UDC 003.26, 519.725, 519.176 DOI 10.17223/20710410/53/5

# CHARACTERISTICS OF HADAMARD SQUARE OF SPECIAL REED — MULLER SUBCODES

V. Vysotskaya

*Moscow State University, Moscow, Russia*
*JSC "NPK Kryptonite", Moscow, Russia*

**E-mail:** vysotskaya.victory@gmail.com

The existence of some structure in a code can lead to the decrease of security of the whole system built on it. Often subcodes are used to "disguise" the code as a "general-looking" one. However, the security of subcodes, whose Hadamard square is equal to the square of the original code, can be reduced to the security of this code. The paper finds the limiting conditions on the number of vectors of degree $r$ whose removing retains this weakness for Reed — Muller subcodes and, accordingly, conditions for it to vanish. For $r = 2$ the exact structure of all resistant subcodes has been found. For an arbitrary code $\mathrm{RM}(r, m)$, the desired number of vectors to remove for providing the security has been estimated from both sides. Finally, the ratio of subcodes with Hadamard square unequal to the square of the original code has been proved to tend to zero if additional conditions on the codimension of the subcode and the parameter $r$ are imposed and $m \to \infty$. Thus, the implementation of checks proposed in the paper helps to immediately filter out some insecure subcodes.

**Keywords:** *post-quantum cryptography, code-based cryptography, Reed — Muller codes, Reed — Muller subcodes, Hadamard product, McEliece cryptosystem.*

## 1. Introduction

The security of most standardized cryptographic algorithms used all around the world is based on the complexity of several number-theoretical problems. They usually are the discrete logarithm or factorization problem. However, in 1994 P. Shor showed [1] that quantum computers could break all schemes constructed in this way. And in 2001 the Shor's algorithm was implemented on a 7-qubit quantum computer. Since then various companies have been actively developing more powerful quantum computers. Progress in this area poses a real threat to modern public-key cryptography.

There are several approaches to build post-quantum cryptographic schemes. One approach is to use error-correcting codes. No successful quantum-computer attacks on "hard" problems from this area are known. Classical examples of code-based schemes are the McEliece cryptosystem [2] and the Niederreiter cryptosystem [3], which are equivalent in terms of security.

The interest in code-based schemes as post-quantum can be noticed while analyzing the works submitted to the contest for prospective public-key post-quantum algorithms which was announced in 2016 by the US National Institute of Standards and Technology (NIST) [4]. The algorithms that win this contest will be accepted as US national standards. 21 of 69 applications filed (that is, almost a third of all works) were based on coding theory. Despite the fact that some of them were attacked, it seems that this approach looks quite promising and deserves further study and development. This interest is also traced in Russian cryptography. Code-based schemes were chosen by the Technical Committee for

Standardization "Cryptographic and Security Mechanisms" (TC 26) as one of directions in developing draft Russian national standards of post-quantum cryptographic algorithms.

When one is facing the challenge to synthesize a new code-based scheme, the first thing to think about is the choice of basic code. Some schemes do not specify the code, thus leaving it to the discretion of the user. Such schemes are usually more reliable since their security is often directly reduced to NP-complete problems. Most often, these problems are decoding and syndrome decoding. However, choosing a special code also has some advantages. For example, such codes provide asymmetric complexity in solving the decoding problem for the legal user and adversary. In addition, due to the structure of the code, the sizes of the public keys can be significantly reduced.

However, the structure can also cause a significant decrease in security of the code, therefore one of the most important tasks is to "disguise" the code as a "general-looking" one. One solution is to use subcodes. This approach allows to "destroy" the structure of the code, retaining the ability to work with the result in mostly the same way as with the original one. Nevertheless, it is worth considering that many of proposed systems based on subcodes turned out to be vulnerable. So in [5, 6] C. Wieschebrink built efficient attacks on some special cases of the Berger — Loidreau cryptosystem [7], which is based on subcodes of the Reed — Solomon code. The McEliece cryptosystem based on subcodes of algebraic geometry codes was attacked in [8]. First version of digital signature pqsigRM [9] based on modified Reed — Muller codes, which was submitted at the NIST contest, was also attacked during the peer review.

One of the mechanisms for analyzing codes with a hidden structure is the use of the technique of Hadamard product of two codes. This method was used by M. Borodin and I. Chizhov [10] to improve Minder — Shokrollahi attack [11] on the McEliece cryptosystem based on Reed — Muller codes. In [12] this technique allowed Chizhov and Borodin to reduce the security of the cryptosystem on subcodes of Reed — Muller codes of codimension one to the security of the scheme on full codes. Recall that codimention means the number of vectors missing in the code basis. In [13] the distinguisher between random codes and Reed — Solomon codes using Hadamard product is described.

In this paper, the mentioned technique is used to analyze Reed — Muller subcodes in standard basis without restriction on codimension. The main question is: which Reed — Muller subcodes do not allow Chizhov — Borodin's approach. Since the reduction can be performed to a subcode whose Hadamard square coincides with the square of the original code, we look for conditions under which this equality ceases to hold. Codes obtaining these conditions will be called *unstable codes*, the others — *stable codes*. In addition, we compute the probability that a randomly chosen Reed — Muller subcode is unstable.

In Section 2, the exact structure of all stable subcodes of $\mathrm{RM}(2, m)$ is found. Thus, to provide the security, it is necessary to choose at least another subcode. To be sure that a subcode of $\mathrm{RM}(2, m)$ is unstable, it is sufficient to exclude $m+1$ monomials of degree 2 from it's standard basis. For an arbitrary Reed — Muller code $\mathrm{RM}(r, m)$, in Section 3 we estimate (both from the above and below) the number of vectors of degree $r$ that must be excluded from the basis of the code in order to distort its square. Finally, in Section 4 we show that the ratio of unstable subcodes tends to zero (as $m \to \infty$) given some additional conditions on the codimension of the subcode and the parameter $r$. Thus, it is not enough to choose an arbitrary Reed — Muller subcode when synthesizing a real scheme. It is necessary to check the property formulated below as Proposition 4. At the same time subcodes satisfying this property require additional consideration since they may have some special structure.

## 2. The structure of stable $\mathbf{RM}(2, m)$ subcodes

Recall that *Reed — Muller code* $\mathrm{RM}(r, m)$ is the set of all Boolean functions $f$ in $m$ variables such that $\deg(f) \leqslant r$. Consider the code $\mathrm{RM}(1, m)$. We look for the minimum number of monomials $f_1, \ldots, f_{w(m,2)}$ of degree 2 such that the code

$$\operatorname{span}\big(\mathrm{RM}(1, m) \cup \{f_1, \ldots, f_{w(m,2)}\}\big) \tag{1}$$

is *stable*, i.e.,

$$\big(\operatorname{span}(\mathrm{RM}(1, m) \cup \{f_1, \ldots, f_{w(m,2)}\})\big)^2 = \mathrm{RM}(4, m). \tag{2}$$

Here, the squaring operation refers to the squaring of Hadamard. *Hadamard product* of two vectors is a vector obtained as a result of component-wise product of coordinates:

$$(a_1, \ldots, a_n) \circ (b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n),$$

and Hadamard product of two codes $A$ and $B$ is the span of all pairwise products of form $a \circ b$, where $a \in A$, $b \in B$. Codes that do not satisfy condition (2) we will denote *unstable*.

We consider Reed — Muller codes spanned by their standard basis. *The standard basis of the Reed — Muller code* $\mathrm{RM}(r, m)$ includes all monomials of $m$ variables of degree from 0 to $r$ inclusively, i.e.,

$$1,\ x_1,\ x_2,\ \ldots,\ x_m,\ x_1x_2,\ \ldots,\ x_{m-1}x_m,\ \ldots,\ x_1 \ldots x_r,\ \ldots,\ x_{m-r-1} \ldots x_m.$$

Obviously, after finding the minimal number $w(m, 2)$ of monomials $f_i$, one can also answer another question: what is the maximum number $q(m, 2)$ of monomials of degree 2 that can be removed from the basis of the code $\mathrm{RM}(2, m)$ so that the code

$$\operatorname{span}\big(\{1, x_1, x_2, \ldots, x_m, x_1x_2, \ldots, x_{m-1}x_m\} \setminus \{g_1, \ldots, g_{q(m,2)}\}\big)$$

is still stable. The relation between these values is given by the following equality:

$$q(m, 2) = \binom{m}{2} - w(m, 2). \tag{3}$$

And so, after removing $q(m, 2) + 1 = \binom{m}{2} - w(m, 2) + 1$ basis vectors, one gets an unstable code. Therefore, we will not dwell on this issue separately.

Now let us proceed to the graph interpretation of this problem. We match a subcode $\mathcal{A} \subset \mathrm{RM}(2, m)$ with a graph $G = (V, E)$ with vertex set $V = \{x_1, \ldots, x_m\}$ and edge set $E$; $\{x_i, x_j\} \in E \Leftrightarrow x_i x_j \in \mathcal{A}$.

Let us denote by $\deg(v)$ the degree of vertex $v$ in the graph. As vertices are monomials, it can be a little embarrassing, but we never use $\deg(\cdot)$ to refer to the degree of a polynomial.

We will say that a graph with $m$ vertices *has the property* $P$ if

1) the degree $\deg(v)$ of any vertex $v$ is not less than $m - 3$;
2) if $\deg(v) = m - 3$ and $\{v, u_1\} \notin E$, $\{v, u_2\} \notin E$, then $\{u_1, u_2\} \in E$.

The case $\deg(x_1) = m - 3$ is shown in Fig. 1, where lines denote graph edges.

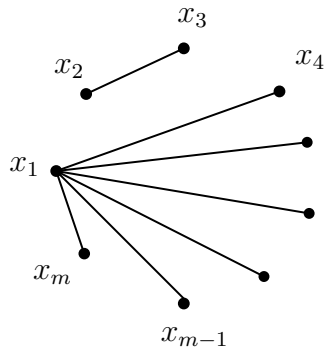


Fig. 1. Case $\deg(x_1) = m - 3$

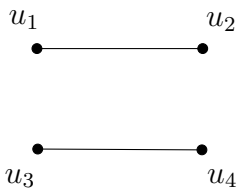


Fig. 2. Graph $H$

**Theorem 1.** For any $m \geqslant 4$ a subcode of $\mathrm{RM}(2, m)$ of the form (1) is stable if and only if the property $P$ holds for the corresponding graph.

***Proof.*** Denote $G = (V, E)$ the graph corresponding to the subcode of form (1). Note that the condition (2) is equivalent to the condition that any induced subgraph of $G$ with 4 vertices has a subgraph isomorphic to the graph $H$ shown in Fig. 2. The edges $\{u_1, u_2\}$ and $\{u_3, u_4\}$ correspond to degree-2 monomial used to produce the monomial $u_1u_2u_3u_4$. Also, note that to show that the subcode (1) is stable it is enough to prove that any monomial of degree 4 can be represented as a product of two monomials from this code. Then the same is automatically true for all monomials of degree 3. Indeed, for any monomial $u_1u_2u_3$ at least one of monomials $u_1u_2$, $u_1u_3$ or $u_2u_3$ lie in the code. Otherwise, no monomial of form $u_1u_2u_3v$ could be obtained after squaring. Degree-1 monomials are in the code by the definition.

To prove the necessity, we fix any vertex $v$. If any three incident edges $\{v, u_j\}$ for $j = 1, 2, 3$ are missing, then the induced subgraph on vertices $v, u_1, u_2, u_3$ would not have the required subgraph $H$. The contradiction proves that $\deg(v) \geqslant m - 3$. If, however, $\deg(v) = m - 3$ and $\{v, u_1\} \notin E$, $\{v, u_2\} \notin E$, then $\{u_1, u_2\} \in E$, as otherwise none of the induced 4-vertex subgraphs containing vertices $v, u_1$ and $u_2$ will have the required subgraph. Thus, the property $P$ holds.

The sufficiency: fix any induced subgraph with 4 vertices (let us denote them $v, u_1, u_2$ and $u_3$). Note that it has the property $P$ for $m = 4$. If vertex $v$ has degree 1, i.e., $\{v, u_1\} \in E$, but $\{v, u_2\} \notin E$, $\{v, u_3\} \notin E$, then by the property $P$ it follows that $\{u_2, u_3\} \in E$. Thus, we have edges $\{v, u_1\}$ and $\{u_2, u_3\}$ necessary for the $H$-isomorphic subgraph.

If all 4 vertices in the subgraph have degree at least 2, then there is a simple cycle of length 3 or 4. If it has length 4, the presence of $H$-isomorphic subgraph is obvious. Otherwise, we have a triangle $\{u_1, u_2, u_3\}$ and, moreover, the fourth vertex $v$ has degree at least 2. Assume (without loss of generality) that $\{v, u_1\} \in E$, then for $H$-isomorphic subgraph we can take the edges $\{v, u_1\}$ and $\{u_2, u_3\}$. ■

From Theorem 1, the minimum number of edges is obtained if a graph has the property $P$ and the degree of each vertex is $m - 3$. It remains to describe such graphs.

**Proposition 1.** Assume $m \geqslant 4$. If the property $P$ holds for some graph $G$ with $m$ vertices such that the degree of each vertex is $m - 3$, then the complementary graph $\overline{G}$ is a union of cycles of length at least 4.

***Proof.*** Since the degree of each vertex of graph $G$ is $m - 3$, the degree of each vertex of $\overline{G}$ is 2. Moreover, from the second item of the property $P$ follows that if $\overline{G}$ contains edges

$\{v, u_1\}$ and $\{v, u_2\}$, then it does not contain the edge $\{u_1, u_2\}$. So graph $\overline{G}$ is triangle-free. Choose an arbitrary vertex $u_1$. It is not isolated, therefore, one can select a vertex adjacent to it. Let us call it $u_2$. As $\deg(u_2) = 2$, there exists some adjacent vertex $u_3 \neq u_1$. Continue in this way until $u_j$ coincides with one of $u_1, \ldots, u_{j-1}$. Note that $u_j$ cannot coincide with $u_i$ for $i > 1$ as it would mean that $\deg(u_i) \geqslant 3$. Thus, $u_1, \ldots, u_{j-1}$ form a simple cycle. Its length is at least 4, as $\overline{G}$ is triangle-free. ■

Thus, we have described the structure of the graph corresponding to the minimal stable subcode of form (1). Now let us describe the complete structure of such codes. Let us denote a *bamboo graph* a tree without branching (having no vertices of degree greater than 2).

**Proposition 2.** Assume $m \geqslant 4$. If the property $P$ holds for some graph $G$ with $m$ vertices, then the complementary graph $\overline{G}$ is a union of cycles of length at least 4 and bamboo graphs.

***Proof.*** We proceed as in Proposition 1 and try to find a cycle in $\overline{G}$. But we can stop in a vertex of degree 1, thus obtaining a bamboo graph. Isolated vertices are bamboo graphs by definition. ■

**Corollary 1.** For any $m \geqslant 4$, it holds that

$$w(m, 2) = m(m-3)/2.$$

***Proof.*** As it was already mentioned after Theorem 1, we need to consider the subcodes corresponding to graphs with property $P$ where degree of each vertex is $m - 3$. From Proposition 1 it follows that $\overline{G}$ has exactly $m$ edges. Thus, $G$ has at least $\binom{m}{2} - m = m(m-3)/2$ edges. Moreover, it means that after removing any $m$ edges from a complete graph (corresponding to the full Reed — Muller code) we still obtain a stable code. ■

Note that, according to (3), removing $m + 1$ or more monomials of degree 2 from the basis of the code $\mathrm{RM}(2, m)$ leads to an unstable code.

## 3. Lower and upper bounds for minimal stable RM$(r, m)$ subcode sizes

In this Section we carry out argument for $r > 2$. That is, we look for the minimum number $w(m, r)$, such that the code

$$\mathrm{span}\big(\mathrm{RM}(r-1, m) \cup \{f_1, \ldots, f_{w(m,r)}\}\big) \tag{4}$$

is stable. Here, $f_i$ is a monomial of degree $r$. We match a subcode $\mathcal{A} \subset \mathrm{RM}(r, m)$ with a hypergraph $G = (V, E)$ with vertex set $V = \{x_1, \ldots, x_m\}$; an $r$-edge $\{x_{i_1}, \ldots, x_{i_r}\}$ is in $E$ if and only if $x_{i_1} \ldots x_{i_r} \in \mathcal{A}$. In the general case, the condition similar to having an $H$-isomorphic subgraph in each 4-vertex induced subgraph is equivalent to condition of the code (4) being stable. Namely, each set of $2r$ vertices must be covered by two disjoint $r$-edges. Let us denote a graph satisfying this condition by *stable graph*. Note about covering monomials of lower degrees is the same as in the case of $r = 2$.

We can also extend relation (3) from Section 2 as:

$$q(m, r) = \binom{m}{r} - w(m, r).$$

And again we will not dwell on this issue separately.

We will use terms "graph" and "hypergraph" interchangeably. Denote $w(r, m)$ the minimal number of degree-$r$ monomials needed to make subcode (4) stable, or, alternatively, minimal number of edges in a stable $r$-hypergraph with $m$ vertices.

**Proposition 3.** For any natural $r$ and $m \geqslant 2r$, it holds that

$$w(m,r) \geqslant \binom{m}{2r} \Big/ \binom{m-r}{r}.$$

***Proof.*** Note that any set of $2r$ vertices in a stable graph contains at least one edge. Moreover, any edge is contained in exactly $\binom{m-r}{r}$ such sets. Thus, total number of edges multiplied by $\binom{m-r}{r}$ is at least number of all sets of $2r$ vertices, which is $\binom{m}{2r}$. This gives the necessary bound. ■

**Corollary 2.** Any stable graph contains at least $1/\binom{2r}{r}$ edges of a complete graph.

***Proof.*** The total possible number of $r$-edges in a graph with $m$ vertices is $C_m^r$. Then

$$\frac{\binom{m}{2r}}{\binom{m-r}{r}\binom{m}{r}} = \frac{(r!)^2}{(2r)!} = 1/\binom{2r}{r}.$$

Corollary 2 is proven. ■

This lower bound can be improved by the following theorem.

**Theorem 2.** For any natural $r$ and $m \geqslant 2r$, it holds that

$$w(m,r) \geqslant \frac{1}{2}\left(\sqrt{(\gamma+1)^2 + 8\binom{m}{2r}} + \gamma + 1\right), \text{ where } \gamma = \sqrt{\sum_{u=\max\{1,3r-m\}}^{r-1} \binom{r}{u}}.$$

***Proof.*** Fix smallest set of edges $E$ such that every $2r$ vertices are covered by two disjoint edges from $E$. By definition, $|E| = w(m,r)$.

Fix any edge $e \in E$. Denote $E_e$ the set of edges from $E$ that intersect $e$ and $P_e$ — the set of unordered pairs $\{e', e''\}$, $e', e'' \in E_e$. Each pair $\{e', e''\}$ corresponds to the subset $B \subset e$, $B = (e' \cup e'') \cap e$. In the similar manner, each edge in $E_e$ corresponds to the subset $B = e' \cap e$. On the other hand, let us fix any subset $B \subset e$ of size

$$\max\{1, 3r-m\} \leqslant |B| \leqslant r-1. \tag{5}$$

As $|B| \geqslant 3r-m$, we have $|V \setminus e| + |B| \geqslant 2r$, and thus there exists a set $S$ such that $|S| = 2r$ and $S \cap e = B$. By the assumption on the edge set $E$, it contains a pair of edges covering $S$. Let us denote these edges $e'$ and $e''$. There are two possible cases: either both $e'$ and $e''$ intersect $e$ or only one of them does. Thus, we can match the subset $B$ with an element of $E_e \cup P_e$. Note that despite that the subset $B$ can match several elements of $E_e \cup P_e$, the inverse mapping is single-valued, as we explained earlier. Thus, we can write

$$|P_e| + |E_e| \geqslant \sum_{u=\max\{1,3r-m\}}^{r-1} \binom{r}{u} = \gamma^2,$$

where the right-hand side is the number of all subsets $B \subset e$ satisfying (5).

Obviously, $|P_e| = \binom{|E_e|}{2}$, and thus

$$\binom{|E_e|}{2} + |E_e| \geqslant \gamma \;\Leftrightarrow\; |E_e|^2 + |E_e| \geqslant 2\gamma^2.$$

From this inequality it follows that $|E_e| \geqslant \gamma$ (technically, we use the fact that $\gamma$ can not lie in the interval $(0, 1)$ following from its definition as a square root of 0 or a natural number).

Now we can estimate the cardinality of the set $P$ of all unordered pairs $\{e', e''\}$ of edges from $E$. Denote $\widehat{P}$ the set of all *disjoint* unordered pairs of edges from $E$. It is clear that

$$P = \widehat{P} \cup \bigcup_{e \in E} \big\{\{e', e\} : e' \in E_e\big\}$$

and, moreover,

$$|P| = |\widehat{P}| + \frac{1}{2}\sum_{e \in E} |E_e|,$$

as $\widehat{P}$ is disjoint with the other set and in the union over all $e \in E$ we count each intersecting unordered pair exactly twice.

From the property that edges from $E$ cover each set of size $2r$ we conclude that $|\widehat{P}| \geqslant \binom{m}{2r}$. Thus,

$$|P| - \frac{1}{2}\sum_{e \in E} |E_e| \geqslant \binom{m}{2r}.$$

As $|P| = \binom{|E|}{2} = \binom{w(m,r)}{2}$, we can write

$$\binom{w(m,r)}{2} - \frac{w(m,r)}{2}\gamma \geqslant \binom{m}{2r}.$$

Solving the square inequality

$$w(m,r)^2 - w(m,r)(\gamma + 1) - 2\binom{m}{2r} \geqslant 0,$$

we obtain the state of the theorem. ■

Now let us proceed to the proof of the upper bound. Let us fix the set of maximal size $\mathcal{S}$ consisting of sets $S_i \subset V$ of size $2r$ such that

$$\max_{i,j} |S_i \cap S_j| \leqslant h.$$

**Lemma 1.** If $h < r/3$, then for any set $Q \notin \mathcal{S}$, $|Q| = 2r$, there are at most two sets from $\mathcal{S}$ such that their intersection with $Q$ have size at least $r$.

***Proof.*** Assume that $Q$ intersects with at least 3 sets such that intersection size is at least $r$. Without loss of generality we assume that the sets are $S_1, S_2$ and $S_3$. Let us denote $Q \cap S_1 = A_1$, $Q \cap S_2 = A_2$, $Q \cap S_3 = A_3$. Since $|Q| = 2r$, then it is obvious that $|A_1 \cup A_2 \cup A_2| \leqslant 2r$. On the other hand, according to the inclusion-exclusion formula,

$$|A_1 \cup A_2 \cup A_2| \geqslant |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3|.$$

Then

$$\sum_{i=1}^{3} |A_i| \leqslant 2r + 3h.$$

By condition $|A_i| \geqslant r$ for any $i \in \{1, 2, 3\}$, therefore

$$\sum_{i=1}^{3} |A_i| \geqslant 3r.$$

Whence $3r \leqslant 2r + 3h$ and $h \geqslant r/3$, which contradicts the condition of the lemma. ■

Let us find the maximum possible number of edges that can be removed from the complete graph using the above arguments such that the graph remains stable.

**Theorem 3.** For any natural $r \geqslant 2$, $m \geqslant 2r$, and $h < r/3$,

$$w(m, r) \leqslant \binom{m}{r} - T(r, m, h)\left(\binom{2r}{r} - 2\right),$$

where

$$T(r, m, h) = \max\big\{t : \exists S_1, \ldots, S_t \big(S_i \subset \{1, \ldots, m\} \;\&$$
$$\& |S_i| = 2r \;\&\; (i \neq j \Rightarrow |S_i \cap S_j| \leqslant h),\; i, j \in \{1, \ldots, t\}\big)\big\}.$$

***Proof.*** Note that two disjoint $r$-edges are sufficient to cover a set of $2r$ vertices. Thus, it is possible to remove $\delta = \left(\binom{2r}{r} - 2\right)$ $r$-edges from the complete graph on the $2r$ vertices and preserve the stability of it. Obviously, no more edges can be removed.

Suppose that $\delta$ edges are removed from each set from $\mathcal{S}$ so that all of them are covered by at least two $r$-edges. It remains to verify that there exists a similar cover for *any* set of $2r$ vertices. Since by construction we can certainly cover any set $S_i$, we will prove that we can also cover any set $Q \notin S$, $|Q| = 2r$.

Note that if the cardinality of the intersection with some $S_i$ does not exceed $(r - 1)$, then removing edges in it does not affect the number of edges in $Q$. At the same time, according to Lemma 1, for $h < r/3$ any set of size $2r$ can have intersection of size at least $r$ with no more than two sets from $\mathcal{S}$. If there is only one such set, say $S_1$, then we have two cases:

1) $|Q \cap S_1| = 2r - 1$. In this case there exists some edge $e_1 \in Q \cap S_1$ not containing vertex $v$, $\{v\} = S_1 \setminus Q$ (as $S_1$ must be covered by two disjoint edges). Thus, we can take $e_2 = Q \setminus e_1$ (note that $e_2 \in E$ as we have removed only edges contained inside sets $S_i$), and $\{e_1, e_2\}$ form the disjoint cover of $Q$.
2) $|Q \cup S_1| < 2r - 1$. In this case there are at least two vertices $v_1$ and $v_2$ inside $Q \setminus S_1$ and the cover can be formed using any two disjoint edges $e_1, e_2 \subset Q$ such that $v_1 \in e_1$, $v_2 \in e_2$.

Now consider the case when there are exactly two sets $S_1$ and $S_2$ intersecting with $Q$ at no less than $r$ vertices. Assume that $|A_1| > r + h$. Then, according to the inclusion-exclusion formula, $|A_1 \cap A_2| = |A_1| + |A_2| - |A_1 \cup A_2| > r + h + r - 2r = h$, that contradicts with $|S_1 \cap S_2| \leqslant h$. Thus, $r \leqslant |A_i| \leqslant r + h$ for $i \in \{1, 2\}$. So there are at most $2\binom{r+h}{r}$ edges removed from $Q$. Note that

$$\binom{2r}{r} \Big/ \left(2\binom{r+h}{r}\right) = \frac{(2r)!\, r!\, h!}{2r!\, r!\, (r+h)!} = \frac{1}{2} \cdot \frac{2r}{r+h} \cdot \frac{2r-1}{r+h-1} \cdots \frac{r+1}{h+1}.$$

The last multiplier is greater than 2 for $r > 3$. For others holds

$$\frac{2r-i}{r+h-i} > \frac{2r}{r+h} > \frac{6}{4}.$$

Thus, for $r > 3$,

$$\binom{2r}{r} \Big/ \left(2\binom{r+h}{r}\right) > \frac{1}{2}\left(\frac{3}{2}\right)^{r-1} \cdot 2 > 2.$$

For $r = 2$ and 3 the inequality can be verified directly.

There are $\binom{2r}{r}\Big/2$ pairs of disjoint edges inside $Q$, so there remains at least one such pair after removal of $2\binom{r+h}{r} < \binom{2r}{r}\Big/2$ edges from $Q$.

So we have obtained a stable graph removing $\delta$ edges from a complete graph for each set from $\mathcal{S}$. It remains to remember that $|\mathcal{S}|$ is the number of sets of size $2r$ whose intersections are not larger than $h$ and thus $|\mathcal{S}| = T(r, m, h)$. ■

**Remark 1.** In [14], P. Erdös and J. Spencer introduce the value $\mathbf{m}(n, k, t)$ (typeset in bold to avoid confusion with $m$). It determines the size of the largest set of $k$-element subsets of $\{1, \ldots, n\}$ such that any two members of this set intersect in less than $t$ elements. Later V. Rödl [15] proves that

$$\lim_{n\to\infty} \mathbf{m}(n, k, t) = \binom{n}{t} \Big/ \binom{k}{t}.$$

That is, in our case, $\lim\limits_{m\to\infty} T(r, m, h) = \lim\limits_{m\to\infty} \mathbf{m}(m, 2r, \lfloor r/3 \rfloor) = \binom{m}{\lfloor r/3 \rfloor} \Big/ \binom{2r}{\lfloor r/3 \rfloor}$.

The upper bound can be also improved, but only empirically. We introduce an algorithm that on input set of vertices $V$ returns a set of edges $E \subset V \times V$ such that in resulting graph each set of $2r$ vertices is covered by two disjoint $r$-edges. Its simplified form is presented in Algorithm 1. You can find the full version at `https://github.com/VysotskayaVictory/StableGraphGreedy/`.

**Algorithm 1.** Greedy $r$-covering

**Input:** set of vertices $V$, edge cardinality $r$.
**Output:** set $E$ of $r$-edges that covers $V$.
**Function** ChooseEdge:
  $e := \varnothing$.
  **For** $i = 1, \ldots, r$:
    $V' := \{v \in V : v \text{ can be added to } e\}$;
    $v := \arg\min\limits_{v \in V'} \deg v$;  $e := e \cup \{v\}$.
  **Return** $e$.
**Function** Main:
  $E := \varnothing$.
  **While** $E$ does not cover all vertices of $V$:
    $e :=$ `ChooseEdge()`;  $E := E \cup \{e\}$.
  **Return** $E$.

To finalize Section 3, we can compare all obtained bounds. On Fig. 3 one can see two lower and two upper bounds that are obtained in Proposition 3, Theorem 2, Theorem 3 and Algorithm 1. The Figure shows that improved bounds are rather tight.

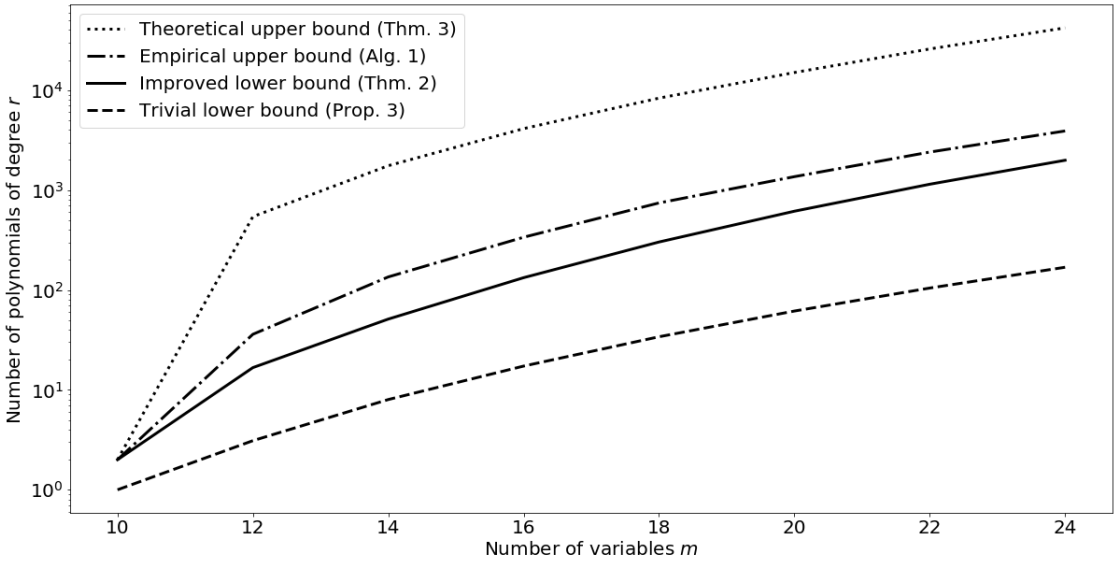


Fig. 3. Comparison of bounds

## 4. The ratio of unstable RM$(r, m)$ subcodes

We consider subcodes of the standard basis of the Reed — Muller code in which $\ell$ vectors are missing.

For the given parameter $s$ and the set $I = \{i_j : j = 1, \ldots, s\}$ we call unordered pairs $\{A, B\}$ *critical partition* if:

$$A \cap B = \varnothing,$$
$$A \cup B = I,$$
$$1 \leqslant |A|, |B| \leqslant r.$$

Then it is impossible to obtain the monomial $x_{i_1} \ldots x_{i_s}$ after squaring a subcode if and only if at least one element of each critical partition is removed. This follows from the fact that if this monomial is present in the square of the code, it should be formed of a pair $\{A, B\}$ from the appropriate critical partition. But by the hypothesis either $A$ or $B$ is absent. This argument proves the following proposition.

**Proposition 4.** A code is an unstable RM$(r, m)$ subcode if and only if at least one element from each critical partition for some monomial $x_{i_1} \ldots x_{i_s}$ is removed.

**Proposition 5.** For the given parameter $s$ and any set $I$ of size $s$ the number of critical partitions of $I$ is

$$v(s) = \frac{1}{2} \sum_{p=\max\{s-r,1\}}^{\min\{r,s-1\}} \binom{s}{p}.$$

***Proof.*** On the one hand, the sizes of the subsets must not exceed $r$. On the other hand, the partition must be non-trivial, that is, partitioning into an empty set and a set, coinciding with $I$, is unacceptable. Finally, when considering all partitions, each pair is counted twice. ■

Let us order in some way (say, lexicographically) the elements of each critical partition and then the critical partitions themselves. Now we consider any set $M$ consisting of elements of critical partitions and having the property that for every critical partition $M$ contains at least one element of this partition. We can encode $M$ with a string $\alpha \in \{1,2,3\}^{v(s)}$, where

$$\alpha_j = \begin{cases} 1 & \Leftrightarrow \text{ the 1st element of the } j\text{-th pair lies in } M, \\ 2 & \Leftrightarrow \text{ the 2nd element of the } j\text{-th pair lies in } M, \\ 3 & \Leftrightarrow \text{ both elements of the } j\text{-th pair lie in } M. \end{cases}$$

We will also write $M(\alpha)$ to denote the set corresponding to a given $\alpha \in \{1,2,3\}^{v(s)}$. It can be seen that

$$|M(\alpha)| = \#_1(\alpha) + \#_2(\alpha) + 2 \cdot \#_3(\alpha),$$

where $\#_c(\alpha)$ is the number of symbols $c$ in the string $\alpha$.

Let us denote $k = \sum\limits_{p=0}^{r} \binom{m}{p}$ the dimension of the original code (or the number of vectors in its standard basis). There are exactly two kinds of unstable subcodes: those containing monomial 1 and those not containing it. There are $\binom{k-1}{\ell-1}$ subcodes of the second kind.

Now we fix $s$, an index set $I$ of size $s$, and a string $\alpha \in \{1,2,3\}^{v(s)}$. Among the subcodes of the first kind there are

$$\binom{k-1-2v(s)}{\ell-|M(\alpha)|}$$

ones that has the property: among the monomials comprising critical partitions for $I$ exactly monomials from $M(\alpha)$ are absent. The reason is that we need to choose $\ell - |M(\alpha)|$ monomials from all monomials of degree more than 0 that do not comprise any critical partition (there are $k-1-2v(s)$ of them).

For a given $s$ there are $\binom{m}{s}$ variants of choosing index set $I$. But some codes may be counted several times. So we can consider the following theorem proved.

**Theorem 4.** The number of unstable $RM(r,m)$ subcodes is

$$\theta \leqslant \sum_{s=2}^{2r} \binom{m}{s} \cdot \sum_{\alpha \in \{1,2,3\}^{v(s)}} \binom{k-1-2v(s)}{\ell-|M(\alpha)|} + \binom{k-1}{\ell-1}.$$

**Theorem 5.** If $\ell = \text{const}$ and $r \geqslant 2\ell+1$, then the ratio of unstable $\mathrm{RM}(r,m)$ subcodes tends to zero as $m \to \infty$.

***Proof.*** Our goal is the asymptotic estimate of the probability of the event that after removing $\ell$ vectors from the standard basis of the code $\mathrm{RM}(r,m)$, the square of the resulting code will differ from $\mathrm{RM}(2r,m)$. The upper bound for it is $\theta \Big/ \binom{k}{\ell}$. We divide this bound into two parts and show the tendency to zero for each of them independently. For one of them it follows immediately from the fact that

$$\binom{k-1}{\ell-1} \Big/ \binom{k}{\ell} = \frac{\ell}{k} \xrightarrow[m\to\infty]{} 0,$$

since $k \to \infty$ as $m \to \infty$.

Now we consider the first part and denote it's numerator by $\gamma$. Notice that

$$\#_\alpha(1) + \#_\alpha(2) + 2 \cdot \#_\alpha(3) = |M(\alpha)| \geqslant v(s) = \#_\alpha(1) + \#_\alpha(2) + \#_\alpha(3).$$

Then the number of removed vectors that are elements of critical partitions for $s$ is $|M(\alpha)| \geqslant v(s)$ and the total number of removed vectors is $\ell$. That is, $v(s) \leqslant \ell$ and we can consider only parameters $s$ satisfying this condition. Then

$$2v(s) = \sum_{p=\max\{s-r,1\}}^{\min\{r,s-1\}} \binom{s}{p} \leqslant 2\ell. \tag{6}$$

We consider separately two cases. If $s \geqslant r+1$, we have $\min\{r, s-1\} = r$ and in the sum (6) there is the element $\binom{s}{r}$. Thus,

$$2\ell \geqslant 2v(s) \geqslant \binom{s}{r} \geqslant s.$$

The last inequality follows from the fact that

$$\binom{s}{r} = \frac{(r+1)}{2} \cdot \frac{(r+2)}{3} \cdot \ldots \frac{(s-1)}{r} \cdot \frac{s}{1}.$$

If, on the other hand, $s < r+1$, we have $\max\{s-r, 1\} = 1$ and there is the element $\binom{s}{1}$ in the sum (6). Hence

$$2\ell \geqslant 2v(s) \geqslant \binom{s}{1} = s.$$

So either way the inequality $s \leqslant 2\ell$ is satisfied.

We simplify the upper bound for $\gamma$ using this inequality and the monotonicity of the binomial coefficient $\binom{n}{k}$ with respect to the parameter $k$, which guarantees the increase of the value $\binom{n}{k}$ with the increase of $k$:

$$\sum_{s=2}^{2r} \binom{m}{s} \sum_{\alpha\in\{1,2,3\}^{v(s)}} \binom{k-1-2v(s)}{\ell-|M(\alpha)|} \leqslant \sum_{s=2}^{2\ell} \binom{m}{2\ell} \sum_{\alpha\in\{1,2,3\}^{v(s)}} \binom{k-1-2v(s)}{\ell-|M(\alpha)|} \leqslant$$
$$\leqslant 2\ell \binom{m}{2\ell} \max_{s\in[2,2\ell]} \left\{ \binom{k-1-2v(s)}{\ell-z} \cdot 3^{v(s)} \right\},$$

where $z = \min\limits_{\alpha\in\{1,2,3\}^{v(s)}} \{|M(\alpha)|\}$.

Note that $\ell = \text{const}$ and $3^{w(s)} \leqslant \text{const}$, since $s \leqslant 2\ell$, and $v(s) < 2^s$. The last is true by virtue of

$$2^s = (1+1)^s = \sum_{p=0}^{s} \binom{s}{k} > \frac{1}{2} \sum_{p=\max\{s-r,1\}}^{\min\{r,s-1\}} \binom{s}{p} = v(s).$$

These considerations, as well as the monotonicity of the binomial coefficient $\binom{n}{k}$ with respect to $n$ and the inequality $|M(\alpha)| \geqslant v(s)$, allow us to obtain the upper bound

$$\text{const} \cdot \binom{m}{2\ell} \binom{k}{\ell - v(s)} \leqslant \text{const} \cdot \binom{m}{2\ell} \binom{k}{\ell - 1} := \psi.$$

We proceed to the ratio estimation:

$$\gamma/\binom{k}{\ell} \leqslant \psi/\binom{k}{\ell} = \text{const}\cdot\binom{m}{2\ell}\binom{k}{\ell-1}\Big/\binom{k}{\ell} = \text{const}\cdot\binom{m}{2\ell}\cdot\ell/(k-\ell+1) =$$
$$= \text{const}\cdot\binom{m}{2\ell}/(k-\ell+1) \leqslant \text{const}\cdot\frac{m^{2\ell}}{2k}.$$

After tending $m$ to infinity we can claim that such $p = 2\ell + 1$ exists, that is, summand $\binom{m}{p} \geqslant m^p$ is an element of the sum representation of $k$. Then

$$\text{const}\cdot\frac{m^{2\ell}}{2k} \leqslant \text{const}\cdot\frac{m^{2\ell}}{m^{2\ell+1}} = \text{const}\cdot\frac{1}{m} \xrightarrow[m\to\infty]{} 0.$$

Theorem 5 is proven. ■

## Future research

More accurate estimates on the minimal stable code sizes for general case are still required, as are better estimates of the ratio of stable subcodes. In addition, an idea for future research could be to find an analogues of the obtained results for an arbitrary basis of the Reed — Muller code.

## Acknowledgments

The author thanks Ivan Chizhov for stating the problem and Lev Vysotsky for valuable help in preparing the text.

## REFERENCES

1. *Shor P. V.* Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM J. Computing, 1997, vol. 26, no. 5, pp. 1484–1509.
2. *McEliece R. J.* A public-key cryptosystem based on algebraic coding theory. DSN Progress Report, 1978, vol. 4244, pp. 114–116.
3. *Niederreiter H.* Knapsack-type cryptosystems and algebraic coding theory. Problems Control Inform. Theory, 1986, vol. 15, no. 2, pp. 159–166.
4. `https://csrc.nist.gov/Projects/post-quantum-cryptography/Post-Quantum-Cryptography-Standardization/Call-for-Proposals` — NIST Call for Proposals, 2016.
5. *Wieschebrink C.* An attack on a modified Niederreiter encryption scheme. LNCS, 2006, vol. 3958, pp. 14–26.
6. *Wieschebrink C.* Cryptanalysis of the Niederreiter public key scheme based on GRS subcodes. LNCS, 2010, vol. 6061, pp. 61–72.
7. *Berger T. P. and Loidreau P.* How to mask the structure of codes. Designs, Codes, Cryptogr., 2005, vol. 35, no. 1, pp. 63–79.
8. *Couvreur A., Marquez-Corbella I., and Pellikaan R.* Cryptanalysis of public-key cryptosystems that use subcodes of algebraic geometry codes. Coding Theory Appl.ications. CIM Ser. Math. Sci., 2015, vol. 3, pp. 133–140.
9. *Lee Y., Lee W., Kim Y. S., and No J.-S.* Modified pqsigRM: RM code-based signature scheme. IEEE Access, 2020, vol. 8, pp. 177506–177518.
10. *Borodin M. A. and Chizhov I. V.* Effective attack on the McEliece cryptosystem based on Reed — Muller codes. Discr. Math. Appl., 2014, vol. 24, no. 5, pp. 273–280.
11. *Minder L. and Shokrollahi A.* Cryptanalysis of the Sidelnikov cryptosystem. LNCS, vol. 4515, pp. 347–360.

12. *Chizhov I. V. and Borodin M. A.* Hadamard products classification of subcodes of Reed — Muller codes codimension 1. Discr. Math. Appl., 2020, vol. 32, no. 1, pp. 115–134.
13. *Couvreur A., Gaborit P., Gauthier-Umãna V., et al.* Distinguisher-based attacks on public-key cryptosystems using Reed — Solomon codes. Designs, Codes, Cryptogr., 2014, vol. 73, no. 2, pp. 641–666.
14. *Erdös P. and Spencer J.* Probabilistic Methods in Combinatorics. Budapest, Akadèmiai Kiadó, 1974. 106 p.
15. *Rödl V.* On a Packing and Covering Problem. European J. Combinatorics, 1985, vol. 6, no. 1, pp. 69–78.