BINARY LINEAR CODES WITH NEAR-EXTREMAL MAXIMUM DISTANCE*

ANDRÁS PONGRÁCZ†

Abstract. Let C denote a binary linear code with length n all of whose coordinates are essential, i.e., for each coordinate there is a codeword that is not zero in that position. Then the maximum distance D is strictly bigger than $n/2$, and the extremum $D = (n+1)/2$ is attained exactly by punctured Hadamard codes. In this paper, we classify binary linear codes with $D = n/2+1$. All of these codes can be produced from punctured Hadamard codes in one of essentially three different ways, each having a transparent description.

Key words. code, anticode, support, maximum distance

AMS subject classifications. 94B05, 94B65, 20B10

1. Introduction.

The present paper is a follow-up to [16], where binary linear codes with near extremal maximum distance were analyzed to obtain classification results for an extremal problem about finite permutation groups. More precisely, the size S of the support of a finite permutation group G is at most $2s-2$, where s denotes the maximum degree of elements in the permutation group G, and a description was given to those G such that S is $2s-2, 2s-3$ or $2s-4$. The dual notion $\mu(G)$, the minimum degree of non-identity elements, is also a central notion in permutation group theory. It was particularly well-studied for primitive permutation groups, see [13] for a recent improvement on the lower bound. Often the results are phrased for the fixity $S - \mu(G)$ of G, see [17, 19, 20].

The main direct motivation is a recent paper [1]. It was shown that an upper estimation to S in terms of s can be applied to obtain results about the asymptotic probability that a finite structure over a given finite relational language has an automorphism group isomorphic to some permutation group H, provided that the automorphism group contains a given permutation group G. It follows that only finitely many H occurs with positive asymptotic probability, and that the probability for any such H is a rational number. This generalizes the well-known theorem that, given a finite relational vocabulary, asymptotically almost all finite structures are rigid; see [4, 6, 7, 10] for further details. In order to compute the family of possible H corresponding to a given G, it is crucial to refine the upper bound on S in terms of s, and study the near extremal cases.

In [16] the cases $S = 2s-2$ and $S = 2s-3$ were fully characterized. The proof relies on a refinement of Burnside's lemma [14], and mainly on the following classification of punctured Hadamard codes up to equivalence in terms of the maximum distance of the code. We say that a coordinate is essential in a code if not all codewords are zero in that position.

*Submitted to the editors on September 19, 2019.

Funding: This work is supported by the EFOP-3.6.2-16-2017-00015 project, which has been supported by the European Union, co-financed by the European Social Fund. The paper was also supported by the National Research, Development and Innovation Fund of Hungary, financed under the FK 124814 and PD 125160 funding schemes, the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, and by the ÚNKP-18-4 and the ÚNKP-19-4 New National Excellence Program of the Ministry of Human Capacities.

†Department of Algebra and Number Theory, University of Debrecen, Debrecen, 4032 Hungary (pongracz.andras@science.unideb.hu, http://math.unideb.hu/pongracz-andras/en).

This manuscript is for review purposes only.

THEOREM 1.1. *Let $n \in \mathbb{N}$ and assume that a binary linear code C of length n has maximum distance $D \leq \frac{n+1}{2}$. Assume that all coordinates of the code are essential. Then $D = \frac{n+1}{2} = 2^{k-1}$ for some $k \geq 1$, and the code is equivalent to the punctured Hadamard code H_k with parameters $[2^k - 1, k, 2^{k-1}]_2$.*

The case $S = 2s - 4$ hinges on a partial result about binary linear codes with length n and maximum distance $D = \frac{n}{2} + 1$ all of whose coordinates are essential (see Theorem 2.2). Some further preliminary results were shown in [16] about codes with these properties, and the description to the above extremal problem $S = 2s - 4$ was reduced to a classification of such codes. The main contribution of the present paper is the complete description of such codes, see Theorem 2.6. Many of these codes are two- or three-weight binary linear codes (and give rise to further constructions like that), a concept actively studied lately, see [5, 11, 12, 23]. We recommend [3, 18] for an introduction to linear codes. An upper bound on the maximum distance is in [2].

2. Constructions and the main result. We recall a construction from [16].

DEFINITION 2.1. *Let H_k be the $[2^k - 1, k, 2^{k-1}]_2$ punctured Hadamard code, and let $m \leq k$. We define $H_{k\times m} := H_k \times H_m$, i.e., producing all concatenations of codewords in H_k and H_m. The code $H_{k|m}$ can be obtained from H_k by picking $2^m - 1$ coordinates such that the restriction of H_k to those is isomorphic to H_m, and repeating those coordinates simultaneously. Any code C with $H_{k|m} \leq C \leq H_{k\times m}$ has length $n = 2^k + 2^m - 2$ and maximum distance $D = 2^{k-1} + 2^{m-1} = \frac{n}{2} + 1$, and moreover, all coordinates of C are essential.*

For example, a generating matrix of $H_{3|2}$ is $M_{3|2}$ below.

$$M_{3|2} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

We also provide a generating matrix $M_{3\times 2}$ of $H_{3\times 2}$.

$$M_{3\times 2} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

It was noted in [16] that the list of codes in Definition 2.1 is not exhaustive. However, the following positive result was shown in [16].

THEOREM 2.2. *Let C be a binary linear code all of whose coordinates are essential with length $n \in \mathbb{N}$ and maximum distance $D = \frac{n}{2} + 1$. Then there exist $1 \leq m \leq k$ such that $n = 2^k + 2^m - 2$, $D = 2^{k-1} + 2^{m-1}$, and $H_{k|m} \leq C$.*

To obtain a full classification of codes with $D = \frac{n}{2} + 1$, we present some further constructions.

DEFINITION 2.3. *As usual, we say that two coordinates i, j are equivalent with respect to a code C, if for all codewords $c \in C$ we have $c_i = c_j$. The equivalence classes of $H_{m|m}$ are pairs. We say that a partition $X \cup X'$ of the coordinates of $H_{m|m}$ is symmetrical if X intersects all these pairs in exactly one element. More generally, for any $k \geq m$ we can talk about symmetrical partitions $X \cup Y \cup X'$ of the set of coordinates of $H_{k|m}$: Y consists of the non-repeated coordinates, and $X \cup X'$ is a symmetrical partition of the code restricted to $X \cup X'$ (which is isomorphic to $H_{m|m}$).*

Note that there are 2^m symmetrical partitions of the coordinates of $H_{k|m}$. In Definition 2.1 we somewhat loosely put $H_{k|m} \le C \le H_{k\times m}$. In order to represent the codes $H_{k|m}$ and $H_{k\times m}$, we need to fix a symmetrical partition $X \cup Y \cup X'$ of the set of coordinates of $H_{k|m}$, so that the supports of H_k and H_m are specified, namely these are $X \cup Y$ and X', respectively. This problem is going to cause some difficulties later on. E.g., if we are looking for nontrivial examples for codes C with $H_{k|m} \le C$ and $D = \frac{n}{2} + 1$, i.e., not of the form $H_{k|m} \le C \le H_{k\times m}$, then we need to make sure that such a containment does not hold with respect to any symmetrical partition.

Definition 2.4. *Let $X \cup X'$ be a symmetrical partition of the coordinates of $H_{m|m}$. We say that a vector c is $H_{m|m}$-balanced (with respect to the partition $X \cup X'$), if there exist $1 \le \ell \le m$ and ℓ independent codewords $c_1, \dots, c_\ell \in H_{m|m}$ such that $\mathrm{supp}(c) = X' \cap \bigcup_{i=1}^{\ell} \mathrm{supp}(c_i)$. Later on (cf. Lemmas 3.1 and 3.5), we are going to see that these are exactly the vectors such that $\langle H_{m|m}, c\rangle$ has the same maximum distance $D = 2^m$ as $H_{m|m}$. Thus it is natural for a code C with $H_{m|m} < C \le H_{m\times m}$ to say that a vector c be C-balanced if $\langle C, c\rangle$ has maximum distance $D = 2^m$.*

Clearly, C -balanced vectors for $H_{m|m} < C \le H_{m\times m}$ are $H_{m|m}$-balanced, thus they are as described in Definition 2.4. It is not hard to find such vectors for a given C, e.g., by solving a system of linear equations over $\mathbb{Q}$. We provide a non-trivial example. The following matrix is a generating matrix of a code C with $H_{3|3} < C \le H_{3\times 3}$.

$$M_{3|3} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Then $(0,0,0,1,0,1,0,1,1,0,0,0,0,0)$ is a C-balanced vector.

Finally, we present an infinite family of codes of the form $\langle H_{m+1|m}, c\rangle$.

Definition 2.5. *Let $X \cup Y \cup X'$ be a symmetrical partition of the coordinates of $H_{m+1|m}$. Let $a, b \in H_{m+1|m}$ be two codewords such that $\mathrm{supp}(a) \cap \mathrm{supp}(b) \cap Y$ is nonempty and the restriction of a and b to X are different nonzero vectors. Let c be the vector whose support is $\mathrm{supp}(c) = ((\mathrm{supp}(a) \cup \mathrm{supp}(b)) \cap X') \cup (\mathrm{supp}(a) \cap \mathrm{supp}(b) \cap Y)$. Then we say that c is $H_{m+1|m}$-balanced (with respect to the partition $X \cup Y \cup X'$).*

As an example, the second and third rows in $M_{3|2}$ can be chosen as a and b. (Here, X and X' are the set of first three and last three coordinates, respectively.) Then the matrix is extended by the row $(0,0,0,0,0,0,1,1,1,1)$. Note that as the construction requires two different nonzero vectors in H_m, such $H_{m+1|m}$-balanced vectors exist iff $2 \le m$. Also note that the definition of $H_{m|m}$- and $H_{m+1|m}$-balanced vectors depend on the symmetrical partition of the coordinates, an issue that causes some difficulties in proofs to come. We are now ready to state the main theorem of the paper.

Theorem 2.6. *Let C be a binary linear code all of whose coordinates are essential with length $n \in \mathbb{N}$ and maximum distance D. Then the following are equivalent.*

1. *The equation $D = \frac{n}{2} + 1$ holds.*
2. *For some $1 \le m \le k$ we have either*
   (a) *$H_{k|m} \le C \le H_{k\times m}$ (with respect to some symmetrical partition), or*
   (b) *$k = m$, $C = \langle C_0, c\rangle$ with $H_{m|m} \le C_0 \le H_{m\times m}$ and a C_0-balanced c not in $H_{m\times m}$ (with respect to any symmetrical partition), or*
   (c) *$2 \le m$, $k = m+1$, $C = \langle H_{m+1|m}, c\rangle$, and c is $H_{m+1|m}$-balanced.*

This manuscript is for review purposes only.

3. Correctness of the constructions and minimal examples. We show the implication 2. ⇒ 1. in Theorem 2.6 in the next two lemmas (3.1 and 3.2).

LEMMA 3.1. *Let* $m \in \mathbb{N}$.

1. *Let* c *be an* $H_{m|m}$*-balanced vector. Then* $C = \langle H_{m|m}, c\rangle$ *has the same length and maximum distance as* $H_{m|m}$ *(and all coordinates are essential in* C*).*
2. *Let* $X \cup X'$ *be a symmetrical partition of the coordinates of* $H_{m|m}$*. Then a vector* c *such that* $\mathrm{supp}(c) \cap X' \neq \emptyset$ *is* $H_{m|m}$*-balanced with respect to the partition* $X \cup X'$*, iff* $\mathrm{supp}(c) = X'$ *or the restriction of* $H_{m|m}$ *to* $X' \setminus \mathrm{supp}(c)$ *is equivalent to a punctured Hadamard code. In particular, there exists a* $0 \leq m' \leq m-1$ *such that for all codewords* $c' \in \langle H_{m|m}, c\rangle \setminus H_{m|m}$*, the number of* $H_{m|m}$*-equivalent pairs of coordinates* (x, x') *such that the value of* c' *in* x *and* x' *coincides is* $2^{m'} - 1$*.*

Proof. We use the notations of Definition 2.4.

For item 1. we need to show that for all $u \in H_{m|m}$ we have $w(c+u) \leq 2^m$. This clearly holds for $u = 0$. Assume that $u \in H_{m|m}$ is not zero. Then u is a concatenation aa', where a and a' are identical maximum weight codewords in the two copies of H_m. If $\mathrm{supp}(a') \subseteq \mathrm{supp}(c)$, then $w(c+u) < w(u) = 2^m$.

Hence, assume that $\mathrm{supp}(a') \not\subseteq \mathrm{supp}(c)$. In particular, $\ell \leq m-1$. By using induction on ℓ, it is easy to show that $\mathrm{supp}(a') \setminus \mathrm{supp}(c) = \mathrm{supp}(a') \setminus \bigcup_{i=1}^{\ell} \mathrm{supp}(c_i)$ has size $2^{m-1-\ell}$ (we note that this fails for $\ell = m$). Consequently, $|\mathrm{supp}(a') \cap \mathrm{supp}(c)|$ is $2^{m-1} - 2^{m-1-\ell}$. It is also clear by using induction on ℓ that $w(c) = 2^m - 2^{m-\ell}$. Thus $w(c+u) = 2^{m-1} + ((2^m - 2^{m-\ell}) + 2^{m-1} - 2 \cdot (2^{m-1} - 2^{m-1-\ell})) = 2^m$.

The only if part in item 2. is trivial by induction on ℓ, as the cancellation of the support of a nonzero codeword from a punctured Hadamard code H_r yields H_{r-1}.

We use induction on m for the if part. It clearly holds if $\mathrm{supp}(c) = X'$ by the definition of an $H_{m|m}$-balanced vector, hence we may assume that $\mathrm{supp}(c) \neq X'$. In particular, the initial step $m = 1$ is trivial. Hence, assume that $m \geq 2$ and the assertion holds for $m-1$.

Let H_r be the punctured Hadamard code obtained as the restriction of $H_{m|m}$ to $X' \setminus \mathrm{supp}(c)$. Then $1 \leq r \leq m-1$ by assumption. Restriction of codewords to $X' \setminus \mathrm{supp}(c)$ is a homomorphism, and as every coordinate of $H_{m|m}$ is essential, the kernel of this homomorphism is nontrivial. Thus there is a nonzero codeword $c_1 \in H_{m|m}$ whose support is disjoint from $X' \setminus \mathrm{supp}(c)$. Let us puncture the code $H_{m|m}$ by omitting $\mathrm{supp}(c_1)$. Then we obtain the code $H_{m-1|m-1}$ with the same properties (the punctured version of c takes the role of c), and then we are done by the induction hypothesis. □

We denote the characteristic vector of Y by 1_Y. Note that $1_Y \in H_{m+1|m}$.

LEMMA 3.2. *Let* $2 \leq m$ *and let* $a, b, c \in H_{m+1|m}$ *as in Definition 2.5. Then* $w(c) = w(c+a) = w(c+b) = w(c+a+b+1_Y) = 2^m$*, and* $w(c+u) = 3 \cdot 2^{m-1}$ *for all other codewords* $u \in H_{m+1|m}$*. In particular, the code* $C = \langle H_{m+1|m}, c\rangle$ *has the same length and maximum distance as* $H_{m+1|m}$ *(and all coordinates are essential in* C*).*

Proof. It is easy to see that if $u \in \langle a, b, 1_Y\rangle$, then we have $w(c+u) = 2^m$ if $u \in \{0, a, b, a+b+1_Y\}$, and $w(c+u) = 3 \cdot 2^{m-1}$ for the other four vectors $u \in \langle a, b, 1_Y\rangle$. So assume that $u \in H_{m+1|m} \setminus \langle a, b, 1_Y\rangle$. Then both $\mathrm{supp}(c) \cap X'$ and $\mathrm{supp}(c) \cap Y$ are cut in half by $\mathrm{supp}(u)$. As $w(c) = 2^m$, $\mathrm{supp}(c) \cap X$ is empty, $|\mathrm{supp}(u) \cap (X' \cup Y)| = 2^m$ and $|\mathrm{supp}(u) \cap X| = 2^{m-1}$, we have $w(c+u) = 2^{m-1} + (2^m + 2^m - 2 \cdot \frac{1}{2} \cdot 2^m) = 3 \cdot 2^{m-1}$. □

Now we turn our attention to the implication 1. ⇒ 2. in Theorem 2.6. According to Theorem 2.2, all codes C that satisfy item 1. of Theorem 2.6 contain some $H_{k|m}$. It is a natural idea to first understand the minimal examples.

Definition 3.3. *Throughout the rest of the paper, we call a binary linear code* C *with* $H_{k|m} \leq C$ *for some* $1 \leq m \leq k$ *(making every coordinate of* C *essential automatically) a minimal example, if* $|C : H_{k|m}| = 2$ *and* $D = \frac{n}{2} + 1$*, where* $n = 2^k + 2^m - 2$ *is the length of* C *and* $D = 2^{k-1} + 2^{m-1}$ *is the maximum distance of* C*. Recall that the union of singleton* $H_{k|m}$ *equivalence classes is denoted by* Y*.*

The next proposition classifies minimal examples as a special case of Theorem 2.6.

Proposition 3.4. *Let* $1 \leq m \leq k$ *and let* $H_{k|m} \leq C$ *be a minimal example (cf. Definition 3.3). Then either*

- $H_{k|m} \leq C \leq H_{k\times m}$ *(with respect to some symmetrical partition), or*
- $k = m$*,* $C = \langle H_{m|m}, c\rangle$ *with some* $H_{m|m}$*-balanced* c *not in* $H_{m\times m}$ *(with respect to any symmetrical partition), or*
- $2 \leq m$*,* $k = m + 1$*,* $C = \langle H_{m+1|m}, c\rangle$*, and* c *is* $H_{m+1|m}$*-balanced.*

For the sake of transparency, we break the proof of Proposition 3.4 down into two cases: $k = m$ and $k > m$. If $k = m$, then the first two items can be merged: note that vectors in $H_{m\times m}$ are $H_{m|m}$-balanced (with $\ell = 1$ in Definition 2.4).

Lemma 3.5. *Let* $m \in \mathbb{N}$ *and let* $H_{m|m} \leq C$ *be a minimal example (cf. Definition 3.3). Then* $C = \langle H_{m|m}, c\rangle$ *for some* $H_{m|m}$*-balanced vector* c*.*

Proof. Assume that a codeword $c \in C \setminus H_{m|m}$ is one in a pair of repeated coordinates. We can pick $c_1, \ldots, c_{m-1} \in H_{m|m}$ so that their supports cover all coordinates except for that pair. Thus all coordinates of $C' = \langle c_1, \ldots, c_{m-1}, c\rangle$ are essential, and $\dim C' = m$. Clearly, the length of C' is $n = 2^{m+1} - 2$, and its maximum distance is $D = 2^m$. Hence, according to Theorem 2.2, C' is equivalent to $H_{m|m}$. In particular, $w(c) = D > \frac{n}{2}$. As the average weight in $C \setminus H_{m|m}$ is $\frac{n}{2}$, this cannot hold for all $c \in C \setminus H_{m|m}$. Thus we can pick a $c \in C \setminus H_{m|m}$ that is zero in at least one position within each pair of repeated coordinates. Then there is a symmetrical partition $X \cup X'$ such that $\mathrm{supp}(c) \subseteq X'$. Let $Z := X' \setminus \mathrm{supp}(c)$ and $r = |Z|$. If $r = 0$ then c is indeed an $H_{m|m}$-balanced vector (with $\ell = m$ in Definition 2.4).

Assume that $r \geq 1$, and pick a codeword $c' \in H_{m|m}$. If c' has t ones in Z, then $w(c+c') = 2^{m-1} + t + ((2^m - 1 - r) - (2^{m-1} - t)) = 2t - (r+1) + 2^m \leq D = 2^m$, thus $t \leq \frac{r+1}{2}$. By Theorem 1.1, the restriction of $H_{m|m}$ to Z is equivalent to the punctured Hadamard code H_r, and the assertion follows from Lemma 3.1. ∎

The rest of this section is all about minimal examples with $k > m$.

Lemma 3.6. *Let* $1 \leq m < k$ *and let* $H_{k|m} \leq C$ *be a minimal example (cf. Definition 3.3). Assume that there is a symmetrical partition* $X \cup Y \cup X'$ *of the coordinates of* $H_{k|m}$ *such that for some* $c \in C \setminus H_{k|m}$ *we have* $\mathrm{supp}(c) \subseteq X'$*. Then* $H_{k|m} \leq C \leq H_{k\times m}$ *(with respect to some symmetrical partition).*

Proof. We need to show that the restriction c_0 of c to X' is in the punctured Hadamard code H_m obtained as the restriction of $H_{m+1|m}$ to X'. Assuming this is not the case, by Theorem 1.1 the code $\langle c_0, H_m\rangle$ contains a codeword that has bigger weight than 2^{m-1}. This codeword cannot be c_0, as otherwise $w(c + c') > D$ for some nonzero $c' \in H_{k|m}$ with $\mathrm{supp}(c') \subseteq Y$, as all such c' have weight 2^{k-1}. Thus such a codeword in H_m is obtained as the restriction of $c + c'$ with some maximum weight $c' \in H_{k|m}$. But then the weight of the restriction of c', and also of $c + c'$ to $X \cup Y$ is $D - 2^{m-1}$, making $w(c + c') > D$, a contradiction. ∎

LEMMA 3.7. *Let* $1 \leq m < k$ *and let* $H_{k|m} \leq C$ *be a minimal example (cf. Definition 3.3). Assume that there is a codeword* $c \in C \setminus H_{k|m}$ *such that* $\mathrm{supp}(c) \cap Y = \emptyset$*. Then* $H_{k|m} \leq C \leq H_{k\times m}$ *(with respect to some symmetrical partition).*

Proof. We have $w(c) \leq 2^{m-1}$, as otherwise $w(c + c') > D$ for some nonzero $c' \in H_{k|m}$ with $\mathrm{supp}(c') \subseteq Y$. If we puncture the code by omitting Y, then we obtain $H_{m|m}$. The punctured version c_0 of c has the same weight as c, and thus $c_0 \notin H_{m|m}$.

If the maximum distance of $\langle H_{m|m}, c_0\rangle$ is larger than 2^m, then there is a nonzero $c' \in H_{k|m}$ such that $|\mathrm{supp}(c + c') \setminus Y| > 2^m$. The support of c' intersects Y in $2^{k-1} - 2^{m-1}$ coordinates, thus $w(c + c') > D$, a contradiction.

Hence, $\langle H_{m|m}, c_0\rangle$ is a minimal example, and then it contains an $H_{m|m}$-balanced vector u_0 with respect to a symmetrical partition $X \cup Y \cup X'$ by Lemma 3.5. By Lemma 3.6 we have $u_0 \neq c_0$, thus u_0 must be the punctured version of $c + c'$ for some nonzero $c' \in H_{k|m}$ with $\mathrm{supp}(c') \not\subseteq Y$. Hence, $\mathrm{supp}(c + c') \cap X = \mathrm{supp}(u_0) \cap X = \emptyset$, which means that c and c' agree on X, and consequently, $|\mathrm{supp}(c) \cap X| = 2^{m-1}$. As $w(c) \leq 2^{m-1}$, we have $\mathrm{supp}(c) \subseteq X$, and then we are done by Lemma 3.6. ∎

LEMMA 3.8. *Let* $1 \leq m < k$ *and let* $H_{k|m} \leq C$ *be a minimal example (cf. Definition 3.3). If* $\mathrm{supp}(c) \cap Y \neq \emptyset$ *for some* $c \in C \setminus H_{k|m}$*, then either* $w(c) = D$ *or* $w(c) = D - 2^{m-1}$*.*

Proof. As c is one in a coordinate of the H_k-component, there are $k - 1$ independent vectors in H_k such that together with the H_k-component of c their supports cover every coordinate of H_k. Let $c_1, \dots, c_{k-1}$ be the corresponding $k - 1$ independent vectors in $H_{k|m}$. As the H_m-component is produced by repetition, the supports of $c_1, \dots, c_{k-1}, c$ cover every coordinate of $H_{k|m}$. Then the code C' generated by these k vectors has dimension k, length $n = 2^k + 2^m - 2$ and maximum distance $D = 2^{k-1} + 2^{m-1}$, and all coordinates of C' are essential. According to Theorem 2.2, C' is equivalent to $H_{k|m}$, all of whose nonzero codewords have weight D or $D-2^{m-1}$. ∎

LEMMA 3.9. *Let* $1 \leq m < k$ *and let* $H_{k|m} \leq C$ *be a minimal example (cf. Definition 3.3). If the support of a codeword* $c \in C\setminus H_{k|m}$ *contains a pair of* $H_{k|m}$*-equivalent coordinates* (x, x')*, then* $w(c) = D$*.*

Proof. There exist $m - 1$ independent vectors in the H_m-component with set of coordinates X' whose total support is $X' \setminus \{x'\}$. Pick extensions $c_1, \dots, c_{m-1} \in H_{k|m}$ of these vectors. Then the supports of $c_1, \dots, c_{m-1}, c$ cover $X \cup X'$. There are $k - m$ independent vectors $c_{m+1}, \dots, c_k \in H_{k|m}$ whose total support is Y. Hence, the code $C' := \langle c_1, \dots, c_{m-1}, c, c_{m+1}, \dots, c_k\rangle$ has dimension k, length $n = 2^k + 2^m - 2$ and maximum distance $D = 2^{k-1} + 2^{m-1}$, and all coordinates of C' are essential. According to Theorem 2.2, C' is equivalent to $H_{k|m}$. As the support of c contains a pair of equivalent coordinates in C', it must be a maximum weight codeword. ∎

In order to finish the proof of Proposition 3.4, we need the following lemma.

LEMMA 3.10. *Let* $1 \leq m < k$ *and let* $H_{k|m} \leq C \not\leq H_{k\times m}$ *(with respect to any symmetrical partition) be a minimal example (cf. Definition 3.3). Then* $2 \leq m$*,* $k = m + 1$ *and* $C = \langle H_{m+1|m}, c\rangle$ *with some* $H_{m+1|m}$*-balanced vector* c*.*

Proof. Let C_0 denote the index 2 subcode in C isomorphic to $H_{k|m}$. For all $c \in C \setminus C_0$ we have $\mathrm{supp}(c) \cap Y \neq \emptyset$ according to the assumption and Lemma 3.7, and $w(c) = 2^{k-1}$ or $w(c) = 2^{k-1} + 2^{m-1}$ by Lemma 3.8. As the average weight in $C \setminus C_0$ is $\frac{n}{2}$, there are 2^{k-m+1} codewords in $C \setminus C_0$ with weight 2^{k-1} and $2^k - 2^{k-m+1}$ with weight $2^{k-1} + 2^{m-1}$. Pick a $c \in C \setminus C_0$ with $w(c) = 2^{k-1}$. By Lemma 3.9 there is a symmetrical partition $X \cup Y \cup X'$ such that $\mathrm{supp}(c) \cap X = \emptyset$.

Let $y \in \operatorname{supp}(c) \cap Y$ be arbitrary, and let $c_1, \dots, c_{k-1} \in C_0$ be such that their supports cover all coordinates except for y. Then with respect to $\langle c_1, \dots, c_{k-1} \rangle$ there are $2^m - 1$ equivalence classes of the coordinates with size three, $2^{k-1} - 2^m$ with size two and 1 with size one. The three-element $\langle c_1, \dots, c_{k-1} \rangle$-classes are obtained from the pairs in $X \cup X'$ by adjoining an element from Y. By Theorem 2.2 we have that $C' := \langle c_1, \dots, c_{k-1}, c \rangle$ is equivalent to $H_{k|m}$, a code with no three-element equivalence classes. Thus c splits all three-element $\langle c_1, \dots, c_{k-1} \rangle$-classes into one with size two and one with size one, and since $w(c) = 2^{k-1}$, the support of c is contained in the singleton coordinates of C'. Thus a three-element $\langle c_1, \dots, c_{k-1} \rangle$-class $\{x, x', z\}$ with $x \in X, x' \in X', z \in Y$ is split by c so that the two-element class obtained is outside $\operatorname{supp}(c)$, and the singleton class obtained is inside $\operatorname{supp}(c)$. As $x \notin \operatorname{supp}(c)$, we have that x is a repeated coordinate in C', and its pair with respect to C' is either x' or z, hence it is outside X. Consequently, if we puncture C' by omitting X, then we obtain a code isomorphic to H_k.

If there is a coordinate $y \in Y$ where some $c' \in C_0$ is zero and c is one, then in the above argument c' can be chosen as one of the generators of C'. In particular, if $w(c') = D$, then $w(c + c') = D$, as the weight of $c + c'$ is 2^{k-1} in the restriction of C' to $X' \cup Y$ (isomorphic to H_k), and inside X the weight of $c + c'$ is 2^{m-1}. Similarly, if $w(c') = 2^{k-1}$, then $w(c + c') = 2^{k-1}$, provided that $c' \in C_0$ has a zero in Y where c is one. As there are $2^{k-m} - 1$ codewords in C_0 with weight 2^{k-1} and there are 2^{k-m+1} codewords in $C \setminus C_0$ with weight 2^{k-1}, there exists a codeword $a \in C_0$ such that $w(a) = 2^{k-1} + 2^{m-1}$ and $w(c + a) = 2^{k-1}$. Then $\operatorname{supp}(c) \cap Y \subseteq \operatorname{supp}(a) \cap Y$, and moreover, as $a \in C_0$ is a maximum weight codeword, we have $|\operatorname{supp}(a) \cap Y| = 2^{k-1} - 2^{m-1}$, and $|\operatorname{supp}(a) \cap X| = |\operatorname{supp}(a) \cap X'| = 2^{m-1}$.

Let K denote the set $\{c_1 \in C_0 \mid w(c_1) = 2^{k-1}\}$. Assume that for all $c_1 \in K$ we have $w(c + c_1) = 2^{k-1}$. Let $C_1 := \langle \{c\} \cup K \} \rangle$, and let n_1 be the number of essential coordinates of C_1. The average weight in C_1 is $\frac{n_1}{2} = \frac{2^{k-m+1}-1}{2^{k-m+1}} \cdot 2^{k-1}$, thus $n_1 = 2^k - 2^{m-1}$. Note that $\bigcup_{c_1 \in K} \operatorname{supp}(c_1) = Y$ with size $2^k - 2^m$. Hence, $|\operatorname{supp}(c) \cap X'| = 2^{m-1}$, and then $|\operatorname{supp}(c) \cap Y| = 2^{k-1} - 2^{m-1} = |\operatorname{supp}(a) \cap Y|$. As $\operatorname{supp}(c) \cap Y \subseteq \operatorname{supp}(a) \cap Y$, we have $\operatorname{supp}(c) \cap Y = \operatorname{supp}(a) \cap Y$, and then $c + a \in C \setminus C_0$ is all zero in Y, a contradiction by Lemma 3.7.

Thus there is a $c_1 \in C_0$ with $w(c_1) = 2^{k-1} = w(c)$ and $w(c + c_1) = 2^{k-1} + 2^{m-1}$, and consequently, $|\operatorname{supp}(c) \setminus \operatorname{supp}(c_1)| = 2^{k-2} + 2^{m-2}$. We have shown above that $w(c_1) = 2^{k-1}$ and $w(c + c_1) = 2^{k-1} + 2^{m-1}$ is not possible if there is a coordinate in Y where c_1 is zero and c is one, thus $\operatorname{supp}(c) \cap Y \subseteq \operatorname{supp}(c_1) \cap Y$. In particular, $\operatorname{supp}(c) \setminus \operatorname{supp}(c_1) \subseteq X'$, thus $2^{k-2} + 2^{m-2} \leq 2^m - 1$, and then $k = m + 1$. Moreover, as $c_1 \in K$, we have $\operatorname{supp}(c_1) \subseteq Y$. Hence, $\operatorname{supp}(c) \setminus \operatorname{supp}(c_1) = \operatorname{supp}(c) \cap X'$. Thus $|\operatorname{supp}(c) \cap X'| = 2^{m-1} + 2^{m-2} = 3 \cdot 2^{m-2}$ and $|\operatorname{supp}(c) \cap Y| = 2^{m-2}$. Moreover, $w(a) = 3 \cdot 2^{m-1}$, $w(c + a) = 2^m$, and $|\operatorname{supp}(a) \cap Y| = 2^{m-1}$. Then we have that $|\operatorname{supp}(c+a) \cap Y| = 2^{m-2}$, $|\operatorname{supp}(c+a) \cap X| = |\operatorname{supp}(a) \cap X| = 2^{m-1}$, and consequently, $|\operatorname{supp}(c+a) \cap X'| = w(c+a) - |\operatorname{supp}(c+a) \cap Y| - |\operatorname{supp}(c+a) \cap X| = 2^{m-2}$. Hence, $|\operatorname{supp}(c) \cap \operatorname{supp}(a) \cap X'| = \frac{1}{2} \cdot (|\operatorname{supp}(c) \cap X'| + |\operatorname{supp}(a) \cap X'| - |\operatorname{supp}(c+a) \cap X'|) = 2^{m-1} = |\operatorname{supp}(a) \cap X'|$, thus $\operatorname{supp}(a) \cap X' \subseteq \operatorname{supp}(c) \cap X'$.

We now revisit the ideas in the first and third paragraphs of the proof, using the additional information that $k = m + 1$. In particular, there is a unique codeword in C_0 with weight $2^{k-1} = 2^m$, namely 1_Y. Thus all the remaining $2^{m+1} - 2$ nonzero codewords in C_0 have maximum weight $3 \cdot 2^{m-1}$. In $C \setminus C_0$, there are $2^{k-m+1} = 4$ codewords with weight $2^{k-1} = 2^m$ and $2^k - 2^{k-m+1} = 2^{m+1} - 4$ codewords with maximum weight $D = 3 \cdot 2^{m-1}$. Recall that $w(c) = 2^{k-1} = 2^m$. As $|\operatorname{supp}(c) \cap Y| =$

2^{m-2} and $|Y| = 2^m$, we have $w(c+1_Y) = w(c) + 2^m - 2 \cdot 2^{m-2} = 3 \cdot 2^{m-1} = D$, thus $c+1_Y$ is one of the $2^{m+1}-4$ maximum weight codewords in $C \setminus C_0$. Hence, out of the $2^{m+1}-2$ maximum weight codewords in $C_0 \setminus \{0, 1_Y\}$, there are exactly three codewords c' with $w(c+c') = 2^m$. One of those three is a, and there are exactly two codewords in C_0 with the same restriction to X' as a, namely a and $a+1_Y$. Thus there must be a codeword $b \in C_0 \setminus \{0, 1_Y\}$ such that $w(c+b) = 2^m$ and the restrictions of a and b to X' are different. In particular, there exist two different nonzero codewords in H_m, thus $m \geq 2$. Moreover, every claim that we have proved about a can be copied to b, namely: $\operatorname{supp}(c) \cap Y \subseteq \operatorname{supp}(b) \cap Y$, $\operatorname{supp}(b) \cap X' \subseteq \operatorname{supp}(c) \cap X'$, $|\operatorname{supp}(b) \cap Y| = 2^{m-1}$, and $|\operatorname{supp}(b) \cap X| = |\operatorname{supp}(b) \cap X'| = 2^{m-1}$. Thus $\operatorname{supp}(c) \cap Y \subseteq \operatorname{supp}(a) \cap \operatorname{supp}(b) \cap Y$, and both have size 2^{m-2}, and consequently, $\operatorname{supp}(c) \cap Y = \operatorname{supp}(a) \cap \operatorname{supp}(b) \cap Y$. Furthermore, $(\operatorname{supp}(a) \cap X') \cup (\operatorname{supp}(b) \cap X') \subseteq \operatorname{supp}(c) \cap X'$, and both have size $3 \cdot 2^{m-2}$, so $(\operatorname{supp}(a) \cup \operatorname{supp}(b)) \cap X' = \operatorname{supp}(c) \cap X'$.

Hence, c is $H_{m+1|m}$-balanced with the choice of a, b as above in Definition 2.5. □

Proof of Proposition 3.4. Done by Lemmas 3.5 and 3.10. □

4. The general case. The next lemma finishes the proof of the classification if $k = m$.

LEMMA 4.1. *Let $H_{m|m} \leq C$ be a code with maximum distance 2^m. Then there exists a $C_0 \leq C$ with index at most two such that $H_{m|m} \leq C_0 \leq H_{m \times m}$.*

Proof. We may assume that $H_{m|m} < C$. Pick $H_{m|m} \leq C_0 \leq C$ together with a symmetrical partition such that $H_{m|m} \leq C_0 \leq H_{m \times m}$ (with respect to that partition) and the dimension of C_0 be maximal. Let $X \cup X'$ be a symmetrical partition such that $H_{m|m} \leq C_0 \leq H_{m \times m}$. Assume indirectly that $|C : C_0| > 2$.

Pick $c_1, c_2 \in C \setminus C_0$ from different cosets of C_0. Then both $C_i = \langle H_{m|m}, c_i \rangle$ are minimal examples (cf. Definition 3.3), and then by Lemma 3.5 we may assume that both c_i are $H_{m|m}$-balanced (with respect to potentially different symmetrical partitions that may also differ from $X \cup X'$). By Lemma 3.1, the number of $H_{m|m}$-equivalent pairs (x, x') such that the value of c_i in x and in x' coincide is $2^{m_i} - 1$ for some $0 \leq m_1 \leq m_2 \leq m - 1$, without loss of generality.

First, assume that $m_2 \leq m - 2$. Then $2^{m_1} - 1 \leq 2^{m_2} - 1 < \frac{1}{4} \cdot (2^m - 1)$, where $2^m - 1$ is the number of all $H_{m|m}$-equivalent pairs. Hence, the number of $H_{m|m}$-equivalent pairs (x, x') such that the value of $c_1 + c_2$ in x and x' differ is less than $\frac{1}{2} \cdot (2^m - 1)$. If $c_1 + c_2 \notin H_{m|m}$ then $\langle H_{m|m}, c_1 + c_2 \rangle$ is a minimal example, and consequently, every codeword in $\langle H_{m|m}, c_1 + c_2 \rangle \setminus H_{m|m}$ differs in more than half of the pairs. Thus $c_1 + c_2 \in H_{m|m}$, and then c_1 and c_2 are in the same C_0-coset, a contradiction.

Hence, $m_2 = m - 1$, and then there exists a symmetrical partition $X_2 \cup X_2'$ such that c_2 is the restriction of a nonzero codeword in $H_{m|m}$ to X_2'. In particular, we have $H_{m|m} < C_0$ by maximality of the dimension of C_0.

Let $c \in C_0 \setminus H_{m|m}$ be any vector with weight 2^{m-1}. If the support of c and c_2 intersect the same pairs of $H_{m|m}$-equivalent coordinates nontrivially, then $c + c_2$ have a symmetrical support: each $H_{m|m}$-equivalent pair is either fully contained or fully not contained in it. Thus the $\langle H_{m|m}, c + c_2 \rangle$-classes coincide with the $H_{m|m}$-classes, and then $\langle H_{m|m}, c + c_2 \rangle$ is the repetition of an index 2 extension of H_m. According to Theorem 1.1 any extension of H_m has larger maximum weight than 2^{m-1}, and thus the code $\langle H_{m|m}, c + c_2 \rangle$ has larger maximum distance than 2^m, a contradiction. Hence, c_2 must be the restriction of a nonzero codeword to X_2' that is different from any codeword whose restriction to X or X' is in C_0. Due to the large degree of symmetry of $H_{m|m}$, it makes no difference which nonzero codeword we choose among those. The illustration below is for $m = 4$.

$$
\begin{array}{ccccccccccccccc|cccccccccccccccc}
\multicolumn{15}{c}{\overbrace{\hphantom{0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0}}^{X}} & \multicolumn{15}{c}{\overbrace{\hphantom{0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0}}^{X'}} & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0& \\
0&0&0&0&0&0&0&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&1&1&1&1&1&1&1&1& e_1 \\
0&0&0&1&1&1&1&0&0&0&0&1&1&1&1&0&0&0&1&1&1&1&0&0&0&0&1&1&1&1& e_2 \\
0&0&0&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&1&1&1&1&1&1&1&1&0&0&0&0& \\
0&1&1&0&0&1&1&0&0&1&1&0&0&1&1&0&1&1&0&0&1&1&0&0&1&1&0&0&1&1& e_3 \\
0&1&1&0&0&1&1&1&1&0&0&1&1&0&0&0&1&1&0&0&1&1&1&1&0&0&1&1&0&0& \\
0&1&1&1&1&0&0&0&0&1&1&1&1&0&0&0&1&1&1&1&0&0&0&0&1&1&1&1&0&0& \\
0&1&1&1&1&0&0&1&1&0&0&0&0&1&1&0&1&1&1&1&0&0&1&1&0&0&0&0&1&1& \\
1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1& e_4 \\
1&0&1&0&1&0&1&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&1&0&1&0&1&0&1&0& \\
1&0&1&1&0&1&0&0&1&0&1&1&0&1&0&1&0&1&1&0&1&0&0&1&0&1&1&0&1&0& \\
1&0&1&1&0&1&0&1&0&1&0&0&1&0&1&1&0&1&1&0&1&0&1&0&1&0&0&1&0&1& \\
1&1&0&0&1&1&0&0&1&1&0&0&1&1&0&1&1&0&0&1&1&0&0&1&1&0&0&1&1&0& \\
1&1&0&0&1&1&0&1&0&0&1&1&0&0&1&1&1&0&0&1&1&0&1&0&0&1&1&0&0&1& \\
1&1&0&1&0&0&1&0&1&1&0&1&0&0&1&1&1&0&1&0&0&1&0&1&1&0&1&0&0&1& \\
1&1&0&1&0&0&1&1&0&0&1&0&1&1&0&1&1&0&1&0&0&1&1&0&0&1&0&1&1&0& \\
\hline
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&1&1&1& c \\
0&0&0&0&0&0&0&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0& \\
0&0&0&1&1&1&1&0&0&0&0&1&1&1&1&0&0&0&1&1&1&1&1&1&1&1&0&0&0&0& \\
0&0&0&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&1&1&1&1&0&0&0&0&1&1&1&1& \\
\hline
0&1&1&0&0&1&1&0&0&1&1&0&0&1&1&0&1&1&0&0&1&1&1&1&0&0&1&1&0&0& \\
0&1&1&0&0&1&1&1&1&0&0&1&1&0&0&0&1&1&0&0&1&1&0&0&1&1&0&0&1&1& \\
0&1&1&1&1&0&0&0&0&1&1&1&1&0&0&0&1&1&1&1&0&0&1&1&0&0&0&0&1&1& \\
0&1&1&1&1&0&0&1&1&0&0&0&0&1&1&0&1&1&1&1&0&0&0&0&1&1&1&1&0&0& \\
1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&1&0&1&0&1&0&1&1&0&1&0&1&0&1&0& \\
1&0&1&0&1&0&1&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1& \\
1&0&1&1&0&1&0&0&1&0&1&1&0&1&0&1&0&1&1&0&1&0&1&0&1&0&0&1&0&1& \\
1&0&1&1&0&1&0&1&0&1&0&0&1&0&1&1&0&1&1&0&1&0&0&1&0&1&1&0&1&0& \\
1&1&0&0&1&1&0&0&1&1&0&0&1&1&0&1&1&0&0&1&1&0&1&0&0&1&1&0&0&1& \\
1&1&0&0&1&1&0&1&0&0&1&1&0&0&1&1&1&0&0&1&1&0&0&1&1&0&0&1&1&0& \\
1&1&0&1&0&0&1&0&1&1&0&1&0&0&1&1&1&0&1&0&0&1&1&0&0&1&0&1&1&0& \\
1&1&0&1&0&0&1&1&0&0&1&0&1&1&0&1&1&0&1&0&0&1&0&1&1&0&1&0&0&1& \\
\multicolumn{7}{c}{} & \multicolumn{8}{c}{\underbrace{\hphantom{0\;0\;0\;0\;0\;0\;0\;0}}_{2^{m-1} \le i \le 2^m - 1}} & \multicolumn{16}{c}{}
\end{array}
$$

Let us represent $H_{m|m}$ in the standard way. That is, we produce the generating matrix by writing the binary representation of all numbers from 1 to $2^m - 1$ in columns, and then by repeating all these columns. Let $e_1, \dots, e_m$ denote the rows of this matrix from top to bottom; this is the standard basis of the code. Then we sort the codewords $\sum_{i=1}^{m} \varepsilon_i e_i$, $\varepsilon_i \in \{0, 1\}$, so that the sequence of coefficients $\overline{\varepsilon_1 \cdots \varepsilon_m}$ corresponding to the r-tk codeword is the binary representation of r (extended by zeros on the right) for $r = 0, \dots, 2^m - 1$. That is, the list of codewords is $0, e_1, e_2, e_2 + e_1, e_3, \dots, e_m + \dots + e_1$. Without loss of generality, we may assume that c is the restriction of e_1 to X', and c_2 is the restriction of e_2 to X'_2. The vectors $c + u \in c + H_{m|m}$ are listed according to the order of the elements $u \in H_{m|m}$. Note that in this coset, every codeword has the same value in the pair of $H_{m|m}$-equivalent coordinates x, x' if $x' \notin \mathrm{supp}(c)$, that is, in the first $2^{m-1} - 1$ pairs from the left. In particular, regardless of the choice of X_2 and X'_2, every codeword of the form $c_2 + c + u \in c_2 + c + H_{m|m}$ with $u \in \{1, e_1, e_2, e_2 + e_1\}$ (i.e., the first four vectors in $H_{m|m}$) has 2^{m-2} ones in the union of the first $2^{m-1} - 1$ pairs, and every codeword of the form $c_2 + c + u \in c_2 + c + H_{m|m}$ with $u \in H_{m|m} \setminus \{1, e_1, e_2, e_2 + e_1\}$ has 2^{m-1} ones in the union of the first $2^{m-1} - 1$ pairs. Let us focus on the latter vectors, i.e., the ones of the form $c_2 + c + u \in c_2 + c + H_{m|m}$ with $u \in H_{m|m} \setminus \{1, e_1, e_2, e_2 + e_1\}$. Note that these are listed in consecutive pairs of vectors that have opposite value in every coordinate from index 2^{m-1} to $2^m - 1$ in both X and X'. Thus in two such rows, the number of ones in those coordinates is 2^m altogether, regardless of the choice of X_2 and X_2'. According to the above observations, the number of ones in the first $2^{m-1} - 1$ pairs of coordinates is also 2^m in such a codeword, making the sum of weights of a consecutive pair of codewords 2^{m+1}. As the maximum weight in the code is 2^m, both codewords have weight exactly 2^m. Thus for all $u \in H_{m|m} \setminus \{1, e_1, e_2, e_2 + e_1\}$, we have $w(c_2 + c + u) = 2^m$.

This gives rise to a system of linear equations over $\mathbb{Q}$. Let us introduce pairs of variables corresponding to the pairs of $H_{m|m}$-equivalent coordinates denoted by x_1, x_1', $x_2, x_2', \cdots, x_{2^m-1}, x_{2^m-1}'$ with $x_1, \ldots, x_{2^m-1}$ corresponding to coordinates in X, such that $x_i = 1$ if the i-th coordinate in X is in $\mathrm{supp}(c_2)$ and zero otherwise, and $x_i' = 1$ if the i-th coordinate in X' is in $\mathrm{supp}(c_2)$ and zero otherwise. Let $y_i := x_i - x_i'$. The above observation that $w(c_2+c+u) = 2^m$ for all $u \in H_{m|m} \setminus \{1, e_1, e_2, e_2+e_1\}$ translates to linear equations, one for each u. We do not pay attention to the first $2^{m-1} - 1$ pairs of variables, as the role of the corresponding coordinates are symmetrical, thus it makes no difference where the ones in c_2 are in those coordinates: we can redefine $X \cup X'$ if need be so that the code $\langle H_{m|m}, c\rangle$ be unaffected. More importantly, we are more interested in showing that there is a very limited number of possibilities for the position of ones in the last 2^{m-1} pairs of coordinates. So we produce a system of linear equations with variables x_i, x_i' where $2^{m-1} \le i \le 2^m - 1$. Note that in all such positions i for all $u \in H_{m|m} \setminus \{1, e_1, e_2, e_2 + e_1\}$, the codeword $c_2 + c + u$ has opposite values in the i-th coordinate of X and that of X'. If the former coordinate is 1 and the latter is 0, then the contribution of the i-th pair of coordinates to the weight of $c_2 + c + u$ is $1 - x_i + x_{i'} = 1 - y_i$, and if the former coordinate is 0 and the latter is 1, then the contribution of the i-th pair of coordinates to the weight of $c_2 + c + u$ is $x_i + 1 - x_{i'} = 1 + y_i$. This can be summarized in the formula $1 + (-1)^{u[i]} y_i$, where $u[i]$ is the i-th coordinate of u in X, which is the same as that in $c + u$ in X. There are altogether 2^{m-1} ones in pairs of coordinates with index $i \le 2^{m-1} - 1$ in $c_2 + c + u$, and the above expressions $1 + (-1)^{u[i]} y_i$ contribute 2^{m-1} summands 1 in the left hand side of the equation. The right hand side of the equation corresponding to u is 2^m, as we have seen above that $w(c_2 + c + u) = 2^m$. Thus for all $u \in H_{m|m} \setminus \{1, e_1, e_2, e_2 + e_1\}$, we obtain the linear equation

$$\sum_{i=2^{m-1}}^{2^m-1} (-1)^{u[i]} y_i = 0$$

by double counting. If we arrange the vector u into consecutive pairs, then inside every pair we obtain essentially the same equation: namely, one can be obtained from the other by multiplication with (-1), since two such consecutive vectors complement each other in the coordinates $2^{m-1} \le i \le 2^m - 1$. Thus we can erase every other equation. Then we obtain an Hadamard matrix with two rows missing as the matrix of coefficients: indeed, if we produce the matrix with entries $(-1)^{u[i]}$ for all $u \in H_{m|m}$, i.e., including the first four vectors as well, where $2^{m-1} \le i \le 2^m - 1$, and delete every other row, then we obtain an Hadamard matrix. Hadamard matrices are invertible, thus the punctured matrix obtained by the omission of the first two rows has co-rank 2. Clearly, all vectors with $y_{2^{m-1}} = \cdots = y_{2^{m-1}+2^{m-2}-1}$ and $y_{2^{m-1}+2^{m-2}} = \cdots = y_{2^m-1}$ satisfy the system of linear equations. As these conditions define a co-rank 2 subspace in $\mathbb{Q}^{2^{m-1}}$, the conditions are equivalent to the system of linear equations. As the vector c_2 has exactly 2^{m-2} ones in pairs of coordinates in $X \cup X'$ with index $2^{m-1} \le i \le 2^m - 1$, exactly one out of the following four possibilities occur:

- $x_{2^{m-1}} = \cdots = x_{2^{m-1}+2^{m-2}-1} = 1$ and the remaining x_i, x_i' are 1 for $2^{m-1} \leq i \leq 2^m - 1$, or
- $x_{2^{m-1}+2^{m-2}} = \cdots = x_{2^m-1} = 1$ and the remaining x_i, x_i' are 1 for $2^{m-1} \leq i \leq 2^m - 1$, or
- $x'_{2^{m-1}} = \cdots = x'_{2^{m-1}+2^{m-2}-1} = 1$ and the remaining x_i, x_i' are 1 for $2^{m-1} \leq i \leq 2^m - 1$, or
- $x'_{2^{m-1}+2^{m-2}} = \cdots = x'_{2^m-1} = 1$ and the remaining x_i, x_i' are 1 for $2^{m-1} \leq i \leq 2^m - 1$, or

By replacing c_2 with its mirror image if necessary (also contained in $\langle H_{m|m}, c_2\rangle$), that is, switching the roles of X_2 and X_2', we may assume that we are in one of the last two possibilities. Note that in particular $|\operatorname{supp}(c_2) \cap \operatorname{supp}(c) \cap X'| = 2^{m-2}$. After a suitable rearrangement of the symmetrical partition $X \cup X'$ to $X_3 \cup X_3'$ that does not affect the code $\langle H_{m|m}, c\rangle$ and only involves potential transposition of pairs of coordinates with index $1 \leq i \leq 2^{m-1} - 1$, we obtain that c_2 is the restriction of e_2 to X_3'. But then $\langle H_{m|m}, c, c_2\rangle$ together with the modified symmetrical partition $X_3 \cup X_3'$ is a code between $H_{m|m}$ and $H_{m\times m}$. By maximality of the dimension of C_0, we have that there must be at least one more $H_{m|m}$-balanced vector $c' \in C_0$ different from c. It cannot be the restriction of e_2 or $e_2 + e_1$ to X': in that case, the support of c' or $c + c'$ would intersect the same pairs of equivalent coordinates nontrivially as the support of c_2, which was earlier shown to be impossible (in the above arguments, c was an arbitrary $H_{m|m}$-balanced vector in C_0). Without loss of generality, c' is the restriction of e_3 to X'. In particular, the transposition of pairs of coordinates with index $1 \leq i \leq 2^{m-1} - 1$ to obtain the new symmetrical partition $X_3 \cup X_3'$ could not have involved pairs with indices $2^{m-2} \leq i \leq 2^{m-1} - 1$, as the same argument as we have applied for c yields that $|\operatorname{supp}(c_2) \cap \operatorname{supp}(c') \cap X'| = 2^{m-2}$. Then we can again rearrange $X_3 \cup X_3'$ to some $X_4 \cup X_4'$ by transposing pairs with indices $1 \leq i \leq 2^{m-2} - 1$, and obtain that c_2 is the restriction of e_2 to X_4'. Again, this means that $\langle H_{m|m}, c', c, c_2\rangle$ is a good candidate for C_0, thus C_0 itself must contain at least one more $H_{m|m}$-balanced vector $c'' \in C_0$ that is the restriction of e_3 to X', without loss of generality. By carrying on in the same fashion, after m steps, we obtain a symmetrical partition $X_{m+2} \cup X'_{m+2}$ such that $\langle H_{m|m}, c^{(m-1)}, \ldots, c', c, c_2\rangle$ is between $H_{m|m}$ and $H_{m\times m}$ with respect to $X_{m+2} \cup X'_{m+2}$, and this has the same dimension as $H_{m\times m}$, which is the biggest dimension that C_0 can possibly have. Thus $C_0 = H_{m\times m}$ (with the symmetrical partition $X_{m+2} \cup X'_{m+2}$), and $c_2 \in C_0$, a contradiction. ∎

Now we focus on the $k > m$ case. According to Proposition 3.4, we only need to show that there are no unknown examples for $2 \leq m$, $k = m + 1$.

Lemma 4.2. *Let $2 \leq m$ and let $H_{m+1|m} \leq C \not\leq H_{(m+1)\times m}$ (with respect to any symmetrical partition) be a minimal example (cf. Definition 3.3 and item 3. of Proposition 3.4). Then C cannot be extended to a code C' with the same length n and maximum distance D.*

Proof. Let C_0 denote the copy of $H_{m+1|m}$ in C. Let $C = \langle C_0, c\rangle$ with some codeword c that is C_0-balanced with respect to the symmetrical partition $X \cup Y \cup X'$. Let $c' \in C' \setminus C$; we may assume that $C' = \langle C, c'\rangle$. Then $\langle C_0, c'\rangle$ is a minimal example, thus either $C_0 \leq \langle C_0, c'\rangle \leq H_{(m+1)\times m}$ with respect to some symmetrical partition (potentially different from $X \cup Y \cup X'$), or $\langle C_0, c'\rangle$ is as in item 3. of Proposition 3.4.

Assume first that $c' \in H_{(m+1)\times m}$ with respect to some symmetrical partition. We may assume that $\operatorname{supp}(c') \cap Y = \emptyset$, and then $w(c') = 2^{m-1}$.

There are four codewords $u \in C \setminus C_0$ with weight 2^m, and all four has 2^{m-2} ones in Y. Thus $w(c' + u + 1_Y) = w(c' + u) + |Y| - 2 \cdot 2^{m-2} = w(c' + u) + 2^{m-1} \leq D$, which makes $w(c'+u) \leq D - 2^{m-1} = w(u)$. Clearly, if $u \in C \setminus C_0$ has weight D, then $w(c'+u) \leq w(u)$. Thus $w(c'+u) \leq w(u)$ for all $u \in C \setminus C_0$, and $\sum_{u \in C \setminus C_0} w(c'+u) = \sum_{u \in C \setminus C_0} w(u)$. Hence, $w(c'+u) = w(u)$ for all $u \in C \setminus C_0$, and consequently, the support of any $u \in C \setminus C_0$ cuts the support of c' in half.

This yields a system of linear equations over $\mathbb{Q}$. Introduce pairs of variables corresponding to the pairs of C_0-equivalent coordinates denoted by $x_1, x'_1,\ x_2, x'_2, \cdots, x_{2^m-1}, x'_{2^m-1}$ with $x_1, \ldots, x_{2^m-1}$ corresponding to coordinates in X, such that $x_i = 1$ if the i-th coordinate in X is in $\operatorname{supp}(c')$ and zero otherwise, and $x'_i = 1$ if the i-th coordinate in X' is in $\operatorname{supp}(c')$ and zero otherwise. Then each $u \in C \setminus C_0$ yields a linear equation by equating the sum of variables corresponding to $\operatorname{supp}(u)$ in X with $\frac{w(c')}{2} = 2^{m-1}$.

Given an $1 \leq i \leq 2^m - 1$, let us add the linear equations corresponding to the 2^m codewords $u \in C \setminus C_0$ such that the value of u is one in the i-th coordinate of X, and subtract the remaining 2^m equations. If we did this with codewords in C_0, then x_i and x'_i would have coefficient 2^m and all the remaining variables would have coefficient 0, thus yielding the equation $2^m \cdot (x_i + x'_i) = 0$. As $C \setminus C_0 = c + C_0$, thus zeros and ones are flipped in the support of c, the equation obtained is of the form $2^m \cdot (x_i + x'_i) = 0$ if i' is not in the support of c, and it is of the form $2^m \cdot (x_i - x'_i) = 0$ if i' is in the support of c. Thus $x_i + x'_i = 0$, or equivalently $x_i = x'_i = 0$ for all the i such that i' is not in the support of c, and $x_i = x'_i$ for all the i such that i' is in the support of c. As c' has nonzero coordinates in $X \cup X'$, the latter possibility occurs with some i such that $x_i = x'_i = 1$. But then there are C_0-equivalent coordinates where c' is one, a contradiction.

Hence, $\langle C_0, c' \rangle$ is as in item 3. of Proposition 3.4 for all $c' \in C' \setminus C_0$. That is, if we partition C' into C_0-cosets $C' = C_0 \cup K \cup K' \cup K''$, then there are C_0-balanced vectors each of $c \in K, c' \in K'$ and $c'' \in K''$ (with respect to possibly different symmetrical partitions), where c and c' have already been chosen along with the symmetrical partition $X \cup Y \cup X'$ corresponding to c. Let $a, b \in C_0$ be as in Definition 2.5 for c.

All nonzero codewords in C' have weight 2^m or $3 \cdot 2^{m-1}$. The four codewords in each of K, K' and K'' with weight 2^m are exactly those u with $|\operatorname{supp}(u) \cap Y| = 2^{m-2}$. In each of K, K' and K'', these four sets of the form $\operatorname{supp}(u) \cap Y$ partition Y. Given the intersection of two maximal weight codewords in Y as in Definition 2.5, if we produce the C_0-balanced vector and its C_0-translates with weight 2^m, the partition obtained either coincides with the above one, or the two partitions bisect each other (i.e., their intersection consists of eight classes with half the size of the original classes). Clearly, these intersections bisect each other, as otherwise there were two vectors $u \in K, u' \in K'$ with the same support inside Y, and then $u + u' \in K''$ would be all zero in Y, a contradiction. In particular, $3 \leq m$. Moreover, we cannot choose the same pair u, v to define c', but we may assume that $\operatorname{supp}(c) \cap Y$ and $\operatorname{supp}(c') \cap Y$ cut each other in half, and in particular that $c'' = c + c'$. However we pick a pair u', v' to define c' so that this condition is met, we obtain equivalent binary linear codes.

So we are going to work in a particular example, for the sake of transparency. First of all, let us represent $C_0 \cong H_{m+1|m}$ in the standard way. The illustration below is for $m = 3$.

	X							Y								X'								
C_0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	1_Y	
	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	a	
	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1		
	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	1	1	0	0	1	1	b	
	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0	0	1	1	0	0	1	1		
	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	0	0		
	0	1	1	1	1	0	0	1	1	0	0	0	0	1	1	0	1	1	1	1	0	0		
	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	b'	
	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1		
	1	0	1	1	0	1	0	0	1	0	1	1	0	1	0	1	0	1	1	0	1	0		
	1	0	1	1	0	1	0	1	0	1	0	0	1	0	1	1	0	1	1	0	1	0		
	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	1	1	0	0	1	1	0		
	1	1	0	0	1	1	0	1	0	0	1	1	0	0	1	1	1	0	0	1	1	0		
	1	1	0	1	0	0	1	0	1	1	0	1	0	0	1	1	1	0	1	0	0	1		
	1	1	0	1	0	0	1	1	0	0	1	0	1	1	0	1	1	0	1	0	0	1		
K	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1	1	1	c	special
	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1		
	0	0	0	1	1	1	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0		
	0	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	0	0	0	0		
	0	1	1	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0		special
	0	1	1	0	0	1	1	1	1	0	0	1	1	1	1	0	0	0	1	1	0	0		
	0	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	1	1		
	0	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1		
	1	0	1	0	1	0	1	0	1	0	1	0	1	1	0	1	1	0	1	0	1	0		
	1	0	1	0	1	0	1	1	0	1	0	1	0	0	1	1	1	0	1	0	1	0		
	1	0	1	1	0	1	0	0	1	0	1	1	0	0	1	1	1	0	0	1	0	1		
	1	0	1	1	0	1	0	1	0	1	0	0	1	1	0	1	1	0	0	1	0	1		
	1	1	0	0	1	1	0	0	1	1	0	0	1	0	1	1	0	1	1	0	0	1		
	1	1	0	0	1	1	0	1	0	0	1	1	0	1	0	1	0	1	1	0	0	1		
	1	1	0	1	0	0	1	0	1	1	0	1	0	1	0	1	0	1	0	1	1	0		
	1	1	0	1	0	0	1	1	0	0	1	0	1	0	1	1	0	1	0	1	1	0		

$2^{m-1} \le i \le 2^m - 1$

Hence, we produce the generating matrix as in Section 2 (see $M_{3|2}$), by writing the binary representation of all numbers from 1 to $2^{m+1}-1$ in columns, and then by repeating the first $2^m - 1$ columns. Let $e_1, \dots, e_{m+1}$ denote the rows of this matrix from top to bottom; this is the standard basis of the code. Then we sort the codewords $\sum_{i=1}^{m+1} \varepsilon_i e_i$, $\varepsilon_i \in \{0,1\}$, so that the sequence of coefficients $\overline{\varepsilon_1 \cdots \varepsilon_{m+1}}$ corresponding to the r-tk codeword is the binary representation of r (extended by zeros on the right) for $r = 0, \dots, 2^{m+1}-1$. That is, the list of codewords is $0, e_1 = 1_Y, e_2, e_2 + e_1, e_3, \dots, e_{m+1} + \dots + e_1$. Let $a = a' = e_2, b = e_3, b' = e_4$ (note that this is possible as $3 \le m$). So $\langle C_0, c\rangle$ is uniquely determined (together with the fixed symmetrical partition of coordinates), and so are the vectors a', b' in C_0 to define c'. The difficulty in showing that this cannot yield an appropriate code is that there are 2^m possibilities for the symmetrical partition corresponding to c'.

Considering the way we represented the code C_0, note that the restriction of

This manuscript is for review purposes only.

codewords to Y have an alternating nature: $\mathrm{supp}(u_{2k-1}) \cap Y = Y \setminus (\mathrm{supp}(u_{2k}) \cap Y)$. Inside X (and symmetrically inside X'), the restrictions to $\mathrm{supp}(a) \cap X$ is a list of identical pairs of vectors, and the restrictions to $(\mathrm{supp}(b) \setminus \mathrm{supp}(a)) \cap X$ is a list of identical quartets of vectors. Since $\mathrm{supp}(c) \cap X = \emptyset$, the same holds for the coset K (whose elements $c + u$ are listed in the same order as the vectors $u \in C_0$ are). That is, the list can be partitioned into consecutive quartets with the same restriction to $(\mathrm{supp}(b) \setminus \mathrm{supp}(a)) \cap X$, and each quartet consists of two consecutive pairs with the same restriction to $\mathrm{supp}(a) \cap X$. It is easy to see that the eight codewords in $c+u \in K$ such that $\mathrm{supp}(c + c' + u) \cap Y$ has size 2^{m-2} or $3 \cdot 2^{m-1}$, which are exactly those codewords in K'' whose weight in $X \cup X'$ is $3 \cdot 2^{m-2}$ rather than 2^m, is the union of two such quartets. Thus the indices of these eight vectors are independent from the choice of the symmetrical partition corresponding to c', as we can find them by only studying the restriction of vectors to Y.

Introduce pairs of variables corresponding to the pairs of C_0-equivalent coordinates denoted by x_1, x_1', x_2, x_2', $\cdots, x_{2^m-1}, x_{2^m-1}'$ with $x_1, \ldots, x_{2^m-1}$ corresponding to coordinates in X, such that $x_i = 1$ if the i-th coordinate in X is in $\mathrm{supp}(c')$ and zero otherwise, and $x_i' = 1$ if the i-th coordinate in X' is in $\mathrm{supp}(c')$ and zero otherwise. As a first step, we are going to simplify the notations, so that it is enough to focus on the variables $x_1, \ldots, x_{2^m-1}$. Fist of all, if $1 \leq i \leq 2^m - 1$ is such that both $\mathrm{supp}(a) \cup \mathrm{supp}(b)$ and $\mathrm{supp}(a) \cup \mathrm{supp}(b')$ are one in the i-th coordinate of X', then $x_i + x_i' = 1$ and every $c + u \in K$ has opposite values in the i-th coordinate in X and the i-th coordinate of X', respectively. Note that this applies exactly to $3 \cdot 2^{m-3} \leq i \leq 2^m - 1$. Thus if the i-th coordinate of $c + u$ in X is $u[i] = 0$, then the sum of the i-th coordinates in X and in X' of $c' + c + u \in K''$ (as rational numbers rather than elements of $\mathbb{Z}_2$) is

- 0 if $x_i = 0$, and
- 2 if $x_i = 1$.

Similarly, if the i-th coordinate of $c + u$ in X is $u[i] = 1$, then the sum of the i-th coordinates in X and in X' of $c' + c + u \in K''$ (as rational numbers rather than elements of $\mathbb{Z}_2$) is

- 2 if $x_i = 0$, and
- 0 if $x_i = 1$.

Hence, the sum of the i-th coordinates in X and in X' of $c' + c + u \in K''$ is $2u[i] + 2(-1)^{u[i]} \cdot x_i$ for all $3 \cdot 2^{m-3} \leq i \leq 2^m - 1$. In case of the remaining values $1 \leq i \leq 3 \cdot 2^{m-3}$, the choice of the symmetrical partition in the definition of c' does not affect the sum of the i-th coordinates in X and in X' of $c' + c + u \in K''$. Let us denote this sum by $s(u, i)$ for $1 \leq i \leq 3 \cdot 2^{m-3} - 1$.

Then each codeword $c + u \in K$ yields a linear equation. Namely, if $c + u$ is one of the eight codewords with either 2^{m-2} or $3 \cdot 2^{m-2}$ ones in Y, then we have $\sum_{i=1}^{3\cdot 2^{m-3}-1} s(u,i) + \sum_{i=3\cdot 2^{m-3}}^{2^m-1} (2u[i] + 2(-1)^{u[i]} \cdot x_i) = 3 \cdot 2^{m-2}$, and in case of the rest of the codewords in K, the equation is $\sum_{i=1}^{3\cdot 2^{m-3}-1} s(u,i) + \sum_{i=3\cdot 2^{m-3}}^{2^m-1} (2u[i] + 2(-1)^{u[i]} \cdot x_i) = 2^m$. After rearranging the equations, we obtain

- $\sum_{i=3\cdot 2^{m-3}}^{2^m-1} 2(-1)^{u[i]} \cdot x_i = 3 \cdot 2^{m-2} - \left(\sum_{i=1}^{3\cdot 2^{m-3}-1} s(u,i) + \sum_{i=3\cdot 2^{m-3}}^{2^m-1} 2u[i] \right)$ in case of the eight special vectors u, and

- $\sum_{i=3\cdot 2^{m-3}}^{2^m-1} 2(-1)^{u[i]} \cdot x_i = 2^m - \left(\sum_{i=1}^{3\cdot 2^{m-3}-1} s(u,i) + \sum_{i=3\cdot 2^{m-3}}^{2^m-1} 2u[i] \right)$ for the rest.

In each quartet, the identical restriction to X yield identical equations. So we obtain two different linear equation from each quartet.

We study the equations corresponding to the first quartet of vectors in K separately, as they are essentially different from the rest. The first equation (obtained from the first two vectors in K) is $\sum_{i=3\cdot 2^{m-3}}^{2^m-1} x_i = 2^{m-1}$, and the second equation is $\sum_{i=3\cdot 2^{m-3}}^{2^{m-1}-1} x_i - \sum_{i=\cdot 2^{m-1}}^{2^m-1} x_i = 2^{m-1} = -2^{m-1}$. By adding up these two linear equations, we obtain $\sum_{i=3\cdot 2^{m-3}}^{2^{m-1}-1} x_i = 0$. As all the x_i are non-negative rational numbers, this is only possible if $x_i = 0$ for all $3 \cdot 2^{m-3} \leq i \leq 2^{m-1} - 1$. Thus it is enough to focus on the variables x_i with $2^{m-1} \leq i \leq 2^m - 1$, and the equations

- $\sum_{i=2^{m-1}}^{2^m-1} (-1)^{u[i]} \cdot x_i = 3 \cdot 2^{m-3} - \left(\sum_{i=1}^{3\cdot 2^{m-3}-1} \frac{s(u,i)}{2} + \sum_{i=3\cdot 2^{m-3}}^{2^m-1} u[i] \right)$ in case of the eight special vectors u, and
- $\sum_{i=2^{m-1}}^{2^m-1} (-1)^{u[i]} \cdot x_i = 2^{m-1} - \left(\sum_{i=1}^{3\cdot 2^{m-3}-1} \frac{s(u,i)}{2} + \sum_{i=3\cdot 2^{m-3}}^{2^m-1} u[i] \right)$ for the rest.

For each remaining quartet, let us subtract the first equation from the second. Fortunately, the right sides of the two equations are equal: in all quartets (other than the first), the number of ones in X in the indices $3 \cdot 2^{m-2} \leq i \leq 2^m - 1$ is the same in all four vectors, and the restriction of the vectors to the first $1 \leq i \leq 2^{m-1}$ coordinates in X is also the same, making $s(u,i)$ independent from u (within a quartet). Thus the right hand side of the difference of equations is 0. On the left hand side, we have all the x_i with opposite sign in the two equations, as there are opposite coordinates in the region $2^{m-1} \leq i \leq 2^m - 1$ in X in the two different vectors of each quartet. After subtracting the two equations and dividing by 2, we obtain the same coefficients as if we simply subtracted the restrictions of the two vectors in K to the coordinates $2^{m-1} \leq i \leq 2^m - 1$ in X (where the 0-1 vectors are considered as rational vectors). If we do this for all quartets, including the first, then the coefficients in the 2^m equations obtained form an Hadamard matrix. On the right hand side, we have 2^{m-2} in the first equation, and 0 everywhere else. Since Hadamard matrices are invertible, this system of linear equations has a unique solution in $\mathbb{Q}^{2^{m-1}}$. As $x_{2^{m-1}} = \cdots = x_{2^m-1} = \frac{1}{2}$ is obviously a solution, this is the unique solution of the system of linear equations obtained. However, each x_i should be 0 or 1, a contradiction. □

The proof of the main theorem is now complete.

Proof of Theorem 2.6. By Theorem 2.2, Proposition 3.4 and Lemmas 4.1, 4.2. □

5. Further comments. Although our sole purpose was to (nearly) minimize the maximum distance of a binary linear code, the codes obtained turn out to have a relatively large minimum distance. According to the Plotkin bound [15], a binary linear code C with length n and minimum distance d such that $n = 2d$ has dimension $\dim C \leq 1 + \lfloor \log_2 n \rfloor$. This upper bound is attained by the codes $C = \langle H_{m|m}, c \rangle$, where c is the $H_{m|m}$-balanced vector that is all one in X', given a symmetrical partition $X \cup X'$ of the coordinates. Indeed, $\dim C = m + 1$, $d = 2^m - 1$ and $n = 2d = 2^{m+1} - 2$, thus $\lfloor \log_2 n \rfloor = m$. Moreover, these codes also meet the Griesmer bound [8]:

$\sum_{i=0}^{m}\left\lceil\frac{2^m-1}{2^i}\right\rceil = 2^m-1+\sum_{i=1}^{m}2^{m-i} = 2^m-1+2^m-1 = n$. We note that the Griesmer bound is also attained by the code $C = \langle H_{m|m}, c, 1\rangle$ where 1 is the all one vector. In that case, $\dim C = m+1$ and the minimum distance is $d = 2^m-2$.

Again, by the Griesmer bound, a binary linear code of length $n = 10$ and dimension $\dim C = 4$ cannot have minimum distance $d \geq 5$. The optimal minimum distance $d = 4$ is attained by $C = \langle H_{3|2}, c\rangle$ with any $H_{3|2}$-compatible c. In fact, we can improve the dimension by once again extend the code by the all one vector 1, to obtain a $[10,5,4]_2$ code. This example cannot be further improved in the sense that there is no $[10,6,4]_2$ code. According to [22], there are exactly four inequivalent binary linear codes with parameters $[10,5,4]_2$; the above example C is Code 2 in that document. It is noted in [22] that C is not self-dual. However, the dual of C has the same weight distribution as C, and thus - as the remaining three examples have different weight distribution - we have $C \cong C^{\perp}$. It is also mentioned in [22] that according to the Assmus-Mattson theorem [9, Theorem 8.4.2], the supports of the weight 4 codewords in C form a $2-(10,4,2)$ block design.

The concepts of two- and three-weight codes are getting more and more popular recently, see [5, 11, 12, 23]. Every code of the form $C = \langle H_{m|m}, c\rangle$, where c is an $H_{m|m}$-balanced vector, is a two-weight code. According to Lemma 3.2, every code of the form $C = \langle H_{m+1|m}, c\rangle$, where c is an $H_{m+1|m}$-balanced vector, is also a two-weight code. Furthermore, for $m = 2$, the latter example can be extended by the all one vector to obtain a three-weight binary linear code. For all $1 \leq m < k$, $H_{k|m}$ is a two-weight code, and the trivial examples $H_{k|m} < C \leq H_{k\times m}$ are three-weight codes.

REFERENCES

[1] O. Ahlman and V. Koponen, *Limit laws and automorphism groups of random nonrigid structures*, J. Log. Anal., 7 (2015), pp. 1–53.

[2] S. M. Ball and A. Blokhuis, *A bound for the maximum weight of a linear code*, SIAM J. Discrete Math., 27 (2013), pp. 575–583.

[3] A. Betten, M. Braun, H. Fripertinger, A. Kerber, A. Kohnert, and A. Wassermann, *Error-Correcting Linear Codes*, Springer, 2006.

[4] P. J. Cameron, *On graphs with given automorphism group*, European J. Combin., 1 (1980), pp. 91–96.

[5] K. Ding and C. Ding, *A class of two-weight and three-weight codes and their applications in secret sharing*, Trans. Inform. Theory, 61 (2015), pp. 5835–5842.

[6] P. Erdős and A. Rényi, *Asymmetric graphs*, Acta Math. Hungar., 14 (1963), pp. 295–315.

[7] R. Fagin, *The number of finite relational structures*, Discrete Math., 19 (1977), pp. 17–21.

[8] J. H. Griesmer, *A bound for error-correcting codes*, IBM Journal of Res. and Dev., 4 (1960), pp. 532–542.

[9] W. C. Huffman and V. Pless, *Fundamentals of Error-Correcting Codes*, Cambridge Universty Press, 2003.

[10] V. Koponen, *Typical automorphism groups of finite nonrigid structures*, Arch. Math. Logic, 54 (2015), pp. 571–586.

[11] C. Li, S. Bae, and S. Yang, *Some two-weight and three-weight linear codes*, Adv. Math. Commun., 13 (2019), pp. 195–211.

[12] C. Li, Q. Yue, and F. W. Fu, *A construction of several classes of two-weight and three-weight linear codes*, Appl. Algebra Engrg. Comm. Comput., 28 (2017), pp. 11–30.

[13] M. W. Liebeck and A. Shalev, *On fixed points of elements in primitive permutation groups*, J. Algebra, 421 (2015), pp. 438–459.

[14] P. M. Neumann, *A lemma that is not Burnside's*, The Mathematical Scientist, 4 (1979), pp. 133–141.

[15] M. Plotkin, *Binary codes with specified minimum distance*, Trans. Inform. Theory, 6 (1960), pp. 445–450.

[16] A. Pongrácz, *Extremal solutions of an inequality concerning supports of permutation groups*

and punctured Hadamard codes. submitted, 2018.
[17] C. Ronse, *On permutation groups of prime power order*, Math. Z., 173 (1980), pp. 211–215.
[18] R. Roth, *Introduction to coding theory*, Cambridge Universty Press, 2006.
[19] J. Saxl and A. Shalev, *The fixity of permutation groups*, J. Algebra, 174 (1995), pp. 1122–1140.
[20] A. Shalev, *On the fixity of linear groups*, Proc. Lond. Math. Soc., 68 (1994), pp. 265–293.
[21] J. J. Sylvester, *Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to newton's rule, ornamental tile-work, and the theory of numbers*, Philosophical Magazine, 34 (1867), pp. 461–475.
[22] H. N. Ward, *The four [10,5,4] binary codes*. The manuscript can be found on the homepage http://www.people.virginia.edu/˜hnw/Four104.pdf.
[23] Z. Zhou, N. Li, C. Fan, and T. Helleseth, *Linear codes with two or three weights from quadratic Bent functions*, Des. Codes Cryptogr., 81 (2016), pp. 283–295.