# Minimum Distance Bounds for Expander Codes

Vitaly Skachek
Claude Shannon Institute
University College Dublin

*Open Problems Session*

Information Theory and Applications Workshop
UCSD

January 28, 2008

# Basic Definitions

**Definition**

*Code* $\mathcal{C}$ is a set of words of length $n$ over an alphabet $\Sigma$.

# Basic Definitions

**Definition**

*Code $\mathcal{C}$* is a set of words of length $n$ over an alphabet $\Sigma$.

**Definition**

- The *Hamming distance* between $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ in $\Sigma^n$, $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y})$, is the number of pairs of symbols $(x_i, y_i)$, $1 \leq i \leq n$, such that $x_i \neq y_i$.

# Basic Definitions

**Definition**

*Code* $\mathcal{C}$ is a set of words of length $n$ over an alphabet $\Sigma$.

**Definition**

- The *Hamming distance* between $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ in $\Sigma^n$, $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y})$, is the number of pairs of symbols $(x_i, y_i)$, $1 \le i \le n$, such that $x_i \neq y_i$.
- The *minimum distance* of a code $\mathcal{C}$ is

$$d = \min_{\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C}, \boldsymbol{x} \neq \boldsymbol{y}} \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}).$$

# Basic Definitions

**Definition**

*Code* $\mathcal{C}$ is a set of words of length $n$ over an alphabet $\Sigma$.

**Definition**

- The *Hamming distance* between $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ in $\Sigma^n$, $\mathsf{d}(\boldsymbol{x}, \boldsymbol{y})$, is the number of pairs of symbols $(x_i, y_i)$, $1 \le i \le n$, such that $x_i \neq y_i$.
- The *minimum distance* of a code $\mathcal{C}$ is

$$d = \min_{\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C}, \boldsymbol{x} \neq \boldsymbol{y}} \mathsf{d}(\boldsymbol{x}, \boldsymbol{y}).$$

- The *relative minimum distance* of $\mathcal{C}$ is defined as $\delta = d/n$.

# Linear Code

**Definition**

- A code $\mathcal{C}$ over field $\mathbb{F} = \mathrm{GF}(q)$ is said to be a *linear* $[n, k, d]$ *code* if there exists a matrix $\mathcal{H}$ with $n$ columns and rank $n - k$ such that
$$\mathcal{H}\boldsymbol{x}^t = \bar{\mathbf{0}} \;\Leftrightarrow\; \boldsymbol{x} \in \mathcal{C}.$$
- The matrix $\mathcal{H}$ is a *parity-check matrix*.
- The value $k$ is the *dimension* of the code $\mathcal{C}$.
- The ratio $r = k/n$ is the *rate* of the code $\mathcal{C}$.

**Definition**

- A code $\mathcal{C}$ over field $\mathbb{F} = \mathrm{GF}(q)$ is said to be a *linear $[n, k, d]$ code* if there exists a matrix $\mathcal{H}$ with $n$ columns and rank $n - k$ such that

$$\mathcal{H}\boldsymbol{x}^t = \bar{\mathbf{0}} \;\Leftrightarrow\; \boldsymbol{x} \in \mathcal{C}.$$

- The matrix $\mathcal{H}$ is a *parity-check matrix*.
- The value $k$ is the *dimension* of the code $\mathcal{C}$.
- The ratio $r = k/n$ is the *rate* of the code $\mathcal{C}$.

**Definition**

- Let $\mathcal{C}$ be a code of minimum distance $d$ over $\Sigma$.

# Linear Code

## Definition

- A code $\mathcal{C}$ over field $\mathbb{F} = \mathrm{GF}(q)$ is said to be a *linear $[n, k, d]$ code* if there exists a matrix $\mathcal{H}$ with $n$ columns and rank $n - k$ such that
$$\mathcal{H}\boldsymbol{x}^t = \bar{\mathbf{0}} \;\Leftrightarrow\; \boldsymbol{x} \in \mathcal{C}.$$
- The matrix $\mathcal{H}$ is a *parity-check matrix*.
- The value $k$ is the *dimension* of the code $\mathcal{C}$.
- The ratio $r = k/n$ is the *rate* of the code $\mathcal{C}$.

## Definition

- Let $\mathcal{C}$ be a code of minimum distance $d$ over $\Sigma$.
- The *unique decoding problem:*

# Linear Code

## Definition

- A code $\mathcal{C}$ over field $\mathbb{F} = \mathrm{GF}(q)$ is said to be a *linear* $[n, k, d]$ *code* if there exists a matrix $\mathcal{H}$ with $n$ columns and rank $n - k$ such that

$$\mathcal{H}\boldsymbol{x}^t = \bar{\mathbf{0}} \;\Leftrightarrow\; \boldsymbol{x} \in \mathcal{C}.$$

- The matrix $\mathcal{H}$ is a *parity-check matrix*.
- The value $k$ is the *dimension* of the code $\mathcal{C}$.
- The ratio $r = k/n$ is the *rate* of the code $\mathcal{C}$.

## Definition

- Let $\mathcal{C}$ be a code of minimum distance $d$ over $\Sigma$.
- The *unique decoding problem:*

  **Input:** $\boldsymbol{y} \in \Sigma^n$.

# Linear Code

## Definition

- A code $\mathcal{C}$ over field $\mathbb{F} = \mathrm{GF}(q)$ is said to be a *linear* $[n, k, d]$ *code* if there exists a matrix $\mathcal{H}$ with $n$ columns and rank $n - k$ such that
$$\mathcal{H}\boldsymbol{x}^t = \bar{\mathbf{0}} \;\Leftrightarrow\; \boldsymbol{x} \in \mathcal{C}.$$
- The matrix $\mathcal{H}$ is a *parity-check matrix*.
- The value $k$ is the *dimension* of the code $\mathcal{C}$.
- The ratio $r = k/n$ is the *rate* of the code $\mathcal{C}$.

## Definition

- Let $\mathcal{C}$ be a code of minimum distance $d$ over $\Sigma$.
- The *unique decoding problem:*

  **Input:** $\boldsymbol{y} \in \Sigma^n$.

  **Find:** $\boldsymbol{c} \in \mathcal{C}$, such that $\mathsf{d}(\boldsymbol{c}, \boldsymbol{y}) < d/2$.

# Gilbert-Varshamov Bound

Let $\mathsf{H}_q : [0,1] \to [0,1]$ be the $q$-ary entropy function:

$$\mathsf{H}_q(x) = x \log_q(q-1) - x \log_q x - (1-x) \log_q(1-x) \; .$$

Let $\mathsf{H}_q : [0,1] \to [0,1]$ be the $q$-ary entropy function:

$$\mathsf{H}_q(x) = x \log_q(q-1) - x \log_q x - (1-x) \log_q(1-x) \ .$$

**Theorem**

*Let $\mathbb{F} = \mathrm{GF}(q)$, and let $\delta \in (0, 1-1/q]$ and $\mathcal{R} \in (0,1)$, such that*

$$\mathcal{R} \leq 1 - \mathsf{H}_q(\delta) \ .$$

*Then, for large enough values of $n$, there exists a linear $[n, \mathcal{R}n, \geq \delta n]$ code over $\mathbb{F}$.*

# Gilbert-Varshamov Bound

Let $\mathsf{H}_q : [0,1] \to [0,1]$ be the $q$-ary entropy function:

$$\mathsf{H}_q(x) = x \log_q(q-1) - x \log_q x - (1-x) \log_q(1-x) \; .$$

**Theorem**

*Let* $\mathbb{F} = \mathrm{GF}(q)$, *and let* $\delta \in (0, 1-1/q]$ *and* $\mathcal{R} \in (0,1)$, *such that*

$$\mathcal{R} \leq 1 - \mathsf{H}_q(\delta) \; .$$

*Then, for large enough values of* $n$, *there exists a linear* $[n, \mathcal{R}n, \geq \delta n]$ *code over* $\mathbb{F}$.

- The above expression is called the *Gilbert-Varshamov bound*.

# Gilbert-Varshamov Bound

Let $\mathsf{H}_q : [0,1] \to [0,1]$ be the $q$-ary entropy function:

$$\mathsf{H}_q(x) = x \log_q(q-1) - x \log_q x - (1-x) \log_q(1-x) \ .$$

**Theorem**

*Let* $\mathbb{F} = \mathrm{GF}(q)$, *and let* $\delta \in (0, 1-1/q]$ *and* $\mathcal{R} \in (0,1)$, *such that*

$$\mathcal{R} \leq 1 - \mathsf{H}_q(\delta) \ .$$

*Then, for large enough values of* $n$, *there exists a linear* $[n, \mathcal{R}n, \geq \delta n]$ *code over* $\mathbb{F}$.

- The above expression is called the *Gilbert-Varshamov bound*.
- Denote $\delta_{GV}(\mathcal{R}) = \mathsf{H}_2^{-1}(1-\mathcal{R})$.

# Concatenated Codes

[Forney '66] Ingredients:

- A linear $[\Delta, k{=}r\Delta, \theta\Delta]$ code $\mathcal{C}$ over $\mathbb{F} = \mathrm{GF}(q)$ (*inner code*).

# Concatenated Codes

[Forney '66] Ingredients:

- A linear $[\Delta, k{=}r\Delta, \theta\Delta]$ code $\mathcal{C}$ over $\mathbb{F} = \mathrm{GF}(q)$ (*inner code*).
- A linear $[N, R_\Phi N, \delta_\Phi N]$ code $\mathbb{C}_\Phi$ over $\Phi = \mathbb{F}^k$ (*outer code*).

# Concatenated Codes

[Forney '66] Ingredients:

- A linear $[\Delta, k{=}r\Delta, \theta\Delta]$ code $\mathcal{C}$ over $\mathbb{F} = \mathrm{GF}(q)$ (*inner code*).
- A linear $[N, R_\Phi N, \delta_\Phi N]$ code $\mathbb{C}_\Phi$ over $\Phi = \mathbb{F}^k$ (*outer code*).
- A linear one-to-one mapping $\mathcal{E} \; : \; \Phi \to \mathcal{C}$.

[Forney '66] Ingredients:

- A linear $[\Delta, k{=}r\Delta, \theta\Delta]$ code $\mathcal{C}$ over $\mathbb{F} = \mathrm{GF}(q)$ (*inner code*).
- A linear $[N, R_\Phi N, \delta_\Phi N]$ code $\mathbb{C}_\Phi$ over $\Phi = \mathbb{F}^k$ (*outer code*).
- A linear one-to-one mapping $\mathcal{E} \, : \, \Phi \to \mathcal{C}$.

Concatenated code $\mathbb{C}$ of length $N = \Delta n$ over $\mathbb{F}$ is defined as

$$\mathbb{C} = \Big\{ (\boldsymbol{c}_1 | \boldsymbol{c}_2 | \cdots | \boldsymbol{c}_n) \in \mathbb{F}^{\Delta n} : \boldsymbol{c}_i = \mathcal{E}(a_i) \; ,$$

$$\text{for } i \in 1, 2, \cdots, n, \quad \text{and } (a_1 a_2 \cdots a_n) \in \mathbb{C}_\Phi \Big\} \; .$$

# Concatenated Codes

[Forney '66] Ingredients:

- A linear $[\Delta, k{=}r\Delta, \theta\Delta]$ code $\mathcal{C}$ over $\mathbb{F} = \mathrm{GF}(q)$ (*inner code*).
- A linear $[N, R_\Phi N, \delta_\Phi N]$ code $\mathbb{C}_\Phi$ over $\Phi = \mathbb{F}^k$ (*outer code*).
- A linear one-to-one mapping $\mathcal{E} \; : \; \Phi \to \mathcal{C}$.

Concatenated code $\mathbb{C}$ of length $N = \Delta n$ over $\mathbb{F}$ is defined as

$$\mathbb{C} = \Big\{ (\boldsymbol{c}_1|\boldsymbol{c}_2|\cdots|\boldsymbol{c}_n) \in \mathbb{F}^{\Delta n} : \boldsymbol{c}_i = \mathcal{E}(a_i) \; ,$$

$$\text{for } i \in 1, 2, \cdots, n, \;\; \text{and } (a_1 a_2 \cdots a_n) \in \mathbb{C}_\Phi \Big\} \; .$$

- The rate of $\mathbb{C}$: $\mathcal{R} = r R_\Phi$.

# Concatenated Codes

[Forney '66]    Ingredients:

- A linear $[\Delta, k{=}r\Delta, \theta\Delta]$ code $\mathcal{C}$ over $\mathbb{F} = \mathrm{GF}(q)$ (*inner code*).
- A linear $[N, R_\Phi N, \delta_\Phi N]$ code $\mathbb{C}_\Phi$ over $\Phi = \mathbb{F}^k$ (*outer code*).
- A linear one-to-one mapping $\mathcal{E} \ : \ \Phi \to \mathcal{C}$.

Concatenated code $\mathbb{C}$ of length $N = \Delta n$ over $\mathbb{F}$ is defined as

$$\mathbb{C} = \Big\{ (\boldsymbol{c}_1 | \boldsymbol{c}_2 | \cdots | \boldsymbol{c}_n) \in \mathbb{F}^{\Delta n} : \boldsymbol{c}_i = \mathcal{E}(a_i) \ ,$$
$$\text{for } i \in 1, 2, \cdots, n, \quad \text{and } (a_1 a_2 \cdots a_n) \in \mathbb{C}_\Phi \Big\} \ .$$

- The rate of $\mathbb{C}$: $\mathcal{R} = r R_\Phi$.
- The relative minimum distance of $\mathbb{C}$: $\delta \geq \theta \delta_\Phi$.

# Concatenated Codes (Cont.)

- *Generalized minimum distance* (GMD) decoder corrects any fraction of errors up to $\frac{1}{2}\delta$.

- *Generalized minimum distance* (GMD) decoder corrects any fraction of errors up to $\frac{1}{2}\delta$.
- [Justesen '72] For a wide range of rates, concatenated codes attain the *Zyablov bound:*

$$\delta \geq \max_{\mathcal{R} \leq r \leq 1} \left(1 - \frac{\mathcal{R}}{r}\right) \mathsf{H}_q^{-1}(1 - r).$$

# Concatenated Codes (Cont.)

- *Generalized minimum distance* (GMD) decoder corrects any fraction of errors up to $\frac{1}{2}\delta$.
- [Justesen '72] For a wide range of rates, concatenated codes attain the *Zyablov bound:*

$$\delta \geq \max_{\mathcal{R} \leq r \leq 1} \left(1 - \frac{\mathcal{R}}{r}\right) \mathsf{H}_q^{-1}(1-r).$$

- [Blokh-Zyablov '82] Multilevel concatenations of codes (almost) attain the *Blokh-Zyablov bound*:

$$\mathcal{R} = 1 - \mathsf{H}_2(\delta) - \delta \int_0^{1-\mathsf{H}_2(\delta)} \frac{dx}{\mathsf{H}_2^{-1}(1-x)} .$$

# Graphs and Eigenvalues

- Consider a $\Delta$-regular graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.

# Graphs and Eigenvalues

- Consider a $\Delta$-regular graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.
- The largest eigenvalue of the adjacency matrix $A_{\mathcal{G}}$ of $\mathcal{G}$ equals $\Delta$.

# Graphs and Eigenvalues

- Consider a $\Delta$-regular graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.
- The largest eigenvalue of the adjacency matrix $A_{\mathcal{G}}$ of $\mathcal{G}$ equals $\Delta$.
- Let $\lambda^*_{\mathcal{G}}$ be the second largest absolute value of eigenvalues of $A_{\mathcal{G}}$.

# Graphs and Eigenvalues

- Consider a $\Delta$-regular graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.
- The largest eigenvalue of the adjacency matrix $A_{\mathcal{G}}$ of $\mathcal{G}$ equals $\Delta$.
- Let $\lambda^*_{\mathcal{G}}$ be the second largest absolute value of eigenvalues of $A_{\mathcal{G}}$.
- Lower ratios of $\lambda^*_{\mathcal{G}}/\Delta$ imply greater values of *expansion* [Alon '86].

# Graphs and Eigenvalues

- Consider a $\Delta$-regular graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.
- The largest eigenvalue of the adjacency matrix $A_{\mathcal{G}}$ of $\mathcal{G}$ equals $\Delta$.
- Let $\lambda^*_{\mathcal{G}}$ be the second largest absolute value of eigenvalues of $A_{\mathcal{G}}$.
- Lower ratios of $\lambda^*_{\mathcal{G}}/\Delta$ imply greater values of *expansion* [Alon '86].
- Expander graphs with

$$\lambda^*_{\mathcal{G}} \leq 2\sqrt{\Delta - 1}$$

are called a *Ramanujan graphs*. Constructions are due to [Lubotsky Philips Sarnak '88], [Margulis '88].

# Graphs and Eigenvalues

- Consider a $\Delta$-regular graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.
- The largest eigenvalue of the adjacency matrix $A_{\mathcal{G}}$ of $\mathcal{G}$ equals $\Delta$.
- Let $\lambda^*_{\mathcal{G}}$ be the second largest absolute value of eigenvalues of $A_{\mathcal{G}}$.
- Lower ratios of $\lambda^*_{\mathcal{G}}/\Delta$ imply greater values of *expansion* [Alon '86].
- Expander graphs with

$$\lambda^*_{\mathcal{G}} \leq 2\sqrt{\Delta - 1}$$

  are called a *Ramanujan graphs*. Constructions are due to [Lubotsky Philips Sarnak '88], [Margulis '88].
- Let $\lambda_{\mathcal{G}}$ be the second largest eigenvalues of $A_{\mathcal{G}}$ and $\gamma_{\mathcal{G}} = \lambda_{\mathcal{G}}/\Delta$.

# Barg-Zémor's Expander Codes '02

- $\mathcal{G}$ is bipartite: $\mathcal{V} = A \cup B$, $A \cap B = \emptyset$, $|A| = |B| = n$.
- Ordering on the vertices and the edges.
- Denote by $(\boldsymbol{z})_{\mathcal{E}(u)}$ the sub-block of $\boldsymbol{z}$ that is indexed by $\mathcal{E}(u)$.
- Let $\mathcal{C}_A$ and $\mathcal{C}_B$ be two linear codes of length $\Delta$ over $\mathbb{F}$.
- Denote $N = |\mathcal{E}| = \Delta n$.

# Barg-Zémor's Expander Codes '02

- $\mathcal{G}$ is bipartite: $\mathcal{V} = A \cup B$, $A \cap B = \emptyset$, $|A| = |B| = n$.
- Ordering on the vertices and the edges.
- Denote by $(\boldsymbol{z})_{\mathcal{E}(u)}$ the sub-block of $\boldsymbol{z}$ that is indexed by $\mathcal{E}(u)$.
- Let $\mathcal{C}_A$ and $\mathcal{C}_B$ be two linear codes of length $\Delta$ over $\mathbb{F}$.
- Denote $N = |\mathcal{E}| = \Delta n$.

The code $\mathbb{C} = (\mathcal{G}, \mathcal{C}_A : \mathcal{C}_B)$:

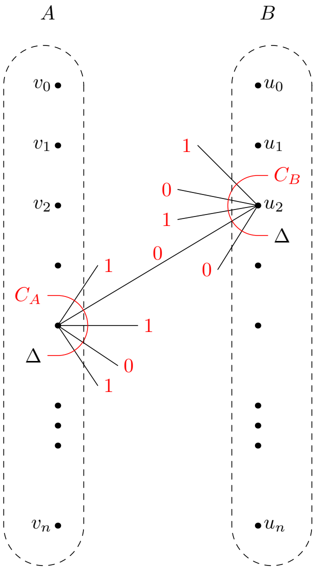
$$\mathbb{C} = \Big\{ \boldsymbol{c} \in \mathbb{F}^N \ : \ (\boldsymbol{c})_{\mathcal{E}(u)} \in \mathcal{C}_A \text{ for } v \in A \text{ and } (\boldsymbol{c})_{\mathcal{E}(v)} \in \mathcal{C}_B \text{ for } u \in B \Big\} .$$

# Barg-Zémor's Expander Codes '02

- $\mathcal{G}$ is bipartite: $\mathcal{V} = A \cup B$, $A \cap B = \emptyset$, $|A| = |B| = n$.
- Ordering on the vertices and the edges.
- Denote by $(\boldsymbol{z})_{\mathcal{E}(u)}$ the sub-block of $\boldsymbol{z}$ that is indexed by $\mathcal{E}(u)$.
- Let $\mathcal{C}_A$ and $\mathcal{C}_B$ be two linear codes of length $\Delta$ over $\mathbb{F}$.
- Denote $N = |\mathcal{E}| = \Delta n$.

The code $\mathbb{C} = (\mathcal{G}, \mathcal{C}_A : \mathcal{C}_B)$:

$$\mathbb{C} = \Big\{ \boldsymbol{c} \in \mathbb{F}^N \; : \; (\boldsymbol{c})_{\mathcal{E}(u)} \in \mathcal{C}_A \text{ for } v \in A \text{ and } (\boldsymbol{c})_{\mathcal{E}(v)} \in \mathcal{C}_B \text{ for } u \in B \Big\} .$$



# Barg-Zémor Expander Codes ’03

- ‘Dangling edges’ are introduced [Barg Zémor ’03].

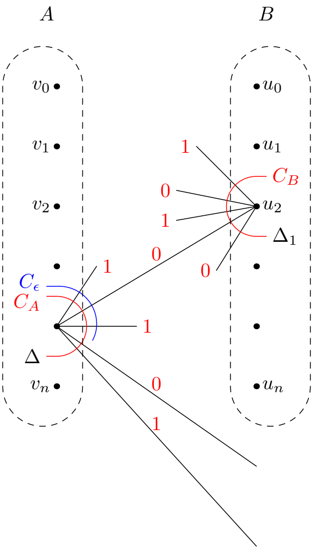
# Barg-Zémor Expander Codes '03

- 'Dangling edges' are introduced [Barg Zémor '03].
- Mimics behavior of concatenated codes.

- ‘Dangling edges’ are introduced [Barg Zémor ’03].
- Mimics behavior of concatenated codes.
- Can be viewed as a concatenation of two codes [Roth Skachek ’04].

# Barg-Zémor Expander Codes '03

- 'Dangling edges' are introduced [Barg Zémor '03].
- Mimics behavior of concatenated codes.
- Can be viewed as a concatenation of two codes [Roth Skachek '04].
- Another construction with similar properties [Guruswami Indyk '02].



# Analysis in [Barg Zémor '04]

Analysis of the codes in [Barg Zémor '02] and [Barg Zémor '03].

# Analysis in [Barg Zémor '04]

Analysis of the codes in [Barg Zémor '02] and [Barg Zémor '03].

**Lower bounds on the relative minimum distance**

(i)

$$\delta(\mathcal{R}) \geq \frac{1}{4}(1-\mathcal{R})^2 \cdot \min_{\delta_{GV}((1+\mathcal{R})/2) < \mathsf{B} < \frac{1}{2}} \frac{g(\mathsf{B})}{\mathsf{H}_2(\mathsf{B})} ,$$

where the function $g(\mathsf{B})$ is defined in the next slides.

(ii)

$$\delta(\mathcal{R}) \geq \max_{\mathcal{R} \leq r \leq 1} \left\{ \min_{\delta_{GV}(r) < \mathsf{B} < \frac{1}{2}} \left( \delta_0(\mathsf{B}, r) \cdot \frac{1 - \mathcal{R}/r}{\mathsf{H}_2(\mathsf{B})} \right) \right\} ,$$

where the function $\delta_0(\mathsf{B}, r)$ is defined in the next slides.

# Definition of the Function $g(\mathsf{B})$

These two families of codes surpass the Zyablov bound.

# Definition of the Function $g(\mathsf{B})$

These two families of codes surpass the Zyablov bound.

Let $\delta_{GV}(\mathcal{R}) = \mathsf{H}_2^{-1}(1-\mathcal{R})$, and let $\mathsf{B}_1$ be the largest root of the equation

$$\mathsf{H}_2(\mathsf{B}) = \mathsf{H}_2(\mathsf{B})\left(\mathsf{B} - \mathsf{H}_2(\mathsf{B}) \cdot \frac{\delta_{GV}(\mathcal{R})}{1-\mathcal{R}}\right) = -\left(\mathsf{B} - \delta_{GV}(\mathcal{R})\right) \cdot \log_2(1-\mathsf{B}) \ .$$

Moreover, let

$$a_1 = \frac{\mathsf{B}_1}{\mathsf{H}_2(\mathsf{B}_1)} - \frac{\delta_{GV}(\mathcal{R})}{\mathsf{H}_2(\delta_{GV}(\mathcal{R}))} \ ,$$

and

$$b_1 = \frac{\delta_{GV}(\mathcal{R})}{\mathsf{H}_2(\delta_{GV}(\mathcal{R}))} \cdot \mathsf{B}_1 - \frac{\mathsf{B}_1}{\mathsf{H}_2(\mathsf{B}_1)} \cdot \delta_{GV}(\mathcal{R})) \ .$$

# Definition of the Function $g(\mathsf{B})$ (Cont.)

The function $g(\mathsf{B})$ is defined as

$$g(\mathsf{B}) = \begin{cases} \dfrac{\delta_{GV}(\mathcal{R})}{1-\mathcal{R}} & \text{if } \mathsf{B} \le \delta_{GV}(\mathcal{R}) \\ \dfrac{\mathsf{B}}{\mathsf{H}_2(\mathsf{B})} & \text{if } \delta_{GV}(\mathcal{R}) \le \mathsf{B} \text{ and } \mathcal{R} \le 0.284 \\ \dfrac{a_1\mathsf{B}+b_1}{\mathsf{B}_1 - \delta_{GV}(\mathcal{R})} & \text{if } \delta_{GV}(\mathcal{R}) \le \mathsf{B} \le \mathsf{B}_1 \text{ and } 0.284 < \mathcal{R} \le 1 \\ \dfrac{\mathsf{B}}{\mathsf{H}_2(\mathsf{B})} & \text{if } \mathsf{B}_1 < \mathsf{B}_1 \le 1 \text{ and } 0.284 < \mathcal{R} \le 1 \end{cases} .$$

# Definition of the Function $\delta_0(\mathsf{B}, r)$

The function $\delta_0(\mathsf{B}, r)$ is defined to be $\omega^{\star\star}(\mathsf{B})$ for $\delta_{GV}(r) \le \mathsf{B} \le \mathsf{B}_1$, where

$$\omega^{\star\star}(\mathsf{B}) = r\mathsf{B} + (1-r)\mathsf{H}_2^{-1}\left(1 - \frac{r}{1-r}\mathsf{H}_2(\mathsf{B})\right) ,$$

and $\mathsf{B}_1$ is the only root of the equation

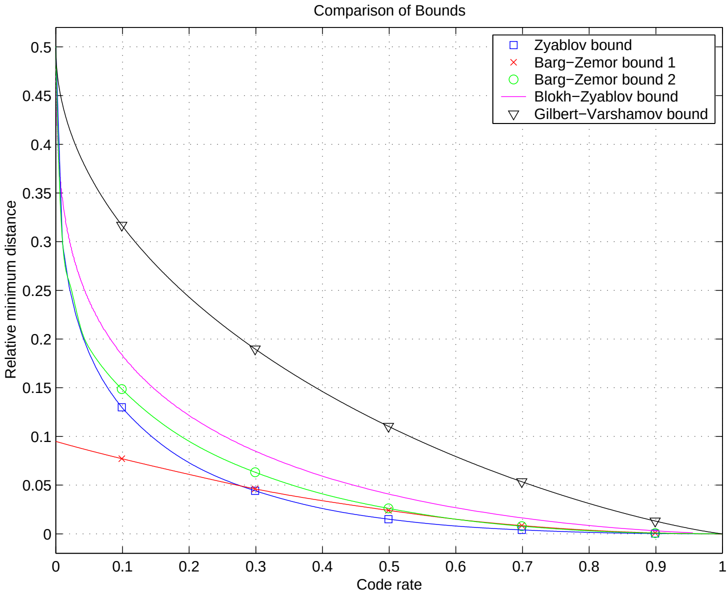
$$\delta_{GV}(r) = w^{\star}(\mathsf{B}) ,$$

where

$$w^{\star}(\mathsf{B}) = (1-r)\left((2^{\mathsf{H}_2(\mathsf{B})/\mathsf{B}}+1)^{-1} + \frac{\mathsf{B}}{\mathsf{H}_2(\mathsf{B})}\left(1 - \mathsf{H}_2\left((2^{\mathsf{H}_2(\mathsf{B})/\mathsf{B}}+1)^{-1}\right)\right)\right) .$$

For $\mathsf{B}_1 \le \mathsf{B} \le \frac{1}{2}$, the function $\delta_0(\mathsf{B}, r)$ is defined to be a tangent to the function $\omega^{\star\star}(\mathsf{B})$ drawn from the point $\left(\frac{1}{2}, \omega^{\star}(\frac{1}{2})\right)$.

# Minimum Distance Bounds



# Generalized Expander Codes

- $\mathcal{G} = (\mathcal{V} = A \cup B, \mathcal{E})$ be a bipartite $\Delta$-regular, as before

# Generalized Expander Codes

- $\mathcal{G} = (\mathcal{V} = A \cup B, \mathcal{E})$ be a bipartite $\Delta$-regular, as before
- $B = B^1 \cup B^2$, $B^1 \cap B^2 = \emptyset$. Let $|B^2| = \eta n$, $|B^1| = (1-\eta)n$, $\eta \in [0,1]$.

# Generalized Expander Codes

- $\mathcal{G} = (\mathcal{V} = A \cup B, \mathcal{E})$ be a bipartite $\Delta$-regular, as before
- $B = B^1 \cup B^2$, $B^1 \cap B^2 = \emptyset$. Let $|B^2| = \eta n$, $|B^1| = (1-\eta)n$, $\eta \in [0,1]$.
- $\mathcal{C}_A$, $\mathcal{C}_1$ and $\mathcal{C}_2$ are linear $[\Delta, r_A\Delta, \delta_A\Delta]$, $[\Delta, r_1\Delta, \delta_1\Delta]$ and $[\Delta, r_2\Delta, \delta_2\Delta]$ codes over $\mathbb{F}$, respectively.

# Generalized Expander Codes

- $\mathcal{G} = (\mathcal{V} = A \cup B, \mathcal{E})$ be a bipartite $\Delta$-regular, as before
- $B = B^1 \cup B^2$, $B^1 \cap B^2 = \emptyset$. Let $|B^2| = \eta n$, $|B^1| = (1-\eta)n$, $\eta \in [0,1]$.
- $\mathcal{C}_A$, $\mathcal{C}_1$ and $\mathcal{C}_2$ are linear $[\Delta, r_A\Delta, \delta_A\Delta]$, $[\Delta, r_1\Delta, \delta_1\Delta]$ and $[\Delta, r_2\Delta, \delta_2\Delta]$ codes over $\mathbb{F}$, respectively.

The code code $\mathbb{C} = (\mathcal{G}, \mathcal{C}_A, \mathcal{C}_1, \mathcal{C}_2)$:

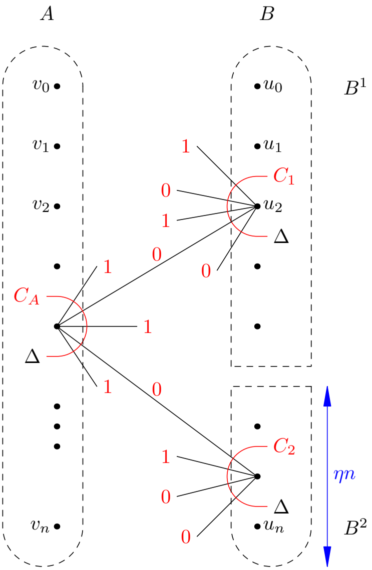
$$\mathbb{C} = \Big\{ \boldsymbol{c} \in \mathbb{F}^N \quad : \ (\boldsymbol{c})_{\mathcal{E}(u)} \in \mathcal{C}_A \text{ for } u \in A, \\ (\boldsymbol{c})_{\mathcal{E}(u)} \in \mathcal{C}_1 \text{ for } u \in B^1 \\ \text{and} \quad (\boldsymbol{c})_{\mathcal{E}(u)} \in \mathcal{C}_2 \text{ for } u \in B^2 \Big\}$$

# Generalized Expander Codes

- $\mathcal{G} = (\mathcal{V} = A \cup B, \mathcal{E})$ be a bipartite $\Delta$-regular, as before
- $B = B^1 \cup B^2$, $B^1 \cap B^2 = \emptyset$. Let $|B^2| = \eta n$, $|B^1| = (1-\eta)n$, $\eta \in [0,1]$.
- $\mathcal{C}_A$, $\mathcal{C}_1$ and $\mathcal{C}_2$ are linear $[\Delta, r_A\Delta, \delta_A\Delta]$, $[\Delta, r_1\Delta, \delta_1\Delta]$ and $[\Delta, r_2\Delta, \delta_2\Delta]$ codes over $\mathbb{F}$, respectively.

The code code $\mathbb{C} = (\mathcal{G}, \mathcal{C}_A, \mathcal{C}_1, \mathcal{C}_2)$:

$$\mathbb{C} = \Big\{ \boldsymbol{c} \in \mathbb{F}^N \; : \; (\boldsymbol{c})_{\mathcal{E}(u)} \in \mathcal{C}_A \text{ for } u \in A, \; (\boldsymbol{c})_{\mathcal{E}(u)} \in \mathcal{C}_1 \text{ for } u \in B^1 \text{ and } (\boldsymbol{c})_{\mathcal{E}(u)} \in \mathcal{C}_2 \text{ for } u \in B^2 \Big\}$$



# Properties of Generalized Expander Codes

- *The rate:* $\mathcal{R} \geq r_A + (1-\eta)r_1 + \eta r_2 - 1$.

# Properties of Generalized Expander Codes

- *The rate:* $\mathcal{R} \geq r_A + (1-\eta) r_1 + \eta r_2 - 1$.
- Assume
$$\eta < \frac{\delta_A - \gamma_{\mathcal{G}} \sqrt{\delta_A / \delta_2}}{1 - \gamma_{\mathcal{G}}} - \gamma_{\mathcal{G}}^{2/3} \; .$$
Then, *the relative minimum distance:*
$$\delta > \delta_A (\delta_1 - \tfrac{1}{2} \gamma_{\mathcal{G}}^{2/3}) \; .$$
$\Rightarrow$ The code $\mathbb{C}$ attains the *Zyablov bound*.

# Properties of Generalized Expander Codes

- *The rate:* $\mathcal{R} \geq r_A + (1-\eta) r_1 + \eta r_2 - 1$.
- Assume
$$\eta < \frac{\delta_A - \gamma_{\mathcal{G}} \sqrt{\delta_A / \delta_2}}{1 - \gamma_{\mathcal{G}}} - \gamma_{\mathcal{G}}^{2/3} \; .$$
Then, *the relative minimum distance:*
$$\delta > \delta_A (\delta_1 - \tfrac{1}{2} \gamma_{\mathcal{G}}^{2/3}) \; .$$
$\Rightarrow$ The code $\mathbb{C}$ attains the *Zyablov bound*.
- *A linear-time decoding algorithm:* if $\delta_1 > 2\gamma_{\mathcal{G}}^{2/3}$ and $\eta$ as above, the decoder corrects any error pattern of size $\mathbb{J}_{\mathbb{C}}$,
$$\mathbb{J}_{\mathbb{C}} \triangleq \frac{\frac{1}{2}\delta_1 - \gamma_{\mathcal{G}}^{2/3} \left( 1 + \sqrt{2 \left( \delta_1 - 2\gamma_{\mathcal{G}}^{2/3} \right)} \right)}{1 - \gamma_{\mathcal{G}}} \cdot \delta_A \Delta n \; .$$
The number of correctable errors is (almost) half of the Zyablov bound.

# Properties of Generalized Expander Codes (cont.)

**Theorem**

*Let $|\mathbb{F}|$ be a power of 2. There exists a polynomial-time constructible family of binary linear codes $\mathbb{C}$ of length $N = n\Delta$, $n \to \infty$, and sufficiently large but constant $\Delta = \Delta(\varepsilon)$, whose relative minimum distance satisfies*

$$\delta(\mathcal{R}) \geq \max_{\mathcal{R} \leq r_A \leq 1} \left\{ \min_{\delta_{GV}(r_A) \leq \beta \leq 1/2} \left( \delta_0(\beta, r_A) \frac{1 - \mathcal{R}/r_A}{\mathsf{H}_2(\beta)} \right) \right\} - \varepsilon \ .$$

# Properties of Generalized Expander Codes (cont.)

**Theorem**

*Let $|\mathbb{F}|$ be a power of 2. There exists a polynomial-time constructible family of binary linear codes $\mathbb{C}$ of length $N = n\Delta$, $n \to \infty$, and sufficiently large but constant $\Delta = \Delta(\varepsilon)$, whose relative minimum distance satisfies*

$$\delta(\mathcal{R}) \geq \max_{\mathcal{R} \leq r_A \leq 1} \left\{ \min_{\delta_{GV}(r_A) \leq \beta \leq 1/2} \left( \delta_0(\beta, r_A) \frac{1 - \mathcal{R}/r_A}{\mathsf{H}_2(\beta)} \right) \right\} - \varepsilon \; .$$

Consider a code $\mathbb{C}$ with parameter $\eta = 0$. Then, $|B^2| = 0$, and the code $\mathbb{C}$ coincides with the code in [Barg Zémor'02].

**Theorem**

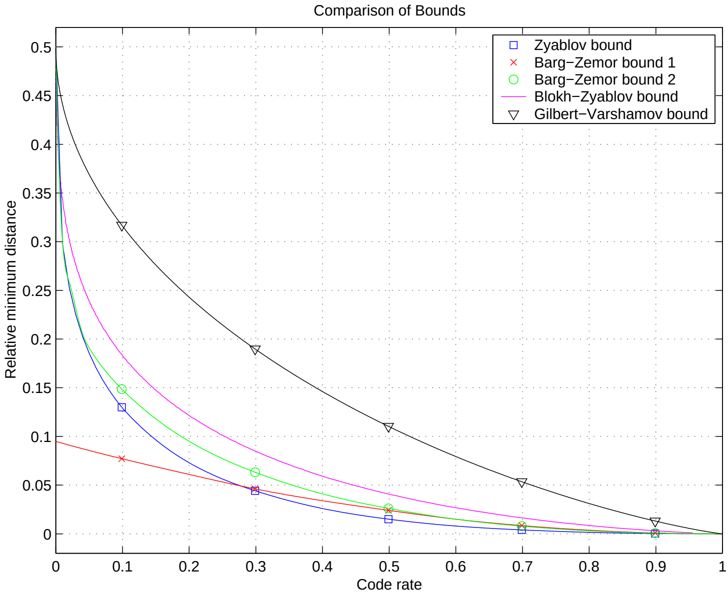
*Let $|\mathbb{F}|$ be a power of 2. There exists a polynomial-time constructible family of binary linear codes $\mathbb{C}$ of length $N = n\Delta$, $n \to \infty$, and sufficiently large but constant $\Delta = \Delta(\varepsilon)$, whose relative minimum distance satisfies*

$$\delta(\mathcal{R}) \geq \max_{\mathcal{R} \leq r_A \leq 1} \left\{ \min_{\delta_{GV}(r_A) \leq \beta \leq 1/2} \left( \delta_0(\beta, r_A) \frac{1 - \mathcal{R}/r_A}{\mathsf{H}_2(\beta)} \right) \right\} - \varepsilon \ .$$

Consider a code $\mathbb{C}$ with parameter $\eta = 0$. Then, $|B^2| = 0$, and the code $\mathbb{C}$ coincides with the code in [Barg Zémor'02]. The minimum distance:

$$\delta(\mathcal{R}) \geq \frac{1}{4}(1 - \mathcal{R})^2 \cdot \min_{\delta_{GV}((1+\mathcal{R})/2) < \mathsf{B} < \frac{1}{2}} \frac{g(\mathsf{B})}{\mathsf{H}_2(\mathsf{B})} \ .$$

# Minimum Distance Bounds



# Open Problems

- Further improvements on the **minimum distance bounds**.

# Open Problems

- Further improvements on the **minimum distance bounds**.
- Bounds on the **error-correcting capabilities** of the decoders.

# Open Problems

- Further improvements on the **minimum distance bounds**.
- Bounds on the **error-correcting capabilities** of the decoders.
- Could **other types of expander graphs** yield better properties?

# Open Problems

- Further improvements on the **minimum distance bounds**.
- Bounds on the **error-correcting capabilities** of the decoders.
- Could **other types of expander graphs** yield better properties?
- Do the **generalized expander codes have any advantage** over the known expander codes?