Covering Problems on Permutations

Amber Estelle Gentle

Bachelor of Mathematics (Honours), University of Queensland, 2020

A thesis submitted for the degree of

Doctor of Philosophy

at Monash University in 2024

School of Mathematics
Monash University
Melbourne, Australia
2024

Copyright notice

Abstract

A perfect sequence covering array, denoted $\mathsf{PSCA}(v,t,\lambda)$ is a set X of permutations of $\{0,\ldots,v-1\}$ in which every sequence of t distinct elements of $\{0,\ldots,v-1\}$ appears as a subpermutation of exactly λ permutations in X. The perfect sequence covering array number, $\mathsf{PSCAN}(v,t)$ is the smallest λ for which a $\mathsf{PSCA}(v,t,\lambda)$ exists. We prove new restrictions on the number of times a given symbol can appear in each column of a $\mathsf{PSCA}(v,t,\lambda)$. These restrictions motivate the development of an exhaustive computational search for $\mathsf{PSCA}(v,t,\lambda)$ which we use to show $\mathsf{PSCAN}(6,3)=\mathsf{PSCAN}(7,3)=\mathsf{PSCAN}(7,4)=2$ and $\mathsf{PSCAN}(8,3)=3$. Additionally, this algorithm is used to generate complete catalogues of $\mathsf{PSCA}(v,t,\lambda)$ for a variety of parameter sets. We perform searches over permutation subgroups of the symmetric group $\mathcal{S}_v$ for $v\leqslant 14$ for $\mathsf{PSCA}(v,t,\lambda)$. These searches yield many different permutation groups that also form perfect sequence covering arrays. Further, we give a construction for perfect sequence covering arrays that can be applied for any choice of v and t that makes use of the group of projectivities of a finite projective plane. The size of the perfect sequence covering arrays yielded by this method grows as a polynomial in v for fixed t.

We also explore a related object called a sequence covering array, denoted $\mathsf{SCA}(N;t,v)$, which is a set X of N permutations of $\{0,\ldots,v-1\}$ such that every sequence of t distinct elements of $\{0,\ldots,v-1\}$ appears as a subpermutation of at least one permutation in X. An $\mathsf{SCA}(N;t,v)$ must have at least $t!$ permutations. Levenshtein conjectured that an $\mathsf{SCA}(t!;t,v)$ exists if and only if $v\in\{t,t+1\}$. The existence of $\mathsf{SCA}(t!;t,v)$ has been completely resolved for $t\leqslant 6$. Chee, Colbourn, Horsley and Zhou proved that if an $\mathsf{SCA}(t!;t,v)$ exists, then $v\leqslant 2t-1$ by drawing connections between sequence covering arrays and a related combinatorial object called a covering array. We further explore these connections and perform exhaustive computational searches for a special kind of covering array. These searches allow us to prove that if an $\mathsf{SCA}(7!;7,v)$ exists, then $v\leqslant 9$. This improves on the previous best known upper bound of $v\leqslant 13$.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Amber Estelle Gentle
31 July 2024

Publications during enrolment

- Chapter 7 is based off a paper published in Annals of Combinatorics [20] with Ian Wanless.
- Chapter 8 is based off a paper published in Algebraic Combinatorics [19].

Acknowledgements

So many people have supported me throughout but in particular I would like to thank

My supervisors Ian and Daniel for allowing me the opportunity to pursue a PhD in the first place and for all the guidance you have given me.

My parents Nathalie and Steve for dozens of late night phone calls and for always having a home for me when I most needed it.

My friends Leah, Imogen, Angus, Em and Michael for making me feel welcome in Melbourne, for marathon Mario Kart sessions and for hours of bad dumb office jokes.

My best friend Audrey for always being there to talk about anything and for teaching me so much about myself.

And my girlfriend Scout for all the love, joy and silliness you give me each day.

Thank you, Amber.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

Contents

Chapter 1

Introduction

The problem of finding minimal set covers is fundamental to many different areas of combinatorics. The subcategory of such problems relating to finding perfect set covers defines much of design theory. For example, Steiner systems can be seen as perfect covers of 2-subsets of a ground set with k-subsets. In this thesis, we consider similar set cover problems in which permutations of $\{0, \ldots, v-1\}$ cover sequences of distinct elements of $\{0, \ldots, v-1\}$.

The main objects we discuss in this thesis are the sequence covering array and its variant, the perfect sequence covering array. A sequence covering array, denoted $\mathsf{SCA}(N; t, v)$, is a set X of N permutations of $\{0, \ldots, v-1\}$ such that each sequence of t distinct elements of $\{0, \ldots, v-1\}$ appears as a subpermutation of at least one permutation in X. Sequence covering arrays have been studied as far back as 1971 [51], but the study of these objects has seen a resurgence in the mathematics and computer science literatures since 2012 because of their natural applications in software testing [34]. A perfect sequence covering array, denoted $\mathsf{PSCA}(v, t, \lambda)$, is a sequence covering array in which each sequence of t distinct elements of $\{0, \ldots, v-1\}$ appears as a subpermutation of exactly λ permutations. Perfect sequence covering arrays were introduced by Yuster [64] in 2020 as a design-theoretical extension of sequence covering arrays.

In Chapter 3, we introduce sequence covering arrays more formally and discuss the problem of finding $\mathsf{SCAN}(t, v)$, the smallest N for which an $\mathsf{SCA}(N; t, v)$ exists for fixed v and t. In Chapter 3.1, we discuss general upper and lower bounds for $\mathsf{SCAN}(t, v)$. As we will see, $t!$ is a basic lower bound for $\mathsf{SCAN}(t, v)$. In Chapter 3.2, we discuss Levenshtein's Conjecture [36] which predicts the conditions under which $\mathsf{SCAN}(t, v) = t!$. We identify Levenshtein's Conjecture as a key motivation in the theory of sequence covering arrays which drives much of our research in Chapter 9. In Chapter 3.3, we discuss different algorithmic constructions of $\mathsf{SCA}(N; t, v)$ for small values of t.

In Chapter 4, we introduce perfect sequence covering arrays and other related combinatorial objects. As with sequence covering arrays, the main problem we consider is finding a perfect sequence covering array with the minimum number of permutations for fixed v and t. Here, the relevant parameter is $\mathsf{PSCAN}(v, t)$ which is the smallest λ for which a $\mathsf{PSCA}(v,t,\lambda)$ exists. In Chapter 4.1, we discuss several upper and lower bounds for $\mathsf{PSCAN}(v,t)$. In Chapter 4.2, we discuss computational methods that have been used to find exact values of $\mathsf{PSCAN}(v,t)$.

In Chapter 5, we introduce covering arrays, a related combinatorial object that also has applications in software testing. Although at first glance it may seem that the relationship between covering arrays and sequence covering arrays is limited to their similar applications, Chee et al. [9] describe how to construct covering arrays from sequence covering arrays and use this connection to show that if an $\mathsf{SCA}(t!;t,v)$ exists, then $v \leqslant 2t-1$. This is currently the best known upper bound on the number of symbols in such a sequence covering array for $t \geqslant 7$. We discuss this construction and the resulting upper bound in Chapter 5.1.

In Chapter 6 we introduce a number of related sequence covering problems. These problems often involve finding sets of permutations covering specific sets of sequences, rather than covering every sequence of a given length. Such problems arise in many different fields of mathematics, including coding theory, graph theory, extremal geometry and the theory of partial orders.

In Chapters 7, 8 and 9, we present original research into different aspects of sequence covering arrays and perfect sequence covering arrays. In Chapter 7, we present new computational methods that can exhaustively search for $\mathsf{PSCA}(v,t,\lambda)$ for any choice of parameters. These search methods are motivated by restrictions we prove in Chapter 7.2 on the number of times a given symbol may appear in each column of a $\mathsf{PSCA}(v,t,\lambda)$. We use these methods to catalogue all possible $\mathsf{PSCA}(v,t,\lambda)$ for different parameter choices. In particular, we are able to show $\mathsf{PSCAN}(6,3) = \mathsf{PSCAN}(7,3) = \mathsf{PSCAN}(7,4) = 2$ and $\mathsf{PSCAN}(8,3) = 3$. We discuss the results of these computations in Chapter 7.3. We also perform searches for perfect sequence covering arrays over permutation subgroups of $\mathcal{S}_v$ for $v \leqslant 14$ and find many permutation groups that form perfect sequence covering arrays. We draw connections between permutation group structure and sequence coverage in Chapter 7.4 and use these ideas to construct $\mathsf{PSCA}(v,3,\lambda)$ for $v \leqslant 32$ from cosets of elementary abelian 2-groups in Chapter 7.5.

In Chapter 8, we present a method of constructing perfect sequence covering arrays that can be applied for any v and t that makes use of projectivities of a finite projective plane. As described in Chapter 8.3, this method involves taking a suitable permutation representation of the group of projectivities and deleting all but a fixed number of symbols from each permutation. The size of the perfect sequence covering arrays built using this method grows polynomially in v for fixed t. This solves an open problem posed by Yuster [64]. Shortly after the publication of the paper Chapter 8 is based on, Iurlano [28] uncovered a connection between perfect sequence covering arrays and t-wise independent sets of permutations. This connection showed that Yuster's problem could be solved using prior research, in particular by a probabilistic argument of Kuperberg, Lovett and Peled [35]. Our construction requires the deletion of symbols from a permutation group. In

Chapter 8.4, we discuss what happens if we do not delete these symbols in the $t = 4$ case and show that particular permutation representations of the group of projectivities of a finite projective plane are capable of covering nearly all 4-sequences a fixed number of times.

In Chapter 9, we further explore the connections between sequence covering arrays and covering arrays outlined in Chapter 5. In Chapter 9.3, we describe exhaustive computational search methods for objects called excess coverage arrays introduced by Chee et al. [9], and we discuss the results of these computations. In particular, we show that an $\mathsf{SCA}(7!; 7, 10)$ does not exist. This result makes substantial progress towards proving the smallest open case of Levenshtein's Conjecture. In Chapter 9.4, we discuss similarities between excess coverage arrays and orthogonal arrays and suggest strategies of extending our computational results in Chapter 9.3 to apply to the $t = 8$ case of Levenshtein's Conjecture. We also consider excess coverage arrays for which the underlying ground set contains two symbols in Chapter 9.5.

We conclude this thesis in Chapter 10 with a discussion of open problems and directions for future research.

Chapter 2

Notation and preliminaries

Let v and t be positive integers with $v \geqslant t$. We let $[v] = \{0, \ldots, v-1\}$. We then let $\mathcal{S}_v$ be the set of permutations of $[v]$. Unless stated otherwise, permutations are assumed to be written in one-line notation with $\pi \in \mathcal{S}_v$ being denoted by $\pi(0)\pi(1)\cdots\pi(v-1)$. We let $\mathcal{S}_{v,t}$ be the set of ordered sequences of t distinct elements of $[v]$. Unless stated otherwise, we write a sequence $s \in \mathcal{S}_{v,t}$ as $(s_0, s_1, \ldots, s_{t-1})$.

Let $k \leqslant v$ and let $s = (s_0, \ldots, s_{k-1}) \in \mathcal{S}_{v,k}$ and $u = (u_0, \ldots, u_{t-1}) \in \mathcal{S}_{v,t}$. Then s *covers* u if there exists $\{\ell_0, \ldots, \ell_{t-1}\} \subseteq [k]$ such that $\ell_i < \ell_{i+1}$ for $0 \leqslant i \leqslant t-2$ and $u_i = s_{\ell_i}$ for $0 \leqslant i \leqslant t-1$. Note that for s to cover u, we require $k \geqslant t$. Typically, we only consider the case where $k = v$. Then, the sequence s is in fact a permutation in $\mathcal{S}_v$ and the following definition of coverage applies. A permutation $\pi \in \mathcal{S}_v$ *covers* $s \in \mathcal{S}_{v,t}$ if $\pi^{-1}(s_i) < \pi^{-1}(s_{i+1})$ for $0 \leqslant i \leqslant t-2$. Here $\pi^{-1}(i)$ is the unique $j \in [v]$ such that $\pi(j) = i$.

A *group* is a set X with a binary operation $\cdot$ that satisfies the following axioms.

(1) *Associativity:* for $a, b, c \in X$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

(2) *Identity:* there is an element $e \in X$ such that $a \cdot e = e \cdot a = a$ for all $a \in X$.

(3) *Inverse:* for each $a \in X$ there exists $a^{-1} \in X$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$.

We often omit the $\cdot$ from our notation and write $a \cdot b$ as ab.

For two permutations, $\pi, \sigma \in \mathcal{S}_v$, the *composition of* π *and* σ, denoted $\pi \circ \sigma$, is the permutation defined by $(\pi \circ \sigma)(i) = \pi(\sigma(i))$ for $0 \leqslant i \leqslant v-1$. Then the set $\mathcal{S}_v$ with the operation $\circ$ forms a group called the *symmetric group with degree* v. Note the identity permutation $e \in \mathcal{S}_v$ is defined as $e(i) = i$ for $0 \leqslant i \leqslant v-1$ and that π^{-1} as defined previously is indeed the inverse of π under $\circ$. A *permutation group* is a subgroup of some symmetric group.

For a group G with subgroup H and for $g \in G$, the *left coset of* H *with representative* g, denoted gH, is the set $\{gh : h \in H\}$. Similarly, the *right coset of* H *with representative* g, denoted Hg, is the set $\{hg : h \in H\}$.

For groups G and H, a *homomorphism from* G *to* H is a function $\phi : G \to H$ such that for all $g, h \in G$, $\phi(gh) = \phi(g)\phi(h)$. For a set X, let $\mathrm{Sym}(X)$ denote the group of permutations of X. Note that when $X = [v]$, $\mathrm{Sym}(X) = \mathcal{S}_v$. An *action* of a group G on X is a homomorphism $\phi : G \to \mathrm{Sym}(X)$. For $g \in G$ and $x \in X$, we use gx to refer to the image of x under the permutation $\phi(g)$. The *orbit* of x is the set $\mathrm{Orb}(x) = \{gx : g \in G\}$. The action of G on X is *transitive* if for all $x, y \in X$, there exists $g \in G$ such that $gx = y$. A permutation group $G \leqslant \mathcal{S}_v$ has the following natural action on $\mathcal{S}_{v,t}$. If $g \in G$ and $s \in \mathcal{S}_{v,t}$, then $gs = (g(s_0), \dots, g(s_{t-1}))$. A permutation group is *t-transitive* if its action on $\mathcal{S}_{v,t}$ is transitive.

Let X be a set. A *partial order on* X is a set $P \subseteq X \times X$ of pairs that satisfies the following properties.

(1) *Reflexivity:* for all $x \in X$, $(x, x) \in P$.

(2) *Antisymmetry:* for all $x, y \in X$, if $(x, y) \in P$ and $(y, x) \in P$, then $x = y$.

(3) *Transitivity:* for all $x, y, z \in X$, if $(x, y) \in P$ and $(y, z) \in P$, then $(x, z) \in P$.

A *linear order on* X is a partial order L such that for all $x, y \in X$, either $(x, y) \in L$ or $(y, x) \in L$. For a partial order P and a linear order L, L is a *linear extension of* P if $P \subseteq L$. A family $\mathcal{L}$ of linear extensions of P *realises* P if for any $x, y \in X$ such that $(x, y), (y, x) \notin P$, there are linear orders $L_1, L_2 \in \mathcal{L}$ such that $(x, y) \in L_1$ and $(y, x) \in L_2$. The *Dushnik-Miller dimension of* P [15] is the smallest cardinality of a family of linear extensions that realises P.

A *graph* is a pair $G = (V, E)$ containing a set of *vertices* V, and a set E of 2-subsets of V called *edges*. A *directed graph* is a pair consisting of a set of vertices V and a set of ordered pairs of distinct vertices called *directed edges*. A *hypergraph* is a pair consisting of a set of vertices V and a set of subsets of V called *hyperedges*. An *r-graph* is a hypergraph in which every hyperedge contains r vertices.

An *orthogonal array* with parameters λ, t, k, v, denoted $\mathsf{OA}_\lambda(t, k, v)$, is a $\lambda v^t \times k$ array with entries from $[v]$ such that in every t columns, every tuple of t elements of $[v]$ appears in exactly λ rows. We often omit the λ from our notation when $\lambda = 1$.

Chapter 3

Sequence Covering Arrays

In this chapter, we introduce sequence covering arrays, one of the main objects that we discuss in this thesis. Informally, a sequence covering array is a set of permutations in which every sequence of a given length appears in at least one permutation. As we will see, these objects are of interest in part because of their applications in software testing. We discuss several aspects of and problems relating to sequence covering arrays including existence results and construction methods that attempt to minimise the number of permutations needed for a sequence covering array.

A *sequence covering array* with *size* N, *order* v and *strength* t, denoted $\mathsf{SCA}(N; t, v)$, is a set X of N permutations in $\mathcal{S}_v$ such that for each sequence $s \in \mathcal{S}_{v,t}$, there exists $\pi \in X$ such that π covers s.

Sequence covering arrays were first introduced by Kuhn et al. [34] in 2012 as a way of generating testing suites for software systems. Kuhn et al. describe a scenario in which v peripheral devices have to be attached and installed into a central computer. It may be that the installation of peripheral B requires that peripheral A has already been installed or that peripheral C needs to be installed after both peripherals D and E. In any case, the operator needs to decide on an order in which to attach all devices, but an improper ordering may lead to faults. It is important to be able to screen the system for such faults so that it can be operated safely.

In general, consider a system in which v tasks are performed in some order where failures can occur due to improper ordering of these tasks. There are $v!$ ways of arranging these tasks so testing all of these orderings quickly becomes infeasible. As a compromise, we may assume that errors occur because of the improper ordering of at most t tasks. Then, a testing scheme based on a sequence covering array can be used to test all ways of ordering each subset of t tasks. In addition to the system described above, examples of testing applications of sequence covering arrays include memory management systems [34], graphical user interfaces [62], web and Android applications [59, 48], and water treatment processes [57]. Their applications in software testing make sequence covering arrays similar to covering arrays, which we discuss in Chapter 5.

Although sequence covering arrays have garnered much recent attention because of their applications, equivalent objects have been studied as far back as 1971 by Spencer [51]. A *completely scrambling set of permutations*, denoted $\mathsf{CSSP}(N;t,v)$ is a set X of N permutations in $\mathcal{S}_v$ such that for each $s \in \mathcal{S}_{v,t}$, there exists a $\pi \in X$ such that $\pi(s_i) < \pi(s_{i+1})$ for $0 \leqslant i \leqslant t-2$. The following lemma highlights the equivalence between sequence covering arrays and completely scrambling sets of permutations.

Lemma 3.1 ([9]). *A* $\mathsf{CSSP}(N;t,v)$ *is equivalent to an* $\mathsf{SCA}(N;t,v)$*.*

Proof. Let $s = (s_0, \ldots, s_{t-1}) \in \mathcal{S}_{v,t}$. By definition, a permutation π has $\pi(s_i) < \pi(s_{i+1})$ for $0 \leqslant i \leqslant t-2$ if and only if the inverse permutation of π covers s. Therefore, by taking the inverse of every permutation in a $\mathsf{CSSP}(N;t,v)$, we obtain an $\mathsf{SCA}(N;t,v)$. Similarly, by taking the inverse of every permutation of an $\mathsf{SCA}(N;t,v)$, we obtain a $\mathsf{CSSP}(N;t,v)$. □

Another way of interpreting Lemma 3.1 is to consider an $\mathsf{SCA}(N;t,v)$ as an $N \times v$ array A where for each permutation π in the SCA, there is a row r of A such that $\mathsf{A}[r,c] = \pi(c)$ for $c \in [v]$. For each entry in A, we assign a triple (r,c,i), where $r \in [N]$ is a row of A, $c \in [v]$ is a column of A, and $i \in [v]$ is the symbol of A such that $\mathsf{A}[r,c] = i$. Then, by swapping c and i in each of these triples, we transform A from an array whose rows are the permutations of an $\mathsf{SCA}(N;t,v)$ to an array whose rows are the permutations of a $\mathsf{CSSP}(N;t,v)$. We can transform a $\mathsf{CSSP}(N;t,v)$ to an $\mathsf{SCA}(N;t,v)$ in the same manner.

One goal, particularly driven by the need to develop efficient testing schemes, is to minimise the number of permutations in a sequence covering array. Let $\mathsf{SCAN}(t,v)$ be the smallest N for which an $\mathsf{SCA}(N;t,v)$ exists. In Chapter 3.1, we discuss upper and lower bounds on $\mathsf{SCAN}(t,v)$. Given a t-subset $T \subseteq [v]$, there are $t!$ ways of arranging the elements of T. In order to cover all of these arrangements, a sequence covering array requires at least $t!$ permutations. Several authors have considered the problem of determining when $\mathsf{SCAN}(t,v) = t!$. We discuss this problem in Chapter 3.2. In Chapter 3.3, we discuss and compare several algorithmic constructions of sequence covering arrays for small values of t.

3.1 Bounds on SCAN

In this section we discuss existing bounds on $\mathsf{SCAN}(t,v)$. As we saw in the previous section, $\mathsf{SCAN}(t,v) \geqslant t!$. This bound can be trivially met when $v = t$ by observing that $\mathcal{S}_v$ is an $\mathsf{SCA}(v!;v,v)$. When $t = 2$, we can build an $\mathsf{SCA}(2;2,v)$ by taking any permutation in $\mathcal{S}_v$ and its reverse. Thus, $\mathsf{SCAN}(2,v) = 2$ for all $v \geqslant 2$. Henceforth, we focus our attention on $\mathsf{SCAN}(t,v)$ for $3 \leqslant t \leqslant v-1$. For $4 \leqslant t \leqslant v-1$ and $s \in \mathcal{S}_{v,t-1}$, there is some $s' \in \mathcal{S}_{v,t}$ that covers s. Therefore, an $\mathsf{SCA}(N;t,v)$ is also an $\mathsf{SCA}(N;t-1,v)$ and so $\mathsf{SCAN}(t,v) \geqslant \mathsf{SCAN}(t-1,v)$. For $v > t$, if the symbol $v-1$ is deleted from every permutation of an $\mathsf{SCA}(N;t,v)$, then we obtain an $\mathsf{SCA}(N;t,v-1)$. Therefore, $\mathsf{SCAN}(t,v) \geqslant \mathsf{SCAN}(t,v-1)$. We begin with upper and lower bounds for $\mathsf{SCAN}(t,v)$ first proved by Spencer [51].

Theorem 3.2 ([51]). *For $v \geqslant t \geqslant 3$,*

$$\mathsf{SCAN}(t,v) \leqslant \frac{t\log_2(v)}{\log_2(\frac{t!}{t!-1})}.$$

Proof. For each sequence $s \in \mathcal{S}_{v,t}$, there are $\binom{v}{t}(v-t)! = v!/t!$ permutations in $\mathcal{S}_v$ that cover s. Hence, a permutation chosen from $\mathcal{S}_v$ uniformly at random covers some sequence in $\mathcal{S}_{v,t}$ with probability $1/t!$. Therefore, it fails to cover that sequence with probability $(t!-1)/t!$. The probability that a sequence is not covered by any of N permutations chosen independently and uniformly at random is thus $((t!-1)/t!)^N$. There are $\binom{v}{t}t! < v^t$ sequences in $\mathcal{S}_{v,t}$ so the expected number of sequences that are uncovered by all of the N permutations is less than $v^t((t!-1)/t!)^N$. When this expected value falls below 1, there must be a set of N permutations for which the number of uncovered sequences is zero. This set of permutations must then be an $\mathsf{SCA}(N;t,v)$. The inequality $v^t((t!-1)/t!)^N \leqslant 1$ holds when

$$N \geqslant \frac{t\log_2(v)}{\log_2(\frac{t!}{t!-1})}.$$ □

The following definition is required for our next theorem and will be important in our discussion of 3-*mixing permutations* later in this chapter. Let $v \geqslant 3$. For a set of permutations $X \subseteq \mathcal{S}_v$, $x \in [v]$ and $A \subseteq X$, we let $L_X(x,A)$ be the set of symbols $y \in [v] \setminus \{x\}$ such that for $\pi \in X$, $\pi^{-1}(y) < \pi^{-1}(x)$ if and only if $\pi \in A$. Note that for $y \in [v] \setminus \{x\}$, there is a unique A such that $y \in L_X(x,A)$.

Theorem 3.3 ([51]). *For $v \geqslant t \geqslant 3$,*

$$\mathsf{SCAN}(t,v) \geqslant \log_2 v.$$

Proof. Let X be an $\mathsf{SCA}(N;t,v)$ and let $x \in [v]$. Suppose there exists $A \subseteq X$ such that $|L_X(x,A)| > 1$. Let y and z be distinct elements of $L_X(x,A)$. Then in each permutation in X, y and z either both appear after x or both appear before x. Hence there is no permutation in X that covers (y,x,z) and so X is not an $\mathsf{SCA}(N;3,v)$. This contradicts X being an $\mathsf{SCA}(N;t,v)$. Therefore, for each $A \subseteq X$, $|L_X(x,A)| \leqslant 1$.

Now fix $\pi \in X$ and let $x = \pi(0)$. Then if $\pi \in A$, $|L_X(x,A)| = 0$. Therefore, for each $y \in [v] \setminus \{x\}$, there is a unique $A \subseteq X \setminus \{\pi\}$ such that $y \in L_X(x,A)$. Since $|L_X(x,A)| \leqslant 1$ for all $A \subseteq X \setminus \{\pi\}$, we have $2^{N-1} \geqslant v-1$. Hence, $N \geqslant 1 + \log_2(v-1) \geqslant \log_2 v$. □

Already, we can see that for fixed t, $\mathsf{SCAN}(t,v)$ grows logarithmically in v. The best currently known upper bound for $\mathsf{SCAN}(t,v)$ is due to Yuster [63] through an application of the Lovász Local Lemma.

Lemma 3.4 (Lovász Local Lemma [16, 49]). *Let $A_1, A_2, \ldots, A_n$ be a sequence of random events, each occurring with probability at most p, such that each event is independent with all but at most d of the other events. If $epd \leqslant 1$, where e is the base of the natural logarithm, then there is a non-zero probability that none of the events occurs.*

Theorem 3.5 ([63])**.** *For $v \geqslant t \geqslant 3$,*

$$\mathsf{SCAN}(t, v) \leqslant \ln(2)t!(t-1)\log_2(v) + \ln(2)t!\log_2(et^2).$$

Proof. Let S be a set of N permutations in $\mathcal{S}_v$, each chosen uniformly at random. For a sequence $s \in \mathcal{S}_{v,t}$, let A_s be the event that s is not covered by any permutation in S. Then the probability that A_s occurs is $p = (1 - 1/t!)^N$. For sequences $s, u \in \mathcal{S}_{v,t}$, the events A_s and A_u are independent if s and u have no elements of $[v]$ in common. Let d be the number of sequences in $\mathcal{S}_{v,t}$ that share an element with s. Then $d < t!t\binom{v-1}{t-1} < t^2v^{t-1}$. By the Lovász Local Lemma, if $epd \leqslant 1$, then there must exist an $\mathsf{SCA}(N; t, v)$. Now, when $N \geqslant \ln(2)t!\log_2(ed)$,

$$epd = ed\left(1 - \frac{1}{t!}\right)^N \leqslant ed\left(\left(1 - \frac{1}{t!}\right)^{t!}\right)^{\ln(2)\log_2(ed)} \leqslant ed\left(\frac{1}{2}^{\log_2(ed)}\right) = 1.$$

Substituting $d < t^2v^{t-1}$, we find that there exists an $\mathsf{SCA}(N; t, v)$ when

$$N \geqslant \ln(2)t!(t-1)\log_2(v) + \ln(2)t!\log_2(et^2). \qquad \square$$

The best currently known lower bound for $\mathsf{SCAN}(t, v)$ is due to Radhakrishnan [47]. The proof of this result makes use of the *binary entropy function* $H : [0, 1] \to [0, 1]$, which is defined by

$$H(x) = \begin{cases} 0, & \text{if } x \in \{0, 1\} \\ -x\log_2 x - (1-x)\log_2(1-x), & \text{if } x \in (0, 1). \end{cases}$$

The proof of Radhakrishnan's bound also requires the *content* of a bipartite graph. Let G be a bipartite graph with n vertices, let W be the set of vertices with positive degree in G, and let $w = |W|$. Let $\chi : W \to \{0, 1\}$ be a 2-colouring of the subgraph of G induced by W that has the smallest colour class among all 2-colourings of the subgraph of G induced by W. If G is non-empty, then the *content* of G is defined by

$$c(G) = \left(\frac{w}{n}\right) H\left(\frac{|\chi^{-1}(0)|}{w}\right).$$

If G is empty, then we set $c(G) = 0$. Note that since $H(x) = H(1-x)$ for $x \in [0, 1]$, it does not matter which colour class of χ is the minimum over all colourings. In particular, regardless of whether $\chi^{-1}(0)$ is the largest or smallest possible colour class, $c(G)$ will still have the same value. The following lemma is due to Fredman and Komlós [17].

Lemma 3.6 ([17])**.** *Let $G_0, G_1, \ldots, G_k$ be bipartite graphs each with vertex set $[n]$ whose* *union is the complete graph on* $[n]$*. Then,*

$$\sum_{i=0}^{k} c(G_i) \geqslant \log_2 n.$$

Theorem 3.7 ([47])**.** *For* $v \geqslant t \geqslant 3$*,*

$$\mathsf{SCAN}(t, v) \geqslant \left(\frac{2}{\log_2 e}\right)(t-1)!\left(\frac{v}{2v-t+1}\right)\log_2(v-t+2).$$

Proof. Let X be an $\mathsf{SCA}(N; t, v)$ and let $s = (s_0, \ldots, s_{t-3}) \in \mathcal{S}_{v,t-2}$. For $\pi \in X$, let $G_s(\pi)$ be the graph whose vertex set is $[v] \setminus \{s_0, \ldots, s_{t-3}\}$ and whose edge set is empty if π does not cover s and is

$$\{\{i, j\} : \pi^{-1}(i) < \pi^{-1}(s_0) < \cdots < \pi^{-1}(s_{t-3}) < \pi(j)\}$$

otherwise. Note that this set is also empty if $\pi^{-1}(s_0) = 0$ or $\pi^{-1}(s_{t-3}) = v - 1$. Suppose π covers s with $\pi^{-1}(s_0) > 0$ and $\pi^{-1}(s_{t-3}) < v - 1$. Then the set W of vertices with positive degree in $G_s(\pi)$ is $W = \{i \in [v] : \pi^{-1}(i) < \pi^{-1}(s_0) \text{ or } \pi^{-1}(i) > \pi^{-1}(s_{t-3})\}$. So $w = |W| = v - 1 - \pi^{-1}(s_{t-3}) + \pi^{-1}(s_0)$. The only way to 2-colour L is to assign all symbols appearing before s_0 in π one colour and assign all symbols appearing after s_{t-3} in π another. Therefore,

$$c\left(G_s(\pi)\right) = \left(\frac{v - 1 - \pi^{-1}(s_{t-3}) + \pi^{-1}(s_0)}{v - t + 2}\right) H\left(\frac{\pi^{-1}(s_0)}{v - 1 - \pi^{-1}(s_{t-3}) + \pi^{-1}(s_0)}\right).$$

Furthermore, as X is an $\mathsf{SCA}(N; t, v)$, for each distinct $i, j \in [v]$, there exists a permutation $\pi \in X$ that covers $(i, s_0, \ldots, s_{t-3}, j)$. Hence, $\{i, j\}$ is an edge of $G_s(\pi)$. Therefore, the union of $G_s(\pi)$ as π ranges over X is the complete graph on $[v] - \{s_0, \ldots, s_{t-3}\}$. Thus, Lemma 3.6 can be applied to give

$$\sum_{\pi \in X} c(G_s(\pi)) \geqslant \log_2(v - t + 2).$$

Then, by performing a series of calculations that bound the expected value of $c(G_s(\pi))$ when s is sampled uniformly at random from $\mathcal{S}_{v,t-2}$, one can obtain the bound in the theorem statement. These calculations can be found in [47] but we omit them here. □

Note that in the proof of Theorem 3.7, we do not actually require the set X to be an $\mathsf{SCA}(N; t, v)$. We can relax that condition to require that for a sequence $s = (s_0, \ldots, s_{t-3}) \in \mathcal{S}_{t-2}$ and for distinct $i, j \in [v] - \{s_0, \ldots, s_{t-3}\}$, there is a permutation in X that covers either $(i, s_0, \ldots, s_{t-3}, j)$ or $(j, s_0, \ldots, s_{t-3}, i)$. If X satisfies this condition and we define $G_s(\pi)$ as in the proof of Theorem 3.7, it is still possible to apply Lemma 3.6. Indeed, this is the way the theorem is proved in [47].

This condition can be seen as a way of generalising 3-*mixing sets of permutations*, which were introduced by Füredi [18]. Let $v \geqslant 3$. A set $X \subseteq \mathcal{S}_v$ is 3-*mixing* if for any distinct $i, j, k \in [v]$, there is a permutation in X that covers either (i, j, k) or (k, j, i). Clearly, an $\mathsf{SCA}(N; 3, v)$ is 3-mixing and if X is 3-mixing, an $\mathsf{SCA}(2|X|; 3, v)$ can be constructed by taking each of the permutations in X and their reverses. For $N \geqslant 3$, we let $g(N)$ be the maximum v for which there exists a 3-mixing set of N permutations in $\mathcal{S}_v$. The restriction $N \geqslant 3$ is applied because there are no 3-mixing sets of N permutations for $N < 3$. We have $g(3) = 3$ [24], $g(4) = 4$ [24] and $g(5) = 7$ [24, 18].

The concept of 3-mixing sets of permutations was introduced as a way of codifying as a problem on permutations a geometric problem on containment of points in $\mathbb{R}^d$ within so-called *orthants* posed by Ishigami [23, 24]. Let X be a set of permutations in $\mathcal{S}_v$. Recall that for $x \in [v]$ and $A \subseteq X$, $L_X(x, A)$ is the set of symbols $y \in [v] \setminus \{x\}$ such that for $\pi \in X$, $\pi^{-1}(y) < \pi^{-1}(x)$ if and only if $\pi \in A$. We then let $L(X) = \max\{|L_X(x, A)| : x \in [v] \text{ and } A \subseteq X\}$ and $l(v, N) = \min\{L(X) : X \subseteq \mathcal{S}_v \text{ and } |X| = N\}$. In the proof of Theorem 3.3, we showed that if X is an $\mathsf{SCA}(N; t, v)$ then $L(X) = 1$. The next lemma shows that $L(X) = 1$ if and only if X is 3-mixing.

Lemma 3.8. *For integers $v, N \geqslant 3$, there exists a 3-mixing set of N permutations in $\mathcal{S}_v$ if and only if $l(v, N) = 1$.*

Proof. Suppose $l(v, N) > 1$ and let X be a set of N permutations in $\mathcal{S}_v$. Then as $l(v, N) > 1$, $L(X) > 1$ and so there exists $x \in [v]$ and $A \subseteq X$ such that $L_X(x, A) > 1$. Let y and z be distinct elements of $L_X(x, A)$. Then in every permutation in X, y and z either both appear before x or both appear after x. Hence, there is no permutation in X that covers either (y, x, z) or (z, x, y). Therefore, X is not 3-mixing.

Now suppose $l(v, N) = 1$ and let X be a set of N permutations in $\mathcal{S}_v$ such that $L(X) = 1$. Let $x, y, z \in [v]$. There exists some $A \subseteq X$ such that $y \in L_X(x, A)$. Since $L(X) = 1$, $z \notin L_X(x, A)$. Therefore, there is a permutation in X in which y and z are not both before or both after x. Hence, this permutation covers either (y, x, z) or (z, x, y). Therefore, X is 3-mixing. □

As a consequence of Lemma 3.8, $g(N) = \max\{v : l(v, N) = 1\}$. Ishigami [24] proved the following relationship between $v, l(v, N)$ and $g(N)$.

Lemma 3.9 ([24]). *For integers $v, N \geqslant 3$,*

$$l(v, N) \leqslant \left\lceil \frac{v}{g(N)} \right\rceil .$$

Recall from Chapter 2 that a *hypergraph with vertex set* V is a collection $\mathcal{F}$ of subsets of V. Here we allow for repetition of subsets in $\mathcal{F}$. For $X \subseteq V$, we define $\mu(X)$ to be the number of times X appears in $\mathcal{F}$. We then let $\mu(\mathcal{F}) = \max_{X \subseteq V} \mu(X)$.

The following lemma was shown by Kleitman, Shearer and Sturtevant [32].

Lemma 3.10 ([32]). *Let $\mathcal{F}$ be a hypergraph with vertex set V. For $x \in V$, let $p(x)$ be the fraction of sets in $\mathcal{F}$ which contain x. Then,*

$$\sum_{x \in V} H(p(x)) \geqslant \log_2 \frac{|\mathcal{F}|}{\mu(\mathcal{F})},$$

where H is the binary entropy function.

Füredi [18] used the previous two lemmas to prove the following theorem.

Theorem 3.11 ([18]). *For all integers $N \geqslant 3$, $g(N) < e^{N/2}$.*

Proof. We prove this theorem by establishing the stronger result that for $v, N \geqslant 3$, $l(v, N) > v \exp(-N/2)$. Then, by substituting this bound into Lemma 3.9 with $v = g(N)$, we obtain $g(N) < \exp(N/2)$ as required.

Let $X = \{\pi_0, \ldots, \pi_{N-1}\}$ be a set of N permutations in $\mathcal{S}_v$ and let $l = l(v, N)$. Fix $x \in [v]$. For $y \in [v] \setminus \{x\}$, let $A_y = \{i \in [N] : \pi_i^{-1}(y) < \pi_i^{-1}(x)\}$. Note that by definition, $y \in L_X(x, A_y)$. Then, we define the hypergraph $\mathcal{F}(x) = \{A_y : y \in [v] \setminus \{x\}\}$ with vertex set $[N]$. We have $|\mathcal{F}(x)| = v - 1$ and $\mu(\mathcal{F}(x)) \leqslant l$. For $i \in [N]$, the number of sets in $\mathcal{F}(x)$ that contain i is $\pi_i^{-1}(x)$. Hence, by Lemma 3.10,

$$\log_2 \frac{v-1}{l} \leqslant \sum_{i \in [N]} H\left(\frac{\pi_i^{-1}(x)}{v-1}\right).$$

By adding the above inequalities for all $x \in [v]$, we obtain

$$\begin{aligned} v \log_2 \frac{v-1}{l} &\leqslant \sum_{x \in [v]} \sum_{i \in [N]} H\left(\frac{\pi_i^{-1}(x)}{v-1}\right) \\ &= \sum_{i \in [N]} \sum_{x \in [v]} H\left(\frac{\pi_i^{-1}(x)}{v-1}\right) \\ &= N \sum_{j=0}^{v-1} H\left(\frac{j}{v-1}\right). \end{aligned}$$

Since the binary entropy function is known to be concave, $H(z) = H(1-z)$ for $z \in [0, 1]$, and $H(1) = 0$,

$$\sum_{j=0}^{v-1} H\left(\frac{j}{v-1}\right) = \sum_{j=0}^{v-2} H\left(\frac{j}{v-1}\right) < (v-1) \int_0^1 H(z) dz = \frac{v-1}{2} \log_2 e.$$

Therefore,

$$\begin{aligned} v \log_2 \frac{v-1}{l} &< \frac{N(v-1)}{2} \log_2 e \\ \frac{v-1}{l} &< \exp\left(\frac{N(v-1)}{2v}\right) \\ l &> (v-1)\exp\left(-\frac{N(v-1)}{2v}\right). \end{aligned}$$

Now, for fixed v and N, the derivative of $x\exp(-Nx/2v)$ is $(1-Nx/2v)\exp(-Nx/2v)$. If $v, N \geqslant 3$, then this derivative is at most 0 for all $x \geqslant v-1$. Therefore, for $v, N \geqslant 3$,

$$l(v,N) > (v-1)\exp\left(-\frac{N(v-1)}{2v}\right) \geqslant v\exp\left(-\frac{N}{2}\right). \qquad \square$$

This result leads to the following bound on $\mathsf{SCAN}(3,v)$.

Theorem 3.12 ([18]). *For all integers $v \geqslant 3$,*

$$\mathsf{SCAN}(3,v) \geqslant \frac{2}{\log_2 e}\log_2 v.$$

Proof. Let X be an $\mathsf{SCA}(N;3,v)$. Then X is 3-mixing and so by Theorem 3.11, $v < e^{N/2}$. Therefore,

$$N > 2\ln v = \frac{2}{\log_2 e}\log_2 v. \qquad \square$$

As an extension of $g(N)$, Tarui [53] defined $f(N)$ to be the largest v such that an $\mathsf{SCA}(N;3,v)$ exists. Then, analogously to Theorem 3.12, bounds on $f(N)$ can be used to bound $\mathsf{SCAN}(3,v)$. Note that since an $\mathsf{SCA}(N;3,v)$ must have at least 6 permutations, we only consider $f(N)$ for $N \geqslant 6$. The following bound on $f(N)$ was proved constructively by Tarui [53].

Theorem 3.13 ([53]). *For $N \geqslant 6$,*

$$f(N) \geqslant \binom{\lfloor N/2 \rfloor}{\lfloor N/4 \rfloor}.$$

Proof. Let $N \geqslant 6$, let $r = \lfloor N/2 \rfloor$ and let

$$w = \binom{\lfloor N/2 \rfloor}{\lfloor N/4 \rfloor}.$$

We then construct an $\mathsf{SCA}(2r;3,w)$. As $2r \leqslant N$, this also implies the existence of an $\mathsf{SCA}(N;3,w)$ which in turn proves the theorem. Let $\mathcal{F} = \{A_0, \ldots, A_{w-1}\}$ be the set of all subsets of $[r]$ with cardinality $\lfloor N/4 \rfloor$. For each $x \in [r]$, we define two permutations, $\pi_x, \sigma_x \in \mathcal{S}_w$. Let $i, j \in [w]$ be distinct. If $x \in A_i$ and $x \notin A_j$, then both $\pi_x^{-1}(i) < \pi_x^{-1}(j)$

and $\sigma_x^{-1}(i) < \sigma_x^{-1}(j)$. If either $x \in A_i, A_j$ or $x \notin A_i, A_j$, then $\pi_x^{-1}(i) < \pi_x^{-1}(j)$ if and only if $i < j$ while $\sigma_x^{-1}(i) < \sigma_x^{-1}(j)$ if and only if $i > j$. Intuitively, in both π_x and σ_x, the indices of the sets containing x appear before the indices of the sets that do not contain x. In π_x, the indices of the sets containing x and the indices of the sets that do not contain x are then arranged in ascending order. In σ_x, the indices of the sets containing x and the indices of the sets that do not contain x are arranged in descending order. We claim $P = \bigcup_{x\in[r]}\{\pi_x, \sigma_x\}$ is an $\mathsf{SCA}(2r; 3, w)$.

By construction, for distinct i, j, $A_i \setminus A_j$ is non-empty. Let $(i, j, k) \in \mathcal{S}_{w,3}$ and let $x \in A_i \setminus A_k$. Suppose $x \in A_j$. If $i < j$, then π_x covers (i, j, k) and if $i > j$, then σ_x covers (i, j, k). Either way, there is a permutation in P that covers (i, j, k). Now suppose $x \notin A_j$. If $j < k$, then π_x covers (i, j, k) and if $j > k$, then σ_x covers (i, j, k). Again, there is a permutation in P that covers (i, j, k). Therefore, in all possible cases there is a permutation in P that covers (i, j, k) and so P is an $\mathsf{SCA}(2r; 3, w)$. □

Theorem 3.14 ([53])**.** *For $v \geqslant 3$, $\mathsf{SCAN}(3, v) \leqslant 2\log_2 v + (1 + o(1))\log_2\log_2 v$.*

Proof. By Theorem 3.13, for $v \geqslant 3$, there exists an $\mathsf{SCA}(N; 3, v)$ with

$$v \geqslant \binom{\lfloor N/2\rfloor}{\lfloor N/4\rfloor}.$$

Now,

$$\binom{\lfloor N/2\rfloor}{\lfloor N/4\rfloor} \geqslant \binom{2\lfloor N/4\rfloor}{\lfloor N/4\rfloor}.$$

By the Binomial Theorem,

$$2^{2\lfloor N/4\rfloor} = \sum_{i=0}^{2\lfloor N/4\rfloor}\binom{2\lfloor N/4\rfloor}{i}.$$

Since

$$\binom{2\lfloor N/4\rfloor}{\lfloor N/4\rfloor} \geqslant \binom{2\lfloor N/4\rfloor}{i}$$

for $0 \leqslant i \leqslant 2\lfloor N/4\rfloor$, we have

$$\binom{2\lfloor N/4\rfloor}{\lfloor N/4\rfloor} \geqslant \frac{2^{2\lfloor N/4\rfloor}}{2\lfloor N/4\rfloor + 1}.$$

Therefore,

$$v \geqslant \frac{2^{2(N/4-1)}}{N/2+1} = \frac{2^{N/2}}{2N+4}.$$

Hence,

$$\log_2 v \geqslant \frac{N}{2} - \log_2(2N+4).$$

From this we can see that $N \leqslant (1 + o(1))2\log_2 v$. Therefore,

$$N \leqslant 2\log_2 v + (1 + o(1))\log_2\log_2 v. \qquad \square$$

We conclude this section by considering the asymptotics of $\mathsf{SCAN}(t, v)$. Given that it was known as far back as 1971 that $\mathsf{SCAN}(t, v)$ grows logarithmically in v for fixed t, a natural question is to ask about the existence of

$$\lim_{v\to\infty} \frac{\mathsf{SCAN}(t, v)}{\log_2 v}$$

for fixed t. This question was first posed by Ishigami [25]. Füredi [18] first proved that $\lim_{N\to\infty} g(N)^{1/N}$ is finite. Let $g^*(N)$ be the largest v such that there exists a set of N permutations that is both 3-mixing and an $\mathsf{SCA}(N; 2, v)$. Then $g^*(N) \leqslant g(N)$. A set X of 3-mixing permutations with $|X| = N$ can be extended to an $\mathsf{SCA}(N + 1; 2, v)$ by adding the reverse of one of the permutations in X. Hence, $g^*(N) \geqslant g(N - 1)$. Through an appropriate recursive construction, Füredi shows that $g^*(N_1 + N_2) \geqslant g^*(N_1)g^*(N_2)$. Then $\log(g^*(N_1 + N_2)) \geqslant \log(g^*(N_1)) + \log(g^*(N_2))$. By Fekete's Theorem, $\lim_{N\to\infty}(1/N)\log(g^*(N))$ exists and is equal to its supremum. By Theorem 3.11, this supremum must be finite. Therefore, $\lim_{N\to\infty} g^*(N)^{1/N}$ must also be finite. Then, since $g^*(N) \leqslant g(N) \leqslant g^*(N + 1)$, $\lim_{N\to\infty} g(N)^{1/N}$ is finite too.

This proof was then generalised by Tarui [53] to show that $\lim_{N\to\infty}(1/N)\log_2 f(N)$ is finite. This proof included the introduction of a new sequence $f^*(N)$ with similar properties to $g^*(N)$. However, given that an $\mathsf{SCA}(N; 3, v)$ is also an $\mathsf{SCA}(N; 2, v)$, the definition of $f^*(N)$ cannot be strictly analogous to $g^*(N)$. Instead, we say an $\mathsf{SCA}(N; 3, v)$ is 2-*reversing* if it can be partitioned into a *red* $\mathsf{SCA}(M_1; 2, v)$ and a *blue* $\mathsf{SCA}(M_2; 2, v)$ with $M_1 + M_2 = N$. Then, let $f^*(N)$ be the largest v such that there exists a 2-reversing $\mathsf{SCA}(N; 3, v)$. Then, $f^*(N) \leqslant f(N) \leqslant g(N)$. An $\mathsf{SCA}(N; 3, v)$ can be extended to a 2-reversing $\mathsf{SCA}(N + 2; 3, v)$ by adding the reverses of two distinct permutations (and assigning one permutation-reverse pair the colour red and the other the colour blue). Therefore, $f^*(N) \leqslant f(N) \leqslant f^*(N + 2)$.

Theorem 3.15 ([53]). *The limit*

$$\lim_{N\to\infty} \frac{\log_2 f(N)}{N}$$

is finite.

Proof. First we show by construction that $f^*(N_1 + N_2) \geqslant f^*(N_1)f^*(N_2)$. Let $P_1 = \{\pi_0, \ldots, \pi_{N_1-1}\}$ be a 2-reversing $\mathsf{SCA}(N_1; 3, v_1)$ and let $P_2 = \{\sigma_0, \ldots, \sigma_{N_2-1}\}$ be a 2-reversing $\mathsf{SCA}(N_2; 3, v_2)$. We then construct a 2-reversing $\mathsf{SCA}(N_1 + N_2; 3, v_1v_2)$ with ground set $[v_1] \times [v_2]$. Let $x = (i_1, j_1)$ and $y = (i_2, j_2)$ be distinct elements of $[v_1] \times [v_2]$. For $k \in [N_1]$, we define $\tilde{\pi}_k$ according to the following rules.

(1) If $i_1 \neq i_2$, then $\tilde{\pi}_k^{-1}(x) < \tilde{\pi}_k^{-1}(y)$ if and only if $\pi_k^{-1}(i_1) < \pi_k^{-1}(i_2)$.

(2) If $i_1 = i_2$ and π_k is red, then $\tilde{\pi}_k^{-1}(x) < \tilde{\pi}_k^{-1}(y)$ if and only if $j_1 < j_2$.

(3) If $i_1 = i_2$ and π_k is blue, then $\tilde{\pi}_k^{-1}(x) < \tilde{\pi}_k^{-1}(y)$ if and only if $j_1 > j_2$.

Moreover, $\tilde{\pi}_k$ is assigned the same colour as π_k. For $k \in [N_2]$, we define $\tilde{\sigma}_k$ according to the following rules.

(1) If $j_1 \neq j_2$, then $\tilde{\sigma}_k^{-1}(x) < \tilde{\sigma}_k^{-1}(y)$ if and only if $\sigma_k^{-1}(j_1) < \sigma_k^{-1}(j_2)$.

(2) If $j_1 = j_2$ and σ_k is red, then $\tilde{\sigma}_k^{-1}(x) < \tilde{\sigma}_k^{-1}(y)$ if and only if $i_1 < i_2$.

(3) If $j_1 = j_2$ and σ_k is blue, then $\tilde{\sigma}_k^{-1}(x) < \tilde{\sigma}_k^{-1}(y)$ if and only if $i_1 > i_2$.

Moreover, $\tilde{\sigma}_k$ is assigned the same colour as σ_k. We claim that $\tilde{P} = \{\tilde{\pi}_0, \dots, \tilde{\pi}_{N_1-1}, \tilde{\sigma}_0, \dots, \tilde{\sigma}_{N_2-1}\}$ is a 2-reversing $\mathsf{SCA}(N_1 + N_2; 3, v_1 v_2)$.

First we show $\tilde{P}$ is a strength 3 sequence covering array. Let $x = (i_1, j_1)$, $y = (i_2, j_2)$ and $z = (i_3, j_3)$ be distinct elements of $[v_1] \times [v_2]$. If i_i, i_2, i_3 are all distinct, then since P_1 is an $\mathsf{SCA}(N_1; 3, v_1)$, we can find six permutations in P_1, each covering a distinct ordering of $\{i_1, i_2, i_3\}$. Therefore, the corresponding six permutations in $\tilde{P}$ will also cover the six possible orderings of $\{x, y, z\}$. A similar argument applied to P_2 shows that there are permutations in $\tilde{P}$ covering each of the orderings of $\{x, y, z\}$ if j_1, j_2, j_3 are all distinct. The only remaining case to consider is if $|\{i_1, i_2, i_3\}| = |\{j_1, j_2, j_3\}| = 2$. Without loss of generality, assume $i_1 = i_2$ and $j_1 = j_3$. The following four cases show that there are permutations in $\tilde{P}$ covering each of the six possible orderings of $\{x, y, z\}$.

(1) (x, y, z) and (y, x, z): we choose a red π_k with $\pi_k^{-1}(i_1) < \pi_k^{-1}(i_3)$ and a blue π_ℓ such that $\pi_\ell^{-1}(i_1) < \pi_\ell^{-1}(i_3)$. One of $\{\tilde{\pi}_k, \tilde{\pi}_\ell\}$ covers (x, y, z) and the other covers (y, x, z).

(2) (z, x, y) and (z, y, x): we choose a red π_k with $\pi_k^{-1}(i_1) > \pi_k^{-1}(i_3)$ and a blue π_ℓ such that $\pi_\ell^{-1}(i_1) > \pi_\ell^{-1}(i_3)$. One of $\{\tilde{\pi}_k, \tilde{\pi}_\ell\}$ covers (z, x, y) and the other covers (z, y, x).

(3) (x, z, y): we choose a red σ_k with $\sigma_k^{-1}(j_1) < \sigma_k^{-1}(j_2)$ and a blue σ_ℓ such that $\sigma_\ell^{-1}(j_1) < \sigma_\ell^{-1}(j_2)$. Then one of $\{\tilde{\sigma}_k, \tilde{\sigma}_\ell\}$ covers (x, z, y).

(4) (y, z, x): we choose a red σ_k with $\sigma_k^{-1}(j_1) > \sigma_k^{-1}(j_2)$ and a blue σ_ℓ such that $\sigma_\ell^{-1}(j_1) > \sigma_\ell^{-1}(j_2)$. Then one of $\{\tilde{\sigma}_k, \tilde{\sigma}_\ell\}$ covers (y, z, x).

Since P_1 and P_2 are 2-reversing, the choice of permutations made in each case is possible. Therefore, $\tilde{P}$ is an $\mathsf{SCA}(N_1 + N_2; 3, v_1 v_2)$.

Now let $x = (i_1, j_1), y = (i_2, j_2)$ be distinct elements of $[v_1] \times [v_2]$. If $i_1 \neq i_2$, then as P_1 is 2-reversing, there are red and blue permutations in $\tilde{P}$ that cover (x, y) and (y, x). The same is true if $j_1 \neq j_2$ as P_2 is 2-reversing. Hence, $\tilde{P}$ is 2-reversing and so $f^*(N_1 + N_2) \geqslant f^*(N_1) f^*(N_2)$. By Theorem 3.11 and since $f^*(N) \leqslant g(N)$, the sequence $((1/N) \log_2 f^*(N))$ is bounded above. Therefore $\lim_{N \to \infty} (1/N) \log_2 f^*(N)$ exists and is finite. Finally, as $f^*(N) \leqslant f(N) \leqslant f^*(N+2)$, $\lim_{N \to \infty} (1/N) \log_2 f(N)$ is also finite. □

Corollary 3.16 ([53]). *The limit*

$$\lim_{v\to\infty} \frac{\mathsf{SCAN}(3,v)}{\log_2 v}$$

is finite.

The value of this limit is undetermined. The finiteness of the corresponding limit for $t \geqslant 4$ is also undetermined.

3.2 Levenshtein's Conjecture

As stated previously, $\mathsf{SCAN}(t,v) \geqslant t!$. As $\mathcal{S}_t$ is an $\mathsf{SCA}(t!;t,t)$, $\mathsf{SCAN}(t,t) = t!$. Similarly, when $t = 2$ we can take any permutation in $\mathcal{S}_v$ and its reverse to form an $\mathsf{SCA}(2;2,v)$. Hence, $\mathsf{SCAN}(2,v) = 2$. For $3 \leqslant t < v$, we can then ask for which other values of t and v does $\mathsf{SCAN}(t,v) = t!$. The following construction of Levenshtein [37] shows that $\mathsf{SCAN}(t,t+1) = t!$ for all $t \geqslant 3$.

Theorem 3.17 ([37]). *For $t \geqslant 3$, $\mathcal{S}_{t+1}$ can be partitioned into $t+1$ $\mathsf{SCA}(t!;t,t+1)$. In particular, $\mathsf{SCAN}(t,t+1) = t!$.*

We describe Levenshtein's construction without proving why it partitions $\mathcal{S}_{t+1}$ into $\mathsf{SCA}(t!;t,t+1)$. For a permutation $\pi = \pi(0)\ldots\pi(v-1) \in \mathcal{S}_v$, let $R(\pi) = r_0 \ldots r_{v-2}$ be a binary string of length $v-1$ where

$$r_i = \begin{cases} 0, & \text{if } \pi(i) < \pi(i+1) \\ 1, & \text{if } \pi(i) > \pi(i+1). \end{cases}$$

For example, if $\pi = 420153$, then $R(\pi) = 11001$. For a binary string $B = b_0 \ldots b_{n-1} \in \{0,1\}^n$ of length n, let

$$W(B) = \sum_{i=0}^{n-1} (i+1) b_i.$$

Varšamov and Tenengolts [58] introduced the following code. Let

$$W^{n,a} = \{B \in \{0,1\}^n : W(B) \equiv a \pmod{n+1}\},$$

where $0 \leqslant a \leqslant n$. Now, for $t \geqslant 3$ and $0 \leqslant a \leqslant t$, let

$$Z_{t,a} = \{\pi \in \mathcal{S}_{t+1} : R(\pi) \in W^{t,a}\}.$$

Referring to our previous example, if $t = 5$ and $\pi = 420153$, then as stated before, $R(\pi) = 11001$. Hence, $W(R(\pi)) = 8 \equiv 2 \pmod 6$. Therefore, $\pi \in Z_{5,2}$. Levenshtein [37] proved that for $0 \leqslant a \leqslant t$, $Z_{t,a}$ is an $\mathsf{SCA}(t!;t,t+1)$ and moreover, $\{Z_{t,a} : 0 \leqslant a \leqslant t\}$

$Z_{3,0}$	$Z_{3,1}$	$Z_{3,2}$	$Z_{3,3}$
0123	0321	0213	0132
1032	1023	0231	0231
2031	1320	1203	1230
2130	2013	1302	2103
3021	2310	2301	3102
3120	3012	3210	3201

Table 3.1: Partition of $\mathcal{S}_4$ into $Z_{3,a}$ for $0 \leqslant a \leqslant 3$. Each of these parts forms an $\mathsf{SCA}(6; 3, 4)$.

π	$R(\pi)$	$W(R(\pi))$
0123	000	0
1032	101	4
2031	101	4
2130	101	4
3021	101	4
3120	101	4

Table 3.2: Properties of the elements of $Z_{3,0}$. Note that the values in the final column are all equivalent to 0 (mod 4), verifying that these permutations are indeed elements of $Z_{3,0}$.

is a partition of $\mathcal{S}_{t+1}$. We give this partition when $t = 3$ in Table 3.1. In Table 3.2, we take the part $Z_{3,0}$ and verify that for each $\pi \in S_{3,0}$, $R(\pi) \in W^{3,0}$. In addition to this construction, Levenshtein [36] made the following conjecture.

Conjecture 3.18 (Levenshtein's Conjecture [36])**.** *For $t \geqslant 3$, an $\mathsf{SCA}(t!; t, v)$ exists if and only if $v \in \{t, t+1\}$.*

Mathon and Tran van Trung [41] considered Conjecture 3.18 for small values of t and proved the following.

Theorem 3.19 ([41])**.** *Conjecture 3.18 is true for $t \in \{3, 5, 6\}$. An $\mathsf{SCA}(24; 4, v)$ exists if and only if $v \in \{4, 5, 6\}$.*

All non-existence proofs of Theorem 3.19 in [41] are computational, except for the proof of the non-existence of an $\mathsf{SCA}(6; 3, 5)$. The algorithm used to demonstrate the non-existence of an $\mathsf{SCA}(120; 5, 7)$ and an $\mathsf{SCA}(720; 6, 8)$ was described by Mathon in [40]. Klein [31] later gave a combinatorial proof of the non-existence of an $\mathsf{SCA}(24; 4, 7)$.

Theorem 3.20 ([31])**.** *An $\mathsf{SCA}(24; 4, 7)$ does not exist.*

Proof. Suppose for a contradiction that P is an $\mathsf{SCA}(24; 4, 7)$ and assume, without loss of generality, that $0123456 \in P$. Let

$$S = \{(0,1,3,2), (0,2,1,3), (0,2,3,1), (1,0,2,3), (1,2,0,3), (1,2,3,0), (2,0,1,3)\}.$$

012345	210435	403512	012345	032514	135204
032154	230514	413205	052431	042153	145023
042513	240153	423150	105342	304512	314205
052431	250341	453021	125430	324150	354021
105342	304152	501432	210435	230154	413025
125430	314025	521340	250341	240513	453201
135024	324510	531204	501432	403152	531024
145203	354201	541023	521340	423510	541203
	$\mathcal{S}_4$			D_8	

Figure 3.1: The two different $\mathsf{SCA}(24; 4, 6)$.

Each of the sequences in S contain the symbols 0, 1, 2, and 3 so they must be covered in seven distinct permutations in P. Furthermore, each sequence in S contains an ascending subsequence of length three that includes the symbol 3. Let π be a permutation covering a sequence in S. If any of the symbols 4, 5, or 6 appears after the symbol 3 in π, π would cover one of the sequences in $\mathcal{S}_{7,4}$ covered by 0123456. Hence, in the seven permutations covering one of the sequences in S, the symbols 4, 5, and 6 must appear before 3. However, there are only 6 ways of permuting $\{4, 5, 6\}$ and so in these seven permutations, there must be a sequence of $\{3, 4, 5, 6\}$ that is covered twice. Thus, P is not an $\mathsf{SCA}(24; 4, 7)$ and therefore, no such array exists. □

The existence of an $\mathsf{SCA}(24; 4, 6)$ is the only currently known counter-example to Conjecture 3.18. Mathon and Tran van Trung [41] found, up to isomorphism, exactly two $\mathsf{SCA}(24; 4, 6)$. These are presented in Figure 3.1. The first of these arrays, labelled $\mathcal{S}_4$, is a permutation subgroup of $\mathcal{S}_6$ isomorphic to $\mathcal{S}_4$. One possible generating set for this group is $\{042513, 135024\}$.

The second $\mathsf{SCA}(24; 4, 6)$ shown in Figure 3.1, labelled D_8, contains a permutation subgroup of $\mathcal{S}_6$ isomorphic to D_8, the dihedral group with 8 elements. One possible generating set for this group is $\{052431, 105342\}$. The other 16 permutations of this array are comprised of two left cosets of this permutation group. These cosets were mistakenly reported in [41] as having representatives 032514 and 304512 and the resulting array is not an $\mathsf{SCA}(24; 4, 6)$. We correct this construction by instead taking the left cosets with representatives 032514 and 135204.

We often consider elements of a coset of a permutation group in terms of a composition of two permutations. The following lemma establishes the sequences that a permutation composition covers.

Lemma 3.21. *For $\pi, \sigma \in \mathcal{S}_v$ and $s = (s_0, \ldots, s_{t-1}) \in \mathcal{S}_{v,t}$, π covers s if and only if $\sigma \circ \pi$ covers $\sigma(s) = (\sigma(s_0), \ldots, \sigma(s_{t-1}))$.*

Proof. By definition, $\sigma \circ \pi$ covers $\sigma(s)$ if and only if

$$(\pi^{-1} \circ \sigma^{-1})(\sigma(s_i)) < (\pi^{-1} \circ \sigma^{-1})(\sigma(s_{i+1})),$$

for $0 \leqslant i \leqslant t-2$. Therefore, $\sigma \circ \pi$ covers $\sigma(s)$ if and only if $\pi^{-1}(s_i) < \pi^{-1}(s_{i+1})$ for $0 \leqslant i \leqslant t-2$. □

In the case where an $\mathsf{SCA}(t!; t, v)$, P, is a permutation group, Mathon and Tran van Trung [41] proved that for $\sigma \in \mathcal{S}_v$, the left coset σP must also be an $\mathsf{SCA}(t!; t, v)$. We observe that the restriction that P must be a group is not necessary and thus extend this result to all $\mathsf{SCA}(t!; t, v)$.

Lemma 3.22. *Let P be an $\mathsf{SCA}(t!; t, v)$ and let $\sigma \in \mathcal{S}_v$. Then, $\sigma P = \{\sigma \circ \pi : \pi \in P\}$ is also an $\mathsf{SCA}(t!; t, v)$.*

Proof. Let $s = (s_0, \ldots, s_{t-1}) \in \mathcal{S}_{v,t}$ and let π be the unique permutation in P that covers $(\sigma^{-1}(s_0), \ldots, \sigma^{-1}(s_{t-1}))$. By Lemma 3.21, $\sigma \circ \pi$ is the only permutation in σP that covers s. Therefore, each sequence in $\mathcal{S}_{v,t}$ is covered by exactly one permutation in σP and hence, σP is an $\mathsf{SCA}(t!; t, v)$. □

We note that the permutations in σP are the permutations in P with the symbols permuted according to σ. Recall that the permutation group generated by $\{042513, 135024\}$ forms an $\mathsf{SCA}(24; 4, 6)$. This group, along with Lemma 3.22, gives us the following theorem.

Theorem 3.23 ([41])**.** *The symmetric group $\mathcal{S}_6$ can be partitioned into* 30 $\mathsf{SCA}(24; 4, 6)$ *by taking the left cosets of the group generated by* $\{042513, 135024\}$.

While Lemma 3.22 says that σP is an $\mathsf{SCA}(t!; t, v)$ when P is an $\mathsf{SCA}(t!; t, v)$, a partition of $\mathcal{S}_v$ into $\mathsf{SCA}(t!; t, v)$ in the vein of Theorem 3.23 is only guaranteed by Lemma 3.22 when P is a permutation group. The only use for Lemma 3.22 in [41] was to demonstrate Theorem 3.23 and so it was only necessary to prove Lemma 3.22 in the case where P is a permutation group. However, the more general form of Lemma 3.22 will be important in Chapter 7 when we discuss isomorphisms of perfect sequence covering arrays. We finish this section by stating the best known upper bound on v when $\mathsf{SCAN}(t, v) = t!$.

Theorem 3.24 ([9])**.** *For $t \geqslant 3$, if $\mathsf{SCAN}(t, v) = t!$, then $v \leqslant 2t - 1$.*

This result relies on a connection between sequence covering arrays and *covering arrays*. We will prove this theorem in Chapter 5.

3.3 Methods of constructing sequence covering arrays

In Chapter 3.1, we saw that for $t \geqslant 4$, the best known upper bounds for $\mathsf{SCAN}(t, v)$ came from probabilistic arguments with general constructions only available for the trivial case

$t = 2$ as well as for $t = 3$. These probabilistic results only guarantee the existence of an $\mathsf{SCA}(N; t, v)$ for some N by outlining a random process that has a non-zero chance of success. However, these processes are inadequate for practical applications which require the efficient generation of suitably small sequence covering arrays. In this section, we discuss some of the algorithms that have been developed to construct sequence covering arrays.

We begin by discussing greedy algorithms. Following the general method for constructing set covers outlined by Johnson [29], greedy algorithms build a sequence covering array one permutation at a time such that the permutation chosen at each step covers a large number of the sequences that have yet to be covered. Kuhn et al. [34] achieve this by randomly sampling a fixed number of permutations at each step and choosing the permutation which covers the most uncovered sequences. Then, after this permutation is added to the array, the reverse of that permutation is also added. For a permutation π, the set of sequences covered by π is disjoint to the set of sequences covered by its reverse. If permutations are added to the array with their reverses, it ensures that at each step, the reverse permutation is just as good a choice as the original permutation.

A derandomised greedy algorithm was presented by Chee et al. [9]. In their algorithm, permutations are built one symbol at a time. After the ith step, a permutation of $[i]$ is built. In the $(i + 1)$th step, the symbol i is inserted into this permutation of $[i]$ such that a random permutation in $\mathcal{S}_v$ that covers this permutation of $[i + 1]$ is more likely to cover a large number of uncovered t-sequences. Then, once a permutation in $\mathcal{S}_v$ is built, it is added to the array. A variant of this algorithm is also proposed in which reverse permutations are added at each step, similar to the algorithm of Kuhn et al. [34].

By encoding the definition of sequence covering arrays as a set of logical constraints, answer set programming can be used to generate sequence covering arrays. Brain et al. [6] develop a greedy algorithm where each permutation chosen to be in the array is selected using answer set programming. Additionally, by encoding logical variables to check the order of two symbols in a given permutation, Brain et al. [6] use answer set programming to construct sequence covering arrays with strengths 3 and 4. A similar answer set programming method is also given by Banbara, Tamura and Inoue [3].

Murray and Colbourn [42] and Torres-Jimenez et al. [57] employ post-optimisation methods to improve upon sequence covering arrays generated using greedy methods. With these methods, changes are made to arrays in order to make certain permutations redundant. These redundant permutations can then be removed, thus reducing the size of the array. Murray and Colbourn [42] take a sequence covering array generated by the greedy algorithm of Chee et al. [9] and, for some ordering of the permutations in the array, compute the sequences each permutation covers that are not covered by any of the previous permutations. Each permutation is then replaced by a random permutation that covers at least those same sequences for the first time. By reordering the permutations and recalculating which sequences are covered for the first time, it is possible that some permutations become redundant and can be removed from the array.

Torres-Jimenez et al. [57] develop a greedy algorithm for generating sequence covering arrays by reinterpreting sequences as edges of a directed graph. Consider a graph G with vertex set $[v]$ and let $s = (x, y, z)$ be a 3-sequence. We can encode s into the graph G by adding directed edges from x to y and from y to z. If G contains no directed cycles, then G can be converted into a permutation π such that a directed edge from x to y in G implies $\pi^{-1}(x) < \pi^{-1}(y)$. The greedy algorithm of Torres-Jimenez et al. involves encoding as many uncovered sequences into a directed acyclic graph as possible and then converting that graph into a permutation. In each permutation of this resulting sequence covering array, each symbol is tagged as either useful or non-useful depending on whether it is part of a sequence that is covered for the first time by that permutation. A reduction step exploits non-useful symbols to remove permutations from the array. A simulated annealing algorithm is then used to efficiently replace permutations that cover sequences that may have become uncovered during the reduction stage.

In addition to the arrays generated algorithmically using the above methods, there are also other constructive methods that can be used to build sequence covering arrays with strength 3. We have already discussed Tarui's construction [53] that establishes the best known upper bound for $\mathsf{SCAN}(3, v)$ in Theorem 3.13. A recursive construction for strength 3 CSSPs (which can then be converted to SCAs by Lemma 3.1) was given by Chee et al. [9]. A *signing* of a $\mathsf{CSSP}(N; 3, v)$ A is an $N \times v$ matrix S with entries from $\{\uparrow, \downarrow\}$. The matrix S is a *proper signing of* A if for each pair of distinct columns $x, y \in [v]$ and for each $s \in \{\uparrow, \downarrow\}$, there is a $\pi \in \mathsf{A}$ such that $\pi(x) < \pi(y)$ and $S[\pi, x] = s$. Chee et al. show that a properly signed $\mathsf{CSSP}(N; 3, v)$ and a properly signed $\mathsf{CSSP}(M; 3, w)$ can be used to construct a properly signed $\mathsf{CSSP}(N + M; 3, vw)$, thus giving a recursive construction for strength 3 sequence covering arrays. Chee et al. provide proper signings for CSSPs generated by Banbara, Tamura and Inoue [3] and Margolit [39] and use these properly signed CSSPs to build $\mathsf{SCA}(N; 3, v)$ for specific values of v. While asymptotically this construction does not outperform Tarui's, for certain values of v, the $\mathsf{SCA}(N; 3, v)$ built using this recursive construction is smaller than the one built using Tarui's method.

Tables 3.3, 3.4 and 3.5 show upper bounds for $\mathsf{SCAN}(3, v)$, $\mathsf{SCAN}(4, v)$ and $\mathsf{SCAN}(5, v)$ respectively using the methods described above. The column labelled TA gives bounds derived from Tarui's construction [53]. The column labelled SC gives bounds derived from the signed CSSP construction of Chee et al. [9]. Bounds derived from implementations of the greedy algorithms of Kuhn et al. [34], Brain et al. [6], Chee et al. [9] and Torres-Jimenez et al. [57] are given in the columns labelled K, BG, D and TJG respectively. The column labelled $\mathrm{D_R}$ gives bounds from the variant of the Chee et al. [9] greedy algorithm that uses reverse permutations. Bounds from the answer set programming methods of Banbara, Tamura and Inoue [3] and Brain et al. [6] are given in the columns labelled BTI and B respectively. Bounds from the post-optimisation algorithms of Murray and Colbourn [42] and Torres-Jimenez et al. [57] are given in the columns labelled MC and

v	TA	SC	K	BG	D	D_R	TJG	BTI	B	MC	TJP
4	8				6	8	6			6	6
5	8		8		8	10	8	7	7	7	7
6	8		10		8	10		8	8	8	
7	10		12		9	12		8	8	8	
8	10		12		10	12	10	8	8	9	8
9	10		14		11	12		9	9	9	
10	10		14	11	12	14	11	9	9	10	9
11	12		14		12	14		10	10		
12	12		16		13	14		10	10		
13	12		16		13	16		10	10		
14	12		16		14	16		10	10		
15	12		18		14	16		10	10	12	
16	12		18		15	16		11	10		
17	12		20		15	18		11	11		10
18	12		20		16	18		12	12		10
19	12		22		16	18	16	12	12		10
20	12	13	22	17	16	18		12	12	13	11
21	14		22		17	18		12	12		11
22	14		22		17	20	17	13	12		11
23	14		24		17	20		13	13		12
24	14		24		17	20		14	13		12
25	14		24		18	20		14	14	14	12
26	14		24		18	20		14	14		12
27	14		26		18	20		14	14		12
28	14		26		18	20		14	14		12
29	14		26		19	22		15	14		12
30	14	14	26	22	19	22	20	15	15	15	12
40	16	15	32	26	21	24	22	17	17	16	13
50	16	16	34	29	23	26	23	19	18	17	14
60	16	16	38	32	24	26	25	21	20	18	15
70	16	17	40	35	25	28		22	22	19	16
80	18	17	42	36	26	30	28	24	23	20	16
90	18	18			27	30	30			21	18

Table 3.3: Comparison of upper bounds for $\mathsf{SCAN}(3, v)$ derived through various construction methods.

TJP respectively.

Note that since $\mathsf{SCAN}(t, v) \geqslant \mathsf{SCAN}(t, v-1)$, it is possible to fill in some of the gaps in these tables. This is done in the TJP column of Table 3.3 for values of v in which the Torres-Jimenez et al. post-optimisation bound is the best known bound. It may also be the case that for certain methods, bounds for $\mathsf{SCAN}(3, v)$ can only currently be derived by finding a bound for $\mathsf{SCAN}(3, w)$ for some $w > v$, even if that value of w does not appear in any of our tables. In particular, the signed CSSP product construction of Chee et al. only gives bounds for specific values of v. The bound $\mathsf{SCAN}(3, 60) \leqslant 16$ is found by using the product construction on a signed $\mathsf{CSSP}(6; 3, 4)$ and a signed $\mathsf{CSSP}(10; 3, 16)$ to build a $\mathsf{CSSP}(16; 3, 64)$. Although not all the available data matches neatly with the values of v in these tables, these were the values first tested by Kuhn et al. [34] and then by several subsequent authors [9, 3, 6, 42]. We also choose these values for v for the sake of easy comparison.

In Table 3.3, we see upper bounds for $\mathsf{SCAN}(3, v)$. Of the greedy algorithms, the smallest arrays are generally built from the Chee et al. algorithm which does not use reversals. However, all of the greedy methods are outperformed by the two different answer set programming methods which give near identical bounds. For $v \geqslant 30$, the constructions of Tarui and Chee et al. are able to build smaller arrays than the ones built by answer set programming. However, for $v \geqslant 17$, it is the post-optimisation method of Torres-Jimenez et al. which is able to give the best upper bounds for $\mathsf{SCAN}(3, v)$. The bounds given by Banbara, Tamura and Inoue and Brain et al. are best possible for $4 \leqslant v \leqslant 10$ [6].

Table 3.4 shows upper bounds for $\mathsf{SCAN}(4, v)$. Now we see that for $9 \leqslant v \leqslant 23$, the answer set programming bounds are no longer smaller than the best greedy bounds. This and the fact that there are no available bounds for $v > 23$ indicate scalability issues for answer set programming methods. Again, the algorithms of Chee et al. give the best greedy bounds although there is a much less clear pattern as to whether including reversals gives smaller or larger arrays. Another interesting difference is the extent to which the Torres-Jimenez et al. greedy algorithm is outperformed by the Chee et al. greedy algorithm, given these two algorithms performed quite similarly when $t = 3$. Again, the post-optimisation methods provide sizeable improvements to the greedy methods. Much less is known about exact values of $\mathsf{SCAN}(4, v)$ than in the $t = 3$ case which makes it difficult to judge how close to optimal each of these bounds are.

Table 3.5 shows upper bounds for $\mathsf{SCAN}(5, v)$. Here we no longer have data from the answer set programming methods or from the Kuhn et al. greedy algorithm. Again, the Chee et al. greedy algorithm regularly produces significantly smaller arrays than the Torres-Jimenez et al. greedy algorithm. Interestingly, in almost all cases, including reversals provides the best greedy bounds. Unsurprisingly, the post-optimisation algorithms make sizeable improvements to the greedy arrays. For $v \geqslant 27$, the simulated annealing step of the Torres-Jimenez et al. algorithm became computationally infeasible.

v	K	BG	D	D_R	TJG	BTI	MC	TJP
5	26		26	24		24	24	
6	36		34	32		24	24	
7	46		41	40	43	38	37	32
8	50		47	44	49	44	42	39
9	58		52	50	54	52	46	43
10	66	55	57	56	60	58	53	47
11	70		61	60	65	65		51
12	78		66	64	70	69		54
13	86		71	70	75	77		58
14	90		73	74	79	81		61
15	96		78	78	83	84	67	63
16	100		81	80	87	89		66
17	108		84	84	91	91		69
18	112		86	86	94	97		71
19	114		91	90	97	100		73
20	120	104	92	92	100	104	80	75
21	126		95	96	104	104		77
22	128		97	98	106	111		80
23	134		99	98	109	112		82
24	136		101	102	111			84
25	140		104	104	114		90	87
26	142		105	106	117			88
27	148		107	108	119			90
28	150		110	110	122			92
29	154		111	112	124			94
30	156	149	113	114	125		98	95
40	182	181	128	132	145		112	112
50	204		141	146	162		123	125
60	222		151	154	174		133	135
70	238		160	166	187		142	145
80	250		168	174	202		150	154
90			176	180	211		162	162

Table 3.4: Comparison of upper bounds for $\mathsf{SCAN}(4, v)$ derived through various construction methods.

v	D	D_R	TJG	MC	TJP
6	149	148	152	122	120
7	198	200	201	175	160
8	243	242	238	218	208
9	284	282	299	261	249
10	322	318	344	294	288
11	356	354	382	330	323
12	386	384	420	360	354
13	419	416	461	390	383
14	448	446	496	418	412
15	475	470	530	448	440
16	501	496	566	474	464
17	521	518	597	496	489
18	547	540	627	520	511
19	570	560	655	541	531
20	590	582	681	568	551
21	610	600	714		595
22	629	622	739		617
23	646	636	762		632
24	665	654	784		650
25	682	674	808	656	687
26	698	688	830		706
27	715	706	850		765
28	732	718	869		782
29	746	734	887		796
30	760	748	908	725	816

Table 3.5: Comparison of upper bounds for $\mathsf{SCAN}(5, v)$ derived through various construction methods.

This means that for these values of v, the only post-optimisation step involves exploiting redundancies in the initial greedy array. This explains the relatively large increase in bounds from $v = 26$ to $v = 27$. Despite this limitation, this less effective post-optimisation algorithm is still able to make significant reductions to the greedy arrays.

Torres-Jimenez et al. [57] also report some upper bounds for $\mathsf{SCAN}(t, v)$ for $t \in \{6, 7, 8\}$ however these bounds are derived purely from greedy methods. Given the extent to which these bounds were improved by post-optimisation methods and that the Torres-Jimenez et al. greedy algorithm regularly produced larger arrays than other greedy methods for smaller values of t, it seems reasonable to suggest that these bounds are far from optimal.

Chapter 4

Perfect Sequence Covering Arrays

In this chapter we consider a stronger form of sequence covering array in which every sequence must be covered the same number of times. A *perfect sequence covering array* with *order* v, *strength* t and *multiplicity* λ, denoted $\mathsf{PSCA}(v,t,\lambda)$, is a multiset X of permutations in $\mathcal{S}_v$ such that each sequence in $\mathcal{S}_{v,t}$ is covered by exactly λ permutations in X. Let $N = |X|$. If T is a subset of $[v]$ containing t symbols, then there are $t!$ ways of arranging the elements of T. Each of these arrangements forms a sequence in $\mathcal{S}_{v,t}$ that must be covered by λ permutations in X. Furthermore, every permutation in X covers exactly one of these sequences. Therefore, we must have $N = t!\lambda$.

An $\mathsf{SCA}(t!;t,v)$ can also be interpreted as a $\mathsf{PSCA}(v,t,1)$. Indeed, we can view the study of perfect sequence covering arrays as a way of generalising Conjecture 3.18. In our discussion of Conjecture 3.18 in Chapter 3.2, we saw that for fixed t there are seemingly very few choices for v for which a $\mathsf{PSCA}(v,t,1)$ exists. For fixed v and t, we can instead consider the smallest possible value for λ for which a $\mathsf{PSCA}(v,t,\lambda)$ exists. We call this value the *perfect sequence covering array number* and denote it by $\mathsf{PSCAN}(v,t)$. We can reinterpret Conjecture 3.18 as a statement on when $\mathsf{PSCAN}(v,t) = 1$. As we will discuss, much of the research into perfect sequence covering arrays focuses on bounding or determining $\mathsf{PSCAN}(v,t)$ for different values of v and t. The following lemma makes two basic observations about $\mathsf{PSCAN}(v,t)$.

Lemma 4.1 ([64])**.** *For $3 \leqslant t < v$, the existence of a $\mathsf{PSCA}(v,t,\lambda)$ implies the existence of a $\mathsf{PSCA}(v-1,t,\lambda)$. For $3 < t \leqslant v$, the existence of a $\mathsf{PSCA}(v,t,\lambda)$ implies the existence of a $\mathsf{PSCA}(v,t-1,t\lambda)$. In particular,*

- *For $3 \leqslant t < v$, $\mathsf{PSCAN}(v-1,t) \leqslant \mathsf{PSCAN}(v,t)$, and*
- *For $3 < t \leqslant v$, $\mathsf{PSCAN}(v,t-1)/t \leqslant \mathsf{PSCAN}(v,t)$.*

Proof. To prove the first statement, take a $\mathsf{PSCA}(v,t,\lambda)$ and remove the symbol $v-1$ from each permutation. We are then left with a $\mathsf{PSCA}(v-1,t,\lambda)$. To prove the second statement, let X be a $\mathsf{PSCA}(v,t,\lambda)$, let $s = (s_0,\ldots,s_{t-2}) \in \mathcal{S}_{v,t-1}$, and let $u \in [v]$ be such that u does not appear in s. Then, there are t sequences of the elements $\{s_0, \ldots, s_{t-2}, u\}$ which cover s. Each of these sequences are covered by λ permutations in X and no permutation can cover two of these sequences. Hence, there are $t\lambda$ permutations in X that cover s. Therefore, X is also a $\mathsf{PSCA}(v, t-1, t\lambda)$. □

Perfect sequence covering arrays were first introduced by Yuster [64] in 2020 as an extension of sequence covering arrays. Yuster saw perfectness as an important design-theoretic property, drawing an equivalence between perfect sequence covering arrays and directed designs. Moreover, imposing a perfectness condition onto sequence covering arrays mirrors the relationship between other combinatorial covering objects. For example, an orthogonal array can be seen as a covering array with an additional perfectness condition.

Other related objects have also previously been studied. First, we introduce the aforementioned directed design. A *directed design* of type t-(v, k, λ) is a family $\mathcal{B}$ of sequences in $\mathcal{S}_{v,k}$ such that each sequence in $\mathcal{S}_{v,t}$ is covered by exactly λ sequences in $\mathcal{B}$. A $\mathsf{PSCA}(v, t, \lambda)$ is then equivalent to a directed design of type t-(v, v, λ). Directed designs have been extensively studied although typically in the cases where $k < v$ [5].

A family of permutations $X \subseteq \mathcal{S}_v$ is *t-wise independent* if for any $a, b \in \mathcal{S}_{v,t}$,

$$\frac{1}{|X|}|\{\pi \in X : \pi(a_i) = b_i, 0 \leqslant i \leqslant t-1\}| = \frac{(v-t)!}{v!}.$$

In other words, in every ordered sequence of t columns of a t-wise independent family of permutations, we see each sequence in $\mathcal{S}_{v,t}$ the same number of times. Let X be a family of t-wise independent permutations. Then the multiset of sequences in $\mathcal{S}_{v,t}$ covered by the permutations in X is

$$\{(\pi(a_0), \pi(a_1), \ldots, \pi(a_{t-1})) : \pi \in X, a_0 < \cdots < a_{t-1}\}.$$

Given that X is t-wise independent and that there are $\binom{v}{t}$ sequences in $\mathcal{S}_{v,t}$ that are in ascending order, the number of times each sequence in $\mathcal{S}_{v,t}$ appears in the above multiset is $\binom{v}{t}|X|(v-t)!/v! = |X|/t!$. Thus, a t-wise independent family of permutations in $\mathcal{S}_v$ is also a $\mathsf{PSCA}(v, t)$. Note that a perfect sequence covering array need not be a t-wise independent family of permutations.

The definition of t-wise independent families of permutations can be seen as a generalisation of the concept of t-transitive permutation groups. While there are infinitely many t-transitive permutation groups when $t \leqslant 3$, it is known [8] that the only t-transitive permutation groups for $t \geqslant 4$ are the symmetric groups $\mathcal{S}_k$ for $k \geqslant t$, the alternating groups A_k for $k \geqslant t+2$, and certain sporadic Mathieu groups. The groups $M_{11} \leqslant \mathcal{S}_{11}$ and $M_{23} \leqslant \mathcal{S}_{23}$ are 4-transitive but not 5-transitive while the groups $M_{12} \leqslant \mathcal{S}_{12}$ and $M_{24} \leqslant \mathcal{S}_{24}$ are 5-transitive but not 6-transitive. By relaxing the definition of t-transitive groups to t-wise independent families of permutations, we can still study sets of permutations with these interesting sequence coverage properties analogous to t-transitive groups for larger values of t.

A family of permutations $X \subseteq \mathcal{S}_v$ is *t-rankwise independent* if for all $s \in \mathcal{S}_{v,t}$,

$$\frac{1}{|X|}|\{\pi \in X : \pi(s_i) < \pi(s_{i+1}), 0 \leqslant i \leqslant t-2\}| = \frac{1}{t!}.$$

Recall the definition of a completely scrambling set of permutations. It can be seen that a t-rankwise independent family of permutations is a kind of *perfect* CSSP. By extending Lemma 3.1, we can see that a t-rankwise independent family of permutations is equivalent to a PSCA of order v and strength t. Itoh, Takei and Tarui [26] first introduced families of t-rankwise independent permutations as a way of relaxing the definition of t-wise independent families of permutations. The connection between t-rankwise independent families of permutations and perfect sequence covering arrays was recently highlighted by Iurlano [28] in 2023.

A family of permutations $X \subseteq \mathcal{S}_v$ is *t-restricted min-wise independent* if for each $T \subseteq [v]$ with $|T| \leqslant t$, and for each $i \in T$,

$$\frac{1}{|X|}\left|\left\{\pi \in X : \min_{j\in T}(\pi(j)) = \pi(i)\right\}\right| = \frac{1}{|T|}.$$

Broder et al. [7] first introduced t-restricted min-wise independence in order to detect duplicate documents on the World Wide Web. The following lemma shows that t-rankwise independence implies t-restricted min-wise independence.

Lemma 4.2. *If a family of permutations is t-rankwise independent, then it is also t-restricted min-wise independent.*

Proof. Suppose $X \subseteq \mathcal{S}_v$ is a t-rankwise independent family of permutations. Then, by Lemma 4.1, it is also t'-rankwise independent for any $t' < t$. Let $Y \subseteq [v]$ with $|Y| = y \leqslant t$ and let $i \in Y$. There are $(y-1)!$ permutations π of Y in which $\pi(i)$ is the smallest element of Y. Therefore, as X is y-rankwise independent,

$$\frac{1}{|X|}\left|\left\{\pi \in X : \min_{j\in Y}(\pi(j)) = \pi(i)\right\}\right| = (y-1)!\frac{1}{y!} = \frac{1}{y}$$

Therefore, X is also t-restricted min-wise independent. ∎

Moreover, as the following lemma shows, 3-rankwise independence is equivalent to 3-restricted min-wise independence.

Lemma 4.3. *If a family of permutations is 3-restricted min-wise independent, then it is also 3-rankwise independent.*

Proof. Let $X \subseteq \mathcal{S}_v$ be 3-restricted min-wise independent with $|X| = N$. Let $i, j, k \in [v]$. As X is 3-restricted min-wise independent, the number of permutations $\pi \in X$ for which

$\pi(j) < \pi(k)$ is exactly $N/2$. As X is 3-restricted min-wise independent, of these $N/2$ permutations, exactly $N/3$ also satisfy $\pi(j) < \pi(i)$. This leaves $N/6$ permutations in which $\pi(i) < \pi(j) < \pi(k)$. Thus, X is 3-rankwise independent. □

So far in this section, we have introduced three different notions of independence relating to sequences in permutations. These notions have been progressively weaker in the sense that t-wise independence implies t-rankwise independence which in turn implies t-restricted min-wise independence. Moreover, as perfect sequence covering arrays are equivalent to t-rankwise independent families of permutations, upper bounds on the size of t-wise independent families of permutations can be re-evaluated as upper bounds on $\mathsf{PSCAN}(v,t)$ and lower bounds on the size of t-restricted min-wise independence can similarly be re-evaluated as lower bounds on $\mathsf{PSCAN}(v,t)$.

We conclude this section by considering the "perfectness" of t-restricted min-wise independent families of permutations. A set $X \subseteq \mathcal{S}_v$ of permutations is *t-suitable* if for every $T \subseteq [v]$ with $|T| = t$ and for every $i \in T$, there exists a permutation $\pi \in X$ such that $\min_{j\in T}(\pi(j)) = \pi(i)$. Then the relationship between t-suitable sets of permutations and t-restricted min-wise independent families of permutations is analogous to the relationship between sequence covering arrays and perfect sequence covering arrays.

The definition of t-suitable sets of permutations was introduced by Dushnik [14] in 1950. Dushnik's main focus was to find the smallest set of t-suitable permutations in $\mathcal{S}_v$. This problem has applications in the theory of partial orders. For integers $1 \leqslant k < t < v$, let $P(v,k,t)$ denote the partial order defined on the set of all t-subsets and k-subsets of $[v]$ and whose ordering relation is set inclusion. It was shown by Dushnik that there exists a t-suitable set of N permutations in $\mathcal{S}_v$ if and only if there exists a set of N linear extensions of $P(v,1,t-1)$ that realises $P(v,1,t-1)$. Therefore, the size of the smallest set of t-suitable permutations in $\mathcal{S}_v$ is equal to the Dushnik-Miller dimension of $P(v,1,t-1)$. Dushnik's study of t-suitable sets of permutations also directly motivated Spencer's [51] introduction of completely scrambling sets of permutations. Although each of these concepts were introduced under different contexts and to answer different mathematical questions, we can see that Itoh, Takei and Tarui's [26] introduction of t-rankwise independence as a stronger version of t-restricted min-wise independence has parallels to Spencer's [51] introduction of completely t-scrambling sets of permutations as a stronger form of t-suitable sets of permutations.

In this section we have considered different notions of perfectness as they relate to sequences in permutations. We have seen that whether intentionally or not, defining perfect combinatorial structures often extends existing problems. Moreover, perfect combinatorial objects often have different applications and rely on different mathematics as their more general forms. We return now to perfect sequence covering arrays. Yuster [64] remarked that perfectness was an important design-theoretic property. Regardless of the fact that equivalent structures had been researched previously, or indeed because such research had taken place, we can see that Yuster's assertion is correct and that perfect sequence covering arrays are an important and interesting extension of sequence covering arrays. In Chapter 4.1, we discuss general bounds for $\mathsf{PSCAN}(v,t)$. In Chapter 4.2, we discuss computational methods that give exact values of $\mathsf{PSCAN}(v,t)$.

4.1 Bounds on PSCAN

In this section, we discuss general bounds on $\mathsf{PSCAN}(v,t)$. Recall from Chapter 3.2 that we can build an $\mathsf{SCA}(2;2,v)$ for any $v \geqslant 2$. Therefore, $\mathsf{PSCAN}(2,v) = 1$ for $v \geqslant 2$. Hence we focus on bounds for $\mathsf{PSCAN}(v,t)$ for $t \geqslant 3$. We first discuss lower bounds for $\mathsf{PSCAN}(v,3)$ and $\mathsf{PSCAN}(v,t)$ for $t \geqslant 4$. These lower bounds indicate that the size of the smallest $\mathsf{PSCA}(v,t,\lambda)$ is considerably larger than the size of the smallest $\mathsf{SCA}(N;t,v)$. We also discuss different probabilistic and constructive upper bounds for $\mathsf{PSCAN}(v,t)$. We begin with a lower bound for $\mathsf{PSCAN}(v,3)$.

Theorem 4.4 ([64])**.** *For $v \geqslant 3$, $\mathsf{PSCAN}(v,3) \geqslant v/6$.*

Proof. Let X be a $\mathsf{PSCA}(v,3,\lambda)$ and let A be the matrix whose columns are indexed by the permutations in X and whose rows are indexed by the sequences in $\mathcal{S}_{v,2}$ with $A[s,\pi] = 1$ if π covers s and $A[s,\pi] = 0$ otherwise. Then, A has 6λ columns and $v(v-1)$ rows. Let $B = AA^T$. Then, the rows and columns of B are indexed by the sequences in $\mathcal{S}_{v,2}$ with $B[s,u]$ being equal to the number of permutations in X that cover both s and u. Moreover, note that since A has 6λ columns, $\text{rank}(B) \leqslant \text{rank}(A) \leqslant 6\lambda$.

Let C be the submatrix of B whose rows and columns are indexed by the sequences $(i, v-1)$ for $0 \leqslant i \leqslant v-2$ and $(v-1,0)$. Then, C is a $v \times v$ matrix. By Lemma 4.1, every sequence in $\mathcal{S}_{v,2}$ is covered by 3λ permutations in X. Therefore, the diagonal entries in C are all equal to 3λ. Consider the entry in row $(i,v-1)$ and column $(j,v-1)$ of C with $i \neq j$. A permutation in X covers both of these sequences if and only if it covers either $(i,j,v-1)$ or $(j,i,v-1)$. Hence, $C[(i,v-1),(j,v-1)] = 2\lambda$. Now consider the entry in row $(i,v-1)$ and column $(v-1,0)$ of C. If $i = 0$, then it is impossible for any permutation to cover both sequences while for $1 \leqslant i \leqslant v-2$, a permutation in X covers both sequences if and only if it covers $(i,v-1,0)$. Therefore, $C[(0,v-1),(v-1,0)] = 0 = C[(v-1,0),(0,v-1)]$ and $C[(i,v-1),(v-1,0)] = \lambda = C[(v-1,0),(i,v-1)]$ for $1 \leqslant i \leqslant v-2$. Let $C^* = C/\lambda$.

Let x be a length v column vector with entries indexed by the sequences $(i,v-1)$ for $0 \leqslant i \leqslant v-2$ and $(v-1,0)$. Let $x[(i,v-1)] = 1$ for $0 \leqslant i \leqslant v-2$ and $x[(v-1,0)] = 1/2$. Let $y = 2x$ and let $D = C^* - xy^T$. Then,

$$D[i,j] = \begin{cases} 1 & \text{if } i = j \text{ and } i \in \{(a,v-1) : 0 \leqslant a \leqslant v-2\} \\ \frac{5}{2} & \text{if } i = j = (v-1,0) \\ -1 & \text{if } \{i,j\} = \{(0,v-1),(v-1,0)\} \\ 0 & \text{otherwise.} \end{cases}$$

Consider the determinant of D. As an example, when $v = 5$, we can see that,

$$\begin{aligned}
\begin{vmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & \frac{5}{2} \end{vmatrix} &= \frac{5}{2}\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} + (-1)^4(-1)\begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{vmatrix} \\
&= \frac{5}{2} - (-1)^3(-1)\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \\
&= \frac{3}{2}.
\end{aligned}$$

In general, $\det(D) = 5/2 + (-1)^{v-1}(-1) \cdot (-1)^{v-2}(-1) = 3/2$. Furthermore,

$$D^{-1}[i,j] = \begin{cases} \frac{5}{3} & \text{if } i = j = (0, v-1) \\ 1 & \text{if } i = j \text{ and } i \in \{(a, v-1) : 1 \leqslant a \leqslant v-2\} \\ \frac{2}{3} & \text{if } i = j = (v-1, 0) \text{ or } \{i,j\} = \{(0, v-1), (v-1, 0)\} \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$x^T D^{-1}[i] = \begin{cases} 2 & \text{if } i = (0, v-1) \\ 1 & \text{otherwise.} \end{cases}$$

Hence, $x^T D^{-1} y = 4 + 2(v-2) + 1 = 2v + 1$. Therefore, by the matrix determinant lemma, $\det(C^*) = 3(1 + 2v + 1)/2 = 3(v+1)$. Therefore, C^* has full rank and so, $\operatorname{rank}(C) = v$. Therefore, $v \leqslant \operatorname{rank}(B) \leqslant 6\lambda$ and thus, $\lambda \geqslant v/6$. □

For higher values of t, we will discuss three different lower bounds for PSCAN(v, t). The first two were proved by Itoh, Takei and Tarui [27] and third was proved by Yuster [64]. Each of these bounds gives the best known lower bound for PSCAN(v, t) for different values of t. We first introduce and derive these bounds and then we discuss the conditions under which each bound performs better than the other two. First we introduce the bounds derived by Itoh, Takei and Tarui [27]. For positive integers n and k, let

$$m(n,k) = \begin{cases} \displaystyle\sum_{j=0}^{k/2} \binom{n}{j}, & \text{if } k \text{ is even, and} \\ \displaystyle\binom{n-1}{\frac{k-1}{2}} + \sum_{j=0}^{(k-1)/2} \binom{n}{j}, & \text{if } k \text{ is odd.} \end{cases}$$

Furthermore, let

$$B(n,k) = \left\lfloor \frac{n}{\ell} \right\rfloor \binom{n - \lfloor \frac{n}{\ell} \rfloor}{\ell - 1},$$

where $\ell = \lfloor k/2 \rfloor$.

Theorem 4.5 ([27]). *For all $v \geqslant t \geqslant 4$,*

$$\mathsf{PSCAN}(v,t) \geqslant \frac{1}{t!} \max\left\{m(v-1,t-1), B(v,t)\right\}.$$

Proof. First we consider $(1/t!)m(v-1,t-1)$. As we will see, $m(v-1,t-1)$ is a lower bound for the size of a t-restricted min-wise independent family of permutations. Since a t-rankwise independent family of permutations is also t-restricted min-wise independent, and since PSCAs are equivalent to t-rankwise independent families of permutations, lower bounds on the size of t-restricted min-wise independent families of permutations are also lower bounds on the size of perfect sequence covering arrays. As $m(v-1,t-1)$ bounds the number of permutations in a $\mathsf{PSCA}(v,t,\lambda)$, to obtain a corresponding lower bound for $\mathsf{PSCAN}(v,t)$, we divide $m(v-1,t-1)$ by $t!$.

Let $\mathcal{F} \subseteq \mathcal{S}_v$ be a family of t-restricted min-wise independent permutations. Define the random variables $Z_1, \ldots, Z_{v-1} : \mathcal{F} \to \{0,1\}$ by $Z_i = 1$ if $\pi(0) < \pi(i)$ and $Z_i = 0$ if $\pi(0) > \pi(i)$ where π is sampled uniformly at random from $\mathcal{F}$. For $J \subset \{1, \ldots, v-1\}$, let $\tilde{Z}_J = 1$ if $Z_j = 1$ for all $j \in J$ and let $\tilde{Z}_J = 0$ otherwise. As $\mathcal{F}$ is t-restricted min-wise independent, for any $J \subset \{1, \ldots, v-1\}$ with $|J| \leqslant t-1$,

$$\Pr[\tilde{Z}_J = 1] = \frac{1}{1+|J|}.$$

Now define the random variables $Y_1, \ldots, Y_{v-1} : \mathcal{S}_v \to \{0,1\}$ by $Y_i = 1$ if $\pi(0) < \pi(i)$ and $Y_i = 0$ if $\pi(0) > \pi(i)$ where π is sampled uniformly at random from $\mathcal{S}_v$. For $J \subset [1, v-1]$, let $\tilde{Y}_J = 1$ if $Y_j = 1$ for all $j \in J$ and let $\tilde{Y}_j = 0$ otherwise. Furthermore, let $Y_J = 1$ if $Y_j = 1$ for all $j \in J$ and $Y_j = 0$ for all $j \notin J$. Otherwise, let $Y_J = 0$. For any subset $J \subset [1, v-1]$, for $Y_J = 1$ we require $\pi(0) = v-1-|J|$ and $\{\pi(j) : j \in J\} = \{v-|J|, \ldots, v-1\}$. Since π is chosen uniformly at random from $\mathcal{S}_v$, $\Pr[Y_J = 1] > 0$. Furthermore, $\Pr[\tilde{Y}_J = 1] = 1/(1+|J|)$. By constructing a suitable Gram matrix and showing that it is non-singular, Itoh, Takei and Tarui show that the existence of the random variables $Z_1, \ldots, Z_{v-1}, Y_1, \ldots, Y_{v-1}$ with these properties implies that $\mathcal{F} \geqslant m(v-1, t-1)$. Therefore, the number of permutations in a $\mathsf{PSCA}(v,t,\lambda)$ must be at least $m(v-1,t-1)$ and so, $\mathsf{PSCAN}(v,t) \geqslant m(v-1,t-1)/t!$.

We complete the proof by showing that $\mathsf{PSCAN}(v,t) \geqslant B(v,t)/t!$. Let $\mathcal{F} \subseteq \mathcal{S}_v$ be a t-rankwise independent family of permutations. Let $k = \lfloor t/2 \rfloor$ and let $\mathcal{I} = \{(i,J) : i \in [\lfloor v/k \rfloor - 1]$ and $J \subset \{\lfloor v/k \rfloor, \ldots v-1\}$ with $|J| = k-1\}$. Observe that $|\mathcal{I}| = B(v,t)$. Let U be a $|\mathcal{I}| \times |\mathcal{F}|$ matrix with rows indexed by the elements of $\mathcal{I}$ and columns indexed by the permutations in $\mathcal{F}$. For $(i, J) \in \mathcal{I}$ and $\pi \in \mathcal{F}$, let

$$U[(i,J),\pi] = \begin{cases} 1/\sqrt{|\mathcal{F}|}, & \text{if } \pi(i) < \pi(j) \text{ for all } j \in J \\ 0, & \text{otherwise.} \end{cases}$$

Now let $W = UU^T$. Then, $|\mathcal{F}| \geqslant \operatorname{rank}(U) \geqslant \operatorname{rank}(W)$. We have that W has $B(v,t)$ rows and columns with these rows and columns being indexed by the elements of $\mathcal{I}$. For $(i,J),(i',J') \in \mathcal{I}$, $W[(i,J),(i',J')]$ is equal to $1/|\mathcal{F}|$ times the number of permutations $\pi \in \mathcal{F}$ satisfying $\pi(i) < \pi(j)$ for all $j \in J$ and $\pi(i') < \pi(j')$ for all $j' \in J'$. We now derive the value of $W[(i,J),(i'J')]$. This value depends on whether $i = i'$ or $i \neq i'$. Let $|J \cap J'| = n$. By definition, $|J| = |J'| = k-1$ so $|J \cup J'| = 2k-2-n$.

Suppose $i = i'$ and let $L = \{i\} \cup J \cup J'$. Then $|L| = 2k-1-n < t$. Since $\mathcal{F}$ is t-rankwise independent, it is also t-restricted min-wise independent. Therefore, the number of permutations $\pi \in \mathcal{F}$ satisfying $\pi(i) < \pi(k)$ for all $k \in J \cup J'$, is $|\mathcal{F}|/(2k-1-n)$. Thus, if $i = i'$, then $W[(i,J),(i'J')] = 1/(2k-1-n)$.

Now suppose $i \neq i'$ and let $L = \{i,i'\} \cup J \cup J'$. Note that $|L| = 2k-n \leqslant t$. For $\pi \in \mathcal{F}$ to satisfy $\pi(i) < \pi(j)$ for all $j \in J$ and $\pi(i') < \pi(j')$ for all $j' \in J'$, either $\pi(i) < \pi(\ell)$ for all $\ell \in L \setminus \{i\}$ and $\pi(i') < \pi(j')$ for all $j' \in J'$, or $\pi(i') < \pi(\ell)$ for all $\ell \in L \setminus \{i'\}$ and $\pi(i) < \pi(j)$ for all $j \in J$. Note that a single permutation cannot satisfy both of these cases. The number of permutations τ of L that satisfy $\tau(i) < \tau(\ell)$ for all $\ell \in L \setminus \{i\}$ and $\tau(i') < \tau(j')$ for all $j' \in J'$ is $(2k-n)!/(k(2k-n))$. As $\mathcal{F}$ is t-rankwise independent, it is also $(2k-n)$-rankwise independent. Therefore, the number of permutations $\pi \in \mathcal{F}$ satisfying $\pi(i) < \pi(\ell)$ for all $\ell \in L \setminus \{i\}$ and $\pi(i') < \pi(j')$ for all $j' \in J'$ is $|\mathcal{F}|/(k(2k-n))$. A similar argument shows that the same number of permutations $\pi \in \mathcal{F}$ satisfy $\pi(i') < \pi(\ell)$ for all $\ell \in L \setminus \{i'\}$ and $\pi(i) < \pi(j)$ for all $j \in J$. Therefore, if $i \neq i'$, then $W[(i,J),(i',J')] = 2/(k(2k-n))$.

Having derived the value of every entry in W, Itoh, Takei and Tarui show that this matrix is non-singular. Therefore, $|\mathcal{F}| \geqslant \operatorname{rank}(U) \geqslant \operatorname{rank}(W) = B(v,t)$. Thus, a $\mathsf{PSCA}(v,t,\lambda)$ must have at least $B(v,t)$ permutations and so, $\mathsf{PSCAN}(v,t) \geqslant B(v,t)/t!$ □

Both $m(v-1,t-1)/t!$ and $B(v,t)/t!$ give lower bounds for $\mathsf{PSCAN}(v,t)$ that are polynomial in v for fixed t. Now we assess the conditions under which each bound is greater. Although the factor of $1/t!$ is necessary in the lower bound in Theorem 4.5, for the purposes of our present analysis, we can ignore it and focus on comparing $m(v-1,t-1)$ and $B(v,t)$ for fixed t. If t is even, the degree of $m(v-1,t-1)$ is $(t-2)/2$ while the degree of $B(v,t)$ is $t/2$. Hence, for even t and sufficiently large v, $B(v,t) > m(v-1,t-1)$. For odd t, the degree of both $m(v-1,t-1)$ and $B(v,t)$ is $(t-1)/2$. To decide which provides the greater bound for sufficiently large v, we can compare the leading coefficient of each polynomial. As $t-1$ is even, the leading term of $m(v-1,t-1)$ comes from the binomial coefficient $\binom{v-1}{(t-1)/2}$. Hence the leading coefficient of $m(v-1,t-1)$ is $1/((t-1)/2)!$. Consider the leading term of $B(v,t)$. Since t is odd, $\lfloor t/2 \rfloor = (t-1)/2$. Let r be the residue when v is divided by $(t-1)/2$ so $\lfloor v/((t-1)/2) \rfloor = (v-r)/((t-1)/2)$. Then,

$$B(v,t) = \frac{v-r}{\frac{t-1}{2}} \binom{\frac{(t-3)v+2r}{t-1}}{\frac{t-3}{2}}.$$

Therefore, the leading coefficient of $B(v,t)$ is

$$\left(\frac{t-3}{t-1}\right)^{\frac{t-3}{2}} \frac{1}{(\frac{t-1}{2})!} < \frac{1}{(\frac{t-1}{2})!}.$$

Therefore, if t is odd then for sufficiently large v, $m(v-1,t-1) > B(v,t)$.

The third lower bound for $\mathsf{PSCAN}(v,t)$ that we discuss here is due to Yuster [64] and applies to values of t for which $t/2$ is a prime. The proof of this bound makes use of a result of Wilson [60] on the rank of set-inclusion matrices over finite fields. Let r,t,v be integers such that $1 \leqslant t \leqslant \min\{r, v-r\}$. The *set-inclusion matrix* $W_{t,r,v}$ is the matrix with rows indexed by all t-subsets of $[v]$, columns indexed by all r-subsets of $[v]$ and with $W_{t,r,v}[T,R] = 1$ if $T \subseteq R$ and $W_{t,r,v}[T,R] = 0$ otherwise. For a matrix A and prime p, let $\text{rank}_p(A)$ be the rank of A over the finite field $GF(p)$.

Lemma 4.6 ([60]). *Let p be a prime, let t,r,v be integers as defined above, and let $D(r,t)$ be the set of integers i such that $0 \leqslant i \leqslant t$ and $\binom{r-i}{t-i} \not\equiv 0 \pmod{p}$. Then,*

$$\text{rank}_p(W_{t,r,v}) = \sum_{i \in D(r,t)} \left(\binom{v}{i} - \binom{v}{i-1}\right) \geqslant \binom{v}{t} - \binom{v}{t-1}.$$

Theorem 4.7 ([64]). *If $t/2$ is a prime, then for all $v \geqslant t$,*

$$\mathsf{PSCAN}(v,t) \geqslant \frac{\binom{v}{t/2} - \binom{v}{t/2-1}}{t!}.$$

Proof. Let $p = t/2$ be a prime and let X be a $\mathsf{PSCA}(v,t,\lambda)$. Let A be the matrix whose columns are indexed by the permutations in X and whose rows are indexed by the sequences in $\mathcal{S}_{v,p}$ with $A[s,\pi] = 1$ if π covers s and $A[s,\pi] = 0$ otherwise. Let $C = AA^T$. Then, the rows and columns of C are indexed by the sequences in $\mathcal{S}_{v,a}$ with $C[s,u]$ being equal to the number of permutations in X that cover both s and u. Moreover, note that since A has $t!\lambda$ columns, $\text{rank}(C) \leqslant \text{rank}(A) \leqslant t!\lambda$.

The proof splits into two cases: when $p = 2$ and when p is an odd prime. We first consider the case when $p = 2$. Let $a,b,c,d \in [v]$ be distinct and let $s = (a,b)$. Then, using the fact that X is a $\mathsf{PSCA}(v,4,\lambda)$, we can determine $C[s,u]$ for any sequence u of two elements of $\{a,b,c,d\}$. This is outlined in Table 4.1.

We can see from Table 4.1 that every entry in C is divisible by 2λ. Let $D = C/2\lambda$ and note that $\text{rank}(D) = \text{rank}(C)$. Consider D over the finite field $GF(2)$. Over this field, $D[(a,b),(c,d)] = 1$ if $\{a,b\} \cap \{c,d\} = \emptyset$ and $D[(a,b),(c,d)] = 0$ otherwise. Now consider the submatrix D' of D whose rows and columns are indexed by sequences $(a,b) \in \mathcal{S}_{v,2}$

	u						
	(a,b)	(b,a)	(a,c) (a,d)	(b,c) (b,d)	(c,a) (d,a)	(c,b) (d,b)	(c,d) (d,c)
$C[s,u]$	12λ	0	8λ	4λ	4λ	8λ	6λ

Table 4.1: Values for $C[s,u]$ in the proof of Theorem 4.7 for the case $p=2$ where $s=(a,b)$.

with $a<b$. We can then relabel the row of D' indexed by (a,b) with the set $\{a,b\}$ and we can relabel the column of D' indexed by (c,d) with the set $[v]\setminus\{c,d\}$. Now, over $GF(2)$, $D'[T,R]=1$ if $T\subset R$ and $D'[T,R]=0$ otherwise. That is, over $GF(2)$, D' is equal to the set-inclusion matrix $W_{2,v-2,v}$. Therefore, by Lemma 4.6,

$$\binom{v}{2}-\binom{v}{1}\leqslant \operatorname{rank}_2(D')\leqslant \operatorname{rank}(D)\leqslant \operatorname{rank}(C)\leqslant 4!\lambda.$$

Therefore, $\lambda\geqslant(\binom{v}{2}-\binom{v}{1})/4!$.

Now let p be an odd prime and let $s,u\in\mathcal{S}_{v,p}$ be such that $s\cap u=\emptyset$ where $s\cap u$ is the set of elements of $[v]$ that appear in both s and u. Let $s\cup u\subset[v]$ be the set of elements of $[v]$ appearing in either s or u and note that since $s\cap u=\emptyset$, $|s\cup u|=t$. Therefore, there are exactly $\binom{t}{p}$ sequences in $\mathcal{S}_{v,t}$ that cover both s and u. Hence, $C[s,u]=\binom{t}{p}\lambda$. As p is an odd prime, $C[s,u]$ is not divisible by $p\lambda$. Now suppose $s\cap u\neq\emptyset$. Let $Y=s\cup u$ and fix $Z\subset[v]$ such that $Z\cap Y=\emptyset$ and $|Z\cup Y|=t$. Let $z=|Z|$. Then, for each permutation π of Y that covers both s and u, there are $\binom{t}{z}z!$ permutations of $Y\cup Z$ that cover π. Therefore, the total number of permutations of $Y\cup Z$ that cover both s and u is divisible by t. Each of these permutations is a sequence in $\mathcal{S}_{v,t}$ which must be covered by λ permutations in X. Therefore, $C[s,u]$ is divisible by $t\lambda$ and thus, $C[s,u]$ is also divisible by $p\lambda$.

Let $D=C/g$ where g is the greatest common divisor of all the entries in C. Note that λ divides g but $p\lambda$ does not. Then, over the field $GF(p)$ and for some non-zero $d\in GF(p)$, $D[s,u]=d$ if $s\cap u=\emptyset$ and $D[s,u]=0$ otherwise. Let D' be the submatrix of D with rows and columns indexed by the increasing sequences in $\mathcal{S}_{v,p}$. Then, with a similar relabelling of rows and columns as in the $p=2$ case, we find that over the field $GF(p)$, the matrix $d^{-1}D'$ is the set-inclusion matrix $W_{p,v-p,v}$. Again, using Lemma 4.6,

$$\binom{v}{p}-\binom{v}{p-1}\leqslant \operatorname{rank}_p(d^{-1}D)=\operatorname{rank}_p(D)\leqslant \operatorname{rank}(D)\leqslant \operatorname{rank}(C)\leqslant t!\lambda.$$

Therefore, $\lambda\geqslant(\binom{v}{p}-\binom{v}{p-1})/t!$. □

When $t/2$ is a prime, Theorem 4.7 gives a lower bound for $\mathsf{PSCAN}(v,t)$ that is polynomial in v with degree $t/2$. We know from our analysis of the bounds in Theorem 4.5 that for even t, the degree of $B(v,t)$ is also $t/2$. The leading coefficient of the bound in Theorem 4.7 is $1/(t/2)!t!$. Using similar arguments as the odd t case, we find for even t, the leading coefficient of $B(v,t)/t!$ is

$$\left(\frac{t-2}{t}\right)^{\frac{t-2}{2}} \frac{1}{(\frac{t}{2})!t!} < \frac{1}{(\frac{t}{2})!t!}.$$

Therefore, when $t/2$ is prime and v is sufficiently large, Theorem 4.7 gives a larger lower bound than Theorem 4.5.

Theorem 4.7 can also be used to give lower bounds for values of t that are not twice a prime. For example, $5/2$ is not prime but a strength 5 PSCA is also a strength 4 PSCA. By Theorem 4.7, $\mathsf{PSCAN}(v,4) \geqslant v(v-3)/48$. By Lemma 4.1, $\mathsf{PSCAN}(v,5) \geqslant \mathsf{PSCAN}(v,4)/5 \geqslant v(v-3)/240$. However, $m(v-1,4)/5! = (v^2-v+2)/240$ gives a better lower bound for $\mathsf{PSCAN}(v,5)$ for all $v \geqslant 5$.

Let $t = 2p+1$ for some prime p. Then, by Theorem 4.7 and Lemma 4.1,

$$\mathsf{PSCAN}(v,t) \geqslant \frac{\binom{v}{p} - \binom{v}{p-1}}{t!}. \tag{4.1}$$

We know that for odd t, the better lower bound for $\mathsf{PSCAN}(v,t)$ given by Theorem 4.5 is $m(v-1,t-1)/t!$. Both Theorems 4.5 and (4.1) give lower bounds that are polynomial in v with degree p. In fact, the leading coefficient of both of these polynomials is $1/p!t!$. However, if we consider the coefficient of v^{p-1} in each polynomial, we find that this coefficient in $m(v-1,t-1)/t!$ is $(1/t!)(1/(p-1)! - \binom{p}{2}/p!)$. The corresponding coefficient in the polynomial given by (4.1) is $(1/t!)(-1/(p-1)! - \binom{p}{2}/p!)$. Therefore, when $t = 2p+1$ for some prime p and for sufficiently large v, $m(v-1,t-1)/t!$ provides a greater lower bound for $\mathsf{PSCAN}(v,t)$ than (4.1).

We could continue to use Theorem 4.7 and Lemma 4.1 to derive lower bounds for $\mathsf{PSCAN}(v,t)$ with $t = 2p+i$ for prime p and $i > 1$. However the resulting polynomial bound would have lower degree than $B(v,t)$ for even i and lower degree than $m(v-1,t-1)$ for odd i. Combining these analyses, we can characterise the lower bounds of Theorems 4.5 and 4.7 in the following way.

Theorem 4.8. *For integers $v \geqslant t \geqslant 4$,*

$$\mathsf{PSCAN}(v,t) \geqslant \begin{cases} m(v-1,t-1)/t! & \text{if } t \text{ is odd,} \\ \frac{\binom{v}{t/2} - \binom{v}{t/2-1}}{t!} & \text{if } t/2 \text{ is prime,} \\ B(v,t)/t! & \text{if } t \text{ is even and } t/2 \text{ is not prime.} \end{cases}$$

Ultimately, we find that the growth of PSCAN is at least polynomial in v for fixed t. The degree of this polynomial lower bound is $t/2$ for even t and $(t-1)/2$ for odd t. This growth is much larger than the logarithmic growth of $\mathsf{SCAN}(t,v)$.

Next we describe three constructions that give upper bounds for $\mathsf{PSCAN}(v,t)$. The first two of these are due to Tarui, Itoh and Takei [54] and are constructions for families of 3- and 4-rankwise independent permutations respectively. Both of these constructions are recursive and each make use of a different finite geometry. The third construction is due to Itoh, Tarui and Takei [26] and can be applied to any choice of v and t. This construction converts polynomials over a finite field with degree less than t into a family of t-rankwise independent permutations.

Theorem 4.9 ([54, 28])**.** *For $v \geqslant 3$,*

$$\mathsf{PSCAN}(v,3) \leqslant \frac{\sqrt{e}}{2}(1+o(1))v\log_2(v)^2.$$

Proof. Let $q = 2^t$ for some integer $t \geqslant 2$. The bound is shown by a recursive construction of a 3-rankwise independent family of permutations in $\mathcal{S}_{q^2}$ from a 3-rankwise independent family of permutations in $\mathcal{S}_q$ and the affine plane $\mathrm{AG}(2,q)$ where the points of $\mathrm{AG}(2,q)$ are associated with the elements of $[q^2]$.

Let G_q be a family of 3-rankwise independent permutations in $\mathcal{S}_q$ and let $\mathcal{C}$ be the set of parallel classes of $\mathrm{AG}(2,q)$. Fix $C \in \mathcal{C}$. Then C is a set $C = \{\ell_0, \dots, \ell_{q-1}\}$ of q lines of $\mathrm{AG}(2,q)$ such that every point is contained in exactly one line of C. For $0 \leqslant i \leqslant q-1$, let $\ell_i = \{p_{i,0}, \dots, p_{i,q-1}\}$. So $\{p_{i,j} : 0 \leqslant i,j \leqslant q-1\} = [q^2]$. For each $\pi \in G_q$ we define the permutation $\sigma_\pi \in \mathcal{S}_{q^2}$ by

$$\sigma_\pi(p_{i,j}) = \pi(i)q + \pi(j).$$

Points on the line ℓ_i are mapped to the interval $\{\pi(i)q, \dots, \pi(i)q + q - 1\}$. The order in which the points are mapped to this interval is determined by π. Additionally, we define a second permutation $\tau_\pi \in \mathcal{S}_{q^2}$ by

$$\tau_\pi(p_{i,j}) = \pi(i)q + q - 1 - \pi(j).$$

In τ_π, the same lines are mapped to the same intervals, but the order of the points on each line has been reversed.

As an example, let $q = 4$ and let C be such that $p_{i,j} = 4i + j$ for $0 \leqslant i,j \leqslant 3$. Then, $C = \{\{0,1,2,3\},\{4,5,6,7\},\{8,9,10,11\},\{12,13,14,15\}\}$. Let $\pi = 2301$. Then,

$$\sigma_\pi = (10,11,8,9,14,15,12,13,2,3,0,1,6,7,4,5)$$

and

$$\tau_\pi = (9,8,11,10,13,12,15,14,1,0,3,2,5,4,7,6).$$

We let $X_C = \bigcup_{\pi \in G_q}\{\sigma_\pi, \tau_\pi\}$ and then let $G_{q^2} = \bigcup_{C \in \mathcal{C}} X_C$. It is shown in [54] that G_{q^2} is a 3-rankwise independent family of permutations in $\mathcal{S}_{q^2}$. For the base of their recursion, Tarui, Itoh and Takei use $\mathcal{S}_4$. As noted by Iurlano [28], a 3-rankwise independent family of six permutations in $\mathcal{S}_4$ can be used instead, reducing the bound reported in [54] by a factor of 4. We also note that Yuster [64] independently describes a very similar construction for $\mathsf{PSCA}(v,3)$ but reports a larger upper bound for $\mathsf{PSCAN}(v,3)$. □

Theorem 4.10 ([54, 28])**.** *For $v \geqslant 4$,*

$$\mathsf{PSCAN}(v,4) \leqslant \frac{e}{8}(1+o(1))v^3\log_2(v)^6.$$

Proof. Let $q = 2^t$ for some integer $t \geqslant 2$. The bound is shown by a recursive construction of a 4-rankwise independent family of permutations in $\mathcal{S}_{q^2+1}$ from a 4-rankwise independent family of permutations in $\mathcal{S}_{q+1}$, the projective plane $\mathrm{PG}(2,q)$ and an orthogonal array $\mathsf{OA}_{q/2}(3,q+1,2)$, the existence of which is guaranteed by Alon, Babai and Itai [1].

Let G_{q+1} be a family of 4-rankwise independent permutations in $\mathcal{S}_{q+1}$ and let A be an $\mathsf{OA}_{q/2}(3,q+1,2)$. Then A has $4q$ rows indexed by $[4q]$ and $q+1$ columns indexed by $[q+1]$. For a point $a \in \mathrm{PG}(2,q)$, there are $q+1$ lines $\ell_0^a,\ldots,\ell_q^a$ containing a and every point in $\mathrm{PG}(2,q)$ distinct from a is contained in exactly one of these lines. Let $\ell_i^a = \{p_{i,0}^a,\ldots,p_{i,q-1}^a,a\}$. For $\pi \in \mathcal{S}_{q+1}$, let $\pi' \in \mathcal{S}_q$ be defined by

$$\pi'(i) = \begin{cases} \pi(i) & \text{if } \pi(i) < \pi(q) \\ \pi(i)-1 & \text{if } \pi(i) > \pi(q). \end{cases}$$

For $a \in \mathrm{PG}(2,q)$, $r \in [4q]$ and $\pi \in G_{q+1}$, we define $\sigma_{a,r,\pi} \in \mathcal{S}_{q^2+q+1}$ by $\sigma_{a,r,\pi}(a) = q^2+q$ and

$$\begin{aligned}\sigma_{a,r,\pi}(p_{i,j}^a) = q(&\mathsf{A}[r,q]\pi(i) + (1-\mathsf{A}[r,q])(q-\pi(i))) \\ &+ \mathsf{A}[r,i]\pi'(j) + (1-\mathsf{A}[r,i])(q-1-\pi'(j)).\end{aligned}$$

Points on the line ℓ_i^a are mapped by $\sigma_{a,r,\pi}$ to points on the interval $\{\pi(i)q,\ldots,\pi(i)q+q-1\}$ if $\mathsf{A}(r,q) = 1$ and are mapped by $\sigma_{a,r,\pi}$ to points on the interval $\{(q-\pi(i))q,\ldots,(q-\pi(i))q+q-1\}$ if $\mathsf{A}(r,q) = 0$. The order of the points on ℓ_i^a is determined by π' if $\mathsf{A}(r,i) = 1$ or $q-1-\pi'$ if $\mathsf{A}(r,i) = 0$. The point a is then mapped to q^2+q. We also define $\tau_{a,r,\pi} \in \mathcal{S}_{q^2+q+1}$ by $\tau_{a,r,\pi}(a) = 0$ and $\tau_{a,r,\pi}(p_{i,j}^a) = \sigma_{a,r,\pi}(p_{i,j}^a)+1$.

As an example, let $q = 4$. The sets $\{\{0+i,1+i,4+i,14+i,16+i\} : i \in \mathbb{Z}_{21}\}$, where addition is performed modulo 21, are known to form the lines of $\mathrm{PG}(2,4)$ [50]. Let $a = 0 \in \mathrm{PG}(2,4)$. Then, $\ell_0^0 = \{1,4,14,16,0\}$, $\ell_1^0 = \{2,7,8,11,0\}$, $\ell_2^0 = \{3,13,15,20,0\}$, $\ell_3^0 = \{5,6,9,19,0\}$ and $\ell_4^0 = \{10,12,17,18,0\}$. Suppose A is an $\mathsf{OA}_2(3,5,2)$ such that $\mathsf{A}[r] = [0,0,0,1,1]$. Let $\pi = 10342 \in \mathcal{S}_5$. Then $\pi' = 1023 \in \mathcal{S}_4$. The following table describes $\sigma_{a,r,\pi}$.

x	0	1	2	3	4	5	6	7	8	9	10
$\sigma_{a,r,\pi}(x)$	20	6	2	14	7	17	16	3	1	18	9

x	11	12	13	14	15	16	17	18	19	20
$\sigma_{a,r,\pi}(x)$	0	8	15	5	13	4	10	11	19	12

The following table describes $\tau_{a,r,\pi}$.

x	0	1	2	3	4	5	6	7	8	9	10
$\tau_{a,r,\pi}(x)$	0	7	3	15	8	18	17	4	2	19	10

x	11	12	13	14	15	16	17	18	19	20
$\tau_{a,r,\pi}(x)$	1	9	16	6	14	5	11	12	20	13

Let $G_{q^2+q+1} = \{\sigma_{a,r,\pi}, \tau_{a,r,\pi} : a \in \mathrm{PG}(2,q), r \in [4q], \pi \in G_{q+1}\}$. It is shown in [54] that G_{q^2+q+1} is a 4-rankwise independent family of permutations in $\mathcal{S}_{q^2+q+1}$. This can be reduced to a family of 4-rankwise independent permutations in $\mathcal{S}_{q^2+1}$ which can then be used for the next step of the recursion. Tarui, Itoh and Takei choose $\mathcal{S}_5$ as the base of this recursion. As noted by Iurlano [28], a 4-rankwise family of 24 permutations in $\mathcal{S}_5$ can be used instead, reducing the bound reported in [54] by a factor of 5. □

Theorem 4.11 ([26]). *For any integers $v \geqslant t \geqslant 1$, there exists a t-rankwise independent family of permutations $X \in \mathcal{S}_v$ such that $|X| = v^{O(t^2/\ln t)}$ if $v \geqslant (t-1)!$ and $|X| = e^{O(t^3)}$ otherwise.*

Proof. Let $v \geqslant t \geqslant 1$ and let $p \geqslant v$ be a prime number. Let $\mathbb{F}_p(\xi, t-1)$ be the set of univariate polynomials over the field with p elements, $\mathbb{F}_p$, with degree between 1 and $t-1$. The key to this construction is to transform each of these polynomials into permutations of the field elements. Of course, a polynomial may map two distinct elements to the same element. Itoh, Tarui and Takei develop t-wise independent random variables that act as tie-breakers, allowing each polynomial in $\mathbb{F}_p(\xi, t-1)$ to be converted to a permutation in $\mathcal{S}_p$. This family of permutations is then shown to be t-rankwise independent. □

In addition to these constructive results, the following result proved by Kuperberg, Lovett and Peled [35] guarantees the existence of a t-wise independent family of permutations in $\mathcal{S}_v$ with size at most $(cv)^{ct}$ for some constant $c > 0$. We do not discuss the proof of this theorem here, other than to say it uses a broader probabilistic framework described in [35].

Theorem 4.12 ([35]). *For all integers $v \geqslant 1$ and $1 \leqslant t \leqslant v$, there exists a t-wise independent family of permutations $T \subseteq \mathcal{S}_v$ with $|T| \leqslant (cv)^{ct}$ for some universal constant $c > 0$.*

If we reinterpret Theorem 4.12 in terms of the perfect sequence covering array number, we obtain the following upper bound on $\mathsf{PSCAN}(v,t)$.

Theorem 4.13. $\mathsf{PSCAN}(v,t) \leqslant (cv)^{ct}/t!$ *for some universal constant $c > 0$.*

In particular, $\mathsf{PSCAN}(v,t)$ grows as a polynomial in v for any fixed t. This answers a question posed by Yuster [64]. At present, Theorem 4.13 gives the best general upper bound for $\mathsf{PSCAN}(v,t)$, however it is unknown how to efficiently generate a $\mathsf{PSCA}(v,t)$ with that multiplicity.

4.2 Exact values of PSCAN

In addition to the bounds discussed in Chapter 4.1, the exact value of $\mathsf{PSCAN}(v,t)$ is known for certain choices of v and t. By Theorem 3.17, we know that a $\mathsf{PSCA}(t+1,t,1)$ exists and hence, $\mathsf{PSCAN}(t+1,t) = 1$ for all $t \geqslant 3$. By Theorem 3.19, we know that $\mathsf{PSCAN}(v,t) > 1$ for many small values for v and t. For these values of v and t, the size of a theoretical $\mathsf{PSCA}(v,t,\lambda)$ may be small enough to support computational searches for these objects. In particular for $t = 3$, where the size of a $\mathsf{PSCA}(v,3,\lambda)$ is 6λ, and the growth of $\mathsf{PSCAN}(v,3)$ is at most quasi-linear in v by Theorem 4.9, computational searches may feasibly uncover exact values of $\mathsf{PSCAN}(v,3)$ for small values of v. Hence, since Yuster [64] introduced perfect sequence covering arrays and defined PSCAN in 2020, authors have designed different computational methods to explicitly construct $\mathsf{PSCA}(v,t,\lambda)$ and hence find $\mathsf{PSCAN}(v,t)$ for small values of v and t. In Chapter 7, we discuss our own computational methods that independently confirm many of the results we discuss in this chapter as well as uncover new values of $\mathsf{PSCAN}(v,t)$. The first to determine exactly a value of $\mathsf{PSCAN}(v,t)$ greater than one was Yuster [64] who proved that $\mathsf{PSCAN}(5,3) = 2$. By Theorem 3.19, it was already known that $\mathsf{PSCAN}(5,3) > 1$, so the proof of this result simply requires the existence of a $\mathsf{PSCA}(5,3,2)$.

Theorem 4.14 ([64]). $\mathsf{PSCAN}(5,3) = 2$

Proof. Let X be the following subset of $\mathcal{S}_5$:

$$X = \{12345, 43215, 35214, 14523, 25413, 53412\}.$$

Yuster notes that X, which covers 56 of the 60 sequences in $\mathcal{S}_{5,3}$, covers the maximum number of sequences in $\mathcal{S}_{5,3}$ for any set of six permutations in $\mathcal{S}_5$. The four sequences that are not covered by X are $(1,3,2)$, $(2,3,1)$, $(1,5,4)$ and $(4,5,1)$. Moreover, the sequences $(1,2,3)$, $(3,2,1)$, $(1,4,5)$ and $(5,4,1)$ are each covered by two permutations in X. Note that the uncovered sequences can be mapped to the sequences covered twice by swapping 2 and 3 and swapping 4 and 5. Therefore, if $\sigma = 13254 \in \mathcal{S}_5$, then $X \cup \sigma X$ is a $\mathsf{PSCA}(5,3,2)$. For concreteness, we list the permutations of σX below.

$$\sigma X = \{13254, 52314, 24315, 15432, 34512, 42513\}.$$ □

In work concurrent with the work we describe in Chapter 7, Na, Jedwab and Li [44] describe a search algorithm for perfect sequence covering arrays that adapts an algorithm of Mathon [40] to $\mathsf{PSCA}(v,t,\lambda)$ with $\lambda > 1$. Let A be the matrix with rows indexed by the permutations in $\mathcal{S}_v$ and columns indexed by the sequences in $\mathcal{S}_{v,t}$ with $A[\pi,s] = 1$ if π covers s and $A[\pi,s] = 0$ otherwise. Then a $\mathsf{PSCA}(v,t,\lambda)$ is a set of permutations X such that for all $s \in \mathcal{S}_{v,t}$,

$$\sum_{\pi \in X} A[\pi,s] = \lambda.$$

Thus, the search for a $\mathsf{PSCA}(v,t,\lambda)$ can be reinterpreted as a search for a set of rows in the matrix A that sum to a uniform vector. Na, Jedwab and Li [44] describe an algorithm to perform such a search. Moreover, this algorithm is exhaustive. If a $\mathsf{PSCA}(v,t,\lambda)$ exists, then the algorithm will output such an array whereas if the algorithm terminates without output, then a $\mathsf{PSCA}(v,t,\lambda)$ does not exist. Using this search algorithm, the following values of PSCAN are found.

Theorem 4.15 ([44])**.** $\mathsf{PSCAN}(6,3) = \mathsf{PSCAN}(7,3) = 2$.

Proof. By Theorem 3.19, $\mathsf{PSCAN}(6,3) > 1$ and $\mathsf{PSCAN}(7,3) > 1$ so the proof of this theorem requires only the existence of a $\mathsf{PSCA}(6,3,2)$ and a $\mathsf{PSCA}(7,3,2)$. The following 12 permutations in $\mathcal{S}_6$ form a $\mathsf{PSCA}(6,3,2)$.

123456	154326	216543	245613	354162	361452
423165	461325	516234	532614	632541	645231

The following 12 permutations in $\mathcal{S}_7$ form a $\mathsf{PSCA}(7,3,2)$.

1234567	1573426	3275641	3617524	4261735	4756123
5164327	5243176	6257314	6345721	7216453	7431625

□

Na, Jedwab and Li [44] also adapt this algorithm to search over sets of cosets of permutation groups rather than just individual permutations. For a permutation group $G \leqslant \mathcal{S}_v$, we define the matrix A_G similarly to the matrix A defined above. The columns of A_G are still indexed by the sequences in $\mathcal{S}_{v,t}$ but the rows are now indexed by the right cosets of G in $\mathcal{S}_v$ with $A_G[G\pi, s]$ being equal to the number of permutations in $G\pi$ that cover s. In other words, the row of A_G corresponding to $G\pi$ can be found by taking the sum of the rows of A indexed by the elements of $G\pi$. Alternatively, one could define A_G by letting the rows be indexed by the left cosets of G. However, for reasons we will soon discuss, Na, Jedwab and Li favoured defining A_G in terms of right cosets. Note that if G is the trivial group, then $A_G = A$. Now, if X is a set of rows of A_G such that for all $s \in \mathcal{S}_{v,t}$,

$$\sum_{r \in X} A_G[r,s] = \lambda,$$

then the union of the cosets represented by the rows in X forms a $\mathsf{PSCA}(v,t,\lambda)$. Hence, the search algorithm performs a similar role as before: finding a set of rows of A_G that sum to a uniform vector. The obvious advantage of this method is that the required depth the algorithm needs to reach to find a PSCA is $t!\lambda/|G|$ instead of $t!\lambda$. However, if this algorithm fails to find a $\mathsf{PSCA}(v,t,\lambda)$, that is not sufficient proof that such an array does not exist.

Recall that we have a choice in deciding whether to search over left cosets or right cosets of G. The following lemma proved by Na, Jedwab and Li [44], allows for a greater simplification of the search algorithm if right cosets are chosen. Two permutation groups $G, H \leqslant \mathcal{S}_v$ belong to the same *conjugacy class* if $H = \sigma^{-1}G\sigma$ for some $\sigma \in \mathcal{S}_v$.

Lemma 4.16 ([44]). *Let $G, H \leqslant \mathcal{S}_v$ be permutation groups belonging to the same conjugacy class. If there exists a $\mathsf{PSCA}(v,t,\lambda)$ that is the union of right cosets of H, then there is also a $\mathsf{PSCA}(v,t,\lambda)$ that is the union of right cosets of G.*

Proof. Let $\mathcal{R}$ be a set of right coset representatives of H such that

$$\bigcup_{\pi \in \mathcal{R}} H\pi$$

is a $\mathsf{PSCA}(v,t,\lambda)$. As H is in the same conjugacy class as G, $H = \sigma^{-1}G\sigma$ for some $\sigma \in \mathcal{S}_v$. Therefore,

$$\bigcup_{\pi \in \mathcal{R}} \sigma^{-1}G\sigma\pi$$

is a $\mathsf{PSCA}(v,t,\lambda)$. By Theorem 3.22,

$$\sigma \bigcup_{\pi \in \mathcal{R}} \sigma^{-1}G\sigma\pi = \bigcup_{\pi \in \mathcal{R}} G\sigma\pi$$

is a $\mathsf{PSCA}(v,t,\lambda)$ that is the union of right cosets of G. □

Lemma 4.16 shows that searching over the right cosets of every permutation subgroup of $\mathcal{S}_v$ is equivalent to searching over the right cosets of a representative from each conjugacy class of permutation subgroups of $\mathcal{S}_v$, greatly reducing the computations needed to perform such an exhaustive search. The following results come from applications of this group-based search algorithm.

Theorem 4.17. $\mathsf{PSCAN}(7,4) = 2$ *while* $\mathsf{PSCAN}(8,3) \in \{2,3\}$, $\mathsf{PSCAN}(9,3) \in \{2,3,4\}$ *and* $\mathsf{PSCAN}(7,5) \in \{2,3,4\}$.

Proof. Again, by Theorem 3.19, the proof of this theorem requires only the existence of suitable $\mathsf{PSCA}(v,t,\lambda)$. First, let $G = \langle 4725621 \rangle \leqslant \mathcal{S}_7$. This group is isomorphic to the cyclic group of order 6. Then, the union of the 8 right cosets of G with the following representatives forms a $\mathsf{PSCA}(7,4,2)$.

1254736	1347256	1352746	1362745
1365472	1462537	1745236	1765234

Next, let $G = \langle 85672341 \rangle \leqslant \mathcal{S}_8$. This group is isomorphic to the cyclic group of order 2. Then, the union of the 9 right cosets of G with the following representatives forms a $\mathsf{PSCA}(8,3,3)$.

12345678	15468237	17624385	27561843	28461573
31864275	32654187	37461528	47218653	

Next, let $G = \langle 768241593 \rangle \leqslant \mathcal{S}_9$. This group is isomorphic to the cyclic group of order 6. Then, the union of the 4 left cosets with the following representatives forms a $\mathsf{PSCA}(9, 3, 4)$.

$$123456897 \quad 154372968 \quad 198426537 \quad 318697425$$

Finally, let $G = \langle 7261354, 4216537 \rangle \leqslant \mathcal{S}_7$. This group is isomorphic to $\mathcal{S}_4$. Then, the union of the 20 right cosets of G with the following representatives forms a $\mathsf{PSCA}(7, 5, 4)$.

1234567	1236574	1324576	1324675	1325476
1325674	1347265	1354267	1357642	1364275
1364572	1367542	1374256	1375624	1623574
1635247	1637425	2137465	2137654	2163457

□

Strategies to limit the search space for perfect sequence covering arrays have been proposed by other authors. For example, Iurlano [28] suggests limiting searches to reflection symmetric matrices. Here, a PSCA is understood to be a matrix where the rows are permutations of $[v]$. A matrix is *reflection symmetric* if in any pair of columns, and for any distinct $a, b \in [v]$, the pair (a, b) appears equally as often as the pair (b, a). Iurlano [28] observes that the $\mathsf{PSCA}(6, 4, 1)$ on the left hand side of Figure 3.1 is reflection symmetric and claims the existence of a reflection symmetric $\mathsf{PSCA}(6, 3, 2)$. Iurlano then suggests that this property may be of use in finding a $\mathsf{PSCA}(9, 3, 3)$. It is as yet unknown whether such an array exists and, as we will see later, $v = 9$ is the smallest value of v for which $\mathsf{PSCAN}(v, 3)$ is undetermined.

In Chapter 7, we will discuss our own exhaustive computational search methods for perfect sequence covering arrays devised during an investigation that ran independently and concurrently with some of the aforementioned authors [44, 28]. We confirm many of the values of $\mathsf{PSCAN}(v, t)$ proved by Na, Jedwab and Li [44] while also establishing new values.

Chapter 5

Covering Arrays

In this chapter we introduce and discuss covering arrays. Of particular interest is a stronger form of a covering array introduced by Chee et al. [9] called an *excess coverage array*. These arrays have connections to sequence covering arrays and have proved to be a useful tool in investigating Conjecture 3.18. The study of these objects motivates our research in Chapter 9, where we use excess coverage arrays to prove that an $\mathsf{SCA}(7!; 7, 10)$ does not exist.

Let C be an $N \times k$ array where each entry in C is a symbol from the alphabet $[v]$. A *t-way interaction* is a set of t pairs $\{(c_i, \nu_i) : 0 \leqslant i \leqslant t-1\}$ where each c_i is a column of C such that $c_i \neq c_j$ for $i \neq j$, and each ν_i is an element of $[v]$. The row ρ of C *covers* the interaction $\{(c_i, \nu_i) : 0 \leqslant i \leqslant t-1\}$ if the entry of C in row ρ and column c_i is ν_i for $0 \leqslant i \leqslant t-1$. The array C is a *covering array of strength* t, denoted by $\mathsf{CA}(N; t, k, v)$, if for each t-way interaction T, there is some row of C that covers T.

Covering arrays can also be described in the language of set partitions. A set X of partitions of the set V is *t-qualitatively independent* if any t parts from t distinct partitions in X have a non-empty intersection. The concept of t-qualitatively independent partitions was introduced by Marczewski [38]. Let A be a $\mathsf{CA}(N; t, k, v)$. For each column $j \in [k]$, we can define a partition F_j of $[N]$ into v parts where $F_j = \{X_{j,0}, \ldots X_{j,v-1}\}$ with $X_{j,a} = \{i : \mathsf{A}[i, j] = a\}$ for $a \in [v]$. Then the t parts $X_{j_0,i_0}, \ldots, X_{j_{t-1},i_{t-1}}$ for distinct $j_0, \ldots, j_{t-1}$ correspond to a t-way interaction. As A is a $\mathsf{CA}(N; t, k, v)$ there must be some row covering this t-way interaction and so the intersection of these t parts must be non-empty. This construction can be reversed to build a $\mathsf{CA}(N; t, k, v)$ from a set of k t-qualitatively independent partitions of $[N]$, each into v parts. Thus covering arrays with strength t are equivalent to sets of t-qualitatively independent partitions with the same number of parts.

Covering arrays have been extensively studied because of their applications in software testing. Consider a system consisting of k factors, each of which can be set to one of v levels. Faults in this system may occur because of the simultaneous interaction between factors set to certain levels. Exhaustively testing all v^k arrangements of factors and levels can become infeasible. Instead, to simplify testing requirements, we may assume that faults occur because of the interaction of at most t factors. By associating each factor with a column of an array and each level with a symbol from $[v]$, we can see that a t-way interaction corresponds to a setting of t factors to certain levels. Thus, a $\mathsf{CA}(N;t,k,v)$ can be used to test all interactions between any t factors of our system. Distinct factors in real world systems may have different numbers of levels. That is, for practical reasons, it may be useful to consider covering arrays in which different columns have different alphabets with different numbers of symbols. Such arrays are called *mixed covering arrays*. For our purposes, we only consider covering arrays in which each column shares the same alphabet.

Let $\mathsf{CAN}(t,k,v)$ be the smallest N for which a $\mathsf{CA}(N;t,k,v)$ exists. As with sequence covering arrays, applications of covering arrays in software testing mean that a key area of research is to determine or bound $\mathsf{CAN}(t,k,v)$. An $\mathsf{OA}(t,k,v)$ is a $\mathsf{CA}(v^t;t,k,v)$ in which every interaction is covered *exactly* once. In particular, we have that v^t is a trivial lower bound for $\mathsf{CAN}(t,k,v)$. The value of $\mathsf{CAN}(2,k,2)$ has been determined exactly [30, 33], while for other parameter sets, explicit constructions of covering arrays give upper bounds for $\mathsf{CAN}(t,k,v)$. Methods of constructing covering arrays include the application of different computational methods, greedy and random algorithms, and recursive constructions (see [11] for details). Tables recording the best known upper bounds for $\mathsf{CAN}(t,k,v)$ for a variety of parameter sets can be found at [10].

5.1 Covering arrays with excess coverage

In this chapter we discuss a stronger version of a covering array called an *excess coverage arrays*. For such arrays, we define a function that determines how many times each interaction needs to be covered. Excess coverage arrays were introduced by Chee et al. [9] as a means of connecting covering arrays and sequence covering arrays. In particular, they use excess coverage arrays to prove that if an $\mathsf{SCA}(t!;t,v)$ exists, then $v \leqslant 2t-1$. We discuss the proof of this result in this chapter.

Let X be an $\mathsf{SCA}(N;t,v)$. For $s \in \mathcal{S}_{v,a}$ with $a \leqslant t$, let X_s be the set of permutations in X that cover s. Ishigami [25] proved that for distinct $s,u \in \mathcal{S}_{v,t-2}$, X_s is not a subset of X_u and used this observation to derive a lower bound for $\mathsf{SCAN}(t,v)$. This bound was later improved upon by Füredi [18] who constructed a $\mathsf{CA}(|X_s|;2,v-t+2,t-1)$, C, from the permutations in X_s for any $s \in \mathcal{S}_{v,t-2}$. Each row of C corresponds to a permutation in X_s and each column corresponds to a symbol in $[v]$ that does not appear in s. The entry in row i and column j of C is the number of symbols of s that precede the symbol corresponding to j in the permutation corresponding to i. There are $(t-1)^2$ interactions involving symbols $x,y \notin s$ that C needs to cover. Since each interaction specifies how many symbols of s appear before each of x and y, each interaction corresponds to at least one sequence in $\mathcal{S}_{v,t}$ that must be covered by some permutation in X_s. Through an argument similar to Theorem 5.1 below, it can be shown that C is indeed a covering array.

This connection between sequence covering arrays and covering arrays was then further explored by Chee et al. [9].

For an interaction $T = \{(c_i, \nu_i) : 0 \leqslant i \leqslant t-1\}$, and for $\sigma \in [v]$, let $\tau_\sigma(T) = \{i : \nu_i = \sigma\}$. Then, let $\mu(T) = \prod_{\sigma=0}^{v-1} |\tau_\sigma(T)|!$. Then C is a *covering array with excess coverage*, denoted by $\mathsf{CA}_X(N; t, k, v)$, if each interaction T is covered by at least $\mu(T)$ different rows of C. We also call covering arrays with excess coverage by the shorter name *excess coverage arrays*. Let $\mathsf{CAN}_X(t, k, v)$ be the smallest N for which a $\mathsf{CA}_X(N; t, k, v)$ exists. The following theorem was proved by Chee et al. [9].

Theorem 5.1 ([9]). *Let v, t and a be integers such that $v \geqslant t \geqslant 3$ and $t > a > 0$. Then* $\mathsf{SCAN}(t, v) \geqslant a!\mathsf{CAN}_X(t-a, v-a, a+1)$.

Proof. Let $N = \mathsf{SCAN}(t, v)$, let X be an $\mathsf{SCA}(N; t, v)$ and let $A \subset [v]$ with $|A| = a$. The sets X_s with s ranging over all possible sequences of the elements of A partition X. Let $u = (x_0, \ldots, x_{a-1})$ be a sequence of the elements of A, let $n = |X_u|$ and let C be an $n \times (v-a)$ array with rows indexed by the permutations in X_u and columns indexed by the symbols in $[v] \setminus A$. The entry of row ρ and column ν of C is the number of symbols of A that precede ν in the permutation ρ. Note then that each entry in C is an element of $[a+1]$.

Let $D = \{y_0, \ldots, y_{t-a-1}\}$ be a subset of $[v]$ that is disjoint with A and let $T = \{(y_i, \nu_i) : 0 \leqslant i \leqslant t-a-1\}$ be a $(t-a)$-way interaction where each ν_i is an element of $[a+1]$. Consider the permutations of $A \cup D$ that cover u and have the elements of $\tau_0(T)$ preceding x_0, the elements of $\tau_a(T)$ succeeding x_{a-1} and the elements of $\tau_i(T)$ lying between x_{i-1} and x_i for $1 \leqslant i \leqslant a-1$. There are $\mu(T)$ such permutations. Each of these corresponds to a sequence in $\mathcal{S}_{v,t}$ which must be covered by at least one permutation in X_u. Moreover, a row of C covers T if and only if the corresponding permutation in X_u covers one of these sequences. Therefore, there are at least $\mu(T)$ rows of C that cover T. Hence, C is a $\mathsf{CA}_X(n; t-a, v-a, a+1)$ and so $n \geqslant \mathsf{CAN}_X(t-a, v-a, a+1)$. As there are $a!$ sequences of the elements of A and there are at least $\mathsf{CAN}_X(t-a, v-a, a+1)$ permutations in X covering each of these sequences, $N \geqslant a!\mathsf{CAN}_X(t-a, v-a, a+1)$. □

This result extends Füredi's in two ways. First, excess coverage more accurately describes the number of times each interaction must be covered in a covering array built from X_u for some $u \in \mathcal{S}_{v,a}$. Second, Füredi only constructed covering arrays with $a = t-2$, i.e. covering arrays with strength 2, whereas Theorem 5.1 extends this construction to apply for all $0 < a < t$. Nevertheless, we focus once again on the case $a = t-2$. By Theorem 5.1, $\mathsf{SCAN}(t, v) \geqslant (t-2)!\mathsf{CAN}_X(2, v-t+2, t+1)$. Let $T = \{(c_1, \nu_1), (c_2, \nu_2)\}$ be a 2-way interaction. If $\nu_1 = \nu_2$, then $\mu(T) = 2$ but if $\nu_1 \neq \nu_2$, then $\mu(T) = 1$. We refer to the former kind of pair as *constant* and the latter as *non-constant*. In other words, a $\mathsf{CA}_X(N; 2, k, v)$ must cover every constant pair at least twice and every non-constant pair at least once. For each pair of columns of a $\mathsf{CA}_X(N; 2, k, v)$, there are v^2 2-way interactions involving those columns, v of which are constant pairs. Therefore,

$\mathsf{CAN}_X(2,k,v) \geqslant 2v + v^2 - v = v(v+1)$. Moreover, a $\mathsf{CA}_X(v(v+1);2,k,v)$ must cover every constant pair exactly twice and every non-constant pair exactly once. Consider the number of times each symbol appears in each column of such an array. In the rows in which symbol ν appears in column c, we must see ν exactly twice and all other $v-1$ symbols exactly once in each of the other columns. Therefore, each symbol appears $v+1$ times in each column of a $\mathsf{CA}_X(v(v+1);2,k,v)$. The following theorem establishes necessary conditions for when $\mathsf{CAN}_X(2,k,v) = v(v+1)$.

Theorem 5.2 ([9]). *For $v \geqslant 4$, if $\mathsf{CAN}_X(2,k,v) = v(v+1)$, then $k \leqslant v+2$.*

Proof. Let A be a $\mathsf{CA}_X(v(v+1);2,k,v)$ with rows indexed by $[v(v+1)]$, columns indexed by $[k]$ and symbols indexed by $[v]$. We build a set of blocks $\mathcal{B}$ on the point set $V = ([k]\times[v]) \cup \infty$. For each $r \in [v(v+1)]$, we add the block $\{(c, \mathsf{A}[r,c]) : c \in [k]\}$ to $\mathcal{B}$. Then, for $c \in [k]$, we add the block $\{(c,i) : i \in [v]\} \cup \{\infty\}$ to $\mathcal{B}$. Thus, $\mathcal{B}$ contains $v(v+1)+k$ blocks on $kv+1$ points. There are k blocks in $\mathcal{B}$ containing ∞. As each symbol in $[v]$ appears in each column of A exactly $v+1$ times, each element of $[k]\times[v]$ will appear in exactly $v+1$ of the blocks in $\mathcal{B}$ that do not contain ∞. Additionally, each of these elements will appear in exactly one block with ∞, so the total number of blocks in $\mathcal{B}$ containing a given element of $[k]\times[v]$ is $v+2$.

We build a $(kv+1)\times(v(v+1)+k)$ matrix B whose rows are indexed by the points of V and whose columns are indexed by the blocks of $\mathcal{B}$. The entry $B[r,c]$ is 1 if $r \in c$ and 0 otherwise. Let $C = BB^T$. Then C is a $(kv+1)\times(kv+1)$ matrix with rows and columns indexed by the points of V. The rank of C is at most the number of columns of B. That is, $\text{rank}(C) \leqslant v(v+1)+k$. Our goal is to find a lower bound on $\text{rank}(C)$ in terms of k and v to compare with this upper bound.

Consider the entries of C. The diagonal entries of C count the number of blocks in $\mathcal{B}$ each element appears in so $C[\infty,\infty] = k$ while $C[i,i] = v+2$ for $i \in [k]\times[v]$. For $i \in [k]\times[v]$, only one block of $\mathcal{B}$ contains both i and ∞ so $C[i,\infty] = C[\infty,i] = 1$. Let $i = (c_1,\nu_1)$ and $j = (c_2,\nu_2)$ be distinct elements of $[k]\times[v]$. If $c_1 = c_2$, then the only block containing both i and j is $\{(c_1,\nu) : \nu \in [v]\} \cup \{\infty\}$. Hence $C[i,j] = C[j,i] = 1$. If $\nu_1 = \nu_2$, then the interaction $\{i,j\}$ is a constant pair and is thus covered by two rows of A. This means i and j appear together in two blocks of $\mathcal{B}$ and thus, $C[i,j] = C[j,i] = 2$. If $c_1 \neq c_2$ and $\nu_1 \neq \nu_2$, then $\{i,j\}$ is a non-constant pair and thus $C[i,j] = C[j,i] = 1$.

Sort the elements of $[k]\times[v]$ lexicographically by the elements of $[v]$ first and then the elements of $[k]$. Without loss of generality, suppose this ordering has been applied to the rows and columns of C with the last row and column of C corresponding to ∞. That is, the first k rows and columns of C correspond to the pairs $(c,0)$, the next k rows and columns correspond to the pairs $(c,1)$, and so on. Now let $D = C - J_{kv+1,kv+1}$ where $J_{m,n}$ is the $m\times n$ matrix containing all ones. For each $0 \leqslant i \leqslant v-1$, consider the submatrix of D containing the rows and columns indexed by (c,i) for $c \in [k]$. Each of these v submatrices have diagonal entries all equal to $v+1$ and off-diagonal entries all equal to 1.

We can thus describe each of these submatrices as being equal to $vI_k + J_{k,k}$ where I_n is the $n \times n$ identity matrix. Consider the other entries of D. We have that $D[\infty, \infty] = k-1$ and all other entries in D are 0. Hence, D is block diagonal, with v blocks equal to $vI_k + J_{k,k}$ and a 1×1 block $D[\infty, \infty] = k - 1$. Therefore, $\det(D) = (k-1)\det(vI_k + J_{k,k})^v$.

Consider $\det(vI_k + J_{k,k})$. As $J_{k,k} = J_{k,1}J_{1,k}$, then by the matrix determinant lemma,

$$\begin{aligned}\det(vI_k + J_{k,k}) &= \left(1 + \frac{1}{v}J_{1,k}I_kJ_{k,1}\right)\det(vI_k)\\ &= \left(1 + \frac{k}{v}\right)v^k\\ &\geqslant 0.\end{aligned}$$

Therefore, $\det(D) \neq 0$ and thus $\operatorname{rank}(D) = kv + 1$. Recall that $D = C - J_{kv+1,kv+1}$. Therefore, $\operatorname{rank}(D) \leqslant \operatorname{rank}(C) + \operatorname{rank}(J_{kv+1,kv+1})$. As $\operatorname{rank}(J_{kv+1,kv+1}) = 1$, $\operatorname{rank}(C) \geqslant kv$. Recall that $v(v+1)+k \geqslant \operatorname{rank}(C)$. Therefore, $kv \leqslant v(v+1)+k$. Thus, $k \leqslant v(v+1)/(v-1)$ and hence, as k must be an integer and as $v \geqslant 4$, $k \leqslant v+2$. □

We remark on the similarities between $\mathsf{CA}_X(v(v+1); 2, k, v)$ and $\mathsf{OA}(2, k, v)$. First we note that if an $\mathsf{OA}(2, k, v)$ exists, then $k \leqslant v + 1$ [46]. This bound is best possible given that an $\mathsf{OA}(2, v+1, v)$ exists whenever v is a prime power [21]. Suppose in the proof of Theorem 5.2 that A was instead an $\mathsf{OA}(2, k, v)$ and suppose we built the block design $\mathcal{B}$ in an analogous way. Then, every pair of symbols from V would appear together in exactly one block. Again, we could define an incidence matrix B for $(V, \mathcal{B})$ and the Gram matrix $C = BB^T$. We could then obtain the same inequality $\operatorname{rank}(C) \geqslant kv$ by subtracting 1 from every entry in C and observing the resulting matrix is diagonal. However, an $\mathsf{OA}(2, k, v)$ has v fewer rows than a $\mathsf{CA}_X(v(v+1); 2, k, v)$, so now B has fewer columns. This in turn gives us a tighter upper bound for $\operatorname{rank}(C)$ which implies $k \leqslant v+1$ for an $\mathsf{OA}(2, k, v)$. So, as compared to an $\mathsf{OA}(2, k, v)$, the increased number of rows of a $\mathsf{CA}_X(v(v+1); 2, k, v)$ leads to the possibility that such an array may have $v+2$ columns.

We now return to Conjecture 3.18 and the question of when $\mathsf{SCAN}(t, v) = t!$. If $\mathsf{SCAN}(t, v) = t!$, then by Theorem 5.1, $t! \geqslant (t-2)!\mathsf{CAN}_X(2, v-t+2, t-1)$. We know that $\mathsf{CAN}_X(2, k, v) \geqslant v(v+1)$. Therefore, by Theorem 5.1, if $\mathsf{SCAN}(t, v) = t!$, then $\mathsf{CAN}_X(2, v-t+1, t-1) = t(t-1)$. The following theorem takes the conditions from Theorem 5.2 and derives conditions that are necessary for $\mathsf{SCAN}(t, v) = t!$.

Theorem 5.3 ([9]). *If* $\mathsf{SCAN}(t, v) = t!$, *then* $v \leqslant 2t - 1$.

Proof. By Theorem 5.1, if $\mathsf{SCAN}(t, v) = t!$, then $\mathsf{CAN}_X(2, v-t+2, t-1) = t(t-1)$. By Theorem 5.2, if $\mathsf{CAN}_X(2, v-t+2, t-1) = t(t-1)$, then $v - t + 2 \leqslant t + 1$. Therefore, if $\mathsf{SCAN}(t, v) = t!$, then $v \leqslant 2t - 1$. □

Chapter 6

Related Problems

In the previous chapters, we have been concerned with universal sequence coverage. That is, sets of permutations that cover every sequence of a given length either at least once or some fixed number of times. In these discussions we have already encountered several related objects. For example, in Chapter 3, we discussed 3-mixing sets of permutations [18, 23] and the close relationship they have to strength 3 sequence covering arrays. Recall that a 3-mixing set of permutations is a set of permutations in $\mathcal{S}_v$ such that for distinct $i, j, k \in [v]$, there is a permutation that places i between j and k. In Chapter 4, we discussed sets of permutations with t-wise independence, t-rankwise independence and t-restricted min-wise independence. These different definitions of independence each described different sequence coverage properties with t-rankwise independence being equivalent to perfect sequence coverage. We also discussed t-suitable sets of permutations introduced by Dushnik [14], the connections these objects have to t-restricted min-wise independent sets of permutations, and the ways t-suitable sets of permutations eventually motivated the study of sequence covering arrays. Recall that $X \subseteq \mathcal{S}_v$ is a t-suitable set of permutations if for any t-subset $T \subseteq [v]$ and for any $i \in T$, there is a permutation $\pi \in X$ such that $\pi(i) < \pi(j)$ for all other $j \in T$. Both 3-mixing sets of permutations and t-suitable sets of permutations describe sets that cover a subset of sequences of a given length. In this section, we describe a number of problems that similarly generalise the concept of sequence coverage to other contexts.

6.1 Deletion correcting codes

In this section, we describe the significance of sequence coverage in the context of coding theory. Here, a *word* is a string of symbols from a given alphabet where repetition of symbols is allowed. Let B_v^n be the set of words of length n over the alphabet $[v]$. A *subword* of a word $x \in B_v^n$ is a word $y \in B_v^{n-s}$ that can be obtained by deleting s symbols from x for some $0 < s \leqslant n$. A code $C \subset B_v^n$ is capable of *correcting s deletions* if every word in B_v^{n-s} is a subword of at most one word in C. If every word in B_v^{n-s} is a subword of exactly one word in C, then C is *perfect*. We can define similar codes over the set $\mathcal{S}_{v,t}$.

When s symbols are deleted from a sequence in $\mathcal{S}_{v,t}$, we obtain a sequence in $\mathcal{S}_{v,t-s}$. So, an s-deletion correcting code in $\mathcal{S}_{v,t}$ is a set $C \subset \mathcal{S}_{v,t}$ such that every sequence in $\mathcal{S}_{v,t-s}$ is covered by at most one element of C. Again, if every sequence in $\mathcal{S}_{v,t-s}$ is covered by exactly one element of C, then C is perfect. A perfect s-deletion correcting code in $\mathcal{S}_v$ is thus equivalent to a $\mathsf{PSCA}(v, v-s, 1)$. This equivalence provides further context for Conjecture 3.18 as a statement on the existence of perfect deletion correcting codes. Indeed, Levenshtein's work on sequence covering arrays [37] is phrased as an investigation into perfect s-deletion correcting codes in $\mathcal{S}_v$.

To draw an equivalence between perfect deletion correcting codes and perfect sequence covering arrays, we must enforce that every codeword contains every symbol in $[v]$ exactly once. If we relax this restriction slightly to allow for codewords of a given length $k < v$ while still requiring that symbols in codewords do not repeat, then a perfect s-deletion correcting code is a $(k-s)$-$(v,k,1)$ directed design.

An *s-burst deletion* from a permutation $\pi \in \mathcal{S}_v$ is the deletion of the symbols $\pi(i), \pi(i+1), \ldots, \pi(i+s-1)$ from π. A code $C \subset \mathcal{S}_v$ is capable of correcting s-burst deletions if every sequence in $\mathcal{S}_{v,v-s}$ can be obtained by an s-burst deletion from at most one permutation in C. See e.g. [52] for more on these types of permutation codes.

6.2 Separation dimension of hypergraphs

Let G be a hypergraph. A *separating family* of G is a set X of permutations of $V(G)$ such that for any disjoint $e, f \in E(G)$, there exists $\sigma \in X$ such that either $\sigma^{-1}(i) < \sigma^{-1}(j)$ for all $i \in e$, $j \in f$, or $\sigma^{-1}(i) > \sigma^{-1}(j)$ for all $i \in e$, $j \in f$. That is, every vertex of one edge appears before every vertex of the other in σ. We say that σ *separates* e and f. The *separation dimension* of G, denoted $\pi(G)$, is the smallest size of a separating family of G.

Separation dimension was introduced by Basavaraju et al. [4] who draw a connection between separation dimension and another graph property called boxicity. A *k-box* is a Cartesian product of k closed real intervals. A *k-box representation* of a graph G is a function f that maps $V(G)$ to k-boxes such that $f(u)$ intersects with $f(v)$ if and only if $\{u,v\} \in E(G)$. The *boxicity* of G is the minimum k for which there exists a k-box representation of G. The *line graph* of a hypergraph G is the graph with vertex set $E(G)$ such that two edges of G are adjacent in the line graph if they intersect. Basavaraju et al. proved that the separation dimension of a hypergraph G is equal to the boxicity of the line graph of G.

Let K_v^r denote the complete r-graph on v vertices. That is K_v^r is the hypergraph with vertex set $[v]$ and whose edge set is the set of all r-subsets of $[v]$. As an $\mathsf{SCA}(N; 2r, v)$ covers all sequences of length $2r$, it must necessarily be a separating family for K_v^r. Given two disjoint edges e, f of K_v^r, there are $2(r!)^2$ permutations of $e \cup f$ which separate e and f as there are $r!$ ways each of arranging the vertices of e and f and we can place the vertices of e before or after the vertices of f. A separating family for K_v^r need only cover one of these permutations. Basavaraju et al. gave the following upper and lower bounds for $\pi(K_v^r)$.

Theorem 6.1 ([4]). *For v sufficiently larger than $r > 2$,*

$$\frac{4^r}{2^7\sqrt{r-2}}\log v \leqslant \pi(K_v^r) \leqslant \frac{e\ln 24^r\sqrt{r}}{\pi\sqrt{2}}\log v.$$

Despite the fact that a separating family for K_v^r is required to cover far fewer sequences than an $\mathsf{SCA}(N; 2r, v)$, we still see the same logarithmic growth in v. The main difference between $\pi(K_v^r)$ and $\mathsf{SCAN}(2r, v)$ appears to be the dependence on r, which is exponential in $\pi(K_v^r)$ and factorial in $\mathsf{SCAN}(2r, v)$. A separating family for G is also a separating family for any subgraph of G, so if G is some r-graph with v vertices, then $\pi(G) \leqslant \pi(K_v^r)$.

Yuster [65] introduced *perfect separating dimension*, analogous to perfect sequence covering arrays. A separating family X for G is perfect if each pair of disjoint edges is separated by exactly λ permutations in X for some positive integer λ. Yuster denotes the size of the smallest perfect separating family for a hypergraph G by $\mathrm{PSD}(G)$. Similarly to before, a $\mathsf{PSCA}(v, 2r, \lambda)$ will also be a perfect separating family for K_v^r. Indeed, we might expect to see some similarities in the asymptotic behaviours of $\mathsf{PSCAN}(v, 4)$ and $\mathrm{PSD}(K_v)$ where $K_v = K_v^2$ is the complete graph on v vertices. However, Yuster [65] adapts a construction in [64] for strength 3 perfect sequence covering arrays to build perfect separating families for K_v. The resulting asymptotics for $\mathrm{PSD}(K_v)$ are thus much more similar to asymptotics for $\mathsf{PSCAN}(v, 3)$ given in [64].

6.3 Covering subgraphs of tournaments

In this section, we describe a problem posed by Yuster [63] on covering subgraphs of tournaments. This problem relates to the problems of finding sequence covering arrays and t-suitable sets of permutations by reinterpreting permutations as transitive tournaments and sequences as directed acyclic graphs. The *complete directed graph* on v vertices, D_v, is the graph with vertex set $[v]$ and edges (i, j) for all distinct $i, j \in [v]$. A *tournament* with v vertices, T, is a directed graph with vertex set $[v]$ such that for distinct $i, j \in [v]$, either $(i, j) \in E(T)$ or $(j, i) \in E(T)$ but not both. A tournament T is *transitive* if for distinct $i, j, k \in [v]$, $(i, j) \in E(T)$ and $(j, k) \in E(T)$ implies $(i, k) \in E(T)$. Given a permutation $\pi \in \mathcal{S}_v$, we can build a corresponding transitive tournament T on v vertices with $(i, j) \in E(T)$ if and only if $\pi^{-1}(i) < \pi^{-1}(j)$. Similarly, one can construct a permutation in $\mathcal{S}_v$ from a transitive tournament with v vertices.

For a tournament T with v vertices and a permutation $\pi \in \mathcal{S}_v$, let $L_\pi(T)$ be the subgraph of T containing exactly the edges $(i, j) \in E(T)$ that satisfy $\pi^{-1}(i) < \pi^{-1}(j)$. A subgraph of T is *covered* by π if it is also a subgraph of $L_\pi(T)$. Let H be some directed acyclic graph. Yuster [63] defines $\tau_H(v)$ to be the smallest integer N such that for any tournament T with v vertices, there is a set of N permutations $X = \{\pi_0, \ldots, \pi_{N-1}\} \in$

$\mathcal{S}_v$ such that each subgraph of T isomorphic to H is covered by at least one of the permutations in X.

Let T_t be the transitive tournament with t vertices. Then an $\mathsf{SCA}(N;t,v)$ can be seen as a set of permutations $X = \{\pi_0, \ldots, \pi_{N-1}\} \in \mathcal{S}_v$ such that each subgraph of D_v isomorphic to T_t is covered by at least one of the permutations in X. Every tournament with v vertices is necessarily a subgraph of D_v. Moreover, if H is a directed acyclic graph with t vertices, then it is a subgraph of some subgraph of D_v isomorphic to T_t. Therefore, if H is a directed acyclic graph with t vertices, then $\tau_H(v) \leqslant \mathsf{SCAN}(t,v)$. However, a tournament with v vertices will possess fewer subgraphs isomorphic to T_t than D_v, so it is feasible that even in the case $H = T_t$, $\mathsf{SCAN}(t,v)$ does not provide tight upper bounds for $\tau_H(v)$.

We can similarly translate t-suitable sets of permutations into the language of directed acyclic graphs. The *out-star with t vertices*, denoted S_t^+ is the directed graph with vertex set $[t]$ and edge set $\{(0,j) : 1 \leqslant j \leqslant t-1\}$. Then a t-suitable set of permutations is equivalent to a set of permutations that covers every subgraph of D_v isomorphic to S_t^+.

6.4 Sequence covering arrays with constraints

In real world systems, certain sequences of events may be impossible while other sequences may be forced. For example, in a file API system in which the four possible commands are *open, read, write* and *close*, the command *open* must proceed all other commands. For these constrained systems, testing using a sequence covering array may be infeasible, as permutations in a sequence covering array may cover forbidden sequences. We note that this problem mirrors that of finding covering arrays that avoid certain forbidden interactions [12].

Kuhn et al. [34] consider systems with v events in which a single sequence of length 2, say (x,y), is forbidden. A testing suite which tests all feasible sequences of events of length t can be developed by constructing a set X of permutations such that no permutation in X covers (x,y) and each sequence in $\mathcal{S}_{v,t}$ that does not cover (x,y) is covered by some permutation in X. The restriction that no permutation in X can cover (x,y) means that the number of permutations that are feasible for testing has been halved. However, since in general many sequences in $\mathcal{S}_{v,t}$ will not contain both x and y, the relative decrease in the number of sequences that need to be tested will be much smaller. Recall the greedy algorithm developed by Kuhn et al. described in Chapter 3.3. Here, permutations covering (x,y) may be screened out when generating candidate permutations. However, it is no longer possible to add both a permutation and its reverse since one of these permutations is guaranteed to cover (x,y). With this potential loss of efficiency, Kuhn et al. find that despite the reduction in the number of sequences that need to be tested, the number of permutations their algorithm generates to test a constrained system is similar to the unconstrained case.

A more concrete and general definition of sequence covering arrays subject to constraints is given by Chee et al. [9]. Let $\mathcal{C}$ be a set of sequences of elements in $[v]$. Then a *constrained* sequence covering array, denoted by $\mathsf{SCA}(N; t, v, \mathcal{C})$, is a set X of N permutations in $\mathcal{S}_v$ such that no permutation in X covers a sequence in $\mathcal{C}$ and if $s \in \mathcal{S}_{v,t}$ such that s does not cover a sequence in $\mathcal{C}$, there is a permutation in X that covers s. The elements of $\mathcal{C}$ are called *constraints*. When $t = 2$, Chee et al. draw a connection between constrained sequence covering arrays and partial orders. Recall the relevant definitions for partial orders from Chapter 2. Let P be a partial order on $[v]$ and let $\mathcal{C} = \{(a, b) : a \neq b \text{ and } (b, a) \in P\}$. Let X be an $\mathsf{SCA}(N; 2, v, \mathcal{C})$. As no permutation in X can cover any sequence in $\mathcal{C}$, each permutation in X must cover every pair in P that contains two distinct elements. That is, every permutation in X is a linear extension of P. If $(a, b) \in \mathcal{S}_{v,t}$ such that $(a, b), (b, a) \notin P$, then there must be a permutation in X that covers (a, b) and a permutation in X that covers (b, a). Therefore, the sequences of length 2 covered by every permutation in X are exactly those in P. Hence, X realises P and so an $\mathsf{SCA}(N; t, v, \mathcal{C})$ exists if and only if the dimension of P is at most N. The problem of determining whether the dimension of P is at most N is NP-complete [61].

In the above case, the complexity arises in finding an $\mathsf{SCA}(N; 2, v, \mathcal{C})$ that minimises N but not necessarily in showing that there is some N for which an $\mathsf{SCA}(N; 2, v, \mathcal{C})$ exists. However, for some types of constraints, even determining whether a constrained sequence covering array exists at all may be a difficult problem. Let $T \subset \mathcal{S}_{v,3}$ and let

$$\mathcal{C} = \bigcup_{(a,b,c) \in T} \{(b, a, c), (b, c, a), (a, c, b), (c, a, b)\}.$$

Then, in each permutation in an $\mathsf{SCA}(t, v, \mathcal{C})$ and for each $(a, b, c) \in T$, b must appear between a and c. Determining whether even a single permutation exists that avoids covering each of the sequences in $\mathcal{C}$ is NP-complete [45].

Duan et al. [13] describe a greedy algorithm that is capable of generating test suites for systems subject to particular constraints. These constraints include forbidden sequences as above, but can also include conditions such as an event x needing to appear immediately before or after an event y. To ensure this algorithm can generate test suites for any set of constraints, the tests generated need not be permutations. In particular, events may be repeated and the length of each test need not be uniform. The number of times an event appears in a test and the length of each test may also be subject to constraints.

Chapter 7

Computations of the Perfect Sequence Covering Array Number

In this chapter we describe exhaustive computations that uncover several new values of $\mathsf{PSCAN}(v, t)$. These computations are guided by restrictions on so-called *distributions*, which we introduce to describe the frequency of a given symbol in each column of a PSCA. We discuss these distributions and prove several results in Chapter 7.2. These results motivate our exhaustive search algorithm for perfect sequence covering arrays, which we describe in Chapter 7.3. Here, we present perfect sequence covering arrays that determine new values for $\mathsf{PSCAN}(v, t)$ for certain parameter choices as well as catalogue the number of such arrays.

In Chapter 7.4, we perform searches for PSCAs over permutation subgroups of $\mathcal{S}_v$ for $v \leqslant 14$. These searches are motivated by a permutation subgroup of $\mathcal{S}_6$ isomorphic to $\mathcal{S}_4$ which Mathon and Tran van Trung [41] found to be a $\mathsf{PSCA}(6, 4, 1)$. We find a variety of permutation groups that also form PSCAs. We pay particular attention to the family of elementary abelian 2-groups in an attempt to design a general construction of strength 3 PSCAs. Although we are unable to design such a construction, we are able to use these groups to construct strength 3 PSCAs for $8 \leqslant v \leqslant 32$ that improve upon current best bounds for $\mathsf{PSCAN}(v, 3)$.

7.1 Introduction

Recall that for positive integers v and t with $v \geqslant t$, we let $[v] = \{0, \ldots, v-1\}$, $\mathcal{S}_v$ be the set of permutations of $[v]$ and $\mathcal{S}_{v,t}$ be the set of ordered sequences of t distinct elements of $[v]$. For $\pi \in \mathcal{S}_v$ and $s = (s_0, \ldots, s_{t-1}) \in \mathcal{S}_{v,t}$ we say that s is *covered* by π if $\pi^{-1}(s_i) < \pi^{-1}(s_{i+1})$ for $0 \leqslant i \leqslant t-2$. A *perfect sequence covering array* with order v, strength t and multiplicity λ, denoted by $\mathsf{PSCA}(v, t, \lambda)$, is a multiset X of permutations in $\mathcal{S}_v$ such that every sequence in $\mathcal{S}_{v,t}$ is covered by exactly λ permutations in X. If we let T be a t-subset of $[v]$, then there are $t!$ orderings of the symbols of T, each of which must be covered by λ permutations in a $\mathsf{PSCA}(v, t, \lambda)$. Furthermore, every permutation in a $\mathsf{PSCA}(v,t,\lambda)$ covers exactly one ordering of T, so a $\mathsf{PSCA}(v,t,\lambda)$ must consist of $t!\lambda$ permutations.

For $v \geqslant t$, $\mathsf{PSCAN}(v,t)$ is defined to be the smallest positive integer λ such that a $\mathsf{PSCA}(v,t,\lambda)$ exists. Observe that $\mathcal{S}_v$ is a $\mathsf{PSCA}(v,t,v!/t!)$ so $\mathsf{PSCAN}(v,t)$ is well defined and $\mathsf{PSCAN}(v,t) \leqslant v!/t!$. Note that if $v > t$ and we remove the symbol $v-1$ from every permutation of a $\mathsf{PSCA}(v,t,\lambda)$, then we obtain a $\mathsf{PSCA}(v-1,t,\lambda)$ and hence $\mathsf{PSCAN}(v,t) \geqslant \mathsf{PSCAN}(v-1,t)$. For $2 \leqslant u \leqslant t$, a $\mathsf{PSCA}(v,t,\lambda)$ is also a $\mathsf{PSCA}(v,u,\lambda t!/u!)$ so $\mathsf{PSCAN}(v,u) \leqslant (t!/u!)\mathsf{PSCAN}(v,t)$.

The question of when $\mathsf{PSCAN}(v,t) = 1$ has received particular attention. Note that $\mathcal{S}_v$ forms a $\mathsf{PSCA}(v,v,1)$. At the other end of the spectrum, if $t = 2$, then we can take any permutation in $\mathcal{S}_v$ and its reverse to form a $\mathsf{PSCA}(v,2,1)$. Therefore, $\mathsf{PSCAN}(v,v) = \mathsf{PSCAN}(v,2) = 1$. Levenshtein [37] proved that $\mathsf{PSCAN}(t+1,t) = 1$ for $t \geqslant 3$. Mathon and Tran van Trung [41] proved that a $\mathsf{PSCA}(5,3,1)$ does not exist (we provide a new proof of this fact in Chapter 7.2). As demonstrated above, $\mathsf{PSCAN}(v,t) \geqslant \mathsf{PSCAN}(v-1,t)$ so it follows that a $\mathsf{PSCA}(v,3,1)$ does not exist for $v \geqslant 5$. Therefore, when $t = 3$, we have $\mathsf{PSCAN}(v,t) > 1$ for $v > t+1$. It was initially conjectured by Levenshtein [36] that this property would hold for any $t \geqslant 3$, however that was later shown to be false for $t = 4$ by Mathon and Tran van Trung [41], who presented a $\mathsf{PSCA}(6,4,1)$. On the other hand, Mathon and Tran van Trung computationally proved that neither a $\mathsf{PSCA}(7,5,1)$ nor a $\mathsf{PSCA}(8,6,1)$ exists, thus confirming Levenshtein's conjecture for $t \in \{5,6\}$. They also found that a $\mathsf{PSCA}(7,4,1)$ does not exist. A combinatorial proof of this last fact was later given by Klein [31]. Chee et al. [9] proved that $\mathsf{PSCAN}(2t,t) > 1$ for $t \geqslant 3$.

Yuster [64] proved that $\mathsf{PSCAN}(5,3) = 2$. In Chapter 7.3, we show that $\mathsf{PSCAN}(6,3) = \mathsf{PSCAN}(7,3) = \mathsf{PSCAN}(7,4) = 2$, $\mathsf{PSCAN}(8,3) = 3$ and $\mathsf{PSCAN}(8,4) \geqslant 3$.

For a multiset X of permutations of $[v]$, a symbol $w \in [v]$, and for $0 \leqslant i \leqslant v-1$, we define

$$d_w(i) := \big|\{\pi \in X : \pi(i) = w\}\big|.$$

We refer to the vector $\boldsymbol{d}_w = \big(d_w(0), \ldots, d_w(v-1)\big)$ as the *distribution vector* of w.

In Chapter 7.2, we derive restrictions on distribution vectors of symbols in PSCAs. These restrictions facilitate the computer searches that we use to exhaustively catalogue $\mathsf{PSCA}(v,t,\lambda)$ for different sets of parameters. These searches and their results are described in Chapter 7.3.

In Chapter 7.4, we explore the relationship between groups and PSCAs and use this relationship to construct PSCAs with strengths 3 and 4, thereby improving upon the best known upper bounds for $\mathsf{PSCAN}(v,3)$ for $9 \leqslant v \leqslant 32$ and providing non-trivial upper bounds for $\mathsf{PSCAN}(v,t)$ for $v \leqslant 24$ and $4 \leqslant t \leqslant 6$. Table 7.1 summarises the improvements to upper bounds on $\mathsf{PSCAN}(v,3)$ while Table 7.2 summarises the new results for $4 \leqslant t \leqslant 6$. The tables also incorporate the exact bounds shown in Chapter 7.3.

v	New bound	Old bound
6–7	2*	8
8	3*	8
9	6	8
10–12	6	160
13–14	7	160
15–16	16	160
17–19	19	160
20–32	96	160

Table 7.1: Improvements to known upper bounds on $\mathsf{PSCAN}(v, 3)$. An asterisk denotes that the new bound is exact.

v	t	New bound
7	4	2*
8–12	4	18
13	4	234
14–21	4	5040
22	4	18 480
23	4	425 040
24	4	10 200 960
7–11	5	66
12	5	792
13–22	5	3696
23	5	85 008
24	5	2 040 192
8–12	6	132
13–24	6	340 032

Table 7.2: New upper bounds for $\mathsf{PSCAN}(v, t)$ for $4 \leqslant t \leqslant 6$. An asterisk denotes that the new bound is exact.

Independently, and using different methods, Na, Jedwab and Li [44] have also considered the problem of determining $\mathsf{PSCAN}(v, t)$. They find that $2 = \mathsf{PSCAN}(6, 3) = \mathsf{PSCAN}(7, 3) = \mathsf{PSCAN}(7, 4)$, while also demonstrating that $3 \geqslant \mathsf{PSCAN}(8, 3)$, $4 \geqslant \mathsf{PSCAN}(9, 3)$ and $4 \geqslant \mathsf{PSCAN}(7, 5)$. They also show that a $\mathsf{PSCA}(v, t, \lambda)$ exists for $(v, t) \in \{(5, 3), (6, 3), (7, 3), (7, 4)\}$ if and only if $\lambda \geqslant 2$ while a $\mathsf{PSCA}(8, 3, \lambda)$ exists for any $\lambda \geqslant 3$. Several of these results were originally reported in Na's Masters thesis [43]; in particular, he reported that $\mathsf{PSCAN}(7, 4) = 2$ before we computed our catalogue of $\mathsf{PSCA}(7, 4, 2)$. While many of our results overlap with those of Na, Jedwab and Li [44], our methods provide unique insights. Our exhaustive searches rule out the existence of a $\mathsf{PSCA}(8, 3, 2)$. Hence, the existence of a $\mathsf{PSCA}(8, 3, 3)$ proves that $\mathsf{PSCAN}(8, 3) = 3$. Our methods are also capable of counting the number of $\mathsf{PSCA}(v, t, \lambda)$ for different parameter sets.

7.2 Distribution vectors

For a multiset X of permutations of $[v]$, a symbol $w \in [v]$, and for $0 \leqslant i \leqslant v-1$, we define

$$d_w(i) := \big|\{\pi \in X : \pi(i) = w\}\big|.$$

We refer to the vector $\boldsymbol{d}_w = \big(d_w(0), \ldots, d_w(v-1)\big)$ as the *distribution vector* of w. The distribution vector of w records the number of times a symbol w appears in each column across a multiset of permutations. In this section, we will derive several restrictions on distribution vectors for symbols in PSCAs. We begin with the following lemma which limits the number of occurrences of each symbol in a PSCA across sets of consecutive columns.

Lemma 7.1. *Let X be a* $\mathsf{PSCA}(v,t,\lambda)$ *with $v \geqslant t \geqslant 2$ and $\lambda \geqslant 1$. Then, for $w \in [v]$ and for $0 \leqslant i \leqslant t-1$,*

$$\frac{\lambda(v-1)!}{(v-t)!} = \sum_{j=0}^{v-1} d_w(j)\binom{j}{i}\binom{v-1-j}{t-1-i}. \tag{7.1}$$

Proof. Let X be a $\mathsf{PSCA}(v,t,\lambda)$, let $i \in \{0,\ldots,t-1\}$ and let $w \in [v]$. Let $S = \{s \in \mathcal{S}_{v,t} : s_i = w\}$. Note that $|S| = (v-1)!/(v-t)!$ and each sequence in S is covered by λ permutations in X. For some $j \in [v]$, let $\pi \in X$ be one of the $d_w(j)$ permutations in X such that $\pi(j) = w$ and consider how many sequences in S are covered by π. There are j symbols that appear before w and $v-1-j$ symbols that appear after w in π. For every sequence in S, there are i symbols appearing before w and $t-1-i$ symbols appearing after w. Hence, π covers $\binom{j}{i}\binom{v-1-j}{t-1-i}$ sequences in S. The lemma follows. □

We can use Lemma 7.1 to prove the following theorem.

Theorem 7.2. *Let X be a* $\mathsf{PSCA}(v,t,\lambda)$ *with $v \geqslant t \geqslant 2$ and $\lambda \geqslant 1$. Then for $w \in [v]$ and for $1 \leqslant s \leqslant t-1$,*

$$\frac{1}{t!\lambda}\sum_{j=0}^{v-1} j^s d_w(j) = \frac{1}{v}\sum_{i=0}^{v-1} i^s. \tag{7.2}$$

Proof. Fix $s \in \{1,\ldots,t-1\}$ and $w \in [v]$. Let $\alpha_k(w) = \sum_{j=0}^{v-1} j^k d_w(j)$ for $k \in \{1,\ldots,s\}$. Now, X is a $\mathsf{PSCA}(v,k+1,\lambda t!/(k+1)!)$ and so, using Lemma 7.1 with $i = k$, we find that $\alpha_k(w)$ is a function of λ, v, t, k and $\alpha_1(w),\ldots,\alpha_{k-1}(w)$. So, proceeding by induction on k, we have that $\alpha_k(w)$ is independent of w for each $k \in \{1,\ldots,s\}$. Thus,

$$v\alpha_s(w) = \sum_{w\in[v]} \alpha_s(w) = \sum_{i=0}^{v-1} i^s \sum_{w\in[v]} d_w(j) = t!\lambda\sum_{i=0}^{v-1} i^s$$

and hence (7.2) holds. □

We call a distribution vector that satisfies (7.2) for parameters (v, t, λ) a (v, t, λ)-*feasible distribution*. We now prove some more facts about distribution vectors when certain restrictions on v and t are imposed. The following theorem demonstrates more stringent restrictions on the distribution vector whenever t is an odd prime. Intuitively, it states that in a PSCA whose strength is an odd prime p and whose order is not divisible by p, the number of occurrences of a symbol across all columns of a given equivalence class modulo p is itself divisible by p.

Theorem 7.3. *Let X be a* PSCA(v, p, λ) *with p an odd prime and with $v \not\equiv 0 \pmod{p}$. For $w \in [v]$ and $0 \leqslant j \leqslant p-1$, let $y_w(j) = \sum_{i \equiv j \pmod{p}} d_w(i)$. Then, $y_w(j) \equiv 0 \pmod{p}$.*

Proof. Let $w \in [v]$. By Theorem 7.2, if X is a PSCA(v, p, λ), then for $1 \leqslant i \leqslant p-1$,

$$\sum_{j=0}^{v-1} j^i d_w(j) = \frac{p!\lambda}{v} \sum_{j=0}^{v-1} j^i.$$

As v is not divisible by p, the right hand side of the equation above must be divisible by p. Therefore, for $1 \leqslant i \leqslant p-1$,

$$0 \equiv \sum_{j=0}^{v-1} j^i d_w(j) \equiv \sum_{j=1}^{p-1} j^i y_w(j) \pmod{p}.$$

This gives a system of $p-1$ linear equations in $p-1$ variables over the field $GF(p)$. We can restate this system as

$$A \begin{pmatrix} y_w(1) \\ y_w(2) \\ \vdots \\ y_w(p-1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where A is a $(p-1) \times (p-1)$ matrix over $\mathbb{F}_p$ with $A_{i,j} = j^i$. Therefore, A is a Vandermonde matrix and thus, A is non-singular. Hence, the only solution to this system is $y_w(j) \equiv 0 \pmod{p}$ for all $j \in \{1, \ldots, p-1\}$. As the number of permutations in X is $p!\lambda \equiv 0 \pmod{p}$, it also follows that $y_w(0) \equiv 0 \pmod{p}$. □

When $v = t + 1$ and t is even, the following lemma proves that all (v, t, λ)-feasible distribution vectors are palindromic.

Lemma 7.4. *Let t be even and let X be a* PSCA$(t+1, t, \lambda)$. *For $w \in [t+1]$ and $0 \leqslant i \leqslant t$, $d_w(t/2 - i) = d_w(t/2 + i)$.*

Proof. Note that by Theorem 3.17, a PSCA$(t+1, t, 1)$ exists. By taking λ copies of a PSCA$(t+1, t, 1)$, we can build a PSCA$(t+1, t, \lambda)$ so a PSCA$(t+1, t, \lambda)$ exists for all $\lambda \geqslant 1$.

With $v = t + 1$, Lemma 7.1 implies that

$$\lambda t! = (t-i)d_w(i) + (i+1)d_w(i+1)$$

for $0 \leqslant i \leqslant t-1$. Therefore, for even t,

$$d_w\left(\frac{t}{2}-i\right) = \frac{\lambda t! - \left(\frac{t}{2}-i+1\right)d_w\left(\frac{t}{2}-i+1\right)}{\frac{t}{2}+i}$$
$$d_w\left(\frac{t}{2}+i\right) = \frac{\lambda t! - \left(\frac{t}{2}-i+1\right)d_w\left(\frac{t}{2}+i-1\right)}{\frac{t}{2}+i}$$

for $1 \leqslant i \leqslant t/2$. Then induction on i shows that $d_w(t/2-i) = d_w(t/2+i)$ for $0 \leqslant i \leqslant t/2$. □

We continue with the case where $v = t+1$. The only possible $\mathsf{PSCA}(t,t,\lambda)$ is a multiset containing λ copies of $\mathcal{S}_v$. Therefore, this is exactly the PSCA we obtain by deleting any symbol from a $\mathsf{PSCA}(t+1,t,\lambda)$. We use this fact to derive further restrictions for a $\mathsf{PSCA}(t+1,t,\lambda)$. Let X be a $\mathsf{PSCA}(v,t,\lambda)$. For $w \in [v]$ and $I \subseteq [v] \setminus \{w\}$ with $|I| = i$, let $d_{I,w}$ be the number of permutations $\pi \in X$ such that $\pi(i) = w$ and $I = \{\pi(j) : 0 \leqslant j \leqslant i-1\}$.

Theorem 7.5. *Let X be a $\mathsf{PSCA}(t+1,t,\lambda)$, let $w \in [v]$, and let $0 \leqslant i \leqslant t$. Then for any i-subset $I \subseteq [v] \setminus \{w\}$,*

$$d_{I,w} = \frac{d_w(i)}{\binom{t}{i}}.$$

Proof. Let $w \in [v]$. We proceed by induction on i. Note that the statement is trivially true for $i = 0$. Suppose the statement is true for some i with $0 \leqslant i \leqslant t-1$ and consider the statement for $i+1$. Let $I = \{u_1, \ldots, u_{i+1}\} \subseteq [v] \setminus \{w\}$ and let $J = I \setminus \{u_{i+1}\}$. The array formed by removing u_{i+1} from each permutation of X is λ copies of $\mathcal{S}_t$. The number of permutations $\tau \in \mathcal{S}_t$ for which $\tau(i) = w$ and $\{\tau(j) : 0 \leqslant j \leqslant i-1\} = J$ is $i!(t-1-i)!$. Therefore, the number of permutations $\pi \in X$ such that either $\pi(i) = w$ and $\{\pi(j) : 0 \leqslant j \leqslant i-1\} = J$ or $\pi(i+1) = w$ and $\{\pi(j) : 0 \leqslant j \leqslant i\} = I$ is $\lambda i!(t-1-i)!$. Thus,

$$\lambda i!(t-1-i)! = d_{J,w} + d_{I,w}.$$

By the inductive hypothesis, $d_{J,w} = d_w(i)/\binom{t}{i}$. Therefore, for any two $(i+1)$-subsets of $[v] \setminus \{w\}$, I and I', $d_w(I) = d_w(I')$. The sum of $d_{I,w}$ as I ranges over all $\binom{t}{i+1}$ possible $(i+1)$-subsets of $[v] \setminus \{w\}$ must be $d_w(i+1)$. Therefore $d_{I,w} = d_w(i+1)/\binom{t}{i+1}$, completing the induction. □

Corollary 7.6. *Let X be a $\mathsf{PSCA}(t+1,t,\lambda)$, and let $0 \leqslant i \leqslant t$. Then $d_w(i)$ is divisible by $\binom{t}{i}$ for all $w \in [v]$.*

In general, if it could be shown that there are no (v,t,λ)-feasible distributions for some choice of v, t and λ, then it would imply that a $\mathsf{PSCA}(v,t,\lambda)$ does not exist. However, it is possible to find (v,t,λ)-feasible distributions for infinitely many choices of v, t, and λ. For example, if $t!\lambda$ is divisible by v, then a distribution vector with $d_w(i) = t!\lambda/v$ for $0 \leqslant i \leqslant v-1$ is (v,t,λ)-feasible.

On the other hand, it is possible to use (v,t,λ)-feasible distributions to disprove the existence of a $\mathsf{PSCA}(v,t,\lambda)$ even when such distributions do exist. For example, consider the $(5,3,1)$-feasible distributions. By Theorem 7.3, for such a distribution, $d_w(2) \in \{0,3,6\}$. If $d_w(2) = 6$, then $d_w = (0,0,6,0,0)$ which violates (7.2) for $s = 2$. Now suppose $d_w(2) = 3$. Again, by Theorem 7.3, $\{d_w(0)+d_w(3), d_w(1)+d_w(4)\} = \{0,3\}$. As the reverse of a PSCA is also a PSCA, we can without loss of generality suppose $d_w(0) + d_w(3) = 0$. Then, for $s = 2$, (7.2) reduces to $d_w(1) + 16d_w(4) = 24$. As $d_w(1)$ and $d_w(4)$ must be nonnegative integers that sum to 3, we find that this equation has no solutions. Therefore, in any $(5,3,1)$-feasible distribution, $d_w(2) = 0$. This means that if a $\mathsf{PSCA}(5,3,1)$ exists, then it would be impossible to place any symbol in column 2. This contradiction provides an alternative proof of the non-existence of a $\mathsf{PSCA}(5,3,1)$. See [41] for an earlier proof.

In the proof of Theorem 7.5, we were able to enforce restrictions on a $\mathsf{PSCA}(t+1,t,\lambda)$ by considering the new array formed by deleting a symbol from this PSCA. We consider this kind of symbol deletion in a more general setting with the following theorem.

Theorem 7.7. *Let X be a $\mathsf{PSCA}(v,t,\lambda)$ and let $\boldsymbol{d}_w = \big(d_w(0),\ldots,d_w(v-1)\big)$ be the distribution vector for $w \in [v]$. Let $\boldsymbol{d}'_w = \big(d'_w(0),\ldots,d'_w(v-2)\big)$ be the distribution vector of w in the PSCA X' obtained by deleting a symbol $w' \neq w$ from X. Then*

$$\delta_k = \sum_{i=0}^{k} \big(d'_w(i) - d_w(i)\big)$$

satisfies $0 \leqslant \delta_k \leqslant d'_w(k)$ for $0 \leqslant k \leqslant v-2$.

Proof. Define $c_i = \big|\{\pi \in X : \pi^{-1}(w) = i < \pi^{-1}(w')\}\big|$ and $c'_i = \big|\{\pi \in X : \pi^{-1}(w) = i > \pi^{-1}(w')\}\big|$ for $0 \leqslant i \leqslant v-1$. Then $c_i + c'_i = d_w(i)$ and $c_i + c'_{i+1} = d'_w(i)$. Now $c'_0 = 0$ and $c'_{i+1} - c'_i = d'_w(i) - c_i - c'_i = d'_w(i) - d_w(i)$. So it follows by induction on i that $\delta_i = c'_{i+1}$ for $0 \leqslant i \leqslant v-2$. The result then follows from the fact that $c_i \geqslant 0$ and $c'_i \geqslant 0$ for each i, by definition. □

We say that $\boldsymbol{d}_w$ and $\boldsymbol{d}'_w$ are *compatible* if they satisfy Theorem 7.7. This test can be used to eliminate some distributions from consideration. If $\boldsymbol{d}_w$ is (v,t,λ)-feasible, it may be the case that there is no $(v-1,t,\lambda)$-feasible distribution $\boldsymbol{d}'_w$ compatible with $\boldsymbol{d}_w$. It may even happen that there is a compatible $\boldsymbol{d}'_w$, but that all such candidates can themselves be ruled out because they are not compatible with a $(v-2,t,\lambda)$-feasible distribution, and so on. A concrete example is that $(2,6,1,1,6,2)$ is a $(6,3,3)$-feasible distribution. The only $(5,3,3)$-feasible distribution that it is compatible with is $(3,6,0,6,3)$. However

$(3, 6, 0, 6, 3)$ is not compatible with any of the four $(4, 3, 3)$-feasible distributions, which are (3,9,0,6), (4,6,3,5), (5,3,6,4) and (6,0,9,3). Hence $(2, 6, 1, 1, 6, 2)$ and $(3, 6, 0, 6, 3)$ can be eliminated from consideration.

Table 7.3 records the number of $(v, t, 1)$-feasible distributions for different values of v and t, as well as incorporating information about how many distributions cannot be ruled out using Theorem 7.7 in the manner just described.

7.3 Exhaustive search algorithm

We have seen in the previous section the relationship between a $\mathsf{PSCA}(v, t, \lambda)$ and the smaller array that results from deleting a symbol from this PSCA. Specifically, we have seen that by deleting a symbol from a $\mathsf{PSCA}(t+1, t, \lambda)$, we are left with λ copies of $\mathcal{S}_t$. We can extend this argument to say that by deleting $v - t$ symbols from a $\mathsf{PSCA}(v, t, \lambda)$, we obtain λ copies of $\mathcal{S}_t$. In this sense, every PSCA contains λ copies of $\mathcal{S}_t$. This relationship between smaller and larger PSCAs with the same strength and multiplicity allows for the design of an algorithm that can exhaustively search for all possible $\mathsf{PSCA}(v, t, \lambda)$ by first cataloguing all possible $\mathsf{PSCA}(v', t, \lambda)$ for $t \leqslant v' < v$. Such an algorithm is further aided by the results proved in the previous section. In order to catalogue all possible PSCAs for a particular choice of parameters, we must first establish a definition of isomorphism for PSCAs. Recall that by Lemma 3.22, if P is a $\mathsf{PSCA}(v, t, 1)$, then for $\sigma \in \mathcal{S}_v$, $\sigma P = \{\sigma \circ \pi : \pi \in P\}$ is also a $\mathsf{PSCA}(v, t, 1)$. It is straightforward to adapt this argument to show σP is a $\mathsf{PSCA}(v, t, \lambda)$ when P is a $\mathsf{PSCA}(v, t, \lambda)$ for any $\lambda \geqslant 1$. Recall that the permutations in σP are exactly the permutations in P with the symbols rearranged according to σ. We also note that if P is a $\mathsf{PSCA}(v, t, \lambda)$ then the array formed by taking the reverse of every permutation in P is also a $\mathsf{PSCA}(v, t, \lambda)$. This can be seen by observing that $\pi \in \mathcal{S}_v$ covers $(s_0, \ldots, s_{t-1})$ if and only the reverse of π covers $(s_{t-1}, \ldots, s_0)$. Thus, we say that two $\mathsf{PSCA}(v, t, \lambda)$ are *isomorphic* if one can be obtained from the other by permuting the symbols and/or reversing every permutation.

In searching for $\mathsf{PSCA}(v, t, \lambda)$ for $v > t$, we employed two different methods. Both of these methods relied on a complete catalogue of isomorphism class representatives of $\mathsf{PSCA}(v-1, t, \lambda)$. For each array in this catalogue, we tested every possible way of inserting a new symbol into each permutation of the array. In the first method, we assigned a (v, t, λ)-feasible distribution for this new symbol and found all possible PSCAs that can be formed when the new symbol obeys that distribution, before moving on to the next (v, t, λ)-feasible distribution. In the second method, we did not fix a distribution. Instead, we maintained a list of (v, t, λ)-feasible distributions that were consistent with the positions so far chosen for the new symbol. If that list ever became empty then we knew the current placements were unviable. Using these two search methods, we were able to independently count the number of isomorphism classes of $\mathsf{PSCA}(v, t, \lambda)$ for different sets of parameters, as shown in Table 7.4. In some cases it was not feasible to

		v							
	t	3	4	5	6	7	8	9	10
$\lambda = 1$	3	1/1	2/2	2/3	0/1	0/3	0/4	0/5	0/9
	4	-	1/1	3/3	6/6	8/13	19/30	36/57	61/119
	5	-	-	1/1	5/5	21/27	117/127	570/689	3359/3620
$\lambda = 2$	3	1/1	3/3	6/8	8/12	16/28	30/55	44/99	67/165
	4	-	1/1	5/5	17/17	59/74	261/291	1034/1128	3940/4235
	5	-	-	1/1	9/9	79/93	900/910	9267/9908	106859/107947
$\lambda = 3$	3	1/1	4/4	11/14	32/37	84/99	224/252	547/609	1315/1409
	4	-	1/1	7/7	35/35	195/221	1246/1296	7243/7341	38781/39486
	5	-	-	1/1	13/13	179/199	2933/2951	46160/48150	790491/793171

Table 7.3: Number of feasible distributions. Each entry r/s indicates that there are s distributions that are (v, k, λ)-feasible, and that r of these cannot be ruled out using Theorem 7.7.

perform an exhaustive enumeration. In such cases, the number of PSCAs that we found before abandoning the search is given with a + symbol indicating that the search was incomplete. In each such case we believe that the true number of PSCAs is much higher than the number that we quote.

In the cases when $(v, t, \lambda) \in \{(5,3,1), (7,4,1), (7,5,1), (8,3,2), (8,4,2)\}$ our enumeration was exhaustive, and demonstrated that no PSCA with these parameters exists. For the first three of these parameter sets this was already known, but the last two are new results. Our computations have discovered several new values of the function PSCAN.

Theorem 7.8. $\mathsf{PSCAN}(6,3) = \mathsf{PSCAN}(7,3) = \mathsf{PSCAN}(7,4) = 2$ *and* $\mathsf{PSCAN}(8,3) = 3$. *Additionally,* $\mathsf{PSCAN}(8,4) > 2$.

Proof. Given the nonexistence results just mentioned, it suffices to present a $\mathsf{PSCA}(7,3,2)$, a $\mathsf{PSCA}(8,3,3)$ and a $\mathsf{PSCA}(7,4,2)$:

PSCA(7,3,2)		PSCA(8,3,3)		PSCA(7,4,2)			
0123465	0642315	04712563	05672341	0123465	0254163	0351264	0432165
1540362	1634052	06432157	07351462	0621435	0634125	0651432	0652341
2405163	2610543	16547203	17453026	1045263	1254063	1432560	1530264
3054261	3625401	17630245	25476301	1632045	1635402	1640253	1652043
4312560	4651230	26751043	27410365	2045361	2103564	2341560	2530164
5231064	5603124	31526074	34675102	2601534	2635104	2643015	2645103
		37206154	42351067	3015462	3214065	3402561	3520461
		46051327	50213476	3604521	3610254	3614520	3625401
		53764201	61234075	4015362	4123065	4351062	4520163
				4610352	4620351	4621530	4653012
				5103462	5214360	5341260	5402361
				5603214	5604123	5612340	5643210

□

There are 260 664 isomorphism classes of $\mathsf{PSCA}(5,3,3)$. We took the one which has the largest automorphism group and extended it in all possible ways. Doing so produced 3072, 481 765 and 51 448 isomorphism classes of PSCAs with parameters $(6,3,3)$, $(7,3,3)$ and $(8,3,3)$ respectively. However, none of these extended to a $\mathsf{PSCA}(9,3,3)$. We also performed a search for all $\mathsf{PSCA}(6,3,3)$ in which every symbol has distribution vector $(3,3,3,3,3,3)$. Using Theorem 7.7, we were able to find all $(5,3,3)$-feasible distributions that are compatible with this uniform distribution and thus could determine the $\mathsf{PSCA}(5,3,3)$ that could potentially extend to such a $\mathsf{PSCA}(6,3,3)$. From them we found 1 053 700 $\mathsf{PSCA}(6,3,3)$ up to isomorphism. These arrays extend to 35 872 460 $\mathsf{PSCA}(7,3,3)$ and 1 992 709 $\mathsf{PSCA}(8,3,3)$ up to isomorphism. Again, none of these arrays extend to a

t	λ	v	PSCAs	groups
3	1	3	1	1
3	1	4	1	0
3	1	5	0	0
3	2	3	1	1
3	2	4	12	1
3	2	5	314	0
3	2	6	1957	5
3	2	7	146	0
3	2	8	0	0
3	3	3	1	1
3	3	4	37	0
3	3	5	260 664	0
3	3	6	29 100 897+	0+
3	3	7	14 943 804+	0+
3	3	8	2 111 540+	0+
4	1	4	1	1
4	1	5	4	0
4	1	6	2	1
4	1	7	0	0
4	2	4	1	1
4	2	5	12 351	0
4	2	6	32 507	2
4	2	7	1826	0
4	2	8	0	0
5	1	5	1	1
5	1	6	3461	0
5	1	7	0	0

Table 7.4: Number of PSCAs generated by adding one symbol at a time.

PSCA$(9, 3, 3)$. We also built some other PSCA$(8, 3, 3)$ via several other routes, but were unable to find a PSCA$(9, 3, 3)$.

The last column of Table 7.4 lists the number of isomorphism classes in our catalogue which contain a PSCA for which the corresponding set of permutations (ignoring multiplicity of repeated permutations) forms a group. To test if a PSCA is isomorphic to a group it suffices to permute the symbols to ensure that one permutation (it does not matter which) is the identity, and then check that the resulting set of permutations is closed under composition. PSCAs that form groups will be studied further in the next section, which will provide details of all of the groups included in Table 7.4 (except the trivial cases when $v = t$).

In Table 7.3 we showed how many distributions might be achieved by symbols in PSCAs. In the "Realised Distributions" column of Table 7.5 we show how many of these

v	t	λ	Realised Distributions	Compatible Distributions
3	3	1	1	1
4	3	1	2	2
3	3	2	1	1
4	3	2	3	3
5	3	2	6	6
6	3	2	4	8
7	3	2	2	16
3	3	3	1	1
4	3	3	4	4
5	3	3	11	11
6	3	3	26	32
4	4	1	1	1
5	4	1	3	3
6	4	1	1	6
4	4	2	1	1
5	4	2	5	5
6	4	2	10	17
7	4	2	16	59
5	5	1	1	1
6	5	1	5	5

Table 7.5: Number of realised distributions for different parameter sets.

distributions are actually present within some PSCA. For comparison, the column headed "Compatible Distributions" repeats the smaller of the two bounds we had computed in Table 7.3. Table 7.5 covers all cases where we computed (non-empty) exhaustive catalogues. It also covers the case $(v, t, \lambda) = (6, 3, 3)$, where we were able to rule out 6 distributions with targeted searches, assisted by Theorem 7.7. The 6 unrealised distributions were (0,9,1,3,0,5), (2,6,0,4,3,3), (3,1,8,0,2,4) and their reverses. The other 26 distributions from Table 7.3 appeared in our partial catalogue.

7.4 PSCAs from permutation groups

In this section, we consider PSCAs which can be constructed from permutation groups. For permutations $f, g \in \mathcal{S}_v$, the composition $f \circ g$ is the permutation $(f \circ g)(x) = f(g(x))$. For a subgroup $H \leqslant G$ and for $g \in G$, the right coset Hg is the set $\{hg : h \in H\}$ whereas the left coset gH is the set $\{gh : h \in H\}$. If H is a subgroup of $\mathcal{S}_v$, then the right coset Hg permutes the columns of H according to g while the left coset gH permutes the symbols of H according to g. Throughout this section, for $s \in \mathcal{S}_{v,t}$, we use the notation $s = (s_0, \dots, s_{t-1})$. Moreover, G will always denote a group such that if G has order v, then the elements of G are $\{0, \dots, v-1\}$, and ψ will denote an injective homomorphism $\psi : G \to \mathcal{S}_v, g \mapsto \psi_g$. We can then consider the action of G on $\mathcal{S}_{v,t}$ where, if $s = (s_0, \dots, s_{t-1}) \in \mathcal{S}_{v,t}$ and $g \in G$, then $gs = (\psi_g(s_0), \dots, \psi_g(s_{t-1}))$.

Mathon and Tran van Trung [41] found that there are exactly two non-isomorphic PSCA$(6, 4, 1)$; one forms a group isomorphic to $\mathcal{S}_4$ while the other forms three cosets of a group isomorphic to D_8. While they note the connection between their PSCAs and groups, their search methods did not focus on building PSCAs from groups (the same is true of our work in Chapter 7.3). However, several connections between PSCAs and groups have been formalised by Na, Jedwab and Li [44] and they found a number of examples of PSCAs based on groups. Note that the permutation composition convention used in [44] differs from the convention used here.

Lemma 7.9. *Let G be a group, $\psi : G \to \mathcal{S}_v$ be an injective homomorphism, T be the image of ψ and let Th be a right coset of T. If x and y are sequences belonging to the same orbit under the action of G on $\mathcal{S}_{v,t}$, then x and y are covered by the same number of permutations in Th.*

Proof. Let x and y be elements of $\mathcal{S}_{v,t}$ that belong to the same orbit under the action of G with $x = (x_0, \dots, x_{t-1})$. As x and y belong to the same orbit, $gx = y$ for some $g \in G$. Let $0 \leqslant c_0 < \dots < c_{t-1} \leqslant v-1$ and suppose $f \in Th$ such that $f(c_i) = x_i$ for $0 \leqslant i \leqslant t-1$. Then, f covers x. Now consider $\psi_g \circ f$. As $f(c_i) = x_i$, $(\psi_g \circ f)(c_i) = \psi_g(x_i)$ for $0 \leqslant i \leqslant t-1$. Therefore, $\psi_g \circ f$ covers y. Therefore, for every permutation in Th that covers x, we can find a corresponding permutation that covers y. So, the number of permutations in Th that cover y is at least the number of permutations in Th that cover x. By reversing the argument, and noting $x = g^{-1}y$, we find that the number of permutations in Th that cover x is at least the number of permutations in Th that cover y. Thus, x and y are covered by the same number of permutations in Th. □

A consequence of Lemma 7.9 is that in a right coset of a permutation group $\psi(G)$, we can determine the number of permutations covering each sequence in the orbit of a sequence x under the action of G on $\mathcal{S}_{v,t}$ by simply finding the number of permutations in the coset that cover x. We will develop this point further in the context of transitive permutation groups in Lemma 7.10. Recall that a set X of permutations in $\mathcal{S}_v$ is *transitive* if for each $i, j \in [v]$, there is a permutation $\pi \in X$ such that $\pi(i) = j$. If for each i and j, the permutation π is unique, then X is *sharply transitive*.

Lemma 7.10. *Let G be a group, let $\psi : G \to \mathcal{S}_v$ be an injective homomorphism such that the image, T, of ψ is a transitive permutation group and let X be an array constructed from right cosets of T. Furthermore, let $w \in [v]$, $0 \leqslant i \leqslant t-1$ and let $S = \{s \in \mathcal{S}_{v,t} : s_i = w\}$. If every sequence in S is covered by λ permutations in X, then X is a* $\mathsf{PSCA}(v, t, \lambda)$.

Proof. Let $s \in \mathcal{S}_{v,t}$. Then, as T is transitive, there is a $g \in G$ such that $\psi_g(s_i) = w$. Therefore, the orbit of s contains a sequence in S. As every orbit of the action of G on $\mathcal{S}_{v,t}$ contains a representative from S, then by Lemma 7.9, if every sequence in S is covered by λ permutations in X, then every sequence in $\mathcal{S}_{v,t}$ is also covered by λ permutations in X. □

7.5 Elementary abelian 2-groups

Throughout this subsection, we use E_v to denote an elementary abelian 2-group on the set $[v]$ with identity 0 and operation $\oplus$. Then for a group E_v, we fix $\psi : E_v \to \mathcal{S}_v$ to be the homomorphism that maps $g \mapsto \psi_g$ where $\psi_g(x) = g \oplus x$. We then let T be the image of ψ. Under this homomorphism, $gs = (g \oplus s_0, \ldots, g \oplus s_{t-1})$ for $g \in E_v$ and $s \in \mathcal{S}_{v,t}$. By construction, T is a sharply transitive set of permutations, a fact that will be critical in what follows. We begin our analysis of elementary abelian 2-groups with an overview of PSCAs built from E_4. Within $\mathcal{S}_4$, there are several subgroups isomorphic to E_4. However, the only one of these subgroups that is sharply transitive is the following:

$$\begin{matrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 3 & 2 \\ 2 & 3 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{matrix}$$

The cosets of this group within $\mathcal{S}_4$ are shown in Figure 7.1. We refer to the cosets on the left as having Type A coverage, the cosets in the middle as having Type B coverage and the cosets on the right as having Type C coverage. Cosets of the same type cover

0	1	2	3	0	2	1	3	0	3	1	2
1	0	3	2	1	3	0	2	1	2	0	3
2	3	0	1	2	0	3	1	2	1	3	0
3	2	1	0	3	1	2	0	3	0	2	1
0	1	3	2	0	2	3	1	0	3	2	1
1	0	2	3	1	3	2	0	1	2	3	0
2	3	1	0	2	0	1	3	2	1	0	3
3	2	0	1	3	1	0	2	3	0	1	2
Type A				Type B				Type C			

Figure 7.1: Three types of cosets of E_4

Type A	Type B	Type C
021	012	013
031	032	023
120	103	102
130	123	132
203	210	201
213	230	231
302	301	310
312	321	320

Table 7.6: Triples uncovered by cosets of Type A, B and C.

the same set of triples. Each coset covers 16 triples of $\mathcal{S}_{4,3}$ exactly once, leaving 8 triples uncovered. These uncovered triples are recorded in Table 7.6.

Observe that the sets of triples uncovered by Type A, Type B and Type C cosets partition $\mathcal{S}_{4,3}$. Suppose X is a $\mathsf{PSCA}(4,3,\lambda)$ which is built from a combination of cosets of our E_4 permutation group. As the number of permutations in X is 6λ, the total number of cosets that make up X is $3\lambda/2$. Consider the triple 012. This triple is covered by Type A and Type C cosets but is not covered by Type B cosets. Given that the number of permutations that cover 012 is λ, there must be $\lambda/2$ Type B cosets. Similar arguments involving other triples (e.g. 021 and 013) demonstrate that X must be built from $\lambda/2$ of each type of coset. Furthermore, because of the coverage properties of each coset type, any combination of $\lambda/2$ Type A cosets, $\lambda/2$ Type B cosets and $\lambda/2$ Type C cosets will form a $\mathsf{PSCA}(4,3,\lambda)$. Therefore, an array built from a combination of cosets of E_4 will form a $\mathsf{PSCA}(4,3,\lambda)$ if and only if the array contains an equal number of each type of coset.

We use this characterisation to aid us in our search for PSCAs from cosets of permutation representations of the elementary abelian 2-group of order v with $v > 4$. Obviously these larger groups contain many subgroups isomorphic to E_4. As in the general case above, we isolate a subset of triples of $\mathcal{S}_{v,3}$ such that balanced coverage on these triples implies balanced coverage for every triple in $\mathcal{S}_{v,3}$.

Lemma 7.11. *Let $\mathcal{H}$ be the set of order 4 subgroups of E_v and let S be the set of triples defined by*

$$S = \{(s_0, s_1, s_2) \in \mathcal{S}_{v,3} : \{s_0, s_1, s_2\} \subset H \text{ for some } H \in \mathcal{H}\}$$

Let X be an array constructed from right cosets of T in $\mathcal{S}_v$. If every triple in S is covered by λ permutations in X, then X is a $\mathsf{PSCA}(v, 3, \lambda)$*.*

Proof. First, we observe that if $\{x, y, z\}$ is a 3-subset of an elementary abelian 2-group, then $\{x, y, z, x \oplus y \oplus z\}$ is a coset of the order 4 subgroup $\{0, x \oplus y, x \oplus z, y \oplus z\}$. Furthermore, $x \oplus y \oplus z$ is the only element we can add to $\{x, y, z\}$ in order to form a coset of some subgroup in $\mathcal{H}$.

Let $(x, y, z) \in \mathcal{S}_{v,3}$. If ψ_x acts on (x, y, z), we obtain the triple $(0, x \oplus y, x \oplus z)$. As per the previous paragraph, $\{0, x \oplus y, x \oplus z\}$ forms a subset of a subgroup in $\mathcal{H}$ so $(0, x \oplus y, x \oplus z) \in S$. Hence, each orbit of $\mathcal{S}_{v,3}$ under the action of E_v contains a triple from S. Therefore, by Lemma 7.9, if every triple in S is covered by λ permutations in X, then X is a $\mathsf{PSCA}(v, 3, \lambda)$. ☐

Let $X \subseteq \mathcal{S}_v$ be a multiset of permutations. For $W \subseteq [v]$, the *reduced array of X on W*, denoted by $X[W]$, is the array we obtain by removing every symbol of $[v] \setminus W$ from X. Let Y be a right coset of T in $\mathcal{S}_v$, let H be an order 4 subgroup of E_v and consider the reduced array $Y[H]$. The symbols of H appear in 4×4 subarrays within T. The rows and columns of each of these subarrays correspond to different cosets of H in E_v. Each subarray will be (up to reordering rows and columns) one of the cosets of the sharply transitive E_4 permutation group with symbol set H. So $Y[H]$ is made up of cosets of this sharply transitive E_4 permutation group. By taking X to be a collection of right cosets of T, we can determine whether $X[H]$ forms a PSCA by analysing the coverage type of each coset of E_4 that appears in $X[H]$. As a result of Lemma 7.11, if the reduced array $X[H]$ is a PSCA of strength 3 for each $H \in \mathcal{H}$, then X will be a PSCA of strength 3.

Lemma 7.12. *Let f be an order n automorphism of E_v and let X be the array*

$$X = \bigcup_{i=0}^{n-1} T f^i.$$

Let H be an order 4 subgroup of E_v. If the reduced array $X[H]$ is a $\mathsf{PSCA}(4, 3, \lambda)$*, then $X[f^i(H)]$ will also be a* $\mathsf{PSCA}(4, 3, \lambda)$ *for $1 \leqslant i \leqslant n - 1$.*

Proof. First we show that $T = f^{-1}Tf$. Let $g \in E_v$. Then, we can consider $\psi_g \in T$ and the composition $f^{-1}\psi_g f$. Let $x \in E_v$. Then, $f^{-1}\psi_g f(x) = f^{-1}(g \oplus f(x))$. As f is an automorphism of E_v, so too is f^{-1}. Hence $f^{-1}(g \oplus f(x)) = f^{-1}(g) \oplus x$. Therefore, $f^{-1}\psi_g f = \psi_{f^{-1}(g)}$ and hence, $f^{-1}Tf \subseteq T$. Now, $\psi_g = \psi_{f^{-1}(f(g))} = f^{-1}\psi_{f(g)} f$ by the above argument. So, $T \subseteq f^{-1}Tf$ and thus, $T = f^{-1}Tf$. Therefore, $fT = Tf$ and so we can consider Tf as being an array in which the symbols of T have been permuted according to f. As a result, the reduced array $T[H]$ is isomorphic to $Tf[f(H)]$. More generally, the reduced array $Tf^i[H]$ is isomorphic to $Tf^{i+1}[f(H)]$ for $0 \leqslant i \leqslant n-1$. Moreover, the isomorphism in each case is the restriction of f to H. Therefore, $X[H]$ is isomorphic to $X[f(H)]$. Applying this argument to the subgroups $f^i(H)$ and $f^{i+1}(H)$ for $0 \leqslant i \leqslant n-1$, we find that $X[H]$ is isomorphic to $X[f^i(H)]$ for $1 \leqslant i \leqslant n-1$. Therefore, if $X[H]$ is a $\mathsf{PSCA}(4,3,\lambda)$, then so is $X[f^i(H)]$ for $1 \leqslant i \leqslant n-1$. □

Essentially, Lemma 7.11 demonstrates that in a collection of right cosets of T, it suffices to check the coverage of triples whose elements form a subset of an order 4 subgroup E_v in order to determine whether the cosets form a PSCA. When these cosets are related by an automorphism of E_v, we are able to further restrict what triples need to be checked by allowing us to consider only certain subgroups, depending on the automorphism f. In each case, the reduced array on any order 4 subgroup H will form a collection of cosets of E_4 and so we can use the characterisation at the start of this section to determine whether these reduced arrays form PSCAs. Using these methods, we have been able to find PSCAs of orders 4, 8, 16 and 32 with strength 3. The following are examples of a $\mathsf{PSCA}(4,3,2)$ and a $\mathsf{PSCA}(8,3,4)$ (note that Na, Jedwab and Li [44] also found a $\mathsf{PSCA}(8,3,4)$).

PSCA(4, 3, 2)	PSCA(8, 3, 4)	
0123	01234567	42671053
1032	10543276	53106742
2301	25076143	60435217
3210	34701652	71342506
0231	43610725	07245316
1320	52167034	16532407
2013	67452301	23061754
3102	76325410	32716045
0312	06253471	45607132
1203	17524360	54170623
2130	24017635	61423570
3021	35760124	70354261

The $\mathsf{PSCA}(4,3,2)$ forms a permutation group isomorphic to the alternating group A_4. The $\mathsf{PSCA}(8,3,4)$ forms a permutation group isomorphic to $A_4 \times C_2$. We also have the following PSCAs of orders 16 and 32.

Theorem 7.13. *$g(v,3) \leqslant 16$ for $v \leqslant 16$ and $g(v,3) \leqslant 96$ for $v \leqslant 32$.*

Proof. To prove the first part of the theorem, we need only present a $\mathsf{PSCA}(16,3,16)$. We let G be the group isomorphic to E_{16} generated by the permutations

$$(0\ 1)(2\ 3)(4\ 5)(6\ 7)(8\ 9)(10\ 11)(12\ 13)(14\ 15),$$
$$(0\ 2)(1\ 3)(4\ 14)(5\ 15)(6\ 12)(7\ 13)(8\ 10)(9\ 11),$$
$$(0\ 4)(1\ 5)(2\ 14)(3\ 15)(6\ 10)(7\ 11)(8\ 12)(9\ 13),$$
$$(0\ 8)(1\ 9)(2\ 10)(3\ 11)(4\ 12)(5\ 13)(6\ 14)(7\ 15).$$

We then let $f =$(1 8 9)(2 4 15 11 5 7)(3 12 6 10 13 14). Then,

$$X = \bigcup_{i=0}^{5} Gf^i$$

forms a $\mathsf{PSCA}(16, 3, 16)$. The 96 permutations of this PSCA also form a group which can be generated by

$$(1\ 8\ 9)(2\ 4\ 15\ 11\ 5\ 7)(3\ 12\ 6\ 10\ 13\ 14),$$
$$(0\ 4\ 7)(1\ 13\ 15)(2\ 3\ 10)(5\ 14\ 8)(6\ 9\ 12).$$

As a result of Lemma 7.11, in order to check whether X forms a PSCA, we need only check that the reduced arrays of X corresponding to the 35 order 4 subgroups of G each form a $\mathsf{PSCA}(4, 3, 16)$. As the cosets of G from which X is constructed are related by an automorphism, we can use Lemma 7.12 to further limit the number of reduced arrays of X that we need to check in order to verify that X is a PSCA. The orbits of the 35 order 4 subgroups of G under f are as follows.

$$\{\{0,1,2,3\},\{0,4,8,12\},\{0,6,9,15\},\{0,1,10,11\},\{0,5,8,13\},\{0,7,9,14\}\}$$
$$\{\{0,1,6,7\},\{0,2,8,10\},\{0,4,9,13\},\{0,1,14,15\},\{0,3,8,11\},\{0,5,9,12\}\}$$
$$\{\{0,2,4,14\},\{0,3,4,15\},\{0,11,12,15\},\{0,5,6,11\},\{0,5,7,10\},\{0,2,7,13\}\}$$
$$\{\{0,2,6,12\},\{0,4,6,10\},\{0,10,13,15\},\{0,11,13,14\},\{0,3,5,14\},\{0,3,7,12\}\}$$
$$\{\{0,1,4,5\},\{0,7,8,15\},\{0,2,9,11\}\}$$
$$\{\{0,1,12,13\},\{0,6,8,14\},\{0,3,9,10\}\}$$
$$\{\{0,2,5,15\},\{0,4,7,11\}\}$$
$$\{\{0,3,6,13\},\{0,10,12,14\}\}$$
$$\{\{0,1,8,9\}\}$$

Hence, by Lemma 7.12, we need only check the reduced array of one subgroup from each of these 9 orbits to verify that X is a PSCA.

For the second part of the theorem, we present a $\mathsf{PSCA}(32, 3, 96)$. We let G_{32} be the group isomorphic to E_{32} generated by the permutations

$$\pi_1 = (0\ 1)(2\ 3)(4\ 5)(6\ 7)(8\ 9)(10\ 11)(12\ 13)(14\ 15)(16\ 17)(18\ 19)(20\ 21)(22\ 23)(24\ 25)(26\ 27)(28\ 29)(30\ 31),$$
$$\pi_2 = (0\ 2)(1\ 3)(4\ 28)(5\ 29)(6\ 30)(7\ 31)(8\ 10)(9\ 11)(12\ 20)(13\ 21)(14\ 22)(15\ 23)(16\ 18)(17\ 19)(24\ 26)(25\ 27),$$

$$\pi_3 = (0\ 4)(1\ 5)(2\ 28)(3\ 29)(6\ 26)(7\ 27)(8\ 14)(9\ 15)(10\ 22)(11\ 23)(12\ 18)(13\ 19)(16\ 20)(17\ 21)(24\ 30)(25\ 31),$$
$$\pi_4 = (0\ 8)(1\ 9)(2\ 10)(3\ 11)(4\ 14)(5\ 15)(6\ 12)(7\ 13)(16\ 24)(17\ 25)(18\ 26)(19\ 27)(20\ 30)(21\ 31)(22\ 28)(23\ 29),$$
$$\pi_5 = (0\ 16)(1\ 17)(2\ 18)(3\ 19)(4\ 20)(5\ 21)(6\ 22)(7\ 23)(8\ 24)(9\ 25)(10\ 26)(11\ 27)(12\ 28)(13\ 29)(14\ 30)(15\ 31).$$

We then let f_1 be the following order 2 automorphism of G_{32}:

$$(2\ 8)(3\ 9)(4\ 6)(5\ 7)(12\ 28)(13\ 29)(14\ 30)(15\ 31)(18\ 24)(19\ 25)(20\ 22)(21\ 23).$$

We then let $G_{64} = G_{32} \cup G_{32}f_1$. Observe that G_{64} also forms a group. Then, we let f_2 be the following order 3 automorphism of G_{32}:

$$(2\ 12\ 24)(3\ 13\ 25)(4\ 6\ 10)(5\ 7\ 11)(8\ 18\ 28)(9\ 19\ 29)(20\ 22\ 26)(21\ 23\ 27).$$

We then let $G_{192} = G_{64} \cup G_{64}f_2 \cup G_{64}f_2^2$. Again, G_{192} forms a group. Finally, we let f_3 be the following order 3 automorphism of G_{32}:

$$(1\ 16\ 17)(3\ 18\ 19)(5\ 20\ 21)(6\ 7\ 23)(9\ 24\ 25)(11\ 26\ 27)(12\ 13\ 29)(15\ 30\ 31).$$

Then $G_{192} \cup f_3G_{192} \cup f_3^2G_{192}$ is a $\mathsf{PSCA}(32, 3, 96)$. Although this construction is not of the form described in Lemma 7.12, it is a collection of right cosets of G_{32}. Therefore, we can use Lemma 7.11 to check that this array is indeed a PSCA. □

We remark that even though G_{32}, G_{64} and G_{192} are groups, the $\mathsf{PSCA}(32, 3, 96)$ described in Theorem 7.13 is not a group. We also note that while f_3 is an automorphism of G_{32}, it is not an automorphism of G_{192}. As such, the shift to left cosets in the final step of the construction is significant as taking right cosets would not form a PSCA.

Motivated by those PSCAs that we had earlier found which turned out to be permutation representations of groups, we decided to search for such objects directly. Fix v, t and λ. We sought a representation in $\mathcal{S}_v$ of some group of order $n = t!\lambda$. We began by deciding on positive integers g_1, g_2 and possibly g_3. We then chose permutations of orders g_1, g_2 (and possibly g_3) and checked whether they generate a group of order n. For each group that we discovered in this way, we then tried to find a conjugate that was a PSCA. This was done by building up the PSCA one column at a time, backtracking whenever some t-sequence would be covered too many times. As the conjugate $h^{-1}Gh$ of a group G is isomorphic in terms of sequence coverage to Gh, searching over all column permutations of G for a PSCA is equivalent to searching over all conjugates of G. Since we checked all conjugates of each group that we found, we were free to insist that the generator of order g_1 that we chose was lexicographically maximal amongst all of its conjugates. In particular, this meant we only had to consider one choice for each possible cycle structure of that generator. Note that this method did not prejudge which group it was going to build. Many non-isomorphic groups of order n may have generators of the specified orders. For example, there are 15 groups of order 24, but they all have a

(v, λ)	Group	Generators
(4,2)	A_4	$\langle(1,2,3),(0,1,2)\rangle$
(6,2)	A_4	$\langle(0,5,4)(1,2,3),(0,5,1)(2,3,4)\rangle$
	D_{12}	$\langle(0,5,4,2,1,3),(0,5)(1,2)(3,4)\rangle$
		$\langle(0,4,5,2,1,3),(0,4)(1,2)(3,5)\rangle$
		$\langle(0,3,1,5,4,2),(0,5)(1,3)(2,4)\rangle$
		$\langle(0,3,1,4,5,2),(0,4)(1,3)(2,5)\rangle$
(6,4)	$C_2 \times A_4$	$\langle(0,5,1,2,3,4),(0,5,4)(1,2,3)\rangle$
	$\mathcal{S}_4$	$\langle(0,2)(1,3)(4,5),(0,4,5)(1,3,2)\rangle$
		$\langle(0,2)(1,3)(4,5),(0,4,5)(1,2,3)\rangle$
		$\langle(0,5)(1,3)(2,4),(0,2,4)(1,3,5)\rangle$
		$\langle(1,3)(4,5),(0,1,5)(2,4,3)\rangle$
		$\langle(1,3)(2,4),(0,1,4)(2,3,5)\rangle$
(8,4)	$SL(2,3)$	$\langle(0,7,4,2)(1,5,3,6),(0,7,1)(2,3,4)\rangle$
		$\langle(0,7,6,4)(1,3,2,5),(0,7,3)(4,5,6)\rangle$
	$\mathcal{S}_4$	$\langle(0,5,4,2)(1,7,3,6),(0,7,4)(1,3,2)\rangle$
		$\langle(0,5,4,2)(1,6,3,7),(0,7,4)(1,3,5)\rangle$
		$\langle(0,5,6,3)(1,4,7,2),(0,7,6)(2,4,5)\rangle$
		$\langle(0,5,7,3)(1,4,6,2),(0,7,6)(2,5,4)\rangle$
		$\langle(0,5,7,4)(1,3,6,2),(0,7,6)(2,5,3)\rangle$
		$\langle(0,5,6,4)(1,3,7,2),(0,7,6)(2,3,5)\rangle$
		$\langle(0,6,4,3)(1,7,5,2),(0,5,4)(2,7,6)\rangle$
		$\langle(0,7,3,6)(1,2,4,5),(0,5,3)(1,4,7)\rangle$
		$\langle(0,7,3,6)(1,5,4,2),(0,5,3)(1,4,6)\rangle$
		$\langle(0,7,4,3)(1,6,5,2),(0,5,4)(2,6,7)\rangle$
		$\langle(0,7,4,6)(1,3,5,2),(0,5,4)(2,3,7)\rangle$
		$\langle(0,7,5,2)(1,6,4,3),(0,5,4)(2,3,6)\rangle$
	$C_2 \times A_4$	$\langle(1,4,7)(2,5,3),(0,1)(2,7)(3,6)(4,5)\rangle$
		$\langle(1,4,6)(2,5,3),(0,1)(2,6)(3,7)(4,5)\rangle$
		$\langle(1,2,5)(4,7,6),(0,1)(2,7)(3,6)(4,5)\rangle$
		$\langle(1,2,5)(4,6,7),(0,1)(2,6)(3,7)(4,5)\rangle$
		$\langle(1,6,7)(2,5,4),(0,2)(1,5)(3,6)(4,7)\rangle$
		$\langle(1,6,7)(2,4,5),(0,2)(1,5)(3,7)(4,6)\rangle$
		$\langle(1,6,7)(2,5,3),(0,2)(1,5)(3,7)(4,6)\rangle$
		$\langle(1,6,7)(2,3,5),(0,2)(1,5)(3,6)(4,7)\rangle$

Table 7.7: Strength 3 PSCAs that are permutation groups for $v \leqslant 8$.

generating set with $(g_1, g_2) \in \{(12,4),(12,2),(8,3),(6,4),(3,2)\}$ or $(g_1, g_2, g_3) = (6,6,2)$. Similarly, the 5 groups of order 18 all have a generating set with $(g_1, g_2) \in \{(9,2),(6,6)\}$ or $(g_1, g_2, g_3) = (3,3,2)$. Of course, groups will typically have many different generating sets with suitable orders, and hence will be built multiple times. But we could be confident that every group of order n that has some representation in $\mathcal{S}_v$ would be built, and thus that our catalogue of PSCAs that are groups is exhaustive for $v \leqslant 14$ and $n \leqslant 42$.

In an alternative computation, we used GAP [56] to generate representatives of conju-

(v, λ)	Group	Generators
(12,6)	$C_6 \times \mathcal{S}_3$	$\langle(0,11,9,10,1,4)(2,7,6,3,8,5),(0,8,9,2,1,6)(3,4,5,11,7,10)\rangle$
	$\mathcal{S}_3 \times \mathcal{S}_3$	$\langle(0,11,9,10,1,4)(2,5,8,3,6,7),(0,8,9,2,1,6)(3,4,7,10,5,11)\rangle$
	$C_3 \times A_4$	$\langle(2,5,9)(3,6,8)(4,10,11),(0,2,4)(1,10,8)(3,9,7)(5,11,6)\rangle$
		$\langle(2,5,9)(3,6,8)(4,11,10),(0,2,4)(1,11,8)(3,9,7)(5,10,6)\rangle$
		$\langle(2,5,9)(3,6,8)(4,7,11),(0,2,4)(1,7,8)(3,9,10)(5,11,6)\rangle$
		$\langle(2,5,9)(3,6,8)(4,7,10),(0,2,4)(1,7,8)(3,9,11)(5,10,6)\rangle$
		$\langle(2,5,9)(3,6,8)(4,11,7),(0,2,4)(1,11,8)(3,9,10)(5,7,6)\rangle$
		$\langle(2,5,9)(3,6,8)(4,10,7),(0,2,4)(1,10,8)(3,9,11)(5,7,6)\rangle$
		$\langle(1,7,9)(2,6,5)(4,10,8),(0,1,8)(2,7,10)(3,5,9)(4,6,11)\rangle$
(14,7)	$C_7 \rtimes C_6$	$\langle(1,4,7)(2,11,5)(3,9,13)(6,8,12),(0,2)(1,11)(3,10)(4,8)(5,13)(6,7)(9,12))\rangle$
		$\langle(1,4,7)(2,11,5)(3,9,12)(6,8,13),(0,2)(1,11)(3,10)(4,8)(5,12)(6,7)(9,13))\rangle$
(16,16)	$(E_{16} \rtimes C_2) \rtimes C_3$	$\langle(1,8,9)(2,4,15,11,5,7)(3,12,6,10,13,14),$ $(0,4,7)(1,13,15)(2,3,10)(5,14,8)(6,9,12)\rangle$
(19,19)	$C_{19} \rtimes C_6$	$\langle(1,11,5,18,15,9)(2,7,12,8,3,13)(4,6,10,14,16,17),$ $(0,1,2,3,4,13,16,5,11,10,17,15,9,6,12,14,7,8,18)\rangle$

Table 7.8: Strength 3 PSCAs that are permutation groups for $12 \leqslant v \leqslant 19$.

(v, λ)	Group	Classes	Example
(6,1)	$\mathcal{S}_4$	1	$\langle(1,3)(4,5),(0,1,4)(2,5,3)\rangle$
(6,2)	$C_2 \times \mathcal{S}_4$	1	$\langle(0,5,2,1),(0,1,3,2,5,4)\rangle$
(7,7)	$PSL(3,2)$	9	$\langle(0,2,3,4,6,5,1),(0,5,4)(2,6,3)\rangle$
(8, 56)	$E_8 \rtimes PSL(3,2)$	22	$\langle(0,7,4,2,3,1,5),(0,1,3,4)(2,6,7,5)\rangle$
(9, 18)	$((E_9 \rtimes Q_8) \rtimes C_3) \rtimes C_2$	38	$\langle(0,8)(1,3)(4,5),(0,3,8)(1,6,4)(2,7,5)\rangle$
(10, 30)	$\mathcal{S}_6$	102	$\langle(0,7)(2,9)(3,4),(0,7,5,9,1)(2,3,6,8,4)\rangle$
	$A_6.C_2$	51	$\langle(0,7,3)(1,2,6)(4,8,5),(0,5,1,6,2,4,7,3)(8,9)\rangle$
(12, 18)	$((E_9 \rtimes Q_8) \rtimes C_3) \rtimes C_2$	24	$\langle(2,8)(3,11)(6,9)(7,10),(0,1,9)(2,4,11)(3,7,5)(6,10,8)\rangle$
(13, 234)	$PSL(3,3)$	77040	$\langle(3,9)(5,7)(8,10)(11,12),(0,1,2,3)(4,11,9,8)(5,12)(6,10)\rangle$
(21, 5040)	$PSL(3,4) \rtimes \mathcal{S}_3$	?	$\langle(0,16,8,9)(1,4,3,20,5,13,18,19)(2,6,10,7,14,17,11,12)$ $(0,17,7,11,10,5,4,19)(1,2,16,14,15,9,12,3)(8,13,18,20)\rangle$

Table 7.9: Strength 4 PSCAs that are permutation groups

gacy classes of subgroups of $\mathcal{S}_v$ and used the backtracking process described above in order to search over each conjugacy class. We have also performed ad hoc computations on some doubly transitive permutation groups. Some of those groups had too many conjugates to search exhaustively, so we randomly sampled conjugates instead. Our results are recorded in three tables. The first two, Table 7.7 and Table 7.8, record permutation groups that are strength 3 PSCAs but not strength 4 PSCAs. The second table, Table 7.9, records permutation groups that are strength 4 PSCAs but not strength 5 PSCAs. In Tables 7.7 and 7.8, a representative of each PSCA-isomorphism class of each group is presented. As a crosscheck, we note that these results agree with those presented in Table 7.4, which were found by a completely separate method. For reasons of space, in Table 7.9 we do not list representatives of each PSCA-isomorphism class. Rather, we just give the number of such classes (or a ? when random sampling of conjugates was used instead of an exhaustive search).

We know of few permutation groups that are PSCAs of strength 5, other than symmetric and alternating groups. These necessarily include the 5-transitive Mathieu groups M_{12} and M_{24}. Perhaps more interestingly, we also found that

$$\langle (1,7)(2,8)(3,4)(6,9), (0,2,10,6)(3,7,5,8) \rangle \tag{7.3}$$

is one of 108 presentations of the (4-transitive) Mathieu group M_{11} in $\mathcal{S}_{11}$ that form PSCAs of strength 5. No subgroup of $\mathcal{S}_{11}$ forms a PSCA of strength 4, other than those isomorphic to M_{11}, A_{11} or $\mathcal{S}_{11}$. Similarly,

$$\langle (2,11,8,6)(3,10,4,5), (0,1,2,3,4,5,11,6,7,10,8), (0,9)(1,8)(2,5)(3,6)(4,7)(10,11) \rangle \tag{7.4}$$

is one of 161 presentations of the (5-transitive) Mathieu group M_{12} in $\mathcal{S}_{12}$ that form PSCAs of strength 6. The presentations of M_{11} are conjugates of each other, and similarly for M_{12}. If we let $r \in \mathcal{S}_v$ be the reverse permutation, i.e. $r(i) = (v-1-i)$ for $0 \leqslant i \leqslant v-1$, then for a permutation group $G \leqslant \mathcal{S}_v$, we will find that G and rGr are isomorphic in terms of sequence coverage. Hence it is plausible that we may find presentations of the same group that are isomorphic as PSCAs. Indeed, this is the case for M_{11} where the 108 presentations that form PSCAs of strength 5 can be reduced to 54 isomorphism classes. Meanwhile, the presentation of M_{12} given in (7.4) is the only one of the 161 strength 6 PSCAs for which conjugation by r leaves the underlying set of permutations unchanged. Thus, these 161 presentations that form PSCAs of strength 6 reduce to 81 isomorphism classes.

For the larger Mathieu groups we were unable to do exhaustive computations and again relied on random sampling. We found that

$$\begin{aligned} \langle &(0,1,20,4,2)(3,8,9,12,13)(5,16,10,11,18)(6,7,15,19,14), \\ &(0,13,16,5,10)(1,14,19,4,2)(3,18,7,12,15)(9,21,11,20,17) \rangle \end{aligned} \tag{7.5}$$

is a presentation of the (3-transitive) Mathieu group M_{22} in $\mathcal{S}_{22}$ that forms a PSCA of strength 5. Also

$$\begin{aligned}&\langle(0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22),\\&(0,23)(1,22)(2,11)(3,15)(4,17)(5,9)(6,19)(7,13)(8,20)(10,16)(12,21)(14,18),\\&(2,16,9,6,8)(3,12,13,18,4)(7,17,10,11,22)(14,19,21,20,15)\rangle\end{aligned} \tag{7.6}$$

is a presentation of the (5-transitive) Mathieu group M_{24} in $\mathcal{S}_{24}$ that forms a PSCA of strength 6. Its point stabilisers provide PSCAs of strength 5 in $\mathcal{S}_{23}$ that are presentations of M_{23}.

Table 7.8 also includes the $\mathsf{PSCA}(16,3,16)$ found earlier in the section, as well as a $\mathsf{PSCA}(19,3,19)$. Exhaustive searches were not undertaken for either of these parameter sets. However, a partial search found 17116 and 232 isomorphism classes, respectively, of $\mathsf{PSCA}(16,3,16)$ and $\mathsf{PSCA}(19,3,19)$ that are conjugate to the examples given in the table. Note that since isomorphism includes the option to freely permute symbols, the only material effect of conjugation in this context is to permute the columns of a PSCA.

A striking feature of results summarised in Table 7.7, Table 7.8 and Table 7.9 is that there are a number of cases of non-isomorphic PSCAs being produced by similar sets of generators. For example, starting from the $\mathsf{PSCA}(6,4,1)$, if we conjugate the generating set by the transposition $(2,5)$ we reach a $\mathsf{PSCA}(6,3,4)$. A similar thing happens if we use the transposition $(4,5)$. Conjugating the generating set by a transposition has the effect of interchanging two columns of the PSCA (and then exchanging two symbols to once again achieve the property of having one row equal to the identity permutation).

Summarising our bounds on $\mathsf{PSCAN}(v,t)$ derived from group presentations, we have:

Theorem 7.14.

- *For* $v \leqslant 11$, *we have* $\mathsf{PSCAN}(v,5) \leqslant 66$.
- *For* $v \leqslant 12$, *we have* $\mathsf{PSCAN}(v,4) \leqslant 18$, $\mathsf{PSCAN}(v,5) \leqslant 792$ *and* $\mathsf{PSCAN}(v,6) \leqslant 132$.
- *For* $v \leqslant 13$, *we have* $\mathsf{PSCAN}(v,4) \leqslant 234$.
- *For* $v \leqslant 21$, *we have* $\mathsf{PSCAN}(v,4) \leqslant 5040$.
- *For* $v \leqslant 22$, *we have* $\mathsf{PSCAN}(v,5) \leqslant 3696$ *and hence* $\mathsf{PSCAN}(v,4) \leqslant 18\,480$.
- *For* $v \leqslant 23$, *we have* $\mathsf{PSCAN}(v,5) \leqslant 85\,008$ *and hence* $\mathsf{PSCAN}(v,4) \leqslant 425\,040$.
- *For* $v \leqslant 24$, *we have* $\mathsf{PSCAN}(v,6) \leqslant 340\,032$ *and hence* $\mathsf{PSCAN}(v,5) \leqslant 2\,040\,192$ *and* $\mathsf{PSCAN}(v,4) \leqslant 10\,200\,960$.

Proof. Examples of a $\mathsf{PSCA}(12,4,18)$, a $\mathsf{PSCA}(13,4,234)$ and a $\mathsf{PSCA}(21,4,5040)$ are given in Table 7.9. Also, we gave a $\mathsf{PSCA}(11,5,66)$ in (7.3), a $\mathsf{PSCA}(12,6,132)$ in (7.4), a $\mathsf{PSCA}(22,5,3696)$ in (7.5) and a $\mathsf{PSCA}(24,6,340\,032)$ in (7.6), from which we derived a $\mathsf{PSCA}(23,5,85\,008)$. □

Chapter 8

Constructing Perfect Sequence Covering Arrays from Projectivities of a Finite Projective Plane

In this chapter, we present a construction that is capable of building a $\mathsf{PSCA}(v, t, \lambda)$ for any $v \geqslant t \geqslant 4$. The main ingredient of this construction is the permutation group $\mathrm{PGL}(n+1, q)$ of projectivities of the finite projective geometry $\mathrm{PG}(n, q)$. The size of the PSCAs built from this construction grows polynomially in v.

We also consider in greater detail permutation representations of the group $\mathrm{PGL}(3, q)$. We explore the connections between difference sets and representations of $\mathrm{PGL}(3, q)$ that give near-perfect 4-sequence coverage. That is, we find permutation representations of $\mathrm{PGL}(3, q)$ for which all but a small fraction of 4-sequences are covered the same number of times. These permutation groups give rise to a set of permutations whose size as a function of v is much smaller than our general construction.

8.1 Introduction

Recall that for positive integers v and t with $v \geqslant t$, we let $[v] = \{0, \ldots, v-1\}$, $\mathcal{S}_v$ be the group of permutations of $[v]$, and $\mathcal{S}_{v,t}$ be the set of ordered sequences of t distinct elements of $[v]$. Unless stated otherwise, permutations are assumed to be written in one-line notation with $\pi \in \mathcal{S}_v$ being denoted by $\pi(0)\pi(1)\cdots, \pi(v-1)$. Additionally, we write $s \in \mathcal{S}_{v,t}$ as $s = (s_0, \ldots, s_{t-1})$. For $\pi \in \mathcal{S}_v$ and $s \in \mathcal{S}_{v,t}$ we say that π *covers* s if $\pi^{-1}(s_i) < \pi^{-1}(s_{i+1})$ for $0 \leqslant i \leqslant t-2$.

A *perfect sequence covering array* with order v, strength t and multiplicity λ, denoted by $\mathsf{PSCA}(v, t, \lambda)$, is a multiset X of permutations in $\mathcal{S}_v$ such that every sequence in $\mathcal{S}_{v,t}$ is covered by exactly λ permutations in X. If T is a t-subset of $[v]$, then there are $t!$ ways of arranging the elements of T, each of which forms a sequence in $\mathcal{S}_{v,t}$ that must be covered by λ permutations in a $\mathsf{PSCA}(v, t, \lambda)$. Furthermore, every permutation in a $\mathsf{PSCA}(v, t, \lambda)$

covers exactly one of these sequences, so a $\mathsf{PSCA}(v,t,\lambda)$ must contain $t!\lambda$ permutations.

For $v \geqslant t$, let $\mathsf{PSCAN}(v,t)$ be the smallest positive integer λ such that a $\mathsf{PSCA}(v,t,\lambda)$ exists. Observe that $\mathcal{S}_v$ is a $\mathsf{PSCA}(v,t,v!/t!)$, so $\mathsf{PSCAN}(v,t)$ exists for all $v \geqslant t$ and $\mathsf{PSCAN}(v,t) \leqslant v!/t!$. Recall that by Lemma 4.1, $\mathsf{PSCAN}(v,t) \geqslant \mathsf{PSCAN}(v-1,t)$ for $v > t$ and $\mathsf{PSCAN}(v,t) \geqslant \mathsf{PSCAN}(v,t-1)/t$ for $t > 3$.

In this chapter, we present an explicit construction of a $\mathsf{PSCA}(v,t,\lambda)$ for all $v \geqslant t \geqslant 4$. The method of this construction involves taking a suitable permutation representation of the group $\mathrm{PGL}(t-1,q)$ of projectivities of the projective geometry $\mathrm{PG}(t-2,q)$. We show that in such a permutation representation, there is a subset of $q+1$ symbols such that any t-sequence of symbols from this subset is covered by λ permutations for a given constant λ. Hence, deleting all but the symbols in this subset forms a $\mathsf{PSCA}(q+1,t,\lambda)$. This construction yields the following upper bound on $\mathsf{PSCAN}(v,t)$.

Theorem 8.1. *For $v \geqslant t \geqslant 4$,*

$$\mathsf{PSCAN}(v,t) < \frac{(2v)^{(t-1)^2}}{t!(v-1)}.$$

This bound is derived through purely constructive means however, a probabilistic upper bound does exist. A *t-wise uniform set of permutations* is a set $T \subseteq \mathcal{S}_v$ such that for any $a,b \in \mathcal{S}_{v,t}$,

$$\frac{1}{|T|}|\{\pi \in T : \pi(a_i) = b_i, 1 \leqslant i \leqslant t\}| = \frac{(v-t)!}{v!}.$$

A t-wise uniform set of permutations $T \subseteq \mathcal{S}_v$ is also a $\mathsf{PSCA}(v,t,|T|/t!)$. Kuperberg, Lovett and Peled [35] proved that for any $t \leqslant v$, there is a t-wise uniform set of permutations $T \subseteq \mathcal{S}_v$ with $|T| \leqslant (cv)^{ct}$ for some universal constant $c > 0$. Although this result gives a tighter bound on $\mathsf{PSCAN}(v,t)$ than Theorem 8.1, it is not yet known how to efficiently construct either a PSCA or a t-wise uniform set of permutations with this size.

While this thesis was being produced, Iurlano [28] established an equivalence between PSCAs of strength t and families of *t-rankwise independent permutations*. Iurlano also uses a construction of Itoh, Takei and Tarui [26] of t-rankwise independent permutations to build PSCAs with $v^{O(t^2/\ln t)}$ permutations.

Our construction can also be applied when $t = 3$. However, an infinite family of $\mathsf{PSCA}(v,3,\lambda)$ built by Yuster [64] established that $\mathsf{PSCAN}(v,3) \leqslant cv(\log v)^{\log 7}$ for an absolute constant c. This result provides a tighter bound on $\mathsf{PSCAN}(v,3)$ than Theorem 8.1 would, were it to be extended to the $t = 3$ case.

In proving Theorem 8.1, we show that sequences of t points of $\mathrm{PG}(t-2,q)$ belonging to a particular family are covered by a constant number of permutations in a representation of $\mathrm{PGL}(t-1,q)$. In the $t = 4$ case, we can choose a particular representation of $\mathrm{PGL}(3,q)$ to ensure that sequences of four points of $\mathrm{PG}(2,q)$ belonging to a separate family are also covered by the same constant number of permutations. Although accounting for this new family of sequences does not provide a substantial improvement to the bound $\mathsf{PSCAN}(v, 4) = O(v^8)$ implied by Theorem 8.1, it does prove the following theorem.

Theorem 8.2. *Let q be a prime power and let $r = q^2+q+1$. Then there is a permutation representation $\Psi \leqslant \mathcal{S}_r$ of $\mathrm{PGL}(3, q)$ such that the number of sequences in $\mathcal{S}_{r,4}$ that are covered by exactly $|\Psi|/4!$ permutations in Ψ is greater than*

$$\left(1 - \frac{1}{q^2 + 5q + 11}\right) |\mathcal{S}_{r,4}|.$$

In building a $\mathsf{PSCA}(v, 4, \lambda)$, we take a prime power q such that $q \geqslant v$, find a suitable permutation representation of $\mathrm{PGL}(3, q)$ in $\mathcal{S}_{q^2+q+1}$ and then delete all but v symbols from each permutation in this group. As a consequence of this symbol deletion, the number of permutations in the resulting PSCA on v symbols is approximately v^8. Theorem 8.2 implies that it is possible to find a set of permutations in $\mathcal{S}_v$ with size approximately v^4 such that the vast majority of 4-sequences are covered by a constant number of permutations. This reduced size is much closer to the lower bound proved by Yuster [64], which says that $\mathsf{PSCAN}(v, 4) \geqslant v(v-3)/48$. However, as the permutation representation presented in Theorem 8.2 may or may not be a PSCA, it is still unclear what the asymptotic behaviour of $\mathsf{PSCAN}(v, 4)$ should be.

8.2 Preliminaries

We begin by recalling some definitions regarding group actions. For a set X, let $\mathrm{Sym}(X)$ denote the group of permutations of X. Note that when $X = [v]$, $\mathrm{Sym}(X) = \mathcal{S}_v$. An *action* of a group G on X is a homomorphism $\phi : G \to \mathrm{Sym}(X)$. For $g \in G$ and $x \in X$, we use gx to refer to the image of x under the permutation $\phi(g)$. The *orbit* of x is the set $\mathrm{Orb}(x) = \{gx : g \in G\}$. The *stabiliser* of x is the set $\mathrm{Stab}(x) = \{g \in G : gx = x\}$. The stabiliser of x forms a subgroup of G. In what follows, we make use of the Orbit-Stabiliser Theorem.

Theorem 8.3. *If G is a group acting on X, then for any $x \in X$,*

$$|\mathrm{Orb}(x)||\mathrm{Stab}(x)| = |G|.$$

A permutation group $G \leqslant \mathcal{S}_v$ has the following natural action on $\mathcal{S}_{v,t}$. If $g \in G$ and $s \in \mathcal{S}_{v,t}$, then $gs = (g(s_0), \dots, g(s_{t-1}))$. Consider an array A with columns indexed by $[v]$ and rows indexed by the elements of G where $\mathsf{A}[g, i] = g(i)$. Let $s \in \mathcal{S}_{v,t}$ and consider the corresponding sequence of columns of A. In row g and in columns $(s_0, \dots, s_{t-1})$ of A, we find the sequence $(g(s_0), \dots, g(s_{t-1})) = gs$. So the sequences that appear in the columns $(s_0, \dots, s_{t-1})$ of A are exactly those in $\mathrm{Orb}(s)$. For $x \in \mathrm{Orb}(s)$, the set of rows of A in which the sequence x appears in the columns $(s_0, \dots, s_{t-1})$ is $\{g : gs = x\}$. This set is a coset of $\mathrm{Stab}(s)$ so it must have the same size as $\mathrm{Stab}(s)$. The permutation g covers gs if and only if $s_0 < \cdots < s_{t-1}$. Let $\mathrm{Asc}(s) = \{x \in \mathrm{Orb}(s) : x_1 < \cdots < x_t\}$. We now have the following lemma.

Lemma 8.4. *If $G \leqslant \mathcal{S}_v$ is a permutation group and $s \in \mathcal{S}_{v,t}$, then the number of permutations in G that cover s is $|\mathrm{Asc}(s)||\mathrm{Stab}(s)|$.*

To conclude this section, we consider deleting symbols from permutations. For $\pi \in \mathcal{S}_v$ and $j \in [v]$, we define $\pi_{[j]}$ to be the permutation in $\mathcal{S}_j$ obtained by deleting the symbols $\{j, j+1, \ldots, v-1\}$ from π. The permutation $\pi_{[j]}$ covers a sequence $s \in \mathcal{S}_{j,t}$ if and only if π also covers s. If X is a multiset of permutations in $\mathcal{S}_v$, then for $j \in [v]$, we define $X_{[j]}$ to be the multiset $\{\pi_{[j]} : \pi \in X\}$. Then, for any sequence $s \in \mathcal{S}_{j,t}$, the number of permutations in $X_{[j]}$ that cover s is equal to the number of permutations in X that cover s. In the next section we construct a PSCA of strength t by deleting symbols from a suitable multiset of permutations.

8.3 Collineations of projective spaces

In this section we prove Theorem 8.1. We begin by introducing some definitions regarding projective spaces. Let $q = p^m$ for some prime p and for some integer $m \geqslant 1$ and let $\mathrm{GF}(q)$ be the field with q elements. Now let $n \geqslant 2$ and let V be an $(n+1)$-dimensional vector space over $\mathrm{GF}(q)$. Then the *n-dimensional projective space* over $\mathrm{GF}(q)$, denoted by $\mathrm{PG}(n, q)$, is the set of all 1-dimensional subspaces of V. The elements of $\mathrm{PG}(n, q)$ are called *points*. If W is a subspace of V, then W forms a set of points in $\mathrm{PG}(n, q)$ with W containing the point X if and only if X is a subspace of W. A 2-dimensional subspace of V forms a *line* in $\mathrm{PG}(n, q)$ and an n-dimensional subspace of V forms a *hyperplane* in $\mathrm{PG}(n, q)$. A *collineation* of $\mathrm{PG}(n, q)$ is a permutation of the points of $\mathrm{PG}(n, q)$ that maps lines to lines. Let $A \in \mathrm{GL}(n+1, q)$ be a non-singular matrix and suppose that $Au = v$ for vectors $u \in X$ and $v \in Y$ where X and Y are points in $\mathrm{PG}(n, q)$. Then $A(cu) = cv$ for $c \in \mathrm{GF}(q)$. Thus, every vector in X is mapped by A to a vector in Y. Hence, A induces a permutation of the points of $\mathrm{PG}(n, q)$. Permutations formed in this way are called *projectivities*. The set of all projectivities of $\mathrm{PG}(n, q)$ forms the group $\mathrm{PGL}(n+1, q)$. A *frame* of $\mathrm{PG}(n, q)$ is an ordered sequence of $n+2$ points in $\mathrm{PG}(n, q)$ such that no $n+1$ of these points lie in the same hyperplane of $\mathrm{PG}(n, q)$. The following theorem is a statement of the Fundamental Theorem of Projective Geometry (see e.g. [22]).

Theorem 8.5. *For any two frames in $\mathrm{PG}(n, q)$, there is a unique projectivity of $\mathrm{PG}(n, q)$ mapping one frame to the other.*

Let r be the number of points in $\mathrm{PG}(n, q)$. Projectivities are defined as permutations of the points of $\mathrm{PG}(n, q)$ but we can view projectivities as permutations of $[r]$ by labelling the points of $\mathrm{PG}(n, q)$. For a bijection $\psi : \mathrm{PG}(n, q) \to [r]$ and a projectivity f, define

$f_\psi \in \mathcal{S}_r$ by $f_\psi(i) = \psi(f(\psi^{-1}(i)))$. Then let $\Psi := \{f_\psi : f \in \mathrm{PGL}(n+1,q)\}$. Note Ψ is a permutation subgroup of $\mathcal{S}_r$. The order of Ψ is given by

$$|\Psi| = |\mathrm{PGL}(n+1,q)| = \frac{\prod_{i=0}^{n}(q^{n+1}-q^i)}{q-1}.$$

By establishing the bijection ψ, points of $\mathrm{PG}(n,q)$ are associated with elements of $[r]$ and so we can treat lines and hyperplanes as subsets of $[r]$ and frames as sequences in $\mathcal{S}_{r,n+2}$.

Lemma 8.6. *For a bijection $\psi : \mathrm{PG}(n,q) \to [r]$ and a frame $s \in \mathcal{S}_{r,n+2}$, the number of permutations in Ψ that cover s is $|\Psi|/(n+2)!$.*

Proof. By Theorem 8.5, every frame in $\mathcal{S}_{r,n+2}$ is part of the same orbit under the action of Ψ. For any frame $s \in \mathcal{S}_{r,n+2}$, any reordering of the points of s will form another frame. Of all the $(n+2)!$ ways of ordering the points of s, only one of these sequences is in ascending order. Thus, $|\mathrm{Asc}(s)| = |\mathrm{Orb}(s)|/(n+2)!$. Therefore, by Lemma 8.4, the number of permutations in Ψ that cover s is

$$|\mathrm{Stab}(s)||\mathrm{Asc}(s)| = \frac{|\mathrm{Stab}(s)||\mathrm{Orb}(s)|}{(n+2)!} = \frac{|\Psi|}{(n+2)!}. \qquad \square$$

A *k-arc* in $\mathrm{PG}(n,q)$ is a set of k points in $\mathrm{PG}(n,q)$, no $n+1$ of which lie in a hyperplane of $\mathrm{PG}(n,q)$.

Theorem 8.7 (e.g. [2])**.** *For $q \geqslant n$, there exists a $(q+1)$-arc in $\mathrm{PG}(n,q)$.*

Lemma 8.8. *For $q \geqslant n+1$, if $\psi : \mathrm{PG}(n,q) \to [r]$ is a bijection such that $\{\psi^{-1}(i) : i \in [q+1]\}$ is a $(q+1)$-arc, then $\Psi_{[q+1]}$ is a $\mathsf{PSCA}(q+1, n+2, |\Psi|/(n+2)!)$.*

Proof. Let $s \in \mathcal{S}_{q+1,n+2}$ and let ψ be as defined in the lemma statement. Then s is a frame. Thus, by Lemma 8.6, s is covered by $|\Psi|/(n+2)!$ permutations in Ψ. Hence, s is covered by $|\Psi|/(n+2)!$ permutations in $\Psi_{[q+1]}$. Therefore, $\Psi_{[q+1]}$ is a $\mathsf{PSCA}(q+1, n+2, |\Psi|/(n+2)!)$. $\square$

Note that for $q \geqslant n+1$, Theorem 8.7 guarantees the existence of a bijection ψ satisfying the condition of Lemma 8.8. We are now ready to prove Theorem 8.1

Proof of Theorem 8.1. Let q be the smallest power of 2 such that $q \geqslant v$. Then $q < 2v$. Let $n = t-2$. Then, by Lemma 8.8,

$$\mathsf{PSCAN}(v,t) \leqslant \frac{|\mathrm{PGL}(n+1,q)|}{(n+2)!} = \frac{\prod_{i=0}^{t-2}(q^{t-1}-q^i)}{t!(q-1)} < \frac{(2v)^{(t-1)^2}}{t!(v-1)}. \qquad \square$$

8.4 Almost perfect 4-sequence covering arrays

We now focus specifically on the case where $n = 2$. That is, we consider sequences of four points in $\mathrm{PG}(2,q)$. In the previous section, we present a construction of a $\mathsf{PSCA}(v, 4, \lambda)$

with $O(v^8)$ permutations. In this section, we adjust this construction such that the size of the new construction is $O(v^4)$. In this new construction we can guarantee that almost all sequences are covered by the same number of permutations (see Theorem 8.2). This shows that we can greatly reduce the size of the construction in Section 3 when $t = 4$ while still ensuring that most sequences are covered by the same number of permutations.

The sequences now under consideration are those containing four points of $\mathrm{PG}(2, q)$. These sequences can divided into three families. The first family contains all sequences of four points such that no three are collinear. As hyperplanes and lines are the same in $\mathrm{PG}(2, q)$, these sequences are frames. By Lemma 8.6, in any permutation representation $\Psi \leqslant \mathcal{S}_r$ of $\mathrm{PGL}(3, q)$ as defined in Chapter 8.3, any frame is covered by $|\Psi|/24$ permutations. The second family contains all sequences with three collinear points but not four. The third family contains all sequences of four collinear points.

Let $r = q^2+q+1$ be the number of points in $\mathrm{PG}(2, q)$. A *planar difference set* of $\mathbb{Z}_r$ is a set $A = \{a_0, a_1, \ldots, a_q\} \subset \mathbb{Z}_r$ such that for any non-zero element $x \in \mathbb{Z}_r$, there are unique i and j such that $a_i - a_j = x$. For any prime power q, it is possible to label the points of $\mathrm{PG}(2, q)$ such that the labelled lineset has the form $\{\{a_i + j : i \in [q + 1]\} : j \in \mathbb{Z}_r\}$, where $A = \{a_0, \ldots, a_q\}$ is a planar difference set of $\mathbb{Z}_r$ and addition is performed modulo r [50]. Let ψ be such a labelling. In this section we show that in the corresponding permutation representation $\Psi \leqslant \mathcal{S}_r$ of $\mathrm{PGL}(3, q)$, each sequence containing at most three collinear points is covered by $|\Psi|/24$ permutations. This in turn proves Theorem 8.2.

The consequence of this result is that for fixed v, we can find a much smaller multiset of permutations than the multiset built in Chapter 8.3 such that the vast majority of sequences are covered by a constant number of permutations. This prompts two immediate questions. The first is whether the bound in Theorem 8.1 can be reduced specifically when $t = 4$. Indeed, in Chapter 8.3, our construction relied on reducing the point set to a subset where no three points lay on the same hyperplane. Here, it seems we can relax that condition to one that requires that no four points lie on the same line. A (k, d)-*arc* in $\mathrm{PG}(2, q)$ is a set of k points in $\mathrm{PG}(2, q)$, no $d + 1$ of which lie on the same line. Thus, we can build a PSCA of strength 4 by taking the representation Ψ and deleting all but a subset of symbols that form a $(k, 3)$-arc. The size of a $(k, 3)$-arc in $\mathrm{PG}(2, q)$ is at most twice the size of the arc used in the construction in Chapter 8.3 (i.e. $2(q+1)$) [55]. Hence, adapting the construction in Chapter 8.3 would only yield an improvement on the bound in Theorem 8.1 by a constant factor of at most 2^8 when $t = 4$.

The second question is whether the permutation representation Ψ may actually form a PSCA by covering sequences containing four collinear points with the same constant number of permutations. A computer search was performed over planar difference sets of $\mathbb{Z}_r$ for prime power orders $4 \leqslant q \leqslant 25$. This search found no representations of $\mathrm{PGL}(3, q)$ that also formed a PSCA of strength 4. However, a representation of $\mathrm{P\Gamma L}(3, 4)$ which forms a PSCA was presented by Gentle and Wanless [20]. The group $\mathrm{P\Gamma L}(3, q)$ is the product of $\mathrm{PGL}(3, q)$ with the group of field automorphisms of $\mathrm{GF}(q)$. In particular,

$\mathrm{P\Gamma L}(3,4)$ is twice as large as $\mathrm{PGL}(3,4)$. Moreover, Gentle and Wanless [20] also present examples of representations of $\mathrm{PGL}(3,2)$ and $\mathrm{PGL}(3,3)$ that form PSCAs of strength 4.

The remainder of this section is devoted to a series of lemmas that collectively prove Theorem 8.2. Consider sequences that contain three collinear points but not four. For $i \in \{0,1,2,3\}$, let T_i be the set of sequences s of four points in $\mathrm{PG}(2,q)$ for which there exists a line ℓ such that s_i does not lie on ℓ, but s_k does for $k \in \{0,1,2,3\} \setminus \{i\}$. By definition, if $s \in T_i$, then the points $\{s_k : k \in \{0,1,2,3\} \setminus \{i\}\}$ are collinear. However, if $j \neq i$ and $s' \in T_j$, then the points $\{s'_k : k \in \{0,1,2,3\} \setminus \{i\}\}$ are not collinear. Therefore, s' is not in the orbit of s under the action of $\mathrm{PGL}(3,q)$. The following lemma addresses the case where s and s' both belong to T_i.

Lemma 8.9. *If s and s' are sequences in T_i for some $i \in \{0,1,2,3\}$ then there exists a projectivity of $\mathrm{PG}(2,q)$ that maps s to s'.*

Proof. Let $\{i,j,k,\ell\} = \{0,1,2,3\}$ and let $s_i = a$, $s_j = b$, $s_k = c$, $s_\ell = d$ and $s'_i = a'$, $s'_j = b'$, $s'_k = c'$, $s'_\ell = d'$. Hence, both $\{a,b,c\}$ and $\{a',b',c'\}$ are sets of non-collinear points. As such, the non-zero vectors $u \in a$, $v \in b$ and $w \in c$ form a basis of V, as do the non-zero vectors $u' \in a'$, $v' \in b'$ and $w' \in c'$. We can find a matrix $A \in \mathrm{GL}(3,q)$ such that $Au = u'$, $Av = v'$ and $Aw = w'$. Let f be the projectivity of $\mathrm{PG}(2,q)$ induced by A. Then $f(a) = a'$, $f(b) = b'$ and $f(c) = c'$. Let $x \in d$. Since b, c and d are collinear, $x = \alpha v + \beta w$ for some non-zero $\alpha, \beta \in \mathrm{GF}(q)$. Hence, $Ax = \alpha v' + \beta w'$. Let $x' \in d'$ and similarly note that $x' = \alpha' v' + \beta' w'$ for some non-zero $\alpha', \beta' \in \mathrm{GF}(q)$. We can then find a matrix $B \in \mathrm{GL}(3,q)$ such that $Bu' = u'$, $Bv' = \alpha^{-1}\alpha' v'$ and $Bw' = \beta^{-1}\beta' w'$. Then, $BAx = \alpha' v' + \beta' w' = x'$. Let g be the projectivity induced by B and observe $g \circ f(a) = a'$, $g \circ f(b) = b'$, $g \circ f(c) = c'$ and $g \circ f(d) = d'$. Therefore, the projectivity $g \circ f$ maps s to s'. □

Now we choose some bijection $\psi : \mathrm{PG}(2,q) \to [r]$ and consider the corresponding permutation group $\Psi \leqslant \mathcal{S}_r$. Let L be the subsets of $[r]$ corresponding to lines of $\mathrm{PG}(2,q)$. Then each line $\ell \in L$ can be represented by $\ell = \{\ell_0, \dots, \ell_q\} \subseteq [r]$ where $\ell_0 < \ell_1 < \cdots < \ell_q$. Our next goal is to find $|\mathrm{Asc}(s)|$ for $s \in T_i$. First consider T_0. Let $\ell \in L$ and let $i \in [q+1]$. There are $\binom{q-i}{2}$ 3-subsets of ℓ that have ℓ_i as their minimal element and there are $\ell_i - i$ points less than ℓ_i that do not lie on ℓ. Therefore, for $s \in T_0$,

$$|\mathrm{Asc}(s)| = \sum_{\ell \in L} \sum_{i=0}^{q} \binom{q-i}{2} (\ell_i - i). \tag{E1}$$

Next, consider T_1. Suppose we build a sequence in T_1 containing three points on ℓ such that ℓ_i is the minimum of these points. Then, for $j > i$, there are $q - j$ points on ℓ greater than ℓ_j and $(\ell_j - j) - (\ell_i - i)$ points between ℓ_i and ℓ_j that do not lie on ℓ.

Therefore, for $s \in T_1$,

$$\begin{aligned}
|\mathrm{Asc}(s)| &= \sum_{\ell \in L} \sum_{0 \leqslant i < j \leqslant q} (q-j)((\ell_j - j) - (\ell_i - i)) \\
&= \sum_{\ell \in L} \left(\sum_{j=1}^{q} j(q-j)(\ell_j - j) - \sum_{i=0}^{q-1} \sum_{j=i+1}^{q} (q-j)\ell_i + \sum_{i=0}^{q-1} \sum_{j=i+1}^{q} (q-j)i \right) \\
&= \sum_{\ell \in L} \left(\sum_{j=1}^{q} \left(j(q-j)\ell_j - j^2(q-j)\right) - \sum_{i=0}^{q-1} \left(\binom{q-i}{2}\ell_i - \binom{q-i}{2} i \right) \right) \\
&= \sum_{\ell \in L} \sum_{i=0}^{q} \left(i(q-i) - \binom{q-i}{2} \right) (\ell_i - i). \qquad \text{(E2)}
\end{aligned}$$

Next, consider T_2. Suppose we build a sequence in T_2 containing three points on ℓ such that ℓ_j is the maximum of these points. Then, for $i < j$, there are i points on ℓ less than ℓ_i and $(\ell_j - j) - (\ell_i - i)$ points between ℓ_i and ℓ_j that do not lie on ℓ. Therefore, for $s \in T_2$,

$$\begin{aligned}
|\mathrm{Asc}(s)| &= \sum_{\ell \in L} \sum_{0 \leqslant i < j \leqslant q} i((\ell_j - j) - (\ell_i - i)) \\
&= \sum_{\ell \in L} \left(\sum_{j=1}^{q} \sum_{i=0}^{j-1} i\ell_j - \sum_{j=1}^{q} \sum_{i=0}^{j-1} ij - \sum_{i=0}^{q-1} \sum_{j=i+1}^{q} i\ell_i + \sum_{i=0}^{q-1} \sum_{j=i+1}^{q} i^2 \right) \\
&= \sum_{\ell \in L} \left(\sum_{j=1}^{q} \left(\binom{j}{2}\ell_j - \binom{j}{2} j \right) - \sum_{i=0}^{q-1} \left(i(q-i)\ell_i - i^2(q-i)\right) \right) \\
&= \sum_{\ell \in L} \sum_{i=0}^{q} \left(\binom{i}{2} - i(q-i) \right) (\ell_i - i). \qquad \text{(E3)}
\end{aligned}$$

Finally, consider T_3. Let $\ell \in L$ and let $i \in [q+1]$. There are $\binom{i}{2}$ 3-subsets of ℓ whose maximum is ℓ_i and there are $q^2 + q - \ell_i - (q - i)$ points greater ℓ_i that do not lie on ℓ. Therefore, for $s \in T_3$,

$$|\mathrm{Asc}(s)| = \sum_{\ell \in L} \sum_{i=0}^{q} \binom{i}{2} (q^2 - (\ell_i - i)). \qquad \text{(E4)}$$

Therefore,

$$\begin{aligned}
\text{(E1)} + \text{(E2)} + \text{(E3)} + \text{(E4)} &= \sum_{\ell \in L} \sum_{i=0}^{q} \binom{i}{2} q^2 \\
&= \frac{(q^2+q+1)q^3(q+1)(q-1)}{6}. \qquad \text{(E5)}
\end{aligned}$$

Later in this section, we will prove the existence of a particular representation Ψ such that (E1) = (E2) = (E3) = (E4). Our first step is to equate (E1) + (E4) and (E2) +

(E3) for which we require the following lemma.

Lemma 8.10. *Let L be the lineset of* PG$(2,q)$ *as subsets of* $[r]$. *Then,*

$$\sum_{\ell\in L}\sum_{i=0}^{q} i\ell_i = \frac{(q^2+q)(q^2+q+1)(2q^2+2q+1)}{6}$$

Proof. Let $j \in [r]$ and let $J = \{(i,\ell) : \ell_i = j\}$. Consider $\sum_{(i,\ell)\in J} i\ell_i$. For each pair $(i,\ell) \in J$, $\ell_i = j$ and i is the number of points less than j that lie on ℓ. There are j elements of $[r]$ less than j, each of which must lie on the same line as j exactly once. Therefore,

$$\sum_{(i,\ell)\in J} i\ell_i = j\sum_{(i,\ell)\in J} i = j^2.$$

Therefore,

$$\sum_{\ell\in L}\sum_{i=0}^{q} i\ell_i = \sum_{j=0}^{q^2+q} j^2 = \frac{(q^2+q)(q^2+q+1)(2q^2+2q+1)}{6}$$ □

Lemma 8.11. *In any representation Ψ,* (E1) + (E4) = (E2) + (E3) = (E5)/2.

Proof. Using the expressions (E2) and (E3),

$$\begin{aligned}(E2)+(E3) &= \sum_{\ell\in L}\sum_{k=0}^{q}\left(\binom{k}{2}-\binom{q-k}{2}\right)(\ell_k-k)\\ &= \frac{1}{2}\sum_{\ell\in L}\sum_{k=0}^{q}\left(2k(q-1)-(q^2-q)\right)(\ell_k-k).\end{aligned}$$

First, consider

$$\sum_{\ell\in L}\sum_{k=0}^{q}(q^2-q)(\ell_k-k).$$

The sum $\sum_{\ell\in L}\sum_{k=0}^{q}\ell_k$ is just the sum of all the points of every line of PG$(2,q)$. Each point appears on $q+1$ lines so this is equal to $\binom{q^2+q+1}{2}(q+1)$. Next, $\sum_{\ell\in L}\sum_{k=0}^{q} k = (q^2+q+1)\binom{q+1}{2}$. Therefore,

$$\begin{aligned}\sum_{\ell\in L}\sum_{k=0}^{q}(q^2-q)(\ell_k-k) &= (q^2-q)\left(\binom{q^2+q+1}{2}(q+1)-(q^2+q+1)\binom{q+1}{2}\right)\\ &= (q^2+q+1)(q^2-q)(q+1)\left(\frac{q^2+q}{2}-\frac{q}{2}\right)\\ &= \frac{1}{2}(q^2+q+1)(q^3-q)q^2.\end{aligned}$$

Next, consider

$$\sum_{\ell\in L}\sum_{k=0}^{q}k(\ell_k-k).$$

By Lemma 8.10, this is equal to

$$\begin{aligned}&\frac{1}{6}\left((q^2+q)(q^2+q+1)(2q^2+2q+1)-(q^2+q+1)q(q+1)(2q+1)\right)\\&=\frac{(q^2+q+1)(q^2+q)q^2}{3}.\end{aligned}$$

Therefore,

$$\begin{aligned}(\text{E2})+(\text{E3})&=\frac{1}{2}\left(\frac{2(q^2+q+1)(q^3-q)q^2}{3}-\frac{(q^2+q+1)(q^3-q)q^2}{2}\right)\\&=\frac{(q^2+q+1)(q^3-q)q^2}{12}\\&=(\text{E5})/2.\end{aligned}$$

As (E1) + (E2) + (E3) + (E4) = (E5), it must also be true that (E1) + (E4) = (E5)/2. □

Lemma 8.11 applies generally to any representation of $\mathrm{PGL}(3,q)$, but now we will choose a specific representation according to a planar difference set of $\mathbb{Z}_r$. Recall that a planar difference set is a set $A=\{a_0,a_1,\ldots,a_q\}\subset\mathbb{Z}_r$ such that for any non-zero element $x\in\mathbb{Z}_r$, there are unique i and j such that $a_i-a_j=x$. For the rest of this section, we choose ψ such that the lines of $\mathrm{PG}(2,q)$ are mapped to sets of the form $\{a_i+j : i\in[q+1]\}$ for $j\in\mathbb{Z}_r$ and a planar difference set $A=\{a_0,\ldots,a_q\}$ where $a_i<a_{i+1}$ for $i\in[q]$. Furthermore, we can assume without loss of generality that $a_0=0$. Let $A+j$ denote the line $\{a_i+j : i\in[q+1]\}$. For any integer k, let $a_{k(q+1)+i}=a_i$.

Our goal is to show that for this particular choice of ψ, in the corresponding permutation subgroup $\Psi\leqslant\mathcal{S}_r$, (E1) = (E2) = (E3) = (E4). This will be done by proving that for $i\in[q+1]$

$$\sum_{\ell\in L}(q^2+q-\ell_{q-i})=\sum_{\ell\in L}\ell_i.$$

To do so, we first require the following lemma.

Lemma 8.12. *For the planar difference set $A=\{a_0,a_1,\ldots,a_q\}\subset\mathbb{Z}_r$, and for $i\in[q+1]$,*

$$\sum_{k=0}^{q}(a_{k+i}-a_k)(a_k-a_{k-1})=\sum_{k=0}^{q}(a_{k+1}-a_k)(a_k-a_{k-i}) \tag{8.1}$$

Proof.

$$\sum_{k=0}^{q}(a_{k+i}-a_k)(a_k-a_{k-1})=\sum_{k=0}^{q}\left(a_{k+i}a_k-a_{k+i}a_{k-1}-a_k^2+a_ka_{k-1}\right)$$
$$\sum_{k=0}^{q}(a_{k+i}-a_k)(a_k-a_{k-1})=\sum_{k=0}^{q}\left(a_ka_{k-i}-a_{k+1}a_{k-i}-a_k^2+a_{k+1}a_k\right)$$

$$\sum_{k=0}^{q}(a_{k+i}-a_k)(a_k-a_{k-1}) = \sum_{k=0}^{q}(a_{k+1}-a_k)(a_k-a_{k-i}). \qquad \square$$

Lemma 8.13. *Let $L = \{A+j : j \in \mathbb{Z}_r\}$ for the planar difference set $A = \{a_0, \dots, a_q\}$. For $i \in [q+1]$*

$$\sum_{\ell \in L}(q^2+q-\ell_{q-i}) = \sum_{\ell \in L}\ell_i. \tag{8.2}$$

Proof. First, we consider $\sum_{\ell \in L}\ell_i$ for $i \in [q+1]$. Consider the values of j for which $a_k + j$ is the smallest element of $A+j$. First, when $j = q^2+q+1-a_k$, the smallest element of $A+j$ is $a_k+j=0$. Then, when $j = q^2+q+1-a_{k-1}-1$, the smallest element of $A+j$ is $a_k+j = a_k - a_{k-1} - 1$ and the largest element of $A+j$ is $a_{k-1}+j = q^2+q$. Hence, the smallest element of $A+(j+1)$ would be $a_{k-1}+(j+1) = 0$. So, for $j \in \{q^2+q+1-a_k, \dots q^2+q-a_{k-1}-1\}$, the smallest point on the line $A+j$ is a_k+j. The ith smallest point on these lines will therefore be $a_{k+i}+j$ and will range in value from $a_{k+i}-a_k$ to $a_{k+i}-a_{k-1}-1$. The sum of the integers in this interval is

$$\binom{a_k - a_{k-1}}{2} + (a_{k+i}-a_k)(a_k-a_{k-1}).$$

Therefore,

$$\sum_{\ell \in L}\ell_i = \sum_{k=0}^{q}\left(\binom{a_k - a_{k-1}}{2} + (a_{k+i}-a_k)(a_k-a_{k-1})\right).$$

Similarly, the range of values for j for which a_k+j is the largest point of the line $A+j$ is $\{q^2+q+1-a_{k+1}, \dots, q^2+q+1-a_k-1\}$. For these lines, the ith largest point is $a_{k-i}+j$ which ranges in value from $q^2+q+1+a_{k-i}-a_{k+1}$ to $q^2+q+a_{k-i}-a_k$. Therefore,

$$\begin{aligned}\sum_{\ell \in L}(q^2+q-\ell_{q-i}) &= \sum_{k=0}^{q}\sum_{j=q^2+q+1-a_{k+1}}^{q^2+q-a_k}(q^2+q-(a_{k-i}+j))\\ &= \sum_{k=0}^{q}\sum_{j=0}^{a_{k+1}-a_k-1}(a_k-a_{k-i}+j)\\ &= \sum_{k=0}^{q}\left(\binom{a_{k+1}-a_k}{2} + (a_{k+1}-a_k)(a_k-a_{k-i})\right).\end{aligned}$$

By Lemma 8.12,

$$\sum_{\ell \in L}(q^2+q-\ell_{q-i}) = \sum_{k=0}^{q}\left(\binom{a_k-a_{k-1}}{2} + (a_{k+i}-a_k)(a_k-a_{k-1})\right) = \sum_{\ell \in L}\ell_i. \qquad \square$$

Now we can prove that for this choice of Ψ, (E1) = (E2) = (E3) = (E4).

Lemma 8.14. *Let $\psi : \mathrm{PG}(2,q) \to [r]$ be a bijection such that the lines of $\mathrm{PG}(2,q)$ are mapped to the sets $\{A+j : j \in \mathbb{Z}_r\}$ for the planar difference set $A = \{a_0, \dots, a_q\}$. Then,* (E1) = (E2) = (E3) = (E4) = (E5)/4.

Proof. We can substitute (8.2) into (E4) and obtain

$$\begin{aligned}
(\text{E4}) &= \sum_{\ell \in L}\left(\sum_{i=0}^{q}\binom{i}{2}(q^2-(\ell_i-i))\right)\\
&= \sum_{\ell \in L}\left(\sum_{i=0}^{q}\binom{i}{2}(q^2+i-(q^2+q-\ell_{q-i}))\right)\\
&= \sum_{\ell \in L}\left(\sum_{i=0}^{q}\binom{i}{2}(\ell_{q-i}-(q-i))\right)\\
&= \sum_{\ell \in L}\left(\sum_{i=0}^{q}\binom{q-i}{2}(\ell_i-i)\right)\\
&= (\text{E1}).
\end{aligned}$$

By Lemma 8.11, (E1) + (E4) = (E5)/2. Therefore, (E1) = (E4) = (E5)/4. We can also substitute (8.2) into (E3) and find

$$\begin{aligned}
(\text{E3}) &= \sum_{\ell \in L}\left(\sum_{0\leqslant i<j\leqslant q} i((\ell_j-j)-(\ell_i-i))\right)\\
&= \sum_{\ell \in L}\left(\sum_{0\leqslant i<j\leqslant q} i((q^2+q-\ell_{q-j}-j)-(q^2+q-\ell_{q-i}-i))\right)\\
&= \sum_{\ell \in L}\left(\sum_{0\leqslant i<j\leqslant q} i((\ell_{q-i}-(q-i))-(\ell_{q-j}-(q-j)))\right)\\
&= \sum_{\ell \in L}\left(\sum_{0\leqslant i<j\leqslant q} (q-j)((\ell_j-j)-(\ell_i-i))\right)\\
&= (\text{E2}).
\end{aligned}$$

Again, by Lemma 8.11, (E2) + (E3) = (E5)/2. Therefore, (E2) = (E3) = (E5)/4. □

We are now ready to prove Theorem 8.2

Proof of Theorem 8.2. If ψ is of the form outlined in Lemma 8.14, then under the action of Ψ on $\mathcal{S}_{r,4}$, and for $i \in \{0,1,2,3\}$ and $s \in T_i$,

$$|\text{Asc}(s)| = \frac{(\text{E5})}{4} = \frac{(q^2+q+1)q^3(q+1)(q-1)}{24}.$$

Now consider Stab(s) and note that $|\text{Stab}(s)| = |\text{PGL}(3,q)|/|T_i|$. There are $(q+1)q(q-1)$ 3-sequences that can be formed from points of a given line ℓ. There are then q^2 points not on ℓ. Therefore,

$$|T_i| = (q^2+q+1)q^3(q+1)(q-1).$$

Thus,

$$|\mathrm{Stab}(s)||\mathrm{Asc}(s)| = \frac{|\Psi|}{(q^2+q+1)q^3(q+1)(q-1)} \frac{(q^2+q+1)q^3(q+1)(q-1)}{24} = \frac{|\Psi|}{24}.$$

Thus, by Lemma 8.4, every sequence in T_i for $i \in \{0,1,2,3\}$ is covered by $|\Psi|/24$ permutations in Ψ. By Lemma 8.6, we also know that every frame in $\mathcal{S}_{r,4}$ is covered by $|\Psi|/24$ permutations in Ψ. Moreover, by Theorem 8.5, the number of frames in $\mathcal{S}_{r,4}$ is exactly

$$|\mathrm{PGL}(3,q)| = (q^2+q+1)q^3(q+1)(q-1)^2.$$

Adding this to the number of sequences in T_i for $i \in \{0,1,2,3\}$ we find that the number of sequences in $\mathcal{S}_{r,4}$ that are covered by exactly $|\Psi|/24$ permutations in Ψ is at least

$$(q^2+q+1)q^3(q+1)(q-1)(q+3).$$

We divide this number by $|\mathcal{S}_{r,4}|$ and conclude

$$\begin{aligned}\frac{(q^2+q+1)q^3(q+1)(q-1)(q+3)}{|\mathcal{S}_{r,4}|} &= \frac{(q^2+q+1)q^3(q+1)(q-1)(q+3)}{(q^2+q+1)(q^2+q)(q^2+q-1)(q^2+q-2)} \\ &= \frac{q^2(q+3)}{(q^2+q-1)(q+2)} \\ &= 1 - \frac{q-2}{q^3+3q^2+q-2}.\end{aligned}$$

Now, $q^2+3q^2+q-2 = (q-2)(q^2+5q+11)+20$. Therefore,

$$1 - \frac{q-2}{q^3+3q^2+q-2} > 1 - \frac{1}{q^2+5q+11}.$$

This completes the proof. □

Chapter 9

Connections Between Excess Coverage Arrays and Levenshtein's Conjecture

In this section, we further explore the connections between excess coverage arrays and Conjecture 3.18. In particular, by performing computer searches for $\mathsf{CA}_X(42; 2, k, 6)$, we are able to make progress on the $t = 7$ case of Conjecture 3.18, the smallest value of t for which the conjecture is open. We also analyse $\mathsf{CA}_X((t+1)!; t, k, 2)$ to see if these arrays may provide further insights on Conjecture 3.18.

9.1 Introduction

Recall Levenshtein's conjecture.

Conjecture 9.1 ([36]). *For $t \geqslant 3$, an $\mathsf{SCA}(t!; t, v)$ exists if and only if $v \in \{t, t+1\}$.*

The following theorem describes computational results of Mathon and Tran van Trung [41] relating to Levenshtein's conjecture for small values of t.

Theorem 9.2 ([41]). *For $t \in \{3, 5, 6\}$, an $\mathsf{SCA}(t!; t, v)$ exists if and only if $v \in \{t, t+1\}$. An $\mathsf{SCA}(24; 4, v)$ exists if and only if $v \in \{4, 5, 6\}$.*

For $t \geqslant 7$, the current best upper bound on the number of symbols in an $\mathsf{SCA}(t!; t, v)$ is due to Chee et al. [9].

Theorem 9.3 ([9]). *For $t \geqslant 3$, if an $\mathsf{SCA}(t!; t, v)$ exists, then $v \leqslant 2t - 1$.*

This theorem is proved by bounding the number of columns in a $\mathsf{CA}_X(v(v+1); 2, k, v)$. In this chapter, we prove the following theorem.

Theorem 9.4. *If an $\mathsf{SCA}(7!; 7, v)$ exists, then $v \in \{7, 8, 9\}$.*

This reduces the best known upper bound on the number of symbols in an $\mathsf{SCA}(7!;7,v)$ from 13 to 9. The set $\mathcal{S}_7$ is an $\mathsf{SCA}(7!;7,7)$ and a general construction due to Levenshtein [37] gives an $\mathsf{SCA}(7!;7,8)$. It is unknown whether an $\mathsf{SCA}(7!;7,9)$ exists. We prove Theorem 9.4 by exploring the connections between sequence covering arrays and covering arrays with excess coverage.

9.2 Covering arrays with excess coverage

Recall the definition of a covering array. Let C be an $N \times k$ array where each entry in C is a symbol from the alphabet $[v]$. A *t-way interaction* is a set of t pairs $\{(c_i, \nu_i) : 0 \leqslant i \leqslant t-1\}$ where each c_i is a column of C such that $c_i \neq c_j$ for $i \neq j$, and each ν_i is an element of $[v]$. The row ρ of C *covers* the interaction $\{(c_i, \nu_i) : 0 \leqslant i \leqslant t-1\}$ if the entry of C in row ρ and column c_i is ν_i for $0 \leqslant i \leqslant t-1$. The array C is a *covering array of strength t*, denoted by $\mathsf{CA}(N;t,k,v)$, if for each t-way interaction T, there is some row of C that covers T. Let $\mathsf{CAN}(t,k,v)$ be the smallest N such that a $\mathsf{CA}(N;t,k,v)$ exists.

For an interaction $T = \{(c_i, \nu_i) : 0 \leqslant i \leqslant t-1\}$, and for $\sigma \in [v]$, let $\tau_\sigma(T) = \{i : \nu_i = \sigma\}$. Then, let $\mu(T) = \prod_{\sigma=0}^{v-1} |\tau_\sigma(T)|!$. Then C is an *excess coverage array*, denoted by $\mathsf{CA}_X(N;t,k,v)$, if each interaction T is covered by at least $\mu(T)$ different rows of C. Let $\mathsf{CAN}_X(t,k,v)$ be the smallest N for which a $\mathsf{CA}_X(N;t,k,v)$ exists.

We begin our analysis of excess coverage arrays with an observation that resembles Theorem 5.1. However, a crucial difference is that we consider all excess coverage arrays that can be built from a sequence covering array, rather than focusing on the existence of a single excess coverage array. Let P be an $\mathsf{SCA}(t!;t,v)$ and let A be a set of a symbols of $[v]$ with $0 < a < t$. We partition the permutations of P into $a!$ parts according to the order in which the symbols of A appear in each permutation. Consider one of these parts, B, and let C be a $t!/a! \times (v-a)$ array with rows indexed by the permutations in B and columns indexed by the symbols in $[v] \setminus A$. The entry of row ρ and column ν of C is the number of symbols of A that precede ν in the permutation ρ. Note then that the alphabet for C is $[a+1]$. This construction can be performed for any choices of A and B so in fact we can construct $\binom{v}{a}a! = v!/(v-a)!$ arrays in this way. We let $\mathcal{C}_a(P)$ be the (multi)set of all the arrays that can be constructed through the removal of a symbols from P.

Theorem 9.5. *Let P be an $\mathsf{SCA}(t!;t,v)$. Then, for any $\mathsf{C} \in \mathcal{C}_a(P)$, each $(t-a)$-way interaction T must be covered by exactly $\mu(T)$ rows of C.*

Proof. Let P be an $\mathsf{SCA}(t!;t,v)$ and let $\mathsf{C} \in \mathcal{C}_a(P)$. Then the rows of C correspond to the permutations in P that cover some sequence $(x_0, \ldots, x_{a-1})$ of the elements of A. Let $D = \{y_0, \ldots, y_{t-a-1}\}$ be a subset of $[v]$ that is disjoint with A and let $T = \{(y_i, \nu_i) : 0 \leqslant i \leqslant t-a-1\}$ be a $(t-a)$-way interaction where each ν_i is an element of $[a+1]$. Consider the permutations of $A \cup D$ that cover $(x_0, \ldots, x_{a-1})$ and have the elements of

$\tau_0(T)$ preceding x_0, the elements of $\tau_a(T)$ succeeding x_{a-1} and the elements of $\tau_i(T)$ lying between x_{i-1} and x_i for $1 \leqslant i \leqslant a-1$. There are $\mu(T)$ such permutations. Each of these corresponds to a sequence in $\mathcal{S}_{v,t}$ which must be covered by exactly one permutation in P. Moreover, a row of C covers T if and only if the corresponding permutation in P covers one of these sequences. Therefore, there are $\mu(T)$ rows of C that cover T. □

In particular, every array in $\mathcal{C}_a(P)$ must be a $\mathsf{CA}_X(t!/a!; t-a, v-a, a+1)$. For a row ρ of an array $\mathsf{C} \in \mathcal{C}_a(P)$, let m_i be the number of times the symbol i appears in ρ for $0 \leqslant i \leqslant a$. We call $(m_0, \dots, m_a)$ the *multiplicity vector* of ρ.

Theorem 9.6. *Let P be an $\mathsf{SCA}(t!; t, v)$ and let $m_0, \dots, m_a$ be non-negative integers that sum to $v-a$. Then, across all the arrays in $\mathcal{C}_a(P)$, there must be exactly $t!$ rows with multiplicity vector $(m_0, \dots, m_a)$.*

Proof. Let $\mathsf{C} \in \mathcal{C}_a(P)$ and let A be the set of symbols that were deleted from P to generate C. In order to generate a row in C with multiplicity vector $(m_0, \dots, m_a)$, the symbols in A must appear in positions $j + \sum_{i=0}^{j} m_i$ for $0 \leqslant j \leqslant a$ in the corresponding permutation in P. Now, there are $t!$ permutations in P and for each permutation, there is one set of symbols in these positions. Therefore, across all the arrays in $\mathcal{C}_a(P)$, there must be exactly $t!$ rows with multiplicity vector $(m_0, \dots, m_a)$. □

We call a row of a $\mathsf{CA}_X(N; t, k, v)$ *constant* if it contains only one symbol (i.e. $m_i = k$ for some i). Since the alphabet for each array in $\mathcal{C}_a(P)$ has $a+1$ symbols, then by Theorem 9.6, there are a total of $(a+1)t!$ constant rows across all the arrays in $\mathcal{C}_a(P)$. Recall that there are $v!/(v-a)!$ arrays in $\mathcal{C}_a(P)$. Therefore, the average number of constant rows per array in $\mathcal{C}_a(P)$ is $(a+1)t!(v-a)!/v!$.

9.3 Covering arrays with strength 2

In this section, we describe computations relating to the existence of excess coverage arrays with strength 2. The results of these computations allow us to prove Theorem 9.4. Consider a strength 2 covering array with excess coverage. Let $T = \{(c_1, \nu_1), (c_2, \nu_2)\}$ be a 2-way interaction. If $\nu_1 = \nu_2$, then $\mu(T) = 2$ but if $\nu_1 \neq \nu_2$, then $\mu(T) = 1$. Following the convention of Chee et al. [9], we refer to the former kind of pair as *constant pairs* and the latter as *non-constant pairs*. Let C be a $\mathsf{CA}_X(N; 2, k, v)$ where every 2-way interaction T is covered exactly $\mu(T)$ times. For distinct columns c_1 and c_2, there are v^2 interactions of the form $\{(c_1, \nu_1), (c_2, \nu_2)\}$. Of these, v are constant pairs which are covered exactly twice by C while the remaining $v^2 - v$ interactions are non-constant pairs and are covered exactly once by C. Therefore, C has $N = v^2 - v + 2v = v(v+1)$ rows. Therefore, $\mathsf{CAN}_X(2, k, v) \geqslant v(v+1)$. By Theorem 9.5, a $\mathsf{CA}_X(v(v+1); 2, k, v)$ is necessary for the existence of an $\mathsf{SCA}((v+1)!; v+1, v+k-1)$. By Theorem 5.2, for $v \geqslant 4$, if a $\mathsf{CA}_X(v(v+1); 2, k, v)$ exists, then $k \leqslant v+2$.

We performed a series of computations to find the maximum number of columns k for which a $\mathsf{CA}_X(v(v+1);2,k,v)$ exists for $v \in \{2,3,4,5,6\}$. The computations for $v=6$ are of particular interest since they relate to the $t=7$ case of Levenshtein's Conjecture. This is the smallest value of t for which the existence of SCAs with $t!$ permutations has not been fully resolved. In particular, the results for $v=6$ allow us to prove Theorem 9.4. Another consequence of these computations is that the upper bound in Theorem 5.2, which was only proved for $v \geqslant 4$, also holds when $v \in \{2,3\}$.

For fixed $v \in \{2,3,4,5\}$ we recursively found all ways of extending a $\mathsf{CA}_X(v(v+1);2,k,v)$ to a $\mathsf{CA}_X(v(v+1);2,k+1,v)$. The base of this recursion was the unique $\mathsf{CA}_X(v(v+1);2,2,v)$ (this array has two columns and simply contains every non-constant pair once and every constant pair twice). At each step of the recursion, we took a $\mathsf{CA}_X(v(v+1);2,k,v)$ and for each symbol found all possible *transversals*. Here, a transversal for symbol ν is a set of $v+1$ rows in the $\mathsf{CA}_X(v(v+1);2,k,v)$ such that in each column, the symbol ν appears twice in these rows and all other symbols appear once each. Then, if we were to add ν to a $(k+1)$th column c of these rows then all 2-way interactions containing the pair (c,ν) would be covered the suitable number of times. We note that because constant pairs must be covered twice, a transversal for symbol i will not be a transversal for symbol j for distinct i and j. Once all transversals were computed, we then found all ways of choosing a transversal for each symbol such that these transversals partition the rows of the $\mathsf{CA}_X(v(v+1);2,k,v)$. Then, to build a $\mathsf{CA}_X(v(v+1);2,k+1,v)$, we take our original array and add a new column c where the symbol in row r is the symbol whose transversal r belongs to.

For $v=6$, we found a large number of $\mathsf{CA}_X(42;2,4,6)$ which made extending individual arrays an expensive task. Thus we employed a slightly different strategy to find all $\mathsf{CA}_X(42;2,5,6)$. Here we took all possible $\mathsf{CA}_X(42;2,3,6)$ and found all ways of adding two columns to each array to generate a $\mathsf{CA}_X(42;2,5,6)$. This was achieved by choosing two transversals for each symbol. One of these transversals would correspond to the rows in which that symbol would appear in column 3 and the other would correspond to the rows in which that symbol would appear in column 4. The two transversals chosen for each symbol had to intersect in exactly two rows to ensure that each constant pair in columns 3 and 4 are covered exactly twice. For two distinct symbols, we must have transversals corresponding to the placement of each symbol in column 3. These transversals must be disjoint. Similarly the transversals corresponding to the placement of each symbol in column 4 must also be disjoint. These restrictions ensure that exactly one symbol appears in every entry of the resulting $\mathsf{CA}_X(42;2,5,6)$. The transversal corresponding to placement of one symbol in column 3 and the transversal corresponding to the placement of the other symbol in column 4 must intersect in one row to ensure every non-constant pair in columns 3 and 4 is covered exactly once.

As part of these computations, it was important to screen for isomorphisms. Here, two arrays are isomorphic if one can be obtained from the other by applying a permutation

v	Maximum number of columns
2	4
3	4
4	5
5	6
6	5

Table 9.1: Maximum number of columns for $\mathsf{CA}_X(v(v+1);2,k,v)$ for $2 \leqslant v \leqslant 6$.

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

Figure 9.1: A $\mathsf{CA}_X(6;2,4,2)$.

to the columns and applying a permutation to the symbols. It is important to note here that any permutation of the symbols must be applied throughout the whole array. Say we were to swap the symbols 0 and 1 in only column 0 of a $\mathsf{CA}_X(v(v+1);2,k,v)$. Then, the constant pair $\{(0,0),(1,0)\}$ would now only be covered once while the non-constant pair $\{(0,0),(1,1)\}$ would be covered twice and so the resulting array would not be an excess coverage array. However, if we were to apply the same symbol permutation to every column, we would obtain a $\mathsf{CA}_X(v(v+1);2,k,v)$. This behaviour contrasts with that of traditional covering arrays in which one can apply different symbol permutations to different columns and still obtain a covering array.

The maximum number of columns k for which a $\mathsf{CA}_X(v(v+1);2,k,v)$ exists is recorded in Table 9.1. We see that the bound of $k \geqslant v+2$ in Theorem 5.2 also holds for $v \in \{2,3\}$. Indeed, when $v = 2$, this bound is met. An example of a $\mathsf{CA}_X(6;2,4,2)$ is given in Figure 9.1. This is the only value of v for which we found a $\mathsf{CA}_X(v(v+1);2,v+2,v)$. We also see that a $\mathsf{CA}_X(42;2,6,6)$ does not exist. By Theorem 9.5, this shows that an $\mathsf{SCA}(7!;7,11)$ does not exist. However, by analysing all possible $\mathsf{CA}_X(42;2,5,6)$, we can prove the stronger Theorem 9.4.

Proof of Theorem 9.4. Our computations found exactly one $\mathsf{CA}_X(42;2,5,6)$ up to isomorphism. This is given in Figure 9.2. We can see that this array contains no constant rows. Our definition of isomorphism preserves the number of constant rows so there are no $\mathsf{CA}_X(42;2,5,6)$ with any constant rows. However, by Theorem 9.6, if an $\mathsf{SCA}(7!;7,10)$ exists, there must be some $\mathsf{CA}_X(42;2,5,6)$ with a constant row. Therefore, an $\mathsf{SCA}(7!;7,10)$ does not exist. □

Each row in the array in Figure 9.2 contains a repeated symbol. One can similarly argue that by Theorem 9.6, if an $\mathsf{SCA}(7!;7,10)$ exists, then there must be a $\mathsf{CA}_X(42;2,5,6)$

0	0	0	0	1	1	0	2	0	5	2	0	1	1	4
0	0	0	2	0	1	1	1	5	1	2	1	0	3	0
0	1	2	1	2	1	1	3	1	1	2	2	2	2	3
0	2	3	3	4	1	2	0	4	2	2	2	4	0	2
0	3	5	0	0	1	3	4	2	4	2	3	3	5	5
0	4	4	5	3	1	4	1	1	0	2	4	5	4	1
0	5	1	4	5	1	5	5	3	3	2	5	2	2	2
3	0	3	3	3	4	0	5	5	2	5	0	4	4	0
3	1	5	2	5	4	1	4	4	4	5	1	1	0	3
3	2	2	5	0	4	2	1	2	1	5	2	5	1	5
3	3	1	3	2	4	3	0	1	3	5	3	2	3	1
3	3	3	4	3	4	4	2	4	4	5	4	3	2	2
3	4	0	0	4	4	4	4	3	5	5	5	0	5	5
3	5	4	1	1	4	5	3	0	0	5	5	5	5	4

Figure 9.2: The unique $\mathsf{CA}_X(42; 2, 5, 6)$.

with a row with multiplicity vector $(m_0, \ldots, m_5)$ such that $m_i \leqslant 1$ for $0 \leqslant i \leqslant 5$. However, our computations show that no such array exists.

We remark that since a $\mathsf{CA}_X(42; 2, 6, 6)$ does not exist, $\mathsf{CAN}_X(2, 6, 6) \geqslant 43$. Then, by Theorem 5.1, $\mathsf{SCAN}(7, 11) \geqslant 5160$. Similarly, using Table 9.1 and Theorem 5.1, we find that $\mathsf{SCAN}(4, 7) \geqslant 26$, $\mathsf{SCAN}(5, 9) \geqslant 126$ and $\mathsf{SCAN}(6, 11) \geqslant 744$. Although we have proved that an $\mathsf{SCA}(7!; 7, 10)$ does not exist, the existence of a $\mathsf{CA}_X(42; 2, 5, 6)$ means we are unable to use Theorem 5.1 to derive an improved lower bound for $\mathsf{SCAN}(7, 10)$ as in the other cases above.

9.4 Strength 2 orthogonal arrays

In this section, we discuss connections between strength 2 excess coverage arrays and strength 2 orthogonal arrays and analyse the computational results of Chapter 9.3 in this context. We begin by observing upper and lower bounds for $\mathsf{CAN}_X(2, k, v)$.

Lemma 9.7. *For integers $v, k \geqslant 2$,*

$$\mathsf{CAN}(2, k, v) \leqslant \mathsf{CAN}_X(2, k, v) \leqslant \mathsf{CAN}(2, k, v) + v.$$

Proof. A $\mathsf{CA}_X(N; 2, k, v)$ is also a $\mathsf{CA}(N; 2, k, v)$ so the first inequality holds. Let $N = \mathsf{CAN}(2, k, v)$ and let $N' = \mathsf{CAN}(2, k, v) + v$. Let C be a $\mathsf{CA}(N; 2, k, v)$ and let D be a $v \times v$ array such that $\mathsf{D}[i, j] = i$ for $i, j \in [v]$. Then, every 2-way interaction is covered at least once in C and every constant pair is covered exactly once in D. Therefore the rows of C and D form a $\mathsf{CA}_X(N'; 2, k, v)$. Thus, the second inequality holds. □

By considering strength 2 orthogonal arrays, we obtain the following corollary.

	k			
v	3	4	5	6
2	1/0	0/1	0/0	0/0
3	3/1	2/1	0/0	0/0
4	15/12	32/6	80/5	0/0
5	283/1 067	2 234/3 805	104 146/348	2 073 801/0
6	190 472/1 666 259	0/39 802 785	0/1	0/0

Table 9.2: Number of $\mathsf{CA}_X(v(v+1); 2, k, v)$ for different values of k and v. Each entry r/s indicates that there are r arrays with those parameters that contain an orthogonal array and s arrays that do not contain an orthogonal array.

Corollary 9.8. *If an* $\mathsf{OA}(2, k, v)$ *exists, then* $\mathsf{CAN}_X(2, k, v) = v(v+1)$.

If v is a prime power, then an $\mathsf{OA}(2, v+1, v)$ exists [21] and so a $\mathsf{CA}_X(v(v+1); 2, v+1, v)$ also exists. Hence, for infinitely many values of v, we can construct a $\mathsf{CA}_X(v(v+1); 2, k, v)$ whose number of columns is one less than the upper bound given by Theorem 5.2. We continue our discussion of orthogonal arrays by characterising exactly when a $\mathsf{CA}_X(v(v+1); 2, k, v)$ contains a subarray that forms an $\mathsf{OA}(2, k, v)$.

Lemma 9.9. *A* $\mathsf{CA}_X(v(v+1); 2, k, v)$ *contains an* $\mathsf{OA}(2, k, v)$ *if and only if it contains a constant row for each symbol in* $[v]$.

Proof. Let C be a $\mathsf{CA}_X(v(v+1); 2, k, v)$ that contains an $\mathsf{OA}(2, k, v)$. In the rows of the OA, every 2-way interaction is covered exactly once so the remaining v rows of C must cover each constant pair once while avoiding all non-constant pairs. The only way to achieve this is if these v rows are all constant and contain distinct symbols of $[v]$. Conversely, if C contains constant rows for each symbol, then the remaining rows must cover every 2-way interaction exactly once and so must form an $\mathsf{OA}(2, k, v)$. □

Let P be an $\mathsf{SCA}(t!; t, v)$. By Theorem 9.6, the average number of constant rows per array in $\mathcal{C}_{t-2}(P)$ is $(t-1)t!(v-t+2)!/v!$. When $t \geqslant 4$ and $v \geqslant t+2$, this average is less than $t-1$. An array in $\mathcal{C}_{t-2}(P)$ that contains an $\mathsf{OA}(2, v-t+2, t-1)$ must have at least $t-1$ constant rows. Therefore, when $t \geqslant 4$ and $v \geqslant t+2$, there must be some array in $\mathcal{C}_{t-2}(P)$ that does not contain an orthogonal array. Hence, it is sensible to consider the existence of covering arrays with excess coverage that do not contain orthogonal arrays. Table 9.4 details the number of $\mathsf{CA}_X(v(v+1); 2, k, v)$ that do and do not contain orthogonal arrays for $2 \leqslant v \leqslant 6$ and $3 \leqslant k \leqslant 6$.

After $v = 6$, the next three values of v are all prime powers. Hence, for these values of v, we will find $\mathsf{CA}_X(v(v+1); 2, k, v)$ containing orthogonal arrays for all $3 \leqslant k \leqslant v+1$. Recall our definition of isomorphic excess coverage arrays. If we were to permute the symbols of some column in an excess coverage array, in order to obtain another excess coverage array we need to apply the same symbol permutation to all columns. However, if we apply different symbol permutations to different columns of an orthogonal array, we will still obtain an orthogonal array. Hence, we have fewer symmetries to exploit when it comes to excess coverage arrays. Ideally, given that certain sequence covering arrays require the existence of excess coverage arrays that do not contain an orthogonal array, we would like a computational search method that avoids having to generate excess coverage arrays containing orthogonal arrays. If we were to employ the same search methods used for $v \in \{2,3,4,5\}$ in Chapter 9.3 for $v > 6$, our catalogues of excess coverage arrays would necessarily include those that contain orthogonal arrays. The relative lack of symmetries of excess coverage arrays would further inflate our catalogue. The following lemma informs one possible strategy of cataloguing excess coverage arrays that do not contain an orthogonal array that we describe below.

Lemma 9.10. *Let $k \geqslant 5$ and let C be a $\mathsf{CA}_X(v(v+1); 2, k, v)$ that does not contain an orthogonal array. Then there is some column of C such that if that column is removed from C, then the resulting $\mathsf{CA}_X(v(v+1); 2, k-1, v)$ does not contain an orthogonal array either.*

Proof. As C does not contain an orthogonal array, then by Lemma 9.9 and without loss of generality, there is no constant row in C containing only the symbol 0. Suppose that by deleting any column from C, we obtain a $\mathsf{CA}_X(v(v+1); 2, k-1, v)$ that does contain an orthogonal array. Then, by deleting any column of C, we create a constant row containing the symbol 0. This means that for $0 \leqslant i \leqslant k-1$, there is a row r of C such that $\mathsf{C}[r,i] \neq 0$ and $\mathsf{C}[r,j] = 0$ for $j \neq i$. However, since $k \geqslant 5$, this means there are at least three rows of C in which the symbols in columns 0 and 1 are both 0. Thus, the 2-way interaction $\{(0,0),(1,0)\}$ is covered at least three times which contradicts the fact that C covers every constant pair twice. Therefore, we must be able to delete some column of C to obtain a $\mathsf{CA}_X(v(v+1); 2, k-1, v)$ that does not contain an orthogonal array. □

The consequence of Lemma 9.10 is that a complete catalogue of $\mathsf{CA}_X(v(v+1); 2, 4, v)$ that do not contain an orthogonal array can be used to generate a complete catalogue of $\mathsf{CA}_X(v(v+1); 2, 5, v)$ that do not contain an orthogonal array which in turn can be used to generate a complete catalogue of $\mathsf{CA}_X(v(v+1); 2, 6, v)$ that do not contain an orthogonal array and so on. The key is to generate this initial catalogue. This may be achieved by using a similar method for the $v = 6$ computations in Chapter 9.3 and adding two transversals for each symbol to the unique $\mathsf{CA}_X(v(v+1); 2, 2, v)$. To ensure that the arrays we build do not contain orthogonal arrays, we can without loss of generality ensure that these arrays do not contain a constant row of 0s. This can be done by choosing the transversals for the symbol 0 such that for each of the two constant rows containing 0s in the $\mathsf{CA}_X(v(v+1); 2, 2, v)$, there is a transversal avoiding that row.

Note that there may exist a $\mathsf{CA}_X(v(v+1); 2, 4, v)$ that does not contain an orthogonal array such that removing any column from this array gives a $\mathsf{CA}_X(v(v+1); 2, 3, v)$ that does contain an orthogonal array. An example of such an array is the $\mathsf{CA}_X(6; 2, 4, 2)$ in

Figure 9.1. This array contains no constant row of 1s but deleting any column will form a constant row of 1s. Thus, it is necessary that we start with a catalogue of arrays with four columns.

9.5 Binary covering arrays

In the previous sections, we have been concerned with excess coverage arrays with strength 2. We can build such arrays using the method described in Theorem 5.1 with $a = t - 2$. However, this method works for any $1 \leqslant a \leqslant t-1$ so we are not bound to looking solely at strength 2 covering arrays. In this section, we consider excess coverage arrays with $v = 2$. That is, $\mathsf{CA}_X(N; t, k, 2)$. These arrays may be built by letting $a = 1$ in Theorem 5.1. In particular, we study these arrays to see whether they may provide additional insight into Levenshtein's Conjecture and the existence of an $\mathsf{SCA}(t!; t, v)$. We consider rows of these arrays as being elements of $\{0, 1\}^k$. For $u \in \{0, 1\}^k$, the *weight* of u is the number of 1's in u.

Lemma 9.11. *If* C *is a* $\mathsf{CA}_X(N; t, k, 2)$ *such that* C *covers each* t*-way interaction* T *exactly* $\mu(T)$ *times, then* $N = (t+1)!$.

Proof. Let $T = \{(c_i, \nu_i) : 0 \leqslant i \leqslant t-1\}$ be a t-way interaction with $|\tau_0(T)| = j$. Then $\mu(T) = j!(t-j)!$. For t distinct columns, there are $\binom{t}{j}$ interactions on those columns with $|\tau_0(T)| = j$. Therefore,

$$N = \sum_{j=0}^{t} \binom{t}{j} j!(t-j)! = \sum_{j=0}^{t} t! = (t+1)!. \qquad \square$$

We next analyse and count the number of $\mathsf{CA}_X((t+1)!; t, t+1, 2)$.

Lemma 9.12. *Let* C *be a* $\mathsf{CA}_X((t+1)!; t, t+1, 2)$. *For* $0 \leqslant i \leqslant t+1$, *there is a constant* x_i *such that each element in* $\{0, 1\}^{t+1}$ *with weight* i *appears as a row of* C *exactly* x_i *times. Furthermore,*

$$x_i + x_{i+1} = i!(t-i)!, \text{ for } 0 \leqslant i \leqslant t. \tag{9.1}$$

Proof. We proceed by induction on i. As there is a single element of $\{0, 1\}^{t+1}$ with weight 0, we can take x_0 to be the number of weight 0 rows in C. Let x_i be the number of times each element of $\{0, 1\}^{t+1}$ with weight i appears as a row of C. Let $u \in \{0, 1\}^{t+1}$ with weight $i+1$ and let T be a t-way interaction with $\tau_1(T) = i$ such that u covers T. Let w be the unique vector in $\{0, 1\}^{t+1}$ with weight i covering T. Then u and w are the only vectors in $\{0, 1\}^{t+1}$ that cover T. As $\tau_1(T) = i$, we must have $\mu(T) = i!(t-i)!$. As there are x_i rows in C equal to w, there must be $i!(t-i)! - x_i$ rows equal to u. Therefore, each element of $\{0, 1\}^{t+1}$ with weight $i+1$ appears as a row of C x_{i+1} times where $x_{i+1} = i!(t-i)! - x_i$. $\square$

Theorem 9.13. *The number of distinct* $\mathsf{CA}_X((t+1)!; t, t+1, 2)$ *is exactly* $\lfloor \frac{t}{2} \rfloor! \lceil \frac{t}{2} \rceil! + 1$ *for* $t \geqslant 1$. *When* t *is odd, no two of these arrays are isomorphic but when* t *is even, there are*

$\lfloor \frac{t}{2} \rfloor! \lceil \frac{t}{2} \rceil!/2$ *isomorphism classes each containing two arrays and one isomorphism class containing a single array.*

Proof. From each non-negative integer solution $x_0, \ldots, x_{t+1}$ to the system in (9.1), we can build a $\mathsf{CA}_X((t+1)!; t, t+1, 2)$ by adding x_i rows corresponding to each vector in $\{0,1\}^{t+1}$ with weight i. Therefore, the number of distinct $\mathsf{CA}_X((t+1)!; t, t+1, 2)$ is equal to the number of non-negative integer solutions to (9.1).

Let $s = \lfloor t/2 \rfloor$. Now consider x_s. By (9.1), $x_s + x_{s+1} = \lfloor \frac{t}{2} \rfloor! \lceil \frac{t}{2} \rceil!$. Let x_s be some integer between 0 and $\lfloor \frac{t}{2} \rfloor! \lceil \frac{t}{2} \rceil!$. Then, in order to satisfy each equation in (9.1), fixing x_s fixes the values for all other x_i. Suppose $0 \leqslant i \leqslant s-1$. Then $i!(t-i)! > (i+1)!(t-(i+1))!$. Therefore, if x_{i+1} is a non-negative integer that is at most $(i+1)!(t-(i+1))!$, then x_i is a non-negative integer that is at most $i!(t-i)!$. As x_s is a non-negative integer that is at most $\lfloor \frac{t}{2} \rfloor! \lceil \frac{t}{2} \rceil!$, then x_i is a non-negative integer for all $i \in \{0, \ldots, s\}$. Similarly, x_{s+1} is a non-negative integer that is at most $\lfloor \frac{t}{2} \rfloor! \lceil \frac{t}{2} \rceil!$ and $i!(t-i)! < (i+1)!(t-(i+1))!$ for $s+1 \leqslant i \leqslant t$, so x_i is a non-negative integer for all $i \in \{s+1, \ldots, t+1\}$. Thus, each non-negative integer choice for x_s gives rise to a non-negative integer solution to (9.1) and thus, a $\mathsf{CA}_X((t+1)!; t, t+1, 2)$. Therefore, the number of $\mathsf{CA}_X((t+1)!; t, t+1, 2)$ is exactly $\lfloor \frac{t}{2} \rfloor! \lceil \frac{t}{2} \rceil! + 1$.

Now we can consider isomorphisms of these arrays. Let C be a $\mathsf{CA}_X((t+1)!; t, t+1, 2)$. Since each vector in $\{0,1\}^{t+1}$ with the same weight appears the same number of times in C, permuting the columns of C will not generate a new array. The only other possible isomorphism to consider then is swapping all 1's and 0's in C. By (9.1), $x_i + x_{i+1} = x_{t+1-i} + x_{t-i}$. Suppose t is odd. Then, $x_{(t+1)/2-1} + x_{(t+1)/2} = x_{(t+1)/2} + x_{(t+1)/2+1}$. Thus, $x_{(t+1)/2-1} = x_{(t+1)/2+1}$ and so $x_i = x_{t+1-i}$ for $0 \leqslant i \leqslant (t+1)/2$. Hence, swapping 1's and 0's in a $\mathsf{CA}_X((t+1)!; t, t+1, 2)$ when t is odd will not generate a new array. Therefore, the $\lfloor \frac{t}{2} \rfloor! \lceil \frac{t}{2} \rceil! + 1$ possible arrays with those parameters must be in distinct isomorphism classes.

Now suppose t is even. Then it can be shown by induction on i and using (9.1) that $x_{t/2-i} = x_{t/2+1+i} + (-1)^i (x_{t/2} - x_{t/2+1})$ for $0 \leqslant i \leqslant t/2$. In particular, $x_i = x_{t+1-i}$ for $0 \leqslant i \leqslant t+1$ if and only if $x_{t/2} = x_{t/2+1}$. Thus, if $x_{t/2} \neq x_{t/2+1}$ swapping 1's and 0's will generate a new array. Therefore, when t is even, there are $\lfloor \frac{t}{2} \rfloor! \lceil \frac{t}{2} \rceil!/2$ isomorphism classes of $\mathsf{CA}_X((t+1)!; t, t+1, 2)$ each containing 2 arrays and 1 isomorphism class containing a single array. □

To conclude this section, we consider whether analysing the existence of $\mathsf{CA}_X((t+1)!; t, k, 2)$ may lead to improvements to the upper bound of Chee et al. [9] that the number of symbols in an $\mathsf{SCA}(t!; t, v)$ is at most $2t-1$. Suppose an $\mathsf{SCA}(t!; t, 2t-1)$ exists. Then from this array, we can obtain a $\mathsf{CA}_X(t!; t-1, 2t-2, 2)$. That is, we obtain an excess coverage array with 2 symbols and where the number of columns is twice the strength. Moreover, this array will cover every interaction T exactly $\mu(T)$ times. Therefore, if we can establish the non-existence of a $\mathsf{CA}_X((t+1)!; t, 2t, 2)$, for some value of t, then we can improve upon the bound of Chee et al.

In building a $\mathsf{CA}_X((t+1)!; t, 2t, 2)$ we essentially need to decide how many times each vector in $\{0,1\}^{2t}$ will appear in the array. For a t-way interaction T, there are 2^t vectors in $\{0,1\}^{2t}$ that cover T. The number of times these vectors collectively appear in a $\mathsf{CA}_X((t+1)!; t, 2t, 2)$ must be $\mu(T)$. If we apply this logic to all possible t-way interactions, we can obtain a system of $\binom{2t}{t}2^t$ equations with 2^{2t} variables. A non-negative integer solution to this system can be used to generate a $\mathsf{CA}_X((t+1)!; t, 2t, 2)$ by adding the appropriate number of each vector as specified by the solution.

We can simplify this system by only considering arrays in which vectors with the same weight appear the same number of times. Note that this restriction is not necessary as it was in the case where $k = t+1$. By making this restriction, we reduce the number of variables to $2t+1$, one for each possible weight, and the number of equations to $t+1$ since all interactions with the same value of τ_1 will generate the same equation. Specifically, if $\tau_1(T) = i$ for some interaction T, then there are $\binom{t}{j}$ vectors with weight $i+j$ that cover T. Let x_i be the number of vectors with weight i. Then, the relevant system of equations is

$$\sum_{j=0}^{t} \binom{t}{j} x_{i+j} = i!(t-i)!, \text{ for } 0 \leqslant i \leqslant t. \tag{9.2}$$

We have found non-negative integer solutions to (9.2) for all $t \leqslant 17$. Therefore, a $\mathsf{CA}_X((t+1)!; t, 2t, 2)$ exists for all $t \leqslant 17$. These arrays mean that at present, we cannot disprove the existence of an $\mathsf{SCA}(t!; t, 2t-1)$ for $t \leqslant 18$. For $t = 18$, no non-negative integer solution to (9.2) exists. However, this does not necessarily rule out the existence of a $\mathsf{CA}_X(19!; 19, 38, 2)$. Hence, we are also at present unable to disprove the existence of an $\mathsf{SCA}(19!; 19, 37)$.

Chapter 10

Conclusion

In this thesis, we have primarily considered sequence covering arrays and perfect sequence covering arrays. We have been able to make original contributions to the theories of each of these objects. In Chapter 7, we devised new methods of exhaustively searching for perfect sequence covering arrays. We used these methods to uncover new values of $\mathsf{PSCAN}(v,t)$. In Chapter 8, we presented a new construction for $\mathsf{PSCA}(v,t)$ that could be applied to any choice of v and t. The size of the corresponding $\mathsf{PSCAN}(v,t,\lambda)$ grew polynomially in v for fixed t. In Chapter 9, we investigated excess coverage arrays and used these objects to rule out the existence of an $\mathsf{SCA}(7!;7,10)$. We conclude by outlining some open problems and potential directions for future research.

In the paper first introducing perfect sequence covering arrays, Yuster [64] poses three problems. The first problem is to find polynomial upper bounds for $\mathsf{PSCAN}(v,t)$. This problem has been addressed in this thesis as well as in the concurrent work of Iurlano [28] and we can indeed say that $\mathsf{PSCAN}(v,t)$ grows polynomially in v with the exponent of this polynomial growing with t. In particular, the best upper bound for $\mathsf{PSCAN}(v,t)$ is given by a probabilistic upper bound on the size of a t-wise independent set of permutations due to Kuperberg, Lovett and Peled [35]. However, as we discussed in Chapter 4, families of t-wise independent permutations are under much stricter conditions than PSCAs. There is clearly much room for improvement for constructions of perfect sequence covering arrays.

The second problem is, given the linear lower bound and quasi-linear upper bound for $\mathsf{PSCAN}(v,3)$ proved in [64], what is the right order of magnitude of $\mathsf{PSCAN}(v,3)$? While an improved upper bound for $\mathsf{PSCAN}(v,3)$ due to Tarui, Itoh and Takei [54] was uncovered by Iurlano [28], this new upper bound is still quasi-linear in v. Hence, this problem remains unresolved.

The third problem asks for more exact values of $\mathsf{PSCAN}(v,t)$. The computational methods we developed in Chapter 7 were able to uncover new values of $\mathsf{PSCAN}(v,t)$. Specifically, we found $\mathsf{PSCAN}(6,3) = \mathsf{PSCAN}(7,3) = \mathsf{PSCAN}(7,4) = 2$ and $\mathsf{PSCAN}(8,3) = 3$. These methods were effective in cataloguing all $\mathsf{PSCA}(v,3,2)$ but could not do the same for $\mathsf{PSCA}(v,3,3)$. The large number of $\mathsf{PSCA}(6,3,3)$ made a complete exhaustive search infeasible. The placement of a new symbol in the first few rows of a $\mathsf{PSCA}(v,3,3)$ is largely unrestricted. In particular, ignoring constraints on this new symbol that may be imposed by fixing its distribution, there is no restriction on the placement of this symbol in the first three rows. Exhaustively finding all ways of extending a $\mathsf{PSCA}(v,3,3)$ to a $\mathsf{PSCA}(v+1,3,3)$ thus took much longer than finding all ways of extending a $\mathsf{PSCA}(v,3,2)$ to a $\mathsf{PSCA}(v+1,3,2)$. The size of the catalogue of $\mathsf{PSCA}(5,3,3)$ also far exceeded the size of any catalogue of $\mathsf{PSCA}(v,3,2)$. These factors help explain the difficulties we had in compiling a complete catalogue of $\mathsf{PSCA}(6,3,3)$. Our incomplete catalogue of these arrays then made finding a complete catalogue of $\mathsf{PSCA}(v,3,3)$ for $v>6$ impossible. While our partial catalogue was enough to find examples of $\mathsf{PSCA}(7,3,3)$ and $\mathsf{PSCA}(8,3,3)$, we were unable to find a $\mathsf{PSCA}(9,3,3)$. Thus, 9 is the smallest value for v for which $\mathsf{PSCAN}(v,3)$ is still unknown. Based off our complete catalogue of $\mathsf{PSCA}(6,3,2)$ in which arrays where each symbol appeared twice in each column were particularly prevalent, we predicted that if a $\mathsf{PSCA}(9,3,3)$ does exist, we may similarly be able to find one in which every symbol appears exactly twice in every column. As discussed in Chapter 7, we performed a search for such $\mathsf{PSCA}(9,3,3)$ but were not able to find any. More sophisticated methods seem to be required to perform a full search for $\mathsf{PSCA}(9,3,3)$ and determine $\mathsf{PSCAN}(9,3)$.

Conjecture 3.18 and the question of when an $\mathsf{SCA}(t!;t,v)$ exists remains an important open problem. In the $t=7$ case, we have shown that an $\mathsf{SCA}(7!;7,10)$ does not exist. Thus, the only value of v for which the existence of an $\mathsf{SCA}(7!;7,v)$ is still unknown is $v=9$. More generally, we can ask whether the two $\mathsf{SCA}(24;4,6)$ are isolated counter-examples to Conjecture 3.18. The group properties of these arrays are compelling and in general, a representation of $\mathcal{S}_t$ in $\mathcal{S}_{t+2}$ seems a strong candidate for building an $\mathsf{SCA}(t!;t,t+2)$. However, given the conjecture is known to hold for some values of t, such a group representation cannot always be used to build an $\mathsf{SCA}(t!;t,t+2)$. One direction for future research may be to explore representations of $\mathcal{S}_t$ in $\mathcal{S}_{t+2}$ to see whether other group-based counter-examples to Conjecture 3.18 exist, or whether there is some demonstrably special property about $\mathcal{S}_4$ that allows a representation of this group in $\mathcal{S}_6$ to form a strength 4 SCA.

In Chapter 9, we saw that excess coverage arrays were a useful tool in addressing Conjecture 3.18. Excess coverage arrays may also be of independent interest, particularly in the case when $t=2$ and $N=v(v+1)$ where such arrays are quite closely related to orthogonal arrays. While stronger non-existence results for these arrays would have meaningful consequences for Conjecture 3.18, general constructions may also provide interesting insights. For $v=2$, we found a strength 2 excess coverage array with $v+2$ columns. Are there other values of v for which a $\mathsf{CA}_X(v(v+1);2,v+2,v)$ exists?

In Chapter 3, we saw that the best upper bound for $\mathsf{SCAN}(3,v)$ came from a construction due to Tarui [53], while for $t\geqslant 4$, the best upper bound for $\mathsf{SCAN}(t,v)$ is due to Yuster [63] and comes from an application of the Lovász Local Lemma. No general constructions of logarithmic size exist for sequence covering arrays with $t\geqslant 4$ with the only constructive results coming from applications of different computational algorithms applied to small values of v. However, comparing these algorithmic bounds to Yuster's upper bound reveals that the algorithmic bounds provide tighter upper bounds for $\mathsf{SCAN}(4, v)$ and $\mathsf{SCAN}(5, v)$ for all values of v for which we have such data (see Table 3.4 and 3.5). Although constructions for perfect sequence covering arrays are necessarily constructions for sequence covering arrays, this is an inefficient method for building sequence covering arrays given that $\mathsf{PSCAN}(v, t)$ is in general much larger than $\mathsf{SCAN}(t, v)$.

Bibliography

[1] N. Alon, L. Babai, and A. Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *J. Algorithms*, 7(4):567–583, 1986.

[2] S. Ball and M. Lavrauw. Arcs in finite projective spaces. *EMS Surv. Math. Sci.*, 6(1-2):133–172, 2019.

[3] M. Banbara, N. Tamura, and K. Inoue. Generating event-sequence test cases by answer set programming with the incidence matrix. In *Technical Communications of the 28th International Conference on Logic Programming*, volume 17 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 86–97. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2012.

[4] M. Basavaraju, L. S. Chandran, M. C. Golumbic, R. Mathew, and D. Rajendraprasad. Separation dimension of graphs and hypergraphs. *Algorithmica*, 75(1):187–204, 2016.

[5] F. E. Bennett and A. Mahmoodi. Directed designs. In C. J. Colbourn and J. H. Dinitz, editors, *Handbook of Combinatorial Designs*, pages 441–445. Chapman & Hall/CRC, New York, second edition, 2006.

[6] M. Brain, E. Erdem, K. Inoue, J. Oetsch, J. Pührer, H. Tompits, and C. Yilmaz. Event-sequence testing using answer-set programming. *International Journal on Advances in Software*, 5(3&4), 2012.

[7] A. Z. Broder, M. Charikar, A. M. Frieze, and M. Mitzenmacher. Min-wise independent permutations. volume 60, pages 630–659. 2000. 30th Annual ACM Symposium on Theory of Computing (Dallas, TX, 1998).

[8] P. J. Cameron. *Permutation groups*, volume 45 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1999.

[9] Y. M. Chee, C. J. Colbourn, D. Horsley, and J. Zhou. Sequence covering arrays. *SIAM J. Discrete Math.*, 27(4):1844–1861, 2013.

[10] C. J. Colbourn. Covering array tables for $t = 2, 3, 4, 5, 6$. https://www.public.asu.edu/~ccolbou/src/tabby/catable.html.

[11] C. J. Colbourn. Covering arrays and hash families. In *Information security, coding theory and related combinatorics*, volume 29 of *NATO Sci. Peace Secur. Ser. D Inf. Commun. Secur.*, pages 99–135. IOS, Amsterdam, 2011.

[12] P. Danziger, E. Mendelsohn, L. Moura, and B. Stevens. Covering arrays avoiding forbidden edges. *Theoret. Comput. Sci.*, 410(52):5403–5414, 2009.

[13] F. Duan, Y. Lei, R. N. Kacker, and D. R. Kuhn. An approach to t-way test sequence generation with constraints. In *2019 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, pages 241–250, 2019.

[14] B. Dushnik. Concerning a certain set of arrangements. *Proc. Amer. Math. Soc.*, 1:788–796, 1950.

[15] B. Dushnik and E. W. Miller. Partially ordered sets. *Amer. J. Math.*, 63:600–610, 1941.

[16] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vols. I, II, III*, volume Vol. 10 of *Colloq. Math. Soc. János Bolyai*, pages 609–627. North-Holland, Amsterdam-London, 1975.

[17] M. L. Fredman and J. Komlós. On the size of separating systems and families of perfect hash functions. *SIAM J. Algebraic Discrete Methods*, 5(1):61–68, 1984.

[18] Z. Füredi. Scrambling permutations and entropy of hypergraphs. *Random Structures Algorithms*, 8(2):97–104, 1996.

[19] A. Gentle. A polynomial construction of perfect sequence covering arrays. *Algebr. Comb.*, 6(5):1383–1394, 2023.

[20] A. Gentle and I. M. Wanless. On perfect sequence covering arrays. *Ann. Comb.*, 27(3):539–564, 2023.

[21] A. S. Hedayat, N. J. A. Sloane, and J. Stufken. *Orthogonal arrays*. Springer Series in Statistics. Springer-Verlag, New York, 1999. Theory and applications, With a foreword by C. R. Rao.

[22] J. W. P. Hirschfeld. *Projective geometries over finite fields*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1979.

[23] Y. Ishigami. An extremal problem of orthants containing at most one point besides the origin. *Discrete Math.*, 132(1-3):383–386, 1994.

[24] Y. Ishigami. Containment problems in high-dimensional spaces. *Graphs Combin.*, 11(4):327–335, 1995.

[25] Y. Ishigami. An extremal problem of d permutations containing every permutation of every t elements. *Discrete Math.*, 159(1-3):279–283, 1996.

[26] T. Itoh, Y. Takei, and J. Tarui. On permutations with limited independence. In *Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms (San Francisco, CA, 2000)*, pages 137–146. ACM, New York, 2000.

[27] T. Itoh, Y. Takei, and J. Tarui. On the sample size of k-restricted min-wise independent permutations and other distributions. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, pages 710–719. ACM, New York, 2003.

[28] E. Iurlano. Growth of the perfect sequence covering array number. *Des. Codes Cryptogr.*, 91(4):1487–1494, 2023.

[29] D. S. Johnson. Approximation algorithms for combinatorial problems. *J. Comput. System Sci.*, 9:256–278, 1974.

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

[31] A. Klein. On perfect deletion-correcting codes. *J. Combin. Des.*, 12(1):72–77, 2004.

[32] D. J. Kleitman, J. Shearer, and D. Sturtevant. Intersections of k-element sets. *Combinatorica*, 1(4):381–384, 1981.

[33] D. J. Kleitman and J. Spencer. Families of k-independent sets. *Discrete Math.*, 6:255–262, 1973.

[34] D. R. Kuhn, J. M. Higdon, J. F. Lawrence, R. N. Kacker, and Y. Lei. Combinatorial methods for event sequence testing. In *2012 IEEE Fifth International Conference on Software Testing, Verification and Validation*, pages 601–609, 2012.

[35] G. Kuperberg, S. Lovett, and R. Peled. Probabilistic existence of regular combinatorial structures. *Geom. Funct. Anal.*, 27(4):919–972, 2017.

[36] V. Levenshtein. Perfect deletion-correcting codes as combinatorial designs. In *II International Workshop Algebraic and Combinatorial Coding Theory*, pages 137–140, 1990.

[37] V. I. Levenshtein. Perfect codes in the metric of deletions and insertions. *Diskret. Mat.*, 3(1):3–20, 1991.

[38] E. Marczewski. Indépendance d'ensembles et prolongement de mesures (résultats et problèmes). *Colloq. Math.*, 1:122–132, 1948.

[39] O. Margalit. Better bounds for event sequencing testing. In *2013 IEEE Sixth International Conference on Software Testing, Verification and Validation Workshops*, pages 281–284, 2013.

[40] R. Mathon. Searching for spreads and packings. In *Geometry, combinatorial designs and related structures (Spetses, 1996)*, volume 245 of *London Math. Soc. Lecture Note Ser.*, pages 161–176. Cambridge Univ. Press, Cambridge, 1997.

[41] R. Mathon and V. T. Tran. Directed t-packings and directed t-Steiner systems. volume 18, pages 187–198. 1999. Designs and codes—a memorial tribute to Ed Assmus.

[42] P. C. Murray and C. J. Colbourn. Sequence covering arrays and linear extensions. In *Combinatorial algorithms*, volume 8986 of *Lecture Notes in Comput. Sci.*, pages 274–285. Springer, Cham, 2015.

[43] J. Na. Perfect sequence covering arrays. Master's thesis, Simon Fraser University, 2021.

[44] J. Na, J. Jedwab, and S. Li. A group-based structure for perfect sequence covering arrays. *Des. Codes Cryptogr.*, 91(3):951–970, 2023.

[45] J. Opatrný. Total ordering problem. *SIAM J. Comput.*, 8(1):111–114, 1979.

[46] C. Radhakrishna Rao. Factorial experiments derivable from combinatorial arrangements of arrays. *Suppl. J. Roy. Statist. Soc.*, 9:128–139, 1947.

[47] J. Radhakrishnan. A note on scrambling permutations. *Random Structures Algorithms*, 22(4):435–439, 2003.

[48] Z. B. Ratliff, D. R. Kuhn, and D. J. Ragsdale. Detecting vulnerabilities in android applications using event sequences. In *2019 IEEE 19th International Conference on Software Quality, Reliability and Security (QRS)*, pages 159–166, 2019.

[49] J. B. Shearer. On a problem of Spencer. *Combinatorica*, 5(3):241–245, 1985.

[50] J. Singer. A theorem in finite projective geometry and some applications to number theory. *Trans. Amer. Math. Soc.*, 43(3):377–385, 1938.

[51] J. Spencer. Minimal scrambling sets of simple orders. *Acta Math. Acad. Sci. Hungar.*, 22:349–353, 1971/72.

[52] Y. Sun, Y. Zhang, and G. Ge. Improved constructions of permutation and multi-permutation codes correcting a burst of stable deletions. *IEEE Transactions on Information Theory*, 69(7):4429–4441, 2023.

[53] J. Tarui. On the minimum number of completely 3-scrambling permutations. *Discrete Math.*, 308(8):1350–1354, 2008.

[54] J. Tarui, T. Itoh, and Y. Takei. A nearly linear size 4-min-wise independent permutation family by finite geometries. In *Approximation, randomization, and combinatorial optimization*, volume 2764 of *Lecture Notes in Comput. Sci.*, pages 396–408. Springer, Berlin, 2003.

[55] J. A. Thas. Some results concerning $\{(q+1)(n-1); n\}$-arcs and $\{(q+1)\ (n-1)+1; n\}$-arcs in finite projective planes of order q. *J. Combinatorial Theory Ser. A*, 19(2):228–232, 1975.

[56] The GAP Group. Gap - groups, algorithms, programming - a system for computational discrete algebra. https://www.gap-system.org/index.html.

[57] J. Torres-Jimenez, D. O. Ramirez-Acuna, B. Acevedo-Juárez, and H. Avila-George. New upper bounds for sequence covering arrays using a 3-stage approach. *Expert Systems with Applications*, 207:118022, 2022.

[58] R. R. Varšamov and G. M. Tenengolts. A code which corrects single asymmetric errors. *Avtomat. i Telemeh.*, 26:288–292, 1965.

[59] W. Wang, S. Sampath, Y. Lei, and R. Kacker. An interaction-based test sequence generation approach for testing web applications. In *2008 11th IEEE High Assurance Systems Engineering Symposium*, pages 209–218. IEEE, 2008.

[60] R. M. Wilson. A diagonal form for the incidence matrices of t-subsets vs. k-subsets. *European J. Combin.*, 11(6):609–615, 1990.

[61] M. Yannakakis. The complexity of the partial order dimension problem. *SIAM J. Algebraic Discrete Methods*, 3(3):351–358, 1982.

[62] X. Yuan, M. B. Cohen, and A. M. Memon. Towards dynamic adaptive automated test generation for graphical user interfaces. In *2009 International Conference on Software Testing, Verification, and Validation Workshops*, pages 263–266. IEEE, 2009.

[63] R. Yuster. Covering small subgraphs of (hyper)tournaments with spanning acyclic subgraphs. *Electron. J. Combin.*, 27(4):Paper No. 4.13, 15, 2020.

[64] R. Yuster. Perfect sequence covering arrays. *Des. Codes Cryptogr.*, 88(3):585–593, 2020.

[65] R. Yuster. Perfect and nearly perfect separation dimension of complete and random graphs. *J. Combin. Des.*, 29(11):786–805, 2021.