On squares of cyclic codes

Ignacio Cascudo*

March 6, 2017

Abstract

The square C^{*2} of a linear error correcting code C is the linear code spanned by the coordinate-wise products of every pair of (non-necessarily distinct) words in C. Squares of codes have gained attention for several applications mainly in the area of cryptography, where typically one is concerned about some of the parameters (dimension, minimum distance) of both C^{*2} and C. In this paper, squares of cyclic codes are considered. General results on the minimum distance of the squares of cyclic codes are obtained and constructions of cyclic codes C with relatively large dimension of C and minimum distance of the square C^{*2} are discussed. In some cases, the constructions lead to codes C such that both C and C^{*2} simultaneously have the largest possible minimum distances for their length and dimensions.

1 Introduction

The m-th power C^{*m} of a linear error correcting code C is the linear code spanned by the coordinate-wise products of every tuple of m (non-necessarily distinct) words in C. When $m = 2$, we speak about the square C^{*2} of C. Powers and especially squares of codes play a relevant role in several recent results in cryptography and in particular in the area of secure multiparty computation, some of which are mentioned below in more detail. In addition, the study of squares of codes is also useful for other applications such as the construction of bilinear multiplication algorithms in finite extensions of finite fields (through the notion of supercode introduced in [30]) or the cryptanalysis of public key encryption schemes based on error correcting codes (see [11] and the references therein). Moreover, the notion of a square of a code is a special case of that of a component-wise product of two codes, which has been studied in connection to the concept of error-correcting pair (also known as error locating pair)[21, 25] upon which some efficient error correction algorithms are based. As a consequence of these applications, properties of products and powers of codes have been analysed in recent years in works such as [5, 23, 24, 26, 27, 28]. More information about applications of squares of codes can be found in [3, 5, 13, 28].

Some of the applications referred to above benefit from using codes C such that, simultaneously, the minimum distance of C^{*2} and the dimension of C are both large in relation to the length of C. For this reason, the relationship between these two parameters has been studied in recent works. A Singleton-like bound relating $\dim C$ and $d(C^{*2})$ (where d denotes the minimal distance and dim denotes the dimension) was established in [27] and later the family of codes attaining this bound was characterized in [24] (both works treat in fact the more general setting of products of codes). In particular, unless one of the two parameters ($\dim C$ or $d(C^{*2})$) is very restricted, Reed-Solomon codes are the only ones which can match this bound (see Section 2.1 for more information about these results).

*Department of Mathematics, Aalborg University. Email:ignacio@math.aau.dk. The author was partially supported by the Danish Council for Independent Research, under grant no. DFF-4002-00367

However, Reed-Solomon codes have a well-known restriction, which is a shortcoming in the cryptographic applications mentioned below; namely, their lengths are upper-bounded by the size of the finite field over which they are defined.[1] Therefore there has been interest in analyzing squares of long codes defined over small finite fields, particularly codes over the binary field. In this vein, the asymptotic behaviour of families of squares of codes has been considered, where the finite field is fixed and the length of the codes in the family grows to infinity. The existence, over every finite field, of asymptotically good families[2] of codes whose squares also form an asymptotically good family was established in [26]. However, [5] showed that families of codes with such asymptotic properties are not very abundant, since choosing codes uniformly at random among all codes of a prescribed dimension will, with high probability, not satisfy the desired properties.

Instead of considering the asymptotic setting, one can focus on specific values for the length of the code C; for example, one can fix a small finite field and certain values for $d(C^{*2})$ and $\dim C$ and attempt to find codes C of as small length as possible for which these parameters can be achieved. This is the case in some of the applications detailed below. In this case, cyclic codes look like a good candidate family of codes to provide a good answer in many cases. Indeed in the usual error-correcting code theory scenario where one is concerned about $\dim C$ and $d(C)$ (rather than $d(C^{*2})$) there are many instances of concrete values for these parameters for which the shortest known codes (sometimes shortest possible) are cyclic, even if cyclic codes are not known to be asymptotically good. In other cases, these shortest codes are not cyclic but can be obtained from cyclic codes by known operations such as puncturing of shortening. The question is, of course, if all of this holds when the distance of the square $d(C^{*2})$ is considered instead of $d(C)$.

Squares of cyclic codes have however not been studied too much so far. In [23] the problem is studied briefly and the square of a cyclic code, which is again cyclic, is described by relating its generator polynomial to that of the original code, and $\dim C^{*2}$ (and in some cases also $d(C^{*2})$) is computed for all cyclic codes of certain specific lengths and dimensions $\dim C$. Moreover, it is suggested that squares of cyclic codes have smaller dimensions than those of random codes. Other related results appeared in [15], who studied error-locating pairs for cyclic codes. While the application considered there is different, some of their intermediate results will be useful in the setting considered here too.

1.1 Applications

Aside from the theoretical interest of the problem, there are two concrete applications in cryptography where the study of squares of cyclic codes is relevant. First, linear codes C such that $d(C^{*2}) \geq t+2$ and $d(C^{\perp}) \geq t+2$, where $C^{\perp}$ denotes the dual code of C, can be used to construct t-strongly multiplicative secret sharing schemes, a notion introduced in [12], and upon which secure multiparty computation protocols can be based. Secure multiparty computation deals with the situation where several distrustful parties want to cooperate in order to carry out computations on private information held by each of them, without each party needing to reveal this information to the others; see [13] for more information about this area. It was shown in [12] that a multiparty computation protocol with unconditional security against an adversary actively corrupting any t parties can be constructed from any

[1] The reason why this is a drawback, in a nutshell, is that codes are used in those works as a way of distributing information among a prescribed number of locations, by associating one coordinate of a codeword to each of them, which results in requiring long codes; but on the other hand, having those codes defined over large finite fields increases both the computational and the communication complexity of the corresponding protocols, in comparison to using, for example, the binary field.

[2] We say that a family of codes $\{C_i\}_{i\in\mathbb{N}}$ with lengths n_i is asymptotically good if $n_i \to \infty$ when $i \to \infty$ and the limits $\lim_{i\to\infty} \dim C_i/n_i$ and $\lim_{i\to\infty} d(C_i)/n_i$ exist and are strictly positive.

t-strongly multiplicative linear secret sharing scheme. This generalized seminal results on unconditionally secure multiparty computation in [1, 9]. The advantage of this generalization is that while [1, 9] are based on Shamir's secret sharing scheme (which in turn is constructed from Reed-Solomon codes), the result in [12] provided more flexibility in the choice of the secret sharing scheme, and in particular allows for reducing the communication complexity of the protocol at the cost of some loss in the corruption tolerance factor t when the number of players participating in the protocol is large. This is because one can replace Shamir's scheme by secret sharing schemes defined over small fields. Families of strongly multiplicative secret sharing schemes with asymptotically good properties, based on algebraic geometric codes, were studied later in [10, 4, 6] and the results of these papers have been exploited for significantly decreasing the asymptotical complexity of a number of cryptographic tools, see [13, Section 12.8] for references.
Second, cyclic codes played a central role in a construction of a cryptographic tool known as additively homomorphic universally composable secure commitment schemes [8]. This result requires binary codes C with certain fixed $\dim C$ and $d(C)$ and, for those values, the shortest known codes are BCH codes, which are a family of cyclic codes. The concrete parameters that are considered in [8], when comparing the performance of their construction with previous alternatives, are $\dim C = 256$ and $d(C) \geq 120$. However, the construction was further improved in [16, 7] and it was shown that the same level of security can be achieved with a modified construction that only needs half the minimum distance. They consider the cases $\dim C = 256$ and $d(C) \geq 40, 60, 80$ which achieve different levels of security. The complexity of the protocol depends on the length of the codes and it is advantageous for the construction that they are short. The constructions from [8, 16, 7] attained several efficiency advantages with respect to prior work [14, 17], but lack one of the useful properties from [14] regarding verifiable commitment multiplication proofs. As suggested in [18], one can recover this property by a small modification of the construction, but this requires replacing the requirement on $d(C)$ by the same one on $d(C^{*2})$. The question is then how much the length of the code (and consequently the complexity of the commitment protocol) needs to grow in order to accommodate this more stringent requirement.

1.2 Overview of the paper

The main goals of this paper are two: first, to give a description of the square C^{*2} of a cyclic code C that facilitates the task of finding tight lower bounds for $d(C^{*2})$. The second goal is to exploit this description in order to find families of cyclic codes with simultaneously "large" (with respect to its length) values of $\dim C$ and $d(C^{*2})$, with special focus on binary codes and on the range of parameters which is interesting for the applications in [8].
The first goal is addressed with Theorem 3.5, where it is shown that some observations from [15] lead to a description of the generator polynomial that seems to present some advantages with respect to the one given in [23]. In particular it gives a direct description of the generating set of C^{*2} in terms of the one for C, which allows for applying the BCH bound easily.
As for the second aim, it is important to understand first that choosing a cyclic code C with a large minimum distance $d(C)$ does not always imply $d(C^{*2})$ is also large, as explained in Section 3.1. Several ways of choosing the generating sets of C are suggested. The two first proposed constructions, described in Section 4, only yield Reed-Solomon codes and punctured Reed-Muller codes. A third approach is described in Section 5. It considers the case of codes of length $n = q^k - 1$ and is based on the notion of restricted weights, which is introduced also in that section. Bounds for $d(C^{*2})$ are given and the dimension of C is determined exactly by counting the number of walks of a given length in a certain graph. This construction is still parametrized by two integers, and in Section 6 certain concrete

values for these integers are fixed and explicit values for the lengths, dimensions and bounds for the minimum distances of binary codes and their squares are given.
It is seen that in some cases, the codes C obtained satisfy the following two simultaneous features: $d(C)$ is the largest minimum distance possible for a code of length n and dimension $\dim C$; and $d(C^{*2})$ is the largest possible for a code of length n and dimension $\dim C^{*2}$. In other cases, "largest possible minimum distance" is replaced by "largest minimum distance achieved by currently known codes (according to the code tables in [19, 29])", since for those sets of parameters currently known lower and upper bounds for the minimum distance of codes do not coincide.
As explained in Section 7, this is not enough to conclude whether these codes attain the largest possible value $d(C^{*2})$ for the corresponding values of $(n, \dim C)$. Nevertheless it seems to be a relevant step in this direction and it is left as an open question whether one can find better constructions in this metric or, on the contrary, to prove them optimal.

2 Preliminaries

2.1 Squares of codes

Throughout this work, q will be a power of a prime and $n > 0$ will be a positive integer. Let $\mathbb{F}_q$ be a finite field of q elements, and let C be a linear code over $\mathbb{F}_q$ of length n, i.e., a $\mathbb{F}_q$-linear subspace of $\mathbb{F}_q^n$. As usual in coding theory, the dimension of C is its dimension as a vector space over $\mathbb{F}_q$ and denoted as $\dim C$; and the minimum distance of C, denoted as $d(C)$, is the smallest Hamming weight of a nonzero word in C.
Moreover given $\mathbf{v}, \mathbf{w} \in \mathbb{F}_q^n$, $\mathbf{v} * \mathbf{w}$ will denote their component-wise product as vectors in $\mathbb{F}_q^n$.

Definition 2.1. *Given two linear codes C and D over $\mathbb{F}_q$, their product $C * D$ is the linear code spanned over $\mathbb{F}_q$ by the set $\{\mathbf{c} * \mathbf{d} : \mathbf{c} \in C, \mathbf{d} \in D\}$.*
The square of C is the linear code $C^{*2} = C * C$, i.e., the linear code spanned over $\mathbb{F}_q$ by the set $\{\mathbf{c} * \mathbf{c}' : \mathbf{c}, \mathbf{c}' \in C\}$.

Similarly one can recursively define the m-th power of C, for $m \geq 2$ as $C^{*m} = C^{*(m-1)} * C$. The primary focus of this paper are however the squares C^{*2}. Some relatively straightforward relations between the dimensions and minimum distances of C and C^{*2} are given next. The proof and generalizations of these result for higher powers and products of different codes can be found in [28].

Proposition 2.2. *The dimension of C^{*2} satisfies $\dim C \leq \dim C^{*2} \leq \frac{(\dim C)\cdot(\dim C+1)}{2}$. The minimum distance of C^{*2} satisfies $d(C^{*2}) \leq d(C)$.*

The two propositions above indicate that lower bounds for $\dim C$ and $d(C^{*2})$ will also be lower bounds for $\dim C^{*2}$ and $d(C)$ respectively. Therefore it is interesting to examine how large $\dim C$ and $d(C^{*2})$ can be simultaneously. The following Singleton-like bound was shown in [27].

Proposition 2.3 ([27]). *It holds that*

$$d(C^{*2}) \leq \max\{1, n - 2\dim C + 2\}.$$

It was later shown in [24] that, unless either $\dim C$ or $d(C^{*2})$ is very small, the only codes that achieve the above bound are the Reed-Solomon codes. More precisely,

Proposition 2.4 ([24]). *Suppose that $d(C^{*2}) > 1$. If $d(C^{*2}) = n - 2\dim C + 2$, then C is either a Reed-Solomon code or a direct sum of self-dual codes, where self-duality is relative to a non-degenerate bilinear form which is not necessarily the standard inner product. Furthermore, if in addition $\dim C \geq 2$ and $d(C^{*2}) \geq 3$, then C is a Reed-Solomon code.*

Squares of Reed-Solomon codes are in fact again Reed-Solomon codes. Given integers $0 \leq m < n$, a finite field $\mathbb{F}$ of cardinality $|\mathbb{F}| \geq n$ and a vector $b = (b_1, b_2, \ldots, b_n) \in \mathbb{F}^n$ of evaluation points under the condition that $b_i \neq b_j$ if $i \neq j$, the Reed-Solomon code $RS_{\mathbb{F},b}(m,n)$ is defined as

$$RS_{\mathbb{F},b}(m,n) = \{(f(b_1), f(b_2), \ldots, f(b_n)) : f \in \mathbb{F}[X], \deg f \leq m\}$$

and it is a code of dimension $m+1$ and minimum distance $n-m$. We have that

$$(RS_{\mathbb{F},b}(m,n))^{*2} = RS_{\mathbb{F},b}(2m,n),$$

as long as $2m < n$. Otherwise $(RS_{\mathbb{F},b}(m,n))^{*2} = RS_{\mathbb{F},b}(n-1,n) = \mathbb{F}^n$.
Similar arguments can be used for other families of evaluation codes: concretely, consider Reed-Muller codes, which consist of evaluations of multivariate polynomials.

Definition 2.5. *A binary Reed-Muller code of length 2^k and order r (where $1 \leq r \leq k$), for short a $RM(r,k)$ code, is a linear code of the form*

$$C = \{(f(\mathbf{b}_1), \ldots, f(\mathbf{b}_{2^k}) : f \in \mathbb{F}_2[X_1, \ldots, X_k], \deg f \leq r\}$$

where $\mathbf{b}_1, \ldots, \mathbf{b}_{2^k}$ are all the distinct elements in $\mathbb{F}_2^k$, in some order; here deg *refers to the total degree of the k-variate polynomial f.*

It is well known that the distance of a $RM(r,k)$ code is 2^{k-r} and its dimension is $\sum_{i=0}^{r} \binom{k}{i}$.
If C is an $RM(r,k)$ code, then C^{*2} is an $RM(2r,k)$ code (if $2r \leq k$; $C^{*2} = \mathbb{F}_2^{2^k}$ otherwise) and consequently its parameters are also straightforward to compute.
In spite of these observations, squaring is a quite "destructive" operation for most codes; indeed it was shown in [5] that for large enough k and n, if a linear code C is chosen uniformly at random among all codes of dimension k and length n then with high probability the dimension of C^{*2} will be very close to the "maximal possible dimension" $\min\{n, k(k+1)/2\}$. See [5] for the precise statements.
This implies that, for a random family of codes, with very high probability either the family itself or the family of their squares will be asymptotically bad. On the other hand, a construction, over every finite field, of asymptotically good families of linear codes whose squares are also an asymptotically good family was shown in [26], using algebraic geometric codes.
A final observation is about the operations of shortening and puncturing, which allow to transform a linear code into a shorter one, and how they act on squares. Puncturing a linear code C at coordinate i consists in "erasing" that coordinate, hence producing the code $C_{pu,i}$ defined as

$$C_{pu,i} = \{(c_1, \ldots c_{i-1}, c_{i+1}, \ldots, c_n) : \exists (c_1, \ldots c_{i-1}, c_i, c_{i+1}, \ldots, c_n) \in C\}.$$

As long as $d(C) > 1$, we have that $\dim C_{pu,i} = \dim C$. Moreover $d(C_{pu,i}) \geq d(C) - 1$. Shortening C at coordinate i consists on puncturing at that coordinate the subcode of C consisting of all codewords whose i-th coordinate is 0. The following code is obtained:

$$C_{sh,i} = \{(c_1, \ldots c_{i-1}, c_{i+1}, \ldots, c_n) : (c_1, \ldots c_{i-1}, 0, c_{i+1}, \ldots, c_n) \in C\}.$$

Clearly $\dim C_{pu,i} \geq \dim C - 1$ and $d(C_{pu,i}) \geq d(C)$.
The immediate observations about how these operations interact with squaring are as follows:

Proposition 2.6. *$(C_{pu,i})^{*2} = (C^{*2})_{pu,i}$ and $(C_{sh,i})^{*2} = (C^{*2})_{sh,i}$*

Consequently the distance of the square of C^{*2} does not decrease when C is shortened, and may decrease in at most one unit when C is punctured. By applying the puncturing and shortening operations repeatedly the following is obtained.

COROLLARY 2.7. *Let C be a linear code of length n. For any a, b non-negative integers with $a + b < n$, and $b < d(C)$, there exists a linear code D of length $n - a - b$ and such that $\dim D \geq \dim C - a$ and $d(D^{*2}) \geq d(C^{*2}) - b$.*

2.2 Cyclic codes

From now on it will always be assumed that n is coprime with q. There are a few different ways of defining a cyclic code, and it will be useful to consider two of them. The most common one is as follows. Consider the $\mathbb{F}_q$-vector space $R = \mathbb{F}_q[X]/(X^n - 1)$. Since R has dimension n it is isomorphic as a $\mathbb{F}_q$-vector space to $\mathbb{F}_q^n$ and an isomorphism $\iota : \mathbb{F}_q^n \to R$ is given by

$$(c_0, c_1, \ldots, c_{n-1}) \mapsto c(X) + \langle X^n - 1 \rangle$$

where $c(X) := \sum_{i=0}^{n-1} c_i X^i$.
R is a ring with the product operation induced by the usual product of polynomials in $\mathbb{F}_q[X]$. From now on, the elements in R are identified with polynomials in $\mathbb{F}_q[X]$ of degree at most $n - 1$, since every class in R has exactly one representative of that form.

DEFINITION 2.8. *Let $g \in \mathbb{F}_q[X]$ be a polynomial dividing $X^n - 1$. The cyclic code generated by g is the ideal generated by g in R.*

LEMMA 2.9. *The dimension of the cyclic code C generated by g is $n - \deg g$, since*

$$C = \{g \cdot h \mid h \in \mathbb{F}_q[X],\ \deg h \leq n - \deg g - 1\}.$$

Let β be a primitive n-root of unity in an algebraic closure of $\mathbb{F}_q$, i.e., $\beta^n = 1$ but $\beta^k \neq 1$ for $1 \leq k \leq n$. Let $\mathfrak{F} = \mathbb{F}_q(\beta)$ be the smallest field containing β and $\mathbb{F}_q$. $\mathfrak{F}$ is in fact a finite field $\mathbb{F}_{q^r}$ of q^r elements, where n divides $q^r - 1$.
Since g divides $X^n - 1$, all roots of g are of the form β^j, for some $j \in \{0, \ldots, n - 1\}$. As a matter of fact, since β is a n-root of unity, we can also define the notation β^j for $j \in \mathbb{Z}/n\mathbb{Z}$.

DEFINITION 2.10. *We call $J := \{j \in \mathbb{Z}/n\mathbb{Z} : g(\beta^j) = 0\}$ and $I := \{j \in \mathbb{Z}/n\mathbb{Z} : g(\beta^j) \neq 0\}$ respectively the defining and generating sets of the cyclic code C generated by g.*

Note that $g = \prod_{j \in J}(X - \beta^j) = (X^n - 1)/\prod_{i \in I}(X - \beta^i)$ and hence $\dim C = |I|$, where $|I|$ denotes the cardinality of I. Since g is in $\mathbb{F}_q[X]$, whenever γ is a root of g, γ^q is a root too, and hence there are some restrictions to J and I:

DEFINITION 2.11. *Let $u \in \mathbb{Z}/n\mathbb{Z}$. The q-cyclotomic coset of u is the set $[u] := \{uq^j : j \geq 0\} \subseteq \mathbb{Z}/n\mathbb{Z}$ (where the products are understood to be in $\mathbb{Z}/n\mathbb{Z}$).*

LEMMA 2.12. *Both I and J are unions of q-cyclotomic cosets.*

A key result in the theory of cyclic codes is the following

PROPOSITION 2.13 (BCH bound). *Suppose that $c, d \in \mathbb{Z}/n\mathbb{Z}$ are such that $\{c, c+1, \ldots, c+d-2\} \subseteq J$. Then the minimum distance of C is at least d.*

This motivates the definition of BCH code.

DEFINITION 2.14. *A BCH code of designed distance d is a cyclic code with generator polynomial $g = lcm\{m_j : j \in \{c, c+1, \ldots, c+d-2\}\}$, where m_j is the minimal polynomial of β^j. That is, the defining set J is the union of the cyclotomic cosets containing the elements $c, c+1, \ldots, c+d-2$.*

LEMMA 2.15. *The minimum distance of a BCH code of designed distance d is at least d. Its dimension is at least $n - m(d-1)$, where m is the smallest integer such that $n|(q^m - 1)$. If $q = 2$ and $c = 1$, then its dimension is at least $n - md/2$.*

The dual of a cyclic code is another cyclic code. In fact, the following holds.

DEFINITION 2.16. *For $I \subseteq \mathbb{Z}/n\mathbb{Z}$, $-I$ denotes the set $\{-i : i \in I\} \subseteq \mathbb{Z}/n\mathbb{Z}$.*

LEMMA 2.17. *Let C be the cyclic code generated by g and let $h = \frac{X^n-1}{g(X)} := \sum_{i=0}^{|I|} h_i X^i$. Then*

- *The dual $C^\perp$ of C is the cyclic code generated by the polynomial*

$$h^{[-1]} := \sum_{i=0}^{|I|} h_{|I|-i} X^i.$$

- *Let J and I be the defining and generating sets of C and let J^* and I^* the defining and generating sets of $C^\perp$. Then $J^* = -I$ and $I^* = -J$.*

It is more useful for the problem in hand to consider the following alternative description of cyclic codes as a subfield subcode of an evaluation code over the field $\mathfrak{F} = \mathbb{F}_q^r$. We now follow the notation from [2].

DEFINITION 2.18. *For a set $M \subseteq \{1, \ldots, n-1\}$, we denote by $\mathcal{P}(M)$ the $\mathfrak{F}$-span of the monomials X^i, $i \in M$, i.e.,*

$$\mathcal{P}(M) := \{\sum_{i \in M} f_i X^i : f_i \in \mathfrak{F}\}.$$

In addition, let $\mathcal{B}(M)$ denote the $\mathfrak{F}$-vector space

$$\mathcal{B}(M) := \{(f(1), f(\beta), \ldots, f(\beta^{n-1})) : f \in \mathcal{P}(M)\} \subseteq \mathfrak{F}^n.$$

Finally as it is usual, for a set $V \subseteq \mathfrak{F}^n$, denote $V|_{\mathbb{F}_q} = V \cap \mathbb{F}_q^n$.

LEMMA 2.19. *Let C be the cyclic code generated by*

$$g = \frac{X^n - 1}{\Pi_{i \in I}(X - \beta^i)} \in \mathbb{F}_q[X].$$

Then

$$C = \mathcal{B}(-I)|_{\mathbb{F}_q}.$$

PROOF. In [2, Section 3], it is established that $C = \mathrm{Tr}(\mathcal{B}(J))^\perp$, where

$$\mathrm{Tr}(V) := \{(\mathrm{Tr}(v_1), \mathrm{Tr}(v_2), \ldots, \mathrm{Tr}(v_n)) : (v_1, v_2, \ldots, v_n) \in V\}$$

and Tr denotes the trace from $\mathfrak{F}$ to $\mathbb{F}_q$. Given that I is the complement of J in $\{1, \ldots, n\}$ and that I is a union of cyclotomic sets, [2, Theorem 6] states that $\mathrm{Tr}(\mathcal{B}(J))^\perp = \mathcal{B}(-I)|_{\mathbb{F}_q}$. △

3 Squares of cyclic codes

In this section, several general results on squares of cyclic codes will be obtained.

3.1 Initial observations

Before adapting a more systematic approach towards the study of cyclic codes, some relatively simple observations can be made. First, the following result shows some achievable parameters for $d(C^{*2})$ if the dimension of C divides its length.

PROPOSITION 3.1. *For every finite field $\mathbb{F}_q$, and every integers $m_1, m_2 > 0$ there exists a cyclic code C over $\mathbb{F}_q$ with $\dim C = m_1$, $d(C^{*2}) = m_2$ and length $n = m_1 \cdot m_2$.*
More precisely, those conditions are met by the code

$$C = \{(\mathbf{a}, \mathbf{a}, \ldots, \mathbf{a}) : \mathbf{a} \in (\mathbb{F}_q)^{m_1}\}.$$

In fact, the code above satisfies $C = C^{*2}$. As a consequence of this result, the interest will primarily be in cyclic codes achieving at least $\dim C \cdot d(C^{*2}) > n$.
In second place, an example is shown below where a pair of cyclic codes have the same dimension and minimum distance, but the minimum distance of theirs squares differs. Moreover, it can also happen that the minimum distance of one of the codes is smaller than the other, but the inequality gets reversed for the squares.

EXAMPLE 3.2. *Let $q = 2$ and $n = 63$. Let β be a root of $X^6 + X^4 + X^3 + X + 1$. Then β is a primitive 63-th root of unity. Let m_j denote the minimal polynomial of β^j. We choose $h_1 := m_1 \cdot m_3$, $h_2 := m_1 \cdot m_5$, $h_3 := m_1 \cdot m_7$ and $h_4 = m_1 \cdot m_{11}$. For $i = 1, 2, 3, 4$, let C_i be the binary cyclic code of length 63 generated by the polynomial $(X^n - 1)/h_i$. Then $\dim C_i = 12$ for all i. Furthermore it can be computed that $d(C_1) = d(C_2) = 24$, $d(C_3) = 14$, $d(C_4) = 20$ and $d(C_1^{*2}) = 6$, $d(C_2^{*2}) = 4$, $d(C_3^{*2}) = 3$, $d(C_4^{*2}) = 2$. Hence, C_1 and C_2 have the same dimension and minimum distance, but different minimum distance of their squares. Moreover, C_3 and C_4 have the same dimension and $d(C_3) < d(C_4)$ but $d(C_3^{*2}) > d(C_4^{*2})$.*

This indicates that not so much information can be extracted about $d(C^{*2})$ given only $\dim C$ and $d(C)$, and that the cyclic codes with the largest minimum distance (among the ones with the same length and dimension) do not necessarily have the best squares, which motivates the need for the more careful analysis in the following section.

3.2 Describing the generator polynomial of C^{*2}

Squares of cyclic codes are described now in more detail. First consider the description of a cyclic code as an ideal in R. Given the identification between R and $\mathbb{F}_q^n$, we can talk about the coordinatewise product of elements in R; more precisely, let

$$h * h' := \sum_{i=0}^{n-1} h_i h'_i X^i.$$

Note that given a cyclic code C with generator polynomial g, C^{*2} consists of all elements of the form

$$\sum_i \lambda_i (a_i \cdot g) * (b_i \cdot g)$$

where $\lambda_i \in \mathbb{F}_q$ and $a_i, b_i \in R$ are such that $\deg a_i, \deg b_i \leq n - \deg g - 1$. Furthermore it was observed in [23] that

LEMMA 3.3 ([23]). *C^{*2} is a cyclic code with generator polynomial*

$$g' = \gcd\{g * g, g * (g \cdot X), g * (g \cdot X^2), \ldots, g * (g \cdot X^{n-\deg g-1})\}$$

However, this description is not too easy to work with. Instead, it seems much more useful to use the interpretation of a cyclic code as an evaluation code, given by Lemma 2.19, and then argue about the squares similarly to how it is done in the case of Reed Solomon codes. Consider the following definition.

DEFINITION 3.4. *For subsets $A, B \subseteq \mathbb{Z}/n\mathbb{Z}$, define $A + B := \{i + j : i \in A, j \in B\} \subseteq \mathbb{Z}/n\mathbb{Z}$.*

Now note that if f and $f' \in \mathcal{P}(-I)$ (notation as in Definition 2.18) then $f \cdot f' (\mod X^n - 1)$ is in $\mathcal{P}(-(I+I))$. Hence $\mathcal{B}(-I)^{*2} = \mathcal{B}(-(I+I))$ and by Lemma 2.19, this yields the theorem below, which is a special case of one of the observations in [15] (stated there for a more general result for the product of two, non-necessarily equal, codes).

THEOREM 3.5. *If $C = \mathcal{B}(-I)|_{\mathbb{F}_q}$, then $C^{*2} = \mathcal{B}(-(I+I))|_{\mathbb{F}_q}$.*
In other words, if C is a cyclic code generated by the polynomial

$$g = \frac{X^n - 1}{\prod_{i \in I}(X - \beta^i)},$$

then C^{*2} is a cyclic code with generator polynomial

$$g' = \frac{X^n - 1}{\prod_{\ell \in I+I}(X - \beta^\ell)}.$$

Now the BCH bound (Proposition 2.13) can be used on the square of the cyclic code. To simplify the exposition, the following notation is introduced.

DEFINITION 3.6. *Let $A \subseteq \mathbb{Z}/n\mathbb{Z}$ be a nonempty set. Its amplitude* $\operatorname{amp} A$ *is*

$$\operatorname{amp} A := \min\{i \in \{1, \ldots, n\} : \exists c \in \mathbb{Z}/n\mathbb{Z} \text{ such that } A \subseteq \{c, c+1, \ldots, c+i-1\}\}$$

(where sums are understood to be in $\mathbb{Z}/n\mathbb{Z}$). That is, $\operatorname{amp} A$ *is the size of the smallest set of consecutive elements in $\mathbb{Z}/n\mathbb{Z}$ that contains A.*

REMARK 3.7. *Note that*

- $\operatorname{amp} A \leq 1 + \max A$ *where* $\max A$ *denote the largest element of A when $\mathbb{Z}/n\mathbb{Z}$ is identified with the set of integers $\{0, \ldots, n-1\}$. This is because $A \subseteq \{0, \ldots, \max A\}$.*
- $n - \operatorname{amp} A$ *is the size of the largest set of consecutive elements that do not belong to A, i.e., the largest set of consecutive elements contained in A^c.*
- *It is then a direct consequence of Proposition 2.13 that the minimum distance of a cyclic code C satisfies $d(C) \geq n - \operatorname{amp} I + 1$ (remember $I^c = J$).*

THEOREM 3.8. *Let C be a cyclic code of length n with generator polynomial $g = (X^n - 1)/f(X)$ where $f = \prod_{i \in I}(X - \beta^i)$.*
Then

- $\dim C = |I|$ *and* $\dim C^{*2} = |I + I|$.
- $d(C) \geq n - \operatorname{amp} I + 1$ *and* $d(C^{*2}) \geq n - \operatorname{amp}(I + I) + 1$.

Thus, finding $I \subseteq \{0, \ldots, n-1\}$ such that I is a union of cyclotomic sets, and $|I|$ is large but $\operatorname{amp}(I + I)$ is relatively small will yield codes C such that $\dim C$ and $d(C^{*2})$ are simultaneously large.

4 Some preliminary constructions

In this section some natural approaches towards constructing the index sets I are analyzed. However, the two approaches in this section will lead respectively to Reed-Solomon and generalized Reed-Muller codes, whose squares are well understood, as discussed in Section 2.1. Nevertheless, they will also provide useful intuitions for the more involved techniques presented in Section 2.1, so it is still interesting to elaborate on them here.

The first approach consists in taking the generator set I to be the union of all cyclotomic sets that are entirely contained in $\{0, \dots, t\}$ for some integer $t < n/2$. The idea is that $I + I$ is then contained in $\{0, \dots, 2t\}$ and therefore its amplitude is (at most) $2t + 1$, which gives a lower bound $d(C^{*2}) \geq n - 2t$. Note that the complement of I, the defining set J, is the smallest union of cyclotomic sets containing $\{t + 1, \dots, n - 1\}$. Hence the generator polynomial is $g := \mathrm{mcm}(m_{t+1}, m_{t+2}, \dots, m_{n-1})$ where m_i is the minimal polynomial in $\mathbb{F}_q[X]$ of β^i. In other words, the code C is a BCH code of designed distance $n - t$.

This immediately suggests the following consequence:

THEOREM 4.1. *Let t, n be positive integers and let k be the smallest integer with $n|(q^k - 1)$. There exists a $\mathbb{F}_q$-linear code C, of length n such that*

- $\dim C \geq \max\{1, n - (n - t - 1)k\}$,
- $d(C) \geq n - t$ *and*
- $d(C^{*2}) \geq n - 2t$.

If in addition $q = 2$, then $\dim C \geq \max\{1, n - \frac{(n-t)k}{2}\}$.

PROOF. Take as generator of C the polynomial $g := \mathrm{mcm}(m_{t+1}, m_{t+2}, \dots, m_{n-1})$ where m_i is the minimal polynomial in $\mathbb{F}_q[X]$ of β^i. The statements about the distance of C and C^{*2} follow from Main Theorem 3.8 and the fact that the amplitudes of I and $I + I$ are at most $t + 1$ and $2t + 1$ respectively. On the other hand, the estimates about the dimension are as in Lemma 2.15.

△

Unfortunately, the result above cannot be used to ensure that $\dim C > 1$ and $d(C^{*2}) > 1$ simultaneously, unless in the case where $k = 1$. However, in that case C is just a Reed Solomon code over $\mathbb{F}_q$.

A different idea will be considered next. Given that the set I generating the code needs to be a union of cyclotomic cosets, we can think of associating to each integer a quantity which is invariant within a cyclotomic coset and at the same time can be “controlled” to a certain extent when two integers are summed.

We will from now on consider the case $n = q^k - 1$. Then we can use as invariant the q-ary weight, defined next.

DEFINITION 4.2. *The q-ary representation of an element $t \in \mathbb{Z}/n\mathbb{Z}$ is the unique vector $(t_{k-1}, t_{k-2}, \dots, t_0)_q \in \{0, \dots, q - 1\}^k$ such that $t = \sum_{i=0}^{k-1} t_i q^i$.*
The q-ary weight of t is defined as $w_q(t) = \sum_{i=0}^{k-1} t_i$.

REMARK 4.3. *Note that the binary ($q = 2$) weight of t is the Hamming weight of its binary representation.*

LEMMA 4.4. *Let $n = q^k - 1$. Let $a, b \in \mathbb{Z}/n\mathbb{Z}$. Then:*

- $w_q(q^j a) = w_q(a)$ *for any $j \geq 0$, i.e., all elements in the same q-cyclotomic coset have the same q-ary weight.*

- $w_q(a+b) \leq w_q(a) + w_q(b)$.

PROOF. The first part of the lemma comes from the fact that multiplying by q simply induces a cyclic shift on the q-ary representation of an element of $\mathbb{Z}/n\mathbb{Z}$, because n is of the form $q^k - 1$. For the second part, let $a = \sum_{i=0}^{k-1} a_i q^i, b = \sum_{i=0}^{k-1} b_i q^i$. If $0 \leq a_i + b_i \leq q-1$ for all i (i.e., if there are no carries in the sum), then $w_q(a+b) = w_q(a) + w_q(b)$. Otherwise, whenever there is a carry, the weight will decrease by $q-1$.

$\triangle$

The first part of the lemma implies that we can talk about the q-ary weight of a q-cyclotomic set (which is the q-ary weight of any of its elements). The second part leads to the following:

PROPOSITION 4.5. *If I is the union of all cyclotomic sets whose q-ary weights are at most*[3] *$(q-1)s$, then $I+I$ is a union of cyclotomic sets of weight at most $2(q-1)s$, and moreover* $\operatorname{amp}(I+I) = 1 + q^k - q^{k-2s}$.

PROOF. The first part of the proposition follows directly from Lemma 4.4. For the second part, note $q^k - q^{k-2s} = (q-1)\sum_{\ell=1}^{2s} q^{k-\ell}$ has q-ary representation $(q-1, q-1, \ldots, q-1, 0, 0, \ldots, 0)_q$. Hence it is obviously the largest integer in $\{0, \ldots, n-1\}$ of weight at most $2(q-1)s$. On the other hand, this integer can be written as the sum of two integers from I, namely

$$(q-1)\sum_{\ell=1}^{2s} q^{k-\ell} = (q-1)\sum_{\ell=1}^{s} q^{k-\ell} + (q-1)\sum_{\ell=s+1}^{2s} q^{k-\ell},$$

so it is indeed in $I+I$. We have shown $\max(I+I) = q^k - q^{k-2s}$. On the other hand $0 \in I+I$, and hence the amplitude of $I+I$ is exactly $1 + q^k - q^{k-2s}$. $\triangle$

COROLLARY 4.6. *Let $n = q^k - 1$ and let C be the cyclic code generated by the polynomial $g = (X^n - 1)/f(X)$, where $f = \prod_{i \in I}(X - \beta^i)$ and $I = \{i : w_q(i) \leq (q-1)s\}$. We have $d(C^{*2}) \geq q^{k-2s} - 1$.*

PROOF. It follows from Main Theorem 3.8 and Proposition 4.5.

$\triangle$

Nevertheless, these cyclic codes are nothing else than generalized Reed-Muller codes punctured in one position, as it is shown next.

REMARK 4.7. *By setting $\alpha = \beta^{n-1}$ (which is again a primitive n-th root of unity), and noticing that $w_q(n-i) = (q-1)k - w_q(i)$, it is easy to see that $g = \prod_{j \in J'}(X - \alpha^j)$ and $1 \leq w_q(j) \leq (q-1)k - (q-1)s$ for all $j \in J'$.*

PROPOSITION 4.8. *In the conditions of the previous theorem C is equivalent to a generalized Reed-Muller code of length q^k and order $(q-1)s$ punctured in one position.*

PROOF. By the Remark 4.7, the polynomial g can be written as $g = \prod_{1 \leq w_q(j) \leq (q-1)k-(q-1)s}(X - \alpha^j)$. Now let C' be the cyclic code generated by the polynomial

$$g' = (X-1)g = \prod_{0 \leq w_q(j) \leq (q-1)k-(q-1)s} (X - \alpha^j)$$

and let D be the code of length q^k spanned by the vectors $\{(c', 0) : c' \in C'\} \cup \{(1, \ldots, 1)\} \subseteq \mathbb{F}_q^{n+1}$. It is known [22] that D is equivalent to the Reed Muller code of length q^k and

[3]There is nothing special about choosing a bound of the form $(q-1)s$. It just leads to a simpler expression for the square distance, but the result can be generalized to other values which are not divisible by $q-1$.

order $(q-1)s$. Puncturing this code in the last position we obtain the code spanned by $C' \cup \{(1,\ldots,1)\} \subseteq \mathbb{F}_q^n$. It is easy to see that this code is C, since $\dim C = \dim C' + 1$, C contains C', and C contains the vector $(1,\ldots,1) \in \mathbb{F}_q^n$ (because $1 + X + \cdots + X^{n-1} = (X^n - 1)/(X-1)$ is clearly a multiple of g).

△

5 Construction of codes based on restricted weights

In this section, a modification of the second approach from the previous section is suggested. As in the last part of the previous section, the length of the codes will be $n = q^k - 1$ for some k. The modification consists on replacing the notion of q-ary weight by a more fine-grained notion of weight which is defined next.

DEFINITION 5.1. *Let $t \in \{0,\ldots,n-1\}$ with q-ary representation $(t_{k-1}, t_{k-2}, \ldots, t_0)_q$, and let $1 \le s \le k$.*
The s-restricted binary weight of t is defined as $w_q^{(s)}(t) = \max_{i\in\{0,\ldots,k-1\}} \sum_{j=0}^{s-1} t_{i+j}$, where the sums $i+j$ are considered modulo k.

That is, the s-restricted binary weight of t is the maximum Hamming weight of a substring of s consecutive digits in the binary representation of t. Here “consecutive” is also meant cyclically, and hence it is clear that this notion is an invariant of a cyclotomic coset, i.e., $w_q^{(s)}(q^i t \mod n) = w_q^{(s)}(t)$ for any $i \ge 0$. Thus, we can speak of the s-restricted binary weight of a cyclotomic coset.

REMARK 5.2. *If $s = k$, then $w_2^{(s)}(t) = w_2(t)$.*

Moreover the notion of restricted weight also satisfies the subadditivity property.

PROPOSITION 5.3. *Let $t, u \in \{0,\ldots,n-1\}$. Let $v := t + u \mod n$. Then $w_q^{(s)}(v) \le w_q^{(s)}(t) + w_q^{(s)}(u)$.*

The proof of this result is somewhat tedious and it is therefore deferred to the appendix.
In view of the proposition above, it is clear that if we take I to contain only elements of s-restricted weight at most m, then all elements in $I+I$ will have s-restricted weight at most $2m$. This motivates the following definitions and results.

DEFINITION 5.4. *We denote:*

$$W_{k,s,m} := \{j \in \{0,\ldots,n-1\} : w_q^{(s)}(j) \le m\}.$$

$$N_{k,s,m} := |W_{k,s,m}|.$$

$$B_{k,s,m} := \max\ W_{k,s,m}.$$

(where, as usual, $n = q^k - 1$).

PROPOSITION 5.5. *We have* $\mathrm{amp}(W_{k,s,m}) \le 1 + B_{k,s,m}$. *Furthermore let $m \le \frac{s-1}{2}$. Then $W_{k,s,m} + W_{k,s,m} \subseteq W_{k,s,2m}$ and consequently* $\mathrm{amp}(W_{k,s,m} + W_{k,s,m}) \le 1 + B_{k,s,2m}$.

PROOF. This is straightforward from the definitions above and Proposition 5.3. △

Nevertheless, it does not necessarily hold always that $W_{k,s,m} + W_{k,s,m} = W_{k,s,2m}$. Indeed, consider the following counterexample for $q = 2$; on the one hand it holds that $W_{5,3,1} = \{0, 1, 2, 4, 8, 16\}$, which are all the binary strings of length 5 and weight 0 and 1; indeed, given any string of weight at least 2, one can find 3 cyclically consecutive positions containing two

1's. By Lemma 4.4 in the previous section, every element in $W_{5,3,1}+W_{5,3,1}$ must have binary weight at most 2. On the other hand, since the binary representation of 26 is 11010, we have $26 \in W_{5,3,2} \setminus (W_{5,3,1}+W_{5,3,1})$.
This observation is generalized below in order to provide a tighter bound for $\mathrm{amp}(W_{k,s,m} + W_{k,s,m})$.

Proposition 5.6. *Let $t \in W_{k,s,m}$. Then $w_q(t) \leq \left\lfloor \frac{mk}{s} \right\rfloor$.*

Proof. The q-ary representation of t contains k different substrings of s cyclically consecutive positions, and each position belongs to s of these strings. Hence the sum S of the weights of these strings is exactly $S = sw_q(t)$. On the other hand, each of these strings has weight at most m, and hence $S \leq km$. Hence $w_q(t) \leq \frac{mk}{s}$ and the result follows from the fact that $w_q(t)$ is an integer. △

Corollary 5.7. *Let $t \in W_{k,s,m} + W_{k,s,m}$. Then $w_q(t) \leq 2\left\lfloor \frac{mk}{s} \right\rfloor$.*

Note that solely from the fact that $t \in W_{k,s,2m}$, one can only guarantee that $w_q(t) \leq \left\lfloor \frac{2mk}{s} \right\rfloor$. This may be larger than $2\left\lfloor \frac{mk}{s} \right\rfloor$.

Definition 5.8. *Let*

$$\widehat{B}_{k,s,2m} := \max\left\{ t \in W_{k,s,2m} : w_2(t) \leq 2\left\lfloor \frac{mk}{s} \right\rfloor \right\}$$

Theorem 5.9. *let C be the cyclic code generated by the polynomial $g = (X^n-1)/f(X)$, where $f = \prod_{i \in W_{k,s,m}} (X-\beta^i)$. Then*

- $\dim C = N_{k,s,m}$.
- $d(C) \geq n - B_{k,s,m}$.
- $d(C^{*2}) \geq n - \widehat{B}_{k,s,2m}$.

A slight variation of this result can be obtained if the index 0 is removed from $W_{k,s,m}$. Let $I := W_{k,s,m} \setminus \{0\}$. Then obviously $|I| = N_{k,s,m} - 1$. On the other hand, if $2\left\lfloor \frac{mk}{s} \right\rfloor < k$, then $0 \notin I + I$. In these conditions, $I + I \subseteq \{1, \ldots, \widehat{B}_{k,s,2m}\}$. Hence

Theorem 5.10. *let C be the cyclic code generated by the polynomial $g = (X^n-1)/f(X)$, where $f = \prod_{i \in W_{k,s,m} \setminus \{0\}} (X-\beta^i)$. In addition, assume $2\left\lfloor \frac{mk}{s} \right\rfloor < k$. Then*

- $\dim C = N_{k,s,m} - 1$.
- $d(C) \geq n - B_{k,s,m} + 1$.
- $d(C^{*2}) \geq n - \widehat{B}_{k,s,2m} + 1$.

The rest of this section is devoted to analyze the numbers $B_{k,s,m}$, $\widehat{B}_{k,s,2m}$, and $N_{k,s,m}$.

5.1 Bounds for the distance of the codes and their squares.

It is not difficult to calculate $B_{k,s,m}$ and $\widehat{B}_{k,s,2m}$. We simply consider their binary representations and determine their bits one by one, going from the highest order bit to the lowest one and assigning, at each step, 1 if this is consistent with the conditions on the weights, or 0 otherwise.
Therefore, assuming $n \geq s$, the binary representation of $B_{k,s,m}$ will begin with $\lfloor k/s \rfloor$ blocks of the form 11...100...0 (m ones and $s-m$ zeros). The remaining $k - s\lfloor k/s \rfloor < s$ positions

should contain 1's until the last $s-m$ positions are reached: these must in any case be all zero, because of the fact that the first m positions are one, and that the definition of restricted weight considers any set of s consecutive positions cyclically. The binary representation of $\widehat{B}_{k,s,2m}$ is obtained from the representation of $B_{k,s,2m}$ by swapping the 1's in the lowest order coordinates to 0's until the weight is at most $2\left\lfloor \frac{mk}{s} \right\rfloor$.

LEMMA 5.11. *Let $1 \le m \le s-1$. Then*

$$B_{k,s,m} = \sum_{i=0}^{\lfloor \frac{k}{s} \rfloor -1} \sum_{j=1}^{m} 2^{k-is-j} + \sum_{i=s-m}^{k-s\lfloor \frac{k}{s} \rfloor -1} 2^i$$

Let $1 \le m \le \frac{s-1}{2}$. Then

$$\widehat{B}_{k,s,2m} = \sum_{i=0}^{\lfloor \frac{k}{s} \rfloor -1} \sum_{j=1}^{2m} 2^{k-is-j} + \sum_{i=u}^{k-s\lfloor \frac{k}{s} \rfloor -1} 2^i$$

where $u = \max\{s-2m, (k - s\lfloor \frac{k}{s} \rfloor) - (2\left\lfloor \frac{mk}{s} \right\rfloor - 2m\left\lfloor \frac{k}{s} \right\rfloor)\}$

PROOF. Let $B_{k,s,m} = \sum_{\ell=0}^{k-1} b_\ell 2^\ell$. We determine all b_ℓ's starting by b_{k-1} and ending with b_0. At every step b_ℓ will be 1 if and only if that selection, together with the already fixed values for $b_{\ell+1}, \ldots, b_{k-1}$, allows $w_2^{(s)}(B_{k,s,m}) \le m$. In the case of $\widehat{B}_{k,s,2m}$ we follow the exact same process but ensuring that the weight is at most $2\left\lfloor \frac{mk}{s} \right\rfloor$. Since the first summand in the expression above has weight $2m\left\lfloor \frac{k}{s} \right\rfloor$, then the second summand must have weight at most $2\left\lfloor \frac{mk}{s} \right\rfloor - 2m\left\lfloor \frac{k}{s} \right\rfloor$.

$\triangle$

REMARK 5.12. *In particular, we have $\widehat{B}_{k,s,2m} < 2^k - 2^{k-2m-1}$, and hence, the distance of C^{*2} where C is defined as in Theorem 5.9 always satisfies $d(C^{*2}) \ge 2^{k-2m-1}$.*

5.2 Determining the dimension of the codes

In this section, a recurrence formula for the numbers $N_{k,s,m}$ with respect to the length k will be found. Remember $N_{k,s,m}$ equals the number of strings in $\{0, \ldots, q-1\}^k$ such that every sequence of s consecutive positions of the string and of its cyclic shifts contains at most m ones.

In the case $q = 2$ and if we remove the cyclic condition (meaning for example that in the case $k = 4$, $s = 3$, $m = 2$, the string 1011 would be included in the counting while its cyclic shift 1110 would not) a solution for certain parameters of s, m can be found in the online encyclopedia of integer sequences [20]. More concretely, the cases $s = 4, m = 2$ and $s = 5, m = 2$ are studied in sequences A118647 and A120118 respectively.

Nevertheless, the cyclic version of this problem does not seem to have been studied anywhere in the literature. The following is an adaptation to our problem of the counting strategy hinted at in the aforementioned references. It is based on counting the number of closed walks of length k in certain graph.

Fix integers $m \ge 1$ and $s \ge 2$ with $m < s$.

DEFINITION 5.13. *Let $V_{(s-1),m}$ be the set of all elements $x \in \{0, \ldots, q-1\}^{s-1}$ of Hamming weight at most m. We define the set $E_{(s-1),m} \subseteq V_{(s-1),m} \times V_{(s-1),m}$ as follows: $(x, y) \in E_{(s-1),m}$ if and only if*

1. $x_2 = y_1,\ x_3 = y_2,\ \ldots,\ x_{s-1} = y_{s-2}$ *and*

2. *the weight of the string* $(x_1, x_2, \dots, x_{s-1}, y_{s-1})$ *is at most* m.

Now we have:

THEOREM 5.14. *Let* $t \in \{0, \dots, q^k - 1\}$ *and let* $(t_0, t_1, \dots, t_{k-1})$ *be its binary representation. In addition let* $t_{k+i} := t_i$ *for* $i = 0, \dots, s-1$. *Then* $w_2^{(s)}(t) \leq m$ *if and only if, for all* j *in* $0, \dots, k-1$,

1. $(t_j, t_{j+1}, \dots, t_{j+s-2}) \in V_{(s-1),m}$ *and*

2. $((t_j, t_{j+1}, \dots, t_{j+s-2}), (t_{j+1}, t_{j+2}, \dots, t_{j+s-1})) \in E_{(s-1),m}$

It is clear that the pair $(V_{(s-1),m}, E_{(s-1),m})$ is a directed graph. In addition it has at most one edge connecting two vertices in a given direction, and it may contain loops (an edge may connect one vertex to itself).
In the following, for the sake of simplicity, a graph will mean a directed graph with the properties just mentioned.

DEFINITION 5.15. *A walk of length* k *in a graph* (V, E) *is a sequence* $(v_0, v_1, \dots, v_k)$ *such that* $v_j \in V$ *for* $j = 0, \dots, k$ *and* $(v_{j-1}, v_j) \in E$ *for* $j = 1, \dots, k$. *Here* v_j *and* $v_{j'}$ *do not need to be different. The vertex* v_0 *is the initial vertex of the walk and the vertex* v_k *is the terminal vertex of the walk. The walk is closed if the initial and terminal vertices coincide, i.e.,* $v_0 = v_k$.

PROPOSITION 5.16. *There is a one to one correspondence between the set* $\{t \in \{0, \dots, n-1\} : w_2^{(s)}(t) \leq m\}$ *and the set of closed walks of length* k *in the graph* $(V_{(s-1),m}, E_{(s-1),m})$.

PROOF. Given an integer $t \in \{0, \dots, n-1\}$ with $(t_0, t_1, \dots, t_{k-1})$ as its q-ary representation, we associate the sequence $v_0, v_1, \dots, v_k \in \{0, \dots, q-1\}^{s-1}$ where $v_j = (t_j, t_{j+1}, \dots, t_{j+s-2})$. As usual the sums in the indices are modulo k. Then a direct consequence of Theorem 5.14 is that $w_2^{(s)}(t) \leq m$ if and only if $(v_0, \dots, v_k)$ is a walk of length k. Moreover $v_0 = v_k$, so it is a closed walk. It is clear that every closed walk of length k corresponds to a unique integer. △

Now it is a well known fact from graph theory that

LEMMA 5.17. *Let* $A \in \mathbb{R}^{g \times g}$ *be the adjacency matrix of the graph. The number of walks of length* k *with initial vertex* v *and terminal vertex* w *is the* (v, w)*-th entry of the matrix* A^k. *In particular, the* (v, v)*-th entry of* A^k *is the number of closed walks of length* k *starting and ending in* v.

COROLLARY 5.18. *We have*

$$N_{k,s,m} = \mathrm{Tr}(A^k)$$

where A *is the adjacency matrix of the graph* $(V_{(s-1),m}, E_{(s-1),m})$ *and* Tr *denotes its trace, i.e., the sum of its diagonal elements.*

Note that the graph $(V_{(s-1),m}, E_{(s-1),m})$ and therefore its adjacency matrix do not depend on k, and hence having fixed s, m the matrix A is completely determined. Furthermore a recurrence formula can be given for the successive powers of A, and hence for their traces.

PROPOSITION 5.19. *Let* $A \in \mathbb{R}^{g \times g}$*-matrix and* $p(X) = \sum_{i=0}^{g} p_i X^i$ *its characteristic polynomial. Then*

$$\sum_{i=0}^{g} p_i \, \mathrm{Tr}(A^{i+j}) = 0.$$

The proposition follows from Cayley-Hamilton theorem, which states that $p(A) = 0$, i.e., $\sum_{i=0}^{g} p_i A^i$ is the all-zero matrix. Multiplying by A^j and using linearity of the trace yields the result.
This leads to a recurrence formula for the numbers $N_{k,s,m}$. Since these are only defined for $k \geq s$, we introduce the following definition.

DEFINITION 5.20. *Let* $N'_{k,s,m} := \mathrm{Tr}(A^k)$*, where* A *is the adjacency matrix of the graph* $(V_{(s-1),m}, E_{(s-1),m})$.

THEOREM 5.21. *For* $k \geq s$, $N_{k,s,m} = N'_{k,s,m}$*; and for all* $k \geq g$ *the numbers* $N'_{k,s,m}$ *satisfy the recurrence*

$$N'_{k,s,m} = -\sum_{j=1}^{g} p_{g-j} N'_{(k-j),s,m}$$

where $p(X) = \sum_{i=0}^{g} p_i X^i$ *is the characteristic polynomial of the graph* $(V_{(s-1),m}, E_{(s-1),m})$.

6 Some concrete values for binary codes

In this section we compute the parameters obtained for $q = 2$ and certain specific choices for s and m in the construction from the previous section.

6.1 Case $m = 1$

Remember that for each integer $k \geq s$, we are considering the cyclic code C of length $n = 2^k - 1$, generated by the polynomial $g = (X^n - 1)/\prod_{i \in W_{k,3,1}}(X - \beta^i)$, where $W_{k,3,1} = \{t \in \{0, \ldots, n-1\} : w_2^{(3)}(t) \leq 1\}$.
We first determine the numbers $B_{k,s,1}$, $\widehat{B}_{k,s,2}$ which, according to Theorem 5.9, yield bounds for the minimum distance of the codes C and C^{*2} respectively. Let $\ell = k \mod s$. The binary representation of $B_{k,s,1}$ starts with $\lfloor k/s \rfloor$ blocks of the form 100...0 (one 1 and $s-1$ zeros). The remaining ℓ bits need to be all zero because otherwise the block of the last $s - 1 \geq \ell$ bits, together with the first one, would create a sequence of s consecutive positions with weight at least 2.
Therefore a recurrence is given by the formula

$$B_{s,s,1} = 2^{s-1} \text{ and } B_{k,s,1} = \begin{cases} 2B_{(k-1),s,1} + 2^{s-1}, & \text{if } \ k = 0 \mod s \\ 2B_{(k-1),s,1}, & \text{if } \ k \neq 0 \mod s \end{cases}, \text{ for } k > s.$$

If we now write $d_k = 2^k - 1 - B_{k,s,1}$ (which is the bound for $d(C)$ promised by Theorem 5.10), then d_k satisfies

$$d_s = 2^{s-1} - 1 \text{ and } d_k = \begin{cases} 2d_{k-1} - 2^{s-1} + 1, & \text{if } \ k = 0 \mod s \\ 2d_{k-1} + 1, & \text{if } \ k \neq 0 \mod s \end{cases}, \text{ for } k > s.$$

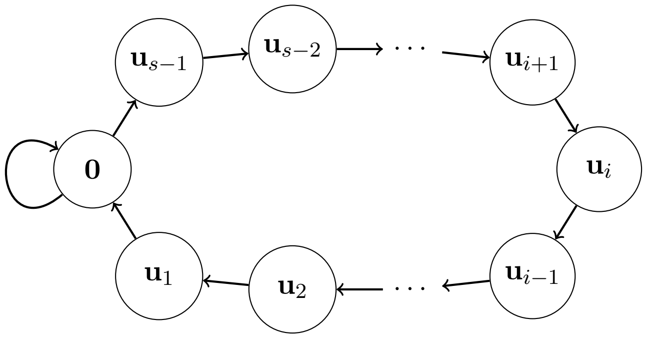
In the case of the numbers $\widehat{B}_{k,s,2}$, observe that, since $m = 1$ and hence $2m \left\lfloor \frac{k}{s} \right\rfloor = 2 \left\lfloor \frac{mk}{s} \right\rfloor$, the second summand of the expression in Lemma 5.11 is always 0.
Therefore, the binary representation of $\widehat{B}_{k,s,2}$ consists of $\lfloor k/s \rfloor$ blocks of the form 110...0 (two 1's and $s - 2$ 0's) followed by ℓ 0's. Hence $\widehat{B}_{k,s,2}$ satisfies the recursion

$$\widehat{B}_{s,s,2} = 2^{s-1} + 2^{s-2} \text{ and } \widehat{B}_{k,s,2} = \begin{cases} 2B_{(k-1),s,2} + 2^{s-1} + 2^{s-2}, & \text{if } \ k = 0 \mod s \\ 2B_{(k-1),s,2}, & \text{if } \ k \neq 0 \mod s \end{cases}, \text{ for } k > s.$$

Moreover, the numbers $\widehat{d}_k = 2^k - 1 - \widehat{B}_{k,s,2}$ satisfy

$\widehat{d}_s = 2^{s-1} - 2^{s-2} - 1$ and $\widehat{d}_k = \begin{cases} 2\widehat{d}_{k-1} - 2^{s-1} - 2^{s-2} + 1, & \text{if} \quad k = 0 \mod s \\ 2\widehat{d}_{k-1} + 1, & \text{if} \quad k \neq 0 \mod s \end{cases}$, for $k > s$.

We now determine the numbers $N'_{k,s,1}$ which for $k \geq s$ yield the dimension of the code (the size of I). For this we use the graph $(V_{(s-1),1}, E_{(s-1),1})$.
The vertex set $V_{(s-1),1}$ consists of the all-zero vector and all the unit vectors in $\{0,1\}^s$. The graph $(V_{(s-1),1}, E_{(s-1),1})$ is



Indeed observe that even though $\mathbf{u}_1$ can be glued with $\mathbf{u}_{s-1}$, the resulting vector 10...01 would have weight 2 and hence $(\mathbf{u}_1, \mathbf{u}_{s-1}) \notin E_{(s-1),1}$.
The adjacency matrix A of the graph is of the following form

$$\begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

It is not difficult to verify that the characteristic polynomial of A is $X^s - X^{s-1} - 1$.
Hence we have the recurrence

$$N'_{k,s,1} = N'_{(k-1),s,1} + N'_{(k-s),s,1},$$

for $k \geq s$. It remains to compute the values $N'_{k,s,1} = \mathrm{Tr}(A^k)$ for $0 \leq k \leq s-1$. Observe that $\mathrm{Tr}(A^0) = s$ and $Tr(A) = 1$ can be observed directly. For the remaining values, one could compute the matrix A^k, but it is just easier to remember that $\mathrm{Tr}(A^k)$ is the number of closed walks of length k in the graph. Clearly, for $k < s$, the only closed walk of length k is the walk $(\mathbf{0}, \mathbf{0}, \dots, \mathbf{0})$, as any walk involving any other vertex will take at least k steps to return to the origin. Hence, $N'_{k,s,1} = 1$ for $1 \leq k \leq s-1$.
The observations in this section are collected in the following theorem.

THEOREM 6.1. *Let $k \geq s \geq 3$. Let C be the cyclic code generated by the polynomial $g = (X^n - 1)/f(X)$, where $f = \prod_{i \in W_{k,s,1}} (X - \beta^i)$. Then*

- $\dim C = N'_{k,s,1}$, *where $N'_{k,s,1}$ is given by the recurrence*

$$N'_{k,s,1} = N'_{(k-1),s,1} + N'_{(k-s),s,1}, \text{ for } k \geq s,$$

$N'_{k,s,1} = 1$, *for $1 \leq k \leq s-1$ and $N'_{0,s,1} = s$.*

- $d(C) \geq d_k$, *where* d_k *is given by the recurrence*

$$d_s = 2^{s-1} - 1 \text{ and } d_k = \begin{cases} 2d_{k-1} - 2^{s-1} + 1, & \text{if } \ k = 0 \mod s \\ 2d_{k-1} + 1, & \text{if } \ k \neq 0 \mod s \end{cases}, \text{ for } k > s.$$

- $d(C^{*2}) \geq \widehat{d}_k$, *where* $\widehat{d}_k$ *is given by the recurrence*

$$\widehat{d}_s = 2^{s-1} - 2^{s-2} - 1 \text{ and } \widehat{d}_k = \begin{cases} 2\widehat{d}_{k-1} - 2^{s-1} - 2^{s-2} + 1, & \text{if } \ k = 0 \mod s \\ 2\widehat{d}_{k-1} + 1, & \text{if } \ k \neq 0 \mod s \end{cases}, \text{ for } k > s.$$

The explicit parameters obtained for the first few values of k in the case $s = 3$ are collected in Table 1. Here $\dim C$ and the bounds for $d(C)$ and $d(C^{*2})$ follow from the explicit formulas above, while the values for $\dim C^{*2}$ have been obtained by direct computation. Moreover, the parameters obtained for both C and C^{*2} are compared with the code tables from [19, 29] which collect lower and upper bounds for the largest possible minimum distance of a (in this case, binary) linear code of a given length and dimension. In the table below the observation "C (resp. C^{*2}) best known" means that, according to [19, 29], no binary code of length n is known that has the same dimension of C (resp. C^{*2}) and larger minimum distance. Furthermore, "optimal" means that no code with the same length and dimension and strictly larger minimum distance can exist.

k	n	$\dim C$	$d(C) \geq$	$\dim C^{*2}$	$d(C^{*2}) \geq$	Observations
3	7	4	3	7	1	Both C and C^{*2} optimal
4	15	5	7	11	3	Both C and C^{*2} optimal
5	31	6	15	16	7	C optimal, C^{*2} not
6	63	10	27	37	9	C best known, C^{*2} not
7	127	15	55	71	19	Both C and C^{*2} best known
8	255	21	111	123	39	Both C and C^{*2} best known
9	511	31	219	232	73	
10	1023	46	439	441	147	
11	2047	67	879	804	295	
12	4095	98	1755	1475	585	

Table 1: Case $m = 1$, $s = 3$.

6.2 Case $s = 5$, $m = 2$

In order to obtain codes with larger dimension (for the same length), one needs to increase the value of m. On the other hand, fixing a value of m, the largest dimensions are obtained when s is as small as possible, and since we are operating under the restriction $m \leq \frac{s-1}{2}$, this suggests to use $s = 2m + 1$. In this section the case $s = 5$, $m = 2$ is analysed.
The first $5\lfloor k/5 \rfloor$ bits of the binary representations of the numbers $B_{k,5,2}$ consist of $\lfloor k/5 \rfloor$ repetitions of the block 11000. The remaining bits must satisfy that the three last bits need to be 0 because of the restricted weight (cyclic) constraint and the fact that the two first bits of $B_{k,5,2}$ are 1. Therefore, these remaining bits are respectively $0, 00, 000, 1000$ for $k = 1, 2, 3, 4 \mod 5$ Hence we have the recurrence

$$B_{5,5,2} = 24 \text{ and } B_{k,5,2} = \begin{cases} 2B_{(k-1),5,2} + 8, & \text{if } \ k = 0, 4 \mod 3 \\ 2B_{(k-1),5,2}, & \text{if } \ k = 1, 2, 3 \mod 3 \end{cases}, \text{ for } k > 5.$$

As for $\widehat{B}_{k,5,4}$, note that the first $5\lfloor k/5\rfloor$ bits of their binary representations are 11110. The conditions on the restricted weight imply that the last bit must be 0 of each of these numbers. Finally, by the definition the binary weight of $\widehat{B}_{k,5,4}$ is at most $2\lfloor 2k/5\rfloor$, which equals $4\lfloor k/5\rfloor$ if $k = 0,1,2 \mod 5$ and $4\lfloor k/5\rfloor + 2$ if $k = 3,4 \mod 5$. Note that the first $5\lfloor k/5\rfloor$ bits of $\widehat{B}_{k,5,4}$ already have weight $4\lfloor k/5\rfloor$. Hence the remaining bits are respectively $0, 00, 110, 1100$ for $k = 1,2,3,4 \mod 5$, and we have the recurrence

$$\widehat{B}_{5,5,4} = 30 \text{ and } \widehat{B}_{k,5,4} = \begin{cases} 2\widehat{B}_{(k-1),5,4} + 6, & \text{if} \quad k = 0,3 \mod 5 \\ 2\widehat{B}_{(k-1),5,4}, & \text{if} \quad k = 1,2,4 \mod 5 \end{cases}, \text{ for } k > 5.$$

We analyze the numbers $N'_{k,5,2}$. The set $V_{4,2}$ consists of the 11 vectors 0000, 0001, 0010, 0011, 0100, 0101, 0110, 1000, 1001, 1010, 1100. The characteristic polynomial of the graph $(V_{4,2}, E_{4,2})$ is $X^{11} - X^{10} - X^8 - 2X^6 + X^3 + X$ and therefore the recurrence

$$N'_{k,5,2} = N'_{(k-1),5,2} + N'_{(k-3),5,2} + 2N'_{(k-5),5,2} - N'_{(k-8),5,2} - N'_{(k-10),5,2}$$

holds for $k \geq 11$. Direct computation yields that the values of $N'_{k,5,2}$ for $k = 1, 2, \ldots, 10$ are $1, 1, 4, 5, 16, 22, 29, 45, 76, 126$ respectively.

Theorem 6.2. *Let $k \geq 5$. Let C be the cyclic code generated by the polynomial $g = (X^n - 1)/f(X)$, where $f = \prod_{i \in W_{k,5,2}}(X - \beta^i)$. Then*

- $\dim C = N'_{k,5,3}$, *where $N'_{k,5,2}$ is given by the recurrence*

$$N'_{k,5,2} = N'_{(k-1),5,2} + N'_{(k-3),5,2} + 2N'_{(k-5),5,2} - N'_{(k-8),5,2} - N'_{(k-10),5,2},$$

  and $N'_{k,5,2} = 1, 1, 4, 5, 16, 22, 29, 45, 76, 126$ for $k = 1, 2, \ldots, 10$.

- *$d(C) \geq d_k$, where d_k is given by the recurrence*

$$d_5 = 7 \text{ and } d_k = \begin{cases} 2d_{k-1} - 7, & \text{if} \quad k = 0,4 \mod 5 \\ 2d_{k-1} + 1, & \text{if} \quad k = 1,2,3 \mod 5 \end{cases}, \text{ for } k > 5.$$

- *$d(C^{*2}) \geq \widehat{d}_k$, where $\widehat{d}_k$ is given by the recurrence*

$$\widehat{d}_5 = 1 \text{ and } \widehat{d}_k = \begin{cases} 2\widehat{d}_{k-1} - 5, & \text{if} \quad k = 0,3 \mod 5 \\ 2\widehat{d}_{k-1} + 1, & \text{if} \quad k = 1,2,4 \mod 5 \end{cases}, \text{ for } k > 5.$$

Concretely, for the first few values of k, we obtain the parameters collected by Table 2. The same comments about the "Observations" column apply as in Table 1.
Note that by Theorem 5.10, for every entry of Tables 1 and 2 and its corresponding code C, another cyclic code C' can be found with the same length, with $\dim C' = \dim C - 1$ and such that the lower bounds for $d(C')$ and $d((C')^{*2})$ are one unit more than in the table. Furthermore, Corollary 2.7 guarantees that if a, b are integers with as long as $a + b < n$ and $b < d(C)$, then by shortening and puncturing we can find a (non-necessarily cyclic) code D with length $n - a - b$, $\dim D \geq \dim C - a$, $d(D) \geq d(C) - b$ and $d(D^{*2}) \geq d(C^{*2}) - b$. A similar remark holds by replacing C by the C' mentioned some lines above.

k	n	$\dim C$	$d(C) \geq$	$\dim C^{*2}$	$d(C^{*2}) \geq$	Observations
5	31	16	7	31	1	C best known, C^{*2} optimal
6	63	22	15	57	3	C^{*2} optimal
7	127	29	31	99	7	
8	255	45	63	223	9	C^{*2} best known
9	511	76	119	430	19	C^{*2} best known
10	1023	126	231	863	33	C^{*2} best known
11	2047	210	463	1695	67	
12	4095	338	927	3293	135	

Table 2: Case $m = 2$, $s = 5$.

7 Final remarks and open questions

In some cases in Tables 1 and 2, both C and C^{*2} are optimal in the sense that both $d(C)$ is the largest possible for a code of length n and dimension $\dim C$, and $d(C^{*2})$ is the largest possible for a code of length n and dimension $\dim C^{*2}$. In other cases, both $d(C)$ and $d(C^{*2})$ match the largest values which are known to be attainable according to the tables of binary codes in [19, 29]. It should be remarked that [19, 29] only contains information about binary linear codes up to certain length (which is 512 in [19] and 1024 under certain restrictions for the dimension in the case of [29]) and hence the parameteres of some of the longer codes obtained here cannot be measured against these tables.
A natural question is whether it also holds that $d(C^{*2})$ is the largest possible given $(n, \dim C)$, since this would be desirable for the applications mentioned in the introduction. In principle this cannot be established from the optimality of C and C^{*2} only, since it is conceivable that there exists another code E of length n such that $\dim E = \dim C$ and $\dim E^{*2} < \dim C^{*2}$; in such a case it would be possible that $d(E^{*2}) > d(C^{*2})$. Hence, it is left as an open question whether the codes in the table do achieve the largest possible value for $d(C^{*2})$ given $(n, \dim C)$.
This observation also motivates the problem of finding tighter upper bounds for $d(C^{*2})$ given $(n, \dim C)$ over small fields. It would be also interesting to find lower bounds for $\dim C^{*2}$ in terms of $(n, \dim C)$ (which would yield upper bounds for $d(C^{*2})$ as well) but it should be remarked that in that case, additional assumptions on the codes are needed to avoid the trivial case where $C^{*2} = C$, as happens for example in Proposition 3.1. For example, [23] showed asymptotical lower bounds for $\dim C^{*2}$ in terms of $\dim C$ and the minimum distance of the dual $d(C^{\perp})$.
In fact, applications to multiplicative secret sharing, as mentioned in the introduction, usually require the dual of the codes to have large minimum distance, too. Therefore it is another interesting open question to modify the constructions in this paper in order to achieve linear codes C with simultaneously large $d(C^{\perp})$ and $d(C^{*2})$, especially over the field of two elements.

References

[1] Michael Ben-Or, Shafi Goldwasser, and Avi Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation (extended abstract). In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 1–10, 1988.

[2] Jürgen Bierbrauer. The theory of cyclic codes and a generalization to additive codes. *Des. Codes Cryptography*, 25(2):189–206, 2002.

[3] Ignacio Cascudo. Powers of codes and applications to cryptography. In *2015 IEEE Information Theory Workshop, ITW 2015, Jerusalem, Israel, April 26 - May 1, 2015*, pages 1–5, 2015.

[4] Ignacio Cascudo, Hao Chen, Ronald Cramer, and Chaoping Xing. Asymptotically Good Ideal Linear Secret Sharing with Strong Multiplication over *Any* Fixed Finite Field. In *Advances in Cryptology - CRYPTO 2009, 29th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2009. Proceedings*, pages 466–486, 2009.

[5] Ignacio Cascudo, Ronald Cramer, Diego Mirandola, and Gilles Zémor. Squares of Random Linear Codes. *IEEE Trans. Information Theory*, 61(3):1159–1173, 2015.

[6] Ignacio Cascudo, Ronald Cramer, and Chaoping Xing. Torsion Limits and Riemann-Roch Systems for Function Fields and Applications. *IEEE Trans. Information Theory*, 60(7):3871–3888, 2014.

[7] Ignacio Cascudo, Ivan Damgård, Bernardo David, Nico Döttling, and Jesper Buus Nielsen. Rate-1, Linear Time and Additively Homomorphic UC Commitments. In *Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part III*, pages 179–207, 2016.

[8] Ignacio Cascudo, Ivan Damgård, Bernardo Machado David, Irene Giacomelli, Jesper Buus Nielsen, and Roberto Trifiletti. Additively Homomorphic UC Commitments with Optimal Amortized Overhead. In *Public-Key Cryptography - PKC 2015 - 18th IACR International Conference on Practice and Theory in Public-Key Cryptography, Gaithersburg, MD, USA, March 30 - April 1, 2015, Proceedings*, pages 495–515, 2015.

[9] David Chaum, Claude Crépeau, and Ivan Damgård. Multiparty unconditionally secure protocols (extended abstract). In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 11–19, 1988.

[10] Hao Chen and Ronald Cramer. Algebraic geometric secret sharing schemes and secure multi-party computations over small fields. In *Advances in Cryptology - CRYPTO 2006, 26th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2006, Proceedings*, pages 521–536, 2006.

[11] Alain Couvreur, Ayoub Otmani, and Jean-Pierre Tillich. Polynomial time attack on wild mceliece over quadratic extensions. *IEEE Trans. Information Theory*, 63(1):404–427, 2017.

[12] Ronald Cramer, Ivan Damgård, and Ueli M. Maurer. General secure multi-party computation from any linear secret-sharing scheme. In *Advances in Cryptology - EUROCRYPT 2000, International Conference on the Theory and Application of Cryptographic Techniques, Bruges, Belgium, May 14-18, 2000, Proceeding*, pages 316–334, 2000.

[13] Ronald Cramer, Ivan Damgård, and Jesper Buus Nielsen. *Secure Multiparty Computation and Secret Sharing.* Cambridge University Press, 2015.

[14] Ivan Damgård, Bernardo Machado David, Irene Giacomelli, and Jesper Buus Nielsen. Compact VSS and efficient homomorphic UC commitments. In *Advances in Cryptology - ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014, Proceedings, Part II*, pages 213–232, 2014.

[15] Iwan M. Duursma and Ralf Kötter. Error-locating pairs for cyclic codes. *IEEE Trans. Information Theory*, 40(4):1108–1121, 1994.

[16] Tore Kasper Frederiksen, Thomas P. Jakobsen, Jesper Buus Nielsen, and Roberto Trifiletti. On the complexity of additively homomorphic UC commitments. In *Theory of Cryptography - 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part I*, pages 542–565, 2016.

[17] Juan A. Garay, Yuval Ishai, Ranjit Kumaresan, and Hoeteck Wee. On the complexity of UC commitments. In *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*, pages 677–694, 2014.

[18] Irene Giacomelli. *New Applications of Secret-Sharing in Cryptography.* PhD thesis, Aarhus University, Denmark, October 2016.

[19] Markus Grassl. Bounds on the minimum distance of linear codes and quantum codes. Available at http://www.codetables.de. Accessed on 01 feb. 2017.

[20] OEIS Foundation Inc. The on-line encyclopedia of integer sequences, https://oeis.org/.

[21] Ralf Kötter. A unified description of an error locating procedure for linear codes. In *Proceedings of Algebraic and Combinatorial Coding Theory, Voneshta Voda, Bulgaria*, pages 113–117, 1992.

[22] Jacobus Hendricus van Lint. *Introduction to Coding Theory.* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1982.

[23] Diego Mirandola. Schur products of linear codes: a study of parameters. Master's thesis, Université de Bordeaux 1 and Stellenbosch University, 2012.

[24] Diego Mirandola and Gilles Zémor. Critical Pairs for the Product Singleton Bound. *IEEE Trans. Information Theory*, 61(9):4928–4937, 2015.

[25] Ruud Pellikaan. On decoding by error location and dependent sets of error positions. *Discrete Mathematics*, 106:369 – 381, 1992.

[26] Hugues Randriambololona. Asymptotically good binary linear codes with asymptotically good self-intersection spans. *IEEE Trans. Information Theory*, 59(5):3038–3045, 2013.

[27] Hugues Randriambololona. An upper bound of Singleton type for componentwise products of linear codes. *IEEE Trans. Information Theory*, 59(12):7936–7939, 2013.

[28] Hugues Randriambololona. On products and powers of linear codes under componentwise multiplication. *Contemp. Math.*, 637:3–78, 2015.

[29] Rudolf Schürer and Wolfgang Ch. Schmid. *MinT: A Database for Optimal Net Parameters.* In Harald Niederreiter, Denis Talay (eds) Monte Carlo and Quasi-Monte Carlo Methods 2004. Springer, Berlin, Heidelberg. Database available at http://mint.sbg.ac.at/. Accessed on 01 Feb. 2017.

[30] Igor E. Shparlinski, Michael A. Tsfasman, and Serge G. Vladut. Curves with many points and multiplication in finite fields. Proceedings of the International Workshop held in Luminy, France, June 1721, 1991. In *Coding Theory and Algebraic Geometry.*, pages 145–169, 1992.

A Proof of Proposition 5.3

PROOF. The case $s = k$ is Lemma 4.4, so we assume $s \leq k-1$, which simplifies the notation. Let $v := t + u \mod n$, and remember $n = q^k - 1$. Since the restricted weights are invariant of cyclotomic cosets, we can assume without loss of generality that the maximum in the definition of $w_q^{(s)}(v)$ is attained for the set of s least significant digits. That is, if we consider v as an integer in $\{0, \ldots, q^k - 2\}$ and write $v = v' + v''q^s$, where $0 \leq v' \leq q^s - 1$ and $0 \leq v'' \leq q^{k-s} - 1$, then we are assuming $w_q^{(s)}(v) = w_q(v')$. We also write $t = t' + t''q^s$, $u = u' + u''q^s$, where $0 \leq t', u' \leq q^s - 1$ and $0 \leq t'', u'' \leq q^{k-s} - 1$. Note that $w_q(t') \leq w_q^{(s)}(t)$ and $w_q(u') \leq w_q^{(s)}(u)$.
We now need to split the proof in different cases, according to whether $t + u$ (summed over the integers) is smaller than, equal to or larger than $q^k - 1$.

Case 1. $t + u \leq q^k - 2$.
In this case $v = t + u$ and therefore $t' + u' = v' + \epsilon q^s$ where $\epsilon = 0$ or 1. Now clearly $w_q^{(s)}(v) = w_q(v') \leq w_q(v') + \epsilon = w_q(v' + \epsilon q^s) = w_q(t' + u') \leq w_q(t') + w_q(u') \leq w_q^{(s)}(t) + w_q^{(s)}(u)$, where the inequality $w_q(t' + u') \leq w_q(t') + w_q(u')$ comes from Lemma 4.4.

Case 2. $t + u = q^k - 1$.
In this case $v = 0$ and the statement follows trivially since $w_q^{(s)}(v) = 0$ and all weights are non-negative.

Case 3. $t + u \geq q^k$.
This case is more involved. Note $v = t + u + 1 - q^k$. Then the q-ary representation of v is obtained by computing the one for $t + u + 1$ and then erasing the 1 in the position corresponding to q^k. It is easy to see then that $t' + u' + 1 = v' + \epsilon q^s$, where $\epsilon = 0$ or 1. We need to further split the proof in these two cases.

Case 3a. $t + u \geq q^k$ and $t' + u' + 1 \geq q^s$.
In this case (since also $t' + u' + 1 \leq 2q^s - 1$) it holds that $t' + u' + 1 = v' + q^s$. Then $w_q^{(s)}(v) + 1 = w_q(v') + 1 = w_q(v' + \epsilon q^s) = w_q(t' + u' + 1) \leq w_q(t') + w_q(u') + 1 \leq w_q^{(s)}(t) + w_q^{(s)}(u) + 1$ and hence $w_q^{(s)}(v) \leq w_q^{(s)}(t) + w_q^{(s)}(u)$.

Case 3b. $t + u \geq q^k$ and $t' + u' + 1 \leq q^s - 1$.
In this case it holds that $t' + u' + 1 = v'$ and we can only show the inequality $w_q(v') \leq w_q(t') + w_q(u') + 1$. In fact, the inequality is tight, i.e., there are cases in which $w_q(v') = w_q(t') + w_q(u') + 1$.
In order to show the theorem, we need to argue the following:

Claim 1. Under the restrictions of case 3b., it holds that $w_q^{(s)}(t) + w_q^{(s)}(u) \geq w_q(t') + w_q(u') + 1$.

Once we prove this claim, the proof is finished, since in that case $w_q^{(s)}(v) = w_q(v') \leq w_q(t') + w_q(u') + 1 \leq w_q^{(s)}(t) + w_q^{(s)}(u)$.

Proof of Claim 1. Clearly $w_q^{(s)}(t) \geq w_q(t')$ and $w_q^{(s)}(u) \geq w_q(u')$. So we need to rule out either $w_q^{(s)}(t) = w_q(t')$ or $w_q^{(s)}(u) = w_q(u')$. Write $t = \sum_{i=0}^{k-1} t_i q^i$, $u = \sum_{i=0}^{k-1} u_i q^i$. Then proving the claim amounts to showing the existence of $j \in \{1, \ldots, k+1\}$ such that either $\sum_{i=0}^{s-1} t_{j+i} > \sum_{i=0}^{s-1} t_i (= w_q(t'))$, or $\sum_{i=0}^{s-1} u_{j+i} > \sum_{i=0}^{s-1} u_i (= w_q(u'))$, where the sums $j+i$ are modulo k.

Suppose towards a contradiction, that this is not true, and hence $w_q^{(s)}(t) + w_q^{(s)}(u) = w_q(t') + w_q(u')$. We now make the following claim.

Claim 2. Under the restrictions of case 3b. and assuming $w_q^{(s)}(t) + w_q^{(s)}(u) = w_q(t') + w_q(u')$, we have $t_{k-j} + u_{k-j} = t_{s-j} + u_{s-j} = q-1$ for all $j \in \{1, \dots, s\}$.

Assuming claim 2, we quickly arrive at a contradiction, since in fact in that case $t' + u' = (q-1)(1 + q + \cdots + q^{s-1}) = q^s - 1$ but we are assuming $t' + u' + 1 \le q^s - 1$. This shows claim 1. Hence we are left to prove claim 2.

Proof of Claim 2. We argue by induction on j.
For the case $j = 1$, note that the condition $t' + u' + 1 \le q^s - 1$ clearly implies that $t_{s-1} + u_{s-1} \le q - 1$. On the other hand since $t = t_{k-1}q^{k-1} + \tilde{t}$, $u = u_{k-1}q^{k-1} + \tilde{u}$ with $\tilde{t}, \tilde{u} \le q^{k-1} - 1$ the condition $t + u \ge q^k$ implies that $t_{k-1} + u_{k-1} \ge q - 1$. However, if $t_{s-1} + u_{s-1} < t_{k-1} + u_{k-1}$, then $\sum_{i=0}^{s-1} t_i + \sum_{i=0}^{s-1} u_i < \sum_{i=0}^{s-1} t_{k-1+i} + \sum_{i=0}^{s-1} u_{k-1+i}$, and we reach a contradiction. So the only possibility is $t_{s-1} + u_{s-1} = t_{k-1} + u_{k-1} = q - 1$.
Now, assume $t_{k-j} + u_{k-j} = t_{s-j} + u_{s-j} = q - 1$ is true for all $j < j_*$. Thus we have $q^s - 1 \ge t' + u' \ge (q-1)(q^{s-1} + \cdots + q^{s-j_*+1}) + (t_{s-j_*} + u_{s-j_*})q^{s-j_*}$. Then it is easy to see that this implies $t_{s-j_*} + u_{s-j_*} \le q - 1$. On the other hand $q^k \le t + u < (q-1)(q^{k-1} + \cdots + q^{k-j_*+1}) + (t_{k-j_*} + u_{k-j_*})q^{k-j_*} + 2q^{k-j_*} = q^k - q^{k-j_*+1} + (t_{k-j_*} + u_{k-j_*} + 2)q^{k-j_*}$. This implies $t_{k-j_*} + u_{k-j_*} > q - 2$, hence $t_{k-j_*} + u_{k-j_*} \ge q - 1$. Finally by the assumption $w_q^{(s)}(t) + w_q^{(s)}(u) = w_q(t') + w_q(u')$, we have $\sum_{i=0}^{s-1} t_i + \sum_{i=0}^{s-1} u_i \ge \sum_{i=0}^{s-1} t_{k-j_*+i} + \sum_{i=0}^{s-1} u_{k-j_*+i}$. However, taking into account the induction assumption $t_{k-j} + u_{k-j} = t_{s-j} + u_{s-j} = q - 1$ for $j < j_*$ and after removing terms that appear in both sides, we have $t_{s-j_*} + u_{s-j_*} \ge t_{k-j_*} + u_{k-j_*}$. Then necessarily $t_{s-j_*} + u_{s-j_*} = t_{k-j_*} + u_{k-j_*} = q - 1$ and we have completed the induction and shown claim 2.

△