# Existence of a maximum balanced matching in the hypercube

Gyula O.H. Katona [*]
Rényi Institute
Budapest, Hungary
ohkatona@renyi.hu

Krisztián Tichler [†]
Eötvös University
Budapest, Hungary
ktichler@inf.elte.hu

June 27, 2012

**Abstract**

We prove, that for $n \neq 2$ the maximum possible $\lfloor 2^n/2n \rfloor$ edges can be chosen simultaneously from each parallel class of the $n$-cube in such a way, that no two edges have a common vertex.

## 1 Introduction

We consider the following problem for the $n$ dimensional hypercube. Select as many edges as possible from each parallel class simultaneously in such a way, that the set of edges form a matching of the hypercube. Here, matching is a subset of the edges, such that no two edges have a common vertex. More precisely, among all matchings of the hypercube maximize the minimum number of edges of the $n$ parallel classes of the edges. Obviously, no more than $\lfloor 2^n/2n \rfloor$ is possible, since each $n$ edges of a matching, one from each parallel class, need $2n$ of the $2^n$ vertices of the hypercube. A matching is called a *maximum balanced matching* if it contains $\lfloor 2^n/2n \rfloor$ edges from each parallel class. Our main result is the following.

[*]The first author was supported by the Hungarian National Foundation for Scientific Research, grant number NK78439.

[†]This article is supported by the European Union and co-financed by the European Social Fund (ELTE TÁMOP-4.2.2/B-10/1-2010-0030).

**Theorem 1.1.** *There exists a maximum balanced matching of the $n$-cube for $n \neq 2$.*

The problem emerged as a possible solution for a question of the authors ([2]) in combinatorial search theory.

There is a similar, well examined problem. List all words of length $n$ over the binary alphabet $\Sigma = \{0, 1\}$ in such a way, that for each word the succeeding word differs only by a single bit, that is for each consecutive pair of words their Hamming distance is 1. (The Hamming distance of words $u = t_1 \cdots t_n$ and $v = t'_1 \cdots t'_n$ over the alphabet $\Sigma$ is defined by $H(u, v) = |\{i \in \{1, \dots, n\} \,|\, t_i \neq t'_i\}|$.) In other formulation, construct a Hamiltonian path (or cycle) in the $n$ dimensional hypercube.

One such Hamiltonian cycle for the $n$-cube is generated recursively from the Hamiltonian cycle for the $(n-1)$-cube. Take the same Hamiltonian path (eliminate an edge from the Hamiltonian cycle for $n - 1$) in two parallel hyperplanes and add two edges, that connect their first and last vertices. This results in a Hamiltonian cycle for the $n$-cube. For the list of words this construction corresponds to the following recursive recipe: take two copies of the list for the words of length $n-1$, add a 0 prefix to each word in the first copy, reflect the order of the words in the second copy of the list and add a 1 prefix to each word, concatenate the two modified lists to get the list for word length $n$.

This list of words is called the *binary-reflected Gray code*. The name "Gray" refers to F. Gray, who patented this list of words as a solution to a communication problem involving digitization of analogue data ([3]).

More generally, any Hamiltonian path (cycle) in the $n$-cube is called a *(cyclic) Gray code*. There are many papers on Grey codes satisfying certain properties, for a survey see [1].

A long standing open problem on Gray codes was to construct a (cyclic) balanced one, i.e., one that contains a balanced number of edges from each of the $n$ parallel classes of edges. Since the number of edges in each parallel class must be even for a cyclic Gray code, the smallest possible positive difference is two. So for word lengths of non-2-powers, a balanced Grey code must have either the smallest even integer larger, or the largest even integer smaller than $2^n/n$ edges in each parallel class. Finally, G. S. Bhat and C. D. Savage ([4]) constructed a balanced Gray code for all $n$ using a proposed construction of J. Robinson and M. Cohn ([5]).

Note, that despite the similarity neither a balanced Grey code, nor a maximum balanced matching imply the existence of the other.

In section 2 we introduce some notations and prove our main lemma in proving Theorem 1.1. We complete its proof in section 3. In section 4 we introduce a generalization of the problem and prove some initial results in section 5. However, the problem remains open in general.

# 2 Balanced cycle cover of the hypercube

First, let us introduce some notations. Let $[n] = \{1, \ldots, n\}$ and $\binom{[n]}{r} = \{S \subseteq [n] \,|\, |S| = r\}$. Furthermore let $[\![x]\!]_r = r\lfloor x/r \rfloor$. If $r = 2$ we write shortly $[\![x]\!]$ instead of $[\![x]\!]_2$. If $\Sigma$ is an alphabet let $\Sigma^n$ denote the set of words of length $n$ over $\Sigma$.

Let $B_n$ be the $n$ dimensional hypercube, $B_n = \langle V(B_n), \mathcal{E}(B_n)\rangle$, where $V(B_n) = \{0,1\}^n$ is the set of binary words of length $n$ and $\mathcal{E}(B_n) = \{\{u,v\} \,|\, H(u,v) = 1\}$.

$\mathcal{E} = \mathcal{E}(B_n)$ has a natural decomposition $\mathcal{E} = \cup_{i=1}^n \mathcal{E}_i$ according to the directions, formally

$$\{b_1 \cdots b_n, b'_1 \cdots b'_n\} \in \mathcal{E}_i \text{ if and only if } b_j = b'_j, j \neq i \text{ and } b_i \neq b'_i.$$

For $\mathcal{E}' \subseteq \mathcal{E}$ and $i \in [n]$ let

$$\lambda_i = \lambda_i(\mathcal{E}') = |\mathcal{E}' \cap \mathcal{E}_i|$$

furthermore let

$$\chi(\mathcal{E}') = (\lambda_1, \ldots, \lambda_n)$$

be the *profile vector* of $\mathcal{E}'$.

For a subgraph $G = \langle V, \mathcal{E}\rangle$ of $B_n$ and $b \in \{0,1\}$ let

$$G^b = \langle \{vb \,|\, v \in V\}, \{\{v_1 b, v_2 b\} \,|\, \{v_1, v_2\} \in \mathcal{E}\}\rangle.$$

If $\mathcal{G} = \{G_1, \ldots, G_k\}$, then let $\mathcal{G}^b = \{G_1^b, \ldots, G_k^b\}$ and $\mathcal{E}(\mathcal{G}) = \bigcup_{i=1}^k \mathcal{E}(G_i)$.

For $v = b_1 \cdots b_n \in V(B_n)$ let

$$\sigma_i(v) = b_1 \cdots b_{i-1} \bar{b}_i b_{i+1} \cdots b_n \quad (\bar{b} = 1 - b).$$

If $E \in \mathcal{E}(B_n)$ let

$$\sigma_i(E) = \begin{cases} \{\sigma_i(u), \sigma_i(v)\} & \text{if } \{u,v\} \notin \mathcal{E}_i \\ \{u,v\} & \text{if } \{u,v\} \in \mathcal{E}_i \end{cases}.$$

Let us introduce the notations $\sigma_i(V') = \{\sigma_i(v) \,|\, v \in V'\}$ for $V' \subseteq V(B_n)$ and $\sigma_i(\mathcal{E}') = \{\sigma_i(E) \,|\, E \in \mathcal{E}'\}$ for $\mathcal{E}' \subseteq \mathcal{E}(B_n)$. Given a subgraph $G = \langle V, \mathcal{E}\rangle$ of $B_n$ let $\sigma_i(G) = \langle \sigma_i(V), \sigma_i(\mathcal{E})\rangle$. So $\sigma_i$ gives nothing else, but the mirror image w.r.t. direction $i$.

We know ([4]) that, there exists a balanced Grey code. On one hand, the following lemma states less, the existence of a balanced cover of cycles instead of a single balanced Hamiltonian cycle. On the other hand, the lemma gives us a small, specific cycle, containing edges in all direction, that will be used for correcting a later specified almost balanced matching.

**Lemma 2.1.** *For $n \geq 3$ there exist a set of cycles $\mathcal{C}_n = \{C_0, C_1, \ldots, C_t\}$ of $B_n$ for some $t = t(n)$ having the following properties.*

*(i)* $\bigcup_{i=0}^{t} V(C_i) = V(B_n)$,

*(ii)* $V(C_i) \cap V(C_j) = \emptyset$ $(i \neq j; 0 \leq i, j \leq t)$,

*(iii)* $C_0 = (v_1, E_1, \ldots, v_{2n}, E_{2n})$, $E_i = \{v_i, v_{i(mod\ 2n)+1}\}$ $(i \in [2n])$, $E_i, E_{2n-i} \in \mathcal{E}_i, (i \in [n-1])$, $E_n, E_{2n} \in \mathcal{E}_n$,

*(iv) let* $\lambda_i = \lambda_i(\mathcal{E}(\mathcal{C}_n))$, *then* $|\lambda_i - \lambda_j| \leq 2$ $(1 \leq i, j \leq n)$.

A set of cycles satisfying $(i) - (iv)$ is called a *balanced cycle cover* (bcc).

Note, that since $B_n$ is a bipartite graph, it has only even cycles so the value of $\lambda_j$ is even as well $(1 \leq j \leq n)$. Furthermore, $\lambda_j(\mathcal{E}(C_i))$ is even, too, for $0 \leq i \leq t, 1 \leq j \leq n$.

Circuits of the form $(v_1, E, v_2, E), v_1, v_2 \in V(B_n), E = \{v_1, v_2\}, E \in \mathcal{E}(B_n)$ are considered to be cycles, as well.

*Proof of Lemma 2.1.* The proof is by induction. It is easy to construct a bcc for $n = 3$ or $n = 4$. Suppose that we have a bcc for $B_n$ and let us construct one for $B_{n+1}$.

The edges of $\mathcal{E}_{n+1}$ connect two disjoint copies of $B_n$ in $B_{n+1}$ since $\mathcal{E}_{n+1} = \{\{u0, u1\} \,|\, u \in \{0,1\}^n\}$. By the induction hypothesis there exist a bcc $\mathcal{C}_n = \{C_0, \ldots, C_t\}$ in $B_n$, so that it has a profile

$$\chi(\mathcal{E}(\mathcal{C}_n)) = (\lambda_1, \ldots, \lambda_n),$$

where $\lambda_1 = \cdots = \lambda_s, \lambda_{s+1} = \cdots \lambda_n, \lambda_{s+1} = \lambda_s + 2$, for some $s \in [n]$ and all $\lambda_i$'s are even.

Then let $\mathcal{C}$ be the following cover of $V(B_{n+1})$ by vertex disjoint cycles $\mathcal{C} = \mathcal{C}_n^0 \cup \mathcal{C}_n^1 = \{C_0^0, \dots, C_t^0, C_0^1 \dots, C_t^1\}$. So $C_0^b = \{v_1 b, E_1^b, \dots, v_{2n} b, E_{2n}^b\}$, where $E_i^b = \{v_i b, v_{i (\mathrm{mod}\ 2n)+1} b\}$ $(b \in \{0,1\})$. By the induction hypothesis $E_i^b, E_{2n-i}^b \in \mathcal{E}_i$, $E_n^b, E_{2n}^b \in \mathcal{E}_n$ $(i \in [n-1], b \in \{0,1\})$.

Observe, that $\mathcal{C}$ has the property

$$C \in \mathcal{C} \Leftrightarrow \sigma_{n+1}(C) \in \mathcal{C}, \tag{1}$$

so

$$E \in \mathcal{E}(\mathcal{C}) \Leftrightarrow \sigma_{n+1}(E) \in \mathcal{E}(\mathcal{C}) \tag{2}$$

holds as well.

$\mathcal{C}$ has properties (i)-(ii), but does not satisfy properties (iii)-(iv). We have

$$\chi(\mathcal{E}(\mathcal{C})) = (2\lambda_1, \dots, 2\lambda_n, 0).$$

Replace $C_0^0$ and $C_0^1$ by two other cycles. Let the set of their edges be

$$\big\{E_1^0, \dots, E_n^0, \{v_{n+1}0, v_{n+1}1\}, E_n^1, \dots, E_1^1, \{v_1 0, v_1 1\}\big\} \tag{3}$$

and

$$\big\{E_{n+2}^0, \dots, E_{2n-1}^0, \{v_{2n}0, v_{2n}1\}, E_{2n-1}^1, \dots, E_{n+2}^1, \{v_{n+2}0, v_{n+2}1\}\big\}.$$

By renaming the cycles we get a set of vertex disjoint cycles $\{C_0, \dots, C_{2t+1}\}$ covering $V(B_{n+1})$, where $\mathcal{E}(C_0)$ equals (3). We use the same notation $\mathcal{C}$ for the new cycle system. Note, that $\mathcal{C}$ satisfies (i)-(iii) and (1). Furthermore,

$$\chi(\mathcal{E}(\mathcal{C})) = (2\lambda_1, \dots, 2\lambda_{n-2}, 2\lambda_{n-1} - 2, 2\lambda_n - 2, 4).$$

The first $n$ components of the profile vector differ by maximum 2 and are at least 4 for $n \geq 4$. Take an edge $E \in \mathcal{E}(\mathcal{C} \setminus \{C_0\})$ of $\mathcal{E}_i$ $(i \in [n])$, where $\lambda_i(\mathcal{E}(\mathcal{C}))$ is at least as large as any other component. W.l.o.g. suppose, that $E = \{u0, v0\}$ $(u, v \in \{0,1\}^n)$. Then $E' = \sigma_{n+1}(E) = \{u1, v1\} \in \mathcal{E}(\mathcal{C})$ holds as well by (2). Replace $E$ and $E'$ by $E'' = \{u0, u1\}$ and $E''' = \{v0, v1\}$ (see Figure 1). This transformation decreases $\lambda_i(\mathcal{E}(\mathcal{C}))$ by 2 and increases $\lambda_{n+1}(\mathcal{E}(\mathcal{C}))$ by 2, while properties (i)-(iii) still hold.

Observe, that if $E$ and $E'$ belong to different cycles

$$C_1 = (w_0, E_0, \dots, w_k, E_k) \text{ and } C_2 = \sigma_{n+1}(C_1) = (w_0', E_0', \dots, w_k', E_k')$$

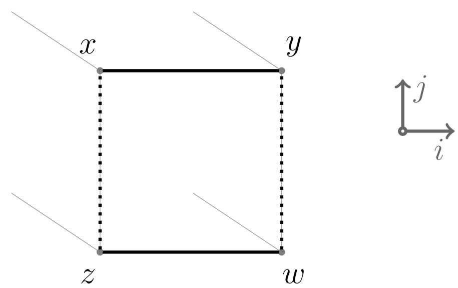


Figure 1: The following basic transformation is used many times. Suppose, that $\{x, y\}, \{z, w\} \in \mathcal{E}$ and both have color (direction) $i$, suppose furthermore, that $\{x, z\}, \{y, w\} \notin \mathcal{E}$ and both have color (direction) $j$, then flipping the pairs of edges decreases $\lambda_i$ by 2 and increases $\lambda_j$ by 2.

where $k \geq 1$, $w_0 = u0, w_1 = v0, E_0 = E, E'_0 = E', w'_i = \sigma_{n+1}(w_i)\,(0 \leq i \leq k)$, then $C_1$ and $C_2$ is replaced by a single, larger cycle

$$C = (w_1, E_1, \ldots, w_k, E_k, w_0, E'', w'_0, E'_k, w'_k, \ldots, E'_1, w'_1, E''').$$

On the other hand if $E$ and $E'$ are edges of the same cycle

$$C = (w_0, E_0, \ldots, w_k, E_k)$$

satisfying $\sigma_{n+1}(C) = C$, where $k \geq 3$, $E_0 = E, E_t = E'$ (for some $2 \leq t \leq k-1$), $w_0 = u0, w_1 = v0, w_t = v1, w_{t+1} = u1$, than $C$ is replaced by two smaller cycles

$$C_1 = (w_1, E_1, \ldots, w_{t-1}, E_{t-1}, w_t, E''') \text{ and } C_2 = (w_{t+1}, E_{t+1}, \ldots, w_k, E_k, w_0, E'').$$

Easy to check, that in both cases also (1) holds for the modified family of cycles. We use the same notation $\mathcal{C}$ for for the new cycle system.

Repeat the previous step until the cycle cover becomes balanced. We can do this, since the preconditions of the transformation (properties (i)-(iii) and (1)) still hold after each execution.

We also need, that there is at least one pair of edges not belonging to $C_0$ to flip. But this is true, since

$$|\mathcal{E}(C_0)| + \lambda_{n+1}(\mathcal{E}(\mathcal{C})) \leq 2n + 2 + \frac{2^{n+1}}{n+1} < 2^{n+1} = |\mathcal{E}(\mathcal{C})| \quad (n \geq 4).$$

For that actual $\mathcal{C}$ let $\mathcal{C}_{n+1} = \mathcal{C}$. Properties (i)-(iv) hold for $\mathcal{C}_{n+1}$. $\square$

# 3 Maximum balanced matching in the hypercube

## 3.1 Case of $n \leq 7$

For $n = 1$ and $n = 2$ the statement is obvious. For $n = 3$ a possible solution is to take the even edges of the following Grey code (Hamiltonian path) $G(3) = v_0, v_1, \ldots, v_7$.

| | | | |
|---|---|---|---|
| $v_0$=000 | $v_2$=011 | $v_4$=101 | $v_6$=110 |
| $v_1$=010 | $v_3$=001 | $v_5$=100 | $v_7$=111 |

| $v_0$ | | $v_1$ | | $v_2$ | | $v_3$ | | $v_4$ | | $v_5$ | | $v_6$ | | $v_7$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | | $\underline{\mathbf{3}}$ | | 2 | | $\underline{\mathbf{1}}$ | | 3 | | $\underline{\mathbf{2}}$ | | 1 | |

For $n = 4$ consider the following cyclic Grey code (Hamiltonian cycle) $G(4) = v_0, v_1, \ldots, v_{15}$.

| | | | |
|---|---|---|---|
| $v_0$=0000 | $v_4$=0110 | $v_8$=1111 | $v_{12}$=1001 |
| $v_1$=1000 | $v_5$=0100 | $v_9$=0111 | $v_{13}$=1011 |
| $v_2$=1010 | $v_6$=1100 | $v_{10}$=0101 | $v_{14}$=0011 |
| $v_3$=1110 | $v_7$=1101 | $v_{11}$=0001 | $v_{15}$=0010 |

| $v_0$ | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ | $v_6$ | $v_7$ | $v_8$ | $v_9$ | $v_{10}$ | $v_{11}$ | $v_{12}$ | $v_{13}$ | $v_{14}$ | $v_{15}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\underline{\mathbf{1}}$ | 3 | $\underline{\mathbf{2}}$ | 1 | $\underline{\mathbf{3}}$ | 1 | $\underline{\mathbf{4}}$ | 3 | $\underline{\mathbf{1}}$ | 3 | $\underline{\mathbf{2}}$ | 1 | $\underline{\mathbf{3}}$ | 1 | $\underline{\mathbf{4}}$ | 3 |

$G(4)$ have some interesting properties, we shall need the following later:

$$\{v_{2s}, v_{2s+1}\} \in \mathcal{E}_{s \pmod 4)+1} \qquad (s = 0, \ldots, 7). \tag{4}$$

So the odd edges give a maximum balanced matching, $\mathcal{M}_4$ for $n = 4$.

For $n = 5, 6, 7$ we consider $B_n$ as $B_2 \times B_3$, $B_3 \times B_3$ and $B_3 \times B_4$, respectively. We use only the edges of the above Grey codes $G(3)$ and $G(4)$ in the corresponding subcubes to construct maximum balanced matchings $\mathcal{M}_5, \mathcal{M}_6, \mathcal{M}_7$. One possible solution for each $n \in \{5, 6, 7\}$ can be seen on Figure 2.

## 3.2 Case of $n \geq 4$, $n$ is a power of 2

For $n \geq 4$, $n$ is a power of 2, we can construct a complete matching with equal number of edges in each parallel class. We construct recursively a cyclic Grey

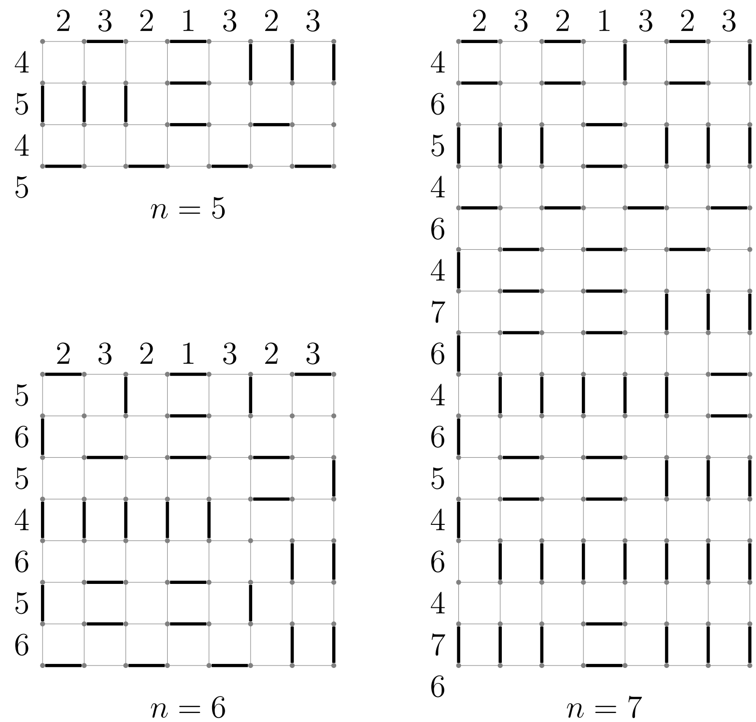


Figure 2: Maximum balanced matchings for $n = 5, 6, 7$. The parallel classes $\mathcal{E}_i (1 \leq i \leq 7)$ are denoted shortly by $1, 2, 3, 4, 5, 6, 7$.

code $G(2^t)$ of $B_{2^t}, (t \geq 2)$, such that its odd edges form the desired complete matching. Furthermore, the Grey code will have the following property:

$$\text{the } i\text{th and the } (i + 2^{2^t-1})\text{th element belong} \\ \text{to the same parallel class } (1 \leq i \leq 2^{2^t-1}). \tag{5}$$

For $n = 4$ we have already constructed a cyclic Grey code. By (4) it has property (5). Suppose, that we have already constructed a Grey code $G(2^t) = v_1, \ldots, v_{2^{2^t}}$ satisfying (5). We construct a Grey code satisfying (5) for $B_{2^{t+1}} = B_{2^t} \times B_{2^t}$. By the induction hypothesis, the following Hamiltonian cycle is

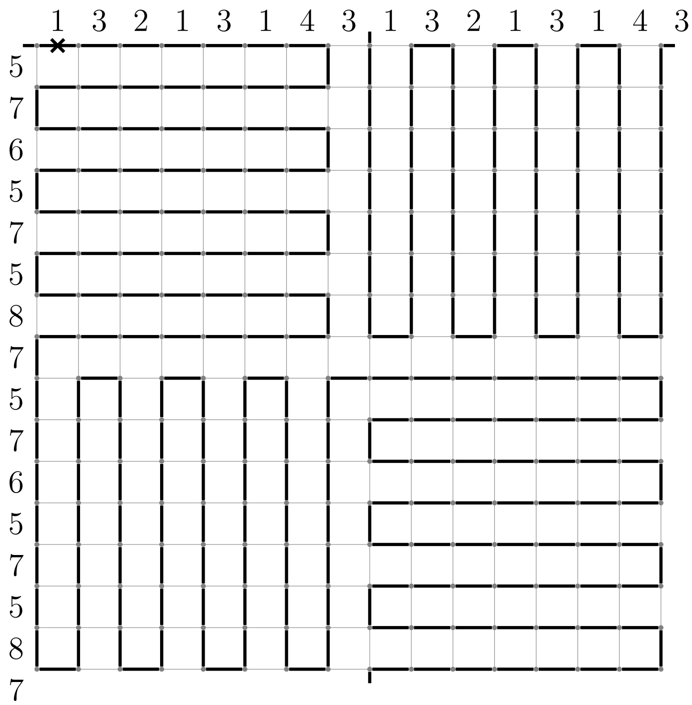


Figure 3: Construction of a cyclic Grey code for the power of 2 ($n = 8$). Taking every second edge from the marked one yields a maximum balanced matching.

appropriate (for $n = 8$, see Figure 3).

$$\begin{aligned}G(2^{t+1}) = {} & (v_1, v_1), (v_1, v_2), \dots, (v_1, v_{2^{2^t-1}}), (v_2, v_{2^{2^t-1}}), \dots, (v_2, v_1), (v_3, v_1), \\ & \dots, (v_3, v_{2^{2^t-1}}), (v_4, v_{2^{2^t-1}}), \dots, \dots, (v_{2^{2^t-1}}, v_1), (v_{2^{2^t-1}+1}, v_1), (v_{2^{2^t-1}+2}, v_1), \\ & \dots, (v_{2^{2^t}}, v_1), (v_{2^{2^t}}, v_2), \dots, (v_{2^{2^t-1}+1}, v_2), (v_{2^{2^t-1}+1}, v_3), \dots, \\ & \dots, (v_{2^{2^t-1}+1}, v_{2^{2^t-1}}), (v_{2^{2^t-1}+1}, v_{2^{2^t-1}+1}), (v_{2^{2^t-1}+1}, v_{2^{2^t-1}+2}), \dots, \\ & (v_{2^{2^t-1}+1}, v_{2^{2^t}}), (v_{2^{2^t-1}+2}, v_{2^{2^t}}), \dots, (v_{2^{2^t-1}+2}, v_{2^{2^t-1}+1}), (v_{2^{2^t-1}+3}, v_{2^{2^t-1}+1}), \\ & \dots, \dots, (v_{2^{2^t}}, v_{2^{2^t-1}+1}), (v_1, v_{2^{2^t-1}+1}), \dots, (v_{2^{2^t-1}}, v_{2^{2^t-1}+1}), (v_{2^{2^t-1}}, v_{2^{2^t-1}+2}), \\ & \dots, (v_1, v_{2^{2^t-1}+2}), (v_1, v_{2^{2^t-1}+3}), \dots, \dots, (v_1, v_{2^{2^t}}).\end{aligned}$$

### 3.3 Case of $n \geq 9$, $n$ is not a power of 2

For $n \geq 9$, $n$ is not a power of 2, we construct a maximum balanced matching using a balanced cycle cover of Lemma 2.1 for $B_{n-4}$. Note, that in this case $2^n - 2n\lfloor 2^n/2n \rfloor \geq 2$ holds, so we can afford not to cover at least 2 vertices.

$B_n = B_{n-4} \times B_4$, so we can assume that the vertices of $B_n$ are of the form $(u_i, v_j), 1 \leq i \leq 2^{n-4}, 0 \leq j \leq 15$, where $G(4) = v_0, \ldots, v_{15}$. Let $\mathcal{C}_{n-4} = \{C_0, C_1, \ldots, C_t\}$ be a balanced cycle cover of $B_{n-4}$, such that

$$\bigcup_{i=0}^{t} V(C_i) = \{u_1, \ldots, u_{2^{n-4}}\}$$

and

$$\begin{gathered} \mathcal{E}(C_0) = \{\{u_1, u_2\}, \{u_2, u_3\}, \ldots, \{u_{2n-9}, u_{2n-8}\}, \{u_{2n-8}, u_1\}\}, \\ \{u_i, u_{i+1}\}, \{u_{2n-8-i}, u_{2n-7-i}\} \in \mathcal{E}_i \quad (1 \leq i \leq n-5), \\ \{u_{n-4}, u_{n-3}\}, \{u_{2n-8}, u_1\} \in \mathcal{E}_{n-4}. \end{gathered} \tag{6}$$

By (4) we have

$$\begin{aligned} \{(u_i, v_{2j}), (u_i, v_{2j+1})\}, \{(u_i, v_{2j+8}), (u_i, v_{2j+9})\} \in \mathcal{E}_{n-3+j}, & \\ j = 0, 1, 2, 3, \quad 1 \leq i \leq 2^{n-4}. & \end{aligned}$$

Let $\mathcal{M}$ be the following matching. If $E = \{u_i, u_j\}$ is an odd edge of $C_i (i \geq 1)$, then let

$$\{(u_i, v_0), (u_j, v_0)\}, \ldots, \{(u_i, v_7), (u_j, v_7)\} \in \mathcal{M}, \tag{7}$$

otherwise let

$$\{(u_i, v_8), (u_j, v_8)\}, \ldots, \{(u_i, v_{15}), (u_j, v_{15})\} \in \mathcal{M}. \tag{8}$$

If $E$ is an odd edge of $C_0$, then let

$$\{(u_i, v_1), (u_j, v_1)\}, \{(u_i, v_3), (u_j, v_3)\}, \ldots, \{(u_i, v_{15}), (u_j, v_{15})\} \in \mathcal{M},$$

otherwise let

$$\{(u_i, v_0), (u_j, v_0)\}, \{(u_i, v_2), (u_j, v_2)\}, \ldots, \{(u_i, v_{14}), (u_j, v_{14})\} \in \mathcal{M}.$$

These edges are called $C_0$-*edges*.

If $\mathcal{C}_{n-4}$ has a profile $(\lambda'_1, \ldots, \lambda'_{n-4})$, $\lambda'_1 = \cdots = \lambda'_s$, $\lambda'_{s+1} = \cdots = \lambda'_{n-4}$, $\lambda'_{s+1} = \lambda'_s + 2$, for some $1 \leq s < n-4$, then we have

$$\chi(\mathcal{M}) = (8\lambda'_1, \ldots, 8\lambda'_{n-4}, 0, 0, 0, 0).$$

Take 2 edges $E = \{(u_i, v_j), (u_{i'}, v_j)\}$ and $E' = \{(u_i, v_{j+1}), (u_{i'}, v_{j+1})\}$, such that $j$ is even and $\{u_i, u_{i'}\} \in \mathcal{E}(\mathcal{C}_{n-4} \setminus \{C_0\})$. Remove $E$ and $E'$ from $\mathcal{M}$ and add $\{(u_i, v_j), (u_i, v_{j+1})\}$ and $\{(u_{i'}, v_j), (u_{i'}, v_{j+1})\}$. So if $E, E' \in \mathcal{M} \cap \mathcal{E}_k$, then we are decreased $\lambda_k(\mathcal{M})$ by 2, while increased one of the last 4 components of $\chi(\mathcal{M})$ by 2 (by (4)).

Repeating the above transformation in an approriate order, we can reach, that all components of $\chi(\mathcal{M})$ differ by either 0 or 2 if there are enough edges initially in $\mathcal{E}(\mathcal{C}_{n-4} \setminus \{C_0\}) \cap \mathcal{E}_k$ $(1 \leq k \leq n-4)$.

In the initial matching there are at least $8[\![2^{n-4}/(n-4)]\!] - 16$ edges in $\mathcal{E}(\mathcal{C}_{n-4} \setminus \{C_0\})$ in each parallel class, while at most $[\![2^n/2n]\!]$ needed. Substituting $n = 9$ the first quantity is larger than the second one. For $n \geq 10$ we have

$$8\left[\!\!\left[\frac{2^{n-4}}{n-4}\right]\!\!\right] - 16 \geq 8\left(\frac{2^{n-4}}{n-4} - 2\right) - 16 \geq \frac{2^n}{2n} \geq \left[\!\!\left[\frac{2^n}{2n}\right]\!\!\right]$$

The middle inequality is equivalent to the inequality $2^{n-4} \geq n(n-4)$, which holds for $n \geq 10$.

So we have a matching $\mathcal{M}$, such that

$$\chi(\mathcal{M}) = (\lambda_1, \ldots, \lambda_n),$$

where $\lambda_{i_1} = \cdots = \lambda_{i_s}, \lambda_{i_{s+1}} = \cdots = \lambda_{i_n}, \lambda_{i_{s+1}} = \lambda_{i_s} + 2$ with $2(n-s) = 2^n - 2n\lfloor 2^n/2n \rfloor$ and all $\lambda_{i_j}$'s are even.

Note, that we can set $\{i_1, \ldots, i_s\}$ to be any specific $s$-subset of $[n]$ and $\mathcal{M}$ still contains all $C_0$-edges. $\lambda_{i_1}$ equals either $\lfloor 2^n/2n \rfloor - 1$ or $\lfloor 2^n/2n \rfloor$. If $\lambda_{i_1} = \lfloor 2^n/2n \rfloor - 1$ then the $C_0$-edges will be used for correction. We distinguish 5 cases (Figure 4).

*Case 1.* If $s \geq n/2$, then we are either ready, since $\lambda_{i_1} = \lfloor 2^n/2n \rfloor$ (if $s > n/2$) or $n$ is a power of 2 (if $s = n/2$), since $2^n/2n = \lfloor 2^n/2n \rfloor$ can not hold otherwise. The case of $n$ is a power of 2 is already discussed.

*Case 2.* Let $s < (n-4)/2$. Assume, that $\lfloor 2^n/2n \rfloor - 1 = \lambda_1 = \lambda_3 = \cdots =$

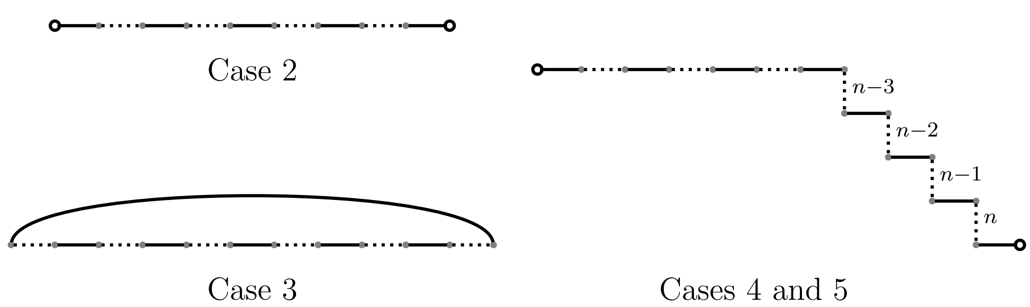


Figure 4: Balanceness correction using the $C_0$-edges. (The original edges are replaced by the dotted ones.)

$\lambda_{2s-1}$. Let us introduce the notation

$$\mathcal{D}_{k,s} = \{\{(u_k, v_0), (u_{k (\mathrm{mod}\ (2n-8))+1}, v_0)\}, \{(u_{k+1 (\mathrm{mod}\ (2n-8))+1}, v_0), (u_{k+2 (\mathrm{mod}\ (2n-8))+1}, v_0)\}, \dots, \{(u_{k+2s-3 (\mathrm{mod}\ (2n-8))+1}, v_0), (u_{k+2s-2 (\mathrm{mod}\ (2n-8))+1}, v_0)\}\}.$$

By (6), (7) and (8) $\mathcal{M} \setminus \mathcal{D}_{2n-8,s+1} \cup \mathcal{D}_{1,s}$ is a maximum balanced matching, since the $\lambda_{2i}$'s are decreased for $i = n-4$ and $i \in [s]$, while the $\lambda_{2i-1}$'s are increased by 1, for $i \in [s]$.

*Case 3.* If $s = (n-4)/2$ and $\lfloor 2^n/2n \rfloor - 1 = \lambda_1 = \lambda_3 = \cdots = \lambda_{n-5}$ then $\mathcal{M} \setminus \mathcal{D}_{2,(n-4)/2} \cup \mathcal{D}_{1,(n-4)/2}$ is a maximum balanced matching.

*Case 4.* $s = \lfloor (n-4)/2 \rfloor + 1$. We can assume, that $\lfloor 2^n/2n \rfloor - 1 = \lambda_3 = \cdots = \lambda_{2\lfloor (n-4)/2 \rfloor - 5} = \lambda_{n-3} = \lambda_{n-2} = \lambda_{n-1} = \lambda_n$, while all other components of $\chi(\mathcal{M})$ equal to $\lfloor 2^n/2n \rfloor + 1$. Let

$$\mathcal{D}_4^- = \{\{(u_{2\lfloor (n-4)/2 \rfloor - 3}, v_1), (u_{2\lfloor (n-4)/2 \rfloor - 2}, v_1)\}, \{(u_{2\lfloor (n-4)/2 \rfloor - 2}, v_2), (u_{2\lfloor (n-4)/2 \rfloor - 1}, v_2)\}, \{(u_{2\lfloor (n-4)/2 \rfloor - 1}, v_3), (u_{2\lfloor (n-4)/2 \rfloor}, v_3)\}, \{(u_{2\lfloor (n-4)/2 \rfloor}, v_8), (u_{2\lfloor (n-4)/2 \rfloor + 1}, v_8)\}\}$$

and

$$\mathcal{D}_4^+ = \{\{(u_{2\lfloor (n-4)/2 \rfloor - 3}, v_0), (u_{2\lfloor (n-4)/2 \rfloor - 3}, v_1)\}, \{(u_{2\lfloor (n-4)/2 \rfloor - 2}, v_1), (u_{2\lfloor (n-4)/2 \rfloor - 2}, v_2)\}, \{(u_{2\lfloor (n-4)/2 \rfloor - 1}, v_2), (u_{2\lfloor (n-4)/2 \rfloor - 1}, v_3)\}, \{(u_{2\lfloor (n-4)/2 \rfloor}, v_3), (u_{2\lfloor (n-4)/2 \rfloor}, v_8)\}\}.$$

Note, that in $G(4)$ $\{v_0, v_1\}, \{v_1, v_2\}, \{v_2, v_3\}, \{v_3, v_8\}$ belong to 4 different classes of edges. $\mathcal{M} \setminus (\mathcal{D}_{2,s-3} \cup \mathcal{D}_4^-) \cup \mathcal{D}_{3,s-4} \cup \mathcal{D}_4^+$ is a maximum balanced matching, since the $\lambda_{2i}$'s for $1 \leq i \leq \lfloor (n-4)/2 \rfloor - 2$ and the $\lambda_i$'s for $2\lfloor (n-4)/2 \rfloor - 3 \leq i \leq 2\lfloor (n-4)/2 \rfloor$ are decreased, while the $\lambda_{2i-1}$'s for $2 \leq i \leq \lfloor (n-4)/2 \rfloor - 2$ and the $\lambda_i$'s for $n-3 \leq i \leq n$ are increased by 1.

*Case 5.* $s = \lfloor (n-4)/2 \rfloor + 2$ and $n$ is odd. (Note, that the case of even $n$ was already considered in Case 1.) We can assume, that $\lfloor 2^n/2n \rfloor - 1 = \lambda_1 = \lambda_3 = \cdots = \lambda_{2\lfloor (n-4)/2 \rfloor - 5} = \lambda_{n-3} = \lambda_{n-2} = \lambda_{n-1} = \lambda_n$, while all other components of $\chi(\mathcal{M})$ equal to $\lfloor 2^n/2n \rfloor + 1$.

$\mathcal{M} \setminus (\mathcal{D}_{2n-8,s-3} \cup \mathcal{D}_4^-) \cup \mathcal{D}_{1,s-4} \cup \mathcal{D}_4^+$ is a maximum balanced matching, since $\lambda_{n-4}$, the $\lambda_{2i}$'s for $1 \leq i \leq \lfloor (n-4)/2 \rfloor - 2$ and the $\lambda_i$'s for $2\lfloor (n-4)/2 \rfloor - 3 \leq i \leq 2\lfloor (n-4)/2 \rfloor$ are decreased, while the $\lambda_{2i-1}$'s for $1 \leq i \leq \lfloor (n-4)/2 \rfloor - 2$ and the $\lambda_i$'s for $n-3 \leq i \leq n$ are increased by 1. (Note, that we have $n - 4 \neq 2\lfloor (n-4)/2 \rfloor$ in this case.)

We could achieve in all the 5 cases, that each of the parallel classes contain at least $\lfloor 2^n/2n \rfloor$ elements. □

# 4 Balanceness of hypergraphs

Let us consider the following generalization of our problem. Let $\mathcal{H} = \langle V, \mathcal{E} \rangle$ be a hypergraph (i.e., $\mathcal{E} \subseteq 2^V$) and $\kappa : \mathcal{E} \to [n]$ be a (total) coloring of the edges. For $i \in [n]$ let

$$\mathcal{E}_i = \{E \in \mathcal{E} \mid \kappa(E) = i\}$$

be the set of those edges that have color $i$, we call $\mathcal{E}_i$ the $i$th color class.

If $\mathcal{E}' \subseteq \mathcal{E}$ and $i \in [n]$ let

$$\lambda_i = \lambda_i(\mathcal{E}') = |\mathcal{E}' \cap \mathcal{E}_i|,$$

furthermore let

$$\chi(\mathcal{E}') = (\lambda_1, \ldots, \lambda_n)$$

be the *profile* of $\mathcal{E}'$. The *balanceness of an edge set* $\mathcal{E}' \subseteq \mathcal{E}$ w.r.t. the coloring $\kappa$ is defined by

$$\mathrm{bal}(\mathcal{E}') = \mathrm{bal}_\kappa(\mathcal{E}') = \min_{i \in [n]} \lambda_i(\mathcal{E}').$$

$\mathcal{M} \subseteq \mathcal{E}$ is called a *matching*, if $E_1, E_2 \in \mathcal{M}$ implies $E_1 \cap E_2 = \emptyset$ (in other formulation $\mathcal{M}$ is a set of independent edges). The *matching balanceness* of the hypergraph $\mathcal{H}$ w.r.t. the coloring $\kappa$ is defined by

$$\mathrm{bal}(\mathcal{H}) = \mathrm{bal}_\kappa(\mathcal{H}) = \max_{\mathcal{M} \text{ is a matching in } \mathcal{H}} \mathrm{bal}(\mathcal{M}).$$

Let $\mathcal{B}_{n,k,d}$ denote the following $k^d$-uniform hypergraph ($k \geq 2, d \in [n]$). The vertices of $\mathcal{B}_{n,k,d}$ are words of length $n$ over the alphabet $\Sigma = \{0, \ldots, k-1\}$. The edges are those $k^d$-sets $E$, called *d-spaces*, that have an index set $I \subseteq [n]$, $|I| = d$, such that for each $u = t_1 \cdots t_n \in E$ and $v = t'_1 \cdots t'_n \in E$ the property $t_j = t'_j$ holds whenever $j \notin I$. For $k = 2$ and $d = 1$, $\mathcal{B}_{n,k,d}$ is nothing else, but the $n$-cube, $B_n$ (the edges are those pair of $n$-bit strings that have Hamming distance 1).

There is a natural coloring $\kappa_{\mathrm{nat}}$ of $\mathcal{B}_{n,k,d}$ with $\binom{n}{d}$ colors, those edges are colored with the same color that have the same $I$ in the definition of the edges of $\mathcal{B}_{n,k,d}$. Each color class contains $k^{n-d}$ edges. As a special case, the edges of $B_n$ are colored by $n$ colors according to the $n$ parallel classes, each color class has $2^{n-1}$ edges.

Let us introduce the short notation

$$b(n, k, d) = \mathrm{bal}_{\kappa_{\mathrm{nat}}}(\mathcal{B}_{n,k,d}).$$

Given an $r$-uniform hypergraph $\mathcal{H} = \langle V, \mathcal{E} \rangle$ and coloring $\kappa : \mathcal{E} \to [n]$ we call a matching $\mathcal{M}$ a *maximum balanced matching* if

$$\mathrm{bal}(\mathcal{M}) = \min \left\{ \min_{i \in [n]} |\mathcal{E}_i|, \left\lfloor \frac{|V|}{rn} \right\rfloor \right\} \tag{9}$$

holds. The balanceness of a matching obviously can not be larger than the RHS of (9). For the case of $B_n$, this RHS is equal to $\lfloor 2^n/2n \rfloor$. So, our main result, Theorem 1.1, can be formulated in the following way.

$$b(n, 2, 1) = \lfloor 2^n/2n \rfloor \quad (n \neq 2).$$

# 5 Balanced $d$-spaces

In this section we prove a general lower bound on $b(n.k, d)$. Note, that this lower bound is an initial result, determining the exact value remains open in most of the cases.

**Lemma 5.1.** *Let $\mathcal{S}$ be the multiset, that contain exactly $s$ copies of each element of $\binom{[n]}{d}$, where $s = d/\gcd(d, n-d+1)$. Then for the multiset $\mathcal{T}$ of $(n-d+1)s/d$ copies of $\binom{[n]}{d-1}$, there exists a bijection $\varphi : \mathcal{S} \to \mathcal{T}$, such that $S \supset \varphi(S)$ holds for all $S \in \mathcal{S}$.*

*Proof.* The bipartite graph $\langle \mathcal{S}, \mathcal{T}, \mathcal{E} \rangle$, where $\{S, T\} \in \mathcal{E} \Leftrightarrow T \subset S$ is $(n-d+1)s$-regular, therefore it has a matching. $\square$

**Corollary 5.1.** *Given $s\binom{n}{d}$ edges (d-spaces) of $\mathcal{B}_{n,k,d}$, where $s = d/\gcd(d, n-d+1)$ and exactly $s$ of the edges have the same color in $\kappa_{nat}$ for each color class. Then we can replace each d-space by $k$ $(d-1)$-spaces of the same color class of $\mathcal{B}_{n,k,d-1}$ in such a way, that there will be exactly $k(n-d+1)s/d$ edges in each of the $\binom{n}{d-1}$ color classes of $\mathcal{B}_{n,k,d-1}$ w.r.t $\kappa_{nat}$.*

*Proof.* Let $\mathcal{S}$ correspond to the color classes of $\mathcal{B}_{n,k,d}$, while $\mathcal{T}$ to the color classes of $\mathcal{B}_{n,k,d-1}$. Replace a $d$-space of color class $S \in \mathcal{S}$ by $k$ $(d-1)$-spaces of the color class $\varphi(S)$. $\square$

The following theorem gives a recursive method to count a general lower bound for $b(n, k, d)$.

**Theorem 5.1.**

$$b(n+1, k, d) \geq kb(n, k, d) - ks \left\lceil \frac{db(n, k, d)}{(n+1)s} \right\rceil, \tag{10}$$

*where $s = d/\gcd(d, n-d+1)$.*

*Proof.* Suppose, that we have a matching $\mathcal{M}_n$ having $b(n, k, d)$ $d$-spaces of each color. $V(\mathcal{B}_{n+1,k,d}) = X_0 \cup \cdots \cup X_{k-1}$, where $X_i = \{ui \mid u \in V(\mathcal{B}_{n,k,d})\}$ $(0 \leq i \leq k-1)$. Let the edge set $\mathcal{D}$ consist of $k$ isomorphic copies of $\mathcal{M}_n$ on the vertex sets $X_i$ $(0 \leq i \leq k-1)$. $\mathcal{D}$ have a profile vector

$$\chi(\mathcal{D}) = (kb(n, k, d), \ldots, kb(n, k, d), 0, \ldots, 0),$$

where we have 0 for those $d$-sets of $[n+1]$, that contain $n+1$ (let these be the last $\binom{n}{d-1}$ components).

Replace $s$ $d$-spaces of each color by $(d-1)$-spaces over $X_0$ according to Corollary 5.1. Each type of $(d-1)$-space will occur $k(n-d+1)s/d$ times. Do exactly the same for $X_1, \ldots X_{k-1}$. Replace each $k$ corresponding $(d-1)$-spaces in $X_0, \ldots, X_{k-1}$ by a single $d$-space. So the first $\binom{n}{d}$ components of $\chi(\mathcal{D})$ are decreased by $ks$, while the last $\binom{n}{d-1}$ one are increased by $k(n-d+1)s/d$.

Repeating this transformation $\ell$ times, we have the following profile for the actual edge set $\mathcal{D}$.

$$\chi(\mathcal{D})=\Big(kb(n,k,d)-\ell ks, \ldots, kb(n,k,d)-\ell ks, \ell k\frac{(n-d+1)s}{d}, \ldots, \ell k\frac{(n-d+1)s}{d}\Big).$$

Let $\ell_0$ be the least integer satisfying

$$kb(n,k,d) - \ell_0 ks \le \ell_0 k\frac{(n-d+1)s}{d},$$

i.e., $\ell_0 = \lceil db(n,k,d)/(n+1)s\rceil$. Then all components of $\chi(\mathcal{D})$ is at least the RHS of (10). $\square$

We omit the elementary, but space and paper consuming counting of the following.

**Corollary 5.2.** *Let $n_0 \ge kd/(k-1)$. Then we have*

$$\left\lfloor \frac{k^{n-d}}{\binom{n}{d}} \right\rfloor \ge b(n,k,d) \ge \frac{k^{n-d}}{\binom{n}{d}} \frac{\binom{n_0}{d}}{k^{n_0-d}} \Big(b(n_0,k,d) - d\frac{n_0+1}{n_0-d+1}\frac{(n_0+2-d)k}{(n_0+2-d)k-n_0-2}\Big).$$

We can see, that there is a big room to improve. For $d=1$ the same inductive argument gives somewhat better.

**Theorem 5.2.** *For $n \ge 4$*

$$\left\lfloor \frac{k^{n-1}}{n} \right\rfloor \ge b(n,k,1) \ge \left[\!\left[ \frac{k^{n-1}}{n} \right]\!\right]_k.$$

*Proof.* There is a maximum balanced matching for $n=4$. Suppose, that we have a matching $\mathcal{M}_n$ ($n \ge 4$) having $[\![k^{n-1}/n]\!]_k$ 1-spaces in each direction. Let $V(\mathcal{B}_{n+1,k,1}) = X_0 \cup \cdots \cup X_{k-1}$, where $X_i = \{ui \,|\, u \in V(\mathcal{B}_{n,k,1})\}$ ($0 \le i \le k-1$). Take isomorphic copies $\mathcal{M}_n^{(i)}$ of $\mathcal{M}_n$ in each $X_i$ and add the 1-spaces that consist of the corresponding vertices of $V(\mathcal{M}_n^{(i)}) - X_i$ ($0 \le i \le k-1$).

A set of $k$ 1-spaces of direction $r$, $E_i = \{t_1 \cdots t_{r-1} x t_{r+1} \cdots t_n i \,|\, 0 \le x \le k-1\}$ ($0 \le i \le k-1$) can be replaced by another $k$ 1-spaces of direction $n+1$, $E_i' = \{t_1 \cdots t_{r-1} i t_{r+1} \cdots t_n x \,|\, 0 \le x \le k-1\}$ ($0 \le i \le k-1$). Consider again the following transformation: replace $kn$ edges, $k$ from each direction and of the above type, by $kn$ edges of direction $n+1$.

Repeat the transformation while the number of edges of direction $i$ ($i \in [n]$) is bigger than $\lfloor k^n/(n+1) \rfloor$. Note, that the initial number of edges of direction $i$ ($i \in [n]$) in $\mathcal{B}_{n+1,k,1}$ is divisible by $k$. The transformations do not change this property, so the statement follows. □

# References


[1] C. D. Savage. A Survey of Combinatorial Gray Codes. SIAM Review 39 (4): 605-629, 1997.

[2] G. O. H. Katona and K. Tichler. Search when the lie depends on the target, submitted.

[3] F. Gray. Pulse Code Communication, U.S Patent No. 2632058, March 1953.

[4] G. S. Bhat and C. D. Savage. Balanced Gray codes. Electronic Journal of Combinatorics 3 (1): R25, 1996.

[5] J. Robinson and M. Cohn. Counting Sequences. IEEE Transactions on Computers, C-30: 17-23, 1981.