

Synchronization and separation in the Johnson schemes

Mohammed Aljohani^a, John Bamberg^b and Peter J. Cameron^{a*}

^a School of Mathematics and Statistics, University of St Andrews, North Haugh, St Andrews,
Fife KY16 9SS, U.K.

^b School of Mathematics and Statistics, University of Western Australia, Crawley, Perth WA
6009, Australia

Abstract

Recently Peter Keevash solved asymptotically the existence question for Steiner systems by showing that $S(t, k, n)$ exists whenever the necessary divisibility conditions on the parameters are satisfied and n is sufficiently large in terms of k and t . The purpose of this paper is to make a conjecture which if true would be a significant extension of Keevash's theorem, and to give some theoretical and computational evidence for the conjecture.

We phrase the conjecture in terms of the notions (which we define here) of synchronization and separation for association schemes. These definitions are based on those for permutation groups which grow out of the theory of synchronization in finite automata. In this theory, two classes of permutation groups (called *synchronizing* and *separating*) lying between primitive and 2-homogeneous are defined. A big open question is how the permutation group induced by S_n on k -subsets of $\{1, \dots, n\}$ fits in this hierarchy; our conjecture would give a solution to this problem for n large in terms of k . We prove the conjecture in the case $k = 4$: our result asserts that S_n acting on 4-sets is separating for $n \geq 10$ (it fails to be synchronizing for $n = 9$).

MSC classification: Primary 20B15; secondary 05E30, 20M35

1 The definitions

The concepts of synchronization and separation were defined for permutation groups about ten years ago (see [1]). We outline the definitions here and extend them to association schemes.

A (finite-state deterministic) automaton is a machine which can be in one of a finite set Ω of internal states. On reading a letter from the alphabet associated with the automaton, it undergoes a change of state. If the machine reads a word (a finite sequence of letters), the corresponding state changes are composed.

*Corresponding author: [pj20@st-andrews.ac.uk](mailto:pjc20@st-andrews.ac.uk)

An automaton can also be regarded as a directed graph whose arcs are labelled with the letters in the alphabet, having the property that there is a unique arc with each label leaving any vertex. (Loops and multiple arcs are allowed.) Alternatively, it can be regarded as a transformation semigroup on the set Ω with a distinguished set of generators (the transformations which correspond to single letters).

An automaton is said to be *synchronizing* if there is a word in its alphabet (called a *reset word*) such that, after reading this word, the automaton is in a known state, regardless of its starting state. In the semigroup interpretation, an automaton is synchronizing if the semigroup contains a transformation of rank 1 (one whose image consists of a single element).

Clearly it is not possible for a permutation group to be synchronizing if $|\Omega| > 1$, since any composition of permutations is a permutation. So we abuse terminology and say that a permutation group G on Ω is *synchronizing* if, for any transformation a of Ω which is not a permutation, the semigroup $\langle G, a \rangle$ is a synchronizing semigroup. (In other words, the automaton whose transitions are generators of G together with an arbitrary non-permutation is synchronizing.)

There is a combinatorial characterisation as follows. A G -*transversal* for a partition P of Ω is a set A , all of whose images under G are transversals for P . Now the group G is non-synchronizing if there is a nontrivial partition for which all images of some set A are G -transversals. (The trivial partitions are the partition into singletons and the partition with a single part.)

This can also be formulated in terms of simple graphs (loopless and without multiple edges). The *clique number* of a graph is the number of vertices in a largest complete subgraph, and the *chromatic number* is the least number of colours required to colour the vertices so that the ends of any edge receive different colours. An *endomorphism* of a graph is a transformation on the vertex set of the graph which maps edges to edges. The endomorphisms of a graph Γ form a transformation semigroup $\text{End}(\Gamma)$. The following theorem can be found in [1].

Theorem 1.1.

- (a) A transformation semigroup S is non-synchronizing if and only if there exists a non-null graph Γ such that $S \leq \text{End}(\Gamma)$ and the clique number and chromatic number of Γ are equal.
- (b) A permutation group G is non-synchronizing if and only if there exists a graph Γ , neither null nor complete, such that $G \leq \text{Aut}(\Gamma)$ and the clique number and chromatic number of Γ are equal.

In terms of the previous characterisation, the partition of Ω is given by a colouring of Γ with the minimum number of colours, and a G -transversal is a clique with size equal to this number.

A related concept is also defined in [1]. A transitive permutation group G on Ω is *non-separating* if there exist subsets A and B of Ω with $|A|, |B| > 1$ such that $|Ag \cap B| = 1$ for all $g \in G$ (this entails $|A| \cdot |B| = |\Omega|$); it is *separating* otherwise. Again, it is known that a permutation group G is non-separating if and only if there is a graph Γ , neither null nor complete, such that the product of its clique number and its coclique number

(the size of the largest coclique) is $|\Omega|$, and with $G \leq \text{Aut}(\Gamma)$. (In the definition, we take A and B to be a clique and a coclique of maximum size.)

It is easy to see that a separating group is synchronizing. (If G is transitive and non-synchronizing, take P to be a nontrivial partition and A a G -transversal for P ; then $|Ag \cap B| = 1$ and $|A| \cdot |B| = |\Omega|$ for every part B of P .) The converse is false, but examples are not easy to come by.

The following theorems are proved in [1].

Theorem 1.2. *Let $n \geq 5$, and let G be the permutation group induced by the symmetric group S_n on the set of 2-element subsets of $\{1, \dots, n\}$. Then the following are equivalent:*

- (a) G is synchronizing;
- (b) G is separating;
- (c) n is odd.

Theorem 1.3. *Let $n \geq 7$, and let G be the permutation group induced by the symmetric group S_n on the set of 3-element subsets of $\{1, \dots, n\}$. Then the following are equivalent:*

- (a) G is synchronizing;
- (b) G is separating;
- (c) n is congruent to 2, 4 or 5 (mod 6) and $n > 8$.

The aim of this paper is to extend these results to larger values of k ; we achieve a complete result for $k = 4$ (using methods based on the work of Delsarte on association schemes, which will hopefully extend to larger values of k) and a conjecture about the situation when n is sufficiently large in terms of k .

2 Steiner systems and the main conjecture

An *association scheme* is a collection $\{A_0, \dots, A_s\}$ of symmetric zero-one matrices, with $A_0 = I$, whose sum is the all-1 matrix, and having the property that the linear span (over the real numbers) of the matrices is an algebra (closed under matrix multiplication), called the *Bose–Mesner algebra* of the scheme. See [2, 7] for further details. (Note that Delsarte uses a slightly more general definition, but his important examples fit the definition given here.)

Given an association scheme as above, each matrix A_i for $i > 0$ is the adjacency matrix of a graph, as indeed are the sums of some of these matrices. For $I \subseteq \{1, \dots, s\}$, we denote by Γ_I the graph whose adjacency matrix is $\sum_{i \in I} A_i$. Such a graph will be called *non-trivial* if it is neither complete nor null, that is, if $I \neq \emptyset$ and $I \neq \{1, \dots, s\}$. Note that the complement of the graph Γ_I is $\Gamma_{\{1, \dots, s\} \setminus I}$. An important property of graphs in an association scheme is that, in any such graph, the product of the clique number and the coclique number is at most the number of vertices. (They share this property with vertex-transitive graphs; neither class of graphs includes the other.)

Adapting the definitions from permutation group theory given above, we say that an association scheme is *non-synchronizing* if there is a non-trivial graph in the scheme

with clique number equal to chromatic number, and is *synchronizing* otherwise; moreover, the scheme is *non-separating* if there is a non-trivial graph in the scheme such that the product of its clique number and its coclique number is equal to the number of vertices, and is *separating* otherwise.

In this paper we are concerned with the *Johnson scheme*. This is the association scheme $J(n, k)$ whose vertices are the k -element subsets of an n -element set (without loss of generality $\{1, \dots, n\}$); for each i with $0 \leq i \leq k$, we take the matrix with rows and columns indexed by the vertices, having (A, B) entry 1 if $|A \cap B| = i$, and 0 otherwise. It is convenient to change the order of the indices by calling this matrix A_{k-i} . Then A_0 is the identity matrix, and the other matrices in the scheme are A_1, \dots, A_k . For $I \subseteq \{1, \dots, k\}$, we let $A_I = \sum_{i \in I} A_i$, and let Γ_I be the graph with adjacency matrix A_I .

While this numbering seems a little odd, our notation is chosen to agree with that of Delsarte [7], who considered this scheme in detail and gave the eigenvalues of the matrices in the scheme in terms of Eberlein polynomials, as we will discuss.

Two particular cases will be very important, so we introduce a different notation for them:

- $\Delta_t(n, k) = \Gamma_{\{k-t+1, \dots, k\}}(n, k)$, the graph in which two k -sets are joined if they intersect in fewer than t points;
- $\Phi_t(n, k) = \Gamma_{\{1, \dots, k-t\}}(n, k)$, the complement of $\Delta_t(n, k)$, in which two k -sets are joined if they intersect in at least t points.

A *Steiner system* $S(t, k, n)$, for $0 < t < k < n$, is a collection \mathcal{B} of k -subsets of $\{1, \dots, n\}$ with the property that any t -subset of $\{1, \dots, n\}$ is contained in a unique member of \mathcal{B} . It is well-known that, for $0 \leq i \leq t-1$, the number of members of \mathcal{B} containing an i -subset of $\{1, \dots, n\}$ is $\binom{n-i}{t-i} / \binom{k-i}{t-i}$, independent of the choice of i -set. Thus necessary conditions for the existence of $S(t, k, n)$ are that

$$\binom{k-i}{t-i} \text{ divides } \binom{n-i}{t-i} \text{ for } 0 \leq i \leq t-1.$$

We refer to these as the *divisibility conditions*. In a remarkable recent result, Keevash [15] showed:

Theorem 2.1. *There exists a function F such that, if $n \geq F(t, k)$ and the divisibility conditions are satisfied, then a Steiner system $S(t, k, n)$ exists.*

The connection with what went before is that the set of blocks of a Steiner system $S(t, k, n)$ is a clique in the graph $\Delta_t(n, k)$, of size $\binom{n}{t} / \binom{k}{t}$. Moreover, there is a coclique in this graph of size $\binom{n-t}{k-t}$, consisting of all the k -sets containing a fixed t -set. (We say that such a coclique is of *EKR type*, after the theorem of Erdős, Ko and Rado, asserting that they are the largest cocliques provided that n is sufficiently large.)

Moreover, it is easily checked that

$$\binom{n-t}{k-t} \cdot \binom{n}{t} / \binom{k}{t} = \binom{n}{k}.$$

So, if a Steiner system $S(t, k, n)$ exists, then the product of the clique number and coclique number in $\Delta_t(n, k)$ is equal to the number of vertices, and the Johnson scheme $J(n, k)$ is non-separating.

Our main conjecture is that, asymptotically, the converse holds:

Conjecture 2.2. There is a function G such that, if $n \geq G(k)$ and the Johnson scheme $J(n, k)$ is non-separating, then there exists a Steiner system $S(t, k, n)$ for some t with $0 < t < k$.

Putting this conjecture together with Keevash's theorem, we can re-formulate it as follows:

Conjecture 2.3. There is a function H such that, if $n \geq H(k)$, then the Johnson scheme is non-separating if and only if the divisibility conditions for $S(t, k, n)$ are satisfied for some t with $0 < t < k$.

What about synchronization? We first observe that, for sufficiently large n , the graph $\Delta_t(k, n)$ cannot have clique number equal to chromatic number. First we note the exact bound in the Erdős–Ko–Rado theorem, proved by Wilson [22]:

Theorem 2.4. Suppose that $n > (t+1)(k-t+1)$. Then a coclique in the graph $\Delta_t(n, k)$ has size at most $\binom{n-t}{k-t}$, with equality if and only if it is of EKR type.

Theorem 2.5. For $t < k$, there is no partition of the set of k -subsets of an n -set into cocliques of EKR type in the graph $\Delta_t(n, k)$.

Proof. First we note that, if $k \geq 2t$, then any two t -sets are contained in some k -set; so sets of EKR type intersect. So we may assume that $k < 2t$.

A set S of EKR type consists of all the k -sets containing a fixed t -set T , which we call its *kernel*; we denote S by $S(T)$. Now, if T_1 and T_2 are t -sets with $|T_1 \cap T_2| \geq 2t - k$, then $|T_1 \cup T_2| \leq k$, and so $S(T_1) \cap S(T_2) \neq \emptyset$. So, in a family of pairwise disjoint sets of EKR type, the kernels are t -sets which intersect in at most $2t - k - 1$ points.

The number of kernels in such a collection is at most $\binom{n}{2t-k} / \binom{t}{2t-k}$; so we are done if we can show that

$$\binom{n}{t} / \binom{k}{t} > \binom{n}{2t-k} / \binom{t}{2t-k}.$$

For this, it is enough to show that the following claims are true:

- (i) $\binom{n}{2t-k} \frac{t!(n-2t+k)!}{k!(n-t)!} = \binom{n}{t} \binom{t}{2t-k} / \binom{k}{t}$;
- (ii) $(t!(n-2t+k)!)/(k!(n-t)!) > 1$.

For (i), we have

$$\begin{aligned} \binom{n}{t} \binom{t}{2t-k} / \binom{k}{t} &= \frac{n!}{(n-2t+k)!(2t-k)!} \frac{t!(n-2t+k)!}{k!(n-t)!} \\ &= \binom{n}{2t-k} \frac{t!(n-2t+k)!}{k!(n-t)!}. \end{aligned}$$

For (ii),

$$\frac{t!(n-2t+k)!}{k!(n-t)!} = \frac{(n-2t+k)(n-2t+k-1)\cdots(n-2t+k-(k-t)+1)}{k(k-1)\cdots(k-(k-t)+1)}$$

And since $n-2t+k > k$ for $n \geq 2k$ we have that

$$\frac{t!(n-2t+k)!}{k!(n-t)!} > 1. \quad \square$$

We also have to consider the possibility that the complementary graph $\Phi_t(n, k)$ has clique number equal to chromatic number.

The existence of cliques of size $\binom{n-t}{k-t}$ in $\Phi_t(n, k)$ shows that the size of a coclique in this graph is at most $\binom{n}{t}/\binom{k}{t}$, a fact that is easily proved directly; equality holds if and only if the coclique is a Steiner system. So the graph has clique number equal to chromatic number if and only if the set of k -subsets of an n -set can be partitioned into block sets of Steiner systems $S(t, k, n)$. A collection of Steiner systems which partitions the set of all k -subsets is a *large set*. In view of this, we further conjecture the following:

Conjecture 2.6. There is a function L such that, if $n \geq L(k)$, then the Johnson scheme $J(n, k)$ is non-synchronizing if and only if there exists a large set of Steiner triple systems $S(t, k, n)$ for some t with $0 < t < k$.

Much less is known about the existence of large sets. The main results are the following:

- (a) For $t = 1$, an $S(t, k, n)$ is simply a partition of $\{1, \dots, n\}$ into sets of size k , which exists if and only if k divides n . A theorem of Baranyai [3] shows that a large set of such partitions exists whenever k divides n .
- (b) For $t = 2$, $k = 3$, Kirkman [16] showed that a $S(2, 3, n)$ exists if and only if $n \equiv 1$ or $3 \pmod{6}$. The smallest example is the *Fano plane* with $n = 7$. Cayley [6] showed that there does not exist a large set of Fano planes; indeed, there do not exist more than two pairwise disjoint Fano planes. However, Lu and Teirlinck [21] showed that, for all “admissible” n greater than 7, a large set of Steiner triple systems exists.
- (c) Kolotoğlu and Magliveras [17] have constructed large sets of projective planes of orders 3 and 4 (that is, $S(2, 4, 13)$ and $S(2, 5, 21)$).

It may be that large sets of $S(t, k, n)$ exist whenever the divisibility conditions are satisfied and n is sufficiently large. If so, then our conjecture above would imply that, for n sufficiently large in terms of k , the Johnson scheme $J(n, k)$ is synchronizing if and only if it is separating. However, we are not sufficiently confident to conjecture this!

3 Small examples

An association scheme is non-separating if there is a graph in the scheme for which the product of the clique and coclique numbers is equal to the number of vertices. Given such

a situation, there are two ways that synchronization can fail: either there is a partition of the vertices into cocliques of maximal size, so that the clique number and chromatic number are equal, or there is a partition into cliques of maximal size, so that this equality holds in the complementary graph.

As we saw in the preceding section, if $n > (t+1)(k-t+1)$, then $\Delta_t(n, k)$ cannot have clique number equal to chromatic number, since the maximum-size cocliques are of EKR type; and complement $\Phi_t(n, k)$ has clique number equal to chromatic number if and only if there is a large set of Steiner systems $S(t, k, n)$.

For smaller n , non-synchronization can arise in one of two ways: either because there are other large cocliques in $\Delta_t(n, k)$, or because one of the other graphs $\Gamma_I(n, k)$ in the association scheme has clique number equal to chromatic number.

Here are five examples. In all but one (the case $n = 9, k = 4$) it is the first of these two ways which occurs.

The case $k = 3, n = 7$. In this case, for $t = 2$, a clique is the block set of the $S(2, 3, 7)$, the Fano plane. Since it is a projective plane, any two lines meet in a point, so it is a clique in the graph $\Gamma_{\{2\}}(3, 7)$. As well as cocliques of EKR type (all 3-sets containing a given two points), there are cocliques of size 5 defined as follows: let L be a line of the Fano plane, and take L together with the four 3-sets disjoint from it. Furthermore, the seven such sets obtained by performing this construction for each line of the Fano plane partition the 35 sets of size 3 into seven cocliques of size 5. So this graph has clique number equal to chromatic number.

The case $k = 3, n = 8$. Again a Fano plane gives us a 7-clique in the graph $\Gamma_{\{2\}}(3, 8)$. Now the eight 3-sets consisting of a line L , three sets each comprising two points of L and the point outside the Fano plane, and the four sets consisting of three of the four points of the Fano plane outside L , form a coclique; doing this for the seven lines we obtain a partition of the 56 sets of size 3 into seven cocliques of size 8, so this graph has clique number equal to chromatic number. Another way of viewing this is to observe that the Fano plane has an extension to a $S(3, 4, 8)$ whose blocks fall into 7 parallel classes with two blocks in each; the eight 3-sets contained in a block of a parallel class form a coclique, and we obtain seven such sets, one for each parallel class.

The case $k = 4, n = 9$. The Steiner system $S(3, 4, 8)$ has 14 blocks, any two meeting in 0 or 2 points. We can construct a set of 9 subsets of size 4, any two meeting in 1 or 3 points, as follows: partition $\{1, \dots, 9\}$ into three sets of size 3, arranged around a circle; now take the 4-subsets consisting of one part and a single point of the next (in the cyclic order).

Breach and Street [4] showed that the 126 4-subsets of a 9-set can be partitioned into a so-called *overlarge set* of nine Steiner systems $S(3, 4, 8)$ (each omitting a point); indeed, this can be done in just two non-isomorphic ways, each admitting a 2-transitive group. This gives a colouring of the graph $\Gamma_{\{1,3\}}(4, 9)$ corresponding to intersections 1 and 3. (Their proof was computational; a more geometric proof involving triality was given by Cameron and Praeger [5].)

The case $k = 5, n = 11$. An example similar to the first is obtained using the Steiner system $S(4, 5, 11)$, any two of whose blocks intersect in 1, 2 or 3 points. Thus the blocks form a 66-clique in the graph $\Gamma_{\{2,3,4\}}(11, 5)$. A block together with the six 5-sets disjoint from it form a coclique of size 7, and the 66 sets obtained in this way form a colouring of the graph.

The case $k = 5, n = 12$. Again the blocks of $S(4, 5, 11)$ form a 66-clique in $\Gamma_{\{2,3,4\}}(12, 5)$. The Steiner system has an extension to a $S(5, 6, 12)$ whose blocks come in 66 parallel classes with two disjoint blocks in each, and the twelve 5-sets contained in a block of a fixed parallel class form a coclique.

4 The case $I = \{1, k\}$

In this section, we deal with the case $I = \{1, k\}$ (or the complement $I = \{2, \dots, k-1\}$) of Conjecture 2.2, and show that these cannot occur if n is sufficiently large. In other words, taking account of the indexing used in the Johnson scheme, we show the following.

Theorem 4.1. *There is a function f such that, if $n \geq f(k)$, and S and T are families of k -subsets of $\{1, \dots, n\}$ with the property that S is $\{0, k-1\}$ -intersecting (that is, any two of its members intersect in 0 or $k-1$ points) and T is $\{1, \dots, k-2\}$ -intersecting, then $|S| \cdot |T| < \binom{n}{k}$.*

Proof. The proof proceeds in three steps.

Step 1 $|S| \leq n$.

To see this, consider first a $(k-1)$ -intersecting family U of k -sets. It is easy to see that there are just two possibilities:

- (a) all members of U contain a fixed $(k-1)$ -set;
- (b) all members of U are contained in a fixed $(k+1)$ -set.

Next we claim that the relation \sim on S defined by $A \sim B$ if $A = B$ or $|A \cap B| = k-1$ is an equivalence relation. It is clearly reflexive and symmetric, so suppose that $A \sim B$ and $B \sim C$. Then $|A \cup B| = |B \cup C| = k+1$, and so $|A \cap C| \geq k-2$, whence $|A \cap C| = k-1$ as required.

Now if two members of S belong to distinct equivalence classes, they are disjoint. So the support of S (the set of points lying in some element of S) is the union of the supports of the equivalence classes, which are pairwise disjoint. We have seen that the number of sets in each equivalence class does not exceed the cardinality of its support; so the same holds for S , and the claimed inequality follows.

For the next step, we note that T is an intersecting family. We split the proof into two subcases.

Step 2 If the intersection of the sets in T is non-empty, then $|T| \leq \binom{n-1}{k-2}/(k-1)$.

For let x be the unique point in the intersection. Then

$$T = \{\{x\} \cup B : B \in T'\},$$

where T' is a $\{0, \dots, k-3\}$ -intersecting family of $(k-1)$ -subsets of $\{1, \dots, n\} \setminus \{x\}$; in other words, a partial $S(k-2, k-1, n-1)$. So $|T| = |T'| \leq \binom{n-1}{k-2}/(k-1)$, the right-hand side being the number of blocks in a hypothetical Steiner system with these parameters.

Step 3 If the intersection of the sets in T is empty, then $|T| \leq \binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1$.

Since T is an intersecting family, this is just the conclusion of the Hilton–Milner theorem [14].

Conclusion of the proof We have

$$\binom{n}{k} = |S| \cdot |T| \leq \begin{cases} n \binom{n-1}{k-2}/(k-1) & \text{if } \bigcap T \neq \emptyset, \\ n \left(\binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1 \right) & \text{if } \bigcap T = \emptyset. \end{cases}$$

But in each case, for fixed k , the left-hand side of the inequality is a polynomial of degree k in n , whereas the right-hand side is a polynomial of degree $k-1$; thus the inequality holds for only finitely many values of n . \square

We remark that, in fact, we know of no examples meeting the bound for this case with $n > 2k$. So as well as extending these techniques to other cases, the problem of deciding whether the bound is always strict remains.

In the next section we use the fact proved in the first part of the above proof:

Corollary 4.2. *For $k \geq 3$ and $n > 2k$, a $\{0, k-1\}$ -intersecting family of k -subsets of $\{1, \dots, n\}$ has size at most n .*

Indeed, it is not too hard to find the precise upper bound; but we do not require this.

5 The case $k = 4$

5.1 Background

To handle the case $k = 4$, we use the results of Delsarte [7] on association schemes. We begin with a brief introduction to this material, but we also refer to [11, Chapter 6] where some of the computations that we omit are done in detail.

Suppose that the matrices A_0, \dots, A_{k-1} (with $A_0 = I$) span the Bose–Mesner algebra of an association scheme on v points. Since this algebra is a commutative algebra of real symmetric matrices, the matrices are simultaneously diagonalisable: that is, there are idempotent matrices E_0, \dots, E_{k-1} spanning the same algebra, with $E_0 = J/v$, where J is the all-1 matrix. Thus, for some coefficients $P_k(i)$ and $Q_k(i)$, we have

$$\begin{aligned} A_j &= \sum_{i=0}^{k-1} P_j(i) E_i, \\ E_j &= v^{-1} \sum_{i=0}^{k-1} Q_j(i) A_i. \end{aligned}$$

Here the numbers $P_j(i)$ for $i = 0, \dots, k-1$ are the eigenvalues of A_j . The matrices with (i, j) entry $P_j(i)$ and $Q_j(i)$ are called the *matrix of eigenvalues* and *dual matrix of eigenvalues* of the scheme.

Delsarte used these matrices to provide bounds on the clique and coclique numbers of graphs in an association scheme:

Theorem 5.1 ([7], Theorem 5.9; see also [11]). *Let \mathcal{A} be an association scheme on v vertices and let Γ be the union of some of the graphs in the scheme. If C is a clique and S is a coclique in Γ , then $|C| \cdot |S| \leq v$. If equality holds and x and y are the respective characteristic vectors of C and S , then $(xE_jx^\top)(yE_jy^\top) = 0$ for all $j > 0$.*

In the above notation, the inner distribution a of C is the vector where $a_i = xA_ix^\top/|C|$ for each $i \in \{0, \dots, d\}$ (and \mathcal{A} has d classes). Now if Q is the dual matrix of eigenvalues of \mathcal{A} , then

$$(aQ)_j = \frac{v}{|C|} xE_jx^\top$$

for all $j \geq 0$. This vector is sometimes known as the *MacWilliams transform* of C .

The *degree set* of a subset X of the vertices of Γ is the set of nonzero indices i for which the i -th coordinate of its inner distribution is nonzero. The *dual degree set* of X is the set of nonzero indices j for which the j -th coordinate of its MacWilliams transform is nonzero. Two subsets X and Y of the vertices of Γ are *design-orthogonal* if their dual degree sets are disjoint. Similarly, X and Y are *code-orthogonal* if their degree sets are disjoint.

Corollary 5.2. *Suppose a clique C and coclique S meet the bound; $|C| \cdot |S| = v$. Then the Schur product of the MacWilliams transforms of C and S equals $(v, 0, \dots, 0)$.*

Corollary 5.3. *Suppose Γ is a graph from an association scheme \mathcal{A} on v vertices, and suppose $\alpha(\Gamma)\omega(\Gamma) = v$. Then $\omega(\Gamma) = 1 - \deg(\Gamma)/\tau$, where τ is the smallest eigenvalue of Γ , and in particular, τ divides $\deg(\Gamma)$.*

Proof. Let τ be the smallest eigenvalue of Γ . By the result of Lovász [19], $\alpha(\Gamma) \leq \frac{v}{1 - \deg(\Gamma)/\tau}$. So since $\alpha(\Gamma)\omega(\Gamma) = v$, we have $v \leq \omega(\Gamma)v/(1 - \deg(\Gamma)/\tau)$ and hence $1 - \deg(\Gamma)/\tau \leq \omega(\Gamma)$; that is, we obtain equality $\omega(\Gamma) = 1 - \deg(\Gamma)/\tau$. \square

We can find a simple expression for the eigenvalues of the Johnson scheme from Section 4.2.1 of Delsarte's PhD thesis. The analogue of the Krawchouk polynomials of the Hamming scheme to the Johnson scheme are the *Eberlein polynomials*, or *dual Hahn polynomials*. Given an integer $0 \leq j \leq k$, we define the Eberlein polynomial $E_j(x)$ in the indeterminate x , as follows:

$$E_j(x) := \sum_{t=0}^j (-1)^{j-t} \binom{k-t}{j-t} \binom{k-x}{t} \binom{n-k+t-x}{t}.$$

The (i, j) -entry of the matrix of eigenvalues of P will be denoted $P_j(i)$. Now Theorem 4.6 of Delsarte [7] asserts:

Theorem 5.4. *The matrix P of eigenvalues and the dual matrix of eigenvalues Q of the Johnson scheme $J(n, k)$ are given by*

$$P_j(i) = E_j(i), \quad Q_i(j) = \frac{\binom{n}{i} - \binom{n}{i-1}}{\binom{k}{j} \binom{n-k}{j}} E_j(i),$$

for $i, j = 0, 1, \dots, k$.

5.2 The main result

Theorem 5.5. *Let $n \geq 10$ and $I \in \{\{1, 3, 4\}, \{1, 3\}, \{1, 4\}, \{1, 2, 4\}\}$, and let $\Delta := \Gamma(n, 4, I)$. Then $\omega(\Delta)\alpha(\Delta) < |V\Delta| = \binom{n}{4}$.*

Proof. A straight-forward calculation shows that the matrix P of eigenvalues, and the dual matrix Q , of the Johnson scheme $J(n, 4)$ are

$$P = \begin{pmatrix} 1 & 4(n-4) & 3(n-5)(n-4) & \frac{2}{3}(n-6)(n-5)(n-4) & \frac{1}{24}(n-7)(n-6)(n-5)(n-4) \\ 1 & 3n-16 & \frac{3}{2}(n-8)(n-5) & \frac{1}{6}(n-16)(n-6)(n-5) & -\frac{1}{6}(n-7)(n-6)(n-5) \\ 1 & 2(n-7) & \frac{1}{2}((n-21)n+92) & -(n-9)(n-6) & \frac{1}{2}(n-7)(n-6) \\ 1 & n-10 & -3(n-8) & 3n-22 & 7-n \\ 1 & -4 & 6 & -4 & 1 \end{pmatrix},$$

$$Q = \begin{pmatrix} 1 & n-1 & \frac{1}{2}(n-3)n & \frac{1}{6}(n-5)(n-1)n & \frac{1}{24}(n-7)(n-2)(n-1)n \\ 1 & \frac{1}{4}(3n-7) - \frac{3}{n-4} & \frac{(n-7)(n-3)n}{4(n-4)} & \frac{(n-10)(n-5)(n-1)n}{24(n-4)} & -\frac{(n-7)(n-2)(n-1)n}{24(n-4)} \\ 1 & \frac{(n-8)(n-1)}{2(n-4)} & \frac{(n-3)n((n-21)n+92)}{12(n-5)(n-4)} & -\frac{(n-8)(n-1)n}{6(n-4)} & \frac{(n-7)(n-2)(n-1)n}{12(n-5)(n-4)} \\ 1 & \frac{(n-16)(n-1)}{4(n-4)} & -\frac{3(n-9)(n-3)n}{4(n-5)(n-4)} & \frac{(n-1)n(3n-22)}{4(n-6)(n-4)} & -\frac{(n-7)(n-2)(n-1)n}{4(n-6)(n-5)(n-4)} \\ 1 & -\frac{4(n-1)}{n-4} & \frac{6(n-3)n}{(n-5)(n-4)} & -\frac{4(n-1)n}{(n-6)(n-4)} & \frac{(n-2)(n-1)n}{(n-6)(n-5)(n-4)} \end{pmatrix}.$$

Case $I = \{1, 3, 4\}$: The complement $\overline{\Delta}$ of Δ is a graph of an association scheme (because it is $\Gamma(n, 4, \{2\})$). Consider the third column of P . The eigenvalues of $\overline{\Delta}$ are

$$3(n-5)(n-4), \frac{3}{2}(n-8)(n-5), \frac{1}{2}(n^2 - 21n + 92), 24 - 3n, 6.$$

The degree of $\overline{\Delta}$ is the first eigenvalue $3(n-5)(n-4)$. Most of the time, the fourth eigenvalue $24 - 3n$ is the smallest eigenvalue, certainly when $n \geq 11$, when it is less than the third eigenvalue. So let us assume that $n \geq 11$. By Corollary 5.3, $24 - 3n$ divides $3(n-5)(n-4)$, and hence $8 - n$ divides $(n-5)(n-4)$. So by elementary number theory, we have $n \in \{11, 12, 14, 20\}$.

The cases $n = 9, 10, 11, 12, 14, 20$ can be settled each in turn. First, for $n = 9$, the minimum eigenvalue of $\overline{\Delta}$ is -8 , and the degree is 65 . So by Corollary 5.3, $\omega(\Delta)\alpha(\Delta) < \binom{n}{4}$. For the remaining cases, we can calculate $\omega(\overline{\Delta})$ with GAP/GRAPE [10, 20]¹ and compare it to $1 - \deg(\overline{\Delta})/\tau$, where τ is the smallest eigenvalue.

n	$\omega(\overline{\Delta})$	$1 - \deg(\overline{\Delta})/\tau$
10	5	11
11	6	15
12	9	15
14	13	16
20	13	21

¹The share package GRAPE for the computer algebra system GAP contains an efficient clique finder.

We find that $\omega(\overline{\Delta})$ is never equal to $1 - \deg(\overline{\Delta})/\tau$ and so by Corollary 5.3, $\omega(\Delta)\alpha(\Delta) < \binom{n}{4}$.

Case $I = \{1, 2, 4\}$: This time, the complement graph $\overline{\Delta}$ is $\Gamma(n, 4, \{3\})$. The eigenvalues of $\overline{\Delta}$ are

$$\frac{2}{3}(n-6)(n-5)(n-4), \frac{1}{6}(n-16)(n-6)(n-5), -(n-9)(n-6), 3n-22, -4$$

For $n \geq 13$, the third eigenvalue of $\overline{\Delta}$ is the smallest in the spectrum of $\overline{\Delta}$. So by Corollary 5.3, $n-9$ divides $\frac{2}{3}(n-5)(n-4)$ and hence $n \in \{13, 14, 17, 19, 29, 49\}$ (by basic elementary number theory). For $n \in \{9, 11\}$, the smallest eigenvalue of $\overline{\Delta}$ does not divide the degree. Thus by Corollary 5.3, we are left with $n \in \{10, 12, 13, 14, 17, 19, 29, 49\}$ to check. The cases $n = 10, 12, 13, 14, 17, 19, 29, 49$ can be settled each in turn with the computer algebra system GAP/GRAPE:

n	$\omega(\overline{\Delta})$	$1 - \deg(\overline{\Delta})/\tau$
10	2	5
12	3	9
13	3	13
14	3	13
17	4	14
19	4	15
29	7	21
49	12	34

Case $I = \{1, 3\}$: Let Q be the dual matrix of eigenvalues. The second column of Q is

$$c = \frac{n-1}{4(n-4)} (4(n-4), 3n-16, 2(n-8), n-16, -16).$$

Let $u = (1, a, 0, x-a-1, 0)$ and $v = (1, 0, b, 0, y-b-1)$ such that $a, b \geq 0$, $x-a-1, y-b-1 \geq 0$ and $xy = \binom{n}{4}$. So u and v are the inner distributions of an arbitrary clique and coclique of Γ_I , respectively, attaining the clique-coclique bound. The second entries of uQ and vQ are

$$(uQ)_1 = uc^\top = \frac{n-1}{4(n-4)} ((2a+x+3)n-16x)$$

$$(vQ)_1 = vc^\top = \frac{n-1}{4(n-4)} 2((b+2)n-8y).$$

By Corollary 5.2, the product of these quantities is zero, which gives us two scenarios: $x = n\frac{2a+3}{16-n}$ or $y = n\frac{1}{8}(b+2)$. We will consider the former, with the additional assumption that $xy = \binom{n}{4}$. This then yields an expression for y :

$$y = -\frac{(n-16)(n-3)(n-2)(n-1)}{48a+72}.$$

Now $a, y \geq 0$, which implies that $n \leq 16$. So we will assume for the moment that $n \geq 16$. Therefore,

$$x = \frac{(n-3)(n-2)(n-1)}{3(b+2)}, \quad y = \frac{1}{8}(b+2)n.$$

We now consider the two equations $(uQ)_2(vQ)_2 = 0$ and $(uQ)_4(vQ)_4 = 0$, with the above values of x and y substituted in:

$$\begin{aligned} (b(n-10) + 6(n-4))(a(b+2)(n-8) + b(2n-13) - (n-4)((n-10)n+7)) &= 0 \\ a(b+2)(n-8) + 18(n-4) - b(n^2 - 12n + 38) &= 0. \end{aligned}$$

The second equation gives us a value for a , which we can substitute into the first equation. This results in the following equation:

$$(n-5)(b(n-10) + 6(n-4))(b-n+4) = 0.$$

However, $n \neq 5$ and $b(n-10) + 6(n-4) \neq 0$ (as $b \geq 0$ and $n > 10$), so it follows that $b = n-4$ and hence

$$x = \frac{(n-3)(n-1)}{3}, \quad y = \frac{1}{8}(n-2)n.$$

Now suppose we have a $\{1, 3\}$ -clique S of size $x = (n-1)(n-3)/3$. Consider the members of S containing a point p , with p removed from each. This is a family of 3-sets of an $(n-1)$ -set, any two intersecting in 0 or 2 elements. By Corollary 4.2, we know that there are at most $n-1$ of them. Now the standard double count gives

$$|S| \leq n(n-1)/4.$$

But this is smaller than $(n-1)(n-3)/3$ so long as n is at least 16. This leaves the cases $10 \leq n \leq 16$ to be considered (since we have excluded $n = 9$).

n	$\omega(\Gamma_I)$	$\alpha(\Gamma_I)$
10	9	14
11	9	14
12	9	15
13	9	15
13	13	15
14	13	21
15	13	21
16	13	28

Case $I = \{1, 4\}$: In this case, Theorem 4.1 shows that the result holds for sufficiently large n ; indeed, the proof there works for $n \geq 46$. However, the intervening values are far too large for computation, so we use the Q-matrix methods.

Let $u = (1, a, 0, 0, x-a-1)$ and $v = (1, 0, b, y-b-1, 0)$ such that $a, b \geq 0$, $x-a-1, y-b-1 \geq 0$ and $xy = \binom{n}{4}$. So u and v are the inner distributions

of an arbitrary clique and coclique of Γ_I , respectively, attaining the clique-coclique bound. The second coordinates of uQ and vQ are

$$(uQ)_1 = uc^\top = \frac{(n-1)}{4(n-4)}((3a+4)n-16x),$$

$$(vQ)_1 = vc^\top = \frac{(n-1)}{4(n-4)}((b+y+3)n-16y).$$

By Corollary 5.2, the product of these quantities is zero, which gives us two scenarios:

- (i) $x = n\frac{3a+4}{16}$, or
- (ii) $y = n\frac{b+3}{16-n}$.

Since $y \geq 0$, Case (ii) does not arise if we assume $n \geq 17$, which we will for now. So suppose we have Case (i). If we now consider the equation $(uQ)_2(vQ)_2 = 0$, we have

$$((n-1)(a(n-11)+2(n-8))+24x)((n-1)(b(n-11)+6n-39)-9(n-9)y) = 0. \quad (1)$$

Assuming (i), that is $x = n(3a+4)/16$, Equation (1) becomes

$$(2(a+2)n-11a-16)((n-1)(b(n-11)+6n-39)-9(n-9)y) = 0. \quad (2)$$

However, if the first term is zero, then $a = (16-4n)/(2n-11)$ which is negative for $n \geq 6$; a contradiction. Therefore, the second term in Equation (2) is zero. In other words, we have an expression for b in terms of n and y :

$$b = \frac{-6n^2 + 9ny + 45n - 81y - 39}{(n-11)(n-1)}.$$

If we now consider the equation $(uQ)_3(vQ)_3 = 0$, upon substitution of our value for b , we obtain an equation relating a , y and n :

$$(a(n^2-18n+68)+4(n-8)(n-4))(2n^4-34n^3-9n^2y+181n^2+123ny-335n-426y+186) = 0. \quad (3)$$

Having the first term in Equation (3) equal to zero leads to a contradiction, since in this case we would have

$$a = -\frac{4(n^2-12n+32)}{n^2-18n+68}$$

which is negative for $n \geq 13$. So let us now assume $n \geq 13$. Then the second term in Equation (3) is zero, which gives us an expression for y in terms of n (and we also obtain an expression for x):

$$x = \frac{n(n-3)(3n^2-41n+142)}{8(2n^2-28n+93)},$$

$$y = \frac{(n-1)(n-2)(2n^2-28n+93)}{9n^2-123n+426}.$$

However, $x \leq n$ (Corollary 4.2), which implies that $n \leq 8$; a contradiction.

This leaves us now to consider by computer the cases where $9 \leq n \leq 16$. In each case, we see that the clique number and coclique number have a product that does not attain $\binom{n}{4}$.

n	$\omega(\Gamma_I)$	$\alpha(\Gamma_I)$
9	6	12
10	10	15
11	10	15
12	10	15
13	10	15
14	11	15
15	15	15
16	15	15

□

Corollary 5.6. *For $k = 4$, Conjectures 2.2 and 2.3 hold for $n \geq 10$.*

This follows from the main theorem together with well-known results of Hanani on the existence of Steiner systems $S(t, 4, n)$ [12, 13]. Note that $n = 9$ is a genuine exception, as we saw in Section 3.

We conclude that, if $n \geq 10$, then the permutation group induced by S_n on 4-sets fails to be separating if and only if one of the following holds:

- $n \equiv 0 \pmod{4}$ (so that $S(1, 4, n)$ exists),
- $n \equiv 1$ or $4 \pmod{12}$ (so that $S(2, 4, n)$ exists),
- $n \equiv 2$ or $4 \pmod{6}$ (so that $S(3, 4, n)$ exists);

summarising, $n \equiv 0, 1, 2, 4, 8, 10 \pmod{12}$. What about synchronization? Wilson's theorem [22] ensures that maximal cocliques are of EKR type, so G is non-synchronizing if and only if a large set of Steiner systems exists. Baranyai's theorem [3] ensures that G is not synchronizing if $n \equiv 0 \pmod{4}$, but existence of large sets of Steiner systems in the other cases is unresolved (except for $t = 2$, $n = 13$, [17]).

In particular, the existence of a Steiner system $S(3, 4, 10)$ shows that the symmetric group S_{10} acting on 4-sets is non-separating. However, it is synchronizing. Our above results show that synchronization could only fail if there were a large set of seven pairwise disjoint $S(3, 4, 10)$ systems. However, Kramer and Mesner [18] showed that there cannot be more than five such systems.

Thus, for $k = 4$, synchronization and separation fail to be equivalent for S_n on k -sets, unlike the cases $k = 2$ and $k = 3$ described earlier.

6 Projective planes

The constructions in Section 3 involved the fact that, in certain Steiner systems, certain cardinalities of block intersection do not occur. There are relatively few examples of such systems: the only ones known are projective planes, $S(3, 4, 8)$, $S(4, 5, 11)$, $S(5, 6, 12)$, $S(3, 6, 22)$, $S(4, 7, 23)$ and $S(5, 8, 24)$.

A *projective plane* of order q is a Steiner system $S(2, q+1, q^2+q+1)$ for some integer $q > 1$. Projective planes of all prime power orders exist, and none are known for other orders.

A projective plane has the property that any two of its blocks meet in a point. Hence it is a clique in either of the graphs $\Gamma_{\{>k-2\}}(n, k)$ or $\Gamma_{k-1}(n, k)$, with $k = q + 1$ and $n = q^2 + q + 1$. In the first of these graphs, cocliques of maximum cardinality are of EKR type, and we have

$$\binom{q^2 + q - 1}{q - 1} \cdot (q^2 + q + 1) = \binom{q^2 + q + 1}{q + 1},$$

so non-separation holds for this graph: these sets (all k -sets containing two given points) also show non-separation for $\Gamma_1(n, k)$. Also, by Theorem 2.5, we cannot partition the k -sets into subsets of EKR type.

In the case $q = 2$, in our example above, we observed that there were other cocliques, so that the possibility of a colouring with $q^2 + q + 1$ colours cannot be ruled out; and indeed we saw that such a colouring exists.

Conjecture 6.1. For $q > 2$, a coclique of maximum size in the graph $\Gamma_q(q^2 + q + 1, q + 1)$ must consist of all the $(q + 1)$ -sets containing two given points; so the chromatic number of this graph is strictly larger than $q^2 + q + 1$.

A simple computation shows that the conjecture is true for $q = 3$ and for $q = 4$.

On the other hand, the truth of this conjecture would probably not give an infinite family of examples which are synchronizing but not separating. Magliveras conjectured that large sets of projective planes of any order $q > 2$ exist; the existence is shown for $q = 3$ and $q = 4$ in [17].

7 Conclusion

The computations with rational functions in Section 5 were performed with Mathematica [23]; computations of clique number in special cases were done with GAP and its package GRAPE, as already noted.

The techniques we used to prove the separation conjecture for $k = 4$ are amenable to the use of traditional computer algebra systems such as Mathematica, and it should be possible to settle several more values of k .

In the arguments for $I = \{1, 3\}$ and $I = \{1, 4\}$, we saw that the sizes of a clique and a coclique whose product is equal to the number of vertices can be determined from the Q-matrix of the association scheme. Is this true in general? (It does hold for two-class association schemes, that is, strongly regular graphs.)

The synchronization question is likely to be harder; general results on the existence or nonexistence of large sets of Steiner systems are likely to require a significant new idea, and even particular cases involve quite large computations with GAP or more specialised software. Perhaps new techniques in hypergraph decomposition will help.

There seems plenty of scope for extending these results to other primitive association schemes.

References

- [1] J. Araújo, P. J. Cameron, and B. Steinberg, Between primitive and 2-transitive: Synchronization and its friends, *Europ. Math. Soc. Surveys*, in press.
- [2] R. A. Bailey, *Association Schemes: Designed Experiments, Algebra and Combinatorics*, Cambridge Studies in Advanced Mathematics **84**, Cambridge Univ. Press, Cambridge, 2004.
- [3] Zs. Baranyai, On the factorization of the complete uniform hypergraph, *Infinite and finite sets* (Colloq., Keszthely, 1973), Vol. I, pp. 91–108, *Colloq. Math. Soc. Janos Bolyai*, **10**, North-Holland, Amsterdam, 1975.
- [4] D. R. Breach and A. P. Street, Partitioning sets of quadruples into designs II, *J. Combinatorial Math. Combinatorial Computing* **3** (1988), 41–48.
- [5] P. J. Cameron and C. E. Praeger, Partitioning into Steiner systems, *Combinatorics '88* (ed. A. Barlotti et al.), Mediterranean Press, Roma, 1992, pp. 61–71.
- [6] A. Cayley, On the triadic arrangements of seven and fifteen things, *Philosophical Mag.* **37** (1850), 50–53.
- [7] P. Delsarte, An algebraic approach to the association schemes of coding theory, *Philips Research Reports Suppl.* **10** (1973).
- [8] P. Erdős, C. Ko and R. Rado, Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford* (2) **12** (1961), 313–320.
- [9] P. Frankl and Z. Füredi, Forbidding just one intersection, *J. Combinatorial Theory* (A) **39** (1983), 160–176.
- [10] The GAP Group, GAP – Groups, Algorithms, and Programming, Version 4.8.7; 2017. <http://www.gap-system.org>
- [11] C. Godsil and K. Meagher, *Erdős–Ko–Rado Theorems: Algebraic Approaches*, Cambridge Studies in Advanced Mathematics **149**, Cambridge Univ. Press, Cambridge, 2016.
- [12] H. Hanani, On quadruple systems, *Canad. J. Math.* **12** (1960), 145–157.
- [13] H. Hanani, The existence and construction of balanced incomplete block designs, *Ann. Math. Stat.* **32** (1961), 361–386.
- [14] A. J. W. Hilton and E. C. Milner, Some intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser. (2)* (1967), 369–384.
- [15] P. Keevash, The existence of designs, <https://arxiv.org/abs/1401.3665>
- [16] T. P. Kirkman, On a problem in combinations, *Cambridge and Dublin Math. J.* **2** (1847), 713–724.

- [17] E. Kolotoğlu and S. S. Magliveras, On large sets of projective planes of orders 3 and 4, *Discrete Math.* **313** (2013), 2247–2252.
- [18] E. S. Kramer and D. M. Mesner, Intersections among Steiner systems, *J. Combinatorial Theory (A)* **16** (1974), 273–285.
- [19] L. Lovász, On the Shannon capacity of a graph, *IEEE Trans. Inform. Theory* **25** (1979), 1–7.
- [20] L. H. Soicher, GRAPE (Version 4.7) Package for GAP, <http://www.maths.qmul.ac.uk/~leonard/grape/>
- [21] L. Teirlinck, A completion of Lu’s determination of the spectrum of large sets of Steiner triple systems, *J. Combinatorial Theory (A)* **57** (1991), 302–305.
- [22] Richard M. Wilson, The exact bound in the Erdős–Ko–Rado theorem, *Combinatorica* **4** (1984), 247–257.
- [23] Wolfram Research, Inc., Mathematica, Version 11.2, Champaign, IL (2017).