should choose B so that $\mathbf{1}$ is an eigenvector of B. Note that this is not equivalent to saying that the graph is regular.

On the other hand, we can relax one of the other conditions on B: a careful reading of the proof of Lemma 3.15.1 reveals that it is sufficient to have $B_{ij} \leq 0$ when $i \not\sim j$.

Let B be some positive semi-definite matrix such that $B_{ij} \leq 0$, and in addition, $B\mathbf{1} = r\mathbf{1}$. If we further define

$$\overline{b}_s = \frac{1}{s} \sum_{i \in S} B_{ii},$$

then we are pleasantly surprised to discover that Lemma 3.15.1 simplifies quite nicely.

3.15.2 Lemma. *Let S be an independent set in some graph, and let B be a positive semi-definite matrix such that $B_{ij} \leq 0$ whenever $i \not\sim j$, and $B\mathbf{1} = r\mathbf{1}$. Let $\overline{b_S}$ be the average of B_{ii} as i runs over S. Then*

$$\frac{s}{n} \leq \frac{\overline{b_S}}{r}.$$ □

Note that we may assume, for instance, that $r = 1$ by an appropriate scaling of B.

We will denote the trace of B by $\mathrm{tr}(B)$, and the sum of all entries of B as $\mathrm{sum}(B)$. We obtain the following bound.

3.15.3 Lemma. *Let S be an independent set in some graph, and let B be a positive semi-definite matrix such that $B_{ij} \leq 0$ whenever $i \neq j, i \not\sim j$, $B\mathbf{1} = r\mathbf{1}$, and $B \circ I$ is a positive multiple of I. Then*

$$\frac{s}{n} \leq \frac{\mathrm{tr}(B)}{\mathrm{sum}(B)}$$

Proof. If $B \circ I$ is a positive multiple of I then $v\overline{b}_S = \mathrm{tr}(B)$. If $B\mathbf{1} = r\mathbf{1}$ then $vr = \mathrm{sum}(B)$. Then using Lemma 3.15.2, we have

$$\frac{s}{n} \leq \frac{\overline{b_S}}{r} = \frac{\mathrm{tr}(B)}{\mathrm{sum}(B)}.$$ □

Furthermore, if equality holds then $z - \frac{s}{n}\mathbf{1}$ is an eigenvector for $B - B \circ I$.

Note that by scaling B we can choose $\mathrm{tr}(B) = 1$, or $\mathrm{sum}(B) = 1$ if we wish,

This result is interesting for the following reason. Godsil [30] has shown this result in the special case where we restrict ourselves to graphs in an association scheme, and restrict B to be in the algebra of the association scheme; furthermore, in this case, it encompasses the linear programming bound of Delsarte that we saw in Lemma 3.2.1. Thus we have extended this result to arbitrary graphs.

We will defer further investigation of Lemma 3.15.1, Lemma 3.15.2 and Lemma 3.15.3 to a future work. In particular, we need to determine the relationship of these generalizations to the ϑ-bound and ϑ'-bound of Section 3.3.

Chapter 4

Partition Graphs

4.1 Introduction

Our motivation for considering partition graphs is an application to covering arrays. Meagher and Stevens asked [42] whether the graph $\mathcal{P}(3^3)$ of Example 2.4.2 is a core. We are able to resolve this question, and our solution gives a general method which we use throughout the remainder of this thesis. We approach the problem in two ways. The first is based on a fairly direct analysis of endomorphisms of this graph, using as a prototype a known argument that shows that the Petersen graph is a core. This is in some sense the obvious approach. The second is less obvious, but yields a method with broader applications. We are able to characterize the maximum independent sets, and leverage this into a proof that the graph is a core. It is, in the end, knowledge of the maximum independent sets that is more fundamental.

We start off by giving a brief introduction to covering arrays, as this is the application that motivated our result. We make no attempt at completeness here, offering only enough to explain our own interest. We then give some standard background on graph cores. We find we are able to extend some of the tools, to obtain new results relating a graph and its core. We record these here as they may find uses elsewhere. Using the results on cores outlined, we then give a proof based on a rather direct analysis that the graph $\mathcal{P}(3^3)$ is a core.

For our second proof we take a somewhat different viewpoint. There is an obvious family of maximum independent sets, and we are able to prove that these are in fact the only ones. This result tells us more about the structure of the graph than the previous argument. For instance, it proves that it is not four-colourable, a fact which had previously only been known as a result of a computer search.

4.2 Covering Arrays

A *covering array* of size n, alphabet size g, and k rows is a $k \times n$ array A with entries in $\mathbb{Z}_g$ that satisfies the following property. For any two rows i_1, i_2, and any two $a, b \in \mathbb{Z}_g$, there exists a column j such that $A_{i_1 j} = a$ and $A_{i_2 j} = b$. In other words, if we take any two rows, then all ordered pairs of elements of $\mathbb{Z}_g$ appear in some column.

This is a combinatorial structure that has applications in software testing.

Consider the following scenario. A software package consists of different pieces. These can be tested individually, but it is also critical that they be tested in combination with one another. Furthermore, each of these pieces may, during its execution, be in several different states, and it is desirable to test different combinations of states as well. At first glance, the obvious solution is to test all modules together in all possible states. This is in some sense ideal, but can be time-consuming and costly. For practical reasons, we might be content with a series of tests such that at some point each pair of modules is tested together in all possible pairwise combinations. This amounts to assuming that there are no multiple interactions between the different components that cannot be explained by some set of pairwise interactions. We can test all the pieces together at once, we just don't have time to examine all the possible combinations of all the pieces.

We then need to run a series of "tests", where each test corresponds to a particular setting of states for each the components. We would like every pair of components to be run together in every possible combination of states.

We can model this with a covering array. The rows of the array correspond to the components to be tested. Each column corresponds to an assignment of states to each component. Considering any two components (rows), we see that as we run through all the tests (columns), we will cover all possible combinations of states (ordered pairs of elements from $\mathbb{Z}_g$). For a given k and g, the goal is to find a covering array of smallest size. Here is an example of a covering array with 3 rows, size 3, on an alphabet of size 3:

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 \end{pmatrix}$$

We have tacitly assumed that all components have the same number of states. If this is not true, we can take g to be the largest number of states among all the components, and then renumber entries that correspond to non-existent states for a particular component. If it turns out that the second component only has two states, we can simply change the 2's in that row to anything else. There may (or may not) be a way to do this so as to make one of the columns redundant, but no matter how we do it, we still have all the possible combinations. For instance, we could change the 2's to 0's:

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 0 & 1 & \mathbf{0} & 0 & 1 & \mathbf{0} & 0 & 1 & \mathbf{0} \\ 0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 \end{pmatrix}$$

Note that in this case we can't possibly reduce the number of columns, even if the second component only had one state. We need at least nine columns to cover all the combinations between the first and third column.

We have a further generalization, studied by Meagher and Stevens [42]. The rows of the covering array correspond to vertices in some graph G. We relax the condition on the array so that for every two rows i_1, i_2 that are adjacent in the graph, and for any $a, b \in \mathbb{Z}_g$, there is a column j such that $A_{i_1 j} = a$ and $A_{i_2 j} = b$. This is known as a *covering array on a graph*.

From the point of view of our application, we are assuming that nonadjacent rows correspond to components that do not need to be tested together. This may be because they are known to not interact, were already tested last year, or maybe because the budget ran short. Note that the definition we gave above for a covering array corresponds to a covering array on a clique. The question now becomes, for which graphs G is a minimum covering array on G smaller than a minimal covering array on a clique on the same number of vertices as G. In other words, does knowing that some combinations need not be tested mean that we can reduce the total number of tests required?

In this context, the following family of graphs arises naturally (they are described in some detail in [42]). They are important because we can use them to obtain bounds on the size of covering arrays.

4.2.1 Definition. Let $n \geq g^2$. The graph $CAG(n, g)$ has as vertex set all vectors in $\mathbb{Z}_g^n$ with the property that each element appears at least g times and the first occurrence of j occurs before the first occurrence of $j + 1$ for $0 \leq j < g - 1$. Two vertices u, v are adjacent when the corresponding vectors have the property that for any $a, b \in \mathbb{Z}_g$ there is a j such that the j-th entry in the vectors corresponding to u and v are, respectively, a and b. □

For our purposes, it will be more useful to think of them in a slightly different way. We consider a related family of graphs.

4.2.2 Definition. Let n, g be given with $n \geq g^2$. Let $n = \sum_i^g n_i$ where $n_i \geq g$. The graph $\mathcal{P}(n_1, \ldots, n_g)$ has as vertex set all partitions of $\{1, \ldots, n\}$ into g cells of sizes $n_1, \ldots, n_g$. Two vertices u, v are adjacent when the set of intersections of one cell from u and one cell from v has size g^2. □

More prosaically, two vertices are adjacent when their partitions are *skew*, that is, each cell of the partition of one vertex intersects each cell of the partition of the other.

The graph $CAG(n, g)$ has as a spanning subgraph the disjoint union of the graphs $\mathcal{P}(n_1, \ldots, n_g)$ over all $n_1, \ldots, n_g$ such that $n = \sum_i^g n_i$ and $n_i \geq g$ ($CAG(n, g)$ has other edges between the different $\mathcal{P}(n_1, \ldots, n_g)$). In particular, the graphs $CAG(g^2, g)$ and $\mathcal{P}(g, \ldots, g)$ are the same; we shall abbreviate the latter as $\mathcal{P}(g^g)$. The condition on first occurrences in the definition of $CAG(n, g)$ may seem somewhat odd; it is to ensure that a given partition is only counted once, and not once for each possible ordering.

We will use the notation "$G \to H$" to mean "there exists a homomorphism from G to H". It is straightforward to show that if $G \to H$ then, for fixed g, the minimal size of a covering array on G is bounded above by the minimal size of a covering array on H. In fact we can say more. The following result is well-known (a proof appears in [42]).

4.2.3 Lemma. *There exists a covering array on G of size n and alphabet size g if and only if there is a homomorphism from G to $CAG(n, g)$.* □

So the question of finding a covering array is closely related to finding a homomorphism into some graph $CAG(n, g)$. Notice that if $CAG(n, g)$ could be mapped homomorphically onto some proper subgraph X, then there is a homomorphism $G \to CAG(n, g)$ if and only if there is a homomorphism $G \to X$. In this case, we would rather look for homomorphisms into X, as X is smaller than $CAG(n, g)$. This leads naturally to the question of whether or not $CAG(n, g)$ is a core, as we will see in the next section.

We will focus principally on the graph $\mathcal{P}(3^3)$ (which is the same as $CAG(9, 3)$). In this case, and more generally, whenever $n = g^2$, skew means that the meet of the two partitions is discrete, or equivalently that every cell of one partition is completely separated by the other.

More details on covering arrays may be found in [49, 42]; further applications of covering arrays are given in [14].

4.3 Graph Cores

A graph X is a *core* if every endomorphism of X is an automorphism; equivalently if there are no homomorphisms from X to a proper subgraph of X. For example, K_2 is easily seen to be a core, as is any odd cycle. Every bipartite graph other than K_2 can be mapped homomorphically onto one of its edges, and hence is not a core.

Two graphs X, Y are *homomorphically equivalent* if there are homomorphisms $X \to Y$ and $Y \to X$. Every bipartite graph is homomorphically equivalent to K_2, while no two odd cycles of different sizes are homomorphically equivalent. For a graph X, consider the set of subgraphs of X to which X is homomorphically equivalent. This list is not empty, since it contains X, and it is finite, so it must have minimal elements. These subgraphs must themselves be cores, and are referred to as *cores of* the graph X.

A *retraction* from X to a subgraph Y is a homomorphism from X to Y that fixes each vertex of Y. Y is a *retract* of X is there is a retraction $X \to Y$. Clearly K_2 is a retract of any bipartite graph.

Known Tools

The following results are basic to the theory of graph cores; proofs can be found in Godsil and Royle [33, Chapter 6] or Hahn and Tardif [36].

4.3.1 Lemma. *Let Y be a core of X. Then Y is a retract of X, and hence an induced subgraph of X.* □

4.3.2 Lemma. *Let X be any graph. Then X has a core which is unique up to isomorphism; we denote the core of X by $X^\bullet$.* □

4.3.3 Lemma. *Any retraction from a graph X to a proper subgraph Y maps two vertices at distance two to a common vertex.* □

For a homomorphism $\psi : X \to X^\bullet$, we define the fibre of a vertex y of $X^\bullet$ to be the set of vertices of X that are mapped to y by ψ, and denote this set by $\mathrm{fib}(y)$. Note that $\mathrm{fib}(y)$ is an independent set in X. The proofs of the following three results are based on Godsil and Royle [33]; we summarize them here for later comparison.

4.3.4 Lemma. *If X is vertex transitive, then so is $X^\bullet$.*

Proof. Let $u, v \in V(X^\bullet)$. Since X is vertex transitive, there is some automorphism of X, ϕ, such that $\phi(u) = v$. By Lemma 4.3.1 there is exists a retraction $\psi : X \to X^\bullet$. Thus $\psi \circ \phi$ is an automorphism of $X^\bullet$ that maps u to v. □

4.3.5 Lemma. *Let X be a vertex transitive graph. Let y be a vertex of $X^\bullet$. If $\psi : X \to X^\bullet$, then $|\mathrm{fib}(y)|$ is independent of choice of y in $X^\bullet$.*

Proof. Let ϕ be an automorphism of X, and let ψ be a retraction $X \to X^\bullet$. Then the image of $X^\bullet$ under ϕ must meet each fibre of $X^\bullet$, or else the composition $\psi \circ \phi$ would be a homomorphism of $X^\bullet$ onto a proper subgraph.

Now let $y \in V(X^\bullet)$ and $y' \in \mathrm{fib}(y)$. The number of automorphisms ϕ of X such that $\phi^{-1}(y') \in X^\bullet$ is independent of y', since X is vertex transitive. But this quantity multiplied by $|\mathrm{fib}(y)|$ gives $|\mathrm{Aut}(X)|$. Thus $|\mathrm{fib}(y)|$ is independent of y. □

4.3.6 Corollary. *If X is vertex transitive, then $|V(X^\bullet)|$ divides $|V(X)|$* □

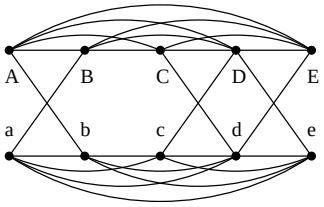
New Extensions

In fact, the above can be generalized to other transitivities using the same proof ideas. Let H be some graph. A graph X is transitive on H if given any two copies H_1, H_2 of H in X, there is an automorphism of X that induces an isomorphism from H_1 to H_2. Furthermore, let (H, h) be a pair where H is some graph and h a distinguished vertex of H, call such a pair a *rooted H*. Two rooted H graphs (H_1, h_1) and (H_2, h_2) are isomorphic if there is an graph isomorphism from H_1 to H_2 that maps h_1 to h_2. A graph is transitive on rooted H if given any two copies $(H_1, h1), (H_2, h_2)$ of (H, h) in X, there is an automorphism of

X that induces an isomorphism from (H_1, h_1) to (H_2, h_2). For example, for a graph to be transitive on K_2 is the same as being edge-transitive; for a graph to be transitive on rooted K_2 is the same as being arc-transitive.

For a homomorphism $\psi : X \to X^\bullet$ define the fibre of a subgraph H of $X^\bullet$ to be the set of copies of H in X that are mapped to H by ψ. So for example, the fibre of a particular pentagon H in $X^\bullet$ is the set of pentagons in X that are mapped to H. This is not the same thing as the set of vertices that get mapped to vertices of H. First of all, the fibre of H is a collection of subgraphs isomorphic to H, not a collection of vertices. Secondly, there may well be vertices of X that get mapped to H without having anything to do with pentagons.

A picture may help to clarify this. The graph of Figure 4.1 has K_5 as its core, with the homomorphism mapping capital letters to lowercase ones an example of a retraction. Let H be the 4-cycle with vertices a, b, c, d, in that order. The fibre of H is the set of 4-cycles that are mapped to H. Thus the 4-cycle on the vertices A, b, c, D is an element of the fibre of H. The 4-cycle on the vertices A, B, a, b is not however, nor is the one on A, C, D, B.



fib($abcd$)	not	
$abcd$	$ACBD$	$AbdC$
$aBCd$	$ACDB$	$aBDc$
$ABCD$	$acbd$	$ABab$
$AbcD$	$acdb$	$CDcd$

Figure 4.1: Many 4-cycles, small fibre.

Similarly, we define the fibre of a rooted subgraph (H, h) of $X^\bullet$ to be the set of copies of H in X rooted at the (fixed) vertex h that are mapped to (H, h) by ψ. If in the previous example we consider H to be the 4-cycle a, b, c, d, rooted at a, then a, B, C, d, rooted at a is in the fibre, but A, B, C, D rooted at A and a, B, C, d rooted at B are not.

4.3.7 Lemma. *If X is transitive on H, then so is $X^\bullet$.*

Proof. The idea is the same as Lemma 4.3.4. Let H_1 and H_2 be copies of H in $X^\bullet$. Then there is an automorphism of X that maps H_1 to H_2; the composition with a retraction ψ gives an automorphism of $X^\bullet$ that maps H_1 to H_2. □

4.3.8 Lemma. *Let H be a subgraph of $X^\bullet$. If $\psi : X \to X^\bullet$, then $|\operatorname{fib}(H)|$ is independent of choice of H in $X^\bullet$.*

Proof. The idea is again similar to Lemma 4.3.5; we omit the details. □

4.3.9 Corollary. *If X is transitive on H, then the number of copies of H in $X^\bullet$ divides the number of copies of H in X.* □

4.3.10 Lemma. *If X is transitive on rooted H, then so is $X^\bullet$.*

Proof. Similar to Lemma 4.3.4. □

4.3.11 Lemma. *Let (H, y) be a rooted subgraph of $X^\bullet$. If $\psi : X \to X^\bullet$, then $|\operatorname{fib}(H, y)|$ is independent of choice of (H, y) in $X^\bullet$.*

Proof. We adopt the argument of Lemma 4.3.5.

Let ϕ be an automorphism of X and let ψ be a retraction $X \to X^\bullet$. Then no two distinct copies of (H, h) can be mapped into the same fibre, or else the composition $\psi \circ \phi$ would be many-to-one, and hence a homomorphism of $X^\bullet$ onto a proper subgraph.

Now let $(H, y) \subseteq X^\bullet$ and $(H', y) \in \operatorname{fib}(H, y)$. The number of $\phi \in \operatorname{Aut}_{\mathrm{y}}(X)$ such that $\phi^{-1}(H', y) \subseteq X^\bullet$ is independent of choice of (H', y). This quantity multiplied by $|\operatorname{fib}(H, y)|$ gives $|\operatorname{Aut}_{\mathrm{y}}(X)|$, which is independent of choice of (H, y). Thus $|\operatorname{fib}(H, y)|$ is independent of (H, y). □

4.3.12 Corollary. *Let x be a vertex of X such that there is a copy of (H, h) in X with $h = x$, and y be a vertex of $X^\bullet$ such that there is a copy of (H, h) in $X^\bullet$ with $h = y$. If X is (H, h) transitive then the number of copies of (H, h) in $X^\bullet$ with $h = y$ divides the number of copies of (H, h) in X with $h = x$.* □

As an application of Corollary 4.3.12, taking $H = K_2$ and h one of its vertices, we see that if X is arc transitive, then the degree of $X^\bullet$ divides the degree of X. We will see that Lemma 4.3.7 and the results following can be applied to the graph $\mathcal{P}(3^3)$.

Note that the divisibility criterion in Corollary 4.3.12 applies only to vertices where there are copies of (H, h). Being transitive on rooted triangles does not mean there is a constant number of triangles at each vertex, it means that there is a constant number of triangles at each vertex where there is a triangle. Thus Corollary 4.3.12 can be applied to graphs that are not even vertex transitive.

In general, many of the parameters of the core are related to those of the graph. We can use these types of results to rule out possibilities for cores. For instance, since the core of a graph is a retract, we can lift any colouring of the core to the graph. As the core is also an induced subgraph, we see that $\chi(X) = \chi(X^\bullet)$. One can also show that a graph and its core have the same clique number and fractional chromatic number.

We mention one further example of a core: the Petersen graph. One way to see this is the following. It is a triangle free graph but it contains pentagons. Thus any endomorphism must map pentagons to pentagons. In particular, the core must contain a pentagon. By Corollary 4.3.6, if it is not a core, then the core must have five vertices, and by Corollary 4.3.12, its degree must divide three. This is impossible, and so the Petersen graph is a core. The odd girth plays a crucial role here, and we will see plays a similar role in $\mathcal{P}(3^3)$.

Establishing that a graph is a core is a difficult question in general. The odd girth can play a role, as in the Petersen graph. There are certain special cases that can also be resolved: as we noted above, a graph and its core have the same chromatic number and so any vertex critical graph is automatically a core. But there is no known general method, aside from exhaustive search. A naive exhaustive search is hopeless, as it would involve checking on the order of v^v candidate homomorphisms.

4.4 $\mathcal{P}(3^3)$ Is a Core

We first collect together some elementary facts about $\mathcal{P}(3^3)$. It has 280 vertices, and is regular of degree $d = 36$. It is arc transitive, by virtue of the fact that the automorphism group contains Sym(9) acting on the underlying 9-set. Each neighbourhood consists of 12 triangles, and thus the maximum clique size is 4. Furthermore, every edge lies in a unique 4-clique.

This last remark is significant. There are many triangles in $\mathcal{P}(3^3)$, but they are all contained in 4-cliques. A *geometric cycle* has no three consecutive vertices in a common maximum clique of the graph. So there are no geometric triangles in $\mathcal{P}(3^3)$. There are geometric pentagons though. A geometric pentagon can be homomorphically mapped into a non-geometric triangle, but this would force all of the cliques on the edges of the pentagon to be mapped into the (unique) 4-clique on the triangle. We will see that this means that $\mathcal{P}(3^3)$ behaves in some ways like a triangle-free graph, forcing pentagons to be mapped to pentagons, much as in the above proof that the Petersen graph is a core.

4.4.1 Theorem. *The graph $\mathcal{P}(3^3)$ is a core.*

Proof. Let ψ be a homomorphism from $\mathcal{P}(3^3)$ to $\mathcal{P}(3^3)^\bullet$. Then by Lemma 4.3.3, there are vertices $z \sim v \sim w$ in $\mathcal{P}(3^3)$ such that $\psi(z) = \psi(w)$. Consider these vertices to be fixed.

Now if there are further vertices x, y such that v, w, x, y, z are the vertices of a geometric pentagon (in that order), then we see that the image of w, x, y, z under ψ is a triangle, which is contained in some unique 4-clique c of $\mathcal{P}(3^3)^\bullet$. Now each of the edges vw, wx, xy, yz is contained in a unique and distinct 4-clique of $\mathcal{P}(3^3)$; call these the basic cliques of the pentagon. The image of each of the basic 4-cliques under ψ meets c in at least an edge, therefore the image of each of these cliques is c.

Were there to be further vertices x', y' such that v, w, x', y', z were also the vertices of a geometric pentagon (in that order), a similar statement would apply

but the image clique in $\mathcal{P}(3^3)^\bullet$ might not be c. However, if we had that x' is a vertex of the clique containing wx, or y' a vertex of the clique containing yz, then the image clique in $\mathcal{P}(3^3)^\bullet$ would be c. Thinking of v, w, z as fixed vertices, call such a pair of geometric pentagons *overlapping*. Thus any overlapping pentagons must have all of their basic cliques mapped onto the same clique by ψ. This can be extended. Call two geometric pentagons P, P' *related* if there is a sequence $P_1, \ldots, P_t$ of geometric pentagons where $P = P_1$ and $P_t = P'$ and P_i overlaps P_{i+1}. Then ψ must map all of the basic cliques of P and P' (and all related P_i too) onto the same clique in $\mathcal{P}(3^3)^\bullet$.

This is now a computational problem. As $\mathcal{P}(3^3)$ is arc transitive, we may arbitrarily fix adjacent vertices v, w. Computer calculations using [39] shows that $\mathrm{Aut}(\mathcal{P}(3^3))_{v,w}$ partitions $N(v) \setminus N(w)$ into three orbits, so we need only consider three cases for z. Recall that $\mathcal{P}(3^3)$ is part of an association scheme with four classes: the three cases correspond to which of the other three relations contain the "edge" wz, more precisely, whether the meet of the partitions of the vertices w and z has size $7, 6, 5$.

Using [39], one finds the following. If the meet of z and w has size seven, then there is a total of 152 cliques contained in related geometric pentagons. Together, these 152 cliques contain a total of 213 vertices of $\mathcal{P}(3^3)$ that all get mapped to a the four vertices of a single 4-clique in the core. If the meet has size six, then there are 144 such cliques containing 169 vertices. If the meet has size five, then there are 96 such cliques containing 100 vertices. Thus, using Corollary 4.3.6, we see that the size of the fibre of a vertex is at least 25 and thus the core must have either four or eight vertices. Eight is impossible since the core must be a connected graph that consists of 4-cliques that pairwise have at most one vertex in common. Four is impossible since this would imply the graph is 4-chromatic, which has been shown not to be the case by computer search [42]. □

It can be shown [41] that $\mathcal{P}(3^3)$ is part of an association scheme. Define five relations $X_0, \ldots, X_4$ on pairs of vertices of $\mathcal{P}(3^3)$ as follows. Vertices are $0, 1, 2, 3, 4$ related if their meet has $3, 9, 7, 6, 5$ cells. Thus X_0 is the identity relation and X_1 is the adjacency relation of $\mathcal{P}(3^3)$. Since it is an association scheme, one determine numbers $p_{ij}(k)$ such that given any two vertices a, b that are k-related, there are $p_{ij}(k)$ vertices that are i-related to a and j-related to b. These numbers are computed in [41], and can be used to simplify many of our computations.

Comment

With the benefit of much later hindsight, we add one comment after proofreading the above proof. If we pick apart the calculations, what we have actually shown is that (in the terminology of the proof) all geometric pentagons are related. It seems likely that one could use the automorphism group of the graph to show this directly. The stabilizer group of an arc has order 36, which suggests that there may be enough flexibility in the automorphism group to construct

the sequence of pentagons mentioned in the proof. This might be a useful observation for extending this proof to more general partition graphs. The technical calculations in the above proof were essentially carried out by drawing out the graph and counting various types of paths. Even with a computer such a procedure can be tedious, not to mention difficult to generalize. Working with the automorphism group might be a way around this. We have not investigated this approach in detail at this time.

4.5 Maximum Independent Sets

We offer a second proof that $\mathcal{P}(3^3)$ is a core, based on maximum independent sets. It turns out that characterizing the maximum independent sets is sufficient to show that $\mathcal{P}(3^3)$ is a core. This has been submitted as [32]; we give here a sketch of the result.

By [41] we know the eigenvalues of $\mathcal{P}(3^3)$. So we can compute the ratio bound: it gives that the maximum size of an independent set is 70. Define the set S_{ij} $(i \neq j)$ to be the set of all vertices that have i and j in the same cell; clearly this is an independent set, and since it has size 70, it is maximum. As we shall see, there are no other maximum independent sets.

Define the matrix H to have as its columns the characteristic vectors of the sets S_{ij}. Thus the rows of H are indexed by vertices of the graph, the columns by pairs of elements, and the $(\alpha, \{i,j\})$-entry of H is one if i and j are in the same cell of α and zero otherwise. We know that the column space of H is contained in the sum of the d-eigenspace and the τ-eigenspace (the d-eigenspace is spanned by $\mathbf{1}$). We will show that equality holds.

4.5.1 Lemma. *The columns of $H - \frac{1}{4}J$ span the τ-eigenspace.*

Proof. We know that the columns of $H - \frac{1}{4}J$ lie in the τ-eigenspace, by Theorem 3.4.1 (or indeed, since we know τ from [41], we could just verify it by direct calculation). We also know from [41] that the multiplicity of τ is 27.

So we need to show that $\mathrm{rk}(H - \frac{1}{4}J) = 27$. Since $\mathbf{1}$ is in the column space of H and is orthogonal to all columns of $H - \frac{1}{4}J$, it will suffice to show that $\mathrm{rk}(H) = 28$. Recall that $\mathrm{rk}(H) = \mathrm{rk}(H^T H)$. Both the rows and columns of $H^T H$ are indexed by pairs of elements, and the $(\{i,j\},\{k,l\})$-entry of $H^T H$ is just $|S_{ij} \cap S_{kl}|$. But this last is determined entirely by $|\{i,j\} \cap \{k,l\}|$:

$$|S_{12} \cap S_{12}| = |S_{12}| = 70, \quad |S_{12} \cap S_{13}| = 10, \quad |S_{12} \cap S_{34}| = 20.$$

We can think of the rows and columns of $H^T H$ as being indexed by the vertices of $L(K_9)$, with the a, b entry being $70, 10, 20$ as the vertices a and b are, respectively, equal, adjacent, or not adjacent. Thus if we let L be the adjacency matrix of $L(K_9)$ then

$$H^T H = 70I + 10L + 20(J - I - L) = 50I + 20J - 10L.$$

Since the eigenvalues of L are known, we can then compute the eigenvalues of $H^T H$ and we see that zero is an eigenvalue of multiplicity eight, meaning that $H^T H$ has rank 28. □

The practical consequence of Lemma 4.5.1 is that, because the ratio bound holds with equality we know that the characteristic vector of any maximum independent set lies in the column space of H. There are the obvious candidates: the columns of H themselves. We will show that these are in fact the only ones, but first we will need one more technical result.

4.5.2 Lemma. *Any independent set of size* 70 *in* $\mathcal{P}(3^3)$ *must contain two vertices whose partitions have a meet of size* 7.

There are several different ways to establish this, one of which is given in [32]. Here, we offer a more direct proof based only on the intersection numbers of the association scheme, computed in [41].

Proof. Recall that $\mathcal{P}(3^3)$ is a graph in an association scheme, with classes $X_1 = \mathcal{P}(3^3), X_2, X_3, X_4$, and that adjacency in X_2 means having a meet of size 7. So we wish to show that it is not possible to have 70 vertices such that the only types of edges between them are edges from the graphs X_3 and X_4. Let S be such a set of vertices. If S induces a clique in X_4, then

$$|S| \leq 2 + p_{44}(4) = 10.$$

So there must be two vertices of S that are adjacent in X_3. But then

$$|S| \leq 2 + \sum_{i,j \in \{3,4\}} p_{ij}(3) = 26.$$ □

Note that any two vertices whose meet has size 7 have exactly two pairs of elements that are in the same cell in both partitions. This means that there are exactly two sets S_{ij} that contain both of these vertices.

4.5.3 Theorem. *The sets* S_{ij} *are the only maximum independent sets in* $\mathcal{P}(3^3)$.

Proof. By Lemma 4.5.1, we need only show that there are no 01-vectors of weight 70 in the column space of H, aside from the columns of H themselves. In fact, we will show that there are no 01-vectors in the column space of H, aside from the columns of H themselves.

Let S be some maximum independent set S. As $\mathcal{P}(3^3)$ is vertex transitive, we may arbitrarily assume that any particular fixed vertex is in S. By Lemma 4.5.2 (and the arc-transitivity of each of the graphs in this association scheme), we may assume that S contains some particular fixed pair of vertices whose meet is 7. Call these vertices a and b, and let M be the submatrix of H consisting of those rows corresponding to neighbours of a or b. Note that neighbours of a or b can not be in S. We may write $z = Hy$ for some vector y, and we now know that $My = 0$. Let N be a matrix whose columns span the kernel of M; then $y = Nx$ for some vector x, and thus $z = MNx$. Let C be the matrix formed by

the non-zero columns of the reduced column echelon form of MN; then $z = Cx'$ for some vector x'.

Since C is in reduced column echelon form, it contains a set of rows that correspond to an identity matrix. As z is a 01-vector, this forces x' to be a 01-vector too. The problem is now finite, as we need only check the $2^{\mathrm{rk}(MN)}$ possibilities for x'.

Using `Maple`, we compute C, finding its rank to be six. Checking the 2^6 possibilities for x, we find that only three 01-vectors result: $\mathbf{0}$ and the characteristic vectors of the two S_{ij} that contain a and b. □

Lemma 4.5.2 is not actually necessary for the above proof: it merely allows us to reduce our computation time. We could have specified the matrix M to correspond to the neighbourhood of a single vertex, rather than the union of the neighbourhoods of two vertices whose meet is of size 7. This would have given a smaller M, and larger C. In particular, we find that the rank of C would then be 15, giving 2^{15} possibilities instead of 2^6. So with a very fast computer at our disposal, we might dispense with Lemma 4.5.2 altogether; with a slower machine we would be motivated to extend it.

In hindsight, this is straightforward. If we momentarily think of $\mathcal{P}(3^3)$ as being a complete graph on 280 vertices, with edges labelled as coming from X_1, X_2, X_3, X_4, then we need only find some edge-labelled subgraph Y that is contained in any maximum independent set (and hence has no X_1-edges). We can then use Y in place of an X_2-edge in the proof; M would now be the union of all neighbourhoods of all vertices of Y. It is not too hard to show that Y can be chosen to be two incident edges from X_2. With a little more effort, one can show that Y can be chosen to be a triangle from X_2, but such triangles are not all equivalent under the automorphism group of the scheme. We have made no attempt to find an "optimal" Y (this would be much more important in extending these results to larger partition graphs). We note that in [32], the exact number of edges of each type in an independent set of size 70 is determined in order to give Lemma 4.5.2 as a corollary.

The following result is a straightforward consequence of a result in Godsil and Royle [33, Lemma 7.5.4].

4.5.4 Lemma. *Let X be vertex transitive, S a maximum independent set in X, and $\psi : X \to X$ a homomorphism. Then the preimage $\psi^{-1}(S)$ is a maximum independent set in X.* □

Using Theorem 4.5.3 and Lemma 4.5.4, one can show that $\mathcal{P}(3^3)$ is a core. The details appear in [32].

4.5.5 Corollary. *The graph $\mathcal{P}(3^3)$ is a core.* □

Unlike the proof given in Section 4.4, this does not depend on having determined the chromatic number by computer. Recall that for that proof we needed to know that $\mathcal{P}(3^3)$ was not 4-colourable. So we have the following bound:

4.5.6 Corollary. $\chi(\mathcal{P}(3^3)) > 4$.

Proof. If it were four colourable, it could be mapped homomorphically into a 4-clique. As $\mathcal{P}(3^3)$ contains a 4-clique, this would mean it was not a core. □

There is a 6-colouring of $\mathcal{P}(3^3)$. Fix a 4-set $C \subseteq \{1, \ldots, 9\}$. The sets S_{ij} where $i, j \in C$ form a cover of $\mathcal{P}(3^3)$ by 6 independent sets. Thus the chromatic number of $\mathcal{P}(3^3)$ is either 5 or 6. Computations by computer [42] show that $\chi(\mathcal{P}(3^3)) = 6$, but it would be nice to have a direct proof of this.

4.6 The Graphs $\mathcal{P}(g^g)$.

Some of the results on the graph $\mathcal{P}(3^3)$ can be extended to the more general case $\mathcal{P}(g^g)$. We have a family of canonical independent sets S_{ij}, defined as before to be the set of partitions that have i and j in the same cell. Some of the arguments used previously extend to the more general case in a straightforward manner.

The graph $\mathcal{P}(g^g)$ has v vertices and is d-regular where

$$v = \frac{1}{g!}\prod_{i=0}^{g-1}\binom{g^2-ig}{g},$$
$$d = (g!)^{g-1}.$$

Again define the matrix H to be the incidence matrix of vertices of the graphs by pairs of elements, so that its columns are the characteristic vectors of the canonical sets S_{ij}. We start by determining its rank. The dot product of any two columns of H corresponding to the pair $\{i, j\}$ and $\{k, l\}$ depends only on $|\{i, j\} \cap \{k, l\}|$. Accordingly, we can compute

$$s_1 := |S_{12}| = s = \frac{1}{(g-1)!}\binom{g^2-2}{g-2}\prod_{i=1}^{g-1}\binom{g^2-ig}{g} = \frac{v}{g+1},$$

$$s_2 := |S_{12} \cap S_{13}| = \frac{1}{(g-1)!}\binom{g^2-3}{g-3}\prod_{i=1}^{g-1}\binom{g^2-ig}{g} = \frac{g-2}{g^2-2}s,$$

$$s_3 := |S_{12} \cap S_{34}| = \frac{1}{(g-1)!}\binom{g^2-4}{g-4}\prod_{i=1}^{g-1}\binom{g^2-ig}{g} + \frac{1}{(g-2)!}\binom{g^2-4}{g-2}\binom{g^2-g-2}{g-2}\prod_{i=2}^{g-1}\binom{g^2-ig}{g}.$$

Let L be the adjacency matrix for the line graph of K_{g^2} (or equivalently, the complement of $K_{g^2:2}$). Then we can determine H^TH as:

$$\begin{aligned} H^TH &= s_1I + s_2L + s_3(J - I - L) \\ &= (s_2 - s_3)L + (s_1 - s_3)I + s_3J \end{aligned}$$

Since the eigenvectors of L are known, this gives the spectrum of H^TH; in particular, we determine the nullity of H^TH to be g^2-1. Recalling that the rank of H is the same as the rank of H^TH, we conclude that

$$\mathrm{rk}(H) = \binom{g^2}{2} - (g^2-1) = \binom{g^2-1}{2}.$$

Now let z be some column of H (i.e., z is the characteristic vector of some S_{ij}). A straightforward computation shows that

$$A\left(z - \frac{s}{v}\mathbf{1}\right) = -\frac{d}{g}\left(z - \frac{s}{v}\mathbf{1}\right).$$

Thus the column space of $H - \frac{s}{v}\mathbf{1}$ is contained in an eigenspace. The sum of the columns of H is $s\mathbf{1}$; thus $\mathbf{1}$ is in the column space of H. Since subtracting $\frac{s}{v}\mathbf{1}$ from a column of H is equivalent to projecting that column onto $\mathbf{1}^\perp$, we see that the column space of $H - \frac{s}{v}\mathbf{1}$ has dimension one less than that of H. We have established the following result:

4.6.1 Lemma. *Let H be a matrix whose columns are the characteristic vectors of the canonical independent sets S_{ij}. Then the column space of $H - \frac{s}{v}\mathbf{1}$ is a subspace of dimension $\binom{g^2-1}{2} - 1$ contained in an eigenspace of eigenvalue $-\frac{d}{g}$.* □

We would like to know that $-\frac{d}{g}$ is in fact the least eigenvalue, and that $\binom{g^2-1}{2} - 1$ is the dimension of the eigenspace. This is the case for $g = 3$, as we have shown previously. We do not know how to resolve this question in general, but we can say something when g is a prime power. In this case, there is a clique of size $g+1$, as this corresponds to an affine plane of order g (or a set of $g+1$ mutually orthogonal Latin squares of order g). Using Lemma 3.2.2, we find that

$$g + 1 \leq \omega \leq 1 - \frac{d}{\tau}$$

But we know that $-\frac{d}{g}$ is an eigenvalue, so we have

$$g + 1 \leq \omega \leq 1 - \frac{d}{\tau} \leq 1 + g.$$

Therefore, both inequalities are tight, which forces $-\frac{d}{g}$ to be the least eigenvalue. We also obtain as a side benefit that these cliques are maximum, although we already knew this from bounds on mutually orthogonal Latin squares.

Chapter 5

Intersecting Systems

5.1 Introduction

Our motivation is the well-known theorem of Erdős, Ko, and Rado dealing with intersecting systems of subsets [22]. This has played a central role in the field of extremal combinatorics. In its base case, it is equivalent to characterizing the maximum independent sets in Kneser graphs; in full generality, it is equivalent to characterizing maximum independent sets in the graphs of the Johnson scheme.

There is a natural generalization from sets to vector spaces: one considers subspaces that intersect in a given dimension, in lieu of subsets that intersect in a given size. This is in fact the natural q-analogue of the EKR Theorem, and we will refer to it as the qEKR Theorem. As we will see, both the statement and conclusion of qEKR are the natural q-analogues to EKR. Again the problem can be rephrased in terms of graphs: characterize the maximum independent sets in the q-Kneser graphs, or in the more general case, the graphs of the Grassmann scheme. So we have a strong motivation for looking for an approach that will encompass both cases.

We will show that a proof based on eigenspaces can in fact characterize independent sets in both the Kneser graphs and the q-Kneser graphs. It is significant that the more well-known approaches to the EKR Theorem have no known generalization to qEKR. We will also give a proof for a case of the qEKR Theorem that has not appeared in the literature.

5.2 Erdős-Ko-Rado and Kneser Graphs

The work of Erdős, Ko, and Rado in the study of intersecting set systems was a major step in the development of extremal combinatorics. The well-known theorem that bears their names was apparently derived in 1938 [21], but not published until 1961 [22].

5.2.1 Theorem. *Let v, k, t be given with $k > t$ and $v > 2k - t$, and let V be a set of size v. Let $\mathcal{F}$ be a collection of k-subsets of V such that the pairwise intersections have cardinality at least t.*

Then for sufficiently large v, $|\mathcal{F}| \leq \binom{v-t}{k-t}$ and equality implies that $\mathcal{F}$ is the set of all k-subsets that contain a fixed t-subset of V. □

Theorem 5.2.1 was first established by Erdős, Ko and Rado in [22]; in particular, they show that for $t = 1$, "sufficiently large" means $v \geq 2k$ for the bound and $v > 2k$ for the characterization.

Recall that the Kneser graph $K_{v:k}$ has as vertex set all k-subsets of a v-set V, with two subsets adjacent if they are disjoint. The case $t = 1$ of Theorem 5.2.1 is equivalent to characterizing maximum independent sets in the Kneser graph $K_{v:k}$. So even without knowing a proof of Theorem 5.2.1, it is clear that for $t = 1$ the range for v cannot be extended: if $v < 2k$ then the Kneser graph $K_{v:k}$ is edgeless, and if $v = 2k$ then $K_{v:k}$ is a perfect matching. For $t > 1$ their proof applies only for sufficiently large v.

The theorem has since been proved by Katona [38] using an approach based on cyclically ordering the elements of V, and also by Daykin [17] using the Kruskal-Katona Theorem (see also Anderson [3] or Bollobás [6]). Frankl [24] showed that the theorem was true for v on the order of a constant times kt. Wilson [51] established the bound for all $v \geq (t+1)(k-t+1)$ and the characterization for $v > (t+1)(k-t+1)$. It was already known that it would not hold for smaller v, and so this was the first proof that established the exact bound in all cases, as opposed to an asymptotic result. His approach uses association schemes. For $2k - t < v < (t+1)(k-t+1)$ the bound given in Theorem 5.2.1 does not apply: these cases are dealt with by Ahlswede and Khachatrian [1].

The Kneser graphs arise naturally as one of the classes in the Johnson Scheme. Recall that the classes in this scheme are the graphs $J(v, k, t)$, with common vertex set the set of k-subsets of V. Two vertices are adjacent in $J(v, k, t)$ if their intersection has size $k - t$ (so $J(v, k, k)$ is the same graph as $K_{v:k}$). Then for general t, Theorem 5.2.1 is equivalent to characterizing the maximum independent sets in the graph $J(v, k, k - t + 1)$.

For historical reasons we have stated Theorem 5.2.1 in terms of intersecting systems. In our proofs, we will mostly take the graph viewpoint; in any case, the two approaches are completely equivalent.

We will present a proof of Theorem 5.2.1 for $t = 1$ based on the ratio bound and the eigenspace corresponding to the least eigenvalue of $K_{v:k}$. Our method has some similarities to [51].

5.3 q-Erdős-Ko-Rado and q-Kneser Graphs

We will need some preliminaries before introducing the q-analogues of Theorem 5.2.1 and the q-Kneser graphs.

For q a prime power, we define

$$[v]_q = \frac{q^v - 1}{q - 1}.$$

Note that $[n]_q$ is actually a polynomial in q. We will generally omit the subscript if it is clear from the context. Furthermore, we define the *q-factorial* of an integer and the *q-binomial coefficients*:

$$\begin{aligned}
[0]! &= 1, \\
[v]! &= \prod_{i=1}^{v} [i], \\
\begin{bmatrix} v \\ k \end{bmatrix} &= \frac{[v]!}{[k]![v-k]!}.
\end{aligned}$$

If we substitute $q = 1$, then

$$[v]_1! = v! \quad \text{and} \quad \begin{bmatrix} v \\ k \end{bmatrix}_1 = \binom{v}{k},$$

and we recover the ordinary binomial coefficients.

Most results on binomial coefficients have analogues for the q-binomial coefficients. For instance, the familiar recursion for the binomial coefficients becomes

$$\begin{bmatrix} v+1 \\ k \end{bmatrix} \quad = \quad q^k \begin{bmatrix} v \\ k \end{bmatrix} + \begin{bmatrix} v \\ k-1 \end{bmatrix} \quad = \quad \begin{bmatrix} v \\ k \end{bmatrix} + q^{v-k+1} \begin{bmatrix} v \\ k-1 \end{bmatrix}$$

in the q-binomial case. It follows that the q-binomial coefficients are polynomials in q; this was perhaps not obvious from the definition given. More relevant to our purposes is the fact that the number of k-subspaces of a v-dimensional vector space over the field of order q is

$$\begin{bmatrix} v \\ k \end{bmatrix},$$

and the number of k-dimensional subspaces of a v-space that have trivial intersection with a fixed t-space is

$$q^{tk} \begin{bmatrix} v-t \\ k \end{bmatrix}.$$

Note that again, these all give the "right" answer if we set $q = 1$. Here, "right" means that subspaces of a vector space of dimension v become subsets of a set of size v, and dimension becomes cardinality. We do not propose to deal explicitly with vector spaces over fields of order 1; we merely observe that the formulae behave in nice ways when we set $q = 1$. This phenomenon is not uncommon in dealing with q-analogues.

We now consider the q-analogue of Theorem 5.2.1.

5.3.1 Theorem. *Let v, k, t be given with $k > t$ and $v > 2k - t$, and let V be a vector space of dimension v over the field of order q. Let $\mathcal{F}$ be a collection of k-subspaces of V such that the pairwise intersections have dimension at least t.*

Then for sufficiently large v, $|\mathcal{F}| \leq \begin{bmatrix} v-t \\ k-t \end{bmatrix}$ and equality implies that $\mathcal{F}$ is the set of all k-subspaces that contain a fixed t-subspace of V. □

Note that qEKR reduces to EKR if we set $q = 1$. This suggests that it is a natural analogue of EKR, and that the two are strongly related.

The q-Kneser graphs are the natural q-analogues of the Kneser graphs. Let V be a v-dimensional vector space over $GF(q)$. The vertex set of $qK_{v:k}$ is the set of all k-subspaces of V. Two vertices are adjacent if their intersection has dimension zero. Furthermore, the q-Kneser graphs arise as one of the classes in the Grassmann scheme. The classes in this scheme are the graphs $J_q(v, k, t)$, with common vertex set the set of k-subspaces of V. Two vertices are adjacent in $J_q(v, k, t)$ if their intersection has dimension t (so $J_q(v, k, k)$ is the same graph as $qK_{v:k}$). Then for general t, Theorem 5.3.1 is equivalent to characterizing the maximum independent sets in the graph $J_q(v, k, k - t + 1)$.

Lest the reader imagine that q-analogues are nothing more than a verbose and esoteric way to redefine 1, we point out that the chromatic number of the q-Kneser graphs, in the few cases that it is known, does *not* reduce to the chromatic number of the Kneser graphs if we set $q = 1$; see [13] for details.

Theorem 5.3.1 was first proved by Hsieh [37], but not for all relevant v. His proof is combinatorial and very technical. Deza and Frankl give a brief sketch of an inductive argument using shifting in [20]. Czabarka and Székely [16] claim that there are counterexamples to Deza and Frankl's proof, and a new proof is advanced, using a generalization of shifting. However this definition of shifting is also flawed and the proof is not valid. At present there is no known proof using a shifting-type argument. There is no known q-analogue of Katona's cyclic proof. Neither is there any known q-analogue of Daykin's proof; indeed, there is no known q-analogue of the Kruskal-Katona theorem to base it on. Finally Frankl and Wilson [26] proved Theorem 5.3.1 using methods similar to [51], although they did not establish the characterization for $v = 2k, t > 1$ (they state it without proof for $v = 2k, t = 1$).

Our result for EKR applies to qEKR also. While there are some similarities between our work and [51, 26], the approach is different. We also give a proof of the characterization for $v = 2k, t = 1$; this has not appeared before in the literature, though the result is stated in [26].

5.4 Proving EKR

We prove the case $t = 1$ of Theorem 5.2.1. The proof we offer here is new, although it turns out to be related to the methods in [51].

The eigenvalues of the Kneser graphs are known. A derivation is given in Godsil and Royle [33], for instance. Alternatively, the eigenvalues are known

because the Kneser graphs lie in the Johnson scheme, whose parameters were derived in [19]. In particular, the degree and the least eigenvalue are:

$$d = \binom{v-k}{k},$$

$$\tau = -\binom{v-k-1}{k-1}.$$

Knowing this, we have the bound we want as an immediate consequence.

5.4.1 Lemma. *The size of an independent set in $K_{v:k}$ is at most $\binom{v-1}{k-1}$.*

Proof. It is simply a matter of plugging the eigenvalues into the bound of Theorem 3.4.1.

$$|S| \leq \frac{\binom{v}{k}}{1 - \frac{\binom{v-k}{k}}{-\binom{v-k-1}{k-1}}} = \binom{v-1}{k-1}. \qquad \square$$

Thus the bound in EKR is a direct consequence of the spectrum of the Kneser graphs. This bound holds with equality whenever S is the set of all k-subsets containing a fixed element. We have a family of maximum independent sets; the task remaining is to show that they are in fact the only ones.

Let H be the matrix whose rows are indexed by k-subsets, whose columns are indexed by elements, and whose (a, i) entry is equal to 1 when $i \in a$ and 0 otherwise. Then the columns of H are precisely the characteristic vectors of the canonical maximum independent sets. We will show that all ratio-tight sets lie in the column space of H. We will then argue that the only 01 vectors in the column space of H are the columns themselves. The critical ingredients are the equality condition of Theorem 3.4.1, and a rank computation of a matrix in an association scheme.

The column space of H is clearly contained in the sum of the τ-eigenspace of A and the d-eigenspace of A (the latter being just the constant vectors). This can be seen by direct computation: by subtracting off the right multiple of $\mathbf{1}$ from each column of H to make the result orthogonal to $\mathbf{1}$, and multiplying by the adjacency matrix A of $K_{v:k}$, we discover that

$$A\left(H - \frac{\binom{v-1}{k-1}}{\binom{v}{k}}J\right) = \tau\left(H - \frac{\binom{v-1}{k-1}}{\binom{v}{k}}J\right).$$

In fact we can say more. Define the matrices W_{ij} to be the incidence matrices of i-sets and j-sets where $(W_{ij})_{ab} = 1$ when the j-set b is contained in the i-set a. So W_{k1} is our matrix H. Let U_j be the column space of W_{kj}, and for convenience set $U_{-1} = \{0\}$. It is well known that the columns of W_{kj} are linearly independent and that the eigenspaces of the Johnson scheme can be ordered so that the j-th eigenspace is exactly the orthogonal complement of U_{j-1} in U_j. This is in fact the cometric ordering of the scheme. Proofs of these statements can be found in Godsil [29].

For $v > 2k$, the eigenspaces of the scheme are exactly the eigenspaces of $K_{v:k}$. This can be seen directly from the matrix of eigenvalues of the Johnson scheme: none of the values in the column corresponding to $K_{v:k}$ is repeated. We also observe that, using the above ordering, the eigenspace corresponding to $j = 0$ consists of the constant vectors, with eigenvalue equal to the degree, and the eigenspace corresponding to $j = 1$ has eigenvalue τ, the least eigenvalue of $K_{v:k}$.

For $v = 2k$, the eigenspaces of the $K_{v:k}$ are sums of eigenspaces of the scheme, as the relevant column of the matrix of eigenvalues has in fact only two distinct values. We will have more to say about this at the end of Section 5.5.

As noted above, for $v > 2k$, the columns of H form a basis for the sum of the τ-eigenspace and the d-eigenspace. Since the ratio bound holds with equality, then by Theorem 3.4.1 the characteristic vector of any maximum independent set lies in the column space of H. We have established the following result:

5.4.2 Lemma. *Let $v > 2k$. Then the characteristic vector of any maximum independent set lies in the column space of H.* □

Any 01-vector in the column space of H that is the characteristic vector of an independent set must be the characteristic vector of an independent set meeting the ratio bound. We need to show that the only such vectors are the columns of H. In fact, we will show something stronger: the only 01-vectors in the column space of H are the columns themselves. To do this we will need to investigate the matrix H a little further.

Let a be some fixed but arbitrary vertex. Recall that the rows of H are indexed by vertices of $K_{v:k}$, and the columns by elements of the ground set V. It will be useful to have notation for the following submatrices.

- Let N be the submatrix of H with rows corresponding to the neighbourhood of a.
- Let M be the submatrix of H with rows corresponding to the neighbourhood of a and columns corresponding to the elements of V not in a.
- Let W be the submatrix of H with columns corresponding to the elements of V in a.

Strictly speaking, these matrices all depend on a, but varying a only permutes the rows and columns. As an example, we give the matrix H for $K_{5,2}$, the Petersen graph:

	1	2	3	4	5
12	1	1	0	0	0
34	0	0	1	1	0
35	0	0	1	0	1
45	0	0	0	1	1
13	1	0	1	0	0
14	1	0	0	1	0
15	1	0	0	0	1
23	0	1	1	0	0
24	0	1	0	1	0
25	0	1	0	0	1

The matrix W corresponds to the first two columns, the matrix N corresponds to the third, fourth and fifth rows, and the matrix M corresponds to the last three columns of N (in general, M is not square). We have labelled the rows and columns with the subsets they represent. Notice that N is just M preceded by a block of zeroes; it is not hard to see that this is true in general. Also notice that M has full column rank. This is true in general, and is the subject of our next result.

5.4.3 Lemma. *Let $v > 2k$. Then the matrix M has full column rank.*

Proof. Recall that M and M^TM have the same non-zero eigenvalues with the same multiplicities. So it suffices to show that M^TM has full rank. The rows and columns of M^TM are indexed by elements of the ground set V. In particular, the (i,j)-entry is the number of neighbours of a that contain i and j. For $i = j$ this number is

$$\binom{v-k-1}{k-1},$$

and for $i \neq j$ it is

$$\binom{v-k-2}{k-2}.$$

Thus we have shown that

$$\begin{aligned} M^TM &= \binom{v-k-1}{k-1} I + \binom{v-k-2}{k-2}(J-I) \\ &= \binom{v-k-2}{k-2}\left(\frac{v-2k}{k-1} I + J\right). \end{aligned} \tag{5.1}$$

Being a linear combination of I and J, M^TM has two eigenspaces: $\langle \mathbf{1} \rangle$ and $\langle \mathbf{1} \rangle^\perp$. We can then compute the eigenvalues to be

$$k\binom{v-k-1}{k-1} \text{ and } \frac{v-2k}{v-k-1}\binom{v-k-1}{k-1}, \tag{5.2}$$

with multiplicities 1 and $v-k-1$, respectively. These are both positive, unless $v = 2k$, in which case the second eigenvalue listed is zero. □

Note that this proof is really just an application of Lemma 2.5.4, on the association scheme with one class. Of course, the association scheme with one class (there is only one such scheme) in a rather simple object, and we don't need much theory at all to work with it. It would be quite possible to give the above proof without knowing anything about association schemes. But the schemes *are* part of what we are doing.

5.4.4 Corollary. *Let $v > 2k$. If y lies in the kernel of N, then Hy lies in the column space of W.*

Proof. Recall that y is indexed by the elements of the ground set V. Let y_M be the subvector of y corresponding to the elements of V not in a. If y lies in the kernel of N then

$$0 = Ny = (0M)\, y = My_M.$$

But by Lemma 5.4.3, M has full rank and therefore $y_M = 0$, so y is zero in all positions corresponding to elements of V not in a. Thus Hy is a linear combination of the columns of W. □

We now prove Theorem 5.2.1.

Proof. First we deal with the case $v > 2k$. Let S be any independent set that meets the ratio bound, and z its characteristic vector. Then by Lemma 5.4.2, $z = Hy$ for some vector y. We may assume that S contains the vertex a. Since no neighbour of a can be contained in S, $Ny = 0$. Then by Corollary 5.4.4, $z = Wx$ for some vector x.

For each $i \in a$, there is a vertex b_i such that $a \cap b_i = i$. The submatrix of W with rows indexed by the vertices b_i is the identity matrix. Thus x is a zero-one vector. For every pair $\{i, j\} \in a$, there is a vertex c_{ij} such that $a \cap c_{ij} = \{ij\}$. Thus x does not have a one at positions i and j, and so has at most one nonzero entry. Therefore z is one of the columns of W.

The case $v = 2k$ is dealt with by elementary methods: in this case $K_{v:k} = K_{2k:k}$ is a perfect matching, and thus the largest independent sets consist of exactly one vertex chosen from each pair of matched vertices. □

5.5 Proving qEKR

The argument we used in the preceding section can also be used to establish the q-analogue of EKR, which we will do for $t = 1$. The proof for $v = 2k$ has not appeared before in the literature, although the result was stated by Wilson [51].

The eigenvalues of the q-Kneser graphs are also known: these graphs are classes in the Grassmann scheme [19]. In particular, the degree and least eigen-

value are:

$$d = q^{k^2}\begin{bmatrix} v-k \\ k \end{bmatrix},$$

$$\tau = -q^{k(k-1)}\begin{bmatrix} v-k-1 \\ k-1 \end{bmatrix}.$$

Again we find that the ratio bound gives the bound we want.

5.5.1 Lemma. *The size of an independent set in $qK_{v:k}$ is at most $\begin{bmatrix} v-1 \\ k-1 \end{bmatrix}$.*

Proof. It is again a simple matter of plugging the new eigenvalues into the ratio bound.

$$|S| \leq \frac{\begin{bmatrix} v \\ k \end{bmatrix}}{1 - \frac{q^{k^2}\begin{bmatrix} v-k \\ k \end{bmatrix}}{q^{k(k-1)}\begin{bmatrix} v-k-1 \\ k-1 \end{bmatrix}}} = \begin{bmatrix} v-1 \\ k-1 \end{bmatrix}. \qquad \square$$

Let H be the matrix whose rows are indexed by k-subspaces, whose columns are indexed by 1-subspaces, and whose (a, i) entry is equal to 1 when $i \in a$ and 0 otherwise. Knowledge of the eigenspaces of the Grassmann scheme tells us that for $v \geq 2k$ the column space of H is exactly the sum of the d-eigenspace and the τ-eigenspace. The constructions are completely analogous to the ordinary case in Section 5.4, with one exception. When we were dealing with sets, the relationship between the column space of H and the eigenspaces of the Kneser graph held only for $v > 2k$. The matrix of eigenvalues of the Grassmann scheme has distinct values in the column corresponding to $qK_{v:k}$, even for $v = 2k$.

Thus the columns of H form a basis for the sum of the d-eigenspace and the τ-eigenspace whenever $v \geq 2k$. As before, since the ratio bound does hold with equality, the characteristic vector of any maximum independent set lies in the column space of H.

5.5.2 Lemma. *Let $v \geq 2k$. Then the characteristic vector of any maximum independent set lies in the column space of H.* $\square$

Let a be some fixed but arbitrary vertex. We define the matrices N, M and W in analogy with the previous section. Let N be the submatrix of H with rows corresponding to the neighbourhood of a. Let M be the submatrix of N with columns corresponding to the 1-dimensional subspaces of V not in a, so $N = (0M)$. Let W be the submatrix of H with columns corresponding to the 1-dimensional subspaces of V in a. Again, note that varying a only permutes the rows and columns of N, M, W.

We will use some of the language of projective geometry, and refer to a 1-dimensional subspace as a point and a $(v-1)$-dimensional subspace as a hyperplane. A point is *on a subspace* if it is contained in it; it is *off a subspace* if it has trivial intersection with the subspace. Hence the columns of W correspond to the points on a, and the columns of M to the points off of a.

5.5.3 Lemma. *Let $v > 2k$. Then the matrix M has full column rank.*

Proof. We will again consider the matrix M^TM, with rows and columns indexed by points off of a. The (i,j) entry of M^TM is the number of neighbours of a that contain both i and j. Consider the space spanned by i and j. This space intersects a non-trivially if and only if i, j and a are all contained in a common $(k+1)$-subspace. In this case, any subspace that contains i and j would intersect a non-trivially: in particular, there would be no vertices adjacent to a but containing both i and j.

So we see that the points off of a are partitioned by the $(k+1)$-subspaces. Fix some $(k+1)$-subspace $\tilde{a}$ containing a; the number of points in $\tilde{a}$ but off of a is $[k+1]-[k] = q^k$. On the other hand, the number of $(k+1)$-subspaces containing a is $[v-k]$. Conveniently, $q^k[v-k] = [v]-[k]$, which is the number of points off of a. If we define a graph on these points with two of them adjacent if they are in the same class, then this graph is $[v-k]$ disjoint copies of K_{q^k}.

We now compute the entries of M^TM. Let i, j be two points off of a.

First, $\left(M^TM\right)_{ii}$ is equal to the number of k-subspaces containing i but disjoint from a. There are

$$q^{k^2}\begin{bmatrix} v-k \\ k \end{bmatrix}$$

k-spaces disjoint from a, each containing $[k]$ points, and there are a total of $[v]-[k]$ points off of a. So by counting in two ways the number of pairs consisting of a point off of a and a k-space on the point but disjoint from a, we calculate the number of k-spaces containing i but not a to be:

$$\frac{q^{k^2}\begin{bmatrix} v-k \\ k \end{bmatrix}[k]}{[v]-[k]} = q^{k(k-1)}\begin{bmatrix} v-k-1 \\ k-1 \end{bmatrix}.$$

If i and j are distinct but in the same class, then every r-subspace containing them intersects a non-trivially and so $\left(M^TM\right)_{ij} = 0$.

If i and j are in different classes, then $\left(M^TM\right)_{ij}$ is equal to the number of k-subspaces containing a given 2-dimensional subspace (i.e., the span of i and j) but disjoint from a. Each k-space disjoint from a contains $[k]([k]-1)$ ordered pairs of points from different classes (any two points in such an k-space are from different classes); there are a total of $([v]-[k])([v]-[k+1])$ ordered pairs of points off of a in different classes. Again, by counting in two ways we calculate the number of k-spaces containing i but not a to be:

$$\frac{q^{k^2}\begin{bmatrix} v-k \\ k \end{bmatrix}[k]([k]-1)}{([v]-[k])([v]-[k+1])} = q^{k(k-2)}\begin{bmatrix} v-k-2 \\ k-2 \end{bmatrix}.$$

Let L be the adjacency matrix of $[v-k]$ disjoint copies of K_{q^k}. We have shown

that

$$M^T M = q^{k(k-1)} \begin{bmatrix} v-k-1 \\ k-1 \end{bmatrix} I + q^{k(k-2)} \begin{bmatrix} v-k-2 \\ k-2 \end{bmatrix} (J - I - L)$$

$$= q^{k(k-2)} \begin{bmatrix} v-k-2 \\ k-2 \end{bmatrix} \left(\left(\frac{q^k [v-k-1] - [k-1]}{[k-1]} \right) I - L + J \right). \quad (5.3)$$

We know the eigenspaces of the scheme generated by I, J and L, so we can compute the eigenvalues of $M^T M$ to be

$$[k] q^{k(k-1)} \begin{bmatrix} v-k-1 \\ k-1 \end{bmatrix},$$

$$\frac{q^{k-1}[v-2k]}{[v-k-1]} q^{k(k-1)} \begin{bmatrix} v-k-1 \\ k-1 \end{bmatrix},$$

$$q^{k(k-1)} \begin{bmatrix} v-k-1 \\ k-1 \end{bmatrix}, \quad (5.4)$$

with multiplicities 1, $[v-k]-1 = q[v-k-1]$, and $\left(q^k - 1\right)[v-k]$, respectively. These are all positive, unless $v = 2k$, in which case the second eigenvalue listed is zero. □

If we compare these results with the ordinary set version, we see that (5.3) and (5.4) reduce to (5.1) and (5.2) if we set $q = 1$. For $q = 1$, the matrix L is identically zero, hence the third eigenvalue listed disappears (its multiplicity becomes zero).

5.5.4 Corollary. *Let $v > 2k$. If y lies in the kernel of N, then Hy lies in the column space of W.* □

We deal now with the case of $v = 2k$. We will need some more information on the structure of the kernel of the matrix N. To do so, it will be useful to have at hand another basis for the column space of H.

We have up until now used the language of vector spaces: our vertices are k-dimensional subspaces of a v-dimensional vector space. We may also take the viewpoint of projective geometry. For $v = 2k$, our vertices are projective subspaces of rank $k-1$ in $PG(2k-1, q)$. But $PG(2k-1, q)$ is self-dual: projective subspaces of rank t get mapped to projective subspaces of rank $2k-2-t$. This duality fixes (as a set) the vertices of $qK_{v:r}$ since these are projective subspaces of rank $k-1$. Lemma 5.5.2 can be loosely described as using the projective subspaces of rank 0 to obtain information about the projective subspaces of rank $k-1$. Duality suggests that we should be able to get the same information

using the projective subspaces of rank $2k-2$. This is by no means a proof, and in fact we shall not make explicit use of duality, but it will be helpful in understanding our motivation in what follows. Note that duality will not be helpful for $v > 2k$, because the dual of vertices of $qK_{v:k}$ are not vertices of $qK_{v:k}$.

Let H' be the matrix whose rows are indexed by k-subspaces, whose columns are indexed by hyperplanes (i.e.,$(2k-1)$-subspaces, and whose (a,p) entry is equal to 1 when $a \in p$ and 0 otherwise (note that the containment is backwards relative to the definition of H). The reader may notice that the columns of H' are characteristic vectors of maximum independent sets that meet the ratio bound, and that they do *not* arise as linear combinations of columns of W. We will keep the same fixed vertex a as before, and define N', M', W' in an analogous fashion: N' is the submatrix of H' corresponding to the neighbourhood of a, M' is the submatrix of N' corresponding to the hyperplanes off of a, and W' is the submatrix of H' corresponding to the hyperplanes on a.

The following result is well-known; we include it as the proof will be useful later.

5.5.5 Lemma. *The column space of H is equal to the column space of H'.*

Proof. It suffices to express the columns of each in terms of the other. Let ϕ_i be the column of H corresponding to the point i, and ψ_p be the column of H' corresponding to the hyperplane p. Then it is straightforward to verify that:

$$[k]\psi_p = \sum_{i:i\in p} \psi_i - \frac{[k-1]}{q^{k-1}} \sum_{i:i\notin p} \psi_i,$$

$$[k]\phi_i = \sum_{p:i\in p} \phi_i - \frac{[k-1]}{q^{k-1}} \sum_{p:i\notin p} \phi_i. \qquad \square$$

5.5.6 Lemma. *Let $v = 2k$. Let B_0 be the matrix indexed by points off of a and hyperplanes on a such that*

$$(B_0)_{i,p} = \begin{cases} \frac{1}{[k]} & i \in p, \\ -\frac{[k-1]}{q^{k-1}[k]} & i \notin p. \end{cases}$$

Then the kernel of the matrix N is spanned by the columns of

$$B = \begin{pmatrix} I_{[k]} & \frac{1}{[k]}J \\ 0 & B_0 \end{pmatrix}.$$

Proof. It is perhaps trite to observe that multiplying H by the first $[k]$ columns of B gives W, but it is also true, by the proof of Lemma 5.5.5, that multiplying

H by the last $[k]$ columns of B gives W'. As the rows of both W and W' corresponding to the neighbourhood of a are zero, B is certainly contained in the kernel of N.

We know from the proof of Lemma 5.5.3 that when $v = 2k$, B has rank $2[k] - 1$. Thus it is suffices to show that the rank of B_0 is $[k] - 1$.

To do this, note that the rank of B_0 is the same as the rank of $B_0^T B_0$. The rows and columns of $B_0^T B_0$ are indexed by $(2k-1)$-spaces containing a. The (i, j)-entry of $B_0^T B_0$ depends only on whether $i = j$ or $i \neq j$.

Let i be a $(2k-1)$-space containing a. The number of 1-subspaces in i but not in a is $q^k[k-1]$, and the number of 1-spaces not in i is q^{2k-1}. Thus the diagonal entries of $B_0^T B_0$ are given by

$$(q^k[k-1])\left(\frac{1}{[k]}\right)^2 + (0)\left(\frac{-[k-1]}{q^{k-1}[k]^2}\right) + (q^{2k-1})\left(\frac{-[k-1]}{q^{k-1}[k]}\right)^2 = \frac{q[k-1]}{[k]}.$$

Let i and j be $(2k-1)$-spaces containing a. The number of 1-subspaces in $i \cap j$ but not in a is $q^k[k-2]$, the number of 1-subspaces in in the symmetric difference of i and j but not in a is $2q^{k-2}$, and the number of 1-spaces not in $i \cup j$ is $q^{2k-1-q^{2k-2}}$. Thus the off-diagonal entries of $B_0^T B_0$ are given by

$$(q^k[k-2])\left(\frac{1}{[k]}\right)^2 + (2q^{k-2})\left(\frac{-[k-1]}{q^{k-1}[k]^2}\right) + (q^{2k-2}(q-1))\left(\frac{-[k-1]}{q^{k-1}[k]}\right)^2 = -\frac{1}{[k]}.$$

Thus we have shown that

$$B_0^T B_0 = \frac{q[k-1]}{[k]} I - \frac{1}{[k]}(J - I).$$

The result follows. □

Note that in the above proof, we showed that $B_0^T B_0$ lies in the Bose-Mesner algebra of the complete graph; using the matrix of eigenvalues of the scheme, we could then easily compute all the eigenvalues of this matrix to obtain the rank. Our proof is similar to our proof of Lemma 5.5.3, and in fact, we can also show (using the same technique) that $B_0 B_0^T$ lies in the Bose-Mesner algebra of $[k]K_{q^k}$, and obtain the rank from the matrix of eigenvalues of this scheme.

The following notation will be useful:

$$\widehat{W} = \begin{pmatrix} W & W' \end{pmatrix}.$$

5.5.7 Corollary. *Let $v = 2k$. If y lies in the kernel of N, then Hy lies in the column space of $\widehat{W}$.*

Proof. Since y lies in the kernel of N, it lies in the column space of B, and so Hy lies in the column space of $\widehat{W}$. □

We now prove Theorem 5.3.1.

Proof. We first deal with the case $v > 2k$. Let S be a maximum independent set, and z its characteristic vector. By Lemma 5.5.2, we know that $z = Hy$. Since we may assume that S contains an arbitrary vertex (and thus does not contain the neighbourhood of that vertex), we see that $Ny = 0$ and so, by Corollary 5.5.4, $z = Wx$ for some vector x.

For each point i in a, there is a vertex b_i such that $a \cap b_i = i$. The submatrix of W with rows indexed by the vertices b_i is the identity matrix. Thus x is a zero-one vector. For every pair of points $i_1, i_2 \in a$, there is a vertex $c_{i_1 i_2}$ such that $a \cap c_{i_1 i_2} = \langle i_1, i_2 \rangle$. Thus x has at most one nonzero entry. Therefore z is one of the columns of W.

We now deal with the case $v = 2k$. Again, let S be a maximum independent set, and z its characteristic vector. Using Lemma 5.5.2 and Corollary 5.5.7, we we that $z = \widehat{W}x$ for some vector x.

For each point i in a and hyperplane p on a, there is a vertex $b_{i,p}$ such that $a \cap b_{i,p} = i$ and $\langle a, b \rangle = p$. This means that there is a row of $\widehat{W}$ that is zero except on the columns corresponding to i and p, where it is one, and this is true for any point i and hyperplane p. (Alternatively, the submatrix of $\widehat{W}$ indexed by the rows corresponding to $b_{i,p}$ is $(\mathbf{1} \otimes I \mid I \otimes \mathbf{1})$).

Since z is a zero-one vector, if $x_i = \gamma$ for some point i in a, then x_p must be either $-\gamma$ or $1 - \gamma$ for *every* hyperplane p on a. Furthermore, if both of these values actually occur, then $x_i = \gamma$ for every point i in a. Thus x is constant on points or constant on hyperplanes. Since $z_a = 1$, we must have that $x_i = 1 - \gamma$ for some point i in a, $x_j = -\gamma$ for all other points j in a, and $x_p = \gamma$ for all hyperplanes p on a (or the dual obtained by inverting points and hyperplanes). It is now straightforward to check that $z = \widehat{W}x$ is exactly the column of $\widehat{W}$ corresponding to the particular choice of i. □

Comparing Corollary 5.5.7 with Corollary 5.5.4, we notice one important difference. The columns of $\widehat{W}$ are not linearly independent, so vectors expressible as linear combinations of $\widehat{W}$ are not uniquely expressible. This is the reason for the apparent extra flexibility in the choice of γ at the end of the proof. This is just a reflection of the fact that we did not bother to find a basis for the column space of $\widehat{W}$. In fact, it would not have been difficult: since $W\mathbf{1} = W'\mathbf{1}$, we can express any column of W' as the sum of the columns of W minus the sum of the other columns of W', and thus removing any one column from $\widehat{W}$ would leave behind a basis.

We have one final observation about our proof of Theorem 5.3.1 for $v = 2k$. Our characterization of the 01-vectors in the column space of $\widehat{W}$ is virtually identical to our proof of the result of Alon *et al.* on graph products, Theorem 3.13.1. In fact, If we set $r = 2$ in Theorem 3.13.1, then there is a set of rows of $\widehat{W}$ that is exactly the matrix H we used in our proof of Theorem 3.13.1. Not only that, but this set of rows is exactly what we used in our proof of Theorem 5.3.1. Our proof of Theorem 5.3.1 preceded our proof of Theorem 3.13.1, and it is tempting to reword them so that one is a corollary of the other. We

will be satisfied with the observation that the tools we are using have a wider applicability than we first intended.

5.6 Why Is $v = 2k$ Special?

The proof of Theorem 5.3.1 in the case $v = 2k$ is not substantially different from the proof of the case $v > 2k$. True, the kernel of N is a little larger, but overall the approach is similar. This was not the case for Theorem 5.2.1. We could have, in the set version, also defined matrices H', N', M', W', and (*almost*...) everything would carry over, with essentially only minor changes in the wording. Things might even seem fairly simple: for instance, for $v = 2k$ in the set version, the matrix M is just $\mathbf{1}^T$. There is an easy answer as to why we did not do this: the answer thus obtained is wrong. So we are motivated to ask: Why does this not work for the set version?

The answer is that Lemma 5.4.2 is false when $v = 2k$. In order to understand why, it will be useful to consider the relationship between the Kneser graphs and the Johnson scheme. There are $k + 1$ classes in the scheme, and therefore there are $k + 1$ eigenspaces of the scheme. Each of these is an invariant subspace of each graph of the scheme, and thus the eigenspaces of each graph are sums of eigenspaces of the scheme. A more pedagogically correct version of Lemma 5.4.2 would have said that the columns of H form a basis for the sum of two eigenspaces of the *scheme*. Then we would have determined that the eigenspaces of $K_{v:k}$ are the same as the eigenspaces of the scheme.

So the question is how to tell which eigenspaces of the scheme combine to give the eigenspaces of a graph in the scheme. To answer this, we return to the matrix of eigenvalues of the scheme. The value of P_{ij} is the eigenvalue of the i-th eigenspace of the j-th graph. So two eigenspaces of the scheme both lie in the same eigenspace of the graph if the corresponding values of P_{ij} are equal.

If all the values in the j-th column of the matrix of eigenvalues are distinct, then clearly the eigenspaces of the j-th graph are exactly the eigenspaces of the scheme. This is the case for the Johnson scheme for $v > 2k$, and also for the Grassmann scheme for $v > 2k$. We presented Lemma 5.4.2 and Lemma 5.5.2 as though we had luckily stumbled on a convenient basis for the sum of the τ-eigenspace and the constant vectors. Not so: what we actually did was to take the standard basis for the "first" two eigenspaces of the scheme, and then observe (i.e., by looking at the matrix of eigenvalues) that these are in fact the right eigenspaces of $K_{v:k}$. The ordering with respect to which they are the "first" is just the standard cometric ordering of the Johnson scheme (see Section 2.7).

To return to our immediate purposes, what happens when $v = 2k$? For the Johnson scheme, the k-th column of the matrix of eigenvalues contains only two distinct values: ± 1. This is expected, since these are the eigenvalues of $K_{2k:k}$, which is just a perfect matching. Now looking back at Lemma 5.4.2, we see that our H is still the sum of two eigenspaces of the scheme, but that these are no longer eigenspaces of the graph. In fact, what we want is the whole of the τ-eigenspace (which corresponds to half of the eigenspaces of the scheme) and

the constant vectors (an eigenspace of the scheme, but no longer an eigenspace of the graph). Substituting this new H would then yield a characterization of the maximum independent sets in $K_{2k:k}$. The new matrix N would have a very large kernel: we know this because there are 2^{2k-1} maximum independent sets that contain a given vertex a.

This makes the behaviour of the q-analogue more interesting. Although the kernel of N was larger than for $v > 2k$, the matrix H was still correct: for both $v > 2k$ and $v = 2k$, we were looking for zero-one vectors in the same subspace. This is because the corresponding column of the matrix of eigenvalues had all values distinct, meaning that even for $v = 2k$, the eigenspaces of the Grassmann scheme are exactly the eigenspaces of $qK_{v:k}$. However, on second glance, they are "barely" distinct: they fall into two groups, and within each group they differ only by a factor of a power of q. They do differ, so H still spans the relevant eigenspaces of $qK_{v:k}$, but upon substitution of $q = 1$, they collapse.

It is a happy accident that $K_{v:k}$ is a perfect matching, otherwise our methods as used would not work. In fact, most classical proofs of Theorem 5.2.1 for $t = 1, v > 2k$ do not give the characterization when $v = 2k$. This has never been regarded as a problem, of course. But seen in the present light, $t = 1, v = 2k$ is the only hard case, because that is where our theoretical machinery, Lemma 5.4.2, fails. From this point of view, the q-analogue is actually simpler: the q-analogue of Lemma 5.4.2, Lemma 5.5.2, is valid for $t > 1, v \geq 2k$, and the only extra work we need to do derives from the fact that the kernel of N is no longer quite as simple.

Chapter 6

An Orthogonality Graph

6.1 Introduction

We turn now to the last of the three main problems of this thesis, and look at a graph colouring question that arises in the study of quantum entanglement.

We will give a brief introduction to quantum theory and quantum computing. We do not assume any physics background on the reader's part, nor is a deep understanding of quantum mechanics necessary to understand our work. But some background on quantum computing is important in understanding the particular problem that arises. We then turn our attention to describing the problem that gives rise to the graph colouring problem that we will then solve. We will delay the physics for a moment in order to describe our main result in graph theoretic terms and some work of others related to it.

The object of our study in this chapter will be the graph $\Omega(n)$. It has as vertex set the set of all ± 1 vectors of length n, and two vectors are adjacent if they are orthogonal. We will be interested in the chromatic number of $\Omega(n)$, and this will lead us into the study of its maximum independent sets. For the purposes of our application to quantum entanglement, we are concerned with n a power of two. We will, along the way, deal with other values of n, but these turn out to be for the most part both simpler and less important. Our main result is the following theorem.

6.1.1 Theorem. *Let n be a power of two. Then $\chi(\Omega(n)) = n$ if $n \leq 8$ and $\chi(\Omega(n)) > n$ otherwise.* □

One way to prove a lower bound on the chromatic number is to find an upper bound on the size of an independent set. If $\alpha(\Omega(n)) < \frac{2^n}{n}$, then $\chi(\Omega(n)) > n$ (it will turn out that $\frac{2^n}{n}$ is the ratio bound). All non-trivial known results on lower bounds for $\chi(\Omega(n))$ are derived from upper bounds on $\alpha(\Omega(n))$. This may seem somewhat surprising at first, but as we shall see in Corollary 6.8.3, the connection between maximum independent sets and vertex colourings is very close for our graphs.

For clarity, we stated Theorem 6.1.1 in a complete form. Some of the cases are trivial, and some have been dealt with by others. The cases $n = 1, 2$ are completely trivial, as the reader may verify in a suitable margin. The case $n = 4$ might require a larger margin, though as a consequence of Corollary 6.9.2 a medium margin is seen to suffice.

The first result on the graph $\Omega(8)$ was due to Gordon Royle, who, motivated by the questions of Viktor Galliard, found an 8-colouring of $\Omega(8)$. This result is described in [31]. Galliard, Tapp, and Wolf [28] showed the following result

6.1.2 Lemma. $\alpha(\Omega(16)) \leq 3912$. □

From this they trivially deduce that $\chi(\Omega(16)) > 16$. Their approach was to partition the vertices into levels according to their weight, and find an upper bound on the size of a maximum independent set in each level. Adding these bounds up yields an upper bound on the size of an independent set in the graph. Their proof involves a somewhat technical case analysis for one level, and the use of the Erdős-Ko-Rado Theorem for the others. We later showed $\Omega(8) = 8$ and $\Omega(16) > 16$, in a different manner than the previous authors.

The papers by Buhr, Cleve, and Wigderson [10], and Brassard, Cleve, and Tapp [7] dealt with the quantum application that motivates us. The relevance of the quantum problem that motivates our chapter is introduced in [10], and further developed in [7]. These authors were almost certainly aware of the relationship between the quantum question and the graph problem, although they do not describe it in that language. The link is made explicitly in [27]. For the purposes of resolving our question about $\Omega(n)$, the result of [10] relies in an essential way on a result of Frankl and Rödl [25].

The paper of Frankl and Rödl deals with intersecting systems. There is no hint of quantum in their paper, nor is their motivation determining the chromatic number of graphs. It predated the quantum application it was used to solve; actually, it predated most of quantum computing. Aside from its mathematical significance, it is perhaps noteworthy as an excellent example of how it is often very difficult to determine in advance which results, methods, and even fields of study will later prove most useful. We quote briefly from [10]:

> ... we need the following strong result of Frankl and Rödl, which seems tailor-made to our needs.

For our purposes, we paraphrase a restricted version of Frankl and Rödl's result [25].

6.1.3 Theorem. *Then there exists a positive constant ϵ such that whenever n is a multiple of four, $\alpha(\Omega(n)) \leq (2 - \epsilon)^n$.* □

We make two comments on this bound. Firstly, it does not directly imply Theorem 6.1.1, as it does not tell us any particular value of n for which the independent sets are too small. Secondly, and much more importantly, it tells

us that the size of a maximum independent set is exponentially smaller than the required $\frac{2^n}{n}$. So in fact, not only does it tell us that the chromatic number is greater than n, it tells us that it is exponential in n.

Our methods address the first comment, as we determine whether $\Omega(n)$ is equal to n, or greater than n for all values n. In fact, we will discover as a by-product the exact characterization of all ratio-tight independent sets. There is some possibility of extending our work to obtain the bound of Frankl and Rödl (in a very different way than they do), but for the moment we are unable to do this.

The next section is a very brief introduction to quantum theory and quantum entanglement. Following this, we discuss quantum computing, and introduce the particular problem that motivates our interest in $\chi(\Omega(n))$. Our goal is to provide enough context to motivate our graph, so our treatment is far from complete. For a comprehensive study of the subject, we refer the reader to Nielsen and Chuang [43]. We will use the standard (for quantum computing) notation and terminology of [43].

6.2 Quantum Entanglement

Imagine that we have, in our idealized laboratory, two particles that have interacted. It is not important what exactly they are nor how they interacted, just that their future behaviour depends on this interaction. We assume that their behaviour is not completely determined (due to physical laws, imprecision in our apparatus, outside effects, or some other reason), but that we can predict their behaviour probabilistically. In principle, we could write down a probability distribution for the outcome of some future measurement.

Now our two laboratory assistants, Alice and Bob, take one of these particles each, and move very far apart. What happens when Alice "measures" her particle at a later date? It depends on what happened back in the lab, but it shouldn't depend on what Bob did with his particle after he left. The behaviours of the two particles are not unrelated, but the outcomes of the two subsequent measurements should be statistically independent.

In quantum physics the situation is more complicated. It is possible for two particles to be "entangled" to some degree. When Alice and Bob measure their particles, the results of these measurements are correlated beyond any similarities in the probability distributions which predict their behaviours. This is true even if the result of the measurement is unpredictable. In an extreme case, the measurement may be an essentially random process when looked at individually, yet the two measurements are perfectly correlated. It is as if they both flip fair coins which, though separated, always come up the same. This last situation can be thought of as a (gross) simplification of a "paradox" described by Einstein, Podolsky, and Rosen; they argued that quantum mechanics (as it was then understood) was an incomplete description of nature due to the "spooky" action at a distance that seems to be occurring.

The overwhelming mass of experimental evidence suggests that the classical

model is wrong: the predictions of quantum mechanics have been largely borne out. Spooky or not, the correlations are quite real.

Alice and Bob are separated and can not interact directly, but they do share something: the entanglement. It turns out that this can be a very useful thing to share. It is possible for Alice and Bob to turn this shared entanglement into a shared private encryption key, for instance. It is possible for them to use this to exchange a particle using so-called "quantum teleportation". If entanglement is such a powerful and useful part of quantum theory, then we might be motivated to ask if it can be simulated, and if so, how? This question will lead us to our graph.

6.3 Qubits

If our goal is to process information in some way, then we are not much interested in the precise details of any physical apparatus. It will be simpler to deal with an abstract particle, which we call a *qubit*. Of course, underneath this, a qubit must correspond to an actual quantum system, but for our purposes it will be sufficient to work at an abstract level.

The simplest particle is one for which there are only two possible measurement outcomes: call them 0 and 1. Quantum mechanical theory describes the state of such a particle as a complex vector of unit length in a 2-dimensional complex vector space. We can label the basis elements of this vector space as $|0\rangle$ and $|1\rangle$, corresponding to the possible measurement outcomes. Then our particle is in a state which we describe as

$$\alpha_{|0\rangle}|0\rangle + \alpha_{|1\rangle}|1\rangle,$$

where $\alpha_{|0\rangle}, \alpha_{|1\rangle} \in \mathbb{C}$. We cannot observe this vector directly (this is a fundamental tenet of quantum mechanics). What we can observe are the results of measurements. Moreover, there is a lack of consensus as to what the state-vector truly means. Rather, we are to interpret $|\alpha_{|0\rangle}|^2$ and $|\alpha_{|1\rangle}|^2$ as the probability that a measurement will result, respectively, in the outcome 0 or 1. So a measurement is fundamentally not deterministic: there is no way of knowing for certain what the outcome will be, even if we did know the quantum state exactly.

Note that 0 and 1 represent the possible outcomes of a measurement. There is no question of vector spaces here: either the answer is 0 or 1, and these are the only possible outcomes of a measurement. The result of a measurement on a qubit can be regarded as one bit (in the computer science sense of the word: hence the etymology of "qubit" is "quantum bit"). On the other hand, $|0\rangle$ and $|1\rangle$ are elements of a vector space, and there are infinitely many linear combinations of these vectors sitting in this vector space which could equally well represent the state of some qubit. It may help to think of a "measurement" as a probabilistic mapping from qubits to bits.

How does quantum mechanics represent the evolution of a physical system over time? Recall that the state of a qubit is a vector. Actual physical trans-

formations correspond exactly to multiplying this vector by a unitary matrix:

$$|\psi\rangle \to U|\psi\rangle.$$

Here we are considering time to be discrete, as opposed to continuous. Strictly speaking, U should be parametrized by time, but for our purposes we can ignore that. The result is that any physical process corresponds to some unitary matrix U, and any unitary matrix corresponds to a realizable physical process.

How do we deal with more than one qubit? The state of two unentangled qubits is the tensor product of their individual states. So if we had two particles, we could describe their state as (we use the shorthand $|ij\rangle$ for $|i\rangle \otimes |j\rangle$):

$$\begin{aligned}&\left(\alpha_{|0\rangle}|0\rangle + \alpha_{|1\rangle}|1\rangle\right) \otimes \left(\beta_{|0\rangle}|0\rangle + \beta_{|1\rangle}|1\rangle\right)\\ &\quad = (\alpha_{|0\rangle}\beta_{|0\rangle})|00\rangle + (\alpha_{|0\rangle}\beta_{|1\rangle})|01\rangle + (\alpha_{|1\rangle}\beta_{|0\rangle})|10\rangle + (\alpha_{|1\rangle}\beta_{|1\rangle})|11\rangle \qquad (6.1)\end{aligned}$$

If we read off from the right hand side the probability that the two qubits are in one of the four possible joint states, we see that this is exactly the product of their individual probabilities. In this way, (6.1) can be read as a statement about joint probabilities of two statistically independent events. Note that two qubits together define a four dimensional complex vector space: this corresponds to the fact that there are four possible measurement outcomes. More precisely, for a system of l qubits, the state vector lives in a complex vector space of dimension 2^l, and we will identify as a basis the elements

$$\left\{|i\rangle : i \in \{0,1\}^l\right\}.$$

How does entanglement come in? Consider the following state, known as a *Bell state*:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle.$$

It is impossible to write it as a tensor product of two qubits! This is the state of two qubits that are maximally entangled, as we shall see. Recall that a quantum state is not the sort of thing we can observe directly. Rather, we observe the result of a measurement. Two qubits, each of which can be in state 0 or 1, can in principle result in the observations 00, 01, 10, or 11. But for our Bell state, the probability of obtaining 01 or 10 is identically zero—it never happens. In other words, if Alice measures her qubit and the result is 0, then it must be the case that Bob would measure 0 for his qubit as well! The word "if" is very important: the fact that the qubit was measured as 0 does *not* mean that it "really was a 0 all along and so was the other one". It *could* have been measured as a 1, in which case so would have been the other. This is in fact one way of describing the Einstein-Podolsky-Rosen argument, and why Bell states are also called EPR pairs.

Our example above was motivated by the idea of one pair of entangled qubits, which is then separated. We can generalize this: start with a set of l pairs of qubits such that each pair is in the state $|\Phi^+\rangle$. Then Alice and Bob each take

one qubit from each pair (keeping track of which ones were which) and go their separate ways. Now they each have l qubits, and the total state is just

$$\left(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\right)^{\otimes l}.$$

It will be useful to reorder the basis elements, so as to write Alice's qubits "first". Then the above state is

$$\frac{1}{\sqrt{2^l}} \sum_{i\in\{0,1\}^l} |i\rangle \otimes |i\rangle. \tag{6.2}$$

Can they use this resource to do anything? Or, to put it another way, what kinds of things can they do with this resource that they could *not* do if they each had taken l unentangled classical bits instead?

One thing they could do is measure them. At first, this may seem rather pointless. If Alice measures one of her particles, she has a 50% chance of getting 0 and a 50% chance of getting 1: why wouldn't she just flip a coin? The answer is that if Alice and Bob both measure corresponding qubits then they are guaranteed to get the same result. Effectively, they have l pairs of perfectly correlated random variables at their disposal. We can actually go a little further: they can first "do" something to their respective qubits, and then measure them. Recall that any physical transformation corresponds to a unitary operator. If they are truly separated, then they can each act only on their qubits, meaning that they each choose a unitary matrix of order 2^l: call these matrices U_A and U_B. So the state would change according to

$$\begin{aligned}\frac{1}{\sqrt{2^l}} \sum_{i\in\{0,1\}^l} |i\rangle \otimes |i\rangle \to (U_A \otimes U_B) \left(\frac{1}{\sqrt{2^l}} \sum_{i\in\{0,1\}^l} |i\rangle \otimes |i\rangle\right) \\ = \frac{1}{\sqrt{2^l}} \sum_{i\in\{0,1\}^l} (U_A|i\rangle) \otimes (U_B|i\rangle)\end{aligned}$$

6.4 Quantum Cooperation

We now turn our attention to the application that motivates our study of $\Omega(2^l)$, and more precisely, how the entangled qubits of the previous section are relevant.

Alice and Bob face a challenge. They are each to be given a 01-vector of length 2^l, and they must each answer with a 01-vector of length l. There are conditions: if the challenge vectors are equal, then their responses must be equal too, and if the challenge vectors are at maximum Hamming distance, then their responses must be distinct. There is a catch: Alice and Bob are physically separated and can't communicate in any way at all. They are allowed to set up whatever strategy they want beforehand, but they must somehow arrange that they will be able to successfully answer every challenge. On the face of it, this seems hopeless. But Alice and Bob are experts on quantum

computing—actually, Alice and Bob are experts on just about everything—and they've read [10].

We will show how Alice and Bob can use entanglement to solve this challenge. Let z be a 01-vector of length 2^l: it may help to think of the components of z as being indexed by the basis elements, i.e., by the 01-vectors of length l. We consider the following family of unitary transformations, parametrized by z. We will define it on the basis elements; there is a unique linear extension to the whole vector space. First, map the basis elements according to:

$$|i\rangle \to (-1)^{z_i}|i\rangle.$$

This is clearly unitary: it corresponds to a diagonal matrix with diagonal entries ± 1. Next, apply the mapping

$$|i\rangle \to \sum_j (-1)^{i \bullet j}|j\rangle.$$

This is also clearly unitary: in fact, it corresponds to the unitary (Hadamard) matrix

$$U = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{\otimes l}.$$

So the combined effect is given by:

$$|i\rangle \to (-1)^{z_i} \sum_j (-1)^{i \bullet j}|j\rangle.$$

Assume that Alice and Bob start off with their entangled qubits in state (6.2), and are now unable to communicate in any way. Let x and y be two 01-vectors of length 2^l. These are the challenges: give x to Alice and y to Bob. If Alice applies the above transformation using x and Bob does the same using y, then the resulting state is:

$$\frac{1}{\sqrt{2^l}} \sum_{i \in \{0,1\}^l} |i\rangle \otimes |i\rangle \to \frac{1}{\sqrt{2^{3l}}} \sum_{i \in \{0,1\}^l} (-1)^{x_i + y_i} \sum_{j \in \{0,1\}^l} \sum_{k \in \{0,1\}^l} (-1)^{i \bullet (j+k)} |j\rangle \otimes |k\rangle$$

$$= \frac{1}{\sqrt{2^{3l}}} \sum_{j,k \in \{0,1\}^l} \sum_{i \in \{0,1\}^l} (-1)^{x_i + y_i} (-1)^{i \bullet (j+k)} |j\rangle \otimes |k\rangle$$

Now what happens when Alice and Bob each measure their qubits? Are they still correlated, and if so, how?

First, consider the case where $x = y$, that is, Alice and Bob both apply the same transformations to their qubits. Then the total state would be:

$$\frac{1}{\sqrt{2^{3l}}} \sum_{j,k \in \{0,1\}^l} \left(\left(\sum_{i \in \{0,1\}^l} (-1)^{i \bullet (j+k)} \right) |j\rangle \otimes |k\rangle \right).$$

Here the inner sum is 2^l if $j = k$ and 0 otherwise. So the state is in fact:

$$\frac{1}{\sqrt{2^l}} \sum_{j \in \{0,1\}^l} |j\rangle \otimes |j\rangle.$$

It is true that this is the same state they started with, but the real point is that Alice and Bob's measurements must (still) be perfectly correlated. This is perhaps not too surprising: they have essentially each applied the same change of basis to the same vector, resulting, oddly enough, in the same answer. But this means that when the two challenge vectors are the same, their answers (the final state of their qubits) will also be the same.

Now what happens when x and y differ in exactly half of their positions, that is, their Hamming distance is 2^{l-1}? Then the sum

$$\sum_{i \in \{0,1\}^l} (-1)^{x_i + y_i} (-1)^{i \bullet (j+k)}$$

is identically zero whenever $j = k$. What does this mean? Simply that the final state vector, when expressed in our chosen basis contains no element of the form $|j\rangle \otimes |j\rangle$. Thus it is *impossible* for Alice and Bob to get the same result when they measure their qubits. This means that when the challenges are at maximum Hamming distance, their answers are guaranteed to be different.

Qubits aside, we can think of Alice and Bob each being asked questions x and y (the 01-vectors of length 2^l) and responding with answers a and b (the 01-vectors of length l that are the results of their measurements). Their answers have the property that $a = b$ if $x = y$ and $a \neq b$ if x and y are at Hamming distance 2^{l-1}. There is no communication or exchange between them, although they of course share some entangled qubits.

6.5 Graph Colouring

The reader might do well to pause and consider how such a thing could occur without quantum effects. That is to say, could they accomplish the same thing if instead of sharing an initial resource of entangled qubits, they had some initial shared non-quantum information? Alice may agree that she will answer in a certain fashion, but this alone will not help, because it is not her answer but the relationship between their two answers that matters.

Buhrman, Cleve and Widgerson [10] have shown no classical scheme can accomplish this without some communication between Alice and Bob. In fact, they show that the amount of communication needed is of the order of a constant times 2^l, for sufficiently large l. Their result depends heavily on the previously mentioned result of Frankl and Rödl [25], which on the surface, has nothing whatsoever to do with quantum anything.

The key is to think of the problem in the following way. We consider a graph with vertex set the 01-vectors of length 2^l. Two vectors are adjacent if they are at Hamming distance 2^{l-1}. Alice and Bob are each given a vertex, and

they must answer such that their answers are the same if they are given the same vertex and different if they are given adjacent vertices. What information could they usefully agree on beforehand? Why, a proper colouring of the graph! Then they would each answer with the colour of the given vertex, and satisfy the given conditions exactly. Note that just as in the quantum scenario, if the two vertices are distinct and non-adjacent, the answers may or may not agree. There is just one small matter left: the answers are to be 01-vectors of length l. So in fact they can simulate the entanglement with a proper colouring of this graph—but only if the chromatic number of this graph is at most 2^l.

Thus we have the graph $\Omega(n)$ that we introduced before, and the motivation for our studying it. If $\chi(\Omega(n)) > n$ when n is a power of two, then we have a scenario where quantum entanglement is demonstrably a more powerful resource than classical information alone. From the point of view of our application, the idea of taking n not a power of two is meaningless since n is the dimension of the vector space obtained by taking the tensor product of l two-dimensional vector spaces. It will be convenient to use the parameter $n = 2^l$ rather than l, and this is why we have labelled our graphs as such. For the remainder of the chapter, we will deal exclusively with the problem as a question about the chromatic number of $\Omega(n)$.

6.6 Structure of $\Omega(n)$

For the purposes of our quantum application, we are only interested in the case where $n = 2^l$. However, there is nothing in the definition of the graph to force this restriction on us, so we start by taking a more general look.

If n is odd, then the graph is in fact edgeless, so the problem of finding a maximum independent set is trivial.

For $n \equiv 2 \pmod 4$, the situation is not much more difficult. Call a vertex *even* if it contains an even number of -1's; otherwise, it is *odd*. When $n \equiv 2 \pmod 4$, the graph is bipartite, with the even and odd vertices forming the bipartition. It follows that there are exactly two maximum independent sets: the even vertices and the odd vertices.

If n is a multiple of four, the situation is now somewhat reversed: no edge joins an even vertex to an odd vertex, so the graph has two components. Moreover, let σ be the map defined by:

$$\sigma : (x_1, x_2, \ldots, x_n) \mapsto (-x_1, x_2, \ldots, x_n).$$

It is straightforward to see that σ is an automorphism, and that σ exchanges even and odd vertices. Thus the even component and the odd component are isomorphic. We are interested in maximum independent sets so we can consider one component only.

Given a vertex $x = (x_1, \ldots, x_n)$, consider the vertex $-x = (-x_1, \ldots, -x_n)$. There is a natural relationship between these vertices: they are nonadjacent and have the same neighbourhood. So any maximum independent set containing x also contains $-x$. More precisely, $\Omega(n)$ is a lexicographic product of a graph

with $\overline{K_2}$. This is in fact true for any value of n. When n is a multiple of four, each component is a lexicographic product of a graph with $\overline{K_2}$. For each pair of even vertices $x, -x$, arbitrarily choose one of them. Call the graph induced by the chosen vertices Y. We have the following result (a similar result holds when $n \equiv 2 \pmod 4$, but we will not need it).

6.6.1 Lemma. *Let n be a multiple of four. Let Y be the graph induced by a set of vertices containing one vertex from each pair of even vertices x and $-x$. Then the graph $\Omega(n)$ is isomorphic to two disjoint copies of $Y[\overline{K_2}]$. Thus any maximum independent set in $\Omega(n)$ is just four copies of a maximum independent set in Y, and $4 \mid \alpha(\Omega(n))$.* □

6.7 Eigenvectors

Our proof of Theorem 6.1.1 rests on a characterization of the ratio-tight sets. In order to demonstrate this, we will need to know the eigenspace corresponding to the least eigenvalue.

We present a derivation of the eigenvectors (and hence eigenvalues) of $\Omega(n)$, based on automorphism of the graph and character theory. This section is based on the presentation of Godsil [31]. On the other hand, what we are really doing is determining the eigenspaces of the Hamming scheme, which are known objects. The method we are applying is in fact a standard approach that can be applied to any Cayley graph of an Abelian group. So the result is well-known, the technique is standard, and the entire section could be replaced by a statement of Corollary 6.7.4 and an appropriate reference to general results. We choose to include the derivation because it is relatively straightforward and demonstrates a method of obtaining eigenvectors for certain Cayley graphs. For more details, see Godsil [29, Section 12.8–12.9].

Let a be a vertex of $\Omega(n)$. Recall that $a \circ b$ means the Schur product of a and b. There is a natural automorphism σ_a associated with a, which we will call the *translation* by a, defined by

$$\sigma_a : x \mapsto a \circ x.$$

Given two vertices x, y, we note that $\sigma_{yx^{-1}}$ maps x to y and fixes no vertex. The set of all translations forms a subgroup of the automorphism group. The group of translations acts transitively on the vertices of $\Omega(n)$, and no non-identity translation has a fixed point: we say that the group of translations acts *regularly* on the graph. As a consequence of this, we can view $\Omega(n)$ as a Cayley graph for the group of translations. The vertex set is the set of all automorphisms σ_a, where $a \in \{\pm 1\}^n$, and the connection set is the set of all σ_a where $a^T \mathbf{1} = 0$. Note that the connection set is inverse-closed, since ${\sigma_a}^{-1} = \sigma_a$, so this a Cayley graph and not a Cayley digraph. We can use this to read off the eigenvectors of the graph: they will be the characters of the group of translations.

So we can view a vertex a of $\Omega(n)$ as a $\{\pm 1\}$ vector of length n, or as the automorphism σ_a. We can also view it as a subset of $\{1, \ldots, n\}$, corresponding

to the indices of the -1 entries. As a quick example, we have the following 3 equivalent ways of viewing a particular vertex in $\Omega(4)$:

$$a = (-1, 1, -1, -1),$$
$$a : (x_1, x_2, x_3, x_4) \mapsto (-x_1, x_2, -x_3, -x_4),$$
$$a = \{1, 3, 4\}.$$

This is not as bad as it seems: when needed, context will allow us to distinguish between a vector, a function, and a subset, and it will be useful to have these three representations in mind. The notation will usually make clear the representation: $a^T b$ would mean dot product of vector-vertices, ab would mean composition of group-element-vertices, and $a \cap b$ would mean intersection of subset-vertices. Representing vertices as ± 1 vectors means adjacency corresponds to orthogonality, which is convenient from a linear algebra perspective. The representation as group elements gives us the structure of a Cayley graph. The representation as sets will be a natural way to express the characters. Moreover, we could have represented vertices as 01-vectors: this is perhaps the traditional way of representing vertices in the Hamming scheme. The Schur product of two vertices as 01-vectors corresponds to the intersection of the two vertices as sets; the dot product, to the size of their intersection. We will not do this, since this would lead to truly ambiguous notation, but the reader might find it helpful to think of intersection this way.

Let S be some subset of $\{1, \ldots, n\}$ (i.e., a vertex of $\Omega(n)$). Define the function ψ, which maps vertices to complex numbers, by:

$$\psi_S(\sigma_a) = (-1)^{|S \cap a|}.$$

On the right hand side, we are of course thinking of the vertex a as a subset. Let $a \Delta b$ denote the symmetric difference of a and b. It is straightforward to show that ψ is in fact a character for the group of translations, since

$$\begin{aligned}
\psi_S(\sigma_a \sigma_b) &= \psi_S(\sigma_{a \Delta b}) \\
&= (-1)^{|S \cap (a \Delta b)|} \\
&= (-1)^{|S \cap a|} (-1)^{|S \cap b|} \\
&= \psi_S(\sigma_a) \psi_S(\sigma_b).
\end{aligned}$$

Any two distinct characters of an Abelian group are orthogonal. We can see that for our example directly, and in fact a general proof is not much more difficult. For clarity, we will use "a" for a vertex rather than "σ_a", even though we are thinking of $\Omega(n)$ as a Cayley graph whose vertices are group elements.

$$\langle \psi_R, \psi_s \rangle = \sum_{a \in V(\Omega(n))} \psi_R(a) \psi_S(a) = \sum_a \psi_{R \Delta S}(a).$$

On the other hand, let g be any vertex of $\Omega(n)$, viewed as a translation. Then

$$\sum_a \psi_{R \Delta S}(a) = \sum_a \psi_{R \Delta S}(ga) = \psi_{R \Delta S}(g) \sum_a \psi_{R \Delta S}(a).$$

If $\psi_{R\Delta S}$ is not the trivial character (i.e., if $R \neq S$), then

$$\langle \psi_R, \psi_s \rangle = \sum_a \psi_{R\Delta S}(a) = 0.$$

We have an orthogonal set of group characters that is equinumerous with the set of group elements, so in fact we have all the characters (see for example [29, Section 12.8]).

For a character ψ and a set $C \subset G$, define

$$\psi(C) = \sum_{c \in C} \psi(c).$$

The framework we have developed so far is useful because of the following result, which we paraphrase from Godsil [29, p.246].

6.7.1 Lemma. *Let X be a Cayley graph for an Abelian group G, with inverse-closed connection set $C \subset G$. Then the characters of G form a complete set of eigenvectors for the adjacency matrix of X. The eigenvalue corresponding to ψ is $\psi(C)$.*

Proof. The proof is straightforward. Let ψ be a character of G and A be the adjacency matrix of X. Then

$$(A\psi)(g) = \sum_{h \sim g} \psi(h) = \sum c \in C \psi(cg) = \psi(C)\psi(g). \qquad \square$$

The consequence is that we know the eigenvectors and eigenvalues of $\Omega(n)$.

6.7.2 Corollary. *Let $n = 2m$. Then a complete set of eigenvectors for $\Omega(n)$ is given by the characters ψ_S, where S ranges over all subsets of $\{1, \ldots, n\}$, and*

$$\psi_S(a) = (-1)^{|S \cap a|}. \qquad \square$$

6.7.3 Lemma. *Let $n = 2m$. Then for $1 \leq r \leq n$, the eigenvalues of $\Omega(n)$ are*

$$\lambda_r = \frac{2^m}{m!}(1-r)(3-r)\cdots(2m-1-r).$$

Proof. We already know a complete set of linearly independent eigenvectors. These are the characters ψ_S. Moreover, we already know that the eigenvalue depends only on $|S|$, so this gives $n+1$ subspaces each contained in an eigenspace. (As a corollary, we see that there are at most $n+1$ eigenvalues: we already knew this of course, since the Hamming scheme has n classes.) It remains only to compute the eigenvalues, which we can do directly.

If $S = \emptyset$, then

$$\psi_\emptyset(C) = \sum_{c \in C} \psi_\emptyset(c) = |C| = \binom{2m}{m}.$$

For general S with $|S| = r$, we can determine the eigenvalues by splitting up the sum:

$$\psi_S(C) = \sum_{c \in C} \psi_S(c) = \sum_{j \geq 0} \sum_{c \,:\, |S \cap c| = j} (-1)^j$$
$$= \sum_{j \geq 0} (-1)^j \binom{r}{j} \binom{2m-r}{m-j}. \tag{6.3}$$

If r is odd, then we find that the j-th term and the $(r-j)$-th term cancel, so the summation is zero. Now viewing r as a variable, the summands in (6.3) are each polynomials in r of degree m, and there are only finitely many summands, since $0 \leq j \leq m$. Moreover, the summands corresponding to $j > 0$ have constant term equal to zero, and the constant term in the $j = 0$ summand is $\binom{2m}{m}$. Thus the summation is a polynomial of degree m, with roots $1, 3, \ldots, 2m-1$, and constant term $\binom{2m}{m}$. This uniquely specifies the polynomial, and thus

$$\sum_{j \geq 0} (-1)^j \binom{r}{j} \binom{2m-r}{m-j} = \frac{2^m}{m!}(1-r)(3-r)\cdots(2m-1-r). \tag{6.4}$$

□

We can glean more information from the previous result. The multiplicity of 0 as an eigenvalue is 2^{n-1} (this is the number of odd subsets of an n-set). Using (6.3), one can show that

$$\lambda_{2i} = (-1)^m \lambda_{n-2i},$$

and thus $\lambda_{2i} = \pm\lambda_{n-2i}$, depending on whether m is even or odd. It follows from (6.4) that the values λ_{2i} are distinct for $2i \leq \frac{n}{2}$; in fact:

$$|\lambda_0| > |\lambda_2| > \cdots > |\lambda_m|$$

and the signs alternate.

When m is odd, the non-zero spectrum consists of the eigenvalues λ_{2i}, for $0 \leq i \leq 2m$, with multiplicities $\binom{2m}{2i}$. The least eigenvalue is $\lambda_{2m} = -\lambda_0$. When m is even, the non-zero spectrum consists of the eigenvalues λ_{2i}, for $0 \leq i < m$, with multiplicities $2\binom{2m}{2i}$, as well as λ_m with multiplicity $\binom{2m}{m}$. The least eigenvalue is $\lambda_2 = \lambda_{2m-2}$.

From a practical perspective, we are mainly interested in the following corollary. As mentioned at the beginning of this section, this can be regarded as a corollary of our knowledge of the eigenspaces of the Hamming scheme.

6.7.4 Corollary. *Let n be a multiple of four. The least eigenvalue of $\Omega(n)$ is*

$$\tau = \lambda_2 = -\frac{1}{n-1}\binom{n}{\frac{n}{2}}.$$

Moreover, let U be the matrix whose rows are indexed by vertices, whose columns are indexed by sets $S \subset \{1, \ldots, n\}$ of size 2 or $n-2$, and whose (a, S) entry is $(-1)^{|a \cap S|}$. Then the columns of U form a basis for the τ-eigenspace. □

6.8 Independent Sets and Chromatic number

Using the results of the previous section, we can apply the ratio bound to our graphs.

6.8.1 Corollary. *Let n be a multiple of four. Then*

$$\alpha(\Omega(n)) \leq \frac{2^n}{n}.$$ □

This means that $\chi(\Omega(n)) \geq n$. It also means that if equality is to hold, then n must divide 2^n. We have previously dealt with the cases where n is not a multiple of four. Now we see that if n is a multiple of four that is not a power of two then $\chi(\Omega(n)) > n$. It only remains to consider the case where n is a power of two, i.e., the case that is relevant to our application. This is not to say that what we have established up till now is irrelevant. Rather, the most interesting case with regard to determining if $\chi(\Omega(n)) = n$ seems to be exactly the hardest case. It will also be the case for which we know the most different derivations of the ratio bound.

There is one other result that will be useful. The first part is well-known; the equality condition is due to Godsil [31]. We state it here (in a restricted form) without proof.

6.8.2 Lemma. *Let X be a Cayley graph for an Abelian group. Then*

$$\alpha(X)\,\omega(X) \leq |V(X)|.$$

Moreover, if equality holds, then $\omega(X) = \chi(X)$. □

Why is this useful? Consider the set of ± 1 vectors corresponding to a clique. This a set of orthogonal (hence linearly independent) vectors of length n, so $\omega(\Omega(n)) \leq n$. If $\omega(\Omega(n)) = n$ then the set of vectors of a clique gives a Hadamard matrix of order n. Hadamard matrices of order n do exist whenever n is a power of two, and thus in these cases $\omega(\Omega(n)) = n$. Combining this observation with Lemma 6.8.2 gives another proof of the ratio bound result of Corollary 6.8.1 for n a power of two. Moreover, we can say something about the case where the ratio bound holds with equality.

6.8.3 Corollary. *Let n be a power of two. Then*

$$\chi(\Omega(n)) = n \iff \alpha(\Omega(n)) = \frac{2^n}{n}.$$ □

We have shown that the colouring problem posed is equivalent to finding the size of a maximum independent set. More precisely, we need to determine when the ratio bound holds with equality. Certainly bounds on maximum independent sets give bounds on chromatic numbers. But in general, one does not expect such a tight relationship.

Independent sets whose size meets the ratio bound are significant objects. Their structure (or absence!) has important consequences for the graphs we are studying. As we will see in the next section, we are not yet done extracting information about what happens in the equality case.

6.9 Recursive Structure

Before we deal with finding actual independent sets, we further explore the structure of $\Omega(n)$. There is a natural recursive structure that will be very useful to us, and conveniently, it applies in exactly the cases where n is a power of two.

The vertex set of $\Omega(2n)$ consists of all the ± 1 vectors of length $2n$. It will be convenient to represent these vertices in a precise fashion. Let "$a|b$" denote concatenation of the two vectors a and b and "$a \circ b$" denote their Schur product, with concatenation having the lowest precedence of any operator. Thus if a, b, c are all vectors of length n, then $a|b \circ c$ means the vector of length $2n$ whose first n entries are exactly a and whose last n entries are exactly the Schur product of b and c. Then every vertex in $\Omega(2n)$ has a unique representation of the form $i|r \circ i$, where i and r are vertices in $\Omega(n)$. We can regard the $i|r \circ i$ representation of vertices of $\Omega(2n)$ as a coordinatization by two copies of $\Omega(n)$, indexed by i and r.

6.9.1 Lemma. *For any fixed $r \in V(\Omega(n))$, let X_r be the subgraph of $\Omega(2n)$ induced by the vertices of the form $x|r \circ x$. Then X_r is isomorphic to $\Omega(n)$. Thus the vertices of $\Omega(2n)$ can be partitioned into 2^n sets, each of which induces a copy of $\Omega(n)$.*

Moreover, every vertex of X_r is adjacent to every vertex of X_{-r}.

Proof. Consider automorphisms of the form

$$\sigma_{\mathbf{1}|t \circ \mathbf{1}} \quad \text{and} \quad \sigma_{t|\mathbf{1} \circ t}.$$

The first permutes copies of X_r, while the second fixes each X_r as a set and permutes vertices within it. So the automorphism group of $\Omega(2n)$ acts transitively on the copies of X_r, and transitively within the copies of X_r fixing each X_r setwise. Informally, we may say that $\Omega(2n)$ is transitive on each coordinate.

Proving the claim is a straightforward matter of checking adjacency conditions between pairs of vertices. By the above remarks on automorphisms, it suffices to show that $X_{\mathbf{1}}$ is isomorphic to $\Omega(n)$, and that every vertex in $X_{\mathbf{1}}$ is joined to every vertex in $X_{-\mathbf{1}}$.

First, let r be fixed, and consider the vertices $\mathbf{1}|\mathbf{1} \circ \mathbf{1}$ and $x|\mathbf{1} \circ x$.

$$\begin{aligned} (\mathbf{1}|\mathbf{1} \circ \mathbf{1})^T (x|\mathbf{1} \circ x) = 0 &\iff \mathbf{1}^T x + (\mathbf{1} \circ \mathbf{1})^T (\mathbf{1} \circ x) = 0 \\ &\iff \mathbf{1}^T x + \mathbf{1}^T x = 0 \\ &\iff \mathbf{1}^T x = 0 \end{aligned}$$

Thus $\mathbf{1}|\mathbf{1}\circ\mathbf{1}$ and $x|\mathbf{1}\circ x$ are adjacent in $\Omega(2n)$ if and only if $\mathbf{1}$ and x are adjacent in $\Omega(n)$. So the set of vertices corresponding to a fixed choice of r gives a copy of $\Omega(n)$.

Now consider the vertices $\mathbf{1}|\mathbf{1}\circ\mathbf{1}$ and $x|r\circ x$.

$$\begin{aligned}(\mathbf{1}|\mathbf{1}\circ\mathbf{1})^T\,(x|r\circ x)=0 &\iff \mathbf{1}^Tx+(\mathbf{1}\circ\mathbf{1})^T(r\circ x)=0\\ &\iff \mathbf{1}^Tx+\mathbf{1}^T(r\circ x)=0\\ &\iff x^T(\mathbf{1}+r)=0 \qquad (6.5)\end{aligned}$$

So if $r=-\mathbf{1}$ then $\mathbf{1}|\mathbf{1}\circ\mathbf{1}$ and $x|r\circ x$ are adjacent in $\Omega(2n)$. □

What we have shown is sufficient for our purposes, but we investigate a little further. Obviously the condition $r=-\mathbf{1}$ given at the end of the proof is not necessary. Viewing vertices as subsets, we may rewrite (6.5) as

$$|x|+|r\Delta x|=n.$$

So a counting argument shows that, for a fixed x, there are

$$\sum_{t=0}^{|x|}\binom{|x|}{t}\binom{2n-|x|}{n-2|x|+t}$$

distinct r such that $\mathbf{1}|\mathbf{1}\circ\mathbf{1}$ and $x|r\circ x$ are adjacent in $\Omega(2n)$. Setting $y=r\circ x$ gives $x=r\Delta y$, so a symmetry argument shows that for a fixed r, there are the same number of x such that $\mathbf{1}|\mathbf{1}\circ\mathbf{1}$ and $x|r\circ x$ are adjacent in $\Omega(2n)$. This can be expanded into a rather technical condition on exactly when $\mathbf{1}|\mathbf{1}\circ\mathbf{1}$ and $x|r\circ x$ are adjacent, but we will not need this. Suffice to say that there are "extra" edges unaccounted for in the statement of this lemma. None of these edges lies within one of the graphs X_r, and none of them lie between X_r and X_{-r} (indeed, there is no more room to put any edges in this last category, since they are all already there). We will have more to say about these "extra" edges.

The immediate usefulness of Lemma 6.9.1 is in the following corollary. The *join* of two graphs X and Y, written as $X\vee Y$, is the graph whose vertex set is $V(X)\cup V(Y)$ and whose edge set is $E(X)\cup E(Y)\cup\{xy : x\in V(X), y\in V(Y)\}$.

6.9.2 Corollary. *Let n be a power of two. The graph $\Omega(2n)$ contains a spanning subgraph isomorphic to the disjoint union of 2^{n-1} copies of $\Omega(n)\vee\Omega(n)$.* □

The restriction of an independent set to a subgraph is still an independent set, so we can use this to bound independent sets in $\Omega(2n)$ in terms of independent sets in $\Omega(n)$. Furthermore, any independent set in $X\vee Y$ is contained entirely in X or Y. Putting this together, we get a bound on independent sets, which is the ratio bound. This time we get structural information when the bound holds with equality.

6.9.3 Lemma. *Let n be a power of two. Then*

$$\alpha(\Omega(n)) \leq \frac{2^n}{n}.$$

If equality holds, and $n > 1$, then the maximum independent sets are necessarily unions of maximum independent sets in $2^{\frac{n}{2}-1}$ vertex-disjoint subgraphs isomorphic to $\Omega(\frac{n}{2})$. Moreover, these $2^{\frac{n}{2}-1}$ disjoint subgraphs are induced subgraphs.

Proof. Let $n = 2^l > 1$. We first observe that the recursive structure of Corollary 6.9.2 gives that

$$\alpha\left(\Omega(2^l)\right) \leq \frac{2^{2^{l-1}}}{2}\alpha\left(\Omega(2^{l-1})\right).$$

Applying this l times, we get

$$\begin{aligned}\alpha\left(\Omega(2^l)\right) &\leq \frac{2^{2^{l-1}}}{2}\frac{2^{2^{l-2}}}{2}\cdots\frac{2^{2^1}}{2}\alpha\left(\Omega(1)\right)\\ &= \frac{2^{2^l-1}}{2^l}\alpha\left(\Omega(1)\right)\\ &= \frac{2^{n-1}}{n}\alpha\left(\Omega(1)\right)\end{aligned}$$

It remains only to show that $\alpha\left(\Omega(1)\right) = 2$, which we leave as an exercise. □

For n a power of two, this is the third derivation of this bound: once by eigenvalues (Corollary 6.8.1), once using cliques (comments following Lemma 6.8.2), and now once using a recursive structure. In each case, we get extra information when the bound holds with equality. It is worth summarizing this information.

6.9.4 Lemma. *Let $n > 1$ be a power of two. Let S be an independent set in $\Omega(n)$, with characteristic vector z. Then the following are equivalent.*

(a) $|S| = \frac{2^n}{n}$

(b) $z - \frac{|S|}{2^n}\mathbf{1} = z - \frac{1}{n}\mathbf{1}$ *is an eigenvector for the least eigenvalue*

(c) $\chi(\Omega(n)) = n$

(d) There are 2^n disjoint copies of $\Omega(n/2)$ such that S is disjoint from half of them, and the restriction of S to each of the remaining copies is a maximum independent set of size $\frac{2^{n/2}}{n/2}$. □

This is perhaps surprising. As an example, assume that $\chi(\Omega(2n)) = 2n$. Then there is a partition Π of $\Omega(2n)$ into 2^n disjoint induced copies of $\Omega(n)$ such that the following condition holds: every 01-vector that is a linear combination of $\mathbf{1}$

and a τ-eigenvector for $\Omega(2n)$ is partitioned by Π into 2^{n-1} zero vectors and 2^{n-1} vectors that are a linear combination of $\mathbf{1}$ and a τ-eigenvector for $\Omega(n)$.

We have made no serious effort to show the equivalence of (b), (c) and (d) directly. This is partly because each of these is easily equivalent to (a), but also partly because we wish to emphasize that they are consequences of the ratio bound being tight.

We will actually make use of all of the information in Lemma 6.9.4. Condition (c) tells us that results about ratio-tight independent sets are exactly answers to the colouring problem. We will use (b) to determine actual independent sets in Section 6.11. Finally, it is (d) that will allow us to leverage our results into an answer for all n. We close this section with that result now, which follows from Lemma 6.9.3.

6.9.5 Corollary. *There exists $l_0 \in \{\mathbb{N} \cup \infty\}$ such that $\chi\left(\Omega(2^l)\right) = 2^l$ for all $0 \leq l \leq l_0$ and $\chi\left(\Omega(2^l)\right) > 2^l$ for all $l > l_0$.* □

While it is true that the result as stated follows from Frankl and Rödl's work [25], we will determine n_0, and indeed, characterize every ratio-tight independent set in all graphs $\Omega(n)$.

6.10 Examples

Before we proceed we will use the ideas of Corollary 6.9.2 to construct some graphs.

We first give the graph $\Omega(1)$. It has two vertices, corresponding to the vectors (1) and (-1). They are not adjacent, meaning that $\Omega(1)$ is just $\overline{K_2}$. Thus $\alpha\left(\Omega(1)\right) = 2$, and there is a unique maximum independent set.

We turn our attention to the marginally less trivial $\Omega(2)$. It contains a spanning subgraph consisting of $2^{\frac{2}{2}-1} = 1$ copy of $\Omega(1) \vee \Omega(1)$. This is in fact the whole graph, because any "extra" edges would have to be between different copies of $\Omega(1) \vee \Omega(1)$, and there aren't different copies. So $\Omega(2) \cong \Omega(1) \vee \Omega(1)$. Thus we see clearly that the maximum independent sets in $\Omega(2)$ are necessarily maximum independent sets in one of the two copies of $\Omega(1)$. (We also observe that $\Omega(2) \cong K_2[\overline{K_2}]$, following the ideas of Lemma 6.6.1 and the paragraph preceding it; alternatively, we could just write $\Omega(2) \cong C_4$ and be done with it.) In any case, there are exactly two maximum independent sets. This corresponds to a choice of which half of $\Omega(1) \vee \Omega(1)$ to put the independent set in.

Now consider $\Omega(4)$. It contains a spanning subgraph consisting of $2^{\frac{4}{2}-1} = 2$ copies of $\Omega(2) \vee \Omega(2)$. But recall that when n is a multiple of four, $\Omega(n)$ has an even component and an odd component: these are the two copies of $\Omega(1) \vee \Omega(1)$, and hence there are no edges between them, and hence $\Omega(4) \cong 2\left(C_4 \vee C_4\right)$. By the recursive structure, we see that the maximum independent sets are always two copies of maximum independent sets in $\Omega(4)$, hence have size four. Furthermore, we know all of them: there are four possible maximum independent sets, corresponding to the two choices for each copy of $\Omega(2) \vee \Omega(2)$.

We come to $\Omega(8)$. As usual we find a spanning subgraph consisting of eight copies of $\Omega(4) \vee \Omega(4)$. There are two components, so each of these contains as a spanning subgraph 4 copies of $\Omega(4) \vee \Omega(4)$. But now within each component, there may well be (in fact, there are) more edges than we have listed so far. The previous three examples were all small enough to do by hand, but they also all follow directly from the recursion. The graph $\Omega(8)$ is a little too big to do with pencil and paper, but more importantly, it does *not* follow trivially from the recursion.

We content ourselves with one further observation at this point. Assume that there is a ratio-tight independent set in $\Omega(8)$. Then it must consist of eight copies of ratio-tight independent sets in $\Omega(4)$. There is a maximum of 16 ratio-tight independent sets in each component, and if there are in fact 16, then it must be the case that for each copy of $\Omega(4) \vee \Omega(4)$ we may arbitrarily choose which half to put the independent set in.

6.11 No Independent Sets

At this point, our approach is straightforward. We want to know when $\chi(\Omega(n)) = n$. It is sufficient to determine when the ratio bound holds with equality. This is equivalent to characterizing the 01-vectors in the matrix U of Corollary 6.7.4. It turns out the methods that we used on the partition graph in Chapter 4 and in proving EKR in Chapter 5 are applicable here also. There is a twist: we don't want to find all ratio-tight independent sets, we want to show there aren't any. Here we are using a structural result (the equality condition of the ratio bound) to show absence.

Before embarking on a matrix algebra adventure, we reduce our computational expense somewhat.

Recall from Corollary 6.7.4 that when n is a multiple of four, there are two kinds of τ-eigenvectors for $\Omega(n)$: those corresponding to ψ_S where S is either a 2-set or an $(n-2)$-set. Let S be a 2-set, and $\overline{S}$ the complementary $(n-2)$-set. Then in fact,

$$\psi_S(a) = (-1)^{|a|}\psi_{\overline{S}}(a).$$

So the eigenvectors corresponding to 2-sets are equal to the eigenvectors corresponding to $(n-2)$-sets on the even vertices, and the negative of those eigenvectors on the odd vertices. Let D be the diagonal matrix indexed by 2-sets whose diagonal elements are equal to -1 in the column corresponding to 2-sets containing the element 1, and equal to $+1$ otherwise. It follows that the matrix U of Corollary 6.7.4 is of the form

$$\begin{pmatrix} U_0 & U_0 \\ U_0 D & -U_0 D \end{pmatrix}.$$

Thus the columns of the matrix U_0 form a basis for the τ-eigenspace for the even component of $\Omega(n)$. We can use the lexicographic product of Lemma 6.6.1 to reduce the size still further; this amounts to taking half of the rows of U_0.

For our computations, we used the following result, which follows from the preceding comments.

6.11.1 Corollary. *Let n be a multiple of four, and let U_0 be the matrix whose rows are indexed by the even vertices, whose columns are indexed by 2-sets, and whose (a,p)-entry is equal to $(-1)^{|a\cap p|}$. Let H be the matrix obtained by adjoining $\mathbf{1}$ to U_0. Then the characteristic vector of every maximum independent set that meets the ratio bound lies in the column space of H.* □

Let S be a maximum independent set in $\Omega(n)$, and z its characteristic vector. As when we characterized the maximum independent sets for the partition graph and the Kneser graphs, we observe that we may arbitrarily assume that any particular vertex—say $\mathbf{1}$—is in S, and thus that the neighbourhood of $\mathbf{1}$ is disjoint from S. So let N be the submatrix of H whose rows correspond to the neighbourhood of $\mathbf{1}$, and let M be the submatrix of U_0 whose rows correspond to the neighbourhood of a (so $N = \begin{pmatrix} M & \mathbf{1} \end{pmatrix}$). We are, as before, interested in the kernel of N.

Let B be the incidence matrix of K_n. Then a simple computation shows that

$$MB^T = -\mathbf{1}.$$

This means that

$$N \begin{pmatrix} B & \mathbf{1} \end{pmatrix}^T = \mathbf{0},$$

and so the column space of $\begin{pmatrix} B & \mathbf{1} \end{pmatrix}^T$ is contained in the kernel of N. We will show in fact that it is the kernel of N. We will need two technical results for this.

6.11.2 Lemma. $\mathrm{rk}(B) = n$.

Proof. Observe that $BB^T = (n-1)I + J$. Since I is positive definite and J is positive semi-definite, BB^T is positive definite, and hence has full rank. So $\mathrm{rk}(B) = \mathrm{rk}(BB^T) = n$.

6.11.3 Lemma. $\mathrm{rk}(M) = \binom{n}{2} - n + 1$.

Proof. We again consider the matrix M^TM. This time we will need to be more explicit and determine the actual eigenvalues, much as we did for Lemma 5.4.3 and Lemma 5.5.3. (We could have done this for Lemma 6.11.2 too, but didn't need to.)

The rows and columns of M^TM are indexed by 2-sets. The (ij,kl)-entry of M^TM is equal to the number of vertices (i.e., $\frac{n}{2}$-sets) whose intersections with $\{i,j\}$ and with $\{k,l\}$ have the same parity, minus the number of vertices whose intersections with $\{i,j\}$ and with $\{k,l\}$ have the opposite parity. This depends only on $|\{i,j\} \cap \{k,l\}|$.

For $|\{i,j\} \cap \{k,l\}| = 2$, this is easy: every vertex intersects $\{i,j\}$ and $\{k,l\}$ in the same parity, since these are the same set. So, defining $c_0 = \left(M^T M\right)_{ij,ij}$, we have:

$$c_0 = \binom{n}{\frac{n}{2}}.$$

For $|\{i,j\} \cap \{k,l\}| = 1$, let $i = k$ and $j \neq l$. Any vertex disjoint from $\{i,j,l\}$ will contribute $+1$. Of the vertices containing exactly one of $\{i,j,l\}$, two of these contribute -1 and one contributes $+1$. Of the vertices containing exactly two of $\{i,j,l\}$, two of these contribute -1 and one contributes $+1$. Vertices containing all of $\{i,j,l\}$ contribute $+1$. If we define $c_1 = \left(M^T M\right)_{ij,il}$ then

$$\begin{aligned} c_1 &= \binom{n-3}{\frac{n}{2}} - \binom{n-3}{\frac{n}{2}-1} - \binom{n-3}{\frac{n}{2}-2} + \binom{n-3}{\frac{n}{2}-3} \\ &= \binom{n}{\frac{n}{2}} - 8\binom{n-3}{\frac{n}{2}-1} \end{aligned}$$

For $|\{i,j\} \cap \{k,l\}| = 0$ we define $c_2 = \left(M^T M\right)_{ij,kl}$, and a similar argument shows that

$$\begin{aligned} c_2 &= \binom{n-4}{\frac{n}{2}} - 4\binom{n-4}{\frac{n}{2}-1} + 6\binom{n-4}{\frac{n}{2}-2} - 4\binom{n-4}{\frac{n}{2}-3} + 6\binom{n-4}{\frac{n}{2}-4} \\ &= \binom{n}{\frac{n}{2}} - 16\binom{n-4}{\frac{n}{2}-1} \end{aligned}$$

Let L be the adjacency matrix of the line graph of K_n and $\overline{L} = J - I - L$ the adjacency matrix of the complement of the line graph of K_n. Then we have shown that

$$M^T M = c_0 I + c_1 L + c_2 \overline{L}$$

The matrices $I, L, \overline{L}$ form an association scheme; more precisely, the line graph of K_n is strongly regular. The matrix of eigenvalues of this scheme is well known. It follows from Lemma 2.5.4, or even Lemma 2.2.3, that we can write down the eigenvalues of $M^T M$.

We could have just "guessed" and multiplied $M^T M$ by a complete set of eigenvectors for L, and watched the eigenvalues of $M^T M$ emerge. This is correct, but it misses the point. Multiplying an eigenvector by $M^T M$ is multiplying a vector of size $\binom{n}{2}$ by a matrix of size $\binom{n}{2}$, whereas using the matrix of eigenvalues, we need only multiply a vector of size 3 by a matrix of size 3. It is computationally much *easier* to compute the eigenvalues of a matrix in the Bose-Mesner algebra than it is to multiply that matrix by a known eigenvector.

We conclude that the eigenvalues of $M^T M$ are:

$$\frac{n}{2(n-1)}\binom{n}{\frac{n}{2}}, \qquad \frac{n(n-2)}{(n-1)(n-3)}\binom{n}{\frac{n}{2}}, \qquad 0.$$

with respective multiplicities 1, $\binom{n}{2} - n$, and $n - 1$.

Since $\mathrm{rk}(M) = \mathrm{rk}(M^T M)$, the result follows. □

Notice that $\mathbf{1}$ is in the column space of M, since

$$M\mathbf{1} = -\frac{n}{2}\mathbf{1}.$$

Thus $\mathrm{rk}(N) = \mathrm{rk}(M)$. Using Lemma 6.11.2 and Lemma 6.11.3, it follows that the kernel of N is the column space of the matrix

$$\widehat{B} = \begin{pmatrix} B & \mathbf{1} \end{pmatrix}^T.$$

We have the following result:

6.11.4 Corollary. *Let n be a multiple of four. Let S be a maximum independent set that meets the ratio bound, and z its characteristic vector. Then $z = Hy$ where $y = \widehat{B}x$, for some vector x.*

Proof. As above, we may assume that the set S contains, say, the vertex $\mathbf{1}$, and thus that it is disjoint from the neighbourhood of $\mathbf{1}$. Thus z is zero on the rows corresponding to neighbours of $\mathbf{1}$, meaning that y is in the kernel of N, meaning that $y = \widehat{B}x$ for some vector x. □

Now let C be the reduced column echelon form of $H\widehat{B}$ (discarding any zero columns). Then z is still a linear combination of the columns of C, say $z = Cx'$. But C contains a set of rows that form an identity matrix. So the fact that z is a 01-vector forces x' to be a 01-vector as well. Thus there are $2^{\mathrm{rk}(C)}$ possible values for x'.

Using `Maple`, we compute the matrix C for $n = 8$ and discover that there are eight ratio-tight independent sets that contain the vertex $\mathbf{1}$. There are four copies of $\Omega(4) \vee \Omega(4)$ in the even component of $\Omega(8)$. So we could have as many as $2^4 = 16$ independent sets; equivalently, as many as 2^3 independent sets that contain $\mathbf{1}$. It follows that there are 16 distinct maximum independent sets in the even component, and thus 2^8 maximum independent sets in $\Omega(8)$. Thus, following the comments in Section 6.10, we see that the recursive construction of an independent set always works for $\Omega(8)$: we are free to choose either half of each copy of $\Omega(4) \vee \Omega(4)$.

Now we repeat the same computations for $n = 16$. This time, however, we discover that there are no 01-vectors at all in the column space of H. This gives the following result:

6.11.5 Theorem. *Let n be a power of two. Then $\alpha(\Omega(n)) = \frac{2^n}{n}$ for $n \leq 8$ and there exists $c < 1$ such that $\alpha(\Omega(n)) < c\frac{2^n}{n}$ for $n > 8$.*

Proof. For $n \leq 8$ we have already characterized the maximum independent sets. Recall from Lemma 6.6.1 that $\alpha(\Omega(n))$ is a multiple of four. Using the

proof of Lemma 6.9.3, we see that

$$\begin{aligned}\alpha((\Omega(n)) &\leq \frac{\alpha(\Omega(16))}{\frac{2^{16}}{16}}\ \frac{2^n}{n} \\ &\leq \frac{\frac{2^{16}}{16}-4}{\frac{2^{16}}{16}}\ \frac{2^n}{n} \\ &\leq \left(1-\frac{1}{2^{10}}\right)\frac{2^n}{n}.\end{aligned}$$

Thus $c \leq 1 - 2^{-10} \approx 0.999$. □

Now using Theorem 6.11.5 and Corollary 6.8.3, Theorem 6.1.1 follows. Note that we can improve the value of c using Lemma 6.1.2; this gives $c \leq 0.955$.

6.12 Recent Developments

This is a brief summary of further results that have been obtained: a computation of De Klerk and Pasechnik that determines $\alpha(\Omega(16))$, and an equivalence between different maximal independent sets.

A very recent unpublished computation by De Klerk and Pasechnik gives an upper bound of 2304 on the size of an independent set in $\Omega(16)$. They used a method of Schrijver [47], which can be seen an extension of Delsarte's linear programming bound, Lemma 3.2.1, to a larger algebra.

Recall that the ratio bound is a special case of Delsarte's linear programming bound, Lemma 3.2.1. However, for the graphs $\Omega(n)$, $n \leq 64$, we have checked that this gives the same value as the ratio bound. We do not yet have a proof of this in general, but it seems likely. The method of Schrijver gives a smaller bound than the ratio bound for $n \geq 12$ (recall that we know the for sufficiently large n ratio bound is exponentially too big). So it is the strongest known tool for computing bounds on independent sets in $\Omega(n)$.

Galliard [27] gave a construction of an independent set in $\Omega(2^k)$ and conjectured that it was maximum. For $n = 16$ his set is of size 2304. At the time, there was little evidence either way, but now we see that $n = 16$ he was correct. According to De Klerk and Pasechnik, for $n = 20$ and $n = 32$ the method of Schrijver yields a bound that is greater than his construction.

Recall that for n a multiple of four, $\Omega(n)$ consists of two disjoint copies of $Y(n)[\overline{K_2}]$. The vertex set of $Y(n)$ can be conveniently obtained by pairing each vertex with its complement. Choosing one vertex from each pair of even vertices gives a copy of $Y(n)$ in the even component; choosing from the odd vertices gives a copy of Y(n) in the odd component. Copies of $Y(n)$ in the even component are of course isomorphic to copies of $Y(n)$ in the odd component. It will be convenient to consider both the odd and even components.

We will choose to regard the vertices of $\Omega(n)$ as subsets of $\{1, \ldots, n\}$. Thus vertices are adjacent if their symmetric difference is $\frac{n}{2}$.

We paraphrase Galliard's construction. He states it for n a power of two; here we assume only that n is a multiple of four. Furthermore, we will state it in terms of a maximum independent set in $Y(n)$, so that the full set would be obtained by taking four isomorphic copies.

For convenience, set $[n] = \{1, \ldots, n\}$. Let $c = \frac{n}{4} - 1$, and let C be a subset of $[n]$ of size c. Let

$$\mathcal{F}(n) = \{F \subseteq [n] : |F| = 2i, 0 \leq i \leq 2c;\, |F \cap C| \geq |F \setminus C|\}.$$

It is not too hard to see that $\mathcal{F}(n)$ is an independent set in the even component $Y(n)$, and that it is not properly contained in any other independent set. Furthermore, since $|\mathcal{F}(16)| = 576 = \frac{2304}{4}$, it gives a maximum independent set.

Galliard, Tapp and Wolf [28] gave an upper bound on the size of an independent set in $\Omega(16)$. They construct a set consisting of all the even vectors of weight 0 or 2, all the vectors of weight 4 that have a one in a specified position, and a set of vectors of weight 6 that has a technical description. If we pause at this point in their construction, what we have is a bound on an independent set in $Y(16)$. However, it turns out that if we omit the vectors of weight 6, the remaining set is equivalent to $\mathcal{F}(16)$ under an automorphism of the graph. Thus it is maximum.

We have also shown that the set of all subsets of weight 1 or 3 is again equivalent to $\mathcal{F}$ under an automorphism of the graph. We can generalize this, and we have the following two sets, both of which are equivalent to $\mathcal{F}(n)$. There are two cases, but they are not fundamentally different.

If $n = 8m$, then the following are both independent sets that are not properly contained in any larger independent set:

$$\{F \subseteq [n] : |F| = 2i + 1, 0 \leq i \leq m - 1\},$$

$$\{F \subseteq [n] : |F| = 2i, 0 \leq i \leq m - 1\} \cup \{F \subseteq [n] : |F| = 2m, 1 \in F\}.$$

The sizes of these are of course equal, and equal to:

$$\sum_{i=0}^{m-1} \binom{8m}{2i+1} = \binom{8m-1}{2m-1} + \sum_{i=0}^{m-1} \binom{8m}{2i}$$

If $n = 8m + 4$, then we have the following analogous constructions:

$$\{F \subseteq [n] : |F| = 2i + 1, 0 \leq i \leq m - 1\} \cup \{F \subseteq [n] : |F| = 2m + 1, 1 \in F\},$$

$$\{F \subseteq [n] : |F| = 2i, 0 \leq i \leq m\}.$$

The sizes of these are:

$$\binom{8m+3}{2m} + \sum_{i=0}^{m-1} \binom{8m+4}{2i+1} = \sum_{i=0}^{m} \binom{8m+4}{2i}$$

We conjecture that for n a multiple of four, these are maximum independent sets. We note that if this conjecture is true, it would imply the previously mentioned result of Frankl and Rödl.

Bibliography

[1] R. Ahlswede and L. H. Khachatrian, *The complete intersection theorem for systems of finite sets*, European J. Combin., **18** (1997), 125–136.

[2] N. Alon, I. Dinur, E. Friedgut, and B. Sudakov, *Graph products, fourier analysis and spectral techniques.* To appear.

[3] I. Anderson, *Combinatorics of Finite Sets*, Oxford University Press, New York, 1987.

[4] W. N. Anderson, Jr. and T. D. Morley, *Eigenvalues of the Laplacian of a graph*, Linear and Multilinear Algebra, **18** (1985), 141–145.

[5] E. Bannai and T. Ito, *Algebraic Combinatorics. I*, Benjamin/Cummings Publishing Co. Inc., Menlo Park, CA, 1984.

[6] B. Bollobás, *Combinatorics*, Cambridge University Press, Cambridge, 1986.

[7] G. Brassard, R. Cleve, and A. Tapp, *Cost of exactly simulating quantum entanglement with classical communication*, Phys. Rev. Lett., **83** (1999), 1874–1877.

[8] A. E. Brouwer, A. M. Cohen, and A. Neumaier, *Distance-Regular Graphs*, Springer-Verlag, Berlin, 1989.

[9] W. G. Brown, *On graphs that do not contain a Thomsen graph*, Canad. Math. Bull., **9** (1966), 281–285.

[10] H. Buhrman, R. Cleve, and A. Widgerson, *Quantum vs. classical communication and computation*, in Proceedings of the 30th Annual ACM Symposium on the Theory of Computing, 1998, 63–68.

[11] F. C. Bussemaker, D. M. Cvetković, and J. J. Seidel, *Graphs related to exceptional root systems*, in Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. I, North-Holland, Amsterdam, 1978, 185–191.

[12] P. J. Cameron and C. Y. Ku, *Intersecting families of permutations*, European J. Combin., **24** (2003), 881–890.

[13] A. Chowdhury, C. D. Godsil, and G. Royle, *Colouring lines in projective space*. Submitted.

[14] C. J. Colbourn, J. H. Dinitz, and D. R. Stinson, *Applications of combinatorial designs to communications, cryptography, and networking*, in Surveys in Combinatorics, 1999 (Canterbury), J. D. Lamb and D. A. Preece, eds., Cambridge Univ. Press, Cambridge, 1999, 37–100.

[15] D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of graphs*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980.

[16] É. Czabarka and L. Székely, *An alternative shifting proof to Hsieh's theorem*, in Proceedings of the Twenty-ninth Southeastern International Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 1998), 1998, 117–122.

[17] D. E. Daykin, *Erdős-Ko-Rado from Kruskal-Katona*, J. Combinatorial Theory Ser. A, **17** (1974), 254–255.

[18] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep. Suppl. (1973), vi+97.

[19] P. Delsarte, *Association schemes and t-designs in regular semilattices*, J. Combinatorial Theory Ser. A, **20** (1976), 230–243.

[20] M. Deza and P. Frankl, *Erdős-Ko-Rado theorem—22 years later*, SIAM J. Algebraic Discrete Methods, **4** (1983), 419–431.

[21] P. Erdős, *My joint work with Richard Rado*, in Surveys in combinatorics 1987 (New Cross, 1987), Cambridge Univ. Press, Cambridge, 1987, 53–80.

[22] P. Erdős, C. Ko, and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math. Oxford Ser. (2), **12** (1961), 313–320.

[23] P. Erdős, A. Rényi, and V. T. Sós, *On a problem of graph theory*, Studia Sci. Math. Hungar., **1** (1966), 215–235.

[24] P. Frankl, *The Erdős-Ko-Rado theorem is true for $n = ckt$*, in Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. I, North-Holland, Amsterdam, 1978, 365–375.

[25] P. Frankl and V. Rödl, *Forbidden intersections*, Trans. Amer. Math. Soc., **300** (1987), 259–286.

[26] P. Frankl and R. M. Wilson, *The Erdős-Ko-Rado theorem for vector spaces*, J. Combin. Theory Ser. A, **43** (1986), 228–236.

[27] V. Galliard, *Classical Pseudo-Telepathy and Colouring Graphs*, Master's thesis, ETH Zurich, 2001.

[28] V. Galliard, A. Tapp, and S. Wolf, *The impossibility of pseudo-telepathy without quantum entanglement*. arXiv:quant-ph/0211011, 2002.

[29] C. D. Godsil, *Algebraic Combinatorics*, Chapman & Hall, New York, 1993.

[30] C. D. Godsil, *Association Schemes*. Unpublished notes, 2002.

[31] C. D. Godsil, *Interesting Graphs and Their Colourings*. Unpublished notes, 2003.

[32] C. D. Godsil and M. W. Newman, *Independent sets in association schemes*. Submitted.

[33] C. D. Godsil and G. Royle, *Algebraic Graph Theory*, Springer-Verlag, New York, 2001.

[34] W. Haemers, *A generalization of the Higman-Sims technique*, Nederl. Akad. Wetensch. Indag. Math., **40** (1978), 445–447.

[35] W. Haemers, *On some problems of Lovász concerning the Shannon capacity of a graph*, IEEE Trans. Inform. Theory, **25** (1979), 231–232.

[36] G. Hahn and C. Tardif, *Graph homomorphisms: structure and symmetry*, in Graph Symmetry (Montreal, PQ, 1996), G. Hahn and G. Sabidussi, eds., Kluwer Acad. Publ., Dordrecht, 1997, 107–166.

[37] W. N. Hsieh, *Intersection theorems for systems of finite vector spaces*, Discrete Math., **12** (1975), 1–16.

[38] G. O. H. Katona, *A simple proof of the Erdős-Chao Ko-Rado theorem*, J. Combinatorial Theory Ser. B, **13** (1972), 183–184.

[39] W. Kocay, *Groups & Graphs*. Computer program, `http://bkocay.cs.umanitoba.ca/G&G/G&G.html`.

[40] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory, **25** (1979), 1–7.

[41] R. Mathon and A. Rosa, *A new strongly regular graph*, J. Combin. Theory Ser. A, **38** (1985), 84–86.

[42] K. Meagher and B. Stevens, *Covering arrays on graphs*. Submitted.

[43] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, Cambridge University Press, Cambridge, 2000.

[44] T. D. Parsons, *Graphs from projective planes*, Aequationes Math., **14** (1976), 167–189.

[45] S. E. Payne and J. A. Thas, *Finite Generalized Quadrangles*, Pitman (Advanced Publishing Program), Boston, MA, 1984.

[46] P. Sarnak, *Some Applications of Modular Forms*, Cambridge University Press, Cambridge, 1990.

[47] A. Schrijver, *New code upper bounds from the terwilliger algebra*. Preprint.

[48] A. Schrijver, *A comparison of the Delsarte and Lovász bounds*, IEEE Trans. Inform. Theory, **25** (1979), 425–429.

[49] B. Stevens, *Transversal Covers and Packings*, Ph.D. thesis, University of Toronto, 1998.

[50] J. Williford, *Constructions in Finite Geometry with Applications to Graphs*, Ph.D. thesis, University of Delaware, 2004.

[51] R. M. Wilson, *The exact bound in the Erdős-Ko-Rado theorem*, Combinatorica, **4** (1984), 247–257.

Index