

WEIGHT DISTRIBUTION AND DECODING OF CODES ON HYPERGRAPHS

ALEXANDER BARG

Dept. of ECE and Institute for Systems Research
University of Maryland, College Park, MD 20742, USA
and
Institute for Problems of Information Transmission
Moscow, Russia

ARYA MAZUMDAR

Department of ECE
University of Maryland, College Park, MD 20742, USA

GILLES ZÉMOR

Institut de Mathématiques de Bordeaux
Université de Bordeaux 1
351 cours de la Libération, 33405 Talence, France

(Communicated by Marcus Greferath)

ABSTRACT. Codes on hypergraphs are an extension of the well-studied family of codes on bipartite graphs. Bilu and Hoory (2004) constructed an explicit family of codes on regular t -partite hypergraphs whose minimum distance improves earlier estimates of the distance of bipartite-graph codes. They also suggested a decoding algorithm for such codes and estimated its error-correcting capability.

In this paper we study two aspects of hypergraph codes. First, we compute the weight enumerators of several ensembles of such codes, establishing conditions under which they attain the Gilbert-Varshamov bound and deriving estimates of their distance. In particular, we show that this bound is attained by codes constructed on a fixed bipartite graph with a large spectral gap.

We also suggest a new decoding algorithm of hypergraph codes that corrects a constant fraction of errors, improving upon the algorithm of Bilu and Hoory.

I. INTRODUCTION

Codes on graphs account for some of the best known code families in terms of their error correction under low-complexity decoding algorithms. They are also known to achieve a very good tradeoff between the rate and relative distance. The most well-studied case is codes defined on a bipartite graph. In this construction, a code of length $N = mn$ is obtained by “parallel concatenation” of $2m$ codes of a small length n which refers to the fact that each bit of the codeword is checked by two independent length- n codes. The arrangement of parity checks is specified

2000 *Mathematics Subject Classification:* Primary: 94B25; Secondary: 94B35, 05C65.

Key words and phrases: Codes on graphs, hypergraphs, weight distribution, decoding.

The first author was supported in part by NSF grants CCF0515124 and CCF0635271 and by NSA grant H98230-06-1-0044.

The second author was supported in part by NSF grant CCF0635271.

by the edges of a bipartite graph which are in one-to-one correspondence with the codeword bits.

Codes on bipartite graphs are known to be asymptotically good, i.e., to have nonvanishing rate R and relative distance δ as the code length N tends to infinity. Constructive families of bipartite-graph codes with the best known tradeoff between R and δ have been found by the present authors [1]. In particular, codes constructed in that paper surpass the product bound on the minimum distance which is a common performance benchmark for concatenated constructions.

Moving from constructive families to existence results obtained by averaging over ensembles of bipartite-graph codes, it is possible to derive even better rate-distance tradeoffs. In particular, bipartite-graph codes with random local codes and random bipartite graphs attain the Gilbert-Varshamov (GV) bound for relatively small code rates and are only slightly below it for higher rates [1].

A natural way to generalize codes on bipartite graphs is to consider concatenations governed by regular t -partite hypergraphs, $t \geq 2$. This code family was studied by Bilu and Hoory in [2]. While constructive families of bipartite-graph codes rely on the expansion property of the underlying graph, expansion is not well defined for hypergraphs. Instead, [2] put forward a property of hypergraphs, called ε -homogeneity, which replaces expansion in the analysis of hypergraph codes. [2] showed that there exist explicit, easily constructible families of ε -homogeneous hypergraphs, and estimated the number of errors corrected by their codes under a decoding algorithm suggested in that paper.

In this paper we study hypergraph codes both from the perspective of weight distributions and their decoding. The results of [1] on weight distributions are advanced in several directions. In Theorem 2 and its corollary we prove that the code ensemble defined by random regular t -partite hypergraphs and random local linear codes contains codes that meet the GV bound. The region of code rates for which this claim holds true extends as t increases from the value $t = 2$. We also show (Theorem 7, Cor. 8) that the ensemble of hypergraph codes contains codes that attain the GV bound even if random hypergraphs are replaced with a *fixed* ε -homogeneous hypergraph. Specializing the last result for $t = 2$, we establish that expander codes of Sipser and Spielman [5] constructed from a fixed graph with a large spectral gap and random local codes with high probability attain the GV bound. Finally, we derive an estimate of the average weight distribution for the ensemble of hypergraph codes with a fixed local code (see Theorem 5) that refines substantially a corresponding result in [1] and generalizes it from $t = 2$ to arbitrary t .

The tradeoff between the rate and relative distance of hypergraph codes shows an improvement over bipartite-graph codes for small values of the distance. On the other hand, the decoding algorithm of [2] does not exploit the full power of their codes; moreover, for small δ the proportion of errors corrected by it vanishes compared to the value of the distance. Motivated by this, we propose a new decoding algorithm of hypergraph codes and estimate its error-correcting capability. We show that it corrects the number of errors which constitutes a fixed proportion of the code's distance.

I-A. CODES ON BIPARTITE GRAPHS. Let $G = (V, E)$ be a balanced, n -regular bipartite graph with the vertex set $V = V_1 \cup V_2$, $|V_1| = |V_2| = m$ and $|E| = N = nm$ edges. Let us choose an arbitrary ordering of edges in E . For a given vertex $v \in V$ this defines an ordering of edges $v(1), v(2), \dots, v(n)$ incident to it. We denote this

subset of edges by $E(v)$. Given a binary vector $x \in \{0, 1\}^N$, let us establish a one-to-one correspondence between the coordinates of x and the edges in E . For a given vertex v let $x(v) = (x_e, e \in E(v))$ be the subvector that corresponds to the edges in $E(v)$. Denote by λ the second largest in the absolute value eigenvalue of the graph G .

Consider a set of binary linear codes $A_v[n, R_1n]$ of length n and rate $R_1 \triangleq \dim(A_v)/n$, where $v \in V$. Define a *bipartite-graph code* as follows:

$$C(G, \{A_v\}) = \{x \in \{0, 1\}^N : \forall v \in V_1 \cup V_2 x(v) \in A_v\}.$$

The rate of the code C is easily seen to satisfy

$$(1) \quad R(C) \geq 2R_1 - 1.$$

If we assume that all the local codes are the same, i.e., $A_v = A$, where $A[n, R_1n, d_1 = \delta_1 n]$ is some linear code, then the distance of the code C can be estimated as follows:

$$d/N \geq \delta_1^2 \left(1 - \frac{\lambda}{d_1}\right)^2$$

(we will write $C(G, A)$ instead of $C(G, \{A\})$ in this case). In particular, if the spectral gap of G is large, i.e., λ is small compared to d_1 , then the relative distance d/N is close to the value δ_1^2 , similarly to the case of the direct product code $C = A \otimes A$.

The weight distribution of bipartite-graph codes constructed from random regular bipartite graphs and a fixed local code A with a known weight distribution was analyzed in [3, 4]. In particular, it was shown that if A is the Hamming code then the ensemble $\mathcal{C} = (C(G, A))$ contains asymptotically good codes. Generalizing these results, paper [1] studied the weight distribution of bipartite-graph codes with fixed and random component codes A . It was shown that for $m \rightarrow \infty$ the ensemble of codes constructed from random regular bipartite graphs and a fixed code A with distance $d_1 \geq 3$ contains asymptotically good codes. It has also been shown [1] that if the local codes are chosen randomly, then the code ensemble \mathcal{C} contains codes that meet the GV bound in the interval of code rates $R(C) \leq 0.202$.

I-B. CODES ON HYPERGRAPHS. Generalizing the above construction, let $H = (V, E)$ be a t -uniform t -partite n -regular hypergraph. This means that the set of vertices $V = V_1 \cup \dots \cup V_t$ of H consists of t disjoint parts of equal size, say, $|V_i| = m, 1 \leq i \leq t$. Every hyperedge $\{v_{i_1}, v_{i_2}, \dots, v_{i_t}\}$ contains exactly t vertices, one from each part, and each vertex is incident to n hyperedges. Below for brevity we say edges instead of hyperedges. The number of edges of H equals $N = mn$ which will also be the length of our hypergraph codes. As above, assume that the edges are ordered in an arbitrary fixed way and denote by $E(v)$ the set of edges incident to a vertex v . For definiteness, let us assume that edges $e_{(i-1)n+j}, j = 1, \dots, n$ are incident to the vertex $v_i \in V_1, 1 \leq i \leq m$.

Given a binary vector $x \in \{0, 1\}^N$ whose coordinates are in a one-to-one correspondence with the edges of H denote by $x(v)$ its subvector that corresponds to the edges in $E(v)$.

Define a *hypergraph code* as follows:

$$C(H, \{A_v\}) = \{x \in \{0, 1\}^N : \forall v \in V x(v) \in A_v\},$$

where $\{A_v, v \in V\}$ is a set of binary linear codes of length n . As above, if all the codes are the same, we write $C(H, A)$. Assume that all the codes A_v have the same

rate R_1 , then the rate of the code C satisfies

$$(2) \quad R(C) \geq tR_1 - (t - 1).$$

Definition 1. [2] A hypergraph H is called ε -homogeneous if for every t sets D_1, D_2, \dots, D_t with $D_i \subseteq V_i$ and $|D_i| = \alpha_i m$,

$$(3) \quad \frac{|E(D_1, D_2, \dots, D_t)|}{N} \leq \prod_{i=1}^t \alpha_i + \varepsilon \min_{1 \leq i < j \leq t} \sqrt{\alpha_i \alpha_j},$$

where $E(D_1, D_2, \dots, D_t)$ denotes the set of edges that intersect all the sets D_i .

This definition quantifies the deviation of the hypergraph H from the expected behavior of a random hypergraph. For $t = 2$ the well-known “expander mixing lemma” asserts that

$$\left| \frac{|E(D_1, D_2)|}{N} - \alpha_1 \alpha_2 \right| \leq \frac{\lambda}{n} \sqrt{\alpha_1 \alpha_2},$$

showing that regular bipartite graphs are λ/n -homogeneous. This inequality is frequently used in the analysis of bipartite-graph codes [5, 6].

Let $A[n, R_1 n, d_1 = \delta_1 n]$ be a binary linear code. The distance of a code $C(H, A)$ where H is ε -homogeneous satisfies [2]

$$(4) \quad d/N \geq \delta_1^{\frac{t}{t-1}} - c_1(\varepsilon, \delta_1, t)$$

where $c_1 \rightarrow 0$ as $\varepsilon \rightarrow 0$.

One of the main results in [2] gives an explicit construction of ε -homogeneous hypergraphs H starting with a regular graph $G(U, E)$ with degree Δ and second eigenvalue λ . Putting $V_i = U, i = 1, 2, \dots, t$ and introducing a hyperedge whenever the t vertices in the graph G are connected by a path of length $t - 1$, that paper shows that the resulting hypergraph is n -regular and ε -homogeneous with $n = \Delta^{t-1}, \varepsilon = 2(t - 1)\lambda/\Delta$. Therefore, starting with a family of Δ -regular bipartite graphs with a large spectral gap, one can construct a family of regular homogeneous hypergraphs with a small value of ε . Paper [2] has also established that random n -regular hypergraphs with high probability are $O(1/\sqrt{n})$ -homogeneous.

II. WEIGHT DISTRIBUTIONS

Below we consider ensembles of random codes on graphs and hypergraphs. In some cases the (hyper)graph will be selected randomly. In the case of bipartite graphs this is done as follows. Connect the edges $e_{(i-1)n+j}, j = 1, \dots, n$ to the vertex $v_i \in V_1, i = 1, \dots, m$. Next choose a permutation on the set E with a uniform distribution and connect the remaining half-edges to the vertices in V_2 using this permutation. Similarly, to construct an ensemble of random hypergraphs, we choose $t - 1$ permutations independently with uniform distribution and use them to connect the parts of H .

Random linear codes are selected from the standard ensemble of length- n codes defined by $n(1 - R_1) \times n$ random binary matrices whose entries are chosen independently with a uniform distribution.

We consider the following three ensembles of hypergraph codes.

Ensemble $\mathcal{C}_1(t)$. A code $C(H, \{A_1, \dots, A_t\}) \in \mathcal{C}_1(t)$ is constructed by choosing a random t -partite hypergraph H and choosing random local linear codes A_i of length n independently for each part $V_i \in V$.

Ensemble $\mathcal{C}_2(t, A)$. A code $C(H, A) \in \mathcal{C}_2$ is constructed by choosing a random t -partite hypergraph H and using the same fixed local code $A[n, R_1n, d_1]$ as a local code at every vertex.

Ensemble $\mathcal{C}_3(t, H)$. A code $C(H, \{A_v\})$ from this ensemble is formed by choosing a fixed, nonrandom hypergraph H and taking random local linear codes A_v independently for each vertex $v \in V$.

Our purpose is to compute ensemble-average asymptotic weight distributions for codes in these ensembles and to estimate the average minimum distance assuming that $m \rightarrow \infty$ and n is a constant. The case $t = 2$ corresponds to ensembles of bipartite-graph codes, some of which were studied in [1, 3, 4]. Below we will cover the remaining cases for the code ensembles $\mathcal{C}_i(t)$, $i = 1, 2, 3$ and any $t \geq 2$. Below $B_w = B_w(C)$ denotes the number of codewords of weight w and $w(x)$ denotes the Hamming weight of the vector x . Before proceeding, we note that upper bounds on the ensemble-average weight distribution in many cases also give a lower bound on the code's distance.

Lemma 1. *Suppose that for an ensemble of codes \mathcal{C} of length N there exists an $\omega_0 > 0$ such that*

$$\lim_{N \rightarrow \infty} \sum_{w \leq \omega_0 N} \mathbb{E}B_w = 0.$$

Then for large N the ensemble contains codes whose relative distance satisfies $d/N \geq \omega_0$.

The proof is almost obvious because

$$\Pr[d(C) \leq \omega_0 N] \leq \sum_{w \leq \omega_0 N} \Pr[B_w(C) \geq 1] \leq \sum_{w \leq \omega_0 N} \mathbb{E}B_w(\mathcal{C}).$$

II-A. ENSEMBLE $\mathcal{C}_1(t)$.

Theorem 2. *For $m \rightarrow \infty$ the average weight distribution over the ensemble of linear codes $\mathcal{C}_1(t)$ of length $N = mn$ and rate (2) satisfies $\mathbb{E}B_{\omega N} \leq 2^{N(F+\gamma)}$, where*

$$(5) \quad F = \begin{cases} \omega t \log_2(2^{(1-R)/t} - 1) - (t-1)h(\omega) & \text{if } 0 \leq \omega \leq 1 - 2^{(R-1)/t}, \\ h(\omega) + R - 1 & \text{if } \omega \geq 1 - 2^{(R-1)/t}, \end{cases}$$

and

$$\gamma \leq tn^{-1}(1 + \log_2 n) + (t/2N) \log_2(2N),$$

$$h(z) = -z \log_2 z - (1 - z) \log_2(1 - z).$$

Proof. The proof is an extension of the corresponding result for $t = 2$ in [1]. Let $C_i, i = 1, \dots, t$ be the set of vectors $x \in \{0, 1\}^N$ that satisfy the linear constraints of part V_i of the hypergraph H so that $C(H, A) = \cap_i C_i$. Let $P_i = \Pr[x \in C_i]$. The events $x \in C_i$ for different i are independent, and therefore

$$\Pr[x \in C] = P_i^t$$

(for any $i = 1, \dots, t$). Let $B_w(C_i)$ be the random number of vectors of weight w in the code C_i . Then

$$\mathbb{E}B_w(C) = \binom{N}{w} \Pr[x \in C] = \binom{N}{w} \prod_{i=1}^t \frac{\mathbb{E}B_w(C_i)}{\binom{N}{w}}.$$

Let $\mathcal{X}_{s,w}$ be the set of vectors of weight $w = \omega N$ whose nonzero coordinates are incident to some vertices $v_{i_1}, \dots, v_{i_s} \in V_1$, $s \geq w/n$. Let $w_j = w(x(v_{i_j}))$, $j = 1, \dots, s$ and let $\omega_j = w_j/n$. We have

$$|\mathcal{X}_{s,w}| = \binom{m}{s} \sum_{\substack{w_1, \dots, w_s \\ \sum w_j = w}} \prod_{j=1}^s \binom{n}{w_j} \leq \binom{m}{s} \sum_{\substack{w_1, \dots, w_s \\ \sum w_j = w}} 2^{n \sum_j h(\omega_j)}.$$

By convexity of the entropy function, the maximum of the last expression on $\omega_1, \dots, \omega_s$ under the constraint $n \sum_j \omega_j = \omega N$ is attained for $\omega_j = \omega m/s$, $j = 1, \dots, s$. Since the sum contains no more than n^s terms, we obtain

$$|\mathcal{X}_{s,w}| \leq 2^{mh(x) + s \log n + snh(\omega m/s)} \leq 2^{N(xh(\omega/x) + \varepsilon)}$$

where $x = s/m$ and $\varepsilon = (1 + \log n)/n$. A vector $x \in \mathcal{X}_{s,w}$ is contained in C_1 with probability $2^{sn(R_1-1)}$. Thus,

$$\mathbb{E}B_w(C_1) = |\mathcal{X}_{s,w}| 2^{sn(R_1-1)},$$

and the same expression is true for $\mathbb{E}B_w(C_i)$, $i = 2, \dots, t$. Therefore,

$$\mathbb{E}B_w(C) \leq \binom{N}{w}^{-(t-1)} 2^{tN(\max_{\omega \leq x \leq 1} (xh(\omega/x) + R_1 - 1) + \varepsilon)}.$$

Since $t(R_1 - 1) \leq R - 1$, we obtain $\mathbb{E}B_w(C) \leq 2^{N(F(\omega) + \gamma)}$, where

$$\begin{aligned} F(\omega) &\leq -(t-1)h(\omega) + t \max_{\omega \leq x \leq 1} (x(R_1 - 1 + h(\omega/x))) \\ &\leq -(t-1)h(\omega) + \max_{\omega \leq x \leq 1} (x(R - 1 + th(\omega/x))). \end{aligned}$$

The maximum on x of $x(R - 1 + th(\omega/x))$ is attained for $x = x_0 = \omega/(1 - z)$ where $t \log_2 z = R - 1$. The two cases in the theorem are obtained depending on whether $x_0 < 1$ or not. If $x_0 < 1$, we substitute x_0 in the expression for $F(\omega)$ and obtain

$$F(\omega) \leq -(t-1)h(\omega) + \omega t \log_2 \frac{z}{1-z}$$

which implies the first case of (5) on account of the identity $R - 1 + th(z) = t(1 - z) \log_2(z/(1 - z))$. If $x_0 \geq 1$, we substitute the value $x = 1$ to obtain the second case of (5). □

Corollary 3. *Let ω^* be the only nonzero root of the equation*

$$\omega \left(R - 1 - t \log_2 \left(1 - 2^{(R-1)/t} \right) \right) = (t-1)h(\omega).$$

Then the average relative distance over ensemble $\mathcal{C}_1(t)$ behaves as

$$\delta(R) \geq \begin{cases} \omega^*, & \text{if } R \leq \log_2(2(1 - \delta_{\text{GV}}(R))^t), \\ \delta_{\text{GV}}(R), & \text{if } R > \log_2(2(1 - \delta_{\text{GV}}(R))^t), \end{cases}$$

where $\delta_{\text{GV}}(x) \triangleq h^{-1}(1 - x)$.

The proof is analogous to the proof of Corollary 4 in [1] and will be omitted.

For $t = 2$ we proved in [1] that ensemble \mathcal{C}_1 contains codes that reach the GV bound if the code rate satisfies $0 \leq R \leq 0.202$. This result forms a particular case

of the above corollary. Increasing t , we find that the ensemble contains codes that reach the GV bound for the values of the rate as shown below:

$$\begin{matrix} t = 3 & 4 & 10 \\ R \leq 0.507 & 0.737 & 0.998. \end{matrix}$$

Thus already for $t = 10$ almost all codes in the ensemble \mathcal{C}_1 attain the GV bound for all but very high rates.

II-B. ENSEMBLE $\mathcal{C}_2(t, A)$. In this case the results depend on the amount of information available for the local codes. Specifically, [1] shows that for $t = 2$ the ensemble contains asymptotically good codes provided that the distance of the local code A is at least 3. In the case when the weight distribution of the code A is known, a better estimate is known from [3, 4].

Theorem 4. *Let A be a linear code of length n with weight enumerator $a(x) = \sum_{i=0}^n a_i x^i$. Let B_w the random number of codewords of weight w of a code $C(H, A) \in \mathcal{C}_2(t, A)$. Then its average value over the ensemble satisfies*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log_2 \mathbb{E} B_{\omega N} \leq - (t - 1)h(\omega) + \frac{t}{\ln 2} \left(\frac{1}{n} \ln a(e^{s^*}) - s^* \omega \right),$$

where s^* is the root of $(\ln a(e^s))'_s = n\omega$.

This theorem enables us to estimate the asymptotics of the mean relative distance $\delta = \lim_{m \rightarrow \infty} \frac{\mathbb{E}d(C)}{N}$ for the ensemble \mathcal{C}_2 . Let us consider several examples.

1. Let $t = 3$ and let A be the Hamming code of length $n = 15$ and rate $R_1 = 11/15$. Then the rate $R(\mathcal{C}_2) \geq 0.2$ and the distance $\delta = 0.2307$. The relative GV distance for this rate is $\delta_{GV}(0.2) = 0.2430$.
2. Let $t = 3$ and let A be the Hamming code of length $n = 31$. Then $R(\mathcal{C}_2) \geq 16/31$ and $\delta \approx 0.0798$. Using the same code with $t = 4$ gives $R(\mathcal{C}_2) \geq 11/31$ and $\delta \approx 0.1607$ while $\delta_{GV}(11/31) \approx 0.1646$.
3. Let $t = 3$ and let A be the 2-error-correcting primitive BCH code of length $n = 31$ and rate $R_1 = 21/31$. Then the rate $R(\mathcal{C}_2) \geq 1/31$ and the value of δ is ≈ 0.3946608 . The relative GV distance for this rate is $\delta_{GV}(1/31) \approx 0.3946614$.

Let us turn to the case when only the minimum distance d_1 of the code A is available. In [1] we addressed the case $t = 2$, proving that as long as $d_1 \geq 3$, there exists an $\epsilon > 0$ such that the ensemble-average relative distance $\delta > \epsilon$ as $m \rightarrow \infty$. In the next theorem this result is extended to arbitrary $t \geq 2$. We also prove a related result which gives an upper bound on the average weight spectrum and provides a way of estimating the value of ω_0 .

Theorem 5. (a) *Let A be the local code of length n and distance d_1 used to construct the ensemble $\mathcal{C}_2(t, A)$ of hypergraph codes. Let $x_0 = x_0(\omega)$ be the positive solution of the equation*

$$(6) \quad \omega n + \sum_{i=d_1}^n \binom{n}{i} (\omega n - i) x_0^i = 0.$$

The ensemble-average weight distribution satisfies

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} B_{\omega N} \leq \frac{t}{n} \log \frac{1 + \sum_{i=d_1}^n \binom{n}{i} x_0^i}{x_0^{\omega n}} - (t - 1)h(\omega).$$

(b) The inequality $d_1 > t/(t - 1)$ gives a sufficient condition for the ensemble to contain asymptotically good codes.

Proof. In the proof we write d instead of d_1 to refer to the distance of the code A .

(a) Let H be a random hypergraph and $C(H, A)$ be the corresponding code. Recall that $C = \cap_i C_i$, where C_i is the set of vectors that satisfy the constraints of part i of the graph. Let $U_i(w, d)$ be the set of vectors $x \in \{0, 1\}^N$ such that $w(x) = w$ and $w(x(v)) = 0$ or $w(x(v)) \geq d$ for all $v \in V_i$. Since the number of such vectors is the same for all i , below we write $|U(w, d)|$ omitting the subscript. Let us choose a vector $x \in \{0, 1\}^N$ randomly with a uniform distribution. Then

$$\Pr[x \in C_1 | w(x) = w] \leq \frac{|U(w, d)|}{\binom{N}{w}}$$

and for $i \geq 2$,

$$\Pr[x \in C_i | w(x) = w, x \in C_1] = \Pr[x \in C_i | w(x) = w].$$

Then

$$\begin{aligned} (7) \quad \mathbb{E}B_w(C) &= \binom{N}{w} \Pr[x \in C | w(x) = w] \\ &= \binom{N}{w} (\Pr[x \in C_1 | w(x) = w])^t \\ &\leq \frac{|U(w, d)|^t}{\binom{N}{w}^{t-1}}. \end{aligned}$$

Given a vector x denote by j_ℓ the number of vertices $v \in V_i$ such that $w(x(v)) = \ell$. Clearly,

$$|U(w, d)| = \sum_{\substack{j_0, j_d, j_{d+1}, \dots, j_n \\ \sum \ell j_\ell = w, j_0 + \sum_{\ell \geq d} j_\ell = m}} \binom{m}{j_0, j_d, \dots, j_n} \prod_{\ell=d}^n \binom{n}{\ell}^{j_\ell}.$$

This sum contains no more than $(m + 1)^n = O(N^n)$ terms, so for $N \rightarrow \infty$ its exponent is determined by the maximum term (which has exponential growth). We obtain

$$\begin{aligned} (8) \quad \frac{1}{N} \log |U(\omega N, d)|^t &\leq \frac{t}{n} \max_{\substack{\nu_0, \nu_d, \dots, \nu_n \\ \sum \ell \nu_\ell = \omega n, \sum \nu_\ell = 1}} \left\{ h(\nu_0, \nu_d, \nu_{d+1}, \dots, \nu_n) \right. \\ &\quad \left. + \sum_{\ell=d}^n \nu_\ell \log \binom{n}{\ell} \right\} + \frac{\log N}{m}, \end{aligned}$$

where $\nu_\ell = j_\ell/m, \ell = 0, d, d + 1, \dots, n$, and $h(\underline{x}) = -\sum_i x_i \log x_i$. The objective function is concave, so the point of extremum is found from the system of equations

$$\begin{aligned} \binom{n}{i} (1 - \sum_{\ell=d}^n \nu_\ell) &= \nu_i \mu^{-i}, \quad i = d, d + 1, \dots, n; \\ \sum_{\ell=d}^n \ell \nu_\ell &= \omega n. \end{aligned}$$

Its solution is given by

$$\nu_i = \frac{\binom{n}{i} \mu^i}{1 + \sum_{\ell=d}^n \binom{n}{\ell} \mu^\ell}, \quad i = d, d + 1, \dots, n,$$

where μ is chosen so as to satisfy the last equation of the system. Evaluating $\sum_i i\nu_i$ and writing x instead of μ , we observe that it should satisfy Eq. (6). This equation has a unique root $x_0 > 0$ because putting $x = p/(1 - p)$, we can write it as

$$\omega n \left(\frac{\Pr[X = 0]}{\Pr[X \geq d]} + 1 \right) = \mathbb{E}[X|X \geq d],$$

where X is a binomial $(p, 1 - p)$ random variable. As p changes from 0 to 1, the left-hand side of the last equation decreases monotonically from $+\infty$ to ωn while the right-hand side increases monotonically from d to n .

Finally, computing the entropy and simplifying, we obtain the estimate

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log |U(\omega N, d)|^t \leq \log \frac{1 + \sum_{i=d}^n \binom{n}{i} x_0^i}{x_0^{\omega n}}.$$

(b) The proof of the second part is analogous to the case of $t = 2$ in [1]. Let $w, 1 \leq w \leq N$ be the weight and let $p = w/d$. We have

$$\begin{aligned} |U(w, d)| &\leq \sum_{i=w/n}^p \binom{n}{i} \binom{n}{d}^i \binom{in}{(p-i)d} \\ &\leq \binom{m}{p} \binom{n}{d}^p \sum_{i=w/n}^p \binom{pn}{(p-i)d} \\ &\leq \binom{m}{p} \binom{n}{d}^p 2^{pn}. \end{aligned}$$

Then

$$\mathbb{E}B_w(C) \leq \left(\binom{m}{p} \binom{n}{d}^p 2^{pn} \right)^t \binom{N}{w}^{1-t}.$$

Using the estimates $\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k$, we compute

$$\begin{aligned} \mathbb{E}B_w(C) &\leq \left(\frac{em}{p}\right)^{pt} n^{dpt} 2^{tpn} \left(\frac{w}{N}\right)^{w(t-1)} \\ &= (sm/w)^{\frac{w}{d}(t-d(t-1))}, \end{aligned}$$

where $s = ((ed2^n)^t n^d)^{\frac{1}{t-d(t-1)}}$. Thus, for any ω satisfying $\omega < s/m$, the average number of vectors of weight ωN tends to 0 as $m \rightarrow \infty$ as long as $d(t - 1) > t$. This proves that under this condition the ensemble contains asymptotically good codes. \square

Examples. Let A be the $[7, 4, 3]$ Hamming code and let $t = 2$. Theorem 5(a) implies a lower bound $\delta \geq 0.01024$ on the average relative distance for the ensemble $\mathcal{C}_2(2, A)$. This improves upon previous results ([3, 4]; also Part (b) of this theorem) which assert only that the ensemble contains asymptotically good codes. Of course, in this case we can use the entire weight distribution of the code A to find the estimate $\delta \geq 0.186$ from Theorem 4; however, in cases when the weight distribution is difficult to find, the last theorem provides new information for the ensemble of graph codes.

Similarly, for $A[23, 12, 7]$ from Theorem 5(a) we obtain the estimate $\delta \geq 0.0234$. Again, using the entire weight distribution, it is possible to obtain a better estimate.

Part (a) of the last theorem implies the following corollary which shows what happens to the average weight spectrum of the ensemble for long local codes.

Corollary 6. *Let $d_1 = \delta_1 n$. Then*

$$\frac{1}{N} \log \mathbb{E} B_{\omega N}(C) \leq \frac{t\omega}{\delta_1} h(\delta_1) - (t-1)h(\omega) + \gamma$$

where $\gamma \leq (\log N)/m + (\log n)/n$.

Proof. In (8) let us bound above $h(\cdot)$ by $\log n$. Then

$$\frac{1}{N} \log |U(w, d_1)|^t \leq \frac{t}{n} \max_{\substack{\nu_{d_1}, \dots, \nu_n \\ \sum \nu_\ell = \omega n, \ell = d_1}} \sum_{\ell = d_1}^n \nu_\ell \log \binom{n}{\ell} + \gamma.$$

Computing the maximum amounts to solving a linear programming problem whose dual is

$$\begin{aligned} \omega n z &\rightarrow \min \\ \ell z &\geq \log \binom{n}{\ell}, \ell = d_1, d_1 + 1, \dots, n; z \geq 0. \end{aligned}$$

Its solution is given by $z^* = \omega n \max_{d_1 \leq \ell \leq n} \log \binom{n}{\ell} / \ell$. We obtain

$$\frac{1}{N} \log |U(w, d_1)|^t \leq t\omega \max_{\delta_1 \leq x \leq 1} \frac{h(x)}{x} + \gamma \leq t\omega h(\delta_1) / \delta_1 + \gamma.$$

Employing (7) now completes the proof. □

II-C. ENSEMBLE $\mathcal{C}_3(t, H)$.

Theorem 7. *Assume that H is ε -homogeneous. For $m \rightarrow \infty$ the average weight distribution over the ensemble of linear codes $\mathcal{C}_3(t, H)$ satisfies $\mathbb{E} B_{\omega N} \leq 2^{N(F+\gamma)}$ where*

$$F = \begin{cases} -x_0(1-R) + x_0^t h\left(\frac{\omega}{x_0^t}\right) & \text{if } x_0 < 1, \\ h(\omega) + R - 1 & \text{if } x_0 \geq 1, \end{cases}$$

where x_0 is the unique positive root of the equation

$$(9) \quad tx^{t-1} \log(x^t / (x^t - \omega)) = 1 - R,$$

$$\gamma = t(n + \log m) / N + \varepsilon.$$

Proof. Let $C \in \mathcal{C}_3(t, H)$ and let $x \in \{0, 1\}^N$ be a nonzero vector. Denote by B_i the set of nonzero vertices of x in the part $V_i, i = 1, \dots, t$. Let $E = |E(B_1, B_2, \dots, B_t)|$. Let $b_i = |B_i|, \beta_i = b_i/m$, then the probability that $x \in C$ equals $2^{-(1-R_1)N \sum_i \beta_i}$. Assume w.l.o.g. that $\beta_1 < \beta_2 < \dots < \beta_t$. The average number of vectors of weight $w = \omega N$ in the code C can be bounded above as

$$\mathbb{E} B_w \leq \sum_{\omega m \leq b_1, b_2, \dots, b_t \leq m} \binom{N \prod_{i=1}^t \beta_i + \varepsilon \sqrt{b_1 b_2}}{w} \prod_{i=1}^t \binom{n}{b_i} 2^{-(1-R_1)N \sum_i \beta_i}.$$

Then

$$\frac{1}{N} \log \mathbb{E} B_{\omega N} \leq \max_{\substack{\omega \leq \beta_i \leq 1 \\ \prod_i \beta_i \geq \omega}} \left\{ \prod_i \beta_i h\left(\frac{\omega}{\prod_i \beta_i}\right) - (1-R_1) \sum_i \beta_i \right\} + \gamma.$$

Let $\phi(\beta_1, \dots, \beta_t)$ be the function in the brackets in the last expression. Let us prove that ϕ is concave in the domain $\mathcal{D} = \prod_i [\omega, 1] \cap \{(\beta_1, \dots, \beta_t) : \prod_i \beta_i \geq \omega\}$. Computing its Hessian matrix, we obtain

$$H_\phi = -\log e \begin{bmatrix} \frac{s_1}{\beta_1^2} & \frac{s_2}{\beta_1\beta_2} & \cdots & \frac{s_2}{\beta_1\beta_t} \\ \frac{s_2}{\beta_2\beta_1} & \frac{s_1}{\beta_2^2} & \cdots & \frac{s_2}{\beta_2\beta_t} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{s_2}{\beta_t\beta_1} & \frac{s_2}{\beta_t\beta_2} & \cdots & \frac{s_1}{\beta_t^2} \end{bmatrix},$$

where

$$s_1 = \frac{\omega \prod_i \beta_i}{\prod_i \beta_i - \omega},$$

$$s_2 = s_1 + \prod_i \beta_i \ln \left(1 - \frac{\omega}{\prod_i \beta_i} \right).$$

The matrix H_ϕ can be written as

$$H_\phi = -\log e (s_2 z z^t + (s_1 - s_2) \text{diag}(\beta_1^{-2}, \dots, \beta_t^{-2})),$$

where $z = (1/\beta_1, \dots, 1/\beta_t)^t$ and $\text{diag}(\cdot)$ denotes a diagonal matrix. We wish to prove that H_ϕ is negative definite for $\beta_i > 0, 0 < \omega < \prod_i \beta_i$. Clearly, $s_1 > s_2$, and therefore the claim will follow if we show that $s_2 > 0$. This is indeed true because letting $Q = \prod_i \beta_i$ and using the inequality $x > \ln(1 + x)$ valid for $x > -1, x \neq 0$, we have

$$s_2 = Q \left(\frac{\omega}{Q - \omega} + \ln \frac{Q - \omega}{Q} \right) > Q \left(\ln \left(1 + \frac{\omega}{Q - \omega} \right) + \ln \frac{Q - \omega}{Q} \right) = 0.$$

We will now show that the maximum of ϕ in \mathcal{D} is attained on the line ℓ given by $\beta_1 = \beta_2 = \dots = \beta_t$. Note that \mathcal{D} is an intersection of convex domains and therefore itself convex. Moreover, the domain \mathcal{D} is also symmetric in the sense that together with any point $p = (\beta_1, \dots, \beta_t)$ it also contains all the points obtained from p by permuting its coordinates, and the value of ϕ at each of these points is the same and equal to $\phi(p)$. Because ϕ is strictly concave, for any point $p \in \mathcal{D}, p \notin \ell$ it is possible to find a point q such that $\phi(q) > \phi(p)$ (any point q on the segment between p and one of its symmetric points will do). This shows that the global maximum of ϕ in \mathcal{D} is attained on ℓ including possibly the point $\beta_1 = \dots = \beta_t = 1$. Thus, we obtain

$$\frac{1}{N} \log \mathbb{E} B_w \leq \max_{\omega^{1/t} \leq x \leq 1} \{-(1 - R)x + x^t h\left(\frac{\omega}{x^t}\right)\} + \gamma.$$

The maximum of this expression on x is attained for x determined from (9). This equation has a unique positive root x_0 because the left-hand side is a falling function of x that takes all positive values for $x \in (\omega^{1/t}, \infty)$. This concludes the proof. \square

This theorem implies the following result.

Corollary 8. *For all values of the code rate satisfying $R \geq \log(2(1 - \delta_{\text{GV}}(R))^t)$, almost all codes in the ensemble $\mathcal{C}_3(t)$ approach the GV bound as $N \rightarrow \infty$.*

Proof. From the previous theorem, the GV bound is met for the first time when x_0 becomes 1. Substituting 1 in (9), we obtain a condition on ω in the form $\omega = 1 - 2^{(R-1)/t}$. As long as this value is less than $\delta_{\text{GV}}(R)$, the ensemble-average relative distance approaches $\delta_{\text{GV}}(R)$ as $N \rightarrow \infty$. \square

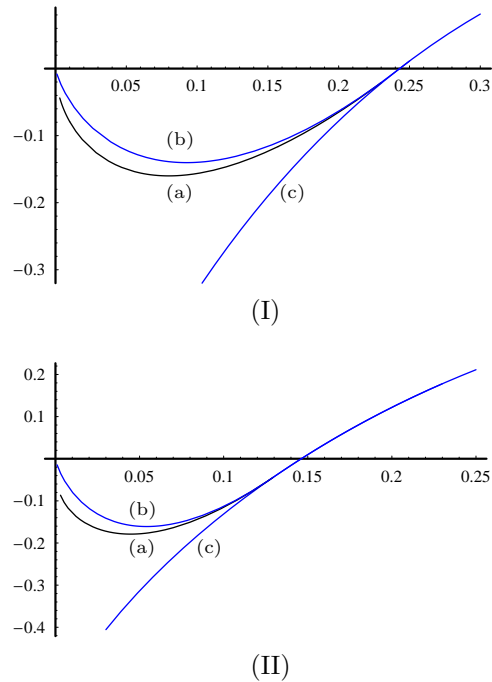


FIGURE 1. Average weight spectra for ensembles of graph codes: (I) $t = 2, R = 0.2$, (II) $t = 3, R = 0.4$; (a) ensemble $\mathcal{C}_3(2, H)$, (b) ensemble $\mathcal{C}_1(2)$, (c) ensemble of random linear codes.

We note that the condition for the attainment of the GV bound turns out to be the same as for the ensemble $\mathcal{C}_1(t)$ constructed from random graphs. The ε -homogeneity condition, and in particular, the expander mixing lemma for bipartite graphs are known to approximate the behavior of random graphs. This approximation turns out to be good enough to ensure that both ensembles contain GV codes in the same interval of code rates. Moreover, for small weights the average number of codewords for the ensemble $\mathcal{C}_3(t, H)$ turns out to be smaller than for the ensemble $\mathcal{C}_1(t)$. This is illustrated in 2 examples in Fig. 1.

For $t = 2$ codes in the ensembles \mathcal{C}_3 and \mathcal{C}_1 reach the GV bound for code rates $R \leq 0.202$. For $R > 0.202$ the codes are still asymptotically good, although slightly below the GV bound. For these values of the rate, the average relative distance for the ensemble \mathcal{C}_3 is greater than for the ensemble \mathcal{C}_1 as shown by the following numerical examples.

R	0.3	0.5	0.7	0.9
$\mathcal{C}_1(2)$	0.18558	0.09276	0.03211	0.00337
$\mathcal{C}_3(2, H)$	0.18605	0.09492	0.03242	0.00380

Similar relations between the weight spectra and distances of the ensembles $\mathcal{C}_1(t), \mathcal{C}_3(t, H)$ hold also for larger values of t .

III. DECODING

For the case of a code $C(G, A)$ on a bipartite graph G , decoding can be performed by a natural algorithm [6] that alternates between parallel decoding of local codes in

the parts V_1 and V_2 until, hopefully, it converges to a fixed point. In this algorithm, the most current value of each edge (bit) is stored at the vertex in the part decoded in the most recent iteration. However, pursuing such an edge-oriented procedure is difficult for $t > 2$. In [2] the following alternative is suggested: starting from the values of the bits stored on the edges of H decode in parallel all local codes in *all* parts of H and for each $v \in V$ form an independent decision about the codeword of A that corresponds to the edges $E(v)$. Next, the values of the bits at every vertex are updated, so that now every vertex stores an independent opinion of its bits' values. For the update, the value of the bit $x_e(v)$ is set to the majority value of the decoded versions of this bit at all the vertices $v' \in e \setminus v$, where $e \ni v$ is an edge (for this to be well-defined, the values of t are assumed to be even). The decoding then iterates, repeating this parallel decoding round until all the vertices agree on all bits.

In [2] this algorithm is shown to correct all patterns of errors provided that their proportion, as a fraction of the blocklength N , is less than

$$(10) \quad \binom{t-1}{t/2}^{-2/t} \left(\frac{\delta_1}{2}\right)^{(t+2)/t} - c_2(\varepsilon, \delta_1, t)$$

where $c_2(\varepsilon, \delta_1, t) \rightarrow 0$ as $\varepsilon \rightarrow 0$. This algorithm consists of $\log N$ iterations, each of which has serial running time linear in the blocklength N . Its analysis relies on the ε -homogeneous property of H .

For fixed values of $t > 2$, if one thinks of δ_1 as a variable quantity, then the number of correctable errors in (10) is not a constant fraction of the designed distance (4). For example, for $t = 4$, (10) gives a decoding radius equal to N times the fraction

$$\frac{\delta_1^{3/2}}{2\sqrt{6}}.$$

For small δ_1 this is a much smaller quantity than the relative designed distance $\delta_1^{4/3}$. This consideration is reinforced by the fact that advantages of hypergraph codes are most pronounced for small values of the distance δ .

Our objective is to propose an alternative decoding strategy that decodes a constant fraction of the designed distance.

For every i , we shall define a *i-th subprocedure* that decodes the subcode A on every vertex belonging to the vertex set V_i . We shall claim that if the initial number of errors is less than a bound that we shall introduce, then *for at least one i*, the *i-th* subprocedure applied to the initial error pattern produces a pattern with a smaller number of errors.

Let us now describe the decoding procedure in more detail. For every vertex v , and the associated subspace $\{0, 1\}^n$ where coordinates are indexed by the edges incident to v , we will use the following *threshold decoding* procedure T_κ of the constituent code A . This means that we introduce a number $\kappa \geq 2$, to be optimized later, and that we decode a vertex subcode *only if* its Hamming distance to the nearest codeword is less or equal to $\theta = d_1/\kappa$. If every codeword of A is at distance more than d_1/κ we leave the subvector untouched. Let $V_i = (v_{i,1}, \dots, v_{i,m})$ be the *i*th component of H . Given an N -vector $z = (z(v_{i,1}), \dots, z(v_{i,m}))$, we can decode each of the m of its subvectors with T_κ , obtaining an N -vector w . Abusing notation, we will write $w = T_\kappa(z)$. The *i-th subprocedure* now consists of applying T_κ to the component V_i .

As mentioned above, we shall claim that one among t of the i -th subprocedures lowers the total number of errors. However the decoding algorithm will not be able to discern which of the i -th subprocedures is successful. So the decoder will apply all t subprocedures in parallel to the received vector, yielding t output vectors. The next decoding iteration will have to be applied to every output of the preceding iteration, so that s iterations of the algorithm will yield t^s output vectors. We will only apply the algorithm for a constant number of iterations however, until we are guaranteed that the number of remaining error for at least one of the t^s outputs has fallen below the error-correcting capability of Bilu and Hoory's decoding procedure. We then let the latter decoder take over and decode all t^s candidates. At least one of them is guaranteed to be the closest codeword, and it can be singled out simply by computing the Hamming distance of every candidate to the initial received vector.

To give a more formal description of the algorithm, suppose that $y \in \{0, 1\}^N$ is the vector received from the channel. In each iteration the processing is done in parallel in all the vertices of H . Let $\mathcal{Y}_i^j = \{y_{i,l}^{(j)}\}$ be the set of N -vectors stored at the vertices of the component V_i before the j th iteration. By the discussion above, $|\mathcal{Y}_i^j(v)| \leq t^{j-1}$.

We begin by setting $\mathcal{Y}_i^1 = \{y\}$ for all i . Iteration $j, j = 1, 2, \dots, s$ consists of running t parallel subprocedures. The i th subprocedure applies decoder T_κ to every vector $y_{i,l}^{(j)}$ in the set \mathcal{Y}_i^j , replacing it with the vector $T_\kappa(y_{i,l}^{(j)})$, $l = 1, \dots, |\mathcal{Y}_i^j|$. The outcome of this step creates t potentially different decodings of every vector $y_{i,l}^{(j)} \in \mathcal{Y}_i^j, i = 1, \dots, t$. In the second part of the iteration we form the sets $\mathcal{Y}_i^{j+1}, i = 1, \dots, t$ by replacing each vector $y_{i,l}^{(j)} \in \mathcal{Y}_i^j$ with its decodings obtained in all the t subprocedures.

Next, we prove that one of the t subprocedures will actually diminish the number of errors. This analysis also relies on ε -homogeneity, although in a way different from [2]. Let \mathcal{E} be the set of coordinates, i.e. the set of edges, that are in error. For every $i = 1 \dots t$, let us partition the set of vertices in V_i that are incident to \mathcal{E} into three subsets, G_i, N_i, B_i . The set G_i is the subset of vertices that will be correctly decoded, N_i is the subset of vertices that are left untouched by the threshold decoder, and B_i is the set of those vertices that are wrongly decoded to a parasite codeword of A . The situation is summarized in Figure 2. From now on by the \mathcal{E} -degree of a vertex we shall mean the degree of this vertex in the subhypergraph induced by the edge set \mathcal{E} . It should be clear that every vertex of G_i has \mathcal{E} -degree not more than d_1/κ , every vertex in N_i has \mathcal{E} -degree at least d_1/κ , and every vertex in B_i has \mathcal{E} -degree at least $(\kappa - 1)d_1/\kappa$.

We use a shorthand notation $\mathcal{E}(G_i)$ to mean the set of edges that has one of its endpoints in G_i . Similarly we shall write $\mathcal{E}(N_i)$ and $\mathcal{E}(B_i)$.

Lemma 9. *If the i -th decoding subprocedure introduces more errors than it removes, then $|\mathcal{E}(G_i)| \leq |\mathcal{E}|/\kappa$. Moreover, if*

$$\mu_i = \frac{|\mathcal{E}(N_i)|}{|\mathcal{E}(N_i) \cup \mathcal{E}(B_i)|}, \quad i = 1, \dots, t$$

then

$$|\mathcal{E}(G_i)| \leq \frac{1 - \mu_i}{\kappa - \mu_i} |\mathcal{E}|.$$

Proof. The first part of the lemma follows from the second part, which is proved as follows. We bound from above $|\mathcal{E}(G_i)|$, the set of edges removed, by the set of

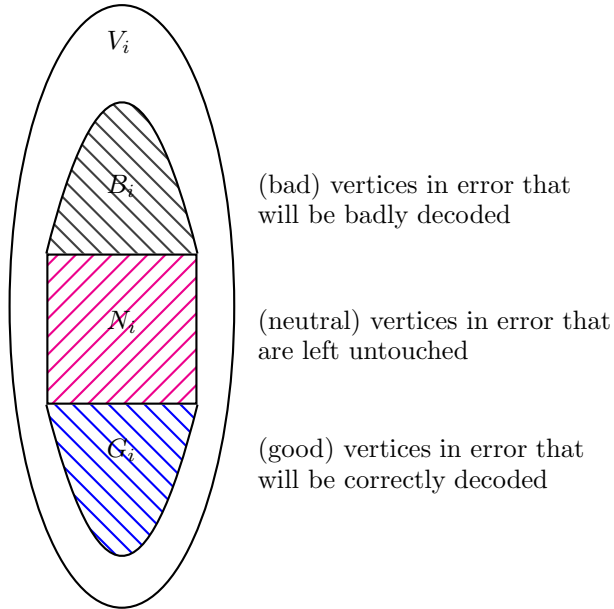


FIGURE 2. Details of the set of vertices incident to edges in error. The max \mathcal{E} -degree in G_i is less than d_1/κ , the min \mathcal{E} -degree in B_i is at least $(\kappa - 1)d_1/\kappa$, the min \mathcal{E} -degree in N_i is at least d_1/κ .

edges added, $|\mathcal{E}(B_i)|$: we get

$$\begin{aligned} |\mathcal{E}(G_i)| &\leq |B_i| \frac{d_1}{\kappa} = |B_i| d_1 \left(1 - \frac{1}{\kappa}\right) \frac{1}{\kappa - 1} \\ &\leq |\mathcal{E}(B_i)| \frac{1}{\kappa - 1}. \end{aligned}$$

The first inequality comes from the definition of κ and the threshold decoder. The second inequality states that $(1 - 1/\kappa)d_1$ is a lower bound on the minimum \mathcal{E} -degree in B_i . We now have

$$\begin{aligned} (11) \quad |\mathcal{E}| &= |\mathcal{E}(G_i)| + |\mathcal{E}(N_i)| + |\mathcal{E}(B_i)| \\ &= |\mathcal{E}(G_i)| + |\mathcal{E}(B_i)| / (1 - \mu_i) \\ &\geq \frac{\kappa - \mu_i}{1 - \mu_i} |\mathcal{E}(G_i)| \end{aligned}$$

which proves the lemma. □

Theorem 10. For any $\alpha > 0$, if the number of errors eN is such that

$$(12) \quad e \leq (1 - \alpha) \frac{\delta_1^{t/(t-1)}}{(t + 1)^{(t+1)/(t-1)}}$$

they can be corrected in time $O(N \log N)$.

Proof. The theorem will follow if we show that at least one subprocedure reduces the error count by a constant fraction. Indeed, in this case a constant number of rounds of the above algorithm will reduce the error count to any positive proportion of the

designed distance whereupon the remaining errors will be removed in $O(\log N)$ steps of Bilu-Hoory's algorithm.

Assume toward a contradiction that *all* the i -th decoding subprocedures, $i = 1 \dots t$, introduce more errors than they remove. Let us introduce the following notation: $|\mathcal{E}| = eN$, $S_i = B_i \cup N_i$, $|S_i| = \sigma_i m$. Note that since the minimum \mathcal{E} -degree in S_i is at least d_1/κ , we have

$$(13) \quad \sigma_i \leq \kappa e / \delta_1.$$

Consider the subset of edges obtained from \mathcal{E} by removing all edges incident to "good" vertices G_i for all i . We are left with a subhypergraph $H_{\mathcal{E}}$ with vertex set S_i , $i = 1 \dots t$. Use Lemma 9 (the first part) for all i to argue that the total fraction of edges in $H_{\mathcal{E}}$ is at least $e(1 - t/\kappa)$. Applying the ε -homogeneous property (3) gives

$$e \left(1 - \frac{t}{\kappa}\right) \leq \sigma_1 \cdots \sigma_t + \varepsilon \min_{1 \leq i < j \leq t} (\sigma_i \sigma_j)^{1/2}.$$

Applying (13) we obtain

$$e \left(1 - \frac{t}{\kappa}\right) \leq \left(\frac{\kappa e}{\delta_1}\right)^t + \varepsilon \frac{\kappa e}{\delta_1}.$$

This inequality does not hold (and therefore our assumption is false) if

$$(14) \quad e < \delta_1^{t/(t-1)} \left(\frac{1 - t/\kappa - \varepsilon \kappa / \delta_1}{\kappa^t}\right)^{1/(t-1)}.$$

Taking $\kappa = t + 1$, rewrite the expression in the brackets on the right as

$$\left(\frac{1}{t+1}\right)^{(t+1)/(t-1)} \left(1 - \frac{(t+1)^2 \varepsilon}{\delta_1}\right)^{\frac{1}{t+1}}.$$

By taking sufficiently large n it is possible to make ε small enough so that for any given $\alpha' > 0$ there holds

$$(1 - (t+1)^2 \varepsilon / \delta_1)^{1/t+1} > 1 - \alpha'.$$

This means that (14) is satisfied for all

$$e < (1 - \alpha') \frac{\delta_1^{t/(t-1)}}{(t+1)^{(t+1)/(t-1)}}.$$

Finally, choosing $\alpha' < \alpha$ guarantees that at least one subprocedure reduces the error count by a constant fraction. \square

We see that the upper bound on the number of correctable errors given by Theorem 10 is a constant proportion γ of the designed distance δN (4), where $\gamma = 1/(t+1)^{(t+1)/(t-1)}$. For example, for $t = 3, 4$ we get $\gamma = 1/16$ and $1/14.2$, respectively.

The next theorem provides a better estimate of γ by refining the above analysis. The way this is done is to rely on the full power of Lemma 9 instead of its first part as above.

Theorem 11. *For any $\alpha > 0$, if the number of errors eN is such that*

$$e \leq (1 - \alpha) \delta_1^{t/(t-1)} \max_{\kappa \geq 2} \min_{0 < \mu < 1} f(\mu, \kappa)$$

with

$$f(\mu, \kappa) = \frac{[1 - t(1 - \mu)/(\kappa - \mu)]^{1/(t-1)}}{\kappa^{t/(t-1)}[\mu + (1 - \mu)/(\kappa - 1)]^{t/(t-1)},}$$

then they can be corrected in time $O(N \log N)$.

Proof. We proceed as in the previous theorem, assuming toward a contradiction that each subprocedure increases the error count. Using the definition of μ_i given above,

$$|\mathcal{E}(S_i)| = \frac{|\mathcal{E}(B_i)|}{1 - \mu_i} = \frac{|\mathcal{E}(N_i)|}{\mu_i}.$$

Recall that the subhypergraph $H_{\mathcal{E}}$ is formed of the edges all of whose vertices are in S_i . To count the total fraction of edges $\beta(H_{\mathcal{E}})$ in the subhypergraph $H_{\mathcal{E}}$ we employ Lemma 9:

$$\beta(H_{\mathcal{E}}) \geq e \left(1 - \sum_{i=1}^t \frac{1 - \mu_i}{\kappa - \mu_i} \right).$$

The \mathcal{E} -degree of a vertex in S_i (resp., B_i) is at least d_1/κ (resp., $d_1(\kappa - 1)/\kappa$). Hence

$$\begin{aligned} |S_i| = |B_i| + |N_i| &\leq \mathcal{E}(N_i) \frac{\kappa}{d_1} + \mathcal{E}(B_i) \frac{\kappa(1 - \mu_i)}{d_1(\kappa - 1)} \\ &\leq \frac{\kappa e}{d_1} \left(\frac{1 - \mu_i}{\kappa - 1} + \mu_i \right) N. \end{aligned}$$

Using the last two inequalities in (3), we obtain

$$e \left(1 - \sum_{i=1}^t \frac{1 - \mu_i}{\kappa - \mu_i} \right) \leq \left(\frac{\kappa e}{\delta_1} \right)^t \prod_{i=1}^t \left(\frac{1 - \mu_i}{\kappa - 1} + \mu_i \right) + \varepsilon \frac{\kappa e}{\delta_1}.$$

To contradict this, let

$$e < \left(\frac{\delta_1}{\kappa} \right)^{t/(t-1)} \left\{ \frac{1 - \sum_{i=1}^t \frac{1 - \mu_i}{\kappa - \mu_i} - \varepsilon \kappa / \delta_1}{\prod_{i=1}^t \left(\frac{1 - \mu_i}{\kappa - 1} + \mu_i \right)} \right\}^{1/(t-1)}.$$

We again bound the terms that involve ε from below by a multiplicative term $1 - \alpha'$. Optimizing on all possible values of μ_i gives $\mu_i = \mu$ for all $i = 1 \dots t$, whereupon the expression on the right can be replaced by $(1 - \alpha)\delta_1^{t/(t-1)} f(\mu, \kappa)$. The proof is thus complete. \square

Numerically, the first values of the decoding radius ρ given by Theorem 11 are

$$\rho \geq \begin{cases} \frac{\delta_1^{3/2}}{5.94} & \text{for } t = 3, \\ \frac{\delta_1^{4/3}}{6.46} & \text{for } t = 4, \end{cases}$$

attained for κ satisfying $(\kappa - 1)^{-t} = 1 - t/\kappa$ and $\mu = 0$ or 1 .

Can one obtain better bounds for the decoding radius? In principle, it is possible to obtain further improvements by introducing *multiple* thresholds instead of the single decoding threshold $\theta = d_1/\kappa$, and approach $\rho = \delta/2$ by increasing their number. However we shall only be able to claim that using one of the multiple thresholds reduces the number of errors for one of the subprocedures, but we shall not be able to discern which decoding threshold achieves that. This will result in yet another layer of parallelism, further increasing the value of the constant in the decoding complexity. We will not pursue this line of research further here. A

remaining challenge is to decode up to half the designed distance with an iterative decoding procedure of reasonable complexity.

REFERENCES

- [1] A. Barg and G. Zémor, *Distance properties of expander codes*, IEEE Trans. Inform. Theory, **52** (2006), 78–90.
- [2] Y. Bilu and S. Hoory, *On codes from hypergraphs*, European. J. Combin., **25** (2004), 339–354.
- [3] J. Boutros, O. Pothier and G. Zémor, *Generalized low density (Tanner) codes*, in “Proc. IEEE ICC,” Vancouver, Canada, **1** (1999), 441–445.
- [4] M. Lentmaier and K. Sh. Zigangirov, *On generalized low-density parity-check codes based on Hamming component codes*, IEEE Communications Letters, **3** (1999), 248–260.
- [5] M. Sipser and D. A. Spielman, *Expander codes*, IEEE Trans. Inform. Theory, **42** (1996), 1710–1722.
- [6] G. Zémor, *On expander codes*, IEEE Trans. Inform. Theory, **47** (2001), 835–837.

Received July 2008; revised October 2008.

E-mail address: abarg@umd.edu

E-mail address: arya@umd.edu

E-mail address: gilles.zemor@math.u-bordeaux1.fr