BOUNDS ON THE PERFORMANCE OF PSK BLOCK CODES

Efraim Laksman

Blekinge Institute of Technology
Licentiate Dissertation Series No. 2010:02

School of Engineering

Bounds on the Performance of PSK Block Codes

Efraim Laksman

Blekinge Institute of Technology Licentiate Dissertation Series
No 2010:02

Bounds on the Performance of PSK Block Codes

Efraim Laksman

Department of Mathematics and Natural Sciences
School of Engineering
Blekinge Institute of Technology
SWEDEN

© 2010 Efraim Laksman
Department of Mathematics and Natural Sciences
School of Engineering
Publisher: Blekinge Institute of Technology
Printed by Printfabriken, Karlskrona, Sweden 2010
ISBN: 978-91-7295-175-4
Blekinge Institute of Technology Licentiate Dissertation Series
ISSN 1650-2140
urn:nbn:se:bth-00457

Abstract

In wireless communication, the minimum Euclidean distance between codewords is a major factor for the ability to correct errors in messages, and it is of interest to maximize the minimum Euclidean distance.

The thesis improves previously established general upper bounds on minimum Euclidean distance of phase shift keying block codes. There are no requirements on structure of codes, as the bound depends only on alphabet size, word length and code size. Prior to this thesis, bounds found by use of a method of Elias, had been improved by generalization of Elias' method. The method used here is an attempt to optimize that generalization.

Acknowledgments

First and foremost I would like to thank my supervisors Dr. Håkan Lennerstad and Dr. Magnus Nilsson for the help and encouragement they have given me during this work. I have learned a lot from discussions with both of them.

I would also like to thank Dr. Robert Nyqvist who have helped me greatly with LaTeX. Without his help compiling this thesis would have taken me a lot more work than it did.

My thanks also goes to all of my colleagues at BTH, in particular those at the department of mathematics and natural sciences, for being good friends and forming a fantastic research enviroment.

Last but not least I would like to thank my family for moral support.

Contents

Abstract v
Acknowledgments vi
Preface 1
1 Introduction 2
1.1 Communication 2
1.2 Codes 3
1.3 Phase Shift Keying (PSK) 5
1.4 Distances and metrics 6
1.5 Hard- and Soft-decision decoding 11
1.6 The research problem 12
2 Presentation of papers 15
2.1 Paper I 15
2.2 Paper II 15
2.3 Paper III 16
Bibliography 16

I Improving bounds on the minimum Euclidean distance for block coded PSK by inner metric optimization 19

II Bounding the minimum Euclidean distance for any PSK block codes of alphabet size 8 37

III Inner distance measure bounds on the minimum Euclidean distance for symmetric PSK block codes 43

List of Figures

1.1 Some symbols and corresponding waves in 8-PSK 6
1.2 Example of waves and noise 10
1.3 Squared Euclidean distance between points on a circle and between waves 12
1.4 Example of the downside of HDD 13

Preface

This work consists of two parts. The first section introduces error-correcting codes and presents the three papers which form the second part.

List of Papers:

I E. Laksman, H. Lennerstad, M. Nilsson, Improving bounds on the minimum Euclidean distance for block coded PSK by inner metric optimization, *Combinatorics 2008*, Costermano, Italy, 2008
Submitted to *Discrete Mathmatics*

II E. Laksman, H. Lennerstad, M. Nilsson, Bounding the Euclidean minimum distance for any PSK block codes of alphabet size 8, *IEEE Information Theory Workshop*, Taormina, Italy, 2009

III E. Laksman, H. Lennerstad, M. Nilsson, Inner distance measure bounds on the minimal Euclidean distance for symmetric PSK block codes, *Fq 9, the 9th international conference on finite fields and their applications*, Dublin, Ireland, 2009
Submitted to *IEEE Transactions on Information Theory*

The following paper is related to but not included in this thesis:

IV M. Nilsson, H. Lennerstad, E. Laksman, A two-metric approach to improve bounds on the minimum Euclidean distance of block codes, *Proceedings of RVK08*, Växjö, Sweden, 2008

Contribution to Papers included in thesis:
The author of this thesis has constructed the main parts of the proofs, though all papers are joint work.

1 Introduction

This introduction starts with a short popular presentation of the general problem of noise in communication, before turning to the research question.

1.1 Communication

Communication occurs whenever a sender transfers a message to a receiver through a channel. A problem common to all forms of communication is noise on the channel, i. e. the message attained by the receiver may differ from the message transmitted by the sender. Some examples: in a conversation some words may be unclear due to sound from the background, when a computer attempts to read from a CD, there may be a scratch or dust on the CD making some symbols on the CD unreadable or when a receiver receives a carrier wave from a transmitting aerial the wave may be disturbed by movement of the receiver or by other signals.

Given a received message, the receiver must find the message that was sent, so error-correcting codes must be used to overcome the noise.

An error-correcting code is a way for the sender to encode a message such that when the coded message is sent and disturbed by noise, the original message will be reconstructed when the receiver decodes it, provided that the noise was small enough. Some examples of where error-correcting codes are used are internet, deep-space telecommunications, satellite broadcasting and data storage. One of the simplest examples of an error-correcting code is the repetition code, see [1], [9], [16] or almost any other literature on error-correcting codes.

Example 1 (Repetition code)**.** *Assume that the message in question is in written form. Now the sender codes the message with a repetition code by transmitting each symbol (including blanks)* $n = 2e + 1$ *times. When the receiver attempts to read the text, the receiver expects to observe* n *instances of every symbol. If the* n *first symbols that the receiver reads are not identical, by deciding that the instance occuring most frequently is the one that was sent, the receiver makes the correct decision, if at most* e *errors occured among the first* n *symbols, leaving the majority of the copies intact. The same decoding procedure is then done for the next* n *symbols, and so on. Note that if more than* e *errors occur the repetition code may fail.*

A repetition code using n *repetitions is said to have rate* $1/n$*, i. e. only one of* n *symbols sent carries information. The other symbols are redundant. Though unnecessary in the case of noiseless communication, they provide the ability to correct errors in the case of appropriate noise.*

A code in which n symbols have to be sent for where the same information could have been passed along using only k symbols in a noiseless channel is said to have rate k/n. Low rate, or high redundancy, results in longer messages which costs more resources and time to send. It would seem to be of interest to for every rate find a code with the best error-correcting capabilities possible, i. e. with the smallest risk of getting a symbol error.

However, as Shannon's Theorem shows, it is not so. A simplified version of Shannon's Theorem is shown:

Theorem 2 (Shannon's Theorem)**.** *Every channel has a capacity such that:*

- *for any rate below the capacity, codes yielding arbitrarily small error probability exists, and*
- *for any rate above the capacity, no code yielding arbitrarily small error probability exists.*

In the case of a channel where all errors are equally likely, the capacity is a convex function of the error probability p, with a minimum at $p = \frac{1}{2}$, and for many channels the capacity is known.

A proof of Shannon's Theorem can be found in [15], where it was first presented.

Remark that Shannon's Theorem only states whether or not codes of certain rate resulting in arbitrarily small error probability exists. It states nothing about how to find such codes. Generally though, to reach lower error probabilities without lowering the rate, one has to code a higher number of symbols simultaneously, i. e. use large k and n. Coding many symbols jointly comes with its own set of troubles, such as making the code impractical for use when transmitting short messages and making the code slow to decode.

It becomes interesting not only to find a relationship between the rate and the error-correcting capability of a code, but to find a relationship between n, k and the error-correcting capability of a code.

1.2 Codes

There are two ways to view a code, leading to two different definitions. The most common way is to define a code as a map from one space to another space of at least the same size as the first space. This is a good model for how encoders work. The idea is that input, the uncoded message, is given to an encoder which responds by giving the coded message as output. Such a definition of codes also renders itself well to defining different types of codes by imposing restrictions on what mappings are acceptable.

Another way to define codes, which has been successful in this research, is by viewing a code as a subset of some space. The idea is to consider the "output" to be the entire code, and to take no notion of the existance of input and encoder. This makes sense when there is no interest in which input corresponds to which output.

Let I be a discrete index set, and let the sequence space W be the direct product $\prod_{i \in I} A_i$, so that different alphabets may be used for symbols, sometimes called letters, on different positions, even though it is common that $A_i = A_j$ for all $i, j \in I$. Now any subset C of W is a *code*. This is the definition of a code used by Forney and Trott [4] and Loeliger and Mittelholzer [8].

The vast majority of codes used in practice are either block codes or convolutional codes. Block codes can be defined as codes with finite I. For a block code, elements of W are commonly called words, while the elements in C are also called codewords. For other codes we have sequences and code sequences.

The definition of convolutional codes that will be given here comes from [8]. Some definitions are necessary to be able to define convolutional codes. First, it is necessary that the index set I is totally ordered. For simplicity in expressing the necessary definitions, assume $I = \mathbf{Z}$ and let x_i be the value of the i:th coordinate of the sequence $\mathbf{x}$.

A code C is *time-invariant* if for any $\mathbf{y} \in C$ there exists elements $\mathbf{x}, \mathbf{z} \in C$ such that $x_{i-1} = y_i = z_{i+1}$ for all $i \in I$. Note that this leads to all alphabets being identical.

A code C is *strongly controllable* if there is a nonnegative integer l such that for any $i \in \mathbf{Z}$ and any $\mathbf{x}, \mathbf{y} \in C$, there exists an element $\mathbf{z} \in C$ such that $z_j = x_j$ for all $j < i$ and $z_j = y_j$ for all $j \geq i + l$.

A code C is *strongly observable* if there is a nonnegative integer l such that for any $i \in \mathbf{Z}$ and any $\mathbf{x}, \mathbf{y} \in C$ with $x_j = y_j$ for all $j \in [i, i + l)$, there exists an element $\mathbf{z} \in C$ such that $z_j = x_j$ for all $j < i$ and $z_j = y_j$ for all $j \geq i$.

A code C is a *group code* if the alphabet A_i is a group for every $i \in I$ and C is a subgroup of W when the group operation acts componentwise on sequences in W.

A code C is a *convolutional code* if it is a strongly controllable and strongly observable time-invariant group code.

If a part of the message which should correspond to a code sequence (codeword) doesn't do so, it can be corrected to a code sequence (codeword). It is desirable to correct it to the code sequence (codeword) which is most likely to have been sent, but this is not always done, see Section 1.5. To achieve fast decoding, one often decodes to a code sequence (codeword) close to what was received with respect to some distance function, $d(\cdot, \cdot)$. Types of distances will be discussed in Section 1.4. The free distance of a code, called minimum distance in case of a block code,

$$\begin{array}{l} d_{\text{free}}(C) = \min_{\mathbf{x},\mathbf{y} \in C} d(\mathbf{x}, \mathbf{y}) \text{ or} \\ d_{\min}(C) = \min_{\mathbf{x},\mathbf{y} \in C} d(\mathbf{x}, \mathbf{y}) \end{array} \tag{1.1}$$

is a crucial quantity for the code's error correcting capability.

Mathematical theory for error-correcting codes has mainly been developed for block codes [16], possibly only due to simplicity in analysis. In this thesis, only block codes are treated, and an interesting way to continue the research presented in this thesis would be to generalize the method used to bound minimum distance for block codes to bound the free distance for convolutional codes. In the papers in this thesis upper bounds on the minimum squared Euclidean distance, see Section 1.4, are found for block codes over alphabets of the form $(\mathbf{Z}_q^n, +)$, based on fixing q and fixing the number of codewords.

1.3 Phase Shift Keying (PSK)

In wireless applications, machines communicate by transmitting and receiving electromagnetic or acoustic waves. The wave carries information as variation of amplitude, frequency or phase (or combinations of these). Error-correcting codes are widely used during communication of such forms.

This thesis considers only PSK. This signifies communication where the carrier waves have equal amplitude and frequency, and only the phase is modulated. PSK is commonly used in applications, where some examples are wireless Local Area Networks, wireless Personal Area Networks and Proximity Cards, where 2 or 4 different phases are used. PSK with 8 phases is sometimes used in cellphones with 3G and in wireless Personal Area Networks.

This thesis employs the common assumption that the carrier waves are sinusoidal, and that phases used to represent symbols are evenly spread out between 0 and 2π. While these assumptions are common, work has also been made with different assumptions, e. g. [3] and [5].

Here, the waves take the form $A\sin(\omega t + \alpha)$ for fixed A and ω, while the value of $\alpha \in \{2\pi\frac{0}{q}, 2\pi\frac{1}{q}, 2\pi\frac{2}{q}, \ldots, 2\pi\frac{q-1}{q}\}$, for a fixed integer q, carries the information. When these particular phases are used, the modulation is called symmetric PSK. In some communication, typically when the transmitter or receiver is moving, information is instead carried in the difference between phase in consecutive signals transmitted, which is known as differential phase shift keying (DPSK). Whether PSK or DPSK is used, high minimum distance is still of utmost importance, so with respect to the further analysis PSK may be assumed without loss of generality.

This thesis will adhere to the common assumption that all messages are equally likely. This assumption is reasonable as data is usually compressed before it is coded error-correctingly and transmitted. There are many different methods for data compression, depending on the data to be compressed, but a common feature among them is that they remove redundancy, which makes all messages roughly equiprobable.

As mentioned already in Section 1.2, the message is coded as a possibly finite sequence of letters $x_1, x_2, \ldots$. Each letter x_i corresponds to a phase $2\pi(x_i \mod q)/q$. An example of this can be seen in Figure 1.1, where each symbol is shown for only one period. In practice, the wave corresponding to each symbol is sent for a number of periods depending on the ratio between the symbol rate and the frequency of the carrier wave. From a mathematical point of view, this may be seen as an "internal" infinite repetition code, as observing the wave for even very short period of time (any open interval) would be enough to determine the phase uniquely. When using symmetric PSK, it is practical to take the alphabet to be the group $(\mathbf{Z}_q, +)$ as it indicates that the distance between the letters i_1 and i_2 is the same as the distance between $i_1 + j$ and $i_2 + j$, which is true for the carrier waves corresponding to the letters, given a reasonable way of measuring distance between the waves.

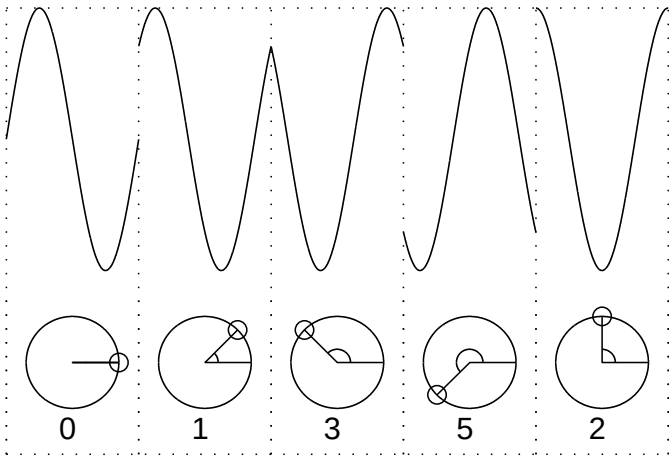

Figure 1.1: For symmetric 8-PSK, the phases and the carrier wave of the message $(0, 1, 3, 5, 2)$ is shown.

In this thesis the minimum distance used to measure how good a code is will be squared Euclidean distance, and Section 1.4 will not only explain what squared Euclidean distance is, but also why it is suitable for many applications.

The issue of distance is an important one in this thesis, and results have been achieved by using different distance measures, even though all results are in squared Euclidean distance.

1.4 Distances and metrics

We define distance measures and metrics as follows:

Definition 3 (Distance measure)**.** *For a set S, let δ be a function*

$$\delta : S \times S \mapsto \mathbb{R} \tag{1.2}$$

which satisfy the following properties:

- $\delta(s_1, s_2) \geq 0$*, with equality if and only if* $s_1 = s_2$ *and*
- $\delta(s_1, s_2) = \delta(s_2, s_1)$.

Then we will say that δ is a distance measure on S.

Definition 4 (Metric)**.** *Let δ be a distance measure on S with the additional property*

- $\delta(s_1, s_3) \leq \delta(s_1, s_2) + \delta(s_2, s_3)$ *(the triangle inequality).*

Then δ is a metric on S.

We next concider some common distances and upper bounds on minimum distance for codes when using these distances.

Hamming distance

The distance measure which is probably most common in mathematical literature on error-correcting codes is the Hamming distance,

$$d_H(\mathbf{x}, \mathbf{y}) = |\{x_i | x_i \neq y_i\}|, \tag{1.3}$$

where $\mathbf{x}$ and $\mathbf{y}$ denotes words and $|S|$ is the number of elements in the set S. Hamming distance, which is also a metric, is the only distance measure used in [1], [9] and [16]. Hamming distance suits best when

$$P(\text{receive } y | x \text{ was sent}) \approx P(\text{receive } z | x \text{ was sent}), \text{ for all } y, z \neq x \in \mathbf{Z}_q \tag{1.4}$$

This is the case for symmetric binary and ternary PSK. As was shown in [7], also symmetric quaternary PSK can be treated successfully with Hamming distance.

An upper bound on the minimum Hamming distance in a block code is the Hamming bound, also known as the sphere-packing bound or the volume bound. It's a well known bound, which can be found e. g. in [16]. It states that for a block code C with length of words n, alphabet size q and minimum Hamming distance $d_{H\,\min}$, the following inequality holds:

$$|C| V_q \left(n, \left\lfloor \frac{d_{H\,\min}}{2} \right\rfloor \right) \leq q^n, \tag{1.5}$$

where

$$V_q(n, r) = \sum_{k=0}^{r} \binom{n}{k} (q-1)^k \tag{1.6}$$

is the volume of a sphere with radius r in the space $\mathbf{F}_q^n$. The bound follows from placing one sphere of radius r centered at every codeword and maximizing r with the restriction that the spheres must be pairwise disjoint. This relation can be taken to be a bound on either $d_{H\,\min}$, $|C|$, n or q, given the other three. Note that while it is an upper bound for $|C|$ or $d_{H\,\min}$, it is a lower bound for n or q. Block codes which meet the Hamming bound are usually called perfect codes.

Other bounds on $d_{H\,\min}$ which are stronger for some n, $|C|$ and q exist. Of these, the Elias bound is presented here, as it is related to the results in this thesis.

Elias' bound states that for given word length n, code size $|C|$, alphabet size q and minimum distance $d_{H\,\min}$, the relationship

$$|C| \leq \frac{(q-1) n d_{H\,\min}}{q r^2 + n(q-1)(d_{H\,\min} - 2r)} \frac{q^n}{V_q(n, r)} \tag{1.7}$$

holds, where r is any integer in the interval $[2, n(q-1)/q]$. The idea is that a sphere of radius r in $\mathbf{F}_q^n$ can be shown to contain at most

$$\frac{(q-1) n d_{H\,\min}}{q r^2 + n(q-1)(d_{H\,\min} - 2r)} \tag{1.8}$$

words if the minimum distance between them is at least $d_{H\min}$, and that $\mathbf{F}_q^n$ can contain at most $\frac{q^n}{V_q(n,r)}$ pairwise disjoint spheres of radius r. See [16] for a complete proof.

Lee distance

Another distance measure which is also a metric, sometimes used when the alphabet is $(\mathbf{Z}_q, +)$, is Lee distance:

$$d_L(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{\lfloor q/2 \rfloor} i|\{j : x_j - y_j \equiv_q i \vee y_j - x_j \equiv_q i\}|, \tag{1.9}$$

where the summation sign is to be taken as addition over $\mathbf{Z}$, not $\mathbf{Z}_q$. For symmetric quaternary PSK, Lee distance coincides with squared Euclidean distance. For $n = 2$, Lee distance is also known as Manhattan distance or Taxicab distance.

Several bounds on the maximum number of words in a subset of $\mathbf{Z}_q^n$ with minimum Lee distance d are presented in [14], where also results for $q = 5, 6, 7$ and $n \in [1, 7]$ are presented in tables, constructed by application of the different bounds presented. The most frequently applied bounds for constructing the tables are the sphere-packing bound (generalization of the sphere-packing bound for Hamming codes), and two improvements of it that can be made for some parameter values.

The version of the sphere-packing bound given is

$$|C| \leq \left\lfloor \frac{q^n}{V_q(n,r)} \right\rfloor, \tag{1.10}$$

where the sphere radius is $r = \lfloor (d_{L\min} - 1)/2 \rfloor$.

An improved version is given for even $d_{L\min}$. If $d_{L\min} = 2e$, then

$$|C| \leq \left\lfloor \frac{q^n}{W_q(n,r)} \right\rfloor, \tag{1.11}$$

where $W_q(n,r) = V_q(n, e-1) + V_q(n-1, e-1)$.

If also q is not less than $d_{L\min}$, the bound can be improved further. If $q \geq d_{L\min} = 2e$, then

$$|C| \leq \left\lfloor \frac{2q^{n-1}}{W_q(n,r)} \left\lfloor \frac{q}{2} \right\rfloor \right\rfloor . \tag{1.12}$$

Euclidean distance

On $\mathbf{R}^n$, Euclidean distance (which is also a metric) is defined as

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} . \tag{1.13}$$

On $\mathbf{Z}_q^n$, Euclidean distance is instead defined as

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} 4 \sin^2 \left(\frac{(x_i - y_i)\pi}{q} \right)}, \tag{1.14}$$

which corresponds to Euclidean distance between q points spread out evenly on the unit circle on $\mathbf{R}^2$. Now follows motivation to why squared Euclidean distance on $\mathbf{Z}_q^n$ (which is not a metric) is suitable for modelling symmetric PSK.

Noise in communication by carrier waves will be assumed to be additive white gaussian noise (AWGN), which is a very common assumption of noise in many applications, not only coding theory. That the noise is *additive* doesn't actually say anything about the noise, it only describes a viewpoint of how to relate the received signal to the transmitted signal and the noise. More strictly, it says that the noise is chosen as the (signed) difference between the received signal and the transmitted signal. However, when stated that the noise is AWGN, something is said about the statistical distribution of the difference between the received signal and the transmitted signal. With a different viewpoint on how the transmitted and received signal relates, the statistical distribution of noise would be different.

A process without any structure, i. e. with an autocorrelation function which is zero except at $t = 0$, is called a *white* process. The received signal is usually filtered in such a way, that even if noise originally was not white, the filtered signal behaves as though it was the transmitted signal plus white noise.

If the noise is comprised of additive white noise from many small sources, then by the central limit theorem it is a reasonable assumption that it will be gaussian. The noise being *gaussian* means that it has the density function

$$f_W(w; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{w^2}{2\sigma^2}}, \tag{1.15}$$

so using the additivity gives the density function

$$f_X(x; m, t, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x - s_m(t))^2}{2\sigma^2}}, \tag{1.16}$$

for the received signal, where $s_m(t)$ is the effect of the signal for the message m at time t, and σ is the standard deviation of the noise. Figure 1.2 illustrates the situation.

Even though the receiver has access to the received signal in continuous time, one may assume that access to the values of the received signal exists only for some discrete set of points, A. As it is possible to let A grow to cover the continuous timespan arbitrarily well, this is a sensible approximation. To find the message which is most likely to have been sent, given the received message, one has to find the message m which maximizes the expression

$$\prod_{a \in A} f_X(x_a; m, t_a, \sigma), \tag{1.17}$$

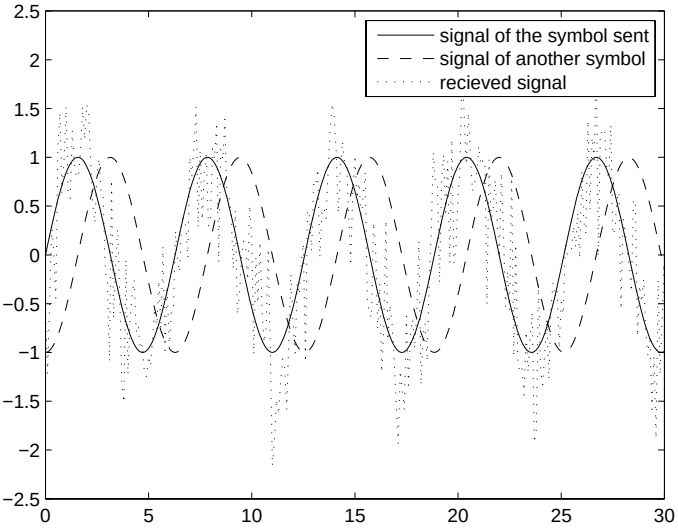

Figure 1.2: An example showing what may be sent and received. Given the received signal, the receiver wishes to find what symbol has the highest likelyhood of having been sent.

where t_a is the time at point a. The expression takes this form as the noise at different points in time is independent when the noise is white.

The expression can be simplified to finding m which minimize the expression

$$\sum_{a \in A} (x_a - s_m(t_a))^2, \tag{1.18}$$

which is the same as deciding which of the possibly sent signals which is closest to the received signal in terms of squared Euclidean distance.

Note that the squared Euclidean distance and Euclidean distance give the same minimum since the function x^2 is increasing for $x \geq 0$.

The squared Euclidean distance is not a metric, but it is additive, i. e.

$$d_E^2((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) = \sum_{i=1}^{n} d_E^2(x_i, y_i), \tag{1.19}$$

so the distance between signals corresponding to letters is of particular interest.

According to the assumptions previously stated, the signals for x_1 and y_1 are $A\sin(2x_1\pi/q + t)$ and $A\sin(2y_1\pi/q + t)$ respectively, where A is the amplitude of the signal. The squared Euclidean distance between the two

signals with phases $2x_1\pi/q$ and $2y_1\pi/q$ is then

$$\int_0^{2\pi} \left(A\sin\left(\frac{2x_1\pi}{q}+t\right) - A\sin\left(\frac{2y_1\pi}{q}+t\right)\right)^2 dt = \\ = 2A^2\pi\left(1-\cos\left(\frac{2(x_1-y_1)\pi}{q}\right)\right) = 4A^2\pi\sin^2\left(\frac{(x_1-y_1)\pi}{q}\right) . \tag{1.20}$$

Note that in this calculation it was assumed that each signal is sent for only one period. Usually each symbol is sent for many periods, but all that does with the distance is to multiply it by a constant. Normalizing the constant, in this thesis the squared Euclidean distance between symbols will be taken to be $4\sin^2((x_1-y_1)\pi/q)$, which is the same as the squared Euclidean distance between the two points

$$\left(\cos\left(\frac{2x_1\pi}{q}\right), \sin\left(\frac{2x_1\pi}{q}\right)\right) \text{ and } \left(\cos\left(\frac{2y_1\pi}{q}\right), \sin\left(\frac{2y_1\pi}{q}\right)\right), \tag{1.21}$$

on the unit circle. Thus, by spreading out the symbols from $\mathbf{Z}_q$ evenly on the unit circle in $\mathbf{R}^2$, or placing them in the positions for the q:th roots of unity on the complex plane, the squared Euclidean distance (as defined on $\mathbf{R}^2$) between the points will be the same as that between the symbols corresponding carrier waves. This is illustrated in Figure 1.3.

1.5 Hard- and Soft-decision decoding

Hard-decision decoding (HDD) means that every signal corresponding to a letter is interpreted as a letter, the one which is closest to the received signal. I. e. initial interpretation of the signal is performed individually for each letter. If a word which is not also a codeword appears, *then* that word is interpreted as the codeword closest to that word.

Soft-decision decoding (SDD) means that the signal corresponding to a word is interpreted as the codeword with signal closest to the received signal. The signal is interpreted as a codeword directly, without first interpreting it as individual symbols.

Observe that closest is here taken to mean closest in the sense of Euclidean distance, but what is stated about HDD and SDD in this thesis also holds for other distance measures, such as Hamming- or Lee distance.

Figure 1.4 shows an example of the problems that can occur with hard-decision decoding.

It would seem that SDD is clearly better than HDD, as can be seen from the example in Figure 1.4, yet HDD occurs in practice. One reason is that it is faster for a computer to work with integers than with real numbers. Another reason is that to find the distance between a signal of the length of one word to the signal of each codeword may be slow since $|C|$ can be very large. However, if the code has some special structure, there may be fast ways to find the codeword closest to a given word. Many such methods require that the signal is first approximated by a word to be usable.

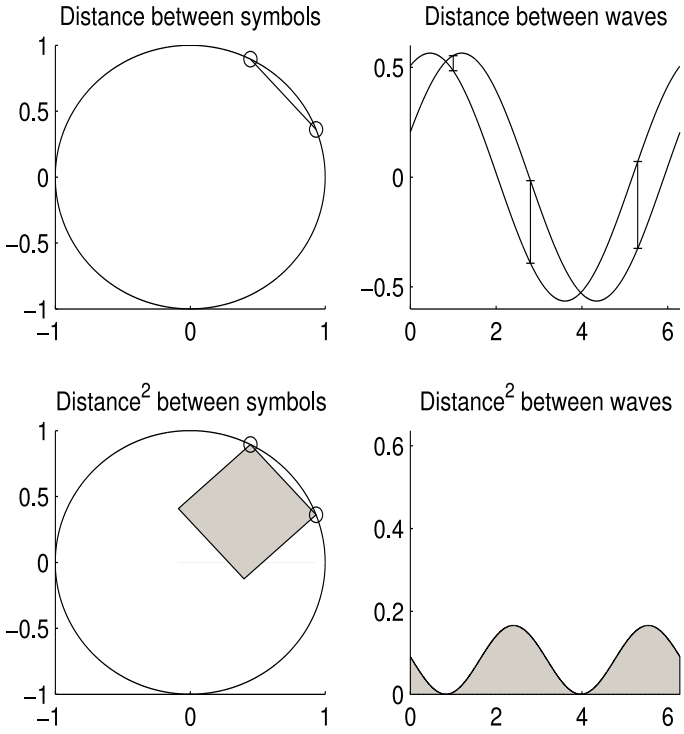

Figure 1.3: The distance between the symbols 1 and 3 in 17-PSK is shown above and to the left. Distance of the corresponding waves is indicated above and to the right. Below are the corresponding squares, i. e. the squared Euclidean distance, which is the same in both graphs.

To be efficient, a block code should have as large minimum Euclidean distance as possible, no matter whether HDD or SDD is used. This thesis is written with the assumption of SDD, with the exception of the repetition code as presented in Section 1.1, but the results hold just as well for HDD.

1.6 The research problem

From what has been stated in this thesis up to this point, it is clear that there is practical interest in finding subsets C of $\mathbf{Z}_q^n$, with a fixed size $|C|$ that have *maximal* minimum Euclidean distance $d_{E\,\min}(C)$. For general q, n and $|C|$, the maximal minimum Euclidean distance $\max_C d_{E\,\min}(C)$ is unknown. There exist both lower and upper bounds on this quantity, and the research problem has been to improve the upper bounds that exist.

Elias proposed a method for finding an upper bound on the minimum distance for any additive distance measure, word length n, code size $|C|$ and any alphabet size q. Elias' bound with respect to Hamming distance

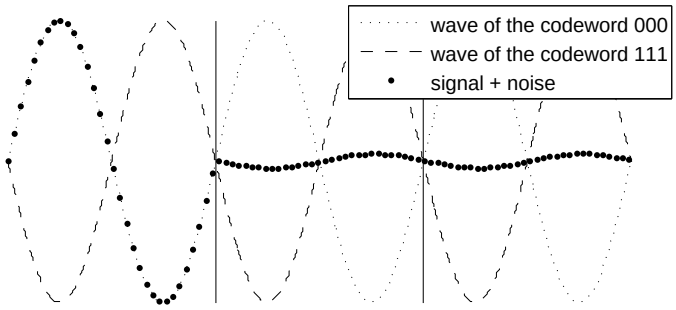

Figure 1.4: Consider a 2-PSK code with only two codewords, 000 and 111. The carrier waves for the two words are shown in the figure, as well as the received signal. With HDD, the first symbol will be approximated to 0 while the second and third symbols will be approximated to 1. Then the whole word will be taken to be 111. With SDD, the entire word is a lot closer to 000 than to 111, which can be seen even without calculation as the first symbol is clearly a 0 while the other symbols are almost as likely to be 0's as 1's.

presented in Section 1.4 was found by applying Elias' method with respect to Hamming distance. Elias did not write of this method himself, but it can be found in [2]. The idea is the following:

1. Form a sphere with radius t around every codeword in C.
2. Set $K = \lceil |C| \cdot |S_t(\mathbf{x})| q^{-n} \rceil$, where $S_t(\mathbf{x})$ is the set of words in a sphere of radius t with center at $\mathbf{x}$. For reasons of symmetry, $|S_t(\mathbf{x})|$ is independent on which word $\mathbf{x}$ is chosen. With asymmetric constellations of symbols, this may not hold true. (This is a limitation of Elias' method, but in accordance with assumptions made in this thesis.)
3. By the pigeon hole principle, there must be at least one word, named $\mathbf{w}$, in $\mathbf{Z}_q^n$ contained in at least K spheres.
4. Form the sphere $S_t(\mathbf{w})$, and observe that it contains at least K codewords. This sphere is called a critical sphere.
5. The minimum distance between codewords in $W = C \cap S_t(\mathbf{w})$ is an upper bound on the minimum distance of C, assuming t was chosen large enough so that K is at least 2.
6. The average distance between codewords in W is an upper bound on the minimum distance of W, hence also on the minimum distance of C.
7. Write the codewords in W as rows in a matrix M.
8. Find the column in M which results in the highest average difference between symbols on the i:th position, with respect to a limitation on the total difference between $\mathbf{w}$ and the codewords of W.

9. This way an upper bound on the average distance of codewords in W can be found, which is an upper bound for the minimum distance of C.

Every value of t, will result in a bound. Notice that K is a stepwise right-continuous function of t, and that increasing t without increasing K cannot make the bound tighter, so that only the values of t where K increases are of interest. On the set of these points, K is an invertible function of t, so that t may also be seen as a function of K.

A contribution to Elias' method in this research is to examine the possibilities of creating a better bound on the minimum Euclidean distance by using different distance measures than Euclidean for defining the shape of the spheres that lead to a critical sphere.

In 2000, Nilsson and Lennerstad [10] considered a different shape of the spheres in Elias method. As can be observed in the description of Elias method above, while it uses spheres of radius t, it doesn't state what distance measure should be used for the radius. Elias originally used the same distance measure for the radius as the one in which he wanted to bound the minimum distance. Nilsson and Lennerstad found that bounds could sometimes be improved by use of a different additive distance measure for the shape of the spheres. In this thesis the additive distance measure being used for determining the shape of the spheres will be called *inner distance measure*, i. e. it will not neccessarily satisfy the triangle inequality.

The restriction to additive distance measures as inner distance measures is used in the last step of Elias' method described above, when considering the matrix column by column.

In [10], bounds were found for sets C with $|C| > (q/3)^n$ by using the additive distance measure λ_1, where λ_1 was defined as

$$\lambda_1(0) = 0, \quad \lambda_1(\pm 1) = 1, \quad \lambda_1(i) = \infty \text{ for } i \in [2, q-1] \ . \tag{1.22}$$

In [11], the inner metric λ_1 was generalized to λ_r, where λ_r was defined as

$$\lambda_r(i) = |i| \text{ for } i \in [-r, r] \text{ and } \lambda_r(i) = \infty \text{ for } i \in (-q/2, q/2] \setminus [-r, r] \ . \tag{1.23}$$

This resulted in bounds also when $|C| \leq (q/3)^n$. Using $r = 2$ and $r = 4$ bounds were found for $q = 8$. The distance measure λ_∞ is Lee-distance, and λ_r is a metric as it satisfies the triangle inequality.

A natural next step was to attempt to maximize the minimum squared Euclidean distance with respect to a general inner distance measure. In [12], additive inner distance measures that were allowed to depend on K were introduced. Numerical testing was used to find an inner distance measure which improved on previous bounds for $q = 8$. With a K-dependent inner distance measure, there appears to be a great deal of interdependence between t, K and the inner distance measure. It is important to note that this poses no problem. If K is chosen first, this leads to an inner distance measure δ_K. For any pair of K and δ_K, the smallest t such that $K \leq \lceil |C| \cdot |S_t(\mathbf{x})| q^{-n} \rceil$ can now be used, so t, K and δ can all be well defined in this manner.

The papers in this thesis attempt to optimize and generalize the idea of a K-dependent inner distance measure.

2 Presentation of papers

2.1 Paper I

In Paper I the restriction $q = 8$ was used. Using Elias method, a bound of the form

$$d^2_{E\,\min}(C) \leq \frac{2t_K(\delta)\max_{\mathbf{x}} f_\delta(\mathbf{x})}{K-1} \tag{2.1}$$

was found, where $\mathbf{x}$ is a column of length K and

$$f_\delta(\mathbf{x}) = \frac{\sum_{1\leq j_1<j_2\leq K} d^2_E(x_{j_1}, x_{j_2})}{\sum_{i=1}^{K}\delta(0,x_i)}\ . \tag{2.2}$$

One of the main results was that only the six columns

$$\begin{array}{lll} \mathbf{y}_1 = (1,-1,0,\ldots,0)^T & \mathbf{y}_2 = (2,0,\ldots,0)^T & \mathbf{y}_3 = (3,0,\ldots,0)^T \\ \mathbf{y}_4 = (4,0,\ldots,0)^T & \mathbf{y}_5 = (1,-2,0,\ldots,0)^T & \mathbf{y}_6 = (2,-2,0,\ldots,0)^T \end{array} \tag{2.3}$$

can maximize f_δ. Which of these columns that maximize f_δ depends on δ.

Using this result, a one-parameter family of inner distance measures was found, from which all inner distance measures minimize the quantity $\max_{\mathbf{x}} f_\delta(\mathbf{x})$. This family can be described as

$$\begin{array}{ll} \delta(0,1) = (K-1)d_E(0,1) - 2 + 2\sqrt{2} - h, & \delta(0,2) = (K-1)d_E(0,2) + h, \\ \delta(0,3) = (K-1)d_E(0,3), & \delta(0,4) = (K-1)d_E(0,4), \end{array} \tag{2.4}$$

$h \in [0, \sqrt{2}-1]$. All members of this family leads to $\mathbf{y}_3$, $\mathbf{y}_4$ and $\mathbf{y}_5$ maximizing f_δ. Then, $t_K(\delta)$ may be minimized with respect to δ's within the given one-parameter family.

Finding an inner distance measure from the given family led to an improved bound on the minimum Euclidean distance for symmetric PSK block codes. However, while it is proven in Paper I that $\max_{\mathbf{x}} f_\delta(\mathbf{x})$ is minimized, it is not proven $t_K(\delta)\max_{\mathbf{x}} f_\delta(\mathbf{x})$ is minimized.

2.2 Paper II

A nuisance in Paper I was that rather than minimizing the upper bound of (2.1), only one of two factors in (2.1) was minimized. In this paper it is proven that minimizing the upper bound of (2.1) is done by using an inner distance measure from the same one-parameter family that was found in Paper I and minimized only one factor. Remark that it is not necessarily the same inner distance measure that is optimal in both cases. It is only shown that both can be found by sampling over the same one-parameter family.

2.3 Paper III

In Paper I and Paper II, the restriction $q = 8$ is used. In Paper III a bound for any q, n and $|C|$ is found. Paper III begins similarly to Paper I, even though slightly more general and requiring more trigonometry. When it comes to the result however, it seems difficult to give a formula for exactly which columns maximize f_δ in (2.2) for any q. Generalizing the method of Paper I columns are found by a method for sorting out columns which can not maximize f_δ. There are however many columns, so in order to make the work of sorting out columns reasonable, a theorem which immediately reduces the set of columns under consideration is given. The theorem states that only columns $\mathbf{y}$ satisfying

$$\sum_{i=1}^{K} d_E^2(0, y_i) \leq 4 \tag{2.5}$$

can maximize f_δ.

While Paper III results in bounds for any q, n and $|C|$, just as in Paper I, it is not proven to be optimal within the problem formulation. Paper II showed that for $q = 8$ this way of finding a bound was indeed optimal, but for general q, no corresponding result has been found.

Bibliography

[1] Andersson K. G., *Ändliga Kroppar och Felrättade Koder*, Lunds Universitet, 1996.

[2] Berlekamp E. R., *Algebraic Coding Theory*, New York: McGraw Hill, 1968.

[3] Divsalar D., Simon M. K., and Yuen J. H., Trellis coding with asymmetric modulations, *IEEE Trans. Commun.*, vol. COM-35, pp.130-141, Feb. 1987.

[4] Forney Jr. G. D., Trott M. D., The Dynamics of Group Codes: State Spaces, Trellis Diagrams, and Canonical Encoders, *IEEE Trans. on Inform. Theory*, vol. 39, no. 9, pp. 1491-1513, Sep. 1993

[5] Isaka M., Fossorier M. P. C., Morelos-Zaragoza R. H., Lin S., Imai H., Multilevel Coded Modulation for Unequal Error Protection and Multistage Decoding – Part II: Asymmetric Constellations, *IEEE Trans. Commun.*, vol. 48, no. 5, May 2000, pp. 774-786.

[6] Kschischang F. R., de Buda P. G., Pasupathy S., Block Coset Codes for M-ary Phase Shift Keying, *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, Aug. 1989.

[7] Lee C. Y., Some Properties of Nonbinary Error-Correcting Codes, *IRE Trans. Inform. Theory* 4 (1958), 77-82.

[8] Loeliger H.-A., Mittelholzer T., Convolutional Codes Over Groups, *IEEE Trans. on Inform. Theory*, vol 42, no. 6, pp. 1660-1686, Nov. 1996

[9] MacKay D. J. C., *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, ISBN: 978-0521642981, Sixth Printing 2007 edition.

[10] Nilsson M., Lennerstad H., An Upper Bound on the Minimum Euclidean Distance for Block Coded Phase Shift Keying, *IEEE Trans. Inform. Theory* 46(2): pp. 656-662, 2000.

[11] Nilsson M., Lennerstad H., Improved Upper Bound on the Minimum Euclidean Distance for Block Coded Phase Shift Keying, in proceedings of RVK05, Linköping, Sweden, 2005.

[12] Nilsson M., Lennerstad H., Laksman E., A two-metric Approach to Improve Bounds on the Minimum Euclidean Distance for Block Codes, proceedings of RVK08, Växjö, Sweden, 2008.

[13] Piret Ph., Bounds for Codes over the unit circle, *IEEE Trans. Inform. Theory*, vol. IT-32, pp.760-767, Nov. 1986.

[14] Quistorff J., New upper bounds on Lee codes, *Discrete Applied Mathematics*, 154, pp. 1510-1521 (2006).

[15] Shannon C. E., A mathematical theory of communication, *Bell Syst. Tech. J.* 27, pp. 379-423, 623-656 (1948).

[16] van Lint J. H., *Introduction to Coding Theory, Second Edition*, Springer Verlag, ISBN: 0-387-54894-7, 1992.

[17] Wyner A. D., Bounds on Communication with Polyphase Coding, *Bell System Technical Journal*, 45, April 1966, pp. 523-559.

Paper I

Improving bounds on the minimum Euclidean distance for block coded PSK by inner metric optimization

Efraim Laksman, Håkan Lennerstad, Magnus Nilsson

Improving bounds on the minimum Euclidean distance for block coded PSK by inner metric optimization

Efraim Laksman, Håkan Lennerstad, Magnus Nilsson

Abstract

The minimum Euclidean distance is a fundamental quantity for block coded PSK. In this paper we improve bounds for this quantity that are explicit functions of the alphabet size q, block length n and code size $|C|$. For $q = 8$ we improve previous results by introducing a general inner distance measure allowing different shapes of a neighborhood for a codeword. By optimizing the parameters of this inner distance measure, we find sharper bounds for the outer distance measure, which is Euclidean.

The proof is built upon the Elias critical sphere argument, which localizes the optimization problem to one neighbourhood. We remark that any code with $q = 8$ that fulfills the bound with equality is best possible in terms of the minimum Euclidean distance, for given parameters n and $|C|$. This is true for many multilevel codes.

1 Introduction

We intend to improve bounds for the minimum squared Euclidean distance for block coded PSK. For error correction with respect to maximum likelihood, when using a channel with additive white Gaussian noise, the squared Euclidean distance of the code is a highly relevant measure of the efficiency of a code for fixed block length n, code size $|C|$ and alphabet size q.

On the set $\mathbf{Z}_q^n$, the squared Euclidean distance is defined as

$$d_E^2(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} d_E^2(x_j, y_j), \tag{1}$$

where $d^2(x_j, y_j)$ is

$$d_E^2(x_j, y_j) = d_E^2(x_j - y_j, 0) = 4 \sin^2 \frac{(x_j - y_j)\pi}{q} \ . \tag{2}$$

Note that this distance measure is translation invariant, so that often the arguments can be written in such a way that one of them is zero. To simplify notation we will write $d(x) = d(x, 0)$ for any distance measure. Now a relevant

model for the words are points in the group $(\mathbf{Z}_q^n, +)$, with the squared Euclidean distance used for measuring distance, see for example [1], [9].

We consider an arbitrary subset C of $\mathbf{Z}_q^n$, corresponding to a block code having $|C|$ codewords $\mathbf{x} = (x_1, \ldots, x_n)$ of length n in an alphabet of q letters. The minimum squared Euclidean distance for the code is then

$$d^2_{E\,\min}(C) = \min_{\substack{\mathbf{x},\mathbf{y} \in C \\ \mathbf{x} \neq \mathbf{y}}} d^2_E(\mathbf{x}, \mathbf{y}) \ . \tag{3}$$

Bounds on the minimum Euclidean distance are fundamental for the geometry of $\mathbf{Z}_q^n$: which is the largest possible distance between the two closest members in a subset of $\mathbf{Z}_q^n$ with $|C|$ members?

As is well known, the minimum Euclidean distance is essential for the error correction capabilities for a code. We define the rate of a block code as

$$R(q, n, |C|) = \frac{\log_q |C|}{n} \ . \tag{4}$$

For several combinations of q, n, and $|C|$, mostly for high rates, there are known codes whose minimum squared Euclidean distances fulfil our bound with equality. For these combinations of q, n and $|C|$, neither the codes nor the bound can be improved.

For other combinations of q, n and $|C|$ there is a gap between the bound and the minimum squared Euclidean distances for the best known codes. The size of this gap differs from case to case. Especially for medium and low rates it is unknown whether there exist better codes to discover, or if it is possible to improve the bound, or a combination of both.

Many of the best known block codes, in the sense of minimum squared Euclidean distance, are constructed as multilevel codes, see for example [4], [11] and [12]. There are also other code constructions providing some of the best known block codes.

The results of this paper are derived by using different kinds of distance measures and metrics, so we next define these concepts. Both a distance measure and a metric are a function $d(\mathbf{x}, \mathbf{y})$ from pairs of codewords to non-negative numbers with the symmetry property $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}$ and $\mathbf{y}$, and $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$. Unlike a distance measure, a metric is also required to satisfy the triangle inequality: $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{y}, \mathbf{z})$ for all $\mathbf{x}$ and $\mathbf{y}$. Note that the *squared* Euclidean distance measure $d^2_E(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n d^2_E(x_j - y_j)$ is not a metric in general. For $q = 8$ we have $2 = d^2_E(2) > 2d^2_E(1) = 2(2 - \sqrt{2})$, which may be the reason why the optimal inner metric of Theorem 4 differs from $d^2_E(i) \cdot (K - 1)$ only for $i = 1, 2$ (K is a constant, defined later).

The quantities $d^2_E(x_j, y_j) = 4 \sin^2 \frac{(x_j - y_j)\pi}{q}$ are Euclidean distances between points when the entries $1, \ldots, q$ are distributed equidistantly on a unit circle. The generalized distance measures considered in this paper will be translation invariant and defined on $\mathbf{Z}_q^n$ so they will be defined by a sequence of non-negative

numbers, $\delta = \delta(1), \delta(2), \ldots, \delta(q)$, without any particular geometrical meaning. The distance is then

$$\delta(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} \delta(x_j - y_j), \tag{5}$$

generalizing

$$d_E^2(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} d^2(x_j - y_j). \tag{6}$$

Some of the numbers $\delta(i)$ may be infinite, prohibiting the corresponding differences. The Lee metric, for example, is represented by $\delta(i) = i$. Also truncated Lee metrics, where $\delta(i) = i$ for $i \leq r$ but $\delta(i) = \infty$ for $i > r$, have been considered [7].

An alternative notation is sometimes useful. For two codewords $\mathbf{x}$ and $\mathbf{y}$, the number of positions where $\mathbf{x}$ and $\mathbf{y}$ differ by i or by $q - i$ is denoted by $c_i(\mathbf{x}, \mathbf{y})$:

$$c_i(\mathbf{x}, \mathbf{y}) = |\{j \in [1, n] : (x_j - y_j) \equiv_q i \text{ or } (x_j - y_j) \equiv_q q - i\}|, \tag{7}$$

where $\equiv_q$ means equal with respect to modulo q. We are still working with a generalization of the closest distance of letters in a unit circle, so two words can in one position differ by at most $\lfloor q/2 \rfloor$. Then an alternative notation for $\delta(\mathbf{x}, \mathbf{y})$ is

$$\delta(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} \delta(x_j - y_j) = \sum_{i=1}^{\lfloor q/2 \rfloor} \delta(i) c_i(\mathbf{x}, \mathbf{y}). \tag{8}$$

2 Previous work

Apart from the presentation in this section, Section 6 contains an overview over the arguments that to some extent requires knowledge of the technical definitions that occurs in the paper. The bounds in this paper and in the previous papers in this line of research are partly based on the arguments leading to the well-known Elias' bound (see also Secton 6, [1] pp. 318-321, and [9] pp. 558-564). Elias' bound arguments have been used by Piret [10], who calculates bounds for the maximum rate, $n^{-1} \ln |C|$, for codes C with given $d_{E\,\min}^2(C)/n$ as $n \to \infty$. Piret's upper bound on the rate becomes

$$\ln q - \max_{\substack{2\beta S \beta^T = d_{E\,\min}^2(C)/n \\ \sum_i \beta_i = 1}} (H(\beta)), \tag{9}$$

where H is the entropy function

$$H(\beta) = -\sum_{i=1}^{q} \beta_i \ln(\beta_i), \tag{10}$$

β is a vector of length q and S is the $q \times q$ matrix with elements $2\sin^2[(i-j)\pi/q]$ in position (i,j).

The maximum rate as $n \to \infty$ is a non-increasing function of $d^2_{E\,\min}(C)/n$. Thus we can get a bound on $d^2_{E\,\min}/n$ as $n \to \infty$ as a function of the rate by reflecting the graph of Piret's bound in the line $\frac{\ln|C|}{d^2_{E\,\min}}$.

Wyner [13] has produced another bound for the same quantity as Piret. It is independent of q, and the q points may be distributed arbitrarily, giving for larger q weaker restrictions and a tighter bound in general. Wyner's bound is

$$\lim_{n\to\infty} \frac{1}{n} \frac{nK\left(\frac{d^2}{2n}\right)(2\pi)^n}{V_n\left(\sqrt{2n\left(1-\sqrt{1-\frac{d^2}{2n}}\right)-1}\right)}, \tag{11}$$

where $V_n(r)$ is defined as the volume of a sphere with radius r in the n-dimensional torus with the Euclidean distance 2π in each dimension. Just as with Piret's bound, it is a bound on $d^2_{E\,\min}(C)/n$ for a given rate as $n \to \infty$.

3 Problem formulation

The first result in the present research is a general and not very explicit bound, which is valid for arbitrary values of the parameters q, n and $|C|$. The second result is an explicit bound valid for $q = 8$ only. We next start the argument and simultaneously present results of previous papers [6], [7] and [8].

Generalizing the argument of the Elias' bound (see Sections 6), we define a neighborhood $S_{\delta,t}(\mathbf{z})$ for a word $\mathbf{z}$ as

$$S_{\delta,t}(\mathbf{z}) = \{\mathbf{y} : \sum_{i=1}^{\lfloor q/2 \rfloor} \delta(i) c_i(\mathbf{z},\mathbf{y}) \le t\} \tag{12}$$

(see Section 1), where δ is a distance measure. The set $S_{\delta,t}(\mathbf{z})$ is in the literature sometimes called a sphere with radius t. Observe that the number of words in a sphere is independent of the word that lies at its center.

We continue the Elias' argument of a critical sphere. If we label all words by their membership in a neighborhood $S_{\delta,t}(\mathbf{z})$, for each $\mathbf{z} \in C$, we distribute in total $|C||S_{\delta,t}|$ labels. Assume that t is large enough so that $|C||S_{\delta,t}|q^{-n} > 1$, and define $K = \lceil |C||S_{\delta,t}|q^{-n} \rceil$.

By the definition $K = \lceil |C||S_{\delta,t}|q^{-n} \rceil$, and by the pigeon hole principle, it follows that there is a word $\mathbf{y}^*$ so that $\mathbf{y}^* \in S_{\delta,t}(\mathbf{z})$ for at least K codewords $\mathbf{z} \in C$. (We may assume that exactly K codewords lie in this sphere, as bounds based on this method are non-inreasing functions of K.) Then these K codewords belong to the neighborhood $S_{\delta,t}(\mathbf{y}^*)$. By subtracting $\mathbf{y}^*$ from all codewords we do not change any distances between codewords, so we may as well assume that $\mathbf{y}^* = \mathbf{0}$. Now let $W = S_{\delta,t}(\mathbf{y}^*) \cap C$, so $|W| = K$. We trivially have

$$d^2_{E\,\min}(C) \le d^2_{E\,\min}(W). \tag{13}$$

The problem is thus localized from a minimum distance bound of C to a minimum distance bound of W. Often this inequality results only in a small or very small slack since the density of codewords in W is $\lceil |C||S_{\delta,t}|q^{-n} \rceil / |S_{\delta,t}|$ – slightly higher than the density of C in $\mathbf{Z}_q^n$.

Elias [1] pp. 318-321, and Nilsson and Lennerstad [6], [7] have found upper bounds on $d^2_{E\min}(W)$ by bounding the average distance between the words in W by the mean distance $d^2_{E\text{mean}}(W)$.

Elias finds the bound

$$\frac{K^2 x(2-x)\bar{D}n}{K(K-1)}, \tag{14}$$

where $\bar{D}$ is the average distance between letters $\sum_{j=0}^{q-1} d(j)$, which when d is d_E^2 results in $\bar{D} = 2$ and t is the radius in the spheres. One has to set $x = t/\bar{D}n$.

In [6] Nilsson and Lennerstad used $\delta(1) = 1$ and $\delta(i) = \infty$ for $i > 1$, allowing at most t non-zeroes in a sphere, giving the bound

$$d^2_{E\min}(C) \leq \frac{t}{K-1} d_E^2(2) + 2(t - \frac{t}{K-1}) d_E^2(1), \tag{15}$$

which is a main result of [6]. This bound is applicable for $|C| > (q/3)^n$ only, so it cannot be used for low rates, but it is tight in many cases for high rates. The tightness happens when all pairs of codewords in a neighborhood are at the same distance, in which case $d^2_{E\min}(W) \leq d^2_{E\text{mean}}(W)$ has zero slack – it is an equality.

The Lee metric is the metric $\delta(i) = i$ for all i. When restricted to the group $(\mathbf{Z}_q^n, +)$ it becomes

$$\delta(\mathbf{x}, \mathbf{y}) = \sum_{i=0}^{\lfloor q/2 \rfloor} i c_i(\mathbf{x}, \mathbf{y}) \ . \tag{16}$$

In [7], for $q = 8$ a two parameter (t, r)-Lee metric $\delta(i) = i$ if $i \leq r$ and $\delta(i) = \infty$ if $i > r$ was tried for $q = 8$. It turns out that $r = 2$ improves the small deviation neighborhood for medium rates, while $r = 4$ is preferable for low rates (see [7]).

The idea of considering a general inner distance measure δ and designing it to optimize the bound for the outer distance measure, which is a squared Euclidean, was first presented in [8]. Here a K-dependent inner distance measure was presented, as well as columns that appear to be extremal by sampling the space of all possible distance measures. Compared to that paper, the present paper presents an improved K-dependent distance measure. Furthermore, in this paper a partial optimality of this distance measure is proven.

4 General results

Continuing the argument of the previous section, we have a sphere $S_{\delta,t}(\mathbf{w})$ containing at least $K = \lceil |C||S_{\delta,t}|q^{-n} \rceil$ codewords, $W = S_{\delta,t}(\mathbf{w}) \cap C$, and we assume $\mathbf{w} = \mathbf{0}$.

Theorem 1 *For any code C in $\mathbf{Z}_q^n$ we have the bound*

$$d^2_{E\min}(C) \leq \min_{K\in[2,|C|]} \min_{\delta} \frac{2\widetilde{t}_K f_\delta(\widehat{\mathbf{y}})}{K-1}, \tag{17}$$

where

$$\widetilde{t}_K(\delta) = \min(\{t : K \leq \lceil |C||S_{\delta,t}|q^{-n}\rceil\}), \tag{18}$$

$$f_\delta(\mathbf{y}) = \frac{\sum_{j_1=1}^{K}\sum_{j_2=1}^{j_1-1} d^2_E(y_{j_1},y_{j_2})}{\sum_{j=1}^{K}\delta(y_j)}, \tag{19}$$

and $\widehat{\mathbf{y}}$ is a vector maximizing $f_\delta(\mathbf{y})$.

Even though $\widetilde{t}_K$ is a function not only of δ, but also of n, $|C|$ and q, we usually omit those parameters as we assume that they are fixed. The same is true for the dependence $f_\delta(\widehat{\mathbf{y}})$ has on q. We also remark that the minimum over t always exists since the sphere $S_{\delta,t}$ is defined with an inclusive inequality.

Proof. We start by representing the codewords in W as rows in a matrix M of type $K\times n$. Then we may write the average distance between the codewords as

$$d^2_{E\text{mean}}(W) = \frac{1}{\binom{K}{2}}\sum_{i=1}^{n}\sum_{j_1=1}^{K}\sum_{j_2=1}^{j_1-1} d^2_E(m_{j_1,i},m_{j_2,i}), \tag{20}$$

with the restriction

$$\sum_{i=1}^{n}\sum_{j=1}^{K}\delta(m_{j,i}) \leq Kt, \tag{21}$$

where $m_{j,i}$ is the element on row j and column i of M. The restriction comes from the weight of each codeword being at most t.

Define

$$f_\delta(\mathbf{y}) = \frac{\sum_{j_1=1}^{K}\sum_{j_2=1}^{j_1-1} d^2_E(y_{j_1},y_{j_2})}{\sum_{j=1}^{K}\delta(y_j)}, \tag{22}$$

where $\mathbf{y}$ is a vector of length K – a column in the matrix M. We then have

$$d^2_{E\text{mean}}(W) = \frac{1}{\binom{K}{2}}\sum_{i=1}^{n}\sum_{j_1=1}^{K}\sum_{j_2=1}^{j_1-1} d^2_E(m_{j_1,i},m_{j_2,i}) = \frac{1}{\binom{K}{2}}\sum_{i=1}^{n} f_\delta(m_{\cdot,i})\sum_{j=1}^{K}\delta(m_{j,i}), \tag{23}$$

where $m_{\cdot,i}$ is the i:th column vector of M.

Let $\widehat{\mathbf{y}}$ be a column vector such that $f_\delta(\widehat{\mathbf{y}})$ achieves its maximum value. We then have

$$\begin{aligned} d^2_{E\text{mean}}(W) &= \frac{1}{\binom{K}{2}}\sum_{i=1}^{n} f_\delta(m_{\cdot,i})\sum_{j=1}^{K}\delta(m_{j,i}) \leq \\ &\leq \frac{1}{\binom{K}{2}} f_\delta(\widehat{\mathbf{y}})\sum_{i=1}^{n}\sum_{j=1}^{K}\delta(m_{j,i}) \leq \frac{Kt}{\binom{K}{2}} f_\delta(\widehat{\mathbf{y}}) = \frac{2tf_\delta(\widehat{\mathbf{y}})}{K-1}, \end{aligned} \tag{24}$$

where the second inequality comes from 21. Note that this holds for any additive distance measure δ and any integer $K \in [2, |C|]$. We have proved the theorem. ■

We next find vectors $\widehat{\mathbf{y}}$ by which δ may be chosen to optimize the bound. We start with a lemma to show that the bound is independent of the scaling of δ.

Lemma 2 *The bound in Theorem 1 is scale invariant in the distance measure δ, i.e. for any $s > 0$, let $\lambda(x, y) = s\delta(x, y)$ for every pair x, y. Then*

$$\frac{2\widetilde{t}_K(\delta) f_\delta(\widehat{\mathbf{y}})}{K-1} = \frac{2\widetilde{t}_K(\lambda) f_\lambda(\widehat{\mathbf{y}})}{K-1} \tag{25}$$

holds.

Proof. Remark that $f_\lambda(\widehat{\mathbf{y}}) = f_\delta(\widehat{\mathbf{y}})/s$ since

$$f_\delta(\mathbf{y}) = \frac{\sum_{j_1=1}^{K} \sum_{j_2=1}^{j_1-1} d_E^2(y_{j_1}, y_{j_2})}{\sum_{j=1}^{K} \delta(y_j)} \ . \tag{26}$$

Furthermore, we have $S_{\lambda,t}(\mathbf{z}) = S_{\delta,t/s}(\mathbf{z})$, since

$$S_{\delta,t}(\mathbf{z}) = \{\mathbf{y} : \sum_{i=1}^{\lfloor q/2 \rfloor} \delta(i) c_i(\mathbf{z}, \mathbf{y}) \leq t\} \ . \tag{27}$$

Hence, $\widetilde{t}_K(\lambda) = s\widetilde{t}_K(\delta)$. It follows that,

$$\widetilde{t}_K(\lambda) f_\lambda(\widehat{\mathbf{y}}) = s\widetilde{t}_K(\delta) f_\delta(\widehat{\mathbf{y}})/s = \widetilde{t}_K(\delta) f_\delta(\widehat{\mathbf{y}}) \ . \tag{28}$$

■

In the proof of the following lemma we will need the so-called mediant addition:

$$\frac{a_1}{b_1} \oplus \frac{a_2}{b_2} = \frac{a_1 + a_2}{b_1 + b_2}, \tag{29}$$

presented in [3]. The number $\frac{a_1+a_2}{b_1+b_2}$ is called the *mediant* of $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$. It is similarly defined for c ratios $\frac{a_1}{b_1}, \ldots, \frac{a_c}{b_c}$, and is a weighted mean value of the ratios as can be seen by the identity

$$\frac{a_1}{b_1} \oplus \ldots \oplus \frac{a_c}{b_c} = \frac{b_1}{b_1 + \ldots + b_c}\frac{a_1}{b_1} + \ldots + \frac{b_c}{b_1 + \ldots + b_c}\frac{a_c}{b_c} \ . \tag{30}$$

As a weighted mean, the weights are strictly between 0 and 1, and are determined by the denominators only. We thus have $\frac{a_1}{b_1} < \frac{a_1+a_2}{b_1+b_2} < \frac{a_2}{b_2}$ if $\frac{a_1}{b_1} < \frac{a_2}{b_2}$, and $\frac{a_1}{b_1} = \frac{a_1+a_2}{b_1+b_2} = \frac{a_2}{b_2}$ if $\frac{a_1}{b_1} = \frac{a_2}{b_2}$.

Next, we intend to find vectors $\widehat{\mathbf{y}}$ such that $f_\delta(\widehat{\mathbf{y}})$ achieves its maximum. This we do only in the case $q = 8$, so from here and onwards the results are restricted to $q = 8$.

5 Results for $q = 8$

Lemma 3 *For any additive distance measure δ and for any K, one of the columns*

$$\begin{array}{lll} \widehat{\mathbf{y}}_1 = (1,-1,0,\ldots,0) & \widehat{\mathbf{y}}_2 = (2,0,\ldots,0), & \widehat{\mathbf{y}}_3 = (3,0,\ldots,0), \\ \widehat{\mathbf{y}}_4 = (4,0,\ldots,0), & \widehat{\mathbf{y}}_5 = (1,-2,0,\ldots,0), & \widehat{\mathbf{y}}_6 = (2,-2,0,\ldots,0) \end{array} \quad . \tag{31}$$

provides a maximum for

$$f(\mathbf{y}) = \frac{\sum_{j=1}^{k}\sum_{i=1}^{j-1} d_E^2(y_i,y_j)}{\sum_{i=1}^{k}\delta(y_i)} \quad . \tag{32}$$

Furthermore, $f(\widehat{\mathbf{y}}_2) = f(\widehat{\mathbf{y}}_6)$.

Proof. Maximization of the function $f_\delta(\mathbf{y})$ is done by a sequence of transformations of the variables. We first introduce the functions $a_i(\mathbf{y})$ that counts the number of occurences of i in the column $\mathbf{y}$. I. e. $a_0(\mathbf{y})$ is the number of zeros, $a_1(\mathbf{y})$ the number of 1:s, $a_7(\mathbf{y})$ the number of -1:s (as $-1 \equiv_8 7$), and so on. Since the length of the column $\mathbf{y}$ is K, we know that $K = \sum_{i=0}^{7} a_i(\mathbf{y})$.

The function $f_\delta(\mathbf{y})$ can then be rewritten as follows:

$$\begin{aligned} f_\delta(\mathbf{y}) &= \frac{\sum_{i=1}^{k}\sum_{j=1}^{i-1} d_E^2(y_i,y_j)}{\sum_{i=1}^{k}\delta(y_i)} = \\ &= \frac{\sum_{i=1}^{3}\sum_{j=0}^{7} d_E^2(i)a_j a_{j+i} \;+ \sum_{i=0}^{3} d_E^2(4)a_i a_{i+4}}{\sum_{i=1}^{3}\delta(i)(a_i + a_{-i}) \;+ \delta(4)a_4} \quad . \end{aligned} \tag{33}$$

The main objective is to maximize f_δ with respect to $a_0,\ldots,a_7$. Note that while there are 8^K different columns $(y_1,\ldots,y_K)$, there are only $\binom{8+K-1}{K}$ different vectors $(a_0,\ldots,a_7)$. This is the number of selections of K objects out of 8 alternatives with repetition but without order, since by going from $(y_1,\ldots,y_K)$ to $(a_0(\mathbf{y}),\ldots,a_7(\mathbf{y}))$ we have removed order changes that are insignificant for the value of f_δ. It is independent of rearrangements as $(y_1,y_2,\ldots,y_K) \to (y_2,y_1,\ldots,y_K)$.

We now proceed to the next transformation. The function can be expanded as

$$f_\delta = \frac{g_1 d_E^2(1) + g_2 d_E^2(2) + g_3 d_E^2(3) + g_4 d_E^2(4)}{\delta(1)\,(a_1+a_7) + \delta(2)\,(a_2+a_6) + \delta(3)\,(a_3+a_5) + \delta(4)a_4}, \tag{34}$$

where

$$\begin{aligned} g_1 &= (a_0a_1 + a_1a_2 + a_2a_3 + a_3a_4 + a_4a_5 + a_5a_6 + a_6a_7 + a_7a_0) \\ g_2 &= (a_0a_2 + a_1a_3 + a_2a_4 + a_3a_5 + a_4a_6 + a_5a_7 + a_6a_0 + a_7a_1) \\ g_3 &= (a_0a_3 + a_1a_4 + a_2a_5 + a_3a_6 + a_4a_7 + a_5a_0 + a_6a_1 + a_7a_2) \\ g_4 &= (a_0a_4 + a_1a_5 + a_2a_6 + a_3a_7) \quad . \end{aligned} \tag{35}$$

Now, rewriting f_δ in terms of the functions α_i, $i = 0, \ldots, 8$ where

$$\begin{array}{ll} \alpha_0 = a_0 & \alpha_4 = a_4 \\ \alpha_1 = a_1 + a_7 & \alpha_5 = a_3 - a_5 \\ \alpha_2 = a_2 + a_6 & \alpha_6 = a_2 - a_6 \\ \alpha_3 = a_3 + a_5 & \alpha_7 = a_1 - a_7 \end{array} \tag{36}$$

exploits the symmetries of the problem, and leaves us with a denominator which is far easier to handle. It has inverse relations

$$\begin{array}{ll} a_0 = \alpha_0 & a_4 = \alpha_4 \\ a_1 = \frac{\alpha_1+\alpha_7}{2} & a_5 = \frac{\alpha_3-\alpha_5}{2} \\ a_2 = \frac{\alpha_2+\alpha_6}{2} & a_6 = \frac{\alpha_2-\alpha_6}{2} \\ a_3 = \frac{\alpha_3+\alpha_5}{2} & a_7 = \frac{\alpha_1-\alpha_7}{2} \ . \end{array} \tag{37}$$

Pure calculation proves that

$$f_\delta = \frac{h_1 d_E^2(1) + h_2 d_E^2(2) + h_3 d_E^2(3) + h_4 d_E^2(4)}{4\left(\delta(1)\alpha_1 + \delta(2)\alpha_2 + \delta(3)\alpha_3 + \delta(4)\alpha_4\right)}, \tag{38}$$

where

$$\begin{array}{l} h_1 = 4\alpha_0\alpha_1 + 2\alpha_1\alpha_2 + 2\alpha_2\alpha_3 + 4\alpha_3\alpha_4 + 2\alpha_5\alpha_6 + 2\alpha_6\alpha_7 \\ h_2 = 4\alpha_0\alpha_2 + \alpha_1^2 + 2\alpha_1\alpha_3 + 4\alpha_2\alpha_4 + \alpha_3^2 - \alpha_5^2 + 2\alpha_5\alpha_7 - \alpha_7^2 \\ h_3 = 4\alpha_0\alpha_3 + 4\alpha_1\alpha_4 + 2\alpha_1\alpha_2 + 2\alpha_2\alpha_3 - 2\alpha_5\alpha_6 - 2\alpha_6\alpha_7 \\ h_4 = 4\alpha_0\alpha_4 + 2\alpha_1\alpha_3 + \alpha_2^2 - 2\alpha_5\alpha_7 - \alpha_6^2 \ . \end{array} \tag{39}$$

Recall that $a_0 = K - \sum_{i=1}^7 a_i$ which gives $\alpha_0 = K - \sum_{i=1}^4 \alpha_i$. Also, since we have the restriction $q = 8$, which is studied here, we have

$$\begin{aligned} & d_E^2(0) = 0, \\ & d_E^2(1) = d_E^2(7) = 4\sin^2\frac{\pi}{8} = 2 - \sqrt{2}, \\ & d_E^2(2) = d_E^2(6) = 4\sin^2\frac{2\pi}{8} = 2, \\ & d_E^2(3) = d_E^2(5) = 4\sin^2\frac{3\pi}{8} = 2 + \sqrt{2}, \text{ and} \\ & d_E^2(4) = 4\sin^2\frac{4\pi}{8} = 4 \ . \end{aligned} \tag{40}$$

This makes it possible to rewrite f_δ as

$$f_\delta = \frac{4K\sum_{i=1}^4 \alpha_i d_E^2(i) - \left(\sum_{i=1}^4 \alpha_i d_E^2(i)\right)^2 - \left(\sqrt{2}\alpha_5 + 2\alpha_6 + \sqrt{2}\alpha_7\right)^2}{4\sum_{i=1}^4 \delta_i\alpha_i} \ . \tag{41}$$

By construction, for any $\mathbf{x}$ the quantities $\alpha_1(\mathbf{x})$, $\alpha_2(\mathbf{x})$, $\alpha_3(\mathbf{x})$ and $\alpha_4(\mathbf{x})$ are non-negative integers, at least one greater than zero, while $\alpha_5(\mathbf{x})$, $\alpha_6(\mathbf{x})$ and $\alpha_7(\mathbf{x})$ are integers, possibly negative. Furthermore, we know that $\alpha_1(\mathbf{x})+\alpha_7(\mathbf{x})$, $\alpha_2(\mathbf{x}) + \alpha_6(\mathbf{x})$ and $\alpha_3(\mathbf{x}) + \alpha_5(\mathbf{x})$ are even numbers.

Next we consider the function

$$\varphi_\delta = \frac{4K\sum_{i=1}^4 \alpha_i d_E^2(i) - \left(\sum_{i=1}^4 \alpha_i d_E^2(i)\right)^2}{4\sum_{i=1}^4 \delta(i)\alpha_i} \ . \tag{42}$$

Obviously we have $f_\delta(\mathbf{x}) \leq \varphi_\delta(\mathbf{x})$ for any $\mathbf{x}$.

Let $\mathbf{y}_1 = (1,0,\ldots,0)$, $\mathbf{y}_2 = (2,0,\ldots,0)$, $\mathbf{y}_3 = (3,0,\ldots,0)$, $\mathbf{y}_4 = (4,0,\ldots,0)$, and observe that we have

$$\begin{array}{ll} f_\delta(\mathbf{y}_1) = \frac{4Kd_E^2(1)-d_E^4-2}{4\delta(1)} & f_\delta(\mathbf{y}_2) = \frac{4Kd_E^2(2)-d_E^4-4}{4\delta(2)} \\ f_\delta(\mathbf{y}_3) = \frac{4Kd_E^2(3)-d_E^4-2}{4\delta(3)} & f_\delta(\mathbf{y}_4) = \frac{4Kd_E^2(4)-d_E^4}{4\delta(4)} \ . \end{array} \tag{43}$$

Now, any column $\mathbf{x}$ where

$$\begin{aligned} f_\delta(\mathbf{x}) \leq \frac{\alpha_1(\mathbf{x})f_\delta(\mathbf{y}_1)}{\alpha_1(\mathbf{x})} \oplus \frac{\alpha_2(\mathbf{x})f_\delta(\mathbf{y}_2)}{\alpha_2(\mathbf{x})} \oplus \frac{\alpha_3(\mathbf{x})f_\delta(\mathbf{y}_3)}{\alpha_3(\mathbf{x})} \oplus \frac{\alpha_4(\mathbf{x})f_\delta(\mathbf{y}_4)}{\alpha_4(\mathbf{x})} \leq \\ \leq \max(f_\delta(\mathbf{y}_1), f_\delta(\mathbf{y}_2), f_\delta(\mathbf{y}_3), f_\delta(\mathbf{y}_4)) \end{aligned} \tag{44}$$

cannot be extremal. There are still many columns that must be compared, so we start by discarding a large set of columns which cannot be extremal.

Since we have $f_\delta \leq \varphi_\delta$, we may first discard all $\mathbf{x}$ where

$$\varphi_\delta(\mathbf{x}) \leq \frac{\alpha_1(\mathbf{x})f_\delta(\mathbf{y}_1)}{\alpha_1(\mathbf{x})} \oplus \frac{\alpha_2(\mathbf{x})f_\delta(\mathbf{y}_2)}{\alpha_2(\mathbf{x})} \oplus \frac{\alpha_3(\mathbf{x})f_\delta(\mathbf{y}_3)}{\alpha_3(\mathbf{x})} \oplus \frac{\alpha_4(\mathbf{x})f_\delta(\mathbf{y}_4)}{\alpha_4(\mathbf{x})} \ . \tag{45}$$

This condition is equivalent to

$$\begin{aligned} &\frac{4K\sum_{i=1}^4 \alpha_i(\mathbf{x}) d_E^2(i) - \left(\sum_{i=1}^4 \alpha_i(\mathbf{x}) d_E^2(i)\right)^2}{4\sum_{i=1}^4 \delta(i)\alpha_i(\mathbf{x})} \leq \\ &\quad \leq \frac{4K\sum_{i=1}^4 \alpha_i(\mathbf{x}) d_E^2(i) - \sum_{i=1}^4 \alpha_i(\mathbf{x}) d_E^4(i) - 2\alpha_1(\mathbf{x}) - 4\alpha_2(\mathbf{x}) - 2\alpha_3(\mathbf{x})}{4\sum_{i=1}^4 \delta(i)\alpha_i(\mathbf{x})}, \end{aligned} \tag{46}$$

which can also be expressed as

$$\sum_{i=1}^4 \alpha_i(\mathbf{x}) d_E^4(i) + 2\alpha_1(\mathbf{x}) + 4\alpha_2(\mathbf{x}) + 2\alpha_3(\mathbf{x}) \leq \left(\sum_{i=1}^4 \alpha_i(\mathbf{x}) d_E^2(i)\right)^2 \ . \tag{47}$$

We may immediately discard all columns $\mathbf{x}$ with $\alpha_1(\mathbf{x}) \geq 7$, $\alpha_2(\mathbf{x}) \geq 3$, $\alpha_3(\mathbf{x}) \geq 2$ or $\alpha_4(\mathbf{x}) \geq 2$, as any of these inequalities being satisfied will make inequality 47 true. This leaves us with $8 \cdot 4 \cdot 3 \cdot 3 = 288$ columns to examine. Testing these with condition 44, we find only 14 columns which may be extremal, namely

$$\begin{array}{ll} \mathbf{y}_1 = (1,0,\ldots,0), & \mathbf{y}_2 = (2,0,\ldots,0), \\ \mathbf{y}_3 = (3,0,\ldots,0), & \mathbf{y}_4 = (4,0,\ldots,0), \\ \mathbf{y}_5 = (1,-2,0,\ldots,0), & \mathbf{y}_6 = (2,-2,0,\ldots,0), \\ \mathbf{y}_7 = (1,-3,0,\ldots,0), & \mathbf{y}_8 = (1,-1,0,\ldots,0), \\ \mathbf{y}_9 = (1,1,-2,0,\ldots,0), & \mathbf{y}_{10} = (1,1,-1,-2,0,\ldots,0), \\ \mathbf{y}_{11} = (1,1,-1,0,\ldots,0), & \mathbf{y}_{12} = (1,1,-1,-1,0,\ldots,0), \\ \mathbf{y}_{13} = (1,1,1,-1,-1,0,\ldots,0), & \mathbf{y}_{14} = (1,1,1,-1,-1,-1,0,\ldots,0) \ . \end{array} \tag{48}$$

Since

$$f_\delta(\mathbf{y}_{10}) = f_\delta(\mathbf{y}_{10}) \oplus f_\delta(\mathbf{y}_{10}) < f_\delta(\mathbf{y}_8) \oplus f_\delta(\mathbf{y}_8) \oplus f_\delta(\mathbf{y}_8) \oplus f_\delta(\mathbf{y}_6) \leq \\ \leq \max(f_\delta(\mathbf{y}_8), f_\delta(\mathbf{y}_6)), \tag{49}$$

we may conclude that $f_\delta(\mathbf{y}_{10})$ cannot be extremal. Continuing to reduce the set of columns in our list in this manner, we end up with only the vectors given in the lemma. Also, $f_\delta(\widehat{\mathbf{y}}_2) = f_\delta(\widehat{\mathbf{y}}_6)$ follows from $f_\delta(\widehat{\mathbf{y}}_2) \oplus f_\delta(\widehat{\mathbf{y}}_2) = f_\delta(\widehat{\mathbf{y}}_6)$. Those six columns are extremal, and the set cannot be reduced further. This follows by considering different distance measures δ such that they all cause f_δ to achieve its maximum value, which we know it does for at least one of the six columns. Considering the distance measure

$$\begin{array}{ll} \delta(1) = (2-\sqrt{2})(K-1) - 3 + 2\sqrt{2}, & \delta(2) = 2(K-1) + \sqrt{2} - 1, \\ \delta(3) = (2+\sqrt{2})(K-1), & \delta(4) = 4(K-1), \end{array} \tag{50}$$

we get that $\widehat{\mathbf{y}}_1$, $\widehat{\mathbf{y}}_3$, $\widehat{\mathbf{y}}_4$ and $\widehat{\mathbf{y}}_5$ are extremal. Considering the distance measure

$$\begin{array}{ll} \delta(1) = (2-\sqrt{2})(K-1) - 4 + 3\sqrt{2}, & \delta(2) = 2(K-1), \\ \delta(3) = (2+\sqrt{2})(K-1), & \delta(4) = 4(K-1), \end{array} \tag{51}$$

we get that $\widehat{\mathbf{y}}_2$, $\widehat{\mathbf{y}}_3$, $\widehat{\mathbf{y}}_4$, $\widehat{\mathbf{y}}_5$ and $\widehat{\mathbf{y}}_6$ are extremal. ■

Equipped with knowledge about which columns are extremal, we can specify which δ that minimize $f_\delta(\widehat{\mathbf{y}})$.

Theorem 4 *Distance measures δ which minimize $f_\delta(\widehat{\mathbf{y}})$ can in the case $q = 8$ be described as*

$$\begin{aligned} \delta(1) &= (2-\sqrt{2})(K-1) - 2 + 2\sqrt{2} - h \\ \delta(2) &= 2(K-1) + h \\ \delta(3) &= (2+\sqrt{2})(K-1) \\ \delta(4) &= 4(K-1), \end{aligned} \tag{52}$$

for any $h \in [0, \sqrt{2}-1]$.

Furthermore, with

$$\sum_{i=1}^{4} \delta(i) = 10(K-1) - 2 + 2\sqrt{2}, \tag{53}$$

the minimum value of $f_\delta(\widehat{\mathbf{y}})$ is 1.

Proof. Let $B = f_\delta(\widehat{\mathbf{y}})$. We then have

$$f_\delta(\widehat{\mathbf{y}}_1) \le B \Leftrightarrow \frac{(2-\sqrt{2})(K-1)-1+\sqrt{2}}{\delta(1)} \le B, \tag{54}$$

$$f_\delta(\widehat{\mathbf{y}}_2) \le B \Leftrightarrow \frac{2(K-1)}{\delta(2)} \le B, \tag{55}$$

$$f_\delta(\widehat{\mathbf{y}}_3) \le B \Leftrightarrow \frac{(2+\sqrt{2})(K-1)}{\delta(3)} \le B, \tag{56}$$

$$f_\delta(\widehat{\mathbf{y}}_4) \le B \Leftrightarrow \frac{4(K-1)}{\delta(4)} \le B, \tag{57}$$

$$f_\delta(\widehat{\mathbf{y}}_5) \le B \Leftrightarrow \frac{(4-\sqrt{2})(K-1)-2+2\sqrt{2}}{\delta(1)+\delta(2)} \le B, \tag{58}$$

where at least one of the inequalities is an equality. From 56, 57 and 58 we get

$$\frac{10(K-1)-2+2\sqrt{2}}{\delta(1)+\delta(2)+\delta(3)+\delta(4)} \le B \tag{59}$$

by use of mediant addition. By the normalization 53 on δ, we thus have $B \ge 1$. By letting 56, 57 and 58 all be equalities, we get $B = 1$, which is the lowest value we can get on B. (Using 54, 55, 56 and 57 in a similar manner only gives a less tight bound on B, and so does not determine B.)

So distance measures δ which minimize B, and give $B = 1$ must give equality in 56, 57 and 58 and satisfy 54 and 55. This is exactly the distance measures described in the theorem. ■

Combining Theorem 1 and Theorem 4 with Lemma 2, we get the following corollary.

Corollary 5 *For any code C in $\mathbf{Z}_8^n$ we have*

$$d_{E\min}^2(C) \le \min_{K\in[2,|C|]} \min_{\delta\in\Delta} \frac{2\widetilde{t}_K}{K-1} \tag{60}$$

holds, where Δ is the set of distance measures with

$$\sum_{i=1}^{4} \delta(i) = 10(K-1)-2+2\sqrt{2} \tag{61}$$

and

$$\widetilde{t}_K(\delta) = \min(\{t : K \le \lceil |C||S_{\delta,t}|q^{-n} \rceil\}) \text{ and} \tag{62}$$

$$f_\delta(\widehat{\mathbf{y}}) = \max_{\mathbf{y}} \left(\frac{\sum_{j_1=1}^{K} \sum_{j_2=1}^{j_1-1} d_E^2(y_{j_1}, y_{j_2})}{\sum_{j=1}^{K} \delta(y_j)} \right) . \tag{63}$$

6 Conclusions

In this paper we have improved previous upper bounds for the minimum Euclidean distance. One of the bounds is valid for any combination of the three parameters $|C|$, n and q, while the other is explicit in the two parameters $|C|$, $n$case $q = 8$.

The results develop the Elias sphere method to provide an improved bound on the minimum Euclidean distance that is non-asymptotic. The proof method is not tied to a certain structure of codes, and applies for any PSK block code with parameters $q = 8$, n and $|C|$. This means that one possible continuation is to investigate other distance measures than a Euclidean by following a similar path starting with the Elias' sphere. It may be an even more challenging task to investigate if similar bounds also can be constructed for PSK Trellis codes, having a different basic structure.

Next we give an overview over the technical method of this paper. First Elias' method, [1] pp. 318-321, was followed. Here the problem was localized to a critical sphere, where codewords are at least as dense as elsewhere in the code, and the minimum distance between codewords in the critical sphere is trivially bounded by the average distance between them. However, in this line of research, the critical sphere is defined in terms of a general distance measure, δ, characterized by its values $\delta(i)$ for integers i, called the *inner distance measure*. Later, the values of the coefficients were chosen to obtain the best possible bound on the *outer distance measure*, which here is the squared Euclidean distance.

Still following Elias, the average distance between the codewords in the sphere is bounded by listing the codewords as rows in a matrix, and seeking columns of the matrix which will maximize the average distance. Elias sought columns that maximized the average distance between codewords and he considered compositions of the columns, allowing columns where each symbol may appear a continuous number of times. Allowing such compositions is an approximation which we avoid. Instead, by fixing $q = 8$, we found all columns which can give maximal average distance between the codewords in the critical sphere, independently of the inner distance measure δ. Such columns are called extremal columns. We then chose the values of $\delta(i)$ to optimize the bound on $d^2_{E\,\min}(C)$, which gave one of the main results.

The bound is a product of two factors, both depending on the shape of the spheres. We minimized one of these factors, namely the factor which intuitively is more sensitive to the inner distance measure. While this is not certain to optimize the bound on the $d^2_{E\,\min}(C)$, consisting of two factors, the new bound is as good as previous bounds, and is strictly better for low-rate codes.

Central in the solution is that for $q = 8$ it turns out that only six columns can be singled out as extremal, independently of the inner distance measure. Of these six, for given n and $|C|$, five are used since two of them give identical values, defining an optimal inner distance measure with respect to one factor of the bound.

While there are many high-rate codes which meet the bounds (older bounds as well as the new bound), only a few medium-rate codes and no low-rate codes

which meet the bound are known. It is thus of interest for low and medium rates either to improve the bound or to find codes C with higher $d^2_{E\,\min}(C)$.

References

[1] Berlekamp E. R., *Algebraic Coding Theory*, New York: McGraw Hill, 1968.

[2] Golomb S. W., Welch L. R., Perfect codes in the Lee Metric and the Packing of Polynominoes, *SIAM Jou. of Appl. Math.*, vol. 18, no. 2, Januari, 1970.

[3] Graham R., Knuth D., Patashnik O., *Concrete Mathematics*, Addison Wesley, ISBN 0-201-14236-8, 1994.

[4] Imai, H., Hirakawa, S., A new multilevel coding method using error-correcting codes, *IEEE Transactions of Information Theory*, Volume 23, Issue 3, May 1997, pp. 371-377.

[5] Laksman E., Lennerstad H., Nilsson M., Improving bounds on the minimum Euclidean distance for block codes by inner metric optimization, Combinatorics 2008, to appear in *Discrete Mathematics*, Verona, Italy, 2008.

[6] Nilsson M., Lennerstad H., An Upper Bound on the Minimum Euclidean Distance for Block Coded Phase Shift Keying, *IEEE Trans. Inform. Theory* 46(2): pp. 656-662, 2000.

[7] Nilsson M., Lennerstad H., Improved Upper Bound on the Minimum Euclidean Distance for Block Coded Phase Shift Keying, *proceedings of RVK05*, Linköping, Sweden, 2005.

[8] Nilsson M., Lennerstad H., Laksman E., A Two-Metric Approach to Improve Bounds on the Minimum Euclidean Distance for Block Codes, *proceedings of RVK08*, Växjö, Sweden, 2008.

[9] MacWilliams F. J., Sloane N. J. A., *The Theory of Error-Correcting Codes*, Amsterdam, The Netherlands: North-Holland, 1977.

[10] Piret Ph., Bounds for Codes Over the Unit Circle, *IEEE Trans. Inform. Theory*, vol. IT-32, pp.760-767, Nov. 1986.

[11] Sayegh, S., A Class of Optimum Block Codes in Signal Space, *IEEE Transactions on Communication*, Volume 34, Issue 10, Oct 1986, Page(s): 1043-1045.

[12] Tanabe, H., Umeda, H., Salam, M. A., A new construction method of multilevel coded modulation with a good Euclidean minimum distance, 1997 *IEEE International Symposium on Information Theory*, 29 June – 4 July 1997, pp. 437.

[13] Wyner, A. D., Bounds on Communication with Polyphase Coding, *Bell System Technical Journal*, 45, April 1966, pp. 523-559.

Paper II

Bounding the minimum Euclidean distance for any PSK block codes of alphabet size 8

Efraim Laksman, Håkan Lennerstad, Magnus Nilsson

Bounding the minimal Euclidean distance for any PSK block codes of alphabet size 8

Efraim Laksman
Blekinge Institute of Technology
MMS
Karlskrona, Sweden
Email: efl@bth.se

Håkan Lennerstad
Blekinge Institute of Technology
MMS
Karlskrona, Sweden
Email: hln@bth.se

Magnus Nilsson
Blekinge Institute of Technology
PTI
Karlskrona, Sweden
Email: nim@bth.se

***Abstract*—We consider a previously known general bound for minimal Euclidean distance for PSK block codes with eight letters, attained by generalizing the method of Elias spheres, that is optimal for many values of the parameters block length and code size, and we show that this bound is stronger than has earlier been proven.**

We also improve this bound somewhat.

I. Introduction

We wish to set a bound for how efficient combined coding and modulation may be. We restrict ourselves to PSK-codes, i. e. codes with symbols on the unit circle, and allow any word length n and any code size, $|C|$, but when it comes to alphabet size q, we deal only with $q = 8$. We measure the distance between code words by Euclidean distance, which is commonly regarded as the most relevant measure of efficiency of this type of codes. The distance between two symbols i and j is measured as

$$d(i-j) = d_{i-j} = 2|\sin\left(\frac{(i-j)\pi}{q}\right)|.$$

The distance between two words $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ is

$$d(x-y) = \sqrt{\sum_{i=1}^{n} d^2(x_i - y_i)},$$

and the minimal Euclidean distance of a code C is defined as

$$d^2_{E\min}(C) = \min_{\substack{x,y \in C \\ x \neq y}} \sum_{j=1}^{n} d^2_{x_j - y_j}.$$

We consider an upper bound on $d^2_{E\min}(C)$ for any block code with $q = 8$, linear or non-linear, that is explicit in the two parameters n and $|C|$, i. e. we bound the largest possble distance between the two members in C that are closest, measured in Euclidean distance.

Many known codes fulfil the bound with equality, proving both that these codes are sharp and that the bound is optimal for these values of n and $|C|$. We consider the same bound as in the previous paper [6], which is derived by means of an Elias sphere argument and an inner metric argument, which was new in that paper. That bound optimizes a part of a certain function that provides a bound of $d^2_{E\min}(C)$. In this paper we prove that the bound presented in [6] is exhaustive in that it actually bounds the entire bounding function.

II. Previous work

The bound in this line of research is partly based on the arguments leading to the well-known Elias bound (see [2] pp. 318-321, [4] pp. 558-564). Elias bound arguments have been used by Wyner [10] to derive an asympthotic bound for $q \to \infty$ and by Piret [9] for an asymthotic bound for $n \to \infty$. One development appeared in the paper [1], where the Elias sphere argument was used to find a non-asymthotic bound – explicit in n, q and $|C|$.

In [8], for the first time the idea of optimizing a general inner metric to improve the bound in the outer metric, which is Euclidean, was presented. An inner metric $\{\delta\}$ defines the Elias sphere $S_t(w) = \{y|\delta(w,y) \leq t\}$, and is optimized to give the best possible bound for the outer metrics. It was considered in the case $q = 8$ only. By optimizing the inner metric of the Elias sphere, a sharper bound could be achieved. This is proved in [6], where a so called the k-dependent metric φ was calculated:

$$\begin{aligned}
\varphi_1 &= d_1^2(k-1) - 2 + 2\sqrt{2} - h,\\
\varphi_2 &= d_2^2(k-1) + h,\\
\varphi_3 &= d_3^2(k-1),\\
\varphi_4 &= d_4^2(k-1),
\end{aligned}$$

for some $h \in [0, \sqrt{2}-1]$. In the present paper we prove that this metric provides an exhaustive solution: it cannot be improved within the problem formulation.

We start with Elias' method of a critical sphere. We form critical spheres $S_t(x) = \{y|\delta(x,y) \leq t\}$, for some metric δ, around each code word x, each of which contains $|S_t|$ words. Some word, w, has to be contained in at least $k = \left\lceil \frac{|C||S_t|}{q^n} \right\rceil$ critical spheres, as can be seen by the use of the pigeon hole principle. We now form a sphere around w, $S_t(w) = \{y|\delta(w,y) \leq t\}$, which must contain k code words. We localize the problem to this sphere by the first inequality, and limit a minimum from above by an average by the second inequality:

$$d^2_{E_{min}}(C) \leq d^2_{E_{min}}(S_t(w)) \leq d^2_{E_{\text{average}}}(S_t(w)).$$

We will use the notation $\delta(i) = \delta_i = \delta(i,0)$, and allow any linear metric, i. e. a metric δ such that $\delta(x,y) = \sum \delta(x_i, y_i)$, with $\delta_0 = 0$. We denote the set of all such inner metrics by Δ.

We next form a k by n matrix M, which contains the k code words inside $S_t(w)$ as rows. We then have

$$d^2_{E_{\text{average}}}(S_t(w)) = \frac{1}{\binom{k}{2}} \sum_{i_1=1}^{k} \sum_{i_2=1}^{i_1-1} \sum_{j=1}^{n} d^2_{m_{i_1,j}-m_{i_2,j}},$$

where $m_{i,j}$ is the element on position (i,j) in the matrix M, i. e. the j:th letter in the i:th word. So the tripple sum involves every pair of code words in the sphere (the two first sums), letter by letter (the last sum).

Annotate that if the number of columns in M increase without violating the maximum sum t on each row, all arguments in the following are valid.

Since we have $\delta(x) \leq t$, for every word x in the sphere, we have that the weight of each of the k rows in the matrix is at most t, so the weight of the entire matrix is

$$|M| = \sum_{i=1}^{k} \sum_{j=1}^{n} \delta_{m_{i,j}} \leq kt.$$

We may swap order of the sums. Since we're looking for an upper bound, we are interested in the worst possible matrix:

$$d^2_{E_{\text{average}}}(S_t(w)) \leq \frac{1}{\binom{k}{2}} \max_{M:|M|\leq kt} \sum_{j=1}^{n} \sum_{i_1=1}^{k} \sum_{i_2=1}^{i_1-1} d^2_{m_{i_1,j}-m_{i_2,j}}.$$

We now consider, and denote, the contribution of a column y in comparison to its weight

$$f_\delta(y) = \frac{\sum_{i_1=1}^{k} \sum_{i_2=1}^{i_1-1} d^2_{y_{i_1}-y_{i_2}}}{\sum_{j=1}^{k} \delta_{y_j}}.$$

Again worst case columns give an upper bound:

$$d^2_{E_{\text{average}}}(S_t(w)) \leq \max_{y\in\mathbf{Z}_q^k} \frac{k f_\delta(y) t(\delta,k)}{\binom{k}{2}} = \max_{y\in\mathbf{Z}_q^k} \frac{2 f_\delta(y) t(\delta,k)}{k-1}.$$

Observe that every $\delta \in \Delta$ and any integer k greater than one gives a bound, but we will find δ and k that makes the bound as tight as possible. So we have

$$\begin{aligned} d^2_{E_{\text{average}}}(S_t(w)) &\leq \min_k \min_{\delta\in\Delta} \max_{y\in\mathbf{Z}_q^k} \frac{2t(\delta,k) f_\delta(y)}{k-1} \leq \\ &\leq \min_{\delta\in\Delta} \max_{y\in\mathbf{Z}_q^k} \frac{2 f_\delta(y) t(\delta,k)}{k-1}. \end{aligned}$$

Since k belongs to a discrete set, the quantity in the right hand side can be calculated for many k, and the smallest result can be used.

In [6] extremal columns $\widehat{y} : f(y) \leq f(\widehat{y})$ were found that are independent on the metric.

Lemma 1: For any metric $\{\delta_i\}_{i=1}^4$ and for any k, one of the columns

$$\begin{aligned} \widehat{y}_1 &= (1,-1,0,...,0) \\ \widehat{y}_2 &= (2,0,...,0), \\ \widehat{y}_3 &= (3,0,...,0), \\ \widehat{y}_4 &= (4,0,...,0), \\ \widehat{y}_5 &= (1,-2,0,...,0) \\ \widehat{y}_6 &= (2,-2,0,...,0). \end{aligned}$$

provides a maximum for

$$f(y) = \frac{\sum_{j=1}^{k} \sum_{i=1}^{j-1} d^2_{y_i-y_j}}{\sum_{i=1}^{k} \delta_{y_i}}.$$

Also, $f(\widehat{y}_2) = f(\widehat{y}_6)$ for any inner metric, so we may disregard from $\widehat{y}_6$.

Using the lemma, we have the bound

$$d^2_{E\,\min}(C) \leq \min_k \min_\delta \max_{i\in\{1,2,3,4,12\}} \frac{2t(\delta,k) f_\delta(\widehat{y}_i)}{k-1}.$$

In [6], we found a specific metric $\delta = \varphi$, called a k-dependent metric, that minimizes $\max_y f(\widehat{y})$.

Theorem 2: A metric that minimizes $\max_y f(\widehat{y})$ is

$$\begin{aligned} \varphi_1 &= d_1^2(k-1) - 2 + 2\sqrt{2} - h, \\ \varphi_2 &= d_2^2(k-1) + h, \\ \varphi_3 &= d_3^2(k-1), \\ \varphi_4 &= d_4^2(k-1), \end{aligned}$$

for some $h \in [0, \sqrt{2}-1]$.

Certainly, this result is incomplete since the quantity which we want to minimize is not $f_\delta(\widehat{y})$, but $t(\delta,k) f_\delta(\widehat{y})$.

III. Results

The result in this paper says that the same metric, φ, actually minimizes $t(\delta,k) f_\delta(\widehat{y})$, we have also two inprovements of this bound. We remark that we do not prove that it is the same value of h that minimizes both $f_\delta(\widehat{y})$ and $t(\delta,k) f_\delta(\widehat{y})$.

Theorem 3: The k-dependent metric $\delta = \varphi$ of Theorem 2 mimimizes

$$\min_\delta \max_{i\in\{1,2,3,4,12\}} \frac{2t(\delta,k) f_\delta(\widehat{y}_i)}{k-1},$$

for any k, giving the bound

$$d^2_{E_{\min}} \leq \min_{k,\delta} \max_{i\in\{1,2,3,4,12\}} \frac{2t(\delta,k) f_\delta(\widehat{y}_i)}{k-1}.$$

Theorem 4: The bound of Theorem 3 can be sharpened to

$$d^2_{E_{\min}} \leq \max_{a\in A} \sum_{i=1}^{4} a_i d_i^2,$$

where A is the set of $a = (a_1, a_2, a_3, a_4)$ such that $a_i \geq 0$ for $i = 1,2,3,4$, $\sum_{i=1}^4 a_i \leq n$ and

$$\sum_{i=1}^{4} a_i d_i^2 \leq \min_{k,\delta} \max_{i\in\{1,2,3,4,12\}} \frac{2t(\varphi,k) f_\varphi(\widehat{y}_i)}{k-1}.$$

Proof: The distance between any two words must be of the form

$$\sum_{i=1}^{4} a_i d_i^2,$$

so that the minimum distance also must be of this form. Since it must adhere to the bound in Theorem 3, which is not necessarily of this form, we may sometimes truncate the bound a bit, as stated in Theorem 4. Table I shows optimality of some codes according to the bound of Theorem 3. ∎

Our next theorem gives a further improvement together with Theorem 4. We here consider differences, and restrict the set of words under consideration by adding extra columns, so that all "words", albeit too long to be actual words, have weight $t(\delta, k)$, or almost $t(\delta, k)$.

Theorem 5: We have the bound

$$d_{E_{\min}}^2 \leq \max \sum_{i=1}^{4} |a_i - b_i| d_i^2$$

over all $a = (a_1, a_2, a_3, a_4)$ and $b = (b_1, b_2, b_3, b_4)$ in A', where A' is a set of elements $a = (a_1, a_2, a_3, a_4)$ such that

$$t(\delta, k) - d_1^2 < \sum_{i=1}^{4} a_i d_i^2 \leq t(\delta, k)$$

and $a_i \geq 0$ for $i = 1, 2, 3, 4$.

Proof: This bound introduces a symmetry that in some cases give a sharper bound. It is a counterpart of Theorem II.2 in [1] where the factor 2 in the second term appeared by a similar argument to the one in this proof – in that paper significantly decreasing the set of possible values of $\sum_{i=1}^{2} a_i d_i^2$, which gave a lower (sharper) bound.

In page 2 the matrix M is defined as a $k \times n$-matrix of k codewords and thus n columns. However, the number of columns of M may be expanded beyond n since the bound relies only on the maximal column and that a codeword has maximum weight t. We do not need to consider a codeword x with weight $\sum_{i=1}^{4} a_i d_i^2 \leq t - d_1^2$, since we then can add a new column with 1 in the row of x (or a larger value i if $\sum_{i=1}^{4} a_i d_i^2 \leq t + d_i^2$), giving the word x'. All other elements in the extra column are zero. Denote the new $k \times (n+1)$-matrix by M'. The maximum column of M' is the same as for M, deduced in page 2, since all codewords, now of length $n+1$, still has maximum weight t. Furthermore, the change cannot decrease any distance between words, since all other column elements are zero. Hence,

$$d_{E_{\text{average}}}^2(M) \leq d_{E_{\text{average}}}^2(M') \leq \min_{k,\delta} \max_{i \in \{1,2,3,4,12\}} \frac{2t(\delta, k) f_\delta(\widehat{y})}{k-1}.$$

By repeating this argument we can successively replace the set of words from the constraint $\sum_{i=1}^{4} a_i \leq n$ to the constraint that all words are close to the border of the sphere. Over this set of words the maximum of $\sum_{i=1}^{4} |a_i - b_i| d_i^2$ may very well become lower since extra cancellation occurs in the difference $a_i - b_i$, as in Theorem II.2 in [1]. ∎

It is worth noting that when we optimize the bound, we take no respect to truncation as in Theorem 4 and 5 that may be possible. This means that "special" metrics, forbidding some symbols in the spheres and thus also in M, may lead to stronger truncation and thus perhaps even to a sharper bound. For an example of how this may be utilized to get a stronger bound, see [1].

Before we prove Theorem 3, we state the scale invariance of the bound that is proven in [6]:

Lemma 6: The bound is scale invariant in the metric $\{\delta_i\}_{i=1}^4$, i.e.

$$\min_{\delta} \frac{2t(s\delta, k) f_{s\delta}(\widehat{y})}{k-1} = \min_{\delta} \frac{2t(\delta, k) f_\delta(\widehat{y})}{k-1}$$

if $s\delta = (s\delta_1, s\delta_2, s\delta_3, s\delta_4)$, for any $s > 0$.

Proof of Theorem 3: Denote

$$\begin{aligned} D_1 &= \frac{1}{2} d_2^2 + (k-2) d_1^2 \\ D_{12} &= d_3^2 + (k-2)(d_1^2 + d_2^2) \\ D_2 &= (k-1) d_2^2 \\ D_3 &= (k-1) d_3^2 \\ D_4 &= (k-1) d_4^2. \end{aligned}$$

Remark that D_i is the contribution of $\widehat{y}_i$ to the matrix M. Remark also that φ is such that the ratio between the contribution and the weight for the columns $\widehat{y}_3$, $\widehat{y}_4$ and $\widehat{y}_{12}$ (and possibly either $\widehat{y}_1$ or $\widehat{y}_2$) are equal.

Involving the five columns of Lemma 1, we want to minimize

$$\max_y t(\delta, k) f(y) = \max t(\delta, k) \left(\frac{D_1}{\delta_1}, \frac{D_{12}}{\delta_1 + \delta_2}, \frac{D_2}{\delta_2}, \frac{D_3}{\delta_3}, \frac{D_4}{\delta_4} \right)$$

over the metric δ.

We thus consider the five functions

$$\begin{aligned} f_j(\delta_1, \delta_2, \delta_3, \delta_4) &= D_j \frac{\sum_{i=1}^{4} \delta_i a_i}{\delta_j}, j = 1, 2, 3, 4, \\ f_{12}(\delta_1, \delta_2, \delta_3, \delta_4) &= D_{12} \frac{\sum_{i=1}^{4} \delta_i a_i}{\delta_1 + \delta_2}, \end{aligned}$$

and we want to find a minimum for

$$\max(f_1, f_2, f_3, f_4, f_{12}).$$

We call a function critical if it fulfils the maximum.

Lemma 7: There exists an inner metric δ which minimizes

$$\frac{2t(\delta, k) f_\delta(\widehat{y})}{k-1}$$

such that $f_3 = f_4 = f_{12}$ are critical.

Proof: When we minimize

$$\frac{2t(\delta, k) f_\delta(\widehat{y})}{k-1}$$

over all δ, at least some function f_i, $i \in \{1, 2, 3, 4, 12\}$ is critical. Let's assume that f_1 is critical. But f_1 being critical is equivalent to

$$\begin{aligned} &D_1 \frac{\sum_{i=1}^{4} a_i \delta_i}{\delta_1} \geq D_j \frac{\sum_{i=1}^{4} a_i \delta_i}{\delta_j}, \text{ for } j = 2, 3, 4 \text{ and} \\ &D_1 \frac{\sum_{i=1}^{4} a_i \delta_i}{\delta_1} \geq D_{12} \frac{\sum_{i=1}^{4} a_i \delta_i}{\delta_1 + \delta_2} \end{aligned}$$

for some $a = (a_1, a_2, a_3, a_4)$. This in turn is equivalent to

$$\delta_j \geq \frac{D_1}{D_j}\delta_1, \text{ for } j = 2, 3, 4 \text{ and}$$
$$\delta_2 \geq \frac{D_{12} - D_1}{D_1}\delta_1.$$

We now decrease δ_2, δ_3 and δ_4 until we have equality in three of the inequalities. The one in which we cannot achieve equality is

$$\delta_2 \geq \frac{D_1}{D_2}\delta_1$$

as $D_{12} - D_1 > D_2$ for all values of k.

As we alter the δ's as stated above, we do not change $f_\delta(\widehat{y})$, after all, f_1 is still critical. Also, $t(\delta, k)$ cannot increase, as decreasing some δ_i for a fixed value of t can only increase the number of words in $S_t(w)$. If k increases in this process, then we were not at minimum to start with, as $t(\delta, k)$ should have been smaller. Also, the factor $1/(k-1)$ in the bound assures that a larger value of k, nothing else changing, would only give a sharper bound.

We assumed that it was f_1 that was critical, but to finish the proof, one must perform similar calculations for starting with the other functions as critical as well, but calculations are so similar that we avoid showing it here. ∎

It was shown in [6] that the metric φ resulted in $f(y_1) \leq f(y_{12}) = f(y_3) = f(y_4) \geq f(y_2)$. But that is just the metric we want! We may conclude that no inner metric can give a tighter bound than $\varphi(h)$ does. ∎

IV. Conclusion

The reasoning for finding this bound may leave slack in several places.

- The localization of the problem to a sphere may cause some slack.
- The approximation of minimum distance between code words in a sphere by the average distance between code words in the sphere may cause slack.
- We assume that the weight of the matrix M is kt, i. e. that all code words in the sphere lie in the outermost layer in the sphere. If this is impossible, then this may lead to slack.

It is still an open question if these origins of slack can be "handled" and whether or not our bound can be improved further.

We conclude this section by giving some examples of known codes which are optimal in the sense that they fulfil the bound presented in this paper with equality. We choose to consider the class of multilevel codes, but there may be other code constructions providing optimal codes as well. Multilevel codes were introduced by Imai and Hirakawa [11] and have later been discussed by other authors. Let B_0, B_1 and B_2 be binary block codes with blocklength n and let b_i be a code word in B_i. Then the set of n-tuples of the form $c = b_0 + 2b_1 + 4b_2$ is a multilevel code over $\mathbf{Z}_8$ with component codes B_0, B_1 and B_2. Let $B2$ be a binary single parity code and let $B4$ be a binary extended Hamming code. Using $B2$, $B4$ and $\mathbf{Z}_2^n$ as component codes we obtain optimal codes over $\mathbf{Z}_8$ with code parameters given in Table I. Most of the codes in Table I can be expurgated (code words removed) to some level without changing the bound.

TABLE I
Optimal multilevel codes

$d^2_{E_{\min}}(C)$	n	Component codes
1.1716	$2 - \infty$	$B_0 = B2$, $B_1 = B_2 = \mathbf{Z}_2^n$
2	4, 8, 16, ..., 2^∞	$B_0 = B4$, $B_1 = B_2 = \mathbf{Z}_2^n$
2.3431	15, 16, $22 - \infty$	$B_0 = B4$, $B_1 = B2$, $B_2 = \mathbf{Z}_2^n$
4	3, 4	No B_0 used, $B_1 = B2$, $B_2 = \mathbf{Z}_2^n$

References

[1] Nilsson M., Lennerstad H., An Upper Bound on the Minimum Euclidean Distance for Block Coded Phase Shift Keying, *IEEE Trans. Inform. Theory* 46(2): pp. 656-662, 2000.

[2] Berlekamp E. R., *Algebraic Coding Theory*, New York: McGraw Hill, 1968.

[3] Golomb S. W., Welch L. R., Perfect codes in the Lee Metric and the Packing of Polynominoes, *SIAM Jou. of Appl. Math.*, vol. 18, no. 2, Januari, 1970.

[4] MacWilliams F. J., Sloane N. J. A., *The Theory of Error-Correcting Codes*, Amsterdam, The Netherlands: North-Holland, 1977.

[5] Graham R., Knuth D., Patashnik O., *Concrete Mathematics*, Addison Wesley, ISBN 0-201-14236-8, 1994.

[6] Laksman E., Lennerstad H., Nilsson M., *Improving bounds on the minimal Euclidean distance for block codes by inner metric optimization,* Combinatorics 2008, Verona, Italy, 2008.

[7] Nilsson M., Lennerstad H., Improved Upper Bound on the Minimum Euclidean Distance for Block Coded Phase Shift Keying, in proceedings of RVK05, Linköping, Sweden, 2005.

[8] Nilsson M., Lennerstad H., Laksman E., A two-metric Approach to Improve Bounds on the Minimum Euclidean Distance for Block Codes, proceedings of RVK08, Växjö, Sweden, 2008.

[9] Piret Ph., Bounds for Codes over the unit circle, *IEEE Trans. Inform. Theory*, vol. IT-32, pp.760-767, Nov. 1986.

[10] Wyner A. D., Bounds on communication with polyphase coding, *Probl. Pered. Inform.*, vol. 23, no. 3, pp. 18-26, July-Sept. 1987.

[11] H. Imai and S. Hirakawa, "A new multilevel coding method using error-correcting codes," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 371-377, May 1977

Paper III

Inner distance measure bounds on the minimum Euclidean distance for symmetric PSK block codes

Efraim Laksman, Håkan Lennerstad, Magnus Nilsson

Inner Distance Measure Bounds on the Minimal Euclidean Distance for Symmetric PSK Block Codes

Efraim Laksman, Håkan Lennerstad, *Member, IEEE,* and Magnus Nilsson

***Abstract*—The minimum Euclidean distance is a fundamental quantity for block-coded PSK. In this paper improvements are made of bounds for this quantity that are explicit functions of the alphabet size q, block length n and code size $|C|$. Earlier work, where the restriction $q = 8$ was used, is continued by a generalisation allowing any q.**

The bound generalizes Elias critical sphere argument, which localizes the optimization problem to one neighbourhood, by use of so called inner distance measure for defining the shape of a sphere. Remark that codes which fulfill the bound with equality exist, and are best possible in terms of minimum Euclidean distance, for given parameters q, n and $|C|$.

***Index Terms*—PSK, block code, upper bound, Elias critical sphere, distance measure.**

I. Introduction

THE bound for minimum squared Euclidean distance for symmetric PSK block codes is improved in this paper. For error correction with respect to maximum likelihood, when using a channel with additive white Gaussian noise, the minimum squared Euclidean distance is a highly relevant measure of the efficiency of a code for fixed block length n, code size $|C|$ and alphabet size q. This work is a generalization of [1], where the restriction $q = 8$ was used, and follows it closely in the first sections.

On the set $\mathbf{Z}_q^n$ we take squared Euclidean distance to be defined as

$$d_E^2(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} d_E^2(x_j, y_j), \tag{1}$$

where $d^2(x_j, y_j)$ is defined as

$$d_E^2(x_j, y_j) = d_E^2(x_j - y_j, 0) = 4 \sin^2 \frac{(x_j - y_j)\pi}{q} \ . \tag{2}$$

Note that this distance measure is translation invariant, so that often the arguments can be written in such a way that one of them is zero. To simplify notation we will write $d(x) = d(x, 0)$ for any distance measure. Now a relevant model for the words are points in the group $(\mathbf{Z}_q^n, +)$, with squared Euclidean distance used for measuring distance, see for example [2], [3].

E. Laksman and H. Lennerstad are with the School of Engineering, Blekinge Institute of Technology, Sweden, Blekinge Tekniska Hgskola SE-371 79 Karlskrona e-mail: (see http://www.bth.se/adressbok).

M. Nilsson is with the School of Communication, Blekinge Institute of Technology, Sweden, Blekinge Tekniska Hgskola SE-371 79 Karlskrona e-mail: (see http://www.bth.se/adressbok).

Manuscript received Month XX, 20XX; revised Month XX, 20XX. This paper was presented in part at the conference Fq9, Dublin 2009

We consider an arbitrary subset C of $\mathbf{Z}_q^n$, corresponding to a block code having $|C|$ codewords $\mathbf{x} = (x_1, \ldots, x_n)$ of length n in an alphabet of q letters. The minimum squared Euclidean distance for the code is then

$$d_{E\,\min}^2(C) = \min_{\substack{\mathbf{x}, \mathbf{y} \in C \\ \mathbf{x} \neq \mathbf{y}}} d_E^2(\mathbf{x}, \mathbf{y}) \ . \tag{3}$$

Bounds on the minimum Euclidean distance are fundamental for the geometry of $\mathbf{Z}_q^n$: which is the largest possible distance between the two closest members in a subset of $\mathbf{Z}_q^n$ with $|C|$ members?

As is well known, the minimum Euclidean distance is essential for the error correction capabilities for a code. We define the rate of a block code as

$$R(q, n, |C|) = \frac{\log_q |C|}{n} \ . \tag{4}$$

For several combinations of q, n, and $|C|$, mostly at high rates, there are known codes whose minimum squared Euclidean distances fulfill our bound with equality. For these combinations of q, n and $|C|$ neither the codes nor the bound can be improved in terms of minimum Euclidean distance.

For other combinations of q, n and $|C|$ there is a gap between the bound and the minimum squared Euclidean distances for the best known codes. The size of this gap differs from case to case. Thus especially for medium and low rates it is unknown whether there exist better codes to discover or if it is possible to improve the bound, or both.

Many of the best known block codes with respect to minimum squared Euclidean distance, are constructed as multilevel codes, see for example [4], [5] and [6]. There are also other code constructions providing some of the best known block codes. E. g. some of the BCH codes from [7], where they are shown to be good with respect to the Lee-metric, meet the bound on Euclidean distance presented here.

The results of this paper are derived by using different kinds of distance measures and metrics, so we next define these concepts. Both a distance measure and a metric are a function $d(\mathbf{x}, \mathbf{y})$ from pairs of codewords to nonnegative numbers with the symmetry property $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}$ and $\mathbf{y}$, and $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$. Unlike a distance measure, a metric is also required to satisfy the triangle inequality: $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{y}, \mathbf{z})$ for all $\mathbf{x}$ and $\mathbf{y}$. Note that the *squared* Euclidean distance measure $d_E^2(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} d_E^2(x_j - y_j)$ is not a metric in general. E. g.

for $q = 8$ and $n = 1$ we have $2 = d_E^2(2) = d_E^2((0),(2)) > d_E^2((0),(1)) + d_E^2((2),(1)) = 2d_E^2(1) = 2(2-\sqrt{2})$.

The quantities $d_E^2(x_j, y_j) = 4\sin^2\frac{(x_j-y_j)\pi}{q}$ are Euclidean distances between points when the entries $0,\ldots,q-1$ are distributed equidistantly on a unit circle. The generalized distance measures considered in this paper will be translation invariant and defined on $\mathbf{Z}_q^n$ so they will be defined by a sequence of nonnegative numbers, $\delta(0), \delta(1), \ldots, \delta(q-1)$, without any particular geometrical meaning. The distance is then

$$\delta(\mathbf{x},\mathbf{y}) = \sum_{j=1}^{n} \delta(x_j - y_j) \ . \tag{5}$$

Some of the numbers $\delta(i)$ may be infinite, prohibiting the corresponding differences. The Lee metric, for example, is represented by $\delta(i) = i$, and truncated Lee metrics, where $\delta(i) = i$ for $i \le r$ but $\delta(i) = \infty$ for $i > r$, have been considered [8].

An alternative notation is sometimes useful. For two codewords $\mathbf{x}$ and $\mathbf{y}$, the number of positions where $\mathbf{x}$ and $\mathbf{y}$ differ by i or by $q-i$ is denoted by $c_i(\mathbf{x},\mathbf{y})$:

$$c_i(\mathbf{x},\mathbf{y}) = |\{j \in [1,n] : (x_j - y_j) \equiv_q i \text{ or } (x_j - y_j) \equiv_q q-i\}|, \tag{6}$$

where $\equiv_q$ means equal with respect to modulo q. We are still working with a generalization of the closest distance of letters in a unit circle, so two words can in one position differ by at most $\lfloor q/2 \rfloor$. Then an alternative notation for $\delta(\mathbf{x},\mathbf{y})$ is

$$\delta(\mathbf{x},\mathbf{y}) = \sum_{j=1}^{n} \delta(x_j - y_j) = \sum_{i=1}^{\lfloor q/2 \rfloor} \delta(i) c_i(\mathbf{x},\mathbf{y}) \ . \tag{7}$$

II. Previous work

The bounds in this paper and in the previous papers in this line of research are partly based on the arguments leading to the well-known Elias bound (see also, [2] pp. 318-321, and [3] pp. 558-564). Elias bound arguments have been used by Piret [9], who calculates bounds for the maximum rate, $n^{-1}\ln|C|$, for codes C with given $d_{E\min}^2(C)/n$ as $n \to \infty$. Piret's upper bound on the rate becomes

$$\ln q - \max_{\substack{2\beta S\beta^T = d_{E\min}^2(C)/n \\ \sum_i \beta_i = 1}} (H(\beta)), \tag{8}$$

where H is the entropy function

$$H(\beta) = -\sum_{i=1}^{q} \beta_i \ln(\beta_i), \tag{9}$$

β is a vector of length q and S is the $q \times q$ matrix with elements $2\sin^2[(i-j)\pi/q]$ in position (i,j).

The maximum rate as $n \to \infty$ is a non-increasing function of $d_{E\min}^2(C)/n$. Thus we can get bound on $d_{E\min}^2/n$ as $n \to \infty$ as a function of the rate by reflecting the graph of Piret's bound in the line $\frac{\ln|C|}{d_{E\min}^2}$.

Wyner [10] has produced another bound for the same quantity as Piret. It is independent of q, and the q points may be distributed arbitrarily, giving for larger q weaker restrictions and tighter bound in general. Wyner's bound is

$$\lim_{n\to\infty} \frac{1}{n} \frac{nK\left(\frac{d_E^2}{2n}\right)(2\pi)^n}{V_n\left(\sqrt{2n\left(1-\sqrt{1-\frac{d_E^2}{2n}}\right)-1}\right)}, \tag{10}$$

where $V_n(r)$ is defined as the volume of a sphere with radius r in the n-dimensional torus with the Euclidean distance 2π in each dimension. Just as with Piret's bound, it may be taken as a bound on $d_{E\min}^2(C)/n$ for given rate as $n \to \infty$.

III. Problem formulation

The bounds in the present research are explicit in the parameters q, n and $|C|$. We next start the argument in the present paper and simultaneously present results of previous papers [1], [8], [11] and [12].

Generalizing the argument of the Elias bound, we next define a neighborhood $S_{\delta,t}(\mathbf{z})$ for a word $\mathbf{z}$ as

$$S_{\delta,t}(\mathbf{z}) = \{\mathbf{y} : \sum_{i=1}^{\lfloor q/2 \rfloor} \delta(i) c_i(\mathbf{z},\mathbf{y}) \le t\} \tag{11}$$

where δ is a distance measure. The set $S_{\delta,t}(\mathbf{z})$ is in the literature sometimes called a sphere with radius t. Observe that the number of words in a sphere is independent on the word that lies at its center.

We continue the Elias' argument of a critical sphere. If we label all words by their membership in a neighborhood $S_{\delta,t}(\mathbf{z})$, for each $\mathbf{z} \in C$, we distribute in total $|C||S_{\delta,t}|$ labels. Assume that t is large enough so that $|C||S_{\delta,t}|q^{-n} > 1$, and define $K = \lceil |C||S_{\delta,t}|q^{-n} \rceil$.

By the definition $K = \lceil |C||S_{\delta,t}|q^{-n} \rceil$, and by the pigeon hole principle, it follows that there is a word $\mathbf{y}^*$ so that $\mathbf{y}^* \in S_{\delta,t}(\mathbf{z})$ for at least K codewords $\mathbf{z} \in C$. (We may assume that exactly K codewords lie in this sphere, as the bound based on this method is a non-increasing function of K.) Then these K codewords belong to the neighborhood $S_{\delta,t}(\mathbf{y}^*)$. By subtracting $\mathbf{y}^*$ from all codewords we do not change any distances between codewords, so we may as well assume that $\mathbf{y}^* = \mathbf{0}$. Now let $W = S_{\delta,t}(\mathbf{y}^*) \cap C$, so $|W| = K$. We trivially have

$$d_{E\min}^2(C) \le d_{E\min}^2(W). \tag{12}$$

The problem is thus localized from a minimum distance bound of C to a minimum distance bound of W. Often this inequality results only in a small or very small slack since the density of codewords in W is $\lceil |C||S_{\delta,t}|q^{-n} \rceil / |S_{\delta,t}|$ – slightly higher than the density of C in $\mathbf{Z}_q^n$.

Elias [2] pp. 318-321, and Nilsson and Lennerstad [8], [11] have found upper bounds on $d_{E\min}^2(W)$ by bounding the average distance between the words in W by the mean distance $d_{E\text{mean}}^2(W)$.

In order to bound the average distance between codewords in W, those codewords are written as rows in a K by n matrix. Then the ratio between the distance between symbols and total weight of the symbols in the columns are found. Call columns

with the highest such ratio extremal columns. Then the average distance between codewords in W will be bounded by total weight of the matrix times the worst possible ratio, divided by the number of distinct pairs of codewords in W.

In this manner Elias found the bound

$$\frac{K^2 x(2-x)\bar{D}n}{K(K-1)}, \tag{13}$$

where $\bar{D}$ is the average distance between letters $\sum_{j=0}^{q-1} d(j)$, which when d is d_E^2 results in $\bar{D}=2$ and t is the radius in the spheres. One has to set $x = t/\bar{D}n$. This result is attained when using $\delta = d_E^2$.

In [11], $\delta(1) = 1$ and $\delta(i) = \infty$ for $i > 1$, was used, allowing at most t non-zeroes in a sphere, giving the bound

$$d^2_{E\min}(C) \leq \frac{t}{K-1} d_E^2(2) + 2(t - \frac{t}{K-1}) d_E^2(1) \ . \tag{14}$$

This bound is applicable for $|C| > (q/3)^n$ only, so it cannot be used for low rates, but it is tight in many cases for high rates. The tightness happens when all pairs of codewords in a neighborhood are at the same distance, in which case $d^2_{E\min}(W) \leq d^2_{E\text{mean}}(W)$ has zero slack – it is an equality.

The Lee metric is the metric $\delta(i) = i$ for all i. When restricted to the group $(\mathbf{Z}_q^n, +)$ it becomes

$$\delta(\mathbf{x}, \mathbf{y}) = \sum_{i=0}^{\lfloor q/2 \rfloor} i c_i(\mathbf{x}, \mathbf{y}) \ . \tag{15}$$

In [8], for $q = 8$ a two parameter (t, r)-Lee metric $\delta(i) = i$ if $i \leq r$ and $\delta(i) = \infty$ if $i > r$ was tried for $q = 8$. It was shown that $r = 2$ improves the small deviation neighborhood for medium rates, while $r = 4$ is preferable for low rates.

The idea of considering a general inner distance measure δ and designing it to optimize the bound for the outer distance measure, which is squared Euclidean, was first presented in [12]. Here a K-dependent inner distance measure was presented, for $q = 8$, as well as columns that appeared to be extremal by sampling the space of all possible distance measures. Compared to that paper, in [1] an improved K-dependent distance measure was presented and a partial optimality was proven, but still only for $q = 8$.

IV. Results

Continuing the argument of the previous section, we have a sphere $S_{\delta,t}(\mathbf{w})$ containing at least $K = \lceil |C||S_{\delta,t}|q^{-n} \rceil$ codewords, $W = S_{\delta,t}(\mathbf{w}) \cap C$, and we assume $\mathbf{w} = \mathbf{0}$.

The following theorem is from [1].

Theorem 1: For any code C in $\mathbf{Z}_q^n$ we have the bound

$$d^2_{E\min}(C) \leq \min_{K \in [2,|C|]} \min_{\delta} \frac{2\widetilde{t}_K f_\delta(\widehat{\mathbf{y}})}{K-1}, \tag{16}$$

where

$$\widetilde{t}_K(\delta) = \min(\{t : K \leq \lceil |C||S_{\delta,t}|q^{-n} \rceil\}), \tag{17}$$

$$f_\delta(\mathbf{y}) = \frac{\sum_{j_1=1}^{K} \sum_{j_2=1}^{j_1-1} d_E^2(y_{j_1}, y_{j_2})}{\sum_{j=1}^{K} \delta(y_j)}, \tag{18}$$

and $\widehat{\mathbf{y}}$ is a vector maximizing $f_\delta(\mathbf{y})$.

Even though $\widetilde{t}_K$ is a function not only of δ, but also of n, $|C|$ and q, we usually omit those parameters as we assume that they are fixed. The same is true for the dependence $f_\delta(\widehat{\mathbf{y}})$ has on q. We also remark that the minimum over t always exist since the sphere $S_{\delta,t}$ is defined with an inclusive inequality.

In [1] it was also shown that the bound is independent of the scaling of δ.

Lemma 2: The bound in Theorem 1 is scale invariant in the distance measure δ, i.e. for any $s > 0$, let $\lambda(x, y) = s\delta(x, y)$ for every pair x, y. Then

$$\frac{2\widetilde{t}_K(\delta) f_\delta(\widehat{\mathbf{y}})}{K-1} = \frac{2\widetilde{t}_K(\lambda) f_\lambda(\widehat{\mathbf{y}})}{K-1} \tag{19}$$

holds.

Next in [1] extremal columns, i. e. vectors $\widehat{\mathbf{y}}$ such that $f_\delta(\widehat{\mathbf{y}})$ achieves its maximum, were found. This however, was done only in the case $q = 8$, and was given by the lemma below.

Lemma 3: Let $q = 8$. Then for any additive distance measure δ and for any K, one of the columns

$$\begin{array}{ll} \widehat{\mathbf{y}}_1 = (1, -1, 0, \ldots, 0) & \widehat{\mathbf{y}}_2 = (2, 0, \ldots, 0), \\ \widehat{\mathbf{y}}_3 = (3, 0, \ldots, 0), & \widehat{\mathbf{y}}_4 = (4, 0, \ldots, 0), \\ \widehat{\mathbf{y}}_5 = (1, -2, 0, \ldots, 0), & \widehat{\mathbf{y}}_6 = (2, -2, 0, \ldots, 0) \end{array} \tag{20}$$

provides a maximum for

$$f(\mathbf{y}) = \frac{\sum_{j=1}^{K} \sum_{i=1}^{j-1} d_E^2(y_i, y_j)}{\sum_{i=1}^{K} \delta(y_i)} \ . \tag{21}$$

Here, working with general q, We will instead write a theorem. In the proof of the theorem we will need the so called mediant addition and two trigonometric lemmas. Mediant addition is defined as

$$\frac{a_1}{b_1} \oplus \frac{a_2}{b_2} = \frac{a_1 + a_2}{b_1 + b_2}, \tag{22}$$

presented in [13]. The number $\frac{a_1+a_2}{b_1+b_2}$ is called the *mediant* of $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$. It is similarly defined for c ratios $\frac{a_1}{b_1}, \ldots, \frac{a_c}{b_c}$, and is a weighted mean value of the ratios as can be seen by the identity

$$\frac{a_1}{b_1} \oplus \ldots \oplus \frac{a_c}{b_c} = \frac{b_1}{b_1 + \ldots + b_c} \frac{a_1}{b_1} + \ldots + \frac{b_c}{b_1 + \ldots + b_c} \frac{a_c}{b_c} \ . \tag{23}$$

As a weighted mean, the weights are strictly between 0 and 1, and are determined by the denominators only. We thus have $\frac{a_1}{b_1} < \frac{a_1+a_2}{b_1+b_2} < \frac{a_2}{b_2}$ if $\frac{a_1}{b_1} < \frac{a_2}{b_2}$, and $\frac{a_1}{b_1} = \frac{a_1+a_2}{b_1+b_2} = \frac{a_2}{b_2}$ if $\frac{a_1}{b_1} = \frac{a_2}{b_2}$.

Now the trigonometric lemmas:

Lemma 4: We have $d_E(2i)d_E(2j) = |d_E^2(j+i) - d_E^2(j-i)|$.

Proof: We observe that

$$\begin{aligned} d_E(2i)d_E(2j) &= 4|\sin(\frac{2i\pi}{q})\sin(\frac{2j\pi}{q})| \\ &= 16|\sin(\frac{i\pi}{q})\cos(\frac{i\pi}{q})\sin(\frac{j\pi}{q})\cos(\frac{j\pi}{q})| \\ &= 4\Big|\left(\sin(\frac{i\pi}{q})\cos(\frac{j\pi}{q})+\cos(\frac{i\pi}{q})\sin(\frac{j\pi}{q})\right)^2 \\ &\quad -\left(\cos(\frac{i\pi}{q})\sin(\frac{j\pi}{q})-\sin(\frac{i\pi}{q})\cos(\frac{j\pi}{q})\right)^2\Big| \\ &= 4|\sin^2(\frac{(j+i)\pi}{q})-\sin^2(\frac{(j-i)\pi}{q})| \\ &= |d_E^2(j+i)-d_E^2(j-i)| \end{aligned} \tag{24}$$

and that we in particular have $d_E(2i)d_E(2j) = d_E^2(j+i) - d_E^2(j-i)$ when either i or j, taken modulo q, is between 0 and $q/2$, but not both, and $d_E(2i)d_E(2j) = -(d_E^2(j+i) - d_E^2(j-i))$ otherwise (which is easily seen by considering the sign of $\sin(i)\cos(i)\sin(j)\cos(j)$) as we only care about $1 \le i, j \le q-1$. ∎

Lemma 5: We have

$$0 = 4d_E^2(i) - d_E^4(i) - d_E^2(2i), \tag{25}$$

when $1 \le i \le q-1$,

$$4d_E^2(i) = 4d_E^2(j) + 4d_E^2(j+i), \tag{26}$$

when $1 \le i \le q-1$ and $j = 0$, and

$$\begin{aligned} 4d_E^2(i) = 4d_E^2(j) + 4d_E^2(j+i) - 2d_E^2(j)d_E^2(j+i) \\ - 2(d_E^2(2j+i) - d_E^2(i)) \end{aligned} \tag{27}$$

when $1 \le i \le (q/2)-1, 1 \le j \le q-1$.

Proof: This lemma can be proven by elementary trigonometric relations, in a way similar to the previous lemma. ∎

Theorem 6: The columns

$$\begin{array}{c} (1,-1,0,\ldots,0), \\ (2,-2,0,\ldots,0), \\ \vdots \\ (\lfloor \frac{q}{4} \rfloor, -\lfloor \frac{q}{4} \rfloor, 0, \ldots, 0), \\ (\lfloor \frac{q}{4} \rfloor + 1, 0, \ldots, 0), \\ \vdots \\ (\lfloor \frac{q}{2} \rfloor, 0, \ldots, 0) \end{array} \tag{28}$$

are extremal, and if a column $\mathbf{y}$ is an extremal column, then $\sum_{j=1}^K d_E^2(y_j) \le 4$.
Note that except for small q, other extremal columns than those explicitly mentioned in the theorem exists.

Proof: Maximization of the function $f_\delta(\mathbf{y})$ is done by a sequence of transformations of the variables. We first introduce the functions $a_i(\mathbf{y})$ that counts the number of occurences of i in the column $\mathbf{y}$. I. e. $a_0(\mathbf{y})$ is the number of zeros, $a_1(\mathbf{y})$ the number of 1:s, $a_{q-1}(\mathbf{y})$ the number of -1:s (as $-1 \equiv_q q-1$), and so on. By $a(\mathbf{y})$ we mean $(a_0(\mathbf{y}), \ldots, a_{q-1}(\mathbf{y}))$. Since the length of the column $\mathbf{y}$ is K, we know that $K = \sum_{i=0}^{q-1} a_i(\mathbf{y})$.

The function $f_\delta(\mathbf{y})$ can then be rewritten as follows:

$$\begin{aligned} f_\delta(\mathbf{y}) &= \frac{\sum_{i=1}^K \sum_{j=1}^{i-1} d_E^2(y_i, y_j)}{\sum_{i=1}^K \delta(y_i)} = \\ &= \frac{\sum_{i=1}^{q/2-1} d_E^2(i) \sum_{j=0}^{q-1} a_j a_{j+i}}{\sum_{i=1}^{q-1} \delta(i) a_i} + \\ &\quad + \frac{d_E^2(q/2) \sum_{j=0}^{q/2-1} a_j a_{j+q/2}}{\sum_{i=1}^{q-1} \delta(i) a_i} \end{aligned} \tag{29}$$

if q is even, and

$$f_\delta(\mathbf{y}) = \frac{\sum_{i=1}^{(q-1)/2} d_E^2(i) \sum_{j=0}^{q-1} a_j a_{j+i}}{\sum_{i=1}^{q-1} \delta(i) a_i} \tag{30}$$

if q is odd.

The main objective is to maximize f_δ with respect to $a_0, \ldots, a_{q-1}$. Note that while there are q^K different columns $(y_1, \ldots, y_K)$, there are only $\binom{q+K-1}{K}$ different vectors $(a_0, \ldots, a_{q-1})$. This is the number of selections of K objects out of q alternatives with repetition but without order, since by going from $(y_1, \ldots, y_K)$ to $(a_0(\mathbf{y}), \ldots, a_{q-1}(\mathbf{y}))$ we have removed order changes that are insignificant for the value of f_δ. It is independent of rearrangements such as $(y_1, y_2, \ldots, y_K) \to (y_2, y_1, \ldots, y_K)$.

We make a new transformation. Define the functions α_i according to

$$\begin{array}{lll} \alpha_0 & = a_0, & \\ \alpha_i & = a_i + a_{q-i} & ,\forall i \in \{1, \ldots, \lceil q/2 \rceil - 1\}, \\ \alpha_i & = a_i - a_{q-i} & ,\forall i \in \{\lfloor q/2 \rfloor + 1, \ldots, q-1\} \text{ and} \\ \alpha_{q/2} & = a_{q/2} & , \text{if } q \text{ is even.} \end{array} \tag{31}$$

By $\alpha(\mathbf{y})$ we mean $(\alpha_0(\mathbf{y}), \ldots, \alpha_{q-1}(\mathbf{y}))$.

It will be shown, for even q, that f_δ may be rewritten as

$$\begin{aligned} f_\delta = \frac{4K \sum_{i=1}^{\lfloor q/2 \rfloor} d_E^2(i)\alpha_i - \left(\sum_{i=1}^{\lfloor q/2 \rfloor} d_E^2(i)\alpha_i\right)^2}{\sum_{i=1}^{\lfloor q/2 \rfloor} \delta(i)\alpha_i} \\ - \frac{\left(\sum_{i=\lfloor q/2 \rfloor + 1}^{q-1} d_E(2i)\alpha_i\right)^2}{\sum_{i=1}^{\lfloor q/2 \rfloor} \delta(i)\alpha_i} . \end{aligned} \tag{32}$$

The same is true for odd q, but we refrain from showing it as it is done in the same way.

The denominators in the right-hand sides of equations 29 and 32 are obviously the same, so we only have to show that the numerators as well are equal. The numerator of the right-hand side of equation 32 can be written as $X_1 - X_2 - X_3$, where

$$X_1 = 4\left(\sum_{i=0}^{q-1} a_i\right)\left(\sum_{j=1}^{q-1} d_E^2(j) a_j\right), \tag{33}$$

$$X_2 = \left(\sum_{i=1}^{q-1} d_E^2(i) a_i\right)^2, \tag{34}$$

$$X_3 = \left(\sum_{i=(q/2)+1}^{q-1} d_E(2i)(a_i - a_{q-i})\right)^2 . \tag{35}$$

Now, according to Lemma 4 we may rewrite

$$X_3 = \sum_{i=1}^{q-1}\sum_{j=1}^{q-1}(d_E^2(j+i) - d_E^2(j-i))a_i a_j \ . \quad (36)$$

We rewrite X_1, X_2 and X_3 by writing out squares and summing up terms in a different order:

$$\begin{aligned} X_1 = {} & 4\sum_{j=1}^{q-1} d_E^2(j)a_j^2 + \\ & + 4\Big(\sum_{i=1}^{(q/2)-1}\sum_{j=0}^{q-1}(d_E^2(j+i) + d_E^2(j))a_j a_{j+i} + \\ & + \sum_{j=0}^{(q/2)-1}(d_E^2(j+(q/2)) + d_E^2(j))a_j a_{j+(q/2)}\Big) \quad (37) \\ X_2 = {} & \sum_{j=1}^{q-1} d_E^4(j)a_j^2 + \\ & + 2\Big(\sum_{i=1}^{(q/2)-1}\sum_{j=1}^{q-1} d_E^2(j+i)d_E^2(j)a_{j+i}a_j + \\ & + \sum_{j=1}^{(q/2)-1} d_E^2(j)d_E^2(j+(q/2))a_j a_{j+(q/2)}\Big) \quad (38) \\ X_3 = {} & \sum_{j=1}^{q-1} d_E^2(2j)a_j^2 + \\ & + 2\Big(\sum_{i=1}^{(q/2)-1}\sum_{j=1}^{q-1}(d_E^2(2j+i) - d_E^2(i))a_{j+i}a_j + \\ & + \sum_{j=1}^{(q/2)-1}(d_E^2(2j+(q/2)) - d_E^2(q/2))a_j a_{j+(q/2)}\Big) \ . \quad (39) \end{aligned}$$

Identifying coefficients, by use of Lemma 5, we find that the numerators on the right-hand sides of equations 29 and 32 are the same, so that f_δ is correctly expressed in equation 32.

Let

$$P(\mathbf{y}) = \frac{4K\sum_{i=1}^{\lfloor q/2\rfloor} d_E^2(i)\alpha_i}{4\sum_{i=1}^{\lfloor q/2\rfloor}\delta(i)\alpha_i}, \quad (40)$$

$$N_1(\mathbf{y}) = \frac{\left(\sum_{i=1}^{\lfloor q/2\rfloor} d_E^2(i)\alpha_i\right)^2}{4\sum_{i=1}^{\lfloor q/2\rfloor}\delta(i)\alpha_i}, \quad (41)$$

$$N_2(\mathbf{y}) = \frac{\left(\sum_{i=\lfloor q/2\rfloor+1}^{q-1} d_E(2i)\alpha_i\right)^2}{4\sum_{i=1}^{\lfloor q/2\rfloor}\delta(i)\alpha_i}, \quad (42)$$

so $f_\delta = P - N_1 - N_2$.

The form of f_δ given in equation 32 is highly useful. E. g. considering inner distance measures of the form $\delta(0) = 0$, $\delta(i) = 1$ and $\delta(j) = \infty$ for all $i \neq j$, one can quickly find that each of the columns $(1, -1, 0, \ldots, 0)$, $(2, -2, 0, \ldots, 0)$, $\ldots$, $(\lfloor \frac{q}{4}\rfloor, -\lfloor \frac{q}{4}\rfloor, 0, \ldots, 0)$, $(\lfloor \frac{q}{4}\rfloor + 1, 0, \ldots, 0)$, $\ldots (\lfloor \frac{q}{2}\rfloor, 0, \ldots, 0)$ are extremal. One sees this by observing that to maximize f_δ, it is necessary to have $\alpha_j = 0$ for j other than 0, i or $q - i$, and that P is constant with respect to α_i, N_1 is growing with respect to α_i and N_2 will be 0 when α_i is even and constant when α_i is odd (since we try to maximize f_δ), which immediately results in α_i being either 1 or 2, whereafter it is just a matter of comparison of the two trigonometric expressions

$$\left(4\sin^2\left(\frac{i\pi}{q}\right)\right)^2 + \left(2\left|\sin\left(\frac{2i\pi}{q}\right)\right|\right)^2 \text{ and } \frac{\left(8\sin^2\left(\frac{i\pi}{q}\right)\right)^2}{2} \ . \quad (43)$$

Consider three columns, $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{z}$ such that the relationship $\alpha(\mathbf{y}_1) + \alpha(\mathbf{y}_2) = \alpha(\mathbf{z})$ holds. Based on the viewing mediant addition as a weighted average, if $f_\delta(\mathbf{y}_1) \oplus f_\delta(\mathbf{y}_2) > f_\delta(\mathbf{z})$, then $\mathbf{z}$ cannot be an extremal column. Of course a larger set of columns can be used in order to be able to discard one column from possibly being extremal, since if $\alpha(\mathbf{y}_1) + \ldots + \alpha(\mathbf{y}_m) = \alpha(\mathbf{z})$ and $f_\delta(\mathbf{y}_1) \oplus \ldots \oplus f_\delta(\mathbf{y}_m) > f_\delta(\mathbf{z})$, then $\mathbf{z}$ cannot be extremal. This strategy for comparisons may be improved further by observing that if $\alpha(\mathbf{y}_1) + \ldots + \alpha(\mathbf{y}_m) = v\alpha(\mathbf{z})$ for some integer v and $f_\delta(\mathbf{y}_1) \oplus \ldots \oplus f_\delta(\mathbf{y}_m) > f_\delta(\mathbf{z})$ (extend the right-hand side by v for ease of comparison), then $\mathbf{z}$ cannot be extremal.

When eliminating non-extremal columns we may use the extremal columns found a few paragraphs earlier to eliminate columns that cannot be extremal for any inner distance measure, as to leave us with only a small set of columns to check against eachother.

We can express the last part of the theorem we are proving as if a column $\mathbf{y}$ is an extremal column, then $\sum_{i=1}^{q-1} d_E^2(i)a_i(\mathbf{y}) \leq 4$.

Assume that we have a fixed inner distance measure δ. Select m such that we have $d_E^2(m)/\delta(m) \geq d_E^2(i)/\delta(i)$ for all $i \neq 0$. Observe that

$$\frac{d_E^2(m)}{\delta(m)} \geq \frac{\sum_{i=1}^{q-1} d_E^2(i)c_i}{\delta(i)c_i} \quad (44)$$

for any non-negative constants c_i, at least one of which is positive. Now assume that there is an extremal column $\mathbf{x}$ such that $\sum_{i=1}^{q-1} d_E^2(i)a_i(\mathbf{x}) > 4$.

We get two cases:

$m \leq \lfloor q/4\rfloor$: Form the column $\mathbf{y}$ with $a_m(\mathbf{y}) = 1$, $a_{q-m}(\mathbf{y}) = 1$, $a_0(\mathbf{y}) = K - 2$ and $a_i(\mathbf{y}) = 0$ for $i \notin \{-m, 0, m\}$.

We observe that $N_2(\mathbf{y}) = 0$ and $N_2(\mathbf{x}) \geq 0$. Thus, it suffice to show that $P(\mathbf{y}) - N_1(\mathbf{y}) > P(\mathbf{x}) - N_1(\mathbf{x})$ to show that $\mathbf{x}$ cannot be an extremal column. But we have

$$P(\mathbf{y}) - N_1(\mathbf{y}) = \frac{2d_E^2(m)\left(4K - 2d_E^2(m)\right)}{8\delta(m)} \quad (45)$$

and

$$P(\mathbf{x}) - N_1(\mathbf{x}) = \frac{\sum_{i=1}^{q-1} d_E^2(i)a_i(\mathbf{x})\left(4K - \sum_{i=1}^{q-1} d_E^2(i)a_i(\mathbf{x})\right)}{4\sum_{i=1}^{q-1}\delta(i)a_i(\mathbf{x})} \ . \quad (46)$$

Since

$$\frac{2d_E^2(m)}{2\delta(m)} \geq \frac{\sum_{i=1}^{q-1} d_E^2(i)a_i(\mathbf{x})}{\sum_{i=1}^{q-1}\delta(i)a_i(x)} \text{ and} \tag{47}$$

$$4K - 2d_E^2(m) \geq 4K - 4 > 4K - \sum_{i=1}^{q-1} d_E^2(i)a_i(\mathbf{x}), \tag{48}$$

it follows that $P(\mathbf{y}) - N_1(\mathbf{y}) > P(\mathbf{x}) - N_1(\mathbf{x})$ and $f_\delta(\mathbf{y}) > f_\delta(\mathbf{x})$, so $\mathbf{x}$ is not an extremal column.

$m > \lfloor q/4 \rfloor$: Form the column $\mathbf{y}$ with $a_m(\mathbf{y}) = 1$, $a_0(\mathbf{y}) = K - 1$ and $a_i(\mathbf{y}) = 0$ for $i \notin \{0, m\}$. Just as in the previous case, $N_2(\mathbf{x}) \geq 0$, so it suffices to show $P(\mathbf{y}) - N_1(\mathbf{y}) - N_2(\mathbf{y}) > P(\mathbf{x}) - N_1(\mathbf{x})$ to get that $\mathbf{x}$ cannot be an extremal column. But we have

$$\begin{aligned} P(\mathbf{y}) - N_1(\mathbf{y}) - N_2(\mathbf{y}) &= \\ &= \frac{4Kd_E^2(m) - d_E^4(m) - d_E^2(2m)}{4\delta(m)} = \\ &= \frac{(4K-4)d_E^2(m)}{4\delta(m)} \end{aligned} \tag{49}$$

and $P(\mathbf{x}) - N_1(\mathbf{x})$ has not changed since the previous case, so we get $f_\delta(\mathbf{y}) > f_\delta(\mathbf{x})$ and $\mathbf{y}$ cannot be an extremal column.

We have now proven the theorem. ∎

This theorem reduces the set of possibly extremal columns greatly, and further reduction can be made by use of the method described in the proof.

V. An example

We will show how to find an inner distance measure δ which minimizes $\max_{\mathbf{y}} f_\delta(\mathbf{y})$ in the case $q = 9$.

We get $d^2(1) \approx 0.4679$, $d^2(2) \approx 1.6527$, $d^2(3) = 3$ and $d^2(4) \approx 3.8794$. Using Theorem 6, there are only 19 columns which must be checked for extremality, only six of which survives further elimination by use of mediant addition. The extremal columns are $(1, -1, 0, 0, \ldots)$, $(1, -2, 0, 0, \ldots)$, $(2, -2, 0, \ldots)$, $(3, 0, 0, \ldots)$, $(1, -3, 0, 0, \ldots)$ and $(4, 0, 0, \ldots)$, taking no respect to permutations of elements or multiplication by -1. We'll call them $\widehat{\mathbf{y}}_{11}$, $\widehat{\mathbf{y}}_{12}$, $\widehat{\mathbf{y}}_2$, $\widehat{\mathbf{y}}_{13}$, $\widehat{\mathbf{y}}_3$ and $\widehat{\mathbf{y}}_4$, respectively.

The column $\mathbf{x} = (1, 1, -2, 0, \ldots, 0)$ for example, is not extremal, as

$$f_\delta(\mathbf{x}) \oplus f_\delta(\mathbf{x}) < f_\delta(\widehat{\mathbf{y}}_{11}) \oplus f_\delta(\widehat{\mathbf{y}}_{12}) \oplus f_\delta(\widehat{\mathbf{y}}_{12}) \ . \tag{50}$$

Note that the denominators and the K-dependent parts of the left-hand side and the right-hand side are equal, so only constants have to be compared.

We thus get

$$\begin{aligned} f_\delta(\widehat{\mathbf{y}}_{11}) &= \tfrac{0.4679K - 0.1095}{\delta(1)}, \\ f_\delta(\widehat{\mathbf{y}}_{12}) &= \tfrac{2.1206K - 1.2412}{\delta(1)+\delta(2)}, \\ f_\delta(\widehat{\mathbf{y}}_2) &= \tfrac{1.6527K - 1.3657}{\delta(2)}, \\ f_\delta(\widehat{\mathbf{y}}_{13}) &= \tfrac{3.4679K - 3.0564}{\delta(1)+\delta(3)}, \\ f_\delta(\widehat{\mathbf{y}}_3) &= \tfrac{3K-3}{\delta(3)}, \\ f_\delta(\widehat{\mathbf{y}}_4) &= \tfrac{3.8794K - 3.8794}{\delta(4)} \ . \end{aligned} \tag{51}$$

Setting B as an upper bound on f_δ, we get the following inequalities:

$$\frac{0.4679K - 0.1095}{\delta(1)} \leq B, \tag{52}$$

$$\frac{2.1206K - 1.2412}{\delta(1) + \delta(2)} \leq B, \tag{53}$$

$$\frac{1.6527K - 1.3657}{\delta(2)} \leq B, \tag{54}$$

$$\frac{3.4679K - 3.0564}{\delta(1) + \delta(3)} \leq B, \tag{55}$$

$$\frac{3K - 3}{\delta(3)} \leq B, \tag{56}$$

$$\frac{3.8794K - 3.8794}{\delta(4)} \leq B, \tag{57}$$

which we must fulfill for as small B as possible. The inequalities 53, 56 and 57 combined with normalization, $\delta(1) + \delta(2) + \delta(3) + \delta(4) = 1$, gives us $B \geq 9K - 8.1206$. This is the greatest lower bound on B we can draw from our inequalities. We may thus set $B = 9K - 8.1206$, but this is possible only if the second, the fifth and the sixth inequalities are all equalities. Thus we get

$$\begin{aligned} \delta(1) &= \tfrac{0.4679K - 0.1095 + h}{9K - 8.1206}, \\ \delta(2) &= \tfrac{1.6527K - 1.2412 + 0.1095 - h}{9K - 8.1206} \\ \delta(3) &= \tfrac{3(K-1)}{9K - 8.1206}, \\ \delta(4) &= \tfrac{3.8794(K-1)}{9K - 8.1206}, \end{aligned} \tag{58}$$

where $h \in [0.0531, 0.2340]$.

VI. Conclusion

What has been presented here is a method for how to find an upper bound on the minimum squared Euclidean distance for PSK-codes for any q, n and $|C|$. The method is a generalization with respect to q of how bounds for $q = 8$ were formed in [1]. In this paper, just as in [1], an inner distance measure which minimizes $\max_{\mathbf{y}} f_\delta(\mathbf{y})$ defined in Theorem 1 was found. This is accomplished by showing that the same transformation which worked for $q = 8$ works for any q. Also, Theorem 6 gives a way of finding all extremal columns, which is more efficient than how it was done in [1].

In [14], it was shown that for $q = 8$, the inner distance measure which optimizes the bound itself, with respect to Theorem 1, can be found in the same one-dimensional space as the inner distance measure which in [1] was shown to minimize $\max_{\mathbf{y}} f_\delta(\mathbf{y})$. This may be taken as an indication that also for general q, the inner distance measures found by the method used here at least gets close to optimizing the bound with restriction to the Elias' method. Indeed, for high rates this bound is tight for several values of the parameters in the sense that there are known codes whose minimum squared Euclidean distance equals the bound. For medium and low rates it is still somewhat unclear how tight the bound is.

This line of research uses a novel method by considering a general inner distance measure and optimizing with respect to the bound in the outer distance measure. It may be seen as the start of an investigation of tight bounds for the minimum

Avstånd: 30,24 mm

distance of a subset C of $\mathbf{Z}_q^n$, measured in some specific distance measure, by the method of designing an inner distance measure to the outer distance measure. It concerns tight bounds that are explicit in the three parameters q, n and $|C|$, and are general in the sense that they assume no structure for the subset C.

REFERENCES

[1] Laksman E., Lennerstad H., Nilsson M., Improving bounds on the minimum Euclidean distance for block codes by inner metric optimization, *Combinatorics 2008*, Verona, Italy, 2008.

[2] Berlekamp E. R., *Algebraic Coding Theory*, New York: McGraw Hill, 1968.

[3] MacWilliams F. J., Sloane N. J. A., *The Theory of Error-Correcting Codes*, Amsterdam, The Netherlands: North-Holland, 1977.

[4] Imai, H., Hirakawa, S., A new multilevel coding method using error-correcting codes, *IEEE Transactions of Information Theory*, Volume 23, Issue 3, pp. 371-377, May 1997.

[5] Sayegh, S., A Class of Optimum Block Codes in Signal Space, *IEEE Transactions on Communication*, Volume 34, Issue 10, pp. 1043-1045, Oct 1986.

[6] Tanabe, H., Umeda, H., Salam, M. A., A new construction method of multilevel coded modulation with a good Euclidean minimum distance, 1997 *IEEE International Symposium on Information Theory*, pp. 437, 29 June – 4 July 1997.

[7] Roth, R., M., Siegel, P, H., Lee-Metric BCH Codes and their Application to Constrained and Partial-Response Channels, *IEEE Transactions on Information Theory*, Vol. 40, No. 4, pp: 1083-1096, July 1994.

[8] Nilsson M., Lennerstad H., Improved Upper Bound on the Minimum Euclidean Distance for Block Coded Phase Shift Keying, *Proceedings of RVK05*, Linköping, Sweden, 2005.

[9] Piret Ph., Bounds for Codes over the unit circle, *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 760-767, Nov. 1986.

[10] Wyner A. D., Bounds on communication with polyphase coding, *Bell. Syst. Tech. J.*, vol. XLV, pp. 523-559, Apr. 1966.

[11] Nilsson M., Lennerstad H., An Upper Bound on the Minimum Euclidean Distance for Block Coded Phase Shift Keying, *IEEE Trans. Inform. Theory* 46(2): pp. 656-662, 2000.

[12] Nilsson M., Lennerstad H., Laksman E., A two-metric Approach to Improve Bounds on the Minimum Euclidean Distance for Block Codes, *Proceedings of RVK08*, Växjö, Sweden, 2008.

[13] Graham R., Knuth D., Patashnik O., *Concrete Mathematics*, Addison Wesley, ISBN 0-201-14236-8, 1994.

[14] Laksman E., Lennerstad H., Nilsson M., *Bounding the minimal Euclidean distance for any PSK block codes of alphabet size 8*, *IEEE Information Theoy Workshop,* Taormina, Italy, 2009.

ABSTRACT

In wireless communication, the minimum Euclidean distance between codewords is a major factor for the ability to correct errors in messages, and it is of interest to maximize the minimum Euclidean distance.

The thesis improves previously established general upper bounds on minimum Euclidean distance of phase shift keying block codes. There are no requirements on structure of codes, as the bound depends only on alphabet size, word length and code size. Prior to this thesis, bounds found by use of a method of Elias, had been improved by generalization of Elias' method. The method used here is an attempt to optimize that generalization.

2010:02

ISSN 1650-2140
ISBN 978-91-7295-175-4