BOUNDS ON UNIQUE-NEIGHBOR CODES

Nathan Linial*[1] and Edan Orzech†[2]

[1]*School of Computer Science and Engineering, Hebrew University, Jerusalem 91904, Israel*
nati@cs.huji.ac.il
[2]*Department of Electrical Engineering and Computer Science, MIT CSAIL, Cambridge MA 02139, U.S.A.*
iorzech@csail.mit.edu

Submitted: Mar 17, 2023; Accepted: Mar 25, 2025; Published: Sep 15, 2025

Abstract. Recall that a *binary linear code of length* n is a linear subspace $\mathcal{C} = \{x \in \mathbb{F}_2^n \mid Ax = 0\}$. Here the *parity check matrix* A is a binary $m \times n$ matrix of rank m. We say that $\mathcal{C}$ has *rate* $R = 1 - \frac{m}{n}$. Its *distance*, denoted δn is the smallest Hamming weight of a non-zero vector in $\mathcal{C}$. The *rate vs. distance problem* for binary linear codes is a fundamental open problem in coding theory, and a fascinating question in discrete mathematics. It concerns the function $R_L(\delta)$, the largest possible rate R for given $0 \leqslant \delta \leqslant 1$ and arbitrarily large length n. Here we investigate a variation of this fundamental question that we describe next.

Clearly, $\mathcal{C}$ has distance δn, if and only if for every $0 < n' < \delta n$, every $m \times n'$ submatrix of A has a row of odd weight. Motivated by several problems from coding theory, we say that A has the *unique-neighbor* property with parameter δn, if every such submatrix has a row *of weight* 1. Let $R_U(\delta)$ be the largest possible asymptotic rate of linear codes with a parity check matrix that has this stronger property. Clearly, $R_U(\cdot), R_L(\cdot)$ are non-increasing functions, and $R_U(\delta) \leqslant R_L(\delta)$ for all δ. Also, $R_U(0) = R_L(0) = 1$, and $R_U(1) = R_L(1) = 0$, so let $0 \leqslant \delta_U \leqslant \delta_L \leqslant 1$ be the smallest values of δ at which R_U resp. R_L vanish. It is well known that $\delta_L = \frac{1}{2}$ and we conjecture that δ_U is strictly smaller than $\frac{1}{2}$, i.e., the rate of linear codes with the unique-neighbor property is more strictly bounded. While the conjecture remains open, we prove here several results supporting it.

The reader is not assumed to have any specific background in coding theory, but we occasionally point out some relevant facts from that area.

Keywords. Unique neighbor, linear code

Mathematics Subject Classifications. 05D99, 94B65

*Supported in part by an ERC Grant 101141253, "Packing in Discrete Domains – Geometry and Analysis" and a NSF-BSF research grant "Global Geometry of Graphs".

†Work was done while at the Hebrew University.

1. Introduction and main problems

We consider here only binary codes $\mathcal{C} \subseteq \{0,1\}^n$ of length n. As usual, we denote the *rate* of $\mathcal{C}$ by $R = R(\mathcal{C}) = \frac{1}{n}\log_2|\mathcal{C}|$ and its distance by $\delta n = \mathrm{dist}(\mathcal{C}) = \min_{x\neq y,\, x,y\in\mathcal{C}} d_H(x,y)$, where d_H stands for the Hamming distance. A fundamental open problem in coding theory seeks the best possible tradeoff between R and $0 \leqslant \delta \leqslant 1$. We refer to this as

Problem 1.1. Determine, or estimate the real function

$$R(\delta) = \limsup_{n\to\infty}\{R(\mathcal{C}) \mid \mathcal{C} \subseteq \{0,1\}^n,\ \ \mathrm{dist}(\mathcal{C}) \geqslant \delta n\}.$$

A *linear code* is a linear subspace of the vector space $\mathbb{F}_2^n$, which we identify with $\{0,1\}^n$. Such a code can be defined in terms of a *parity check matrix* A which is a $\lceil(1-R)n\rceil \times n$ binary matrix. Namely, $\mathcal{C} = \{x \in \mathbb{F}_2^n \mid Ax = 0\}$.

Let S be a nonempty set of columns in a binary matrix and let z be the sum *over the integers* of the columns in S. We say that the set S is 1-*free* if no entry of z equals 1, and we call S *even*, if all entries of z are even integers. The number of 1-entries in a binary vector is said to be its *weight* or *sum*.

Here are some basic comments:

Remark 1.2.

- A given linear code $\mathcal{C}$ can clearly have various distinct parity check matrices A. While the distance of $\mathcal{C}$ does not depend on the choice of A, different parity check matrices of the same code may have different 1-free sets.
- If we delete from A any row that linearly depends on the other rows, the code $\mathcal{C}$ does not change. Consequently, all parity check matrices herein can, with no loss in generality be assumed to have full row rank.
- Coding theorists refer to a 1-free set as a *stopping set*, see Section 1.1.
- We consider below also vanishingly small rates $R = o_n(1)$, in which case we may speak of an $n \times (n + f(n))$ parity check matrix with $f(n) = o(n)$. Although this deviates from our standard of considering length-n codes, the asymptotic conclusions remain unchanged.

We are ready now to introduce the protagonists of this article.

Definition 1.3. Let A be a binary $m \times n$ matrix of rank m. Define:

- The smallest cardinality of a nonempty even set of columns in A, i.e., the *distance* of the binary linear code with parity check A is denoted by $\varepsilon(A)$.
- The smallest cardinality of a nonempty 1-free set of columns in A is denoted by $u(A)$.
- The maximum value of $\varepsilon(A)$ over all binary $m \times n$ matrices is denoted $\varepsilon(m,n)$.
- The maximum value of $u(A)$ over all binary $m \times n$ matrices is denoted $u(m,n)$.

The best available estimates of $\varepsilon(m,n)$ are extensively tabulated, e.g., in [Gra07]. We seek to likewise determine or estimate the values of $u(m,n)$. Note the following:

Remark 1.4.
- The claim that $T \leqslant u(m,n)$ means that there exists a binary $m \times n$ matrix A of rank m where every nonempty set of at most T columns has a row of weight 1.
- Proving that $u(m,n) \leqslant T$ entails showing that in every binary $m \times n$ matrix A of rank m there is a set of T or fewer columns for which no row weighs 1.

In analogy with Problem 1.1, the following asymptotic question suggests itself:

Problem 1.5. Determine, or estimate the real function

$$R_L(\delta) \coloneqq \limsup_{n\to\infty}\{R \mid \text{there exists a } \lceil (1-R)n\rceil \times n \text{ binary matrix } A \text{ with } \varepsilon(A) \geqslant \lfloor \delta n \rfloor\}.$$

In other words, $R_L(\delta)$ is the smallest real R such that

$$\text{For any } \rho > R \text{ and large enough } n, \text{ every } \lceil (1-\rho)n\rceil \times n \text{ binary matrix}$$

$$\text{has an even set of } \leqslant \delta n \text{ columns.}$$

And for 1-free sets:

Problem 1.6. Determine, or estimate the real function

$$R_U(\delta) \coloneqq \limsup_{n\to\infty}\{R \mid \text{there exists a } \lceil (1-R)n\rceil \times n \text{ binary matrix with } u(A) \geqslant \lfloor \delta n \rfloor\}.$$

In other words, $R_U(\delta)$ is the smallest real R such that

$$\text{For any } \rho > R \text{ and large enough } n, \text{ every } \lceil (1-\rho)n\rceil \times n \text{ binary matrix}$$

$$\text{has a 1-free set of } \leqslant \delta n \text{ columns.}$$

Clearly,

$$R_U(\delta) \leqslant R_L(\delta) \leqslant R(\delta) \text{ for all } 0 \leqslant \delta.$$

At present, we cannot even rule out the possibility that all these three functions are, in fact, identical. It is easily verified that (i) All three are non-increasing functions of δ, and (ii) $R(0) = R_L(0) = R_U(0) = 1$. It is also well known that $R(\delta), R_L(\delta)$ are positive for $\delta < \frac{1}{2}$ (by the Gilbert–Varshamov bound for linear codes [Gil52, Var57]) and $R(\delta), R_L(\delta) = 0$ for $\frac{1}{2} \leqslant \delta$ (e.g., by the Plotkin bound [Plo60]).

We believe that the strict inequality $R_U(\delta) < R_L(\delta)$ holds for at least some of the range $0 < \delta < \frac{1}{2}$. More specifically that R_U vanishes already at some $\delta_0 < \frac{1}{2}$. Concretely, we state

Conjecture 1.7. There exists some $\epsilon_0 > 0$ such that function $R_U(\delta)$ vanishes already at $\delta = \frac{1}{2} - \epsilon_0$.

In other words, for every $R > 0$ and large enough n, every $\lceil (1-R)n\rceil \times n$ binary matrix has a 1-free set of at most $(\frac{1}{2} - \epsilon_0)n$ columns.

It is conceivable that the same conclusion holds even for $R = o_n(1)$ and with a specified $o(1)$-term.

Conjecture 1.8. There exists some $\epsilon_0 > 0$ and a positive function $\eta(n) = o(n)$ such that for every large enough n, every $\lceil n - \eta(n) \rceil \times n$ binary matrix has a 1-free set of at most $(\frac{1}{2} - \epsilon_0)n$ columns.

Let A be a parity check matrix of a linear code $\mathcal{C} \subseteq \{0,1\}^m$. Of course $\mathcal{C}$ remains invariant under elementary row operations on A. Also, distances among vectors in $\mathcal{C}$ remain unchanged as A's columns get permuted. Consequently, in the study of $R_L(\delta)$ as in Problem 1.5, there is no loss of generality in assuming that A is in *standard form*, i.e., its first n columns form an order-n identity matrix. We pose:

Problem 1.9. Let n, k be positive integers, and let A be a binary $n \times (n + k)$ matrix whose first n columns form the order-n identity matrix. How large can $u(A)$ be?

Below, we often use the invariance of u and ε under row and column permutations.

1.1. Unique-neighbor codes in other contexts

The function u was previously (e.g., [DPT+02, KV03, OVZ05, SV06]) defined under the name of the *stopping distance* of a matrix. This notion arises in the study of binary erasure channels (see, e.g., [Gur06]). For further algorithmic aspects of stopping sets see [JXF10, JXF11, LMSS01, PH17, Rat06].

The closely related notion of *unique-neighbor expansion* of codes appears in the study of message-passing algorithms and expander codes (see [HLW06] for a survey). Such algorithms offer an approach to the *decoding* of linear codes $\mathcal{C} = \{x \in \mathbb{F}_2^m \mid Ax = 0\}$. Here A is viewed as the bipartite adjacency matrix $(U, V; E)$ of the code's *factor graph* (or Tanner graph [Tan81]). Here U, V are the sets of A's rows and columns, and edges correspond to 1-entries in A. Message-passing algorithms such as *belief propagation* [Gal62] work by iteratively passing messages between vertices in U and those in V.

When the factor graph is a bounded-degree expander graph, we say that $\mathcal{C}$ is an *expander code*. Such codes belong to the class of low density parity check (LDPC) codes introduced by Gallager [Gal62]. An important feature of such codes is that they can be efficiently decoded, using message-passing algorithms [Gal62, RU01, ZP75]. In a highly influential paper [SS96], Sipser and Spielman showed that message-passing algorithms can efficiently decode expander codes even when linearly many (in n) errors occur. The performance of the algorithm depends on the unique-neighbor expansion of the graph's bipartite adjacency matrix. For more on this subject, see [DG17, Gur06, HLW06, RU08, Sho04, Vid13].

Unique-neighbor expanders and unique-neighbor codes are still not sufficiently well understood. Alon and Capalbo [AC02] found explicit constructions of bipartite graphs which are (α, β)-unique-neighbor expander graphs, where α, β are some positive absolute constants and $\frac{|V|}{|U|}$ is bounded away from 1. See also [Bec16] and [BSV09] for more. In a recent paper [HLM+24], explicit unique-neighbor expanders are constructed with α bounded away from 0 and $\beta \approx \frac{3}{5}$, which surpasses spectral methods that achieve $\beta \leqslant \frac{1}{2}$.

2. Our new results

Our work addresses Problems 1.6 and 1.9. Problems 1.1 and 1.5 are mentioned here for context only.

1. Theorem 3.1 shows that matrices A in which all row sums are 9 or more, satisfy Conjecture 1.7 with $\epsilon_0 = 0.03$ even when $R = 0$.

2. Theorem 6.1 shows that the conclusion of Conjecture 1.8 fails when $\eta(n) < \log_2(n)$.

3. Theorem 4.1 answers Problem 1.9: $u_I(n, n+k) = \frac{n}{H_k} \pm O_n(1)$. Here H_k is the k-th harmonic sum. In particular we prove Conjecture 1.7 for matrices in standard form.

4. Clearly $u(m,n) \leqslant \varepsilon(m,n)$ for all m and n. Theorem 7.1 shows that $u = \varepsilon$ when $n-m \leqslant 3$ and that $(m,n) = (4,8)$ is the first case where $u < \varepsilon$.

3. With lower bounds on row weights

As we show next, matrices of sufficiently large row weights satisfy Conjecture 1.7. In contrast, small 1-free sets seem harder to find in sparse matrices. The case of row weights 3 may be the hardest. The theorem below, as well as Theorem 6.1 and Theorem 7.1 exhibit matrices with row weights all either 3 or 4 for which the conclusion of Conjecture 1.7 fails to hold.

Theorem 3.1. *1. If every row in a binary $n \times n$ matrix A has weight at least 9, then $u(A) < 0.47n$. Namely, A must have an $n \times n'$ submatrix with $n' < 0.47n$ in which no row has weight 1.*

2. On the other hand, for every n there exists a binary $n \times n$ matrix A where every row has weight 4, such that $u(A) = \lceil \frac{n}{2} \rceil$. Namely, every $n \times n'$ submatrix of A with no row of weight 1 satisfies $n' \geqslant \frac{n}{2}$.

3. For every n there also exists a binary $n \times n$ matrix B where every row has weight 3, such that $u(B) = \lceil \frac{2n}{3} \rceil$.

Proof. Item 1: Let $Z^{(1)}$ be a random set of columns in A that is obtained by picking every column independently with probability ρ. Let $B^{(1)}$ be the resulting $[n] \times Z^{(1)}$ submatrix of A, and let $X_0^{(1)}, X_1^{(1)}$ the set of rows in $B^{(1)}$ of weight zero, resp. one. For every row $i \in X_1^{(1)}$ add some column to $Z^{(1)}$ whose i-th entry is 1, and let $Z^{(2)} \supseteq Z^{(1)}$ be the resulting, extended column set. We denote the $[n] \times Z^{(2)}$ submatrix of A by $B^{(2)}$. Let $X_0^{(2)}, X_1^{(2)}$ be the set of rows in $B^{(2)}$ of weight 0, resp. 1. Note that $X_0^{(2)}, X_1^{(2)} \subseteq X_0^{(1)}$. We proceed in the same way with $X_0^{(2)}, X_1^{(2)}$ until no row in the submatrix has weight 1, and let $Z = \cup_k Z^{(k)}$ be the column set of the resulting matrix. Note that

$$|Z| \leqslant |Z^{(1)}| + |X_0^{(1)}| + |X_1^{(1)}|,$$

because every column in $Z \setminus Z^{(1)}$ accounts for some row in $X_0^{(1)} \cup X_1^{(1)}$ that no other column in $Z \setminus Z^{(1)}$ does. Therefore

$$\mathbb{E}(|Z|) \leqslant \mathbb{E}(|Z^{(1)}| + |X_0^{(1)}| + |X_1^{(1)}|).$$

We have

$$\mathbb{E}(|Z^{(1)}|) = \rho n\,; \quad \mathbb{E}(|X_0^{(1)}|) \leqslant n(1-\rho)^c + o(n)\,; \quad \mathbb{E}(|X_1^{(1)}|) \leqslant nc\rho(1-\rho)^{c-1} + o(n),$$

if the Hamming weight of every row in A is at least c. The claim follows by observing that

$$\mathbb{E}(|Z|) < 0.47n \text{ for } c = 9,\ \rho = 0.3757.$$

Item 2: Let A be the $n \times n$ binary matrix whose rows are comprised of all n cyclic rotations of the vector $1^4 0^{n-4}$. Given a vector $x \in \{0,1\}^n \setminus \{\mathbf{0}\}$ (here $\mathbf{0}$ is the all-zero vector of length n), let the vector Ax be defined by integer arithmetic. Clearly, $Ax \in \{0,1,2,3,4\}^n$. Multiply on the left by the all-1 vector to conclude that $\|Ax\|_1 = 4\|x\|_1$. Note next that if $(Ax)_i = 0$, then $(Ax)_{i+1 \bmod n}$ is either 0 or 1. Therefore, if Ax has no 1 coordinates, then all its coordinates are at least 2, so that $4\|x\|_1 = \|Ax\|_1 \geqslant 2n$ and $\|x\|_1 \geqslant \frac{n}{2}$, so $u(A) \geqslant \frac{n}{2}$. A 1-free set of size $\lceil \frac{n}{2} \rceil$ is the set $\{1, 3, \ldots, 2\lceil \frac{n}{2} \rceil - 1\}$. Let $x \in \{0,1\}^n$ be its indicator vector. Since every two neighboring columns in the selected set have distance at most 2 (where the distance between 1 and n is 1), then for every row i, $\|(Ax)_i\|_1 \geqslant 2$ as claimed.

Item 3: The matrix B is similar to A but with row weights of 3, and a similar argument yields $u(B) = \lceil \frac{2n}{3} \rceil$. To show that $u(B) \geqslant \frac{2n}{3}$, let $x \in \{0,1\}^n$ be the indicator vector of a 1-free set of columns in B. As before, $3\|x\|_1 = \|Bx\|_1 \geqslant 2n$, so that $\|x\|_1 \geqslant \frac{2n}{3}$ as claimed. To show that $u(B) \leqslant \lceil \frac{2n}{3} \rceil$, notice that the set $\{j \in [n] \mid j \not\equiv 0 \bmod 3\}$ is 1-free column set with $\lceil \frac{2n}{3} \rceil$ columns. The conclusion follows. □

Item 2 of Theorem 3.1 reflects on the validity of Conjecture 1.7. It shows that to guarantee the existence of small 1-free sets of columns, we must consider matrices with more columns than rows. This statement is made quantitative in Theorem 6.1.

We suspect that Item 1 of Theorem 3.1 remains valid even when all row weights are at least 5. However, this seems to require a substantial new idea.

4. Matrices in standard form

We denote by $u_I(m,n)$ the maximum of $u(A)$ for a binary $m \times n$ matrix in standard form $A = [I_m|B]$, where I_m stands for the identity $m \times m$ matrix. Answering Problem 1.9, we give an upper bound on $u_I(m,n)$ that is tight in infinitely many cases.

Theorem 4.1. *For every positive integer k and $n \to \infty$, every binary $n \times (n+k)$ matrix of the form $A = [I_n|B]$ has a 1-free set of at most $\frac{n}{H_k} + k$ columns where $H_k = \sum_{\ell=1}^{k} \frac{1}{\ell}$ is the k-th harmonic sum. For fixed k, the bound is tight, i.e., $\frac{n}{H_k} - \exp((1+o_k(1))k) \leqslant u_I(n, n+k)$.*

Proof. The vector $z \in \{0,1\}^{n+k}$ is the characteristic vector of a 1-free set of columns in A if and only if the vector Az (computed over the reals) has no 1-coordinates. It is convenient to express z as the concatenation of $x \in \{0,1\}^n$ and $y \in \{0,1\}^k$. Clearly, $Az = x + By$, and Az has no 1-coordinates if and only if x is the characteristic vector of a set that contains every 1-coordinate in By, and none of its 0-coordinates. Therefore, the smallest size of a 1-free set of columns in A is

$$u(A) = \min\{\|y\|_1 + \text{ the number of 1-coordinates in } By\} \text{ over all } \mathbf{0} \neq y \in \{0,1\}^k. \quad (4.1)$$

In this view, the order at which B's rows appear is immaterial, and all we care is, given $v \in \{0,1\}^k$, how many rows in B equal v. We denote this number by c_v. If a row of B equals v, then the corresponding coordinate in By equals $\langle v, y\rangle$, where $\langle\cdot,\cdot\rangle$ stands for inner product over the reals. Therefore, we can rewrite Equation (4.1) as

$$u(A) = \min\left\{\|y\|_1 + \sum_{v,\,\langle v,y\rangle=1} c_v\right\} \text{ over all } \mathbf{0} \neq y \in \{0,1\}^k. \quad (4.2)$$

It transpires that the theorem seeks the maximum of the expression in Equation (4.2), and this over the choice of the matrix B. However, as we saw, all we need to know about B are the nonnegative integers c_v, where clearly $\sum_{v\in\{0,1\}^k} c_v = n$. This means that the answer to our problem can be expressed as the optimal value of an integer linear program.

To simplify matters, we neglect the term $\|y\|_1$ in Equation (4.2). This causes only a minor loss, since $1 \leqslant \|y\|_1 \leqslant k$, while $n \to \infty$. Furthermore, without loss of generality, to maximize the expression in Equation (4.2) we can set $c_{\mathbf{0}} = 0$. Having made these simplifications, we seek the smallest integer $m = m(k,n)$ for which the following statement holds true:

$$\text{Given any } 2^k - 1 \text{ integers } c_v \geqslant 0 \text{ indexed by } v \in \{0,1\}^k \setminus \{\mathbf{0}\} \text{ with } \textstyle\sum_{\mathbf{0}\neq v\in\{0,1\}^k} c_v = n,$$

$$\text{there is some } \mathbf{0} \neq y \in \{0,1\}^k \text{ with } \textstyle\sum_{v,\langle v,y\rangle=1} c_v \leqslant m.$$

Let M be the $(2^k-1)\times(2^k-1)$ binary matrix that is indexed by $\{0,1\}^k \setminus \{\mathbf{0}\}$, whose (u,v) entry equals 1 if and only if $\langle u, v\rangle = 1$. The above statement can be restated as follows:

$$\forall \text{ integer vector } c \geqslant \mathbf{0} \text{ with } \textstyle\sum_{\mathbf{0}\neq w\in\{0,1\}^k} c_w = n, \text{ some coordinate in } Mc \text{ corresponding to a}$$

$$\mathbf{0} \neq y \in \{0,1\}^k \text{ is } \leqslant m.$$

In other words, $m = m(k,n)$ is the largest integer t for which there is some c as above such that

$$Mc \geqslant \mathbf{1} \cdot t.$$

It follows that

$$\Phi \leqslant u_I(n, n+k) \leqslant \Phi + k,$$

(the $+k$ slack reflects the neglection of the term $\|y\|_1$) where

$$\begin{aligned} \Phi = \max t, \\ \text{subject to} \quad Mc \geqslant \mathbf{1} \cdot t, \\ \langle c, \mathbf{1} \rangle = n \text{ and } c \geqslant \mathbf{0} \text{ is a vector of integers,} \end{aligned} \tag{4.3}$$

where $\mathbf{1} \in \{0,1\}^{2^k}$ is the all-1 vector.

We turn to solve the rational relaxation of the above ILP.

$$\begin{aligned} \max t, \\ \text{subject to} \quad Mc \geqslant \mathbf{1} \cdot t, \\ \langle c, \mathbf{1} \rangle = n \text{ and } c \geqslant \mathbf{0}. \end{aligned} \tag{4.4}$$

Clearly, this optimum is an upper bound on Φ. Using the idea of LP duality, we left-multiply Inequality (4.4) by the following 2^k-dimensional vector w:

$$w_v = \frac{1}{\binom{k-1}{|v|-1}} \text{ for every } v \neq \mathbf{0} \text{ in } \{0,1\}^k.$$

Clearly, if $v \in \{0,1\}^k$ with $|v| = j$ for some $1 \leqslant j \leqslant k$ then

$$(w^T M)_v = \sum_{i=1}^{k} \frac{1}{\binom{k-1}{i-1}} j \binom{k-j}{i-1} = \frac{j!(k-j)!}{(k-1)!} \sum_{i=1}^{k} \binom{k-i}{j-1} = \frac{j!(k-j)!}{(k-1)!} \binom{k}{j} = k. \tag{4.5}$$

The first equality follows from the definition. The second only involves reorganizing terms. The third one uses the standard and easy fact that for all positive integers $s \leqslant N$ it holds that

$$\sum_{s \leqslant r \leqslant N} \binom{r}{s} = \binom{N+1}{s+1}.$$

As we left-multiply Inequality (4.4) by w, Equation (4.5) yields $w^T M = \mathbf{1} \cdot k$.

Also

$$\langle w, \mathbf{1} \rangle = \sum_{i=1}^{k} \frac{\binom{k}{i}}{\binom{k-1}{i-1}} = kH_k.$$

Therefore,

$$kH_k t = w^T \mathbf{1} \cdot t \leqslant w^T Mc \leqslant k \cdot \mathbf{1}^T c \leqslant kn \;\Rightarrow\; t \leqslant \frac{n}{H_k}.$$

The optimum of the LP is an upper bound on the optimum of the ILP. Therefore

$$u_I(n, n+k) \leqslant \frac{n}{H_k} + k.$$

For the lower bound in the theorem we need a lower bound on the ILP (4.3). To this end we define

$$z \coloneqq n\frac{w}{kH_k},$$

and observe that the above calculations yield $\langle z, \mathbf{1}\rangle = n$, and $Mz = \mathbf{1}\frac{n}{H_k}$. So $\frac{n}{H_k}$ is a lower bound on the LP (4.4). We now use this to lower bound the ILP.

Let $H_k \coloneqq \frac{a_k}{b_k}$ written as a reduced rational. If n is divisible by a_k, then $u_I(n, n+k) = \frac{n}{H_k}+k'$, for some $0 \leqslant k' \leqslant k$, because in this case the optimal solutions to our LP and the ILP coincide. In the general case, say $n = n_1a_k+n_2$ where $0 \leqslant n_2 < a_k$, then by the structure of our matrices,

$$u_I(n, n+k) \geqslant u_I(n, n+k) - u_I(n_2, n_2+k) \geqslant n_1b_k - (n_2+1) \geqslant n_1b_k - a_k.$$

There holds

$$\begin{aligned} a_k = H_kb_k \leqslant H_k \operatorname{lcm}(1,\dots,k) &= (\ln k + \gamma + O_k(1/k)) \exp((1+o_k(1))k) \\ &= \exp((1+o_k(1))k), \end{aligned}$$

where we employed the following two well known facts:

- $H_k = \frac{a_k}{b_k} = \ln k + \gamma + O_k(1/k)$, where γ is Euler's constant.
- $b_k = \operatorname{lcm}(1,\dots,k) = \exp((1+o_k(1))k)$, where $\ln(\operatorname{lcm}(1,\dots,k))$ is Chebyshev's second function.

The lower bound on $u_I(n, n+k)$ now follows. □

This implies Conjecture 1.7 for these matrices:

Corollary 4.2. *For every $n \geqslant 400$ and $k \geqslant 4$ there holds $u_I(n, n+k) \leqslant 0.49n$.*

Proof. When $k \geqslant 4$, $u_I(n, n+k) \leqslant u_I(n, n+4) \leqslant \frac{n}{H_4} + 4 = 0.48n + 4 \leqslant 0.49n$ for all $n \geqslant 400$. □

5. Some useful constructions

In what follows we make occasional comments that pertain to some well-known families of codes, that the reader may find interesting. However, no familiarity with coding theory is assumed. All the relevant background information may be found in standard texts such as [RU08] and [vL98].

In this section we introduce several constructions of binary matrices. The building blocks of these constructions are binary $(2^k - 1 - k) \times (2^k - 1)$ matrices called U_k that we define for $k = 2, 3, \dots$. Further building blocks are $U_{k,m}$, which are $((2^k-1)m - k) \times ((2^k-1)m)$ binary matrices. We define U_k both recursively and directly. It is easy to verify by a simple inductive argument that the two definitions coincide. Here is the recursive definition:

$$U_2 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}, \tag{5.1}$$

$$U_{k+1} = \begin{pmatrix} I_{2^k-1} & \mathbf{1}_{2^k-1\times 1} & I_{2^k-1} \\ \mathbf{0} & \mathbf{0} & U_k \end{pmatrix}, \tag{5.2}$$

where $\mathbf{1}_{p\times q}$ is the all-1 matrix of dimensions $p \times q$.

In the direct definition of U_k we index its columns by all integers $2^k-1, 2^k-2, \ldots, 1$, in this order. The rows are indexed by the subsequence of the above excluding the powers of 2. Each row of U_k has weight 3. If the integer $m \in \{1, \ldots, 2^k-1\}$ is not a power of 2, such that $2^t < m < 2^{t+1}$ for an integer t, then the three 1 entries in row m appear in columns $m, m-2^t$ and 2^t.

For example,

$$U_3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix},$$

with rows called $7, 6, 5, 3$ in this order and columns called $7, \ldots, 1$.

We note that U_k is, by definition, a generator matrix of the $[2^k-1, 2^k-1-k, 3]_2$ Hamming code, and a parity check matrix of the corresponding simplex code (shortened Hadamard code). Its rows are linearly independent, since it has an upper-triangular square submatrix with 1's on its diagonal. This matrix, called T_k, is attained by erasing those columns of U_k that correspond to powers of two. The remaining submatrix is called R_k, namely the columns of U_k whose index is a power of 2. For example,

$$T_3 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \; R_3 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

An interesting note is that this construction exploits the duality between Hamming codes and simplex codes. A Hamming code is generated by a matrix with row weights 3, so it is relatively easy to derive a lower bound b on $u(U_k)$. On the other hand, all non-zero codewords in the order-N simplex code have weight $\frac{N+1}{2}$. So for $N = 2^k-1$ this duality yields an upper bound of 2^{k-1} on $\varepsilon(U_k)$. The resulting inequality reads $b \leqslant u(U_k) \leqslant \varepsilon(U_k) \leqslant 2^{k-1}$, and when $b = 2^{k-1}$ holds true, they are both equalities.

We next construct the following $((2^k-1)m-k) \times ((2^k-1)m)$ binary matrix for every $k \geqslant 2, m \geqslant 1$:

$$U_{k,m} = \left(\begin{array}{ccc|c} T_k & & & R_k \\ & \ddots & & \vdots \\ & & T_k & R_k \\ \hline & & & I_k \\ & & I_{(m-1)k} & \vdots \\ & & & I_k \end{array}\right). \tag{5.3}$$

The numbers of T_k and R_k blocks are m, and empty blocks are all-zero blocks.

6. Conjecture 1.8 fails for sub-logarithmic η

We prove next that $u(n-\log_2 n-1,n-1)=\varepsilon(n-\log_2 n-1,n-1)=\frac{n}{2}$ for infinitely many integers n. Concretely,

Theorem 6.1. *For every integer $k\geqslant 2$ it holds that*

$$u(2^k-1-k,2^k-1)=\varepsilon(2^k-1-k,2^k-1)=2^{k-1}.$$

Proof. We proceed by showing that

$$\varepsilon(2^k-1-k,2^k-1)\leqslant 2^{k-1} \text{ and } u(U_k)\geqslant 2^{k-1}.$$

We start with the upper bound on ε:

Proposition 6.2. *Every $n\times(n+k)$ binary matrix A has an even set of at most $\left(1+\frac{1}{2^k-1}\right)\frac{n+k}{2}$ columns.*

In particular, $\varepsilon(2^k-1-k,2^k-1)\leqslant 2^{k-1}$.

Proof. We provide two proofs, one linear algebraic and one that appeals to known bounds in coding theory. Consider the linear code $\mathcal{C}=\{x\in\mathbb{F}_2^{n+k}\mid Ax=0\}$. As usual, we may and do assume that A has rank n, so that $|\mathcal{C}|=2^k$. Clearly, $\mathcal{C}$ is invariant under addition of any given vector $x\in\mathcal{C}$. Consequently, for any $1\leqslant i\leqslant n+k$ either the i-th coordinate is identically zero in $\mathcal{C}$ or there are exactly 2^{k-1} vectors in $\mathcal{C}$ whose i-th coordinate is zero, resp. one. Consequently, the average Hamming weight of vectors in $\mathcal{C}$ is at most $\frac{n+k}{2}$. It follows that the average weight of non-zero codewords in $\mathcal{C}$ is at most $\frac{2^{k-1}(n+k)}{2^k-1}=\left(1+\frac{1}{2^k-1}\right)\frac{n+k}{2}$. The conclusion follows.

Alternatively we can appeal to Griesmer's bound [Gri60] which says that the length of a k-dimensional binary linear code of minimum distance d is at least $\sum_{i=0}^{k-1}\left\lceil\frac{d}{2^i}\right\rceil$. As we remove ceilings we conclude that this expression is at least $d(2-\frac{1}{2^{k-1}})$. After simplifying terms we arrive at $\frac{2^k-1}{2^{k-1}}\varepsilon(A)\leqslant n+k$, as claimed.

The second statement follows by letting $n=2^k-1-k$. □

We turn to prove that $u(U_k)\geqslant 2^{k-1}$. The proof proceeds by induction on $k\geqslant 2$. For $k=2$ the claim clearly holds. For the induction step we use the recursive description of U_{k+1} in Section 5. Consider a 1-free set of U_{k+1}. If it contains column 2^k, then it must include at least 2^k-1 additional columns (from either side of the column), for a total of at least 2^k columns, as claimed.

Thus it suffices to consider a 1-free set of columns of the form $L\sqcup R$, where R,L are the subsets of columns from the 2^k-1 rightmost, leftmost ones, respectively. Using the recursive definition in Equation (5.2), we have $|R|\geqslant 2^{k-1}$ by the induction hypothesis. Furthermore, for every $r\in R$, one of its first (upper) 2^k-1 coordinates contains a 1. Since column 2^k is absent from $L\sqcup R$, the (unique) matching column from the 2^k-1 leftmost columns should be included in L, in order for $L\sqcup R$ to be a 1-free set. It follows that $|L\sqcup R|\geqslant 2^k$, completing the proof. □

Notice that the bound in Proposition 6.2 holds with equality for the parity check matrices of simplex codes. In this case, the weight of all non-zero codewords coincides with the upper bound in Proposition 6.2.

The following extension of Theorem 6.1 yields further cases where ε and u coincide.

Theorem 6.3. *For every integers $k \geqslant 2$ and $m \geqslant 1$ it holds that*

$$u((2^k-1)m-k,(2^k-1)m) = \varepsilon((2^k-1)m-k,(2^k-1)m) = 2^{k-1}m.$$

Proof. Again we bound u from below and ε from above. The bound on u uses the matrices $U_{k,m}$ from Section 5 and the bound on ε follows from Proposition 6.2. To show that $u(U_{k,m}) \geqslant 2^{k-1}m$, we consider 1-free sets in $U_{k,m}$. The matrix $U_{k,m}$ with its last k columns removed has no nonempty 1-free sets, since it is an upper-triangular, full-rank matrix. So consider a 1-free set that includes $t > 0$ columns among the last k columns of $U_{k,m}$. By Theorem 6.1 at least $(2^{k-1}-t)m$ additional columns are needed, specifically at least $2^{k-1}-t$ from every T_k in the direct sum. The lower part of $U_{k,m}$ necessitates adding exactly $(m-1)t$ columns from the columns that contain the block $I_{(m-1)k}$. In total the cardinality of the 1-free set at hand is at least $t + (2^{k-1}-t)m + (m-1)t = 2^{k-1}m$, as claimed. □

Theorem 6.3 and the monotonicity of u, ε (see Proposition 7.2 below) yield

Corollary 6.4. *For every k, n it holds that $u(n, n+k) \geqslant \varepsilon(n, n+k) - 2^{k-1}$.*

7. Between u and ε when $n-m$ is bounded

In this section we compare $u(m,n)$ and $\varepsilon(m,n)$ when $n-m \geqslant 1$ is bounded from above. A most helpful resource in studying these problems is [Gra07], which records best possible linear codes of small lengths. Below we refer to the notion of an elementary collapse that we now define. Say that A is a binary matrix and $A(i,j) = 1$. If the i-th row has Hamming weight of 1, then *elementary collapse* with *pivot* (i,j) is the matrix that is obtained upon removing row i and column j from A.

Here is our main result:

Theorem 7.1. *1. $u(4,8) = 3$ whereas $\varepsilon(4,8) = 4$. This is the lexicographically first case where $u < \varepsilon$, i.e., the ordering that compares $n-m, m, n$ in this order.*

2. $u(n,n+1) = \varepsilon(n,n+1) = n+1$. The case of equality is fully characterized.

3. $u(n,n+2) = \varepsilon(n,n+2) = \left\lfloor \frac{2n+4}{3} \right\rfloor$.

4. If $n \not\equiv -1 \bmod 7$, then $u(n,n+3) = \varepsilon(n,n+3) = \left\lfloor \frac{4n+12}{7} \right\rfloor$. Also, $u(7m-1,7m+2) = 4m$ for every positive integer m.

Proof. We start with several simple observations:

Proposition 7.2. *1. $u(m,n) \leqslant \varepsilon(m,n) \leqslant m+1$.*

2. Both $u(m,n)$ and $\varepsilon(m,n)$ increase with m and decrease with n.

3. $u(m,n) \leqslant u(m+1,n+1)$, $\varepsilon(m,n) \leqslant \varepsilon(m+1,n+1)$.

4. The functions $u(\cdot), \varepsilon(\cdot)$ are invariant under elementary collapses.

Proof. 1. $u \leqslant \varepsilon$ by definition. $\varepsilon(m,n) \leqslant m+1$ because the columns of every $m \times (m+1)$ matrix are linearly dependent.

2. This is because every 1-free/even set of columns in an $(m+1) \times n$ matrix A is 1-free/even for every $m \times n$ submatrix of A, and these properties are preserved when any column is added to A.

3. Observe that $u(A) = u(B),\ \varepsilon(A) = \varepsilon(B)$ for every $m \times n$ matrix B and the matrix

$$A = \begin{pmatrix} B & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}.$$

4. If A has an elementary collapse with pivot (i,j), then no even resp. 1-free set of columns of A can contain j. Therefore u, ε remain unchanged under an elementary collapse with pivot (i,j). □

We now prove the 4 items of the theorem in order.

7.1. Proof of Item 1: $u(4,8) = 3 < 4 = \varepsilon(4,8)$.

As recorded in [Gra07], there holds $\varepsilon(4,8) = 4$ (see reference for a linear code that attains this bound). We now show that $u(4,8) = 3$. The following matrix yields $u(4,8) \geqslant 3$, since every nonempty set of 2 or fewer columns has a row of weight 1.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}. \tag{7.1}$$

Next we show that $u(A) \leqslant 3$ for every binary 4×8 matrix A. We reduce to the case that every column of A has weight at least 2. If A has a zero column, then clearly $u(A) = 1$. If some column of A has weight 1, say $a_{1,1} = 1$ and $a_{i,1} = 0$ for $i = 2,3,4$, consider the submatrix B of A that is obtained by erasing its first row and column. If B has an all-zero column, then $u(A) \leqslant 2$, and if B has two equal columns, then $u(A) \leqslant 3$. In the only remaining case

$$B = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}, \tag{7.2}$$

up to permutations of the rows and columns. We index the columns of B with $2, \ldots, 8$. Consider the weight $w = \sum_j a_{1,j}$ of row 1 in A. If $w = 1$, then $a_{1,j} = 0$ for all $j \geqslant 2$. Consequently, $u(A) = u(B) = 3$, since every 1-free set in B is also 1-free in A. If $w \geqslant 3$ there are at least two indices $1 < \beta < \alpha$ such that $a_{1,\alpha} = a_{1,\beta} = 1$. If columns α, β of B are the vectors u resp. v, then some column γ corresponds to $u \oplus v \pmod 2$ sum). Columns α, β, γ form a 1-free set in A. Finally, if $w = 2$, there is exactly one index $\delta > 1$ such that $a_{1,\delta} = 1$. Then we can find a triplet of columns of the form $u, v, u \oplus v$ in B none of which is column δ.

We can now assume that every column of A has weight $2, 3$ or 4. A has no repeated columns, or else $u(A) = 2$. Also A can have at most two columns of weight at least 3, for any three such distinct vectors form a 1-free set. Consequently A has exactly two columns of weight at least 3 and each of the six columns of weight 2. The latter 6-tuple contains a 1-free set of three columns. This establishes Item 1 of Theorem 7.1.

Note that $4, 8$ are the *minimal* m, n for which $u(m, n) < \varepsilon(m, n)$, with respect to the lexicographical ordering that first compares $n - m$, then m, n: the other parts of the present theorem show that equality holds when $n - m \leqslant 3$. One can verify that $u(m, n) = \varepsilon(m, n)$ for $(m, n) \in \{(1, 5), (2, 6), (3, 7)\}$ using the table in [Gra07].

7.2. Proof of Item 2

It is clear that $u(n, n+1) = \varepsilon(n, n+1) = n+1$. Also $\varepsilon(A) = n+1$ for an $n \times (n+1)$ matrix A if and only if its rank is n and all its row weights are even.

An example of an $n \times (n+1)$ matrix A with $u(A) = n+1$ is obtained by taking the edges vs. vertices incidence matrix A_T of a tree with $n+1$ vertices. As we show, no other examples exist.

Proposition 7.3. *If $u(A) = n+1$ for some $n \times (n+1)$ binary matrix A, then $A = A_T$ for some tree T with $n+1$ vertices.*

Proof. A cannot have a zero row, or else $u(A) \leqslant \varepsilon(A) \leqslant \frac{2(n+1)}{3}$, by Proposition 6.2.

As in Proposition 7.2, Item 4, any row of weight 1 in A can be collapsed, without changing ε and u. So we may and will assume that every row of A weighs at least 2. Let us view A as the edges vs. vertices incidence matrix of a hypergraph $G = (V, E)$. An edge in E of size 2 (resp. $\geqslant 3$) is called *light* (resp. *heavy*). Let $L \subseteq E$ be the set of light edges. If all edges in E are heavy, we can omit any single column of A and obtain a matrix in which all rows weigh at least 2, contrary to our assumption that $u(A) = n + 1$.

The graph (V, L) has no isolated vertices, for if $v \in V$ is incident with no light edge, then $V \setminus \{v\}$ is a 1-free set, contrary to our assumption. If $L = E$, then the vertex set of any connected component of G is a 1-free set. Therefore G is a connected graph with $n+1$ vertices and n edges, i.e., a tree, as claimed. On the other hand, if $L \neq E$, the graph (V, L) must be disconnected, since it has $n+1$ vertices and at most $n-1$ edges. In this case, let $(V_1, L_1), \ldots, (V_k, L_k)$ be the connected components of (V, L). By the above $\sum_i |V_i| = n + 1$, $|L_i| \geqslant |V_i| - 1$, so that $|L| = \sum_i |L_i| \geqslant n + 1 - k$, with equality if and only if (V, L) is a forest with no isolated vertices. Consequently, at most $k - 1$ edges in E are heavy.

Let B be the edges vs. vertices matrix of the hypergraph that results from G by shrinking each V_i to a single new vertex v_i. Since $L \neq \varnothing$ this actually reduces the size of the matrix and we can use induction to prove the proposition. Every 1-free set S in B yields a 1-free set in A by inflating each $v_i \in S$ to V_i. In particular $u(B) < k$ would imply $u(A) \leqslant n$. Consequently, B is a $(k - 1) \times k$ matrix with $u(B) = k$. By induction it is the edge-vertex matrix of K, a tree with vertex set $\{v_1, \ldots, v_k\}$. Say that v_1 is a leaf of K, and let e be the single edge of K that is incident with v_1. We claim that either V_1 or $V \setminus V_1$ comprise a 1-free set in A. Indeed, only the row corresponding to e may have weight 1 in the submatrix of A corresponding to either V_1 or $V \setminus V_1$. But it is impossible that both cases occur, for that would mean that the edge e has

size 2 contrary to the fact that e is a heavy edge. The set of vertices (or columns) V_1 or $V \setminus V_1$ in which e has weight at least 2 comprises a 1-free set of columns in A. Since both $V_1, V \setminus V_1$ are nonempty, this contradicts our assumption that $u(A) = n + 1$. This establishes the proposition and Item 2 of Theorem 7.1. □

7.3. Proof of Items 3 and 4

The proof for $k = 2$ splits to cases according to the value of $n \bmod 3$. When $n \equiv 1 \bmod 3$ we have $u(3m-2, 3m) = \varepsilon(3m-2, 3m) = 2m$ by Theorem 6.3. By Proposition 6.2, u, ε do not change as we move to $n = 3m - 1$. Finally, for $n = 3m$ we introduce the matrix

$$A \coloneqq \begin{pmatrix} U_{2,m} & \mathbf{0} \\ \mathbf{0} \quad I_2 & I_2 \end{pmatrix},$$

with $U_{2,m}$ as defined in Equation (5.3). It is easy to see that $u(A) = \varepsilon(A) = 2m + 1$. By Proposition 6.2 this is also the upper bound on $u(3m, 3m+2), \varepsilon(3m, 3m+2)$. We conclude that $u(n, n+2) = \varepsilon(n, n+2) = \left\lfloor \frac{2n+4}{3} \right\rfloor$, establishing Item 3.

The analysis when $k = 3$ is somewhat more involved and proceeds according to the value of $n \bmod 7$. We start with the upper bound: By Proposition 6.2, $\varepsilon(n, n+3) \leqslant \left\lfloor \frac{4n+12}{7} \right\rfloor$. This bound is shown to be tight, except if $n \equiv -1 \bmod 7$, when it can be reduced by 1 due to Griesmer's bound [Gri60]: indeed, our general upper bound is $\varepsilon(7m-1, 7m+2) \leqslant \left\lfloor \frac{28m+8}{7} \right\rfloor = 4m + 1$, but by Griesmer's bound if the code's distance is $4m + 1$, then its length is at least $4m + 1 + \left\lceil \frac{4m+1}{2} \right\rceil + \left\lceil \frac{4m+1}{4} \right\rceil = 7m + 3$. So $\varepsilon(7m-1, 7m+2) \leqslant 4m$.

We proceed to deal with the lower bounds. The case $k = 3$ of Theorem 6.3 gives $u(7m-3, 7m) = \varepsilon(7m-3, 7m) = 4m$. Namely, $u = \varepsilon$ when $n \equiv 4 \bmod 7$.

Item 3 of Proposition 7.2 and Proposition 6.2 yield

$$u(n-1, n+2) \leqslant u(n, n+3) \leqslant \varepsilon(n, n+3) \leqslant \left\lfloor \frac{4}{7}(n+3) \right\rfloor.$$

So if $u(n-1, n+2) = \left\lfloor \frac{4}{7}(n+3) \right\rfloor$, this trivially allows to derive the case $n \equiv r + 1 \bmod 7$ after establishing $u = \varepsilon$ in the case $n \equiv r \bmod 7$. This works verbatim for $r = 1, 4$. When $n \equiv 5 \bmod 7$, a similar argument establishes $u(7m-1, 7m+2) = \varepsilon(7m-1, 7m+2) = 4m$. It is left to establish the cases $n \equiv 0, 1, 3 \bmod 7$. Here, an additional argument is needed. To this end, we extend $U_{3,m}$ from Section 5 to an $n \times (n+3)$ matrix for the appropriate n. This resembles the construction of $U_{k,m}$ from U_k, and the construction in the case $k = 2$. In all three cases, these matrices show that $u(n, n+3)$ attains the upper bound on $\varepsilon(n, n+3)$, namely $\left\lfloor \frac{4}{7}(n+3) \right\rfloor$. Hence we get in each case a matrix U such that $\varepsilon(n, n+3) \leqslant \left\lfloor \frac{4}{7}(n+3) \right\rfloor \leqslant u(U) \leqslant u(n, n+3)$. For illustration, when $n = 7m$, we use the matrix $U_{3,m}$ to construct

$$U \coloneqq \begin{pmatrix} U_{3,m} & \mathbf{0} \\ \mathbf{0} \quad I_3 & I_3 \end{pmatrix}.$$

Note that $4m+1 = u(U) \leqslant u(7m, 7m+3)$ and $\varepsilon(7m, 7m+3) \leqslant 4m+1$ from Proposition 6.2, so $u(7m, 7m+3) = \varepsilon(7m, 7m+3) = 4m + 1$.

We note that Item 4 holds as well when $n = 1, 2, 3$, but we skip this verification.

This concludes the proof of Theorem 7.1. □

8. Open problems

Problem 8.1. The most obvious question is Conjecture 1.7 which remains open.

Problem 8.2. What is the smallest c for which the conclusion of Theorem 3.1 holds? Is it 5?

Problem 8.3. The proof of Theorem 3.1 suggests a more general setup. We seek a 1-free set of columns in a binary matrix A. Having committed to some subset of columns, the rows of A are split into: $I_0 \sqcup I_1 \sqcup I_*$, those of weight $0, 1$ and $\geqslant 2$, respectively. To extend our initially chosen set into a 1-free set, we need an additional set of columns J, the weight of whose I_0 and I_1 rows differ from $1, 0$ respectively. Under what conditions is it possible to pre-specify which row sums we wish to be $\neq 0$ and which $\neq 1$?

Problem 8.4. Let $u_3(m, n)$ denote $\max u(A)$ of an $m \times n$ binary matrix A where every row has weight 3. Proposition 7.3 implies that $u_3(n, n+1) < u(n, n+1)$, but perhaps $u_3(m, n) = u(m, n)$ when $n + 1 < m$. Some supportive evidence for this is that $u_3(4, 8) = u(4, 8)$, $u_3(2^k - 1 - k, 2^k - 1) = u(2^k - 1 - k, 2^k - 1)$. We note that more generally, $u_3((2^k - 1)m - 1, (2^k - 1)m) = u((2^k - 1)m - 1, (2^k - 1)m)$ holds, because the matrices $U_{k,m}$ can be modified so all rows have weight 3 without changing u, ε.

Remark 8.5. If indeed, Conjecture 1.7 is true, then its proof would require some substantial new ideas. Because, as Theorem 6.1 shows, methods that work for square matrices and matrices with only a few more columns than rows as in Theorem 3.1 and Theorem 6.3 are not likely to deliver a full answer.

References

[AC02] Noga Alon and Michael Capalbo. Explicit unique-neighbor expanders. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 73–79. IEEE, 2002. doi:10.1109/SFCS.2002.1181884.

[Bec16] Oren Becker. Symmetric unique neighbor expanders and good LDPC codes. *Discrete Applied Mathematics*, 211:211–216, 2016. doi:10.1016/j.dam.2016.04.022.

[BSV09] Eli Ben-Sasson and Michael Viderman. Tensor products of weakly smooth codes are robust. *Theory of Computing*, 5:239–255, 2009. doi:10.4086/toc.2009.v005a012.

[DG17] Michael Dowling and Shuhong Gao. Fast decoding of expander codes. *IEEE Transactions on Information Theory*, 64(2):972–978, 2017. doi:10.1109/TIT.2017.2726064.

[DPT+02] Changyan Di, David Proietti, I Emre Telatar, Thomas J Richardson, and Rüdiger L Urbanke. Finite-length analysis of low-density parity-check codes on the binary erasure channel. *IEEE Transactions on Information Theory*, 48(6):1570–1579, 2002. doi:10.1109/TIT.2002.1003839.

[Gal62] Robert Gallager. Low-density parity-check codes. *IRE Transactions on Information Theory*, 8(1):21–28, 1962. doi:10.1109/TIT.1962.1057683.

[Gil52] Edgar N Gilbert. A comparison of signalling alphabets. *The Bell System Technical Journal*, 31(3):504–522, 1952. doi:10.1002/j.1538-7305.1952.tb01393.x.

[Gra07] Markus Grassl. Bounds on the minimum distance of linear codes and quantum codes, 2007. Accessed on 2023-03-16. URL: http://www.codetables.de.

[Gri60] James H Griesmer. A bound for error-correcting codes. *IBM Journal of Research and Development*, 4(5):532–542, 1960. doi:10.1147/rd.45.0532.

[Gur06] Venkatesan Guruswami. Iterative decoding of low-density parity check codes (a survey). 2006. arXiv:cs/0610022.

[HLM+24] Jun-Ting Hsieh, Ting-Chun Lin, Sidhanth Mohanty, Ryan O'Donnell, and Rachel Yun Zhang. Explicit two-sided vertex expanders beyond the spectral barrier. 2024. arXiv:2411.11627.

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006. doi:10.1090/S0273-0979-06-01126-8.

[JXF10] Yong Jiang, Shu-Tao Xia, and Fang-Wei Fu. Stopping set distributions of some linear codes. 2010. arXiv:1003.0367.

[JXF11] Yong Jiang, Shu-Tao Xia, and Fang-Wei Fu. Stopping set distributions of some Reed-Muller codes. *IEEE transactions on Information Theory*, 57(9):6078–6088, 2011. doi:10.1109/TIT.2011.2162181.

[KV03] Navin Kashyap and Alexander Vardy. Stopping sets in codes from designs. In *IEEE International Symposium on Information Theory, 2003. Proceedings.*, page 122. IEEE, 2003. doi:10.1109/ISIT.2003.1228136.

[LMSS01] Michael G Luby, Michael Mitzenmacher, Mohammad Amin Shokrollahi, and Daniel A Spielman. Efficient erasure correcting codes. *IEEE Transactions on Information Theory*, 47(2):569–584, 2001. doi:10.1109/18.910575.

[OVZ05] Alon Orlitsky, Krishnamurthy Viswanathan, and Junan Zhang. Stopping set distribution of LDPC code ensembles. *IEEE Transactions on Information Theory*, 51(3):929–953, 2005. doi:10.1109/TIT.2004.842571.

[PH17] Aiden Price and Joanne Hall. A survey on trapping sets and stopping sets. 2017. arXiv:1705.05996.

[Plo60] Morris Plotkin. Binary codes with specified minimum distance. *IRE Transactions on Information Theory*, 6(4):445–450, 1960. doi:10.1109/TIT.1960.1057584.

[Rat06] Vishwambhar Rathi. On the asymptotic weight and stopping set distribution of regular LDPC ensembles. *IEEE Transactions on Information Theory*, 52(9):4212–4218, 2006. doi:10.1109/TIT.2006.880065.

[RU01] Thomas J Richardson and Rüdiger L Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Transactions on Information Theory*, 47(2):599–618, 2001. doi:10.1109/18.910577.

[RU08] Tom Richardson and Ruediger Urbanke. *Modern coding theory*. Cambridge university press, 2008. doi:10.1017/CBO9780511791338.

[Sho04] Amin Shokrollahi. LDPC codes: An introduction. In *Coding, cryptography and combinatorics*, pages 85–110. Springer, 2004. doi:10.1007/978-3-0348-7865-4_5.

[SS96] Michael Sipser and Daniel A Spielman. Expander codes. *IEEE Transactions on Information Theory*, 42(6):1710–1722, 1996. doi:10.1109/18.556667.

[SV06] Moshe Schwartz and Alexander Vardy. On the stopping distance and the stopping redundancy of codes. *IEEE Transactions on Information Theory*, 52(3):922–932, 2006. doi:10.1109/TIT.2005.864441.

[Tan81] R Tanner. A recursive approach to low complexity codes. *IEEE Transactions on Information Theory*, 27(5):533–547, 1981. doi:10.1109/TIT.1981.1056404.

[Var57] Rom Rubenovich Varshamov. Estimate of the number of signals in error correcting codes. *Docklady Akad. Nauk, SSSR*, 117:739–741, 1957. URL: https://www.mathnet.ru/eng/dan22571.

[Vid13] Michael Viderman. Linear-time decoding of regular expander codes. *ACM Transactions on Computation Theory (TOCT)*, 5(3):1–25, 2013. doi:10.1145/2493252.2493255.

[vL98] Jacobus Hendricus van Lint. *Introduction to coding theory*, volume 86 of *Graduate Texts in Mathematics*. Springer Science & Business Media, 1998. doi:10.1007/978-3-642-58575-3.

[ZP75] Victor Vasilievich Zyablov and Mark Semenovich Pinsker. Estimation of the error-correction complexity for Gallager low-density codes. *Problemy Peredachi Informatsii*, 11(1):23–36, 1975. URL: https://www.mathnet.ru/eng/ppi1568.