THE UNIVERSITY OF CHICAGO

COHERENT CONFIGURATIONS: MOTION, SPECTRAL PROPERTIES, ROBUSTNESS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

BY
BOHDAN KIVVA

CHICAGO, ILLINOIS
JUNE 4, 2022

To Mom and Dad

and to those protecting my family, my home, my country, while I am writing this dissertation

Epigraph Text

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am indebted to László Babai for being my PhD advisor, for his guidance and his unlimited support I had during my studies. I am grateful to him for teaching me how to do and write mathematics properly during numerous long meetings. His true passion for mathematics, computer science, research has been very inspiring and it is a tremendous luck to have a chance to learn from him.

I am grateful to my co-advisor, Aaron Potechin, for introducing me to the world of Sum-of-Squares, for the projects we were working on, and for many insightful comments. I am grateful for his careful guidance and constant support.

I am grateful to Bryon Aragam for supervising me in Machine Learning and Causality projects, for his support, and for many valuable insights. I am also grateful for teaching me how to write papers in Machine Learning area.

Additionally, I would like to thank all those who taught me mathematics prior to joining the University of Chicago: Ievgen Bondarenko, Vitaliy Polonskiy, Ivan Feshchenko, Andriy Arman, Vitaliy Lishunov, Bohdan Rublyov, my parents, and many-many others.

I also want to thank all those who created welcoming and comfortable environment to study, to do research, and to be productive: professors, fellow students and friends, and the department staff. The list of those people is very long, including Madhur Tulsiani, Alexander Razborov, Robert Fefferman, Jitka Stehnova, John Boller, Greg Shakhnarovich, Wenjun Cai, Shuo Pang, Tim Black, Goutham Rajendran, Leo Coregliano, Chris Jones, Nat Mayer, Jon DeWitt, Liam Mazurowski, Dylan Quintana, Bingjin Liu, Boming Jia, Peter Morfe, and many others.

Finally, I am deeply grateful to those who constantly supported me – my parents, grandparents, my brother, my girlfriend, my friends and my advisers.

Without all these people this journey would not be possible.

ABSTRACT

The principal theme of this thesis is the interplay between symmetry and regularity in discrete structures. The most general class of structures we consider are *coherent configurations,* certain highly regular colorings of complete graphs. This class includes such diverse structures as the orbital configurations of permutation groups and association schemes originating from the design of experiments in statistics. Metric schemes, a subclass of association schemes, are derived from distance-regular graphs. Johnson, Hamming, and Grassman schemes are special classes of great importance among metric schemes. We study structural and spectral properties of coherent configurations with special attention to the subclasses mentioned. As a culmination of this analysis, we confirm Babai's conjecture on the minimal degree of the automorphism group for distance-regular graphs of bounded diameter and for primitive coherent configurations of rank 4.

The minimal degree of a permutation group G is the minimum number of points not fixed by non-identity elements of G. Lower bounds on the minimal degree have strong structural consequences on G. Babai conjectured that for some constant $c > 0$ the automorphism group of a primitive coherent configuration on n vertices has minimal degree $\geq cn$ with known exceptions[1]. If confirmed, this conjecture gives a CFSG[2]-free proof of the Liebeck-Saxl classification of primitive groups with sublinear minimal degree. Moreover, if confirmed, this conjecture would point to potential simplification of some steps in Babai's quasipolynomial-time algorithm for the Graph Isomorphism problem.

In this thesis we confirm Babai's conjecture for distance-regular graphs (metric schemes) of bounded diameter and for primitive coherent configurations of rank 4.

Central to our approach is the study of spectral parameters of distance-regular graphs,

1. Recent work by Sean Eberhard expanded the class of known exceptions, but (i) it does not affect the implication in the next sentence about CFSG-free proof of the Liebeck-Saxl classification; (ii) the conjecture for distance-regular graphs is not affected.

2. Classification of Finite Simple Groups

such as spectral gap and smallest eigenvalue.

The spectral gap of a graph is known to be tightly related to expansion properties of the graph. Hence, lower bounds on the spectral gap are widely applicable in various areas of mathematics and theoretical computer science. In this thesis we prove that a distance-regular graph with a dominant distance is a spectral expander. Our lower bound on the spectral gap depends only on the diameter of the graph. The key ingredient of the proof is a new inequality on the intersection numbers.

At the same time, graphs of which the smallest eigenvalue has small absolute value are known to enjoy a rich geometric structure (see, e.g., celebrated results of Hoffman, Seidel, Neumaier, and Cameron et al.).

In this thesis we characterize Hamming graphs as distance-regular graphs of diameter d with smallest eigenvalue $-d$ and[3] $\mu \leq 3$, under mild additional assumptions.

We also characterize Johnson and Hamming graphs as geometric distance-regular graphs satisfying certain inequality constraints on the spectral gap and the smallest eigenvalue. Classical characterizations of Hamming graphs $H(d, q)$ assume equality constraints on certain parameters such as the assumption $\theta_1 = b_1 - 1$ on the second largest eigenvalue or the assumption $n = (\lambda + 2)^d$ on the number of vertices (see, e.g., results of Enomoto and Egawa). The principal novelty of our result is that we make no such tight assumptions.

Finally, in this thesis we study robustness properties of certain classes of coherent configurations. For instance, we show that the family of Johnson schemes is robust in the following sense. If a homogeneous coherent configuration $\mathfrak{X}$ on n vertices or its fission contains a Johnson scheme $J(s, d)$ as a subconfiguration on at least $5n/6$ vertices and $s > 250d^4$, then $\mathfrak{X}$ itself is a Johnson scheme. This result strengthen a 1972 theorem of Kaluzhnin and Klin that corresponds to the case where the subconfiguration itself has n vertices.

Our result is also related to Babai's "Split-or-Johnson lemma" and in particular to the

3. Here μ denotes the number of common neighbours of (every) pair of vertices at distance 2.

philosophy in the theory of Graph Isomorphism testing that we can either find structure or find efficiently verifiable asymmetry. The result represents a step in the direction of simplifying the conclusion of the "Split-or-Johnson" lemma.

We also show that similar robustness results hold for Hamming and Grassmann schemes.

CHAPTER 1

INTRODUCTION

1.1 Symmetry vs. Regularity

A central theme of this thesis is the interplay between symmetry and regularity in combinatorial structures, a subject that has been studied for several decades. The "Symmetry vs. Regularity" framework builds bridges between Group Theory and Combinatorics. Additionally, the framework is related to multiple developments in Theoretical Computer Science, including Babai's quasipolynomial-time Graph Isomorphism test (Babai [2016a,b]) and the study of the complexity of the matrix multiplication (Cohn and Umans [2003, 2013]). Families of coherent configurations which naturally arise in the "Symmetry vs. Regularity" framework, such as the Johnson schemes or the Hamming schemes, due to their nice properties, also arise in numerous other contexts. For instance, Meka et al. [2015] used the eigenspaces of the Johnson schemes in the context of the planted clique problem and the "Sum-of-Squares" hierarchy. Recent progress on the Unique Games conjecture is closely related to the study of the expansion properties of Johnson and Grassmann schemes (Khot et al. [2018], Bafna et al. [2020], Hopkins et al. [2020], Dinur et al. [2021]).

In the "Symmetry vs. Regularity" framework one aims to transition from studying symmetry conditions, such as distance-transitivity, to regularity conditions, such as distance-regularity. This transition is desirable as symmetry is a global, hard-to-detect property of an object, while regularity is local and is usually easy to test. In the opposite direction, one may hope to apply Group Theory to algorithmic and combinatorial problems. For instance, the central piece of Babai's Graph Isomorphism test is a group-theoretic "Unaffected Stabilizer Theorem" which relies on the Classification of Finite Simple Groups (CFSG) through Schreier's Hypothesis.

The vehicle for this transition is *Coherent Configurations* (CCs) which are highly regular

colorings of the edges of the complete directed graphs. They were first introduced by I. Schur [1933] who used them to study permutation groups through their orbital configurations. Later, Bose and Shimamoto [1952] studied a special class of coherent configurations, called *association schemes*, in connection with combinatorial designs. Coherent configurations in their full generality were independently introduced by Weisfeiler and Leman [1968] (see Weisfeiler [1976]), and D. Higman [1967, 1970]. Higman developed the representation theory of coherent configurations and applied it to permutation groups. At the same time, a related algebraic theory of coherent configurations, called "cellular algebras," was introduced by Weisfeiler and Leman, motivated by the algorithmic problems of Graph Isomorphism and Graph Canonization. Special classes of association schemes such as strongly regular graphs and, more generally, distance-regular graphs have been the subject of intensive study in algebraic combinatorics.

A combinatorial study of coherent configurations was initiated by Babai [1981]. Coherent configurations play an important role in the study of the Graph Isomorphism problem, adding combinatorial divide-and-conquer tools to the arsenal. This approach was used by Babai [2016a,b]. Also, recently, the representation theory of coherent configurations found applications to the complexity of matrix multiplication in the work of Cohn and Umans [2013].

Let Ω be a finite set. A permutation group $G \leq \mathrm{Sym}(\Omega)$ defines an equivalence relation on $\Omega \times \Omega$ by $(x, y) \sim (gx, gy)$ for $x, y \in \Omega$ and $g \in G$. This relation can be viewed as a coloring $\mathfrak{c}$ of the pairs $(x, y) \in \Omega$ in which two pairs have the same color if and only if they belong to the same orbit of the induced action of G on $\Omega \times \Omega$. It is not hard to see that $\mathfrak{c}$ has several simple combinatorial properties; these have been abstracted by Schur to define a purely combinatorial object.

Definition 1.1.1. Let Ω be a finite set. A pair $\mathfrak{X} = (\Omega, \mathfrak{c})$ is called a *coherent configuration (CC)* if the coloring $\mathfrak{c} : \Omega \times \Omega \to \{\text{colors}\}$ has the following properties.

(i) $\mathfrak{c}(x, y) \neq \mathfrak{c}(z, z)$ for all $x, y, z \in \Omega$ with $x \neq y$ ("edge-colors"$\neq$"vertex-colors").

(ii) The color of the pair (x, y) uniquely defines the color of (y, x), for all $(x, y) \in \Omega \times \Omega$.

(iii) for all colors i, j, t there is an *intersection number* $p^t_{i,j}$ such that, for all $u, v \in \Omega$, if $\mathfrak{c}(u, v) = t$, then there exist exactly $p^t_{i,j}$ vertices $w \in \Omega$ with $\mathfrak{c}(u, w) = i$ and $\mathfrak{c}(w, v) = j$.

The *rank* of a CC is the number of (non-empty) color classes defining it.

The coherent configurations defined by the group action of $G \leq \mathrm{Sym}(\Omega)$ on $\Omega \times \Omega$, as described above, are called *Schurian configurations*. We note that not all coherent configurations are Schurian, i.e., a coloring $\mathfrak{c}$ satisfying (i)-(iii) may not have any group action defining it.

A coherent configuration $\mathfrak{X} = (\Omega, \mathfrak{c})$ is called *homogeneous* if $\mathfrak{c}(x) = \mathfrak{c}(y)$ for all $x, y \in \Omega$, and it is called an *association scheme* if $\mathfrak{c}(x, y) = \mathfrak{c}(y, x)$ for all $x, y \in \Omega$. A coherent configuration is called *primitive* if the digraph defined by every edge color is weakly connected.

We will be especially interested in a special well-studied case of coherent configurations, $(\Omega, \mathfrak{c})$, in which $\mathfrak{c}(x, y) = i$ if x and y are at distance i in the graph defined by edges of color 1. Such coherent configurations are called *metric schemes* and the corresponding color-1 graph is called a *distance-regular graph* (DRG).

We say that a coherent configuration of rank 2 is *trivial*.

1.2 Babai's conjectures on primitive coherent configurations

1.2.1 Cameron's classification of primitive permutation groups

Many questions on permutation groups reduce to the case of primitive permutation groups.

Definition 1.2.1. A permutation group $G \leq \mathrm{Sym}(\Omega)$ is called *transitive* if for all $x, y \in \Omega$ there exists an element $g \in G$ that maps x to y.

Definition 1.2.2. A *primitive* permutation group is a non-trivial transitive permutation

group whose only invariant partitions are trivial (the entire set, and the partition into singletons).

Relying on the Classification of Finite Simple Groups (CFSG), Cameron [1981] classified all primitive permutation groups whose order is at least $n^{c \log n}$ for some $c > 0$ (see Chapter 4). He showed that such groups G act on $\binom{[k]}{t}^{\ell}$ for some t, k, ℓ and satisfy $\left(A_k^{(t)}\right)^{\ell} \leq G \leq S_k^{(t)} \wr S_\ell$ (with the product action). Here, $A_k^{(t)}$ and $S_k^{(t)}$ are the alternating group A_k and the symmetric group S_k acting on $\binom{[k]}{t}$. Such primitive groups G are called *Cameron groups*.

In the wake of Cameron's classification, Babai initiated several projects with the aim of finding combinatorial relaxations of Cameron's results. Babai conjectured several such relaxations in terms of key parameters of permutation groups: order, minimal degree, thickness.

1.2.2 Minimal degree of a permutation group. Liebeck-Saxl's classification

One of the key contributions of this thesis confirms Babai's conjecture on the minimal degree for metric schemes of bounded rank (corresponding to distance-regular graphs of bounded diameter) and for coherent configurations of rank 4. (Babai settled the rank-3 case which corresponds to strongly regular graphs.)

Let σ be a permutation of a set Ω. The number of points not fixed by σ is called the *degree* of the permutation σ. Let G be a permutation group on the set Ω. The minimum of the degrees of non-identity elements in G is called the *minimal degree*[1] of G and is denoted by $\mathrm{mindeg}(G)$. One of the classical problems in the theory of permutation groups is to classify the primitive permutation groups whose minimal degree is small (see Wielandt [1964]). The study of minimal degree goes back to works of Jordan [1871] and Bochert [1892] in 19th century. In particular, Bochert [1892] proved that a doubly transitive permutation group of

1. For the identity permutation group on the set Ω, we define its minimal degree to be ∞, i.e., the minimum of the empty set.

degree n has minimal degree $\geq n/4 - 1$ with trivial exceptions.

Lower bounds on the minimal degree of a group imply strong constraints on the structure of the group. A result of Wielandt [1934] shows that a linear (in $|\Omega|$) lower bound on $\mathrm{mindeg}(G)$ implies a logarithmic upper bound on the degree of every alternating group involved in G as a quotient of a subgroup (see Theorem 4.3.1).

Similarly to Cameron's classification of large primitive permutation groups, using CFSG, Liebeck [1984], Liebeck and Saxl [1991] characterized primitive permutation groups of degree n with minimal degree $< n/3$ (see Theorem 4.2.2). In fact, they showed that those are Cameron groups.

1.2.3 Babai's combinatorial relaxations of Liebeck-Saxl's and Cameron's classifications

We define *Cameron schemes* as Schurian configurations obtained from Cameron groups. Below we discuss the combinatorial relaxation of the Liebeck-Saxl classification conjectured by Babai.

Definition 1.2.3. Following Russell and Sundaram [1998], for a combinatorial structure $\mathcal{X}$ we use term *motion* to refer to the minimal degree of the automorphism group $\mathrm{Aut}(\mathcal{X})$:

$$\mathrm{motion}(\mathcal{X}) = \mathrm{mindeg}(\mathrm{Aut}(\mathcal{X})). \tag{1.1}$$

For distance-regular graphs Babai conjectured the following relaxation of the Liebeck-Saxl classification.

Conjecture 1.2.4 (Babai)**.** *There exists $\gamma > 0$ such that for every primitive distance-regular graph X of diameter d on n vertices either*

$$\mathrm{motion}(X) \geq \gamma n,$$

or X is a Johnson graph, or a Hamming graph, or their complement.

Babai confirmed this conjecture for distance-regular graphs of diameter ≤ 2 (i.e., for connected strongly regular graphs).

Theorem 1.2.5 (Babai [2014, 2015])**.** *For every primitive distance-regular graph X of diameter 2 on $n \geq 29$ vertices either*

$$\text{motion}(X) \geq n/8,$$

or X, or its complement, is a Johnson graph $J(s, 2)$ or a Hamming graph $H(2, s)$.

In this thesis we confirm this conjecture for distance-regular graphs of bounded diameter.

Theorem 1.2.6 (**Main I**)**.** *For every $d \geq 3$ there exists $\gamma_d > 0$, such that for every primitive distance-regular graph X of diameter d on n vertices either*

$$\text{motion}(X) \geq \gamma_d n,$$

or X is a Johnson graph, or a Hamming graph.

We prove this theorem in Chapter 8. Additionally, we show that if the primitivity assumption is dropped then one more family of exceptions arises, the family of crown graphs (see Theorem 8.4.1).

In the general case, Babai made the following conjecture.

Conjecture 1.2.7 (Babai)**.** *There exists $\gamma > 0$ such that for every primitive coherent configuration $\mathfrak{X}$ on n vertices either*

$$\text{motion}(\mathfrak{X}) \geq \gamma n,$$

or $\mathfrak{X}$ is a Cameron scheme.

For primitive coherent configurations of rank 3 this conjecture follows from Theorem 1.2.5 and Babai [1981]. In this thesis we confirm this conjecture for rank-4 primitive coherent configurations. However, as we discuss below, recently Eberhard [2022] found a counterexample of rank 28 and suggested a slightly modified version of Conjecture 1.2.7 (see Conj. 1.2.13).

Theorem 1.2.8 (**Main II**). *There exists an absolute constant $\gamma_4 > 0$ such that for every primitive coherent configuration $\mathfrak{X}$ of rank 4 on n vertices either*

$$\mathrm{motion}(\mathfrak{X}) \geq \gamma_4 n,$$

or $\mathfrak{X}$ is a Johnson scheme, or a Hamming scheme.

This theorem is proved in Chapter 9 (see Theorem 9.5.1).

A version of Conjecture 1.2.7 in terms of the order of a group says that Cameron schemes are the only primitive coherent configurations with more than quasipolynomial number of automorphisms. A slightly weaker version has the following form.

Conjecture 1.2.9 (Babai). *Let $\varepsilon > 0$. Primitive coherent configurations, other than Cameron schemes, have at most $\exp(O(n^{\varepsilon}))$ automorphisms.*

The first step towards this conjecture was made by Babai [1981]. He proved that a non-trivial primitive coherent configuration on n vertices has at most $\exp(O(n^{1/2}\log^2 n))$ automorphisms. As a byproduct, he solved a then 100-year-old problem on primitive, but not doubly transitive groups, giving a nearly tight bound on their order. After more than 30 years, Sun and Wilmes [2015a,b] made the second step, proving that the only non-trivial primitive coherent configurations on n vertices that have more than $\exp(O(n^{1/3}\log^{7/3} n))$ automorphisms are Johnson and Hamming schemes.

1.2.4 Eberhard's version of Babai's conjectures

In a recent surprising result, Eberhard showed that in fact Conjectures 1.2.9 and 1.2.7 do not hold as stated. His result does not affect Conjecture 1.2.4, Conjecture 1.2.7 for configurations of rank at most 7, and Conjecture 1.2.9 for $\varepsilon > 1/8$.

Theorem 1.2.10 (Eberhard [2022])**.** *For each $m \geq 3$, there is a non-schurian primitive association scheme $\mathfrak{X}$ of rank 28 on $n = m^8$ vertices, such that $\mathrm{Aut}(\mathfrak{X})$ is imprimitive and $|\mathrm{Aut}(\mathfrak{X})| \geq \exp(n^{1/8})$.*

However, Eberhard [2022] proposed a variant of Conjectures 1.2.9 and 1.2.7 that may still hold.

Definition 1.2.11. We say that a configuration $\mathfrak{Y} = (\Omega, \mathfrak{c}_{\mathfrak{Y}})$ is a *fusion* of a configuration $\mathfrak{X} = (\Omega, \mathfrak{c}_{\mathfrak{X}})$ if there is a map $\eta : \mathrm{Range}(\mathfrak{c}_{\mathfrak{X}}) \to \mathrm{Range}(\mathfrak{c}_{\mathfrak{Y}})$ such that $\mathfrak{c}_{\mathfrak{Y}}(u, v) = \eta(\mathfrak{c}_{\mathfrak{X}}(u, v))$ for all $u, v \in \Omega$. In this case, $\mathfrak{X}$ is called a *fission* of $\mathfrak{Y}$.

For configurations $\mathfrak{X}$ and $\mathfrak{X}'$ on Ω, define a partial order by writing $\mathfrak{X} \preceq \mathfrak{X}'$ if $\mathfrak{X}$ is a fission of $\mathfrak{X}'$.

Definition 1.2.12. A primitive coherent configuration $\mathfrak{Y}$ defined on $\binom{[m]}{k}^d$ is called a *Cameron sandwich* if

$$\mathfrak{X}\left(\left(A_m^{(k)}\right)^d\right) \preceq \mathfrak{Y} \preceq \mathfrak{X}\left(S_m^{(k)} \wr S_d\right).$$

Conjecture 1.2.13 (Eberhard's version of Babai's conjecture)**.** *There exist $c, \gamma > 0$, such that for every primitive coherent configuration $\mathfrak{X}$ on n vertices either*

$$|\mathrm{Aut}(\mathfrak{X})| \leq \exp(\log^c n) \quad \textit{and} \quad \mathrm{motion}(\mathfrak{X}) \geq \gamma n, \tag{1.2}$$

or $\mathfrak{X}$ is a Cameron sandwich.

Remark 1.2.14. If confirmed, Conjecture 1.2.13 would still provide a CFSG-free proof of the Cameron classification and the Liebeck-Saxl classification. Additionally, if confirmed, it

would point to potential simplification of Babai's quasipolynomial Graph Isomorphism test as mentioned in [Babai, 2016b, Remark 6.1.3].

1.3 Robustness of coherent configurations

1.3.1 Individualization and refinement

In algorithmic applications, the interplay between symmetry and regularity frequently arises in the context of *individualization/refinement* technique. This is a standard and widely used practical technique for solving tasks related to symmetry computations of graphs and other combinatorial objects, which include computing automorphism groups, isomorphism tests, canonical labeling tools. In particular, individualization/refinement is central to Babai's Graph Isomorphism test (Babai [2016a,b]).

In this technique, one breaks the symmetry of, say, a graph by assigning unique colors to a small subset of its vertices (*individualization*). After that, one propagates the asymmetry, created by individualizing these vertices, using a refinement step.

A classical example of a refinement was introduced by Weisfeiler and Leman [1968]. The *Weisfeiler-Leman refinement* proceeds in rounds. In each round it takes a configuration $\mathfrak{X} = (\Omega, \mathfrak{c})$ of rank r and for each pair $(x, y) \in \Omega \times \Omega$ it encodes in a new color $\mathfrak{c}'(x, y)$ the following information: the color $\mathfrak{c}(x, y)$, and for every $i, j \leq r$ the number of vertices z with $\mathfrak{c}(x, z) = i, \mathfrak{c}(z, y) = j$. It is easy to see that for the refined coloring $\mathfrak{c}'$, the structure $\mathfrak{X}' = (\Omega, \mathfrak{c}')$ is a configuration as well. The refinement process applied to a configuration $\mathfrak{X}$ takes $\mathfrak{X}$ as an input on the first round, and on every subsequent round in takes as an input the output of the previous round. The refinement process stops when it reaches a stable configuration (i.e, $\mathfrak{Y}' = \mathfrak{Y}$). It is easy to see that the process will always stop. Moreover, one can check that the configurations that are stable under this refinement process are precisely the coherent configurations. Therefore, the Weisfeiler-Leman refinement process

takes any configuration and refines it to a coherent configuration.

Clearly, the result of a (non-trivial) individualization and the Weisfeiler-Leman refinement is a (non-homogeneous) fission of the original configuration.

Importantly, the Weisfeiler-Leman refinement is *canonical* in the following sense. Let $\mathfrak{X}$, $\mathfrak{Y}$ be configurations and let $\mathfrak{X}^*$, $\mathfrak{Y}^*$ be the corresponding outputs of the Weisfeiler-Leman refinement simultaneously applied to $\mathfrak{X}$ and $\mathfrak{Y}$. Then the sets of isomorphisms for $\mathfrak{X}$, $\mathfrak{Y}$ and for $\mathfrak{X}^*$, $\mathfrak{Y}^*$ are the same

$$\mathrm{Iso}(\mathfrak{X}, \mathfrak{Y}) = \mathrm{Iso}(\mathfrak{X}^*, \mathfrak{Y}^*) \tag{1.3}$$

1.3.2 Babai's "Split-or-Johnson" Lemma. Robustness of Johnson schemes

The key combinatorial partitioning tool of the Graph Isomorphism algorithm of Babai [2016a,b], the "Split-or-Johnson" lemma, states that one can either find a specific structure or significantly break the symmetry of a coherent configuration after individualizing a logarithmic number of points and applying the Weisfeiler-Leman refinement.

Theorem 1.3.1 (Babai [2016b], "Split-or-Johnson"). *Let $\mathfrak{X} = (\Omega, \mathfrak{c})$ be a primitive coherent configuration of rank ≥ 3 on n vertices and let $2/3 \leq \gamma < 1$ be a threshold parameter. Then by individualizing $O(\log n)$ vertices of $\mathfrak{X}$ and by applying the Weisfeiler-Leman refinement process one can get a coherent configuration $\mathfrak{Y} = (\Omega, \mathfrak{c}_{\mathfrak{Y}})$ that satisfies one of the following.*

1. *No color is assigned by $\mathfrak{c}_{\mathfrak{Y}}$ to $\geq \gamma|\Omega|$ vertices.*
2. *$\mathfrak{c}_{\mathfrak{Y}}$ induces a non-trivial equipartition of the vertex color class of size $\geq \gamma|\Omega|$.*
3. *$\mathfrak{Y}$ contains a homogeneous fission of a Johnson scheme on $\geq \gamma|\Omega|$ vertices as a subconfiguration.*

Babai conjectured that for a sufficiently large γ in the latter case $\mathfrak{X}$ is either a Johnson scheme itself, or $\mathfrak{X}$ has a quasipolynomial number of automorphisms. In this thesis we make

a step towards confirming this conjecture. This is also a step in the direction of simplifying the conclusion of the "Split-or-Johnson" lemma.

Theorem 1.3.2 (**Main III**, Babai and Kivva [2022])**.** *Let $\mathfrak{Y}'$ be a homogeneous coherent configuration of rank ≥ 3 on Ω'. Assume that $\mathfrak{Y}'$ is a fusion of a configuration $\mathfrak{X}'$. Let $\Omega \subseteq \Omega'$, with $n' \leq (6/5)n$. Suppose that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is the Johnson scheme $\mathfrak{J}(s,d)$ with $s \geq 250d^4$. Then $\mathfrak{Y}'$ is a Johnson scheme itself, of the same rank as $\mathfrak{X}$.*

We present the proof of this Theorem in Section 11.4.2.

1.3.3 Robustness of Hamming and Grassmann schemes

Theorem 1.3.2 can also be seen as an answer to a special case of the following question.

Question 1.3.3. *Let $\alpha \geq 0$ and $\Omega \subseteq \Omega'$ be finite sets, such that $|\Omega'| \leq (1+\alpha)|\Omega|$. Assume that $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ and $\mathfrak{X} = (\Omega, \mathfrak{c})$ are homogeneous coherent configurations. Suppose that $\mathfrak{X}$ is "nicely embedded" in $\mathfrak{X}'$ and, moreover, $\mathfrak{X}$ belongs to some class of configurations $\mathcal{A}$.*

For which α and $\mathcal{A}$ can we deduce that $\mathfrak{X}'$ also belongs to $\mathcal{A}$?

In Chapters 10 and 11 we study this question in the following interpretations of "nicely embedded" for various properties $\mathcal{A}$.

(A) $\mathfrak{X}$ is a subconfiguration of $\mathfrak{X}'$.

(B) $\mathfrak{X}$ is a subconfiguration of a fission of $\mathfrak{X}'$.

In particular, we show that analogs of Theorem 1.3.2 hold for Hamming and Grassmann schemes, another two families of schemes that are of interest to several areas of mathematics and theoretical computer science.

Theorem 1.3.4. *Let $\mathfrak{Y}'$ be a homogeneous coherent configuration of rank ≥ 3 on Ω'. Assume that $\mathfrak{Y}'$ is a fusion of a configuration $\mathfrak{X}'$. Let $\Omega \subseteq \Omega'$, with $|\Omega'| \leq (6/5)|\Omega|$. Suppose that*

$\mathfrak{X} = \mathfrak{X}'[\Omega]$ *is the Hamming scheme* $\mathfrak{H}(d,s)$ *with* $s \geq 200d^4 \ln(d)$. *Then* $\mathfrak{Y}'$ *is a Hamming scheme, of the same rank as* $\mathfrak{X}$.

Theorem 1.3.5. *Let* $\mathfrak{Y}'$ *be a homogeneous coherent configuration of rank* ≥ 4 *on* Ω'. *Assume that* $\mathfrak{Y}'$ *is a fusion of a configuration* $\mathfrak{X}'$. *Let* $\Omega \subseteq \Omega'$, *with* $|\Omega'| \leq (5/4)|\Omega|$. *Suppose that* $\mathfrak{X} = \mathfrak{X}'[\Omega]$ *is the Grassmann scheme* $\mathfrak{J}_q(s,d)$ *with* $s \geq 6d+5$. *Then* $\mathfrak{Y}'$ *is a Grassmann scheme, of the same rank as* $\mathfrak{X}$, *and for the same prime power* q.

For Question 1.3.3 in interpretation (A) we prove the following.

Theorem 1.3.6. *Let* $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ *be a homogeneous coherent configuration. Let* $\Omega \subseteq \Omega'$ *with* $|\Omega'| < (3/2)|\Omega|$. *Assume that* $\mathfrak{X} = \mathfrak{X}'[\Omega]$ *is*

- *(Babai and Kivva [2022]) the Johnson scheme* $\mathfrak{J}(d,s)$ *with* $d \geq 2$, $s \geq 288d^2 + d$; *or*
- *the Hamming scheme* $\mathfrak{H}(d,s)$ *with* $d \geq 2$, $s \geq 200d^4 \ln d$; *or*
- *the Grassmann scheme* $\mathfrak{J}_q(s,d)$ *with* $d \geq 3$ *and* $s \geq 3d+7$.

Then $\mathfrak{X}'$ *is a Johnson scheme, or a Hamming scheme, or a Grassmann scheme, respectively.*

These three theorems are proved in Sections 11.4.3, 11.4.4, and 10.4-10.6.

1.3.4 Group theory view on Question 1.3.3: Galois correspondence

Question 1.3.3 has been studied in the following version of "nicely embedded".

(C) $\Omega = \Omega'$ and $\mathfrak{X}$ is a fission of $\mathfrak{X}'$.

For this interpretation of "nicely embedded", the question takes the following form.

Question 1.3.7. *Assume that* $\mathfrak{X}' = (\Omega, \mathfrak{c}')$ *and* $\mathfrak{X} = (\Omega, \mathfrak{c})$ *are homogeneous coherent configurations and* $\mathfrak{X}$ *is a fission of* $\mathfrak{X}'$. *Suppose that* $\mathfrak{X}$ *belongs to some class of configurations* $\mathcal{A}$. *For which* $\mathcal{A}$ *can we deduce that* $\mathfrak{X}'$ *also belongs to* $\mathcal{A}$?

For a finite permutation group $G \leq \mathrm{Sym}(\Omega)$ let $\mathfrak{X}(G)$ be the corresponding Schurian configuration. Note that several groups may define the same Schurian configuration $\mathfrak{X}(G)$. Such groups are called *2-equivalent*. The *2-closure* of the group G is defined as $\mathrm{Aut}(\mathfrak{X}(G))$, which is the maximal element of the 2-equivalence class of G. The group is called *2-closed* if it coincides with its 2-closure.

It is easy to see that if $G \leq G' \leq \mathrm{Sym}(\Omega)$, then $\mathfrak{X}(G)$ is a fission of $\mathfrak{X}(G')$. And vice versa, if $\mathfrak{X}$ is a fission of $\mathfrak{X}'$, then $\mathrm{Aut}(\mathfrak{X}) \leq \mathrm{Aut}(\mathfrak{X}')$. Recall, that for configurations $\mathfrak{X}$ and $\mathfrak{X}'$ on Ω, we define a partial order by writing $\mathfrak{X} \preceq \mathfrak{X}'$, if $\mathfrak{X}$ is a fission of $\mathfrak{X}'$. One can check that there is a *Galois correspondence* between the coherent configurations on Ω with the $\preceq$ relation and the 2-closed permutation groups on Ω with the subgroup relation.

In view of this Galois correspondence, results on the fission/fusion of coherent configurations (Question 1.3.3 in interpretation (C)) can be translated into results on the subgroups/supergroups of 2-closed permutation groups.

Recall that $S_t^{(d)} \leq \mathrm{Sym}\left(\binom{[t]}{d}\right)$ is the permutation group defined by the induced action of S_t on d-element subsets of $[t]$. Kaluzhnin and Klin [1972] showed that the Johnson group is a maximal 2-closed subgroup of the symmetric group $\mathrm{Sym}\left(\binom{[t]}{d}\right)$ when $t \geq c(d)$ for a sufficiently large $c(d)$. They proved this by showing that the corresponding Johnson scheme has no nontrivial fusion. In his PhD thesis, Klin [1974] showed that one can take $c(d) = O(d^4)$. Later, Muzychuk [1992a] improved bound to $c(d) = 3d + 4$ and Uchida [1992] made another slight improvement to $c(d) = 2d + \sqrt{(d - 7/2)^2 + 6} + 3/2$.

Our Theorem 1.3.2 generalizes Kaluzhnin-Klin's theorem.

Similarly, Muzychuk [1992b] proved that the Hamming scheme $\mathfrak{H}(d, s)$ with $s > 4$ does not admit a non-trivial fusion that is a coherent configuration, and he classified the fusion schemes for $s = 4$. The case of $s = 2$ was studied in Muzychuk [1995]. Our Theorem 1.3.4 is as a generalization of Muzychuk [1992b] for $s \geq 200d^4 \ln(d)$.

1.4 Spectral gap and classifications of distance-regular graphs

In order to prove Theorems 1.2.6, 1.2.8 and 1.3.6 which we discussed in Sections 1.2 and 1.3.3, we study spectral and combinatorial properties of distance-regular graphs and coherent configurations. Along the way, we prove several results for distance-regular graphs which fit into several other well-studied frameworks. In particular, we study the spectral gap of distance-regular graphs, the parameter that is closely related to the expansion properties of the graph, and which plays an important role in various applications in combinatorics and theoretical computer science. Additionally, we provide new characterizations of Johnson and Hamming graphs in terms of their smallest eigenvalue and spectral gap. These characterizations can be seen as a contribution to the program that aims to classify sufficiently regular graphs based on their smallest eigenvalue (see, e.g., Hoffman [1970b, 1977], Seidel [1968], Neumaier [1979], Cameron et al. [1991], Bang and Koolen [2014]).

1.4.1 Spectral gap of distance-regular graphs

We say that a k-regular graph is a *spectral η-expander* for $\eta > 0$, if every non-principal eigenvalue ξ_i of its adjacency matrix satisfies $|\xi_i| \leq k(1-\eta)$. We say that a graph on n vertices has $(1-\varepsilon)$-*dominant distance* t, if among the $\binom{n}{2}$ pairs of distinct vertices at least $(1-\varepsilon)\binom{n}{2}$ are at distance t.

In our main result on spectral expansion we show that distance-regular graphs of bounded diameter are spectral expanders if they have $(1-\varepsilon)$-dominant distance for sufficiently small $\varepsilon > 0$, depending only on the diameter. This result is one of the key components in the proof of Theorem 1.2.6.

Theorem 1.4.1. *For every $d \geq 2$ there exist $\epsilon = \epsilon(d) > 0$ and $\eta = \eta(d) > 0$ such that the following holds. If a distance-regular graph X of diameter d has a $(1-\epsilon)$-dominant distance, then X is a spectral η-expander.*

The key ingredient in the proof of Theorem 1.4.1 is the following new inequality on the intersection numbers of the distance-regular graphs. Essentially, this inequality claims that, if for some j, b_j is large (and therefore, by monotonicity, so are b_i for $i \leq j$) and c_{j+1} is small, then b_{j+1} and c_{j+2} cannot be small simultaneously. In particular, if c_d is sufficiently small, then this inequality shows that b_i do not decrease too fast.

Theorem 1.4.2 (Growth-induced tradeoff)**.** *Let X be a distance-regular graph of diameter $d \geq 2$. Let $0 \leq j \leq d-2$. Assume $b_j > c_{j+1}$ and let $C = b_j/c_{j+1}$. Then for every $1 \leq s \leq j+1$ we have*

$$b_{j+1}\left(\sum_{t=1}^{s}\frac{1}{b_{t-1}}+\sum_{t=1}^{j+2-s}\frac{1}{b_{t-1}}\right)+c_{j+2}\sum_{t=1}^{j+1}\frac{1}{b_{t-1}}\geq 1-\frac{4}{C-1}. \tag{1.4}$$

We prove this inequality in Section 7.2.

In a distance-regular graph, denote by λ and μ the number of common neighbours of a pair of adjacent vertices, and a pair of vertices at distance 2, respectively. We mention, that a result of Terwilliger [1986], as strengthened in [Brouwer et al., 1989, Theorem 4.3.3], shows that every non-principal eigenvalue of a k-regular distance-regular graph X has absolute value at most $k-\lambda$ if $\mu > 1$ and X is not the icosahedron. This result assures that X is a spectral η-expander, if $\lambda \geq \eta k$. We note that while both our result and Terwilliger's result provide simple sufficient combinatorial conditions for being spectral expanders, they are incomparable. In fact, our primary motivation for a spectral gap bound is an application of Lemma 4.5.11, where Terwilliger's gap is not sufficient.

Additionally, we note that in Theorem 1.4.1 we do not exclude the elusive case $\mu = 1$, for which almost no classification results are known, and which is known to be a difficult case in various circumstances. A remarkable example is the Bannai-Ito conjecture, where the case $\mu = 1$ was the only obstacle for 30 years, and was resolved only recently in the breakthrough paper by Bang et al. [2015].

Combining Theorem 1.4.1 with the Metsch characterization of geometric graphs (The-

orem 3.1.3), and Babai's Spectral tool for motion lower bounds (Theorem 4.5.11), in Theorem 8.1.7 we reduce Theorem 1.2.6 to the case of geometric graphs. By exploiting rich structure of geometric graphs, we show that the only such graphs with sublinear motion are Johnson and Hamming graphs. This step relies on the new characterizations of these families of graphs that we discuss below.

1.4.2 New characterizations of Johnson and Hamming graphs

A result of Terwilliger [1986] (see [Brouwer et al., 1989, Theorem 4.4.3]) implies that the icosahedron is the only distance-regular graph, for which the second largest eigenvalue θ_1 (of the adjacency matrix) satisfies $\theta_1 > b_1 - 1$ and a pair of vertices at distance 2 has $\mu \geq 2$ common neighbors. Another classical result gives the classification of distance-regular graphs with $\mu \geq 2$ and $\theta_1 = b_1 - 1$.

Theorem 1.4.3 ([Brouwer et al., 1989, Theorem 4.4.11])**.** *Let X be a distance-regular graph of diameter $d \geq 3$ with second largest eigenvalue $\theta_1 = b_1 - 1$. Assume $\mu \geq 2$. Then one of the following holds:*

1. *$\mu = 2$ and X is a Hamming graph, a Doob graph, or a locally Petersen graph (and all such graphs are known).*
2. *$\mu = 4$ and X is a Johnson graph.*
3. *$\mu = 6$ and X is a half cube.*
4. *$\mu = 10$ and X is a Gosset graph $E_7(1)$.*

We consider the case $\theta_1 \geq (1 - \varepsilon)b_1$ for a sufficiently small $\varepsilon > 0$. The relaxation of the assumption on the second largest eigenvalue comes at the cost of requiring additional structural constraints. Our main structural assumption is that X is a geometric distance-regular graph, meaning that there exists a collection of Delsarte cliques (see Sec. 3.1) $\mathcal{C}$

such that every edge of X belongs to a unique clique in $\mathcal{C}$. Additional technical structural assumptions depend on whether the neighborhood graphs of X are connected. We note that for a geometric distance-regular graph X either the neighborhood graph $X(v)$ is connected for every vertex v, or $X(v)$ is disconnected for every vertex v (see Lemma 3.2.4). We give the following characterizations.

Theorem 1.4.4 (**Main IV**). *There exists an absolute constant $\varepsilon^* > 0.0065$ such that the following is true. Let X be a geometric distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-m$. Suppose that $\mu \geq 2$ and $\theta_1 + 1 > (1 - \varepsilon^*)b_1$. Moreover, assume that the vertex degree satisfies $k \geq \max(m^3, 29)$ and the neighborhood graph $X(v)$ is connected for some vertex v of X.*

Then X is a Johnson graph $J(s, d)$ with $s = (k/d) + d$.

Theorem 1.4.5 (**Main V**). *Let X be a geometric distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-m$. Consider an arbitrary $0 < \varepsilon < 1/(6m^4 d)$. Suppose that $\mu \geq 2$ and $\theta_1 \geq (1 - \varepsilon)b_1$. Moreover, assume $c_t \leq \varepsilon k$ and $b_t \leq \varepsilon k$ for some $t \leq d$, and the neighborhood graph $X(v)$ is disconnected for some vertex v of X.*

Then X is a Hamming graph $H(d, s)$ with $s = 1 + k/d$.

Remark 1.4.6. If $s > 6d^5 + 1$, then the Hamming graph $H(d, s)$ satisfies the assumptions of this theorem with $1/(s - 1) \leq \varepsilon < 1/(6d^5)$ and $t = d$.

We present the proof of these theorems in Sections 6.2 and 6.3. These characterizations will be used in Section 8.1 to prove Theorem 1.2.6.

The assumption that a distance-regular graph is geometric excludes only finitely many graphs with $\mu \geq 2$, if the smallest eigenvalue of the graph is assumed to be bounded, as proved by Koolen and Bang [2010].

Theorem 1.4.7 (Koolen and Bang [2010]). *Fix an integer $m \geq 2$. There are only finitely many non-geometric distance-regular graphs of diameter ≥ 3 with $\mu \geq 2$ and smallest eigenvalue at least $-m$.*

However, in the context of Theorem 1.2.6 we do not have a bound on the smallest eigenvalue in the non-geometric case, so we do not use the above theorem in the proof.

1.4.3 A characterization of Hamming schemes by smallest eigenvalue

A number of classification results is known under the assumption of bounded smallest eigenvalue.

For strongly regular graphs, Neumaier [1979] showed that if the smallest eigenvalue is $-m$ (for $m \geq 2$), then it is a Latin square graph $LS_m(n)$, a Steiner graph $S_m(n)$, a complete multipartite graph or one of finitely many other graphs. A classification of the strongly regular graphs with smallest eigenvalue -2 was known earlier (Seidel [1968]) . Moreover, Cameron et al. [1991] gave a complete classification of all graphs with smallest eigenvalue -2. They proved that all but finitely many of such graphs have rich geometric structure (they are generalized line graphs).

Koolen and Bang [2010] proved that all but finitely many distance-regular graphs with smallest eigenvalue $-m$ and $\mu \geq 2$ are geometric. For geometric distance-regular graphs with smallest eigenvalue ≥ -3 and $\mu \geq 2$ Bang [2013] and Bang and Koolen [2014] gave a complete classification. Moreover, they conjectured [Koolen and Bang, 2010, Conjecture 7.4] that for every integer m all but finitely many geometric distance-regular graphs with smallest eigenvalue $-m$ and $\mu \geq 2$ are known.

Conjecture 1.4.8 (Koolen and Bang [2010])**.** *For a fixed integer $m \geq 2$, every geometric distance-regular graph with smallest eigenvalue $-m$, diameter ≥ 3 and $\mu \geq 2$ is either a Johnson graph, or a Hamming graph, or a Grassmann graph, or a bilinear forms graph, or the number of vertices is bounded above by a function of m.*

In this thesis we show that distance-regular graphs of diameter d with smallest eigenvalue $-d$, $\mu \leq 3$, an induced quadrangle, and sufficiently large degree k are Hamming graphs.

Theorem 1.4.9 (**Main VI**). *Let X be a distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-d$. Suppose that X contains an induced quadrangle, $\mu \leq 3$, and $k \geq \left(100d^3 \ln d\right) \cdot c_d$. Then X is the Hamming graph $H(d, k/d+1)$.*

The proof of this theorem is discussed in Section 5.2. This characterization also plays a crucial role in our proof of the robustness under extension for Hamming schemes (Theorem 1.3.6, see Section 10.5).

1.5 Acknowledgement of collaborations

Some of the results of this thesis originally appeared in joint papers with László Babai.

In particular, Theorem 1.3.2 and most of the results of Chapters 10 and 11 are a result of joint work by Babai and Kivva [2022]. Only the results on Hamming and Grassmann schemes from these chapters are not a part of this work by Babai and Kivva [2022].

Additionally, the discussion in Section 4.4 is a part of Babai and Kivva [2020].

Most of other original results of this thesis appeared in Kivva [2021a,b,c, 2022].

More precisely, Theorems 1.4.1 and the results of Chapter 7, Section 8.1.3 and 8.4 appeared in Kivva [2021b]. Theorems 1.4.4, 1.4.5 and 1.2.6 and the results of Chapter 6, Section 8.1 and 8.2 first appeared in Kivva [2021c]. The results of Chapter 9 and Theorem 1.2.8 were proved in Kivva [2021a]. Finally, Theorem 1.4.9 and the results of Chapters 10 and 11 related to Hamming and Grassmann schemes are from Kivva [2022].

1.6 Organization of the thesis

We now outline the structure of this thesis. In Chapter 2 we give definitions and discuss basic properties of graphs, groups, coherent configurations and distance-regular graphs. In Chapter 3 we outline preliminaries on geometric distance-regular graphs, a class of a great interest to our analysis.

In Chapter 4 we discuss the classification of large primitive groups by Cameron [1981] and the classification of primitive group with sublinear minimal degree by Liebeck and Saxl [1991]. Additionally, in this chapter, we outline the combinatorial and spectral tools for bounding the order and the minimal degree of primitive permutation groups developed by Babai.

In Chapters 5 and 6 we prove our characterizations of Johnson and Hamming graphs, which are used in the proof of Theorem 1.2.6. In this chapter, we also briefly discuss how these results are related to the study of regular graphs with bounded eigenvalue and representation theory of distance-regular graphs.

We prove Theorem 1.4.1 in Chapter 7.

We study motion of distance-regular graphs in Chapter 8 and of primitive coherent configurations of rank-4 in Chapter 9.

Finally, we present our results on robustness of Johnson, Hamming and Grassmann schemes in Chapters 10 and 11.

CHAPTER 2

PRELIMINARIES: COHERENT CONFIGURATIONS, DISTANCE-REGULAR GRAPHS

2.1 Graphs and digraphs

Denote $[m] = \{1, 2, ..., m\}$. For a set S and a positive integer k, $\binom{S}{k}$ denotes the set of the subsets of S of size k.

Definition 2.1.1. In this thesis, by a *graph* we mean a pair $X = (V, E)$, where $E \subseteq \binom{V}{2}$. The set $V = V(X)$ is called the *set of vertices* of X and $E = E(X)$ is called the *set of edges* of X. That is, we do not allow loops or repeated edges.

Definition 2.1.2. In this thesis by a *digraph* we mean a pair $X = (V, E)$, where $E \subseteq V \times V$. A pair $(u, v) \in E$ is called a *directed edge* (from u to v). Note, that we allow loops for digraphs.

We think of every graph as also being a digraph by replacing every edge $\{u, v\}$ by a pair of directed edges (u, v) and (v, u).

Definition 2.1.3. In a digraph X a *walk* is a sequence of vertices $u_0, u_1, \ldots, u_t$, such that $(u_{i-1}, u_i) \in E(X)$ for every $i \in [t]$. A *path* is a walk that consists of distinct vertices.

For vertices $v, w \in V(X)$ we define the distance $\mathrm{dist}(v, w)$ from v to w to be the length of the shortest path from v to w. If no such path exists, we define $\mathrm{dist}(v, w) = \infty$. For a non-empty subset $C \subseteq V(X)$ and a vertex $v \in V(X)$ we define $\mathrm{dist}(v, C) = \min_{u \in C} \mathrm{dist}(v, u)$.

Definition 2.1.4. For a digraph X, we define the *diameter* to be the largest distance between a pair of distinct vertices of X.

Let X be a graph. We always denote by n the number of vertices of X and for a regular graph X we denote by k its degree. The diameter of a graph is the largest distance between a

pair of vertices of the graph. We denote the diameter of X by d. If the graph is disconnected, then its diameter is defined to be ∞.

Definition 2.1.5. A regular graph is called *edge-regular* if every pair of adjacent vertices has the same number $\lambda = \lambda(X)$ of common neighbors.

Definition 2.1.6. A regular graph is called *co-edge-regular* if every pair of non-adjacent vertices has the same number $\mu = \mu(X)$ of common neighbors.

Definition 2.1.7. An edge-regular graph is called *amply regular* if every pair of vertices at distance 2 has the same number $\mu = \mu(X)$ of common neighbors.

Definition 2.1.8. A graph is called *strongly regular* if it is edge-regular and co-edge regular.

Denote by $q(X)$ the maximum number of common neighbors of two distinct vertices in X.

Definition 2.1.9. For a graph $X = (V, E)$ and a subset of vertices $S \subseteq V$, the *induced subgraph* on S is the graph $X[S] = (S, E_S)$, where $E_S = E \cap \binom{S}{2}$.

Let $N(v)$ denote the set of neighbors of vertex v and $N_i(v) = \{w \in V(X) | \operatorname{dist}(v, w) = i\}$ the set of vertices at distance i from v in X. By $X(v)$ we denote the *neighborhood graph* of v, i.e., the graph induced by X on $N(v)$.

Definition 2.1.10. Let X be a graph. The *line graph* of X is the graph $L(X)$ with $E(X)$ as its set of vertices, where distinct $e_1, e_2 \in E(X)$ are adjacent if they (as edges of X) share a vertex.

For a digraph X on n vertices the *adjacency matrix* is the $n \times n$ matrix A indexed by the set $V(X)$, in which $A_{uv} = 1$ if $(u, v) \in E(X)$ and $A_{uv} = 1$ otherwise. In particular, if X is a graph, then A is symmetric and all the diagonal entries are 0.

By the *eigenvalues of a graph* we mean the eigenvalues of its *adjacency matrix*.

Let A be the adjacency matrix of the graph X. Suppose that X is k-regular. Then the all-ones vector is an eigenvector of A with eigenvalue k. We call them the *trivial eigenvector*

and the *trivial eigenvalue*. All other eigenvalues of A have absolute value not greater than k. We call them *non-trivial* eigenvalues.

Definition 2.1.11. A *crown graph* is a graph that is obtained from a complete bipartite graph by deleting one perfect matching.

Definition 2.1.12. A *complete multipartite graph* $K_{a_1,a_2,\ldots,a_t}$ is defined to be the graph whose set of vertices is $V = V_1 \sqcup V_2 \sqcup \ldots \sqcup V_t$, with $|V_i| = a_i$, in which there is an edge between $x \in V_i$ and $y \in V_j$ if and only if $i \neq j$. We use notation $K_{t\times a}$ to denote the complete multipartite graph $K_{a_1,a_2,\ldots,a_t}$ in which $a_i = a$ for all i.

Definition 2.1.13. We say that a graph has an *induced quadrangle* if this graph has an induced 4-cycle C_4.

2.2 Groups

For a pair of groups G and H we use notation $H \leq G$ to say that H is a subgroup of G. For a finite set Ω we use $\mathrm{Sym}(\Omega)$ to denote the symmetric permutation group on the set Ω. We use S_n and A_n to denote the symmetric group and the alternating group on n elements, respectively.

Definition 2.2.1. A group $G \leq \mathrm{Sym}(\Omega)$ is called a *permutation group* on Ω.

For a permutation group $G \leq \mathrm{Sym}(\Omega)$ we say that $|\Omega|$ is the *degree* of G.

Definition 2.2.2. A permutation group $G \leq \mathrm{Sym}(\Omega)$ is called *transitive* if for every $x, y \in \Omega$ there exists an element $g \in G$ that maps x to y.

Let G be a transitive permutation group on the set Ω. A G-invariant partition $\Omega = B_1 \sqcup B_2 \sqcup \ldots \sqcup B_t$ is called a *system of imprimitivity* of G. Every permutation group $G \leq \mathrm{Sym}(\Omega)$ admits two trivial G-invariant partitions: the partition consisting of Ω only, and the partition of Ω into singletons.

Definition 2.2.3. A non-trivial transitive permutation group is called *primitive* if it does not admit any non-trivial system of imprimitivity.

Definition 2.2.4. For a group G define the *socle* to be the product of its minimal normal subgroups.

2.3 Coherent configurations

Our terminology follows Babai [2016b].

Let V be a finite set, elements of which will be called vertices of a configuration.

Definition 2.3.1. A *configuration* $\mathfrak{X}$ of rank r on the set V is a pair $(V, \mathfrak{c})$, where $\mathfrak{c}$ is a surjective map $\mathfrak{c} : V \times V \to \{0, 1, ..., r-1\}$ such that

(i) $\mathfrak{c}(v, v) \neq \mathfrak{c}(u, w)$, for every $v, u, w \in V$ with $u \neq w$,

(ii) for every $i < r$, there is $i^* < r$, such that $\mathfrak{c}(u, v) = i$ implies $\mathfrak{c}(v, u) = i^*$, for all $u, v \in V$.

The value $\mathfrak{c}(u, v)$ is called the *color* of a pair (u, v). The color $\mathfrak{c}(u, v)$ is a *vertex color* if $u = v$, and is an *edge color* if $u \neq v$. Then condition (i) says that edge colors are different from vertex colors, and condition (ii) says that the color of a pair (u, v) determines the color of (v, u).

Definition 2.3.2. For every $i < r$ consider the set $R_i = \{(u, v) : \mathfrak{c}(u, v) = i\}$ of pairs of color i and consider the digraph $X_i = (V, R_i)$. We refer to both R_i and X_i as the color-i *constituent* of $\mathfrak{X}$.

There are two possibilities: if $i = i^*$, then color i and the corresponding constituent X_i are called *undirected*; if $i \neq i^*$, then $(i^*)^* = i$ and color i together with the corresponding constituent X_i are called *oriented*.

Clearly, $\{R_i\}_{i<r}$ forms a partition of $V \times V$.

We denote the adjacency matrix of the digraph X_i by A_i. The adjacency matrices of the constituents satisfy

$$\sum_{i=0}^{r-1} A_i = J_{|V|} = J, \tag{2.1}$$

where J denotes the all-ones matrix.

Note that conditions (i) and (ii) of Definition 2.3.1 in the matrix language mean the following. There exists a set $\mathcal{D}$ of colors, such that the identity matrix can be represented as a sum $\sum\limits_{i\in\mathcal{D}} A_i = I$. And for every color i, $A_i^T = A_{i^*}$.

For a set of colors $\mathcal{I}$ we denote by $X_{\mathcal{I}}$ the digraph on the set of vertices V, where an arc (x, y) is in $X_{\mathcal{I}}$ if and only if $\mathfrak{c}(x, y) \in \mathcal{I}$. For small sets we omit braces, for example, $X_{1,2}$ will be written in place of $X_{\{1,2\}}$.

Definition 2.3.3. A configuration $\mathfrak{X}$ is *homogeneous* if $\mathfrak{c}(u, u) = \mathfrak{c}(v, v)$ for every $u, v \in V$.

Unless specified otherwise, we always assume that 0 is the vertex color of a homogeneous configuration. The constituent which corresponds to the vertex color is also referred as the *diagonal constituent*.

Definition 2.3.4. A configuration $\mathfrak{X} = (V, \mathfrak{c})$ is called *symmetric* if $\mathfrak{c}(u, v) = \mathfrak{c}(v, u)$ for all $u, v \in V$.

Definition 2.3.5. We say that a homogeneous symmetric configuration $\mathfrak{X}$ is *regular* if every off-diagonal constituent is a regular graph.

Definition 2.3.6. A configuration $\mathfrak{X}$ is *coherent* if

(iii) for every $i, j, t < r$, there is an *intersection number* $p_{i,j}^t$ such that, for all $u, v \in V$, if $\mathfrak{c}(u, v) = t$, then there exist exactly $p_{i,j}^t$ vertices $w \in V$ with $\mathfrak{c}(u, w) = i$ and $\mathfrak{c}(w, v) = j$.

The definition of a coherent configuration has several simple, but important, consequences.

Observation 2.3.7. Let $\mathfrak{X}$ be a coherent configuration. Then every edge color is aware of the colors of its tail and head. That is, for every edge color i, there exist vertex colors i_- and i_+ such that if $\mathfrak{c}(u,v) = i$, then $\mathfrak{c}(u,u) = i_-$ and $\mathfrak{c}(v,v) = i_+$.

Proof. Indeed, they are the only colors for which p^i_{i,i_+} and $p^i_{i_-,i}$ are non-zero. □

Observation 2.3.8. For every color i its *in-degree* and *out-degree* are well-defined as $k_i^- = p^{i_+}_{i^*,i}$ and $k_i^+ = p^{i_-}_{i,i^*}$, respectively.

In a homogeneous coherent configuration we have $k_i^+ = k_i^-$ for every color i. We denote this common value by k_i.

Definition 2.3.9. Let A_i denote the adjacency matrix of the color i constituent of the coherent configuration $\mathfrak{X}$. Define the *Bose-Mesner algebra* of $\mathfrak{X}$ to be to be the algebra generated by A_i.

Observe that the existence of the intersection numbers is equivalent to the following conditions on the adjacency matrices of the constituent digraphs.

$$A_i A_j = \sum_{t=0}^{r-1} p^t_{i,j} A_t \quad \text{for all } i,j < r. \tag{2.2}$$

Hence, the following observation follows.

Observation 2.3.10. $\{A_i : 0 \le i \le r-1\}$ form a basis of the Bose-Mesner algebra of $\mathfrak{X}$ with structure constants $p^t_{i,j}$. In particular, this algebra is r-dimensional and every A_i has minimal polynomial of degree at most r.

Observation 2.3.11. The intersection numbers of a homogeneous coherent configuration satisfy the following relations.

$$\sum_{j=0}^{r-1} p^t_{ij} = k_i \quad \text{and} \quad p^s_{i,j} k_s = p^i_{s,j^*} k_i. \tag{2.3}$$

Let $i, j < r$ be edge colors.

Definition 2.3.12. Take $u, v \in V$ with $\mathfrak{c}(u, v) = j$. Define $\mathrm{dist}_i(u, v)$ to be the length ℓ of a shortest walk $u_0 = u, u_1, ..., u_\ell = v$ such that $\mathfrak{c}(u_{t-1}, u_t) = i$ for every $t \in [\ell]$.

Observation 2.3.13. $\mathrm{dist}_i(j) = \mathrm{dist}_i(u, v)$ is well defined, i.e., does not depend on the choice of u, v, but only on the colors j and i.

Proof. Let $\mathfrak{c}(u, v) = \mathfrak{c}(u', v') = j$ and suppose there exist a walk $u_0 = u, u_1, ..., u_\ell = v$ of length ℓ, such that $\mathfrak{c}(u_{t-1}, u_t) = i$. Denote by $e_t = \mathfrak{c}(u_t, v)$. Then we know that $p_{i,e_t}^{e_{t-1}} \neq 0$ for $t \in [\ell - 1]$. Let $u'_0 = u'$. Then, as $p_{i,e_t}^{e_{t-1}} \neq 0$, by induction, there exists a u'_t such that $\mathfrak{c}(u'_{t-1}, u'_t) = i$ and $\mathfrak{c}(u_t, v) = e_t$ for all $t \in [\ell - 1]$. Hence, $\mathrm{dist}(u', v') \leq \mathrm{dist}(u, v)$ and similarly $\mathrm{dist}(u, v) \leq \mathrm{dist}(u', v')$. Therefore, $\mathrm{dist}_i(j)$ is well-defined. □

Observation 2.3.14. *If* $\mathrm{dist}_i(j)$ *is finite, then* $\mathrm{dist}_i(j) \leq r - 1$.

Proof. Suppose that $\mathrm{dist}_i(j)$ is finite, then for $\mathfrak{c}(u, v) = j$ there exists a shortest walk $u_0 = u, u_1, ..., u_\ell = v$ with $\mathfrak{c}(u_{t-1}, u_t) = i$. Denote by $e_t = \mathfrak{c}(u_t, v)$ for $0 \leq t \leq \ell - 1$. Then, all e_t are distinct edge colors, or the walk can be shortened. Thus $\ell \leq r - 1$. □

Definition 2.3.15. A coherent configuration is called an *association scheme* if $\mathfrak{c}(u, v) = \mathfrak{c}(v, u)$ for all $u, v \in V$.

Corollary 2.3.16. *Every association scheme is a homogeneous configuration.*

Proof. Since in a coherent configuration color of every edge is aware of the colors of its head and tail vertices, these vertices have the same color for every edge. □

Note, for an association scheme every constituent digraph is a graph. Thus, for an association scheme and $i \neq 0$, the i-th constituent X_i is a k_i-regular graph with $\lambda(X_i) = p_{i,i}^i$. Moreover, it is clear that $p_{i,j}^s = p_{j,i}^s$ for association schemes.

Definition 2.3.17. A homogeneous coherent configuration is called *primitive* if every non-diagonal constituent is weakly connected.

It is not hard to check that every non-diagonal constituent of a homogeneous coherent configuration is weakly connected if and only if it is strongly connected.

Note, that by Observation 2.3.14, we have $\mathrm{dist}_i(j) \leq r-1$ for all edge colors i, j of a primitive coherent configuration.

The following definition will be useful.

Definition 2.3.18. We say that an association scheme has *diameter* d if every non-diagonal constituent has diameter at most d and there exists a non-diagonal constituent of diameter d.

Note, that if an association scheme has a finite diameter, then, in particular, it is primitive. Moreover, every primitive association scheme of rank r has diameter $\leq r-1$.

Observe, that for every undirected color i the constituent X_i is an edge-regular graph. We also introduce the following definition.

Definition 2.3.19. We say that a homogeneous coherent configuration $\mathfrak{X}$ of rank r *has constituents ordered by degree*, if color 0 corresponds to the diagonal constituent and the degrees of non-diagonal constituents satisfy $k_1 \leq k_2 \leq ... \leq k_{r-1}$.

2.4 Distance-regular graphs

Definition 2.4.1. A connected graph X is called distance-transitive is for every four vertices $x_1, x_2, y_1, y_2 \in V(X)$ if $\mathrm{dist}(x_1, y_1) = \mathrm{dist}(x_2, y_2)$, then there exists an automorphism $\sigma \in \mathrm{Aut}(X)$, such that $\sigma(x_1) = x_2$ and $\sigma(y_1) = y_2$.

Definition 2.4.2. A connected graph X of diameter d is called *distance-regular* if for every $0 \leq i \leq d$ there exist integers a_i, b_i, c_i such that for all $v \in V(X)$ and all $w \in N_i(v)$ the number of edges between w and $N_i(v)$ is a_i, between w and $N_{i-1}(v)$ is c_i, and between w and $N_{i+1}(v)$ is b_i. The sequence

$$\iota(X) = \{b_0, b_1, \ldots, b_{d-1}; c_1, c_2, \ldots, c_d\}$$

is called the *intersection array* of X.

Clearly, every distance-transitive graph is distance-regular. Also, note that every distance-regular graph is amply regular with $\lambda = a_1$ and $\mu = c_2$.

By simple counting, the following relations hold among the parameters of distance-regular graphs.

(E1) $a_i + b_i + c_i = k$ for every $0 \leq i \leq d$,

(E2) $b_{i+1} \leq b_i$ and $c_{i+1} \geq c_i$ for $0 \leq i \leq d-1$.

(E3) $|N_i(v)|b_i = |N_{i+1}(v)|c_{i+1}$, for $0 \leq i \leq d-1$ and every vertex v.

Thus, in particular, (E3) implies that the numbers $k_i = |N_i(v)|$ do not depend on the vertex $v \in V(X)$, and we can rewrite the last property as

(E3') $k_i b_i = k_{i+1} c_{i+1}$, for $0 \leq i \leq d-1$.

With every graph of diameter d we can naturally associate matrices A_i, where rows and columns are indexed by vertices, with $(A_i)_{u,v} = 1$ if and only if $\mathrm{dist}(u,v) = i$, and $(A_i)_{u,v} = 0$, otherwise. That is, A_i is the adjacency matrix of the distance-i graph X_i of X. For a distance-regular graph these matrices satisfy the following relations

$$A_0 = I, \quad A_1 =: A, \quad \sum_{i=0}^{d} A_i = J, \tag{2.4}$$

$$AA_i = c_{i+1}A_{i+1} + a_i A_i + b_{i-1}A_{i-1} \quad \text{for } 0 \leq i \leq d, \tag{2.5}$$

where $c_{d+1} = b_{-1} = 0$ and $A_{-1} = A_{d+1} = 0$. Clearly, Eq. (2.5) implies that for every $0 \leq i \leq d$ there exists a polynomial ν_i of degree exactly i, such that $A_i = \nu_i(A)$. Moreover, the minimal polynomial of A has degree exactly $d+1$. Hence, since A is symmetric, A has exactly $d+1$ distinct real eigenvalues. Additionally, we conclude that for every $0 \leq i, j, s \leq d$

there exist *intersection numbers* $p^s_{i,j}$, such that

$$A_i A_j = \sum_{s=0}^{d} p^s_{i,j} A_s. \tag{2.6}$$

With this notation, $a_i = p^i_{i,1}$, $b_i = p^i_{i+1,1}$ and $c_i = p^i_{i-1,1}$.

Recalling the definition of A_i, this implies that for all $u, v \in V(X)$ with $\mathrm{dist}(u, v) = s$ there exist exactly $p^s_{i,j}$ vertices at distance i from u and distance j from v, i.e.,

$$|N_i(u) \cap N_j(v)| = p^s_{i,j}. \tag{2.7}$$

Therefore, every distance-regular graph X of diameter d induces an association scheme $\mathfrak{X}$ of rank $d+1$, where vertices are connected by an edge of color i in $\mathfrak{X}$ if and only if they are at distance i in X, for $0 \leq i \leq d$. Hence, we get the following statement.

Lemma 2.4.3. *If a graph X is distance-regular of diameter d, then the distance-i graphs X_i form constituents of an association scheme $\mathfrak{X}$ of rank $d+1$. Moreover, if $\mathfrak{X}$ is primitive, then it has diameter d. In the opposite direction, if an association scheme of rank $d+1$ has a constituent of diameter d, then this constituent is distance-regular.*

Lemma 2.4.4. *Let X be a distance-regular graph of diameter $d \geq 2$. Then $2\lambda \leq k + \mu$.*

Proof. Denote $N(x, y) = N(x) \cap N(y)$ for vertices x and y of X. The inequality above follows from the two obvious inclusions below, applied to vertices v and w at distance 2 in X, and their common neighbor u.

$$N(u, v) \cup N(u, w) \subseteq N(u), \qquad N(u, v) \cap N(u, w) \subseteq N(v, w). \qquad \square$$

In Section 5.2 we will need the following inequality by Terwilliger [1985]. Recall that a graph has an induced quadrangle if it has an induced 4-cycle.

Theorem 2.4.5 (Terwilliger [1985], see [Brouwer et al., 1989, Theorem 5.2.1])**.** *Let X be a distance-regular graph. If X contains an induced quadrangle, then*

$$c_i - b_i \geq c_{i-1} - b_{i-1} + \lambda + 2, \quad \textit{for } i = 1, 2, \ldots, d.$$

We also need the following bound on the difference $c_i - c_{i-1}$ for $i = 3$.

Theorem 2.4.6 ([Brouwer et al., 1989, Theorem 5.4.1])**.** *Let X be a distance-regular graph of diameter $d \geq 3$. If $\mu \geq 2$, then either $c_3 \geq 3\mu/2$, or $c_3 \geq \mu + b_2$ and $d = 3$.*

Corollary 2.4.7. *Let X be a distance-regular graph of diameter $d \geq 3$. If $\mu \geq 2$, then $c_3 > \mu$.*

A distance-regular graph X of diameter d has precisely $d+1$ distinct eigenvalues. Denote these eigenvalues by $\theta_0 = k > \theta_1 > \ldots > \theta_d$. They are the eigenvalues of the tridiagonal *intersection matrix*

$$L_1 = \begin{pmatrix} a_0 & b_0 & 0 & 0 & \ldots \\ c_1 & a_1 & b_1 & 0 & \ldots \\ 0 & c_2 & a_2 & b_2 & \ldots \\ \ldots & & \vdots & & \ldots \\ \ldots & & 0 & c_d & a_d \end{pmatrix}.$$

For an eigenvalue θ, consider the sequence $(u_i(\theta))_{i=0}^d$ defined by the relations

$$u_0(\theta) = 1, \qquad u_1(\theta) = \frac{\theta}{k},$$

$$c_i u_{i-1}(\theta) + a_i u_i(\theta) + b_i u_{i+1}(\theta) = \theta u_i(\theta), \quad \text{for } i = 1, 2, \ldots, d-1,$$

$$c_d u_{d-1}(\theta) + a_d u_d(\theta) = \theta u_d(\theta).$$

The vector $u = (u_0(\theta), u_1(\theta), \ldots, u_d(\theta))^T$ is an eigenvector of L_1 corresponding to θ.

Definition 2.4.8. The sequence $(u_i(\theta))_{i=0}^d$ is called the *standard sequence* of X corresponding to the eigenvalue θ.

We denote by f_i the *multiplicity* of the eigenvalue θ_i of X. Since X is a connected graph, $f_0 = 1$. In general, the multiplicities f_i can be computed using the Biggs formula.

Theorem 2.4.9 (Biggs [1971], see [Brouwer et al., 1989, Theorem 4.1.4])**.** *The multiplicity of the eigenvalue θ of the distance-regular graph X can be expressed as*

$$f(\theta) = \frac{n}{\sum\limits_{i=0}^{d} k_i u_i(\theta)^2}.$$

2.5 Imprimitive distance-regular graphs

Here we briefly describe some basic properties of imprimitive distance-regular graphs that we will need later (in Section 8.4). Recall that a distance-regular graph X is imprimitive if for some $1 \leq i \leq d$ the distance-i graph X_i is disconnected. Smith's theorem states that there are only two types of imprimitive distance-regular graphs.

Definition 2.5.1. A distance-regular graph X of diameter d is called *antipodal* if being at distance d in X is an equivalence relation, that is, if X_d is a disjoint union of cliques.

Theorem 2.5.2 (D. H. Smith [1971])**.** *An imprimitive distance-regular graph of degree $k > 2$ is bipartite or antipodal (or both).*

If X is a bipartite graph, then X_2 has two connected components X^+ and X^-, which are called the *halved graphs* of X and are denoted $\frac{1}{2}X$.

For an antipodal graph X of diameter d, define the graph $\widetilde{X}$ which has the equivalence classes of X_d as vertices and two equivalence classes are adjacent if there is an edge between them in X. The graph $\widetilde{X}$ is called the *folded graph* of X.

In the next proposition we state some properties of the intersection numbers of halved and folded graphs, which we will need later (in Section 8.4).

Proposition 2.5.3 (Biggs and Gardiner [1974], see [Brouwer et al., 1989, Proposition 4.2.2]). *Let X be a distance-regular graph with intersection array $\iota(X) = \{b_0, b_1, \ldots, b_{d-1}; c_1, c_2, \ldots, c_d\}$ and diameter $d \in \{2t, 2t+1\}$.*

1. *The graph X is bipartite if and only if $b_i + c_i = k$ for $i = 0, 1, \ldots, d$. In this case the halved graphs are distance-regular of diameter t with intersection array*

$$\iota(X^{\pm}) = \left\{ \frac{b_0 b_1}{\mu}, \frac{b_2 b_3}{\mu}, \ldots, \frac{b_{2t-2} b_{2t-1}}{\mu}; \frac{c_1 c_2}{\mu}, \frac{c_3 c_4}{\mu}, \ldots, \frac{c_{2t-1} c_{2t}}{\mu} \right\}.$$

2. *The graph X is antipodal if and only if $b_i = c_{d-i}$ for $i \neq t$. In this case X is an antipodal r-cover of its folded graph $\widetilde{X}$, where $r = 1 + b_t/c_{d-t}$. If $d > 2$, then $\widetilde{X}$ is distance-regular of diameter t with intersection array*

$$\iota(\widetilde{X}) = \{b_0, b_1, \ldots b_{t-1}; c_1, c_2, \ldots, c_{t-1}, \gamma c_t\},$$

where $\gamma = r$, if $d = 2t$; and $\gamma = 1$, if $d = 2t+1$.

It is not hard to show that given a distance-regular graph of degree $k > 2$ one may obtain a primitive distance-regular graph after halving at most once and folding at most once. More precisely, the following is true.

Proposition 2.5.4 (see [Brouwer et al., 1989, Sec. 4.2.A]). *Let X be a distance-regular graph of degree $k > 2$.*

1. *If X is a bipartite graph, then its halved graph is not bipartite.*

2. *If X is bipartite and either has an odd diameter or is not antipodal, then its halved graph is primitive.*

3. *If X is antipodal and either has an odd diameter or is not bipartite, then its folded graph is primitive.*

4. *If X has an even diameter and is both antipodal and bipartite, then the halved graphs $\frac{1}{2}X$ are antipodal, the folded graph $\widetilde{X}$ is bipartite and the graphs $\widetilde{\frac{1}{2}X} \cong \frac{1}{2}\widetilde{X}$ are primitive.*

CHAPTER 3

PRELIMINARIES: CLIQUE GEOMETRIES AND GEOMETRIC DISTANCE-REGULAR GRAPHS

3.1 Definition and Metsch's sufficient condition

Let X be a distance-regular graph, and $\theta_{\min}$ be its smallest eigenvalue. Delsarte [1973] proved that every clique C in X satisfies $|C| \leq 1 - \dfrac{k}{\theta_{\min}}$. A clique in X of size $1 - \dfrac{k}{\theta_{\min}}$ is called a *Delsarte clique*.

Definition 3.1.1. A *clique geometry* $\mathcal{C}_0$ for a graph X is a set of maximal cliques of X such that every edge of X is contained in exactly one member of $\mathcal{C}_0$.

Definition 3.1.2. A distance-regular graph X is called *geometric* if it admits a clique geometry consisting of Delsarte cliques.

Metsch proved that, under simple assumptions, a graph admits a clique geometry.

Theorem 3.1.3 ([Metsch, 1995, Result 2.2])**.** *Let $\mu \geq 1$, $\lambda^{(1)}, \lambda^{(2)} \geq 0$ and $m \geq 1$ be integers. Assume that X is a connected graph with the following properties.*

1. *Every pair of adjacent vertices has at least $\lambda^{(1)}$ and at most $\lambda^{(2)}$ common neighbors.*
2. *Every pair of distinct non-adjacent vertices has at most μ common neighbors.*
3. $2\lambda^{(1)} - \lambda^{(2)} > (2m-1)(\mu-1) - 1$.
4. *Every vertex has degree less than $(m+1)(\lambda^{(1)}+1) - \frac{1}{2}m(m+1)(\mu-1)$.*

Define a line *to be a maximal clique C of size $|C| \geq \lambda^{(1)} + 2 - (m-1)(\mu-1)$. Then every vertex belongs to at most m lines, and every pair of adjacent vertices belongs to a unique line. Therefore, the lines form a clique geometry for X.*

The following lemma shows that the existence of a clique geometry in a graph imposes strong lower bound on the smallest eigenvalue of the graph.

Lemma 3.1.4 ([van Dam et al., 2016, Prop. 9.8])**.** *Suppose that X satisfies the conditions of the previous theorem. Then the smallest eigenvalue of X is at least $-m$.*

Proof. Let $\mathcal{C}$ be the collection of lines of X. Consider $|V| \times |\mathcal{C}|$ vertex-clique incidence matrix N. That is, $(N)_{v,C} = 1$ if and only if $v \in C$ for $v \in X$ and $C \in \mathcal{C}$. Since every edge belongs to exactly one line, we get $NN^T = A + D$, where A is the adjacency matrix of X and D is a diagonal matrix. Moreover, $(D)_{v,v}$ equals to the number of lines that contain v. By the previous theorem, $D_{v,v} \leq m$ for every $v \in X$. Thus,

$$A + mI = NN^T + (mI - D)$$

is positive semidefinite, so all eigenvalues of A are at least $-m$. □

The following sufficient condition for being geometric is a slightly reformulated version of a result from van Dam et al. [2016].

Proposition 3.1.5 ([van Dam et al., 2016, Proposition 9.8])**.** *Let X be a distance-regular graph of diameter $d \geq 2$. Assume there exist a positive integer m and a clique geometry $\mathcal{C}$ of X such that every vertex u is contained in exactly m cliques of $\mathcal{C}$. If $k \geq m^2$, then X is geometric with smallest eigenvalue $-m$.*

Corollary 3.1.6 ([van Dam et al., 2016, Proposition 9.9])**.** *Let $m \geq 2$ be an integer, and let X be a distance-regular graph with $(m-1)(\lambda+1) < k \leq m(\lambda+1)$ and diameter $d \geq 2$. If $\lambda \geq \frac{1}{2}m(m+1)\mu$, then X is geometric with smallest eigenvalue $-m$.*

Proof. Directly follows from Theorem 3.1.3 and Proposition 3.1.5. □

We note that Corollary 3.1.6 will be used several times throughout the thesis. In particular, it is used in Chapter 8 in the proof of Theorem 8.1.7 to reduce the general case of Theorem 1.2.6 to geometric case, and in Chapter 10 in the proof of Theorem 10.3.3.

The converse holds without the $k \geq m^2$ assumption.

Lemma 3.1.7 (Godsil [1993b])**.** *Let X be a geometric distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-m$. Let $\mathcal{C}$ be a Delsarte clique geometry. Then m is an integer and every vertex belongs to precisely m Delsarte cliques in $\mathcal{C}$.*

Proof. By the definition of a Delsarte clique, its size is $1 + k/m$. Let $C_1, C_2, \ldots, C_t$ be the cliques in $\mathcal{C}$ which contain a vertex v. Since $\mathcal{C}$ is a clique geometry, all distinct C_i and C_j for $i, j \in [t]$ have only v in their intersection, and every vertex adjacent to v belongs to one of the cliques $C_1, C_2, \ldots, C_t$. Therefore, $k = t(|C_i| - 1) = tk/m$. □

In the case, when the smallest eigenvalue of a geometric distance-regular graph is -2, it is easy to deduce that the graph is a line graph. This also follows from a more general statement, Theorem 5.1.10, by Cameron, Goethals, Seidel and Shult Cameron et al. [1991].

Lemma 3.1.8. *Let X be a geometric distance-regular graph with smallest eigenvalue -2. Then X is the line graph $L(Y)$ for some graph Y.*

Proof. Let $\mathcal{C}$ be a Delsarte clique geometry of X. Define the graph Y with the set of vertices $V(Y) = \mathcal{C}$, in which a pair of distinct vertices $C_1, C_2 \in \mathcal{C}$ in Y is adjacent if and only if $C_1 \cap C_2 \neq \emptyset$. We claim that $L(Y) \cong X$. Indeed, since every edge of X is in exactly one clique from $\mathcal{C}$, $|C_1 \cap C_2| \leq 1$ for all distinct $C_1, C_2 \in \mathcal{C}$. So there is a well-defined map $f : E(Y) \to V(X)$. Moreover, by Lemma 3.1.7, every vertex of X is in exactly two cliques from $\mathcal{C}$, so f is bijective. Additionally, a pair of edges in Y share a vertex if and only if there is an edge between the corresponding vertices in X. Hence, $L(Y) \cong X$. □

Existance of a clique geometry provides the following useful bound on the number of common neighbors for a pair of vertices at distance 2, which will be used in Chapters 5-10.

Lemma 3.1.9 (see Sun and Wilmes [2015a]). *Let X be a graph. Let $\mathcal{C}$ be a collection of cliques in X, such that every edge lies in a unique clique from $\mathcal{C}$ and every vertex is in at most m cliques from $\mathcal{C}$. Then every pair of vertices at distance 2 has at most m^2 common neighbors.*

Proof. Let $u, v \in V(X)$ be a pair of vertices at distance 2. By the assumptions of the lemma we can write $N(u) = \bigcup_{i=1}^{m_u} C_i^u$ and $N(v) = \bigcup_{i=1}^{m_v} C_i^v$, where $C_i^u, C_j^v \in \mathcal{C}$. Since $\mathrm{dist}(u, v) = 2$, all cliques are distinct. Observe, that every pair of distinct cliques in $\mathcal{C}$ intersect each other in at most one vertex. Hence, $N(u) \cap N(v) \leq m_u m_v \leq m^2$. □

3.2 Vertex-clique intersection parameters

Suppose that X is a geometric distance-regular graph with a Delsarte clique geometry $\mathcal{C}$. Consider a Delsarte clique $C \in \mathcal{C}$. Assume $x \in V(X)$ satisfies $\mathrm{dist}(x, C) = i$. Define

$$\psi_i(C, x) := |\{y \in C \mid d(x, y) = i\}|. \tag{3.1}$$

By Bang et al. [2007], the numbers $\psi_i(C, x)$ do not depend on C and x, but only on the distance $\mathrm{dist}(x, C) = i$. Thus, we may define $\psi_i := \psi(C, x)$.

For $x, y \in V(X)$ with $\mathrm{dist}(x, y) = i$ define

$$\tau_i(x, y; \mathcal{C}) = |\{C \in \mathcal{C} \mid x \in C,\ d(y, C) = i - 1\}|. \tag{3.2}$$

Again, in Bang et al. [2007] it is shown that for a geometric distance-regular graph X the number $\tau_i(x, y; \mathcal{C})$ does not depend on the pair x, y, but only on the distance $\mathrm{dist}(x, y) = i$. Therefore, we may define $\tau_i := \tau_i(x, y; \mathcal{C})$.

Lemma 3.2.1 ([Bang et al., 2007, Proposition 4.2]). *Let X be a geometric distance-regular graph of diameter d, with smallest eigenvalue $-m$. Then*

1. $c_i = \tau_i \psi_{i-1}$, *for* $1 \leq i \leq d$;

2. $b_i = (m - \tau_i)\left(\frac{k}{m} + 1 - \psi_i\right)$, *for* $1 \leq i \leq d-1$.

Lemma 3.2.1 is of great importance to our analysis of geometric distance-regular graphs, as it gives a nice control of the intersection numbers in terms of τ_i and ψ_i. More specifically, we use this lemma in the proof of our characterizations of Hamming graphs in Sections 5.2 and 6.3. We also use vertex-clique intersection numbers τ_i and ψ_i in Chapter 10.

Lemma 3.2.2 (Koolen and Bang [2010])**.** *Let X be a geometric distance-regular graph of diameter $d \geq 2$. Then $\tau_2 \geq \psi_1$.*

Proof. Let $C \in \mathcal{C}$ be a Delsarte clique of X and let v be a vertex with $\mathrm{dist}(v, C) = 1$. Since C is a maximal clique, there exists a vertex $y \in C$ non-adjacent to v. Then $\mathrm{dist}(v, y) = 2$. Let $u_1, u_2, \ldots, u_{\psi_1}$ be the neighbors of v in C. Denote by $C_i \in \mathcal{C}$ a Delsarte clique that contains v and u_i for $i \in [\psi_1]$. Note that since C intersects each of C_i in at most one vertex, all C_i are distinct. Moreover, $\mathrm{dist}(y, C_i) = 1$, while $\mathrm{dist}(v, y) = 2$. Therefore, $\tau_2 \geq \psi_1$. □

Corollary 3.2.3. *Let X be a geometric distance-regular graph of diameter $d \geq 2$, with smallest eigenvalue $-m$. Then $\mu \leq m^2$.*

Proof. $\mu = \tau_2 \psi_1 \leq \tau_2^2 \leq m^2$. □

Lemma 3.2.4 ([Bang et al., 2007, Theorem 5.3])**.** *Let X be a geometric distance-regular graph of diameter $d \geq 2$.*

1. *If $\psi_1 = 1$, then for each vertex $v \in V(X)$ its neighborhood graph $X(v)$ is a disjoint union of m cliques, where $-m$ is the smallest eigenvalue of X.*

2. *If $\psi_1 \geq 2$, then for each vertex $v \in X$ its neighborhood graph $X(v)$ is connected.*

Thus, in particular, either each neighborhood graph of X is connected, or each neighborhood graph of X is disconnected.

Proof. Fix a Delsarte clique geometry $\mathcal{C}$ of X. Let $C_1, C_2, \ldots, C_m \in \mathcal{C}$ be the cliques which contain v. Take $w \in N(v)$ and let C_i be the clique which contains w. If $\psi_1 = 1$, then v is the only neighbor of w in C_j for $j \neq i$. Thus $X(v)$ is a disjoint union of m cliques, where by Lemma 3.1.7, $-m$ is the smallest eigenvalue of X. If $\psi_1 \geq 2$, then w is adjacent with at least one vertex in C_j distinct from u for every $j \neq i$. Thus, in this case, $X(v)$ is a connected graph. □

Theorem 3.2.5 (Bang et al. [2007])**.** *Let X be a geometric distance-regular graph of diameter $d \geq 2$. Assume $\psi_1 > 1$. Then $1 < \psi_1 < \psi_2 < \ldots < \psi_{d-1}$.*

3.3 Dual graphs

Let X be a distance-regular graph which has a clique geometry $\mathcal{C}$.

Definition 3.3.1. By a *dual graph of X (that corresponds to $\mathcal{C}$)* we mean the graph $\widetilde{X}$ with the vertex set $\mathcal{C}$, in which C_i and C_j are adjacent if and only if $|C_i \cap C_j| = 1$.

We will use dual graphs in the analysis of the elusive case $\mu = 1$ for geometric graphs in Section 8.2.

Lemma 3.3.2. *Let X be a geometric distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-m$. Then its dual graph $\widetilde{X}$ is an edge-regular graph of diameter $d-1$ with the vertex degree $\widetilde{k} = (m-1)\left(\frac{k}{m} + 1\right)$ and $\widetilde{\lambda} = (m-2) + (\psi_1 - 1)\frac{k}{m}$.*

Proof. It is known that every Delsarte clique is completely regular, with covering radius $d-1$, see [Godsil, 1993a, Lemma 7.2]. In particular, this implies that the diameter of $\widetilde{X}$ is $d-1$. Every clique in the Delsarte clique geometry $\mathcal{C}$ of X has size $1 + k/m$ and, by Lemma 3.1.7, every vertex is in precisely m cliques from $\mathcal{C}$. Since every pair of non-disjoint cliques intersects in precisely one vertex, we get $\widetilde{k} = (m-1)(1 + k/m)$.

Now, assume that C_1 and C_2 are distinct cliques in $\mathcal{C}$ that share a vertex v. Let $u \in C_1$ be a vertex distinct from v. Then u has ψ_1 neighbors in C_2. Let u' be one of such neighbors

distinct from v. Then the edge $\{u, u'\}$ belongs to a clique C which is distinct from C_1 and C_2 and intersects both of them. Thus, C is a common neighbor of C_1 and C_2. Note, that a common neighbor $C \in \mathcal{C}$ of C_1 and C_2, which does not contain v, is uniquely determined by $u \in C_1$ and its neighbor u' in C_2. Finally, note that every clique from $\mathcal{C}$ which contains v and is distinct from C_1 and C_2 is their common neighbor. Hence, $\widetilde{\lambda} = (m-2)+(\psi_1-1)k/m$. $\square$

Let A and $\widetilde{A}$ be the adjacency matrices of X and $\widetilde{X}$. Denote by $\mathrm{spec}(X)$ and $\mathrm{spec}(\widetilde{X})$ the sets of eigenvalues of A and $\widetilde{A}$, respectively.

Lemma 3.3.3. *Let X be a geometric distance-regular graph with smallest eigenvalue $-m$. If $k \geq m^2$, then* $\mathrm{spec}(\widetilde{X}) \subseteq \{\theta - \dfrac{k}{m} + m - 1 \mid \theta \in \mathrm{spec}(X)\}$.

Proof. Let $\mathcal{C}$ be a Delsarte clique geometry of X. Note that, by Lemma 3.1.7, every vertex of X belongs to precisely m cliques of $\mathcal{C}$. Define N to be an $n \times |\mathcal{C}|$ vertex-clique incidence matrix, i.e, $N_{i,j} = 1$ if the vertex v_i belongs to the clique C_j, and $N_{i,j} = 0$ otherwise. Then

$$A = NN^T - mI \qquad \text{and} \qquad \widetilde{A} = N^T N - \left(\frac{k}{m} + 1\right) I. \tag{3.3}$$

From linear algebra it is known that non-zero eigenvalues of NN^T and $N^T N$ coincide. Since $|\mathcal{C}|(1 + k/m) = nm$, we get $|\mathcal{C}| < n$, so 0 is an eigenvalue of NN^T. Therefore, $\mathrm{spec}(N^T N) \subseteq \mathrm{spec}(NN^T)$ and the statement of the lemma follows from Eq. (3.3). $\square$

3.4 Johnson graphs

Definition 3.4.1. Let $d \geq 2$ and Ω be a set of $s \geq 2d$ points. The *Johnson graph* $J(s, d)$ is a graph on the set $V(J(s, d)) = \binom{\Omega}{d}$ of $n = \binom{s}{d}$ vertices, where a pair of distinct vertices is adjacent if and only if the corresponding subsets $U_1, U_2 \subseteq \Omega$ differ by exactly one element, i.e., $|U_1 \setminus U_2| = |U_2 \setminus U_1| = 1$.

It is not hard to check that $J(s,d)$ is a distance-regular graph of diameter d with intersection numbers

$$b_i = (d-i)(s-d-i) \quad \text{and} \quad c_{i+1} = (i+1)^2, \quad \text{for } 0 \le i < d. \tag{3.4}$$

In particular, $J(s,d)$ is regular of degree $k = d(s-d)$ with $\lambda = s-2$ and $\mu = 4$. The eigenvalues of $J(s,d)$ are

$$\xi_j = (d-j)(s-d-j) - j \quad \text{with multiplicity} \quad \binom{s}{j} - \binom{s}{j-1}, \text{ for } 0 \le j \le d. \tag{3.5}$$

Using Lemma 3.2.1, it is easy to see that for the Johnson graph $J(s,d)$

$$\tau_i = i \quad \text{and} \quad \psi_{i-1} = i, \quad \text{for } 1 \le i \le d.$$

For $s \ge 2d+1$, the automorphism group of $J(s,d)$ is the induced symmetric group $S_s^{(d)}$, which acts on $\binom{\Omega}{d}$ via the induced action of S_s on Ω. Indeed, it is clear, that $S_s^{(d)} \le \mathrm{Aut}(J(s,d))$. The opposite inclusion can be derived from the Erdős-Ko-Rado theorem.

Thus, for a fixed d and $s \ge 2d+1$, the order is $|\mathrm{Aut}(J(s,d)| = s! = \Omega(\exp(n^{1/d}))$, the thickness satisfies $\theta(\mathrm{Aut}(J(s,d))) = s = \Omega(n^{1/d})$ and

$$\mathrm{motion}(J(s,t)) = O(n^{1-1/d}). \tag{3.6}$$

Theorem 3.4.2 (see [Brouwer et al., 1989, Theorem 9.1.3])**.** *Let X be a connected graph s.t.*

1. *for each vertex v of X the graph $X(v)$ is the line graph of $K_{s,t}$;*

2. *if $\mathrm{dist}(x,y) = 2$, then x and y have at most 4 common neighbors.*

Then X is a Johnson graph or is doubly covered by a Johnson graph. More precisely, in the latter case X is the quotient of the Johnson graph $J(2d,d)$ by an automorphism of the form

$\tau\omega$, where τ is the automorphism sending each d-set to its complement, and ω is an element of order at most 2 in $\mathrm{Aut}(X)$ with at least 8 fixed points.

3.5 Hamming graphs

Definition 3.5.1. Let Ω be a set of $s \geq 2$ points. The *Hamming graph* $H(d,s)$ is a graph on the set $V(H(d,s)) = \Omega^d$ of $n = s^d$ vertices, for which a pair of vertices is adjacent if and only if the corresponding d-tuples v_1, v_2 differ in precisely one position. In other words, v_1 and v_2 are adjacent if the Hamming distance $d_H(v_1, v_2)$ equals 1.

Again, it is not hard to check that $H(d,s)$ is a distance-regular graph of diameter d with intersection numbers

$$b_i = (d-i)(s-1) \quad \text{and} \quad c_{i+1} = i+1 \quad \text{for } 0 \leq i \leq d-1. \tag{3.7}$$

In particular, $H(d,s)$ is regular of degree $k = d(s-1)$ with $\lambda = s-2$ and $\mu = 2$. The eigenvalues of $H(d,s)$ are

$$\xi_j = d(s-1) - js \quad \text{with multiplicity} \quad \binom{d}{j}(s-1)^j, \text{ for } 0 \leq j \leq d. \tag{3.8}$$

Using Lemma 3.2.1 it is easy to see that for the Hamming graph $H(d,s)$

$$\tau_i = i \quad \text{and} \quad \psi_{i-1} = 1, \quad \text{for } 1 \leq i \leq d.$$

The automorphism group of $H(d,s)$ is isomorphic to the wreath product $S_s \wr S_d$. Hence, its order is $|\mathrm{Aut}(H(d,s))| = (s!)^d d!$, the thickness satisfies $\theta(H(d,s)) \geq s = n^{1/d}$ and

$$\mathrm{motion}(H(d,s)) \leq 2s^{d-1} = 2n^{1-1/d}. \tag{3.9}$$

In Section 5.2 we use the classification of distance-regular graphs with the same intersection array as Hamming graphs. In the case of diameter 2, the unique non-Hamming graph that has the intersection array of the Hamming graph is the *Shrikande graph*. It has 16 vertices and has the same parameters as $H(2,4)$.

Definition 3.5.2. The direct product of the Hamming graph $H(t,4)$ with $\ell \geq 1$ copies of the Shrikande graph is called a *Doob graph*.

One can check that the Doob graphs are distance-regular and have the same intersection numbers as the Hamming graph $H(t+2\ell,4)$. Yoshimi Egawa [1981] proved that the Doob graphs are the only graphs with this property.

Theorem 3.5.3 (Egawa [1981], see [Brouwer et al., 1989, Corollary 9.2.5])**.** *A distance-regular graph of diameter d with intersection numbers given by Eq.* (3.7) *is a Hamming graph or a Doob graph.*

3.6 Grassmann graphs

Definition 3.6.1. Let $\mathbb{F}_q$ be a finite field. For $2 \leq d \leq 2s$ we define a graph $J_q(s,d)$ whose vertices are d-dimensional subspaces of $\mathbb{F}_q^s$ over $\mathbb{F}_q$, such that a pair of vertices is adjacent if and only if the intersection of the corresponding subspaces is a subspace of dimension $d-1$. The graph $J_q(s,d)$ is called the *Grassmann* graph.

Define

$$\binom{n}{d}_q = \frac{(q^n-1)(q^{n-1}-1)\dots(q^{n-d+1}-1)}{(q^d-1)(q^{d-1}-1)\dots(q-1)} \quad \text{and} \quad [n]_q = \binom{n}{1}_q = q^{n-1}\dots+1. \tag{3.10}$$

The parameters of the Grassmann graph are:

$$\begin{gathered} b_i = q^{2i+1}[d-i]_q[s-d-i]_q = \left([d]_q - [i]_q\right)\left(q[s-d]_q + 1 - [i+1]_q\right), \quad \text{and} \\ c_i = \left([i]_q\right)^2 \end{gathered} \tag{3.11}$$

This implies,

$$k = q[d]_q[s-d]_q \quad \text{and} \quad k_j = q^{j^2}\binom{d}{i}_q\binom{s-d}{i}_q, \tag{3.12}$$
$$\mu = (q+1)^2, \quad \lambda = q[s-d]_q + q[d]_q + q,$$

Using Lemma 3.2.1 it is easy to see that for the Grassmann graph $J_q(s,d)$

$$m = [d]_q, \quad \tau_i = [i]_q \quad \text{and} \quad \psi_{i-1} = [i]_q, \quad \text{for } 1 \le i \le d.$$

Theorem 3.6.2 (Ray-Chaudhuri and Sprague [1976], see also [Brouwer et al., 1989, Thm. 9.3.9])**.** *Let X be a geometric distance-regular graph with the smallest eigenvalue $-m$. Assume that*

(i) each clique has at least q^2+q+1 vertices, i.e., $k/m+1 \ge q^2+q+1$,

(ii) each vertex belongs to $> q+1$ cliques, i.e., $m > q+1$,

(iii) $\psi_1 = \tau_2 = q+1$.

Then X is a Grassmann graph $J_q(s,d)$ for some s.

CHAPTER 4

CAMERON'S CLASSIFICATION AND COMBINATORIAL APPROACH

4.1 O'Nan-Scott theorem. Cameron classification

Let Ω be a finite set of points, and let $G \leq \mathrm{Sym}(\Omega)$ be a permutation group on Ω.

We start by stating the O'Nan-Scott theorem, one of the central results in the theory of primitive permutation groups, as it brings the Classification of Finite Simple Groups into play.

Definition 4.1.1. Let $G, G_1, \ldots, G_t$ be isomorphic groups and $\phi_i : G \to G_i$ be isomorphisms for $i \in [t]$. The diagonal subgroup of $S_1 \times S_2 \times \ldots \times S_d$ is the group $D = \{(\phi_1(s), \phi_2(s), \ldots, \phi_t(s)) \mid s \in S\}$.

Theorem 4.1.2 (Scott [1980], Aschbacher and Scott [1985], see [Cameron, 1981, Thm. 4.1])**.** *Let G be a primitive permutation group on the set Ω. Let n denote the degree of G and N denote its socle (see Def. 2.2.4). Then one of the following holds:*

1. *N is elementary abelian of order p^d, where p is prime and $d \geq 1$.*

2. *$N = T_1 \times T_2 \times \ldots \times T_m$, where $T_1, T_2, \ldots, T_m$ are isomorphic to a fixed simple group T. Moreover, either*

   (a) T is a socle of a primitive group G_0 of degree n_0, and $G \leq G_0 \wr S_m$ (with the product action), where $n = n_0^m$; or

   (b) $N \cap G_\alpha = D_1 \times D_2 \times \ldots \times D_\ell$, where G_α is a stabilizer of a point $\alpha \in \Omega$, $m = k\ell$ for some k, D_i is the diagonal supgroup of $T_{(i-1)k+1} \times \ldots \times T_{ik}$ and $n = |T|^{(k-1)\ell}$.

Relying on the O'Nan-Scott theorem, and on the CFSG through "Schreier's Hypothesis" Cameron [1981] gave a classification of primitive permutation groups of large order.

Theorem 4.1.3 ("Shreier's Hypothesis", depends on CFSG)**.** *A non-abelian finite simple group is an alternating group, a group of Lie type, or one of finitely many sporadic groups.*

Theorem 4.1.4 (Cameron [1981])**.** *There exists a constant $c > 0$ such that for every primitive permutation group G of degree n one of the following holds.*

1. *G has an elementary abelian regular normal subgroup.*
2. *$T^\ell \leq G \leq \mathrm{Aut}(T) \wr S_\ell$, where T is either an alternating group acting on k-element subsets, or a classical simple group acting on an orbit of subspaces or pairs of subspaces of complementary dimension (in the case $PSL(d, q)$), and the wreath product has the product action.*
3. *$|G| \leq n^{c \log \log n}$.*

Remark 4.1.5 (Cameron [1981])**.** Groups under 1 and those under 2 for which T is not an alternating group satisfy $|G| \leq n^{c' \log n}$ for some $c' > 0$.

Hence, one can immediately get the following corollary.

Theorem 4.1.6 (Cameron [1981])**.** *There exists a constant $c > 0$ such that for every primitive permutation group G of degree n either*

1. *$\left(A_k^{(t)}\right)^\ell \leq G \leq S_k^{(t)} \wr S_\ell$ (with the product action), for $k \geq 7$, $n = \binom{k}{t}^\ell$, or*
2. *$|G| \leq n^{c \log n}$.*

Definition 4.1.7. We say that a permutation group G in a *Cameron group*, if

$$\left(A_k^{(t)}\right)^\ell \leq G \leq S_k^{(t)} \wr S_\ell \quad \text{(with the product action) for some } k, t, \ell. \tag{4.1}$$

4.2 Permutation groups with small minimal degree

As we discussed in Introduction, the minimal degree is one of the key parameters of permutation groups, study of which goes back to Jordan [1871]. Furthermore, as we discuss below, by Wielandt's result, lower bounds on the minimal degree imply strong structural constraints on the group.

Using the Classification of Finite Simple Groups, Liebeck and Saxl [1991] gave the following classification of primitive permutation groups based on their minimal degree.

Theorem 4.2.1 (Liebeck and Saxl [1991])**.** *Let G be a primitive permutation group on a set Ω of size n. Then one of the following holds.*

1. $\mathrm{mindeg}(G) \geq n/2$.

2. $A_m^r \lhd G \leq S_m \wr S_r$, *where $m \geq 5$ and the wreath product acts on $\Omega = \Delta^r$ and S_m acts on Δ. Moreover, either $\Delta = \binom{[m]}{k}$ is the set of k-subsets of $[m]$ and $n = \binom{m}{k}$, for $k < m/2$, or $m = |\Delta| = 6$ and $n = 6^r$.*

3. $L^r \lhd G \leq L_1 \wr S_r$, *where L is a simple group of Lie type over $\mathbb{F}_2$, L is a socle of L_1 and $\Omega = \Delta^r$ for some set Δ on which L_1 acts primitively. In this case,* $\mathrm{mindeg}(G) \geq n/3$.

As an immediate corollary one gets the following claim.

Theorem 4.2.2 (Liebeck and Saxl [1991])**.** *If G is a primitive permutation group of degree n, then one of the following is true.*

1. *G is a Cameron group.*

2. $\mathrm{mindeg}(G) \geq n/3$.

The next lemma shows that in a certain range of parameters Cameron groups have sublinear minimal degree.

Lemma 4.2.3. *Let G be a Cameron group with $(A_m^{(k)})^d \leq G \leq S_m^{(k)} \wr S_d$ which acts on $n = \binom{m}{k}^d$ points, where $k \leq m/2$. Then as $m \to \infty$, the following holds uniformly in d: we have* $\mathrm{mindeg}(G) = o(n)$ *if and only if $k = o(m)$.*

Proof. It is not hard to see that the minimal degree of G is realized by the induced action of a cycle of length 2 or 3 (in S_m or A_m, respectively) on k-subsets in just one of the d coordinates. If there is a 2-cycle action in a coordinate, then the minimal degree of G is

$$\left(\binom{m}{k} - \binom{m-2}{k} - \binom{m-2}{k-2}\right)\binom{m}{k}^{d-1},$$

otherwise, the minimal degree of G is

$$\left(\binom{m}{k} - \binom{m-3}{k} - \binom{m-3}{k-3}\right)\binom{m}{k}^{d-1}.$$

As $m \to \infty$ these expressions are equal to

$$n \cdot \left(1 - \frac{(m-k)^2 + k^2}{m^2} + o(1)\right) \quad \text{and} \quad n \cdot \left(1 - \frac{(m-k)^3 + k^3}{m^3} + o(1)\right),$$

respectively. Clearly, each of these expressions is $o(n)$ if and only if $k = o(m)$. □

4.3 Wielandt's upper bound for thickness

The *thickness* $\theta(G)$ of a group G is the greatest t for which the alternating group A_t is involved as a quotient group of a subgroup of G (the term was coined in Babai [2014]). Babai et al. [1982] proved that primitive permutation groups with bounded thickness have polynomially bounded order.

Wielandt [1934] proved that a linear lower bound for the minimal degree of a permutation group implies a logarithmic upper bound for the thickness of the group.

Theorem 4.3.1 (Wielandt [1934], see [Babai, 1982, Theorem 6.1])**.** *Let $n > k > \ell$ be positive integers, $k \geq 7$, and let $0 < \alpha < 1$. Suppose that G is a permutation group of degree n and minimal degree at least αn. If*

$$\ell(\ell-1)(\ell-2) \geq (1-\alpha)k(k-1)(k-2),$$

and $\theta(G) \geq k$, then $n \geq \binom{k}{\ell}$.

Corollary 4.3.2. *Let G be a permutation group of degree n. Suppose* $\mathrm{mindeg}(G) \geq \alpha n$. *Then the thickness $\theta(G)$ of G satisfies $\theta(G) \leq \dfrac{3}{\alpha}\ln(n)$.*

4.4 Motion of distance-transitive graphs

As an application of the Liebeck-Saxl classification, in this section we prove the following result on motion of distance-transitive graphs.

Theorem 4.4.1 (Babai and Kivva [2020])**.** *Let X be a primitive distance-transitive graph. Assume that $G \leq \mathrm{Aut}(X)$ acts on X distance-transitively. Then*

$$\mathrm{mindeg}(G) \geq n/3,$$

or X is a Johnson graph, a Hamming graph, a complement to $J(m,2)$, a complement to $H(2,m)$, the Sylvester graph or the line graph of Tutte's 8-cage.

It is easy to see that a group acting distance-transitively on a primitive distance-regular graph is primitive.

Definition 4.4.2. Let $\mathfrak{X} = (V, \mathfrak{c})$ be a coherent configuration. We say that G acts color-transitively on $\mathfrak{X}$ if for every $x, y, x', y' \in V$ with $\mathfrak{c}(x,y) = \mathfrak{c}(x',y')$ there exists $g \in G$ such that $g(x) = x'$ and $g(y) = y'$.

Lemma 4.4.3 (D. Higman)**.** *Let $\mathfrak{X} = (V, \mathfrak{c})$ be a primitive coherent configuration. Assume that group $G \leq \mathrm{Aut}(\mathfrak{X})$ acts color-transitively. Then, G is a primitive group.*

Proof. Let R be a non-diagonal orbital of G and let $(u, v) \in R$. Let $i = \mathfrak{c}(u, v)$. Since $\mathfrak{X}$ is primitive, between every pair of vertices $x, y \in V$ there is a path in X_i. The group G acts color-transitively, so each edge of such path is in R. Therefore, G is primitive. □

Corollary 4.4.4. *Let X be a primitive distance-regular graph. Assume that group $G \leq \mathrm{Aut}(X)$ acts distance-transitively. Then, G is a primitive group.*

Therefore, by the Liebeck-Saxl classification, if the group G acts distance-transitively on a primitive distance-regular graph and has $\mathrm{mindeg}(G) < n/3$, then G should be one of the groups described in case 2 of Theorem 4.2.1.

The distance-transitive graphs with the automorphism group described in case 2 of the Liebeck-Saxl theorem are known via the following results of Praeger et al. [1987] and Liebeck et al. [1987].

Definition 4.4.5. A group G is called *almost simple* if there exists a simple non-abelian group such that $S \leq G \leq \mathrm{Aut}(S)$.

Theorem 4.4.6 (Praeger, Saxl, Yokoyama [1987])**.** *Let X be a primitive distance-transitive graph with $d \geq 2$ and $k \geq 3$. Assume that G acts on X distance-transitively. Then one of the following is true.*

1. *X is a Hamming graph $H(d, m)$, or a complement to $H(2, m)$.*

2. *G is almost simple.*

3. *G is affine.*

Theorem 4.4.7 (Liebeck, Praeger, Saxl [1987])**.** *Let X be a distance-transitive graph with $G = \mathrm{Aut}(X)$. Assume that $A_m \triangleleft G \leq \mathrm{Aut}(A_m)$, for $m \geq 5$. Then $G = S_m$ and X is a*

Johnson graph $J(m,k)$, a complement to $J(m,2)$, an Odd graph, a folded Johnson graph $\overline{J}(2k,k)$ or $G = \mathrm{Aut}(A_6)$ and X is a Sylvester graph or a line graph of Tutte's 8-cage.

Lemma 4.4.8. *If X is an Odd graph $O(k)$ or a folded Johnson graph $\overline{J}(2k,k)$, then*

$$\mathrm{motion}(X) \geq |V(X)|/3.$$

Proof. If X is an Odd graph $O(k)$, it has $n = \binom{2k-1}{k-1}$ vertices. It is not hard to check that every non-identity automorphism of $O(k)$ moves at least $2\binom{2k-3}{k-2}$ vertices. Therefore,

$$\mathrm{motion}(X) \geq 2\frac{k(k-1)}{(2k-1)(2k-2)}n = \frac{k}{2k-1}n \geq \frac{1}{2}n.$$

If X is a folded Johnson graph $\overline{J}(2k,k)$, then it has $n = \frac{1}{2}\binom{2k}{k}$ vertices. It is not hard to check that every non-identity automorphism of $\overline{J}(2k,k)$ moves at least $\binom{2k-2}{k-1}$ vertices. Therefore,

$$\mathrm{motion}(X) \geq \frac{2k^2}{2k(2k-1)}n = \frac{k}{2k-1}n \geq \frac{1}{2}n. \qquad \square$$

Before, giving a proof of Theorem 4.4.1 we make the following simple observation.

Observation 4.4.9. *Let $H \leq G \leq \mathrm{Sym}(\Omega)$. Then* $\mathrm{mindeg}(H) \geq \mathrm{mindeg}(G)$.

Proof of Theorem 4.4.1. Assume that X is not complete. By Observation 4.4.9, it is sufficient to prove the theorem for $G = \mathrm{Aut}(X)$. By Corollary 4.4.4, the group G is primitive. Assume, $\mathrm{mindeg}(G) < n/3$. Then G is one of the groups described in case 2 of Theorem 4.2.1. Therefore, by Theorem 4.4.6, either X is a Hamming graph, or a complement to $H(2,m)$, or $A_m \triangleleft \mathrm{Aut}(X) \leq \mathrm{Aut}(A_m)$. Hence, in the view of Lemma 4.4.8, Theorem 4.4.7 implies that X is a Johnson graph, or a complemet to $J(m,2)$, or the Sylvester graph or the line graph of Tutte's 8-cage. $\square$

4.5 Combinatorial tools to bound the order and the minimal degree of automorphism groups

4.5.1 Distinguishing numbers

In a seminal paper, Babai [1981] introduced a combinatorial technique to bound the order of primitive permutation groups.

Definition 4.5.1 (Babai [1981])**.** In a configuration $\mathfrak{X} = (\Omega, \mathfrak{c})$ a pair of vertices $u, v \in \Omega$ is *distinguished* by a vertex $x \in \Omega$ if the colors $\mathfrak{c}(x, u)$ and $\mathfrak{c}(x, v)$ are distinct. Define

$$D(u, v) = |\{x \mid u,\ v \text{ are distinguished by } x\}|\,.$$

It is easy to see that for a homogeneous coherent configuration $\mathfrak{X}$, the number $D(u, v)$ of vertices which distinguish u and v depends only on the color i between u and v. So one can define $D(i) = D(u, v)$. Babai [1981] showed the following relation for distinguishing numbers of two different colors.

Lemma 4.5.2 ([Babai, 1981, Proposition 6.4])**.** *Let $\mathfrak{X}$ be a homogeneous coherent configuration of rank r. Then for all colors $1 \leq i, j \leq r - 1$ the inequality $D(j) \leq \mathrm{dist}_i(j) D(i)$ holds.*

Proof. Note that for all vertices u, v, w we have $D(u, v) \leq D(v, w) + D(w, u)$. If $\mathrm{dist}_i(j)$ is finite, then the statement follows from this triangle inequality. If $\mathrm{dist}_i(j)$ is infinite, the statement is trivial. □

Definition 4.5.3. Define the *minimal distinguishing number* $D_{\min}(\mathfrak{X})$ of the configuration $\mathfrak{X} = (V, c)$ to be

$$D_{\min}(\mathfrak{X}) = \min_{u \neq v \in V} D(u, v).$$

4.5.2 Order

Babai showed that the minimal distinguishing number of $\mathfrak{X}$ can be used to bound the order of the automorphism group of $\mathfrak{X}$.

Definition 4.5.4. A set S of vertices of a configuration $\mathfrak{X}$ is *distinguishing* if every pair of distinct vertices in $\mathfrak{X}$ is distinguished by at least one element of S.

Note that the pointwise stabilizer of a distinguishing set is trivial. Thus, if S is a distinguishing set of $\mathfrak{X} = (V, \mathfrak{c})$, then $|\operatorname{Aut}(\mathfrak{X})| \leq n^{|S|}$, where $|V| = n$.

Lemma 4.5.5 ([Babai, 1981, Lemma 5.4])**.** *Let $\mathfrak{X}$ be a primitive coherent configuration. Then there exists a distinguishing set of size at most $(2n \log n / D_{\min}(\mathfrak{X})) + 1$. Moreover, we have $|\operatorname{Aut}(\mathfrak{X})| \leq n^{3+2n(\log(n)-5/9)/D_{\min}(\mathfrak{X})}$.*

In the same paper, he proved that for non-trivial primitive coherent configurations the minimal distinguishing number is at least $\Omega(\sqrt{n})$.

Theorem 4.5.6 (Babai [1981])**.** *Let $\mathfrak{X}$ be a non-trivial primitive coherent configuration on n vertices. Then $D_{\min}(\mathfrak{X}) > (\sqrt{n} - 1)/2$.*

Combining these two results, one immediately gets the following corollary.

Theorem 4.5.7 (Babai [1981])**.** *Let $\mathfrak{X}$ be a non-trivial primitive coherent configuration on n vertices. Then $|\operatorname{Aut}(\mathfrak{X})| \leq n^{4\sqrt{n}\log n}$.*

As a byproduct, Babai resolved a then 100-year-old problem on primitive but not doubly transitive permutation groups.

Theorem 4.5.8 (Babai [1981])**.** *Let G be a primitive but not doubly transitive permutation group of degree n. Then $|G| \leq n^{4\sqrt{n}\log n}$.*

Note that Theorems 4.5.7, 4.5.8 are tight up to a logarithmic factor, as automorphism groups of the Johnson schemes $\mathfrak{J}(s, 2)$ and Hamming schemes $\mathfrak{H}(2, s)$ show.

After more than 30 years, Sun and Wilmes [2015a] overcame the first layer of exceptions and proved the following result, which again is tight up to a logarithmic factor.

Theorem 4.5.9 (Sun and Wilmes [2015a])**.** *Let $\mathfrak{X}$ be a non-trivial primitive coherent configuration on n vertices. If $\mathfrak{X}$ is not a Johnson scheme $\mathfrak{J}(s,2)$ or a Hamming schemes $\mathfrak{H}(2,s)$ for some s, then* $|\operatorname{Aut}(\mathfrak{X})| \leq \exp\left(O\left(n^{1/3}\log^{7/3} n\right)\right)$.

4.5.3 Minimal degree. Spectral tool

Babai [2014] developed a combination of combinatorial and spectral tool to prove lower bounds on the motion of coherent configurations. The combinatorial tool uses the minimal distinguishing number.

Observation 4.5.10. *Let $\mathfrak{X}$ be a configuration with n vertices. Then*

$$\operatorname{motion}(\mathfrak{X}) \geq D_{\min}(\mathfrak{X}).$$

Proof. Indeed, let $\sigma \in \operatorname{Aut}(\mathfrak{X})$ be any non-trivial automorphism of $\mathfrak{X}$. Let u be a vertex not fixed by σ. No fixed point of σ distinguishes u and $\sigma(u)$, so the degree of σ is at least $D(u,\sigma(u)) \geq D_{\min}(\mathfrak{X})$. □

For a k-regular graph X, let $k = \xi_1 \geq \xi_2 \geq \ldots \geq \xi_n$ denote the eigenvalues of the adjacency matrix of X. We call $\xi = \xi(X) = \max\{|\xi_i| : 2 \leq i \leq n\}$ the *zero-weight spectral radius* of X. The second tool is based on the Expander Mixing Lemma.

Lemma 4.5.11 ([Babai, 2014, Proposition 12])**.** *Let X be a regular graph of degree k on n vertices with the zero-weight spectral radius ξ. Suppose every pair of distinct vertices in X has at most q common neighbors. Then*

$$\operatorname{motion}(X) \geq n \cdot \frac{(k-\xi-q)}{k}.$$

Note that this spectral tool gives a trivial bound for bipartite graphs, as $\xi(X) = k$ for a k-regular bipartite graph X. We prove a bipartite version of this lemma in Theorem 8.4.3.

Using this pair of tools Babai proved a linear lower bound on the minimal degree of strongly regular graphs with known exceptions (see Theorem 1.2.5). Using the same pair of tools we extended his result to primitive coherent configurations of rank 4 (Theorem 1.2.8) and distance-regular graphs of bounded diameter (Theorem 1.2.6).

CHAPTER 5

CLASSIFICATION OF GRAPHS WITH BOUNDED SMALLEST EIGENVALUE

5.1 Prior work and our contribution

Hoffman [1967] observed that there is a relation between conbinatorial properties of a graph and spectral properties of its adjacency matrix. He initiated the program of classifying graphs and graph properties that can be characterized by spectrum. A particularly interesting direction that received a lot of attention is the program of classifying graphs by the least eigenvalue of their adjacency matrix.

Many interesting graphs with certain degree of regularity are constructed from geometric objects or using geometric intuition. These include incidence graphs of partial linear spaces, line graphs, Grassmann graphs, Hamming graphs, Johnson graphs, etc. Vaguely speaking, the absolute value of the smallest eigenvalue for many of these families is "small." Moreover, a long line of work, which we briefly discuss in this chapter, essentially shows that if a "sufficiently regular" graph has the smallest eigenvalue of a "sufficiently small" absolute value, then it comes from a certain "geometric family".

For instance, if a connected graph is regular and the smallest eigenvalue $\theta_{\min}$ is greater than -2, one gets the following complete characterization.

Theorem 5.1.1 (Doob and Cvetković [1979], see [Brouwer et al., 1989, Corollary 3.12.3])**.** *Let X be a connected regular graph with smallest eigenvalue > -2. Then X is a complete graph, or X is an odd polygon.*

Remark 5.1.2. Doob and Cvetković [1979] gave complete characterization of all graphs (not necessarily regular) with smallest eigenvalue > -2.

Remark 5.1.3. Note that the complete graph K_n is the line graph of the complete bipartite graph $K_{n,1}$, and the cycle C_n is the line graph of itself.

5.1.1 Classification of graphs with smallest eigenvalue -2

If one allows the smallest eigenvalue to be ≥ -2, even within regular graphs there are much more examples than in Theorem 5.1.1. In particular, for every graph, its line graph has smallest eigenvalue ≥ -2.

Seidel [1968] classified strongly regular graphs with smallest eigenvalue -2.

Theorem 5.1.4 (Seidel [1968])**.** *Let X be a strongly regular graph with smallest eigenvalue -2. Then X is either the Johnson graph $J(s,2)$ for $s \geq 4$, the Hamming graph $H(2,s)$ for $s \geq 2$, a complete multipartite graph $K_{n\times 2}$, or a X has at most 28 vertices.*

Remark 5.1.5. We note that the Johnson graph $J(s,2)$ is the line graph of K_s. $J(s,2)$ is also known as the *triangular graph* $T(s)$; and the Hamming graph $H(2,s)$ is the line graph of $K_{s,s}$, and is also known as the *lattice graph* $L_2(s)$.

Remark 5.1.6. Seidel [1968] also showed that the graphs with ≤ 28 vertices that are not explicitly mentioned in Theorem 5.1.4 are the Peterson, Clebsch, Schläfli, Shrikande, or Chang graphs.

Hoffman [1970b, 1977] gave a characterization of graphs whose least eigenvalue is ≥ -2 dropping all regularity assumptions.

Definition 5.1.7. For an integer $a \geq 2$, a *cocktail-party graph* $CP(a)$ is a complete graph on $2a$ vertices with one perfect matching deleted.

Definition 5.1.8. Let Y be a graph whose vertex set is $[v]$, and let $a_1, a_2, \ldots, a_v$ be a set of non-negative integers. A generalized line graph $L(Y; a_1, a_2, \ldots, a_v)$ is a graph obtained from the disjoint union of the line graph $L(Y)$ and the cocktail-party graphs

$CP(a_1), CP(a_2), \ldots, CP(a_n)$ by adding edges between vertex $\{i, j\}$ in $L(Y)$ and vertices of $CP(a_i)$ and $CP(a_j)$ for every distinct $i, j \in [v]$.

Theorem 5.1.9 (Hoffman [1970b, 1977]). *Let X be a connected graph with $n \geq 37$ vertices and smallest eigenvalue ≥ -2. Then X is a generalized line graph.*

Alternative simplified proof of this result was given by Cameron et al. [1991], see [Brouwer et al., 1989, Theorem 3.12.2]. Their results also imply the following refined classification in the case of regular graphs.

Theorem 5.1.10 (Cameron et al. [1991], see [Brouwer et al., 1989, Theorem 3.12.2]). *Let X be a connected regular graph with n vertices of degree k and smallest eigenvalue ≥ -2. Then one of the following holds.*

1. *X is the line graph of a regular or a bipartite semiregular connected graph.*
2. *$n = 2(k+2) \leq 28$ and X is a subgraph of $E_7(1)$.*
3. *$n = (3/2)(k+2) \leq 27$ and X is a subgraph of the Schläfli graph.*
4. *$n = (4/3)(k+2) \leq 16$ and X is a subgraph os the Clebsch graph.*
5. *$n = k+2$ and X is a complete s-partite graph $K_{s\times 2}$ for some $s \geq 3$*

We also note the following generalization of Seidel's classification for edge-regular and co-edge-regular graphs.

Theorem 5.1.11 ([Brouwer et al., 1989, Theorem 3.12.4]). *Let X be a connected regular graph on n vertices with smallest eigenvalue -2.*

(i) If X is edge-regular, then X is strongly regular or the line graph of a regular triangle-free graph.

(ii) If X is co-edge-regular, then X is strongly regular, an $m_1 \times m_2$-grid, or one of the two regular subgraphs of the Clebsh graph with 8 and 12 vertices, respectively.

5.1.2 Gap in possible values of the smallest eigenvalue

Another important discovery made by Hoffman [1977] is that in fact one can also say a lot about graphs whose least eigenvalue is in the range $(-1-\sqrt{2}, -2)$.

Definition 5.1.12. Define ϑ_k to be the supremum of the smallest eigenvalues of graphs with minimal degree k and smallest eigenvalue < -2.

Theorem 5.1.13 (Hoffman [1970a], see [Brouwer et al., 1989, Theorem 3.12.5])**.** *The sequence $(\vartheta_k)_k$ forms a monotone decreasing sequence with limit $-1-\sqrt{2}$.*

Theorem 5.1.14 (Bussemaker and Neumaier [1992], see [Brouwer et al., 1989, Theorem 3.12.5])**.** *ϑ_1 (≈ -2.006594) is the smallest root of the equation*

$$\theta^2(\theta^2-1)^2(\theta^2-3)(\theta^2-4) = 1.$$

These results immediately imply that for $\delta \approx -0.006594$ the smallest eigenvalue of a graph is never in the interval $(-2-\delta, -2)$. This will be a crucial ingredient in our characterization of Johnson graphs (see Sec 6.2).

5.1.3 Strongly regular graphs with fixed smallest eigenvalue

Much less is known about general graphs whose least eigenvalue is $< -1-\sqrt{2}$. However, strong classifications were established for special classes of graphs, such as strongly regular and distance-regular graphs.

Extending Seidel's classification, Neumaier classified strongly regular graphs with fixed smallest eigenvalue $-m$ (which was also established in an unplushid work of Sims [1968]).

Prior to stating Neumaier's classification we define the families of graphs that appear in his classification.

Definition 5.1.15. A *Steiner system* $S(2, m, v)$ is a collection of m-element subsets of $[v]$, called *blocks*, such that every pair of distinct elements from $[v]$ is contained in exactly one block.

Definition 5.1.16. For a Steiner system $S(2, m, v)$ define a graph whose vertices are blocks of the Steiner system and two blocks are adjacent if they share an element. Such a graph is called a *Steiner graph*.

It is not hard to verify that the Steiner graph constructed from $S(2, m, v)$ is a geometric strongly regular graph with smallest eigenvalue $-m$.

Definition 5.1.17. An *orthogonal array* $OA(v, m)$ with parameters m and v is an array of size $m \times v^2$ with entries from $[v]$, such that the v^2 ordered pairs in every pair of distinct rows are all different.

Definition 5.1.18. For an orthogonal array $OA(v, m)$ define a graph whose vertices are columns of the array and two columns are adjacent if they have the same entry in exactly one position. We call such a graph a *Latin square graph*.

It is not hard to verify that the Latin square graph constructed from $OA(v, m)$ is a geometric strongly regular graph with smallest eigenvalue $-m$.

Theorem 5.1.19 (Neumaier [1979])**.** *Let $m \geq 0$. Let X be a strongly regular graph with smallest eigenvalue $-m$, then X is one of the following*

1. *a Steiner graph defined by $S(2, m, s)$;*
2. *a Latin square graph defined by $OA(s, m)$;*
3. *a complete multipartite graph $K_{s \times m}$*
4. *a conference graph;*
5. *a union of disjoint cliques;*
6. *one of finitely many exceptions.*

5.1.4 Case of distance-regular graphs

In the case of distance-regular graphs, all but finitely many graphs with bounded smallest eigenvalue are geometric (see Sec. 3).

Theorem 5.1.20 (Koolen and Bang [2010])**.** *Fix an integer $m \geq 2$. Then there are only finitely many coconnected non-geometric distance-regular graphs with smallest eigenvalue at least $-m$, and intersection number $\mu \geq 2$.*

As pointed out in the survey by [van Dam et al., 2016, below Thm. 9.10], one can remove condition $\mu \geq 2$ if one imposes $k \geq 3$ and $d \geq 3$ instead. This follows from the Bannai-Ito conjecture, confirmed by Bang et al. [2015].

Theorem 5.1.21 (Bannai-Ito conjecture, Bang et al. [2015])**.** *For every fixed $k \geq 3$ there are finitely many connected distance-regular graphs with degree k.*

Indeed, when $\mu = 1$, $k \geq 3$, $d \geq 3$ and graph is not geometric, the degree is bounded by $O(m^4)$, see [van Dam et al., 2016, below Thm. 9.10]. So there are only finitely many such graphs for a fixed m.

Moreover, Koolen and Bang conjectured that in fact all but finitely many geometric distance-regular graphs with bounded smallest eigenvalue are known.

Conjecture 5.1.22 (Koolen and Bang [2010])**.** *For a fixed integer $m \geq 2$, every geometric distance-regular graph with smallest eigenvalue $-m$, diameter ≥ 3 and $\mu \geq 2$ is a Johnson graph, or a Hamming graph, or a Grassmann graph, or a bilinear forms graph, or the number of vertices is bounded by a function of m.*

If confirmed, this conjecture combined with Theorem 5.1.20, classifies all infinite families of distance-regular graphs with bounded smallest eigenvalue and can be seen as an extension of Neumaier's classification (Theorem 5.1.19). Bang and Koolen confirmed their conjecture in the case $m = 3$.

Theorem 5.1.23 (Bang [2013], Bang and Koolen [2014]). *Let X be a geometric distance-regular graph of diameter $d \geq 3$ with smallest eigenvalue -3 and $\mu \geq 2$. Then X is one of the following:*

- *The Hamming graph $H(3, s)$ for $s \geq 3$.*
- *The Johnson graph $J(s, 3)$ for $s \geq 6$.*
- *The collinearity graph of the generalized quadrangle of order $(s, 3)$ deleting the edges in a spread, where $s \in \{3, 5\}$.*

We also note that in fact, for geometric distance-regular graphs of diameter d under mild assumptions the diameter and the smallest eigenvalue $-m$ satisfy $d \leq m$.

Theorem 5.1.24 (Bang [2018]). *Suppose that X is a geometric distance-regular graph with diameter $d \geq 3$ and smallest eigenvalue $-m$. If X contains an induced $K_{2,1,1}$, then $d \leq m$.*

Moreover, if $d \geq \max(3, m-1)$, then X is a Johnson graph.

5.1.5 Our contribution

We characterize Hamming graphs as distance-regular graphs of diameter d with smallest eigenvalue $-d$, $\mu \leq 3$, an induced quadrangle and sufficiently large degree k.

Theorem 5.1.25. *Let X be a distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-d$. Suppose that X contains an induced quadrangle, $\mu \leq 3$, and $k \geq \left(100d^3 \ln d\right) \cdot c_d$.*

Then X is the Hamming graph $H(d, k/d + 1)$.

Remark 5.1.26. A distance-regular graph that does not contain an induced quadrangle is called a *Terwilliger graph*. Such distance-regular graphs were studied, e.g., in Terwilliger [1985], Gavrilyuk et al. [2008], Gavrilyuk [2010].

This characterization also plays a crucial role in our proof of the robustness under extension for Hamming schemes (part 2 of Theorem 1.3.6, see Section 10.5).

5.2 Characterization of Hamming graphs by smallest eigenvalue

We note that for a sufficiently large k the assumption that X contains an induced quadrangle implies that X is geometric.

Observation 5.2.1. Let X be a distance-regular graph of diameter $d \geq 3$ with smallest eigenvalue $-d$. If $k \geq 5d^3 c_d$ and X contains an induced quadrangle, then X is geometric.

Proof. The Terwilliger inequality (Theorem 2.4.5) implies that $\lambda+2 \geq (k+c_d)/d$. Therefore, the inequality $k \geq 5d^3 c_d$ implies $2d(\lambda+1) \geq k$ and $\lambda \geq d(2d+1)c_d \geq d(2d+1)\mu$. Hence, by Corollary 3.1.6, X is geometric. □

Hence, Theorem 5.1.25 may be reformulated in the following equivalent way.

Theorem 5.2.2. *Let X be a geometric distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-d$. Suppose that $2 \leq \mu < 4$, and $k \geq \left(100d^3 \ln d\right) \cdot c_d$.*

Then X is the Hamming graph $H(d, k/d+1)$.

The proof of Theorem 5.2.2 (and of Theorem 1.4.5 in Sec. 6) uses the following result of Terwilliger.

Theorem 5.2.3 (Terwilliger [1986])**.** *Let X be a distance-regular graph of diameter $d \geq 2$. Assume that the second largest eigenvalue θ_1 has multiplicity $f_1 < k$. Then each local graph $X(v)$ has eigenvalue $-1-b^+$ with multiplicity at least $k-f_1$, where $b^+ = \dfrac{b_1}{\theta_1+1}$.*

We prove that if X is not a Hamming graph, then the assumptions of Theorem 5.2.2 imply that the second largest eigenvalue has multiplicity at most $k-1$. Therefore, by the theorem above, each neighborhood graph of X has an eigenvalue less than -1. This contradicts the fact that each neighborhood graph is a disjoint union of cliques.

5.2.1 A bound on the second eigenvalue of geometric DRGs

In this section we prove a bound on the eigenvalues of geometric distance-regular graphs.

Lemma 5.2.4 ([Brouwer et al., 1989, p.130])**.** *Let X be a (k-regular) distance-regular graph of diameter d. The eigenvalues of X, distinct from k, are equal to the eigenvalues of the following matrix*

$$T = \begin{pmatrix} -c_1 & b_1 & 0 & 0 & ... \\ c_1 & k-b_1-c_2 & b_2 & 0 & ... \\ 0 & c_2 & k-b_2-c_3 & b_3 & ... \\ ... & & \vdots & & ... \\ ... & & 0 & c_{d-1} & k-b_{d-1}-c_d \end{pmatrix}. \tag{5.1}$$

To get a bound on the second largest eigenvalue of a distance-regular graph we also need the Perron-Frobenius theorem in the following form.

Theorem 5.2.5 (Perron, Frobenius)**.** *Let A be a non-negative matrix. Then there exists a non-negative number θ and a non-negative vector v, such that $Av = \theta v$. Moreover, every other eigenvalue θ_i of A satisfies $|\theta_i| \leq \theta$.*

As a simple corollary we can deduce the following claim.

Lemma 5.2.6. *Let X be a distance-regular graph of diameter d. Then the second largest eigenvalue θ_1 of X satisfies*

$$\min_{i\in[d]} (k - b_{i-1} + b_i + c_{i-1} - c_i) \leq \theta_1 \leq \max_{i\in[d]} (k - b_{i-1} + b_i + c_{i-1} - c_i)$$

Proof. Let T be the matrix defined in Lemma 5.2.4. Then, the matrix $T' = T + c_d I_d$ is a non-negative matrix. Clearly, $\theta_1 + c_d$ is the largest eigenvalue of T', and by the Perron-Frobenius theorem, there exists a corresponding eigenvector v with non-negative entries. Since all entries with indices $(i, i+1)$ and $(i+1, i)$ are positive, it is easy to see that v must

have only strictly positive entries. Hence, the desired bounds follow from the inequality

$$\left(\min_{j\in[n]} v_j\right)\cdot\left(\max_{i\in[n]}\sum_{j=1}^{n} T'_{i,j}\right) \le (\theta_1 + c_d)v_i = \sum_{j=1}^{n} T'_{i,j}v_j \le \left(\max_{j\in[n]} v_j\right)\cdot\left(\max_{i\in[n]}\sum_{j=1}^{n} T'_{i,j}\right),$$

by picking i that minimizes (respectively, maximizes) the entry v_i. (Note that $b_0 = k$.) □

In particular, in the case of geometric-distance regular graphs we get the following.

Theorem 5.2.7. *Let X be a geometric distance-regular graph of diameter d with smallest eigenvalue $-m$. Let $\tau_\Delta = \min_i(\tau_i - \tau_{i-1})$ and $\tau^\Delta = \max_i(\tau_i - \tau_{i-1})$. Then*

$$k - \frac{\tau^\Delta}{m}k - mc_d \le \theta_1 \le k - \frac{\tau_\Delta}{m}k + mc_d.$$

Proof. Recall that, by Lemma 3.2.1, for geometric distance-regular graphs we have

$$c_{i+1} = \tau_{i+1}\psi_i \quad \text{and} \quad b_i = (m-\tau_i)\left(\frac{k}{m} + 1 - \psi_i\right) \quad \text{for } i = 0, 1, \ldots d-1. \tag{5.2}$$

Therefore,

$$(m-\tau_i)\left(\frac{k}{m} - c_d\right) \le b_i \le (m-\tau_i)\frac{k}{m}, \tag{5.3}$$

and so, for $i \in [d]$,

$$-(\tau_i-\tau_{i-1})\frac{k}{m}-(m-\tau_i+1)c_d \le -b_{i-1}+b_i+c_{i-1}-c_i \le -(\tau_i-\tau_{i-1})\frac{k}{m}+(m-\tau_{i-1})c_d. \tag{5.4}$$

Hence, the desired inequality follows from Lemma 5.2.6. □

5.2.2 Bounds on parameters of geometric DRGs

Recall that Theorem 3.2.5 shows that under mild assumptions the vertex-clique intersection parameters ψ_i form an increasing sequence. Below we prove that under mild assumptions τ_i

also form an increasing sequence.

Lemma 5.2.8. *Let X be a geometric distance-regular graph of diameter d, with smallest eigenvalue $-m$. Assume that $\mu \geq 2$, and $k \geq m^2 c_d$. Then*

$$\tau_i < \tau_{i+1}, \quad \textit{for all } i = 1, 2, \ldots, d-1.$$

Proof. Recall, by Lemma 3.2.1,

$$c_i = \tau_i \psi_{i-1}, \qquad b_i = (m - \tau_i)\left(\frac{k}{m} + 1 - \psi_i\right).$$

Hence, in particular, $\psi_{i-1} \leq c_i \leq c_d$, for $i \leq d$. So for $i \leq d-1$

$$(m - \tau_i)\left(\frac{k}{m} - c_d\right) < b_i \leq \frac{m - \tau_i}{m} k. \tag{5.5}$$

A geometric distance-regular graph with $\mu \geq 2$ contains an induced quadrangle. Thus, by the Terwilliger inequality (see Theorem 2.4.5) we have

$$b_i \geq b_{i+1} + \lambda + 2 + c_i - c_{i+1}, \text{ for } i = 0, 1, \ldots, d-1.$$

Therefore, for $i \leq d-2$, using Eq. (5.5),

$$\frac{m - \tau_i}{m} k > (m - \tau_{i+1})\left(\frac{k}{m} - c_d\right) + \lambda + 2 - c_d.$$

Since $\lambda \geq k/m - 1$, for $i \leq d-2$, we get

$$(m - \tau_i) > (m - \tau_{i+1}) - m^2 c_d/k + 1 \quad \Rightarrow \quad \tau_{i+1} > \tau_i + 1 - m^2 c_d/k.$$

Therefore, $\tau_{i+1} > \tau_i$, for $i = 0, 1, \ldots, d-2$. Moreover, $\tau_{d-1} < m$ and $\tau_d = m$. Hence, the

claim of the lemma follows. □

Corollary 5.2.9. *Let X be a geometric distance-regular graph of diameter d with smallest eigenvalue $-d$. If $k \geq d^2 c_d$ and $\mu \geq 2$, then $\tau_i = i$ for every $i \in [d]$.*

Corollary 5.2.10. *Let $\delta = k/(dc_d)$. If the assumptions of Lemma 5.2.8 hold for $m = d$, then*

$$(1-\delta)\frac{d-i}{d}k \leq b_i \leq \frac{d-i}{d}k, \quad \text{for } 1 \leq i \leq d-1.$$

Proof. By Corollary 5.2.9, we have $\tau_i = i$ for $i \leq d-1$. So the desired inequality directly follows from Eq. (5.5). □

5.2.3 A lower bound on the standard sequence of geometric DRGs

Next, for a geometric distance-regular graph, we prove a lower bound on the standard sequence of its second largest eigenvalue.

Lemma 5.2.11. *Let X be a geometric distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-d$. Let θ_1 be its second largest eigenvalue and $(u_i)_{i=0}^d$ be the corresponding standard sequence. Let $\delta < 1/30$. Assume that $\mu \geq 2$ and $k \geq c_d \cdot \delta^{-1} d^2$.*

Then, for $1 \leq j \leq d-1$

$$u_j \geq \frac{\theta_1}{k} \cdot \left(\frac{d-j}{d-1}\right) d^{-10\delta}.$$

Proof. Recall that the standard sequence corresponding to the eigenvalue θ_1 satisfies

$$u_0 = 1, \qquad u_1 = \frac{\theta_1}{k}, \qquad c_i u_{i-1} + a_i u_i + b_i u_{i+1} = \theta_1 u_i, \text{ for } i = 1, \ldots, d-1.$$

We can rewrite this as

$$u_{i+1} = u_i \frac{(\theta_1 + b_i + c_i - k)}{b_i} - u_{i-1}\frac{c_i}{b_i} \geq u_i \left(1 - \frac{k-\theta_1}{b_i}\right) - u_{i-1}\frac{c_i}{b_i}. \tag{5.6}$$

By the assumptions of the lemma,

$$\psi_{i-1} \leq c_i \leq c_d \leq \delta \frac{k}{d^2} \leq \delta \frac{k}{2d}. \tag{5.7}$$

Additionally, since $d^2 c_d < k$, by Corollary 5.2.9, $\tau_i = i$ for $i \in [d]$.

So, using Lemma 3.2.1, Theorem 5.2.7 and Eq. (5.7), we get the following bounds

$$k - \theta_1 \leq \frac{k}{d} + dc_d \leq (1+\delta)\frac{k}{d}. \tag{5.8}$$

$$b_i \geq (d-i)\left(\frac{k}{d} - c_d\right) \geq (1-\delta)\frac{(d-i)}{d}k, \quad \text{for } i \leq d-1. \tag{5.9}$$

For the convenience of the future computations we first show that for $1 \leq i \leq d-2$, the inequality $3u_{i+1} \geq u_i \geq 0$ holds. Indeed, by Eq. (5.8), $u_1 \geq u_0/3$. For $i \leq d-2$, by Eq. (5.8) and Eq. (5.9),

$$1 - \frac{k-\theta_1}{b_i} \geq 1 - \frac{1+\delta}{2-2\delta} \geq \frac{1}{2} - 2\delta. \tag{5.10}$$

Additionally, using Eq. (5.7) and Eq. (5.9), we get

$$\frac{c_i}{b_i} \leq \frac{dc_d}{(1-\delta)k} \leq \frac{2dc_d}{k} \leq \delta. \tag{5.11}$$

Thus, by induction, for $\delta < 1/30$, Eq. (5.6), Eq. (5.10) and Eq. (5.11) yield

$$u_{i+1} \geq \left(\frac{1}{2} - 2\delta\right) u_i - \delta u_{i-1} \geq \left(\frac{1}{2} - 5\delta\right) u_i \geq \frac{1}{3} u_i.$$

Hence, for $i \leq d-2$, we can rewrite Eq. (5.6)

$$u_{i+1} \geq u_i \frac{(\theta_1 + b_i + c_i - k)}{b_i} - 3u_i \frac{c_i}{b_i} \geq u_i \left(1 - \frac{k - \theta_1 + 2c_d}{b_i}\right).$$

Thus, using Eq. (5.7), Eq. (5.8) and Eq. (5.9), for $i \le d-2$,

$$\begin{aligned} u_{i+1} &\ge u_i \left(1 - \frac{(1+\delta)k + 2dc_d}{(1-\delta)(d-i)k}\right) \ge u_i \left(1 - \frac{1}{(d-i)}\frac{(1+2\delta)}{(1-\delta)}\right) \ge \\ &\ge u_i \left(1 - \frac{1}{(d-i)}(1+4\delta)\right) = u_i \left(\frac{d-i-1}{d-i}\right)\left(1 - \frac{4\delta}{d-i-1}\right). \end{aligned} \tag{5.12}$$

We note that for $0 < x < 1/3$ we have $(1-x) \ge e^{-(5/4)x}$. Thus, for $j \le d-1$,

$$\prod_{i=1}^{j-1}\left(1 - \frac{4\delta}{d-i-1}\right) \ge \prod_{i=1}^{d-2}\left(1 - \frac{4\delta}{i}\right) \ge \exp\left(-5\delta \sum_{i=1}^{d-2}\frac{1}{i}\right) \ge \exp\left(-10\delta \ln d\right). \tag{5.13}$$

Therefore, for all $1 \le j \le d-1$,

$$u_j \ge u_1 \prod_{i=1}^{j-1}\frac{d-i-1}{d-i}\prod_{i=1}^{j-1}\left(1 - \frac{4\delta}{d-i-1}\right) \ge \frac{\theta_1}{k}\cdot\left(\frac{d-j}{d-1}\right)d^{-10\delta}. \qquad \square$$

5.2.4 Proof of Theorem 5.2.2

Proof of Theorem 5.2.2. Assumptions of the theorem, and Corollary 5.2.10, imply that

$$\frac{c_i}{b_{i-1}} \le \frac{c_d}{b_{d-1}} \le \frac{2dc_d}{k} \le \frac{1}{50d^2} \quad \text{for every } 1 \le i \le d. \tag{5.14}$$

This means that $50d^2 \cdot k_{i-1} \le k_i$ for every $1 \le i \le d$. Thus

$$k_d \ge n\left(1 - \frac{1}{50d^2}\right). \tag{5.15}$$

Define $\delta = 1/(100d\ln d) < 1/30$. Then $k \ge c_d \cdot \delta^{-1}d^2$.

Using Lemma 5.2.11 for $j = d-1$, Eq. (5.8) and Corollary 5.2.10

$$\begin{aligned} k_{d-1}u_{d-1}^2 &\geq k_d \frac{c_d}{b_{d-1}} \cdot \left(\frac{\theta_1}{k}\right)^2 \cdot \left(\frac{1}{d-1}\right)^2 d^{-20\delta} \geq \\ &\geq k_d \cdot \frac{dc_d}{k} \cdot \left(\frac{d-1-\delta}{d}\right)^2 \left(\frac{1}{d-1}\right)^2 \cdot d^{-20\delta} \geq \\ &\geq n\left(1-\frac{1}{50d^2}\right) \cdot \frac{1}{k} \cdot \frac{c_d}{d} \cdot \left(\frac{d-1-\delta}{d-1}\right)^2 \cdot d^{-20\delta} \geq \\ &\geq \frac{n}{k} \cdot \frac{c_d}{d} \cdot \left(\left(1-\frac{1}{50d^2}\right) \cdot \left(1-\frac{2\delta}{d}\right)^2 \cdot d^{-20\delta}\right). \end{aligned} \tag{5.16}$$

Now, for $\delta = 1/(100d\ln d)$,

$$\begin{aligned} &\left(\frac{d+1}{d}\right) \cdot \left(1-\frac{1}{50d^2}\right) \cdot \left(1-\frac{2\delta}{d}\right)^2 \cdot d^{-20\delta} \geq \\ &\geq \exp\left(\frac{1}{2d} - 40\delta\ln d - \frac{8\delta}{d} - \frac{1}{50d^2}\right) > 1 \end{aligned} \tag{5.17}$$

where we use the inequalities $e^{x/2} \leq 1+x$ and $e^{-2x} \leq 1-x$ for $0 < x \leq 1/2$.

Therefore, if $c_d \geq d+1$, then $k_{d-1}u_{d-1}^2 > n/k$. Using the Biggs formula, this immediately implies that the second largest eigenvalue θ_1 has multiplicity $f_1 \leq k-1$. Therefore, $-1-\dfrac{b_1}{\theta_1+1} < -1$ is an eigenvalue of every neighborhood graph of X. This implies that $\psi_1 > 1$ (as for $\psi_1 = 1$, every neighborhood graph is a union of disjoint cliques). So, by Lemma 3.2.2, $\mu \geq 4$. We get a contradiction with the assumptions of the theorem.

Hence, $c_d \leq d$. At the same time, $\tau_i = i$, for every $1 \leq i \leq d$, so $\psi_{i-1} = 1$ and $c_i = i$ for every $1 \leq i \leq d$. Therefore, X has the same intersection array as the Hamming graph $H(d, 1+k/d)$. Therefore, by Theorem 3.5.3, X is a Hamming graph or a Doob graph. Note that X may be a Doob graph, only if $1+k/d = 4$, which is not possible. Therefore, X is a Hamming graph. □

CHAPTER 6

CHARACTERIZATION OF DISTANCE-REGULAR GRAPHS WITH BOUNDED SMALLEST EIGENVALUE AND LARGE SPECTRAL GAP

6.1 Prior work and our contributions

In Chapter 5, we discussed characterization of graphs based on their least eigenvalue. Another direction that received a lot of attention is the problem of characterizing graphs using their second largest eigenvalue (of the adjacency matrix). Here we focus on such characterizations for distance-regular graphs of diameter at least 3 and $\mu \geq 2$.

These characterizations will be used in Section 8.1 to prove Theorem 1.2.6.

6.1.1 Prior work

A result of Terwilliger [1986] (see [Brouwer et al., 1989, Theorem 4.4.3]) implies that the icosahedron is the only distance-regular graph, for which the second largest eigenvalue θ_1 satisfies $\theta_1 > b_1 - 1$ and $\mu \geq 2$. One of the central results in representation theory of distance-regular graphs gives the classification of distance-regular graphs with $\mu \geq 2$ and $\theta_1 = b_1 - 1$.

Theorem 6.1.1 ([Brouwer et al., 1989, Theorem 4.4.11])**.** *Let X be a distance-regular graph of diameter $d \geq 3$ with second largest eigenvalue $\theta_1 = b_1 - 1$. Assume $\mu \geq 2$. Then one of the following holds:*

1. *$\mu = 2$ and X is a Hamming graph, a Doob graph, or a locally Petersen graph (and all such graphs are known).*

2. *$\mu = 4$ and X is a Johnson graph.*

3. $\mu = 6$ *and* X *is a half cube.*

4. $\mu = 10$ *and* X *is a Gosset graph* $E_7(1)$.

6.1.2 Our contribution

We consider the case $\theta_1 \geq (1-\varepsilon)b_1$ for a sufficiently small $\varepsilon > 0$. The relaxation of the assumption on the second largest eigenvalue comes at the cost of requiring additional structural constraints. Our main structural assumption is that X is a geometric distance-regular graph. Additional technical structural assumptions depend on whether the neighborhood graphs of X are connected. We note that for a geometric distance-regular graph X either the neighborhood graph $X(v)$ is connected for every vertex v, or $X(v)$ is disconnected for every vertex v (see Lemma 3.2.4). We give the following characterizations.

Theorem 6.1.2. *There exists an absolute constant* $\varepsilon^* > 0.0065$ *such that the following is true. Let* X *be a geometric distance-regular graph of diameter* $d \geq 2$ *with smallest eigenvalue* $-m$. *Suppose that* $\mu \geq 2$ *and* $\theta_1 + 1 > (1-\varepsilon^*)b_1$. *Moreover, assume that the vertex degree satisfies* $k \geq \max(m^3, 29)$ *and the neighborhood graph* $X(v)$ *is connected for some vertex* v *of* X.

Then X *is a Johnson graph* $J(s,d)$ *with* $s = (k/d) + d$.

Remark 6.1.3. We give the exact definition of ε^* in Lemma 6.2.2 (see also Def. 5.1.12 and Theorem 5.1.14). We note that ε^* can be set to be as large as $2/7$, if we additionally assume k to be sufficiently large (see Remark 6.2.9).

Theorem 6.1.4. *Let* X *be a geometric distance-regular graph of diameter* $d \geq 2$ *with smallest eigenvalue* $-m$. *Consider an arbitrary* $0 < \varepsilon < 1/(6m^4 d)$. *Suppose that* $\mu \geq 2$ *and* $\theta_1 \geq (1-\varepsilon)b_1$. *Moreover, assume* $c_t \leq \varepsilon k$ *and* $b_t \leq \varepsilon k$ *for some* $t \leq d$, *and the neighborhood graph* $X(v)$ *is disconnected for some vertex* v *of* X.

Then X *is a Hamming graph* $H(d,s)$ *with* $s = 1 + k/d$.

Remark 6.1.5. Theorem 6.1.4 is closely related to our characterization of Hamming graphs by smallest eigenvalue (Theorem 5.2.2). Theorem 6.1.4 makes much weaker assumption on the smallest eigenvalue of the graph, however, in return it requires a bound on the second largest eigenvalue and mild additional assumptions on the intersection numbers. We believe that it should be possible to prove the characterization that has weaker assumption on the smallest eigenvalue than Theorem 5.2.2, and at the same time which makes weaker assumption on the second largest eigenvalue than Theorem 6.1.4.

As we discuss in Section 5.1.4 (see Theorem 5.1.20), the assumption that a distance-regular graph is geometric excludes only finitely many graphs with $\mu \geq 2$, if the smallest eigenvalue of the graph is assumed to be bounded.

Even though the assumptions of Theorem 6.1.2 seem weaker than those of Theorem 6.1.4 (for instance, ε is absolute), we believe that, in comparison with known results, Theorem 6.1.4 brings more novelty. The known characterization of Johnson graphs in terms of the local structure (Theorem 3.4.2) seems to be more easily applicable than the known characterizations of Hamming graphs. All characterizations of Hamming graphs known to the author, in terms of intersection numbers, eigenvalues or local structure, make strong equality constraints either on the number of vertices, or on the eigenvalues. In contrast, Theorem 6.1.4 makes no assumptions of such flavor and therefore might be more broadly applicable.

Finally, we note that our characterizations confirm Conjecture 5.1.22 in rather special cases.

6.2 Characterization of Johnson graphs via spectral gap

In this section, we prove Theorem 1.4.4, our characterization of Johnson graphs. Specifically, we prove that a distance-regular graph with $\theta_1 + 1 > (1 - \varepsilon^*)b_1$ and connected neighborhood graphs is a Johnson graph (for a sufficiently large k). We also show that the inequality $\theta_1 + 1 > (1 - \varepsilon^*)b_1$ can hold for a distance-regular graph with disconnected neighborhood

graphs only if $\mu \leq 2$ (see Proposition 6.2.7).

The main idea of the proofs is to use the fact that for $b^+ = \dfrac{b_1}{\theta_1 + 1}$ the expression $-1-b^+$ is a lower bound on the smallest eigenvalue of a neighborhood graph $X(v)$. More precisely, we use the following result of Terwilliger [1986].

Theorem 6.2.1 (Terwilliger [1986], see [Brouwer et al., 1989, Theorem 4.4.3])**.** *Let X be a distance-regular graph of diameter $d \geq 2$ with distinct eigenvalues $k = \theta_0 > \theta_1 > \ldots > \theta_d$, and let $b^+ = \dfrac{b_1}{\theta_1 + 1}$, $b^- = \dfrac{b_1}{\theta_d + 1}$. Then each neighborhood graph $X(v)$ has the smallest eigenvalue $\geq -1 - b^+$, and the second largest eigenvalue $\leq -1 - b^-$.*

Recall, we assume that the second largest eigenvalue of X satisfies $\theta_1 + 1 \geq (1 - \varepsilon)b_1$. In this case the smallest eigenvalue of the neighborhood graph $X(v)$ is at least $-2 - \delta$, for $\delta = \varepsilon/(1 - \varepsilon)$. We also observe that if X is an edge-regular graph, its neighborhood graph $X(v)$ is regular for every vertex $v \in X$.

First, we note that if the diameter d of a distance-regular graph X is at least 2, $\lambda > 2$ and the neighborhood graph $X(v)$ is connected, then the smallest eigenvalue of $X(v)$ is at most -2. Indeed, if a regular connected graph has the smallest eigenvalue > -2, then, by Theorem 5.1.1, it is a complete graph or an odd polygon. The neighborhood graph $X(v)$ cannot be complete as $d \geq 2$, and $X(v)$ cannot be an odd polygon as $\lambda > 2$.

The graphs for which the smallest eigenvalue is precisely -2 were classified by Cameron et al. [1991]. We use their classification in the case of connected regular graphs (Theorem 5.1.10).

Combining the above discussion with results of Section 5.1.2 (Theorem 5.1.14) we get the following claim. Recall, that in Theorem 5.1.14 we use ϑ_1 (≈ -2.006594) to denote the smallest root of the equation $\theta^2(\theta^2 - 1)^2(\theta^2 - 3)(\theta^2 - 4) = 1$.

Lemma 6.2.2. *Let X be a distance-regular graph of diameter $d \geq 2$. Assume that the second largest eigenvalue of X satisfies $\theta_1 + 1 > (1 - \varepsilon^*)b_1$, for $0 < \varepsilon^* = \dfrac{-2 - \vartheta_1}{-1 - \vartheta_1}$. Then for every*

vertex v of X, the neighborhood graph $X(v)$ is a regular graph with smallest eigenvalue at least -2.

Moreover, if $X(v)$ is connected, $\lambda > 2$, and the vertex degree of X is at least 29, then $X(v)$ is the line graph of a regular or of a bipartite semiregular connected graph.

Remark 6.2.3. One can compute that $\varepsilon^* \approx 0.0065504$. Observe that, the neighborhood graph $X(v)$ is regular of degree λ, and by Theorem 5.1.13, $\lim_{\lambda\to\infty} \vartheta_\lambda = -1-\sqrt{2}$. Thus, we can replace ε^* with an arbitrary number which is less than $1-1/\sqrt{2} \approx 0.29289$, if we additionally require λ to be sufficiently large.

Next we analyze the structure of the local graph $X(v)$ in the case when X is geometric.

Lemma 6.2.4. *Let X be a geometric distance distance-regular graph with smallest eigenvalue $-m$. Suppose that $X(v)$ is the line graph of a regular or a bipartite semiregular connected graph. Assume that the vertex degree $k \geq \max(m^3, 3)$. Then, $X(v)$ is the line graph of a complete bipartite graph $K_{s,t}$ for each vertex v of X, where $\{s,t\} = \{m, k/m\}$.*

Proof. Fix a Delsarte clique geometry $\mathcal{C}$ of X. Let $C_1, C_2, \ldots, C_m \in \mathcal{C}$ be the cliques that contain a vertex v. Since every edge of X is contained in precisely one clique, every vertex of $N(v)$ is contained in precisely one of $C_1, C_2, \ldots, C_m$. Let $u \in C_1 \setminus \{v\}$, by the definition of ψ_i (see Section 3.2), u is adjacent with precisely ψ_1 vertices of C_i for all $i = 2, 3, \ldots, m$. Therefore, the degree of every vertex u in $X(v)$ equals $k/m - 1 + (\psi_1 - 1)(m-1)$.

Assume that Y is the line graph of a regular graph Z with vertex degree t. Then the degree of a vertex in Y is equal $2(t-1)$. Moreover, the size of a maximal clique in Y is t, if $t \geq 3$. Since Y contains a clique of size k/m, we get $t \geq k/m$. Therefore, if $k/m \geq 3$ and $(k/m - 1) > (\psi_1 - 1)(m-1)$, then $X(v)$ is not a line graph of a regular graph. In particular this is true, if $k \geq \max(m^3, 3)$, as $\psi_1 \leq \tau_2 \leq m$ by the definition of τ_2 and Lemma 3.2.2.

Hence, for every v, the neighborhood graph $X(v)$ is the line graph of a complete bipartite graph $K_{s,t}$. The size of the maximal clique in the line graph of $K_{s,t}$ is $\max(s,t)$. Thus,

$\max(s,t) = k/m$. There are k vertices in $X(v)$ and st vertices in the line graph of $K_{s,t}$, so $\{s,t\} = \{m, k/m\}$. $\square$

In the case when $X(v)$ is the line graph of a complete bipartite graph $K_{s,t}$ and $1+\theta_1 \geq (1-\varepsilon)b_1$, we show that X is a Johnson graph. Our goal is to use the characterization of the Johnson graphs by local structure stated in Theorem 3.4.2. The only condition we still need to verify is $\mu \leq 4$. We prove that if $\mu > 4$, then X contains an induced subgraph $K_{3,2}$ and so we can use the inequality provided by the theorem below.

Theorem 6.2.5 ([Brouwer et al., 1989, Theorem 4.4.6])**.** *Let X be a distance-regular graph of diameter $d \geq 2$ with eigenvalues $k = \theta_0 > \theta_1 > \ldots > \theta_d$ and put $b^+ = b_1/(\theta_1+1)$. If X contains a non-empty induced complete bipartite subgraph $K_{s,t}$, then*

$$\frac{2st}{s+t} \leq b^+ + 1.$$

Lemma 6.2.6. *Let X be a geometric distance-regular graph of diameter $d \geq 2$.*

1. *Assume that $\psi_1 = 1$, then X contains an induced $K_{\tau_2,2}$.*

2. *Assume that $\mu \geq 2$, then X contains an induced $K_{2,2}$ (a quadrangle).*

Proof. Let u and v be two vertices at distance 2 in X. By the definition of τ_2 there exist distinct cliques $C_1, C_2, \ldots, C_{\tau_2}$ which contain u and have non-trivial intersection with $N(v)$.

1. Each C_i has precisely $\psi_1 = 1$ common vertices with $N(v)$. Denote $w_i = C_i \cap N(v)$. Note that w_i is at distance 1 from C_j for $i \neq j$, moreover, w_i is adjacent to u, while $u \in C_j$ and $u \neq w_j$. Thus w_j is not adjacent to w_i for $i \neq j$. Therefore, X contains an induced $K_{\tau_2,2}$ (on vertices $\{w_1, w_2, \ldots, w_{\tau_2}, u, v\}$).

2. By Lemma 3.2.2, $\tau_2 \geq 2$, if $\mu \geq 2$. Take $w \in N(v) \cap C_1$. Assume there are no induced $K_{2,2}$ in X. Then w is adjacent to each vertex in $T = C_2 \cap N(v)$. Note that $|T| = \psi_1$,

$u \notin T$ and w is adjacent to u, so w has at least $\psi_1 + 1$ neighbors in C_2. This gives a contradiction with the definition of ψ_1. $\square$

Using the lemma above we obtain the following corollary to the Theorem 6.2.5.

Proposition 6.2.7. *Let X be a geometric distance-regular graph of diameter $d \geq 2$. Assume that the neighborhood graphs of X are disconnected. If $\mu \geq 3$, then the second largest eigenvalue of X satisfies $\theta_1 + 1 \leq \frac{5}{7} b_1$.*

Proof. Since X is geometric and $X(v)$ is disconnected, by Lemma 3.2.4, $\psi_1 = 1$. Moreover, if $\mu \geq 3$, by Lemma 6.2.6, there is an induced $K_{3,2}$. Therefore, by Theorem 6.2.5,

$$\frac{b_1}{\theta_1 + 1} \geq \frac{12}{5} - 1 = \frac{7}{5}. \qquad \square$$

In the next lemma we show the existence of an induced complete bipartite subgraph $K_{\tau_2,2}$ in the case when a neighborhood graph is the line graph of a triangle-free graph.

Lemma 6.2.8. *Let X be a geometric distance-regular graph. Assume that for each vertex v of X the induced subgraph $X(v)$ is the line graph of a triangle-free graph. Then $\psi_1 = 2$ and X contains induced $K_{\tau_2,2}$.*

Proof. Observe that if the line graph Y of a triangle-free graph Y' contains a triangle, then the three corresponding edges of the base graph Y' are incident to the same vertex of Y'.

Fix a Delsarte clique geometry $\mathcal{C}$ of X . Let v be a vertex of X, and $C \in \mathcal{C}$ be a Delsarte clique which contains v, and let $w \in N(v) \setminus C$. By Lemma 6.2.6, since $X(v)$ is connected, $\psi_1 \geq 2$. Assume that $\psi_1 \geq 3$. Then w is adjacent to at least two vertices v_1 and v_2 in C distinct from v. Since w, v_1, v_2 form a triangle in $X(v)$, the corresponding edges in the base graph are incident to the same vertex. Similarly, for every vertex $x \in C \setminus \{v\}$, the edges of the base graph that correspond to x, v_1 and v_2 are incident to the same vertex. Therefore, $\{w\} \cup C$ is a clique in X, which contradicts maximality of C. Therefore, $\psi_1 = 2$.

Let u and v be a pair of vertices at distance 2 in X. There exist distinct cliques $C_1, C_2, \ldots, C_{\tau_2}$ which contain u and have non-trivial intersection with $N(v)$, and distinct cliques $C'_1, C'_2, \ldots, C'_{\tau_2}$ which contain v and have non-trivial intersection with $N(u)$.

Let $T = N(u) \cap N(v)$. Assume that $w_1, w_2 \in T$ are adjacent, but $\{w_1, w_2\}$ is not a subset of C_i or C'_i for every $i \in [\tau_2]$. Since $\psi_1 = 2$ there exists a vertex $w \in T$ such that $\{w, w_1\}$ is a subset of some C_i. Similarly there is $w' \in T$ with $\{w', w_1\} \subseteq C'_j$ for some j. Assume that $X(v)$ is the line graph of a triangle-free graph Y. Assume that edges of Y that correspond to w_1, w' are incident with a vertex x of Y. If corresponding to w_2 edge is incident with x, then by the argument similar to the one above, $w_2 \in C'_j$. This contradicts the choice of w_2. Let y be the vertex of Y incident to the edges in Y corresponding to w_2 and w_1. Since $|C_i \cap C'_j| \leq 1$, we have $w \notin C'_j$, so w is not incident to x. Thus, w is incident to y. Hence, w, w_1 and w_2 form a triangle.

Since $\{w, w_1\} \subseteq C_i$, and $X(u)$ is the line graph of a triangle-free graph, we similarly get that $w_2 \in C_i$. This gives a contradiction with the choice of w_1, w_2. Therefore, every pair of distinct vertices in T are adjacent if and only if they share the same clique C_i or C'_i.

We obtain that an edge between a pair of vertices in T is between vertices in $T \cap C_i$ or between vertices in $T \cap C'_j$ for some i, j. We refer to them as edges of type 1 and edges of type 2, respectively. Since $|T \cap C_i| = |T \cap C'_i| = \psi_1 = 2$, every vertex in T is incident with precisely one edge of type 1 and precisely one edge of type 2. Therefore, the subgraph induced on T is a union of even cycles. Hence, we may choose an independent set $S \subseteq T$ of size $\frac{1}{2}|T| = \frac{1}{2}\psi_1\tau_2 = \tau_2$ in X. The graph induced on $S \cup \{u, v\}$ is K_{2,τ_2}. $\square$

Now we are ready to combine the arguments above into a proof of Theorem 1.4.4.

Proof of Theorem 1.4.4. Denote $b^+ = \dfrac{b_1}{\theta_1 + 1}$. Since $\varepsilon^* = \dfrac{-2 - \vartheta_1}{-1 - \vartheta_1}$, the assumptions of the theorem imply that $b^+ < -1 - \vartheta_1$. Since X is not complete, by Lemma 3.1.7, $m \geq 2$. Thus the inequalities $k \geq m^3$ and $\lambda \geq k/m - 1$ imply $\lambda \geq 3$. Hence, by Lemma 6.2.2 and Lemma 6.2.4, either $X(v)$ is a disconnected graph for some vertex v, or $X(v)$ is the line graph of the

complete bipartite graph $K_{s,t}$ for every vertex v of X.

In the latter case, by Lemma 6.2.8, X contains $K_{\tau_2,2}$ and $\psi_1 = 2$. Since we assumed that $b^+ \leq -1 - \vartheta_1 < 7/5$, by Theorem 6.2.5, we get that $\tau_2 \leq 2$ (as otherwise there is an induced $K_{2,3}$ subgraph). Hence, $\mu \leq 4$. By Theorem 3.4.2, we get that X is the Johnson graph $J(s,d)$, or a graph which can be double covered by $J(2d,d)$. The latter case is not possible because $k \geq m^3$. □

Remark 6.2.9. We note that in the light of Remark 6.2.3, if we additionally assume that k is large enough, then we can replace ε^* with $2/7$ in Theorem 1.4.4. Indeed, since $\lambda \geq (k/m) - 1$ and $k \geq m^3$, the assumption that k is large enough guarantees that λ is large enough. Hence, the proof above will work since the inequality $\varepsilon^* < 2/7$ implies that $b^+ < 7/5$.

6.3 Characterization of Hamming graphs via spectral gap

In this section, we prove a characterization of Hamming graphs in terms of the spectral gap and local parameters. As in the previous section, we assume that the second largest eigenvalue of X satisfies $\theta_1 \geq (1-\varepsilon)b_1$. We show that if additionally for each vertex the neighborhood graph is a disjoint union of cliques, $\mu = 2$, and there is a dominant distance, then X is a Hamming graph.

We start by showing that for a geometric distance-regular graph the sequence $(\tau_i)_{i=1}^{t-1}$ is increasing if $\mu \geq 2$ and c_t is sufficiently small. This can be seen as a version of Lemma 5.2.8

Lemma 6.3.1. *Let X be a geometric distance-regular graph of diameter d, with smallest eigenvalue $-m$. Assume that $\mu \geq 2$ and $c_t \leq \varepsilon k$, where $t \leq d$ and $0 < \varepsilon < 1/m^2$. Then*

$$\tau_i < \tau_{i+1}, \quad \textit{for all } i = 1, 2, \ldots, t-2.$$

Proof. Recall, by Lemma 3.2.1,

$$c_i = \tau_i \psi_{i-1}, \qquad b_i = (m - \tau_i)\left(\frac{k}{m} + 1 - \psi_i\right).$$

Hence, in particular, $\psi_{i-1} \le c_i \le c_t \le \varepsilon k$, for $i \le t$. So for $i \le t-1$

$$(m - \tau_i)\left(\frac{1}{m} - \varepsilon\right) k \le b_i \le \frac{m - \tau_i}{m} k. \tag{6.1}$$

By Lemma 6.2.6, a geometric distance-regular graph with $\mu \ge 2$ contains an induced quadrangle. Thus, by the Terwilliger inequality (see Theorem 2.4.5) we have

$$b_i \ge b_{i+1} + \lambda + 2 + c_i - c_{i+1}, \text{ for } i = 0, 1, \ldots, d-1.$$

Therefore, for $i \le t-2$, using Eq. (6.1),

$$\frac{m - \tau_i}{m} k \ge (m - \tau_{i+1})\left(\frac{1}{m} - \varepsilon\right) k + \lambda + 2 - \varepsilon k.$$

Since $\lambda \ge k/m - 1$, for $i \le t-2$, we get

$$(m - \tau_i) \ge (m - \tau_{i+1}) - m^2\varepsilon + 1 \quad \Rightarrow \quad \tau_{i+1} \ge \tau_i + 1 - m^2\varepsilon. \qquad \square$$

Corollary 6.3.2. *If the assumptions of Lemma 6.3.1 hold for $t = d$, then $\tau_i < \tau_{i+1}$ for $i \le d-1$.*

Proof. By Lemma 6.3.1, $\tau_i < \tau_{i+1}$ for $i \le d-2$. Observe, that by the definition of τ_i, we have $\tau_d = m$ and $\tau_i \le m-1$ for all $i \le d-1$. $\square$

Corollary 6.3.3. *If the assumptions of Lemma 6.3.1 hold for $t = d$, then*

$$(d - i)\left(\frac{1}{m} - \varepsilon\right) k \le b_i \le \frac{m - i}{m} k, \quad \textit{for } 1 \le i \le d-1.$$

Proof. Since $\tau_1 = 1$ and $\tau_{d-1} \leq m-1$, by Lemma 6.3.1, we have $i \leq \tau_i \leq m-d+i$ for $i \leq d-1$. So the desired inequality directly follows from Eq. (6.1). $\square$

To get a bound on the multiplicity of the second largest eigenvalue θ_1 of X we first prove lower bounds on the elements of the standard sequence corresponding to θ_1 (see Sec. 2.4).

Lemma 6.3.4. *Let X be a geometric distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-m$. Let θ_1 be its second largest eigenvalue and $(u_i)_{i=0}^d$ be the corresponding standard sequence. Assume that $\mu \geq 2$, $\theta_1 \geq (1-\varepsilon)b_1$, and $c_t \leq \varepsilon k$ for some $2 \leq t \leq d$, where $0 < \varepsilon < 1/(24m^2)$.*

Then, for $1 \leq j \leq t-1$

$$u_j \geq (1-3m^2\varepsilon)^{j-1}\left(\frac{m-\tau_j}{m-\tau_j+j-1}\right)\frac{\theta_1}{k}.$$

Proof. Recall that the standard sequence corresponding to the eigenvalue θ_1 satisfies

$$u_0 = 1, \qquad u_1 = \frac{\theta_1}{k}, \qquad c_i u_{i-1} + a_i u_i + b_i u_{i+1} = \theta_1 u_i, \text{ for } i = 1, \ldots, d-1.$$

We can rewrite this as

$$u_{i+1} = u_i \frac{(\theta_1 + b_i + c_i - k)}{b_i} - u_{i-1}\frac{c_i}{b_i} \geq u_i\left(1 - \frac{k-\theta_1}{b_i}\right) - u_{i-1}\frac{c_i}{b_i}. \tag{6.2}$$

For $2 \leq i \leq t$, by the assumptions of the lemma, $\psi_{i-1} \leq c_i \leq c_t \leq \varepsilon k$. So, by Lemma 3.2.1,

$$k - \theta_1 \leq k - (1-\varepsilon)b_1 \leq (k-b_1) + \varepsilon k \leq \frac{k}{m} + (m-1)\psi_1 + \varepsilon k \leq \frac{k}{m} + m\varepsilon k; \tag{6.3}$$

$$(m-\tau_i)\left(\frac{1}{m} - \varepsilon\right)k \leq b_i, \quad \text{for } i \leq t-1. \tag{6.4}$$

For the convenience of the future computations we first show that for $1 \leq i \leq t-2$, the inequality $3u_{i+1} \geq u_i \geq 0$ holds. Indeed, by Eq. (6.3), $u_1 \geq u_0/3$. Moreover, by

Lemma 6.3.1, $\tau_i \leq \tau_{t-1} - 1 \leq m-2$, so by Eq. (6.3) and Eq. (6.4),

$$1 - \frac{k-\theta_1}{b_i} \geq 1 - \frac{1+m^2\varepsilon}{2-2m\varepsilon} \geq \frac{1}{2} - m^2\varepsilon.$$

Thus, using that $\tau_i \leq m-2$ for $i \leq t-2$, by induction, we get from Eq. (6.2) and Eq. (6.4)

$$u_{i+1} \geq \left(\frac{1}{2} - m^2\varepsilon\right) u_i - m\varepsilon u_{i-1} \geq \left(\frac{1}{2} - 4m^2\varepsilon\right) u_i \geq \frac{1}{3} u_i.$$

Hence, for $i \leq t-2$, we can rewrite Eq. (6.2)

$$u_{i+1} \geq u_i \frac{(\theta_1 + b_i + c_i - k)}{b_i} - 3u_i \frac{c_i}{b_i} \geq u_i \left(1 - \frac{k - \theta_1 + 2\varepsilon k}{b_i}\right).$$

Thus, using Eq. (6.3) and Eq. (6.4), for $i \leq t-2$,

$$u_{i+1} \geq u_i \left(1 - \frac{k + (m^2+2m)\varepsilon k}{m b_i}\right) \geq u_i \left(1 - \frac{1}{(m-\tau_i)} \frac{(1+2m^2\varepsilon)}{(1-m\varepsilon)}\right) \geq u_i \left(1 - \frac{(1+3m^2\varepsilon)}{m-\tau_i}\right).$$

By Lemma 6.3.1, $\tau_i \leq \tau_j - (j-i)$ for $i \leq j \leq t-1$. Thus, for $\delta = 3m^2\varepsilon$ and $i+1 \leq j \leq t-1$,

$$u_{i+1} \geq (1-\delta)\left(1 - \frac{1}{m-\tau_i}\right) u_i \geq (1-\delta) \frac{m - \tau_j + j - i - 1}{m - \tau_j + j - i} u_i.$$

Therefore, for every $1 \leq j \leq t-1$,

$$u_j \geq (1-\delta)^{j-1} \prod_{i=1}^{j-1} \frac{m - \tau_j + j - i - 1}{m - \tau_j + j - i} u_1 = (1-\delta)^{j-1} \left(\frac{m-\tau_j}{m - \tau_j + j - 1}\right) \frac{\theta_1}{k}. \qquad \square$$

Theorem 6.3.5. *Let X be a geometric distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-m$. Take an arbitrary $0 < \varepsilon < 1/(6m^4 d)$. Suppose that $\mu \geq 2$, $c_t \leq \varepsilon k$ and $b_t \leq \varepsilon k$ for some $2 \leq t \leq d$. Assume, moreover, that the second largest eigenvalue of X satisfies $\theta_1 \geq (1-\varepsilon) b_1$.*

Then either the multiplicity f_1 of θ_1 satisfies $f_1 \leq k-1$, or $m = d$, $t = d$ and $c_d = d$.

Proof. Let $(u_i)_{i=0}^d$ be the standard sequence of X corresponding to θ_1. Then, by the Biggs formula, the multiplicity of θ_1 can be computed as

$$f_1 = \frac{n}{\left(\sum_{i=0}^{d} k_i u_i^2\right)}.$$

Note, as in Eq. (6.1), $b_{i-1} \geq b_{t-1} \geq \left(\frac{1}{m} - \varepsilon\right) k \geq \frac{1}{2m} k$ and $c_i \leq c_t$ for all $1 \leq i \leq t$. So

$$k_{i-1} = \frac{c_i}{b_{i-1}} k_i \leq \frac{c_t}{b_{t-1}} k_i \leq 2m\varepsilon k_i, \quad \text{for } i \leq t. \tag{6.5}$$

For $d-1 \geq i \geq t$, by Lemma 3.2.1, the inequality $b_i \leq b_t \leq \varepsilon k$ implies

$$\psi_i \geq \left(\frac{1}{m} - \varepsilon\right) k \geq \frac{1}{2m} k, \qquad \text{so} \qquad c_{i+1} = \tau_{i+1}\psi_i \geq \frac{1}{2m} k.$$

Hence, for $t \leq i \leq d-1$ we deduce,

$$k_{i+1} = \frac{b_i}{c_{i+1}} k_i \leq \frac{\varepsilon k}{k/(2m)} k_i = 2m\varepsilon k_i. \tag{6.6}$$

Combining Eq. (6.5) and Eq. (6.6) we obtain

$$n = \sum_{i=0}^{d} k_i \leq k_t \left(\sum_{i=0}^{t} (2m\varepsilon)^i + \sum_{i=1}^{d-t} (2m\varepsilon)^i \right) \leq \frac{1}{1-4m\varepsilon} k_t \qquad \Rightarrow \qquad k_t \geq (1-4m\varepsilon)n.$$

As in Eq. (6.3), $\theta_1/k \geq (m-1)/m - m\varepsilon$. So, by Lemma 6.3.4 and Eq.(6.1), for $t \geq 2$,

$$k_{t-1}u_{t-1}^2 \geq k_t \frac{c_t}{b_{t-1}} (1-3m^2\varepsilon)^{2t-4} \left(\frac{m-\tau_{t-1}}{m-\tau_{t-1}+t-2}\right)^2 \left(\frac{\theta_1}{k}\right)^2 \geq$$

$$\geq k_t \frac{c_t}{b_{t-1}} (1-3m^2\varepsilon)^{2t-4} \left(\frac{m-\tau_{t-1}}{m-\tau_{t-1}+t-2}\right)^2 (1-2m\varepsilon)^2 \left(\frac{m-1}{m}\right)^2 \geq$$

$$\geq (1-4m\varepsilon)n \cdot \frac{mc_t}{(m-\tau_{t-1})k} \cdot (1-3m^2\varepsilon)^{2d-1} \left(\frac{m-\tau_{t-1}}{m-\tau_{t-1}+t-2}\right)^2 \left(\frac{m-1}{m}\right)^2 \geq$$

$$\geq \frac{n}{k} \cdot (1-3m^2\varepsilon)^{2d} \cdot \frac{c_t}{m} \cdot \frac{(m-\tau_{t-1})(m-1)^2}{(m-\tau_{t-1}+t-2)^2}.$$

Our goal is to deduce from this inequality that $k_{t-1}u_{t-1}^2 > n/k$, unless $c_t = t = m = d$. We start by giving a bound on c_t. Observe that $\psi_{t-2} \geq 1$, and $\tau_{t-1} \geq t-1$, by Lemma 6.3.1. So, we obtain

$$c_t \geq c_{t-1} \geq \tau_{t-1}\psi_{t-2} \geq t-1. \tag{6.7}$$

Case 1. First, assume that $c_t = t-1$.

Then, Eq. (6.7) implies $\tau_{t-1} = t-1$ and $c_t = c_{t-1}$. Thus, in particular, we can simplify

$$\frac{c_t}{m} \cdot \frac{(m-\tau_{t-1})(m-1)^2}{(m-\tau_{t-1}+t-2)^2} = \frac{t-1}{m} \cdot \frac{(m-t+1)(m-1)^2}{(m-1)^2} = \frac{(t-1)(m-t+1)}{m}.$$

Also, observe that the constraint $c_t = c_{t-1}$ implies $t > 2$ as $1 = c_1 < 2 \leq \mu = c_2$. Moreover, by Corollary 2.4.7, $c_3 > c_2$ for $\mu \geq 2$, so we should have $t \geq 4$ in this case.

At the same time, $c_t = c_{t-1}$, using Terwilliger's inequality (see Theorem 2.4.5), implies

$$b_{t-1} \geq c_{t-1} - c_t + b_t + \lambda + 2 \geq \lambda + 2 \geq \frac{k}{m} + 1.$$

Recall that

$$b_{t-1} = (m-\tau_{t-1})\left(\frac{k}{m} + 1 - \psi_{t-1}\right) \leq (m-\tau_{t-1})\frac{k}{m},$$

which yields $\tau_{t-1} \leq m-2$. So, $t \leq m-1$, as $\tau_{t-1} = t-1$.

Thus, we have $4 \leq t \leq m-1$, which implies $m \geq 5$, and we get

$$\frac{c_t}{m} \cdot \frac{(m-\tau_{t-1})(m-1)^2}{(m-\tau_{t-1}+t-2)^2} = \frac{(t-1)(m-t+1)}{m} \geq \frac{2(m-2)}{m} \geq 2 - \frac{4}{m} \geq \frac{6}{5}.$$

Since, $3m^2\varepsilon < 1$, by Bernoulli's inequality

$$(1 - 3m^2\varepsilon)^{2d} \geq (1 - 6dm^2\varepsilon) > \frac{5}{6}.$$

Therefore, in this case,

$$k_{t-1}u_{t-1}^2 > \frac{n}{k} \quad \Rightarrow \quad f_1 \leq \frac{n}{k_{t-1}u_{t-1}^2} < k \qquad \Rightarrow \qquad f_1 \leq k - 1.$$

Case 2. Else, we have $c_t \geq t$.

Lemma 6.3.1 implies $t \leq \tau_{t-1} + 1 \leq m$. It is not hard to check (see Appendix 6.A), that

$$\frac{(m - \tau_{t-1})(m-1)^2}{(m - \tau_{t-1} + t - 2)^2} \geq \frac{m-1}{t-1}. \tag{6.8}$$

Hence, applying the inequality from Eq. (6.8),

$$k_{t-1}u_{t-1}^2 \geq \frac{n}{k}(1 - 3m^2\varepsilon)^{2d}\frac{c_t}{m}\left(\frac{m-1}{t-1}\right) \geq \frac{n}{k}(1 - 3m^2\varepsilon)^{2d}\left(\frac{t}{t-1}\right)\left(\frac{m-1}{m}\right).$$

If $2 \leq t < m$, then

$$\left(\frac{t}{t-1}\right)\left(\frac{m-1}{m}\right) \geq \left(\frac{m-1}{m-2}\right)\left(\frac{m-1}{m}\right) = 1 + \frac{1}{m^2 - 2m} \geq 1 + \frac{1}{m^2 - 1}.$$

If $2 \leq t = m$, and $c_t > t$, then

$$\left(\frac{c_t}{t-1}\right)\left(\frac{m-1}{m}\right) \geq \left(\frac{t+1}{t-1}\right)\left(\frac{m-1}{m}\right) = \frac{m+1}{m} \geq 1 + \frac{1}{m^2 - 1}.$$

In each of these two cases, we get

$$k_{t-1}u_{t-1}^2 \geq \frac{n}{k} \cdot (1 - 3m^2\varepsilon)^{2d}\left(1 + \frac{1}{m^2 - 1}\right).$$

By Bernoulli's inequality, since $3m^2\varepsilon < 1$,

$$(1-3m^2\varepsilon)^{2d} \geq (1-6dm^2\varepsilon) > 1 - \frac{1}{m^2} = \left(1 + \frac{1}{m^2-1}\right)^{-1}.$$

Therefore, if $c_t > t$ or $m > t$, then $k_{t-1}u_{t-1}^2 > n/k$ and so

$$f_1 \leq \frac{n}{k_{t-1}u_{t-1}^2} < k \qquad \Rightarrow \qquad f_1 \leq k-1.$$

Finally, assume $\tau_t\psi_{t-1} = c_t = t$ and $m = t$. We know from Lemma 6.3.1 that $\tau_{t-1} \geq t-1$. If $t < d$, then $b_t \geq 1$. So, by Terwilliger's inequality and Lemma 3.2.1,

$$\frac{k}{m} \geq (m - \tau_{t-1})\left(\frac{k}{m} + 1 - \psi_{t-1}\right) = b_{t-1} \geq c_{t-1} - c_t + b_t + \lambda + 2 \geq \lambda + 2 \geq \frac{k}{m} + 1,$$

which gives a contradiction with the assumption $t < d$. Therefore, $t = d$ and $c_d = d$, $m = d$. □

Now we are ready to prove Theorem 1.4.5 in the following equivalent form (see Lemma 3.2.4).

Theorem 6.3.6. *Let X be a distance-regular graph of diameter $d \geq 2$. Suppose that every neighborhood graph $X(v)$ is a disjoint union of m cliques. Moreover, assume $\mu \geq 2$, $c_t \leq \varepsilon k$ and $b_t \leq \varepsilon k$ for some $t \leq d$ and $\theta_1 \geq (1-\varepsilon)b_1$, where $0 < \varepsilon < 1/(6m^4 d)$.*

Then X is a Hamming graph $H(d,s)$, for $s = 1 + k/d$.

Proof. Pick some vertex v and let $X(v) = \bigcup_{i=1}^m C_i$, where C_i is a clique for every i. Since X is distance-regular, all C_i are of the same size $\lambda + 1$. Note that $\{v\} \cup C_i$ is a maximal clique in X of the size $k/m + 1$. Since $k \geq 1/\varepsilon > m^2$, by Proposition 3.1.5, X is geometric with smallest eigenvalue $-m$.

Hence, by Theorem 6.3.5, either we have $f_1 \leq k-1$, or $c_d = m = d$. If $f_1 \leq k-1$, by Theorem 5.2.3, $-1 - \dfrac{b_1}{\theta_1 + 1}$ is an eigenvalue of $X(v)$. However, $b_1 > 0$ and $\theta_1 > 0$, so $X(v)$

has an eigenvalue less than -1. This gives a contradiction with the assumption that $X(v)$ is a disjoint union of cliques.

Therefore, $c_d = m = d$. By Lemma 6.3.1, we get $\tau_i \geq i$ for all $i \in [d]$. At the same time, $d = c_d = \tau_d \psi_{d-1}$, so $\tau_d = d$ and $\psi_{d-1} = 1$. We immediately deduce $\tau_i = i$ for all $i \in [d]$. Assume that $\psi_{i-1} \geq 2$, while $\psi_i = 1$ for some $2 \leq i \leq d-1$. Then, we get a contradiction with

$$i + 1 = \psi_i \tau_{i+1} = c_{i+1} \geq c_i = \psi_{i-1} \tau_i \geq 2i.$$

Thus, $\psi_i = 1$ for every i. This means, that the intersection array of X coincides with the intersection array of the Hamming graph $H(d, 1 + k/d)$,

$$c_i = i \qquad \text{and} \qquad b_i = (d - i)\frac{k}{d}.$$

Using the characterization of Hamming graphs by their intersection array (Theorem 3.5.3), X is a Hamming graph or a Doob graph. Note that X may be a Doob graph, only if $1 + k/d = 4$, which is not possible as $k \geq 1/\varepsilon \geq 6d$. Therefore, X is a Hamming graph. □

Corollary 6.3.7. *Let X be a geometric distance-regular graph of diameter $d \geq 2$. Suppose that X has $\mu = 2$ and smallest eigenvalue $-m$. Take $0 < \varepsilon < 1/(6m^4 d)$. Assume, $c_t \leq \varepsilon k$ and $b_t \leq \varepsilon k$ for some $t \in [d]$, and $\theta_1 \geq (1 - \varepsilon) b_1$. Then X is a Hamming graph $H(d, s)$.*

Proof. By Lemma 3.2.1, $\mu = \tau_2 \psi_1$, and by Lemma 3.2.2, $\tau_2 \geq \psi_1$, so $\mu = 2$ implies $\tau_2 = 2$ and $\psi_1 = 1$. This means that the neighborhood graph of every vertex of X is a disjoint union of $-\theta_d = m$ cliques (Lemma 3.2.4). Therefore, the statement follows from the theorem above. □

6.A Appendix: Proof of inequality (6.8)

Below we prove the inequality used in the proof of Theorem 6.3.5.

Lemma 6.A.1. *Let $2 \le t \le x+1 \le m$ be integers, then*

$$\frac{(m-x)(m-1)^2}{(m-x+t-2)^2} \ge \frac{m-1}{t-1}. \tag{6.9}$$

Proof. Note that when $x = m-1$ the inequality is true, as $m-1 \ge t-1$. We can rewrite inequality (6.9) as

$$(m-x)(m-1)(t-1) \ge (m-x+t-2)^2,$$

$$m(m-1)(t-1) - x(m-1)(t-1) \ge x^2 - 2x(m+t-2) + (m+t-2)^2,$$

$$m(m-1)(t-1) - (m+t-2)^2 \ge x(x+m(t-3)-3t+5). \tag{6.10}$$

If $t \ge 4$, then $x \ge 3t-5-m(t-3)$. Indeed, for $t \ge 5$ this is true as

$$x \ge t-1 \ge 3t-5-2m \ge 3t-5-m(t-3),$$

and for $t = 4$ this holds as $x \ge t-1 = 3 \ge 7-m$. Thus, for $t \ge 4$ the maximal value of the RHS of inequality (6.10) is achieved at maximal value of x, i.e., when $x = m-1$. But as noted above, inequality (6.9) holds for $x = m-1$, and inequality (6.10) is equivalent to it.

The statement of the lemma is obvious if $t = 2$. Therefore, the only case we still need to check is $t = 3$. Since the desired inequality holds for $x = m-1$, we can assume $x \le m-2$. In this case inequality (6.10) follows from

$$2m(m-1) - (m+1)^2 = m^2 - 4m + 1 > (m-2)^2 - 4 \ge x^2 - 4 \ge x^2 - 4x. \qquad \square$$

CHAPTER 7

SPECTRAL GAP OF A DISTANCE-REGULAR GRAPH

In this section we give a bound on the spectral gap of a distance-regular graph in terms of its intersection numbers. The spectral gap bound will be used in Sections 8.1 and 8.4 to achieve motion lower bounds through Lemma 4.5.11 (the Spectral tool).

7.1 Approximation of the spectrum by the intersection numbers

Note, that if b_i and c_i are simultaneously small, then by monotonicity, b_j for $j \geq i$ and c_t for $t \leq i$ are small. Hence, the intersection matrix L_1

$$L_1 = \begin{pmatrix} a_0 & b_0 & 0 & 0 & ... \\ c_1 & a_1 & b_1 & 0 & ... \\ 0 & c_2 & a_2 & b_2 & ... \\ ... & & \vdots & & ... \\ ... & & 0 & c_d & a_d \end{pmatrix}. \tag{7.1}$$

is a small perturbation of a block diagonal matrix N, where one block is upper triangular and the other block is lower triangular. So the eigenvalues of N are just the diagonal entries.

To relate the eigenvalues of N to eigenvalues of L_1 we rely on the following result.

Theorem 7.1.1 ([Ostrowski, 1967, Appendix K]). *Let $A, B \in M_n(\mathbb{C})$. Let $\lambda_1, \lambda_2, ..., \lambda_n$ be the roots of the characteristic polynomial of A and $\mu_1, \mu_2, ..., \mu_n$ be the roots of the characteristic polynomial of B. Consider*

$$M = \max\{|(A)_{ij}|, |(B)_{ij}| : 1 \leq i, j \leq n\}, \quad \delta = \frac{1}{nM} \sum_{i=1}^{n} \sum_{j=1}^{n} |(A)_{ij} - (B)_{ij}|.$$

Then, there exists a permutation $\sigma \in S_n$ such that

$$|\lambda_i - \mu_{\sigma(i)}| \leq 2(n+1)^2 M \delta^{1/n}, \quad \textit{for all } 1 \leq i \leq n.$$

Lemma 7.1.2. *Let X be a distance-regular graph of diameter d. Denote by $\theta_0 > \theta_1 > \ldots > \theta_d$ all distinct eigenvalues of X. Suppose that $b_i \leq \varepsilon k$ and $c_i \leq \varepsilon k$ for some $i \leq d$ and $\varepsilon > 0$. Then*

$$|\theta_i - a_i| \leq 2(d+2)^2 \varepsilon^{\frac{1}{d+1}} k.$$

In particular, if furthermore $b_{i-1} \geq \alpha k$ and $c_{i+1} \geq \alpha k$, for some $\alpha > 0$ (here we define $c_{d+1} = k$), then the zero-weight spectral radius ξ of X satisfies

$$\xi \leq k(1 - \alpha + 2(d+2)^2 \varepsilon^{\frac{1}{d+1}}). \tag{7.2}$$

Proof. Let N be a matrix obtained from L_1 (see Eq. (7.1)) by replacing all b_s and c_t with 0 for $s \geq i$ and $t \leq i$. As in Theorem 7.1.1, consider

$$M = \max\{|(T)_{sj}|, |(N)_{sj}| : 1 \leq s, j \leq d\} = k,$$

$$\delta = \frac{1}{(d+1)M} \sum_{s=1}^{d+1} \sum_{j=1}^{d+1} |(T)_{sj} - (N)_{sj}| \leq \frac{(d+1)\varepsilon k}{(d+1)M} = \varepsilon.$$

Observe, that the diagonal entry a_i is the only non-zero entry in the i-th row of N. Furthermore, the upper-left $i \times i$ submatrix is upper triangular and $(d-i) \times (d-i)$ lower-right submatrix is lower triangular. Then the eigenvalues of N are equal to a_j for $0 \leq j \leq d$. Thus, the first part of the statement follows from Theorem 7.1.1.

The inequalities $b_{i-1} \geq \alpha k$ and $c_{i+1} \geq \alpha k$ imply that $a_j \leq k(1-\alpha)$ for $j \neq i$, while $a_i \geq (1-2\varepsilon)k$. Hence, since k is an eigenvalue of multiplicity 1 of X, the zero-weight spectral radius of X satisfies Eq. (7.2). □

We are going to use this result in Section 7.3 to prove Theorem 7.3.8.

7.2 A growth-induced tradeoff for the intersection numbers

Observation 7.2.1. *Let X be a graph. Denote by $\deg(v)$ the degree of a vertex v in X, and denote by $N(u,v)$ the set of common neighbors of vertices u and v in X. Then for all vertices u, v, w we have*

$$|N(u,v)| + |N(u,w)| \leq \deg(u) + |N(v,w)|.$$

Proof. The inequality above follows from the two obvious inclusions below

$$N(u,v) \cup N(u,w) \subseteq N(u), \qquad N(u,v) \cap N(u,w) \subseteq N(v,w). \qquad \square$$

Next, we prove the growth-induced tradeoff for the intersection numbers. Essentially, the theorem below claims that, if for some j, b_j is large (and therefore, by monotonicity, so are b_i for $i \leq j$) and c_{j+1} is small, then b_{j+1} and c_{j+2} cannot be small simultaneously. We use this inequality in Proposition 7.3.5 (and later in Theorem 7.3.8) to establish a lower bound on the spectral gap of a distance-regular graph under mild assumptions.

Theorem 7.2.2 (Growth-induced tradeoff)**.** *Let X be a distance-regular graph of diameter $d \geq 2$. Let $0 \leq j \leq d-2$. Assume $b_j > c_{j+1}$ and let $C = b_j/c_{j+1}$. Then for all $1 \leq s \leq j+1$ we have*

$$b_{j+1}\left(\sum_{t=1}^{s} \frac{1}{b_{t-1}} + \sum_{t=1}^{j+2-s} \frac{1}{b_{t-1}}\right) + c_{j+2}\sum_{t=1}^{j+1} \frac{1}{b_{t-1}} \geq 1 - \frac{4}{C-1}. \tag{7.3}$$

Remark 7.2.3. In applications we require the right-hand side to be bounded away from zero, i.e., C to be greater than some constant > 5. In the case when $b_j \geq \alpha k$ for some constant $\alpha > 0$, each reciprocal $1/b_t$ for $t \leq j$ is at most $1/(\alpha k)$. Thus, if the RHS is

bounded away from zero and d is bounded, we get a lower bound on b_{j+1} or c_{j+2} that is linear in k. We also note that $b_j/c_{j+1} = k_{j+1}/k_j$, where k_j is the size of the sphere of radius j in X. So the assumption says that significant growth occurs from radius j to radius $j+1$.

Remark 7.2.4. Graphs for which the lemma above gives a trivial bound, i.e., when the fraction b_j/c_{j+1} is bounded from above by a (small) constant, were studied in Park et al. [2013]. In particular, in this case one can prove upper bounds for the diameter of a graph.

Proof of Theorem 7.2.2. Consider the graph Y with the set of vertices $V(Y) = V(X)$, where a pair of distinct vertices u, v is adjacent if they are at distance $\mathrm{dist}(u, v) \leq j+1$ in X. We want to find the restriction on the parameters of X implied by Observation 7.2.1 applied to graph Y and vertices w, v at distance $j+2$ in X. Let λ_i^Y denote the number of common neighbors in Y for a pair of vertices u, v at distance i in X for $i \leq j+1$. Let μ_{j+2}^Y denote the number of common neighbors in Y for a pair of vertices u, v at distance $j+2$ in X. The monotonicity of sequences (b_i) and (c_i) implies $k_{i+1} \geq Ck_i$ for $i \leq j$. Thus, the degree of every vertex in Y satisfies

$$k^Y = \sum_{i=1}^{j+1} k_i \leq k_{j+1} \sum_{i=0}^{j} C^{-i} \leq k_{j+1} \frac{C}{C-1}.$$

Note, that $\sum_{t=0}^{d} p_{s,t}^i = k_s$. Hence, we have

$$\mu_{j+2}^Y = \sum_{1 \leq s,t \leq j+1} p_{s,t}^{j+2} \leq 2 \sum_{i=1}^{j} k_i + p_{j+1,j+1}^{j+2} \leq \frac{2}{C-1} k_{j+1} + p_{j+1,j+1}^{j+2},$$

$$\lambda_i^Y = \sum_{1 \leq s,r \leq j+1} p_{r,s}^i \geq \sum_{1 \leq s \leq j+1} p_{j+1,s}^i = k_{j+1} - \sum_{j+2 \leq s \leq d} p_{j+1,s}^i - p_{j+1,0}^i.$$

Now we are going to get some bounds on $\sum_{j+2 \leq s \leq d} p_{j+1,s}^i$. We use the following observation.

Suppose, that x, y are two vertices at distance i. Then there are exactly $\prod_{t=1}^{i} c_t$ paths of length i between x and y. Thus, $\left(\prod_{t=1}^{i} c_t\right) \sum_{s=j+2}^{d} p_{i,s}^{j+1}$ equals the number of paths of length i starting at v and ending at distance at least $j+2$ from u and at distance i from v, where $\mathrm{dist}(u,v) = j+1$ in X. We count such paths by considering possible choices of edges for a path at every step. At step t every such path should go from $N_{t-1}(v)$ to $N_t(v)$, hence there are at most b_{t-1} possible choices for a path at step t for $1 \leq t \leq i$. Moreover, since path should end up at distance at least $j+2$ from u, then for some $1 \leq t \leq i$ path should go from $N_{j+1}(u)$ to $N_{j+2}(u)$. Therefore, the number of paths that go from $N_{j+1}(u)$ to $N_{j+2}(u)$ at step t is at most $\left(\prod_{s=1}^{i} b_{s-1}\right) \frac{b_{j+1}}{b_{t-1}}$. Hence,

$$\sum_{s=j+2}^{d} p_{j+1,s}^{i} = \frac{k_{j+1}}{k_i} \sum_{s=j+2}^{d} p_{i,s}^{j+1} \leq \frac{k_{j+1}}{k_i} \sum_{t=1}^{i} \left(\prod_{s=1}^{i} b_{s-1}\right) \frac{b_{j+1}}{b_{t-1}} \left(\prod_{t=1}^{i} c_t\right)^{-1} = k_{j+1} \sum_{t=1}^{i} \frac{b_{j+1}}{b_{t-1}}.$$

Thus, in particular,

$$\lambda_i^Y \geq k_{j+1} \left(1 - \sum_{t=1}^{i} \frac{b_{j+1}}{b_{t-1}}\right) - p_{j+1,0}^{i}. \tag{7.4}$$

Similarly,

$$p_{j+1,j+1}^{j+2} \leq k_{j+1} \sum_{t=1}^{j+1} \frac{c_{j+2}}{b_{t-1}}.$$

Hence,

$$\mu_{j+2}^Y \leq k_{j+1} \left(\frac{2}{C-1} + \sum_{t=1}^{j+1} \frac{c_{j+2}}{b_{t-1}}\right). \tag{7.5}$$

By applying Observation 7.2.1 to vertices u, v, and w in Y, that satisfy $\mathrm{dist}(v,w) = j+2$, $\mathrm{dist}(u,v) = s$ and $\mathrm{dist}(w,u) = j+2-s$ in X, we get

$$k_{j+1} + \mu_{j+2}^Y \geq \lambda_s^Y + \lambda_{j+2-s}^Y. \tag{7.6}$$

The desired inequality (7.3) follows from Eq. (7.4), (7.5) and (7.6), as $p^i_{j+1,0} \leq 1 \leq k_{j+1}/C$.

□

7.3 Spectral gap bound

In this section we prove a lower bound on the spectral gap of distance-regular graphs of fixed diameter d with a dominant distance. We prove Theorem 1.4.1 in the equivalent formulation as Theorem 7.3.10.

Our key tool is the growth-induced tradeoff proven in the previous section, which will be applied in the following setup. Assume we know lower bounds of the form $b_i \geq \alpha_i k$ for the intersection numbers b_i with $i \leq j$. Our goal is to get a lower bound of the similar form either for b_{j+1}, or for c_{j+2}. We will argue, that if either $c_{j+2} \leq \varepsilon k$, or $b_{j+1} \leq \varepsilon k$, for a sufficiently small $\varepsilon > 0$, then either the second or the first summand of the LHS in inequality (7.3) is at most a δ-fraction of the LHS. Hence, the other summand is at least $(1-\delta)$-fraction of the LHS and we get a linear in k lower bound on either b_{j+1}, or c_{j+2}. The two sequences we define below are the coefficients in front of k in the bounds we get from inequality (7.3).

Definition 7.3.1. Let $0 \leq \delta < 1$. We say that $(\alpha_i)_{i=0}^{\infty}$ is the $FE(\delta)$ sequence, if $\alpha_0 = 1$ and for $j \geq 1$ the element α_j is defined by the recurrence

$$\alpha_{j+1} = (1-\delta)\left(\sum_{t=1}^{\left\lceil \frac{j+2}{2} \right\rceil} \frac{1}{\alpha_{t-1}} + \sum_{t=1}^{\left\lfloor \frac{j+2}{2} \right\rfloor} \frac{1}{\alpha_{t-1}}\right)^{-1}. \tag{7.7}$$

Let $\widehat{\alpha} = (\alpha_i)_{i=0}^{s}$ be a sequence. We say that $\widehat{\beta} = (\beta_i)_{i=2}^{s+2}$ is the $BE(\delta, \widehat{\alpha})$ sequence, if for $j \geq 2$ the element β_j is defined as

$$\beta_j = (1-\delta)\left(\sum_{t=0}^{j-2} \frac{1}{\alpha_t}\right)^{-1}. \tag{7.8}$$

If additionally, $\widehat{\alpha}$ is a prefix of $FE(\delta)$ sequence, then we will say that $\widehat{\beta}$ is the $BE(\delta)$ sequence.

Remark 7.3.2. FE stands for "forward expansion" and BE stands for "backward expansion".

Now we specify how small we expect ε to be in the argument above, so that one of the summands in the LHS of (7.3) is at most a δ-fraction of the LHS.

Definition 7.3.3. Let $\widehat{\alpha} = (\alpha)_{i=0}^{s}$ be a decreasing sequence of positive real numbers with $\alpha_0 = 1$. Let $0 < \delta < 1$, and $\widehat{\beta} = (\beta_i)_{i=2}^{s+2}$ be the corresponding $BE(\delta, \widehat{\alpha})$ sequence. We say that $\varepsilon > 0$ is $(\delta, j, \widehat{\alpha}, d)$-*compatible* for $j \leq s \leq d-2$, if ε satisfies

$$\left(\frac{\alpha_j - 5\varepsilon}{\alpha_j - \varepsilon} - 2\varepsilon \sum_{t=1}^{j+1} \frac{1}{\alpha_{t-1}} \right) > (1-\delta) \quad \text{and} \quad 2(d+2)^2 \varepsilon^{\frac{1}{d+1}} \leq \beta_{j+2}\delta. \tag{7.9}$$

Note that if ε is $(\delta, j, \widehat{\alpha}, d)$-compatible for $j \geq 1$, then it is $(\delta, (j-1), \widehat{\alpha}, d)$-compatible as well. Note also that the second condition on ε implies that $\delta > \varepsilon$ and $\beta_{j+2} > \varepsilon$, $\alpha_j > \varepsilon$.

Definition 7.3.4. We say that $\varepsilon > 0$ is (δ, d)-compatible, if it is $(\delta, d-2, \widehat{\alpha}, d)$-compatible for $FE(\delta)$ sequence $\widehat{\alpha}$. We introduce notation

$$EPS_\delta(d) = \sup\{\varepsilon \mid \varepsilon \text{ is } (\delta, d)\text{-compatible}\}.$$

In the proposition below we provide a formal version of the discussion at the beginning of this subsection.

Proposition 7.3.5. *Let X be a distance-regular graph of diameter $d \geq 2$. Fix an arbitrary $0 < \delta < 1$. Let $0 \leq j \leq d-2$ and $\widehat{\alpha} = (\alpha_i)_{i=0}^{j}$ be a decreasing sequence of positive real numbers. Consider corresponding $BE(\delta, \widehat{\alpha})$ sequence $\widehat{\beta}$ and $(\delta, j, \widehat{\alpha}, d)$-compatible $\varepsilon > 0$. Assume that the intersection numbers of X satisfy $c_{j+1} \leq \varepsilon k$ and $b_i \geq \alpha_i k$ for all $0 \leq i \leq j$. Then one of the following is true.*

1. *$b_{j+1} \geq \varepsilon k$ and $c_{j+2} \geq \varepsilon k$.*

2. *The zero-weight spectral radius ξ of X satisfies*

$$\xi \leq k(1-(1-\delta)\beta_{j+2}), \quad \textit{and} \quad c_{j+2} \geq \varepsilon k.$$

3. *Let $\alpha_{j+1} = (1-\delta)\left(\sum_{t=1}^{\lceil \frac{j+2}{2} \rceil} \frac{1}{\alpha_{t-1}} + \sum_{t=1}^{\lfloor \frac{j+2}{2} \rfloor} \frac{1}{\alpha_{t-1}}\right)^{-1}$, then $b_{j+1} \geq \alpha_{j+1}k$ and $c_{j+2} \leq \varepsilon k$.*

Proof. **Case 1.** Assume that $c_{j+2} \geq \beta_{j+2}k$.

If $b_{j+1} \geq \varepsilon k$, then statement 1 holds. Thus, suppose that $b_{j+1} \leq \varepsilon k$. Then we fall into the assumptions of Lemma 7.1.2 with $i = j+1$. Hence, the zero-weight spectral radius ξ of X satisfies

$$\xi \leq k(1-\min(\alpha_j, \beta_{j+2}) + 2(d+2)^2\varepsilon^{\frac{1}{d+1}}) \leq k(1-(1-\delta)\beta_{j+2}).$$

Note, that by definition, $\beta_{j+2} < \alpha_j$, so $\min(\alpha_j, \beta_{j+2}) = \beta_{j+2}$.

Case 2. Assume $\varepsilon k \leq c_{j+2} \leq \beta_{j+2}k$. Then, by Eq. (7.8) and Eq. (7.9),

$$\beta_{j+2} \leq \left(\frac{\alpha_j - 5\varepsilon}{\alpha_j - \varepsilon}\right)\left(\sum_{t=1}^{j+1} \frac{1}{\alpha_{t-1}}\right)^{-1} - 2\varepsilon.$$

Then, by Lemma 7.2.2 for $C = \alpha_j/\varepsilon$, we get $b_{j+1} \geq \varepsilon k$.

Case 3. Finally, assume that $c_{j+2} \leq \varepsilon k$. Then, since by Eq. (7.9),

$$0 < \alpha_{j+1} \leq \left(\frac{\alpha_j - 5\varepsilon}{\alpha_j - \varepsilon} - \varepsilon\sum_{t=1}^{j+1} \frac{1}{\alpha_{t-1}}\right)\left(\sum_{t=1}^{\lceil \frac{j+2}{2} \rceil} \frac{1}{\alpha_{t-1}} + \sum_{t=1}^{\lfloor \frac{j+2}{2} \rfloor} \frac{1}{\alpha_{t-1}}\right)^{-1},$$

Lemma 7.2.2 for $C \geq \alpha_j/\varepsilon$ implies $b_{j+1} \geq \alpha_{j+1}k$. $\square$

As an immediate corollary we get a lower bound on b_i for $i \leq t$ if $c_t \leq \varepsilon k$ is small. In the Appendix (Sec. 7.A) we give explicit lower bounds for elements of $BE(\delta)$ and $FE(\delta)$ sequences. Another corollary states that we can bound each eigenvalue of X if c_d is small.

Corollary 7.3.6. *Let X be a distance-regular graph of diameter $d \geq 2$. Fix an arbitrary $0 < \delta < 1$. Let $\widehat{\alpha} = (\alpha_i)_{i=0}^{\infty}$ be the $FE(\delta)$ sequence and ε be (δ, d)-compatible.*

Assume that $c_t \leq \varepsilon k$ for some $t \leq d$, then $b_i \geq \alpha_i k$ for all $0 \leq i \leq t-1$.

Corollary 7.3.7. *Fix any $0 < \delta < 1$. Let X be a distance-regular graph of diameter $d \geq 2$. Denote by $\theta_0 > \theta_1 > \ldots > \theta_d$ all distinct eigenvalues of X. Let $\widehat{\alpha} = (\alpha_i)_{i=0}^{\infty}$ be the $FE(\delta)$ sequence and ε be (δ, d)-compatible.*

Assume that $c_d \leq \varepsilon k$, then $\theta_i \leq (1-(1-\delta)\alpha_{d-i})k$ for all $1 \leq i \leq d$.

Proof. Follows from Corollary 7.3.6 and Lemma 7.1.2. $\square$

The next theorem is one of the key ingredients of the proof of our main result on the motion of distance-regular graphs (Theorem 8.3.1). It says that for a primitive distance-regular graph either the minimal distinguishing number is linear, or the spectral gap is large.

Theorem 7.3.8. *For every $d \geq 2$ there exist $\varepsilon = \varepsilon(d) > 0$ and $\eta = \eta(d) > 0$ such that for every distance-regular graph X of diameter d one of the following is true.*

1. *For some $0 \leq i \leq d-1$, we have $b_i \geq \varepsilon k$ and $c_{i+1} \geq \varepsilon k$.*

2. *The zero-weight spectral radius of X satisfies $\xi \leq k(1-\eta)$.*

Proof. Fix $\delta \in (0,1)$. Let $\widehat{\alpha} = (\alpha_i)_{i=0}^{\infty}$ be the $FE(\delta)$ sequence and $\widehat{\beta} = (\beta_i)_{i=2}^{\infty}$ be the $BE(\delta)$ sequence and ε be (δ, d)-compatible. Set $\eta = (1-\delta)\min(\alpha_{d-1}, \beta_d)$.

Define $c_{d+1} = k$. Let i be the unique index such that $c_{i+1} > \varepsilon k$, while $c_i \leq \varepsilon k$. If $b_i \geq \varepsilon k$, then first statement is true. So assume that $b_i \leq \varepsilon k$. By Corollary 7.3.6, for every $j \leq i-1$ we have $b_j \geq \alpha_j k$. Thus, by Proposition 7.3.5, if $i \leq d-1$, then

$$\xi \leq k(1-(1-\delta)\beta_{i+1}) \leq k\,(1-(1-\delta)\beta_d) \leq k(1-\eta).$$

If $i = d$, using that $b_j \geq \alpha_j k \geq \alpha_{d-1} k$ for $j \leq d-1$, we get

$$\xi \leq k(1-(1-\delta)\alpha_{d-1}) \leq k(1-\eta). \qquad \square$$

Remark 7.3.9. In the theorem above one can set

$$\eta = \frac{1}{4} d^{-(1+\log_2 d)} \quad \text{and} \quad \varepsilon = 200^{-(d+1)} d^{-(d+1)(\log_2(d)+3)}.$$

Proof. The proof is based on the explicit bound on the elements of $FE(\delta)$, $BE(\delta)$ and EPS_δ sequences given in the Appendix (Lemmas 7.A.1 and 7.A.2). Note that in Theorem 7.3.8 η is chosen as $\eta = (1-\delta)\min(\alpha_{d-1}, \beta_d)$. Fix $\delta = 1/9$. Using Lemma 7.A.1 we get

$$\alpha_{d-1} \geq \frac{(1-\delta)^2}{2}(d-1)^{-\log_2(d-1)} \quad \text{and} \quad \beta_d \geq \frac{(1-\delta)^3}{2(d-1)}(d-2)^{-\log_2(d-2)}.$$

Hence, we obtain $\eta \geq \frac{1}{4} d^{-(1+\log_2 d)}$. Moreover, by Lemma 7.A.2, $\varepsilon = 200^{-(d+1)} d^{-(d+1)(\log_2(d)+3)}$ is (δ, d)-compatible. $\square$

Finally, we prove our main theorem on spectral expansion.

Theorem 7.3.10. *For every $d \geq 2$ there exist $\epsilon = \epsilon(d) > 0$ and $\eta = \eta(d) > 0$ such that the following holds. Let X be a distance-regular graph of diameter d. If $k_t \geq (1-\epsilon)n$ for some $t \in [d]$, then the zero-weight spectral radius of X satisfies $\xi \leq k(1-\eta)$.*

Proof. Let $\varepsilon = \varepsilon(d)$ and $\eta = \eta(d)$ be constants provided by Theorem 7.3.8. Assume that for

some i we have $b_i \geq \varepsilon k$ and $c_{i+1} \geq \varepsilon k$. Let j be the smallest index for which $c_{j+1} \geq \varepsilon k$, then by monotonicity $b_j \geq \varepsilon k$.

If $t \leq j-1$, then $k_{t+1} = \dfrac{b_t}{c_{t+1}} k_t \geq \dfrac{b_{j-1}}{c_j} k_t > \dfrac{\varepsilon k}{\varepsilon k} k_t = k_t$. Therefore, if k_s is maximal, then $s \geq j$. Observe that $k_{j+1} \geq b_j k_j / c_{j+1} \geq \varepsilon k k_j / k = \varepsilon k_j$. Moreover, if $t \geq j$, then

$$k_{t+1} = \frac{b_t}{c_{t+1}} k_t \leq \frac{k}{\varepsilon k} k_t = \frac{k_t}{\varepsilon}. \tag{7.10}$$

Let k_s be the maximal distance degree. Note that $j < d$ as $b_j \geq \varepsilon k$. Thus, if $s = j$, then $k_{s+1} \geq \varepsilon k_s$, else $s > j$ and $k_{s-1} \geq \varepsilon k_s$, by Eq. (7.10). Define $\epsilon = \epsilon(d) = \varepsilon/(1+\varepsilon)$. Hence,

$$k_s = n - \sum_{t \neq s} k_t < n - \varepsilon k_s \quad \Rightarrow \quad k_s < \left(1 - \frac{\varepsilon}{1+\varepsilon}\right) n = (1-\epsilon) n.$$

Therefore, if $k_s \geq (1-\epsilon)n$ for some s, then there is no i such that $b_i \geq \varepsilon k$ and $c_{i+1} \geq \varepsilon k$. Hence, by Theorem 7.3.8, $\xi \leq k(1-\eta)$. □

We note that we do not exclude the elusive case $\mu = 1$, for which almost no classification results are known, and which is known to be a difficult case in various circumstances. We use this theorem to handle $\mu = 1$ case in the proof of Theorem 1.2.6.

Remark 7.3.11. Note that the conclusion of the theorem above is that X is a spectral η-expander. Recall that a combinatorial edge expansion of a graph is measured by the Cheeger constant, which for a graph $X = (V, E)$ is defined as

$$h(X) = \min \left\{ \frac{E(S, V \setminus S)}{|S|} \mid \text{for } S \subset V,\ 2|S| \leq |V| \right\},$$

where $E(S, T)$ denotes the number of edges between the sets of vertices S and T in X.

For a k-regular graph the Cheeger inequality shows that the Cheeger constant is controlled

by the second largest eigenvalue

$$(k-\xi_2)/2 \leq h(X) \leq \sqrt{2k(k-\xi_2)}.$$

Thus, if X is a spectral η-expander, then $h(X) \geq \eta k$, and so X is a good edge expander.

Note also that in the case when the difference $(k-\xi_2)$ is $o(k)$ the lower and upper bound given by the Cheeger inequality asymptotically differ by more than a constant. Recently, Qiao et al. [2020] conjectured that for distance-regular graphs the lower bound is always tight up to a constant. More precisely, they conjectured that

$$(k-\xi_2)/2 \leq h(X) \leq (k-\xi_2)$$

for all distance-regular graphs. They verified this claim for the known infinite families of distance-regular graphs, for strongly regular graphs and many other special cases.

7.A Appendix: Explicit bounds for $FE(\delta)$ and $BE(\delta)$

In this section we compute expilicit lower bounds for $BE(\delta)$, $FE(\delta)$ and EPS_δ sequences.

Lemma 7.A.1. *Fix $0 < \delta \leq \frac{1}{9}$. Let $(\alpha_i)_{i=0}^{\infty}$ be the $FE(\delta)$ sequence and the corresponding $BE(\delta)$ sequence $(\beta_i)_{i=2}^{\infty}$. Then for $j \geq 1$*

$$\alpha_j \geq \frac{(1-\delta)^2}{2} j^{-\log_2(j)} \quad \textit{and} \quad \beta_{j+2} \geq \frac{(1-\delta)^3}{2(j+1)} j^{-\log_2(j)}.$$

Proof. We prove the statement of the lemma by induction. Indeed, for $j = 1, 2$ we have $\alpha_1 = \frac{1-\delta}{2}$ and $\alpha_2 \geq \frac{(1-\delta)^2}{4}$, so the inequality is true. For $j \geq 2$, we have

$$\alpha_{j+1} = (1-\delta)\left(\sum_{t=1}^{\lceil \frac{j+2}{2} \rceil} \frac{1}{\alpha_{t-1}} + \sum_{t=1}^{\lfloor \frac{j+2}{2} \rfloor} \frac{1}{\alpha_{t-1}}\right)^{-1} \geq \frac{(1-\delta)}{j+2} \alpha_{\lceil \frac{j}{2} \rceil} \geq$$

$$\geq \frac{(1-\delta)^3}{2(j+2)}\left(\frac{j+1}{2}\right)^{-\log_2\left(\frac{j+1}{2}\right)} = \frac{(1-\delta)^3 2^{\log_2\left(\frac{j+1}{2}\right)}}{2(j+2)}(j+1)^{-\log_2(j+1)+1} =$$

$$= \frac{(1-\delta)(j+1)^2}{2(j+2)}\frac{(1-\delta)^2}{2}(j+1)^{-\log_2(j+1)} \geq \frac{(1-\delta)^2}{2}(j+1)^{-\log_2(j+1)}.$$

Thus,

$$\beta_{j+2} = (1-\delta)\left(\sum_{t=0}^{j}\frac{1}{\alpha_t}\right)^{-1} \geq \frac{1-\delta}{j+1}\alpha_j \geq \frac{(1-\delta)^3}{2(j+1)}j^{-\log_2(j)}.$$

□

Lemma 7.A.2. *Let* $0 < \delta \leq 1/9$ *and* $d \geq 3$. *Then*

$$EPS_\delta(d) \geq \left(\frac{\delta}{22}\right)^{(d+1)} d^{-(d+1)(3+\log_2 d)}.$$

Proof. Note that for the inequality $2(d+2)^2\varepsilon^{\frac{1}{d+1}} \leq \beta_{d+2}\delta$ to be satisfied it is enough to have

$$\varepsilon \leq \left(\frac{2^7\delta d^{-\log_2 d}}{9^3(d+1)(d+2)^2}\right)^{d+1} \leq \left(\frac{\delta\beta_{d+2}}{2(d+2)^2}\right)^{d+1}.$$

In particular, this is true if

$$0 < \varepsilon \leq \left(\frac{\delta}{22}\right)^{(d+1)} d^{-(d+1)(3+\log_2 d)}.$$

To check that the other condition on ε is satisfied, note that such choice of ε satisfies $\varepsilon < \alpha_{d-2}/2$. Thus we have

$$\left(\frac{\alpha_{d-2}-5\varepsilon}{\alpha_{d-2}-\varepsilon} - 2\varepsilon\sum_{t=1}^{d-1}\frac{1}{\alpha_{t-1}}\right) \geq 1 - 10\alpha_{d-2}^{-1}\varepsilon - 2d\alpha_{d-2}^{-1}\varepsilon \geq$$

$$\geq 1 - \frac{2\varepsilon}{(1-\delta)^2}(2d+10)d^{\log_2 d} \geq 1 - 22d^{1+\log_2 d}\varepsilon > (1-\delta).$$

□

CHAPTER 8

MOTION OF DISTANCE-REGULAR GRAPHS

In this chapter we confirm the following conjecture on motion of distance-regular graphs of bounded diameter (the metric case of the motion part of Conjecture 1.2.13).

Conjecture 8.0.1 (Babai)**.** *For every $d \geq 3$ there exists $\gamma_d > 0$ such that for every primitive distance-regular graphs X of diameter d on n vertices either*

$$\text{motion}(X) \geq \gamma_d n,$$

or X is a Johnson graph, or a Hamming graph.

8.1 Motion of primitive non-geometric distance-regular graphs

Prior to proving our main result on motion of primitive distance-regular graphs (Theorem 8.3.1) in Section 8.3, we study the minimal distinguishing number of distance-regular graphs. In the cases, when either there is no dominant distance (Proposition 8.1.6), or when the degree of a vertex is linear in the number of vertices (Proposition 8.1.4), we show a lower bound on the minimal distinguishing number that is linear in the number of vertices.

8.1.1 Case of a large vertex degree

In this section we study distance-regular graphs with a large vertex degree.

Lemma 8.1.1. *Let X be a distance-regular graph of diameter $d \geq 2$.*

1. *The parameters of X satisfy $k - \mu \leq 2(k - \lambda)$.*

2. *If $a_2 \neq 0$, then they also satisfy $k - \lambda \leq 2(k - \mu)$.*

Proof. The first statement follows from Observation 7.2.1 applied to vertices v and w at distance 2 in X and their common neighbor u.

Suppose, that $a_2 \neq 0$, then for a vertex u there exist two adjacent vertices v and w at distance 2 from u. So, the second statement follows from Observation 7.2.1 as well. □

Every pair of distinct vertices in a distance-regular graph has λ, or μ, or 0 common neighbors, if distance between them is 1, 2, or at least 3, correspondingly. Therefore, every pair of distinct vertices in a distance-regular graph is distinguished by at least $2(k - \max(\lambda, \mu))$ vertices. Combining this with the previous lemma we get the following bound.

Lemma 8.1.2. *Let X be a distance-regular graph of diameter $d \geq 2$. Then every pair of distinct vertices is distinguished by at least $k - \mu$ vertices.*

Proof. Every pair of vertices $u, v \in X$ is distinguished by at least $|N(u) \triangle N(v)| = 2(k - |N(u) \cap N(v)|)$ vertices. Thus, by Lemma 8.1.1, we get $2(k - \max(\lambda, \mu)) \geq k - \mu$. □

Next, we bound μ and λ away from k.

Lemma 8.1.3 ([Brouwer and Koolen, 2009, Lemmas 3.1, 3.14])**.** *Let X be a distance-regular graph of diameter $d \geq 3$. Then $\lambda \leq 2k/3$. Additionally, if X is primitive or $d \geq 4$, then $\mu \leq \frac{k}{2}$.*

Proposition 8.1.4. *Let X be a distance-regular graph of diameter $d \geq 3$. Suppose $k > n\gamma > 2$ for some $\gamma > 0$. If X is not a bipartite graph, then* $\mathrm{motion}(X)$ *is at least $\gamma n/3$.*

Proof. Suppose, that X is primitive, or $d \geq 4$. Then, by Lemma 8.1.3, $\mu \leq k/2$, and the result follows from Lemma 8.1.2. If $d = 3$ and X is antipodal not bipartite, then $a_2 \neq 0$ (see e.g. [Brouwer et al., 1989, p. 431], so the result follows from Lemma 8.1.1 and Lemma 8.1.3. □

Finally, we show that if k is small compared to n, then μ is small compared to k.

Lemma 8.1.5. *The parameters of a distance-regular graph of diameter $d \geq 2$ satisfy*

$$\mu < k \cdot \max\left(\frac{d-1}{r-1}, \left(\frac{d}{r}\right)^{\frac{1}{d-1}}\right),$$

where $r = \dfrac{n-1}{k}$.

Proof. Recall that the sequences (b_i) and (c_i) are monotone, so $b_i \leq b_1 = k - \lambda - 1$ and $c_{i+1} \geq \mu$ for $1 \leq i \leq d-1$. Thus, $k_{i+1} \leq k_i \dfrac{k-\lambda-1}{\mu}$. Hence,

$$n = \sum_{i=0}^{d} k_i \leq 1 + \sum_{i=0}^{d-1} k \left(\frac{k-\lambda-1}{\mu}\right)^i.$$

If $k - \lambda - 1 \leq \mu$, then $n \leq 1 + k + (d-1)k\dfrac{k-\lambda-1}{\mu}$, so $\mu < \dfrac{d-1}{r-1}k$.

Otherwise, we have $n \leq 1 + dk\left(\dfrac{k-\lambda-1}{\mu}\right)^{d-1}$, so $\mu < \left(\dfrac{d}{r}\right)^{\frac{1}{d-1}} k$. □

8.1.2 Motion of primitive distance-regular graphs

Recall that a distance-regular graph X is primitive if the distance-i graph X_i is connected for every $1 \leq i \leq d$. Recall also that by Definition 4.5.3,

$$D_{\min}(X) = \min_{u \neq v \in V} D(u, v),$$

where $D(u, v)$ denotes the number of vertices that distinguish u and v.

Proposition 8.1.6. *Let X be a primitive distance-regular graph of diameter $d \geq 2$ on n vertices. Fix some positive real number $\alpha > 0$. Suppose that for some $1 \leq j \leq d-1$ inequalities $b_j \geq \alpha k$ and $c_{j+1} \geq \alpha k$ hold. Then*

$$D_{\min}(X) \geq \frac{\alpha}{d} n.$$

Proof. Since the sequence (b_i) is non-increasing, if $t \leq j$, then $a_t = k - b_t - c_t \leq (1-\alpha)k$. Similarly, the sequence (c_i) is non-decreasing, so if $t > j$, then $a_t = k - b_t - c_t \leq (1-\alpha)k$.

Consider any pair of adjacent vertices $u, v \in X$. If vertex x does not distinguish u and v, then $\mathrm{dist}(u,x) = \mathrm{dist}(v,x) = t$ for some $1 \leq t \leq d$. Note, that for a given t there are $p^1_{t,t}$ such vertices x and

$$p^1_{t,t} = p^t_{t,1}\frac{k_t}{k} = k_t\frac{a_t}{k} \leq (1-\alpha)k_t.$$

Clearly, $\sum\limits_{i=1}^{d} k_i = n-1$. Hence, every pair of adjacent vertices is distinguished by at least

$$n - \sum_{t=1}^{d}(1-\alpha)k_t \geq n - (1-\alpha)n = \alpha n$$

vertices. Finally, the result follows from Lemma 4.5.2. □

8.1.3 Reduction to geometric graphs

In the theorem below we prove our main result on the motion of non-geometric distance-regular graphs.

Theorem 8.1.7. *For any $d \geq 3$ there exist $\gamma_d > 0$ and a positive integer m_d, such that for every primitive distance-regular graph X of diameter d with n vertices either*

$$\mathrm{motion}(X) \geq \gamma_d n,$$

or X is geometric with smallest eigenvalue at least $-m_d$.

Furthermore, one can set $m_d = \left\lfloor 5d^{\log_2 d+1} \right\rfloor$.

Proof. By Theorem 7.3.8, there exist constants $\varepsilon > 0$ and $\eta > 0$, which depend only on d, such that

- either $b_i \geq \varepsilon k$ and $c_{i+1} \geq \varepsilon k$,

- or the zero-weight spectral radius of X satisfies $\xi \leq k(1-\eta)$.

In the first case, by Proposition 8.1.6, we obtain

$$\text{motion}(X) \geq \frac{\varepsilon}{d} n.$$

Hence, assume that $\xi \leq k(1-\eta)$. For convenience, we additionally assume $\eta \leq 1/7$.

Case 1. Suppose that $\mu > \eta^3 k$. Then, by Lemma 8.1.5, $n \leq \max\left(d, 2\left(\eta^{-3} d\right)^{d-1}\right) k + 1$. Therefore, by Proposition 8.1.4,

$$\text{motion}(X) \geq \frac{1}{7}\left(\eta^3 d^{-1}\right)^{d-1} n.$$

Case 2. Suppose that $\lambda < \frac{9}{10}\eta k$ and $\mu \leq \eta^3 k$. Then every pair of distinct vertices in X has at most $q(X) = \max(\lambda, \mu) \leq 9\eta k/10$ common neighbors. Therefore, by Lemma 4.5.11,

$$\text{motion}(X) \geq \frac{\eta}{10} n.$$

Case 3. Suppose that $\lambda \geq \frac{9}{10}\eta k$ and $\mu \leq \eta^3 k$. Let m be the integer that satisfies

$$(m-1)(\lambda+1) < k \leq m(\lambda+1).$$

The assumption on λ implies $m - 1 \leq \frac{10}{9}\eta^{-1}$. We additionally assumed $\eta \leq 1/7$, so

$$\frac{1}{2} m(m+1)\mu \leq \frac{1}{2}\left(\frac{10}{9}\eta^{-1} + 1\right)\left(\frac{10}{9}\eta^{-1} + 2\right)\mu \leq \frac{9}{10}\eta^{-2}\mu \leq \frac{9}{10}\eta k \leq \lambda.$$

Thus, by Corollary 3.1.6, the graph X is a geometric distance-regular graph with smallest eigenvalue $-m$.

Finally, we note that we can take $m_d = \left\lfloor \frac{10}{9}\eta^{-1} + 1 \right\rfloor$ and

$$\gamma_d = \min\left(\frac{\varepsilon}{d}, \frac{1}{7}\left(\eta^3 d^{-1}\right)^{d-1}, \frac{\eta}{10}\right).$$

Furthermore, by Remark 7.3.9, m_d can be taken as $m_d = \left\lfloor 5d^{\log_2 d+1} \right\rfloor$. □

Remark 8.1.8. A bit more careful computations show that one can in fact set

$$m_d = \lceil \max\left(2(d-1)(d-2)^{\log_2(d-2)}, 2(d-1)^{\log_2(d-1)}\right) \rceil.$$

In particular, for $d = 3$ this estimate gives upper bound $m_d \leq 4$.

8.2 Geometric distance-regular graphs with $\mu = 1$

In this section we discuss the motion of geometric distance-regular graphs with $\mu = 1$.

8.2.1 Case of $m \geq 3$

In the case of $\mu = 1$ our strategy is to show that the dual graph $\widetilde{X}$ of X has motion linear in the number of vertices of $\widetilde{X}$. After this we deduce that X itself has motion linear in the number of its vertices.

Let X be a geometric distance-regular graph with $\mu = 1$ and let $\widetilde{X}$ be its dual graph. By Lemma 3.3.2, every vertex of $\widetilde{X}$ has degree $\widetilde{k} = k\frac{m-1}{m} + (m-1)$ and every pair of adjacent vertices of $\widetilde{X}$ has $\widetilde{\lambda} = m - 2$ common neighbors. Every pair of vertices at distance two in $\widetilde{X}$ has $\widetilde{\mu} = 1$ common neighbours. Indeed, if there are at least two edges between a pair of cliques C_1 and C_2, that do not share a vertex, then either $\psi_1 \geq 2$ for X, or there is an induced quadrangle. In both cases, we get $\mu \geq 2$, and we reach a contradiction.

Since $q(\widetilde{X}) := \max(\widetilde{\mu}, \widetilde{\lambda})$ is small, we are going to show that Lemma 4.5.11 can be applied efficiently. For this, it is sufficient to show that $\widetilde{X}$ has a linear in $\widetilde{k}$ spectral gap $\widetilde{k} - \xi(\widetilde{X})$.

First, we bound the zero-weight spectral radius of a geometric distance-regular graph using Theorem 7.3.8.

Using the relationship between the spectrum of the geometric graph X and its dual graph $\widetilde{X}$ we get the following corollary.

Lemma 8.2.1. *Let X be a geometric distance-regular graph of diameter $d \geq 2$ with smallest eigenvalue $-m$, where $m \geq 3$. Let $\widetilde{X}$ be its dual graph. Let $\varepsilon = \varepsilon(d)$ and $\eta = \eta(d) \leq 1/2$ be constants provided by Theorem 7.3.8. Assume $k \geq m^2$, $c_t \leq \varepsilon k$ and $b_t \leq \varepsilon k$ for some $t \in [d]$. Then the zero-weight spectral radius of $\widetilde{X}$ satisfies*

$$\xi(\widetilde{X}) \leq \widetilde{k}(1-\eta).$$

Proof. The assumption $k \geq m^2$ implies $\widetilde{k} < k$. Let $\widetilde{\theta}_1$ and $\widetilde{\theta}_{\min}$ denote the second largest and the smallest eigenvalues of $\widetilde{X}$. Then the statement of the lemma follows from the following two inequalities implied by Theorem 7.3.8 and Lemma 3.3.3,

$$\widetilde{\theta}_1 \leq (1-\eta)k - \frac{k}{m} + m - 1 = \widetilde{k} - \eta k \leq \widetilde{k}(1-\eta),$$

$$\widetilde{\theta}_{\min} \geq -m - \frac{k}{m} + m - 1 = -\frac{k}{m} - 1 = -\frac{\widetilde{k}}{m-1} \geq -\widetilde{k}(1-\eta).$$

□

Thus, using Lemma 4.5.11 we get a linear in $|V(\widetilde{X})|$ lower bound on the motion of $\widetilde{X}$.

Proposition 8.2.2. *Let X be a geometric distance-regular graph of diameter $d \geq 2$ with $\mu = 1$ and smallest eigenvalue $-m$, where $m \geq 3$. Let $\varepsilon = \varepsilon(d)$ and $\eta = \eta(d) \leq 1/2$ be constants provided by Theorem 7.3.8. Assume $c_t \leq \varepsilon k$ and $b_t \leq \varepsilon k$ for some $t \in [d]$, and $k \geq \max(4m/\eta, m^2)$. Let $\widetilde{X}$ be the dual graph of X. Then*

$$\text{motion}(\widetilde{X}) \geq \frac{\eta}{2}|V(\widetilde{X})|.$$

Proof. Since $\mu = 1$, by the discussion after Lemma 4.5.11, the maximal number of common neighbors of a pair of distinct vertices of $\widetilde{X}$ is equal $q(\widetilde{X}) = \max(\widetilde{\lambda}, \widetilde{\mu}) = \max(m-2, 1)$. Note that $\widetilde{k} \geq \dfrac{m-1}{m}k \geq \dfrac{k}{2}$. So $\eta\widetilde{k} \geq 2m \geq 2q(\widetilde{X})$. Hence,

$$\xi(\widetilde{X}) + q(\widetilde{X}) \leq (1-\eta)\widetilde{k} + \frac{\eta}{2}\widetilde{k} = \left(1 - \frac{\eta}{2}\right)\widetilde{k}.$$

Therefore, the statement of the proposition follows from Lemma 4.5.11. □

We show that this implies that motion(X) is linear in $n = |V(X)|$.

Lemma 8.2.3. *Let $\mathcal{F}$ be a collection of size-s subsets of a set Ω such that every element of Ω is in m sets from $\mathcal{F}$ and every pair of distinct sets in $\mathcal{F}$ intersects in at most one element of Ω. Let σ be a permutation of Ω which respects $\mathcal{F}$, i.e., for every $C \in \mathcal{F}$ its image $\sigma(C)$ is in $\mathcal{F}$, too. Assume that at most $\alpha|\mathcal{F}|$ sets $C \in \mathcal{F}$ are fixed by σ, then at most $\left(\alpha + \dfrac{1-\alpha}{s}\right)|\Omega|$ elements of Ω are fixed by σ.*

Proof. Note that if $C \in \mathcal{F}$ is not fixed as a set by σ, then $|\sigma(C) \cap C| \leq 1$, as $\sigma(C) \in \mathcal{F}$ too. Hence, at most one element $x \in \Omega$ of C is fixed by σ.

Now let us count the number of pairs (C, v), such that $v \in C$ and $\sigma(v) \neq v$. We just argued that each of at least $(1-\alpha)|\mathcal{F}|$ sets in $\mathcal{F}$ has at least $(s-1)$ elements that are not fixed by σ. Therefore, the desired number of pairs is at least $(1-\alpha)|\mathcal{F}|(s-1)$. Note that every element of Ω belongs to m sets in $\mathcal{F}$. Therefore, the number of elements of Ω not fixed by σ is at least $(1-\alpha)|\mathcal{F}|(s-1)/m$.

Using that every set in $\mathcal{F}$ has s elements and every element belongs to m sets in $\mathcal{F}$, we deduce that $s|\mathcal{F}| = m|\Omega|$. Therefore, the number of elements of Ω not fixed by σ is at least $(1-\alpha)\dfrac{(s-1)}{s}|\Omega|$. □

Corollary 8.2.4. *Let X be a non-complete geometric distance-regular graph on n vertices and let $\widetilde{X}$ be its dual graph on $\widetilde{n}$ vertices. Assume that* motion$(\widetilde{X}) \geq \gamma\widetilde{n}$, *then* motion$(X) \geq \dfrac{\gamma}{2}n$.

Proof. Let σ be a non-identity element of $\mathrm{Aut}(X)$. Then σ maps Delsarte cliques to Delsarte cliques. Thus σ induces an automorphism $\widetilde{\sigma}$ of $\widetilde{X}$. Note that if X is non-complete and geometric, then every vertex in X is uniquely determined by the set of Delsarte cliques that contain it. Hence, if σ is non-identity, then $\widetilde{\sigma}$ is non-identity as well. So by assumptions of the corollary, $\widetilde{\sigma}$ fixes at most $(1-\gamma)\widetilde{n}$ vertices of $\widetilde{X}$. Using that every Delsarte clique is of size at least 2, by the previous lemma, we get that σ fixes at most $(1-\gamma/2)n$ vertices. □

We summarize the discussion of this section in the following theorem.

Theorem 8.2.5. *Let X be a geometric distance-regular graph of diameter $d \geq 2$ on n vertices. Suppose $\mu = 1$ and the smallest eigenvalue of X is $-m$, where $m \geq 3$. Let $\varepsilon = \varepsilon(d)$ and $\eta = \eta(d) < 1/2$ be constants provided by Theorem 7.3.8. Assume $c_t \leq \varepsilon k$ and $b_t \leq \varepsilon k$ for some $t \in [d]$, and $k \geq \max(4m/\eta, m^2)$. Then*

$$\mathrm{motion}(X) \geq \frac{\eta}{4} n.$$

Proof. Let $\widetilde{X}$ be a dual graph of X on $\widetilde{n}$ vertices. Then, Proposition 8.2.2 implies $\mathrm{motion}(\widetilde{X}) \geq (\eta/2) \cdot \widetilde{n}$. Therefore, the statement of the theorem follows from Corollary 8.2.4. □

8.2.2 Distance-regular line graphs with $\mu = 1$

Note that, by Lemma 3.1.8, geometric distance-regular graphs with smallest eigenvalue -2 are line graphs. Thus, we can use the following result of Mohar and Shawe-Taylor [1985].

Definition 8.2.6. A distance-regular graph of diameter d with parameters

$$k = s(t+1),\ \lambda = s-1,\ c_i = 1 \text{ and } b_i = k - s \text{ (for } i = 1, \ldots, d-1),\ c_d = t+1$$

is called a *generalized $2d$-gon of order (s,t)*.

Theorem 8.2.7 ([Mohar and Shawe-Taylor, 1985, Theorem 3.4]). *Suppose the line graph $L(Y)$ of a graph Y is distance-regular. Then, either Y is a Moore graph, or Y is a generalized $2d$-gon of order $(1, s)$ for some $s \geq 1$, or $Y = K_{1,s}$ for $s \geq 1$.*

By the Hoffman-Singleton theorem, it is known that a Moore graph is either a complete graph, a polygon, or it is the Petersen graph ($k = 3$), the Hoffman-Singleton graph ($k = 7$), or it has degree $k = 57$ and diameter $d = 2$.

Note that a generalized $2d$-gon of order $(1, s)$ has intersection numbers $a_i = 0$ for all $i \in [d]$. Thus it is bipartite. Recall, that each of two connected components of the distance-2 graph X_2 of a bipartite distance-regular graph X is called a *halved graph*.

Fact 8.2.8 (see [Brouwer et al., 1989, Theorem. 6.5.1]). If X is a generalized $2d$-gon of order $(1, s)$, then d is even and its halved graph is a generalized d-gon of order (s, s).

A celebrated theorem of Feit and Higman [1964] asserts that apart from polygons, generalized $2d$-gons exist only for $2d \in \{4, 6, 8, 12\}$.

Theorem 8.2.9 (Feit and Higman [1964]). *A generalized $2d$-gon of order (s, t) exists only for $2d \in \{4, 6, 8, 12\}$ unless $s = t = 1$. If $s > 1$, then $2d \neq 12$.*

Finally, we use the following bound on the zero-weight spectral radius of generalized $2d$-gon of order (s, s) for $2d \leq 6$.

Fact 8.2.10 ([Brouwer et al., 1989, Table 6.4]). Let X be a generalized $2d$-gon of order (s, s) for $2d \leq 6$, $s > 1$. Then the zero-weight spectral radius of X satisfies $\xi(X) \leq 2s$.

Proposition 8.2.11. *Let X be a geometric distance-regular graph of diameter $d \geq 2$ on n vertices. Suppose $\mu = 1$, $k > 4$ and the smallest eigenvalue of X is -2. Then*

$$\text{motion}(X) \geq \frac{1}{16} n.$$

Proof. By Lemma 3.1.8, X is a line graph. Let $\widetilde{X}$ be the dual graph of X. By Theorem 8.2.7, $\widetilde{X}$ is a Moore graph or a generalized $2d$-gon of order $(1, s)$ for $s = k/2 > 2$.

If $\widetilde{X}$ is a Moore graph, then $\mu = 1$ implies that $\widetilde{X}$ is not complete, and $k > 4$ implies $\widetilde{X}$ is not a polygon. Thus $\widetilde{X}$ is a strongly regular graph in this case. Hence, Theorem 1.2.5 implies that $\text{motion}(\widetilde{X}) \geq n/8$ and the desired bound on the motion of X follows from Corollary 8.2.4.

Therefore, we may assume that $\widetilde{X}$ is a generalized $2d$-gon of order $(1, s)$ for $s > 2$. Then, by Fact 8.2.8, a halved graph Y of $\widetilde{X}$ is a generalized d-gon of order (s, s) (and d is even). Moreover, by Theorem 8.2.9, $d \leq 6$ and by Fact 8.2.10, $\xi(Y) \leq 2s$. Note that every pair of distinct vertices of Y has at most $q(Y) = s - 1$ common neighbors. Therefore, by Lemma 4.5.11,

$$\text{motion}(Y) \geq \frac{s(s+1) - 3s}{s(s+1)}|V(Y)| \geq \frac{s-2}{s+1}|V(Y)| \geq \frac{1}{4}|V(Y)|.$$

We note that $|V(\widetilde{X})| = 2|V(Y)|$ and $\text{motion}(\widetilde{X}) \geq \text{motion}(Y)$. Therefore, the statement of the proposition follows from Corollary 8.2.4. □

Remark 8.2.12. We note that one can show a linear lower bound (with a worse constant) on motion in this case without using the Feit-Higman classification theorem. Since a dual graph $\widetilde{X}$ is bipartite or of diameter 2, one can use Theorem 1.2.5 and the bounds on the motion of bipartite graphs, which we prove in Section 8.4.

8.3 Combining all pieces together: Proof of Babai's conjecture

Finally, we combine the above results to prove Babai's conjecture on motion of distance-regular graphs (Conjecture 8.0.1).

Theorem 8.3.1 (Confirming Conj. 8.0.1)**.** *For every $d \geq 3$ there exists $\gamma_d > 0$, such that*

for every primitive distance-regular graph X of diameter d on n vertices either

$$\text{motion}(X) \geq \gamma_d n,$$

or X is the Hamming graph $H(d,s)$ or the Johnson graph $J(s,d)$.

Proof. Recall, Theorem 8.1.7 implies that either $\text{motion}(X) \geq \gamma'_d n$ for some $\gamma'_d > 0$, or X is geometric with smallest eigenvalue $\geq -m_d$, for some $m_d \geq 3$.

Let $0 < \varepsilon' = \varepsilon(d)$ and $0 < \eta_d = \eta(d) < 1/2$ be the constants given by Theorem 7.3.8. Set

$$0 < \varepsilon = \frac{1}{2} \min\left(\frac{1}{6m_d^4 d}, \varepsilon', \frac{1}{200}\right).$$

Case A. X is not geometric or the smallest eigenvalue of X is less than $-m_d$.

Then, by Theorem 8.1.7, $\text{motion}(X) \geq \gamma'_d n$.

Case B. There exists $t \in [d]$ such that $c_{t+1} \geq \varepsilon k$ and $b_t \geq \varepsilon k$.

Then, by Proposition 8.1.6, $\text{motion}(X) \geq \varepsilon n/d$.

Case C. X is geometric with smallest eigenvalue at least $-m_d$ and there exists $t \in [d]$ such that $c_t < \varepsilon k$ and $b_t < \varepsilon k$.

By Theorem 7.3.8, the zero-weight spectral radius of X satisfies $\xi(X) \leq (1-\eta_d)k$.

Case C.1. $k < \max(29, 2m_d^3, 4m_d/\eta_d)$.

Then X has at most $N_d = \max(29, 2m_d^3, 4m_d/\eta_d)^d + 1$ vertices. Moreover, every non-trivial automorphism moves at least 2 points, so $\text{motion}(X) \geq \dfrac{2}{N_d} n$.

Case C.2. $k \geq \max(2m_d^3, 29)$ and $\mu \geq 2$.

Case C.2.i. $\theta_1 < (1-\varepsilon)b_1$.

Using Corollary 3.2.3 we obtain, $\lambda \geq \dfrac{k}{m_d} - 1 \geq m_d^2 \geq \mu$. By Lemma 2.4.4, we

have $2\lambda \leq \mu + k$, so $b_1 \geq (k - \mu - 2)/2 \geq k/4$. Thus, for $q(X) = \max(\lambda, \mu)$,

$$\xi(X) + q(X) \leq k - \varepsilon b_1 \leq \left(1 - \frac{\varepsilon}{4}\right) k.$$

Hence, by Lemma 4.5.11, $\text{motion}(X) \geq \frac{\varepsilon}{4} n$.

Case C.2.ii. $\theta_1 \geq (1 - \varepsilon) b_1$ and $\mu \geq 3$.

Since $\varepsilon < \frac{1}{200} < \varepsilon^*$, by Theorem 1.4.4 and Proposition 6.2.7, X is a Johnson graph.

Case C.2.iii. $\theta_1 \geq (1 - \varepsilon) b_1$ and $\mu = 2$.

By Corollary 6.3.7, X is a Hamming graph.

Case C.3. $\mu = 1$ and $k \geq \max(4m_d/\eta_d, m_d^2)$.

Case C.3.i. The smallest eigenvalue $-m$ of X satisfies $-m \leq -3$.

Then by Theorem 8.2.5, $\text{motion}(X) \geq \frac{\eta_d}{4} n$.

Case C.3.ii. The smallest eigenvalue $-m$ of X satisfies $-m > -3$.

Since, by Lemma 3.1.7, m is an integer, we get that $m \leq 2$. Hence, $m = 2$, and by Proposition 8.2.11 $\text{motion}(X) \geq n/16$.

Therefore, the statement of the theorem holds with $\gamma_d = \min\left(\frac{\eta_d}{4}, \frac{\varepsilon}{d}, \frac{2}{N_d}, \gamma'_d, \frac{1}{16}\right)$. □

8.4 Extending the motion result to imprimitive graphs

In this section we analyze the motion of imprimitive distance-regular graphs. Our main result is the following.

Theorem 8.4.1. *For every $d \geq 3$ there exists $\gamma_d > 0$, such that for every distance-regular graph X of diameter d on n vertices either*

$$\text{motion}(X) \geq \gamma_d n,$$

or X is a Johnson graph, or a Hamming graph, or a crown graph.

We start by establishing a version of the Spectral tool (see Lemma 4.5.11) in the bipartite case. Later we show that the motion of the antipodal graphs is controlled by the motion of their folded graphs. After that we prove motion lower bounds for bipartite graphs and for imprimitive graphs of diameter 3 and 4. A separate analysis for an imprimitive graph of diameter 3 and 4 is needed due to the fact that its folded or halved graph may be a complete graph, and in this case different arguments are required.

8.4.1 Spectral tool for bipartite graphs

To prove an analog of the Spectral tool (Lemma 4.5.11) for the case of bipartite graphs we need a version of the Expander Mixing Lemma for regular bipartite graphs.

Theorem 8.4.2 (Expander Mixing Lemma: bipartite version, [Haemers, 1995, Thm 5.1])**.** *Let X be a biregular bipartite graph with parts U and W of sizes n_U and n_W. Denote, by d_U and d_W the degrees of the vertices in parts U and W, respectively. Let λ_2 be the second largest eigenvalue of the adjacency matrix A of X. Then for every $S \subseteq U$, $T \subseteq W$*

$$\left(E(S,T)\frac{n_U}{|S|} - |T|d_W\right)\left(E(S,T)\frac{n_W}{|T|} - |S|d_U\right) \leq \lambda_2^2(n_U - |S|)(n_W - |T|),$$

which, using $d_U n_U = d_W n_W = E(U,W)$, implies

$$\left||E(S,T)| - \frac{d_W|S||T|}{n_U}\right| \leq |\lambda_2|\sqrt{|S||T|},$$

where $E(S,T)$ is the set of edges between S and T in X.

Next lemma is an analog of Lemma 4.5.11 for bipartite graphs.

Lemma 8.4.3. *Let X be a k-regular bipartite graph with parts U and W of size $n/2$ each. Let λ_2 be the second largest eigenvalue of A. Moreover, suppose that every pair of distinct*

vertices in X have at most q common neighbors. Then

$$\text{motion}(X) \geq \frac{k - |\lambda_2| - q}{2k} n.$$

Proof. Take any non-trivial automorphism σ of X. Consider $S_1 \subseteq U$ and $S_2 \subseteq W$, such that $S_1 \cup S_2 = \text{supp}(\sigma) = \{x \in X | x^\sigma \neq x\}$ be the support of σ. Without lost of generality, we may assume that $|S_1| \geq |S_2|$. Denote $S = S_1$ and let $T \subset W$ be a set which satisfies $S_2 \subseteq T$ and $|T| = |S|$. By the Expander Mixing Lemma we get

$$\frac{|E(S,T)|}{|S|} \leq |\lambda_2| + k\frac{2|S|}{n}.$$

Hence, there exists a vertex x in S which has at most $|\lambda_2| + k\dfrac{2|S|}{n}$ neighbors in T. Thus, x has at least $k - \left(|\lambda_2| + k\dfrac{2|S|}{n}\right)$ neighbors in $W \setminus T$, and they all are fixed by σ. Therefore, they all are common neighbors of x and $x^\sigma \neq x$. We get the inequality $q \geq k - \left(|\lambda_2| + k\dfrac{2|S|}{n}\right)$, which is equivalent to $\left(\dfrac{|\lambda_2| + q}{k}\right)\dfrac{n}{2} \geq \dfrac{n}{2} - |S|$. By the definition of S and T the number of fixed points of σ is at most

$$n - |S_1| - |S_2| \leq n - |S| \leq \left(\frac{1}{2} + \frac{|\lambda_2| + q}{2k}\right) n. \qquad \square$$

8.4.2 Reduction results

We show that the motion of an imprimitive distance-regular graph is controlled by the motion of its folded or halved graph.

Proposition 8.4.4. *Let X be an antipodal distance-regular graph of diameter $d \geq 3$ on n vertices and $\widetilde{X}$ be its folded graph on $\widetilde{n}$ vertices. Suppose* $\text{motion}(\widetilde{X}) \geq \alpha\widetilde{n}$. *Then* $\text{motion}(X) \geq \alpha n$.

Proof. Assume that X is an r-cover of $\widetilde{X}$ and let $\phi : X \to \widetilde{X}$ be a cover map. Let σ be an

automorphism of X. Note that by the definition of antipodal and folded graphs, vertices of $\widetilde{X}$ are maximal cliques (connected components) of X_d. Since σ is an automorphism of X, it preserves the relation of being at distance d, so σ respects preimages of ϕ. Hence, it induces an automorphism $\widetilde{\sigma}$ of $\widetilde{X}$ defined as $\widetilde{\sigma}(x) = \phi(\sigma(\phi^{-1}(x)))$.

If $\widetilde{\sigma}$ is non-identity, then by the assumptions of the lemma, the degree of $\widetilde{\sigma}$ is at least $\alpha\widetilde{n}$. Suppose that $x \in V(\widetilde{X})$ is not fixed by $\widetilde{\sigma}$, then $\phi^{-1}(x)$ is disjoint from $\sigma(\phi^{-1}(x))$. Thus, all vertices in $\phi^{-1}(x)$ are not fixed by σ. Therefore, the degree of σ is at least $r \cdot \alpha\widetilde{n} = \alpha n$.

Assume that $\widetilde{\sigma}$ is the identity map. Suppose that σ is a non-identity map. Let x be a vertex such that $\sigma(x) \neq x$ and let y be adjacent to x. Note that $\sigma(x)$ is at distance d from x as $\widetilde{\sigma}$ is the identity map. Thus $\sigma(y) \neq y$, as otherwise y is adjacent to $\sigma(x)$ and we get a contradiction with the assumption $d \geq 3$. Therefore, every vertex of X which is adjacent to a vertex not fixed by σ is itself not fixed by σ. Since X is connected, we get that the degree of σ is n in this case. □

Remark 8.4.5. If X is antipodal of diameter $d = 2$, then X is a complete multipartite graph and its folded graph is a complete graph. The motion of X is 2 in this case, so statement of the proposition above does not hold.

Definition 8.4.6. A pair of distinct vertices u and v in a graph X is called *twins* if the transposition (u, v) is an automorphism of X, i.e., $N(u) \cup \{u, v\} = N(v) \cup \{u, v\}$.

Lemma 8.4.7. *Let X be a distance-regular graph. Assume that X is not a complete graph or a complete multipartite graph. Then X has no twins.*

Proof. Assume that u and v are twins in X. Then $\mathrm{dist}(u, v) \leq 2$. If u and v are adjacent, then $\lambda = k - 1$, so X is a complete graph. If u and v are not adjacent, then $N(u) = N(v)$, so we obtain $\mu = k$. Thus, the diameter of X is 2 and every pair of distinct vertices at distance 2 are twins. This implies that X is complete multipartite. □

Proposition 8.4.8. *Let X be a bipartite distance-regular graph of diameter $d \geq 3$. Let X^+ and X^- be the halved graphs of X. Then*

$$\text{motion}(X) \geq \text{motion}(X^+) + \text{motion}(X^-).$$

Proof. Let σ be an automorphism of X. If $\sigma(X^+) = X^-$, then the degree of σ is $|V(X)| = n$. Otherwise, σ induces $\sigma^+ \in \text{Aut}(X^+)$ and $\sigma^- \in \text{Aut}(X^-)$.

Assume, that σ^+ is trivial, while σ^- is non-trivial. And let $\sigma(v) = u \neq v$, for $u, v \in X^-$. All neighbors of u and v are in X^+, and X^+ is fixed, so u and v are twins. We get a contradiction with Lemma 8.4.7.

Therefore, if σ is non-trivial, then both σ^+ and σ^- are non-trivial. Hence, the statement of the proposition follows. □

8.4.3 Bipartite graphs of diameter at least 4

Theorem 8.4.9. *Let X be a bipartite graph of diameter $d \geq 4$ on n vertices. If a halved graph of X is primitive, then*

$$\text{motion}(X) \geq \gamma'_d n, \quad \textit{where } \gamma'_d = (2d)^{-2d-5}.$$

Proof. Let $\iota(X) = \{b_0, b_1, \ldots, b_{d-1}; c_1, c_2, \ldots, c_d\}$ be the intersection array of X. Denote by Y a halved graph of X. By Proposition 2.5.3, the intersection array of Y is

$$\iota(Y) = \left\{ \frac{b_0 b_1}{\mu}, \frac{b_2 b_3}{\mu}, \ldots, \frac{b_{2t-2} b_{2t-1}}{\mu}; \frac{c_1 c_2}{\mu}, \frac{c_3 c_4}{\mu}, \ldots, \frac{c_{2t-1} c_{2t}}{\mu} \right\},$$

where $d \in \{2t, 2t+1\}$. Note that since X is bipartite, $b_i + c_i = k$, so in particular the degree of Y is equal to

$$\widehat{k} = \frac{b_0 b_1}{\mu} = \frac{k(k - c_1)}{\mu} = \frac{k(k-1)}{\mu}.$$

For convenience, define $c_i = k$ for every $i > d$. Take $1 \le j \le t$ such that $c_{2j-1} \le \varepsilon k$ and $c_{2j+1} \ge \varepsilon k$, where $\varepsilon = (2d)^{-d-2}$. Then for all $i \le j$

$$\frac{c_{2i-1}c_{2i}}{\mu} \le \frac{c_{2j-1}c_{2j}}{\mu} \le 2\varepsilon\widehat{k} \quad \text{and} \quad \frac{b_{2i-2}b_{2i-1}}{\mu} \ge \frac{b_{2j-1}^2}{\mu} = \frac{(k - c_{2t-1})^2}{\mu} \ge (1-\varepsilon)^2\widehat{k}. \tag{8.1}$$

Case 1. Assume that $j = 1$ and $b_{2j+1} \le \varepsilon k$. Then

$$\lambda(Y) = \widehat{k} - \frac{b_2 b_3}{\mu} - 1 \ge \widehat{k} - 2\varepsilon\widehat{k}.$$

By Lemma 8.1.1 we obtain that $\mu(Y) \ge \widehat{k} - 4\varepsilon\widehat{k} > \widehat{k}/2$. By Lemma 8.1.5, we get $|V(Y)| < t2^t k$. Therefore, as halved graph is not bipartite, by Proposition 8.1.4, we obtain

$$\text{motion}(Y) \ge \frac{1}{3t2^t}|V(Y)|.$$

Case 2. Assume that $j \ge 2$ and $b_{2j+1} \le \varepsilon k$. Then, for every $i \ge j+1$

$$\frac{c_{2i-1}c_{2i}}{\mu} \ge \frac{c_{2j+1}^2}{\mu} = \frac{(k - b_{2j+1})^2}{\mu} \ge (1-\varepsilon)^2\widehat{k}. \tag{8.2}$$

Hence, combining Eq. (8.1) and Eq. (8.2), by Lemma 7.1.2, the zero-weight spectral radius of Y satisfies

$$\xi(Y) \le \left(1 - (1-\varepsilon)^2 + 2(t+2)^2\varepsilon^{\frac{1}{t+1}}\right)\widehat{k}.$$

Since $j \ge 2$, using Eq. (8.1), we can estimate

$$\lambda(Y) \le \widehat{k} - \frac{b_2 b_3}{\mu} \le \widehat{k} - (1-\varepsilon)^2\widehat{k} \le 2\varepsilon\widehat{k} \quad \text{and} \quad \mu(Y) = \frac{c_3 c_4}{\mu} \le 2\varepsilon\widehat{k}.$$

Therefore, by Lemma 4.5.11 and the choice of ε,

$$\text{motion}(Y) \geq \left((1-\varepsilon)^2 - 2(t+2)^2\varepsilon^{\frac{1}{t+1}} - 2\varepsilon\right)|V(Y)| \geq \frac{1}{3}|V(Y)|.$$

Case 3. Assume that $b_{2j+1} > \varepsilon k$. Then

$$\frac{c_{2j+1}c_{2j+2}}{\mu} \geq \frac{c_{2j+1}^2}{\mu} \geq \varepsilon^2\widehat{k} \quad \text{and} \quad \frac{b_{2j}b_{2j+1}}{\mu} \geq \frac{b_{2j+1}^2}{\mu} \geq \varepsilon^2\widehat{k}.$$

Since Y is primitive, by Proposition 8.1.6,

$$\text{motion}(Y) \geq \frac{\varepsilon^2}{t}|V(Y)|.$$

Finally, since $|V(Y)| = n/2$, the inequality $\text{motion}(X) \geq \gamma n$ follows from Proposition 8.4.8 for $\gamma = \min\left(\frac{1}{3t2^t}, \frac{1}{3}, \frac{(2d)^{-2d-4}}{t}\right) \geq (2d)^{-2d-5}$. □

8.4.4 Bipartite antipodal graphs of diameter 4

Fact 8.4.10 (Brouwer et al. [1989], p. 425)**.** Let X be a bipartite antipodal distance-regular graph of diameter $d = 4$. Then there exist μ and m such that the number of vertices is $n = 2m^2\mu$, the degree is $k = m\mu$, and the intersection array is

$$\iota(X) = \{m\mu, m\mu - 1, (m-1)\mu, 1; 1, \mu, m\mu - 1, m\mu\}.$$

Moreover, the spectrum of X consists of k and $-k$ of multiplicity 1, $\sqrt{k}$ and $-\sqrt{k}$ of multiplicity $(m-1)k$, and 0 of multiplicity $(2k-2)$.

Proposition 8.4.11. *Let X be a bipartite antipodal distance-regular graph of diameter $d = 4$ on n vertices. Then*

$$\text{motion}(X) \geq 0.15n.$$

Proof. Consider a pair of distinct vertices u, v of X. If $\mathrm{dist}(u,v) > 2$, then they are distinguished by at least $D(u,v) \geq 2k$ vertices. Since X is bipartite, if $\mathrm{dist}(u,v) = 1$, then $D(u,v) \geq 2k$ as well. Clearly, for u, v at distance 2, we have $D(u,v) \geq 2(k-\mu)$. Thus $D_{\min}(X) \geq 2(k-\mu)$.

By Fact 8.4.10, $k = m\mu$ and $n = 2m^2\mu$ for some integer $m \geq 2$. Therefore, by Lemma 4.5.10,

$$\mathrm{motion}(X) \geq D_{\min}(X) \geq \frac{m-1}{m^2}n. \tag{8.3}$$

At the same time, by Fact 8.4.10, we know that the second largest eigenvalue of X equals $\sqrt{k}$. Then, by Lemma 8.4.3,

$$\mathrm{motion}(X) \geq \frac{k-\sqrt{k}-\mu}{2k}n = \frac{m\mu-\sqrt{m\mu}-\mu}{2m\mu}n \geq \frac{m-\sqrt{m}-1}{2m}n. \tag{8.4}$$

Using the bound given by Eq. (8.3) for $m \leq 4$, and the bound given by Eq. (8.4) for $m > 4$, we get the desired inequality. □

8.4.5 Bipartite graphs of diameter 3

Fact 8.4.12 (Brouwer et al. [1989], p. 432)**.** Let X be a bipartite distance-regular graph of diameter 3. Then the number of vertices of X is $n = 2 + 2k(k-1)/\mu$ and X has the intersection array

$$\iota(X) = \{k, k-1, k-\mu; 1, \mu, k\}.$$

The eigenvalues of X are k, $-k$ with multiplicity 1, and $\pm\sqrt{k-\mu}$ with multiplicity $\frac{n}{2}-1$.

Proposition 8.4.13. *Let X be a bipartite distance-regular graph of diameter* 3*. If X is not a crown graph, then*

$$\mathrm{motion}(X) \geq \frac{1}{6}n.$$

Proof. Denote the parts of the bipartite graph X by U and W. Let $Y = X_3$ be a distance-3

graph of X. We consider 2 cases.

Case 1. Suppose that Y is disconnected. Then there exists a pair of vertices u, v in one of the parts, so that u and v lie in different connected components of Y. Clearly, $\mathrm{dist}(u, v) = 2$, so $p^2_{3,3} = 0$. Hence, $k_3 = 1$, and the pairs of vertices at distance 3 form a perfect matching. Therefore, X is a regular complete bipartite graph with one perfect matching deleted.

Case 2. Y is connected and so is itself a distance-regular graph of diameter 3. Note, that $k + k_3 = n/2$, so if necessary, by passing to Y, we may assume that the degree of X satisfies $k \leq n/4$. Fact 8.4.12 implies $\mu \leq k/2$. The graph X is bipartite, so $\lambda = 0$ and by Lemma 4.5.10,

$$\mathrm{motion}(X) \geq D_{\min}(X) \geq 2(k - \mu) \geq k. \tag{8.5}$$

If $\mu \geq k/3$, then by Fact 8.4.12, $n \leq 6k$, so $\mathrm{motion}(X) \geq n/6$. If $k/4 \leq \mu < k/3$, then $n \leq 8k$, and Eq. (8.5) implies

$$\mathrm{motion}(X) \geq 2(k - \mu) \geq \frac{4}{3}k \geq \frac{n}{6}.$$

Finally, assume $\mu < k/4$. By Fact 8.4.12, the second largest eigenvalue is $\lambda_2 = \sqrt{k - \mu}$. Using that $k > 4\mu \geq 4$ and the function $x - \sqrt{x}$ is increasing for $x \geq 1$, by Lemma 8.4.3,

$$\mathrm{motion}(X) \geq \frac{k - \mu - \sqrt{k - \mu}}{2k} n \geq \frac{3k - 2\sqrt{3k}}{8k} n \geq \frac{15 - 2\sqrt{15}}{40} n \geq \frac{n}{6}. \qquad \square$$

8.4.6 Antipodal graphs of diameter 3

Fact 8.4.14 (see [Brouwer et al., 1989, p. 431])**.** Let X be an antipodal distance-regular graph of diameter $d = 3$ on n vertices. There exist integers $m \geq 2$, $r \geq 2$ and $t \geq 1$ such that the following holds.

- If $\lambda \neq \mu$, then $n = r(k + 1)$, $k = mt$, $\mu = (m - 1)(t + 1)/r$, $\lambda = \mu + t - m$. Moreover, the distinct eigenvalues of X are k, t, -1 and $-m$, with multiplicities 1,

$m(r-1)(k+1)/(m+t)$, k, $t(r-1)(k+1)/(m+t)$, respectively.

- If $\lambda = \mu$, then $n = r(k+1)$, $k = r\mu + 1$. The distinct eigenvalues of X are k, $\sqrt{k}$, -1 and $-\sqrt{k}$.

Proposition 8.4.15. *Let X be an antipodal distance-regular graph of diameter $d = 3$ on n vertices. If X is not a crown graph, then*

$$\text{motion}(X) \geq \frac{1}{13}n.$$

Proof. **Case 1.** Suppose that $\lambda \neq \mu$ and $t > m$. Then $\lambda > \mu$ and so

$$k - q(X) - \xi(X) = tm - \frac{(m-1)(t+1)}{r} - t + m - t \geq t\left(m - 2 - \frac{m-1}{r}\right).$$

If $m \geq 3$ and $r \geq 3$, then

$$k - q(X) - \xi(X) \geq t\left(m - 2 - \frac{m-1}{3}\right) = t\left(\frac{2}{3}m - \frac{5}{3}\right) \geq \frac{1}{9}tm = \frac{k}{9}.$$

If $r = 2$ and $m \geq 4$, then

$$k - q(X) - \xi(X) \geq t\left(m - 2 - \frac{m-1}{2}\right) = t\left(\frac{1}{2}m - \frac{3}{2}\right) \geq \frac{1}{8}tm = \frac{k}{8}.$$

Therefore, in both of these situations, by Lemma 4.5.11,

$$\text{motion}(X) \geq \frac{n}{9}.$$

If $r = 2$ and $m = 3$, then $n = 6t + 2$, $k = 3t$, $\mu = t + 1$ and $\lambda = 2t - 2$. Note that by Lemma 4.5.10, in this case

$$\text{motion}(X) \geq D_{\min}(X) \geq \min(2(k-\lambda), 2(k-\mu)) = 2(t+2) \geq \frac{n}{3}.$$

Finally, if $m = 2$, then $k = 2t$, $\mu = (t+1)/r$ is integer, and the multiplicity of t as an eigenvalue is an integer number equal to $2(r-1)(2t+1)/(t+2)$. Thus we may conclude that

$$(t+2) \mid 2(r-1)(2t+4-3) \quad \Rightarrow \quad (t+2) \mid 6(r-1) \quad \Rightarrow$$

$$\Rightarrow \quad (t+2) \mid 6\left(\frac{t+1}{\mu} - 1\right) \quad \Rightarrow \quad (t+2) \mid 6(t+1-\mu) \quad \Rightarrow \quad (t+2) \mid 6(\mu+1).$$

Hence, in particular,

$$(t+2) \leq 6\left(\frac{t+1}{r} + 1\right), \quad \text{so} \quad (t-4)r \leq 6t+6.$$

If $t \geq 10$, we get $r \leq 11$. If $t < 10$, then $r = (t+1)/\mu \leq t+1 < 11$. Therefore, by Lemma 4.5.10,

$$\text{motion}(X) \geq D_{\min}(X) \geq 2(k-\lambda) = 2\left(t+2-\frac{t+1}{r}\right) \geq \frac{r-1}{r^2}n \geq \frac{n}{13}.$$

Case 2. Suppose that $\lambda \neq \mu$ and $m > t$. Then $\lambda < \mu$ and so

$$k - q(X) - \xi(X) = tm - \frac{(m-1)(t+1)}{r} - m \geq m\left(t-1-\frac{t+1}{r}\right).$$

If $r \geq 4$ and $t \geq 2$, we get

$$k - q(X) - \xi(X) \geq m\left(t-1-\frac{t+1}{4}\right) = m\left(\frac{3}{4}t - \frac{5}{4}\right) \geq \frac{1}{8}mt = \frac{k}{8}.$$

Therefore, by Lemma 4.5.11,

$$\text{motion}(X) \geq \frac{n}{8}.$$

If $r \leq 4$ and $t \geq 2$, then $n \leq 4(k+1)$, and by Lemma 4.5.10,

$$\text{motion}(X) \geq D_{\min}(X) \geq \min(2(k-\lambda), 2(k-\mu)) \geq$$

$$\geq 2\left(mt - \frac{(m-1)(t+1)}{2}\right) \geq 2\left(mt - \frac{m(t+1)}{2}\right) \geq 2\left(mt - \frac{3mt}{4}\right) = \frac{k}{2} \geq \frac{n}{12}.$$

Finally, if $t = 1$, then $\lambda \geq 0$ implies $r = 2$. Hence, we obtain $n = 2(k+1)$, $\mu = k-1$ and $\lambda = 0$. It follows that X is a crown graph in this case.

Case 3. Assume $\lambda = \mu$. Then by Fact 8.4.14, $n = r(k+1)$, $k = r\mu + 1$ and $\xi(X) = \sqrt{k}$. By Lemma 4.5.11, for $r \geq 4$

$$\text{motion}(X) \geq \frac{k - \mu - \sqrt{k}}{k} n \geq \frac{(r-1)\mu + 1 - \mu\sqrt{r+1}}{r\mu + 1} n \geq \frac{r - \sqrt{r+1} - 1}{r} n \geq \frac{n}{6}.$$

At the same time, since $\lambda = \mu$, for $2 \leq r \leq 3$, by Lemma 4.5.10,

$$\text{motion}(X) \geq D_{\min}(X) \geq 2(k-\mu) \geq \frac{r-1}{r} k \geq \frac{2(r-1)}{3r^2} n \geq \frac{4}{27} n. \qquad \square$$

8.4.7 Collecting together the analysis for imprimitive graphs

In this section we collect analysis of the imprimitive case into Theorem 8.4.1. In the proof we use the following result about antipodal covers proved by van Bon and Brouwer [1988].

Theorem 8.4.16 (van Bon and Brouwer [1988])**.**

1. *The Hamming graph $H(d,s)$ has no distance-regular antipodal covers, except for $H(2,2)$, the quadrangle, which is covered by the octagon.*

2. *The Johnson graph $J(s,d)$ has no distance-regular antipodal covers for $d \geq 2$.*

3. *The complement $\overline{J(s,2)}$ has no distance-regular antipodal covers for $s \geq 8$.*

4. *The complement $\overline{H(2,s)}$ has no distance-regular antipodal covers for $s \geq 4$.*

Now we are ready to extend Theorem 8.3.1 to imprimitive distance-regular graphs.

Proof of Theorem 8.4.1. If X is primitive, by Theorem 8.3.1, there exists $\gamma_d > 0$ such that the claim holds. If X is bipartite and not antipodal of diameter $d \geq 4$, then by Theorem 8.4.9 $\text{motion}(X) \geq \gamma'_d n$. If X is bipartite (possibly antipodal) graph of diameter $d = 3$, then by Theorem 8.4.13, X is either a crown graph, or $\text{motion}(X) \geq n/6$.

If X is bipartite and antipodal of even diameter $d \geq 6$, then by Proposition 2.5.4, folded graph $\widetilde{X}$ is bipartite (and not antipodal) of diameter $d/2$. So

$$\text{motion}(\widetilde{X}) \geq \min\left(\gamma'_{d/2}, \frac{1}{6}\right) |V(\widetilde{X})|$$

(we use that crown graph is antipodal). Therefore, by Proposition 8.4.4,

$$\text{motion}(X) \geq \min\left(\gamma'_{d/2}, \frac{1}{6}\right) n.$$

In the case when X is bipartite and antipodal of diameter $d = 4$, by Proposition 8.4.11,

$$\text{motion}(X) \geq 0.15n.$$

We still need to analyze the cases when X is antipodal, but not bipartite, or when X is antipodal of odd diameter. By Proposition 2.5.4, in these cases, folded graph of X is primitive. If diameter of X is 3, then by Proposition 8.4.15, $\text{motion}(X) \geq n/13$. If diameter of X is $d \geq 4$, then by Proposition 2.5.4, folded graph $\widetilde{X}$ is primitive of diameter $\widetilde{d} = \lfloor d/2 \rfloor \geq 2$. Since $\widetilde{X}$ is primitive, $\widetilde{X}$ is not the complement to a disjoint union of cliques. If $\widetilde{X}$ has at least 29 vertices, and X is not a Johnson graph $J(s, d)$, the Hamming graph $H(d, s)$, or the complement of $J(s, 2)$ or $H(2, s)$, then by Theorem 1.2.5 and Theorem 8.3.1,

$$\text{motion}(\widetilde{X}) \geq \min\left(\gamma_{\widetilde{d}}, \frac{1}{8}\right) |V(\widetilde{X})|.$$

Therefore, in this case, by Theorem 8.4.4,

$$\text{motion}(X) \geq \min\left(\gamma_{\widetilde{d}}, \frac{1}{8}\right) n.$$

Finally we note, that by Theorem 8.4.16, if Y is the Johnson graph $J(s,d)$, the Hamming graph $H(d,s)$, or the complement of $J(s,2)$ or $H(2,s)$ and Y has at least 29 vertices, then Y has no antipodal covers. In the case when Y has at most 28 vertex, $\text{motion}(Y) \geq |Y|/14$. So by Proposition 8.4.4, if the graph X on n vertices is a distance-regular antipodal cover of such Y, then $\text{motion}(X) \geq n/14$.

Taking $\widetilde{\gamma} = \min\left(\gamma_d, \gamma'_d, \gamma'_{d/2}, \gamma_{\lfloor d/2 \rfloor}, \frac{1}{14}\right)$ we get the desired statement. $\square$

CHAPTER 9

MOTION OF RANK-4 PRIMITIVE COHERENT CONFIGURATIONS

In this chapter we prove Babai's conjecture on motion of primitive coherent configurations in the case when rank equals 4.

9.1 Outline of the proof of Theorem 1.2.8

Primitive coherent configurations of rank 4 naturally split into three classes: configurations induced by a primitive distance-regular graph of diameter 3, association schemes of diameter 2 (see Definition 2.3.18), and primitive coherent configurations with one undirected color and two oriented colors. The case of distance-regular graphs of diameter 3 follows from Theorem 8.3.1 we proved in Chapter 8.

So we need to deal with the other two classes. It is not hard to see that in the case of two oriented colors, the undirected constituent is strongly regular. Thus, by Babai's result on motion of strongly regular graphs (Theorem 1.2.5), if the number of vertices $n \geq 29$, the only possibility for $\mathfrak{X}$ to have motion less than $n/8$ is when the undirected constituent is the triagular graph $T(s)$, or the lattice graph $L_2(s)$, or their complement. In the latter case, we prove that the motion is linear in n using the generalization of an argument appearing in the proof of [Sun and Wilmes, 2015a, Lemma 3.5] (see Lemma 9.4.1 in this thesis).

Hence, we need to concentrate on the case of primitive association schemes of rank 4 with constituents of diameter 2. As the first step, we show that either we have a constituent with a $(1-\delta)$-dominant degree, or every pair of vertices can be distinguished by εn vertices (see Lemma 9.3.1). The latter directly implies that the motion is at least εn. On the other hand, the fact that one of the constituents, say X_3, has large degree implies that some intersection numbers are quite small (see Proposition 9.3.3). This allows us to approximate

the eigenvalues of the constituents X_1 and X_2, and so to approximate their zero-weight spectral radii with simple expressions involving the intersection numbers (see Lemma 9.2.3). We aim to apply Lemma 4.5.11 to the constituents X_1 and X_2. Considering cases how their degrees k_1 and k_2 can differ, we obtain that either the motion of $\mathfrak{X}$ is linear in n, or one of the graphs X_J is a line graph, where $J \in \{1, 2, \{1, 2\}\}$. By definition, $X_{1,2}$ is the complement of X_3. Since X_1 and X_2 are edge-regular, we use the classification of edge-regular and co-edge-regular graphs with smallest eigenvalue -2 (see Theorem 5.1.11). The classification tells us that either X_i is strongly regular with smallest eigenvalue -2, or it is the line graph of a triangle-free regular graph (see Theorem 9.3.10 for a more precise statement). If X_J is strongly regular with smallest eigenvalue -2, then X_J is a triangular graph $T(s)$, or a lattice graph $L_2(s)$, or has at most 28 vertices.

If one of the constituents is a line graph, this allows us to obtain more precise bounds on the intersection numbers. In particular, we approximate the zero-weight spectral radius of the graph $X_{1,2}$ with a relatively simple expression as well. At this point, our main goal becomes to get constraints on the intersection numbers, that will allow us to apply Lemma 4.5.11 effectively to one of the graphs X_1, X_2 or $X_{1,2}$. We consider four cases. Three of them are defined by which of the graphs X_1, X_2 or $X_{1,2}$ is strongly regular. In the fourth case, one of the constituents is the line graph of a triangle-free regular graph. For the ranges of the parameters when Lemma 4.5.11 cannot be used effectively we use Lemma 9.4.1. Roughly speaking, it says that if a triangular graph $T(s)$ is a union of several constituents of a coherent configuration $\mathfrak{X}$, then $\mathrm{Aut}(\mathfrak{X})$ is small if the following holds for every Delsarte clique: if we look on the configuration induced on the Delsarte clique, then each pair of vertices is distinguished by a constant fraction of the vertices of the clique. The hardest case in the analysis is the case when the constituent of the smallest degree, X_1, is strongly regular. This case is settled in Theorem 9.4.8 and requires preparatory work with several new ideas. In particular, we use an analog of the argument from the proof of Metsch's criteria to get a

constant upper bound on the fraction k_2/k_1 in a certain range of parameters.

9.2 Approximation of the eigenvalues of the constituents

First, we provide technical lemmas that allow us to approximate the zero-weight spectral radius of constituents X_1, X_2 and $X_{1,2}$ under quite modest assumptions.

Lemma 9.2.1. *Let $\mathfrak{X}$ be an association scheme of rank 4. Let η be a non-trivial eigenvalue of A_1. Then η satisfies cubic polynomial equation $\eta^3 + a_1\eta^2 + a_2\eta + a_3 = 0$, where*

$$a_1 = -(p^1_{1,1} + p^2_{1,2} - p^3_{1,1} - p^3_{1,2}) \qquad a_3 = \Big((p^2_{1,2} - p^3_{1,2})(k_1 - p^3_{1,1}) + (p^2_{1,1} - p^3_{1,1})p^3_{1,2}\Big)$$

$$a_2 = \Big((p^2_{1,2} - p^3_{1,2})(p^1_{1,1} - p^3_{1,1}) - (p^2_{1,1} - p^3_{1,1})(p^1_{1,2} - p^3_{1,2}) - (k_1 - p^3_{1,1})\Big)$$

Proof. By Eq. (2.2) for intersection numbers we have

$$A_1^2 = p^1_{1,1}A_1 + p^2_{1,1}A_2 + p^3_{1,1}A_3 + k_1 I.$$

We can eliminate A_3 using Eq. (2.1).

$$A_1^2 = (p^1_{1,1} - p^3_{1,1})A_1 + (p^2_{1,1} - p^3_{1,1})A_2 + (k_1 - p^3_{1,1})I + p^3_{1,1}J. \tag{9.1}$$

Let us multiply previous equation by A_1 and use Eq. (2.2).

$$\begin{aligned} A_1^3 = (p^1_{1,1} - p^3_{1,1})A_1^2 + (k_1 - p^3_{1,1})A_1 + p^3_{1,1}k_1 J + \\ +(p^2_{1,1} - p^3_{1,1})((p^1_{1,2} - p^3_{1,2})A_1 + (p^2_{1,2} - p^3_{1,2})A_2 + p^3_{1,2}J - p^3_{1,2}I). \end{aligned} \tag{9.2}$$

Combining Eq. (9.1) and (9.2) we eliminate A_2 as well.

$$A_1^3 - (p^2_{1,2} - p^3_{1,2})A_1^2 = (p^1_{1,1} - p^3_{1,1})A_1^2 + (k_1 - p^3_{1,1})A_1 + p^3_{1,1}k_1 J -$$

$$-(p^2_{1,2}-p^3_{1,2})(p^1_{1,1}-p^3_{1,1})A_1-(p^2_{1,2}-p^3_{1,2})((k_1-p^3_{1,1})I+p^3_{1,1}J)+$$

$$+(p^2_{1,1}-p^3_{1,1})(p^1_{1,2}-p^3_{1,2})A_1+(p^2_{1,1}-p^3_{1,1})(p^3_{1,2}J-p^3_{1,2}I).$$

Suppose that v is an eigenvector of A_1, which is different from the all-ones vector, and let η be the corresponding eigenvalue. Then $Jv=0$ and $A_1v=\eta v$, so the non-trivial eigenvalue η is a root of the polynomial $\eta^3+a_1\eta^2+a_2\eta+a_3$. □

We use the following result from the approximation theory which allows us to estimate the roots of a perturbed polynomial.

Theorem 9.2.2 ([Ostrowski, 1967, Appendix A])**.** *Let $n\geq 1$ be an integer. Consider a pair of polynomials of degree n*

$$f(x)=a_0x^n+...+a_{n-1}x+a_n,\quad g(x)=b_0x^n+...+b_{n-1}x+b_n,$$

where $a_0=b_0=1$. Denote $M=\max\{|a_i|^{1/i},|b_i|^{1/i}:0\leq i\leq n\}$ and

$$\varepsilon=2n\left(\sum_{i=1}^{n}|b_i-a_i|(2M)^{n-i}\right)^{1/n}.$$

Let $x_1,x_2,...,x_n$ denote the roots of f and $y_1,y_2,...,y_n$ denote the roots of g. Then, there exists a permutation $\sigma\in S_n$ such that for every $1\leq i\leq n$

$$|x_i-y_{\sigma(i)}|\leq\varepsilon.$$

Proposition 9.2.3. *Fix $\varepsilon>0$. Let $\mathfrak{X}$ be an association scheme of rank 4. Suppose that the parameters of $\mathfrak{X}$ satisfy $1/\varepsilon\leq k_1$ and $p^3_{1,i}\leq\varepsilon k_1$ for $i=1,2$. Then the zero-weight spectral radius $\xi(X_1)$ of X_1 satisfies*

$$\xi(X_1)\leq\frac{p^1_{1,1}+p^2_{1,2}+\sqrt{(p^1_{1,1}-p^2_{1,2})^2+4p^2_{1,1}p^1_{1,2}}}{2}+25\varepsilon^{1/3}k_1.$$

Proof. By Lemma 9.2.1, every non-trivial eigenvalue of X_1 is a root of the polynomial

$$\eta^3 + a_1\eta^2 + a_2\eta + a_3,$$

where a_1, a_2 and a_3 are as in Lemma 9.2.1. Observe, that for

$$b_1 = -(p^1_{1,1} + p^2_{1,2}), \quad b_2 = p^2_{1,2}p^1_{1,1} - p^2_{1,1}p^1_{1,2}, \quad b_3 = 0,$$

the following inequalities are true

$$|a_1 - b_1| \le 2\varepsilon k_1, \quad |a_2 - b_2| \le \left(4\varepsilon + 2\varepsilon^2 + \frac{1}{k_1}\right) k_1^2, \quad |a_3 - b_3| \le 2k_1^2 \le 2\varepsilon k_1^3.$$

Denote by ν_1, ν_2, ν_3 the non-trivial eigenvalues of A_1. By Theorem 9.2.2, we can arrange the roots x_1, x_2, x_3 of $x^3 + b_1x^2 + b_2x + b_3$ so that $|\nu_i - x_i| \le \delta$, where

$$\delta = 6\left(2\varepsilon k_1(4k_1)^2 + 6\varepsilon k_1^2(4k_1) + 2\varepsilon k_1^3\right)^{1/3} \le 25\varepsilon^{1/3}k_1. \qquad \square$$

Proposition 9.2.4. *Fix $\varepsilon > 0$. Let $\mathfrak{X}$ be an association scheme of rank 4. Suppose that the intersection numbers of $\mathfrak{X}$ satisfy $1/\varepsilon \le k_1$, $p^2_{1,1} \le \varepsilon k_1$ and $p^3_{i,j} \le \varepsilon \min(k_i, k_j)$ for $\{i, j\} = \{1, 2\}$. Then the zero-weight spectral radius of $X_{1,2}$ satisfies*

$$\xi(X_{1,2}) \le \frac{p^1_{1,1} + p^2_{1,2} + p^2_{2,2} + \sqrt{(p^2_{2,2} + p^2_{1,2} - p^1_{1,1})^2 + 4p^2_{1,2}p^1_{2,2}}}{2} + 25\varepsilon^{1/3}(k_1 + k_2).$$

Notation 9.2.5. We use the non-asymptotic notation $y = \square(x)$ to say that $|y| \le x$.

Proof. The proof is similar to the proofs of Proposition 9.2.3 and Lemma 9.2.1. Denote $k = k_1 + k_2$. By Eq. (2.2) we have

$$(A_1+A_2)^2 = (p^1_{1,1}+2p^1_{1,2}+p^1_{2,2})A_1+(p^2_{1,1}+2p^2_{1,2}+p^2_{2,2})A_2+(p^3_{1,1}+2p^3_{1,2}+p^3_{2,2})A_3+kI. \quad (9.3)$$

Note that, by assumptions of this proposition

$$0 \le p^3_{1,1} + 2p^3_{1,2} + p^3_{2,2} \le 2\varepsilon k \quad \text{and} \quad 0 \le p^1_{1,2} = \frac{k_2}{k_1} p^2_{1,1} \le \varepsilon k.$$

Using Eq. (2.1), we eliminate A_3.

$$\begin{aligned}(A_1 + A_2)^2 &= (p^1_{1,1} + p^1_{2,2} + 2\square(\varepsilon k))A_1 + (2p^2_{1,2} + p^2_{2,2} + 2\square(\varepsilon k))A_2 + \\ &\quad + (k + 2\square(\varepsilon k))I + 2\square(\varepsilon k)J = \\ &= (p^1_{1,1} + p^1_{2,2} + 2\square(\varepsilon k))(A_1 + A_2) + 2\square(\varepsilon k^2)I + 2\square(\varepsilon k)J + \\ &\quad + \left(2p^2_{1,2} + p^2_{2,2} - p^1_{1,1} - p^1_{2,2} + 4\square(\varepsilon k)\right) A_2.\end{aligned} \tag{9.4}$$

Denote by $R = 2p^2_{1,2} + p^2_{2,2} - p^1_{1,1} - p^1_{2,2} + 4\square(\varepsilon k)$ the last coefficient in Eq. (9.4). Multiplying Eq. (9.4) by $(A_1 + A_2)$ we get

$$\begin{aligned}(A_1 + A_2)^3 &= (p^1_{1,1} + p^1_{2,2} + 2\square(\varepsilon k))(A_1 + A_2)^2 + 2\square(\varepsilon k^2)(A_1 + A_2) + 2\square(\varepsilon k^2)J + \\ &\quad + R\left((p^1_{1,2} + p^1_{2,2} + \square(\varepsilon k))(A_1 + A_2) + (k_2 + \square(\varepsilon k))I + \square(\varepsilon k)J\right) + \\ &\quad + R(p^2_{1,2} + p^2_{2,2} - p^1_{1,2} + p^1_{2,2})A_2 = \\ &= (p^1_{1,1} + p^1_{2,2} + 2\square(\varepsilon k))(A_1 + A_2)^2 + 9\square(\varepsilon k^2)(A_1 + A_2) + 5\square(\varepsilon k^2)J + \\ &\quad + (2p^2_{1,2} + p^2_{2,2} - p^1_{1,1} - p^1_{2,2})p^1_{2,2}(A_1 + A_2) + 3\square(\varepsilon k^2)I \\ &\quad + R(p^2_{1,2} + p^2_{2,2} - p^1_{1,2} - p^1_{2,2})A_2.\end{aligned} \tag{9.5}$$

Let us multiply Eq. (9.4) by $p^2_{1,2} + p^2_{2,2} - p^1_{1,2} - p^1_{2,2} = p^2_{1,2} + p^2_{2,2} - p^1_{2,2} + \square(\varepsilon k)$ to eliminate A_2 from Eq. (9.5). Observe first, that

$$(2p^2_{1,2} + p^2_{2,2} - p^1_{1,1} - p^1_{2,2})p^1_{2,2} - (p^1_{1,1} + p^1_{2,2})(p^2_{1,2} + p^2_{2,2} - p^1_{2,2}) = p^2_{1,2}p^1_{2,2} - p^1_{1,1}p^2_{1,2} - p^1_{1,1}p^2_{2,2}.$$

Thus,

$$\begin{aligned}(A_1+A_2)^3-(p^1_{1,1}+p^2_{2,2}+p^2_{1,2})(A_1+A_2)^2-\\ -(p^2_{1,2}p^1_{2,2}-p^1_{1,1}p^2_{1,2}-p^1_{1,1}p^2_{2,2})(A_1+A_2)+\\ +3\square(\varepsilon k)(A_1+A_2)^2+13\square(\varepsilon k^2)(A_1+A_2)+5\square(\varepsilon k^3)I+8\square(\varepsilon k^2)J=0.\end{aligned} \tag{9.6}$$

Consider

$$b_1=-(p^1_{1,1}+p^2_{2,2}+p^2_{1,2}),\quad b_2=p^1_{1,1}(p^2_{1,2}+p^2_{2,2})-p^2_{1,2}p^1_{2,2},\quad b_3=0.$$

Then, Eq. (9.6) implies that every non-trivial eigenvalue η of $X_{1,2}$ satisfies the polynomial equation $\eta^3+a_1\eta^2+a_2\eta+a_3=0$, where

$$|a_1-b_1|\le 3\varepsilon k,\quad |a_2-b_2|\le 13\varepsilon k^2,\quad |a_3-b_3|\le 5\varepsilon k^3.$$

Denote by ν_1,ν_2,ν_3 the non-trivial eigenvalues of A_1+A_2. By Theorem 9.2.2, we can permute the roots x_1,x_2,x_3 of $x^3+b_1x^2+b_2x+b_3$ so that $|\nu_i-x_i|\le\delta$, where

$$\delta=6\left(3\varepsilon k(2k)^2+13\varepsilon k^2(2k)+5\varepsilon k^3\right)^{1/3}\le 25\varepsilon^{1/3}k.$$

Here we use that the inequalities $b_1=p^1_{1,1}+(p^2_{2,1}+p^2_{2,2})\le k$ and $b_2\le k^2$ hold, by Eq. (2.3). $\square$

9.3 Reduction to the case of a constituent with a clique geometry

In this section we show that the motion of a rank-4 association scheme of diameter 2 is linear in the number of vertices, unless one of its constituents, or its complement, has a clique geometry.

First, we show that one can assume that some intersection numbers are small.

Lemma 9.3.1. *Let $\mathfrak{X}$ be an association scheme of rank 4 and diameter 2 with the constituents ordered by degree. If $k_2 \geq \gamma k_3$, then every pair of distinct vertices is distinguished by at least $\gamma n/6$ vertices.*

Proof. Since $\mathfrak{X}$ has diameter 2 it is enough to show that some pair of vertices is distinguished by at least $\gamma n/3$ vertices, as then result follows by Lemma 4.5.2. Observe that vertices u, v, connected by an edge of color i, are distiguished by at least $|N_2(u) \triangle N_2(v)| = 2(k_2 - p^i_{2,2})$ vertices. At the same time, we have

$$k_2(k_2 - 1) = \sum_{i=1}^{3} k_i p^i_{2,2} \geq k_2 p^2_{2,2} + k_3 p^3_{2,2}.$$

Thus, $k_3 \geq k_2$ implies $k_2 - 1 \geq p^2_{2,2} + p^3_{2,2}$. So $\min(p^2_{2,2}, p^3_{2,2}) \leq (k_2 - 1)/2$. Hence, a pair of vertices connected by an edge of color i, which minimizes $p^i_{2,2}$, is distinguished by at least $k_2 + 1 \geq \gamma k_3 + 1 \geq \gamma n/3$ vertices. □

Remark 9.3.2. Note that the result of the lemma can also be derived directly from Proposition 6.3 proven by Babai [1981].

Lemma 9.3.3. *Let $\mathfrak{X}$ be an association scheme of rank 4 and diameter 2 with the largest degree equal k_3. Fix some $\varepsilon > 0$. Assume $\max(k_1, k_2) \leq \varepsilon k_3/2$. Then*

$$p^3_{1,2} \leq \varepsilon k_1, \quad p^3_{1,1} \leq \varepsilon k_1, \quad p^3_{2,2} \leq \varepsilon k_2, \quad \textit{and} \tag{9.7}$$

$$p^1_{3,3} \geq k_3(1 - \varepsilon), \quad p^2_{3,3} \geq k_3(1 - \varepsilon). \tag{9.8}$$

Proof. Note that for $i = 1, 2$,

$$k_i(k_i - p^i_{i,i} - 1) \geq k_3 p^3_{i,i}, \quad \text{so} \quad p^3_{i,i} \leq \varepsilon k_i/2.$$

Additionally, $p^1_{2,3} \leq k_2 \leq \varepsilon k_3/2$. Thus, by Eq. (2.3), $p^3_{1,2} = k_1 p^1_{2,3}/k_3 \leq \varepsilon k_1/2$.

Finally, by Eq. (2.3), for $j \in \{1, 2\}$, we have

$$p^i_{3,3} + p^i_{3,2} + p^i_{3,1} = k_3 \quad \text{and} \quad p^i_{3,j} \leq k_j \leq \varepsilon k_3/2.$$

Therefore, $k^i_{3,3} \geq (1 - \varepsilon)k_3$. □

Remark 9.3.4. Note that the inequalities (9.7) are still true if we replace ε by $\varepsilon/2$.

We need the following lemma, corollaries of which will be used several times.

Lemma 9.3.5. *Let $\mathfrak{X}$ be an association scheme. Suppose that there exists a triangle with sides of colors (s, r, t). Then*

$$p^s_{i,j} + p^r_{j,l} \leq k_j + p^t_{i,l}.$$

Proof. Apply the inclusion

$$N_i(u) \setminus N_l(w) \subseteq (N_i(u) \setminus N_j(v)) \cup (N_j(v) \setminus N_l(w)),$$

to vertices u, v, w, where $c(u, v) = s$, $c(v, w) = r$, and $c(u, w) = t$. □

Corollary 9.3.6. *Suppose that an association scheme $\mathfrak{X}$ of rank 4 satisfies $\max(k_1, k_2) \leq \varepsilon k_3/2$. Suppose also that there exists a triangle with sides $(s, t, 3)$. Then, $p^s_{i,j} \leq p^t_{i,3} + \varepsilon k_j$, where $i, j \in \{1, 2\}$.*

Proof. Take $r = l = 3$ in Lemma 9.3.5, then by Lemma 9.3.3,

$$p^3_{j,3} = k_j p^j_{3,3}/k_3 \geq (1 - \varepsilon)k_j.$$ □

Corollary 9.3.7. *Suppose that an association scheme $\mathfrak{X}$ has rank 4 and diameter 2. Moreover, assume $\max(k_1, k_2) \leq \varepsilon k_3/2$. Then, $p^s_{i,j} \leq p^s_{i,3} + \varepsilon k_j$, where $i, j, s \in \{1, 2\}$.*

Proof. Take $r = l = 3$ and $s = t$ in Lemma 9.3.5. Observe, that a triangle with sides of colors $(s, s, 3)$ exists, as diameter of $\mathfrak{X}$ is 2 and $s \neq 3$. □

Corollary 9.3.8. *Suppose an association scheme $\mathfrak{X}$ has rank 4 and diameter 2. Assume* $\max(k_1, k_2) \leq \varepsilon k_3/2$. *Then,* $2p^s_{i,j} \leq k_j + \varepsilon k_i$, *where* $i, j, s \in \{1, 2\}$. *Moreover, if* $k_1 \leq k_2$, *then* $2p^2_{1,2} \leq (1+\varepsilon)k_1$.

Proof. Take $t = 3$, $s = r$ and $l = i$ in Lemma 9.3.5 and we use Lemma 9.3.3. Take $s = 1$ and $i = j = 2$, then $2p^1_{2,2} \leq (1+\varepsilon)k_2$, so $2p^2_{1,2} \leq \dfrac{k_1}{k_2}(1+\varepsilon)k_2 = (1+\varepsilon)k_1$. □

We state the following simple corollary of Metsch's criteria (Theorem 3.1.3) and of the classification of graphs with the smallest eigenvalue ≥ -2 (Theorems 5.1.1 and 5.1.11), which will be used in the proof of Theorem 9.3.10.

Lemma 9.3.9. *Let $\mathfrak{X}$ be an association scheme of rank $r \geq 4$ and diameter 2 on n vertices.*

1. *Assume that for some i the constituent X_i satisfies the assumptions of Theorem 3.1.3 for $m = 2$. Then X_i is a strongly regular graph with smallest eigenvalue -2, or is the line graph of a regular triangle-free graph.*

   Furthermore, X_i satisfies the assumptions of Theorem 3.1.3 for $m = 2$, if we have one of the following

   (a) $\lambda(X_i) = p^i_{i,i} \geq \dfrac{2}{5}k_i$ *and* $\mu(X_i) = \max\{p^j_{i,i} : 0 < j \neq i\} \leq \dfrac{1}{30}k_i$, *or*

   (b) $\lambda(X_i) = p^i_{i,i} \geq \left(\dfrac{1}{2} - \dfrac{1}{20}\right)k_i$ *and* $\mu(X_i) = \max\{p^j_{i,i} : 0 < j \neq i\} \leq \dfrac{1}{11}\left(1 + \dfrac{1}{100}\right)k_i$.

2. *Assume for some i the complement $\overline{X_i}$ of X_i satisfies the assumptions of Theorem 3.1.3 for $m = 2$. If $n \geq 12$, then graph $\overline{X_i}$ is strongly regular with smallest eigenvalue -2.*

Proof. 1. First, we check that X_i satisfies the conditions of Theorem 3.1.3 for $m = 2$. It is sufficient to verify that $\lambda(X_i) > 3\mu(X_i)$ and $3\lambda(X_i) - 3\mu(X_i) > k_i$.

(a) We compute

$$\lambda(X_i) \geq \frac{2}{5}k_i > \frac{3}{30}k_i \geq 3\mu(X_i) \quad \text{and} \quad 3\lambda(X_i) - 3\mu(X_i) \geq \left(\frac{6}{5} - \frac{1}{10}\right)k_i > k_i.$$

(b) We compute

$$\lambda(X_i) - 3\mu(X_i) \geq \left(\frac{1}{2} - \frac{1}{20} - \frac{3}{11}\left(1 + \frac{1}{100}\right)\right) k_i = \frac{48}{275} k_i > 0,$$

$$3\lambda(X_i) - 3\mu(X_i) \geq k_i + \left(\frac{1}{2} - \frac{3}{20} - \frac{3}{11}\left(1 + \frac{1}{100}\right)\right) k_i = k_i + \frac{41}{550} k_i > k_i.$$

Now, if X_i satisfies the conditions of Theorem 3.1.3 for $m = 2$, by Lemma 3.1.8, it is a line graph, and, by Lemma 3.1.4, the smallest eigenvalue of X_i is at least -2. Moreover, recall that X_i is edge-regular. If the smallest eigenvalue is > -2, by Theorem 5.1.1, X_i is a complete graph or an odd polygon. This is impossible, since $\mathfrak{X}$ has diameter 2 and at least three non-empty constituents. If the smallest eigenvalue is -2, then by Theorem 5.1.11, we get that X_i is a strongly regular graph, or is the line graph of a regular triangle-free graph.

2. Since $\overline{X_i}$ satisfies the conditions of Theorem 3.1.3, by Lemmas 3.1.8 and 3.1.4, graph $\overline{X_i}$ is a line graph and its smallest eigenvalue is at least -2. Note also that $\overline{X_i}$ is co-edge-regular. If the smallest eigenvalue is > -2, by Theorem 5.1.1, $\overline{X_i}$ is complete graph or an odd polygon. This is impossible, since $\mathfrak{X}$ has diameter 2 and at least three non-empty constituents. If the smallest eigenvalue is -2, then by Theorem 5.1.11, we get that $\overline{X_i}$ is strongly regular, an $m_1 \times m_2$-grid or one of the two regular subgraphs of the Clebsh graph with 8 or 12 vertices.

   Assume $\overline{X_i}$ is a $m_1 \times m_2$-grid with $m_1 \neq m_2$ and $m_1, m_2 > 1$. That is, $\overline{X_i}$ is the line graph of K_{m_1,m_2}. Denote the parts of K_{m_1,m_2} by U_1 and U_2 with $|U_i| = m_i$. By symmetry, we can assume $m_1 < m_2$. We compute $n = m_1 m_2$, $k_1 + k_2 = m_1 + m_2 - 2$ and $\mu = 2$. Observe that, two edges of K_{m_1,m_2} that share a vertex in U_i have $m_i - 2$ common neighbors. Since $m_1 \neq m_2$, two pairs of edges in K_{m_1,m_2} that share a vertex in U_1 and U_2, respectively, cannot be colored in the same color in $\mathfrak{X}$. Thus, in particular,

$\mathfrak{X}$ is not primitive.

Therefore, $\overline{X_i}$ is strongly regular. $\square$

Theorem 9.3.10. *Let $\mathfrak{X}$ be an association scheme of rank 4 on n vertices with diameter 2 and with constituents ordered by degree. Recall that $q(X_i) = \max\{p^j_{i,i} : j \in [3]\}$ is the maximal number of common neighbours of two distinct vertices in X_i. Fix $\varepsilon = 10^{-16}$. Then one of the following is true.*

1. *Every pair of distinct vertices is distinguished by at least $\varepsilon n/12$ vertices.*
2. *The zero-weight spectral radius $\xi(X_i)$ of X_i satisfies $q(X_i)+\xi(X_i) \leq (1-\varepsilon)k_i$ for $i = 1$ or $i = 2$.*
3. *The graph X_1 is either strongly regular with smallest eigenvalue -2, or the line graph of a connected regular triangle-free graph.*
4. *The graph X_2 is either strongly regular with smallest eigenvalue -2, or the line graph of a connected regular triangle-free graph. Moreover, $k_2 \leq \frac{101}{100}k_1$.*
5. *If $n \geq 12$, then the graph $X_{1,2}$ is strongly regular with smallest eigenvalue -2 and $k_2 \leq \frac{101}{100}k_1$.*

Proof. We may assume that parameters of $\mathfrak{X}$ satisfy $\max(k_1, k_2) \leq \varepsilon k_3/2$, as otherwise, by Lemma 9.3.1 every pair of distinct vertices is distinguished by at least $\varepsilon n/12$ vertices. Therefore, all the inequalities provided by Lemma 9.3.3 hold.

Thus, by Proposition 9.2.3,

$$\xi(X_1) \leq \frac{p^1_{1,1} + p^2_{1,2} + \sqrt{(p^1_{1,1} - p^2_{1,2})^2 + 4p^2_{1,1}p^1_{1,2}}}{2} + \varepsilon_1 k_1, \text{ so}$$

$$\xi(X_1) \leq \max(p^1_{1,1}, p^2_{1,2}) + \sqrt{p^2_{1,1}p^1_{1,2}} + \varepsilon_1 k_1, \tag{9.9}$$

where $\varepsilon_1 = 25\varepsilon^{1/3}$. Similarly,

$$\xi(X_2) \leq \max(p^2_{2,2}, p^1_{2,1}) + \sqrt{p^1_{2,2}p^2_{1,2}} + \varepsilon_1 k_2. \tag{9.10}$$

We note that $k_1 \geq p^3_{1,1}/\varepsilon \geq 1/\varepsilon$, by Eq. (9.7).

Case 1. Assume $\gamma k_2 > k_1$, where $\gamma = \frac{1}{900}$. Then, using Lemma 9.3.3,

$$p^2_{1,1} = \frac{k_1}{k_2}p^1_{1,2} \leq \gamma p^1_{1,2}, \quad \text{so} \quad \mu(X_1) \leq \max(\gamma, \varepsilon)k_1 = \frac{1}{900}k_1.$$

Note that, by Corollary 9.3.8, inequality $\max(p^1_{1,1}, p^2_{1,2}) \leq \dfrac{1+\varepsilon}{2}k_1$ holds. Hence, by Eq. (9.9),

$$q(X_1) + \xi(X_1) \leq ((\varepsilon+\gamma)k_1 + p^1_{1,1}) + \left(\frac{1+\varepsilon}{2}k_1 + \sqrt{\gamma}k_1 + \varepsilon_1 k_1\right). \tag{9.11}$$

If $p^1_{1,1} < \dfrac{2}{5}k_1$, then Eq. (9.11) implies $q(X_1)+\xi(X_1) \leq (1-\varepsilon)k_1$. Otherwise, if $p^1_{1,1} \geq \dfrac{2}{5}k_1$, by Lemma 9.3.9, the graph X_1 satisfies the statement 3 of this proposition.

Case 2. Assume $k_1 = \gamma k_2$, where $(1+\varepsilon_3)^{-1} \geq \gamma \geq \frac{1}{900}$, for $\varepsilon_3 = \frac{1}{100}$.

We consider two subcases.

Case 2.1. Suppose $p^1_{2,3} = 0$.

So, by Corollary 9.3.7,

$$p^1_{2,2} \leq \varepsilon k_2 + p^1_{2,3} = \frac{\varepsilon}{\gamma}k_1, \quad p^2_{1,2} = \frac{k_1}{k_2}p^1_{2,2} \leq \varepsilon k_1, \tag{9.12}$$

$$p^1_{1,2} \leq \varepsilon k_1 + p^1_{2,3} = \varepsilon k_1, \quad p^2_{1,1} = \frac{k_1}{k_2}p^1_{1,2} \leq \varepsilon k_1. \tag{9.13}$$

Then,

$$\max_{i\in[3]}(p^i_{1,1}) + \max(p^1_{1,1}, p^2_{1,2}) + \sqrt{p^2_{1,1}p^1_{1,2}} \leq 3\varepsilon k_1 + 2p^1_{1,1}. \tag{9.14}$$

Thus, by Eq. (9.14), $q(X_1) + \xi(X_1) < (1-\varepsilon)k_1$ if $p^1_{1,1} < \dfrac{2}{5}k_1$. Alternatively, if $p^1_{1,1} \geq \dfrac{2}{5}k_1$,

by Lemma 9.3.9, graph X_1 satisfies the statement 3 of this proposition, as $\mu(X_1) \leq \varepsilon k_1$ by Eq. (9.13).

Case 2.2. Suppose $p^1_{2,3} \neq 0$.

Case 2.2.1. Assume that $p^1_{1,1} \geq p^2_{1,1}$.

By Corollary 9.3.6,

$$p^2_{1,2} \leq p^1_{1,3} + \varepsilon k_2 = p^1_{1,3} + \frac{\varepsilon}{\gamma} k_1 \quad \text{and} \quad p^1_{1,1} \leq p^1_{1,3} + \varepsilon k_1.$$

Then

$$\max_{i\in[3]}(p^i_{1,1}) + p^2_{1,2} + \sqrt{\gamma} p^1_{1,2} \leq p^1_{1,1} + p^1_{1,3} + p^1_{1,2} - (1-\sqrt{\gamma}) p^1_{1,2} + \left(\varepsilon + \frac{\varepsilon}{\gamma}\right) k_1, \quad (9.15)$$

$$\max_{i\in[3]}(p^i_{1,1}) + p^1_{1,1} + \sqrt{\gamma} p^1_{1,2} \leq p^1_{1,3} + p^1_{1,1} + p^1_{1,2} - (1-\sqrt{\gamma}) p^1_{1,2} + \left(\varepsilon + \frac{\varepsilon}{\gamma}\right) k_1. \quad (9.16)$$

Case 2.2.2. Assume that $p^2_{1,1} \geq p^1_{1,1}$.

By Corollary 9.3.7,

$$p^2_{1,1} = \gamma p^1_{1,2} \leq \gamma(p^1_{1,3} + \varepsilon k_2) \leq \gamma p^1_{1,3} + \varepsilon k_1.$$

This implies

$$\max_{i\in[3]}(p^i_{1,1}) + p^1_{1,1} + \sqrt{\gamma} p^1_{1,2} \leq \gamma p^1_{1,3} + p^1_{1,1} + \sqrt{\gamma} p^1_{1,2} + \varepsilon k_1, \quad (9.17)$$

$$\max_{i\in[3]}(p^i_{1,1}) + p^2_{1,2} + \sqrt{\gamma} p^1_{1,2} \leq p^2_{1,1} + p^2_{1,2} + p^2_{1,3} - (1-\sqrt{\gamma}) p^1_{1,2} + \left(\varepsilon + \frac{\varepsilon}{\gamma}\right) k_1, \quad (9.18)$$

where in Eq. (9.18) we use the inequality $p^1_{1,2} \leq p^2_{1,3} + \varepsilon k_2$ given by Corollary 9.3.6.

Therefore, using Eq. (2.3), in both subcases by Eq. (9.15) - (9.18) we get

$$\max_{i\in[3]}(p^i_{1,1}) + \left(\max(p^1_{1,1}, p^2_{1,2}) + \sqrt{p^2_{1,1}p^1_{1,2}}\right) \leq k_1 - (1-\sqrt{\gamma})p^1_{1,2} + \left(\varepsilon + \frac{\varepsilon}{\gamma}\right)k_1, \text{ so}$$

$$q(X_1) + \xi(X_1) \leq k_1 - (1-\sqrt{\gamma})p^1_{1,2} + \left(\varepsilon + \frac{\varepsilon}{\gamma}\right)k_1 + \varepsilon_1 k_1. \tag{9.19}$$

We again consider two subcases.

Case 2.2.a. Suppose $p^1_{1,2} > \varepsilon_2 k_1$ for $\varepsilon_2 = \frac{1}{30}$.

Observe that

$$\varepsilon_2(1-\sqrt{\gamma}) - \varepsilon\left(2 + \frac{1}{\gamma}\right) - \varepsilon_1 \geq 10^{-4} - 902\varepsilon - 25\varepsilon^{1/3} > 0,$$

so by Eq. (9.19),

$$q(X_1) + \xi(X_1) \leq (1-\varepsilon)k_1. \tag{9.20}$$

Case 2.2.b. Suppose $p^1_{1,2} \leq \varepsilon_2 k_1$.

This implies $p^2_{1,1} \leq \varepsilon_2 k_1$, so $\mu(X_1) \leq \frac{1}{30}k_1$. Recall, that by Corollary 9.3.8, the inequality $\max(p^2_{1,2}, p^1_{1,1}) \leq \dfrac{1+\varepsilon}{2}k_1$ holds. Then, by Eq (9.9), we have

$$q(X_1) + \xi(X_1) \leq (\varepsilon + \varepsilon_2)k_1 + p^1_{1,1} + \frac{1+\varepsilon}{2}k_1 + \varepsilon_2 k_1 + \varepsilon_1 k_1. \tag{9.21}$$

Thus, either Eq. (9.21) implies the inequality $q(X_1)+\xi(X_1) < (1-\varepsilon)k_1$, or $p^1_{1,1} \geq \dfrac{2}{5}k_1$. In the latter case, by Lemma 9.3.9, the statement 3 of this proposition holds for X_1.

Case 3. Suppose that $k_2 \leq (1+\varepsilon_3)k_1$, where $\varepsilon_3 = \frac{1}{100}$.

In this case, we work with both X_1 and X_2 in the same way. Additionally, we need to consider the graph $X_{1,2}$ with the set of vertices $V(X_{1,2}) = V(X_1) = V(X_2)$ and set of edges $E(X_{1,2}) = E(X_1) \cup E(X_2)$. The graph $X_{1,2}$ is regular of degree $k_1 + k_2$, and every pair of

non-adjacent vertices has

$$\mu(X_{1,2}) = p^3_{1,1} + 2p^3_{1,2} + p^3_{2,2} \leq 2\varepsilon(k_1 + k_2) \leq 4\varepsilon(1 + \varepsilon_3)k_1 \tag{9.22}$$

common neighbors. Every pair of vertices connected by an edge of color i has

$$\lambda_i = p^i_{1,1} + 2p^i_{1,2} + p^i_{2,2} \tag{9.23}$$

common neighbors, for $i = 1, 2$. We apply the inequality

$$|N(u) \cap N(v)| + |N(v) \cap N(w)| \leq |N(v)| + |N(u) \cap N(w)| \tag{9.24}$$

to the graph $X_{1,2}$ and vertices u, v, w with $c(u, v) = c(v, w) = i$ and $c(u, w) = 3$, where $i \in \{1, 2\}$. We get $2\lambda_i \leq k_1 + k_2 + \mu(X_{1,2})$, so by Eq. (9.22),

$$\lambda_i = p^i_{1,1} + 2p^i_{1,2} + p^i_{2,2} \leq \frac{1 + 2\varepsilon}{2}(k_1 + k_2) \leq k_1(1 + \varepsilon_3 + 2\varepsilon). \tag{9.25}$$

Let $\{i, j\} = \{1, 2\}$, then by Eq. (9.9)-(9.10),

$$\begin{aligned} q(X_i) + \xi(X_i) &\leq q(X_i) + \max(p^i_{i,i}, p^j_{i,j}) + \sqrt{p^i_{i,j} p^j_{i,i}} + \varepsilon_1 k_i \leq \\ &\leq \Big(\max(p^i_{i,i}, p^j_{i,i}) + \varepsilon k_i\Big) + \max(p^i_{i,i}, p^j_{i,j}) + p^i_{i,j}(1 + \varepsilon_3) + \varepsilon_1 k_i. \end{aligned} \tag{9.26}$$

Consider all possible ways of opening the maximums in Eq. (9.26) (we only write terms without epsilons).

1. $2p^i_{i,i} + p^i_{i,j}$,
2. $p^i_{i,i} + p^j_{i,i} + p^i_{i,j} \leq (1 + \varepsilon_3)(p^i_{i,i} + 2p^i_{i,j}) = (1 + \varepsilon_3)(\lambda_i - p^i_{j,j}) \leq (1 + \varepsilon_3)\lambda_i$,
3. $p^j_{i,j} + p^i_{i,i} + p^i_{i,j} \leq (1 + \varepsilon_3)(p^i_{j,j} + p^i_{i,i} + p^i_{i,j}) = (1 + \varepsilon_3)(\lambda_i - p^i_{i,j}) \leq (1 + \varepsilon_3)\lambda_i$,

4. $p^j_{i,j} + p^j_{i,i} + p^i_{i,j} \le (1+\varepsilon_3)(p^i_{j,j} + 2p^i_{i,j}) = (1+\varepsilon_3)(\lambda_i - p^i_{i,i}) \le (1+\varepsilon_3)\lambda_i$.

Hence, by Corollary 9.3.8 applied to $p^i_{i,j}$, Eq. (9.26) implies

$$q(X_i) + \xi(X_i) \le \max(2p^i_{i,i} + p^i_{i,j}, (1+\varepsilon_3)\lambda_i) + k_i\left(\varepsilon_1 + \frac{2}{3}\varepsilon_3 + \varepsilon\right). \tag{9.27}$$

Case 3.1 Suppose $\lambda_t \ge (2/3 + \frac{1}{300})k_t$ for both $t = 1, 2$.

Then in notation of Theorem 3.1.3

$$\lambda^{(1)} \ge \left(\frac{2}{3} + \frac{1}{300}\right)k_1, \quad \text{and by Eq. (9.25)}, \quad \lambda^{(2)} \le \frac{11}{10}k_1.$$

We check that

$$2\lambda^{(1)} - \lambda^{(2)} \ge 20\varepsilon(1+\varepsilon_3)k_1 \ge 5\mu, \quad \text{and} \quad 3\lambda^{(1)} - 3\mu \ge 2k_1 + \frac{1}{100}k_1 \ge k_1 + k_2.$$

Thus, $X_{1,2}$ satisfies conditions of Theorem 3.1.3 for $m = 2$, so the statement 5 of this proposition holds by Lemma 9.3.9.

Case 3.1 Suppose that $\lambda_i \le (2/3 + \frac{1}{300})k_i$ for some $i \in \{1, 2\}$.

If $2p^i_{i,i} + p^i_{i,j} \le k_i - k_i(\varepsilon_3 + 2\varepsilon + \varepsilon_1)$, then Eq. (9.27) implies

$$q(X_i) + \xi(X_i) \le (1-\varepsilon)k_i.$$

Hence, we can assume $2p^i_{i,i} + p^i_{i,j} \ge k_i - k_i(\varepsilon_3 + 2\varepsilon + \varepsilon_1)$. Recall, that

$$2p^i_{i,i} + p^i_{i,j} = \lambda_i + (p^i_{i,i} - p^i_{i,j} - p^i_{j,j}) \le \lambda_i + \frac{1+\varepsilon}{2}k_i - (p^i_{i,j} + p^i_{j,j}).$$

Thus,

$$p^i_{i,j} + p^i_{j,j} \le \lambda_i + \frac{1+\varepsilon}{2}k_i - k_i(1 - \varepsilon_3 - 2\varepsilon - \varepsilon_1) \le k_i\frac{51}{300} + k_i(\varepsilon_3 + 3\varepsilon + \varepsilon_1) \le \frac{2}{11}k_i. \tag{9.28}$$

This implies,

$$\min(p^i_{j,j}, p^j_{i,i}) \leq (1+\varepsilon_3)\min(p^i_{i,j}, p^i_{j,j}) \leq \frac{1+\varepsilon_3}{11}k_1. \tag{9.29}$$

Take $\{s,t\} = \{1,2\}$, so that $p^t_{s,s} \leq \frac{1+\varepsilon_3}{11}k_1$. Then

$$\mu(X_s) \leq \max\left(\varepsilon k_s, \frac{1+\varepsilon_3}{11}k_1\right) = \frac{1+\varepsilon_3}{11}k_1.$$

We consider two possibilities. First, assume that $p^s_{s,s} \geq \left(\frac{1}{2} - \frac{1}{20}\right)k_s$. Then, by Lemma 9.3.9 graph X_s satisfies the statement 3 or 4 of this proposition.

Assume now that $p^s_{s,s} \leq \left(\frac{1}{2} - \frac{1}{20}\right)k_s$, then

$$2p^s_{s,s} + p^s_{s,t} \leq k_s - \frac{1}{10}k_s + \frac{(1+\varepsilon_3)^2}{11}k_s \leq \left(1 - 2\varepsilon - \frac{2}{3}\varepsilon_3 - \varepsilon_1\right)k_s, \tag{9.30}$$

and by Eq. (9.28),

$$\begin{aligned}(1+\varepsilon_3)\lambda_s &\leq p^s_{s,s} + 2p^s_{s,t} + 2p^s_{t,t} + 2\varepsilon_3 k_s \leq \\ &\leq \left(\frac{1}{2} - \frac{1}{20} + \frac{4+4\varepsilon_3}{11} + 2\varepsilon_3\right)k_s < \left(1 - 2\varepsilon - \frac{2}{3}\varepsilon_3 - \varepsilon_1\right)k_s.\end{aligned} \tag{9.31}$$

Thus, by Eq. (9.27), equations (9.30) and (9.31) imply

$$q(X_s) + \xi(X_s) \leq \max(2p^s_{s,s} + p^s_{s,t}, (1+\varepsilon_3)\lambda_s) + \left(\varepsilon + \frac{2}{3}\varepsilon_3 + \varepsilon_1\right)k_s \leq (1-\varepsilon)k_s. \qquad \square$$

9.4 Case of a constituent with a clique geometry

In the previous subsection, Theorem 9.3.10 reduces the diameter 2 case of Theorem 1.2.8 to the case when one of the constituents is a strongly regular graph with smallest eigenvalue -2, or is the line graph of a triangle-free regular graph. In this subsection we resolve the

remaining cases.

In the case when the dominant constituent X_3 is strongly regular we introduce an additional tool (Lemma 9.4.1), which allows us to bound the order of the group and its minimal degree, when vertices inside a clique are well-distinguished.

In the cases when the constituent X_1 or X_2 is strongly regular, we prove upper bounds on the quantity $q(X_J) + \xi(X_J)$ for $J \in \{1, 2, \{1, 2\}\}$ with the consequence that the spectral tool (Lemma 4.5.11) can be applied effectively.

The hardest case in our analysis is the case when the constituent with the smallest degree, X_1 , is strongly regular. This case is settled in Theorem 9.4.8 (Sec. 9.4.4) and it requires considerable preparatory work to establish a constant upper bound on the quotient k_2/k_1 in certain range of parameters.

9.4.1 Triangular graph with well-distinguished cliques

In the case when the union of some constituents of a homogeneous coherent configuration is a triangular graph we prove the following statement inspired by Lemma 3.5 in Sun and Wilmes [2015a].

Lemma 9.4.1. *Let $\mathfrak{X}$ be a homogeneous coherent configuration on n vertices. Let I be a set of colors, such that if $i \in I$, then $i^* \in I$. Suppose that graph X_I is the triangular graph $T(s)$ for some s. Let $\mathcal{C}$ be a Delsarte clique geometry in X_I. Assume there exists a constant $0 < \alpha < \frac{1}{2}$, such that for every clique $C \in \mathcal{C}$ and every pair of distinct vertices $x, y \in C$ there exist at least $\alpha|C|$ elements $z \in C$ which distinguish x and y, i.e., $c(z, x) \neq c(z, y)$. Then*

1. *There exists a set of vertices of size $O(\log(n))$ that completely splits $\mathfrak{X}$. Hence, $|\operatorname{Aut}(\mathfrak{X})| = n^{O(\log(n))}$,*

2. $\operatorname{motion}(\mathfrak{X}) \geq \alpha n/2$.

Proof. Consider a clique $C \in \mathcal{C}$. Since every pair of distinct vertices $x, y \in C$ is distinguished by at least $\alpha|C|$ vertices of C, by Lemma 4.5.5, there is a set of size at most $\frac{2}{\alpha}\log(|C|) + 1$ that splits C completely.

Take any vertex $x \in \mathfrak{X}$. By the assumptions of the lemma, $\{x\} \cup N_I(x) = C_1 \cup C_2$ for some $C_1, C_2 \in \mathcal{C}$. Then there exists a set S of size $\frac{4}{\alpha}\log(|C|) + 2 \leq \frac{4}{\alpha}\log(n) + 2$, that splits both C_1 and C_2 completely. Note that every clique $C \in \mathcal{C}$, distinct from C_1 and C_2, intersects each of them in exactly one vertex, and is uniquely determined by $C \cap C_1$ and $C \cap C_2$. Therefore, the pointwise stabilizer $\mathrm{Aut}(\mathfrak{X})_{(S)}$ fixes every clique $C \in \mathcal{C}$ as a set. At the same time, every vertex v is uniquely defined by the collection of cliques in $\mathcal{C}$ that contain v. Therefore, S splits X completely.

Suppose $\sigma \in \mathrm{Aut}(\mathfrak{X})$ and $|\operatorname{supp}(\sigma)| < \frac{\alpha}{2}n$. Then, by the pigeonhole principle, there exists a vertex x, such that σ fixes at least $\left\lceil \left(1 - \frac{\alpha}{2}\right)(|N_I(x)| + 1)\right\rceil$ vertices in $N_I(x) \cup \{x\}$. Since $X_I = T(s)$, we have $\{x\} \cup N_I(x) = C_1 \cup C_2$ for some $C_1, C_2 \in \mathcal{C}$ and $|C_1| = |C_2| = 1 + \frac{|N_I|}{2}$. Thus σ fixes more than $(1-\alpha)|C_i|$ vertices in C_i for every $i \in \{1, 2\}$. This means that every pair of vertices $x, y \in C_i$ is distinguished by at least one vertex fixed by σ. Hence, $\sigma(x) \neq y$. At the same time, since $(1-\alpha)|C_i| > 1$, $\sigma(x) \in C_i$ for every $x \in C_i$. Therefore, σ fixes poinwise both C_1 and C_2. Finally, by the argument in the previous paragraph, we get that σ fixes every vertex, so σ is the identity. □

9.4.2 Constituent X_3 is strongly regular

In the following theorem we consider the case when the constituent X_3 is a strongly regular graph (case of the statement 5 in Theorem 9.3.10).

Theorem 9.4.2. *Let $\mathfrak{X}$ be an association scheme of rank 4 and diameter 2 on $n \geq 29$ vertices. Assume that the constituents of $\mathfrak{X}$ are ordered by degree and $k_2 \leq \varepsilon k_3/2$ holds for $\varepsilon < \frac{1}{100}$. Suppose that $k_2 \leq \frac{11}{10}k_1$ and $X_{1,2}$ is strongly regular with smallest eigenvalue -2. Then neither X_1, nor X_2 is strongly regular with smallest eigenvalue -2, and one of the*

following is true.

1. *The association scheme $\mathfrak{X}$ satisfies the assumptions of Lemma 9.4.1 for $I = \{1, 2\}$ and $\alpha = 1/16$.*

2. *X_1 or X_2 is the line graph of a regular triangle-free graph.*

Proof. By the assumptions of the theorem, all inequalities from Lemma 9.3.3 hold.

Since $X_{1,2}$ is strongly regular with smallest eigenvalue -2, by Seidel's classification (see Theorem 5.1.11), $X_{1,2} = T(s)$ or $X_{1,2} = L_2(s)$ for some s. Suppose that $X_{1,2}$ is $L_2(s)$, then $n = s^2$, $k_1 + k_2 = 2(s-1)$, so $k_1 \leq (s-1)$. At the same time, since X_1 has diameter 2, degree k_1 should satisfy $k_1^2 \geq n - 1$, which gives us a contradiction. Therefore, $X_{1,2} = T(s)$. Consider 2 cases.

Case 1. Assume $p^2_{1,1} \geq k_1/30$ and $p^1_{2,2} \geq k_2/30$.

We can rewrite the assumptions of this case in the form $p^1_{1,2} = p^2_{1,1}\dfrac{k_2}{k_1} \geq \dfrac{k_2}{30}$ and $p^2_{1,2} \geq \dfrac{k_1}{30}$. We know that $X_{1,2} = T(s)$ for some s. Let $\mathcal{C}$ be a Delsarte clique geometry of $X_{1,2}$. Then every clique $C \in \mathcal{C}$ has size

$$1 + p^i_{1,1} + p^i_{2,2} + 2p^i_{1,2} = 1 + \lambda_i(X_{1,2}) = \frac{k_1 + k_2}{2} + 1 \leq \frac{21}{20}k_1 + 1$$

for $i \in \{1, 2\}$. Every pair of distinct vertices $x, y \in C$ with $c(x, y) = i \in \{1, 2\}$ is distinguished by at least $|C| - p^i_{1,1} - p^i_{2,2} = 2p^i_{1,2} + 1 \geq k_1/15 + 1 \geq |C|/16$ vertices in C.

Case 2. Assume $p^j_{i,i} < k_i/30$ for $\{i, j\} = \{1, 2\}$.

Using Corollary 9.3.8 and the inequality $k_1 \leq k_2 \leq \frac{11}{10}k_1$, we get

$$\frac{k_i + k_j}{2} = \lambda_i(X_{1,2}) = p^i_{i,i} + 2p^i_{i,j} + p^i_{j,j} \leq p^i_{i,i} + 2\frac{11}{10}\frac{k_i}{30} + \frac{(1+\varepsilon)}{2}k_j.$$

Thus,

$$\frac{2}{5}k_i \leq \left(\frac{1}{2} - \frac{11}{150} - \frac{11\varepsilon}{20}\right)k_i \leq p^i_{i,i}.$$

Therefore, by Lemma 9.3.9, the graph X_i is strongly regular with smallest eigenvalue -2, or X_i is the line graph of a regular triangle-free graph.

Assume that for some $i \in \{1, 2\}$ the graph X_i is strongly regular with smallest eigenvalue -2, then X_i, as well as $X_{1,2}$, is either $T(s)$ or $L_2(s)$. Since X_i and $X_{1,2}$ have the same number of vertices, the only possibility is $X_{1,2} = T(s_1)$ and $X_i = L_2(s_2)$. Then $s_1(s_1 - 1)/2 = s_2^2$, so $\sqrt{2}s_2 > (s_1 - 1)$. This implies

$$k_i + k_j = 2(s_1 - 2) \leq 2\sqrt{2}(s_2 - 1) + 1 = \sqrt{2}k_i + 1,$$

so $k_j \leq (\sqrt{2} - 1)k_i + 1$ and we get a contradiction with $k_i \leq \frac{11}{10}k_j$ and $k_i \geq \sqrt{n-1} > 3$. $\square$

Remark 9.4.3. Observe that the argument in the last paragraph of the proof shows that X_1 and $X_{1,2}$ cannot be simultaneously strongly regular with smallest eigenvalue -2 even if the assumption that $k_2 \leq \frac{11}{10}k_1$ does not hold (we assume that all other assumptions of the theorem are satisfied).

9.4.3 Constituent X_2 is strongly regular

Next, we consider the case when X_2 is strongly regular, i.e., we assume that the assumptions of the statement 4 of Theorem 9.3.10 hold.

Theorem 9.4.4. *Let $\mathfrak{X}$ be an association scheme of rank 4 and diameter 2 on $n \geq 29$ vertices. Assume additionally, that the constituents of $\mathfrak{X}$ are ordered by degree and the inequality $k_2 \leq \varepsilon k_3/2$ holds for some $\varepsilon < 10^{-11}$. Suppose that $k_2 \leq \frac{101}{100}k_1$ and X_2 is strongly regular with smallest eigenvalue -2. Then*

$$q(X_{1,2}) + \xi(X_{1,2}) \leq \frac{99}{100}(k_1 + k_2).$$

Proof. The assumptions of the theorem imply that the inequalities from Lemma 9.3.3 hold.

Since X_2 is strongly regular with smallest eigenvalue -2 and $n \geq 29$, by Seidel's classification (Theorem 5.1.11), by Lemma 9.3.3, we conclude that

$$k_2/2 \geq p^2_{2,2} \geq k_2/2 - 1 \quad \text{and} \quad p^1_{2,2} = p^3_{2,2} \leq \varepsilon k_2. \tag{9.32}$$

By Proposition 9.2.4, for $\varepsilon_1 = 25\varepsilon^{1/3}$, we have

$$\xi(X_{1,2}) \leq \frac{p^2_{2,2} + p^1_{2,1} + p^1_{1,1} + \sqrt{(p^1_{2,1} + p^1_{1,1} - p^2_{2,2})^2 + 4p^2_{1,1}p^1_{2,1}}}{2} + \varepsilon_1(k_1 + k_2), \text{ so} \tag{9.33}$$

$$\begin{aligned} \xi(X_{1,2}) &\leq \max(p^2_{2,2}, p^1_{2,1} + p^1_{1,1}) + \sqrt{p^2_{1,1}p^1_{1,2}} + \varepsilon_1(k_1 + k_2) \leq \\ &\leq \max(p^2_{2,2} + p^1_{2,1}, 2p^1_{2,1} + p^1_{1,1}) + \varepsilon_1(k_1 + k_2) \leq \\ &\leq \max(p^2_{2,2} + p^1_{2,1}, \lambda_1(X_{1,2})) + \varepsilon_1(k_1 + k_2). \end{aligned} \tag{9.34}$$

Recall also, that similarly as in Eq. (9.22) and Eq. (9.25), we have

$$\mu(X_{1,2}) \leq 2\varepsilon(k_1 + k_2) \quad \text{and} \quad \max(\lambda_1(X_{1,2}), \lambda_2(X_{1,2})) \leq \frac{1 + 2\varepsilon}{2}(k_1 + k_2). \tag{9.35}$$

Case 1. Assume $p^1_{1,2} > 2k_1/5$.

Then, using that $k_2 \leq \frac{101}{100}k_1 \leq \frac{11}{10}k_1$, and using Eq. (9.32),

$$\lambda_1(X_{1,2}) = p^1_{1,1} + 2p^1_{1,2} + p^1_{2,2} \geq \frac{4}{5}k_1 \geq \frac{4}{11}(k_1 + k_2), \text{ and}$$

$$\lambda_2(X_{1,2}) = p^2_{2,2} + 2p^2_{1,2} + p^2_{1,1} \geq \frac{k_2}{2} + \frac{2}{5}\frac{10}{11}k_1 \geq \frac{4}{11}(k_1 + k_2).$$

Since Eq. (9.35) holds, $X_{1,2}$ satisfies the assumptions of Theorem 3.1.3 for $m = 2$. So, by Lemma 9.3.9, it is strongly regular with smallest eigenvalue -2. However, by Theorem 9.4.2, under the assumptions of this proposition $X_{1,2}$ and X_2 cannot be strongly regular with smallest eigenvalue -2 simultaneously. Hence, this case is impossible.

Case 2. Assume $2k_1/5 \geq p^1_{1,2} \geq k_1/8$.

Case 2.a Suppose $\lambda_2(X_{1,2}) \geq \lambda_1(X_{1,2})$.

Then, since $k_2 \leq \frac{11}{10}k_1$, using Eq. (9.32), the inequality

$$q(X_{1,2}) \leq \lambda_2(X_{1,2}) = p^2_{2,2} + 2p^2_{1,2} + p^2_{1,1} \leq \frac{k_2}{2} + 2\varepsilon k_1 + \frac{2}{5}k_1 \leq \frac{49}{100}(k_1 + k_2)$$

holds. At the same time, by Eq. (9.34), we get

$$\xi(X_{1,2}) \leq \max\left(\frac{1}{2}k_2 + \frac{2}{5}k_1, \lambda_1(X_{1,2})\right) + \varepsilon_1(k_1 + k_2) \leq$$

$$\leq \max\left(\frac{49}{100}(k_1 + k_2), \lambda_2(X_{1,2})\right) + \varepsilon_1(k_1 + k_2) \leq \frac{1}{2}(k_1 + k_2).$$

Therefore, $q(X_{1,2}) + \xi(X_{1,2}) \leq \frac{99}{100}(k_1 + k_2)$.

Case 2.b Suppose $\lambda_1(X_{1,2}) \geq \lambda_2(X_{1,2})$.

We may assume that $\lambda_1(X_{1,2}) \geq \frac{49}{100}(k_1 + k_2)$. Otherwise, by Eq. (9.34),

$$q(X_{1,2}) + \xi(X_{1,2}) \leq \lambda_1(X_{1,2}) + \max\left(\frac{k_2}{2} + \frac{2k_1}{5}, \lambda_1(X_{1,2})\right) + \varepsilon_1(k_1 + k_2) \leq \frac{99}{100}(k_1 + k_2).$$

Let $p^1_{1,2} = \alpha k_1$, then $1/8 \leq \alpha \leq 2/5$. The inequality

$$p^1_{1,1} + p^1_{2,2} + 2p^1_{1,2} = \lambda_1(X_{1,2}) \geq \frac{49}{100}(k_1 + k_2)$$

and Eq. (9.32) imply that

$$p^1_{1,1} + p^1_{1,2} \geq \frac{49}{100}(k_1 + k_2) - \varepsilon k_2 - \alpha k_1 \geq \left(\frac{49}{50} - 2\varepsilon - \alpha\right)k_1 \geq \frac{28}{50}k_1 \geq \frac{28}{55}k_2 > p^2_{2,2}. \quad (9.36)$$

On the other hand, Eq. (9.35) implies

$$p^1_{1,1} + p^1_{1,2} \le \lambda_1(X_{1,2}) - p^1_{1,2} \le \left(\frac{1}{2} + \varepsilon\right)(k_1 + k_2) - \alpha k_1 \le \left(\frac{101}{100} - \alpha\right) k_1. \tag{9.37}$$

Hence, as $p^2_{2,2} \ge (1/2 - \varepsilon)k_2 \ge (1/2 - \varepsilon)k_1$, Eq. (9.36) and (9.37) imply that

$$\left|p^1_{1,1} + p^1_{1,2} - p^2_{2,2}\right| \le \left(\frac{101}{100} - \alpha\right) k_1 - \left(\frac{1}{2} - \varepsilon\right) k_1 \le \left(\frac{52}{100} - \alpha\right) k_1.$$

Therefore, using Eq. (9.33), Eq. (9.37) and $p^2_{2,2} \le k_2/2 \le \frac{101}{200}k_1$, we get for $1/8 \le \alpha \le 2/5$

$$\xi(X_{1,2}) \le \frac{\frac{101}{200} + \left(\frac{101}{100} - \alpha\right) + \sqrt{\left(\frac{52}{100} - \alpha\right)^2 + 4\alpha^2}}{2} k_1 + \frac{201}{100}\varepsilon_1 k_1 \le \frac{195}{200}k_1, \tag{9.38}$$

Thus,

$$q(X_{1,2}) + \xi(X_{1,2}) \le \frac{1 + 2\varepsilon}{2}(k_1 + k_2) + \frac{195}{200}k_1 \le \frac{99}{100}(k_1 + k_2).$$

Case 3. Assume $p^1_{1,2} < k_1/8$.

Then, using Eq. (9.34), Corollary 9.3.8 and inequality $k_2 \le \frac{101}{100}k_1$,

$$\xi(X_{1,2}) \le \max(p^2_{2,2} + p^1_{2,1}, 2p^1_{2,1} + p^1_{1,1}) + \varepsilon_1(k_1 + k_2) \le$$

$$\le \max\left(\frac{1}{2}k_2 + \frac{1}{8}k_1, \frac{1}{4}k_1 + \frac{1 + \varepsilon}{2}k_1\right) + \varepsilon_1(k_1 + k_2) \le \frac{2}{5}(k_1 + k_2).$$

Combining this with Eq. (9.35) we get $q(X_{1,2}) + \xi(X_{1,2}) \le \frac{99}{100}(k_1 + k_2)$. □

9.4.4 Constituent X_1 is strongly regular

The common strategy of our proofs is to prove a good spectral gap for a certain union of the constituents, or to apply Metsch's criteria (Theorem 3.1.3) to a certain union of the constituents. The next lemma covers the range of parameters for which spectral gap is hard

to achieve, and the conditions of Metsch's criteria are not satisfied for X_2 and $X_{1,2}$. However, in this range of parameters, we are still able to use the idea of Metsch's proof to show that k_2 does not differ much from k_1. This will suffice for our purposes.

Definition 9.4.5. For a homogeneous configuration $\mathfrak{X}$ and disjoint non-empty sets of edge colors I and J we say that vertices $x, y_1, y_2, ..., y_t$ form a *t-claw* (claw of size t) in colors (I, J) if $c(x, y_i) \in I$ and $c(y_i, y_j) \in J$ for all distinct $1 \leq i, j \leq t$.

Lemma 9.4.6. *Let $\mathfrak{X}$ be an association scheme of rank 4 and diameter 2 with constituents ordered by degree. Suppose that the inequality $k_2 \leq \varepsilon k_3/2$ holds for some $0 < \varepsilon \leq \frac{1}{100}$. Assume additionally, that for some $0 < \delta \leq \frac{1}{100}$ we have*

$$p^2_{2,2} \geq \frac{1-\delta}{2}k_2 \quad \textit{and} \quad \frac{1}{8}k_2 \leq p^1_{2,2} \leq \frac{1}{3}k_2.$$

Then $k_2 \leq 20k_1$.

Proof. The assumptions of the lemma ensure that the inequalities from Lemma 9.3.3 hold.

First, we show that under the assumptions of the lemma there are no 3-claw in colors $(2, 3)$ in $\mathfrak{X}$. That is, for $x, y_1, y_2, y_3 \in \mathfrak{X}$ it is not possible that $c(x, y_i) = 2$ and $c(y_i, y_j) = 3$ for all distinct $i, j \in [3]$. Indeed, suppose such x, y_i exist. Let $U_i = N_2(x) \cap N_2(y_i)$. Then

$$|U_i| = p^2_{2,2} \geq \frac{1-\delta}{2}k_2, \qquad |U_i \cap U_j| \leq |N_2(y_i) \cap N_2(y_j)| = p^3_{2,2} \leq \varepsilon k_2 \quad \text{and}$$

$$|U_1 \cup U_2 \cup U_3| \leq |N_2(x)| = k_2.$$

Therefore, we should have $k_2 \geq 3\frac{(1-\delta)}{2}k_2 - 3\varepsilon k_2$, a contradiction. Hence, the size of a maximal claw in colors $(2, 3)$ is 2.

Now, we claim that every edge of color 2 lies inside a clique of size at least $p^2_{2,2} - p^3_{2,2} - p^3_{2,1}$ in $X_{1,2}$. Consider any edge $\{x, y\}$ of color 2. Let z be a vertex which satisfies $c(x, z) = 2$

and $c(y,z)=3$. Define

$$C(x,y)=\{x,y\}\cup\{w : c(z,w)=3 \text{ and } c(x,w)=2,\ c(y,w)=2\}. \tag{9.39}$$

Observe that

$$|C(x,y)|\geq 2+p^2_{2,2}-p^3_{2,2}-p^3_{2,1}. \tag{9.40}$$

At the same time, if $z_1,z_2\in C(x,y)$ satisfy $c(z_1,z_2)=3$, then x,z,z_1,z_2 form a 3-claw in colors $(2,3)$, which contradicts our claim above. Hence, $C(x,y)$ is a clique in $X_{1,2}$.

Assume that there is an edge $\{y_1,y_2\}$ in $C(x,y)$ of color 1 for some x,y. Then

$$2k_1+\frac{1}{3}k_2\geq 2\sum_{i=1}^{3}p^1_{1,i}+p^1_{2,2}\geq p^1_{1,1}+2p^1_{1,2}+p^1_{2,2}\geq |C(x,y)|-2\geq \frac{1-\delta-2\varepsilon}{2}k_2-k_1.$$

Therefore, $k_2\leq 20k_1$.

Assume now that all edges in $C(x,y)$ are of color 2 for all x,y, that is, $C(x,y)$ is a clique in X_2. Let $\mathcal{C}$ be the set of all maximal cliques in X_2 of size at least $p^2_{2,2}-p^3_{2,2}-p^3_{2,1}$. Then we have proved that every edge of color 2 is covered by at least one clique in $\mathcal{C}$. Consider, two distinct cliques $C_1,C_2\in\mathcal{C}$. There is a pair of vertices $v\in C_1\setminus C_2$ and $u\in C_2\setminus C_1$ with $c(v,u)\neq 2$. Thus,

$$|C_1\cap C_2|\leq \max(p^1_{2,2},p^3_{2,2})\leq k_2/3. \tag{9.41}$$

Suppose first that some pair of distinct cliques $C_1,C_2\in\mathcal{C}$ satisfies $|C_1\cap C_2|\geq 2$ and let $\{x,y\}\subseteq C_1\cap C_2$. Then $c(x,y)=2$ and every vertex in $C_1\cup C_2$ is adjacent to both x and y by an edge of color 2. Thus,

$$p^2_{2,2}\geq |C_1\cup C_2|-2=|C_1|+|C_2|-|C_1\cap C_2|-2\geq 2(p^2_{2,2}-p^3_{2,2}-k_1)-\frac{1}{3}k_2,$$

so, using Lemma 9.3.3,

$$\left(\frac{1}{3}+2\varepsilon\right)k_2+2k_1 \geq p^2_{2,2} \geq \frac{1-\delta}{2}k_2.$$

Hence, $k_2 \leq 20k_1$.

Finally, if for every pair of distinct cliques $C_1, C_2 \in \mathcal{C}$ we have $|C_1 \cap C_2| \leq 1$, then every edge of color 2 lies in at most one clique of $\mathcal{C}$. Above we proved that every edge of color 2 lies in at least one clique of $\mathcal{C}$, so it lies in exactly one.

Therefore, since $p^2_{2,2} \geq \frac{1-\delta}{2}k_2$, we get that either $p^3_{2,1} \geq k_2/10$, and so $k_2 \leq 10k_1$; or, by Eq. (9.40), $|C| > k_2/3+1$ for every $C \in \mathcal{C}$, and so every vertex lies in at most 2 cliques from $\mathcal{C}$. In the latter case, by Lemma 3.1.9, we get that $p^1_{2,2} \leq 4$, which contradicts $p^1_{2,2} \geq k_2/8$, since $k_2 \geq p^3_{2,2}/\varepsilon \geq 1/\varepsilon$ by Lemma 9.3.3. □

Furthermore, we can get a linear inequality between k_1 and k_2 if we know that $X_{1,2}$ has a clique geometry.

Lemma 9.4.7. *Let $\mathfrak{X}$ be an association scheme of rank 4 on $n \geq 29$ vertices, with diameter 2 and constituents ordered by degree. Assume the inequality $k_2 \leq \varepsilon k_3/2$ holds for some $\varepsilon < \frac{1}{10}$. Suppose $X_{1,2}$ has a clique geometry such that every vertex belongs to at most m cliques. Then*

$$p^1_{2,3} \leq \frac{m^2-2}{2}k_1 \quad \textit{and} \quad k_2 \leq \frac{3}{2-4\varepsilon}(m^2-2)k_1.$$

If, additionally, X_1 is a strongly regular graph with smallest eigenvalue -2, then

$$p^1_{2,3} \leq \frac{m^2-4}{8}k_1 \quad \textit{and} \quad k_2 \leq \frac{3}{8(1-2\varepsilon)}(m^2-4)k_1.$$

Proof. By Lemma 3.1.9 applied to $X_{1,2}$, we know

$$p^3_{1,1}+2p^3_{1,2}+p^3_{2,2} = \mu(X_{1,2}) \leq m^2. \tag{9.42}$$

Since $\mathfrak{X}$ is of diameter 2, we have that $p^3_{1,1} \geq 1$ and $p^3_{2,2} \geq 1$, and $k_1(k_1-1) \geq k_3$. Thus

$$p^3_{1,2} \leq \frac{m^2-2}{2}, \text{ so } p^1_{2,3} \leq \frac{m^2-2}{2}\frac{k_3}{k_1} \leq \frac{m^2-2}{2}k_1. \tag{9.43}$$

By Eq. (2.3), $p^1_{2,1} + p^1_{2,2} + p^1_{2,3} = k_2$, and Corollary 9.3.7 implies that $p^1_{2,3} + \varepsilon k_2 \geq \max(p^1_{2,2}, p^1_{2,1})$. Thus, combining with Eq. (9.43), we get

$$\frac{1-2\varepsilon}{3}k_2 \leq p^1_{2,3} \leq \frac{m^2-2}{2}k_1.$$

If X_1 is strongly regular with smallest eigenvalue -2, we can get better estimates. By Seidel's classification, X_1 is either $T(s)$ or $L_2(s)$ for some s. Thus, either $n = s(s-1)/2$ and $k_1 = 2(s-2)$, or $n = s^2$ and $k_1 = 2(s-1)$. In any case, $4k_3 \leq k_1^2$. Observe that Corollary 9.3.7 implies $p^2_{2,3} + \varepsilon k_2 \geq \max(p^2_{2,2}, p^2_{2,1})$. Hence, $p^2_{2,3} \geq \dfrac{1-2\varepsilon}{3}k_2$, so

$$p^3_{2,2} \geq \frac{(1-2\varepsilon)(k_2)^2}{3k_3} \geq \frac{4(1-2\varepsilon)}{3}.$$

At the same time, $p^3_{1,1} = \mu(X_1) \geq 2$ for $X_1 = T(s)$, or $X_1 = L_2(s)$. Thus, $p^3_{i,i} \geq 2$ for $i = 1$ and $i = 2$. Therefore, as in Eq. (9.43), by Eq. (9.42),

$$p^1_{2,3} \leq \frac{m^2-4}{2} \cdot \frac{k_3}{k_1} \leq \frac{m^2-4}{8}k_1.$$

Again, $p^1_{2,3} \geq \dfrac{1-2\varepsilon}{3}k_2$ implies the desired inequality between k_1 and k_2. □

Now, we are ready to consider the case when the constituent X_1 is strongly regular (case of statement 3 of Theorem 9.3.10).

Theorem 9.4.8. *Let $\mathfrak{X}$ be an association scheme of rank 4 on $n \geq 29$ vertices with diameter 2 and constituents ordered by degree. Assume additionally, that the parameters of $\mathfrak{X}$ satisfy $k_2 \leq \varepsilon k_3/2$ for $\varepsilon = 10^{-26}$. Suppose that X_1 is a strongly regular graph with smallest*

eigenvalue -2. *Then*

$$q(Y) + \xi(Y) \leq (1-\varepsilon)k_Y, \tag{9.44}$$

where either $Y = X_2$ *and* $k_Y = k_2$, *or* $Y = X_{1,2}$ *and* $k_Y = k_1 + k_2$.

Proof. The assumptions of the lemma ensure that the inequalities from Lemma 9.3.3 hold.

Since X_1 is strongly regular with smallest eigenvalue -2, Seidel's classification and Lemma 9.3.3 implies that

$$\frac{1}{2}k_1 \geq p^1_{1,1} \geq \frac{1}{2}k_1 - 1 \geq \left(\frac{1}{2} - \varepsilon\right)k_1 \quad \text{and} \quad p^2_{1,1} = p^3_{1,1} \leq \varepsilon k_1. \tag{9.45}$$

By Lemma 9.2.4, for $\varepsilon_1 = 25\varepsilon^{1/3} \leq \frac{2}{3}10^{-7}$ we have

$$\xi(X_{1,2}) \leq \frac{p^1_{1,1} + p^2_{1,2} + p^2_{2,2} + \sqrt{(p^2_{1,2} + p^2_{2,2} - p^1_{1,1})^2 + 4p^1_{2,2}p^2_{1,2}}}{2} + \varepsilon_1(k_1 + k_2). \tag{9.46}$$

Since $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$,

$$\xi(X_{1,2}) \leq \max(p^1_{1,1}, p^2_{2,2} + p^2_{2,1}) + \sqrt{p^1_{2,2}p^2_{2,1}} + \varepsilon_1(k_1 + k_2). \tag{9.47}$$

Using that $\lambda_2(X_{1,2}) \geq p^2_{2,2} + p^2_{2,1}$ (see Eq. (9.23)) and $p^2_{2,1} = p^1_{2,2}k_1/k_2 \leq p^1_{2,2}$, we can simplify it even more

$$\xi(X_{1,2}) \leq \max\left(\frac{k_1}{2}, \lambda_2(X_{1,2})\right) + p^1_{2,2} + \varepsilon_1(k_1 + k_2). \tag{9.48}$$

Recall that as in Eq. (9.22) and Eq. (9.25), we have

$$\mu(X_{1,2}) \leq 2\varepsilon(k_1 + k_2) \quad \text{and} \quad \max(\lambda_1(X_{1,2}), \lambda_2(X_{1,2})) \leq \frac{1+2\varepsilon}{2}(k_1 + k_2). \tag{9.49}$$

Case A. Suppose $p^2_{2,2} \geq (2-2\delta)p^1_{2,2}$ for $\delta = 10^{-7}$.

Using Corollary 9.3.8 for $p^2_{2,2}$, we get

$$p^1_{2,2} \le \frac{1+\varepsilon}{4(1-\delta)} k_2 \quad \text{and} \quad p^2_{2,1} = \frac{k_1}{k_2} p^1_{2,2} \le \frac{1+\varepsilon}{4(1-\delta)} k_1. \tag{9.50}$$

Note, by Corollary 9.3.7 and Eq. (9.45), $\varepsilon k_1 + p^1_{1,3} \ge p^1_{1,1} \ge (1/2 - \varepsilon)\, k_1$. So, Eq. (2.3) implies $p^1_{1,2} \le 3\varepsilon k_1$. Therefore, by Eq. (9.45),

$$\lambda_1(X_{1,2}) = p^1_{1,1} + p^1_{2,2} + 2p^1_{1,2} \le \frac{1}{2} k_1 + p^1_{2,2} + 6\varepsilon k_1. \tag{9.51}$$

Assume that Eq. (9.44) is not satisfied, then

$$\begin{aligned} &(1-\varepsilon)(k_1+k_2) \le q(X_{1,2}) + \xi(X_{1,2}) \le \\ &\quad \le \max\left(\lambda_2(X_{1,2}), \frac{1}{2}k_1 + p^1_{2,2} + 6\varepsilon k_1\right) + \max\left(\lambda_2(X_{1,2}), \frac{k_1}{2}\right) + p^1_{2,2} + \varepsilon_1(k_1+k_2) \end{aligned} \tag{9.52}$$

Observe, that if $\lambda_2(X_{1,2}) \le k_1/2 + p^1_{2,2} + 6\varepsilon k_1$, then using Eq. (9.50) we get a contradiction

$$(1-\varepsilon)(k_1+k_2) \le k_1 + 3\frac{1+\varepsilon}{4(1-\delta)} k_2 + \varepsilon_1(k_1+k_2) + 6\varepsilon k_1.$$

Otherwise, Eq. (9.52) implies

$$(1-\varepsilon-\varepsilon_1)(k_1+k_2) \le 2\lambda_2(X_{1,2}) + \frac{1+\varepsilon}{4(1-\delta)} k_2, \text{ so } \quad \lambda_2(X_{1,2}) \ge \frac{5}{6} k_1.$$

We estimate the expression under the root sign in Eq. (9.46), using that $\lambda_2(X_{1,2}) \ge \frac{5}{6}k_1$, using Eq. (9.50), Eq. (9.45), and inequality $\frac{1+\varepsilon}{4(1-\delta)^2} \le \frac{1+\varepsilon}{4} + \delta$ for $0 \le \varepsilon, \delta \le \frac{1}{2}$.

$$(p^2_{1,2} + p^2_{2,2} - p^1_{1,1})^2 + 4p^1_{2,2}p^2_{1,2} = (p^2_{1,2} + p^2_{2,2})^2 - 2p^1_{1,1}(p^2_{1,2} + p^2_{2,2}) + (p^1_{1,1})^2 + 4p^1_{2,2}p^2_{1,2} \le$$

$$\leq (p_{1,2}^2 + p_{2,2}^2)^2 - 2p_{1,1}^1 p_{1,2}^2 - (1-2\varepsilon)k_1 p_{2,2}^2 + \frac{(k_1)^2}{4} + \frac{1+\varepsilon}{2(1-\delta)^2} k_1 p_{2,2}^2 \leq$$

$$\leq (p_{1,2}^2 + p_{2,2}^2)^2 - k_1 p_{1,2}^2 + \left(\frac{(k_1)^2}{4} - \frac{1}{2} k_1 p_{2,2}^2\right) + (3\varepsilon + 2\delta) k_1 k_2 \leq$$

$$\leq (p_{1,2}^2 + p_{2,2}^2)^2 - \frac{k_1}{2}\left(\lambda_2(X_{1,2}) - \varepsilon k_1 - \frac{k_1}{2}\right) + (3\varepsilon + 2\delta) k_1 k_2 \leq$$

$$\leq (p_{1,2}^2 + p_{2,2}^2)^2 - \frac{1}{6}(k_1)^2 + (4\varepsilon + 2\delta) k_1 k_2.$$

Thus, since $\sqrt{x}$ is concave, using Eq. (9.49), and inequalities $\sqrt{y^2 - x^2} \leq y - x^2/(2y)$ and $p_{2,1}^2 + p_{2,2}^2 \leq \lambda_2(X_{1,2})$, we obtain

$$\sqrt{(p_{1,2}^2 + p_{2,2}^2 - p_{1,1}^1)^2 + 4p_{2,2}^1 p_{1,2}^2} \leq p_{1,2}^2 + p_{2,2}^2 - \frac{2(k_1)^2}{13(k_1+k_2)} + \sqrt{(4\varepsilon + 2\delta) k_1 k_2}. \quad (9.53)$$

Denote $\varepsilon_4 = \sqrt{4\varepsilon + 2\delta} < 2^{-1} \cdot 10^{-3}$. Hence, by Eq. (9.45), Eq. (9.46) and Eq. (9.53),

$$\xi(X_{1,2}) \leq \frac{k_1}{4} + (p_{1,2}^2 + p_{2,2}^2) - \frac{(k_1)^2}{13(k_1+k_2)} + \left(\frac{1}{2}\varepsilon_4 \sqrt{k_1 k_2} + \varepsilon_1 (k_1 + k_2)\right).$$

Using Corollary 9.3.8 for $p_{2,2}^2$ and Eq. (9.50), we get

$$\begin{aligned} \xi(X_{1,2}) \leq \frac{k_1}{4} + \left(\frac{(1+\varepsilon)}{4(1-\delta)} k_1 + \frac{(1+\varepsilon)}{2} k_2\right) - \frac{(k_1)^2}{13(k_1+k_2)} + \left(\frac{\varepsilon_4}{4} + \varepsilon_1\right)(k_1+k_2) \leq \\ \leq \frac{1}{2}(k_1+k_2) - \frac{(k_1)^2}{13(k_1+k_2)^2}(k_1+k_2) + \varepsilon_5 (k_1+k_2), \end{aligned} \quad (9.54)$$

where $\varepsilon_5 = \varepsilon_1 + \frac{1}{4}\varepsilon_4 + \delta + \varepsilon < 6^{-1} \cdot 10^{-3}$. Thus, we want either $q(X_{1,2})$ to be bounded away from $(k_1 + k_2)/2$, or to have $k_1 \leq c k_2$ for some absolute constant c.

Observe, by Eq. (9.50), Eq. (9.51) and Eq. (9.54),

$$\begin{aligned}\lambda_1(X_{1,2}) + \xi(X_{1,2}) &\leq \left(\frac{k_1}{2} + p^1_{2,2} + 2\varepsilon k_1\right) + \frac{k_1+k_2}{2} + \varepsilon_5(k_1+k_2) \leq \\ &\leq k_1 + \frac{3}{4}k_2 + (3\varepsilon+\delta)(k_1+k_2) + \varepsilon_5 k_2 \leq (1-\varepsilon)(k_1+k_2),\end{aligned} \tag{9.55}$$

$$\lambda_2(X_{1,2}) + \xi(X_{1,2}) \leq \lambda_2(X_{1,2}) + \frac{k_1+k_2}{2} + \varepsilon_5(k_1+k_2). \tag{9.56}$$

Clearly,

$$\mu(X_{1,2}) + \xi(X_{1,2}) \leq 2\varepsilon(k_1+k_2) + \xi(X_{1,2}) \leq (1-\varepsilon)(k_1+k_2).$$

Thus, either we have $q(X_{1,2})+\xi(X_{1,2}) \leq (1-\varepsilon)(k_1+k_2)$, or $\lambda_2(X_{1,2}) \geq (1/2-\varepsilon_5-\varepsilon)(k_1+k_2)$. Suppose that $\lambda_2(X_{1,2}) \geq (1/2-\varepsilon_5-\varepsilon)(k_1+k_2)$. By the assumption of Case A, we have $p^2_{2,2} \geq (2-2\delta)p^2_{1,2}k_2/k_1$, so Eq (9.45) implies

$$\lambda_2(X_{1,2}) = p^2_{2,2} + 2p^2_{1,2} + p^2_{1,1} \leq p^2_{2,2} + \frac{1}{1-\delta}\frac{k_1}{k_2}p^2_{2,2} + \varepsilon k_1 \leq \left(\frac{1}{1-\delta}\frac{p^2_{2,2}}{k_2} + \varepsilon\right)(k_1+k_2).$$

Hence, in this case

$$p^2_{2,2} \geq \left(\frac{1}{2} - \varepsilon_5 - 2\varepsilon\right)(1-\delta)k_2. \tag{9.57}$$

Case A.1 Assume $p^1_{2,2} < k_2/8$.

Then $\mu(X_2) \leq k_2/8$ and $\lambda(X_2) = p^2_{2,2}$. Therefore, X_2 satisfies the assumptions of Theorem 3.1.3 for $m = 2$. Thus, by Lemma 3.1.9 for graph X_2, we get $p^3_{2,2} \leq m^2 = 4$. At the same time, by Crorllary 9.3.7 and Eq. (2.3) we have $p^2_{2,3} \geq \frac{1-2\varepsilon}{3}k_2$. Therefore,

$$4k_3 \geq p^3_{2,2}k_3 = p^2_{2,3}k_2 \geq k_2\frac{(1-2\varepsilon)}{3}k_2 \geq \frac{1}{4}(k_2)^2.$$

Combining with $(k_1)^2 \geq k_3$, we obtain $k_2 \leq 4k_1$.

Case A.2 Assume $p^1_{2,2} \geq k_2/8$.

Then, since Eq. (9.50) and Eq. (9.57) hold, by Lemma 9.4.6, we get that $k_2 \leq 20k_1$.

Hence, Eq. (9.54) and Eq. (9.49) imply

$$\lambda_2(X_{1,2}) + \xi(X_{1,2}) \leq (1-\varepsilon)(k_1+k_2).$$

Therefore, using bound on $\mu(X_{1,2})$ and Eq. (9.55), we get $q(X_{1,2})+\xi(X_{1,2}) \leq (1-\varepsilon)(k_1+k_2)$.

Case B. Suppose $p^2_{2,2} \leq (2-2\delta)p^1_{2,2}$.

In this case, in several ranges of parameters we will show that $X_{1,2}$ has a clique geometry. We first establish the following bounds.

By Corollary 9.3.8, $2p^2_{2,1} \leq (1+\varepsilon)k_1$. Assume $p^1_{2,2} \geq k_2/5$ and $m \leq 5$, then Eq. (9.49) and Eq. (9.45) implies

$$2\lambda_1(X_{1,2}) - \lambda_2(X_{1,2}) \geq k_1 + 2p^1_{2,2} - p^2_{2,2} - 2p^2_{21} - 3\varepsilon k_1 \geq 2\delta p^1_{2,2} - 4\varepsilon k_1 \geq (2m-1)\mu(X_{1,2}). \tag{9.58}$$

Suppose that $\lambda_2(X_{1,2}) \geq (1/4 + 2m\varepsilon)(k_1+k_2)$, then Eq. (9.49) implies

$$2\lambda_2(X_{1,2}) - \lambda_1(X_{1,2}) \geq (2m-1)\mu(X_{1,2}). \tag{9.59}$$

Case B.1. Assume $p^1_{2,2} \geq k_2/3$.

Then, by Eq. (9.45),

$$\lambda_1(X_{1,2}) \geq p^1_{1,1} + p^1_{2,2} \geq \left(\frac{1}{2} - \varepsilon\right)k_1 + \frac{1}{3}k_2 \geq \frac{1}{3}(k_1+k_2). \tag{9.60}$$

Case B.1.a. Suppose $p^2_{2,2} \geq k_2/3$.

Then $\lambda_2(X_{1,2}) \geq (k_1+k_2)/3$. Thus, in notations of Theorem 3.1.3 we get for $X_{1,2}$ that $4\lambda^{(1)} - 6\mu(X_{1,2}) \geq k_1+k_2$, and by Eq. (9.58)-(9.59), inequality $2\lambda^{(1)} - \lambda^{(2)} \geq 5\mu(X_{1,2})$ holds. Hence, by Theorem 3.1.3, the graph $X_{1,2}$ has a clique geometry with $m = 3$.

Thus, by Lemma 9.4.7, we have $k_2 \leq \frac{15}{8(1-2\varepsilon)} k_1 \leq 2k_1$. Therefore,

$$\lambda_1(X_{1,2}) \geq p^1_{1,1} + p^1_{2,2} \geq \left(\frac{1}{2} - \varepsilon\right) k_1 + \frac{1}{3} k_2 > \left(\frac{1}{3} + 4\varepsilon\right)(k_1 + k_2),$$

$$\lambda_2(X_{1,2}) \geq p^2_{2,2} + 2p^2_{1,2} \geq \frac{k_2}{3} + \frac{2k_1}{3} > \left(\frac{1}{3} + 4\varepsilon\right)(k_1 + k_2).$$

Therefore, $X_{1,2}$ satisfies Theorem 3.1.3 for $m = 2$, and so by Lemma 9.3.9, it is strongly regular with smallest eigenvalue -2. However, by Theorem 9.4.2 and Remark 9.4.3, under the assumptions of this theorem the graphs X_1 and $X_{1,2}$ cannot be simultaneously strongly regular with smallest eigenvalue -2.

Case B.1.b. Suppose $p^2_{2,2} < k_2/3$.

Then, in particular, $p^1_{2,2} \geq p^2_{2,2}$, so $q(X_2) = p^1_{2,2}$. Take $0 \leq \alpha \leq \frac{1+\varepsilon}{2} \leq \frac{51}{100}$, and $0 \leq \gamma \leq 1$, so that $p^1_{2,2} = \alpha k_2$ and $k_1 = \gamma k_2$. Using Eq. (9.10) and Eq. (9.45), compute

$$q(X_2) + \xi(X_2) \leq p^1_{2,2} + p^2_{2,2} + \varepsilon k_2 + \sqrt{p^1_{2,2} p^2_{1,2}} + \varepsilon_1 k_2 = p^2_{2,2} + (\alpha + \alpha\sqrt{\gamma} + \varepsilon + \varepsilon_1) k_2 \quad (9.61)$$

If $p^2_{2,2} \leq (1 - \alpha(1+\sqrt{\gamma}) - \varepsilon_1 - 2\varepsilon)k_2$, then $q(X_2) + \xi(X_2) \leq (1-\varepsilon)k_2$ and we reached our goal. So, assume that $p^2_{2,2} \geq (1 - \alpha(1+\sqrt{\gamma}) - \varepsilon_1 - 2\varepsilon)k_2$. We compute

$$\begin{aligned} \lambda_2(X_{1,2}) = p^2_{2,2} + 2p^2_{1,2} + p^2_{1,1} &\geq (1 - \alpha(1+\sqrt{\gamma}) - \varepsilon_1 - 2\varepsilon)k_2 + 2\alpha\gamma k_2 \geq \\ &\geq \left(\frac{1 - \alpha(1+\sqrt{\gamma} - 2\gamma) - \varepsilon_1 - 2\varepsilon}{1+\gamma}\right)(k_1 + k_2) \geq \frac{3}{10}(k_1 + k_2), \end{aligned} \quad (9.62)$$

where we use that $1 + \sqrt{\gamma} - 2\gamma \geq 0$ for $0 \leq \gamma \leq 1$, so expression is minimized for $\alpha = (1+\varepsilon)/2$ and after that we compute the minimum of the expression for $0 \leq \gamma \leq 1$. Thus, by Eq. (9.58)-(9.60), the graph $X_{1,2}$ has a clique geometry for $m = 3$. Hence,

by Lemma 9.4.7, we have $k_2 \leq 2k_1$. This implies that $\frac{1}{2} \leq \gamma \leq 1$. We compute,

$$\begin{aligned} &\min_{1/2\leq\gamma\leq 1} \min_{0\leq\alpha\leq\frac{51}{100}} \left(\frac{1-\alpha(1+\sqrt{\gamma}-2\gamma)-\varepsilon_1-2\varepsilon}{1+\gamma} \right) = \\ &= \min_{1/2\leq\gamma\leq 1} \left(\frac{1-\frac{51}{100}(1+\sqrt{\gamma}-2\gamma)-\varepsilon_1-2\varepsilon}{1+\gamma} \right) \geq \frac{9}{25} > \frac{1}{3}+2\varepsilon. \end{aligned} \tag{9.63}$$

Therefore, using also Eq. (9.58)-(9.60), we get that $X_{1,2}$ satisfies conditions of Theorem 3.1.3 for $m = 2$, so by Lemma 9.3.9, the graph $X_{1,2}$ is strongly regular with smallest eigenvalue -2. However, by Theorem 9.4.2 and Remark 9.4.3, this is impossible, since X_1 is also strongly regular with smallest eigenvalue -2.

Case B.2. Assume $k_2/3 \geq p^1_{2,2} \geq k_2/5$.

Then

$$\lambda_1(X_{1,2}) \geq p^1_{1,1} + p^1_{2,2} \geq \left(\frac{1}{2}-\varepsilon\right)k_1 + \frac{1}{5}k_2. \tag{9.64}$$

If $p^2_{2,2} \leq (1/3-\varepsilon-\varepsilon_1)k_2$, then by Eq. (9.10),

$$\begin{aligned} q(X_2)+\xi(X_2) &\leq \max(p^2_{2,2}, p^1_{2,2}) + p^2_{2,2} + \sqrt{p^1_{2,2}p^2_{1,2}} + \varepsilon_1 k_2 \leq \\ &\leq \frac{k_2}{3} + \left(\frac{1}{3}-\varepsilon-\varepsilon_1\right)k_2 + \frac{k_2}{3} + \varepsilon_1 k_2 \leq (1-\varepsilon)k_2. \end{aligned} \tag{9.65}$$

Else, $p^2_{2,2} \geq (1/3-\varepsilon-\varepsilon_1)k_2 \geq (1/4+10\varepsilon)k_2$, so

$$\lambda_2(X_{1,2}) \geq p^2_{2,2} + 2p^2_{1,2} \geq \left(\frac{1}{4}+10\varepsilon\right)k_2 + \frac{2}{5}k_1. \tag{9.66}$$

Thus, Eq. (9.58)-(9.59) and Eq. (9.64)-(9.66) imply, using Theorem 3.1.3, that the graph $X_{1,2}$ has a clique geometry with $m = 5$. Therefore, using Eq. (9.45) and Eq. (2.3), by Lemma 9.4.7,

$$\left(\frac{2}{3}-\varepsilon\right)k_2 \leq (1-\varepsilon)k_2 - p^1_{2,2} \leq p^1_{2,3} \leq \frac{m^2-4}{8}k_1, \text{ so } k_2 \leq 4k_1.$$

Hence, in fact, Eq. (9.64) implies

$$\lambda_1(X_{1,2}) \geq \frac{1}{5}k_2 + \left(\frac{1}{2} - \varepsilon\right) k_1 \geq \left(\frac{1}{4} + 6\varepsilon\right)(k_1 + k_2).$$

Thus, using Eq. (9.66) and Eq. (9.58) - (9.59), by Theorem 3.1.3, we get that $X_{1,2}$ has a clique geometry for $m = 3$. Thus, we can get a better estimate, as

$$\left(\frac{2}{3} - \varepsilon\right) k_2 \leq \frac{m^2 - 4}{8} k_1, \quad \text{implies} \quad k_2 \leq \frac{15}{16(1 - 2\varepsilon)} k_1 < k_1.$$

However, this contradicts our assumption that $k_2 \geq k_1$, so $p^2_{2,2} \geq (1/3 - \varepsilon - \varepsilon_1)k_2$ is impossible in this case.

Case B.3. Assume $p^1_{2,2} \leq k_2/5$.

Then, by the assumption of Case B, $p^2_{2,2} \leq (2 - 2\delta)p^1_{2,2} \leq (2 - 2\delta)k_2/5$, so

$$\begin{aligned} q(X_2) + \xi(X_2) &\leq \max(p^2_{2,2}, p^1_{2,2}, p^3_{2,2}) + p^2_{2,2} + \sqrt{p^1_{2,2}p^2_{1,2}} + \varepsilon_1 k_2 \leq \\ &\leq 2(2 - 2\delta)\frac{k_2}{5} + \frac{k_2}{5} + \varepsilon_1 k_2 \leq \left(1 - \frac{4}{5}\delta + \varepsilon_1\right) k_2 \leq (1 - \varepsilon)k_2. \quad \square \end{aligned} \tag{9.67}$$

9.4.5 Constituent that is the line graph of a triangle-free regular graph

Finally, we consider the case of the last possible outcome provided by Theorem 9.3.10, the case when one of the constituents is the line graph of a regular triangle-free graph and is not strongly regular.

First recall the following classical result due to Whitney.

Theorem 9.4.9 (Corollary to Whitney [1992])**.** *Let X be a connected graph on $n \geq 5$ vertices. Then the natural homomorphism $\phi : \mathrm{Aut}(X) \to \mathrm{Aut}(L(X))$ is an isomorphism $\mathrm{Aut}(L(X)) \cong \mathrm{Aut}(X)$.*

Observe, that the restriction on the diameter of the line graph gives quite strong bound on the degree of the base graph, as stated in the following lemma.

Lemma 9.4.10. *Let X be a k-regular graph on n vertices. If the line graph $L(X)$ has diameter 2, then $k \geq n/8$.*

Proof. Recall that $L(X)$ has $kn/2$ vertices and degree $2(k-1)$. Since $L(X)$ has diameter 2, the degree of the graph satisfies $4k^2 \geq 4(k-1)^2 + 2(k-1) + 1 \geq kn/2$, i.e., $k \geq n/8$. □

Theorem 9.4.11. *Let X be a connected k-regular triangle-free graph on $n \geq 5$ vertices, where $k \geq 3$. Suppose $\mathfrak{X}$ is an association scheme of rank 4 and diameter 2 on $V(L(X)) = E(X)$, such that one of the constituents is equal to $L(X)$ and is not strongly regular. Then every pair of vertices $u, v \in X$ is distinguished by at least $n/8$ vertices. Therefore,* $\mathrm{Aut}(L(X))$ *has order $n^{O(\log(n))}$ and the motion of $L(X)$ is at least $|V(L(X))|/16$.*

Proof. Denote the constituents of $\mathfrak{X}$ by Y_i, $0 \leq i \leq 3$, where Y_0 is the diagonal constituent and $Y_1 = L(X)$.

Since Y_1 has diameter 2, every induced cycle of X has length at most 5. The graph X is triangle-free, so every induced cycle in X has length 4 or 5, and every cycle of length 4 or 5 is induced.

Case 1. Suppose that there are no cycle of length 5 in X, i.e., it is bipartite.

Then for $v \in X$ there are no edges between vertices in $N_2(v)$. The graph X is regular, and every induced cycle has length 4, so for every vertex $w \in N_2(v)$ the neighborhoods $N(w)$ and $N(v)$ coincide. Hence, as X is connected, X is a complete regular bipartite graph. However, in this case, $L(X)$ is strongly regular.

Case 2. Suppose there is a cycle of length 5.

Let $v_1v_2v_3v_4v_5$ be any cycle of length 5. Take u different from v_2, v_5 and adjacent to v_1. Since the constituent Y_1 has diameter 2, the edges v_1u and v_3v_4 are at distance 2 in $L(X)$, thus there is one of the edges uv_3 or uv_4 in X. Again, X is triangle free, so exactly one of them is in X. Without loss of generality, assume that uv_3 is in X. In particular, we get that there is a cycle of length 4 $uv_1v_2v_3$. Denote by $r_{i,j}$ the number of common neighbors of v_i

and v_j. Then, our argument shows that $r_{i,i+2} + r_{i,i+3} = k$ for every i, where indices are taken modulo 5. Thus, $r_{i,i+2} = k/2$ for every $1 \leq i \leq 5$.

Observe, that v_1v_2 and v_3v_4 have exactly one common neighbor in $L(X)$. At the same time, for every cycle $u_1u_2u_3u_4$ edges u_1u_2 and u_3u_4 have exactly two common neighbors in $L(X)$. Thus, the pairs (u_1u_2, u_3u_4) and (v_1v_2, v_3v_4) belong to different constituents of the association scheme, say Y_2 and Y_3, respectively. Note, that the triple of edges v_1v_2, v_2v_3, v_3v_4 shows that $p^1_{1,3}$ is non-zero.

Take any $v \in X$ and $u \in N_2(v)$. Suppose that there is no $w \in N_2(v)$ adjacent to u. Then by regularity of X we get $N(v) = N(u)$. For every $x, y \in N(v)$ the triple vx, xu, uy form a triangle with side colors $(1, 1, 2)$ and we get a contradiction with $p^1_{1,3} \neq 0$.

Hence, for every $u \in N_2(v)$ there exists $w \in N_2(v)$ adjacent to u. Take $x \in N(v) \cap N(u)$ and $y \in N(v) \cap N(w)$. Consider the cycle $vxuwy$, then as shown above, vertices v and u have exactly $k/2$ common neighbors. Thus, they are distinguished by at least $|N(u) \triangle N(v)| = 2(k - k/2) = k$ vertices.

Every pair of adjacent vertices has no common neighbors, so they are distinguished by at least $2k$ vertices. Thus, every pair of distinct vertices is distinguished by at least k vertices. Therefore, by Lemma 9.4.10, every pair of distinct vertices is distinguished by at least $n/8$ vertices.

By Lemma 9.4.9, $\mathrm{Aut}(X) \cong \mathrm{Aut}(L(X))$ via natural inclusion ϕ. Thus, bound on the order of $\mathrm{Aut}(L(X))$ follows from Lemma 4.5.5. Let W be the support of $\sigma \in \mathrm{Aut}(X) \cong \mathrm{Aut}(L(X))$. We show that every vertex in W is incident to at most one edge fixed by σ. Consider an edge e with ends w_1, w_2, where $w_1 \in W$. Since $\sigma(w_1) \neq w_1$ the only possibility for e to be fixed is $\sigma(w_1) = w_2$ and $\sigma(w_2) = w_1$. This, in particular implies that $w_2 \in W$ as well. Every edge incident with w_1 and different from e is sent by σ to an edge incident with w_2, so is

not fixed. Therefore, the support of $\phi(\sigma) \in \mathrm{Aut}(L(X))$ is at least

$$\frac{|W|(k-1)}{2} \geq \frac{n}{8} \cdot \frac{(k-1)}{2} \geq \frac{nk}{32} = \frac{|V(L(X))|}{16}. \qquad \square$$

9.5 Putting it all together

Finally, we are ready to combine the preceding results into our main theorem on the motion of rank-4 primitive coherent configurations.

Theorem 9.5.1. *There exists an absolute constant $\gamma_4 > 0$ such that for every primitive coherent configuration $\mathfrak{X}$ of rank 4 on n vertices either*

$$\mathrm{motion}(\mathfrak{X}) \geq \gamma_4 n,$$

or $\mathfrak{X}$ is a Cameron scheme.

Proof. By taking $\gamma_4 < 1/100$ we may assume that $n > 100$.

First, assume that there is an oriented color. Since the rank of $\mathfrak{X}$ is 4, the only possibility is to have two oriented colors $i, j = i^*$ and one undirected color t. It is easy to see that X_t is a strongly regular graph. For $n \geq 29$, by Babai's theorem (Theorem 1.2.5), $\mathrm{motion}(X_t) \geq n/8$, or X_t is a triangular graph $T(s)$, a lattice graph $L_2(s)$, for some s, or their complement.

The constituent X_t cannot be the complement of $L_2(s)$, since the oriented diameter of X_i should be 2, which contradicts $k_i^2 \geq n-1$. Indeed, in this case, $2k_i = k_i + k_i^* = 2(s-1)$, while $n = s^2$.

Now, observe that $p^i_{i,i^*} = p^i_{i^*,i} = p^i_{i,i}$. Moreover, by Eq. (2.3),

$$k_i + k_{i^*} = \left(p^i_{i,i} + p^i_{i,i^*} + p^i_{i,t} + p^i_{i,0}\right) + \left(p^i_{i^*,i} + p^i_{i^*,i^*} + p^i_{i^*,t}\right).$$

Thus, using Eq. (2.3) again, $p^i_{i,i} + p^i_{i^*,i^*} \geq (2k_i - k_t - 1)/3$. If X_t is either $T(s)$ or $L_2(s)$, then

$k_i = k_{i^*} > n/3$ and $k_t < n/3$ for $n > 100$. Thus every pair of vertices connected by an edge of color i is distinguished by at least $k_i/3 \geq n/9$ vertices. Hence, by primitivity of $\mathfrak{X}$ and Lemma 4.5.2, the motion of $\mathfrak{X}$ is at least $n/18$. In the last case, when X_t is a complement of $T(s)$, the result follows from Lemma 9.4.1 and the inequality $p^i_{i,i} + p^i_{i^*,i^*} \geq k_i/3$.

Next, assume that all colors in $\mathfrak{X}$ are undirected, i.e., $\mathfrak{X}$ is an association scheme. Every constituent of $\mathfrak{X}$ has diameter at most 3, as rank of $\mathfrak{X}$ is 4. Moreover, as discussed in Lemma 2.4.3, if there is a constituent of diameter 3, then $\mathfrak{X}$ is induced by a distance-regular graph. In this case the statement follows from Theorem 8.3.1. None of the components can have diameter 1 as the rank is not 2.

Finally, if $\mathfrak{X}$ is an association scheme of rank 4 and diameter 2, then the statement of the theorem follows from Lemma 9.3.1, Theorems 9.3.10, 9.4.2, 9.4.4, 9.4.8 and Theorem 9.4.11, Observation 4.5.10 and Lemma 4.5.11. □

9.6 Open questions

A significant obstacle for our approach, in the case of general primitive coherent configurations of rank $r \geq 5$, is the difficulty of spectral analysis for the constituents of the coherent configuration. Namely, for configurations of rank 4 we analyzed the spectral gap "by hand" through Propositions 9.2.3 and 9.2.4. For coherent configurations of higher rank we need more general techniques.

Problem 9.6.1. *Do there exist $\varepsilon, \delta > 0$ such that the following statement holds? If the minimal distinguishing number $D_{\min}(\mathfrak{X})$ of a primitive coherent configuration $\mathfrak{X}$ satisfies $D_{\min}(\mathfrak{X}) < \varepsilon n$, then the spectral gap for the symmetrization of one of the constituents X_i of $\mathfrak{X}$ is $\geq \delta k_i$. What δ can be achieved?*

We would like to point out, that even δk_i spectral gap for one of the constituents is not sufficient for an efficient application of the spectral tool (Lemma 4.5.11). However, we expect that a result of this flavor would introduce important techniques to the analysis.

We also would like to mention that there should be a reasonable hope to prove Conjecture 1.2.13 for the case when no color is overwhelmingly dominant. The following result easily follows from minimal distinguishing number analysis. In particular, in the case of bounded rank it gives an $\Omega(n)$ bound on the motion.

Proposition 9.6.2. *Fix $0 < \delta < 1$ and an integer $r \geq 3$. Let $\mathfrak{X}$ be a primitive coherent configuration of rank r on n vertices. Assume that each constituent X_i has degree $k_i \leq \delta n$. Then*

$$\text{motion}(\mathfrak{X}) \geq D_{\min}(\mathfrak{X}) \geq \frac{\min(\delta, 1-\delta)}{6(r-1)} n.$$

Proof. The condition $k_i \leq \delta n$, for all i, implies that there exists a set I of colors such that $\sum_{i \in I} k_i = \alpha n$ for some $\min(\delta, 1-\delta)/2 \leq \alpha \leq 1/2$. Fix any vertex u of $\mathfrak{X}$. We want to show that for some vertex v the inequality $D(u,v) \geq \alpha n/3$ holds.

Assume this is not true. Denote $N_I(u) = \{z | \, c(u,z) \in I\}$. Let us count the number of pairs (v,z) with $c(u,z) \in I$ and $c(v,z) \in I$ in two different ways. Since $\sum_{i^* \in I} k_i = \alpha n$ and $z \in N_I(u)$, there are $\alpha^2 n^2$ such pairs. On the other hand, for every v we have $D(u,v) \leq \alpha n/3$, so at least $2\alpha n/3$ vertices $z \in N_I(u)$ are paired with v. Therefore, the number of pairs in question is at least $n \cdot \frac{2\alpha}{3} n$. This contradicts the condition $0 < \alpha \leq \frac{1}{2}$.

Therefore, there exists a pair of vertices with $D(u,v) \geq \alpha n/3$. Finally, the configuration $\mathfrak{X}$ is primitive, so by Lemma 4.5.2 we get that $\text{motion}(\mathfrak{X}) \geq D_{\min}(\mathfrak{X}) \geq \frac{\alpha}{3(r-1)} n$. □

However, when the rank is unbounded, this seemingly simple case of Conjecture 1.2.13 (every constituent has degree $\leq \delta n$) is still open. To avoid exceptions we relax the conjectured lower bound to $\Omega(n/\log(n))$.

Conjecture 9.6.3. *Fix $0 < \delta < 1$. Let $\mathfrak{X}$ be a primitive coherent configuration on n vertices. Assume that every constituent has degree $\leq \delta n$. Then* $\text{motion}(\mathfrak{X}) = \Omega(n/\log(n))$.

Next, we observe that Cameron schemes satisfy Conjecture 9.6.3.

Proposition 9.6.4. *Fix $0 < \delta < 1$. Consider a Cameron group $(A_m^{(k)})^d \leq G \leq S_m \wr S_d$ acting on $n = \binom{m}{k}^d$ points and let $\mathfrak{X} = \mathfrak{X}(G)$ be the corresponding Cameron scheme. Assume that every constituent of $\mathfrak{X}$ has degree $\leq \delta n$. Then* $\mathrm{motion}(\mathfrak{X}) = \Omega(n/\log(n))$.

Proof. We can assume $k \leq m/2$. Note that then the rank of $\mathfrak{X}$ is equal to $kd + 1$.

Case 1. Suppose that $k \leq m/3$. Then

$$n = \binom{m}{k}^d \geq \left(\frac{m-k}{k}\right)^{kd} \geq \left(\frac{2k}{k}\right)^{kd} = 2^{kd}$$

Thus, $kd \leq \log(n)$ in this case, and the statement follows from Proposition 9.6.2.

Case 2. Suppose that $m/3 < k \leq m/2$. By Lemma 4.2.3, we have that as $m \to \infty$ the inequality $\mathrm{motion}(\mathfrak{X}) \geq \alpha n$ holds for some $\alpha > 0$. At the same time, by the proof of Lemma 4.2.3 we know that the motion of $\mathfrak{X}$ does not depend on d. Thus as $\mathrm{motion}(X) \geq \alpha n$ is violated just by finite number of pairs (m, k), we still have $\mathrm{motion}(\mathfrak{X}) = \Omega(n)$ in this case. □

We observe that the bound in Conjecture 9.6.3, if true, is nearly tight, for $\delta \in (1/e, 1)$, as the example of Hamming schemes $\mathfrak{H}(tm, m)$ with $t = -\lfloor \log(\delta)m \rfloor / m$ shows. Note that for $m \geq 3$ the Hamming scheme $\mathfrak{H}(k, m)$ is primitive.

Proposition 9.6.5. *Consider Hamming scheme $\mathfrak{H}(tm, m)$ with $t = -\dfrac{\lfloor \log(\delta)m \rfloor}{m}$ on $n = m^{tm}$ points, for $\delta \in (1/e, 1)$. Then its maximum constituent degree satisfies $k_{\max} \leq \delta n$ and the motion satisfies*

$$\mathrm{motion}(\mathfrak{H}(tm, m)) = O\left(\frac{n \log\log(n)}{\log(n)}\right).$$

Proof. Note that since $tm < m$ the maximum degree is $k_{\max} = (m-1)^{tm}$. Then

$$k_{\max} = \left(\frac{m-1}{m}\right)^{mt} n \leq \mathrm{e}^{-t} n \leq \delta n.$$

The motion of $\mathfrak{H}(tm, m)$ is realized by a 2-cycle in the first coordinate, and is equal to

$2n/m$. The number of vertices is $n = m^{mt} = \mathrm{e}^{tm\log(m)}$, so $m\log(m) = \log(n)/t$. Thus $m > \log(n)/(t\log\log(n))$. Hence,

$$\mathrm{motion}(\mathfrak{H}(tm, m)) \leq \frac{2n\log\log(n)t}{\log(n)} = O\left(\frac{n\log\log(n)}{\log(n)}\right).$$

$\square$

CHAPTER 10

ROBUSTNESS UNDER EXTENSION

10.1 Introduction

The material of this Chapter (except for Sections 10.5 and 10.6) is a result of a joint work by Babai and Kivva [2022].

Throughout this chapter we use the following notation.

We use Ω and Ω' to denote sets of vertices of configurations and we always assume $\Omega \subseteq \Omega'$. We denote $n = |\Omega|$, $n' = |\Omega'|$. If $\mathfrak{X}$ is a regular configuration on Ω, we use k_i to denote the degree of the i-th constituent X_i. If $\mathfrak{X}$ is a coherent configuration, $p^t_{i,j}$ stands for an intersection number. Whenever we have a configuration on Ω' we add $'$ to our standard notation.

We remind that if the graph is regular, we use k to denote its degree. If the graph is edge-regular, λ stands for the number of the common neighbors of a pair of adjacent vertices. If the graph is co-edge-regular, μ denotes the number of the common neighbors of a pair of non-adjacent vertices.

For a metric scheme $\mathfrak{X}$, we use k, λ, μ and other notation introduced for distance-regular graphs to refer to the corresponding parameters of the color-1 constituent of $\mathfrak{X}$.

Definition 10.1.1. We say that a metric scheme is *geometric*, if its underlying distance-regular graph (color-1 constituent) is geometric.

In this chapter we study the following version of Question 1.3.3.

Question 10.1.2. *Assume that $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ is a homogeneous coherent configuration and $\mathfrak{X} = (\Omega, \mathfrak{c})$ is its coherent subconfiguration with $|\Omega'| \leq (1+\alpha)|\Omega|$ for some $\alpha > 0$. Assume that $\mathfrak{X}$ belongs to a certain "nice" class of configurations $\mathcal{A}$.*

For which $\alpha > 0$, can we deduce that $\mathfrak{X}'$ also belongs to $\mathcal{A}$?

Definition 10.1.3. We say that the property $\mathcal{A}$ is *robust under extension with parameter* $\alpha > 0$ if in Question 10.1.2 for this α we can deduce that $\mathfrak{X}' \in \mathcal{A}$.

In this chapter we show that the following classes are robust under extension:

(i) for every $r \geq 2$, $\mathcal{A} = \{\mathfrak{X} \mid \mathfrak{X} \text{ is of rank } r\}$ for $\alpha < 1$.

(ii) $\mathcal{A} = \{\mathfrak{X} \mid \mathfrak{X} \text{ is symmetric}\}$ for $\alpha < 1$.

(iii) $\mathcal{A} = \{\mathfrak{X} \mid \mathfrak{X} \text{ is primitive}\}$ for $\alpha < 1$.

(iv) $\mathcal{A} = \{\mathfrak{X} \mid \mathfrak{X} \text{ is metric}\}$ for $\alpha < 1/2$.

(v) $\mathcal{A} = \{\mathfrak{X} \mid \mathfrak{X} \text{ is geometric, } k \geq (5/2)(\lambda + 1), \text{ and } k \geq 100\mu|\theta_{\min}|^3\}$ for $\alpha < 1/2$.

(vi) $\mathcal{A} = \{\mathfrak{X} \mid \mathfrak{X} \text{ is the Johnson scheme, } \mathfrak{J}(s, d) \text{ with } s \geq 300d^2\}$ for $\alpha < 1/2$.

(vii) $\mathcal{A} = \{\mathfrak{X} \mid \mathfrak{X} \text{ is the Hamming scheme, } \mathfrak{H}(d, s) \text{ with } s \geq 200d^4 \ln d\}$ for $\alpha < 1/2$.

(viii) $\mathcal{A} = \{\mathfrak{X} \mid \mathfrak{X} \text{ is the Grassmann scheme, } \mathfrak{J}_q(s, d) \text{ with } s \geq 3d + 7,\ d \geq 3\}$ for $\alpha < 1/2$.

10.2 Basic properties preserved under extension

Lemma 10.2.1. *Let $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ be a homogeneous coherent configuration. Let $\Omega \subseteq \Omega'$ be a subset of size $|\Omega| > |\Omega'|/2$ and assume that the subconfiguration $\mathfrak{X} = \mathfrak{X}'[\Omega] = (\Omega, \mathfrak{c})$ is coherent. Then $\mathrm{rk}(\mathfrak{X}) = \mathrm{rk}(\mathfrak{X}')$. In particular, if $\mathfrak{X}$ is an association scheme then so is $\mathfrak{X}'$. Moreover, if $\mathfrak{X}$ is primitive then so is $\mathfrak{X}'$.*

Proof. Let $|\Omega'| = n'$ and $|\Omega| = n$. If $\mathrm{rk}(\mathfrak{X}) < \mathrm{rk}(\mathfrak{X}')$ then there is a color $i \in \mathrm{Range}(\mathfrak{c}') \setminus \mathrm{Range}(\mathfrak{c})$, and so $i^* \in \mathrm{Range}(\mathfrak{c}') \setminus \mathrm{Range}(\mathfrak{c})$ as well (see Def. 2.3.1). Denote $\widetilde{R'_i} = R'_i \cup (R'_{i^*})$. Let d be the degree of the regular graph $(\Omega', \widetilde{R'_i})$. Then Ω emits dn edges from $\widetilde{R'_i}$ and $\Omega' \setminus \Omega$ absorbs at most $(n' - n)d$ of these edges, so $n \leq n' - n$, contrary the assumption that $n > n'/2$. This proves the first statement.

Now assume $\mathfrak{X}$ is an association scheme, i.e., every constituent of $\mathfrak{X}$ is undirected. But, by the first statement, every constitutent of $\mathfrak{X}'$ has an edge in $\mathfrak{X}$; therefore every constituent of $\mathfrak{X}$ is also undirected.

Now let $X'_j = (\Omega', R'_j)$ be an off-diagonal constituent of $\mathfrak{X}'$. If X'_j were disconnected, its connected components (being of equal size) would have size $\leq n'/2$. But Ω induces a connected subdigraph of X'_j of order $n > n'/2$, a contradiction proving the last statement.

□

Hence, if $\mathfrak{X}$ is an association scheme and $|\Omega| > |\Omega'|/2$, we will assume that $\mathfrak{X}'$ is an association scheme.

Lemma 10.2.2. *Assume that both $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ and $\mathfrak{X} = \mathfrak{X}'[\Omega]$ are regular configurations (see Def. 2.3.5). Let k_i and k'_i be the degrees of color i in $\mathfrak{X}$ and $\mathfrak{X}'$, respectively. Then*

$$k'_i(2n - n') \leq nk_i.$$

Proof. Every vertex $x \in \Omega$ is incident with $(k'_i - k_i)$ edges of color i with the second endpoint within $\Omega' \setminus \Omega$. At the same time, each vertex in $\Omega' \setminus \Omega$ is incident with only k'_i edges. Hence,

$$(k'_i - k_i)n \leq (n' - n)k'_i, \quad \text{so}$$

$$k'_i(2n - n') \leq nk_i.$$

□

Lemma 10.2.3. *Let $\Omega \subseteq \Omega'$. Assume that $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ and $\mathfrak{X} = \mathfrak{X}'[\Omega]$ are association schemes. Suppose that $|\Omega'| < 3|\Omega|/2$ and $\mathfrak{X}$ is a metric scheme. Then $\mathfrak{X}'$ is metric as well of rank $\mathrm{rk}(\mathfrak{X}) = \mathrm{rk}(\mathfrak{X}')$. Moreover, if the color-1 constituent X_1 of $\mathfrak{X}$ is a distance-regular graph, then X'_1 is distance-regular too.*

Proof. We prove that if in $\mathfrak{X}'$ there exists a triangle with sides of colors (i, j, t), then a triangle with sides of the same colors exists in $\mathfrak{X}$. Since $\mathfrak{X}'$ is a coherent configuration there exists a

constant $C_v = p^t_{i,j}$ such that for a given pair (u, v) of color t there exists precisely C_v vertices w such that $\mathfrak{c}'(u, w) = i$ and $\mathfrak{c}'(w, v) = j$. Similarly, there exists a constant C_e, such that for every vertex w there exists precisely C_e pairs (u, v) such that $\mathfrak{c}'(u, v) = t$, $\mathfrak{c}'(u, w) = i$ and $\mathfrak{c}'(w, v) = j$. Trivial double counting yields

$$n' k'_t C_v = n' C_e \qquad \Rightarrow \qquad k'_t C_v = C_e.$$

For every pair $u, v \in \Omega$ there are C_v verices w, such that (v, u, w) is a triangle with side colors (i, j, t). Assume that $\mathfrak{X}$ does not contain a triangle with sides of colors (i, j, t). Then $w \in \Omega' \setminus \Omega$. Therefore, the total number of such triangles with $u, v \in \Omega$ is at most $(n' - n)C_e$. There are nk_t pairs $(u, v) \in \Omega \times \Omega$ with $\mathfrak{c}(u, v) = t$. Hence

$$nk_t \frac{C_e}{k'_t} = nk_t C_v \leq (n' - n)C_e.$$

At the same time, Lemma 10.2.2 gives $k'_t(2n - n') \leq nk_t$. Thus,

$$k'_t(2n - n')/k'_t \leq n' - n \qquad \Rightarrow \qquad 3n \leq 2n'.$$

We get a contradiction with the assumption $n' < 3n/2$. Therefore, if $\mathfrak{X}'$ contains a triangle with sides of colors (i, j, t), then $\mathfrak{X}$ contains such triangle as well.

Now, assume that Ω is a metric scheme. Then by Lemma 2.4.3, there exist colors i and j such that $\mathrm{dist}_i(j) = r - 1$. What we proved above shows that dist_i in $\mathfrak{X}'$ does not decrease, as otherwise in $\mathfrak{X}'$ there is a triangle, that is not in $\mathfrak{X}$. Moreover, as shown in Lemma 10.2.1, $\mathrm{rk}(\mathfrak{X}) = \mathrm{rk}(\mathfrak{X}') = r$, thus, by Lemma 2.4.3, $\mathfrak{X}'$ is metric as well. $\square$

10.3 Extension of geometric metric schemes

Recall that for a metric scheme $\mathfrak{X}$, we use k, λ, μ and other notation introduced for distance-regular graphs to refer to the corresponding parameters of the color-1 constituent of $\mathfrak{X}$.

Lemma 10.3.1. *Let $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ be an association scheme. Let $\Omega \subseteq \Omega'$ with $|\Omega'| = |\Omega|(1+\alpha)$ for $\alpha < 1/2$. Assume that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is a metric scheme and $X = X_1$ be the underlying distance-regular graph. Assume that $(5/2)(\lambda+1) \leq k$ in X. Then*

$$\mu \leq \mu' \leq \frac{\mu}{(1-\alpha)(1-5\alpha/3)} < 12\mu.$$

(Note that, by Lemma 10.2.3, X' is distance-regular, so μ' is well-defined).

Proof. By Lemma 10.2.3, $\mathfrak{X}'$ is a metric scheme as well. Let X' be the corresponding underlying distance-regular graph of $\mathfrak{X}'$, that is, the constituent of the same color, as X. Recall that we add $'$ to denote the parameters of X'.

Clearly, $\lambda' \geq \lambda$ and Lemma 10.2.2 gives $(1-\alpha)k' \leq k$. Denote, $b_1 = k - (\lambda+1)$, and

$$b_1' := k' - (\lambda'+1) \leq k' - (\lambda+1).$$

Then,

$$(1-5\alpha/3)b_1' \leq (1-5\alpha/3)(k' - (\lambda+1)) =$$
$$= (1-\alpha)k' - (\lambda+1) - (2\alpha/3)(k' - (5/2)(\lambda+1)) \leq k - (\lambda+1) = b_1.$$

Note, that

$$\frac{k'b_1'}{\mu'} = k_2' \geq k_2 = \frac{kb_1}{\mu}.$$

This implies

$$\frac{k'b_1'}{\mu'} \geq \frac{(1-\alpha)(1-5\alpha/3)k'b_1'}{\mu}.$$

Therefore, $\mu \leq \mu' \leq \dfrac{\mu}{(1-\alpha)(1-5\alpha/3)} < 12\mu.$ $\square$

Remark 10.3.2. The Johnson scheme $\mathfrak{J}(s,d)$ with $d=3$ and $s \geq 23$, or $d \geq 4$ and $s \geq 12$ satisfies $(5/2)(\lambda+1) \leq k$.

Theorem 10.3.3. *Let $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ be an association scheme. Let $\Omega \subseteq \Omega'$ with $|\Omega'| = |\Omega|(1+\alpha)$ for $0 < \alpha < 1/2$. Assume that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is a geometric metric scheme of diameter $d \geq 2$ with smallest eigenvalue $-m$. Assume additionally that*

$$k \geq \left(\frac{\mu'-1}{2}\right) m \cdot \left\lceil \frac{m}{1-\alpha} \right\rceil \left(\left\lceil \frac{m}{1-\alpha} \right\rceil + 1 \right) + m. \tag{10.1}$$

Then $\mathfrak{X}'$ is a geometric metric scheme with smallest eigenvalue $\geq -\lceil m/(1-\alpha) \rceil$.

Proof. Since $\mathfrak{X}$ is a metric scheme and $\alpha < 1/2$, by Lemma 10.2.3, $\mathfrak{X}'$ is also a metric scheme. Let X and X' be color-1 constituents of $\mathfrak{X}$ and $\mathfrak{X}'$, respectively. We need to prove that X' is geometric.

By Corollary 3.1.6, it suffices to show that there exists an integer m' such that

$$m'(\lambda'+1) \geq k' \quad \text{and} \quad \lambda' \geq \left(\frac{\mu'-1}{2}\right) m'(m'+1).$$

Recall that $\lambda' \geq \lambda$, $\lambda \geq k/m - 1$, and so by Lemma 10.2.2,

$$k' \leq \frac{1}{1-\alpha} k = \frac{1}{1-\alpha} m(\lambda+1) \leq \frac{1}{1-\alpha} m(\lambda'+1).$$

Thus, it is sufficient to have

$$m' \geq \left(\frac{1}{1-\alpha}\right) m \quad \text{and} \quad k \geq \left(\frac{\mu'-1}{2}\right) m \cdot m'(m'+1) + m.$$

Note that $m' = \lceil m/(1-\alpha) \rceil$ satisfies both the inequalities above, since Eq.(10.1) holds. $\square$

Remark 10.3.4. Eq. (10.1) is satisfied if $k \geq (5/2)(\lambda+1)$ and $k \geq \dfrac{3\mu m^3}{2(1-\alpha)^3(1-5\alpha/3)}$.

Proof. Observe that for $m \geq 2$ we have $m^3/4 \geq m$ and $(m+1)(m+2) \leq 3m^2$. Thus

$$\left(\frac{\mu'-1}{2}\right) m \cdot \left\lceil \frac{m}{1-\alpha} \right\rceil \left(\left\lceil \frac{m}{1-\alpha} \right\rceil + 1 \right) + m \leq \left(\frac{\mu'-1}{2}\right) \cdot \frac{3m^3}{(1-\alpha)^2} + m^3/4 \leq \frac{3\mu' m^3}{2(1-\alpha)^2}.$$

Therefore, it is sufficient to have $k \geq \dfrac{3\mu' m^3}{2(1-\alpha)^2}$. Finally, we can apply Lemma 10.3.1. □

Lemma 10.3.5. *Let $X = X'[\Omega]$ be an induced subgraph of a geometric distance-regular graph X' with smallest eigenvalue $-m'$. Assume that X is a geometric distance-regular graph with smallest eigenvalue $-m$. If $n' \leq (1+\alpha)n$ for $\alpha < 1/2$ and $k \geq 2m(m')^2$, then $m = m'$ and every Delsarte clique of X is a subclique of a Delsarte clique of X'.*

Proof. Fix Delsarte clique geometries $\mathcal{C}$ and $\mathcal{C}'$ in X and X', respectively. Let $C' \in \mathcal{C}'$ and $C \in \mathcal{C}$ be cliques that have a common edge. Then

$$\lambda' \geq |C'| + |C| - |C \cap C'| - 2 = k'/m' + k/m - |C \cap C'|. \tag{10.2}$$

If C is not a subset of C', then $\mu' \geq |C \cap C'|$. Recall that,

$$\lambda' = \frac{k'}{m'} + (m'-1)(\psi_1' - 1) - 1.$$

Combining this with Eq. (10.2), we get

$$\mu' + (m'-1)(\psi_1' - 1) - 1 \geq k/m.$$

By Lemma 3.1.9, $\mu' \leq m'^2$ and $\psi_1' \leq \tau_2' \leq m'$. Hence, we get a contradiction with $k \geq 2m(m')^2$. So $C \subseteq C'$.

Now, assume that $m' > m$, then for every vertex $x \in \Omega = V(X)$ there is a Delsarte clique $C' \in \mathcal{C}'$ such that $(C' \setminus \{x\}) \subseteq (\Omega' \setminus \Omega)$. Moreover, since $\mathcal{C}'$ is a clique geometry, by definition, every edge of X' belongs to precisely one clique in $\mathcal{C}'$. All cliques in the Delsarte

clique geometry $\mathcal{C}'$ are of the same size. Hence there are at least

$$n \cdot (|C'| - 1)(|C'| - 2)/2 \geq n\frac{(k')^2}{4(m')^2}$$

edges between vertices in $(\Omega' \setminus \Omega)$. In total, there are $(n'k'/2 - nk/2 - n(k' - k))$ edges between vertices in $(\Omega' \setminus \Omega)$. Hence,

$$n'k'/2 - nk' + nk/2 \geq n\frac{(k')^2}{4(m')^2} \quad \Rightarrow \quad (1+\alpha) - 2 + 1 \geq \frac{2k'}{4(m')^2},$$

which gives a contradiction with $k \geq 2m(m')^2$ and $\alpha < 1/2$. Therefore, $m' = m$. □

Corollary 10.3.6. *Let $\mathfrak{X}$ and $\mathfrak{X}'$ be association schemes such that $\mathfrak{X} = \mathfrak{X}'[\Omega]$. Assume X_1' and X_1 are geometric distance-regular graph with smallest eigenvalue $-m'$, and $-m$, respectively. If $n' \leq (1+\alpha)n$ for $\alpha < 1/2$ and $k \geq 2m(m')^2$, then*

$$\tau_i = \tau_i' \quad \textit{and} \quad \psi_i \leq \psi_i' \quad \textit{for all} \quad i \in [d-1].$$

Proof. Let $\mathcal{C}$ and $\mathcal{C}'$ be Delsarte clique geometries of X_1 and X_1', respectively.

Let $C \in \mathcal{C}$ be a Delsarte clique of X_1 and $v \in V(X_1)$ be a vertex at distance $i \in [d-1]$ from C. Since $i \leq d-1$ there exists $w \in C$ with $\mathrm{dist}(v, w) = i+1$ in X_1. By Lemma 10.3.5, there exists a Delsarte clique in X_1' such that $C \subseteq C' \in \mathcal{C}'$.

Let $T \subset C$ be the set of vertices at distance i from v. By Lemmas 10.2.1 and 10.2.3, they are still at distance i from v in X_1'. Additionally, by Lemmas 10.2.1 and 10.2.3, $\mathrm{dist}(v, w) = i+1$ in X_1', so C' is at distance i from C' in X_1'. Therefore, $\psi_i \leq \psi_i'$.

Let $\mathrm{dist}(u, v) = i$ in X_1. Now, let $C_1, C_2, \ldots, C_{\tau_i} \in \mathcal{C}$ be Delsarte cliques in X_1 which contain u and which are at distance $i-1$ from v. By Lemma 10.3.5, there exist Delsarte cliques $C_1', C_2', \ldots, C_{\tau_i}' \in \mathcal{C}'$ in X_1' such that $C_j \subseteq C_j'$ for every $j \in [\tau_i]$. Moreover, since every C_j is a maximal clique in X_1, all C_k' are distinct for $k \in [\tau_i]$. An argument as above

shows that every C'_j is at distance $i-1$ from v. Therefore, $\tau_i \le \tau'_i$.

At the same time, if $C \in \mathcal{C}$ is a Delsarte clique in X_1 which is not at distance $i-1$ from v with $u \in C$, then C is at distance i from u. By Lemma 10.3.5, there exists a Delsarte clique $C' \in \mathcal{C}'$ in X'_1 such that $C \subseteq C'$. If $i < d$, there is a vertex $w \in C$ with $\mathrm{dist}(v,w) = i+1$ in X_1. By Lemmas 10.2.1 and 10.2.3, $\mathrm{dist}(v,w) = i+1$ in X'_1, so C' does not contain vertices at distance $i-1$ from v. Thus, $\tau_i = \tau'_i$. □

Corollary 10.3.6 shows that $\psi_i \le \psi'_i$. Next we show how to upper bound ψ'_1 in terms of ψ_1.

Lemma 10.3.7. *Let $\mathfrak{X}$ and $\mathfrak{X}'$ be association schemes such that $\mathfrak{X} = \mathfrak{X}'[\Omega]$. Assume X'_1 and X_1 are geometric distance-regular graph with smallest eigenvalue $-m'$, and $-m$, respectively. If $n' < (1+\alpha)n$ for $\alpha < 1/2$ and $k \ge 2m(m')^2$, then $\psi'_1 \le 1 + \left(\frac{1-\alpha}{1-2\alpha}\right)(\psi_1 - 1)$.*

Proof. Let C_1 and C_2 be Delsarte cliques of X containing a vertex v. Let C'_1 and C'_2 be the corresponding cliques of X' that contain C_1 and C_2, respectively, guaranteed by Corollary 10.3.6. Then

$$|C_1| = |C_2| = k/m + 1, \qquad |C'_1| = |C'_2| = k'/m + 1, \quad \text{and} \quad (1-\alpha)k' \le k.$$

For vertex w distinct from v in C_1 there is a vertex $u \in C_2$ non-adjacent with it, since C_2 is a maximal clique. Clearly $\mathrm{dist}(u,w) = 2$. Thus, there are exactly $\psi_1 - 1$ vertices in $C_2 \setminus \{v\}$ adjacent with w in X. Similarly, every vertex in $C_1 \subseteq C'_1$ is adjacent with exactly $\psi'_1 - 1$ vertices in $C'_2 \setminus \{v\}$ in X'_1. Therefore,

$$\begin{aligned} (\psi'_1 - 1)\cdot(k/m) \le (\psi_1 - 1)\cdot(k/m) + (\psi'_1 - 1)\cdot(k'-k)/m \quad &\Leftrightarrow \\ (\psi'_1 - 1)(2k - k') \le (\psi_1 - 1)k, \qquad \text{so} \qquad \psi'_1 \le 1 + \left(\frac{1-\alpha}{1-2\alpha}\right)(\psi_1 - 1). \quad \square \end{aligned} \tag{10.3}$$

Corollary 10.3.8. *If the assumptions of Lemma 10.3.7 hold and $\psi_1 = 1$, then $\psi'_1 = 1$. In this case also $\mu = \mu'$.*

Proof. The first part follows from Lemma 10.3.7. By Corollary 10.3.6, $\tau_2 = \tau'_2$, so $\mu = \mu'$. □

10.4 Robustness of the Johnson schemes under extension

Theorem 10.4.1. *Let $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ be an association scheme. Let $\Omega \subseteq \Omega'$ with $|\Omega'| \leq |\Omega|(1+\alpha)$, for $0 < \alpha < 1/2$. Assume that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is the Johnson scheme $\mathfrak{J}(s,d)$, with $d \geq 3$ and $k \geq \dfrac{6d^3}{(1-\alpha)^3(1-5\alpha/3)}$. Then $\mathfrak{X}'$ is a Johnson scheme as well.*

Proof. We want to apply the classification result stated in Theorem 3.4.2.

To do this, we first verify that X' is geometric. Recall, that in the Johnson scheme $J(s,d)$ we have $\mu = 4$, and by Remark 10.3.2, $(5/2)(\lambda+1) \leq k$ for $d \geq 3$ and $s \geq 23$. The inequality $s \geq 23$ follows from the assumptions of the theorem, as $s \geq k/d \geq 6d^2 \geq 54$. Then, by Theorem 10.3.3 and Remark 10.3.4, X' is geometric.

The smallest eigenvalue of the Johnson graph $J(s,d)$ is $-m = -d$.

Note that by Theorem 10.3.3, the smallest eigenvalue of X' is at least $-m'$ for an integer $m' \leq m/(1-\alpha) + 1$. In particular, $(m')^2 < 3m^2/(1-\alpha)^2$. Thus, $k > 2m(m')^2$.

By Corollary 10.3.6, $\tau'_2 = \tau_2 = 2$. Thus, by Lemma 3.2.2, $4 = \mu \leq \mu' \leq \tau_2^2 = 4$. Hence, $\psi'_1 = 2$. Therefore, the assumptions of Theorem 3.4.2 hold. □

Remark 10.4.2. The inequality $k \geq \dfrac{6d^3}{(1-\alpha)^3(1-5\alpha/3)}$ holds for the Johnson scheme $\mathfrak{J}(s,d)$ when $s \geq \dfrac{6d^2}{(1-\alpha)^3(1-5\alpha/3)} + d$. In particular, for $\alpha = 1/2$, it holds if $s \geq 288d^2 + d$; and for $\alpha = 1/3$ and $d \geq 3$, it holds if $s \geq 46d^2$.

In the case of strongly regular graphs, i.e., primitive coherent configurations of rank 3, we prove that a weaker assumption on the constant c is sufficient.

Theorem 10.4.3. *If X' is a SRG with $n' < v(v-2)$ vertices for some v and $n' \geq 29$ and X' has an induced subgraph X which is a $J(v,2)$ then X' is a $J(v',2)$ for some v'.*

Proof. The parameters of X are $n = \binom{v}{2}$ vertices, degree $k = 2(v-2)$, $\lambda = v-2$, and $\mu = 4$. The eigenvalues of X are k, $r = v-4$, and $s = -2$, with respective multiplicities $m_0 = 1$, $m_1 = v-1$, and $m_2 = \binom{v}{2} - v = v(v-3)/2$. Let $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_n$ denote the eigenvalues of X and $\theta'_1 \geq \theta'_2 \geq \cdots \geq \theta'_{n'}$ the eigenvalues of X'. Then by interlacing we have, for every $0 \leq t \leq n-1$,

$$\theta_{n'-t} \leq \theta'_{n'-t} \leq \theta_{n-t} \tag{10.4}$$

Setting $t = n-v-1$ we have $\theta_{n-t} = \theta_{v+1} = -2$. We also have $n-t \leq n'-t \leq n$ and therefore $\theta_{n'-t} = -2$. It follows by Eq. (10.4) that $\theta'_{n'-t} = -2$, so -2 is one of the eigenvalues of X'. By interlacing, we have $\theta'_2 \geq \theta_2 = v-4 > 0$, so -2 is the smallest eigenvalue of X'. Therefore, by Seidel, X' is a SR line graph and therefore a Johnson graph $J(v',2)$ or a Hamming graph $H(2,v')$ for some v'. But for $H(2,v')$, the number of common neighbors of a non-adjacent pair of vertices is $\mu' = 2 < \mu$, which is impossible for a supergraph of X. So X' is a Johnson graph. □

Corollary 10.4.4. *Question 10.1.2 has positive answer for $\alpha \in (0,1)$ when $\mathfrak{X} = \mathfrak{J}(v,2)$ and $v \geq \max\left(\frac{2}{1-\alpha}, 8\right)$.*

Proof. By Lemma 10.2.1, one of the constituents of $\mathfrak{X}'$ is a (primitive) SRG X' of which the Johnson graph $X = J(v,2)$ is an induced subgraph for some $v \geq 8$. Note that for $v \geq 8$ we have $n \geq 28$. Then the claim follows from Theorem 10.4.3 as

$$v(v-2) > \binom{v}{2}(1+c) \quad \Leftrightarrow \quad 2(v-2) > (v-1)(1+c) \quad \Leftrightarrow \quad (1-c)v > 3-c. \qquad \square$$

Corollary 10.4.5. *Question 10.1.2 has positive answer for $\alpha = 1/2$ when $\mathfrak{X} = \mathfrak{J}(v,2)$ and $v \geq 8$.*

Proof of part 1 of Theorem 1.3.6. follows from Theorem 10.4.1, Remark 10.4.2 and Corollary 10.4.5. □

10.5 Robustness of the Hamming schemes under extension

Lemma 10.5.1. *Let $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ be an association scheme. Let $0 < \alpha < 1/2$ and $\Omega \subseteq \Omega'$ with $|\Omega'| \leq (1+\alpha)|\Omega|$. Assume that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is a Hamming scheme $\mathfrak{H}(d,s)$. Then, for all $i \in [d]$,*

$$c'_i \leq c_i \cdot \frac{1}{(1-\alpha)^2}\left(1 + \alpha\frac{i-1}{d-i+1}\right).$$

In particular,

$$\mu' \leq 2(1+\alpha)(1-\alpha)^{-2} < 12 \quad \textit{and} \quad c'_d \leq d + 3\alpha d^2 \leq d + 3d^2/2.$$

Proof.

$$k' = a'_i + b'_i + c'_i \geq b'_i + c_i + a_i = b'_i + k - b_i \tag{10.5}$$

By Lemma 10.2.2, $(1-\alpha)k' \leq k$ and for the Hamming graph $b_i = \dfrac{d-i}{d}k$, so

$$b'_i \leq b_i + (k' - k) \leq b_i + \frac{\alpha}{1-\alpha}\left(\frac{d}{d-i}\right)b_i = b_i\left(\frac{(d-i)+\alpha i}{(1-\alpha)(d-i)}\right) \tag{10.6}$$

Next we deduce

$$c'_{i+1} = \frac{b'_i k'_i}{k'_{i+1}} \leq \frac{b_i k_i}{k_{i+1}} \cdot \frac{(d-i+\alpha i)}{(1-\alpha)(d-i)} \cdot \frac{1}{(1-\alpha)} = c_{i+1} \cdot \frac{1}{(1-\alpha)^2}\left(1 + \alpha\frac{i}{d-i}\right) \tag{10.7}$$

Finally, we can deduce $c'_d \leq (1 + (d+3)\alpha)c_d \leq d + 3\alpha d^2$. □

Now we are ready to prove the extension theorem for Hamming schemes.

Theorem 10.5.2. *Let $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ be an association scheme. Let $\Omega \subseteq \Omega'$ with $|\Omega'| \leq |\Omega|(1+\alpha)$ for $0 < \alpha < 1/2$. Assume that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is a Hamming scheme $\mathfrak{H}(d,s)$ and*

$$k \geq 100d^3\left(3d^2/2 + d\right)\ln d.$$

Then $\mathfrak{X}'$ is a Hamming scheme itself.

Proof. By Lemma 10.5.1, $\mu' \leq 11$. The Hamming graph $H(d,s)$ is geometric with smallest eigenvalue $-m = -d$, and $k \geq 10m^2(2m+1)+m$, so by Theorem 10.3.3, the scheme $\mathfrak{X}'$ is geometric metric scheme with smallest eigenvalue $-m' \geq -2d$.

Since $k \geq 8d^3$, by Lemma 10.3.5 and Corollary 10.3.6, for $\mathfrak{X}'$ we get $m' = d$, $\tau'_i = \tau_i = i$ for $i \leq d-1$. Moreover, by Corollary 10.3.8, we get that $\mu' = 2$.

By Lemma 10.5.1, $c'_d \leq d + (3/2)d^2$, and by assumptions of the theorem,

$$k' \geq k \geq \Big(100d^3 \ln d\Big) \cdot c'_d.$$

Therefore, since $\mu' = 2$, by Theorem 5.2.2, $\mathfrak{X}'$ is a Hamming scheme. □

Remark 10.5.3. In the Hamming graph $H(d,s)$ we have $k = d(s-1)$, so the desired inequlity on k is satisfied if $s \geq 100d^2\left(3d^2/2 + d\right)\ln d + 1 = \Omega(d^4 \ln d)$.

Hence, we can state the extension theorem for Hamming schemes in the following form (confirming part 2 of Theorem 1.3.6).

Theorem 10.5.4. *Let $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ be an association scheme. Let $\Omega \subseteq \Omega'$ with $|\Omega'| < (3/2)|\Omega|$. Assume that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is the Hamming scheme $\mathfrak{H}(d,s)$ with $s \geq 200d^4 \ln d$.*

Then $\mathfrak{X}'$ is the Hamming scheme $\mathfrak{H}(d,s')$ for some $s' \geq s$.

Proof. Immediately follows from Theorem 10.5.2 and Remark 10.5.3. □

10.6 Robustness of the Grassmann schemes under extension

We rely on the characterization of Grassmann graphs by Ray-Chaudhuri and Sprague [1976] stated in a weaker form in Theorem 3.6.2.

Theorem 10.6.1. *Let $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ be an association scheme. Let $\Omega \subseteq \Omega'$ with $|\Omega'| \leq |\Omega|(1+\alpha)$ for $0 < \alpha < 1/2$. Assume that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is a Grassmann scheme $\mathfrak{J}_q(s,d)$ with $d \geq 3$ and*

$$k \geq 26\mu m^3. \tag{10.8}$$

Then $\mathfrak{X}'$ is a Grassmann scheme $\mathfrak{J}_q(s',d)$ for some $s' \geq s$.

Proof. By Lemma 10.2.3, $\mathfrak{X}'$ is a metric scheme.

In a Grassmann scheme $\mathfrak{J}_q(s,d)$ we have $k = q[d]_q[s-d]_q$ and $\lambda = q[s-d]_q + q[d]_q + q$. Since $[s-d]_q \geq [d]_q > 4$ we have $k \geq (5/2)(\lambda+1)$. Hence, by Lemma 10.3.1, $\mu' \leq 12\mu$.

For $d \geq 3$, we have $m \geq q^2+q+1 \geq 7$, so $26m^3 \geq 24m^3 + 12m^2$

$$k \geq 26\mu m^3 \geq (12\mu - 1)m^2(2m+1) + m. \tag{10.9}$$

Since X_1 is a geometric distance-regular graph and Eq. (10.9) holds, by Theorem 10.3.3, X_1' is also a geometric distance-regular graph with smallest eigenvalue $-m' \geq -2m$. By Lemma 10.3.5 and Corollary 10.3.6, $m = m'$ and $\tau_i' = \tau_i$ and $\psi_i' \geq \psi_i$. Moreover, by Lemma 3.2.2, $\tau_2' \geq \psi_1'$. In a Grassmann graph $\tau_2 = \psi_1 = q+1$. Therefore, $\tau_2' = \psi_1' = q+1$.

Since $d \geq 3$ and $m' = m = [d]_q$ we get $m' > q+1$. Finally, $k'/m' + 1 = k'/m + 1 \geq k/m + 1 \geq q[s-d]_q + 1 \geq q^2+q+1$.

Hence, by Theorem 3.6.2, X_1' is the Grassmann graph $J_q(s',d)$ for some s'. □

Remark 10.6.2. The inequality (10.8) holds for $J_q(s,d)$ if $s \geq 3d+7$.

Proof. Recall that for $J_q(s,d)$ we have $m = [d]_q$, $\mu = (q+1)^2$ and $k = q[d]_q[s-d]_q$. Hence (10.8) is equivalent to

$$q\left(\frac{q^{s-d}-1}{q-1}\right) \geq 26(q+1)^2\left(\frac{q^d-1}{q-1}\right)^2 \tag{10.10}$$

Finally, note that $q^{2d}-1 \geq (q^d-1)^2$ and $q^8(q-1) \geq 2^7 q(q-1) \geq 26(q+1)^2$ for $q \geq 2$. □

Hence, we can state the extension theorem for Grassmann schemes in the following form (confirming part 3 of Theorem 1.3.6).

Theorem 10.6.3. *Let $\mathfrak{X}' = (\Omega', \mathfrak{c}')$ be an association scheme. Let $\Omega \subseteq \Omega'$ with $|\Omega'| < (3/2)|\Omega|$. Assume that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is the Grassmann scheme $\mathfrak{J}_q(s,d)$ with $d \geq 3$ and $s \geq 3d+7$. Then $\mathfrak{X}'$ is the Grassmann scheme $\mathfrak{J}_q(s',d)$ for some $s' \geq s$.*

Proof. Immediately follows from Theorem 10.6.1 and Remark 10.6.2. □

10.7 A universality result

In this section we prove the following result.

Theorem 10.7.1 (Affine superconfigurations)**.** *Every symmetric configuration of rank r on n vertices is a subconfiguration of some primitive 2-dimensional affine association scheme of rank r on $O(n^4)$ vertices.*

(See the relevant definitions below.)

10.7.1 2-dimensional affine association schemes

We define the *2-dimensional affine association schemes.* For simplicity we shall refer to them as *affine schemes.*

Let $\mathbb{F}_q$ be a finite field. Consider the affine plane $\mathbb{F}_q^2$. For every pair of distinct points $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$ in the plane there exists a unique (affine) line passing through these points. The *slope* of this line is

$$\operatorname{slope}(p_1, p_2) = \begin{cases} \infty & \text{if } x_1 = x_2 \text{ and} \\ \dfrac{y_2 - y_1}{x_2 - x_1} \in \mathbb{F}_q & \text{if } x_1 \neq x_2 \end{cases} \tag{10.11}$$

Definition 10.7.2 (Affine configurations)**.** Let $\mathfrak{c}^* : \mathbb{F}_q \cup \{\infty\} \to [r-1]$ be a surjection. Consider the configuration $\mathfrak{X}(\mathbb{F}_q, \mathfrak{c}^*) := (\Omega, \mathfrak{c})$ defined as follows:

- $\Omega := \mathbb{F}_q \times \mathbb{F}_q$
- for $u \in \Omega$ set $\mathfrak{c}(u,u) = 0$
- for $u, v \in \Omega$, $u \neq v$, set $\mathfrak{c}(u,v) = \mathfrak{c}^*(\text{slope}(u,v))$

Lemma 10.7.3. *The affine configuration $\mathfrak{X} := \mathfrak{X}(\mathbb{F}_q, \mathfrak{c}^*) = (\Omega, \mathfrak{c})$ based on the surjection $\mathfrak{c}^* : \mathbb{F}_q \cup \{\infty\} \to [r-1]$ is an association scheme of rank r. Moreover, $\mathfrak{X}$ is primitive if and only if $|(\mathfrak{c}^*)^{-1}(s)| \geq 2$ for every $s \in [r-1]$.*

Proof. Take any $u, v \in \mathbb{F}_q^2$ and any pair of distinct slopes $s_1 \neq s_2$. Then the line through u with slope s_1 intersects the line through v with slope s_2 in precisely one point. The lemma easily follows from this observation. For for $i, j, k \in \{1, \ldots, r-1\}$, the intersection numbers are

$$p^i_{j,k} = |(\mathfrak{c}^*)^{-1}(j)| \cdot |(\mathfrak{c}^*)^{-1}(k)| \,. \tag{10.12}$$

For $i = 0$ and $j = k$ we have $p^0_{j,j} = (q-1)|(\mathfrak{c}^*)^{-1}(j)|$. For $j = 0$ or $k = 0$ we have $p^i_{0,i} = p^i_{i,0} = 1$. In all other cases we have $p^0_{j,k} = 0$. □

Remark 10.7.4. Affine configurations are a special case of "translation association schemes" discussed in [Brouwer et al., 1989, Section 2.10].

10.7.2 Sidon sets in finite fields

Motivated by Simon Sidon's results on lacunary Fourier series (Sidon [1932]), Erdős and Turán [1941] defined a subset $S \subset \mathbb{N}$ to be a *Sidon set* if for every four elements $x, y, z, w \in S$, if $x + y = z + w$ then $\{x, y\} = \{z, w\}$. This concept was generalized to groups in Babai and Sós [1985]. We need the definition in the special case when the group is Abelian.

Definition 10.7.5 ([Babai and Sós, 1985, Def. 1.1.]). We call a subset S of an Abelian group G a *Sidon set* if for every $x, y, z, w \in S$, of which at least three are distinct, we have

$$x + y \neq z + w\,. \tag{10.13}$$

(For groups of odd order, we could just have taken the Erdős–Turán definition. However, if $a \in G$ has order 2 then the identity $x + x = (x + a) + (x + a)$ would kill any reasonable theory.)

We need the following observation from Babai and Sós [1985]. Finite *elementary Abelian groups* are the additive groups of finite fields.

Proposition 10.7.6 ([Babai and Sós, 1985, Prop. 5.1]). *Let q be a prime power and G an elementary Abelian group of order q^2. Then G contains a Sidon set of size q.*

(This result is best possible for odd q.) The proof is a simple adaptation of the method of Erdős and Turán [1941] who studied the maximum size of a Sidon set in the interval $[n]$.

10.7.3 Proof of the universality result

In this section we prove Theorem 10.7.1.

Lemma 10.7.7. *Let $q = p^{2e}$ be an even power of a prime number. Then there exists a subset of points, $X \subset \mathbb{F}_q^2$, of size $|X| \geq \sqrt{q}$, such that all slopes* $\mathrm{slope}(x, y)$ *for $x \neq y \in X$ are distinct.*

Proof. Let S be a Sidon set of size $\sqrt{q}$ in $\mathbb{F}_q$ (such a set exists by by Prop. 10.7.6). Let $X = \{(x, x^2) \mid x \in S\}$. So we have $|X| = |S| = \sqrt{q}$. We claim that all slopes among pairs of points in X are distinct.

Indeed, for $x, y \in X$, $x \neq y$, we have

$$\mathrm{slope}((x, x^2), (y, y^2)) = \frac{x^2 - y^2}{x - y} = x + y\,, \tag{10.14}$$

and these numbers are all distinct by the definition of S. $\square$

Theorem 10.7.8. *Every symmetric configuration $\mathfrak{X}$ with n vertices and rank r is a subconfiguration of some primitive affine scheme $\mathfrak{X}'$ of rank r derived from the affine plane $\mathbb{F}_q^2$ where q is any even power of a prime satisfying $q \geq n^2$. So $n' = q^2$. In particular, there exists such $\mathfrak{X}'$ with $n' \leq 16n^4$ vertices for every n, and asymptotically, $n' \leq n^4(1 + o(1))$ vertices suffice.*

Proof. Assume q is an even power of a prime and $q \geq n^2$. Take a subset $X \subset \mathbb{F}_q^2$ of size $|X| = n$ with all slopes distinct (Lemma 10.7.7). Consider any bijection from vertices of $\mathfrak{X}$ to X. This bijection defines colors of $\mathrm{slope}(x, y)$ for all $x \neq y \in X$. Since all these $n(n-1)/2$ slopes are distinct, their colors are well defined. We define the colors of the other slopes arbitrarily from $[r-1]$, subject to the condition that every color in $[r-1]$ must be the color of at least two distinct slopes. Since the number of remaining slopes to which we need to assign colors is $q+1-n(n-1)/2 > n(n+1)/2$, we have room to duplicate all colors occurring among the points in X.

The affine scheme so constructed is primitive of rank r. By construction, $\mathfrak{X} = \mathfrak{X}'[X]$.

Now taking q to be an even power of 2, we can achieve $q < 4n^2$. Alternatively, taking p to be the smallest prime $p \geq n$ and setting $q = p^2$ we have $q = n^2(1 + o(1))$. $\square$

Corollary 10.7.9. *The Johnson scheme $\mathfrak{J}(s, d)$ on n vertices is a subconfiguration of a non-Johnson primitive association scheme $\mathfrak{X}'$ of the same rank $d + 1$ on $\leq 16n^4$ vertices.*

Proof. We claim that affine schemes are never Johnson. The result below says that they cannot even have the same number of vertices. $\square$

The following result is certainly well known, but we could not find a convenient reference.

Proposition 10.7.10. *Let $2 \leq k \leq n - 2$. Then $\binom{n}{k}$ is not a prime power.*

Rather than deriving this fact from considerably deeper results such as the fact that binomial coefficients are almost never full powers, we give a straightforward and self-contained proof.

Proof. Assume for a contradiction that $\binom{n}{k} = p^\ell$ where p is a prime and $\ell \geq 1$.

It is an easy exercise that if a prime power p^ℓ divides the binomial coefficient $\binom{n}{k}$ then $p^\ell \leq n$. (This is implicit in the proof of Bertrand's postulate by Erdős [1932] and follows from the well-known formula for the exponent of p in $n!$.) So in our case $\binom{n}{k} \leq n$, a contradiction. $\square$

CHAPTER 11

ROBUSTNESS UNDER FUSION AND EXTENSION

11.1 Introduction

The material of this Chapter (except for Sections 11.4.3 and 11.4.4) is a result of a joint work by Babai and Kivva [2022].

In this Chapter we consider the following setup. Let $\mathfrak{Y}' = (\Omega', \mathfrak{c}')$ be the fusion of a configuration $\mathfrak{X}' = (\Omega', \mathfrak{c}'_0)$ via a color map η_0. Let $\Omega \subseteq \Omega'$. Assume, that $\mathfrak{Y}'$ and $\mathfrak{X} = \mathfrak{X}'[\Omega]$ are homogeneous coherent configurations.

We follow the notation of Chapter 10.

We prove the following result.

Theorem 11.1.1. *If a homogeneous coherent configuration $\mathfrak{X}$ on n vertices or its fission contains as a subscheme*

- *a Johnson scheme $J(s,d)$ with $s \geq 250d^4$, $d \geq 2$, on $\geq (5/6)n$ vertices,*
- *a Hamming scheme $H(d,s)$ with $s \geq 200d^4 \ln d$, $d \geq 2$, on $\geq (5/6)n$ vertices,*
- *or a Grassmann scheme $J_q(s,d)$ with $s \geq 6d+5$, $d \geq 3$, on $\geq (4/5)n$ vertices,*

then $\mathfrak{X}$ is itself a Johnson, or a Hamming, or a Grassman scheme, respectively.

11.2 Kaluzhnin-Klin's approach to show non-existence of a non-trivial fusion of a Johnson scheme

An important question in the study of the Johnson schemes is whether they admit a non-trivial fusion or fission. The only infinite families of examples of non-trivial fusion are known for $J(2d,d)$ and $J(2d+1,d)$. The only other known examples are "sporadic" examples for $J(10,3)$, $J(11,4)$ and $J(13,6)$ found by Klin.

Kaluzhnin and Klin [1972] proved the following theorem.

Theorem 11.2.1 (Kaluzhnin and Klin [1972])**.** *There exists a function $c(d)$ such that for every $s > c(d)$ there is no non-trivial fusion of the Johnson scheme $\mathfrak{J}(s,d)$.*

In his PhD thesis Klin [1974] showed that one can take $c(d) = O(d^4)$. Later, Muzychuk [1992a] improved bound to $c(d) = 3d + 4$ and Uchida [1992] proved another slight improvement to $c(d) = 2d + \sqrt{(d - 7/2)^2 + 6} + 3/2$.

About fission of $\mathfrak{J}(s,d)$ almost nothing is known. In particular, to the authors knowledge it is even open if there exists a rank-4 fission of $J(s,2)$ for sufficiently large s. Non-existence of such fission, for instance, will substantially simplify some proofs in Kivva [2021a].

In this section we briefly outline the ideas of the proof by Kaluzhnin and Klin. Their proof consists of several important observations. We denote by $p^t_{i,j}$ and $\widetilde{p}^{\ell}_{h,g}$ intersection numbers of $\mathfrak{X}$ and $\mathfrak{Y}$ respectively, whenever they are well defined.

Observation 11.2.2. Assume that a coherent configuration $\mathfrak{Y}$ is a fusion of $\mathfrak{X}$ via a color map η. For colors i, j, t, ℓ of $\mathfrak{X}$, if

$$\widetilde{p}^{\eta(t)}_{\eta(i),\eta(j)} = \sum_{a\in\eta^{-1}(i)} \sum_{b\in\eta^{-1}(j)} p^t_{a,b} \neq \sum_{a\in\eta^{-1}(i)} \sum_{b\in\eta^{-1}(j)} p^{\ell}_{a,b} = \widetilde{p}^{\eta(\ell)}_{\eta(i),\eta(j)},$$

then $\eta(t) \neq \eta(\ell)$.

In our case, intersection numbers of a fission configuration $\mathfrak{X}$ are well-known, as by assumption it is a Johnson scheme. So to make use of the observation above note that the following inequalities hold.

Observation 11.2.3. Assume that a coherent configuration $\mathfrak{Y}$ is a fusion of coherent configuration $\mathfrak{X}$ via a color map η. Then the following inequality holds.

$$p^t_{i,j} \leq \widetilde{p}^{\eta(t)}_{\eta(i),\eta(j)} = \sum_{a\in\eta^{-1}(i)} \sum_{b\in\eta^{-1}(j)} p^t_{a,b}.$$

The main technical observation of Kaluzhnin and Klin [1972] can be reformulated as follows. For sufficiently large $c(d)$ and $s > c(d)$ intersection numbers of the Johnson scheme $J(s,d)$ satisfy the next lemma.

Lemma 11.2.4. *Let $s > c(d)$. Then for every $1 \le j \le d$, intersection numbers of $\mathfrak{J}(s,d)$ satisfy*

1. $p^i_{j,j} \ge \sum_{\ell,t \le j} p^{j+1}_{\ell,t}$ *for all* $1 \le i \le j$.
2. $p^{j+1}_{j,j} \ge \sum_{i,t \le j} p^{\ell}_{i,t}$ *for all* $j+2 \le \ell \le d$.

Remark 11.2.5. There exists suitable $c(d) = O(d^4)$ by parts 5, 7 and 8 of Prop. 11.A.3.

These inequalities, combined with the two observations above immediately give the main lemma of Kaluzhnin-Klin's proof.

Lemma 11.2.6 (Kaluzhnin and Klin [1972])**.** *Assume that $s > c(d)$. Let $\mathfrak{Y}$ be a fusion of $\mathfrak{J}(s,d)$ via map η. Let $I = \eta^{-1}(h)$ for some color h of $\mathfrak{Y}$ and let t be the largest element of I. Assume that $t < d$. Then $\eta^{-1}(\eta(t+1)) = \{t+1\}$, that is, $t+1$ is not merged with any other color.*

Proof. By Observation 11.2.2 it is enough to show that $\widetilde{p}^{\eta(t+1)}_{\eta(t),\eta(t)}$ is distinct from $\widetilde{p}^{\eta(i)}_{\eta(t),\eta(t)}$ for every $i \ne t+1$. This is straightforward from Observation 11.2.3 and Lemma 11.2.4. □

To finish the proof Theorem 11.2.1 one needs to observe that if color d of $\mathfrak{J}(s,d)$ is not merged with any other color, then no colors are merged. This again follows from Observation 11.2.2 and the fact that $p^i_{d,d}$ are all distinct for the Johnson scheme $\mathfrak{J}(s,d)$.

11.3 A generalization of Kaluzhnin-Klin's approach

Let $\mathfrak{Y}' = (\Omega', \mathfrak{c}')$ be the fusion of a configuration $\mathfrak{X}' = (\Omega', \mathfrak{c}'_0)$ via a color map η_0. Let $\Omega \subseteq \Omega'$. Assume, that $\mathfrak{Y}'$ and $\mathfrak{X} = \mathfrak{X}'[\Omega]$ are homogeneous coherent configurations.

Notation 11.3.1. Let $\mathfrak{c}' : \Omega' \times \Omega' \to S'$ be a coloring. Let $\mathfrak{c} = \mathfrak{c}'|_{\Omega\times\Omega}$. For $\eta : S' \to S'_0$ we denote $\eta|_\Omega := \eta|_{\mathrm{Range}(\mathfrak{c})}$.

Observation 11.3.2. Let $\mathfrak{Y}' = (\Omega', \mathfrak{c}'_y)$ be the fusion of a configuration $\mathfrak{X}' = (\Omega', \mathfrak{c}'_x)$ via a color map η. Let $\Omega \subseteq \Omega'$. Then the subconfiguration $\mathfrak{Y}[\Omega]$ is the fusion of $\mathfrak{X}[\Omega]$ via the color map $\eta|_\Omega$, defined as above.

Remark 11.3.3. Notice, since $\mathfrak{X}'[\Omega]$ is a homogeneous coherent configuration, its fusion $\mathfrak{Y}'[\Omega]$ is a regular configuration. Thus, by the Lemma 10.2.2, for $|\Omega'| < 2|\Omega|$, rank of $\mathfrak{Y}'$ is equal to the rank of $\mathfrak{Y}'[\Omega]$. In particular, if $\mathfrak{Y}'[\Omega]$ is a symmetric configuration, so is $\mathfrak{Y}'$ itself. Therefore, in Theorem 1.3.2 we only need to consider the case when $\mathfrak{Y}'$ is an association scheme.

First, we make the following observation that allows us to focus on association schemes.

Observation 11.3.4. It is sufficient to prove Theorem 11.1.1 under the assumption that $\mathfrak{Y}'$ is an association scheme.

Proof. Notice, since $\mathfrak{X}'[\Omega]$ is a homogeneous coherent configuration, its fusion $\mathfrak{Y}'[\Omega]$ is a regular configuration. Thus, by the Lemma 10.2.2, for $|\Omega'| < 2|\Omega|$, rank of $\mathfrak{Y}'$ is equal to the rank of $\mathfrak{Y}'[\Omega]$. In particular, if $\mathfrak{Y}'[\Omega]$ is a symmetric configuration, so is $\mathfrak{Y}'$. So, in Theorem 11.1.1 we only need to consider the case when $\mathfrak{Y}'$ is an association scheme. $\square$

We are going use the following strategy:

1. First, we show that if $\mathfrak{Y}'$ is an association scheme and $\mathfrak{X}'[\Omega]$ is a Johnson, or a Hamming, or a Grassmann scheme, then under mild assumptions $\mathfrak{Y}'[\Omega] \equiv \mathfrak{X}'[\Omega]$.

2. Next, using Theorems 10.5.2 and 10.6.1, we can deduce that $\mathfrak{Y}'$ is itself a Johnson, or a Hamming or a Grassmann scheme, respectively.

To prove the first claim we follow the strategy, which can be seen as a generalization of approach of Kaluzhnin and Klin [1972]. In this section we develop technical tools to implement this approach.

This approach can be applied to association schemes that satisfy the following simple inequalities.

Definition 11.3.5. We say that an association scheme is *fusion-robust* with parameter $\gamma > 0$ if it satisfies the following properties.

1. For every $j \leq d, \quad k_{j-1} \leq \gamma k_j$.
2. For every $2 \leq j+1 \leq t \leq d-1, \quad p^{t+1}_{j,j} \leq \gamma p^t_{j,j}$.
3. For every $2 \leq j+1 \leq t \leq d, \quad \sum_{i,h=1}^{j} p^t_{i,h} \leq (1+\gamma) p^t_{j,j}$.
4. For $1 \leq t \leq j \leq d-1, \quad p^{j+1}_{j,j} \leq \gamma p^t_{j,j}$.

We observe that in a wide range of parameters Johnson, Hamming and Grassmann schemes are fusion-robust.

Lemma 11.3.6. *Let* $0 < \gamma < 1/2$.

1. *The Johnson scheme* $\mathfrak{J}(s,d)$ *is fusion-robust with parameter* γ *if* $s \geq 6d^4/\gamma + 3d$.
2. *The Hamming scheme* $\mathfrak{H}(d,s)$ *is fusion-robust with parameter* γ *if* $s \geq 10d^3/\gamma$.
3. *The Grassmann scheme* $\mathfrak{J}_q(s,d)$ *is fusion-robust with par.* γ *if* $s \geq 6d - 2 + \log_q(32/\gamma)$.

Proof. Follows from Propositions 11.A.3, 11.B.3 and 11.C.3, respectively, in Sections 11.A, 11.B and 11.C. □

We show that for fusion-robust schemes, the intersection numbers of an association scheme $\mathfrak{X}$ can be used to bound some intersection numbers of a fusion of its extension. The next pair of lemmas are the main technical lemmas of this section.

Lemma 11.3.7. *Assume that an association scheme $\mathfrak{Y}' = (\Omega', \mathfrak{c}')$ is the fusion of a configuration $\mathfrak{X}'$ via the color map η_0. Let $\Omega \subseteq \Omega'$, with $n' < 2n$. Denote $\eta = \eta_0|_\Omega$. Moreover, assume that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is an association scheme. Let h and g be colors of $\mathfrak{Y}'$ and $I \subseteq \eta^{-1}(h)$. Then*

$$n\left(\sum_{i\in I} k_i\right)(p')^h_{g,g} \leq (n'-n)k'_h(p')^h_{g,g} + n\sum_{i\in I} k_i \sum_{m,\ell\in\eta^{-1}(g)} p^i_{\ell,m}.$$

Proof. We count the number M of ordered triples (u, v, w) with $u, v \in \Omega$, $w \in \Omega'$ such that $\mathfrak{c}(u, v) \in I$, $\mathfrak{c}'(u, w) = \mathfrak{c}'(w, v) = g$. On one hand, we have $n\left(\sum_{i\in I} k_i\right)$ ways to choose a pair (u, v) and for every such pair there are $(p')^h_{g,g}$ ways to pick w. Here, we used that $I \subseteq \eta^{-1}(h)$. Therefore,

$$M = n\left(\sum_{i\in I} k_i\right)(p')^h_{g,g}.$$

On the other hand, we may first pick vertex w and then count the number of pairs (u, v) which form a desired triple with w. We distinguish 2 cases: $w \in \Omega$ and $w \notin \Omega$.

Since $I \subseteq \eta^{-1}(h)$, for every $w \in \Omega'$ the number of desired pairs (u, v) is at most the number of pairs (u', v') such that $\mathfrak{c}'(u', v') = h$ and $\mathfrak{c}'(u', w) = \mathfrak{c}'(v', w) = g$. This number of pairs is equal $k'_g(p')^g_{g,h} = k'_h(p')^h_{g,g}$. We use this estimate when $w \notin \Omega$.

When $w \in \Omega$ the number M_0 of desired pairs (u, v) for w does not depend on w, and can be expressed in terms of the intersection numbers of $\mathfrak{X}$. Indeed,

$$M_0 = \sum_{i\in I}\sum_{m,\ell\in\eta^{-1}(g)} k_\ell p^\ell_{i,m} = \sum_{i\in I}\sum_{m,\ell\in\eta^{-1}(g)} k_i p^i_{\ell,m} = \sum_{i\in I} k_i \sum_{m,\ell\in\eta^{-1}(g)} p^i_{\ell,m}.$$

Therefore, we get an inequality

$$n\left(\sum_{i\in I} k_i\right)(p')^h_{g,g} = M \leq (n'-n)k'_h(p')^h_{g,g} + nM_0. \tag{11.1}$$

□

Lemma 11.3.8. *Suppose that $\mathfrak{X}$ satisfies assumptions of Lemma 11.3.7. Additionally, suppose that for $0 < \gamma \leq \varepsilon$, $\mathfrak{X}$ satisfies inequalities 1-3 of Def. 11.3.5. Let $t \in \eta^{-1}(h)$ and let j be the maximal element of the set $\eta^{-1}(g)$. Suppose that $j < t$. Then*

$$\left(\frac{3n - 2n' - \gamma(n' - n)}{2n - n'}\right)(p')^h_{g,g} \leq (1+\gamma)p^j_{t,t}.$$

In particular, for $n' \leq \left(\frac{3}{2} - \varepsilon\right)n$ and $\gamma \leq \varepsilon < 1$, we have $(p')^h_{g,g} \leq \frac{1+\gamma}{\gamma}p^j_{t,t}$.

Proof. Define $I = \{i \in \eta^{-1}(h) \mid i \geq j\}$. By Lemma 10.2.2,

$$k'_h(2n - n') \leq n \sum_{i \in \eta^{-1}(h)} k_i.$$

By the definition of the set I, and by inequality 1 of Def. 11.3.5, we have

$$\sum_{i \in \eta^{-1}(h)} k_i \leq \sum_{i=1}^{j-1} k_i + \sum_{i \in I} k_i \leq \gamma k_j + \sum_{i \in I} k_i \leq (1+\gamma)\sum_{i \in I} k_i.$$

Since, for every $i \in I$ and every $\ell \in \eta^{-1}(g)$ the inequality $i > t \geq \ell$ holds, by inequality 3 of Def. 11.3.5, we obtain

$$\sum_{m,\ell \in \eta^{-1}(g)} p^i_{\ell,m} \leq (1+\gamma)p^i_{t,t}.$$

Now, since $t < j \leq i$ for all $i \in I$, inequality 2 of Def. 11.3.5 implies $p^i_{t,t} \leq p^j_{t,t}$. Hence

$$M_0 \leq (1+\gamma)p^j_{t,t}\sum_{i \in I} k_i.$$

Combining all together, inequality (11.1) becomes

$$n(p')^h_{g,g}\left(\sum_{i \in I} k_i\right) \leq (n' - n)\frac{n}{2n - n'}(p')^h_{g,g}(1+\gamma)\left(\sum_{i \in I} k_i\right) + n(1+\gamma)p^j_{t,t}\left(\sum_{i \in I} k_i\right), \text{ so}$$

$$\left(\frac{3n-2n'-\gamma(n'-n)}{2n-n'}\right)(p')^h_{g,g} \leq (1+\gamma)p^j_{t,t}. \qquad \square$$

Using this bound on the intersection numbers, we get the following lemma for fusion-robust configurations.

Lemma 11.3.9. *Let $\mathfrak{Y}'$ be an association scheme on Ω'. Assume that $\mathfrak{Y}'$ is the fusion of a configuration $\mathfrak{X}'$ via a color map η_0. Let $\Omega \subseteq \Omega'$, with $n' \leq (3/2-\varepsilon)\,n$, where $0<\varepsilon<1/2$. Denote $\eta = \eta_0|_\Omega$. Suppose that $\mathfrak{X}'[\Omega]$ is fusion-robust for $\gamma = \varepsilon/3$. Let h be a color of $\mathfrak{Y}'$ and t be the maximal element of $\eta^{-1}(h)$. Assume $t<d$. Then $\eta^{-1}(\eta(t+1)) = \{t+1\}$, i.e., color $t+1$ is not merged with any other color of $\mathfrak{X}'[\Omega]$.*

In particular, the colors of $\mathfrak{Y}'$ can be renamed, so that for some $1 \leq x \leq d$, we have $\eta^{-1}(1) = \{1,2,\ldots,x\}$ and $\eta^{-1}(i) = \{i+x-1\}$ for all $2 \leq i \leq d-x+1$.

Proof. To prove the claim of the lemma it is enough to show that $(p')^{\eta(t+1)}_{h,h}$ is distinct from all $(p')^{\eta(i)}_{h,h}$, for $i \neq t+1$.

Since $\mathfrak{X}'[\Omega]$ is fusion-robust for $\gamma = \varepsilon/3$, by Lemma 11.3.8, for $j \geq t+1$, $(p')^{\eta(j)}_{h,h} < (2/\varepsilon)p^j_{t,t}$.

Using part 3 of Def. 11.3.5 , we get

$$(p')^{\eta(j)}_{h,h} \leq \frac{2}{\varepsilon}p^j_{t,t} \leq \gamma \cdot \frac{2}{\varepsilon}p^{t+1}_{t,t} < p^{t+1}_{t,t}.$$

Since, $(p')^{\eta(t+1)}_{h,h} \geq p^{t+1}_{t,t}$, we immediately obtain $\eta(t+1) \neq \eta(j)$, for all $j \geq t+2$.

The inclusion $t \in \eta^{-1}(h)$, implies $(p')^{\eta(j)}_{h,h} \geq p^j_{t,t}$. Combining this with Lemma 11.3.8 and part 4 of Def. 11.3.5, for $j \leq t$, we get

$$(p')^{\eta(t+1)}_{h,h} < (2/\varepsilon)p^{t+1}_{t,t} \leq p^j_{t,t} \leq (p')^{\eta(j)}_{h,h}.$$

Thus, $\eta(t+1) \neq \eta(j)$ for all $j \neq t+1$. $\square$

11.4 Robustness of Johnson, Hamming and Grassmann schemes

Now our goal is to show that in Lemma 11.3.9 under mild assumptions for Johnson, Hamming and Grassmann schemes x must be 1. In order to do this, we will need the following inequality.

11.4.1 Technical lemmas

Lemma 11.4.1. *Let $\mathfrak{Y}'$ be an association scheme on Ω' of rank ≥ 3. Assume that $\mathfrak{Y}'$ is a fusion of a configuration $\mathfrak{X}'$ via a color map η_0. Let $\Omega \subseteq \Omega'$, with $n' \leq (3/2-\alpha)n$, where $0 < \alpha < 1/2$. Let $\eta = \eta_0|_\Omega$. Suppose $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is an association scheme and $\eta^{-1}(1) = \{1, 2, \ldots, t\}$. Then*

$$\sum_{i=1}^{t} k_i \sum_{\ell,j=1}^{t} p^i_{\ell,j} \geq 2\alpha k_t p^1_{t,t}$$

Proof. To prove this, we count the number of triangles with all sides being of color 1 in $\mathfrak{Y}'[\Omega]$.

Let Y_1' be the color-1 constituent of $\mathfrak{Y}'$ and let $Y_1 = Y_1'[\Omega]$ be the induced on Ω graph. The total number of triangles in Y_1' equals $n'k_1'(p')^1_{1,1}/6$ and there are at most $(n'-n)k_1'(p')^1_{1,1}/2$ triangles in Y_1 with at least one vertex in $\Omega' \setminus \Omega$. Therefore, there are at least

$$M' = \frac{1}{6}n'k_1'(p')^1_{1,1} - \frac{1}{2}(n'-n)k_1'(p')^1_{1,1} = \left(\frac{n}{2} - \frac{n'}{3}\right) k_1'(p')^1_{1,1}$$

triangles in Y_1.

Clearly, $k_1' \geq k_t$, and

$$(p')^1_{1,1} \geq \sum_{\ell,j=1}^{t} p^1_{\ell,j} \geq p^1_{t,t}, \quad \text{so} \quad 6M' \geq 2\alpha n k_t p^1_{t,t}. \tag{11.2}$$

At the same time, we can count the number of triangles in Y_1 precisely. Recall that $\eta^{-1}(1) = \{1, 2, \dots t\}$, i.e., each edge in Y_1 is an edge with color in $[t]$ in $\mathfrak{X}$. There are nk_i ways to choose an ordered pair of vertices (u, v) joined by an edge of color i in $\mathfrak{X}$ and there are $\sum_{\ell,j=1}^{t} p^i_{\ell,j}$ ways to pick a vertex w with $c(u, w) \in [t]$ and $c(w, v) \in [t]$. Therefore, the number of triangles in Y_1 is

$$M = \frac{n}{6} \sum_{i=1}^{t} k_i \sum_{\ell,j=1}^{t} p^i_{\ell,j}.$$

Thus, the desired inequality is implied by $M \geq M'$. □

Finally, before proving that $\mathfrak{Y}'[\Omega] \equiv \mathfrak{X}'[\Omega]$ for Hamming or Grassmann schemes, we need the following inequality.

Lemma 11.4.2. *Let $0 < \gamma < 1/2$. Let $1 \leq t \leq \ell \leq j$ and $2 \leq j \leq d-1$ be integers.*

1. *For the Johnson scheme $\mathfrak{J}(s, d)$ with $s \geq 2d^4/\gamma + 3d$, we have*

$$p^{\ell}_{j,j} \leq (1+\gamma) p^{t}_{j,j}, \quad \text{and} \quad (d-1) p^{j}_{j,j} \leq (1+\gamma) p^{1}_{j,j}. \tag{11.3}$$

2. *For the Hamming scheme $\mathfrak{H}(d, s)$ with $s \geq 10\gamma^{-1} d^3$ we have*

$$p^{\ell}_{j,j} \leq (1+\gamma) p^{t}_{j,j}, \quad \text{and} \quad (d-1) p^{j}_{j,j} \leq (1+\gamma) p^{1}_{j,j}. \tag{11.4}$$

3. *For the Grassmann scheme $\mathfrak{J}_q(s, d)$ with $s \geq 6d - 4 + \log_q(32/\gamma)$ we have*

$$p^{\ell}_{j,j} \leq (1+\gamma) p^{t}_{j,j}, \quad \text{and} \quad [d-1]_q \cdot p^{j}_{j,j} \leq (1+\gamma) p^{1}_{j,j}. \tag{11.5}$$

Proof. See part 9 of Prop. 11.A.3, part 8 of Prop. 11.B.3 and part 8 of Prop. 11.C.3 in Appendices 11.A, 11.B and 11.C. □

11.4.2 Robustness of Johnson schemes

Theorem 11.4.3. *Let $\mathfrak{Y}'$ be an association scheme on Ω' of rank ≥ 3. Assume that $\mathfrak{Y}'$ is a fusion of a configuration $\mathfrak{X}'$ via a color map η_0. Let $\Omega \subseteq \Omega'$, with $n' \leq (3/2-\alpha)n$, where $0<\alpha<1/2$. Suppose that $\mathfrak{X}=\mathfrak{X}'[\Omega]$ is the Johnson scheme $\mathfrak{J}(s,d)$, with $s>2d^4/(\varepsilon\alpha)+3d$, for $0<\varepsilon\leq 1/18$. If $\alpha \geq (1+6\varepsilon)/(2d-2)$, then $\mathfrak{Y}'[\Omega]=\mathfrak{J}(s,d)$.*

Proof. Let $\eta=\eta_0|_\Omega$. By Lemma 11.3.9, we may assume that $\eta^{-1}(1)=\{1,2,\dots t\}$ and $\eta^{-1}(i)=\{i+t-1\}$ for all $2\leq i\leq d-t+1$. We are going to prove that $t=1$. By Lemma 11.4.1,

$$\sum_{i=1}^{t} k_i \sum_{\ell,j=1}^{t} p^i_{\ell,j} \geq 2\alpha k_t p^1_{t,t} \tag{11.6}$$

We are going to bound the expression in the left-hand side of the inequality. Using parts 2 (for $\gamma=1/2$) and 3 of Proposition 11.A.3, for $s>2d^4/\varepsilon+3d$,

$$\sum_{\ell,j=1}^{t} p^i_{\ell,j} \leq p^i_{t,t} + 2\sum_{j=1}^{t-1} k_j \leq (1+\varepsilon)\binom{d-i}{d-t}\binom{s-d-i}{t} + 4k_{t-1}. \tag{11.7}$$

Note also that for $1\leq i\leq t-1$ part 4 of Proposition 11.A.3 implies

$$\begin{gathered}\binom{d-i}{d-t}\binom{s-d-i}{t} \leq \binom{d-1}{d-t}\binom{s-d-1}{t} \leq p^1_{t,t}, \quad \text{and} \\ p^t_{t,t} \leq (1+\varepsilon)\binom{s-d-t}{t} \leq (1+\varepsilon)\binom{d-1}{d-t}^{-1} p^1_{t,t}.\end{gathered} \tag{11.8}$$

In particular, if $2\leq t\leq d-1$, the last inequality implies

$$p^t_{t,t} \leq (1+\varepsilon)p^1_{t,t}/(d-1).$$

Combining Eq. (11.7) and (11.8), we get

$$\sum_{\ell,j=1}^{t} p^{i}_{\ell,j} \leq (1+\varepsilon)p^{1}_{tt} + 4k_{t-1} \quad \text{and} \quad \sum_{\ell,j=1}^{t} p^{t}_{\ell,j} \leq p^{t}_{tt} + 4k_{t-1}. \tag{11.9}$$

Hence, using part 2 (with $\gamma = 1/2$) of Proposition 11.A.3,

$$k_t \sum_{\ell,j=1}^{t} p^{t}_{\ell,j} + \sum_{i=1}^{t-1} k_i \sum_{\ell,j=1}^{t} p^{i}_{\ell,j} \leq k_t p^{t}_{t,t} + 4k_t k_{t-1} + 2k_{t-1}((1+\varepsilon)p^{1}_{t,t} + 4k_{t-1}) \tag{11.10}$$

Recall that $p^{1}_{t,t} \geq k_t/(2d)$ by Eq. (11.22) and Eq.(11.8) and $k_{t-1} \leq \varepsilon k_t/(2d^2)$, by part 2 of Prop. 11.A.3 for $s \geq 2d^4/\varepsilon + 3d$. So, $\varepsilon p^{1}_{1,1}/d \geq k_{t-1}$ Thus, for $t \geq 2$,

$$k_t p^{t}_{t,t} + 4k_t k_{t-1} + 8k^2_{t-1} + 3k_{t-1}p^{1}_{t,t} \leq \left(\frac{1+\varepsilon}{d-1}\right) k_t p^{1}_{t,t} + \left(\frac{5\varepsilon}{d}\right) k_t p^{1}_{t,t} < \left(\frac{1+6\varepsilon}{d-1}\right) k_t p^{1}_{t,t},$$

This gives a contradiction with Eq. (11.6) for $t \geq 2$ and $\alpha \geq (1+6\varepsilon)/(2d-2)$.
Therefore, $t = 1$ and so $\mathfrak{Y}'[\Omega] = \mathfrak{J}(s,d)$. □

Finally, we prove Theorem 1.3.2

Proof of Theorem 1.3.2. If $d = 2$, clearly $\mathfrak{Y}'[\Omega] = \mathfrak{J}(s,d)$. For $d \geq 3$, we apply Theorem 11.4.3 with $\varepsilon = 1/30$, $\alpha = 3/10$. Hence, the result follows from Theorem 10.4.1. □

11.4.3 Robustness of Hamming schemes

Next, we prove similar robustness results for Hamming schemes.

Theorem 11.4.4. *Let $\mathfrak{Y}'$ be an association scheme on Ω' of rank ≥ 3. Assume that $\mathfrak{Y}'$ is a fusion of a configuration $\mathfrak{X}'$ via a color map η_0. Let $\Omega \subseteq \Omega'$, with $n' \leq (3/2-\alpha)n$, where $0 < \alpha < 1/2$. Suppose that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is the Hamming scheme $\mathfrak{H}(d,s)$, with $s \geq 30d^3/\alpha$, for $0 < \varepsilon \leq 1/18$. If $\alpha \geq (1+2\varepsilon)/(2d-2)$, then $\mathfrak{Y}'[\Omega] = \mathfrak{X} = \mathfrak{H}(d,s)$.*

Proof. Let $\eta = \eta_0|_\Omega$. By Lemma 11.3.9, we may assume that $\eta^{-1}(1) = \{1, 2, \dots t\}$ and $\eta^{-1}(i) = \{i + t - 1\}$ for all $2 \le i \le d - t + 1$. We are going to prove that $t = 1$. By Lemma 11.4.1,

$$\sum_{i=1}^{t} k_i \sum_{\ell,j=1}^{t} p^i_{\ell,j} \ge 2\alpha k_t p^1_{t,t} \tag{11.11}$$

We are going to bound the expression in the left-hand side of the inequality.

Recall, that by part 1 of Prop. 11.B.3, for $2 \le t \le d-1$, we have $k_t \ge (s/d)k_{t-1} \ge 2k_{t-1}$. Using Eq. (2.3),

$$\sum_{\ell,j=1}^{t} p^i_{\ell,j} \le p^i_{t,t} + 2\sum_{j=1}^{t-1} k_j \le p^i_{t,t} + 4k_{t-1}. \tag{11.12}$$

By Lemma 11.4.2, for $2 \le i + 1 \le t \le d - 1$, we have

$$p^i_{t,t} \le (1+\varepsilon)p^1_{t,t} \quad \text{and} \quad p^t_{t,t} \le (1+\varepsilon)p^1_{t,t}/(d-1). \tag{11.13}$$

Combining Eq. (11.12) and Eq. (11.13), we get, for $1 \le i \le t - 1$,

$$\sum_{\ell,j=1}^{t} p^i_{\ell,j} \le (1+\varepsilon)p^1_{t,t} + 4k_{t-1} \quad \text{and} \quad \sum_{\ell,j=1}^{t} p^t_{\ell,j} \le \frac{(1+\varepsilon)}{(d-1)}p^1_{t,t} + 4k_{t-1}. \tag{11.14}$$

Hence,

$$\begin{aligned}\sum_{i=1}^{t} k_i \sum_{\ell,j=1}^{t} p^i_{\ell,j} &\le \left(\sum_{i=1}^{t-1} k_i\right)\left((1+\varepsilon)p^1_{t,t} + 4k_{t-1}\right) + k_t\left(\frac{(1+\varepsilon)}{(d-1)}p^1_{t,t} + 4k_{t-1}\right) \le \\ &\le \frac{(1+\varepsilon)}{(d-1)}k_t p^1_{t,t} + 3k_{t-1}p^1_{t,t} + 4k_t k_{t-1} + 8k^2_{t-1} \le \\ &\le k_t p^1_{t,t}\left(\frac{(1+\varepsilon)}{(d-1)} + \frac{\varepsilon}{2d^2} + \frac{\varepsilon}{2d^2}\right) < \frac{(1+2\varepsilon)}{(d-1)}k_t p^1_{t,t}.\end{aligned} \tag{11.15}$$

Therefore, we get a contradiction with the inequality (11.11) for $\alpha \ge \dfrac{1+2\varepsilon}{2d-2}$. □

Proof of Thm. 1.3.4. If $\mathfrak{Y}$ is of rank 3, the claim follows from Theorem 10.5.2. If $\mathrm{rk}(\mathfrak{Y}') \ge 4$,

the claim follows from Theorem 11.4.4 applied with $\varepsilon = 1/10$ and $\alpha = 3/10$ and Theorem 10.5.2. $\square$

11.4.4 Robustness of Grassmann schemes

Finally, we establish similar claims for the Grassmann schemes.

Theorem 11.4.5. *Let $\mathfrak{Y}'$ be an association scheme on Ω' of rank ≥ 4. Assume that $\mathfrak{Y}'$ is a fusion of a configuration $\mathfrak{X}'$ via a color map η_0. Let $\Omega \subseteq \Omega'$, with $n' \leq (3/2 - \alpha)n$, where $0 < \alpha < 1/2$. Suppose that $\mathfrak{X} = \mathfrak{X}'[\Omega]$ is the Grassmann scheme $\mathfrak{J}_q(s,d)$, with $s \geq 6d - 2 + \log_q(32/\varepsilon)$, for $0 < \varepsilon < 1/2$. If $\alpha \geq (1+2\varepsilon)/(2[d-1]_q)$, then $\mathfrak{Y}'[\Omega] = \mathfrak{X} = \mathfrak{J}_q(s,d)$.*

Proof. Let $\eta = \eta_0|_\Omega$. By Lemma 11.3.9, we may assume that $\eta^{-1}(1) = \{1, 2, \dots t\}$ and $\eta^{-1}(i) = \{i + t - 1\}$ for all $2 \leq i \leq d - t + 1$. We are going to prove that $t = 1$. By Lemma 11.4.1,

$$\sum_{i=1}^{t} k_i \sum_{\ell,j=1}^{t} p^i_{\ell,j} \geq 2\alpha k_t p^1_{t,t} \tag{11.16}$$

We are going to bound the expression in the left-hand side of the inequality.

By part 1 of Prop. 11.C.3, for $2 \leq t \leq d-1$, we have $k_t \geq (q^{s-2d}/2)k_{t-1} \geq 2k_{t-1}$. Using Eq. (2.3),

$$\sum_{\ell,j=1}^{t} p^i_{\ell,j} \leq p^i_{t,t} + 2\sum_{j=1}^{t-1} k_j \leq p^i_{t,t} + 4k_{t-1}. \tag{11.17}$$

By Lemma 11.4.2, for $2 \leq i+1 \leq t \leq d-1$, we have

$$p^i_{t,t} \leq (1+\varepsilon)p^1_{t,t} \quad \text{and} \quad p^t_{t,t} \leq (1+\varepsilon)p^1_{t,t}/[d-1]_q. \tag{11.18}$$

Combining Eq. (11.17) and Eq. (11.18), we get, for $1 \leq i \leq t-1$,

$$\sum_{\ell,j=1}^{t} p^i_{\ell,j} \leq (1+\varepsilon)p^1_{t,t} + 4k_{t-1} \quad \text{and} \quad \sum_{\ell,j=1}^{t} p^t_{\ell,j} \leq \frac{(1+\varepsilon)}{[d-1]_q} p^1_{t,t} + 4k_{t-1}. \tag{11.19}$$

Hence,

$$\begin{aligned}\sum_{i=1}^{t} k_i \sum_{\ell,j=1}^{t} p^i_{\ell,j} &\le \left(\sum_{i=1}^{t-1} k_i\right)\Big((1+\varepsilon)p^1_{t,t} + 4k_{t-1}\Big) + k_t\left(\frac{(1+\varepsilon)}{[d-1]_q}p^1_{t,t} + 4k_{t-1}\right) \le \\ &\le \frac{(1+\varepsilon)}{[d-1]_q}k_t p^1_{t,t} + 3k_{t-1}p^1_{t,t} + 4k_t k_{t-1} + 8k^2_{t-1} \le \\ &\le k_t p^1_{t,t}\left(\frac{(1+\varepsilon)}{[d-1]_q} + \frac{\varepsilon}{[2d]_q} + \frac{\varepsilon}{[2d]_q}\right) < \frac{(1+2\varepsilon)}{[d-1]_q}k_t p^1_{t,t}.\end{aligned} \tag{11.20}$$

Therefore, we get a contradiction with the inequality (11.16) for $\alpha \ge \dfrac{1+2\varepsilon}{2[d-1]_q}$. □

Proof of Theorem 1.3.5. If $\mathrm{rk}(\mathfrak{Y}') \ge 4$, the claim follows from Theorem 11.4.5 applied with $\varepsilon = 1/4$ and $\alpha = 1/4$ and Theorem 10.6.1. □

11.A Appendix: Inequalities on the intersection numbers of the Johnson schemes

In this section we derive inequalities for the intersection numbers of the Johnson schemes. We use these inequalities in Section 11.4.2.

Fact 11.A.1. The intersection numbers of the Johnson scheme $\mathfrak{J}(s,d)$ can be computed as

$$p^j_{i,t} = \sum_{a=t_1}^{t_2} \binom{d-j}{a}\binom{j}{d-a-i}\binom{j}{d-a-t}\binom{s-d-j}{i+t+a-d}, \text{ for } t_1 \le t_2, \tag{11.21}$$

$$\text{and} \quad p^j_{i,t} = 0, \text{ for } t_1 > t_2,$$

where $t_1 = \max\{d-i-t, d-j-t, d-i-j, 0\}$ and $t_2 = \min\{d-t, d-i, d-j, s-j-i-t\}$. The degree of the j-th constituent is

$$k_j = \binom{d}{j}\binom{s-d}{j}. \tag{11.22}$$

Before we prove the desired inequalities for the intersection numbers it is convenient to make the following observation.

Observation 11.A.2. Let $d \geq 1$ and $0 \leq i \leq d-1$. Then

$$\frac{1}{d}\binom{d}{i} \leq \binom{d}{i+1} \leq d\binom{d}{i} \qquad \text{and} \qquad \binom{d-1}{i} \leq \binom{d}{i+1} \leq d\binom{d-1}{i}.$$

Proposition 11.A.3. *Let $\mathfrak{J}(s,d)$ be a Johnson scheme, $s \geq 2d+1$, $d \geq 2$. The following inequalities hold.*

1. $\binom{d}{j}\frac{(s-2d)^j}{j!} \leq k_j \leq \binom{d}{j}\frac{(s-d)^j}{j!}$, *for* $1 \leq j \leq d$.
2. *If* $s > (d^2/\gamma) + 2d$, *for* $\gamma > 0$, *then*

$$k_{j-1} \leq \gamma k_j. \tag{11.23}$$

3. *For* $\gamma \in (0,1)$, $1 \leq j \leq t \leq d$ *and* $s \geq 2d^4/\gamma + 3d$

$$\binom{d-j}{d-t}\binom{s-d-j}{t} \leq p^j_{t,t} \leq (1+\gamma)\binom{d-j}{d-t}\binom{s-d-j}{t}. \tag{11.24}$$

4. *For* $2 \leq t+1 \leq j \leq d$ *and* $s > 2d^4 + 3d$.

$$\binom{j}{j-t}^2\binom{s-d-j}{2t-j} \leq p^j_{t,t} \leq 2\binom{j}{j-t}^2\binom{s-d-j}{2t-j}. \tag{11.25}$$

5. *Let* $\gamma > 0$. *Then for* $2 \leq t+1 \leq j \leq d-1$ *and* $s \geq (d^3/\gamma) + 2d$

$$p^{j+1}_{t,t} \leq \gamma p^j_{t,t}. \tag{11.26}$$

6. *Let* $\gamma > 0$. *Then for* $1 \le \max(i,h) \le j < \min(i+h,d)$, *and* $s > (d^2/\gamma) + 3d$

$$p^j_{i-1,h} \le \gamma p^j_{i,h}. \tag{11.27}$$

7. *Let* $\gamma \in (0,1)$. *Then for* $2 \le t+1 \le j \le d$ *and* $s > \dfrac{(2+2\gamma)}{\gamma} d^2 + 3d$

$$\sum_{a,b=1}^{t} p^j_{a,b} \le (1+\gamma) p^j_{t,t}. \tag{11.28}$$

8. *Let* $\gamma > 0$. *For* $s > \max\{(6d^4/\gamma) + 3d, d^2\}$ *and* $1 \le j \le t \le d-1$

$$p^{t+1}_{t,t} \le \gamma p^j_{t,t}. \tag{11.29}$$

9. *Let* $0 < \gamma < 1/2$. *Let* $1 \le t \le \ell \le j$ *and* $2 \le j \le d-1$ *be integers*

$$p^\ell_{j,j} \le (1+\gamma) p^t_{j,j}, \quad \textit{and} \quad (d-1) p^j_{j,j} \le (1+\gamma) p^1_{j,j}. \tag{11.30}$$

Proof. 1. Follows from Eq. (11.22).

2. $k_{j-1} = \dfrac{j^2}{(d-j+1)(s-d-j+1)} k_j \le \dfrac{d^2}{s-2d} k_j \le \gamma k_j$.

Denote the summand with an index a in the sum (11.21) for $p^j_{i,t}$ by

$$z^a_{i,t,j} := \binom{d-j}{a}\binom{j}{d-a-i}\binom{j}{d-a-t}\binom{s-d-j}{i+t+a-d}.$$

Clearly, $z^a_{i,t,j} \ge 0$. Also, for convenience, we write down Eq. (11.21) for $i = t$

$$p^j_{t,t} = \sum_{a=t_1}^{t_2} \binom{d-j}{a}\binom{j}{d-t-a}^2\binom{s-d-j}{2t+a-d}, \tag{11.31}$$

where t_1 and t_2 are defined as in Fact 11.A.1.

3. For $j \le t$, in the sum (11.31) for $p^j_{t,t}$, we have $t_1 = \max\{d-t-j, 0\}$, $t_2 = d-t$. Since all $z^a_{t,t,j} \ge 0$,

$$p^j_{t,t} = \sum_{a=t_1}^{d-t} z^a_{t,t,j} \ge z^{d-t}_{t,t,j} = \binom{d-j}{d-t}\binom{s-d-j}{t}.$$

To prove the upper bound, it is enough to show that $\gamma z^a_{t,t,j} \ge 2z^{a-1}_{t,t,j}$ for $0 < a \le d-t$. Note that $2t+a-d \le d$, so

$$\binom{s-d-j}{2t+a-d} \ge \frac{s-3d}{d}\binom{s-d-j}{2t+a-1-d}.$$

Thus, using Observation 11.A.2,

$$\binom{d-j}{a}\binom{j}{d-a-t}^2\binom{s-d-j}{2t+a-d} \ge \frac{s-3d}{d^4}\binom{d-j}{a-1}\binom{j}{d-a+1-t}^2\binom{s-d-j}{2t+a-1-d},$$

which implies

$$z^a_{t,t,j} \ge \frac{s-3d}{d^4} z^{a-1}_{t,t,j} \ge 2z^{a-1}_{t,t,j}.$$

4. For $j \ge t+1$, in the sum (11.31) for $p^j_{t,t}$ we have $t_1 = d-2t$, $t_2 = d-j$. Thus

$$p^j_{t,t} = \sum_{a=d-2t}^{d-j} z^a_{t,t,j} \ge z^{d-j}_{t,t,j} = \binom{j}{j-t}^2\binom{s-d-j}{2t-j}.$$

Similarly, as in the previous part, one can verify that $s > 2d^4+3d$ implies $z^a_{t,t,j} \ge 2z^{a-1}_{t,t,j}$ for $d-2t+1 \le a \le d-j$. Hence, the desired upper bound follows.

5. We need to show that

$$\gamma \sum_{a=d-2t}^{d-j} z^a_{t,t,j} = \gamma p^j_{t,t} \ge p^{j+1}_{t,t} = \sum_{a=d-2t}^{d-j-1} z^a_{t,t,j+1}.$$

It is enough to show that $\gamma z^a_{t,t,j} \geq z^{a-1}_{t,t,j+1}$, i.e., we need to prove

$$\gamma \binom{d-j}{a} \binom{j}{d-t-a}^2 \binom{s-d-j}{2t+a-d} \geq \binom{d-j-1}{a-1} \binom{j+1}{d-t-a+1}^2 \binom{s-d-j-1}{2t+a-1-d}.$$

By Observation 11.A.2, it is enough to have

$$\gamma \frac{1}{(j+1)^2} \frac{s-d-j}{2t+a-d} \geq 1,$$

which is true for $s \geq d^3/\gamma + 2d$.

6. Since $\max(i,h) \leq j \leq i+h-1$ we need to verify

$$\sum_{a=d-h-i+1}^{d-j} z^a_{i-1,h,j} \leq \gamma \sum_{a=d-h-i}^{d-j} z^a_{i,h,j}.$$

It is enough to check that $z^a_{i-1,h,j} \leq \gamma z^a_{i,h,j}$, which is equivalent to

$$\gamma \binom{j}{d-a-i} \binom{s-d-j}{i+h+a-d} \geq \binom{j}{d-a-i+1} \binom{s-d-j}{i+h+a-d-1},$$

which follows from Observation 11.A.2 for $\gamma(s-3d) \geq d^2$.

7. Denote by $I_h = \{(a,b) \ : \ 1 \leq a \leq t, \ 1 \leq b \leq t, \ a+b=h\}$. We can write

$$\sum_{a,b=1}^{t} p^j_{a,b} = \sum_{h=j}^{2t} \sum_{(a,b)\in I_h} p^j_{a,b}.$$

Denote $\gamma_0 = \gamma/(2+2\gamma)$. Since $s > d^2/\gamma_0 + 3d$, using previous part, we deduce that

$$\sum_{(a,b)\in I_{h-1}} p^j_{a,b} \leq 2\gamma \sum_{(a,b)\in I_h} p^j_{a,b}.$$

Therefore,

$$\sum_{a,b=1}^{t} p_{a,b}^{j} \leq \sum_{h=0}^{2t-j} (2\gamma_0)^h \sum_{(a,b)\in I_{2t}} p_{a,b}^{j} \leq \sum_{h=0}^{\infty} (2\gamma_0)^h p_{t,t}^{j} = \frac{1}{1-2\gamma_0} p_{t,t}^{j} = (1+\gamma) p_{t,t}^{j}.$$

8. By part 3, for $j \leq t$,

$$p_{t,t}^{j} \geq \frac{(s-2d)^t}{t!}.$$

Part 4 implies

$$p_{t,t}^{t+1} \leq 2(t+1)^2 \binom{s-d-t-1}{t-1} \leq 2d^3 \frac{(s-d)^{t-1}}{t!}.$$

Therefore, for $s > 6d^4/\gamma + 3d$ we get

$$\frac{1}{\gamma} p_{t,t}^{t+1} \leq \frac{2}{\gamma} d^3 \frac{(s-d)^{t-1}}{t!} < \frac{2e}{\gamma} d^3 \frac{(s-2d)^{t-1}}{t!} < \frac{(s-2d)^t}{t!} \leq p_{t,t}^{j}.$$

9. Using parts 2 (for $\gamma = 1/2$) and 3 of Proposition 11.A.3, for $s > 2d^4/\varepsilon + 3d$,

$$\sum_{\ell,j=1}^{t} p_{\ell,j}^{i} \leq p_{t,t}^{i} + 2\sum_{j=1}^{t-1} k_j \leq (1+\varepsilon) \binom{d-i}{d-t} \binom{s-d-i}{t} + 4k_{t-1}. \tag{11.32}$$

Note also that for $1 \leq i \leq t-1$ part 4 of Proposition 11.A.3 implies

$$\begin{gathered} \binom{d-i}{d-t} \binom{s-d-i}{t} \leq \binom{d-1}{d-t} \binom{s-d-1}{t} \leq p_{t,t}^{1}, \quad \text{and} \\ p_{t,t}^{t} \leq (1+\varepsilon) \binom{s-d-t}{t} \leq (1+\varepsilon) \binom{d-1}{d-t}^{-1} p_{t,t}^{1}. \end{gathered} \tag{11.33}$$

In particular, if $2 \leq t \leq d-1$, the last inequality implies

$$p_{t,t}^{t} \leq (1+\varepsilon) p_{t,t}^{1}/(d-1). \qquad \square$$

11.B Appendix: Inequalities on the intersection numbers of the Hamming schemes

In this section we derive inequalities for the intersection numbers of the Hamming schemes. We use these inequalities in Section 11.4.3.

Fact 11.B.1. The intersection numbers of the Hamming graph $H(d,s)$ are

$$p^t_{i,j} = \sum_{a=a_1}^{a_2} (s-1)^a (s-2)^{i+j-t-2a} \binom{d-t}{a} \binom{t}{i+j-t-2a} \binom{2t-i-j+2a}{t-j+a}, \quad (11.34)$$

where $a_1 = \max(0, j-t, i-t)$ and $a_2 = \min\left(\left\lfloor \frac{i+j-t}{2} \right\rfloor, d-t\right)$.

Observation 11.B.2. The following inequalities hold for all $0 \le i \le d-1$

$$\frac{1}{d}\binom{d}{i} \le \binom{d}{i+1} \le d\binom{d}{i} \quad \text{and} \quad \binom{2d+2}{d+1} \le 4\binom{2d}{d}.$$

Proposition 11.B.3. *The intersection numbers of the Hamming scheme $\mathfrak{H}(d,s)$ satisfy*

1. $k_i = \binom{d}{i}(s-1)^i$ *for every* $i \in [d]$.
2. *For* $s > 10\gamma^{-1}d^3$ *and* $j+1 \le t \le d$ *we have*

$$(s-2)^{2j-t}\binom{2t-2j}{t-j}\binom{t}{2j-t} \le p^t_{j,j} \le (1+\gamma)(s-2)^{2j-t}\binom{2t-2j}{t-j}\binom{t}{2j-t}. \quad (11.35)$$

3. *For* $s > 10\gamma^{-1}d^3$ *and* $1 \le t \le j$ *we have*

$$(s-1)^{j-t}(s-2)^t\binom{d-t}{j-t} \le p^t_{j,j} \le (1+\gamma)(s-1)^{j-t}(s-2)^t\binom{d-t}{j-t}. \quad (11.36)$$

4. *For* $s > 10\gamma^{-1}d^3$ *for* $2 \le j+1 \le t \le d-1$,

$$p_{j,j}^{t+1} \le \gamma p_{j,j}^{t}. \tag{11.37}$$

5. *Let* $\gamma > 0$. *Then for* $1 \le \max(i,h) \le t \le \min(i+h,d)$ *and* $s > \gamma^{-1}d^2 + 2$,

$$p_{i-1,h}^{t} \le \gamma p_{i,h}^{t}. \tag{11.38}$$

6. *Let* $\gamma \in (0,1/2)$. *Then for* $2 \le j+1 \le t \le d$ *and* $s \ge 3\gamma^{-1}d^2 + 2$,

$$\sum_{i,h=1}^{j} p_{i,h}^{t} \le (1+\gamma) p_{j,j}^{t}. \tag{11.39}$$

7. *Let* $\gamma > 0$. *For* $1 \le t \le j \le d-1$ *and* $s \ge 10\gamma^{-1}d^3$

$$p_{j,j}^{j+1} \le \gamma p_{j,j}^{t}. \tag{11.40}$$

8. *For* $s > 10\gamma^{-1}d^3$ *and* $1 \le t \le \ell \le j$

$$p_{j,j}^{\ell} \le (1+\gamma) p_{j,j}^{t}, \quad \textit{and} \quad (d-1)p_{j,j}^{j} \le (1+\gamma)p_{j,j}^{1}, \quad \textit{if} \;\; 2 \le j \le d-1. \tag{11.41}$$

Proof. Define

$$z_{j,j,t}^{a} = (s-1)^a (s-2)^{2j-t-2a} \binom{d-t}{a} \binom{t}{2j-t-2a} \binom{2t-2j+2a}{t-j+a}. \tag{11.42}$$

We can compute

$$z_{j,j,t}^{a+1} \le z_{j,j,t}^{a} \cdot \frac{s-1}{(s-2)^2} \cdot d \cdot d^2 \cdot 4 \le z_{j,j,t}^{a} \frac{5d^3}{s}. \tag{11.43}$$

2-3. By Eq. (11.43) for $s > 10d^3/\gamma$ we have $z^{a+1}_{j,j,t} \le (\gamma/2) \cdot z^{a}_{j,j,t}$. Therefore,

$$z^{a_1}_{j,j,t} \le p^t_{j,j} \le (1+\gamma) z^{a_1}_{j,j,t}.$$

In the part 2, $a_1 = 0$ and in the part 3 $a_1 = j - t$.

4. By part 2, we have

$$p^t_{j,j}/p^{t+1}_{j,j} \ge (s-2) \cdot \frac{1}{4(1+\gamma)} \cdot \frac{(2t-2j+1)(2t-2j+2)}{(t+1)(2j-t)} \ge \frac{s}{5d^2} \ge \frac{1}{\gamma}. \tag{11.44}$$

5. In the expression (11.34) for $p^t_{i-1,h}$ the bounds are $a_1 = 0$ and $a_2 = \min\left(\left\lfloor \frac{i+h-t-1}{2} \right\rfloor, d-t\right)$, and for $p^t_{i,h}$ the bounds are $a'_1 = 0$ and $a'_2 = \min\left(\left\lfloor \frac{i+h-t}{2} \right\rfloor, d-t\right)$. Therefore, it is sufficient to check that for every $a \le a_2$,

$$\begin{aligned} \gamma(s-2)\binom{t}{i+h-t-2a}\binom{2t-i-h+2a}{t-h+a} \ge \\ \ge \binom{t}{(i-1)+h-t-2a}\binom{2t-(i-1)-h+2a}{t-h+a}. \end{aligned} \tag{11.45}$$

Since $2t-(i-1)-h+2a \le d$, by Observation 11.B.2, this inequality is satisfied if $s \ge \gamma^{-1}d^2 + 2$.

6. Denote by $I_\ell = \{(i,h) : \ 1 \le i,h \le j, \ i+h = \ell\}$. Then,

$$\sum_{i,h=1}^{j} p^t_{i,h} = \sum_{\ell=t}^{2j} \sum_{(i,h)\in I_\ell} p^t_{i,h}.$$

Let $\gamma' = \gamma/3 < 1/6$. Using part 5,

$$\sum_{(i,h)\in I_{\ell-1}} p^t_{i,h} \le 2\gamma' \sum_{(i,h)\in I_\ell} p^t_{i,h} \tag{11.46}$$

Therefore,

$$\sum_{i,h=1}^{j} p_{i,h}^{t} \leq \sum_{\ell=0}^{2j-t} (2\gamma')^{\ell} \sum_{(i,h)\in I_{2j}} p_{i,h}^{t} = \left(\sum_{\ell=0}^{2j-t} (2\gamma')^{\ell} \right) p_{j,j}^{t} \leq \frac{1}{1-2\gamma'} p_{j,j}^{t} \leq (1+3\gamma') p_{j,j}^{t}.$$

7. By parts 2 and 3, we have

$$p_{j,j}^{j+1} \leq (1+\gamma)(s-2)^{j-1} \binom{2}{1} \binom{j+1}{j-1} \leq (1+\gamma)(s-2)^{j-1} d^2, \quad \text{and}$$
$$p_{j,j}^{t} \geq (s-1)^{j-t}(s-2)^{t} \binom{d-t}{j-t} \geq (s-2)^{j}.$$

Hence, the desired inequality holds since $s-2 \geq 2d^2$.

8. For $1 \leq t \leq \ell \leq j \leq d$,

$$p_{j,j}^{\ell} \leq (1+\gamma)(s-1)^{j-\ell}(s-2)^{\ell} \binom{d-\ell}{d-j} \quad \text{and} \quad (s-1)^{j-t}(s-2)^{t} \binom{d-t}{d-j} \leq p_{j,j}^{t},$$

Note that for $t \leq \ell \leq j$ we have $\binom{d-\ell}{d-j} \leq \binom{d-t}{d-j}$, so $p_{j,j}^{\ell} \leq (1+\gamma) p_{j,j}^{t}$. Moreover, if $2 \leq j \leq d-1$, then $(d-1)p_{j,j}^{j} \leq (1+\gamma)p_{j,j}^{1}$, as $\binom{d-1}{d-j} \geq d-1$.

□

11.C Appendix: Inequalities on the intersection numbers of the Grassmann schemes

In this section we derive inequalities for the intersection numbers of the Grassmann schemes. We use these inequalities in Section 11.4.4.

Fact 11.C.1. The intersection numbers of the Grassmann scheme $\mathfrak{J}_q(s,d)$ are

$$p_{i,j}^{t} = \sum_{a=t_1}^{t_2} \binom{d-t}{a}_q \binom{t}{d-a-i}_q \binom{t}{d-a-j}_q \binom{s-d-t}{i+j+a-d}_q, \text{ for } t_1 \leq t_2, \tag{11.47}$$

$$\text{and} \quad p^j_{i,t} = 0, \text{ for } t_1 > t_2,$$

where $t_1 = \max\{d-i-t, d-j-t, d-i-j, 0\}$ and $t_2 = \min\{d-t, d-i, d-j, s-t-i-j\}$.

The degree of the j-th constituent is

$$k_j = q^{j^2} \binom{d}{j}_q \binom{s-d}{j}_q. \tag{11.48}$$

Before we prove the desired inequalities for the intersection numbers it is convenient to make the following observation.

Observation 11.C.2. Let $q \geq 2$, $1 \leq k \leq n$ be integers. Then

$$\frac{1}{2} q^{n-2k+1} \binom{n}{k-1}_q \leq \binom{n}{k}_q \leq 2q^{n-2k+1} \binom{n}{k-1}_q, \quad \text{and}$$
$$\frac{1}{2} q^{n-k} \binom{n-1}{k-1}_q \leq \binom{n}{k}_q \leq 2q^{n-k} \binom{n-1}{k-1}_q.$$

Proof. The proof follows from the fact that

$$\binom{n}{k}_q = \binom{n}{k-1}_q \left(\frac{q^{n-k+1}-1}{q^k-1} \right) \quad \text{and} \quad \binom{n}{k}_q = \binom{n-1}{k-1}_q \left(\frac{q^n-1}{q^k-1} \right). \qquad \square$$

Proposition 11.C.3. *Consider the Grassmann scheme $\mathfrak{J}_q(s,d)$, for $d \geq 2$. Then the following inequalities hold.*

1. *If $s \geq 2d + x$, then $k_{j-1} \leq \dfrac{(q-1)(q^{x+1}-1)}{q} \cdot k_j$.*
2. *For $\gamma \in (0,1)$, $1 \leq t \leq j \leq d$ and $s \geq 6d - 4 + \log_q(32/\gamma)$*

$$\binom{d-t}{d-j}_q \binom{s-d-t}{j}_q \leq p^t_{j,j} \leq (1+\gamma) \binom{d-t}{d-j}_q \binom{s-d-t}{j}_q.$$

3. *For* $\gamma \in (0,1)$, $2 \le j+1 \le t \le d$ *and* $s \ge 6d - 4 + \log_q(32)$.

$$\binom{t}{t-j}_q^2 \binom{s-d-t}{2j-t}_q \le p_{j,j}^t \le (1+\gamma)\binom{t}{t-j}_q^2 \binom{s-d-t}{2j-t}_q.$$

4. *Let* $\gamma > 0$. *Then for* $2 \le t+1 \le j \le d-1$ *and* $s \ge 5d + \log_q(16/\gamma)$

$$p_{t,t}^{j+1} \le \gamma p_{t,t}^j.$$

5. *Let* $\gamma > 0$. *Then for* $1 \le \max(i,h) \le t < \min(i+h,d)$, *and* $s \ge d + \log_q(1/\gamma)$

$$p_{i-1,h}^t \le \gamma p_{i,h}^t.$$

6. *Let* $\gamma \in (0,1)$. *Then for* $2 \le j+1 \le t \le d$ *and* $s \ge d + \log_q(4/\gamma)$

$$\sum_{a,b=1}^{j} p_{a,b}^t \le (1+\gamma) p_{j,j}^t.$$

7. *Let* $\gamma > 0$. *For* $s \ge 6d - 2 + \log_q(32/\gamma)$ *and* $1 \le t \le j \le d-1$

$$p_{j,j}^{j+1} \le \gamma p_{j,j}^t.$$

8. *Let* $\gamma > 0$. *For* $s \ge 6d - 4 + \log_q(32/\gamma)$ *and* $1 \le t \le \ell \le j$

$$p_{j,j}^{\ell} \le (1+\gamma) p_{j,j}^t, \quad \textit{and} \quad [d-1]_q \cdot p_{j,j}^j \le (1+\gamma) p_{j,j}^1, \quad \textit{if} \;\; 2 \le j \le d-1.$$

Proof. Denote the summand with an index a in the sum (11.47) for $p_{i,j}^t$ by

$$z_{i,j,t}^a := \binom{d-t}{a}_q \binom{t}{d-a-i}_q \binom{t}{d-a-j}_q \binom{s-d-t}{i+j+a-d}_q. \tag{11.49}$$

Clearly, $z^a_{i,j,t} \geq 0$. By Observation 11.C.2, for $a \leq d - \max(j,t)$, for

$$y = (d - t - 2a + 1) - 2(t - 2d + 2a + 2j - 1) + (s - d - t - 4j - 2a + 2d + 1) =$$
$$= s + 6d - 4t - 8j - 8a + 4 \geq s - 2d + 8\max(j,t) - 4t - 8j + 4 \geq s - 6d + 4$$

we have

$$z^a_{j,j,t} \geq (q^y/16) \cdot z^{a-1}_{j,j,t} = (q^{s-6d+4}/16) \cdot z^{a-1}_{j,j,t}. \tag{11.50}$$

Additionally, by Observation 11.C.2, for $a \leq d - \max(j,t)$, for

$$w = (d - t - a) - 2(t - d + a + j) + (s - d - t - 2j - a + d) =$$
$$= s + 3d - 4a - 4t - 4j \geq s - d - 4\min(j,t) \geq s - 5d$$

we have

$$z^a_{j,j,t} \geq (q^w/16) \cdot z^{a-1}_{j,j,t+1} = (q^{s-5d}/16) \cdot z^{a-1}_{j,j,t}. \tag{11.51}$$

Also, for convenience, we write down Eq. (11.47) for $i = t$

$$p^t_{j,j} = \sum_{a=t_1}^{t_2} \binom{d-t}{a}_q \binom{t}{d-j-a}_q^2 \binom{s-d-t}{2j+a-d}_q, \tag{11.52}$$

where t_1 and t_2 are defined as in Fact 11.C.1.

1. For $s \geq 2d + x$, we have

$$b_i = q^{2i+1}[d-i]_q[s-d-i]_q \geq q^{2i+1}[x]_q \quad \text{and} \quad c_i = \big([i]_q\big)^2 \leq q^{2i}/(q-1)^2.$$

Hence

$$k_j = \frac{b_{j-1}}{c_j} k_{j-1} \geq q^{2j-1} \frac{(q^{x+1}-1)}{(q-1)} \frac{(q-1)^2}{q^{2j}} k_{j-1} = \frac{(q^{x+1}+1)(q-1)}{q} k_{j-1} \tag{11.53}$$

2. For $t \le j$, in the sum (11.52) for $p^t_{j,j}$, we have $t_1 = \max\{d-t-j, 0\}$, $t_2 = d-j$. Since all $z^a_{j,j,t} \ge 0$,

$$p^t_{j,j} = \sum_{a=d-t-j}^{d-j} z^a_{j,j,t} \ge z^{d-j}_{j,j,t} = \binom{d-t}{d-j}_q \binom{s-d-t}{j}_q. \tag{11.54}$$

Observe that for $s \ge 6d - 4 + \log_q(32/\gamma)$, $\gamma z^a_{j,j,t} \ge 2z^{a-1}_{j,j,t}$ for $0 < a \le d-j$. Hence, the desired upper bound follows.

3. For $t \ge j+1$, in the sum (11.52) for $p^t_{j,j}$ we have $t_1 = \max\{d-2j, 0\}$, $t_2 = d-t$. Thus

$$p^t_{j,j} = \sum_{a=t_1}^{d-t} z^a_{j,j,t} \ge z^{d-j}_{j,j,t} = \binom{t}{t-j}_q^2 \binom{s-d-t}{2j-t}_q.$$

Eq. (11.50) implies that for $s \ge 6d - 4 + \log_q(32/\gamma)$, $\gamma z^a_{j,j,t} \ge 2z^{a-1}_{j,j,t}$ for $0 \le a \le d-t$. Hence, the desired upper bound follows.

4. We need to show that for $t_1 = \max\{d-2j, 0\}$

$$\gamma \sum_{a=t_1}^{d-t} z^a_{j,j,t} = \gamma p^t_{j,j} \ge p^{t+1}_{j,j} = \sum_{a=t_1}^{d-t-1} z^a_{j,j,t+1}.$$

It is enough to show that $\gamma z^a_{j,j,t} \ge z^{a-1}_{j,j,t+1}$. This inequality follows from Eq. (11.51) for $s \ge 5d + \log_q(16/\gamma)$.

5. Since $\max(i,h) \le t \le i+h-1$ we need to verify

$$\sum_{a=d-h-i+1}^{d-t} z^a_{i-1,h,t} \le \gamma \sum_{a=d-h-i}^{d-t} z^a_{i,h,t}.$$

It is enough to check that $z^a_{i-1,h,t} \le \gamma z^a_{i,h,t}$. By Obs. 11.C.2, for $a \le d - \max(i,h,t)$,

for

$$y = (s - d - t - 2(i + h + a - d) + 1) - (t - 2(d - a - i) - 1) =$$

$$= s + 3d - 2t - 2h - 4i - 4a + 2 \geq s - d + 2$$

we have

$$z^a_{i,h,t} \geq (q^y/4)\, z^a_{i-1,h,t} \geq \left(q^{s-d+2}/4\right) z^a_{i-1,h,t}. \tag{11.55}$$

Thus, $z^a_{i-1,h,t} \leq \gamma z^a_{i,h,t}$ holds for $s \geq d + \log_q(1/\gamma)$.

6. Denote by $I_h = \{(a,b) \ : \ 1 \leq a \leq j, \ 1 \leq b \leq j, \ a + b = h\}$. We can write

$$\sum_{a,b=1}^{j} p^t_{a,b} = \sum_{h=t}^{2j} \sum_{(a,b)\in I_h} p^t_{a,b}.$$

Denote $\gamma_0 = \gamma/(2 + 2\gamma)$. Since $s \geq d + \log_q(1/\gamma_0)$, using previous part, we deduce

$$\sum_{(a,b)\in I_{h-1}} p^t_{a,b} \leq 2\gamma \sum_{(a,b)\in I_h} p^t_{a,b}.$$

Therefore,

$$\sum_{a,b=1}^{j} p^t_{a,b} \leq \sum_{h=0}^{2j-t} (2\gamma_0)^h \sum_{(a,b)\in I_{2j}} p^t_{a,b} \leq \sum_{h=0}^{\infty} (2\gamma_0)^h p^t_{j,j} = \frac{1}{1 - 2\gamma_0} p^t_{j,j} = (1 + \gamma) p^t_{j,j}.$$

7. By part 2, for $t \leq j$, and by Observation 11.C.2,

$$p^t_{j,j} \geq \binom{s-d-t}{j}_q \geq \binom{s-d-j}{j}_q \geq \frac{1}{2} q^{s-d-2j} \binom{s-d-j-1}{j-1}_q.$$

Part 3 implies

$$p_{j,j}^{j+1} \leq 2\binom{j+1}{1}_q^2 \binom{s-d-j-1}{j-1}_q \leq 8q^{2j}\binom{s-d-j-1}{j-1}_q.$$

Therefore, since $s \geq 5d + \log_q(16/\gamma)$ we get $p_{j,j}^{j+1} \leq \gamma p_{j,j}^t$.

8. By part 2,

$$p_{j,j}^{\ell} \leq (1+\gamma)\binom{d-\ell}{d-j}_q \binom{s-d-\ell}{j}_q \quad \text{and} \quad \binom{d-t}{d-j}_q \binom{s-d-t}{j}_q \leq p_{j,j}^t.$$

Thus, for $t \leq \ell$ we get $p_{j,j}^{\ell} \leq (1+\gamma)p_{j,j}^t$. Moreover, for $d-1 \geq j \geq 2$ we have

$$\binom{d-1}{d-j}_q \geq \binom{d-1}{1}_q = [d-1]_q = q^{d-2} + \ldots + q + 1. \tag{11.56}$$

Thus, for $d-1 \geq j \geq 2$, we obtain $[d-1]_q \cdot p_{j,j}^j \leq (1+\gamma)p_{j,j}^1$.

□

REFERENCES

Michael Aschbacher and Leonard Scott. Maximal subgroups of finite groups. *J. Algebra*, 92 (1):44–80, 1985. doi: 10.1016/0021-8693(85)90145-0.

László Babai. On the Order of Uniprimitive Permutation Groups. *Ann. Math.*, 113(3):553, 1981. doi: 10.2307/2006997.

László Babai. On the order of doubly transitive permutation groups. *Inventiones Mathematicae*, 65(3):473–484, 1982. doi: 10.1007/BF01396631.

László Babai. On the automorphism groups of strongly regular graphs I. In *Proc. 5th Conf. on Innov. in Th. Comp. Sc. (ITCS'14)*, pages 359–368, 2014. doi: 10.1145/2554797.2554830.

László Babai. On the automorphism groups of strongly regular graphs II. *J. Algebra*, 421: 359–368, 2015. doi: 10.1016/j.jalgebra.2014.09.007.

László Babai. Graph isomorphism in quasipolynomial time. In *Proc. 48th ACM Sym. on Theory of Computing (STOC'16)*, pages 684–697. ACM Press, 2016a. doi: 10.1145/2897518.2897542.

László Babai. Graph isomorphism in quasipolynomial time. *arXiv:1512.03547*, 2016b. 89 pages.

László Babai and Bohdan Kivva. On the motion of distance-transitive graphs. *Manuscript*, 2020.

László Babai and Bohdan Kivva. Robustness of the Johnson scheme under fusion and extension. *Manuscript*, 2022.

László Babai and Vera T. Sós. Sidon sets in groups and induced subgraphs of Cayley graphs. *Europ. J. Combinat.*, 6(2):101–114, 1985. doi: 10.1016/S0195-6698(85)80001-9.

László Babai, Peter J. Cameron, and Péter Pál Pálfy. On the orders of primitive groups with restricted nonabelian composition factors. *J. Algebra*, 79(1):161–168, 1982. doi: 10.1016/0021-8693(82)90323-4.

Mitali Bafna, Boaz Barak, Pravesh Kothari, Tselil Schramm, and David Steurer. Playing Unique Games on Certified Small-Set Expanders, 2020.

Sejeong Bang. Geometric distance-regular graphs without 4-claws. *Linear Algebra Appl.*, 438(1):37–46, 2013. doi: 10.1016/j.laa.2012.07.021.

Sejeong Bang. Diameter bounds for geometric distance-regular graphs. *Discrete Mathematics*, 341(1):253–260, 2018. doi: 10.1016/j.disc.2017.08.036.

Sejeong Bang and Jack H. Koolen. On geometric distance-regular graphs with diameter three. *European J. Comb.*, 36:331–341, 2014. doi: 10.1016/j.ejc.2013.06.044.

Sejeong Bang, Akira Hiraki, and Jack H. Koolen. Delsarte clique graphs. *European J. Comb.*, 28(2):501–516, 2007. doi: 10.1016/j.ejc.2005.04.015.

Sejeong Bang, Arturas Dubickas, Jack H. Koolen, and Vincent Moulton. There are only finitely many distance-regular graphs of fixed valency greater than two. *Adv. Math.*, 269: 1–55, 2015. doi: 10.1016/j.aim.2014.09.025.

Norman L. Biggs. Intersection matrices for linear graphs. *Comb. Math. and Appl.*, pages 15–23, 1971.

Norman L. Biggs and Anthony Gardiner. The classification of distance transitive graphs. Manuscript, 1974.

Alfred Bochert. Ueber die Classe der transitiven Substitutionengruppen. *Math. Annalen*, 40 (2):176–193, 1892. doi: 10.1007/BF01443562.

Ray C. Bose and T Shimamoto. Classification and Analysis of Partially Balanced Incomplete Block Designs with Two Associate Classes. *J. Amer. Stat. Assoc.*, 47(258):151–184, 1952. doi: 10.1080/01621459.1952.10501161.

Andries E. Brouwer and Jack H. Koolen. The vertex-connectivity of a distance-regular graph. *European J. Comb.*, 30(3):668–673, 2009. doi: 10.1016/j.ejc.2008.07.006.

Andries E. Brouwer, Arjeh M. Cohen, and Arnold Neumaier. *Distance-Regular Graphs*. Springer, 1989. doi: 10.1007/978-3-642-74341-2.

Frans C. Bussemaker and Arnold Neumaier. Exceptional Graphs with Smallest Eigenvalue -2 and Related Problems. *Math. of Computation*, 59(200):583, 1992. doi: 10.2307/2153076.

Peter J. Cameron. Finite Permutation Groups and Finite Simple Groups. *Bull. London Math. Soc.*, 13(1):1–22, 1981. doi: 10.1112/blms/13.1.1.

Peter J. Cameron, Jean-Marie Goethals, Johan Jacob Seidel, and Ernest E. Shult. Line graphs, root systems, and elliptic geometry. *J. Algebra*, pages 208–230, 1991. doi: 10.1016/b978-0-12-189420-7.50021-9.

Xi Chen, Xiaorui Sun, and Shang-Hua Teng. Multi-stage design for quasipolynomial-time isomorphism testing of steiner 2-systems. In *Proc. 45th ACM Symp. Theory of Computing (STOC '13)*, page 271. ACM Press, 2013. doi: 10.1145/2488608.2488643.

Henry Cohn and Christopher Umans. A group-theoretic approach to fast matrix multiplication. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 438–449. IEEE, 2003.

Henry Cohn and Christopher Umans. Fast matrix multiplication using coherent configurations. In *Proc. 24th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'13)*, pages 1074–1087, 2013. doi: 10.1137/1.9781611973105.77.

Charles W. Curtis, William M. Kantor, and Gary M. Seitz. The 2-Transitive Permutation Representations of the Finite Chevalley Groups. *Trans. of AMS*, 218:1, 1976. doi: 10.2307/1997427.

Philippe Delsarte. An algebraic approach to the association schemes of coding theory. *Philips Res. Reports Suppls.*, 10:vi+–97, 1973.

Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. On non-optimally expanding sets in grassmann graphs. *Israel Journal of Mathematics*, 243(1):377–420, 2021.

Michael Doob and Dragoš Cvetković. On spectral characterizations and embeddings of graphs. *Linear Algebra Appl.*, 27:17–26, 1979. doi: 10.1016/0024-3795(79)90028-4.

Sean Eberhard. Hamming sandwiches. *arXiv preprint arXiv:2203.03687*, pages 1–19, 2022.

Yoshimi Egawa. Characterization of H(n, q) by the parameters. *J. Comb. Theory, Ser. A*, 31(2):108–125, 1981. doi: 10.1016/0097-3165(81)90007-8.

Hikoe Enomoto. Characterization of families of finite permutation groups by the subdegrees. II. *J. of the Faculty of Science, the Univ. of Tokyo. Sect. 1 A*, 20(1):2–11, 1973.

Paul Erdős. Beweis eines satzes von tschebyschef. *Acta Litt. Sci. Szeged*, 5:194–198, 1932. Rényi Inst.

Paul Erdős and Pál Turán. On a problem of Sidon in additive number theory, and on some related problems. *J. London Math. Soc.*, s1-16(4):212–215, 1941. doi: 10.1112/jlms/s1-16.4.212.

Walter Feit and Graham Higman. The nonexistence of certain generalized polygons. *J. Algebra*, 1(2):114–131, 1964. doi: 10.1016/0021-8693(64)90028-6.

Dmitry G. Fon-Der-Flaass. Embedding arbitrary graphs into strongly regular and distance-regular graphs. *Siberian Electron. Math. Reports*, 2:218–221, 2005. (link).

A Gavrilyuk, Wenbin Guo, and A Makhnev. Automorphisms of Terwilliger graphs with $\mu = 2$. *Algebra & Logic*, 47(5), 2008.

Alexander L Gavrilyuk. On the Koolen–Park inequality and Terwilliger graphs. *Electronic J. Comb.*, pages R125–R125, 2010. doi: 10.37236/397.

Chris D. Godsil. *Algebraic Combinatorics*. Chapman and Hall Math. Ser., 1993a. doi: 10.1201/9781315137131.

Chris D. Godsil. Geometric distance-regular covers. *New Zealand J. Math.*, 22:31–38, 1993b. (link).

Willem H. Haemers. Interlacing eigenvalues and graphs. *Linear Algebra Appl.*, 226-228(228): 593–616, 1995. doi: 10.1016/0024-3795(95)00199-2.

Donald G. Higman. Intersection matrices for finite permutation groups. *J. Algebra*, 6:22–42, 1967. 10.1016/0021-8693(67)90011-7.

Donald G. Higman. Coherent configurations I. *Rend. Sem. Mat. Univ. Padova*, 44:1–25, 1970. (link).

Alan J. Hoffman. The eigenvalues of the adjacency matrix of a graph. *Research note, Thomas J. Watson Research Center, Yorktown Heights, New York*, page 698, 1967.

Alan J. Hoffman. On eigenvalues and colourings of graphs. *Graph theory and appl.*, pages 79–91, 1970a. doi: 10.1142/9789812796936_0041.

Alan J. Hoffman. $-1-\sqrt{2}$? *Comb. Struct. and Appl.*, pages 173–176, 1970b.

Alan J. Hoffman. On graphs whose least eigenvalue exceeds $-1-\sqrt{2}$. *Linear Algebra Appl.*, 16(2):153–165, 1977. doi: 10.1016/0024-3795(77)90027-1.

Max Hopkins, Tali Kaufman, and Shachar Lovett. High dimensional expanders: Random walks, pseudorandomness, and unique games. *arXiv preprint arXiv:2011.04658*, 2020.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge university press, Cambridge, 2012. doi: 10.1017/cbo9781139020411.

Robert Jajcay and Dale Mesner. Embedding arbitrary finite simple graphs into small strongly regular graphs. *J. Graph Theory*, 34(1):1–8, 2000. doi: 10.1002/(SICI)1097-0118(200005)34:1<1::AID-JGT1>3.0.CO;2-E.

Camille Jordan. Théorèmes sur les groupes primitifs. *J. Math Pures Appl.*, 17:351–367, 1871. (link).

Lev Arkad'evich Kaluzhnin and Mikhail Klin. On certain maximal subgroups of symmetric and alternating groups. *Mat. Sbornik*, 129(1):91–121, 1972. doi: 10.1070/SM1972v016n01ABEH001352, (link, in Russian).

Subhash Khot, Dor Minzer, Dana Moshkovitz, and Muli Safra. Small Set Expansion In The Johnson Graph. *ECCC*, 25:78, 2018. (link).

Bohdan Kivva. On the automorphism groups of rank-4 primitive coherent configurations. *arXiv preprint arXiv:2110.13861*, 2021a.

Bohdan Kivva. On the spectral gap and the automorphism group of distance-regular graphs. *J. Comb. Theory, Ser. B*, 149:161–197, 2021b. doi: 10.1016/j.jctb.2021.02.003.

Bohdan Kivva. A characterization of Johnson and Hamming graphs and proof of Babai's conjecture. *J. Comb. Theory, Ser. B*, 151:339–374, 2021c. doi: 10.1016/j.jctb.2021.07.003.

Bohdan Kivva. Robustness of the Hamming and the Grassmann schemes under fusion and extension. *Manuscript*, 2022.

Mikhail Klin. *Investigation of algebras of invariant relations for certain classes of permutation groups*. PhD thesis, Kyiv State University, 1974. (in Russian).

Jack H. Koolen and Sejeong Bang. On distance-regular graphs with smallest eigenvalue at least $-m$. *J. Comb. Theory, Ser B*, 100(6):573–584, 2010. doi: 10.1016/j.jctb.2010.04.006.

Martin W. Liebeck. On minimal degrees and base sizes of primitive permutation groups. *Archiv der Mathematik*, 43(1):11–15, 1984. doi: 10.1007/BF01193603.

Martin W. Liebeck and Jan Saxl. Minimal Degrees of Primitive Permutation Groups, with an Application to Monodromy Groups of Covers of Riemann Surfaces. *Proc. London Math. Soc.*, s3-63(2):266–314, 1991. doi: 10.1112/plms/s3-63.2.266.

Martin W. Liebeck, Cheryl E. Praeger, and Jan Saxl. Distance transitive graphs with symmetric or alternating automorphism group. *Bull. Aust. Math. Soc.*, 35(1):1–25, 1987. doi: 10.1017/S0004972700012995.

Attila Maróti. On the orders of primitive groups. *J. Algebra*, 258(2):631–640, 2002. doi: 10.1016/S0021-8693(02)00646-4.

Raghu Meka, Aaron Potechin, and Avi Wigderson. Sum-of-Squares lower bounds for Planted Clique. In *Proc. 47th ACM Sym. on Theory of Computing (STOC'15)*, pages 87–96, 2015. doi: 10.1145/2746539.2746600.

Klaus Metsch. A characterization of Grassmann graphs. *European J. Comb.*, 16(6):639–644, 1995. 10.1016/0195-6698(95)90045-4.

Bojan Mohar and John Shawe-Taylor. Distance-biregular graphs with 2-valent vertices and distance-regular line graphs. *J. Comb. Theory, Ser B*, 38(3):193–203, 1985. doi: 10.1016/0095-8956(85)90065-6.

Mikhail Muzychuk. Subschemes of the Johnson scheme. *European J. Comb.*, 13:187–193, 1992a. doi: 10.1016/0195-6698(92)90023-S.

Mikhail Muzychuk. Subschemes of Hamming association schemes $H(n, q)$, $q \geq 4$. *Acta Applicandae Mathematica*, 29(1-2):119–128, 1992b. doi: 10.1007/BF00053381.

Mikhail Muzychuk. The Subschemes of the Hamming Scheme. *Investigations in Alg. Theory of Comb. Objects (Soviet Series)*, 84, 1995. doi: 10.1007/978-94-017-1972-8_4.

Arnold Neumaier. Strongly regular graphs with smallest eigenvalue $-m$. *Archiv der Mathematik*, 33(1):392–400, 1979. doi: 10.1007/BF01222774.

Alexander M. Ostrowski. *Solution of Equations and Systems of Equations*, volume 21 of *Pure and Applied Mathematics*. Elsevier, 1967. doi: 10.2307/2005025.

Jongyook Park, Jack H. Koolen, and Greg Markowsky. There are only finitely many distance-regular graphs with valency k at least three, fixed ratio $k2/k$ and large diameter. *J. Comb. Theory, Ser. B*, 103(6):733–741, 2013. doi: 10.1016/j.jctb.2013.09.005.

Cheryl E. Praeger, Jan Saxl, and Kazuhiro Yokoyama. Distance transitive graphs and finite simple groups. *Proc. London Math. Soc.*, 3(1):1–21, 1987. doi: 10.1112/plms/s3-55.1.1.

László Pyber. On the orders of doubly transitive permutation groups, elementary estimates. *J. Comb. Theory, Ser A*, 62(2):361–366, 1993a. doi: 10.1016/0097-3165(93)90053-B.

László Pyber. Asymptotic results for permutation groups. *DIMACS series in Discrete Math. and Theoretical Computer Science*, 11:197–219, 1993b.

Zhi Qiao, Jack H. Koolen, and Greg Markowsky. On the Cheeger constant for distance-regular graphs. *J. Comb. Theory, Ser. A*, 173:105–227, 2020. doi: 10.1016/j.jcta.2020.105227.

Dijen K. Ray-Chaudhuri and Alan P. Sprague. Characterization of projective incidence structures. *Geometriae Dedicata*, 5(3):361–376, 1976.

Alexander Russell and Ravi Sundaram. A Note on the Asymptotics and Computational Complexity of Graph Distinguishability. *Electronic J. Comb.*, 5(1):23–30, 1998. doi: 10.37236/1361.

Issai Schur. Zur Theorie der einfach transitiven Permutationsgruppen. *Sitzungsb. Preuss. Akad. Wiss., Phys.–Math. Kl.*, pages 598–623, 1933. In: Issai Schur, Gesammelte Abhandlungen, Vol. III (Alfred Brauer and Hans Rohrbach, eds.), Springer, 1973.

Leonard L Scott. Representations in characteristic p. In *The Santa Cruz Conference on Finite Groups (Univ. California, Santa Cruz, Calif., 1979)*, volume 37, pages 319–331. Amer. Math. Soc. Providence, 1980.

Johan Jacob Seidel. Strongly regular graphs with (—1, 1, 0) adjacency matrix having eigenvalue 3. *Linear Algebra Appl.*, 1(2):281–298, 1968. doi: 10.1016/B978-0-12-189420-7.50010-4.

Simon Sidon. Ein Satz über trigonometrische Polynome und seine Anwendung in der Theorie der Fourier-Reihen. *Mathematische Annalen*, 106(1):536–539, 1932. doi: 10.1007/BF01455900.

Charles C Sims. On graphs with rank 3 automorphism groups. *Manuscript*, 1968.

James Singer. A theorem in finite projective geometry and some applications to number theory. *Trans. Amer. Math. Soc.*, 43(3):377–385, 1938. doi: 10.1090/S0002-9947-1938-1501951-4.

Derek H. Smith. Primitive and imprimitive graphs. *Quarterly J. Math.*, 22(4):551–557, 1971. 10.1093/qmath/22.4.551.

Xiaorui Sun and John Wilmes. Structure and automorphisms of primitive coherent configurations. *arXiv:1510.02195*, 2015a.

Xiaorui Sun and John Wilmes. Faster canonical forms for primitive coherent configurations. In *Proc. 47th ACM Sym. on Theory of Computing (STOC'15)*, pages 693–702, 2015b. doi: 10.1145/2746539.2746617.

Paul Terwilliger. Distance-regular graphs with girth 3 or 4: I. *J. Comb. Theory, Ser. B*, 39 (3):265–281, 1985. doi: 10.1016/0095-8956(85)90054-1.

Paul Terwilliger. A new feasibility condition for distance-regular graphs. *Discrete Math.*, 61 (2-3):311–315, 1986. doi: 10.1016/0012-365X(86)90102-0.

Daiyu Uchida. On the subschemes of the Johnson scheme $J(v, d)$. *Memoirs of the Faculty of Science, Kyushu University Ser. A*, 46(1):85–92, 1992. (link).

John van Bon and Andries E. Brouwer. The distance-regular antipodal covers of classical distance-regular graphs. In *Colloq. Math. Soc. János Bolyai, Proc. Eger, 1987*, volume 1988, pages 141–166, 1988. (link).

Edwin R. van Dam, Jack H. Koolen, and Hajime Tanaka. Distance-Regular Graphs. *Electronic J. Comb.*, DS22, 2016. doi: 10.37236/4925.

Boris Weisfeiler. *On Construction and Identification of Graphs*, volume 558 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, 1976. doi: 10.1007/BFb0089374.

Boris Weisfeiler and Andrey A. Leman. A reduction of a graph to a canonical form and an algebra arising during this reduction. *Nauchno-Technicheskaya Informatsia*, 2(9):12–16, 1968.

Hassler Whitney. Congruent Graphs and the Connectivity of Graphs. In *Hassler Whitney Collected Papers*, pages 61–79. Birkhäuser Boston, 1992. doi: 10.1007/978-1-4612-2972-8_4.

Helmut Wielandt. *Abschätzungen für den Grad einer Permutationsgruppe von vorgeschriebenem Transitivitätsgrad*. Dissertation, Berlin, 1934. doi: 10.1515/9783110863383.23.

Helmut Wielandt. *Finite Permutation Groups*. Academic Press, 1964.