New Methods in Coding Theory. Error-Correcting Codes and the Shannon Capacity
S.C. Polak

Summary

New methods in coding theory: error-correcting codes and the Shannon capacity

Error-correcting codes have been studied since 1948, when Claude Shannon published his influential paper *A Mathematical Theory of Communication* [85]. Fix three positive integers q, n, d and let $Q := \{0, \ldots, q-1\}$ be our *alphabet*. We identify elements of Q^n with *words* of length n consisting of letters (elements) from Q. A *code of length* n is any subset of Q^n. For two words u, v, their *Hamming distance* is the number of i with $u_i \neq v_i$. The *minimum distance* of a code C is the minimum Hamming distance between any two distinct elements of C. The central question in coding theory is the following.

$$\text{What is the maximum size of a code } C \subseteq Q^n \text{ with minimum distance at least } d? \quad \text{(S.1)}$$

The maximum size in (S.1) is denoted by $A_q(n, d)$. A code C with minimum distance $d := 2e + 1$ (for some integer e) is called *e-error-correcting*. If a codeword from C is distorted in at most e positions, we can recover the original codeword by taking the codeword that is closest to the distorted codeword in Hamming distance. This principle is used in communication systems for the correction of transmission errors.

The numbers $A_q(n, d)$ are hard to compute in general. For many q, n, d, only upper and lower bounds are known. Explicit codes yield lower bounds on $A_q(n, d)$ and they can be used for error-correction as explained. A classical upper bound on $A_q(n, d)$ is Delsarte's linear programming upper bound [31], which bound can be interpreted as a semidefinite programming (SDP) bound based on pairs of codewords.

In this thesis we try to improve upper bounds on $A_q(n, d)$. We give an SDP bound based on quadruples of codewords, which a priori has size exponential in n. The optimization problem is highly symmetric: it can be assumed that the optimal solution is invariant under the group of distance preserving permutations of Q^n. This symmetry group is the wreath product $S_q^n \rtimes S_n$. By the symmetry of the problem, the SDP can be reduced to a size bounded by a polynomial in n.

In Chapter 3 we give a general method for symmetry reduction, based on representation theory. If G is a finite group acting on a finite set Z and $n \in \mathbb{N}$, we give a reduction of $Z^n \times Z^n$-matrices which are invariant under the simultaneous action of the group $G^n \rtimes S_n$ on their rows and columns. In the reduction, we assume that a reduction is known of $Z \times Z$-matrices which are invariant under the simultaneous action of G on their rows and columns.

In Chapter 4 we apply this general method to reduce the mentioned SDP based on quadruples of codewords for computing upper bounds on $A_q(n, d)$. With the method, we sharpen known upper bounds for five triples (q, n, d).

In Chapter 5, we explore other methods of finding upper bounds on $A_q(n, d)$, based on combinatorial divisibility arguments. The methods yield new upper bounds for four

triples (q, n, d). Our most prominent result in this direction is the following bound, which gives in certain cases a strengthening of a bound implied by the Plotkin bound [67].

Theorem. *Suppose that q, n, d, m are positive integers with $q \geq 2$, such that $d = m(qd - (q-1)(n-1))$, and such that $n-d$ does not divide $m(n-1)$. If $r \in \{1, \ldots, q-1\}$ satisfies*

$$n(n-1-d)(r-1)r < (q-r+1)(qm(q+r-2) - 2r),$$

then $A_q(n, d) < q^2 m - r$.

In Chapter 6 we consider (binary) *constant weight* codes. Here the alphabet is $\{0, 1\}$. The *weight* of a word is the number of 1's it contains. For $n, d, w \in \mathbb{N}$, the number $A(n, d, w)$ denotes the maximum size of a code $C \subseteq \{0, 1\}^n$ with minimum distance at least d and in which every codeword has weight w. (Such a code is called a 'constant weight code' with weight w.) With SDP based on quadruples of codewords and a symmetry reduction with the method of Chapter 3, we find several new upper bounds on $A(n, d, w)$. Two upper bounds matching the known lower bounds are obtained, so that we know the value of $A(n, d, w)$ exactly: $A(22, 8, 10) = 616$ and $A(22, 8, 11) = 672$.

In Chapter 7, we prove with the SDP-output, using 'complementary slackness', that the optimal constant weight codes achieving $A(23, 8, 11) = 1288$, $A(22, 8, 10) = 616$ and $A(22, 8, 11) = 672$ are unique up to coordinate permutations. The mentioned unique constant weight codes can be obtained from the binary Golay code —a famous code with good error-correcting properties— by taking subcodes and deleting coordinates ('shortening').

1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0	0	1	0	1
0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	1	0	1	1
0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	1	1	1
0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	1	1	0	1
0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1	1	0	1	1
0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	1	1	0	1	1	1
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1	1
0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	1	0	1	1	1	0	1
0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	1	0	1	1	1	0	0	1
0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	1	1	1	0	0	0	1
0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	1	1	1	0	0	0	1	1
0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0

Figure S.1: A generator matrix of the extended binary Golay code (i.e., the 2^{12} codewords are sums mod 2 of the rows of this matrix). The extended binary Golay code was used for error-correction in the *Voyager* missions to Jupiter and Saturn [94].

For unrestricted (non-constant weight) binary codes, the bound $A_2(20, 8) \leq 256$ was obtained by Gijswijt, Mittelmann and Schrijver [37], implying that the quadruply shortened extended binary Golay code of size 256 is optimal. Two unrestricted codes $C, D \subseteq \{0, 1\}^n$ are equivalent if there is a $g \in S_2^n \rtimes S_n$ such that $g \cdot C = D$. Up to equivalence the optimal binary codes attaining $A_2(24-i, 8) = 2^{12-i}$ for $i = 0, 1, 2, 3$ are unique, namely they are the i times shortened extended binary Golay codes [20]. We show that there exist several nonequivalent optimal codes achieving $A_2(20, 8) = 256$. We classify such codes under the additional condition that all distances are divisible by 4, and find 15 such codes. We also show that there exist such codes with not all distances divisible by 4.

In Chapter 8 we consider *Lee codes*. Fix three integers $q, n, d \in \mathbb{N}$ and define $Q := \mathbb{Z}_q$ (the cyclic group of order q). For two words $u, v \in Q^n$, their *Lee distance* is $\sum_{i=1}^n \min\{|u_i - v_i|, q - |u_i - v_i|\}$. The *minimum Lee distance* of a code $C \subseteq Q^n$ is the minimum Lee distance between any two distinct elements of C. Let $A_q^L(n, d)$ denote the maximum

size of a code $C \subseteq Q^n$ with minimum Lee distance at least d. We give an SDP upper bound based on triples of codewords and show that it can be computed efficiently, using the symmetry reduction method of Chapter 3. This finally yields new upper bounds on $A_q^L(n,d)$ for several triples (q,n,d).

Chapter 9 is about the Shannon capacity of circular graphs. For any graph $G = (V,E)$ and $n \in \mathbb{N}$, the *n-th strong product power* $G^{\boxtimes n}$ is the graph with vertex set V^n in which two distinct vertices $(u_1,\ldots,u_n)$ and $(v_1,\ldots,v_n)$ of $G^{\boxtimes n}$ are adjacent if and only if for each $i \in \{1,\ldots,n\}$ one has either $u_i = v_i$ or $u_i v_i \in E$. The *Shannon capacity* of G is defined as

$$\Theta(G) := \sup_{n\in\mathbb{N}} \sqrt[n]{\alpha(G^{\boxtimes n})},$$

where for any graph G, the maximum cardinality of an independent set in G (a set of vertices, no two of which are adjacent) is denoted by $\alpha(G)$. The circular graph $C_{d,q}$ is the graph with vertex set $\mathbb{Z}_q$, in which two distinct vertices are adjacent if and only if their distance (mod q) is strictly less than d. The value of $\alpha(C_{d,q}^{\boxtimes n})$ (for fixed n) and $\Theta(C_{d,q})$ can be seen to only depend on the quotient q/d. We show that the function $q/d \mapsto \Theta(C_{d,q})$ is continuous at *integer* points $q/d \geq 3$. It implies that also the function $q/d \mapsto \vartheta(C_{d,q})$, Lovász's upper bound on $\Theta(C_{d,q})$ [53], is continuous at these points — see Figure S.2.

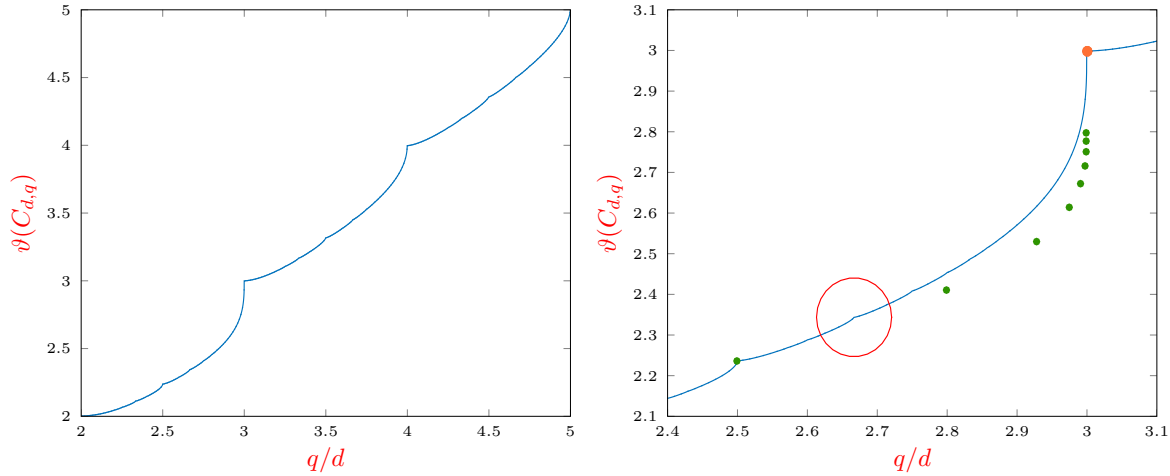

Figure S.2: Two graphs of the function $q/d \mapsto \vartheta(C_{d,q})$. The green points (converging to the orange point $(3,3)$) are some of our lower bounds on $\Theta(C_{d,q})$, which are used to prove left-continuity of $q/d \mapsto \Theta(C_{d,q})$ at integers ≥ 3.

Left-continuity we derive from the following result (proved using an explicit construction).

Theorem. *For each $r,n \in \mathbb{N}$ with $r \geq 3$, we have*

$$\max_{\frac{q}{d}<r} \alpha(C_{d,q}^{\boxtimes n}) = \frac{1+r^n(r-2)}{r-1}.$$

We also prove that the independent set achieving $\alpha(C_{5,14}^{\boxtimes 3}) = 14$, one of the independent sets used in our proof, is unique up to Lee equivalence. Here two sets $C, D \subseteq \mathbb{Z}_q^n$ are *Lee equivalent* if there is a $g \in D_q^n \rtimes S_n$ with $g \cdot C = D$, where D_q is the dihedral group of order $2q$. We adapt our SDP upper bound for Lee codes to compute upper bounds

on $\alpha(C_{d,q}^{\boxtimes n})$. Finally, we give a new lower bound of $367^{1/5} > 3.2578$ on the Shannon capacity of the 7-cycle.

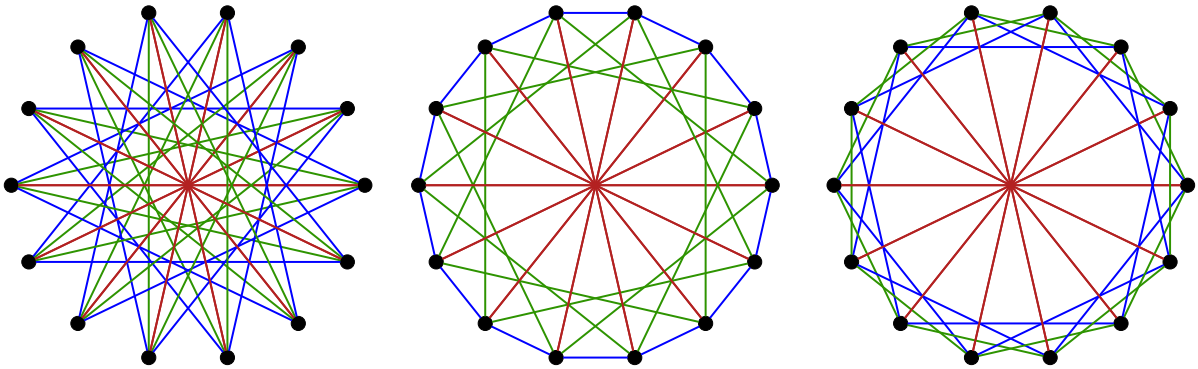

Figure S.3: Three graphs used in the proof to show that the independent set achieving $\alpha(C_{5,14}^{\boxtimes 3}) = 14$ is unique up to Lee equivalence.

An overarching theme: independent sets in graph products.

For any graph $G = (V, E)$, define the number

$$\alpha_d(G) := \max\{|U| \mid U \subseteq V,\ d_G(u, v) \geq d \text{ for all distinct } u, v \in U\}. \tag{S.2}$$

Here $d_G(u, v)$ denotes the smallest length (in edges) of a path between u and v in G. So $\alpha_2(G) = \alpha(G)$. Let K_q denote the complete graph on q vertices, and let C_q be the circuit on q vertices. Then

$$\begin{aligned} A_q(n, d) &= \alpha_d(K_q^{\square n}), \\ A_q^L(n, d) &= \alpha_d(C_q^{\square n}), \\ A_q^{L_\infty}(n, d) := \alpha(C_{d,q}^{\boxtimes n}) &= \alpha_d(C_q^{\boxtimes n}). \end{aligned}$$

Here $G^{\square n}$ denotes the *n-th Cartesian product power of G*: the graph with vertex set $V(G)^n$ in which two distinct vertices $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ are adjacent if and only if there is an $i \in \{1, \ldots, n\}$ such that $u_i v_i \in E$, and $u_j = v_j$ for all $j \neq i$.

So the main objects studied in this thesis are of the form $\alpha_d(G^n)$, where $G \in \{C_q, K_q\}$, and where G^n denotes either $G^{\boxtimes n}$ or $G^{\square n}$. Moreover, we have $A(n, d, w) = \alpha_d(H)$, where H is the subgraph of $K_2^{\square n}$ induced by the vertices $(u_1, \ldots, u_n)$ with $u_i = 1$ for exactly w indices $i \in \{1, \ldots, n\}$ (where the vertices of K_2 are labeled with 0 and 1).