Number of errors correctable with codes on graphs

Alexander Barg
University of Maryland

joint work with

Arya Mazumdar
University of Maryland

ITA workshop, Feb. 9, 2009

Linear binary codes constructed from regular sparse graphs

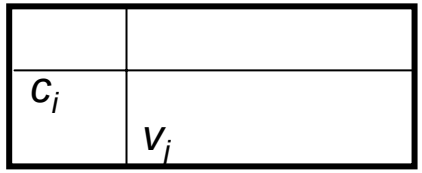
(a) LDPC codes

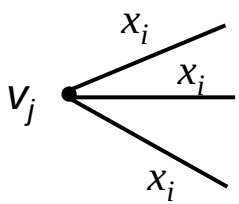
(b) bipartite-graph codes

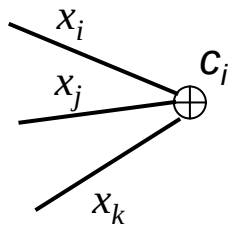
(c) hypergraph codes

LDPC codes

H =





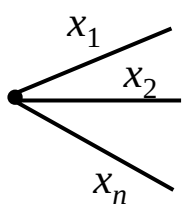


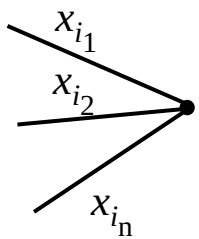
$$x_i + x_j + x_k = 0$$

repetition code

single parity check code

Generalization: Bipartite-graph codes (Tanner)



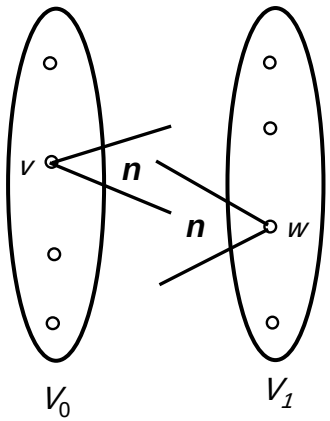


A, B – linear codes of length n

$(x_1,x_2,\ldots,x_n) \in A$

$(x_{i_1},x_{i_2},\ldots,x_{i_n}) \in B$

Bipartite-graph codes



$G=(V_0 \cup V_1, E)$ balanced, n-regular, bipartite

$|V_0|=|V_1|=m$, $|E|=N=nm$

$A[n, R_0 n]$
$B[n, R_1 n]$

binary linear codes

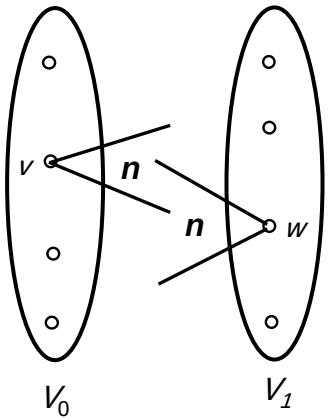
$C\,[N, RN, D]$ = parallel concatenation of A and B:

$$\mathbf{x} \in C \quad \Leftrightarrow \quad \forall_{v \in V_0}\, \mathbf{x}(v) \in A,\ \forall_{w \in V_1}\, \mathbf{x}(w) \in B$$

$R \geq R_0 + R_1 - 1$

Decoding

In the analysis we assume that the local codes are the same: A=B $[n,k,d_0=2t+1]$



Decoding

Left decoding round L ($\boldsymbol{y}$): in parallel decode all subvectors of the N – vector $\boldsymbol{y}$ that appear in the neighborhoods of the left vertices with the code A

Right decoding round R ($\boldsymbol{y}$): in parallel decode all subvectors of the N – vector $\boldsymbol{y}$ that appear in the neighborhoods of the right vertices with the code A

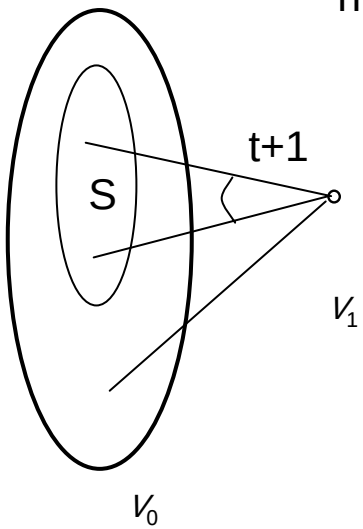
Iterative procedure: $\boldsymbol{y}$ received from the channel. Perform

$$\ldots R\,(L\,(R\,(L\,(R\,(L(\boldsymbol{y}))))))\ \ldots$$

Proceed until a fixed point is encountered, or stop after $O(\log n)$ iterations (Zémor 2001)

Why this works: Property $P_1(\sigma)$. For any subset $S \subset V_0$ (V_1), $|S|=\sigma m$ the number of vertices in V_1 (V_0) whose S-degree exceeds t is

$$\leq(\sigma-\varepsilon)m$$

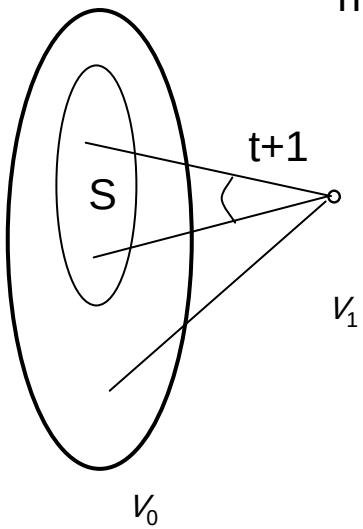


This enables one to prove convergence of decoding

In previous works this property was controlled by the spectral gap of the graph.

However this approach imposes a constraint on $d_0=d(A)$ of the form $d_0 \geq \sqrt{n}$ .Our purpose in this work is to lift this constraint.

Why this works: Property $P_1(\sigma)$. For any subset $S \subset V_0$ (V_1), $|S|=\sigma m$ the number of vertices in V_1 (V_0) whose S-degree exceeds t is $\leq(\sigma-\varepsilon)m$



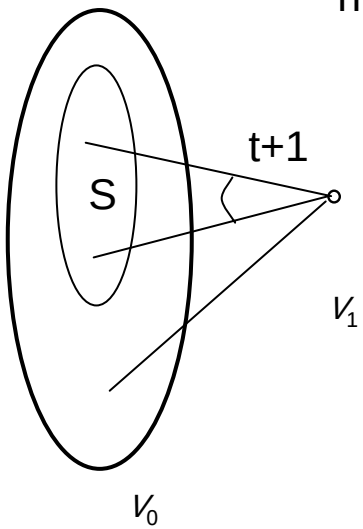
This enables one to prove convergence of decoding

In previous works this property was controlled by the spectral gap of the graph.

However this approach imposes a constraint on $d_0=d(A)$ of the form $d_0 \geq \sqrt{n}$ .Our purpose in this work is to lift this constraint.

Replace graphs with spectral gap with random graphs.
It is known that code ensembles constructed using random regular graph have the same weight distribution as codes obtained from graphs with large spectral gap.

Why this works: Property $P_1(\sigma)$. For any subset $S \subset V_0$ (V_1), $|S|=\sigma m$ the number of vertices in V_1 (V_0) whose S-degree exceeds t is $\leq(\sigma-\varepsilon)m$



This enables one to prove convergence of decoding

In previous works this property was controlled by the spectral gap of the graph.

However this approach imposes a constraint on $d_0=d(A)$ of the form $d_0 \geq \sqrt{n}$ .Our purpose in this work is to lift this constraint.

Replace graphs with spectral gap with random graphs.
It is known that code ensembles constructed using random regular graphs have the same weight distribution as codes obtained from graphs with large spectral gap.

We estimate the guaranteed number of errors in ensembles of codes constructed on random regular bipartite graphs with local codes $A[n,k,d_0]$ for any given d_0.
The estimate is proved by estimating maximum σ for which the ensemble of graphs contains graphs that have Property $P_1(\sigma)$. .

Theorem: Local code A[n,R_0n,d_0=2t+1]. The proportion of correctable errors is at least σ t/n, as defined below

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

$$F_{n,t}(\sigma) = \max_{z \in \mathscr{M}(t,\sigma)} \left(h(z) + \sum_{i=1}^{n} z_i \log \binom{n}{i} \right)$$

where

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

$$\mathcal{Z}_n = \{ z \in [0,1]^{n+1} : \sum_{i=0}^{n} z_i = 1 \}$$

Example: A[23,12,7] binary Golay code. Proportion of errors 0.00063.

What happens when n becomes large, t=τn ?

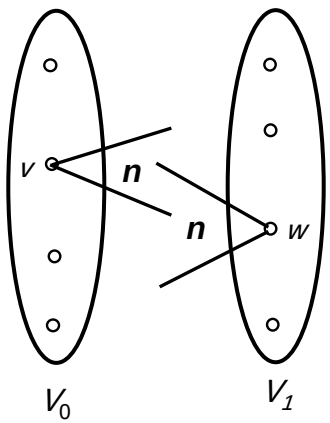
Ensemble $\mathscr{G}(A,m)$ contains codes that correct $\sigma\tau N$ errors for any $\sigma \leq \sigma^*$, where σ^* is given by

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

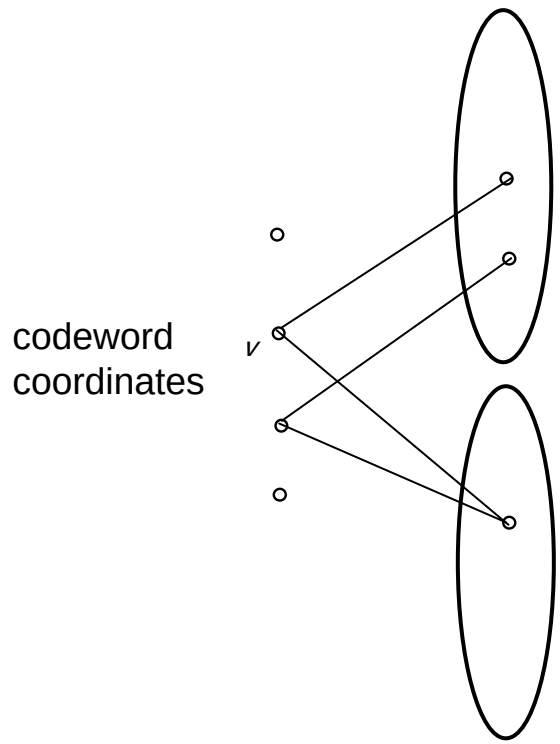
If n→∞, the proportion of correctable errors approaches τ^2

Another view of bipartite-graph codes:

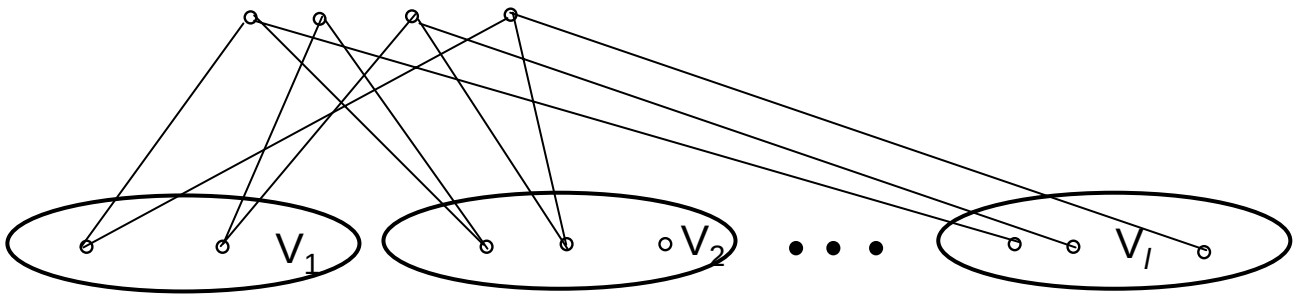
instead of



consider the edge-vertex graph of G:



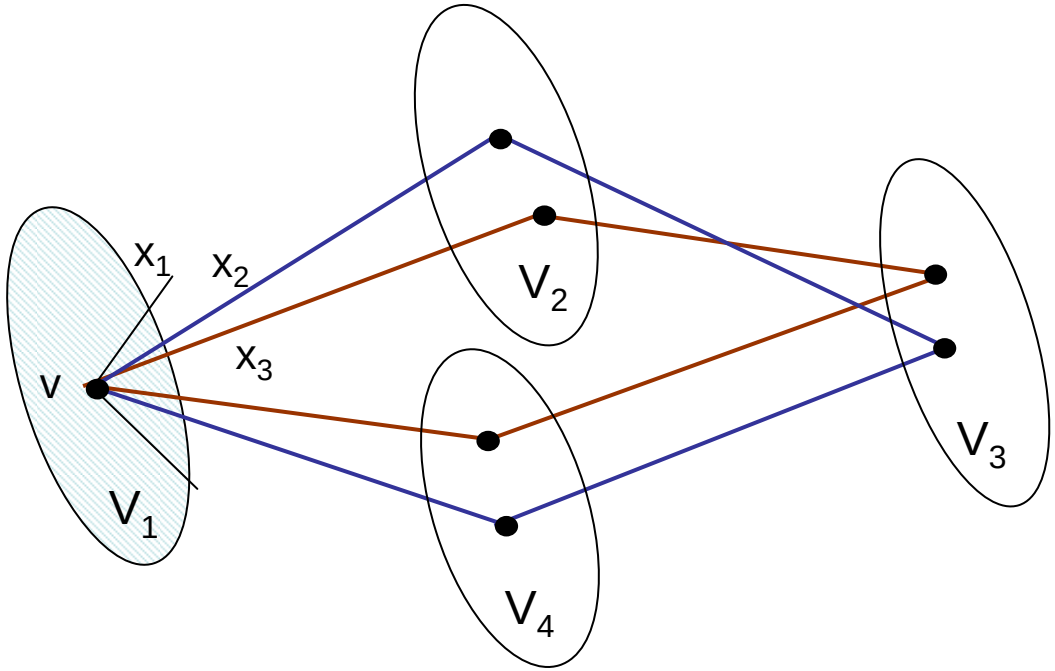
Each variable node is connected to $l \geq 3$ different check vertices



This ensemble was introduced by Gallager; generalized by Bilu and Hoory ’04. under the name “hypergraph codes”

Hypergraph codes

l-partite, n-regular, uniform hypergraph $H=(\{V_1,V_2,...,V_l\},E)$; $|V_1|=...=|V_l|=m$



Bits of the codeword correspond to (hyper)edges
Every vertex imposes a set of local constraints on the bits adjacent to it
$A[n,k=R_0n,d_0]$ linear local code

$$R(C) \geq lR_0-(l-1)$$

Y. Bilu and S.Hoory, On codes from hypergraphs, European J. Comb. 2004, pp. 339-354

A decoding algorithm was suggested in our previous work:
A.B., A. Mazumdar, and G. Zémor, (*Adv. Math. Commun., '08*)

Here we consider the ensemble of codes *H*(A,*l*,m) defined by random regular (hyper)graphs and a local code A[n,R_0n,d_0].

Using the above algorithm and probabilistic analysis, we estimate the proportion of errors correctable with codes from the ensemble *H*(A,*l*,m) as γmn, where

$$\gamma^* = \sup\{x > 0 : \forall_{0<\gamma\leq x}\, (l/n)F_{n,t}(\gamma) < (l-1)h(\gamma)\}.$$

$$\mathscr{M}(t,\gamma) = \Big\{\mathbf{z} \in \mathcal{Z}_n : \sum_{i=1}^{n} iz_i = \gamma n, \sum_{i=1}^{t} iz_i = \sum_{i=d_0-t}^{n} tz_i\Big\}.$$

We compute examples with local Hamming codes.

We also estimate the proportion of correctable errors for large length of the local code A