Department of Mathematical Sciences, Clemson University

http://www.ces.clemson.edu/ keyj/

Designs and codes: partial permutation decoding

J. D. Key

keyj@ces.clemson.edu

Abstract

Codes defined through the row-span over finite fields of incidence matrices of designs have many properties that can be deduced from the combinatorial properties of the designs. In particular those codes that come from designs defined by finite geometries have been used both in applications and for classification purposes within design theory.

We look here at their applicability to a method of decoding that was introduced by MacWilliams in the early 60s, viz. permutation decoding. Her notion of PD-sets is generalized to that of partial PD-sets that can be used to correct some number of errors possibly less than the full capability of the code. These have been found for some infinite classes of codes from finite planes and graphs.

Design theory background

An incidence structure $\mathcal{D} = (\mathcal{P}, \mathcal{B}, \mathcal{I})$, with point set $\mathcal{P}$, block set $\mathcal{B}$ and incidence $\mathcal{I}$ is a t-(v, k, λ) design, if $|\mathcal{P}| = v$, every block $B \in \mathcal{B}$ is incident with precisely k points, and every t distinct points are together incident with precisely λ blocks.

E.g. A 2-$(n^2 + n + 1, n + 1, 1)$ is a projective plane of order n;
a $2 - (16, 6, 2)$ is a biplane.

The code C_F of the design $\mathcal{D}$ over the finite field F is the space spanned by the incidence vectors of the blocks over F, i.e. the row span over F of an incidence matrix: see [AK92, AK96] for more.

Permutation decoding

Permutation decoding was first developed by Jessie MacWilliams [Mac64]. It can be used when a code has sufficiently many automorphisms to ensure the existence of a set of automorphisms called a PD-set. We extend the definition of PD-sets to s-PD-sets for s-error-correction [KMM]:

Definition 1 *If C is a t-error-correcting code with information set $\mathcal{I}$ and check set $\mathcal{C}$, then a PD-set for C is a set $\mathcal{S}$ of automorphisms of C which is such that every t-set of coordinate positions is moved by at least one member of $\mathcal{S}$ into the check positions $\mathcal{C}$.*

For $s \leq t$ an s-PD-set is a set $\mathcal{S}$ of automorphisms of C which is such that every s-set of coordinate positions is moved by at least one member of $\mathcal{S}$ into $\mathcal{C}$.

More specifically, if $\mathcal{I} = \{1, \ldots, k\}$ are the information positions and $\mathcal{C} = \{k+1, \ldots, n\}$ the check positions, then every s-tuple from $\{1, \ldots, n\}$ can be moved by some element of $\mathcal{S}$ into $\mathcal{C}$.

Coding theory terminology

- A **linear code** is a subspace of a finite-dimensional vector space over a finite field.
- The **weight** of a vector is the number of non-zero coordinate entries. If a code has smallest non-zero weight d then the code can correct up to $\lfloor \frac{d-1}{2} \rfloor$ errors by nearest-neighbour decoding.
- If a code C over a field of order q is of length n, dimension k, and minimum weight d, then we write $[n, k, d]_q$ to show this information.
- A **generator matrix** for the code is a $k \times n$ matrix made up of a basis for C.
- The **dual** code $C^\perp$ is the orthogonal under the standard inner product $(,)$, i.e. $C^\perp = \{v \in F^n | (v, c) = 0 \text{ for all } c \in C\}$.
- A **check** matrix for C is a generator matrix H for $C^\perp$.

Coding theory terminology continued

- Two linear codes of the same length and over the same field are **isomorphic** if they can be obtained from one another by permuting the coordinate positions.
- An **automorphism** of a code C is an isomorphism from C to C.
- Any code is isomorphic to a code with generator matrix in **standard form**, i.e. the form $[I_k \mid A]$; a check matrix then is given by $[-A^T \mid I_{n-k}]$. The first k coordinates are the **information symbols** and the last $n-k$ coordinates are the **check symbols**.

Algorithm for permutation decoding

C is a q-ary t-error-correcting $[n, k, d]_q$ code; $d = 2t + 1$ or $2t + 2$.

$k \times n$ generator matrix for C: $G = [I_k|A]$.

Any k-tuple v is encoded as vG. The first k columns are the information symbols, the last $n - k$ are check symbols.

$(n - k) \times n$ check matrix for C: $H = [-A^T|I_{n-k}]$.

Suppose x is sent and y is received and at most t errors occur.

$\mathcal{S} = \{g_1, \ldots, g_m\}$ is a PD-set for C.

- For $i = 1, \ldots, m$, compute yg_i and the syndrome $s_i = H(yg_i)^T$ until an i is found such that the weight of s_i is t or less;
- if $u = u_1u_2 \ldots u_k$ are the information symbols of yg_i, compute the codeword $c = uG$;
- decode y as cg_i^{-1}.

Why permutation decoding works

Result 1 *Let C be an $[n, k, d]_q$ t-error-correcting code. Suppose H is a check matrix for C in standard form, i.e. such that I_{n-k} is in the redundancy positions. Let $y = c+e$ be a vector, where $c \in C$ and e has weight $\leq t$. Then the information symbols in y are correct if and only if the weight of the syndrome Hy^T of y is $\leq t$.*

Minimum size for a PD-set

Counting shows that there is a minimum size a PD-set can have; most the sets known have size larger than this minimum. The following is due to Gordon [Gor82], using a result of Schönheim [Sch64]:

Result 2 *If $\mathcal{S}$ is a PD-set for a t-error-correcting $[n, k, d]_q$ code C, and $r = n - k$, then*

$$|\mathcal{S}| \geq \left\lceil \frac{n}{r} \left\lceil \frac{n-1}{r-1} \left\lceil \cdots \left\lceil \frac{n-t+1}{r-t+1} \right\rceil \cdots \right\rceil \right\rceil \right\rceil.$$

(Proof in Huffman [Huf98].)

This result can be adapted to s-PD-sets for $s \leq t$ by replacing t by s in the formula.

Single error

Correcting a single error is, in fact, simply done by using syndrome decoding, since in that case multiples of the columns of the check matrix will give the possible syndromes. Thus the syndrome of the received vector need only be compared with the columns of the check matrix, by looking for a multiple.

So we look for s-PD-sets for $s \geq 2$.

Classes of codes having s-PD-sets

- If $\mathrm{Aut}(C)$ is k-transitive then $\mathrm{Aut}(C)$ itself is a k-PD-set, in which case we attempt to find smaller sets;
- existence of a k-PD-set is not invariant under isomorphism (not equivalence) of codes;
- codes from the row span over a finite field of an incidence matrix of a design or geometry, or from an adjacency matrix of a graph;
- using Result 2 it follows that many classes of designs and graphs where the minimum-weight and automorphism group are known, cannot have PD-sets for full error-correction for length beyond some bound; for these we look for s-PD-sets with $2 \leq s < \lfloor \frac{d-1}{2} \rfloor$: e.g. finite planes, Paley graphs;
- for some classes of regular and semi-regular graphs with lasrge automorphism groups, PD-sets exist for all lengths: e.g. triangular graphs, lattice graphs, line graphs of multi-partite graphs.

Some infinite classes of codes having PD-sets

In all of these, suitable information sets had to be found.

1. Triangular graphs

For any n, the triangular graph $T(n)$ is the line graph of the complete graph K_n, and is strongly regular. The row span over $\mathbb{F}_2$ of an adjacency matrix gives codes:
$[\frac{n(n-1)}{2}, n-1, n-1]_2$ for n odd and
$[\frac{n(n-1)}{2}, n-2, 2(n-1)]_2$ for n even
where $n \geq 5$.

The automorphism group is S_n acting naturally (apart from $n = 5$) and get PD-sets of size n for n odd and $n^2 - 2n + 2$ for n even, by [KMR04b].

(The computational complexity of the decoding by this method may be quite low, of the order $n^{1.5}$ if the elements of the PD-set are appropriately ordered. The codes are low density parity check (LDPC) codes.)

2. Lattice graphs

The (square) lattice graph $L_2(n)$ is the line graph of the complete bipartite graph $K_{n,n}$, and is strongly regular. The row span over $\mathbb{F}_2$ of an adjacency matrix gives codes: $[n^2, 2(n-1), 2(n-1)]_2$ for $n \geq 5$ with $S_n \wr S_2$ as automorphism group, and PD-sets of size n^2 in $S_n \times S_n$ were found in [KS].

(The lower bound from Result 2 is $O(n)$.)

A similar result holds for the lattice graph $L_2(m,n)$: the codes are
$[mn, m+n-2, 2m]_2$ for $m+n$ even,
$[mn, m+n-1, m]_2$ for $m+n$ odd.
PD-sets of size mn in $S_m \times S_n$ can be found.

More generally for the line graphs of multi-partite graphs, with automorphism group $S_{n_1} \times S_{n_2} \times \ldots \times S_{n_m}$: work in progress, [Sen].

3. Graphs on triples

Define three graphs with vertex set the subsets of size three of a set of size n and adjacency according to the size of the intersection of the 3-subsets. Properties od these codes are in [KMR04a]. Again S_n in its natural action is the automorphism group.

If C is the binary code in the case of adjacency if the 3-subsets intersect in two elements, then the dual $C^{\perp}$ is a $[\binom{n}{3}, \binom{n-1}{2}, n-2]_2$ code and a PD-set of n^3 can be found by [KMR].

Some infinite classes of codes only having partial PD-sets

1. Finite planes

If $q = p^e$ where p is prime, the code of the desarguesian projective plane of order q has parameters: $C = [q^2 + q + 1, (\frac{p(p+1)}{2})^e + 1, q + 1]_p$. For the affine plane the code is $[q^2, (\frac{p(p+1)}{2})^e, q]_p$. The codes are subfield subcodes of the generalized Reed-Muller codes, and the automorphism groups are the semi-linear groups and doubly transitive.

Thus 2-PD-sets always exist but the bound of Result 2 is greater than the size of the group (see [KMM]) in the projective desarguesian case when:

$q = p$ prime and $p > 103$;

$q = 2^e$ and $e > 12$;

$q = 3^e$ and $e > 6$;

$q = 5^e$ and $e > 4$;

$q = 7^e$ and $e > 3$;

$q = 11^e$ and $e > 2$;

$q = 13^e$ and $e > 2$;

$q = p^e$ for $p > 13$ and $e > 1$.

Similar results hold for the affine and dual cases.

Small 2-PD-sets in prime-order planes

2-PD-sets exist for any information set (since the group is 2-transitive); for prime order, using a Moorhouse [Moo91] basis,
2-PD-sets of 37 elements for desarguesian affine planes of any prime order p and
2-PD-sets of 43 elements for desarguesian projective planes of any prime order p were constructed in [KMM].

Also 3-PD-sets for the code and the dual code in the affine prime case of sizes $2p^2(p-1)$ and p^2, respectively, were found.

Other orders q and other codes from geometries will likely yield similar results.

2. Paley graphs

If n is a prime power with $n \equiv 1 \pmod 4$, the **Paley graph** ,$P(n)$, has $\mathbb{F}_n$ as vertex set and two vertices x and y are adjacent if and only if $x - y$ is a non-zero square in $\mathbb{F}_n$. The row span over a field $\mathbb{F}_p$ of an adjacency matrix gives an interesting code (quadratic residue codes) if and only if p is a square in $\mathbb{F}_n$.

For any $\sigma \in Aut(\mathbb{F}_n)$ and $a, b \in \mathbb{F}_n$ with a a non-zero square, the group of maps $\tau_{a,b,\sigma} : x \mapsto ax^\sigma + b$ is the automorphism group of the code, and for $n \geq 1697$ and prime or $n \geq 1849$ and a square, PD-sets cannot exist since the bound of Result 2 is bigger than the order of the group (using the square root bound for the minimum weight, and the actual minimum weight $q + 1$ when $n = q^2$ and q is a prime power).

For the case where n is prime and $n \equiv 1 \pmod 8$, the code of $P(n)$ over $\mathbb{F}_p$ is $C = [n, \frac{n-1}{2}, d]_p$ where $d \geq \sqrt{n}$,
(the square-root bound) for p any prime dividing $\frac{n-1}{4}$.

C has a 2-PD-set of size 6 by [KL]. (The automorphism group is not 2-transitive.)

For the dual code in this case, a 2-PD-set of size 10 for all n was found. Further results in [Lim].

References

[AK92] E. F. Assmus, Jr. and J. D. Key. *Designs and their Codes*. Cambridge: Cambridge University Press, 1992. Cambridge Tracts in Mathematics, Vol. 103 (Second printing with corrections, 1993).

[AK96] E. F. Assmus, Jr. and J. D. Key. Designs and codes: an update. *Des. Codes Cryptogr.*, 9:7–27, 1996.

[Gor82] D. M. Gordon. Minimal permutation sets for decoding the binary Golay codes. *IEEE Trans. Inform. Theory*, 28:541–543, 1982.

[Huf98] W. Cary Huffman. Codes and groups. In V. S. Pless and W. C. Huffman, editors, *Handbook of Coding Theory*, pages 1345–1440. Amsterdam: Elsevier, 1998. Volume 2, Part 2, Chapter 17.

[KL] J. D. Key and J. Limbupasiriporn. Permutation decoding of codes from Paley graphs. Submitted.

[KMM] J. D. Key, T. P. McDonough, and V. C. Mavron. Partial permutation

decoding of codes from finite planes. *European J. Combin.*, To appear.

[KMR] J. D. Key, J. Moori, and B. G. Rodrigues. Permutation decoding of binary codes from graphs on triples. *Ars Combin.* To appear.

[KMR04a] J. D. Key, J. Moori, and B. G. Rodrigues. Binary codes from graphs on triples. *Discrete Math.*, 282/1-3:171–182, 2004.

[KMR04b] J. D. Key, J. Moori, and B. G. Rodrigues. Permutation decoding for binary codes from triangular graphs. *European J. Combin.*, 25:113–123, 2004.

[KS] J. D. Key and P. Seneviratne. Permutation decoding of binary codes from lattice graphs. *Discrete Math.*, To appear.

[Lim] J. Limbupasiriporn. Ph.D. thesis, Clemson University, 2004.

[Mac64] F. J. MacWilliams. Permutation decoding of systematic codes. *Bell System Tech. J.*, 43:485–505, 1964.

[Moo91] G. Eric Moorhouse. Bruck nets, codes, and characters of loops. *Des. Codes Cryptogr.*, 1:7–29, 1991.

[Sch64] J. Schönheim. On coverings. *Pacific J. Math.*, 14:1405–1411, 1964.

[Sen] P. Seneviratne. In preparation.