Erasure Correction and Locality of Hypergraph Codes

Allison Beemer, Carolyn Mayer*, and Christine A. Kelley

University of Nebraska-Lincoln Department of Mathematics

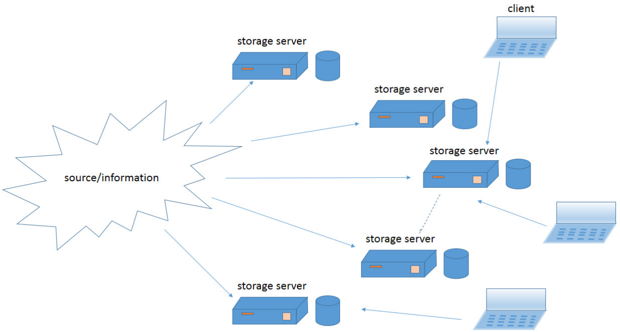
5th International Castle Meeting on Coding Theory and Applications

August 28, 2017

Outline

- Cooperative Locality and Availability in Distributed Storage Systems
- Low Density Parity-Check Codes
- Hypergraph Codes
- Bounds on Regular and Biregular Hypergraph Codes
- Conclusions and Future Work

Coding for Distributed Storage Systems (DSS)



Goal: Store large amounts of data across many servers so that multiple users can access the data reliably and efficiently.

Cooperative Locality and Availability

Cooperative Locality and Availability

- A code C has (r, ℓ)-cooperative locality if for any $\mathbf{y} \in C$, any set of ℓ symbols in $\mathbf{y}$ are functions of at most r other symbols [Rawat, Mazumdar, Vishwanath '16].

Cooperative Locality and Availability

- A code C has (r, ℓ)-cooperative locality if for any $\mathbf{y} \in C$, any set of ℓ symbols in $\mathbf{y}$ are functions of at most r other symbols [Rawat, Mazumdar, Vishwanath '16].

- A code C has (r, τ)-availability if any symbol in $\mathbf{y}$ can be recovered using any of τ disjoint sets of symbols each of size at most r [Rawat, Papailiopoulos, Dimakis, Vishwanath '16].

Low Density Parity-Check (LDPC) Codes

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

$\mathbf{v} = (v_1, v_2, v_3, v_4)$ is a codeword in $\mathcal{C}$ if and only if $Hv^T = 0$.

$$v_1 + v_2 = 0$$
$$v_1 + v_3 = 0$$
$$v_2 + v_3 + v_4 = 0.$$

Low Density Parity-Check (LDPC) Codes

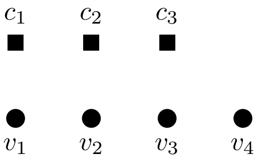
$$H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

$\mathbf{v} = (v_1, v_2, v_3, v_4)$ is a codeword in $\mathcal{C}$ if and only if $Hv^T = 0$.

$$v_1 + v_2 = 0$$
$$v_1 + v_3 = 0$$
$$v_2 + v_3 + v_4 = 0.$$

Low Density Parity-Check (LDPC) Codes

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

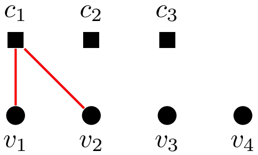


$\mathbf{v} = (v_1, v_2, v_3, v_4)$ is a codeword in $\mathcal{C}$ if and only if $Hv^T = 0$.

$$v_1 + v_2 = 0$$
$$v_1 + v_3 = 0$$
$$v_2 + v_3 + v_4 = 0.$$

Low Density Parity-Check (LDPC) Codes

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

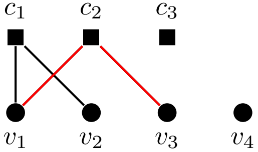


$\mathbf{v} = (v_1, v_2, v_3, v_4)$ is a codeword in $\mathcal{C}$ if and only if $Hv^T = 0$.

$$v_1 + v_2 = 0$$
$$v_1 + v_3 = 0$$
$$v_2 + v_3 + v_4 = 0.$$

Low Density Parity-Check (LDPC) Codes

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

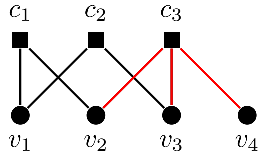


$\mathbf{v} = (v_1, v_2, v_3, v_4)$ is a codeword in $\mathcal{C}$ if and only if $Hv^T = 0$.

$$v_1 + v_2 = 0$$
$$v_1 + v_3 = 0$$
$$v_2 + v_3 + v_4 = 0.$$

Low Density Parity-Check (LDPC) Codes

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

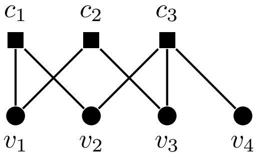


$\mathbf{v} = (v_1, v_2, v_3, v_4)$ is a codeword in $\mathcal{C}$ if and only if $Hv^T = 0$.

$$v_1 + v_2 = 0$$
$$v_1 + v_3 = 0$$
$$v_2 + v_3 + v_4 = 0.$$

Low Density Parity-Check (LDPC) Codes

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$



$\mathbf{v} = (v_1, v_2, v_3, v_4)$ is a codeword in $\mathcal{C}$ if and only if $Hv^T = 0$.

$$v_1 + v_2 = 0$$
$$v_1 + v_3 = 0$$
$$v_2 + v_3 + v_4 = 0.$$

Generalized LDPC codes have more sophisticated constraint nodes.

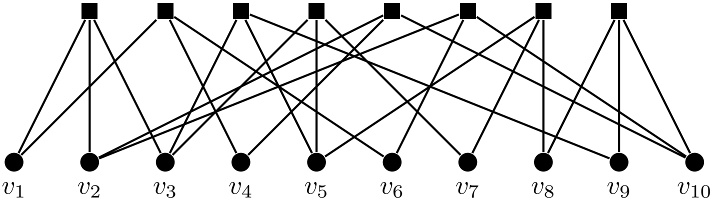
Stopping Sets for Generalized LDPC codes

A stopping set S in the Tanner graph of a generalized LDPC code is a subset of variable nodes such that each neighbor of S connects to S at least $d_{\min}(\text{subcode})$ times.

Stopping Sets for Generalized LDPC codes

A stopping set S in the Tanner graph of a generalized LDPC code is a subset of variable nodes such that each neighbor of S connects to S at least $d_{\min}(\text{subcode})$ times.

Example:

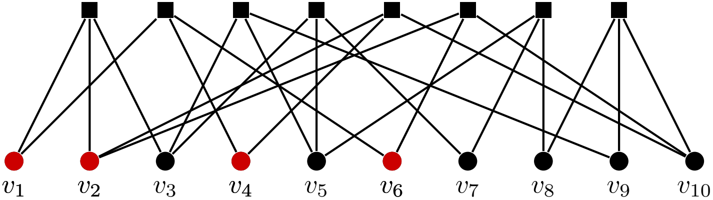


If the constraint nodes represent a subcode C with $d_{\min}(C) = 2$, then $S = \{v_1, v_2, v_4, v_5\}$ is a stopping set.

Stopping Sets for Generalized LDPC codes

A stopping set S in the Tanner graph of a generalized LDPC code is a subset of variable nodes such that each neighbor of S connects to S at least $d_{\min}$(subcode) times.

Example:

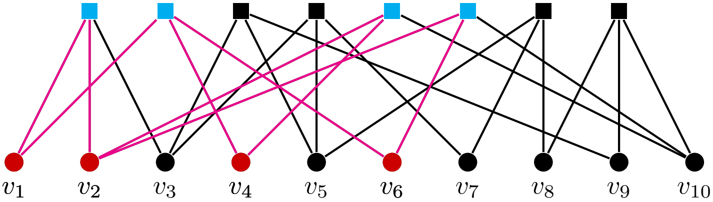


If the constraint nodes represent a subcode C with $d_{\min}(C) = 2$, then $S = \{v_1, v_2, v_4, v_5\}$ is a stopping set.

Stopping Sets for Generalized LDPC codes

A stopping set S in the Tanner graph of a generalized LDPC code is a subset of variable nodes such that each neighbor of S connects to S at least $d_{\min}(\text{subcode})$ times.

Example:

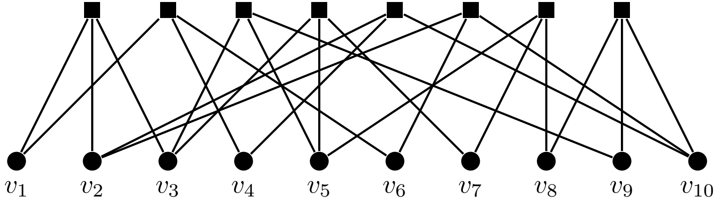


If the constraint nodes represent a subcode C with $d_{\min}(C) = 2$, then $S = \{v_1, v_2, v_4, v_5\}$ is a stopping set.

Stopping Sets for Generalized LDPC codes

A stopping set S in the Tanner graph of a generalized LDPC code is a subset of variable nodes such that each neighbor of S connects to S at least $d_{\min}$(subcode) times.

Example:



If the constraint nodes represent a subcode C with $d_{\min}(C) = 2$, then $S = \{v_1, v_2, v_4, v_5\}$ is a stopping set.

$\ell \leq s_{\min} - 1$, where $s_{\min}$ is the size of the smallest stopping set.

Expander codes

Expander graphs are graphs in which small sets of vertices have large sets of neighbors.

Expander codes

Expander graphs are graphs in which small sets of vertices have large sets of neighbors.

LDPC codes have improved error correction capabilities when the underlying graph is an expander [Sipser, Spielman '96].

Expander codes

Expander graphs are graphs in which small sets of vertices have large sets of neighbors.

LDPC codes have improved error correction capabilities when the underlying graph is an expander [Sipser, Spielman ‘96].

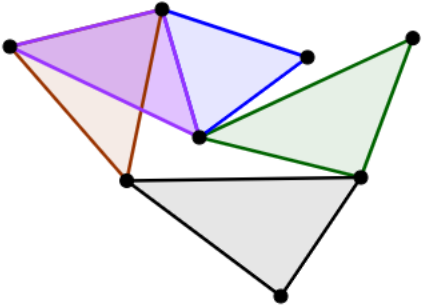
An analogous notion was introduced for hypergraphs [Bilu, Hoory ‘04].

Hypergraphs

A hypergraph $\mathcal{H}$ is a set of vertices V along with a set of edges E, where E is a set of subsets of V.

Hypergraphs

A hypergraph $\mathcal{H}$ is a set of vertices V along with a set of edges E, where E is a set of subsets of V.



Hypergraphs

A hypergraph $\mathcal{H}$ is a set of vertices V along with a set of edges E, where E is a set of subsets of V.

A hypergraph is:

Hypergraphs

A hypergraph $\mathcal{H}$ is a set of vertices V along with a set of edges E, where E is a set of subsets of V.

A hypergraph is:

- t-uniform if every edge contains exactly t vertices.

Hypergraphs

A hypergraph $\mathcal{H}$ is a set of vertices V along with a set of edges E, where E is a set of subsets of V.

A hypergraph is:

- t-uniform if every edge contains exactly t vertices.
- t-partite if $V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_t$ such that no edge contains more than one vertex from any part.

Hypergraphs

A hypergraph $\mathcal{H}$ is a set of vertices V along with a set of edges E, where E is a set of subsets of V.

A hypergraph is:

- t-uniform if every edge contains exactly t vertices.
- t-partite if $V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_t$ such that no edge contains more than one vertex from any part.
- Δ-regular if every vertex belongs to Δ edges.

Hypergraphs

A hypergraph $\mathcal{H}$ is a set of vertices V along with a set of edges E, where E is a set of subsets of V.

A hypergraph is:

- t-uniform if every edge contains exactly t vertices.
- t-partite if $V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_t$ such that no edge contains more than one vertex from any part.
- Δ-regular if every vertex belongs to Δ edges.

$\mathcal{H} = (V_1, V_2, \ldots, V_t; E)$ denotes a t-uniform, t-partite hypergraph in which each edge uses exactly one vertex from each part.

Expansion in Hypergraphs

Let $\mathcal{H} = (V_1, V_2, \ldots, V_t; E)$ be a t-uniform, t-partite, Δ-regular hypergraph with n vertices in each part.

[Bilu, Hoory '04]

Expansion in Hypergraphs

Let $\mathcal{H} = (V_1, V_2, \ldots, V_t; E)$ be a t-uniform, t-partite, Δ-regular hypergraph with n vertices in each part.

$\mathcal{H}$ is ϵ-homogeneous if for every choice of $A_1, A_2, \ldots, A_t$ with $A_i \subseteq V_i$ and $|A_i| = \alpha_i n$,

$$\frac{|E(A_1, A_2, \ldots, A_t)|}{n\Delta} \leq \prod_{i=1}^{t} \alpha_i + \epsilon\sqrt{\alpha_{\sigma(1)}\alpha_{\sigma(2)}},$$

where σ is a permutation on $[t]$ such that $\alpha_{\sigma(i)} \leq \alpha_{\sigma(i+1)}$ for each $i \in [t-1]$ and $E(A_1, \ldots, A_t)$ is the set of edges intersecting all the A_i's.

[Bilu, Hoory '04]

Hypergraph Codes

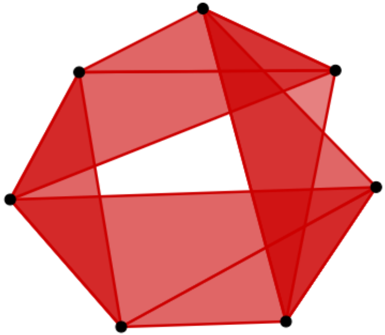
t-uniform, t-partite Δ-regular hypergraphs can be used to design codes:

[Bilu, Hoory '04],[Barg, Zemor '08]

Hypergraph Codes

t-uniform, t-partite Δ-regular hypergraphs can be used to design codes:

- Edges represent code symbols.

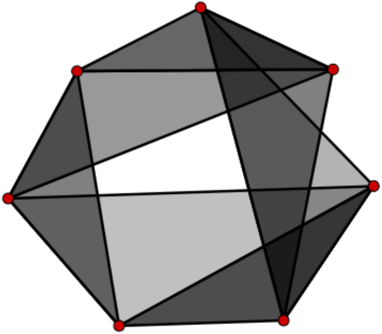


[Bilu, Hoory '04],[Barg, Zemor '08]

Hypergraph Codes

t-uniform, t-partite Δ-regular hypergraphs can be used to design codes:

- Edges represent code symbols.
- Vertices represent constraints.



[Bilu, Hoory ‘04],[Barg, Zemor ‘08]

Hypergraph Code Rate and Minimum Distance

A code $\mathcal{Z}$ from an ϵ-homogeneous t-uniform t-partite Δ-regular hypergraph with n vertices in each part and a $[\Delta, \Delta R, \Delta\delta]$ subcode C at each constraint node has:

$$\text{rate}(\mathcal{Z}) \geq tR - (t-1)$$
$$d_{\min}(\mathcal{Z}) \geq n\Delta\left(\delta^{\frac{t}{t-1}} - c(\epsilon, \delta, t)\right)$$

where $c(\epsilon, \delta, t) \to 0$ as $\epsilon \to 0$.

[Bilu, Hoory '04]

Stopping Sets for Hypergraph Codes

Let $\mathcal{Z}$ be a code on $\mathcal{H} = (V_1, \ldots, V_t; E)$ with subcode C at each vertex. Then a stopping set S is a subset of the edges of $\mathcal{H}$ such that every vertex contained in an element of S is contained in at least $d_{\min}(C)$ elements of S.

Stopping Sets for Hypergraph Codes

Let $\mathcal{Z}$ be a code on $\mathcal{H} = (V_1, \ldots, V_t; E)$ with subcode C at each vertex. Then a stopping set S is a subset of the edges of $\mathcal{H}$ such that every vertex contained in an element of S is contained in at least $d_{\min}(C)$ elements of S.

The size of a minimum stopping set for a code from hypergraph $\mathcal{H}$ with subcode C is denoted $s_{\min}(\mathcal{H})$.

Bounds on Regular Hypergraphs

Theorem

Let $\mathcal{H}$ be a t-uniform t-partite Δ-regular hypergraph. If the vertices of $\mathcal{H}$ represent constraints of a subcode C with minimum distance $d_{\min}(C)$ and block length Δ, then the size of the minimum stopping set, $s_{\min}(\mathcal{H})$, is bounded by

$$s_{\min}(\mathcal{H}) \geq d_{\min}(C)^{t/(t-1)}.$$

Bounds on Regular Hypergraphs

Proof Sketch:

Bounds on Regular Hypergraphs

Proof Sketch:

- Each edge in S contains exactly one constraint node from each part, so each part of $\mathcal{H}$ has $|S| = s_{\min}(\mathcal{H})$ incident edges in S.

Bounds on Regular Hypergraphs

Proof Sketch:

- Each edge in S contains exactly one constraint node from each part, so each part of $\mathcal{H}$ has $|S| = s_{\min}(\mathcal{H})$ incident edges in S.
- Each vertex in an edge in S must be in at least $d_{\min}(C)$ edges in S.

Bounds on Regular Hypergraphs

Proof Sketch:

- Each edge in S contains exactly one constraint node from each part, so each part of $\mathcal{H}$ has $|S| = s_{\min}(\mathcal{H})$ incident edges in S.
- Each vertex in an edge in S must be in at least $d_{\min}(C)$ edges in S.
- No part can have more than $s_{\min}(\mathcal{H})/d_{\min}(C)$ vertices incident to S in a single part, else some vertex must have fewer than $s_{\min}(\mathcal{H})/(s_{\min}(\mathcal{H})/d_{\min}(C)) = d_{\min}(C)$ incident edges from S

Bounds on Regular Hypergraphs

Proof Sketch:

- Each edge in S contains exactly one constraint node from each part, so each part of $\mathcal{H}$ has $|S| = s_{\min}(\mathcal{H})$ incident edges in S.
- Each vertex in an edge in S must be in at least $d_{\min}(C)$ edges in S.
- No part can have more than $s_{\min}(\mathcal{H})/d_{\min}(C)$ vertices incident to S in a single part, else some vertex must have fewer than $s_{\min}(\mathcal{H})/(s_{\min}(\mathcal{H})/d_{\min}(C)) = d_{\min}(C)$ incident edges from S
- There are at most $(s_{\min}(\mathcal{H})/d_{\min}(C))^t$ edges in S.

Bounds on Regular Hypergraphs

Proof Sketch:

- Each edge in S contains exactly one constraint node from each part, so each part of $\mathcal{H}$ has $|S| = s_{\min}(\mathcal{H})$ incident edges in S.
- Each vertex in an edge in S must be in at least $d_{\min}(C)$ edges in S.
- No part can have more than $s_{\min}(\mathcal{H})/d_{\min}(C)$ vertices incident to S in a single part, else some vertex must have fewer than $s_{\min}(\mathcal{H})/(s_{\min}(\mathcal{H})/d_{\min}(C)) = d_{\min}(C)$ incident edges from S
- There are at most $(s_{\min}(\mathcal{H})/d_{\min}(C))^t$ edges in S.

$$\left(\frac{s_{\min}(\mathcal{H})}{d_{\min}(C)}\right)^t \geq s_{\min}(\mathcal{H})$$

Bounds on Regular Hypergraphs

Proof Sketch:

- Each edge in S contains exactly one constraint node from each part, so each part of $\mathcal{H}$ has $|S| = s_{\min}(\mathcal{H})$ incident edges in S.
- Each vertex in an edge in S must be in at least $d_{\min}(C)$ edges in S.
- No part can have more than $s_{\min}(\mathcal{H})/d_{\min}(C)$ vertices incident to S in a single part, else some vertex must have fewer than $s_{\min}(\mathcal{H})/(s_{\min}(\mathcal{H})/d_{\min}(C)) = d_{\min}(C)$ incident edges from S
- There are at most $(s_{\min}(\mathcal{H})/d_{\min}(C))^t$ edges in S.

$$\left(\frac{s_{\min}(\mathcal{H})}{d_{\min}(C)}\right)^t \geq s_{\min}(\mathcal{H}) \Rightarrow s_{\min}(\mathcal{H}) \geq d_{\min}(C)^{t/(t-1)}.$$

Bounds on Regular Hypergraphs

Proof Sketch:

- Each edge in S contains exactly one constraint node from each part, so each part of $\mathcal{H}$ has $|S| = s_{\min}(\mathcal{H})$ incident edges in S.
- Each vertex in an edge in S must be in at least $d_{\min}(C)$ edges in S.
- No part can have more than $s_{\min}(\mathcal{H})/d_{\min}(C)$ vertices incident to S in a single part, else some vertex must have fewer than $s_{\min}(\mathcal{H})/(s_{\min}(\mathcal{H})/d_{\min}(C)) = d_{\min}(C)$ incident edges from S
- There are at most $(s_{\min}(\mathcal{H})/d_{\min}(C))^t$ edges in S.

$$\left(\frac{s_{\min}(\mathcal{H})}{d_{\min}(C)}\right)^t \geq s_{\min}(\mathcal{H}) \Rightarrow s_{\min}(\mathcal{H}) \geq d_{\min}(C)^{t/(t-1)}.$$

□

Hypergraph Code Locality and Availability

Corollary

If the subcodes C of the regular hypergraph code $\mathcal{Z}$ have r_1 locality, then $\mathcal{Z}$ has (r, ℓ)-cooperative locality where

$$r = r_1 t s_{\min}(\mathcal{H}) / d_{\min}(C)$$
$$s_{\min}(\mathcal{H}) - 1 \geq \ell \geq d_{\min}(C)^{t/(t-1)} - 1.$$

Hypergraph Code Locality and Availability

Corollary

If the subcodes C of the regular hypergraph code $\mathcal{Z}$ have r_1 locality, then $\mathcal{Z}$ has (r, ℓ)-cooperative locality where

$$r = r_1 t s_{\min}(\mathcal{H})/d_{\min}(C)$$
$$s_{\min}(\mathcal{H}) - 1 \geq \ell \geq d_{\min}(C)^{t/(t-1)} - 1.$$

If C has (r, τ)-availability, then $\mathcal{Z}$ has at least (r, τ)-availability.

Bound on ϵ-homogeneous Regular Hypergraphs

Theorem

Let $\mathcal{H} = (V_1, V_2, \ldots, V_t; E)$ be a t-uniform t-partite Δ-regular ϵ-homogeneous hypergraph with n vertices in each part, each with subcode C. Then

$$s_{\min}(\mathcal{H}) \geq \left(\left(1 - \frac{\epsilon \Delta}{d_{\min}(C)}\right) \frac{n^{t-1} d_{\min}(C)^t}{\Delta} \right)^{1/(t-1)} .$$

Bound on ϵ-homogeneous Regular Hypergraphs

Theorem

Let $\mathcal{H} = (V_1, V_2, \ldots, V_t; E)$ be a t-uniform t-partite Δ-regular ϵ-homogeneous hypergraph with n vertices in each part, each with subcode C. Then

$$s_{\min}(\mathcal{H}) \geq \left(\left(1 - \frac{\epsilon \Delta}{d_{\min}(C)}\right) \frac{n^{t-1} d_{\min}(C)^t}{\Delta} \right)^{1/(t-1)} .$$

For $\epsilon < \frac{d_{\min}(C)(n^{t-1} - \Delta)}{\Delta n^{t-1}}$, this is an improvement on the previous bound.

Bound on ϵ-homogeneous Regular Hypergraphs

Theorem

Let $\mathcal{H} = (V_1, V_2, \ldots, V_t; E)$ be a t-uniform t-partite Δ-regular ϵ-homogeneous hypergraph with n vertices in each part, each with subcode C. Then

$$s_{\min}(\mathcal{H}) \geq \left(\left(1 - \frac{\epsilon\Delta}{d_{\min}(C)}\right)\frac{n^{t-1}d_{\min}(C)^t}{\Delta}\right)^{1/(t-1)}.$$

For $\epsilon < \frac{d_{\min}(C)(n^{t-1}-\Delta)}{\Delta n^{t-1}}$, this is an improvement on the previous bound.

Corollary

Iterative decoding of code $\mathcal{Z}$ based on a $\mathcal{H}$ corrects up to

$$\left(\left(1 - \frac{\epsilon\Delta}{d_{\min}(C)}\right)\frac{n^{t-1}d_{\min}(C)^t}{\Delta}\right)^{1/(t-1)} - 1$$

erasures.

Biregular Hypergraphs

- A t-uniform t-partite hypergraph $\mathcal{H} = (V_1, \ldots, V_t; E)$ is (Δ_1, Δ_2)-biregular if the parts can be labeled such that each vertex in an odd (resp., even) part is contained in Δ_1 (resp., Δ_2) edges.

Biregular Hypergraphs

- A t-uniform t-partite hypergraph $\mathcal{H} = (V_1, \ldots, V_t; E)$ is (Δ_1, Δ_2)-biregular if the parts can be labeled such that each vertex in an odd (resp., even) part is contained in Δ_1 (resp., Δ_2) edges.
- A stopping set S for a code on $\mathcal{H}$ with subcode C_1 (resp., C_2) in odd (resp., even) index parts is a subset of the edges of $\mathcal{H}$ such that every vertex in an edge in S is in at least $d_{\min}(C_1)$ (resp., $d_{\min}(C_2)$) elements of S if the vertex is in an odd (resp., even) part.

Biregular Hypergraphs

- A t-uniform t-partite hypergraph $\mathcal{H} = (V_1, \ldots, V_t; E)$ is (Δ_1, Δ_2)-biregular if the parts can be labeled such that each vertex in an odd (resp., even) part is contained in Δ_1 (resp., Δ_2) edges.

- A stopping set S for a code on $\mathcal{H}$ with subcode C_1 (resp., C_2) in odd (resp., even) index parts is a subset of the edges of $\mathcal{H}$ such that every vertex in an edge in S is in at least $d_{\min}(C_1)$ (resp., $d_{\min}(C_2)$) elements of S if the vertex is in an odd (resp., even) part.

- $\mathcal{H}$ is ϵ-homogeneous if for every $\{A_1, A_2, \ldots, A_t\}$, with $A_i \subseteq V_i$,

$$\frac{|E(A_1, A_2, \ldots A_t)|}{\Delta_1 n_1} \leq \prod_{i=1}^{t} \alpha_i + \epsilon\sqrt{\alpha_{\sigma(1)}\alpha_{\sigma(2)}},$$

  where n_1 is the number of vertices in each odd part and α_i is the fraction of vertices of V_i in A_i.

Biregular Hypergraph Code Rate and Distance

An $n_1\Delta_1$ code $\mathcal{Z}$ from an ϵ-homogeneous t-uniform t-partite (Δ_1, Δ_2)-regular hypergraph with n_1 (resp., n_2) vertices in each odd (resp., even) index part and $[\Delta_1, \Delta_1 R_1, \Delta_1 \delta_1]$ subcodes C_1 (resp., $[\Delta_2, \Delta_2 R_2, \Delta_2 \delta_2]$ subcodes C_2) has:

$$\text{rate}(\mathcal{Z}) \geq R_1 \lceil \frac{t}{2} \rceil + R_2 \lfloor \frac{t}{2} \rfloor - (t-1) \quad (1)$$

$$d_{\min}(\mathcal{Z}) \geq n_1 \Delta_1 \left((\delta_1^{\lceil \frac{t}{2} \rceil} \delta_2^{\lfloor \frac{t}{2} \rfloor})^{\frac{1}{t-1}} - c(\epsilon, \delta_1, \delta_2, t) \right) \quad (2)$$

where $c(\epsilon, \delta_1, \delta_2, t) \to 0$ as $\epsilon \to 0$.

Bounds on Biregular Hypergraphs

Theorem

Let $\mathcal{H}$ be a t-uniform t-partite (Δ_1, Δ_2)-biregular hypergraph. If the vertices in an odd (resp., even) index part of $\mathcal{H}$ represent constraints of a subcode C_1 (resp., C_2) with block length Δ_1 (resp., Δ_2), then the size of the minimum stopping set, $s_{\min}(\mathcal{H})$, is bounded by

$$s_{\min}(\mathcal{H}) \geq \left(d_{\min}(C_1)^{\lceil \frac{t}{2} \rceil} d_{\min}(C_2)^{\lfloor \frac{t}{2} \rfloor} \right)^{1/(t-1)} .$$

Bounds on Biregular Hypergraphs

Theorem

Let $\mathcal{H}$ be a t-uniform t-partite (Δ_1, Δ_2)-biregular hypergraph. If the vertices in an odd (resp., even) index part of $\mathcal{H}$ represent constraints of a subcode C_1 (resp., C_2) with block length Δ_1 (resp., Δ_2), then the size of the minimum stopping set, $s_{\min}(\mathcal{H})$, is bounded by

$$s_{\min}(\mathcal{H}) \geq \left(d_{\min}(C_1)^{\lceil \frac{t}{2} \rceil} d_{\min}(C_2)^{\lfloor \frac{t}{2} \rfloor}\right)^{1/(t-1)}.$$

Corollary

If the subcodes C_1 (resp., C_2) of the biregular hypergraph code $\mathcal{Z}$ have r_1 (resp., r_2) locality then $\mathcal{Z}$ has (r, ℓ)-cooperative locality where

$$r = r_1 \lceil \frac{t}{2} \rceil \frac{s_{\min}(\mathcal{H})}{d_{\min}(C_1)} + r_2 \lfloor \frac{t}{2} \rfloor \frac{s_{\min}(\mathcal{H})}{d_{\min}(C_2)}$$

$$s_{\min}(\mathcal{H}) - 1 \geq \ell \geq \left(d_{\min}(C_1)^{\lceil \frac{t}{2} \rceil} d_{\min}(C_2)^{\lfloor \frac{t}{2} \rfloor}\right)^{1/(t-1)} - 1.$$

Bound on ϵ-homogeneous Biregular Hypergraphs

Theorem

Let $\mathcal{H} = (V_1, \ldots, V_t; E)$ be a t-uniform t-partite (Δ_1, Δ_2)-regular ϵ-homogeneous hypergraph where there are n_1 (resp., n_2) vertices in each of the odd (resp., even) index parts. Let C_1 and C_2 be the subcodes of the odd and even index parts, respectively. Then

$$s_{\min}(\mathcal{H}) \geq \left(\frac{(n_1 d_{\min}(C_1))^{\lceil \frac{t}{2} \rceil} (n_2 d_{\min}(C_2))^{\lfloor \frac{t}{2} \rfloor}}{n_1 \Delta_1} \left(1 - \frac{\epsilon n_1 \Delta_1}{\min_{i=1,2}\{n_i d_{\min}(C_i)\}} \right) \right)^{\frac{1}{t-1}}.$$

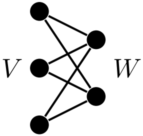
For $\epsilon < \left(1 - \frac{n_1 \Delta_1}{n_1^{\lceil \frac{t}{2} \rceil} n_2^{\lfloor \frac{t}{2} \rfloor}}\right) \frac{\min_{i=1,2}\{n_i d_{\min}(C_i)\}}{n_1 \Delta_1}$, this in an improvement on the previous bound.

Biregular Hypergraph Construction

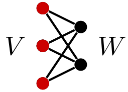
Let $G = V \cup W$ be a (c, d)-regular bipartite graph with $|V| \geq |W|$.

Biregular Hypergraph Construction

Let $G = V \cup W$ be a (c, d)-regular bipartite graph with $|V| \geq |W|$.

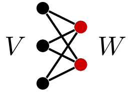


Biregular Hypergraph Construction

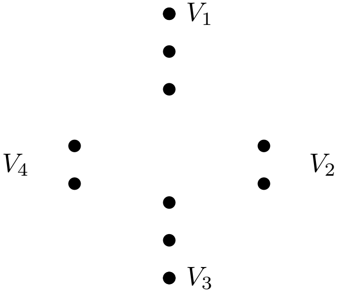


Let $G = V \cup W$ be a (c, d)-regular bipartite graph with $|V| \geq |W|$.
For odd i, let V_i be a copy of V.

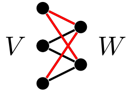
Biregular Hypergraph Construction



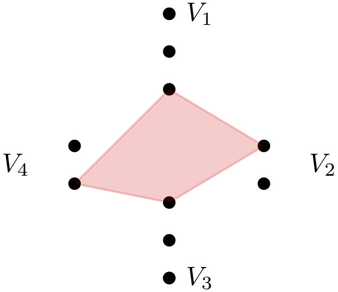
Let $G = V \cup W$ be a (c, d)-regular bipartite graph with $|V| \geq |W|$.
For odd i, let V_i be a copy of V. For even i, let V_i be a copy of W.



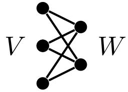
Biregular Hypergraph Construction



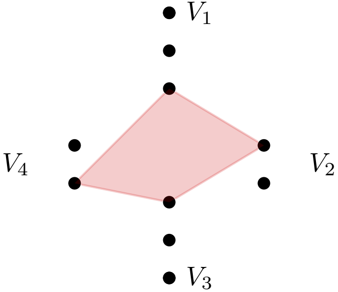
Let $G = V \cup W$ be a (c, d)-regular bipartite graph with $|V| \geq |W|$.
For odd i, let V_i be a copy of V. For even i, let V_i be a copy of W.
Take $E(\mathcal{H})$ to be the set of edges corresponding to walks of length $t - 1$.



Biregular Hypergraph Construction



Let $G = V \cup W$ be a (c, d)-regular bipartite graph with $|V| \geq |W|$.
For odd i, let V_i be a copy of V. For even i, let V_i be a copy of W.
Take $E(\mathcal{H})$ to be the set of edges corresponding to walks of length $t - 1$.



Idea: If G is a (c, d)-regular expander, then $\mathcal{H}$ should have ϵ-homogeneity.

Conclusions and Future Work

We showed:

- Bounds on minimum stopping set size, cooperative locality of regular hypergraph codes.
- Bounds on minimum stopping set size, cooperative locality of biregular hypergraph codes.
- Improved bounds for ϵ-homogeneous hypergraph codes.

We are interested in:

- Constructing explicit families of hypergraph codes that optimize locality properties.
- Showing the existence of a family of ϵ-homogeneous biregular hypergraphs.

Thank you!

References

- Y. Ishai, E. Kushilevitz, R. Ostrovsky and A. Sahai, Batch codes and their applications, *ACM Symp. on Thy. of Comp.*, pp. 262-271, June 2004, Chicago, IL.
- A.G. Dimakis, K. Ramchandran, Y. Wu and C. Suh, A Survey on Network Codes for Distributed Storage,*Proc. of IEEE*, vol. 99, no. 3, Mar. 2011.
- G.M. Kamath, N. Prakash, V. Lalitha and P.V. Kumar, Codes with local regeneration, *IEEE Info. Thy. and App. Workshop*, pp. 1-5, Feb. 2013, San Diego, CA.
- T. Ernvall, T. Westerbäck, R. Freij-Hollanti, and C. Hollanti. A Connection Between Locally Repairable Codes and Exact Regenerating Codes. Proceedings of IEEE Intl. Symp. on Info. Thy. Barcelona, July 2016.
- V. Skachek. Batch and PIR Codes and Their Connections to Locally Repairable Codes. axXiv:1611.09914v2, March 2017.
- Y. Bilu and S. Hoory, On codes from hypergraphs, *European Journal of Combinatorics*, **25** (2004), 339–354.
- A. Barg and G. Zémor, Codes on hypergraphs, *IEEE Intl. Symp. on Info. Thy.*, pp. 156 - 160, July 2008, Toronto, ON.
- M. Sipser and D. A. Spielman, Expander codes, *IEEE Trans. on Info. Thy.*, **42** (1996), 1710–1722.
- H.L. Janwa and A. K. Lal, On expander graphs: parameters and applications, preprint, arXivcs/0406048.
- G. Zémor, On expander codes, *IEEE Trans. on Info.Theory*, **47** (2001), no. 2, pp. 835–837.
- R. M. Tanner, A recursive approach to low complexity codes, *IEEE Trans. on Info. Thy.*, **27** (1981), 533–547.
- H. Janwa and A. K. Lal, On Tanner codes: minimum distance and decoding, *Proceedings of AAECC*, 13 (2003), pp. 335 - 347.
- A.S. Rawat, A. Mazumdar and S. Vishwanath, Cooperative Local Repair in Distributed Storage, arXiv1409.3900.
- A.S. Rawat, D.S. Papailiopoulos, A.G. Dimakis and S. Vishwanath, Locality and availability in distributed storage, IEEE Trans. on Info. Thy., Vol. 62, no. 8, pp. 4481-4493. Aug. 2016.
- C.A. Kelley and D. Sridhara, Eigenvalue bounds on the pseudocodeword weight of expander codes, *Advances in Mathematics of Communications*, **1** (2007), 287 – 307.
- N. Miladinovic and M. P. C. Fossorier, Generalized LDPC codes and generalized stopping sets, *IEEE Trans. on Communications*, **56** (2008), 201 – 212.
- C. Di, D. Proietti, T. Richardson, E. Teletar and R. Urbanke, Finite-length analysis of low-density parity-check codes on the binary erasure channel, *IEEE Trans. on Info. Thy.*, **48** (2002), 1570 – 1579.
- A. Lubotzky, R. Phillips and P. Sarnak, Ramanujan graphs, *Combinatorica*, **8** (1988), 261–277.
- A. W. Marcus, D. A. Spielman and N. Srivastava, Interlacing families I: bipartite Ramanujan graphs of all degrees, *Annals of Mathematics*, **182** (2015), 307–325.