

On the Number of Errors Correctable with Codes on Graphs

Alexander Barg and Arya Mazumdar

Abstract—We study ensembles of codes on graphs (generalized low-density parity-check, or LDPC codes) constructed from random graphs and fixed local constrained codes, and their extension to codes on hypergraphs. It is known that the average minimum distance of codes in these ensembles grows linearly with the code length. We show that these codes can correct a linearly growing number of errors under simple iterative decoding algorithms. In particular, we show that this property extends to codes constructed by parallel concatenation of Hamming codes and other codes with small minimum distance. Previously known results that proved this property for graph codes relied on graph expansion and required the choice of local codes with large distance relative to their length.

Index Terms—Graph codes, hypergraph codes, iterative decoding, parallel concatenation of codes.

I. INTRODUCTION

Considerable attention in recent years was devoted to the study of error correction with codes on graphs. In this paper we are interested in estimating the number of errors correctable with codes on graphs constructed as generalizations of LDPC codes. LDPC codes are constructed on a bipartite graph $G(V, E)$, $V = V_1 \cup V_2$ by associating code's coordinates with the vertices in one part of G , replicating the values of each vertex on the edges incident to it, and imposing a parity-check constraint at each vertex of the other part of G . The generalization that we have in mind is concerned with replacing the repetition and single-parity-check codes as local codes at the graph's vertices with other error-correcting codes.

Error correction with codes on graphs has been studied along two lines, namely, by computing the average number of errors correctable with some decoding algorithm by codes from a certain random ensemble of graph codes, or by examining explicit code families whose construction involves graphs with a large spectral gap. The first direction originates in the works of Gallager [7] and Zyablov and Pinsker [15] who showed that random LDPC codes of growing length can correct a nonvanishing fraction of errors. Recently the decoding algorithm of [15] was studied by Burshtein [6] who derived an improved estimate of the number of correctable

errors compared to [15] and by Zyablov et al. [14] who provided estimates of the number of errors under the assumption of local single error-correcting (Hamming) codes. The second line of work, initiated in Tanner's paper [12] and in Sipser and Spielman's [10], pursues estimates of error correction with codes on regular graphs with a small second eigenvalue and ensuing expansion properties. Presently it is known that such codes under iterative decoding can correct the number of errors equal to a half of the designed distance of graph codes [2]. This estimate fits in a series of analogous results for various "concatenated" coding schemes and has prompted a view of graph codes as parallel concatenations of the local codes [2]. However, this result relies on certain restrictive assumptions discussed below.

An extension of Tanner's construction from graphs to hypergraphs was proposed by Bilu and Hoory [4] who showed that such codes (for high code rates) can have minimum distance greater than the best known bipartite-graph constructions. Interestingly, the codes considered in [4] are a direct extension of a construction in [7] in the same way as Tanner's graph codes extend LDPC codes.

As is well known, graphs with high expansion and random graphs share many properties that can be used to prove estimates of error correction. This similarity in the coding theory context was emphasized in our recent work [1] where we showed that ensembles of codes on random graphs and explicit expander-like constructions share many common features such as properties of the minimum distance and weight distribution.

Regarding the proportion of errors corrected by graph codes under iterative decoding, we note one difference between (generalized) LDPC codes on random graphs and explicit constructions based on the graph spectrum. The explicit constructions based on regular graphs depend on the difference between the largest and the second largest eigenvalue of the graph (the "spectral gap"). For this reason, one is forced to rely on local codes with rather large minimum distance d_0 , for instance, d_0 greater than the square root of the degree n of the graph. Even though in the construction of [10] and later works n is kept constant, this effectively rules out of consideration local codes with small minimum distance such as the Hamming codes and the like. The square root restriction is implied by the spectral gap of regular bipartite graphs, and is the best possible owing to the Alon-Boppana bound for graph spectra [9]. The purpose of the present work is to lift this limitation on the distance d_0 by switching from graphs with a large spectral gap to random graphs.

In this paper we obtain new estimates of the number of correctable errors for random ensembles of bipartite-graph and

The results of this paper were presented in part at the 2009 IEEE International Symposium on Information Theory, Seoul, Korea, July 2009.

Alexander Barg is with the Department of Electrical and Computer Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742 and Institute for Problems of Information Transmission, Moscow, Russia (e-mail: abarg@umd.edu).

Arya Mazumdar is with the Department of Electrical and Computer Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742 (e-mail: arya@umd.edu).

Research supported in part by NSF grants CCF0830699, CCF0635271, DMS0807411, CCF0916919.

hypergraph codes under iterative decoding. The first part of the paper is devoted to codes on regular bipartite graphs. To construct long graph codes, we assume that the degree of the graph is fixed and the number of vertices in both parts approaches infinity. Assuming that local constraint codes are used to correct 2 or more errors, we show that almost all codes in the ensemble of graph codes are capable of correcting all error patterns of weight that forms a constant fraction of the code length. This is a much less restrictive assumption on the local codes than the one taken in earlier works on decoding of graph codes [2], [13]. The proof of this result employs some ideas of [1] introduced there for the analysis of the weight distribution of graph codes.

We then observe that if the degree of the graph is allowed to increase then graph codes with local codes of constant distance do not correct a linearly growing number of errors under the proposed iterative decoding. This motivates us to study graph codes with long local codes correcting a growing number of errors that forms a fixed proportion of the degree. The results obtained in this case parallel earlier theorems for product codes and graph codes based on the spectral gap.

In the second part of the paper we establish similar results for codes on hypergraphs, showing that a constant proportion of errors is corrected by an iterative decoding algorithm that combines some ideas of [1] with the results proved for codes on bipartite graphs in the first part of the paper. Constructing the code ensemble based on regular hypergraphs of a fixed degree, we show that they contain codes capable of correcting a constant proportion of errors. The proof involves no assumptions on the distance of the local codes; in particular, we show that networks of Hamming codes correct a fixed proportion of errors under iterative decoding. This fact was previously proved by Tanner [12] under the assumption that the underlying graph is a tree. This assumption is not needed in our results. As in the case of the graph ensemble, we also perform the analysis of the decoding algorithm for the case of growing degree, finding the proportion of errors correctable with hypergraph codes based on long local codes.

This paper is dedicated to the memory of Ralf Koetter. The first-named author discussed the problem of estimating the performance of codes on graphs with Ralf in the beginning of 2004. Ralf's idea at that point was to investigate the error correcting capability of codes defined on some distance-regular graph, with local constraints imposed at the vertices of the graph. Presently it is understood that the setting most amenable to analysis is that of codes defined on a regular bipartite graph. Ralf himself made an initial attempt to analyze such codes in a joint paper with Xiangyu Tang [11]. The emphasis in [11] is on the estimation of the largest channel error rate tolerated by graph codes under such decoding. In the present paper, similarly to [10], [13] and later works, the local codes are decoded up to their correction radius guaranteed by the minimum distance.

II. CODE ENSEMBLES

An $[N, K]$ binary linear code is a linear subspace of $\{0, 1\}^N$ of dimension K . To construct an $[N, RN]$ binary linear graph

code C , consider an n -regular bipartite graph $G(V = V_1 \cup V_2, E)$, where the set of vertices V consists of two disjoint parts V_1, V_2 of size m each, all the edges are of the form $(u, v), u \in V_1, v \in V_2$, and the degree of every vertex v in V is n . Let $A[n, R_0n, d_0]$ be a linear binary code of length n called the local code below. We identify the coordinates of C with the set E and for a vertex $v \in V$ denote by $\mathbf{x}(v) \in \{0, 1\}^n$ the projection of a vector $\mathbf{x} \in \{0, 1\}^N, N = nm$, on the edges incident to v . A graph code $C(G)$ is defined as follows:

$$C = \{\mathbf{x} \in \{0, 1\}^N : \forall v \in V \mathbf{x}(v) \in A\}. \quad (1)$$

The ensemble of codes $\mathcal{G}(A, m)$ is constructed by associating a code $C(G)$ with a graph G sampled from the set of graphs defined by a random permutation on N elements which establishes how the edges originating in V_1 are connected to the vertices in V_2 .

Generalizing this construction, consider an l -partite n -regular uniform hypergraph $H = (V, E)$ i.e., a finite set $V = V_1 \cup \dots \cup V_l$, where $|V_1| = \dots = |V_l| = m$, and a collection E of l -subsets (hyperedges) of V such that every $e \in E$ intersects each $V_i, 1 \leq i \leq l$ by exactly one element and each vertex $v \in V$ appears in exactly n different subsets of E . Aiming at constructing an $[N, RN]$ binary linear code C by imposing local constraints at the vertices, we again identify the coordinates of C with the (hyper)edges of H . By definition, the code C is formed of the vectors \mathbf{x} that satisfy condition (1) for every vertex in V . The ensemble of codes $\mathcal{H}(A, l, m)$ in this case is constructed by sampling a random hypergraph from the set of hypergraphs defined by $l - 1$ independent random permutations on N elements. For $i = 1, 2, \dots, l - 1$, the i th permutation accounts for the placement of edges between parts V_1 and V_{i+1} of H . Of course, $\mathcal{H}(A, l, m)$ becomes $\mathcal{G}(A, m)$ for $l = 2$.

The following is known about the parameters of codes in the graph and hypergraph ensembles considered here. It is easy to see that the rate R of the codes $C \in \mathcal{H}(A, l, m)$ satisfies $R \geq lR_0 - (l - 1), l = 2, 3, \dots$. Denote by $d(\mathcal{H}) = d(\mathcal{H}(A, l, m))$ the average value of the minimum distance of codes in the hypergraph ensemble and let

$$\delta = \delta(\mathcal{H}) \triangleq \liminf_{N \rightarrow \infty} \frac{d(\mathcal{H})}{N}. \quad (2)$$

A way to bound the value of δ below using the distribution of distances in the local code A was suggested in [5], [8]. More explicit results in this direction were obtained in [1], [3]. In particular, [1] shows that $\delta(\mathcal{H}) > 0$ if the local distance d_0 satisfies $d_0 > l/(l - 1)$. For the bipartite graph ensemble $\mathcal{G}(A, m)$ (i.e., for $l = 2$) this implies that $d_0 \geq 3$, i.e., with high probability codes in ensemble are asymptotically good (have nonvanishing rate and relative distance) when the local codes correct one or more errors. For hypergraphs with $l = 3$ or more parts any local codes (without repeated vectors) account for an asymptotically good ensemble. An explicit lower bound for $\delta(\mathcal{H})$ that depends only on l and d_0 is given by [1], see Theorem 5.5 below. For the case when n is large and $d_0 = \delta_0 n$, a lower estimate of $\delta(\mathcal{H})$ is given by the

solution for x of the following equation [1, Cor. 6]:

$$\frac{h(x)}{x} = \frac{l}{l-1} \frac{h(\delta_0)}{\delta_0}. \quad (3)$$

Finally, if the local codes are chosen randomly as opposed to a fixed code A used at every vertex of H , then the codes in the (hyper)graph ensemble match the best known linear codes, i.e., reach the asymptotic Gilbert-Varshamov bound on the minimum distance [1].

Remarks. 1. An equivalent description of the bipartite code ensemble is obtained by considering an edge-vertex incidence graph of the graph $G(V, E)$, i.e., a bipartite graph $D = (D_1 \cup D_2, \bar{E})$ where $D_1 = E, D_2 = V_1 \cup V_2$, each vertex in D_1 is connected to one vertex in V_1 and to one vertex in V_2 , and there are no other edges in \bar{E} . Thus, for all $v \in D_1$, $\deg(v) = 2$ and for all $v \in D_2$, $\deg(v) = n$. The local code constraints are imposed on the vertices in D_2 . By increasing the number of parts in D_2 from two to l , we then obtain the hypergraph codes defined above. This gives an alternate description of the hypergraph code presented in Fig. 1.

The ensemble of hypergraph codes with local constraints given by single parity-check codes was introduced by Gallager [7, p.12]. The proportion of errors correctable with these codes using the so-called ‘‘flipping’’ algorithm was estimated in [15]. Several generalizations of this ensemble were studied in [1], [4].

2. The derivations of this paper are not specific to binary codes: any local linear codes such as Reed-Solomon codes can be used in the construction with no conceptual changes to the analysis and the conclusions.

III. DECODING ALGORITHMS FOR GRAPH (GENERALIZED LDPC) CODES

Even though the ensemble $\mathcal{G}(A, m)$ forms a particular case of the ensemble $\mathcal{H}(A, l, m)$, in our analysis we employ different decoding algorithms for the cases $l = 2$ and $l \geq 3$. The reason for this is that edge-oriented procedures commonly used for bipartite-graph codes do not generalize well to hypergraphs.

A. Decoding for the ensemble $\mathcal{G}(A, m)$. In our estimates of the number of correctable errors for the ensemble \mathcal{G} we rely upon the algorithm of [13] which iterates between decoding all the vertices in parts V_1 and V_2 in parallel using some decoding algorithm of the code A . Let $C \in \mathcal{G}(A, m)$ be a code. For the ease of analysis we assume that the local codes are decoded to correct up to t errors, where $t \geq 0$ is an integer that satisfies $2t + 1 \leq d_0$ and d_0 is the distance of the code A . Formally, define a mapping $\psi_{A,t} : \{0, 1\}^n \rightarrow \{0, 1\}^n$ such that $\psi_{A,t}(z) = x \in A$ if x is the unique codeword that satisfies $d(z, x) \leq t$ and $\psi_{A,t}(z) = z$ otherwise. Let $\mathbf{y}^{(i)}$ be the estimate of the transmitted vector before the i th iteration, $i \geq 1$, where $\mathbf{y} = \mathbf{y}^{(1)}$ is the received vector. The next steps are repeated for a certain number of iterations.

Algorithm I ($\mathbf{y}^{(1)}$)

- i odd: for all $v \in V_1$ put $\mathbf{y}^{(i+1)}(v) = \psi_{A,t}(\mathbf{y}^{(i)}(v))$;
- i even: for all $v \in V_2$ put $\mathbf{y}^{(i+1)}(v) = \psi_{A,t}(\mathbf{y}^{(i)}(v))$.

B. Decoding for the ensemble $\mathcal{H}(A, l, m)$. For the hypergraph ensemble $\mathcal{H}(A, l, m)$ we use the decoding algorithm

proposed in [1]. It proves to be the best choice in terms of the number of correctable errors among several possible algorithms for these codes such as the one in [4] and procedures analogous to Algorithm I above.

Let $C \in \mathcal{H}(A, l, m)$ be a code and let $H(V, E), V = V_1 \cup \dots \cup V_l$ be the graph associated with it. For every $i = 1, 2, \dots, l$ we will define an i -th subprocedure that decodes the local code A on every vertex in the part V_i . Suppose that a vector $\mathbf{u} \in \{0, 1\}^N$ is associated with the edges $e \in E$. Let $v_{i,1}, \dots, v_{i,m}$ be the vertices in the part V_i of H and let $\mathbf{u}_{i,1} = \mathbf{u}(v_{i,1}), \dots, \mathbf{u}_{i,m} = \mathbf{u}(v_{i,m})$ be the m subvectors obtained from \mathbf{u} upon permuting its coordinates according to the order of edges in V_i and projecting it on the vertices $v_{i,1}, \dots, v_{i,m}$. In other words, the vector $(\mathbf{u}_{i,1}, \dots, \mathbf{u}_{i,m})$ is obtained from \mathbf{u} using the permutation that establishes edge connections between parts V_1 and V_i . The i th subprocedure replaces the vector $(\mathbf{u}_{i,1}, \dots, \mathbf{u}_{i,m})$ with the vector $(\psi_{A,t}(\mathbf{u}_{i,1}), \dots, \psi_{A,t}(\mathbf{u}_{i,m}))$.

The algorithm proceeds in iterations. Let $\mathbf{y} \in \{0, 1\}^N$ be the received vector. Denote by $Y^{(j)}$ the set of estimates of the transmitted codeword (i.e., the set of N -vectors) stored at the vertices of H before the j th iteration $j = 1, 2, \dots$. After each iteration, this set is formed as the union of the vectors obtained upon decoding of the vertices in the i th part, $i = 1, \dots, l$. Decoding begins with setting $Y^{(1)} = \{\mathbf{y}\}$. After the first iteration we obtain l potentially different vectors (one for each subprocedure) which form the current estimates of the transmitted vector. These vectors form the sets $Y_i^{(2)}, i = 1, \dots, l$. In the next iteration each subprocedure will have to be applied to each of the l outcomes of the preceding iteration. Proceeding in this way, we observe that $|Y_i^{(j)}| \leq l^{j-1}$.

This algorithm, called Algorithm II below, will only be applied for a constant number s of iterations until we can guarantee that at least one subprocedure has reduced the number of errors to a specified proportion, say from $\gamma_0 N$ to some $\gamma_1 N, \gamma_1 < \gamma_0$. We then let another algorithm take over and decode all the l^s candidates. Any low-complexity decoder of graph codes that removes an arbitrarily small positive fraction of errors γ_1 will do at this stage. This is because taking the proportion of errors from γ_0 to $\gamma_1 > 0$ can be accomplished in a constant number s of steps, so the number of candidates that this decoder has to handle is at most l^s and does not depend on N .

For the case of local codes correcting $t \geq 2$ errors we let this algorithm to be the decoding algorithm of bipartite-graph codes (Algorithm I), making sure that γ_1 is below the proportion of errors that are necessarily corrected by this algorithm for the ensemble $\mathcal{G}(A, m)$. This is possible because, leaving any two parts of the original hypergraph H to form a bipartite graph G , we obtain a random code from the ensemble $\mathcal{G}(A, m)$ which with high probability (over the ensemble) will remove all the residual errors from at least one candidate estimate. For $t = 1$ this approach fails for the reasons discussed in the next section, so we resort to a procedure in [14] that corrects a small linear fraction of errors for single-error-correcting Hamming codes.

Upon performing the described procedure we obtain a list of at most l^s candidate codewords of the code C . The final

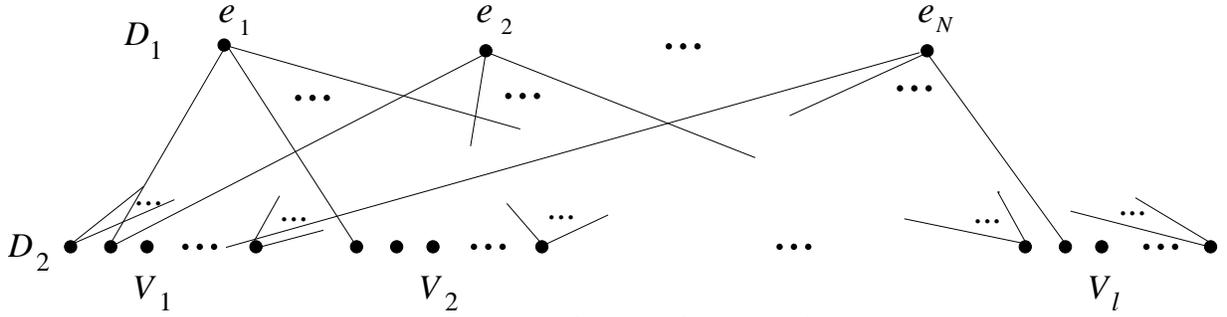


Fig. 1. Alternate construction of the hypergraph code: The set $D_1 = \{e_1, \dots, e_N\}$, where $\deg(e_i) = l$ for all i , represents the coordinates of the code (hyperedges of H); the sets V_1, \dots, V_l , where $|V_j| = m$ for all j , represent the vertices of the hypergraph H . Each vertex $v_{i,j}$, $1 \leq i \leq l, 1 \leq j \leq m$ carries a codeword of the local code A of length n .

decoding result is found by choosing the codeword from this list closest to \mathbf{y} by the Hamming distance.

Though the last step of the decoding algorithm described is different from [1], the main idea is similar to that paper, so we refer to it for a more detailed description and a discussion of the algorithm.

IV. NUMBER OF CORRECTABLE ERRORS FOR THE ENSEMBLE $\mathcal{G}(A, m)$

Let $C \in \mathcal{G}(A, m)$ be a code and let $G(V, E)$ be the graph associated with it. For a given subset of vertices $S \subset V_i, i = 1, 2$ and a vertex v denote by $\deg_S(v)$ the number of edges between v and S . Let $T_r(S) = \{v \in V : \deg_S(v) \geq r + 1\}$, where $r \in \{0, \dots, n - 1\}$ is an integer.

Below $h(z)$ denotes the entropy of the probability vector $z \in \mathbb{R}^{n+1}$. In the particular case of $n = 1$ we write $h(z)$ instead of $h(z, 1 - z)$.

Let $t \geq 0$ be any integer such that $2t + 1 \leq d_0$. The calculation in this section is based on the following simple observation.

Proposition 4.1: Suppose that for all $S \subset V_i, i = 1, 2, |S| \leq \sigma m, \sigma \in (0, 1)$, there exists $\epsilon > 0$ such that $|T_t(S)| \leq |S| - \epsilon m$. Then any $\sigma t m = \sigma t(N/n)$ errors will be corrected by Algorithm I in $O(\log m)$ iterations.

Proof: Suppose that no more than $\sigma t m$ errors occurred in the channel. Let S_i be the set of vertices that are decoded incorrectly in iteration i of Algorithm I. The assumption of the proposition implies that $|S_{i+1}| \leq |S_i|(1 - \epsilon/\sigma)$, so $O(\log m)$ iterations suffice to remove all the errors. ■

Define

$$F_{n,t}(\sigma) = h(\sigma) - \sigma n \log x + \sigma \log \sum_{i=t+1}^n \binom{n}{i} x^i + (1 - \sigma) \log \sum_{i=0}^t \binom{n}{i} x^i, \quad (4)$$

where $x > 0$ is found from the equation

$$\sum_{i=0}^t \sum_{j=t+1}^n \binom{n}{i} \binom{n}{j} (\sigma(n-j) - i(1-\sigma)) x^{i+j-t-1} = 0. \quad (5)$$

Let $\mathcal{Z}_n = \{z \in [0, 1]^{n+1} : \sum_{i=0}^n z_i = 1\}$ be the $(n + 1)$ -dimensional probability simplex.

The main result of this section is given by the next theorem.

Theorem 4.2: Let $A[n, R_0 n, d_0]$ be the local code, let $m \rightarrow \infty$, and let $2 \leq t < d_0/2$. All codes in the ensemble $\mathcal{G}(A, m)$ except for an exponentially small (in N) proportion of them correct any combination of errors of weight $\sigma t m$ in $O(\log m)$ iterations of Algorithm I, where $0 < \sigma < \sigma_0$ and σ_0 is the smallest positive root of the equation

$$F_{n,t}(\sigma) = (n - 1)h(\sigma).$$

Remark. The case of local codes with $t = 1$ is excluded from this theorem because G with high probability contains a large number of 4-cycles, which means that correcting single error at every vertex does not ensure overall convergence of the decoding. Indeed, if two vertices are affected by two errors each, and the corresponding 4 edges form a cycle, then the decoder will loop indefinitely without approaching the correct decision. The theorem is still valid in this case, but gives $\sigma_0 = 0$.

Proof: We need to verify the assumption of Proposition 4.1. Let $S \subset V_1, |S| = \sigma m$ and let $m_i = |\{v \in V_2 : \deg_S(v) = i\}|, i = 1, \dots, n$. Clearly,

$$\sum_{i=1}^n m_i \leq m, \quad \sum_{i=t+1}^n m_i = |T_t(S)|, \quad \sum_{i=1}^n i m_i = |S|n.$$

Let us compute the probability (over the choice of G) that $|T_t(S)| \geq (\sigma - \epsilon)m$. Let $\boldsymbol{\mu} = (m_1, \dots, m_n)$ be a vector with nonnegative integer components, let

$$M_\epsilon(t, \sigma) = \{\boldsymbol{\mu} : \sum_{i=1}^n m_i \leq m, \sum_{i=1}^n i m_i = \sigma N, \sum_{i=t+1}^n m_i \geq (\sigma - \epsilon)m\},$$

and let $\binom{m}{\boldsymbol{\mu}}$ denote the number of choices of subsets of size m_1, \dots, m_n out of a set of size m . We have

$$P(|T_t(S)| \geq (\sigma - \epsilon)m) = \frac{1}{\binom{N}{\sigma N}} \sum_{\boldsymbol{\mu} \in M_\epsilon(t, \sigma)} \binom{m}{\boldsymbol{\mu}} \prod_{i=1}^n \binom{n}{i}^{m_i}. \quad (6)$$

Let $\mathcal{L}_1(s)$ denote the event that V_1 contains a subset $S, |S| = s$ for which $|T_t(S)| \geq |S| - \epsilon m$. We have

$$P(\mathcal{L}_1(\sigma m)) \leq \binom{m}{\sigma m} P(|T_t(S)| \geq |S| - \epsilon m)$$

and

$$P\left(\bigcup_{i=1}^{\sigma m} \mathcal{L}_1(i)\right) \leq mP(\mathcal{L}_1(\sigma m)).$$

Denote by $\mathcal{L}_2(\sigma)$ an analogous event with respect to V_2 . Then

$$P\left(\bigcup_{i=1}^{\sigma m} (\mathcal{L}_1(i) \cup \mathcal{L}_2(i))\right) \leq \frac{2m\binom{m}{\sigma m}}{\binom{N}{\sigma N}} \sum_{\mu \in \mathcal{M}_\epsilon(t, \sigma)} \binom{m}{\mu} \prod_{i=1}^n \binom{n}{i}^{m_i}. \quad (7)$$

Letting L to be the logarithm of the left-hand side divided by m and omitting $o_m(1)$ terms, we obtain the estimate $L \leq n^{-1}\bar{F}_{n,t}(\sigma)$, where

$$\bar{F}_{n,t}(\sigma) = -(n-1)h(\sigma) + \max_{z \in \mathcal{M}'_\epsilon(t, \sigma)} \left(h(z) + \sum_{i=1}^n z_i \log \binom{n}{i} \right),$$

where

$$\mathcal{M}'_\epsilon(t, \sigma) = \left\{ z \in \mathcal{Z}_n : \sum_{i=1}^n iz_i = \sigma n, \sum_{i=t+1}^n z_i \geq \sigma - \epsilon \right\}$$

and $z_i = m_i/m, z_0 = (m - \sum m_i)/m$.

The rest of the proof is concerned with the evaluation of the above maximum. Define

$$g(z) = h(z) + \sum_{i=1}^n z_i \log \binom{n}{i} \quad (8)$$

$$\bar{\sigma} = \sup\{\sigma > 0 : \bar{F}_{n,t}(y) < 0 \text{ for all } 0 \leq y < \sigma\}.$$

As long as $\sigma < \bar{\sigma}$, the probability of not being able to correct σtm errors with a random code from the considered ensemble approaches zero. Thus, we need to find the maximum $\max_{z \in \mathcal{M}'_\epsilon(t, \sigma)} g(z)$ for all $\sigma \in [0, \bar{\sigma}]$. The proof will be accomplished in the next three steps. Since ϵ will be assumed arbitrarily small, we will omit it from our considerations and write \mathcal{M}' instead of \mathcal{M}'_ϵ .

1. We find the point z^* that gives the maximum of $g(z)$ without the constraint $\sum_{i=t+1}^n z_i \geq \sigma$.

2. Next we show that for $0 \leq \sigma < \bar{\sigma}$, the point $z^* \notin \mathcal{M}'$, and therefore the maximum over \mathcal{M}' is attained on the boundary, i.e., we can replace \mathcal{M}' with

$$\mathcal{M}(t, \sigma) = \left\{ z \in \mathcal{Z}_n : \sum_{i=1}^n iz_i = \sigma n, \sum_{i=t+1}^n z_i = \sigma \right\}.$$

3. Finally we compute the value of the maximum.

Step 1. Without the constraint $\sum_{i=t+1}^n z_i \geq \sigma$ the maximum is easily computed. Indeed, the proportion of edges incident to the vertices in S out of the N edges of G is σ , so the fraction of vertices with S -degree i should be close to $z_i^*(\sigma) = \binom{n}{i} \sigma^i (1-\sigma)^{n-i}$. Thus, the coordinates of the maximizing point $z^* = z^*(\sigma)$ are $z_i^*, i = 1, \dots, n; z_0 = 1 - \sum_i z_i^*$, and

$$g(z^*) = nh(\sigma).$$

Slightly more formally, note that z^* is the unique stationary point of the function $g(z)$, and that this function is strictly concave in z . Therefore, z^* is a unique maximum of $g(z)$ on \mathcal{Z}_n , and the function $g(z)$ grows in the direction $z^* - z$ for any $z \in \mathcal{Z}_n$.

Step 2. Suppose that $0 \leq \sigma \leq \bar{\sigma}$. Observe that $p(\sigma) \triangleq \sum_{i=t+1}^n z_i^* = P(X \geq t+1)$, where X is a $(\sigma, 1-\sigma)$ binomial

random variable. This probability is monotone increasing on σ for $\sigma \in [0, 1]$, and $p(0) = p'(0) = 0$. Thus for $\sigma \in [0, \alpha)$ where α is the smallest positive root of $\sum_{i=t+1}^n z_i^*(\sigma) = \sigma$, we have

$$\sum_{i=t+1}^n z_i^* = \sum_{i=t+1}^n \binom{n}{i} \sigma^i (1-\sigma)^{n-i} < \sigma,$$

and so the point $z^*(\sigma) \notin \mathcal{M}'(t, \sigma)$. Our claim will follow if we show that $\bar{\sigma} < \alpha$. This is indeed the case because for $0 \leq \sigma < \bar{\sigma}$,

$$\max_{z \in \mathcal{M}'(t, \sigma)} g(z^*(\sigma)) < (n-1)h(\sigma).$$

On the other hand, $g(z^*(\alpha)) = nh(\alpha)$. This establishes that the maximum of $g(z)$ on $z \in \mathcal{M}'$ is attained on the hyperplane $\sum_{i=t+1}^n z_i = \sigma$.

Step 3. To compute the maximum of $g(z)$ on z , let us form the Lagrangian

$$U(z, \tau_1, \tau_2) = h(z) + \sum_{i=1}^n z_i \log \binom{n}{i} + \tau_1 \left(\sum_{i=1}^n iz_i - \sigma n \right) + \tau_2 \left(\sum_{i=t+1}^n z_i - \sigma \right).$$

Setting $\nabla U = 0$ and $\tau_1 = \log x, \tau_2 = \log y$, we find that

$$z_i = \begin{cases} \binom{n}{i} x^i D & \text{if } 0 \leq i \leq t \\ \binom{n}{i} y x^i D & \text{if } t < i \leq n, \end{cases}$$

where we have denoted

$$D = \left[\sum_{i=0}^t \binom{n}{i} x^i + y \sum_{i=t+1}^n \binom{n}{i} x^i \right]^{-1}.$$

Adding these equations together, we find conditions for x and y :

$$\sigma = Dy \sum_{i=t+1}^n \binom{n}{i} x^i$$

$$\sigma n = D \left(\sum_{i=0}^t i \binom{n}{i} x^i + y \sum_{i=t+1}^n i \binom{n}{i} x^i \right).$$

Once y is eliminated from the last two equations, we obtain the condition (5) for x . Finally, substituting the found values of $z_i, i = 1, \dots, n$ into $g(z)$, we find that the maximum evaluates to the expression $F_{n,t}(\sigma)$ given in (4) (and therefore, $\bar{\sigma} = \sigma_0$). Since we seek to obtain a value $L < 0$, the boundary condition for the proportion of correctable errors is obtained by setting $L = 0$. This concludes the proof. ■

Example 1: Using Theorem 4.2 together with (4) we can compute the proportion of errors corrected by codes in the ensemble $\mathcal{G}(A, m), m \rightarrow \infty$ for several choices of the local code A . For instance, taking A to be the binary Golay code of length $n = 23$ we find $\sigma_0 \approx 0.0048586$ and therefore, the proportion of correctable errors is $\frac{\sigma_0 t}{n} \approx 0.00063$. Similarly,

for the 2-error-correcting $[n = 31, k = 21]$ BCH code we find $\sigma_0 \approx 0.000035$ and $\frac{\sigma_0 t}{n} \approx 0.0000023$.

To underscore similarities with the results obtained for product codes and their later variations including graph codes (e.g., [13]) we compute the proportion of errors correctable with codes from the ensemble $\mathcal{G}(A, m)$ in the case of large n .

Proposition 4.3: Let $t = \tau n$. Then the ensemble $\mathcal{G}(A, m)$ contains codes that correct $\sigma \tau N$ errors for any $\sigma \leq \sigma_0$, where σ_0 is given by

$$\sigma_0 = \sup \left\{ \sigma > 0 : \forall_{0 < x < \sigma} (1-x)h\left(\frac{x(1-\tau)}{1-x}\right) + xh(\tau) + \varepsilon_n < h(x) \right\}$$

where $\varepsilon_n = (1 + \log n)/n$.

Proof: Referring to the notation of the previous proof, let us evaluate the asymptotic behavior of the exponent L of the probability in (7). Since $h(z) \leq \log n$, we have

$$n^{-1} \bar{F}_{n,t}(\sigma) \leq -h(\sigma) + n^{-1} \max_{\mathbf{z} \in \mathcal{M}(\tau n, \sigma)} \sum_{i=0}^n z_i \log \binom{n}{i} + n^{-1}(1 + \log n).$$

Next,

$$\begin{aligned} \frac{1}{n} \sum_{i=0}^n z_i \log \binom{n}{i} &\leq \sum_i z_i h\left(\frac{i}{n}\right) \\ &= (1-\sigma) \sum_{i=0}^t \frac{z_i}{1-\sigma} h\left(\frac{i}{n}\right) + \sigma \sum_{i=t+1}^n \frac{z_i}{\sigma} h\left(\frac{i}{n}\right) \\ &\leq (1-\sigma) h\left(\frac{\sum_{i=1}^t i z_i}{(1-\sigma)n}\right) + \sigma h\left(\frac{\sum_{i=t+1}^n i z_i}{\sigma n}\right). \end{aligned}$$

Let $y = n^{-1} \sum_{i=t+1}^n i z_i$, then for any $\mathbf{z} \in \mathcal{M}(\tau n, \sigma)$ we have

$$\frac{1}{n} \sum_{i=0}^n z_i \log \binom{n}{i} \leq \max_{\tau \sigma \leq y \leq \sigma} \left\{ (1-\sigma) h\left(\frac{\sigma-y}{1-\sigma}\right) + \sigma h\left(\frac{y}{\sigma}\right) \right\}.$$

The function on the right-hand side of this inequality is concave. Its global maximum equals $h(\sigma)$ and is attained for $y = \sigma^2$. Thus, assuming that $\sigma < \tau$, we conclude that the constrained maximum occurs for $y = \tau \sigma$, which gives the following bound on $n^{-1} \bar{F}_{n,t}(\sigma)$:

$$n^{-1} \bar{F}_{n,t}(\sigma) \leq -h(\sigma) + (1-\sigma) h\left(\frac{\sigma(1-\tau)}{1-\sigma}\right) + \sigma h(\tau) + \varepsilon_n.$$

As long as the right-hand side of this inequality is negative, the previous proof implies that the code corrects all errors of multiplicity up to $\sigma \tau N$. ■

From the expression of this proposition we observe that (as $n \rightarrow \infty$) the value of σ_0 approaches τ , so the ensemble \mathcal{G} contains codes that correct up to a τ^2 proportion of errors, where $\tau n = d_0/2$ is the error-correcting capability of the code A . This result parallels the product bound on the error-correcting radius of direct product codes. As in the case of product and expander codes (e.g., [2]), the proportion of correctable errors can be improved from $\tau^2 = (d_0/(2n))^2$ by using a more powerful decoding algorithm.

V. NUMBER OF CORRECTABLE ERRORS FOR THE ENSEMBLE $\mathcal{H}(A, l, m)$

In this section we first state a sufficient condition for the existence of at least one subprocedure within each step of Algorithm II that reduces the number of errors, and then perform the analysis of random hypergraphs to show that with high probability this condition is satisfied. Overall this will show that the number of errors in at least one of the candidates in the list generated after a few iterations is reduced to a desired level.

Denote by $E(v)$ the set of edges incident to a vertex $v \in V$. Let $C \in \mathcal{H}(A, l, m)$ be a code and let $H(V, E)$ be its associated graph. Let $\mathcal{E} \subset E$ be the set of errors at the start of some iteration of the algorithm. The next set of arguments will refer to this iteration. Let $G_i = \{v \in V_i : |E(v) \cap \mathcal{E}| \leq t\}$ be the set of vertices such that each of them is incident to no more than t edges from \mathcal{E} (such errors will be corrected upon one decoding). Let $B_i = \{v \in V_i : |E(v) \cap \mathcal{E}| \geq d_0 - t\}$ be the set of vertices that can introduce errors after one decoding iteration. Note that each of such vertices introduces at most t errors.

The main condition for successful decoding is given in the next lemma.

Lemma 5.1: Assume that for every $\mathcal{E} \subset E, |\mathcal{E}| \leq \gamma N$ there exists $i = i(\mathcal{E}), 1 \leq i \leq l$ such that $|\mathcal{E}(G_i)| \geq t|B_i| + \epsilon N$, where $\mathcal{E}(G_i)$ is the set of edges of \mathcal{E} incident to the vertices of G_i and $\epsilon > 0$. Then for any $0 < \beta < \gamma$, Algorithm II will reduce any γN errors in the received vector to at most βN errors in $c(\beta, \gamma, \epsilon)$ iterations where c is a constant independent of N .

Proof: We need to prove that at least one of the subprocedures will find a vector with no more than βN errors after a constant number of iterations. In any given iteration by the assumption of the lemma there exists a component V_i for which the i th subprocedure will decrease the count of errors by $|\mathcal{E}(G_i)| - t|B_i| \geq \epsilon N$. Thus, in each iteration there exists a subprocedure that reduces the number of errors by a positive fraction. ■

Next we show that the assumption of Lemma 5.1 holds with high probability over the ensemble. Consider the function

$$\tilde{F}_{n,t}(\gamma) = \max_{\mathbf{z} \in \mathcal{M}(t, \gamma)} \left(h(\mathbf{z}) + \sum_{i=0}^n z_i \log \binom{n}{i} \right),$$

where in this section the region $\mathcal{M}(t, \gamma)$ will be as follows:

$$\mathcal{M}(t, \gamma) = \left\{ \mathbf{z} \in \mathcal{Z}_n : \sum_{i=1}^n i z_i = \gamma n, \sum_{i=1}^t i z_i = \sum_{i=d_0-t}^n t z_i \right\}. \quad (9)$$

Lemma 5.2: Let $m \rightarrow \infty$ and let

$$\gamma_0 = \sup \{ x > 0 : \forall_{0 < \gamma \leq x} (l/n) \tilde{F}_{n,t}(\gamma) < (l-1)h(\gamma) \}. \quad (10)$$

A hypergraph from the ensemble of l -partite uniform n -regular hypergraphs with probability $1 - 2^{-\Omega(N)}$ has the property that for all $\mathcal{E} \subset E, |\mathcal{E}| < \gamma_0 N$, and some $\epsilon > 0$, the inequality $|\mathcal{E}(G_i)| \geq t|B_i| + \epsilon N$ holds for at least one $i \in \{1, \dots, l\}$.

Proof: Let $\mathcal{E} \subset E, |\mathcal{E}| = \gamma N$. Let $m_i = |\{v \in V_1 : |E(v) \cap \mathcal{E}| = i\}|, i = 1, \dots, n$. Clearly $|\mathcal{E}(G_1)| = \sum_{i=0}^t im_i$ and $|B_1| = \sum_{i=d_0-t}^n m_i$. We have

$$p \triangleq P(|\mathcal{E}(G_i)| \leq t|B_i| + \epsilon N) \\ = \frac{1}{\binom{N}{\gamma N}} \sum_{\mu \in M_\epsilon(t, \gamma)} \binom{m}{\mu} \prod_{i=0}^n \binom{n}{i}^{m_i},$$

where $\mu = \{m_1, \dots, m_n\}$,

$$M_\epsilon(t, \gamma) = \{\mu \in (\mathbb{Z}_+ \cup 0)^n : \sum_{i=1}^n m_i \leq m, \\ \sum_{i=1}^n im_i = \gamma N, \sum_{i=1}^t im_i \leq \sum_{i=d_0-t}^n tm_i + \epsilon N\}.$$

Denote by $\mathcal{L}(\mathcal{E})$ the event that for a given subset $\mathcal{E} \subset E, |\mathcal{E}| = \gamma N$ no part V_i of H satisfies the assumption of Lemma 5.1. Then $P(\mathcal{L}(\mathcal{E})) = p^l$ and

$$P\{\exists \mathcal{E} : (|\mathcal{E}| \leq \gamma N) \wedge (\mathcal{L}(\mathcal{E}))\} \leq N \binom{N}{\gamma N} p^l.$$

Letting L to be the logarithm of the left-hand side of this inequality divided by N and omitting $o_N(1)$ terms, we obtain

$$L \leq -(l-1)h(\gamma) + \frac{l}{n} \max_{z \in \mathcal{M}'(t, \gamma)} g(z), \quad (11)$$

where $g(z)$ is defined in (8),

$$\mathcal{M}'(t, \gamma) = \{z \in \mathcal{Z}_n : \sum_{i=1}^n iz_i = \gamma n, \sum_{i=1}^t iz_i \leq \sum_{i=d_0-t}^n tz_i\}$$

and $z_i = m_i/m$ (as in the previous section, we have omitted ϵ which can be made arbitrarily small).

The proof will be complete if we show that the optimization region \mathcal{M}' can be replaced by \mathcal{M} . For that we follow the logic of the second part of the proof of Theorem 4.2. As before, the maximum of $g(z)$ without the constraint $\sum_{i=1}^t iz_i \leq \sum_{i=d_0-t}^n tz_i$ is attained at the point $z^*(\gamma) = (z_0^*, z_1^*, \dots, z_n^*) \in \mathcal{Z}_n$, where

$$z_i^* = z_i^*(\gamma) = \binom{n}{i} \gamma^i (1-\gamma)^{n-i}, \quad i = 1, \dots, n.$$

We need to show that as long as $0 \leq \gamma < \gamma_0$, the point $z^* \notin \mathcal{M}'(t, \gamma)$. By concavity of the objective function and the optimization region, this will imply that the maximum is on the boundary. As before, it is possible to show that in the neighborhood of $\gamma = 0$,

$$\sum_{i=1}^t iz_i^* > \sum_{i=d_0-t}^n tz_i^*.$$

and thus for $\gamma < \beta$, where β is the smallest positive root of $\sum_{i=1}^t iz_i^* = \sum_{i=d_0-t}^n tz_i^*$, the point $z^*(\gamma) \notin \mathcal{M}'(t, \gamma)$. Let

$$\bar{\gamma} = \sup\{\gamma : \forall 0 < x < \gamma, \text{ rhs of (11)} < 0\}.$$

We note that for all $\gamma \leq \bar{\gamma}$,

$$\max_{z \in \mathcal{M}'(t, \sigma)} g(z) < (l-1)nh(\gamma).$$

On the other hand, $g(z^*(\beta)) = nh(\beta)$. This implies that $\bar{\gamma} < \beta$, and so for all $\gamma < \bar{\gamma}$, the point $z^*(\gamma) \notin \mathcal{M}'(t, \gamma)$. Thus the region \mathcal{M}' in the maximization can be replaced with \mathcal{M} (and $\bar{\gamma} = \gamma_0$). ■

This lemma establishes that the number of errors in at least one of the candidates in the list generated after a few iterations is reduced to a desired level. After that the residual errors can be removed by another procedure as described above. In this situation we say that the errors are correctable by Algorithm II, without explicitly mentioning the second stage.

In the next theorem, which is the main result of this section, δ refers to the lower estimate of the average relative distance of the hypergraph code ensemble \mathcal{H} from Theorem 5.5 below.

Theorem 5.3: Let $t \geq 2$ be the number of errors correctable by the local code A . Algorithm II corrects any combination of up to $N(\min(\gamma_0, \delta/2))$ errors for any code $C \in \mathcal{H}(A, l, m)$ except for a proportion of codes that declines exponentially with the code length $N = nm, m \rightarrow \infty$.

Proof: With high probability over the ensemble of hypergraphs considered, for a given hypergraph $H(V, E)$ a constant number s of iterations of the algorithm will decrease the weight of error from $\gamma_0 N$ to any given positive proportion β for at least one of the l^s candidates in the list $Y_1^{(s+1)}$. Take $\beta = \sigma_0$, where σ_0 is the quantity given by Theorem 4.2. Next consider the bipartite graph $G(V_G = V_1 \cup V_2, E_G)$ where V_1, V_2 are the parts of H and where $(v_1, v_2) \in E_G$ if $v_1, v_2 \in e$ for some edge $e \in E$. By the previous section, with high probability these $\sigma_0 N$ errors can be corrected with $O(\log m)$ iterations of Algorithm I. Finally, the correct codeword will be selected from the list of candidates because the proportion of errors is assumed not to exceed $N\delta/2$. ■

The complexity of this decoding is $O(N \log N)$ where the implicit constant depends on the code A .

In the following theorem we extend the results of this section to the case of A being a perfect single-error correcting Hamming code of length $n = 2^r - 1$ for some $r = 3, 4, \dots$. In this case the maximum on z in the above proof can be computed in a closed form. As remarked above, in this case in the last part of the error correction procedure we use the decoding algorithm of [14] to remove residual errors from the candidate vectors.

Theorem 5.4: Suppose that the local codes A are taken to be one-error-correcting Hamming codes and let $\delta = \delta(\mathcal{H})$ be the relative average distance (2) of the ensemble $\mathcal{H}(A, l, m)$. Then almost all codes in the ensemble $\mathcal{H}(A, l, m)$ can be decoded to correct $N \min(\gamma_0, \delta/2)$ errors, where γ_0 is given by (10) and

$$\tilde{F}_{n,1}(\gamma) = -\gamma n \log x + \log \left(1 + 2 \sqrt{n \sum_{i=2}^n \binom{n}{i} x^{i+1}} \right) \quad (12)$$

where x is the only positive root of the equation

$$\frac{\sum_{i=2}^n (i+1) \binom{n}{i} x^{i+1}}{2n \sum_{i=2}^n \binom{n}{i} x^{i+1} + \sqrt{n \sum_{i=2}^n \binom{n}{i} x^{i+1}}} = \gamma.$$

Proof: It is obtained by maximizing the function $g(\mathbf{z})$ over the region

$$\mathcal{M}(1, \gamma) = \left\{ \mathbf{z} \in \mathcal{Z}_n : \sum_{i=1}^n iz_i = \gamma n, z_1 = \sum_{i=2}^n z_i \right\}.$$

The Lagrangian takes the form

$$h(\mathbf{z}) + \sum_{i=2}^n z_i \left(\log n + \log \binom{n}{i} \right) + \lambda \left(\sum_{i=2}^n (i+1)z_i - \gamma n \right),$$

where $\mathbf{z} = (z_1, z_2, \dots, z_n, 1 - \sum_i z_i)$ and $z_1 = \sum_{i=2}^n z_i$ and λ is an arbitrary multiplier. Setting the partial derivatives to zero, we find the value λ to satisfy $2^x = \lambda$, where x is given above. The calculations are tedious but straightforward and will be omitted. ■

The last theorem enables us to find the proportion of correctable errors for the case when A is the Hamming code of length $n = 2^r - 1, t = 1$. Since the examples below rely on the value of the ensemble-average distance, we quote the corresponding result from [1].

Theorem 5.5: [1, Thm.5] Let $\delta(\mathcal{H})$ be the asymptotic average relative distance of codes in the l -hypergraph ensemble constructed from the local code A of length n and distance d_0 . Then

$$\delta(\mathcal{H}) \geq \sup_{\omega > 0} \left\{ \omega : \frac{l}{n} \log \frac{1 + \sum_{i=d_0}^n \binom{n}{i} \omega^i}{\omega_0^{i \omega^n}} < (l-1)h(\omega) \right\}$$

where $x_0 = x_0(\omega)$ is the positive solution of the equation

$$\omega n + \sum_{i=d_0}^n \binom{n}{i} (\omega n - i) \omega^i = 0.$$

For instance, for the case $n = 31, l = 5$ this theorem gives the value of the relative distance $\delta(\mathcal{H}) \geq 0.01618$ (the rate of codes $R \geq 6/31$). Performing the calculation in (12), we find that the average code from the ensemble $\mathcal{H}(A, 5, m)$ the proportion of errors correctable by codes in the ensemble using Algorithm II to be at least $\gamma_0 = 1.2 \times 10^{-5}$.

We include some more examples. In the following table $n = 2^9 - 1$.

Example 2:

l	17	23	28	34
Rate	0.7006	0.5949	0.5069	0.4012
γ_0	0.000235	0.000401	0.000521	0.000644
$\delta(\mathcal{H})$	0.00415	0.00504	0.00558	0.00608

l	40	45	51
Rate	0.2955	0.2074	0.1018
γ_0	0.000747	0.000821	0.000898
$\delta(\mathcal{H})$	0.00648	0.00676	0.00704

It is also of interest to compute the values of γ_0 for code rate $R(C) \approx 0.5$.

n	127	255	511	1023
l	9	16	28	51
Rate	0.5039	0.4980	0.5068	0.5015
γ_0	0.0002012	0.0004873	0.0005207	0.0004227
$\delta(\mathcal{H})$	0.01157	0.008658	0.005581	0.003394

These estimates are at least an order of magnitude better than the corresponding results in [6], [14] obtained for LDPC

codes and their generalizations based on the ‘‘flipping’’ algorithm of [15].

The case of large n . As in the previous section, it is interesting to examine the case of long local codes A because it reveals some parallels with the analysis of the decoding algorithm in the case of nonrandom hypergraphs [1]. We begin with the observation that the proportion γ_0 of correctable errors for the ensemble $\mathcal{H}(A, t, m)$ computed above is a function of the number of errors t that each local code corrects in each iteration.

Lemma 5.6: Let $t = \tau n, d_0 = \delta_0 n$. The ensemble $\mathcal{H}(A, t, m)$ contains codes that correct γN errors for any $\gamma < \gamma_0(\tau) \triangleq \min(\tau, x_0(\tau))$ where

$$x_0(\tau) = \sup \left\{ x > 0 : \left(1 - \frac{x}{\delta_0} \right) h \left(\frac{x\tau}{\delta_0 - x} \right) + \frac{x}{\delta_0} h(\delta_0 - \tau) + \varepsilon_n < (1 - 1/l)h(x) \right\}$$

and $\varepsilon_n = \log n/n$.

Proof: Referring to the proof of Lemma 5.2, we aim at establishing conditions for the exponent L of the event $\mathcal{L}(\mathcal{E})$ to be negative as m approaches infinity. We assume that $\gamma \leq \tau$ (otherwise our estimates do not imply that the convergence condition of Lemma 5.1 holds with high probability over the graph ensemble).

From (11), (8) we have

$$L \leq -(l-1)h(\gamma) + l \max_{\mathbf{z} \in \mathcal{M}(t, \gamma)} \sum_{i=0}^n z_i h \left(\frac{i}{n} \right) + \frac{l \log n}{n},$$

where $\mathcal{M}(t, \gamma)$ is defined in (9). Next, write

$$\sum_{i=0}^t z_i h \left(\frac{i}{n} \right) \leq \lambda h \left(\frac{\sum_{i=1}^t iz_i}{\lambda n} \right) = \lambda h \left(\frac{\mu_1}{\lambda} \right), \quad (13)$$

where we have denoted $\sum_{i=0}^t z_i = \lambda, \sum_{i=1}^t iz_i = \mu_1 n$. In addition let us put $\sum_{i=d_0-t}^n iz_i = \mu_2 n$, then the values of the sums $\sum_i z_i$ and $\sum_i iz_i$ over each of the three intervals $I_1 = [0, t], I_2 = [t+1, d_0-t-1], I_3 = [d_0-t, n]$ can be found from the following table:

	I_1	I_2	I_3
$\sum z_i$	λ	$1 - \lambda - \mu_1/\tau$	μ_1/τ
$\sum \frac{i}{n} z_i$	μ_1	$\gamma - \mu_1 - \mu_2$	μ_2

The variables introduced above depend on the point \mathbf{z} and satisfy the following natural constraints: for any $\mathbf{z} \in \mathcal{M}(t, \gamma)$,

$$\begin{aligned} \mu_1 &\leq \tau \lambda \\ \tau \left(1 - \lambda - \frac{\mu_1}{\tau} \right) &\leq \gamma - \mu_1 - \mu_2 \leq (\delta_0 - \tau) \left(1 - \lambda - \frac{\mu_1}{\tau} \right) \\ (\delta_0 - \tau) \frac{\mu_1}{\tau} &\leq \mu_2 \leq \frac{\mu_1}{\tau}. \end{aligned} \quad (14)$$

Proceeding as in (13), we can estimate the sum on z_i in L as follows:

$$\sum_{i=0}^n z_i h \left(\frac{i}{n} \right) \leq f(\lambda, \mu_1, \mu_2) \quad (15)$$

where

$$f(\lambda, \mu_1, \mu_2) = \lambda h\left(\frac{\mu_1}{\lambda}\right) + \left(1 - \lambda - \frac{\mu_1}{\tau}\right) h\left(\frac{\gamma - \mu_1 - \mu_2}{1 - \lambda - (\mu_1/\tau)}\right) + \frac{\mu_1}{\tau} h\left(\frac{\mu_2\tau}{\mu_1}\right).$$

Our plan is to prove that some of the inequalities in (14) can be replaced by equalities, thereby expressing the variables λ, μ_1, μ_2 as functions of γ, τ . We will rely on the fact that the function f is concave in its domain, proved in the end of this section.

Note that for all $z \in \mathcal{Z}_n$ the sum

$$\sum_{i=0}^n z_i h\left(\frac{i}{n}\right) \leq h(\gamma)$$

and that it equals $h(\gamma)$ at the point \tilde{z} such that $z_i = 1$ for $i = \lceil \gamma n \rceil$ and $z_i = 0$ elsewhere. Also note that since $\gamma < \tau$, the point \tilde{z} is outside the region $\mathcal{M}(t, \gamma)$ and thus, by concavity,

$$a := \max_{z \in \mathcal{M}(t, \gamma)} \sum_{i=0}^n z_i h\left(\frac{i}{n}\right) < h(\gamma).$$

Let z_1 be the point at which this maximum is attained, and let $\mathbf{x}_1 = (\lambda, \mu_1, \mu_2)$ be the corresponding point for the arguments of f . By construction, the point \mathbf{x}_1 satisfies the inequalities of (14). At the same time, consider the function $f(\cdot)$ on the line $\lambda = \mu_1 = \mu_2$. As the variables approach 0 along this line, the value $f(\lambda, \mu_1, \mu_2)$ approaches $h(\gamma)$.

To summarize, we have found two points, \mathbf{x}_1 and $\mathbf{x}_2 = (0, 0, 0)$ that are located on different sides of the hyperplane

$$\tau\left(1 - \lambda - \frac{\mu_1}{\tau}\right) = \gamma - \mu_1 - \mu_2$$

such that $f(\mathbf{x}_1) \geq a, f(\mathbf{x}_2) > a$. Invoking concavity of the function f , we now conclude that there is a feasible point \mathbf{x}' on this hyperplane such that $f(\mathbf{x}') \geq a$.

Therefore, put $\mu_2 = \gamma - \tau(1 - \lambda)$ and write

$$f_1(\lambda, \mu_1) = \lambda h\left(\frac{\mu_1}{\lambda}\right) + \left(1 - \lambda - \frac{\mu_1}{\tau}\right) h(\tau) + \frac{\mu_1}{\tau} h\left(\frac{\tau(\gamma - \tau(1 - \lambda))}{\mu_1}\right)$$

where the variables are constrained as follows: for any $z \in \mathcal{M}(t, \gamma)$,

$$\mu_1 \leq \tau\lambda$$

$$\tau(1 - \lambda) - \mu_1 \geq 0 \quad (16)$$

$$(\delta_0 - \tau)\frac{\mu_1}{\tau} \leq \gamma - \tau(1 - \lambda) \leq \frac{\mu_1}{\tau}. \quad (17)$$

Since f_1 is a restriction of f to a hyperplane, it is still concave. Now notice that $f_1(1, \tau) = h(\gamma)$ and that the point $(1, \tau)$ does not satisfy inequality (16) and the left of the inequalities (17). Repeating the above argument, we claim that the function f in (15) can be further restricted to the intersection of the planes $\tau(1 - \lambda) = \mu_1$ and $(\delta_0 - \tau)(\mu_1/\tau) = \gamma - \tau(1 - \lambda)$. Altogether this gives:

$$\lambda = 1 - \gamma/\delta_0, \quad \mu_1 = \gamma\tau/\delta_0.$$

Let us substitute these values into the expression for f_1 and rewrite (15) as follows: for any $0 \leq \gamma < \tau$,

$$\max_{z \in \mathcal{M}(t, \gamma)} \sum_{i=0}^n z_i h\left(\frac{i}{n}\right) \leq \left(1 - \frac{\gamma}{\delta_0}\right) h\left(\frac{\gamma\tau}{\delta_0 - \gamma}\right) + \frac{\gamma}{\delta_0} h(\delta_0 - \tau). \quad (18)$$

Thus if the condition in the statement is fulfilled then $L < 0$. This concludes the proof. \blacksquare

Remark. The main part of the proof is estimating the solution of the following linear program

$$\max_z \sum_{i=1}^n z_i h\left(\frac{i}{n}\right) \\ z = (z_0, z_1, \dots, z_n) \in \mathcal{M}(t, \gamma)$$

where the variables define a probability distribution on $\{0, 1, \dots, n\}$. It is clear from concavity that the maximum is attained at the point where among all the indices $i \in I_1$ at most one value z_i is nonzero, and the same applies to I_2 and I_3 . We have shown that the value of the program is bounded above by the right-hand side of (18). The following point gives this value and is therefore a maximizing point:

$$z_{i_1} = 1 - \frac{\gamma}{\delta_0}, \quad z_{i_2} = \frac{\gamma}{\delta_0}, \quad z_i = 0 \text{ otherwise,}$$

where $i_1 = n\gamma\tau/(\delta_0 - \gamma), i_2 = n(\delta_0 - \tau)$. Since

$$\frac{\gamma\tau}{\delta_0 - \gamma} \leq \tau,$$

this shows that the worst-case allocation of errors to vertices in a given part of the graph assigns no edges to vertices that are neither good nor bad. This also confirms the intuition suggested by Lemma 5.1 that bad vertices (vertices assumed to add errors) should each be assigned the smallest possible number of error edges $d_0 - t$.

The next proposition is now immediate.

Proposition 5.7: The ensemble $\mathcal{H}(A, l, m)$ with long local codes contains codes that can be decoded using Algorithm II to correct all error patterns whose weight is less than $\gamma_0 N$, where

$$\gamma_0 = \max_{0 < \tau \leq \delta_0/2} \gamma_0(\tau). \quad (19)$$

Estimating the number of correctable errors for the ensemble $\mathcal{H}(A, l, m)$ from Proposition 5.7 analytically is difficult because it involves optimization on τ (generally, the local codes should be used to correct a smaller than $\delta_0/2$ proportion of errors). We note that in the particular case of $\tau = \delta_0/2$ the proof of Lemma 5.6 can be considerably simplified, although the resulting value of γ is not always optimal.

Example 3. Let $l = 3$. Using local codes with $\delta_0 = 0.05$ we can construct hypergraph codes of rate $R \geq 0.19$. From [1, Cor. 6], the ensemble-average relative distance is at least $\delta \approx 0.0112$ and the proportion of errors correctable by Algorithm II is found from (19) to be $\gamma_0 \approx 0.0035$.

Example 4. Let $\delta_0 = 0.01$ and $l = 10$. In this case, we find from [1, Cor. 6] the value of the relative distance $\delta \approx 0.00599$. The code rate satisfies $R \geq 0.14$. Performing the computations in (19) and Lemma 5.6 we find the estimate of the proportion of correctable errors to be $\gamma_0 \approx 0.002198$.

Proof that $f(\lambda, \mu_1, \mu_2)$ is concave. First we prove that the function

$$\phi(x, y) = (1-x)h\left(\frac{\gamma-y}{1-x}\right)$$

is concave (not necessarily in the strict sense) for $0 < x, y < 1, 0 < \gamma - y < 1 - x$. For that, let us compute its Hessian matrix:

$$H = \frac{1}{\ln 2} \begin{pmatrix} \frac{\gamma-y}{(1-x)(\gamma-y+x-1)} & -\frac{1}{\gamma-y+x-1} \\ -\frac{1}{\gamma-y+x-1} & \frac{1-x}{(\gamma-y)(\gamma-y+x-1)} \end{pmatrix}$$

The eigenvalues of H are

$$0, \quad \frac{(\gamma-y)^2 + (1-x)^2}{(1-x)(\gamma-y)(\gamma-y-(1-x))} < 0,$$

so $H \preceq 0$, and so ϕ is concave. Next observe that the function

$$\left(1 - \lambda - \frac{\mu_1}{\tau}\right)h\left(\frac{\gamma - \mu_1 - \mu_2}{1 - \lambda - (\mu_1/\tau)}\right)$$

can be obtained from ϕ by a linear change of variables

$$x = \lambda + \mu_1/\tau, \quad y = \mu_1 + \mu_2$$

and therefore is also concave. Finally, the functions $\lambda h(\mu_1/\lambda)$ and $(\mu_1/\tau)h(\mu_2\tau/\mu_1)$ are also concave, and thus so is the function $f(\lambda, \mu_1, \mu_2)$.

VI. CONCLUSION

We have estimated the proportion of errors correctable by codes from ensembles defined by random l -partite graphs, $l \geq 2$. In contrast to the case of expander codes [10], [13], [2], [4], [1] our calculations cover the case of local codes of arbitrary given length and distance, including small values of the distance. The behavior of code ensembles considered here was examined from a different perspective in [1] where we computed estimates of the expected distance and weight distribution of these codes. The paper [1] and the present work together provide answers to the set of basic questions regarding random networks of short linear binary codes and extend our perspective of concatenated code constructions to the case of sparse regular graphs.

Acknowledgment. The authors are grateful to Jørn Justesen and to an anonymous reviewer for useful comments on this work.

REFERENCES

- [1] A. Barg, A. Mazumdar, and G. Zémor, "Weight distribution and decoding of codes on hypergraph," *Advances in Mathematics of Communication*, vol. 2, no. 4, pp. 433–450, 2008.
- [2] A. Barg and G. Zémor, "Concatenated codes: Serial and parallel," *IEEE Trans. Inform. Theory*, vol. 51, pp. 1625–1634, 2005.
- [3] A. Barg and G. Zémor, "Distance properties of expander codes," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 78–90, 2006.
- [4] Y. Bilu and S. Hoory, "On codes from hypergraphs," *European Journal of Combinatorics*, vol. 25, pp. 339–354, 2004.
- [5] J. Boutros, O. Potier and G. Zémor, Generalized low-density (Tanner) codes, in Proc. IEEE ICC, Vancouver, Canada, pp. 441–445, 1999.
- [6] D. Burshtein, "On the error correction of regular LDPC codes using the flipping algorithm," *IEEE Trans. Inform. Theory*, vol. 54, no. 2, pp. 517–530, 2008.
- [7] R. G. Gallager, *Low-Density Parity-Check Codes*, MIT Press, 1963.

- [8] M. Lentmaier and K. Sh. Zigangirov, *On generalized low-density parity-check codes based on Hamming component codes*, IEEE Communications Letters **3** (1999), no. 8, 248–250.
- [9] A. Nilli, "On the second eigenvalue of a graph," *Discrete Math.*, vol. 91, no. 2, 207–210, 1991.
- [10] M. Sipser and D. A. Spielman, "Expander codes," *IEEE Trans. Inform. Theory*, vol. 42, no. 6, 1710–1722, 1996.
- [11] X. Tang and R. Koetter, "Performance of iterative algebraic decoding of codes defined on graphs: An initial investigation," Proc. 2007 IEEE Information Theory Workshop, Lake Tahoe, CA, Sept. 2–6, 2007, pp. 254–259.
- [12] M. Tanner, "A recursive approach to low-complexity codes," *IEEE Trans. Inform. Theory*, vol. 27, no. 5, pp. 1710–1722, 1981.
- [13] G. Zémor, "On expander codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, 835–837, 2001.
- [14] V. V. Zyablov, R. Johannesson, and M. Lončar, "Low-complexity error correction of Hamming-code-based LDPC codes," *Probl. Inform. Trans.*, vol. 45, no. 2, pp. 95–109, 2009.
- [15] V. V. Zyablov and M. S. Pinsker, "Estimation of the error-correcting complexity of Gallager low-density codes," *Probl. Inform. Trans.*, vol. 11, no. 1, pp. 18–28, 1975.