COVERING ARRAYS ON GRAPHS AND EXTREMAL SET THEORY

IISER PUNE

A thesis submitted towards partial fulfillment of
BS-MS Dual Degree Programme
by

Navi Prasad

Under the guidance of

Dr. Soumen Maity
(Assistant Professor in Dept. of Mathematics)
IISER Pune

Department of Mathematics
Indian Institute of Science Education and Research Pune

Certificate

This is to certify that this dissertation entitled "**Covering Arrays on Graphs and Extremal Set Theory**" towards the partial fulfillment of the BS-MS Dual Degree programme at Indian Institute of Science Education and Research Pune, represents original research carried out by **Navi Prasad** at IISER Pune under the supervision of **Dr. Soumen Maity** during the academic year 2010-2011.

Dated: 11 April, 2011

Name of the Student: Navi Prasad

Signature of the Student:

Supervisor: Dr. Soumen Maity	Head (Mathematical Science):
Signature:	Signature:
Date:	Date:
Place:	Place:

Table of Contents

Abstract

A lot of research has been done on Covering Arrays in the past two decades and is still an active area of research. The work in constructions, applications and generalizations have given new insights into this field. It has led to wider knowledge of the covering array in particular and the usage of concepts from various mathematical fields like Algebra, Set theory etc. in this field, in general.

Among the various ways to help humanity, the main contribution of this field has been to be able to efficiently test systems and networks using the concepts from this field. In return, this gave us newer ways to understand the deeper concepts of this field itself. The construction of new covering array either from the existing ones or completely new ones has led to widening of scope of this field. We can, now, construct covering arrays much near to our needs.

Two vectors x and y in $\mathbb{Z}_k^n$ are said to be qualitatively independent if for all ordered pairs $(u, v) \in \mathbb{Z}_k \times \mathbb{Z}_k$, there is a position j in the vectors such that $(x_j, y_j) = (u, v)$. Any pair of rows in a covering array are qualitatively independent. If this array has been defined on a graph then only those pairs of rows are qualitatively independent which correspond to adjacent vertices. A covering array is said to be optimal if it has the minimum number of possible columns for a given number of rows.

The main focus of this thesis is to study the generalization of simple covering arrays to the one defined on graphs. The addition of graph structure enables to study the covering arrays by making good use of the principles of graph theory. The qualitative independence graphs are defined in this thesis as they are closely

related to covering arrays. Good bounds can be obtained on the size of an optimal covering array by using several results in set theory like Sperner's theorem, Erdös-Ko-Rado theorem etc. The core of a binary qualitative independence graphs can be generalized to uniform qualitative independence graphs. Cliques in a uniform qualitative independence graphs are closely related to balanced covering arrays. Using these graphs, bounds on the size of a balanced covering array can be obtained. Also, some aspects of the basic graph theory given in the thesis will aid our study.

Acknowledgements

With high regards and profound respect, I express my deepest gratitude to Dr. Soumen Maity, Assistant Professor, Dept. of Mathematics, IISER Pune who remained my thesis supervisor for the entire project. I am specially very thankful to him for his many valuable suggestions, guidance and continous support during this project.

At the same time I am also grateful to other faculty members and staff of IISER Pune. In particular, I must thank Dr. Anupam Kumar Singh who is my official faculty advisor.

I can't be more grateful to my family for their love, patience and support throughout this time. Finally, I wish to thank all the Hall of Residence-1 residents, specially my roommate - Bedartha Goswami, for encouraging me and making my stay in the hostel interesting.

At last, I would like to extend my gratitude towards each and everyone who has been a help in this project.

Navi Prasad

IISER Pune

Chapter 1

Covering Arrays: An Introduction

1.1 Introduction

Covering arrays, which are also known as *qualitatively independent families* and *surjective arrays*, are a generalization of the well-known and well-studied orthogonal arrays. They are mathematically rich design with many applications.

Definition 1.1.1. Let n, r, k, t be positive integers with $t \leq r$. A covering array, $t - CA(n, r, k)$, with strength t and alphabet size k is an $r \times n$ array with entries from $\{0, 1, ..., k-1\}$ and the property that any $t \times n$ subarray has all k^t possible t-tuples occurring at least once.

Example: An example of a covering array is a $2 - CA(11, 5, 3)$ is

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

Definition 1.1.2. The number of columns, n, in a $t - CA(n, r, k)$ is the **size** of the covering array. The smallest possible size of a covering array is denoted by $t - CAN(r, k)$, *i.e.*,

$$t - CAN(r, k) = min\{n \in \mathbb{N} : \exists\ t - CA(n, r, k)\}.$$

Definition 1.1.3. A covering array $t - CA(n, r, k)$ with $n = t - CAN(r, k)$ is said to be **optimal**.

Definition 1.1.4. The maximum number of rows possible in a covering array with n columns on a given alphabet is denoted by $t - N(n, k)$, that is

$$t - N(n, k) = max\{r \in \mathbb{N} : \exists\ t - CA(n, r, k)\}.$$

Throughout this study, we consider only strength-2 covering arrays. So, the t will be dropped from the notation, so that $CA(n, r, k)$, $CAN(r, k)$ and $N(n, k)$ will be used to denote $2 - CA(n, r, k)$, $2 - CAN(r, k)$ and $2 - N(n, k)$ respectively.

Definition 1.1.5. A covering array $CA(n, r, k)$ with the property that each letter occurs exactly n/k times in every row is a **balanced covering array**.

Definition 1.1.6. Let k, n be positive integers. Two vectors $u, v \in \mathbb{Z}_k^n$ are **qualitatively independent** if for each one of the possible k^2 ordered pairs $(a, b) \in \mathbb{Z}_k \times \mathbb{Z}_k$, there is an index i so that $(u_i, v_i) = (a, b)$. A set of vectors is qualitatively independent if any two distinct vectors in the set are qualitatively independent.

The set of rows in a covering array $CA(n, r, k)$ is a set of r pairwise qualitatively independent vectors from $\mathbb{Z}_k^n$.

Amongst the many applications of covering arrays, testing systems like software, networks etc. is an important one. If we assume that the system has r parameters to be tested and that each parameter is capable of taking k different values, then each row of the associated covering array can be thought of as representing a parameter (which means that the covering array needs to have r rows). The entries of the covering array which come from $\mathbb{Z}_k$ represent the different values that the parameter can take. If each column of the covering array represents the various values of the parameters in a test run for the system then all possible combinations of the values of any two parameters can be checked (as the test suite covered by all the columns in covering array will, for every pair of parameters, test all k^2 possible values of those two parameters). This implies that any two parameters can be tested completely against one another.

1.2 Construction of Covering Arrays

It is a subject of constant research to find newer construction of covering arrays *e.g.* construction of covering arrays with the fewest possible columns. Finding $CAN(r, k)$, for a given r and k, is difficult in general. A construction for covering arrays gives an upper bound on $CAN(r, k)$. We, now, see some of these constructions.

1.2.1 Finite Field Construction

Lemma 1.2.1. *[6] Let k be a prime power, then $CAN(k+1,k) = k^2$.*

Proof. Let C be an array with $(k+1)$ rows and k^2 columns. We index the rows and columns of C starting from 0. Let $GF[k]$ be a finite field of order k with an ordering on its elements as $\{f_0, f_1, f_2, ..., f_{k-1}\}$ where $f_0 = 0$ and $f_1 = 1$.

Let the first row of C be each element in the field repeated k times in the fixed order so that the entry in column x of the first row is f_l where $l = \lfloor x/k \rfloor$. Now, for $i = \{0, 1, 2, ..., k-1\}$, we set the entry in row $i+1$ and column x of the covering array to be $f_i f_l + f_j$ where $l = \lfloor x/k \rfloor$ and $j \equiv x \ mod \ k$.

Suppose the 1^{st} and the n^{th} rows are not qualitatively independent. Then for some distinct columns x and y, a 2-tuple/pair is repeated between them (since there are only k^2 columns, a repetition of any pair would hinder the covering of all possible pairs (a, b) such that $a, b \in \{f_0, f_1, ..., f_{k-1}\}$). In particular,

$$(C_{(1,x)}, C_{(n,x)}) = (C_{(1,y)}, C_{(n,y)}).$$

Thus,

$$f_{\lfloor x/k \rfloor} = f_{\lfloor y/k \rfloor} \tag{1.2.1}$$

and

$$f_{n-1} f_{\lfloor x/k \rfloor} + f_{j_1} = f_{n-1} f_{\lfloor y/k \rfloor} + f_{j_2} \tag{1.2.2}$$

where $j_1 = x \ mod \ k$ and $j_2 = y \ mod \ k$. As $GF[k]$ is a field, from Equation 1.2.1, we have

$$\lfloor x/k \rfloor = \lfloor y/k \rfloor \tag{1.2.3}$$

and from Equation 1.2.2, we have

$$f_{j_1} = f_{j_2}$$

or,

$$j_1 = j_2$$

which means that

$$x = y \ mod \ k \tag{1.2.4}$$

But, Equations 1.2.3 and 1.2.4 can't be true simultaneously if $x \neq y$. Hence, we arrive at a contradiction and the hypothesis that row 1 and row n are not qualitatively independent is proved incorrect. This means that the first row is qualitatively independent to all other rows.

Next, assume that any two distinct rows, say, row r and s (where $r, s \in \{1, 2, ..., k\}$) are not qualitatively independent. Then for some distinct columns x and y, a pair is repeated between them. In particular,

$$(C_{(r,x)}, C_{(s,x)}) = (C_{(r,y)}, C_{(s,y)}). \tag{1.2.5}$$

Thus,

$$f_{r-1}f_{\lfloor x/k \rfloor} + f_{j_3} = f_{r-1}f_{\lfloor y/k \rfloor} + f_{j_4} \tag{1.2.6}$$

and

$$f_{s-1}f_{\lfloor x/k \rfloor} + f_{j_3} = f_{s-1}f_{\lfloor y/k \rfloor} + f_{j_4} \tag{1.2.7}$$

where $j_3 \equiv x \ mod \ k$ and $j_4 \equiv y \ mod \ k$. By subtracting Equation 1.2.7 from 1.2.6, we get

$$f_{\lfloor x/k \rfloor}(f_{r-1} - f_{s-1}) = f_{\lfloor y/k \rfloor}(f_{r-1} - f_{s-1})$$

or,

$$f_{\lfloor x/k \rfloor} = f_{\lfloor y/k \rfloor}$$

or,

$$\lfloor x/k \rfloor = \lfloor y/k \rfloor. \tag{1.2.8}$$

From Equations 1.2.7 and 1.2.8, we get

$$f_{j_3} = f_{j_4}$$

or,

$$j_3 = j_4$$

or,

$$x = y \ mod \ k \tag{1.2.9}$$

But, Equations 1.2.8 and 1.2.9 can't be true simultaneously if $x \neq y$. Hence, we arrive at a contradiction and the hypothesis that row r and row s are not qualitatively independent is proved incorrect. This means that the any row is qualitatively independent to all other rows. In this way, we have constructed a $CA(k^2, k+1, k)$ which proves that $CAN(k+1, k) = k^2$ for k a prime power. □

Example: The following array is an example of $CA(16, 5, 4)$ from the the finite field construction on the finite field with four elements, namely $\{f_0 = 0, f_1 = 1, f_2 = 2, f_3 = 3\}$.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 3 & 3 & 3 & 3 \\ 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 & 1 & 0 & 3 & 2 & 2 & 3 & 0 & 1 & 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 & 2 & 3 & 0 & 1 & 3 & 2 & 1 & 0 & 1 & 0 & 3 & 2 \\ 0 & 1 & 2 & 3 & 3 & 2 & 1 & 0 & 1 & 0 & 3 & 2 & 2 & 3 & 0 & 1 \end{pmatrix}.$$

Definition 1.2.1. In a covering array, two columns are **disjoint** if they have different entries for each row.

A column of all 0's is disjoint to a column of all 1's. A covering array with m disjoint columns has a set of at least m columns that are pairwise disjoint.

Corollary 1.2.2. *[6] For any prime power k, there exists a covering array $CA(k^2, k, k)$ with k disjoint columns.*

Proof. Let C be the covering array $CA(k^2, k+1, k)$ built by the finite field construction. We construct a new covering array C_1 by removing the first row of C. From the finite field construction, we know that, for columns $j = \{0, 1, 2, ..., k-1\}$, the entry on row i of C_1 is $f_i 0 + f_j = f_j$. Thus the first k columns of C_1 are disjoint. $\square$

Example: Based on the above result, the following array $CA(16, 4, 4)$ can be constructed with its first four columns disjoint.

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 & 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 & 1 & 0 & 3 & 2 & 2 & 3 & 0 & 1 & 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 & 2 & 3 & 0 & 1 & 3 & 2 & 1 & 0 & 1 & 0 & 3 & 2 \\ 0 & 1 & 2 & 3 & 3 & 2 & 1 & 0 & 1 & 0 & 3 & 2 & 2 & 3 & 0 & 1 \end{pmatrix}.$$

1.2.2 Block-size Recursive Construction

Block-size Recursive Construction uses two covering arrays with the same alphabet.

Theorem 1.2.3. *([8],[10]) If there exists a $CA(m,s,k)$ and a $CA(n,r,k)$, then there exists a $CA(m+n,sr,k)$.*

Proof. Let $A = CA(n,r,k)$ and $B = CA(m,s,k)$ be two covering arrays with the same alphabet. Let the rows of A and B be denoted by a_i and b_j respectively for $i = \{0,1,2,...,r-1\}$ and $j = \{0,1,2,...,s-1\}$. We can construct a $CA(n+m,rs,k)$ by the following method.

The first s rows of the $CA(n+m,rs,k)$ are formed by concatenating row a_0 of A with row b_j of B for $j = \{0,1,...,s-1\}$. The next s rows of $CA(n+m,rs,k)$ are formed by concatenating row a_1 with row b_j of B for $j = \{0,1,...,s-1\}$ and so on. Hence, in general, row l of $CA(n+m,rs,k)$ is formed by concatenating row a_i of A, where $i = \lfloor l/s \rfloor$ with row b_j of B, where $j \equiv l \ mod \ s$.

Clearly, any two distinct rows of this array $CA(m+n,rs,k)$ are of the form $a_x b_u$ and $a_y b_v$. If $x = y$, then by construction we have $u \neq v$ and hence all the possible pairs in the k-alphabet occur in the last m columns between b_u and b_v. But, if $x \neq y$ then all possible pairs in the k-alphabet occur in the first n columns itself between a_x and a_y. Thus, $CA(n+m,rs,k)$ constructed in this way is indeed a covering array.

$$A = \overbrace{\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{r-1} \end{pmatrix}}^{n\ columns}, B = \overbrace{\begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{s-1} \end{pmatrix}}^{m\ columns} \ and\ CA(n+m, rs, k) = \overbrace{\left(\begin{array}{c|c} a_0 & b_0 \\ a_0 & b_1 \\ \vdots & \vdots \\ a_0 & b_{s-1} \\ \hline a_1 & b_0 \\ a_1 & b_1 \\ \vdots & \vdots \\ a_1 & b_{s-1} \\ \hline \vdots & \vdots \\ \hline a_{r-1} & b_0 \\ a_{r-1} & b_1 \\ \vdots & \vdots \\ a_{r-1} & b_{s-1} \end{array}\right)}^{n+m\ columns}.$$

$\square$

Let C be a $CA(n+m, sr, k)$ built by the block-size recursive construction. All the distinct k pairs $(a, a) \in \mathbb{Z}_k \times \mathbb{Z}_k$ are covered between any two rows of C. It is because each pair of rows has either the first n columns the same or the first n columns are a pair of distinct rows from a covering array. As a result, any pair $(a, a) \in \mathbb{Z}_k \times \mathbb{Z}_k$ is covered either way. Similarly, all the distinct k pairs $(a, a) \in \mathbb{Z}_k \times \mathbb{Z}_k$ also occur in the last m columns for any two rows of C.

From the above discussion, we conclude that every pair $(a, a) \in \mathbb{Z}_k \times \mathbb{Z}_k$ is covered at least twice in the covering array $CA(n+m, sr, k)$. However, when constructing a covering array, it is unnecessary to cover these pairs twice. So, if we could remove some of these pairs then we would improve the block-size recursive construction.

Lemma 1.2.4. *([8],[10]) For a prime power k, there exists a $CA(2k^2-k, k(k+1), k)$. Equivalently, for any prime power k and any integer $r \leq k(k+1)$,*

$$CAN(r, k) \leq 2k^2 - k.$$

Proof. For a prime power k, if the block-size recursive construction is used with the covering array $CA(k^2, k+1, k)$ from the finite field construction and the $CA(k^2, k, k)$ with k disjoint columns, all the distinct k pairs $(a, a) \in \mathbb{Z}_k \times \mathbb{Z}_k$ are covered in the first k^2 columns of the thus formed final covering array $CA(2k^2, k(k+1), k)$. So, the first k columns of $CA(k^2, k, k)$ need not to be included in the final covering array. Hence, after removing the columns (k^2+1) through (k^2+k) from the covering array $CA(2k^2, k(k+1), k)$, we get $CA(2k^2-k, k(k+1), k)$. $\square$

Lemma 1.2.5. *For k a prime power and any positive integer i, there exists a covering array $CA(k^2 + i(k^2-k), k^i(k+1), k)$. Thus,*

$$CAN(k^i(k+1), k) \leq k^2 + i(k^2 - k).$$

Proof. We use the principle of mathematical induction to prove this result.

$$For\ i = 1,\ CA(k^2 + i(k^2-k), k^i(k+1), k) = CA(2k^2-k, k(k+1), k).$$

From Lemma 1.2.4, we conclude that the result holds for $i = 1$.

Now, we assume that the result holds for $i = n$, *i.e.*, there exists a $CA(k^2+n(k^2-k), k^n(k+1), k)$. If the block-size recursive construction is used with the covering array $CA(k^2+n(k^2-k), k^n(k+1), k)$ and the $CA(k^2, k, k)$ with k disjoint columns, all the distinct k pairs $(a, a) \in \mathbb{Z}_k \times \mathbb{Z}_k$ are covered in the first $k^2 + n(k^2 - k)$ columns, so the first k columns of $CA(k^2, k, k)$ need not to be included in the final covering array. Hence, after removing the columns $(k^2+n(k^2-k)+1)$ through $(k^2+n(k^2-k)+k)$, we get a $CA(k^2+(n+1)(k^2-k), k^{n+1}(k+1), k)$. So, we conclude that the result holds for $i = n+1$ whenever it holds for $i = n$. From the principle of mathematical induction, we conclude that there exists a covering array $CA(k^2+i(k^2-k), k^i(k+1), k)$ for any positive integer i when k is a prime power. □

The block-size recursive construction method applied to the covering array $CA(2k^2-k, k(k+1), k)$ and the covering array $CA(k^2, k, k)$ with k disjoint columns produces a covering array $CA(3k^2 - 2k, k^2(k+1), k)$. If we apply the block-size recursive construction to the new covering array with the array $CA(k^2, k, k)$ with k disjoint columns for i times, we get a $CA(k^2 + i(k^2 - k), k^i(k+1), k)$. In this covering array, each letter occurs exactly $k + i(k-1)$ times in each row and is hence a balanced covering array.

1.3 Conclusion

In this chapter, we learned that any two rows of a covering array are qualitatively independent vectors and the number of columns of a covering array can't be reduced arbitrarily, rather it can attain a minimum value denoted by $CAN(r, k)$ for a covering array with parameters n, r and k.

We also learned that covering arrays have industrial applications to software and circuit testing, drying screening and data compression.

We studied two methods for constructing covering arrays such as finite field construction and block-size recursive construction. Although covering designs constructed here are usually very small yet they are not always optimal. Therefore, establishing a more efficient approach still deserves further research.

Chapter 2

Extremal Set Theory

One of the central problems in extremal set theory is the problem of finding a system of sets with the largest cardinality given some restriction on the sets of the system. Here, we discuss two such problems. One of them is to find the maximum cardinality of a set system, over a finite ground set, satisfying the constraint that any two distinct sets in the system are incomparable and the other is to find the set system with the largest cardinality satisfying the constraint that any two distinct sets in the system are intersecting.

2.1 Set Systems

Let $X = \{1, 2, ..., n\}$ be an n-set for some positive integer n. The **power set** of X is the collection of all subsets of X and is denoted by $\mathcal{P}(X)$. A **set system** on an n-set is a collection of sets from $\mathcal{P}(X)$. For a positive integer $k \leq n$, a **k-set** is a set $A \in \mathcal{P}(X)$ with $|A| = k$. The collection of all k-sets of an n-set is denoted by $\binom{[n]}{k}$. A **k-uniform set system** on an n-set is a collection of sets from $\binom{[n]}{k}$. It is possible to arrange $\mathcal{P}(X)$ in a poset ordered by inclusion. In this poset, the sets $\binom{[n]}{k}$ are the **level sets** for every positive integer $k \leq n$.

2.2 Set Partitions

A **set partition** of an n-set is a set of disjoint non-empty subsets (called **classes**) of the n-set whose union is the n-set. A partition P is called a **k-partition** if it contains k classes *i.e.* $P = \{P_1, P_2, ..., P_k\}$. For positive integers k and n, let $\mathcal{P}_k^n$ denote the set of all k-partitions of an n-set. The values $S(n, k) = |\mathcal{P}_k^n|$ are called the **Stirling number of the second type**.

A partition $P \in \mathcal{P}_k^n$ is **uniform** if every class $P_i \in P$ has the same cardinality *i.e.* $|P_i| = \frac{n}{k} \ \forall \ P_i \in P$. If $n = ck$, we denote the set of all uniform k-partition in $\mathcal{P}_k^n$ by $\mathcal{U}_k^n$.

$$U(n, k) = |\mathcal{U}_k^n| = \frac{1}{k!}\binom{n}{c}\binom{n-c}{c}\cdots\binom{n-(k-1)c}{c}. \tag{2.2.1}$$

If k does not divide n, it is not possible for a partition in $\mathcal{P}_k^n$ to be uniform. If $n = ck + r$ where $0 \leq r < k$, a partition $P \in \mathcal{P}_k^n$ is **almost-uniform** if every class of P has cardinality c or $c+1$. In an almost-uniform partition, there are r classes of cardinality $(c+1)$ and $(k-r)$ classes of cardinality c. We denote the set of all almost-uniform partition in $\mathcal{P}_k^n$ by $\mathcal{AU}_k^n$.

$$\begin{aligned} AU(n, k) = \quad & \frac{1}{r!(k-r)!}\binom{n}{c}\binom{n-c}{c}\cdots\binom{n-(k-r-1)c}{c} \\ & \binom{n-(k-r)c}{c+1}\binom{n-(k-r)c-(c+1)}{c+1}\cdots\binom{c+1}{c+1}. \end{aligned} \tag{2.2.2}$$

If k divides n, then $\mathcal{AU}_k^n = \mathcal{U}_k^n$ and $U(n, k) = AU(n, k)$.

2.3 Sperner Theory

An important class of problems in extremal set theory deals with the maximum cardinality of a set system with some restriction on the sets in the system. The restriction that we consider here is that any two distinct sets from the system must be *incomparable*. Two subsets A and B of an n-set are **comparable** if $A \subseteq B$ or $B \subseteq A$. If A and B are not comparable then they are **incomparable**.

Definition 2.3.1. Sperner Set System: Let n be a positive integer. A *Sperner set system* $\mathcal{A}$ on an n-set is a set system on an n-set with the property that any two distinct sets in $\mathcal{A}$ are incomparable.

Definition 2.3.2. Matching: A matching in a graph is a set of edges such that no two of them share a vertex in common.

The size of a matching is the number of edges in it. A vertex contained in an edge of M is said to be covered by M.

Definition 2.3.3. Perfect matching or 1-factor: A matching that covers every vertex of X is called a perfect matching or a 1-factor.

Definition 2.3.4. Maximum matching: A maximum matching is a matching with the maximum possible number of edges.

Definition 2.3.5. Partial Order: A relation " $\leq$ " is a *partial order* on a set S if it has:

1. Reflexivity: $a \leq a \; \forall \; a \in S$.
2. Antisymmetry: $a \leq b$ and $b \leq a$ implies $a = b$.
3. Transitivity: $a \leq b$ and $b \leq c$ implies $a \leq c$.

Definition 2.3.6. Partially Ordered Set: A *partially ordered set* (or *poset*) is a set taken together with a partial order on it. Formally, a *partially ordered set* is defined as an ordered pair $P = (X, \leq)$, where X is called the *ground set* of P and $\leq$ is the partial order of P.

Definition 2.3.7. A set $\mathcal{A}$ is a subset of a set $\mathcal{B}$ if $\mathcal{A}$ is contained inside $\mathcal{B}$. The relationship of one set being a subset of another is called **inclusion**.

For any set S, the inclusion relation $\subseteq$ is a partial order on the set $\mathcal{P}(S)$ of all subsets of S. ($\mathcal{P}(S)$ refers to the power set of S)

Definition 2.3.8. Totally Ordered Set: For a, b distinct elements of a partially ordered set P, if $a \leq b$ or $b \leq a$, then a and b are comparable. Otherwise they are incomparable. If every two elements of a poset are comparable, the poset is called a *totally ordered set* or *chain* (*e.g.* the natural numbers under order). A poset in which every two distinct elements are incomparable is called an *antichain*.

Definition 2.3.9. Chain: Let P be a finite partially ordered set. A *chain* in P is a set of pairwise comparable elements w.r.t the partial order (i.e. a chain is a totally ordered subset). In other words, a chain in a poset is a collection of sets in the poset with the property that any two distinct sets in the chain are ordered in the poset.

Definition 2.3.10. A graph X is called **bipartite** if its vertex set can be partitioned into two parts V_1 and V_2 such that every edge has one end in V_1 and one in V_2.

Let the neighborhood of a set of vertices S, denoted by $N(S)$, is the union of the neighborhoods of the vertices of S.

Theorem 2.3.1. *[3]* ***Hall's Theorem****: Let G be a bipartite graph with partite sets X and Y, not necessarily equally sized. X can be matched into Y if and only if $\mid N(S) \mid \geq \mid S \mid$ for all subsets S of X.*

In other words, a bipartite graph with parts X and Y admits a matching that covers every vertex of X if and only if for every set $S \subseteq X$, the number of vertices of Y with a neighbor in S is at least $|S|$.

Theorem 2.3.2. *[9]* ***Sperner's Theorem****: Let n be a positive integer. If $\mathcal{A}$ is a Sperner set system on an n-set, then $|\mathcal{A}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.*

Proof. We construct a bipartite graph as follows. Let r be a positive integer with $r < \lfloor \frac{n}{2} \rfloor$. For every r-set of the n-set, there is a corresponding vertex in the first part of the graph and for each $(r+1)$-set in the n-set, there is a corresponding vertex in the second part of the graph. Two vertices are adjacent in this bipartite graph if and only if one of the corresponding sets is contained in the other set. To construct an $(r+1)$-set adjacent to an r-set, one element (other than the r elements of the r-set) should be added to the r-set. Since this can be done in $(n-r)$ ways, we conclude that $(n-r)$ number of $(r+1)$-sets exist which are adjacent to the initial r-set and hence all the vertices in the first part of the graph have degree $(n-r)$. To construct an r-set adjacent to an $(r+1)$-set, one of the $(r+1)$ elements of the $(r+1)$-set should be removed. Since this can be done by removing any one of the $(r+1)$ elements, we conclude that $(r+1)$ number of r-sets exist which are adjacent to the initial $(r+1)$-set and hence all the vertices in the second part have degree $(r+1)$. Let S be a set of vertices from the first part of the graph and let $N(S)$ be the set of vertices in the second part of the graph adjacent to any vertex in S. An edge through a vertex in S always ends up on a vertex in $N(S)$. So, the number of edges through all the vertices in S is always smaller than or equal to the number of edges through the vertices of $N(S)$. Hence,

$$|S|(n-r) \leq |N(S)|(r+1). \tag{2.3.1}$$

Since $r < \lfloor \frac{n}{2} \rfloor$, we have $\binom{n}{r} \leq \binom{n}{r+1}$ which implies that

$$(r+1) \leq (n-r). \tag{2.3.2}$$

From the above two equations, we infer that

$$\mid S \mid \quad \leq \quad \mid N(S) \mid . \tag{2.3.3}$$

Thus, by Hall's Theorem, there is a one-to-one matching from $\binom{[n]}{r}$ to $\binom{[n]}{r+1}$ for $r < \lfloor \frac{n}{2} \rfloor$. Similarly, for any positive integer r with $r > \lfloor \frac{n}{2} \rfloor$, there is a one-to-one matching from $\binom{[n]}{r}$ to $\binom{[n]}{r-1}$. Two sets are in the same chain if they are matched in one of these matchings. Then these matching define $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ disjoint chains which partition the poset, since each chain has exactly one set in the set $\binom{n}{\lfloor \frac{n}{2} \rfloor}$. Thus, the poset formed by $\mathcal{P}(\{1, 2, ..., n\})$ ordered by inclusion can be decomposed into $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ disjoint chains. Any Sperner system can intersect such a chain in at most one set, and thus, has cardinality no more than $\binom{n}{\lfloor \frac{n}{2} \rfloor}$. □

Moreover, $\mid \mathcal{A} \mid = \binom{n}{\lfloor \frac{n}{2} \rfloor}$ if and only if $\mathcal{A} = \binom{[n]}{k}$ where $k = \lfloor \frac{n}{2} \rfloor$ or $k = \lceil \frac{n}{2} \rceil$. Based on the above theorem, we have the following matchings for the two different cases.

For n even:

1. Case 1: $r < \lfloor n/2 \rfloor$. Let $r = 2m$ which implies that $r \leq m - 1$. From the above result, we have the following matching (not a one-to-one matching).

$$\binom{[n]}{1} \longrightarrow \binom{[n]}{2} \longrightarrow \text{.........} \longrightarrow \binom{[n]}{m-1} \longrightarrow \binom{[n]}{m}.$$

2. Case 2: $r > \lfloor n/2 \rfloor$. Let $r = 2m$ which implies that $r \geq m + 1$. From the above result, we have the following matching (not a one-to-one matching).

$$\binom{[n]}{n} \longrightarrow \binom{[n]}{n-1} \longrightarrow \text{.........} \longrightarrow \binom{[n]}{m+1} \longrightarrow \binom{[n]}{m}.$$

For n odd:

1. Case 1: $r < \lfloor n/2 \rfloor$. Let $r = 2m+1$ which implies that $r \leq m-1$. From the above result, we have the following matching (not a one-to-one matching).
$$\binom{[n]}{1} \longrightarrow \binom{[n]}{2} \longrightarrow .... \longrightarrow \binom{[n]}{m-1} \longrightarrow \binom{[n]}{m}.$$
2. Case 2: $r > \lfloor n/2 \rfloor$. Let $r = 2m+1$ which implies that $r \geq m+1$. From the above result, we have the following matching (not a one-to-one matching).
$$\binom{[n]}{n} \longrightarrow \binom{[n]}{n-1} \longrightarrow .... \longrightarrow \binom{[n]}{m+1} \longrightarrow \binom{[n]}{m}.$$

Illustrative example for the above result: Based on the above theorem, below are 10 disjoint chains for $n = 5$.

1. $\{3\} \to \{3,5\} \to \{3,4,5\} \to \{2,3,4,5\} \to \{1,2,3,4,5\}$
2. $\{1\} \to \{1,2\} \to \{1,2,4\} \to \{1,2,4,5\}$
3. $\{4\} \to \{3,4\} \to \{1,3,4\} \to \{1,3,4,5\}$
4. $\{5\} \to \{1,5\} \to \{1,2,5\} \to \{1,2,3,5\}$
5. $\{2\} \to \{2,4\} \to \{2,3,4\} \to \{1,2,3,4\}$
6. $\{1,3\} \to \{1,3,5\}$
7. $\{2,5\} \to \{2,3,5\}$
8. $\{1,4\} \to \{1,4,5\}$
9. $\{2,3\} \to \{1,2,3\}$
10. $\{4,5\} \to \{2,4,5\}$.

Hence, two of the *Sperner set systems* are $\mathcal{A}$=$\{\{3,4,5\}$, $\{1,2,4\}$, $\{1,3,4\}$, $\{1,2,5\}$, $\{2,3,4\}$, $\{1,3,5\}$, $\{2,3,5\}$, $\{1,4,5\}$, $\{1,2,3\}$, $\{2,4,5\}\}$ and $\mathcal{B}$=$\{\{3,5\}$, $\{1,2\}$, $\{3,4\}$, $\{1,5\}$, $\{2,4\}$, $\{1,3\}$, $\{2,5\}$, $\{1,4\}$, $\{2,3\}$, $\{4,5\}\}$.

2.4 Intersecting Set Systems

For positive integers $t,\ k,\ n$, let $I(t,k,n)$ denote the collection of all set systems $\mathcal{A}$ on an n-set with the following properties: $\forall\ A \in \mathcal{A},\ |A| \leq k;\ \forall\ A,B \in \mathcal{A},\ A \not\subseteq B$ *and* $|A \cap B| \geq t$.

The set systems in $I(t,k,n)$ are known as **t-intersecting set systems** and if $t=1$, these are also called **intersecting set systems**. Since $A \not\subseteq B$ for any distinct sets $A, B \in I(t,k,n)$, all the set systems in $I(t,k,n)$ are Sperner set systems. If $2k - t \geq n$, then any two k-sets from the n-set have at least t elements in common.

Definition 2.4.1. For positive integers n, k, t, a **k-Uniform t-Intersecting Set System** is a k-uniform set system, $\mathcal{A}$, on an n-set with the property that $\forall\ A, B \in \mathcal{A}$, we have $A \not\subseteq B$ and $\forall\ A, B \in \mathcal{A}$, we have $|A \cap B| \geq t$.

Lemma 2.4.1. *[1] If $2k \leq n$, then for any $\mathcal{A} \in I(t,k,n)$ there exists a k-uniform t-intersecting set system $\mathcal{A}' \in I(t,k,n)$ with $|\mathcal{A}| \leq |\mathcal{A}'|$.*

Proof. Let $\mathcal{A} \in I(t,k,n)$. We construct a set system $\mathcal{A}'$ as follows: Let $A \in \mathcal{A}$ be such that $|A| < k \leq \lfloor \frac{n}{2} \rfloor$. Now, since A is a subset of an n-set, it can be matched to the $\lfloor n/2 \rfloor$-set of the chain of the poset formed by $\mathcal{P}(\{1,2,...,n\})$ (ordered by inclusion) which it belongs to. Let this unique $\lfloor n/2 \rfloor$-set be A_1.

Since any $B \in \mathcal{A}$ distinct from A is such that $A \not\subseteq B$ and $B \not\subseteq A$, so A and B never belong to the same chain of the poset formed by $\mathcal{P}(\{1,2,...,n\})$ (ordered by inclusion) and thus,

$$A_1 \not\subseteq B \textit{ and } B \not\subseteq A_1.$$

Further,

$$|A_1 \cap B| \geq |A \cap B| \geq t.$$

In this way, we can replace all the $A \in \mathcal{A}$ (for which $|A| < k = \lfloor n/2 \rfloor$) with A_1 (for which $|A_1| = k = \lfloor n/2 \rfloor$) to get $\mathcal{A}' \in I(t, k, n)$ which is a k-uniform t-intersecting set system. From the construction itself, it is clear that $|\mathcal{A}| \leq |\mathcal{A}'|$. □

In other words, for n sufficiently large, if there exists a t-intersecting set system on an n-set, then there exists a k-uniform t-intersecting set system that has at least the same cardinality. In particular, it is possible to replace each set of size less than k in a t-intersecting set system on an n-set by a set of size k, which is t-intersecting with all the sets in the system.

Definition 2.4.2. For positive integers n, k, t with $t \leq k \leq n$, a set system on an n-set is a **k-uniform trivially t-intersecting set system** if it is equal, up to a permutation on $\{1, 2, ..., t\}$, to

$$\mathcal{A} = \{A \in \binom{[n]}{k} \ : \ \{1, 2, ..., t\} \subseteq A\}.$$

The cardinality of a k-uniform trivially t-intersecting set system is $\binom{n-t}{k-t}$. If $t = 1$ then a k-uniform trivially t-intersecting set system is simply called a **k-uniform trivially intersecting set system**.

2.5 Qualitatively Independent Subsets

Definition 2.5.1. Two subsets A and B of an n-set are **qualitatively independent subsets** if

$$A \cap B \neq \phi, \ A \cap \overline{B} \neq \phi, \ \overline{A} \cap B \neq \phi, \ \overline{A} \cap \overline{B} \neq \phi$$

A collection of subsets of an n-set are said to be qualitatively independent if each pair of members of it are qualitatively independent of each other.

The definition of qualitatively independent sets is equivalent to the definition of qualitatively independent binary vectors. To see this, let A and B be qualitatively independent subsets of an n-set. We define the vector corresponding to A to be $a \in \mathbb{Z}_2^n$, with $a = (a_1, a_2, ..., a_n)$, $a_k = 1$ if $k \in A$ and $a_k = 0$ otherwise. Similarly, let $b \in \mathbb{Z}_2^n$ be the vector corresponding to the set B. Since A and B are qualitatively independent, $A \cap B \neq \phi$. So there exists some $i \in A \cap B$ for which $(a_i, b_i) = (1, 1)$. Further, since $A \cap \overline{B} \neq \phi$, there exists some $j \in A \cap \overline{B}$ for which $(a_j, b_j) = (1, 0)$. Similarly, for $k \in \overline{A} \cap B$ we get $(a_k, b_k) = (0, 1)$ and for $l \in \overline{A} \cap \overline{B}$ we get $(a_l, b_l) = (0, 0)$. Thus, vectors a and b are qualitatively independent.

Conversely, if $a, b \in \mathbb{Z}_2^n$ are qualitatively independent, then the sets A and B, defined by $i \in A$ if and only if $a_i = 1$ and $j \in B$ if and only if $b_j = 1$, are also qualitatively independent.

2.6 The Erdős-Ko-Rado Theorem

The Erdős-Ko-Rado Theorem proves that the set system in $I(t, k, n)$ with the largest cardinality is a k-uniform trivially t-intersecting set system provided that n is sufficiently large.

Theorem 2.6.1. *[1]* ***Erdős-Ko-Rado Theorem****: Let k and t be positive integers, with $0 < t < k$. There exists a function $f(t, k)$ such that if n is a positive integer with $n > f(t, k)$, then for any $\mathcal{A} \in I(t, k, n)$*

$$|\mathcal{A}| \leq \binom{n-t}{k-t}.$$

Moreover, equality holds if and only if $\mathcal{A}$ is a k-uniform trivially t-intersecting set system.

2.6.1 Application of the Erdős-Ko-Rado Theorem

Theorem 2.6.2. *([4],[5]) If $\mathcal{A} = \{A_1, A_2, ..., A_k\}$ is a qualitatively independent set system of an n-set, then*

$$|\mathcal{A}| \leq \binom{n-1}{\lfloor n/2 \rfloor - 1}.$$

Further, this bound is attained by a $\lfloor n/2 \rfloor$-uniform trivially 1-intersecting set system.

Proof. The proof can be divided into following two cases:

1. Let n be **even**: We define a set system $\mathcal{A}^*$ to be consisting of all the sets from $\mathcal{A}$ along with their complements, *i.e.*,

   $$\mathcal{A}^* = \{A_i,\ \overline{A_i}\ :\ A_i \in \mathcal{A}\}.$$

   Clearly, the set system $\mathcal{A}^*$ is a Sperner set system which implies that $|\mathcal{A}^*| \leq \binom{n}{n/2}$. Hence,

   $$|\mathcal{A}| \leq \frac{1}{2}\binom{n}{n/2} = \binom{n-1}{\frac{n}{2}-1}.$$

   Clearly, this bound is attained by the set system

   $$\mathcal{A} = \{A \in \binom{[n]}{n/2}\ :\ 1 \in A\}.$$

2. Let n be **odd**: If $A_i \in \mathcal{A}$ and $|A_i| \geq n/2$ then replace A_i with $\overline{A_i}$. This does not affect the pairwise qualitative independence of $\mathcal{A}$. For this reason, we can assume that each $A_i \in \mathcal{A}$ has $|A_i| \leq \lfloor n/2 \rfloor$.

By the definition of qualitative independence, $\mathcal{A}$ is a 1-intersecting set system and by the Erdős-Ko-Rado Theorem,

$$|\mathcal{A}| \leq \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}.$$

Clearly, this bound is attained by the set system

$$\mathcal{A} = \{A \in \binom{[n]}{\frac{n-1}{2}} \ : \ 1 \in A\}.$$

$\square$

Theorem 2.6.3. *[5] Let r be a positive integer, then*

$$CAN(r,2) = min\{n \ : \ \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} \geq r\}. \tag{2.6.1}$$

The set $\mathcal{B}$ of n-size binary vectors corresponding to a qualitatively independent set system on an n-set is itself a qualitatively independent set, *i.e.* any two vectors in set $\mathcal{B}$ are qualitatively independent and such sets have cardinality at most $\binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}$, or,

$$|\mathcal{B}| \leq \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} \ (from\ Theorem\ 2.6.2).$$

Hence, a qualitatively independent set of n-size binary vectors can have cardinality at most $\binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}$. Since the rows of a $CA(n,r,2)$ form a set of r qualitatively independent n-size binary vectors, we conclude that if there exists a $CA(n,r,2)$, then

$$r \leq \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}.$$

This also verifies that

$$CAN(r,2) = min\{n : \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} \geq r\}.$$

2.7 Conclusion

In this chapter, we studied the properties of system of sets especially set partitions, intersecting-set systems etc.

Sperner set system is a set system in which no two sets are comparable. We saw the proof for the fact that there is an upper bound to the cardinality of such sets if the sets have been formed from the elements of a finite set. This upper bound is calculated with the help of Hall's theorem.

Intersecting-set system is an important part of set theory. We learnt that for n sufficiently large, if there exists a t-intersecting set system on an n-set, then there exists a k-uniform t-intersecting set system that has at least the same cardinality (which is possible by replacing each set of size less than k in a t-intersecting set system by a set of size k without affecting the property of the set system being t-intersecting).

The concept of qualitatively independent binary vectors to qualitatively independent subsets has been generalized here. We studied the Erdös-Ko-Rado theorem which provides an upper bound to the cardinality of the members of the set system $I(t, k, n)$. This theorem is applied to a qualitatively independent set system on an n-set to get an upper bound on this set system.

Chapter 3

Graph Theory

To study the generalization of covering arrays to add a graph structure to it, the concepts of graph theory is required.

3.1 Basic Graph Theory

Definition 3.1.1. A **graph** G is defined as an ordered pair $(V(G), E(G))$, consisting of a set of vertices, $V(G)$, and a set of edges, $E(G)$, joining pairs of vertices.

In this text, a finite, simple graph is denoted by G unless otherwise stated. The vertex set of G is denoted by $V(G)$ and the edge set is denoted by $E(G)$ which is a subset of the set of all unordered pairs of distinct elements of $V(G)$.

Definition 3.1.2. If x and y are vertices of G and $\{x, y\} \in E(G)$, then x and y are said to be **adjacent**. Adjacency is a symmetric and anti-reflexive relation in case of simple graphs.

3.2 Graph Homomorphism

Definition 3.2.1. Let G and H be any two graphs. A mapping ϕ from $V(G)$ to $V(H)$ is a **graph homomorphism** if vertices $\phi(x)$ and $\phi(y)$ are adjacent in H whenever x and y are adjacent in G.

Definition 3.2.2. Let G and H be any two graphs. A map ϕ from $V(G)$ to $V(H)$ is a **graph isomorphism** if ϕ is a bijection such that x, $y \in V(G)$ are adjacent in G if and only if $\phi(x)$ and $\phi(y)$ are adjacent in H. If there exists an isomorphism between two graphs, then we say the graphs are **isomorphic**.

A homomorphism from a graph G to itself is a **graph endomorphism**. An isomorphism from a graph G to itself is a **graph automorphism**. The *automorphism group* for a graph G is the group of all automorphisms of G denoted by $Aut(G)$.

Definition 3.2.3. A **bipartite graph** $G = G(X, Y)$ is a graph in which the vertex set $V(G)$ can be decomposed into two disjoint sets X and Y such that no two graph vertices within the same set are adjacent. Any edge $e \in E(G)$ can only connect a vertex $x \in X$ with a vertex $y \in Y$.

Definition 3.2.4. A **retraction** is a homomorphism f from a graph X to a subgraph Y of itself such that the restriction $f|_Y$ of f to $V(Y)$ is the identity map. If there is a retraction from X to a subgraph Y, then we say that Y is a **retract** of X.

Definition 3.2.5. A **fibre** of a homomorphism $\phi : G \longrightarrow H$ is a preimage $\phi^{-1}(y)$ of some vertex $y \in V(H)$, that is, $\phi^{-1}(y) = \{x \in V(G) \ : \ \phi(x) = y\}$.

3.3 Coloring, Cliques and Independent Sets

The **complete graph** on *n vertices*, K_n, is the graph with n vertices and with an edge between any two distinct vertices.

Definition 3.3.1. A **proper coloring** of G with n colors is a map from $V(G)$ to a set of n colors such that no two adjacent vertices are assigned the same color.

Definition 3.3.2. The least value of k for which X can be properly k-colored is the **chromatic number** of X and is denoted by $\chi(X)$.

Lemma 3.3.1. *A proper coloring of a graph G with n colors is equivalent to a homomorphism from G to K_n.*

Proof. Firstly, let $\phi : G \longrightarrow K_n$ be a proper coloring of G with n colors and let x_1, $x_2 \in V(G)$ be any two adjacent vertices of G. Thus, $\phi(x_1) \neq \phi(x_2)$. Being distinct vertices of K_n, $\phi(x_1)$ and $\phi(x_2)$ are adjacent. Thus, ϕ maps adjacent vertices to adjacent vertices and hence is a homomorphism.

Conversely, let $\psi : G \longrightarrow K_n$ be a homomorphism and let y_1, $y_2 \in V(G)$ be any two adjacent vertices of G. Thus, $\psi(y_1) \neq \psi(y_2)$. This means that ψ doesn't map any adjacent vertices of G to the same vertex of K_n. Hence, ψ is a proper coloring. □

Thus, the chromatic number of a graph G is the smallest n such that there exists a homomorphism $G \longrightarrow K_n$.

Definition 3.3.3. A **clique** in a graph G is a set of vertices from $V(G)$ in which any two distinct vertices are adjacent in G. A **maximum clique** in a graph G is a maximum set of pairwise adjacent vertices. The maximum clique number of a graph

G is defined to be the size of a maximum clique and denoted by $\omega(G)$. The maximum clique number is also called the **clique number**.

Lemma 3.3.2. *If G has a clique of size n, then there exists a homomorphism*

$$K_n \longrightarrow G.$$

Proof. Let X be a clique of the graph G of size n and let there be a map f from K_n to G which maps the n vertices of K_n to the n vertices of X. Suppose x and y are two distinct vertices of the graph K_n. Since $f(x)$ and $f(y)$ are the vertices of clique X, they are adjacent. Hence, every pair of adjacent vertices of the graph K_n are mapped to a pair of adjacent vertices of the graph G making the map f a homomorphism. □

Thus, the size of a maximum clique in G is the largest n for which there exists a homomorphism from K_n to G. An n-clique is a clique of size n.

Lemma 3.3.3. *For graphs G and H, if there is a homomorphism $G \longrightarrow H$, then*

$$\omega(G) \leq \omega(H).$$

Proof. If possible, let $\omega(H) < \omega(G)$. We know that the size of a maximum clique in G is the largest n for which there exists a homomorphism from K_n to G. So, there exist homomorphisms

$$K_{\omega(G)} \longrightarrow G \longrightarrow H$$

and hence, there exists a homomorphism

$$K_{\omega(G)} \longrightarrow H.$$

But, $\omega(H) < \omega(G)$ and still there exists a homomorphism from $K_{\omega(G)}$ to H which leads us to a contradiction. So, the hypothesis that $\omega(H) < \omega(G)$ is false. Thus, $\omega(G) \leq \omega(H)$ if there is a homomorphism $G \longrightarrow H$. □

Lemma 3.3.4. *For graphs G and H, if there is a homomorphism $G \longrightarrow H$, then*

$$\chi(G) \leq \chi(H).$$

Proof. If possible, let $\chi(H) < \chi(G)$. We know that the chromatic number of a graph G is the smallest n such that $G \longrightarrow K_n$. So, there exist homomorphisms

$$G \longrightarrow H \longrightarrow K_{\chi(H)}$$

and hence, there exists a homomorphism

$$G \longrightarrow K_{\chi(H)}.$$

But, $\chi(H) < \chi(G)$ and still there exists a homomorphism from G to $K_{\chi(H)}$ which leads us to a contradiction. So, the hypothesis that $\chi(H) < \chi(G)$ is false. Thus, $\chi(G) \leq \chi(H)$ if there is a homomorphism $G \longrightarrow H$. □

Lemma 3.3.5. *For any graph G, $\omega(G) \leq \chi(G)$.*

Proof. We know that there exist homomorphisms

$$K_{\omega(G)} \longrightarrow G \longrightarrow K_{\chi(G)}$$

and hence, there exists a homomorphism

$$K_{\omega(G)} \longrightarrow K_{\chi(G)}.$$

If possible, let $\chi(G) < \omega(G)$. So, a map f from $K_{\omega(G)}$ to $K_{\chi(G)}$ can't be an injective map. As a result, at least one pair of distinct vertices in $K_{\omega(G)}$ will be mapped to a single vertex in $K_{\chi(G)}$. Since any two distinct vertices of $K_{\omega(G)}$ are adjacent, at least there exists a pair of adjacent vertices in $K_{\omega(G)}$ which are mapped to the same vertex in $K_{\chi(G)}$. This renders the map f from being a homomorphism. So, there can't be any homomorphism from $K_{\omega(G)}$ to $K_{\chi(G)}$ which is a contradiction to the fact that there exists at least one. So, the hypothesis that $\chi(G) < \omega(G)$ is proved false and we conclude that $\omega(G) \leq \chi(G)$. $\square$

Hence, a proper coloring must always contain at least as many colors as the size of a maximum clique.

Definition 3.3.4. An **independent set** in a graph G is a set of vertices from $V(G)$ in which no two vertices are adjacent in G. The size of a largest independent set in a graph G is denoted by $\alpha(G)$.

The vertices which are assigned the same color in a proper coloring form an independent set. In fact, a proper coloring on a graph G partitions the vertices of G into independent sets called **color classes**. A proper coloring corresponds to a binary function on the independent sets of a graph: Each independent set that is a color class in the proper coloring is assigned a value of 1 and all other independent sets are assigned a value of 0 by the function. Further, each vertex is in exactly one independent set which has an assigned value of 1.

Definition 3.3.5. The **distance** $d_G(x, y)$ between two vertices x and y in a graph G is the length of the shortest path from x to y. The **diameter** of a graph G is the maximum distance over all pairs of vertices in G.

Definition 3.3.6. An **edge cover** of a graph is a set of edges so that each vertex is the terminus for some edge in the set.

We denote the number of edges incident to a vertex v by $d(v)$ and the minimum value of this over all vertices as $\delta(G)$.

Definition 3.3.7. The **degree** of a vertex $v \in V(G)$ is the number of vertices in G which are adjacent to v. If every vertex in G has the same degree, then we say that G is **regular**. In particular, if every vertex in G has degree k, we say that G is **k-regular**.

3.4 Vertex-Transitive Graphs

Definition 3.4.1. A graph is **vertex transitive** if its automorphism group acts transitively on the set of it's vertices. This means that for any two distinct vertices, there is an automorphism on the graph that maps one vertex to the other.

A vertex-transitive graph is necessarily regular.

3.5 Core of a Graph

Definition 3.5.1. Core: A graph G is a *core* if any endomorphism on G is an automorphism.

If a graph G is a core, then there is no homomorphism from G to a proper subgraph of G. For any positive integer n, the complete graph K_n is a core.

Definition 3.5.2. Core of a Graph: A *core* of a graph G is a subgraph $G^\bullet$ of G such that $G^\bullet$ is a core and there is a homomorphism $G \longrightarrow G^\bullet$.

Definition 3.5.3. Core of a Graph: Alternately, a *core* of a graph G is the minimal induced subgraph $G^\bullet$ such that there exist homomorphisms

$$G^\bullet \longrightarrow G \longrightarrow G^\bullet.$$

Corollary 3.5.1. *The above two definitions of the core of a graph are equivalent.*

Proof. Let there be a minimal induced subgraph $G^{\bullet}$ of a graph G such that there exist homomorphisms

$$G^{\bullet} \longrightarrow G \longrightarrow G^{\bullet}. \tag{3.5.1}$$

Let f be an endomorphism on $G^{\bullet}$. Suppose that f is a not surjective map. Then the range of f, $R(f) = G_1$ is a proper subset of $G^{\bullet}$ (*i.e.* $G_1 \subset G^{\bullet}$). Let f_1 be a map from $G^{\bullet}$ to G_1 such that $f_1(x) = f(x)\ \forall\ x \in G^{\bullet}$. Clearly, f_1 is also a homomorphism. So,

$$f_1 : G^{\bullet} \longrightarrow G_1 \tag{3.5.2}$$

is a homomorphism. Since $G_1 \subset G^{\bullet}$, an identity map

$$I : G_1 \longrightarrow G^{\bullet} \tag{3.5.3}$$

is also a homomorphism. Now, from equations 3.5.1, 3.5.2 and 3.5.3, we conclude that there exist homomorphisms

$$G_1 \longrightarrow G^{\bullet} \longrightarrow G \longrightarrow G^{\bullet} \longrightarrow G_1$$

and thus, there exist homomorphisms

$$G_1 \longrightarrow G \longrightarrow G_1. \tag{3.5.4}$$

But this contradicts the assumption that $G^{\bullet}$ is the minimal induced subgraph of the graph G to satisfy the above homomorphisms. So, the hypothesis that any endomorphism on $G^{\bullet}$ can be a non-surjective map is proved false, hence proving that any endomorphism on $G^{\bullet}$ is always a surjection. But a surjective endomorphism has to be an injective map (as a many-to-one function can't map n vertices to n vertices).

So, any endomorphism on $G^\bullet$ is a bijection and hence an isomorphism. Hence, $G^\bullet$ is a core in itself and thus a core of the graph G by definition 3.5.2. So, the minimal induced subgraph $G^\bullet$ such that there exist homomorphisms $G^\bullet \longrightarrow G \longrightarrow G^\bullet$, is a core of the graph G by definition 3.5.2 as well.

Conversely, let a subgraph $G^\bullet$ of a graph G be such that $G^\bullet$ is a core and there exists a homomorphism

$$G \longrightarrow G^\bullet. \tag{3.5.5}$$

Since $G^\bullet \subseteq G$, an identity map

$$I : G^\bullet \longrightarrow G \tag{3.5.6}$$

is a homomorphism. From equations 3.5.5 and 3.5.6, we conclude that there exist homomorphisms

$$G^\bullet \longrightarrow G \longrightarrow G^\bullet. \tag{3.5.7}$$

Now, if possible, let $G_1 \subset G^\bullet$ be the minimal induced subgraph of G such that there exist homomorphisms

$$G_1 \longrightarrow G \longrightarrow G_1. \tag{3.5.8}$$

From equations 3.5.7 and 3.5.8, we conclude that there exist homomorphisms

$$G^\bullet \longrightarrow G \longrightarrow G_1 \tag{3.5.9}$$

and thus, there exists a homomorphism

$$G^\bullet \longrightarrow G_1. \tag{3.5.10}$$

But, $G^\bullet$ is a core in itself and hence there can't be any homomorphism from $G^\bullet$ to a proper subgraph of $G^\bullet$ (here $G_1 \subset G^\bullet$). Hence, we conclude that no such G_1 exists and $G^\bullet$ is a core of the graph G by definition 3.5.3 as well. □

Lemma 3.5.2. *For any graph G, $\chi(G) = \chi(G^\bullet)$ and $\omega(G) = \omega(G^\bullet)$.*

Proof. We know that there exist homomorphisms

$$G^\bullet \longrightarrow G \longrightarrow G^\bullet$$

and thus, we have

$$\omega(G^\bullet) \leq \omega(G) \leq \omega(G^\bullet) \textit{ and } \chi(G^\bullet) \leq \chi(G) \leq \chi(G^\bullet)$$

$$\textit{or, } \omega(G) = \omega(G^\bullet) \textit{ and } \chi(G) = \chi(G^\bullet).$$

□

Lemma 3.5.3. *[2] Let $G_1^\bullet$ and $G_2^\bullet$ be cores. Then $G_1^\bullet$ and $G_2^\bullet$ are homomorphically equivalent if and only if they are isomorphic.*

Proof. Suppose $G_1^\bullet$ and $G_2^\bullet$ are homomorphically equivalent and that $f : G_1^\bullet \longrightarrow G_2^\bullet$ and $g : G_2^\bullet \longrightarrow G_1^\bullet$ are the homomorphisms between them. Then the maps $f \circ g$ and $g \circ f$ are endomorphisms on $G_2^\bullet$ and $G_1^\bullet$ respectively *i.e.* $f \circ g : G_2^\bullet \longrightarrow G_2^\bullet$ and $g \circ f : G_1^\bullet \longrightarrow G_1^\bullet$ are both homomorphisms since they are composition of two homomorphisms. Since $G_1^\bullet$ and $G_2^\bullet$ are cores, any endomorphism on them is an isomorphism. Hence, both $f \circ g$ and $g \circ f$ are isomorphisms which makes them a surjective map. Since $f \circ g$ is a surjective map, we conclude that f is a surjective map too (as it is the function f which gives the final values of the entire function). Similarly, since $g \circ f$ is a surjective map, we conclude that g is a surjective map too.

Now, let $|G_1^\bullet| = n$ and $|G_2^\bullet| = m$. We know that $f : G_1^\bullet \longrightarrow G_2^\bullet$ is a surjective map which means that $n \geq m$. We also know that $g : G_2^\bullet \longrightarrow G_1^\bullet$ is a surjective map which means that $m \geq n$. Combining the two, we get $n = m$ and hence $|G_1^\bullet| = |G_2^\bullet|$. Thus, the surjective maps f and g are injective as well and hence bijective overall. Hence, we conclude that f and g are isomorphisms. $\square$

Lemma 3.5.4. *[2] If $G_1^\bullet$ and $G_2^\bullet$ are both cores of a graph G, then $G_1^\bullet$ and $G_2^\bullet$ are isomorphic.*

Proof. Suppose that $G_1^\bullet$ and $G_2^\bullet$ are core of a graph G. Then there are homomorphisms

$$f_1 : G \longrightarrow G_1^\bullet \textit{ and } f_2 : G \longrightarrow G_2^\bullet$$

Then, the function f_1 restricted to $G_2^\bullet$, $f_1|_{G_2^\bullet}$ is a homomorphism from $G_2^\bullet$ to $G_1^\bullet$ and the function f_2 restricted to $G_1^\bullet$, $f_2|_{G_1^\bullet}$ is a homomorphism from $G_1^\bullet$ to $G_2^\bullet$. Therefore, by Lemma 3.5.3, we conclude that $G_1^\bullet$ and $G_2^\bullet$ are isomorphic. $\square$

Corollary 3.5.5. *A core of any graph X is a retract of the graph X.*

Proof. Let $X^\bullet$ be a core of X and f be a homomorphism from X to $X^\bullet$. Then f restricted to $X^\bullet$, *i.e.*, $f|_{X^\bullet}$ is an automorphism of $X^\bullet$. Let this automorphism be denoted by f_1. So,

$$f|_{X^\bullet} = f_1.$$

Let $a, b \in X^\bullet$ and let

$$f : a \longmapsto b \tag{3.5.11}$$

$$\Longrightarrow f_1 : a \longmapsto b \tag{3.5.12}$$

From Equations 3.5.11 and 3.5.12, we conclude that

$$f_1^{-1} \circ f : a \longmapsto a \tag{3.5.13}$$

and hence $f_1^{-1} \circ f$ is an identity map on $X^\bullet$. Thus, $X^\bullet$ is a retract. $\square$

Theorem 3.5.6. *[2] If X is a vertex-transitive graph, then its core $X^\bullet$ is also vertex-transitive.*

Proof. Let X be a vertex-transitive graph and let x and y be two distinct vertices of $X^\bullet$. Then there is an automorphism of X that maps x to y. The composition of this automorphism with a retraction from X to $X^\bullet$ is a homomorphism f from X to $X^\bullet$. The restriction $f|_{X^\bullet}$ is an automorphism of $X^\bullet$ mapping x to y. □

Theorem 3.5.7. *[2] If X is a vertex-transitive graph, then $|V(X^\bullet)|$ divides $|V(X)|$.*

Proof. We show that the fibres of any homomorphism from X to $X^\bullet$ have the same size. Let f be a homomorphism from X to X whose image Y is a core of X. For any element $g \in Aut(X)$, the translate Y^g is mapped onto Y by f, and therefore Y^g has one vertex in each fibre of f.

Now, suppose $v \in V(X)$ and let F be the fibre of f that contains v. Since X is vertex transitive, the number of automorphisms g such that Y^g contains v is independent of our choice of v. If we denote this number by N, then since every image Y^g of Y meets F,

$$|Aut(X)| = |F| \times N.$$

Since N does not depend on F, this implies that all fibres of f have the same size. □

3.6 Kneser Graphs

We, now, study a family of vertex-transitive graphs, the Kneser Graphs.

Definition 3.6.1. For positive integers r, n with $r \leq n$, the **Kneser graph** $K_{n:r}$ is the graph whose vertex set is $\binom{[n]}{r}$ and r-subsets are adjacent if and only if they are disjoint.

A proper coloring is considered to be a binary function on the independent sets of a graph. A generalization of proper coloring is fractional coloring which is also a function on the independent sets of a graph but the function only needs to be non-negative instead of being binary.

For a graph G with $v \in V(G)$, let $\mathcal{I}(G)$ be the set of all independent sets in G and let $\mathcal{I}(G, v)$ be the set of all independent sets in G that contain the vertex v.

Definition 3.6.2. A **fractional coloring** of a graph G is a non-negative function f on the independent sets of G with the property that for any vertex $v \in V(G)$,

$$\sum_{S \in \mathcal{I}(G,v)} f(S) \geq 1.$$

Definition 3.6.3. Let G be a graph and f a fractional coloring on G. The **weight** of f is the sum of the values of f over all independent sets in G, *i.e.*,

$$\sum_{S \in \mathcal{I}(G)} f(S).$$

Definition 3.6.4. The **fractional chromatic number** of a graph G is the minimum possible weight of a fractional coloring. It is denoted by $\chi^*(G)$.

In other words, the fractional chromatic number of a graph G is the minimum weight over all fractional coloring of G.

Kneser graphs are very closely related to fractional coloring. Like the chromatic number being determined by homomorphisms to complete graphs, the fractional chromatic number is determined by homomorphisms to Kneser graphs.

Theorem 3.6.1. *[2] For any graph G,*

$$\chi^*(G) = min\{\frac{n}{r} : \exists \ a \ homomorphism \ G \ \longrightarrow \ K_{n:r}\} \tag{3.6.1}$$

Corollary 3.6.2. *[2] For any graph G,*

$$\omega(G) \leq \omega^*(G) \leq \chi^*(G) \leq \chi(G). \tag{3.6.2}$$

3.7 Remarks

This chapter is devoted to learning some concepts in graph theory which are used in the later chapter. The chapter concerns graph theory concepts like homomorphism, coloring, clique number, chromatic number, independent set, color classes, edge cover, vertex-transitive graph, core, Kneser graph, fractional coloring etc.

Chapter 4

Covering Array on Graphs

Covering arrays have been widely studied for quite some time now. Various aspects of covering arrays like bounds and constructions have made use of concepts pertaining to fields like design theory, intersecting codes, sperner systems, set systems, algebra etc. Covering arrays find usage in various fields in industry and academia. Some of the industrial applications include data compression, drug screening, switching networks, circuit testing, software testing etc whereas academic applications relate to construction of difference matrices, truth functions, search theory etc. In this text, we extend the definition of covering array to include a graph structure. So, covering arrays on graphs are extensions of the standard covering arrays.

4.1 Definition

Definition 4.1.1. t-Qualitative independence: A set of vectors with entries from $\mathbb{Z}_g$ are $t - \textit{qualitatively independent}$ if for any t-subset, $\{v_i\}$, of vectors and any ordered t-tuple of elements $(g_1, g_2, ..., g_t) \in \mathbb{Z}_g^t$ there exists a j such that for each vector v_i the jth coordinate $v_{ij} = g_i$.

Definition 4.1.2. Covering Array: A t-covering array with alphabet size g, k

rows and size n is a $k \times n$ array on $\mathbb{Z}_g$ with the property that any set of t rows is t-qualitatively independent. It is denoted by $t - CA(n, k, g)$.

Here, we only study 2-covering array and hence shall simply call them covering arrays and denote them by $CA(n, k, g)$. Similarly, any pair of 2-qualitatively independent vectors will simply be called qualitatively independent vectors.

Definition 4.1.3. Two vectors v, w in $\mathbb{Z}_g^n$ are qualitatively independent if for all pairs $(a, b) \in \mathbb{Z}_g \times \mathbb{Z}_g$ there is a position i in the vectors where $(a, b) = (v_i, w_i)$.

The smallest possible size of a covering array is denoted by $CAN(k, g)$. So,

$$CAN(k, g) = min\{l \in \mathbb{N} : \exists\ CA(l, k, g)\}.$$

Example: An example of a covering array $CA(5, 4, 2)$ is

$$\begin{matrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{matrix}.$$

For testing any application, every row of the covering array represents a particular parameter of the system and every column represents a test on the system taking different values of the parameters to check for every possible combinations of the parameters. The goal is to check for any deviation from the expected behavior of the system when burdened with all kinds of input. The covering arrays provide a way to produce an array with the fewest number of columns/tests by removing the unnecessary combinations of the parameters (if it was known that certain pairs don't interact) and hence improving on the total number of tests required to check the system thoroughly. We will study strength-2 covering arrays which test all pairwise interactions.

Definition 4.1.4. Covering arrays on graphs: A *covering array on a graph* G, with alphabet size g, is a $|V(G)| \times n$ array on $\mathbb{Z}_g$. Each row in the array corresponds to a vertex in the graph G and the pairs of rows which correspond to adjacent vertices in the graph are qualitatively independent. A covering array on a graph G is denoted by $CA(n, G, g)$.

The size of the smallest possible covering array on a graph G is denoted by $CAN(G, g)$, that is,

$$CAN(G, g) = min\{l \in \mathbb{N} : \exists\, CA(l, G, g)\}.$$

Thus, for the complete graph K_l, we have

$$CAN(K_l, g) = CAN(l, g).$$

Definition 4.1.5. A $CA(n, G, g)$ with $n = CAN(G, g)$ is referred to as an **optimal** covering array on G.

4.2 Bounds from Homomorphisms

Complete graphs can be used to get bounds on $CAN(G, g)$.

Lemma 4.2.1. *[6] Let G and H be graphs. If $f : G \longrightarrow H$ is a graph homomorphism then*

$$CAN(G, g) \leq CAN(H, g).$$

Proof. Let $CAN(H, g) = n$ and suppose B is a matrix with $|V(G)|$ rows and n columns. Set row l of the matrix B to be identical to the row corresponding to $f(a_l)$ in $CA(n, H, g)$ $\forall\, a_l \in G$ where $l \in \{1, 2, ..., |V(G)|\}$. Let a_i and a_j be any two adjacent vertices in graph G. Since f is a graph homomorphism, vertices $f(a_i)$ and $f(a_j)$ are adjacent in graph H and hence the rows in $CA(n, H, g)$ corresponding to $f(a_i)$ and $f(a_j)$ are qualitatively independent. Thus, if we associate rows i and j of matrix B with the pair of adjacent vertices a_i and a_j in G, then matrix B is a $CA(n, G, g)$ (since a pair of qualitatively independent vectors represent a pair of adjacent vertices in the matrix B), implying that

$$CAN(G, g) \leq n$$

or,

$$CAN(G, g) \leq CAN(H, g).$$

□

Hence, the covering array number is monotonically increasing on graphs ordered by homomorphism.

Corollary 4.2.2. *For any graph G,*

$$CAN(K_{\omega(G)}, g) \leq CAN(G, g) \leq CAN(K_{\chi(G)}, g).$$

Proof. We know that there exist homomorphisms

$$f : K_{\omega(G)} \longrightarrow G \text{ and } g : G \longrightarrow K_{\chi(G)}.$$

Using Lemma 4.2.1, we get

$$CAN(K_{\omega(G)}, g) \leq CAN(G, g)$$

and

$$CAN(G, g) \leq CAN(K_{\chi(G)}, g)$$

which combine to give

$$CAN(K_{\omega(G)}, g) \leq CAN(G, g) \leq CAN(K_{\chi(G)}, g).$$

□

4.3 Qualitatively Independent Graphs

Definition 4.3.1. A **g-partition** of an n-set is a set of g disjoint non-empty classes whose union is the n-set. The set of all g-partitons of an n-set is denoted by $\mathcal{P}_g^n$.

Definition 4.3.2. Qualitatively Independent Partitions: Let n and g be positive integers with $n \geq g^2$. Let $A,\ B \in \mathcal{P}_g^n$ be two g-partitions of an n-set. Assume $A = \{A_1, A_2, ..., A_g\}$ and $B = \{B_1, B_2, ..., B_g\}$. The partitions A and B are *qualitatively independent* if

$$A_i \cap B_j \neq \phi \ for\ all\ i\ and\ j.$$

If g-partitions $A = \{A_1, A_2, ..., A_g\}$ and $B = \{B_1, B_2, ..., B_g\}$ are qualitatively independent, then for each $i \in \{1, 2, ..., g\}$, $|A_i| \geq g$ and $|B_i| \geq g$.

Definition 4.3.3. Let n and g be positive integers with $n \geq g^2$. The **qualitative independence graph** $QI(n, g)$ is defined to be the graph whose vertex set is the set of all g-partitions of an n-set with the property that every class of the partition has size at least g. Vertices are adjacent if and only if the corresponding partitions are qualitatively independent.

Illustrative Examples:

1. $QI(4, 2)$=$\{\{12|34\}, \{13|24\}, \{14|23\}\}$.

2. $QI(5,2)$={{123|45}, {124|35}, {134|25}, {125|34}, {135|24}, {145|23}, {12|345}, {13|245}, {14|235}, {15|234}}.

There is a natural correspondence between length n g-ary vectors and partitions of an n-set into g classes: The indices of the vector with a common letter are in the same class of the corresponding partition.

Then, we can alternatively define the qualitative independence graphs as follows.

Definition 4.3.4. For two positive integers g and n, where $n \geq g^2$, we define the **qualitative independence graph** $QI(n,g)$. The vertices are all length n vectors over $\mathbb{Z}_g$, in which each letter occurs at least g times. Additionally, the vectors have 0 in their first position and the first appearances of each letter are in lexicographic order. Two vertices are adjacent if the two vectors that they represent are qualitatively independent.

Illustrative Examples:

1. $QI(4,2)$={(0011), (0101), (0110)}.

2. $QI(5,2)$={(00011), (00101), (01001), (00110), (01010), (01100), (00111), (01011), (01101), (01110)}.

Note: In graph $QI(4,2)$, all the vertices are pairwise qualitatively independent to each other and hence is isomorphic to K_3. Any covering array with alphabet size 2 and having four columns can only be written in terms of these three vectors, thus any graph G with a covering array $CA(4,G,2)$ must have a homomorphism to this graph K_3.

Definition 4.3.5. The **weight** of an n-bit binary vector is the number of 1's in the vector. It is denoted by $w(v)$ for a vector v.

In $QI(n,g)$, we can assume that the first letter of every vector is 0 and further, when the vector is not binary, that the first appearances of each letter are in lexicographic order (i.e. the first instance of each letter in a row appears in the same order as the natural ordering of the letter). If a vector is not of this form, we can permute the letters so that it becomes one. This action does not change the partition. For n-bit binary vectors we can alternatively assume that the most frequent letter is 0 which means the weight of the vector is no more than $\lfloor n/2 \rfloor$. If a vector has a larger weight, we can exchange the 0's and 1's which is equivalent to using the complement of the corresponding set and thus does not change the 2-partition.

If two binary vectors correspond to sets which have the property that one contains the other, then the vectors are not qualitatively independent. Similarly, if the sets are disjoint, the vectors are again not qualitatively independent. Vectors are qualitatively independent if and only if the corresponding sets intersect each other but neither is completely contained in the other. Finally, if we consider only the vectors of weight not greater than $\lfloor n/2 \rfloor$, then two distinct vectors with the same weight will be qualitatively independent if and only if their corresponding sets intersect.

Definition 4.3.6. A vector u **dominates** a vector v iff vector v has zero in at least all the entries where u has a zero. It is denoted by writing $v \subseteq u$.

In the binary case, the poset with the dominance relation on the vectors coincides with the poset of subsets of n ordered by inclusion. Since we can assume that $w(s) \leq \lfloor n/2 \rfloor$, the binary vectors correspond to the lower half of the poset of subsets of n. Vectors that are related in the poset are not qualitatively independent.

Theorem 4.3.1. *[7] For a graph G and non-negative integers g and n there exists a $CA(n,G,g)$ if and only if there exists a graph homomorphism from $G \longrightarrow QI(n,g)$.*

Proof. Assume that there exists a $CA(n, G, g)$, call this C. Consider a mapping $f : V(G) \longrightarrow V(QI(n, g))$ that takes a vertex $v \in V(G)$ to the vertex in $QI(n, g)$ which corresponds to the vector in C for v. Call this vector C_v. With this mapping, consider two adjacent vertices v and w in G so that $\{v, w\} \in E(G)$. The vectors C_v and C_w, being the rows of $CA(n, G, g)$ corresponding to adjacent vertices, are qualitatively independent implying that $\{C_v, C_w\} \in E(QI(n, g))$. Hence, the map f is a homomorphism.

Conversely, let there be a homomorphism f from G to $QI(n, g)$ that takes a vertex $a_l \in V(G)$ to the vertex $f(a_l) \in V(QI(n, g))$ where $l \in \{1, 2, ..., |V(G)|\}$. And let B be a matrix with $\mid G \mid$ rows and n columns, having its row l identical to $f(a_l)$. Now, for any two adjacent vertices a_i and a_j in G, if we associate rows i and j of matrix B to them respectively, then the matrix B is a $CA(n, G, g)$ (since rows i and j are essentially $f(a_i)$ and $f(a_j)$ respectively which are qualitatively independent). Hence, there exists a $CA(n, G, g)$. $\square$

This gives a bound on $\chi(G)$ for all graphs with $CAN(G, g) \leq n$.

Corollary 4.3.2. *[7] Let G be any graph and g and n integers such that there exists a $CA(n, G, g)$. Then $\chi(G) \leq \chi(QI(n, g))$.*

Proof. From the above theorem, there exists a $CA(n, G, g)$ if and only if there exists a graph homomorphism from $G \longrightarrow QI(n, g)$. So, $\chi(G) \leq \chi(QI(n, g))$. $\square$

As $QI(4, 2) = K_3$, a graph G has a $CA(4,\ G,\ 2)$ if and only if there is a homomorphism of G to K_3. This means that deciding if a graph has a covering array of size 4 and alphabet size 2 is exactly the same as deciding if it is 3 colorable.

The size of the vertex set for $QI(n, 2)$ is $2^{n-1} - n - 1$, this is the number of subsets of n of size at least 2 and no more than $\lfloor n/2 \rfloor$. It is also equal to $(x - y - z)$, where $x = 2^{n-1}$ = total number of length n vectors which have 0 as their first entry, $y = n$ = total number of length n vectors which have 0 as their first entry and with at most one 1 in the rest of the $(n-1)$ entries, $z = 1$ = total number of length n vectors which have 0 in their first entry but no 0 in the rest of the $(n-1)$ entries. This is because y and z account for the counting of all the length n vectors which have weight either less than 2 or greater than $(n-2)$.

Definition 4.3.7. An **edge cover** of a graph is a set of edges so that each vertex is the terminus for some edge in the set.

We denote the number of edges incident to a vertex v by $d(v)$ and the minimum value of this over all vertices as $\delta(G)$.

Definition 4.3.8. The *cover index* $\kappa(G)$ of G is the largest number k such that $E(G)$ can be partitioned into k edge covers.

Theorem 4.3.3. *[5] For any bipartite graph G,*

$$\kappa(G) = \delta(G).$$

Lemma 4.3.4. *[7] The graphs $QI(n, g)$ are connected and have diameter 2.*

Proof. Construct a bipartite graph C with g vertices in both the parts using elements from $\mathbb{Z}_g$. Let u and v be any two vertices of $QI(n, g)$ which are not qualitatively independent. Denote the entry i of the vector u and v by u_i and v_i respectively. For each of the $1 \leq i \leq n$, we add the edge (u_i, v_i) to $E(C)$. Since every letter occurs at least g times in both the vectors, we have $\delta(C) \geq g$. By Thm 4.3.3, C can be decomposed into $\delta(C) \geq g$ edge disjoint edge-covers. Let us call these C_l. To the first

g edge-covers we assign a unique letter from the g-ary alphabets (i.e. 0 to $g-1$), to the rest we assign 0. Now, let $w \in QI(n,g)$ be such that for each i, w_i is the letter assigned to the C_l that contains the edge (u_i, v_i). With this construction of w, we take any pair $(g_i, g_j) \in \mathbb{Z}_g^2$. Consider the edge-cover C_{g_i}. In this edge-cover, there is an edge which has g_j as the terminus on the left side (the side representing the entries of vector u). This edge corresponds to a position in the vectors, say position n_{ij}. Since this edge comes from the edge-cover C_{g_i}, which has been assigned the letter g_i by definition, we get $w_{n_{ij}} = g_i$. Also, this edge has g_j as the terminus at the position n_{ij} of the vector u, so $u_{n_{ij}} = g_j$. So, this pair is covered between w and u. Since, this is an arbitrary pair, we conclude that w and u are qualitatively independent. Similarly, w and v are qualitatively independent. Hence, given any two vertices u and v, it is possible to construct a vertex w which is qualitatively independent to both u and v and hence adjacent to both in $QI(n,g)$, thus proving that the graphs $QI(n,g)$ are connected and have diameter 2. $\square$

Illustrative example for the above result :

Let $u = (0,2,1,2,1,0,0,2,1,0,2,1)$ and $v = (0,0,1,1,2,0,2,2,1,0,1,2)$ with $g = 3$ and $n = 12$. So, there are 12 edges which enables the construction of 4 edge disjoint edge-covers shown as follows:

1. C_0: $\{0 \longmapsto 0$ *by* $\{u_1, v_1\}$, $1 \longmapsto 1$ *by* $\{u_3, v_3\}$, $2 \longmapsto 2$ *by* $\{u_8, v_8\}\}$
2. C_1: $\{0 \longmapsto 0$ *by* $\{u_6, v_6\}$, $1 \longmapsto 2$ *by* $\{u_5, v_5\}$, $2 \longmapsto 1$ *by* $\{u_4, v_4\}\}$
3. C_2: $\{0 \longmapsto 2$ *by* $\{u_7, v_7\}$, $2 \longmapsto 0$ *by* $\{u_2, v_2\}$, $1 \longmapsto 1$ *by* $\{u_9, v_9\}\}$
4. C_0: $\{0 \longmapsto 0$ *by* $\{u_{10}, v_{10}\}$, $1 \longmapsto 2$ *by* $\{u_{12}, v_{12}\}$, $2 \longmapsto 1$ *by* $\{u_{11}, v_{11}\}\}$.

Based on the above construction, $w = (0,2,0,1,1,1,2,0,2,0,0,0)$. Clearly, w and u

are qualitatively independent and so are w and v.

Elements of $QI(n, 2)$ are the 2-partitions of the n-set such that every class of any partition has size at least 2 and at most $(n - 2)$. So, every partition in $QI(n, 2)$ is of the form $P = \{P_1,\ P_2\}$ such that $|P_i| \geq 2$ and $\overline{P_1} = P_2$ (or, $\overline{P_2} = P_1$). Thus, we can uniquely represent every partition $P \in QI(n, 2)$ by the class P_1 or P_2 whichever is smaller in size (*i.e.* which has size less than or equal to $\lfloor \frac{n}{2} \rfloor$).

Corollary 4.3.5. *[5] For all positive integers n,*

$$\omega(QI(n, 2)) = \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}. \tag{4.3.1}$$

Proof. Let S be an n-element set.

1. Let n be even. Then the $n/2$-element subsets of the n-set containing one particular element of S form an $\binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}$ member collection which is qualitatively independent, which can be seen as follows.
   Let $F = \{A \in S\ :\ |A| = \frac{n}{2}\ and\ 1 \in A\}$ and let $A, B \in F$.
   Clearly, $1 \in A \cap B$. So,
   $$A \cap B \neq \phi. \tag{4.3.2}$$
   Since $A \neq B$ and $|A| = \frac{n}{2} = |B|$, there exists some $x \in A$ such that x doesn't belong to B. So, $x \in A \cap \overline{B}$ and hence
   $$A \cap \overline{B} \neq \phi.$$
   Similarly, $\overline{A} \cap B \neq \phi$. Now, if possible assume that $\overline{A} \cap \overline{B} = \phi$ which implies that
   $$\overline{B} \subseteq A. \tag{4.3.3}$$

But we know that $|A| = \frac{n}{2} = |\overline{B}|$ and hence from Equation 4.3.3, we conclude that

$$\overline{B} = A \tag{4.3.4}$$

which is not true as $A \cap B \neq \phi$ (from Equation 4.3.2). Thus, our assumption that $\overline{A} \cap \overline{B} = \phi$ is proven false and we conclude that,

$$\overline{A} \cap \overline{B} \neq \phi. \tag{4.3.5}$$

From above discussion, we conclude that the set F is qualitatively independent.

Now, to show that no qualitatively independent set F can be larger, we note that if A and B are in F, then neither A nor $\overline{A}$ can be contained in B or $\overline{B}$. Thus the members of F and their complements form a family F^* of subsets of S wherein no member is contained in any other. So, F^* is a Sperner set system and from Theorem 2.3.2, we conclude that F^* can have no more than $\binom{n}{\frac{n}{2}}$ members, so that F can have no more than $\frac{1}{2}\binom{n}{n/2}$ or $\binom{n-1}{\frac{n}{2}-1}$ members.

2. Suppose now that n is odd. We can construct (as above) an $\binom{n-1}{\lfloor\frac{n}{2}\rfloor-1}$ member qualitatively independent collection of subsets of S by choosing all $\lfloor\frac{n}{2}\rfloor$ element subsets containing some particular element.

   To show that $\binom{n-1}{\lfloor\frac{n}{2}\rfloor-1}$ is an upper bound as well, we consider the members of F^* (defined as above) which have $\lfloor\frac{n}{2}\rfloor$ or fewer elements. Clearly, there are $|F|$ of them. Every one of these must intersect every other and hence by Theorem 2.6.1, there are therefore at most $\binom{n-1}{\lfloor\frac{n}{2}-1}$ of them.

$\square$

The set of all vectors with weight $n/2$ starting with 0 is a maximum clique when n is even. The set of all vectors with weight $(n-1)/2$ starting with 0 is a maximum clique when n is odd.

Theorem 4.3.6. *[2] Let X be a connected vertex-transitive graph. Then X has a matching that misses at most one vertex, and each edge is contained in a maximum matching.*

It implies that a connected vertex-transitive graph on an even number of vertices has a perfect matching, and that each vertex in a connected vertex-transitive graph on an odd number of vertices is missed by a matching that covers all the remaining vertices.

Theorem 4.3.7. *[7] For all positive integers n, $\chi(QI(n,2)) = \lceil \frac{1}{2}\binom{n}{\lfloor n/2 \rfloor} \rceil$.*

Proof. Consider all the vertices in $QI(n,2)$ as subsets of n of size no more than $\lfloor n/2 \rfloor$ as described above. We know that the poset of subsets of an n-set ordered by inclusion can be decomposed into $\binom{n}{\lfloor n/2 \rfloor}$ disjoint chains and each chain contains exactly one set of size $\lfloor n/2 \rfloor$. Lets call these chains C_i, where $i \in \{1, 2, ..., \binom{n}{\lfloor n/2 \rfloor}\}$. For any i, if the sets $A, B \in C_i$, then A and B are not qualitatively independent (and hence non-adjacent) *i.e.* each chain corresponds to an independent set in $QI(n,2)$.

It is possible to pair the $\binom{n}{\lfloor n/2 \rfloor}$ chains so that any subset of size no more than $n/2$ in one chain is disjoint from any subset of size no more than $n/2$ in the other chain. This is possible if the two chains being paired are such that their $\lfloor n/2 \rfloor$ sized sets are disjoint. To see this, lets consider them case wise as following.

1. Assume that n is even. For each chain C_i, let A_i be the set of size $n/2$. Match two chains C_i and C_j if $A_i = \overline{A_j}$.

2. Assume that n is odd. For each chain C_i, let A_i be the set of size $(n-1)/2$. The sets A_i to be matched should be disjoint and hence form the vertices of the Kneser graph $K_{n:\frac{n-1}{2}}$. The Kneser graph $K_{n:\frac{n-1}{2}}$ is vertex transitive, so there exists a matching that is perfect or is missing just one vertex. So, each set A_i (except possibly one set) is matched to another set of size $(n-1)/2$, lets call it $A_i^{'}$. The set $A_i^{'} \subset \overline{A_i}$. Match the chain C_i which contains A_i with the chain $C_{i'}$ that contains the set $A_i^{'}$.

Any two sets in a matched pair of chains have the property that either one set contains the other or one set contains the complement of the other (since we are restricted to vertices of $QI(n,2)$ which have size at most $\lfloor n/2 \rfloor$ only). In either case, the vertices are not qualitatively independent, hence all the vertices in the paired chains can be assigned the same color in a proper coloring of $QI(n,2)$. This produces a proper $\lceil \frac{1}{2}\binom{n}{\lfloor n/2 \rfloor} \rceil$-coloring on the graph $QI(n,2)$.

To see that this is the smallest possible coloring of $QI(n,2)$, consider the vertices of $QI(n,2)$ that correspond to $\lfloor n/2 \rfloor$-sets. Two such vertices may be assigned the same color if and only if the subsets are disjoint. It is clear that there can't be three mutually disjoint subsets of an n-set with size $\lfloor n/2 \rfloor$ (since $3\lfloor n/2 \rfloor > n$). So it is not possible to properly color these vertices with fewer than $\lceil \frac{1}{2}\binom{n}{\lfloor n/2 \rfloor} \rceil$ colors. □

Corollary 4.3.8. *For n even,* $\lceil \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor} \rceil = \frac{1}{2}\binom{n}{n/2} = \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}$.

Proof. Since n is even, let $n = 2m$. Hence, $\lceil \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor} \rceil = \lceil \frac{1}{2}\binom{2m}{m} \rceil = \lceil \frac{1}{2}\frac{(2m)}{m}\binom{2m-1}{m-1} \rceil = \lceil \binom{2m-1}{m-1} \rceil = \binom{2m-1}{m-1} = \frac{1}{2}\binom{n}{n/2}$. Also, $\binom{2m-1}{m-1} = \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}$ and hence proved. □

Theorem 4.3.9. *[7] For n even, the core of $QI(n,2)$ is* $K_{\frac{1}{2}\binom{n}{n/2}}$.

Proof. We know that $\omega(QI(n,2)) = \binom{n-1}{\frac{n}{2}-1} = \frac{1}{2}\binom{n}{n/2}$ for n even and $\chi(QI(n,2)) = \lceil \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor} \rceil = \frac{1}{2}\binom{n}{n/2}$ for n even. Hence, there exist homomorphisms

$$K_{\frac{1}{2}\binom{n}{n/2}} \longrightarrow QI(n,2) \longrightarrow K_{\frac{1}{2}\binom{n}{n/2}}.$$

Suppose there exists a $G_1 \subset K_{\frac{1}{2}\binom{n}{n/2}}$ such that there exist homomorphisms

$$G_1 \longrightarrow QI(n,2) \longrightarrow G_1.$$

Then we have,

$$\omega(G_1) = \omega(QI(n,2)). \tag{4.3.6}$$

Since $G_1 \subset K_{\frac{1}{2}\binom{n}{n/2}}$, we have

$$\omega(G_1) < \omega(K_{\frac{1}{2}\binom{n}{n/2}}) = \frac{1}{2}\binom{n}{n/2}. \tag{4.3.7}$$

Combining equations 4.3.6 and 4.3.7, we get $\omega(QI(n,2)) < \frac{1}{2}\binom{n}{n/2}$ which is a contradiction to the fact that $\omega(QI(n,2)) = \frac{1}{2}\binom{n}{n/2}$. Hence, no such G_1 exists and $K_{\frac{1}{2}\binom{n}{n/2}}$ is the minimal subgraph of G such that there exist homomorphisms

$$K_{\frac{1}{2}\binom{n}{n/2}} \longrightarrow QI(n,2) \longrightarrow K_{\frac{1}{2}\binom{n}{n/2}}$$

making it the core of the graph $QI(n,2)$.

Alternatively, the complete graph $K_{\frac{1}{2}\binom{n}{n/2}}$ is itself a core and there exists a homomorphism from $QI(n,2)$ to $K_{\frac{1}{2}\binom{n}{n/2}}$ (as discussed above) making it the core of the graph $QI(n,2)$. $\square$

We denote by **F(n,2)** the induced subgraph of $QI(n,2)$ containing the vertices with weight $\lfloor n/2 \rfloor$.

Corollary 4.3.10. *For n odd, there exists a graph homomorphism from $QI(n,2)$ to $F(n,2)$.*

Proof. Consider all the vectors in $QI(n,2)$ as subsets of an n-set of size no more than $\lfloor n/2 \rfloor$. The posets of subsets of the n-set ordered by inclusion can be decomposed into $\binom{n}{\lfloor n/2 \rfloor}$ disjoint chains and each chain contains exactly one set of size $\lfloor n/2 \rfloor$. A function f from $QI(n,2)$ to $F(n,2)$, mapping any vector a of a chain to the unique vector of size $\lfloor n/2 \rfloor$, $a^{'}$, contained in that chain, is a homomorphism. This can be proved as follows. Let a and b be any two adjacent vectors mapped to $a^{'}$ and $b^{'}$ respectively. Since a and b are adjacent, there exists at least one $j \in \{1,2,...,n\}$ such that $(a_j, b_j) = (1,1)$ and hence $(a_j^{'}, b_j^{'}) = (1,1)$ (this is because $\forall\ j \in \{1,2,...,n\}$ for which $a_j = 1$, we have $a_j^{'} = 1$). Now, since the vectors $a^{'}$ and $b^{'}$ are not the same, there should be at least one $l \in \{1,2,...,n\}$ such that $(a_l^{'}, b_l^{'})$=$(1,0)$ or $(0,1)$. Let it be $(0,1)$ that is covered between $a^{'}$ and $b^{'}$ at least once. Then the pair $(1,0)$ is also covered otherwise the size of vector $b^{'}$ will exceed that of $a^{'}$ (because the absence of $(1,0)$ gives more 1's entries to $b^{'}$ than $a^{'}$). Now, suppose that there is no $k \in \{1,2,...,n\}$ such that $(a_k^{'}, b_k^{'}) = (0,0)$. Since the weight of $a^{'}$ is $\lfloor n/2 \rfloor = (n-1)/2$, there must be $(n - \frac{(n-1)}{2}) = (\frac{n+1}{2})$ zero entries in $a^{'}$. For every i for which $a_i^{'} = 0$, we have $b_i^{'} = 1$ (since $(0,0)$ is not covered between $a^{'}$ and $b^{'}$ by assumption). But this means that the weight of $b^{'}$ is greater than $\lfloor n/2 \rfloor$ which is a contradiction and hence the assumption that $(0,0)$ is not covered between $a^{'}$ and $b^{'}$ is not true. From the above discussion, we conclude that all the possible pairs $(a,b) \in \mathbb{Z}_2^2$ are covered between $a^{'}$ and $b^{'}$ and hence they are adjacent. So, the function f maps adjacent vertices a and b to adjacent vertices $a^{'}$ and $b^{'}$ and hence is a homomorphism from $QI(n,2)$ to $F(n,2)$. $\square$

Corollary 4.3.11. *For n odd, we have*

$$\chi(F(n,2)) = \chi(QI(n,2)) \textit{ and } \omega(F(n,2)) = \omega(QI(n,2))$$

Proof. From Corollary 4.3.10, for n odd, there exists a homomorphism

$$QI(n,2) \longrightarrow F(n,2) \tag{4.3.8}$$

and since $F(n,2) \subseteq QI(n,2)$, there exists another homomorphism namely the identity map

$$I \ : \ F(n,2) \longrightarrow QI(n,2). \tag{4.3.9}$$

Combining Equations 4.3.8 and 4.3.9, we conclude that there exists homomorphisms

$$QI(n,2) \longrightarrow F(n,2) \longrightarrow QI(n,2). \tag{4.3.10}$$

From Equation 4.3.10, we have

$$\chi(F(n,2)) = \chi(QI(n,2)) \textit{ and } \omega(F(n,2)) = \omega(QI(n,2))$$

□

Theorem 4.3.12. *[7] For n odd, the core of $QI(n,2)$ is $F(n,2)$.*

Proof. Let $F^{\bullet}(n,2)$ be the core of $F(n,2)$. Since $F(n,2)$ is vertex transitive, $F^{\bullet}(n,2)$ is also vertex transitive (from Theorem 3.5.6) and

$$|V(F^{\bullet}(n,2))| \textit{ divides } |V(F(n,2))| = \binom{n}{(n-1)/2} \ (\textit{from Thm } 3.5.7). \tag{4.3.11}$$

Also, as $F^{\bullet}(n,2)$ is the core of the graph $F(n,2)$,

$$\chi(F^{\bullet}(n,2)) = \chi(F(n,2)) = \lceil \frac{1}{2}\binom{n}{\frac{n-1}{2}} \rceil. \tag{4.3.12}$$

Since the number of vertices in a graph is always greater than or equal to the chromatic number of the graph, we have

$$|V(F^{\bullet}(n,2))| \geq \lceil \frac{1}{2}\binom{n}{\frac{n-1}{2}} \rceil. \tag{4.3.13}$$

From Equations 4.3.11 and 4.3.13, there are only two possibilities for $|V(F^{\bullet}(n,2))|$, either it is $\frac{1}{2}\binom{n}{(n-1)/2}$ or $\binom{n}{(n-1)/2}$. If $|V(F^{\bullet}(n,2))| = \frac{1}{2}\binom{n}{(n-1)/2}$ then $F^{\bullet}(n,2)$ would have to be a complete graph which is not the case since

$$\omega(F^{\bullet}(n,2)) = \omega(F(n,2)) = \omega(QI(n,2)) = \binom{n-1}{\frac{n-3}{2}} < \frac{1}{2}\binom{n}{(n-1)/2}.$$

Thus,

$$|V(F^{\bullet}(n,2))| = \binom{n}{(n-1)/2} \tag{4.3.14}$$

and as $F^{\bullet}(n,2)$ is an induced subgraph of $F(n,2)$, we have $F^{\bullet}(n,2) = F(n,2)$. This means that $F(n,2)$ is a core. Also, since there exists a homomorphism from $QI(n,2)$ to $F(n,2)$ (from Corollary 4.3.10), we conclude that $F(n,2)$ is a core of $QI(n,2)$ for n odd. □

Corollary 4.3.13. *For n odd, $F(n,2)$ is isomorphic to the set of vectors of $QI(n,2)$ with weight $(n+1)/2$.*

Proof. Let n be odd so that $\lfloor n/2 \rfloor = (n-1)/2$. And let a and b be any two vectors from $F(n,2)$. Define a map f from $F(n,2)$ to the set of vectors of $QI(n,2)$ with weight $(n+1)/2$, such that any vector a goes to it's complement $\overline{a}$ making f a one-to-one map. So, the function f maps a vector of weight $(n-1)/2$ to a vector of weight $(n+1)/2$. Clearly, if a and b are adjacent then so are $\overline{a}$ and $\overline{b}$ and vice versa. Hence, f maps adjacent vertices to adjacent vertices and hence is a homomorphic map. Since the number of vectors of weight $(n-1)/2$ is equal to the number of vectors of weight

$(n+1)/2$, the one-to-one function f is also a bijection. So, an inverse map f^{-1} can be defined which maps $\overline{a}$ to a which again is a homomorphism, thus proving that the map f is an isomorphic map. $\square$

Corollary 4.3.14. *For n even, $K_{\frac{1}{2}\binom{n}{n/2}}$ is isomorphic to the set of vectors of $QI(n,2)$ with their first entry equal to zero and weight $n/2$.*

Proof. Let a and b be any two vectors from $QI(n,2)$ with their first entry equal to zero and weight $n/2$. Let entry i of the vectors a and b be denoted by a_i and b_i respectively. Thus, $(a_1,b_1)=(0,0)$. Now, suppose that there exists no i such that $(a_i,b_i)=(1,1)$. So, the only values which (a_i,b_i) can attain are $(0,0)$, $(1,0)$ and $(0,1)$. Since the weight of the vector a is $n/2$, there should be exactly $n/2$ j's such that $(a_j,b_j)=(1,0)$. Because $(a_1,b_1)=(0,0)$, the weight of vector b can be at most $n-1-(n/2)=(n/2)-1$, which is a contradiction to the fact that the weight of vector b is $n/2$. Hence, there exists a j such that $(a_j,b_j)=(1,1)$. Again, let us suppose that there exists no i for which $(a_i,b_i)=(0,1)$. So, the only values which (a_i,b_i) can attain, in principle, are $(0,0)$, $(1,0)$ and $(1,1)$. There has to be at least one i for which $(a_i,b_i)=(1,0)$ otherwise the two vectors a and b will be the same. But, this means that vector a has more 1's than b which creates a difference in the weight of the two vectors leading to a contradiction. Hence, there always exists an i for which $(a_i,b_i)=(0,1)$. Using the same argument, it can be proved that there always exists an i for which $(a_i,b_i)=(1,0)$. From the above discussion, we can conclude that vectors a and b are qualitatively independent. Since they are two arbitrary vectors from $QI(n,2)$ with their first entry equal to zero and weight $n/2$, this proves that all the vectors in this set are qualitatively independent to each other and they are $\frac{1}{2}\binom{n}{n/2}$ in number. Hence, there exists an isomorphism between $K_{\frac{1}{2}\binom{n}{n/2}}$ and the set of vectors of $QI(n,2)$ with their first entry equal to zero and weight $n/2$. $\square$

We denote the core of the graph $QI(n,g)$ by $QI^{\bullet}(n,g)$.

Theorem 4.3.15. *[7] If there exists a $CA(n,G,2)$ it is always possible to find a covering array $CA(n,G,2)$ in which the rows are vectors with weight $\lfloor n/2 \rfloor$. Moreover, if n is even, it is possible to find such a covering array with the rows all beginning with 0.*

Proof. The core of $QI(n,2)$ for n even is $K_{\frac{1}{2}\binom{n}{n/2}}$ and $F(n,2)$ when n odd. For any graph G, if there exists a $CA(n,G,2)$ then there exists a graph homomorphism

$$G \longrightarrow QI(n,2)\ (from\ Theorem\ 4.3.1) \tag{4.3.15}$$

and there exists another homomorphism

$$QI(n,2) \longrightarrow QI^{\bullet}(n,2) \tag{4.3.16}$$

from the definition of the core of a graph. Hence, from Theorem 4.3.9, Corollary 4.3.14 and Equations 4.3.15 and 4.3.16, for n even, we have homomorphisms

$$G \longrightarrow QI(n,2) \longrightarrow K_{\frac{1}{2}\binom{n}{n/2}} \longrightarrow A. \tag{4.3.17}$$

(where A is the set of vectors of $QI(n,2)$ with their first entry equal to zero and weight $n/2$). From Theorem 4.3.12 and Equations 4.3.15 and 4.3.16, for n odd, we have homomorphisms

$$G \longrightarrow QI(n,2) \longrightarrow F(n,2). \tag{4.3.18}$$

The desired covering array on G can be pulled back through these homomorphisms. $\square$

Lemma 4.3.16. *For all $n \geq 4$,*

$$\chi(QI^{\bullet}(n-1,2)) < \omega(QI^{\bullet}(n,2)) \leq \chi(QI^{\bullet}(n,2)) < \omega(QI^{\bullet}(n+1,2)).$$

Proof. We already know that $\omega(G) \leq \chi(G)$ for any graph G, so we only need to prove that $\chi(QI^{\bullet}(n-1,2)) < \omega(QI^{\bullet}(n,2))$ and $\chi(QI^{\bullet}(n,2)) < \omega(QI^{\bullet}(n+1,2))$.

1. **Proving** $\chi(QI^{\bullet}(n-1,2)) < \omega(QI^{\bullet}(n,2))$:

   We know that $\chi(G) = \chi(G^{\bullet})$ and $\omega(G) = \omega(G^{\bullet})$. Also that,

   $$\chi(QI(n,2)) = \lceil \frac{1}{2}\binom{n}{\lfloor n/2 \rfloor} \rceil \; \textit{and} \; \omega(QI(n,2)) = \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}.$$

   Case 1: When n is **even**:

   $$\begin{aligned} &\chi(QI^{\bullet}(n-1,2)) < \omega(QI^{\bullet}(n,2)) \\ \Leftrightarrow\ &\lceil \frac{1}{2}\binom{n-1}{\lfloor (n-1)/2 \rfloor} \rceil < \binom{n-1}{\lfloor n/2 \rfloor - 1} \\ \Leftrightarrow\ &\lceil \frac{1}{2}\binom{n-1}{(n-2)/2} \rceil < \binom{n-1}{\frac{n}{2} - 1} \\ \Leftrightarrow\ &\lceil \frac{1}{2}\binom{n-1}{\frac{n}{2} - 1} \rceil < \binom{n-1}{\frac{n}{2} - 1}, \; \textit{which is true for } n \geq 4. \end{aligned}$$

   Case 2: When n is **odd**:

   $$\begin{aligned} &\chi(QI^{\bullet}(n-1,2)) < \omega(QI^{\bullet}(n,2)) \\ \Leftrightarrow\ &\lceil \frac{1}{2}\binom{n-1}{\lfloor (n-1)/2 \rfloor} \rceil < \binom{n-1}{\lfloor n/2 \rfloor - 1} \\ \Leftrightarrow\ &\lceil \frac{1}{2}\binom{n-1}{(n-1)/2} \rceil < \binom{n-1}{\frac{n-1}{2} - 1} \\ \Leftrightarrow\ &\lceil \binom{n-2}{(n-3)/2} \rceil < \binom{n-1}{(n-3)/2} \\ \Leftrightarrow\ &\binom{n-2}{(n-3)/2} < \binom{n-1}{(n-3)/2}, \; \textit{which is true for } n \geq 4. \end{aligned}$$

2. **Proving** $\chi(QI^{\bullet}(n,2)) < \omega(QI^{\bullet}(n+1,2))$:

   Case 1: When n is **even**:

$$\chi(QI^{\bullet}(n,2)) < \omega(QI^{\bullet}(n+1,2))$$

$$\Leftrightarrow \lceil \frac{1}{2}\binom{n}{\lfloor n/2 \rfloor} \rceil < \binom{n}{\lfloor \frac{n+1}{2} \rfloor - 1}$$

$$\Leftrightarrow \lceil \frac{1}{2}\binom{n}{n/2} \rceil < \binom{n}{\frac{n}{2} - 1}$$

$$\Leftrightarrow \lceil \binom{n-1}{\frac{n}{2} - 1} \rceil < \binom{n}{\frac{n}{2} - 1}$$

$$\Leftrightarrow \binom{n-1}{\frac{n}{2} - 1} < \binom{n}{\frac{n}{2} - 1}, \; \textit{which is true for } n \geq 4.$$

   Case 2: When n is **odd**:

$$\chi(QI^{\bullet}(n,2)) < \omega(QI^{\bullet}(n+1,2))$$

$$\Leftrightarrow \lceil \frac{1}{2}\binom{n}{\lfloor n/2 \rfloor} \rceil < \binom{n}{\lfloor \frac{n+1}{2} \rfloor - 1}$$

$$\Leftrightarrow \lceil \frac{1}{2}\binom{n}{\frac{n-1}{2}} \rceil < \binom{n}{\frac{n+1}{2} - 1}$$

$$\Leftrightarrow \lceil \frac{1}{2}\binom{n}{\frac{n-1}{2}} \rceil < \binom{n}{\frac{n-1}{2}}, \; \textit{which is true for } n \geq 4.$$

$\square$

Lemma 4.3.17. *[7] For all $n \geq 4$, $CAN(QI^{\bullet}(n,2),2) = n$.*

Proof. From the definition of the core of a graph, there exists a homomorphism $QI^{\bullet}(n,2) \longrightarrow QI(n,2)$. But we know that there exists a graph homomorphism $QI^{\bullet}(n,2) \longrightarrow QI(n,2)$ if and only if there exists a $CA(n, QI^{\bullet}(n,2), 2)$ and hence

$CAN(QI^{\bullet}(n,2),2) \leq n$. Let us assume that $CAN(QI^{\bullet}(n,2),2) < n$ which means $CAN(QI^{\bullet}(n,2),2) \leq n-1$, so that there is a homomorphism

$$QI^{\bullet}(n,2) \longrightarrow QI(n-1,2).$$

But there exists another homomorphism

$$QI(n-1,2) \longrightarrow QI^{\bullet}(n-1,2)$$

by the definition of the core of a graph. Thus, there exists a homomorphism

$$QI^{\bullet}(n,2) \longrightarrow QI^{\bullet}(n-1,2).$$

But this contradicts the fact that

$$\chi(QI^{\bullet}(n-1,2)) < \chi(QI^{\bullet}(n,2)). \tag{4.3.19}$$

Hence, $CAN(QI^{\bullet}(n,2),2) = n$. $\square$

Now, we know that there exists a $CA(n,G,g)$ if and only if there exists a graph homomorphism from $G \longrightarrow QI(n,g)$. Putting $g=2$ and $G = QI(m,2)$, there exists a homomorphism $QI(m,2) \longrightarrow QI(n,2)$ if and only if there exists a $CA(n, QI(m,2),2)$. Hence, the minimum of n for which $CA(n, QI(m,2),2)$ exists is the same as the minimum of n for which there exists a homomorphism

$$QI(m,2) \longrightarrow QI(n,2), \textit{ and hence,}$$

$$CAN(QI(m,2),2) = min\{n \in \mathbb{Z} : \exists \textit{ a homomorphism } QI(m,2) \longrightarrow QI(n,2)\}$$

$$or, \; CAN(QI(m,2),2) = m. \tag{4.3.20}$$

Corollary 4.3.18. *[5]* $CAN(K_{\chi(G)},2) = min\{n \; : \; \binom{n-1}{\lfloor \frac{n}{2} \rfloor -1} \geq \chi(G)\}$.

We know that there exists a $CA(n, K_{\chi(G)}, 2)$ if and only if there exists a graph homomorphism $K_{\chi(G)} \longrightarrow QI(n, 2)$. And if there is a graph homomorphism

$$K_{\chi(G)} \longrightarrow QI(n, 2)$$

then,

$$\begin{aligned} \omega(K_{\chi(G)}) \leq \omega(QI(n,2)) &= \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} \\ or,\ \omega(K_{\chi(G)}) &\leq \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} \\ or,\ \chi(G) &\leq \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} \end{aligned}$$

which verifies Corollary 4.3.18.

Now, we know that

$$CAN(K_{\chi(G)}, 2) = min\{l\ :\ \binom{l-1}{\lfloor \frac{l}{2} \rfloor - 1} \geq \chi(G)\}. \tag{4.3.21}$$

Putting $G = QI(n,\ 2)$ in Equation 4.3.21, we get

$$CAN(K_{\chi(QI(n,2))}, 2) = min\{l\ :\ \binom{l-1}{\lfloor \frac{l}{2} \rfloor - 1} \geq \chi(QI(n,2))\}.$$

Because $\chi(QI(n,2)) = \lceil \frac{1}{2}\binom{n}{\lfloor n/2 \rfloor} \rceil$, we conclude that for **n odd**,

$$CAN(K_{\chi(QI(n,2))}, 2) = min\{l\ :\ \binom{l-1}{\lfloor \frac{l}{2} \rfloor - 1} \geq \lceil \frac{1}{2}\binom{n}{\frac{n-1}{2}} \rceil\}.$$

Since $l = n+1$ satisfies the equation

$$\binom{n}{\lfloor \frac{n+1}{2} \rfloor - 1} = \binom{n}{\frac{n-1}{2}} > \lceil \frac{1}{2}\binom{n}{\frac{n-1}{2}} \rceil$$

and $l = n$ does not, we conclude that

$$CAN(K_{\chi(QI(n,\ 2))},\ 2) = n+1. \tag{4.3.22}$$

From Equations 4.3.20 and 4.3.22, we conclude that for n odd,

$$CAN(QI(n,\ 2),\ 2) < CAN(K_{\chi(QI(n,\ 2))}, 2). \tag{4.3.23}$$

Corollary 4.3.19. *[7] For any graph G,*

$$CAN(K_{\chi(G)}, 2) - 1 \leq CAN(G, 2) \leq CAN(K_{\chi(G)}, 2).$$

Moreover, if $CAN(K_{\chi(G)}, 2)$ is odd then

$$CAN(G, 2) = CAN(K_{\chi(G)}, 2).$$

Proof. Since there exists a homomorphism $G \longrightarrow K_{\chi(G)}$, we have

$$CAN(G, 2) \leq CAN(K_{\chi(G)}, 2).$$

Let $CAN(K_{\chi(G)}, 2)$ be odd and equal to $(m + 1)$. So, m is even. Assume that $CAN(G, 2) \leq m$ and hence a $CA(m, G, 2)$ exists which means that there exists a homomorphism $G \longrightarrow QI(m, 2)$. For m even, we know that

$$\chi(QI(m,2)) = \lceil \frac{1}{2}\binom{m}{\lfloor \frac{m}{2} \rfloor} \rceil = \binom{m-1}{\lfloor \frac{m}{2} \rfloor - 1} = \binom{m-1}{\frac{m}{2} - 1} = \binom{m-1}{m/2}$$

and thus $\chi(G) \leq \chi(QI(m, 2)) = \binom{m-1}{\lfloor \frac{m}{2} \rfloor - 1}$. We also know that

$$CAN(K_{\chi(G)}, 2) = min\{n \; : \; \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} \geq \chi(G)\}.$$

Since $\chi(G) \leq \binom{m-1}{\lfloor \frac{m}{2} \rfloor - 1}$, the minimum occurs when $n \leq m$. This means that

$$CAN(K_{\chi(G)}, 2) \leq m < m + 1 = CAN(K_{\chi(G)}, 2)$$

which is a contradiction and hence $CAN(G, 2) = CAN(K_{\chi(G)}, 2)$.

Next, we assume that $m = CAN(K_{\chi(G)}, 2)$ is even and that $CAN(G, 2) \leq m - 2$. Then there exists a $CA(m-2, G, 2)$ and hence a homomorphism $G \longrightarrow QI(m-2, 2)$. Thus,

$$\chi(G) \leq \chi(QI(m-2, 2)) = \binom{m-3}{\frac{m-2}{2}}.$$

Since

$$CAN(K_{\chi(G)}, 2) = min\{n \ : \ \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} \geq \chi(G)\},$$

we have $CAN(K_{\chi(G)}, 2) \leq m - 1$. This is because of the following.

$$\begin{aligned} &\binom{m-1-1}{\lfloor \frac{m-1}{2} \rfloor - 1} \geq \binom{m-3}{\frac{m-2}{2}} \\ \Leftrightarrow &\binom{m-2}{\frac{m-2}{2} - 1} \geq \binom{m-3}{\frac{m-2}{2}} \\ \Leftrightarrow &\binom{m-2}{\frac{m}{2} - 2} \geq \binom{m-3}{\frac{m}{2} - 1} \\ \Leftrightarrow &\, m \geq 4 \end{aligned}$$

which is always true as m is even and $0 < CA(G, 2) \leq m - 2$. Having proved this, we get $CAN(K_{\chi(G)}, 2) \leq m - 1 < m = CAN(K_{\chi(G)}, 2)$ which is a contradiction and hence $CAN(K_{\chi(G)}, 2) - 1 \leq CAN(G, 2) \leq CAN(K_{\chi(G)}, 2)$. □

4.4 Larger alphabet size

When $g > 2$, the problem of finding standard covering array and covering array on graphs becomes much difficult. We see a simple bound on both the chromatic number and clique number of the graphs $QI(n,\ g)$ and calculate an upper bound for the chromatic number of the graphs $QI(g^2,\ g)$.

Corollary 4.4.1. *For any g and $n_1 \leq n_2$,*

$$\omega(QI(n_1, g)) \leq \omega(QI(n_2, g)) \textit{ and } \chi(QI(n_1, g)) \leq \chi(QI(n_2, g)).$$

Proof. Let $n_1 < n_2$ which means that $(n_2 - n_1) > 0$. Now, let there be a function $f : QI(n_1, 2) \longrightarrow QI(n_2, 2)$ such that any vector $u \in QI(n_1, 2)$ is mapped to a vector

$v \in QI(n_2, 2)$ having its first n_1 entries identical to that of u and its rest $(n_2 - n_1)$ entries equal to zero. Let this function maps two adjacent vertices u_1 and u_2 in $QI(n_1, 2)$ to vertices v_1 and v_2 in $QI(n_2, 2)$ respectively. The first n_1 entries of vectors v_1 and v_2 together cover all the 2-tuples $(a, b) \in \mathbb{Z}_2^2$ between them as they are a pair of adjacent vertices themselves. Thus, all the adjacent vertices are mapped to adjacent vertices. Hence, f is a homomorphism from $QI(n_1, 2)$ to $QI(n_2, 2)$ wherefrom we conclude that $\omega(QI(n_1, g)) \leq \omega(QI(n_2, g))$ and $\chi(QI(n_1, g)) \leq \chi(QI(n_2, g))$. □

Theorem 4.4.2. *[7] For any integer g, $\chi(QI(g^2, g)) \leq \binom{g+1}{2}$.*

Proof. Lets pick any $(g+1)$ positions in the vectors, say the last $(g+1)$. Since there are g letters in our alphabet, for each vector at least one letter will occur twice in these chosen $(g+1)$ places. As there are only g^2 letters/entries in each vector, any pair of vectors that have letters repeated in the same positions have one pair from $\mathbb{Z}_g^2$ repeated and hence will not be qualitatively independent i.e. for any two vectors u and v, if there exist at least two i's such that $(u_i, v_i) = (a, b)$, where $a, b \in \{0, 1, ..., g-1\}$, then u and v are not qualitatively independent. Now, let i and j be any two positions out of the chosen $(g+1)$ positions. Let all the vectors be assigned the same color which have the same letter/entry in these two positions i and j. As seen earlier, the vectors in any such color class are not qualitatively independent and hence no adjacent vectors have been assigned the same color. This gives a proper coloring to $QI(g^2, g)$ and because the number of color classes in this case is $\binom{g+1}{2}$, we conclude that $\chi(QI(g^2, g)) \leq \binom{g+1}{2}$. □

4.5 Conclusion

In this chapter, we extend the definition of a covering array to include a graph structure. This action associates a graph structure to the usual covering arrays. If there is a homomorphism from G to H, then we can obtain bounds on $CAN(G,k)$ *i.e.* $(CAN(G,k) \leq CAN(H,k))$.

We studied qualitatively independence graph and observed that

$$\exists\ CA(n,G,g)\ \Leftrightarrow\ \exists\ a\ homomorphism\ f\ :\ G \longrightarrow QI(n,g)$$

which proves that the problem of finding covering arrays on graphs is equivalent to determining homomorphisms to the family of graphs $QI(n,g)$. The chromatic number and core of the graph $QI(n,2)$ are established here. We observed that from a covering array $CA(n,G,2)$, it is always possible to construct a new covering array $CA(n,G,2)$ in which the rows are vectors with weight $\lfloor n/2 \rfloor$. Moreover, if n is even, it is possible to find such a covering array with the rows all beginning with 0. Finally, we presented that the chromatic number and clique number of graphs $QI(n,g)$ and their cores increase monotonically with increase in n. An upper bound of the chromatic number of $QI(n,g)$ for $n = g^2$ is also established.

Bibliography

[1] P. Erdős, C. Ko, and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math. Oxford Ser **12(2)** (1961), 313–320.

[2] Chris Godsil and Gordon Royle, *Algebraic graph theory*, Springer-Verlag, New York, 2001.

[3] John Michael Harris, Jeffry L. Hirst, and Michael J. Mossinghoff, *Combinatorics and graph theory*, Springer, New York, 2008.

[4] G. Katona, *Two applications (for search theory and truth functions) of sperner type theorems*, Period. Math. Hungar **3** (1973), 19–26.

[5] D. J. Kleitman and J. Spencer, *Families of k-indepenent sets*, Discrete Math **6** (1973), 255–262.

[6] Karen Meagher, *Covering arrays on graphs: Qualitative independence graphs and extremal set partition theory*, Ph.D. thesis, Ottawa-Carleton Institute for

Graduate Studies and Research in Mathematics and Statistics, Department of Mathematics & Statistics, 2005.

[7] Karen Meagher and Brett Stevens, *Covering arrays on graphs*, Journal of Combinatorial Theory, Series B **95** (2005), 134–151.

[8] S. Poljak, A. Pultr, and V. Rödl, *On qualitatively independent partitions and related problems*, Discrete Appl. Math **6(2)** (1983), 193–205.

[9] E. Sperner, *Ein satz über untermengen einer endlichen menge*, Math. Z. **27** (1928), 544–548.

[10] B. Stevens and E. Mendelsohn, *New recursive methods for transversal covers*, J. Combin. Des. **7(3)** (1999), 185–203.