EGYÉNI KUTATÓMUNKA 2.

PÁRONKÉNT DISZJUNKT VEKTORPÁROK KERESÉSE $\mathbb{F}_2^n$-BEN ELŐÍRT KÜLÖNBSÉGSOROZATTAL

KOVÁCS BENEDEK
TÉMAVEZETŐ: CSIKVÁRI PÉTER

1. Kérdésfelvetés, korábban elért eredmények

A 2020-as őszi félévi kutatómunkámban Balister, Győri és Schelp ([1]) alábbi 2008-as sejtésével foglalkoztam:

1.1. Sejtés. Legyen $n \geq 2$ egész és $N = 2^n$. Ha adottak $\mathbb{F}_2^n$-ben a d_1, d_2, ..., $d_{\frac{1}{2}N}$ nemnulla különbségvektorok (nem feltétlenül különbözők) úgy, hogy $\sum_{i=1}^{\frac{1}{2}N} d_i = 0$, akkor $\mathbb{F}_2^n$ felosztható diszjunkt $\{a_i, b_i\}$ párokra ($1 \leq i \leq \frac{1}{2}N$) úgy, hogy minden i-re $a_i - b_i = d_i$ legyen.

Ezen sejtésnek az alábbi gyengített változatát vizsgáltam, melyre az igenlő választ igyekeztem minél nagyobb M értékekre belátni:

1.2. Probléma. Legyenek $n \geq 2$, $N = 2^n$ és $M \leq \frac{1}{2}N - 2$ előre megadott egészek. Igaz-e, hogy tetszőleges $d_1, d_2, \ldots, d_M \in \mathbb{F}_2^n$ nemnulla különbségértékek esetén léteznek olyan csupa különböző $a_1, \ldots, a_M, b_1, \ldots, b_M \in \mathbb{F}_2^n$ vektorok, melyekre teljesül, hogy minden $1 \leq i \leq M$-re $a_i - b_i = d_i$?

Kutatásom fő eredménye az volt, hogy az igenlő választ beláttam $M \leq \frac{5}{18}N$ esetén egy továbbfejlesztett mohó módszer használatával.

1.3. Megjegyzés. Gráfelméletileg a problémát úgy képzeljük el, hogy van M db diszjunkt élünk, az i-edik élre a d_i különbségérték van ráírva, és mi az élek végpontjaiba szeretnénk $\mathbb{F}_2^n$-beli vektorokat írni úgy, hogy minden él két végpontjába olyan vektorok kerülnek, melyeknek különbsége az élen szereplő vektor.

2. Új eredmények összefoglalása

A 2021-es tavaszi félévben az alábbi két fő eredményt értem el:

- Az előző féléves eredményt megjavítva az 1.2. probléma igazságát beláttam $M \leq \frac{9}{32}N$ megadott különbségre is.
- Vizsgáltam az 1.2. problémát abban a speciális esetben is, amikor minden d_i különbségérték csupa különböző. Ebben az esetben beláttam, hogy létezik $c > 0$ konstans és $n_0 \in \mathbb{N}$, hogy a probléma ezen speciális esetben megoldható $M \leq \frac{1}{2}N - cN^{5/6}$ esetén, ha $n \geq n_0$.

3. A probléma megoldása $M \leq \frac{9}{32}N$ esetén

Idézzük fel az előző félév kutatási eredményeit tartalmazó cikkből ([3]) az alábbi definíciót:

3.1. Definíció. Legyen $0 \leq k, l, d \leq 1$ úgy, hogy $k + l \leq 1$. Ekkor jelölje $T(k, l, d)$ azt az állítást, hogy tetszőleges $n \geq 2$-re, $M \in \mathbb{N}^+$-ra és d_1, d_2, ..., $d_M \in \mathbb{F}_2^n$ nemnulla vektorokra teljesül, hogy vagy legalább dM azonos van a vektorok között, vagy létezik egy olyan $\alpha : \mathbb{F}_2^n \to \mathbb{F}_2^n$ automorfizmus, melyekre az $\alpha(d_1), ..., \alpha(d_M)$ vektorok között legalább kM 0-val és legalább lM 1-gyel kezdődik.

Az előző félévben beláttam, hogy $0 < l \leq k \leq \frac{1}{3}$ esetén $T(k, l, 1 - 2k - 2l)$ teljesül. Ezen állítást bizonyító korábbi módszeremen tudtam javítani a $k = l$ esetben, a "szegény" és "gazdag" indexek szerepének teljesen szimmetrikussá tételével:

3.2. Tétel. *Ha $0 < k \leq \frac{1}{3}$, akkor $T(k, k, 1 - 2k)$ teljesül.*

A bizonyítás előtt idézzünk fel néhány definíciót az előző félév beszámolójából. Legyen $0 < k \leq \frac{1}{3}$ rögzített, és legyenek adottak a d_1, ..., $d_M \in \mathbb{F}_2^n$ nemnulla vektorok. A d_i vektor j-edik koordinátáját $d_{i,j}$-vel fogjuk jelölni ($1 \leq i \leq M$, $1 \leq j \leq n$, $d_{i,j} \in \{0, 1\}$).

3.3. Definíció. Ha $K \subseteq \{1, 2, ..., n\}$ egy nemüres halmaz, akkor legyen $A_K = \{x \in \mathbb{F}_2^n : \sum_{i \in K} x_i = 0\}$. (Könnyen látható, hogy A_K egy $n - 1$ dimenziós altere $\mathbb{F}_2^n$-nek.)

3.4. Definíció. Egy K halmazt *szegénynek* nevezünk, ha a $d_1, d_2, ..., d_M$ vektorok közül kevesebb mint kM db esik bele az A_K altérbe.

3.5. Definíció. Egy K halmazt *gazdagnak* nevezünk, ha a $d_1, d_2, ..., d_M$ vektorok közül több mint $(1 - k)M$ esik bele az A_K altérbe.

3.6. Definíció. Egy $1 \leq j \leq n$ indexet *szegénynek*, illetve *gazdagnak* nevezünk, ha a $\{j\}$ halmaz szegény, illetve gazdag.

3.7. Definíció. Egy K halmaz *0-gyarmata*: $Gyar_0(K) = \{1 \leq i \leq M : (\exists j \in K) d_{i,j} = 0\}$, és *1-gyarmata*: $Gyar_1(K) = \{1 \leq i \leq M : (\exists j \in K) d_{i,j} = 1\}$.

Továbbá szükségünk lesz az alábbi definícióra és lemmára:

3.8. Definíció. Legyen $B = \{B_1, B_2, \ldots, B_t\}$ egy multihalmaz, ahol B_1, B_2, ..., B_t tetszőleges halmazok, és $t \in \mathbb{N}$. Ekkor $\triangle B = \{x \in \bigcup_{i=1}^{t} B_i : |\{i : \ x \in B_i\}| \equiv 1 \ (mod\ 2)\}$.

3.9. Lemma. *Legyen $n \in \mathbb{N}^+$ és $c \in \mathbb{R}^+$. Ha a $B_1, B_2, \ldots, B_n$ halmazokra teljesül, hogy minden $S \subseteq \{1, 2, \ldots, n\}$-re $|\triangle \{B_i : i \in S\}| < c$, akkor $|\bigcup_{i=1}^{n} B_i| < 2c$.*

Bizonyítás. Ha egy x elem a B_1, B_2, ..., B_n halmazok közül u-ban van benne, ahol $1 \leq u \leq n$, akkor azon $H \subseteq [n]$ részhalmazok száma, melyekre $\triangle\{B_i : i \in H\}$ tartalmazza x-et, $2^{u-1}2^{n-u} = 2^{n-1}$. (Mivel H-nak az x-et tartalmazó halmazok indexei közül páratlan sokat kell tartalmaznia, a többi index közül pedig tetszőleges számút.)

Számoljuk össze azon (H, x) párokat, ahol $H \subseteq [n]$ nemüres részhalmaz, és $x \in \triangle\{B_i : i \in H\}$. Az előző bekezdés alapján ezen párok száma $2^{n-1} \left|\bigcup_{i=1}^{n} B_i\right|$. Másrészt

viszont minden nemüres $H \subseteq [n]$-re tudjuk, hogy a hozzá tartozó szimmetrikus differencia mérete kisebb c-nél, ezért a keresett párok száma összesen kisebb $c(2^n-1)$-nél. Tehát

$$2^{n-1}\left|\bigcup_{i=1}^{n} B_i\right| < c(2^n-1)$$

$$\left|\bigcup_{i=1}^{n} B_i\right| < c\left(2-\frac{1}{2^{n-1}}\right) < 2c$$

□

A 3.2. tétel bizonyítása. Tegyük fel, hogy nincs olyan automorfizmus, mely az M vektor közül legalább kM-et 0-val kezdődőbe és legalább kM-et 1-gyel kezdődőbe visz. Ekkor tehát minden nemüres $K \subseteq \{1,2,\ldots,n\}$ halmaz vagy szegény, vagy gazdag. Speciálisan minden $1 \leq j \leq n$ index vagy szegény, vagy gazdag: legyen $\{s_1, s_2, \ldots, s_a\}$ a szegény, $\{g_1, g_2, \ldots, g_b\}$ pedig a gazdag indexek halmaza.

Minden $1 \leq j \leq a$-ra legyen $S_j = \{1 \leq i \leq M : d_{is_j} = 1\}$, és minden $1 \leq j \leq b$-re legyen $G_j = \{1 \leq i \leq M : d_{ig_j} = 0\}$. Be fogjuk látni, hogy az így kapott $S_1, \ldots, S_a, G_1, \ldots, G_b$ halmazok uniója $2kM$-nél kevesebb elemet tartalmaz. Ehhez a 3.9. lemmát fogom használni, ennek használatához pedig az szükséges, hogy a halmazok közül akárhogyan is választunk ki néhányat, a kiválasztottak szimmetrikus differenciájának elemszáma kisebb kM-nél.

Ez utóbbi állítást a kiválasztott halmazok száma szerinti indukcióval látom be. Ha csak 0 vagy 1 halmazt választunk ki, akkor az állítás triviális, mert a gazdag, illetve szegény indexek definíciója alapján minden S_j és minden G_j elemszáma kisebb kM-nél. Most válasszuk ki WLOG az S_1, S_2, …, S_d, G_1, …, G_e halmazokat, ahol $0 \leq d \leq a$, $0 \leq e \leq b$ és $d+e \geq 2$. Ha a $d+e$ db halmaz közül egyet elhagyunk, akkor az indukciós feltevés miatt a megmaradó halmazok szimmetrikus differenciája kM-nél kevesebb elemet tartalmaz. Maga a kimaradó halmaz is kM-nél kevesebb elemet tartalmaz. Mivel tetszőleges A, B halmazokra $|A \triangle B| \leq |A| + |B|$, ezt használva a $d+e$ db halmazunk szimmetrikus differenciájának elemszáma $2kM$-nél kisebb.

Másrészt pedig ha $K = \{s_1, \ldots, s_d, g_1, \ldots, g_e\}$, akkor $S_1 \triangle \cdots \triangle S_d \triangle G_1 \triangle \cdots \triangle G_e$ vagy megegyezik az $\{1 \leq i \leq M : d_i \in A_K\}$ halmazzal, vagy pedig annak komplementerével. (Hogy a két lehetőség közül melyik áll fenn, az d paritásától függ.) Mivel a $K = \{s_1, \ldots, s_d, g_1, \ldots, g_e\}$ halmaz vagy szegény, vagy gazdag, ezért $|S_1 \triangle \cdots \triangle S_d \triangle G_1 \triangle \cdots \triangle G_e|$ vagy kM-nél kisebb, vagy $(1-k)M$-nél nagyobb. Azt már láttuk, hogy $2kM$-nél kisebb, tehát nem lehet $(1-k)M$-nél nagyobb (mert $k \leq \frac{1}{3}$). Tehát kM-nél kisebb, azaz a kívánt állítást beláttuk.

Tehát a 3.9. lemma miatt $|S_1 \cup \cdots \cup S_a \cup G_1 \cup \cdots \cup G_b| < 2kM$. Tehát az unión kívül eső i indexekre $d_{is_j} = 1$ és $d_{ig_j} = 0$ minden j-re, vagyis ezekre az i-kre a d_i vektorok mind megegyeznek. Vagyis a vektorok közül több mint $(1-2k)M$ mind azonos. □

Ezen javítás segítségével a cikk fő eredményét is meg tudtam javítani:

3.10. Tétel. *Az 1.2. probléma állítása igaz $M \leq \frac{9}{32}N$ esetén.*

Bizonyítás. Használjuk az előző beszámolóban ([3]) szereplő 5.1. tételt $k = l = d = \frac{1}{3}$ és $\alpha = \frac{9}{32}$ paraméterekkel. A (6) feltétel ellenőrzéséhez az előző beszámolóbeli 3.7. következményt használjuk $\lambda = \frac{1}{4}$-del. □

4. Csupa különböző különbségek

Abban a speciális esetben, amikor minden megadott d_i különbség csupa különböző, beláttam, hogy az 1.2. probléma megoldható $\frac{1}{2}N - o(N)$ db különbségre.

Ehhez előkészítésként a következő állítást látjuk be. Az állítás a Charbit et al. cikkében ([2]) található 1. tételnek egyfajta általánosabb változata, és a bizonyítás is az ott szereplőhöz hasonló módszert használ.

4.1. Állítás. *Legyenek $k \geq 1$ és $M > 2^{2k} - 2^k$ egészek, és $u_1, \ldots, u_M \in \mathbb{F}_2^n$ páronként különböző nemnulla vektorok. Ekkor léteznek olyan $a_1, \ldots, a_k \in \mathbb{F}_2^n$ lineárisan független vektorok, melyekre minden $c_1, ..., c_k$ esetén $\left|d_{c_1,\ldots,c_k}(a_1, \ldots, a_k) - \frac{1}{2^k}M\right| \leq \sqrt{M}\sqrt{1 - \frac{1}{2^k}}$, ahol $d_{c_1,\ldots,c_k}(a_1, \ldots, a_k) := |\{i : 1 \leq i \leq M, \langle u_i, a_j\rangle = c_j \ \forall 1 \leq j \leq k\}|$.*

Bizonyítás. Vezessük be az alábbi jelölést: $u, a_1, ..., a_k \in \mathbb{F}_2^n$ vektorok és $c_1, ..., c_k \in \mathbb{F}_2$ értékek esetén legyen

$$f_{c_1,\ldots,c_k}(u, a_1, ..., a_k) = \begin{cases} 1 & \text{ha } \langle a_i, u\rangle = c_i \text{ minden } i\text{-re,} \\ 0 & \text{különben.} \end{cases}$$

Ekkor a fenti állításban szereplő $d_{c_1,\ldots,c_k}(a_1, ..., a_k)$ jelölésre teljesül ez:

$$d_{c_1,\ldots,c_k}(a_1, ..., a_k) = \sum_{i=1}^{M} f_{c_1,\ldots,c_k}(u_i, a_1, ..., a_k)$$

Most rögzítsük $c_1, ..., c_k \in \mathbb{F}_2$ értékeit, és tekintsük az alábbi összegeket:

$$S := \sum_{a_1,\ldots,a_k \in \mathbb{F}_2^n} d_{c_1,\ldots,c_k}(a_1, ..., a_k)$$

$$S_2 := \sum_{a_1,\ldots,a_k \in \mathbb{F}_2^n} d_{c_1,\ldots,c_k}(a_1, ..., a_k)^2$$

Először is

$$S = \sum_{a_1,\ldots,a_k \in \mathbb{F}_2^n} \sum_{i=1}^{M} f_{c_1,\ldots,c_k}(u_i, a_1, ..., a_k)$$

A szummázás sorrendjét megcseréljük:

$$S = \sum_{i=1}^{M} \sum_{a_1,\ldots,a_k \in \mathbb{F}_2^n} f_{c_1,\ldots,c_k}(u_i, a_1, ..., a_k)$$

Kihasználva, hogy minden u_i nemnulla, tehát az $\mathbb{F}_2^n$-ben minden u_i-re 2^{n-1} db merőleges és 2^{n-1} db nem merőleges vektor létezik:

$$S = \sum_{i=1}^{M} (2^{n-1})^k = \frac{1}{2^k} M N^k$$

Most számítsuk ki S_2 értékét:

$$S_2 = \sum_{a_1,\ldots,a_k \in \mathbb{F}_2^n} \sum_{i,j=1}^{M} f_{c_1,\ldots,c_k}(u_i, a_1, ..., a_k) f_{c_1,\ldots,c_k}(u_j, a_1, ..., a_k)$$

$$S_2 = \sum_{i,j=1}^{M} \sum_{a_1,...,a_k \in \mathbb{F}_2^n} f_{c_1,...,c_k}(u_i, a_1, ..., a_k) f_{c_1,...,c_k}(u_j, a_1, ..., a_k)$$

Ha $i = j$, akkor az ezen esethez tartozó tag értéke az előbbiek alapján $\sum_{a_1,...,a_k \in \mathbb{F}_2^n} f_{c_1,...,c_k}(u_i, a_1, ..., a_k) = (2^{n-1})^k$.

Ha pedig $i \neq j$, akkor az u_i és u_j különböző nemnulla vektorok, így tetszőleges $c, c' \in \mathbb{F}_2$ esetén $\mathbb{F}_2^n$-ben $\frac{1}{4}N$ db olyan vektor van, melynek u_i-vel vett skaláris szorzata c, és az u_j-vel vett skaláris szorzata pedig c'. Ezért minden a_l-re $\frac{1}{4}N$ féle választási lehetőségünk van, vagyis

$$\sum_{a_1,...,a_k \in \mathbb{F}_2^n} f_{c_1,...,c_k}(u_i, a_1, ..., a_k) f_{c_1,...,c_k}(u_j, a_1, ..., a_k) = \left(\frac{1}{4}N\right)^k$$

és így

$$S_2 = \sum_{i=1}^{M} \left(\frac{1}{2}N\right)^k + \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} \left(\frac{1}{4}N\right)^k = M\left(\frac{1}{2}N\right)^k + M(M-1)\left(\frac{1}{4}N\right)^k$$

$$S_2 = N^k \left(\left(\frac{1}{2^k} - \frac{1}{4^k} \right) M + \frac{1}{4^k} M^2 \right)$$

Ezek alapján

$$\sum_{a_1,...,a_k \in \mathbb{F}_2^n} \left(d_{c_1,...,c_k}(a_1, ..., a_k) - \frac{1}{2^k}M \right)^2 = S_2 - \frac{2}{2^k} MS + N^k \frac{1}{4^k} M^2 = N^k M \left(\frac{1}{2^k} - \frac{1}{4^k} \right)$$

Most ezt összegezzük minden $c_1, ..., c_k \in \mathbb{F}_2$-re, értelemszerűen ekkor az előző érték 2^k-szorosát kapjuk:

$$\sum_{a_1,...,a_k \in \mathbb{F}_2^n} \sum_{c_1,...,c_k \in \mathbb{F}_2} \left(d_{c_1,...,c_k}(a_1, ..., a_k) - \frac{1}{2^k}M \right)^2 = N^k M \left(1 - \frac{1}{2^k} \right)$$

Ebből az következik, hogy léteznek olyan $a_1, ..., a_k \in \mathbb{F}_2^n$ vektorok, hogy

$$\sum_{c_1,...,c_k \in \mathbb{F}_2} \left(d_{c_1,...,c_k}(a_1, ..., a_k) - \frac{1}{2^k}M \right)^2 \leq M \left(1 - \frac{1}{2^k} \right)$$

Emiatt ezen $a_1, ..., a_k \in \mathbb{F}_2^n$-re minden $c_1, ..., c_k \in \mathbb{F}_2$ esetén

$$\left| d_{c_1,...,c_k}(a_1, ..., a_k) - \frac{1}{2^k}M \right| \leq \sqrt{M}\sqrt{1 - \frac{1}{2^k}} \qquad (*)$$

Annak a belátása maradt, hogy $a_1, ..., a_k \in \mathbb{F}_2^n$ lineárisan függetlenek. Ha lenne közöttük lineáris összefüggés, akkor megadhatóak lennének olyan $c_1, ..., c_k$ értékek, hogy semmilyen $u \in \mathbb{F}_2^n$-re ne teljesüljön $\langle u, a_i \rangle = c_i$ $(\forall 1 \leq i \leq k)$. Ekkor $d_{c_1,...,c_k}(a_1, ..., a_k) = 0$, tehát $(*)$ alapján $\frac{1}{2^k}M \leq \sqrt{M}\sqrt{1 - \frac{1}{2^k}}$, melyből $M \leq 2^{2k} - 2^k$ adódik, ez pedig ellentmond az állítás feltételének. □

4.2. Tétel. *Létezik olyan $a > 0$ konstans és $n_0 \in \mathbb{N}$, melyre $M \leq \frac{1}{2}N - aN^{5/6}$ és $n \geq n_0$ esetén az 1.2. probléma állítása igaz, ha minden d_i páronként különböző.*

Bizonyítás. Adva vannak a $d_1, d_2, ..., d_M \in \mathbb{F}_2^n$ csupa különböző nemnulla különbségvektoraink. Nyilvánvalóan feltehető $M > \frac{9}{32}N$, hiszen különben a 3.10. tétel miatt készen vagyunk. Rögzítsünk egy $2 \leq k \leq \frac{1}{2}n - 1$ egész számot. Ekkor a 4.1. állítás alkalmazható az M db vektorra, mert $2^{2k} - 2^k < 2^{2k} \leq \frac{1}{4}N < M$, így létezik olyan $\phi : \mathbb{F}_2^n \to \mathbb{F}_2^n$ vektortér-automorfizmus, amelyre teljesül, hogy minden $s \in \mathbb{F}_2^k$ esetén $\phi(d_1), ..., \phi(d_M)$ közül legfeljebb $\frac{1}{2^k}M + \sqrt{1 - \frac{1}{2^k}}\sqrt{M}$ db vektornak egyezik meg s-sel az első k koordinátája.

Elegendő az automorfizmussal áttranszformált változatot megoldanunk, így mostantól feltesszük, hogy minden $s \in \mathbb{F}_2^k$-ra a d_i vektorok közül legfeljebb $\frac{1}{2^k}M + \sqrt{1 - \frac{1}{2^k}}\sqrt{M}$ db vektor kezdődik s-sel.

Most a korábbi beszámolóban ismertetett "háromrészes mohó módszert" fogjuk általánosítani. Az $\mathbb{F}_2^n$ vektorteret most osszuk a vektorok első k koordinátája alapján 2^k kupacba (így minden kupacba 2^{n-k} vektor kerül; minden kupac címkéje legyen a benne lévő vektorok első k koordinátája). Ha $r \in \mathbb{F}_2^n$, és r első k koordinátája $s \in \mathbb{F}_2^k$, és $s \neq 0$, és $a, b \in \mathbb{F}_2^k$-ra teljesül $a - b = s$, akkor az a és b indexű kupacok között van egy teljes párosítás, melyben minden él két végpontja közti különbség r.

Ha pedig $r \in \mathbb{F}_2^n$, és r első k koordinátája 0, akkor tetszőleges $a \in \mathbb{F}_2^k$-ra az a indexű kupac felosztható 2^{n-k-1} db párba úgy, hogy minden párban a két elem különbsége r legyen.

Ha $a, b \in \mathbb{F}_2^k$ és $c \in \mathbb{N}$, akkor nevezzük (a, b, c) típusú lépésnek az alábbi folyamatot: a megadott d_i különbségeink közül c darab olyanhoz, amelyhez még nem rendeltünk a_i és b_i vektorokat, és amelynek kezdete $a - b \in \mathbb{F}_2^k$, sorban egyesével hozzárendelünk ilyen a_i-t és b_i-t tetszőleges módon úgy, hogy a_i kezdete a, b_i kezdete pedig b legyen. Ez a mohó folyamat minden esetben elvégezhető, ha teljesül az alábbi feltétel:

- $a \neq b$ esetén $f_a + f_b + 2c \leq 2^{n-k}$,
- $a = b$ esetén pedig $f_a + 2c \leq 2^{n-k-1}$,

ahol (a, b, c) típusú lépés megkezdése előtt felhasznált a kezdetű vektorok számát f_a, a b kezdetűekét f_b jelöli. (Ez a fenti párosítások használatával látható mind $a = b$, mind pedig $a \neq b$ esetén.)

Az $\mathbb{F}_2^k$ vektorteret azonosítsuk a 2^k elemű testtel, és ilyen módon definiáljuk az elemein a szorzást. Legyen g generátora a 2^k elemű test multipikatív csoportjának. Ekkor $\mathbb{F}_2^k = \{1, g, g^2, ..., g^{2^k-2}, 0\}$. Legyen $1 \leq d \leq 2^k - 2$ egy konstans, melynek értékét (k-val együtt) szintén majd később rögzítjük le.

Tegyük meg az alábbi lépéseket, felsorolásuk sorrendjében, ahol c szintén később kiválasztandó konstans:

- $(g^0, g^1, c), (g^0, g^2, c), ..., (g^0, g^d, c)$,
- $(g^1, g^2, c), (g^1, g^3, c), ..., (g^1, g^{d+1}, c)$,
- $(g^2, g^3, c), (g^2, g^4, c), ..., (g^2, g^{d+2}, c)$,
- ...
- $(g^{2^k-d-2}, g^{2^k-d-1}, c), (g^{2^k-d-2}, g^{2^k-d}, c), ..., (g^{2^k-d-2}, g^{2^k-2}, c)$,
- $(g^{2^k-d-1}, g^{2^k-d}, c), (g^{2^k-d-1}, g^{2^k-d+1}, c), ...(g^{2^k-d-1}, g^{2^k-2}, c)$,
- ...
- $(g^{2^k-3}, g^{2^k-2}, c)$.

(Ebben a felsorolásban ha valamelyik lépésnél az adott kezdetű különbségekből már csak c-nél kevesebb maradt, akkor csak a megmaradó különbségekhez keressünk vektorpárt.)

Minden $s \in \mathbb{F}_2^k$-ra az s-sel kezdődő különbségeket dc-szer szeretnénk az eljárás során összesen felhasználni, és ha $c \ll 2^k$, akkor ez már teljesül is, kivéve néhány kimaradó különbséget: mivel minden $1 \leq i \leq d$-re $g^i - g^0$, $g^{i+1} - g^1$, $g^{i+2} - g^2$, ..., $g^{2^k-2} - g^{2^k-i-2}$ csupa különböző elemei $\mathbb{F}_2^k$-nak, ezért egy adott i-re $2^k - i - 1$ különféle különbséget használunk fel (mindegyiket c-szer), így lesz i-szer c db kimaradó különbség. Összesen ha minden $1 \leq i \leq d$-re tekintjük a dolgot, akkor $\frac{d(d+1)}{2}$ db c elemből álló kimaradó azonos kezdetű különbségcsoport van (melyeket még pluszban fel kell használnunk ahhoz, hogy minden nemnulla különbségkezdetet pontosan dc-szer használjunk).

Továbbá 0-val kezdődő különbséget sosem használunk az eljárás során, így abból dc különbség kimaradt.

A most kimaradónak tekintett különbségeket használjuk fel a fenti lépések előtt az alábbi módon:

- A 0-val kezdődő különbségeket mohó módon használjuk el úgy, hogy a lehető legegyenlőbben osszuk szét őket a 2^k db kupac között. Ekkor minden kupacban legfeljebb $\lceil \frac{dc}{2^k} \rceil \leq \frac{dc}{2^k} + 1$ db különbséget használunk, melyek kupaconként legfeljebb $\frac{dc}{2^{k-1}} + 2$ vektort fognak érinteni.
- A többi $\frac{d(d+1)}{2}$ db c elemből álló különbségcsoportot (minden csoport elemei ugyanazzal a vektorral kezdődnek) csoportonként használjuk fel mohó módon úgy, hogy minden csoport elemeit szétosztjuk a 2^{k-1} db lehetséges vektorkezdet-pár között a lehető legegyenlőbben. Minden csoportnál így legfeljebb $\frac{c}{2^{k-1}} + 1$ db vektort használunk fel minden sorban, tehát az összes csoportot tekintve legfeljebb $\frac{d(d+1)}{2}\left(\frac{c}{2^{k-1}} + 1\right)$ darabot.

A fenti plusz párosítások tehát összesen legfeljebb $H := \frac{dc}{2^{k-1}} + 2 + \frac{d(d+1)}{2}\left(\frac{c}{2^{k-1}} + 1\right)$ vektort használnak el minden kupacból. Nézzük most meg, hogy mi szükséges ahhoz, hogy a most leírt előkészítő lépések, és a fenti általános lépések is működjenek.

Az előkészítő mohó lépések működéséhez elégséges, ha minden vektorkupacnak legfeljebb a felét használjuk el ezek alatt, tehát ha $H \leq 2^{n-k-1}$.

Egy általános lépésnél pedig (mint láttuk), ha az éppen a és b kezdetű kupacok között történik, elégséges a lépés működéséhez, ha $f_a + f_b + 2c \leq 2^{n-k}$.

Könnyen látható, hogy az eljárásunk során minden lépésben teljesül, hogy $f_a + f_b + 2c \leq 2H + (2d+1)c$.

Ezek alapján tehát elégséges feltétel az egész eljárás működésére, hogy $2H + (2d+1)c \leq 2^{n-k}$ teljesüljön.

H behelyettesítése után ez az alábbi feltételt adja c-re:

$$c \leq \frac{\frac{1}{2^k}N - (d(d+1)+4)}{(2d+1) + \frac{1}{2^{k-2}}\left(\frac{d(d+1)}{2} + d\right)}$$

Természetesen c-nek egész számnak kell lennie, így a lehető legnagyobb érték, amit c-nek választhatunk (válasszuk is ezt c-nek):

$$c = \left\lfloor \frac{\frac{1}{2^k}N - (d(d+1)+4)}{(2d+1) + \frac{1}{2^{k-2}}\left(\frac{d(d+1)}{2} + d\right)} \right\rfloor$$

A módszerünk tehát képes minden olyan felállás megoldására, ahol mind a 2^k db kezdet legfeljebb dc-szer fordul elő. Nekünk arra kell felkészülnünk, hogy minden kezdetből akár $\frac{1}{2^k}M + \sqrt{1-\frac{1}{2^k}}\sqrt{M}$ db is lehet, tehát a módszer akkor működik, ha

$$\frac{1}{2^k}M + \sqrt{1-\frac{1}{2^k}}\sqrt{M} \leq dc \qquad (*)$$

Legyen $d = \left[2^{n/6}\right]$ és $k = \left[\frac{1}{3}n\right]$.

Ekkor belátjuk, hogy $(*)$ valóban teljesülni fog elég nagy n esetén, ha $M \leq \frac{1}{2}N - aN^{5/6}$ valamilyen $a > 0$ konstansra.

Tudjuk, hogy

$$c \geq \frac{\frac{1}{2^k}N - (d(d+1)+4)}{(2d+1) + \frac{1}{2^{k-2}}\left(\frac{d(d+1)}{2} + d\right)} - 1$$

tehát

$$dc \geq \frac{2^{n-k}d - (d^3+d^2+4d)}{2d+1+\frac{1}{2^{k-2}}\left(\frac{1}{2}d^2+\frac{3}{2}d\right)} - d$$

Elég tehát belátni, hogy

$$\frac{2^{n-k}d - (d^3+d^2+4d)}{2d+1+\frac{1}{2^{k-2}}\left(\frac{1}{2}d^2+\frac{3}{2}d\right)} - d \geq \frac{1}{2^k}M + \sqrt{1-\frac{1}{2^k}}\sqrt{M}$$

azaz

$$\frac{2^n d - 2^k(d^3+d^2+4d)}{2d+1+\frac{1}{2^{k-2}}\left(\frac{1}{2}d^2+\frac{3}{2}d\right)} - 2^k d \geq M + \sqrt{2^{2k}-2^k}\sqrt{M} \qquad (**)$$

Tudjuk, hogy $M \leq \frac{1}{2}N - aN^{5/6}$, és ebből adódóan $\sqrt{M} \leq \sqrt{\frac{1}{2}N - aN^{5/6}} \leq \frac{1}{\sqrt{2}}N^{1/2} - \frac{1}{\sqrt{2}}aN^{1/3}$.

Továbbá $\sqrt{2^{2k}-2^k} \leq 2^k \leq 2^{\frac{1}{3}n}$, tehát $(**)$ jobb oldalának értéke legfeljebb $\frac{1}{2}N - aN^{\frac{5}{6}} + 2^k\sqrt{M} \leq \frac{1}{2}\cdot 2^n + \left(\frac{1}{\sqrt{2}} - a\right)\cdot 2^{\frac{5}{6}n} - \frac{1}{\sqrt{2}}\cdot a \cdot 2^{\frac{2}{3}n}$.

Másrészt pedig elég nagy n esetén $(**)$ bal oldalára teljesül az alábbi egyenlőtlenségsorozat:

$$\frac{2^n d - 2^k(d^3+d^2+4d)}{2d+1+\frac{1}{2^{k-2}}\left(\frac{1}{2}d^2+\frac{3}{2}d\right)} - 2^k d =$$

$$= \frac{2^n d - 2^k(d^3+d^2+4d) - 2^{k+1}d^2 - 2^k d - 4\left(\frac{1}{2}d^3+\frac{3}{2}d^2\right)}{2d+1+\frac{1}{2^{k-2}}\left(\frac{1}{2}d^2+\frac{3}{2}d\right)} \geq$$

$$\geq \frac{2^{\frac{7}{6}n} - 2^n - 2^k(d^3+3d^2+5d) - 2d^3 - 6d^2}{2d+1+\frac{1}{2^{\frac{1}{3}n-3}}d^2} \geq$$

$$\geq \frac{2^{\frac{7}{6}n} - 2^n - 2^k\cdot\frac{3}{2}d^3}{2d+1+\frac{1}{2^{\frac{1}{3}n-3}}\cdot 2^{\frac{1}{3}n}} \geq \frac{2^{\frac{7}{6}n} - 2^n - \frac{3}{2}\cdot 2^{\frac{5}{6}n}}{2\cdot 2^{\frac{1}{6}n}+9}$$

Legyen $a = 100 + \frac{1}{\sqrt{2}}$. Már csak azt kell belátnunk, hogy

$$\frac{2^{\frac{7}{6}n} - 2^n - \frac{3}{2} \cdot 2^{\frac{5}{6}n}}{2 \cdot 2^{\frac{1}{6}n} + 9} \geq \frac{1}{2} \cdot 2^n + \left(\frac{1}{\sqrt{2}} - a\right) \cdot 2^{\frac{5}{6}n} - \frac{1}{\sqrt{2}} a \cdot 2^{\frac{2}{3}n}$$

Itt nagy n-re $\frac{2^{\frac{7}{6}n} - 2^n - \frac{3}{2} \cdot 2^{\frac{5}{6}n}}{2 \cdot 2^{\frac{1}{6}n} + 9} \geq \frac{2^{\frac{7}{6}n} - \frac{3}{2} \cdot 2^n}{2 \cdot 2^{\frac{1}{6}n} + 9}$, így elég belátni, hogy

$$\frac{2^{\frac{7}{6}n} - \frac{3}{2} \cdot 2^n}{2 \cdot 2^{\frac{1}{6}n} + 9} \geq \frac{1}{2} \cdot 2^n - 100 \cdot 2^{\frac{5}{6}n}$$

Ez pedig igaz, mert

$$2^{\frac{7}{6}n} - \frac{3}{2} \cdot 2^n \geq 2^{\frac{7}{6}n} - 200 \cdot 2^n + \frac{9}{2} \cdot 2^n - 900 \cdot 2^{\frac{5}{6}n}$$

nyilván teljesül, ha n elég nagy. □

5. Irodalomjegyzék

[1] Balister, P.N., Győri, E. és Schelp, R.H., 2011. Coloring vertices and edges of a graph by nonempty subsets of a set. European Journal of Combinatorics, 32(4), pp.533-537.

[2] Pierre Charbit, Emmanuel Jeandel, Pascal Koiran, Sylvain Perifel, Stéphan Thomassé. Finding a Vector Orthogonal to Roughly Half a Collection of Vectors. 2007. (ensl-00153736)

[3] Kovács Benedek (témavezető: Csikvári Péter), Diszjunkt vektorpárok keresése $\mathbb{F}_2^n$-ben (2020-ban beadott TDK-dolgozat, elérhető ezen a linken.)