Engineering school TELECOM Bretagne
Supervisor: Professor **Frédéric Guilloud**

University Royal Institute of Technology
Supervisor: Professor **Lars K. Rasmussen**
Co-Supervisor: Doctor **Vishwambhar Rathi**

Rate-Distortion Bounds for Sparse-Graph Codes

Author: **Grégory Demay**

Brest,
Master Thesis Presentation, 2010.

Overview

Overview

KTH

The Royal Institute of Technology

Kungliga Tekniska Högskolan (KTH) in facts

- founded in 1827 in Stockholm;
- Swedish largest university for technical/engineering education;
- 12,000 undergraduate students;
- 1,400 postgraduate;
- 2,800 employees;

Figure 1: KTH
main campus Valhallavägen.

KTH

The Communication Theory Laboratory

A Few Facts

- Commun. Theory Lab ← School of Electrical Engineering ← KTH
- founded in 2003;
- 2 full professors;
- 3 professor assistants;
- 20 PhD students

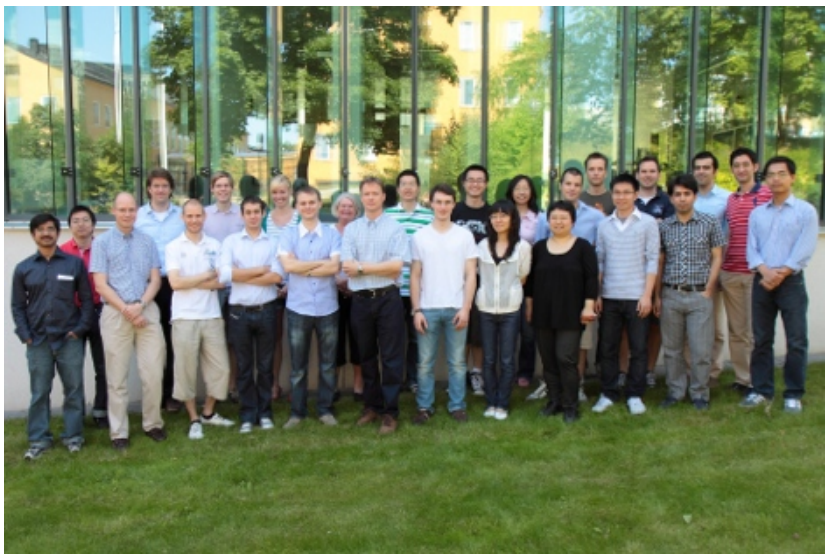


Figure 2: People at Communication Theory Laboratory.

Overview

Background

Digital Communication System Model

Point-to-Point Communication

Background

Digital Communication System Model

Point-to-Point Communication

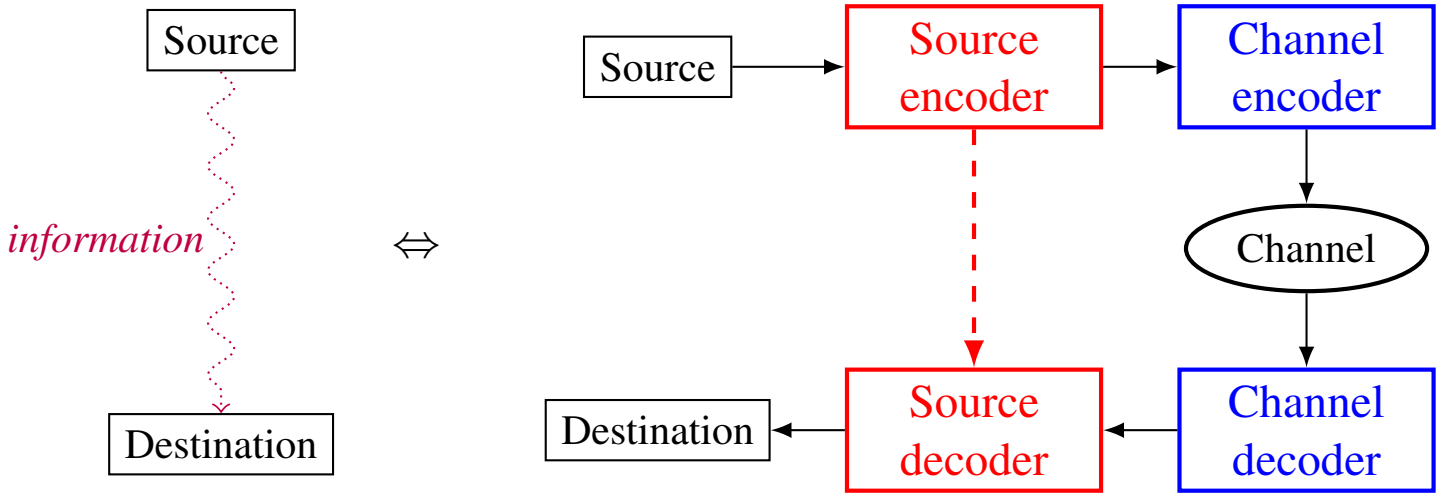


Figure 3: Point-to-point communication problem.

Background

Digital Communication System Model

Point-to-Point Communication

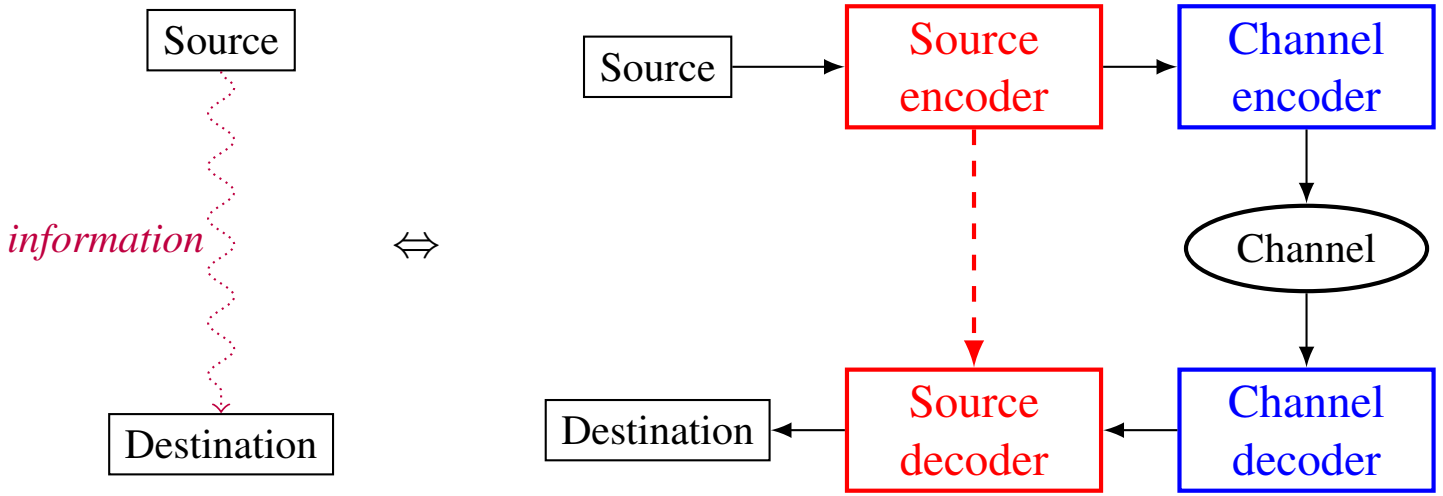


Figure 3: Point-to-point communication problem.

- What is the best possible performance?

Background

Digital Communication System Model

Point-to-Point Communication

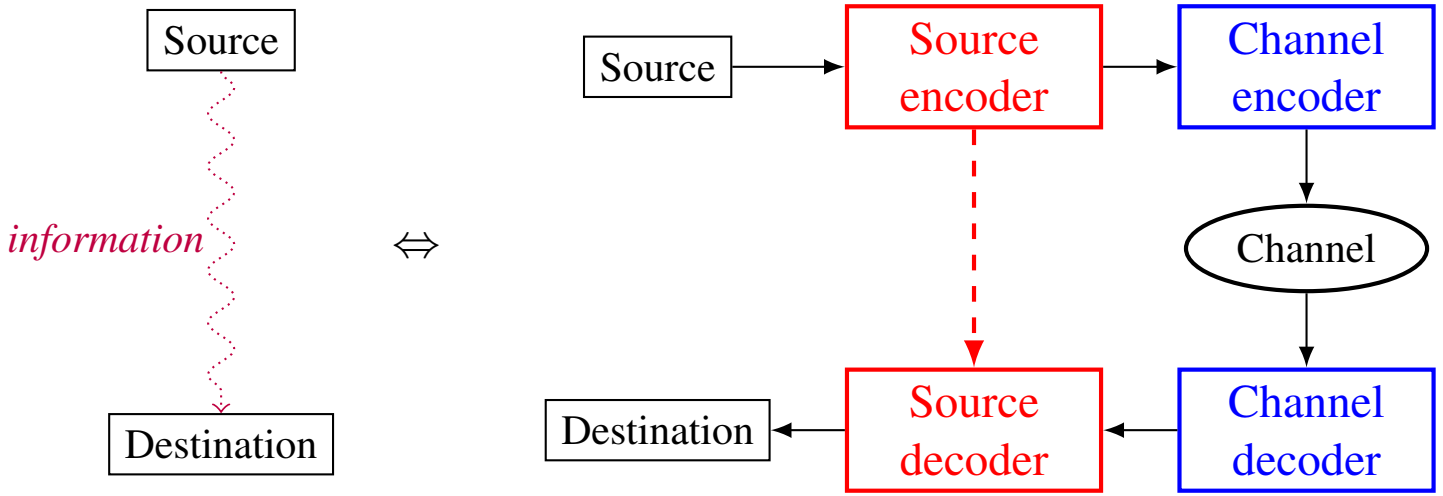


Figure 3: Point-to-point communication problem.

- What is the best possible performance? $\longrightarrow$ *Capacity* C.

Background

Digital Communication System Model

Point-to-Point Communication

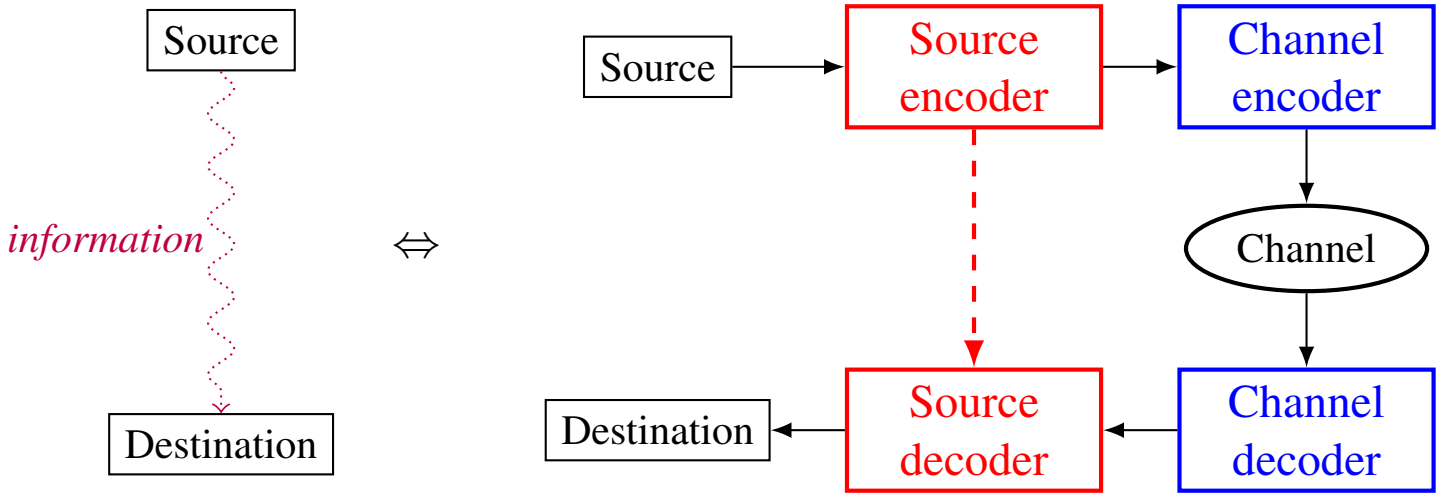


Figure 3: Point-to-point communication problem.

- What is the best possible performance? $\longrightarrow$ *Capacity* C.
- How to achieve it?

Background

Digital Communication System Model

Point-to-Point Communication

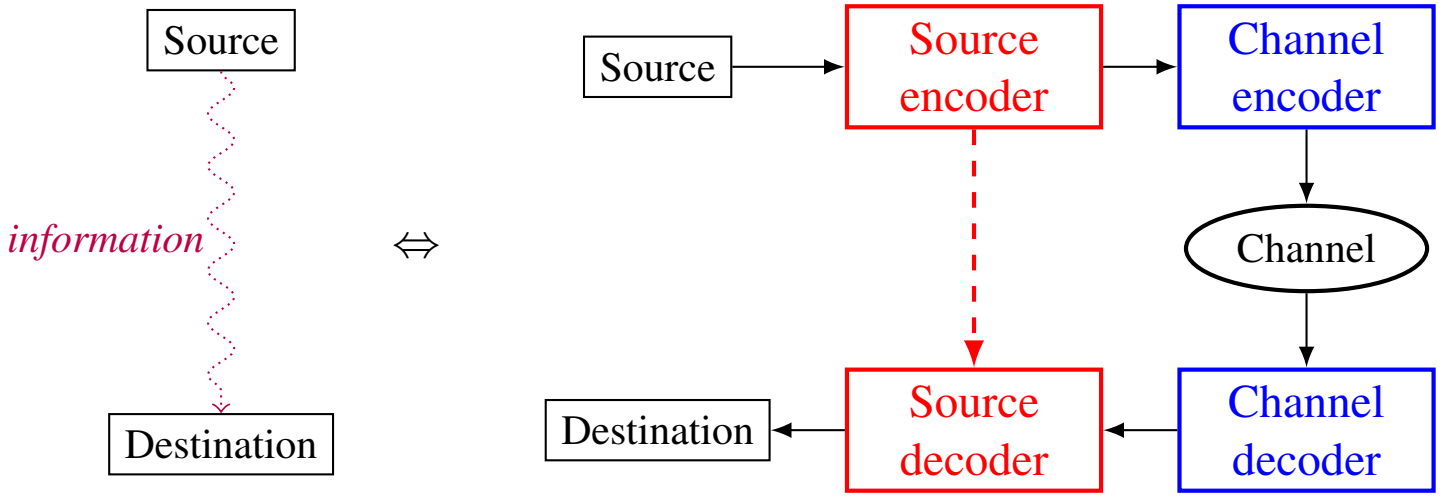


Figure 3: Point-to-point communication problem.

- What is the best possible performance? $\longrightarrow$ *Capacity* C.
- How to achieve it? $\longrightarrow$ *by coding*: source coding / channel coding.

Background

Digital Communication System Model

Point-to-Point Communication

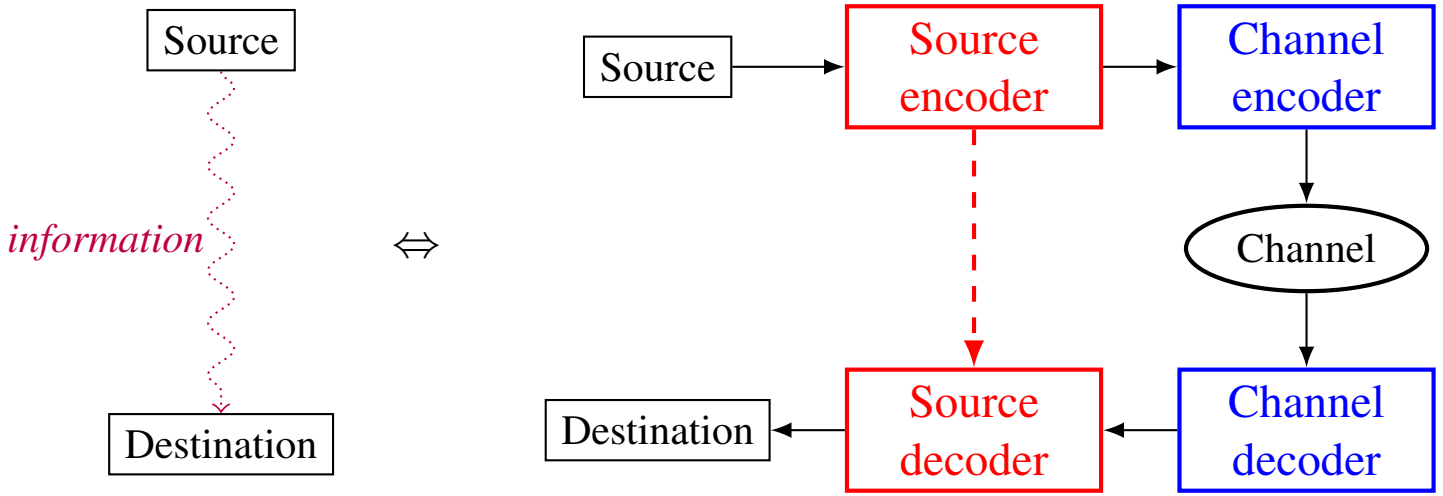


Figure 3: Point-to-point communication problem.

- What is the best possible performance? $\longrightarrow$ *Capacity* C.
- How to achieve it? $\longrightarrow$ *by coding*: source coding / channel coding.

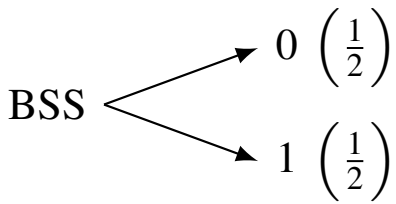
Definition - *Discrete Memoryless Source*

Sequence of i.i.d. RVs $\{S_i\}_{i=1}^{\infty}$ belonging to a finite set $\mathcal{S}$, called the *source alphabet*.

Background

Digital Communication System Model

Binary Symmetric Source (BSS)

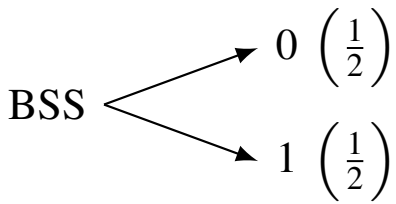


- $\mathcal{S} = \{0, 1\}$
- $S^n = \{S_1, \cdots, S_n\}, S^n \in \mathcal{S}^n$
- $\mathbb{P}\{S_i = 0\} = \mathbb{P}\{S_i = 1\} = \frac{1}{2}, i \in \{1, \cdots, n\}$

Background

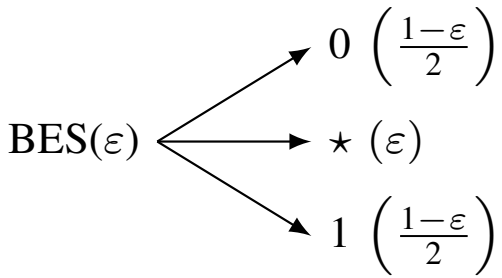
Digital Communication System Model

Binary Symmetric Source (BSS)



- $\mathcal{S} = \{0, 1\}$
- $S^n = \{S_1, \cdots, S_n\}, S^n \in \mathcal{S}^n$
- $\mathbb{P}\{S_i = 0\} = \mathbb{P}\{S_i = 1\} = \frac{1}{2}, i \in \{1, \cdots, n\}$

Binary Erasure Source (BES) [1]



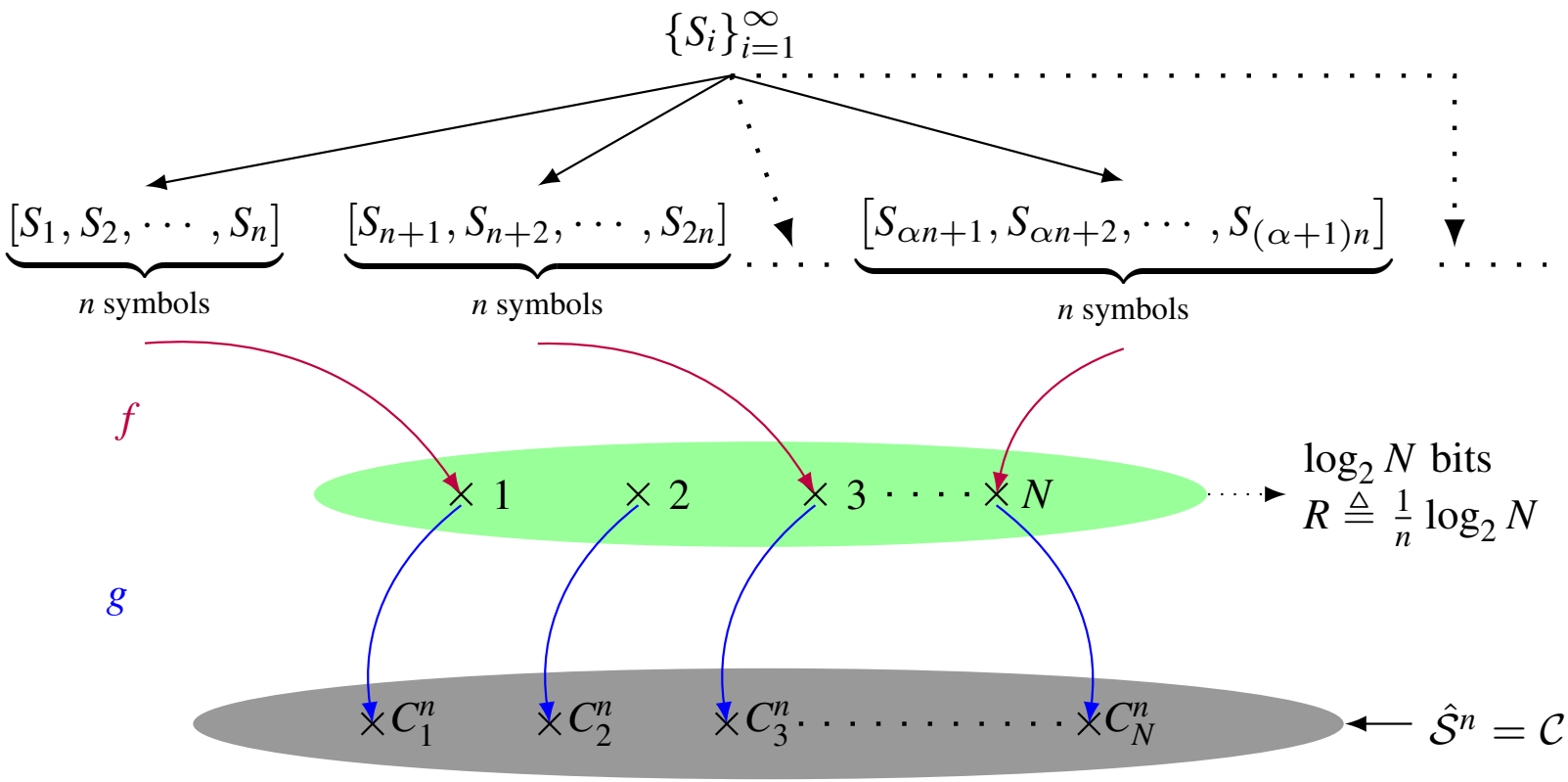
- $\mathcal{S} = \{0, 1, \star\}$
- $S^n = \{S_1, \cdots, S_n\}, S^n \in \mathcal{S}^n$
- $\mathbb{P}\{S_i = \star\} = \epsilon,$
  $\mathbb{P}\{S_i = 0\} = \mathbb{P}\{S_i = 1\} = \frac{1-\epsilon}{2}, i \in \{1, \cdots, n\}$

[1] E. Martinian and J. Yedidia, "Iterative Quantization Using Codes On Graphs," in *Proc. 35th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, 2003.

Background

Rate-Distortion Theory

(n, N) Rate-Distortion Code



Background

Rate-Distortion Theory

Distortion Measure for the BES

$$d(s^n, \hat{s}^n) = \frac{1}{n}\sum_{i=1}^{n} d(s_i, \hat{s}_i), \quad \text{with} \quad d(s_i, \hat{s}_i) = \begin{cases} 0, & \text{if } s_i = \star \text{ or } s_i = \hat{s}_i \\ 1, & \text{otherwise.} \end{cases} \tag{1}$$

- Expected distortion $D = \mathbb{E}\left[d\left(S^n, g\left(f\left(S^n\right)\right)\right)\right]$
- If there any rate-distortion code (f, g) with compression rate R and which achieves an expected distortion D?

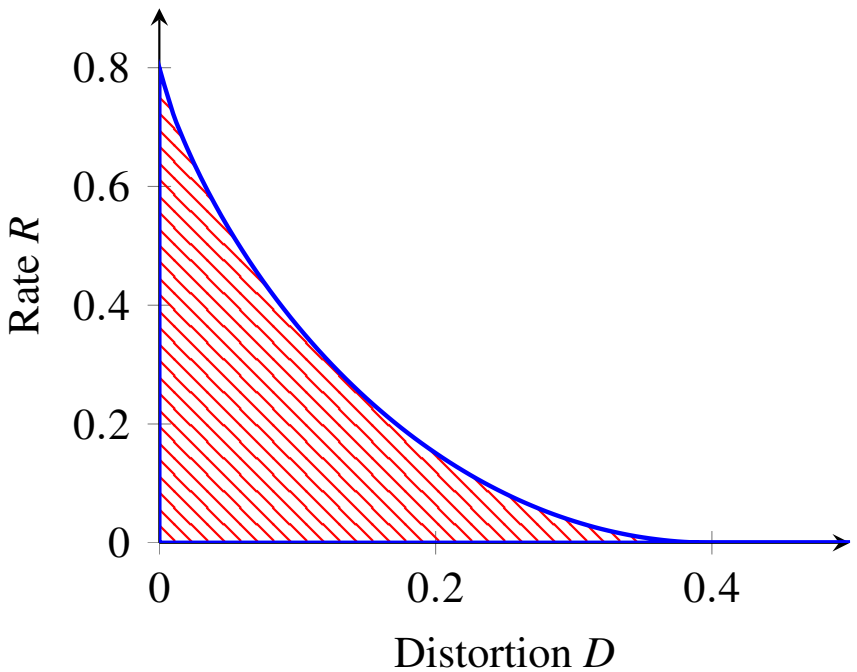
Background

Rate-Distortion Theory

Shannon Rate-Distortion Function for the BES

$$R_{\varepsilon}^{\mathrm{sh}}(D) = \begin{cases} (1-\varepsilon)\left[1 - h\left(\frac{D}{1-\varepsilon}\right)\right], & \text{if } D < \frac{1-\varepsilon}{2} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Figure 4: Shannon rate-distortion function $R_{\varepsilon}^{\mathrm{sh}}(D)$ for a BES(ε) with $\varepsilon = 0.2$.



Background

Linear Block Codes Considered

Definition - *Binary Linear Block Codes*

Collection of binary sequences of length n which span a linear subspace of $\{0,1\}^n$.

- $\mathcal{C}$ linear block code $\Rightarrow \exists$ generator matrix $\mathbf{G}$;
- $\mathcal{C}$ linear block code $\Rightarrow \exists$ parity-check matrix $\mathbf{H}$, generator matrix of $\mathcal{C}^{\perp}$.

Background

Linear Block Codes Considered

Definition - *Binary Linear Block Codes*

Collection of binary sequences of length n which span a linear subspace of $\{0,1\}^n$.

- $\mathcal{C}$ linear block code $\Rightarrow \exists$ generator matrix $\mathbf{G}$;
- $\mathcal{C}$ linear block code $\Rightarrow \exists$ parity-check matrix $\mathbf{H}$, generator matrix of $\mathcal{C}^{\perp}$.

2 ensembles of linear block codes considered [1]

- Check Regular Poisson (CRP) LDGM ensemble $\mathfrak{L}_P\left(d_c, m, n\right)$;
- Compound LDGM-LDPC ensemble $\mathfrak{C}\left(d_c, d_v, d_c', m, n\right)$.

[1] M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, Mar. 2009.

Background

Linear Block Codes Considered

Low-Density Generator Matrix (LDGM) Codes

$$\mathbf{G} = \begin{array}{c} \begin{matrix} c_1 & c_2 & c_3 & c_4 & c_5 & c_6 & c_7 & c_8 & c_9 & c_{10} & c_{11} & c_{12} \end{matrix} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array} \begin{matrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \\ w_8 \\ w_9 \end{matrix}$$

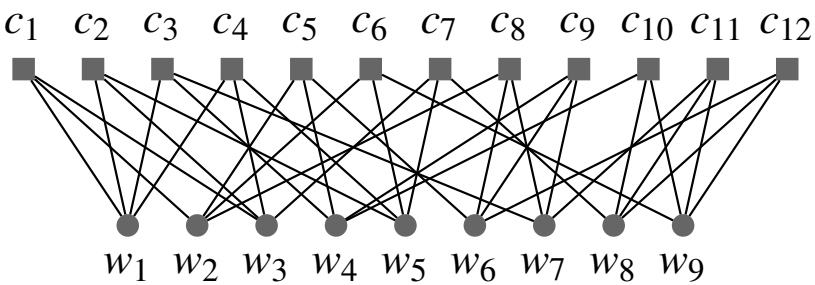


Figure 5: Tanner graph of a $(10, 3, 4)$- regular LDGM code.

$$\mathbb{L}(\mathbf{G}) = \{c^n \in \{0,1\}^n : \exists w^m \in \mathbb{F}^m \text{ s.t. } c^n = w^m \mathbf{G}\} \tag{3}$$

Background

Linear Block Codes Considered

Low-Density Generator Matrix (LDGM) Codes

$$\mathbf{G} = \begin{array}{c} \begin{matrix} c_1 & c_2 & c_3 & c_4 & c_5 & c_6 & c_7 & c_8 & c_9 & c_{10} & c_{11} & c_{12} \end{matrix} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array} \begin{matrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \\ w_8 \\ w_9 \end{matrix}$$

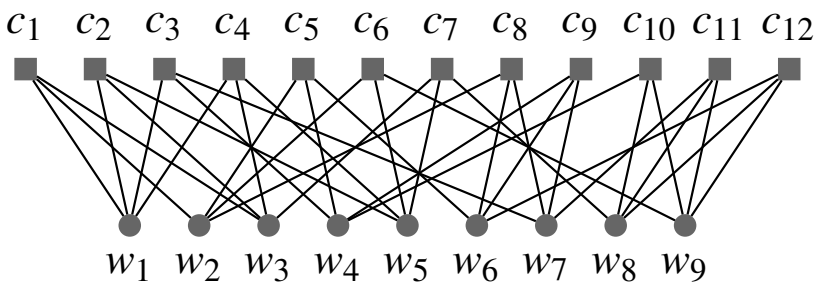


Figure 5: Tanner graph of a $(10, 3, 4)$- regular LDGM code.

$$\mathbb{L}(\mathbf{G}) = \{c^n \in \{0,1\}^n : \exists w^m \in \mathbb{F}^m \text{ s.t. } c^n = w^m \mathbf{G}\} \tag{3}$$

The CRP LDGM Ensemble $\mathfrak{L}_P(d_c, m, n)$

Contains every $\mathbb{L}(\mathbf{G})$, where $\mathbf{G} \in \{0,1\}^{m \times n}$ and is generated by the following procedure. Each check node is connected to d_c information bits chosen uniformly at random and with replacement.

Background

Linear Block Codes Considered

Low-Density Parity-Check (LDPC) Codes

$$\mathbf{H} = \begin{array}{c} \begin{matrix} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 & v_8 & v_9 & v_{10} \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{matrix} \end{array}$$

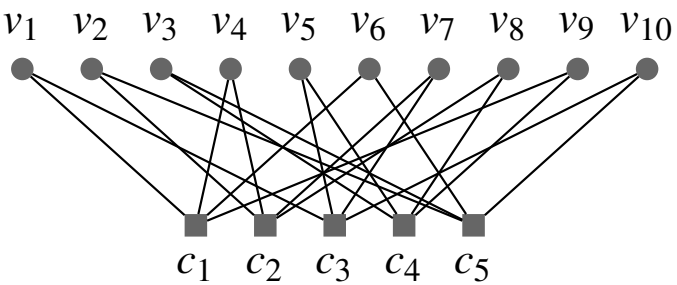


Figure 6: Tanner graph of a $(10, 2, 4)$- regular LDPC code.

$$\mathbb{M}(\mathbf{H}) = \left\{ c^n \in \mathbb{F}^n : \; c^n \mathbf{H}^T = 0 \right\} \tag{4}$$

Background

Linear Block Codes Considered

Compound LDGM-LDPC Codes

Consider $\mathbb{L}(\mathbf{G})$, $\mathbb{M}(\mathbf{H})$, where $\mathbf{G} \in \{0,1\}^{m \times n}$ and $\mathbf{H} \in \{0,1\}^{k \times m}$. Then,

$$\mathbb{C}(\mathbf{G},\mathbf{H}) = \left\{ c^n \in \{0,1\}^n : \exists w^m \in \{0,1\}^m \text{ s.t. } c^n = w^m \mathbf{G} \text{ and } w^m \mathbf{H}^T = \mathbf{0} \right\}. \quad (5)$$

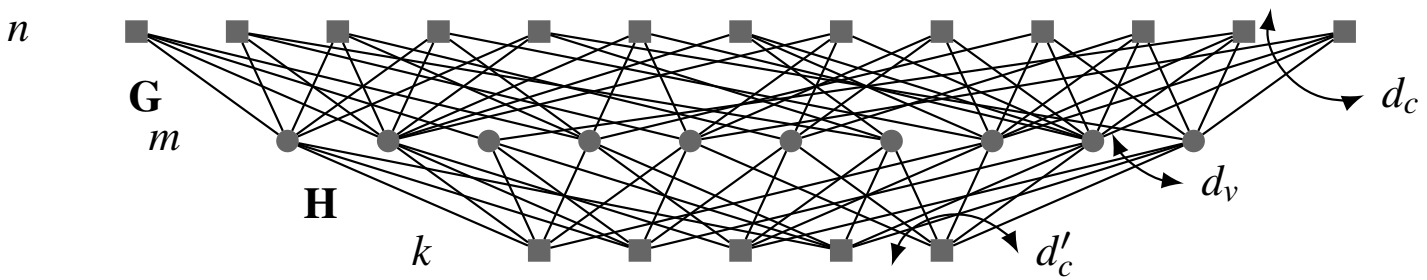


Figure 7: Compound LDGM-LDPC code. Top layer: $\mathbb{L}(\mathbf{G})$; Bottom layer: $\mathbb{M}(\mathbf{H})$.

Background

Linear Block Codes Considered

Compound LDGM-LDPC Codes

Consider $\mathbb{L}(\mathbf{G})$, $\mathbb{M}(\mathbf{H})$, where $\mathbf{G} \in \{0,1\}^{m \times n}$ and $\mathbf{H} \in \{0,1\}^{k \times m}$. Then,

$$\mathbb{C}(\mathbf{G},\mathbf{H}) = \left\{ c^n \in \{0,1\}^n : \exists w^m \in \{0,1\}^m \text{ s.t. } c^n = w^m \mathbf{G} \text{ and } w^m \mathbf{H}^T = \mathbf{0} \right\}. \quad (5)$$

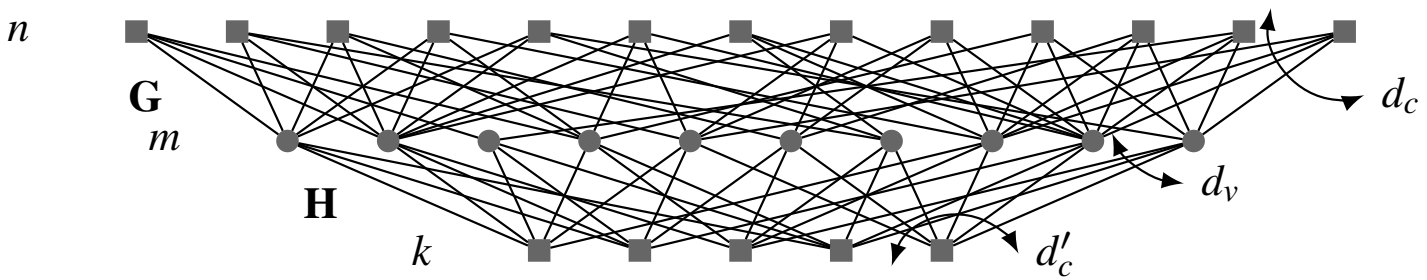


Figure 7: Compound LDGM-LDPC code. Top layer: $\mathbb{L}(\mathbf{G})$; Bottom layer: $\mathbb{M}(\mathbf{H})$.

The Compound LDGM-LDPC Ensemble $\mathfrak{C}(d_c, d_v, d'_c, m, n)$

Set of $\mathbb{C}(\mathbf{G},\mathbf{H})$ s.t.

- $\mathbf{G} \in \mathfrak{L}_P(d_c, m, n)$
- $\mathbf{H} \in (d_v, d'_c)$-regular LDPC ensemble

Background

Motivation

Motivation

- success of sparse-graph codes for channel coding $\longrightarrow$ source coding?
- Past research mainly focused on BSS case:
  - Performance bounds for *ensemble* [1] and *individual* LDGM codes [2];
  - optimality of the compound construction [3].
- BES(ε) is a generalization of a BSS
- Better insight into the behavior of sparse-graph codes

[1] A. Dimakis, M. J. Wainwright, and K. Ramchandran, "Lower bounds on the rate-distortion function of LDGM Codes," in *Proc. of the IEEE Inf. Theory Workshop*, 2007.

[2] Shrinivas Kudekar and Rudiger Urbanke, "Lower bounds on the rate-distortion function of individual LDGM Codes," in *5th Int. Symp. Turbo Codes and Related Topics*, Lausanne, Switzerland, 2008.

[3] M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, Mar. 2009.

Background

Main Achievements

Main Achievements

Sep. to Feb.	• DCC (Snowbird, Utah) conference paper [1]
Mar. to Aug.	• Presentation of [1] • ISITA (Taichung, Taiwan) conference paper [2]

[1] G. Demay, V. Rathi, and L. K. Rasmussen, "Rate Distortion Bounds for Binary Erasure Source Using Sparse Graph Codes," in *Proc. of the Data Compression Conference*, Snowbird, UT, Mar. 2010.

[2] G. Demay, V. Rathi, and L. K. Rasmussen, "Optimality of LDGM-LDPC Compound Codes for Lossy Compression of Binary Erasure Source," accepted to the *Int. Symp. Inf. Theory and its Applications*, Taichung, Taiwan, Oct. 2010.

Background

Main Achievements

Main Achievements

Sep. to Feb.	• DCC (Snowbird, Utah) conference paper [1]
Mar. to Aug.	• Presentation of [1] • ISITA (Taichung, Taiwan) conference paper [2]

[1] G. Demay, V. Rathi, and L. K. Rasmussen, "Rate Distortion Bounds for Binary Erasure Source Using Sparse Graph Codes," in *Proc. of the Data Compression Conference*, Snowbird, UT, Mar. 2010.

[2] G. Demay, V. Rathi, and L. K. Rasmussen, "Optimality of LDGM-LDPC Compound Codes for Lossy Compression of Binary Erasure Source," accepted to the *Int. Symp. Inf. Theory and its Applications*, Taichung, Taiwan, Oct. 2010.

Results to be presented

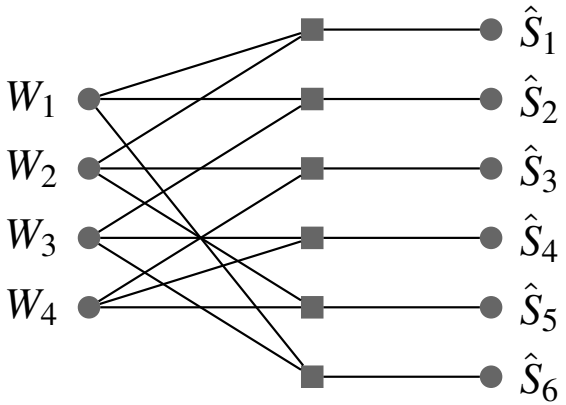
- Upper bounds on $R(D)$ for BES(ε) using
  - CRP LDGM Codes
  - Compound LDGM-LDPC Codes
- Optimality of the compound construction for lossy compression of BES

Overview

(Lower Bounds)

Lossy Compression using LDGM Codes

LDGM Codes as Lossy Compressor



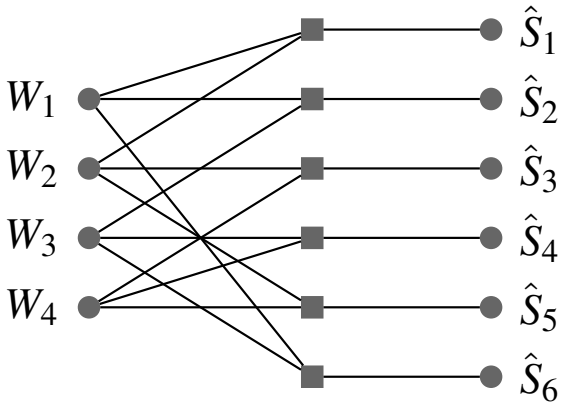
$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Figure 8: LDGM code used for lossy compression $n = 6$, $R = \frac{2}{3}$, and $L(x) = x^3$.

(Lower Bounds)

Lossy Compression using LDGM Codes

LDGM Codes as Lossy Compressor



$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Figure 8: LDGM code used for lossy compression $n = 6$, $R = \frac{2}{3}$, and $L(x) = x^3$.

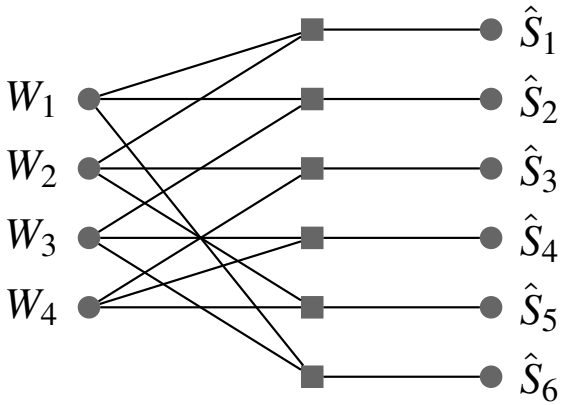
Compression/Reconstruction Process

- $s^n \in \mathcal{S}^n$

(Lower Bounds)

Lossy Compression using LDGM Codes

LDGM Codes as Lossy Compressor



$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Figure 8: LDGM code used for lossy compression $n = 6$, $R = \frac{2}{3}$, and $L(x) = x^3$.

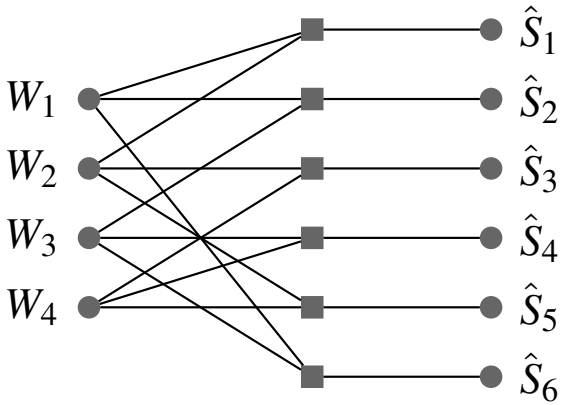
Compression/Reconstruction Process

- $s^n \in \mathcal{S}^n \xrightarrow{f} w^{nR} \in \mathcal{W}^{nR} = \mathbb{F}_2^{nR}$

(Lower Bounds)

Lossy Compression using LDGM Codes

LDGM Codes as Lossy Compressor



$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Figure 8: LDGM code used for lossy compression $n = 6$, $R = \frac{2}{3}$, and $L(x) = x^3$.

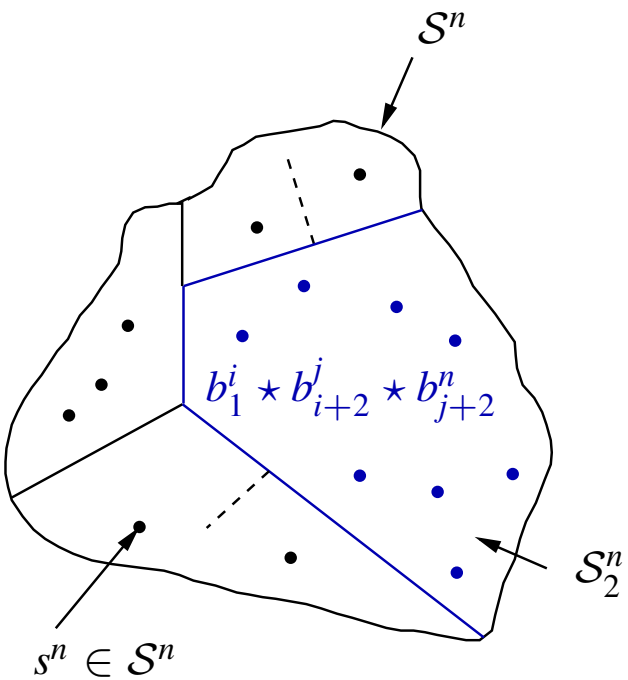
Compression/Reconstruction Process

- $s^n \in \mathcal{S}^n \xrightarrow{f} w^{nR} \in \mathcal{W}^{nR} = \mathbb{F}_2^{nR} \xrightarrow{g} \hat{s}^n \in \hat{\mathcal{S}}^n = \mathbb{F}_2^m$, s.t. $\hat{s}^n = w^{nR}\mathbf{G}$

(Lower Bounds)

Preliminaries

General Ideas - Partition of the Set of Source Sequences $\mathcal{S}^n$



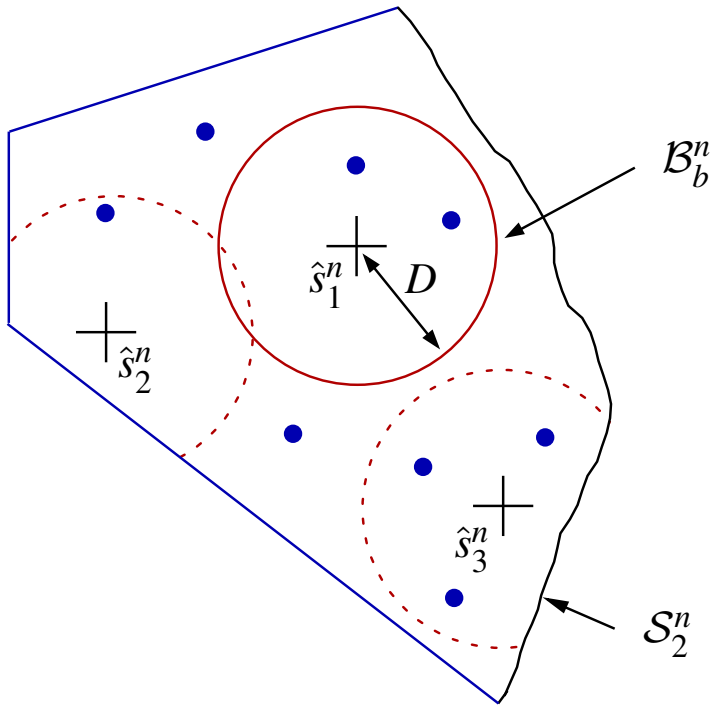
$$H_E(s^n) = \text{number of erasures in } s^n, \ \forall s^n \in \mathcal{S}^n$$

$$\mathcal{S}_b^n = \{s^n \in \mathcal{S}^n : \ H_E(s^n) = b\} \qquad (6)$$

(Lower Bounds)

Preliminaries

General Ideas - "Hamming" Balls in $\mathcal{S}_b^n$

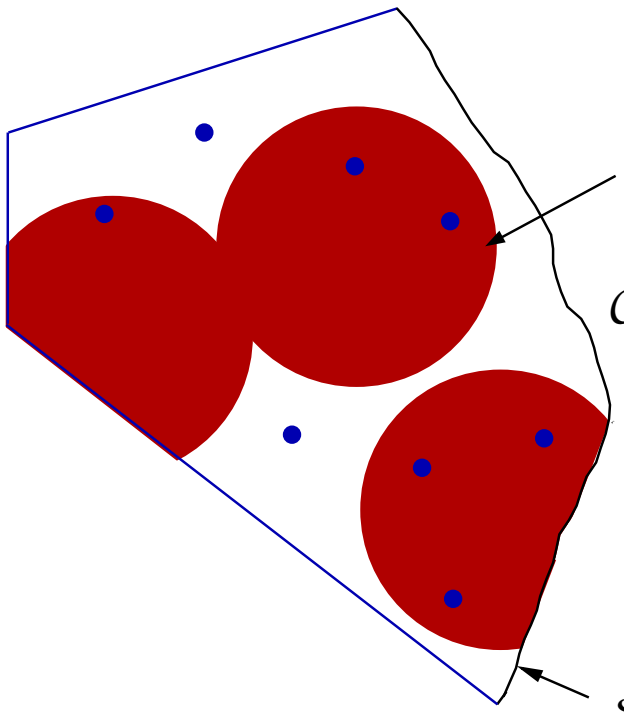


$$\forall \hat{s}^n \in \hat{\mathcal{S}}^n,$$
$$\mathcal{B}_b^n(\hat{s}^n, D) = \{s^n \in \mathcal{S}_b^n : \; d(s^n, \hat{s}^n) \leq D\} \quad (7)$$

(Lower Bounds)

Preliminaries

General Ideas - Union of "Hamming" Balls in $\mathcal{S}_b^n$



$$\mathcal{C}_b^n(D) = \bigcup_{\hat{s}^n \in \hat{\mathcal{S}}^n} \mathcal{B}_b^n(\hat{s}^n, D) \tag{8}$$

$$= \left\{ s^n \in \mathcal{S}_b^n : \exists \hat{s}^n \in \hat{\mathcal{S}}^n, d(s^n, \hat{s}^n) \leq D \right\}$$

(Lower Bounds)

Preliminaries

General Ideas - Summary

- $\mathcal{S}_b^n = \{s^n \in \mathcal{S}^n : H_E(s^n) = b\}$
- $\mathcal{B}_b^n(\hat{s}^n, D) = \{s^n \in \mathcal{S}_b^n : d(s^n, \hat{s}^n) \leq D\}$
- $\mathcal{C}_b^n(D) = \bigcup_{\hat{s}^n \in \hat{\mathcal{S}}^n} \mathcal{B}_b^n(\hat{s}^n, D)$
- Consider only $\mathcal{C}_{\varepsilon n}^n(D)$ and **upperbound** $|\mathcal{C}_{\varepsilon n}^n(D)|$

(Lower Bounds)

Preliminaries

Simplifications

Prove the Results for

- Regular generator node degree $L(x) = x^l$.
- Limit of infinite block-lengths

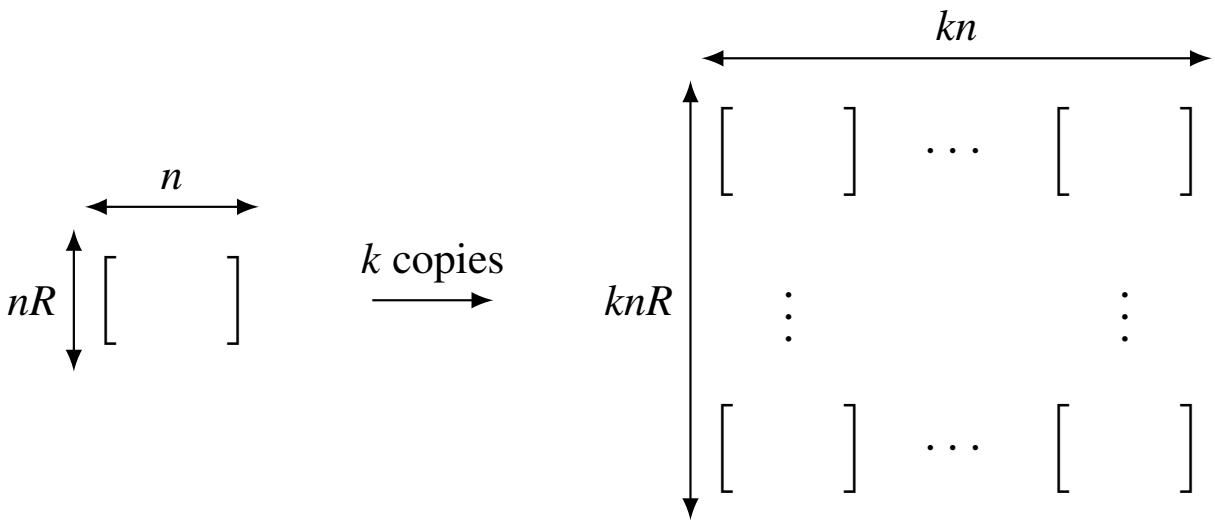


Figure 9: Construction of an arbitrarily large code with same R, D, and $L(x)$.

(Lower Bounds)

Preliminaries

Lemma I: Average Distortion and $|\mathcal{C}_{\varepsilon n}^{n}(D)|$

$$\lim_{n\to\infty} \frac{1}{n} \log \left(\varepsilon^{\varepsilon n} \left(\frac{1-\varepsilon}{2} \right)^{n-\varepsilon n} |\mathcal{C}_{\varepsilon n}^{n}(D)| \right) < 0, \tag{9}$$

then

$$\mathbb{E}\left[d(S^n, g(f(S^n)))\right] \geq D(1+o(1)) \tag{10}$$

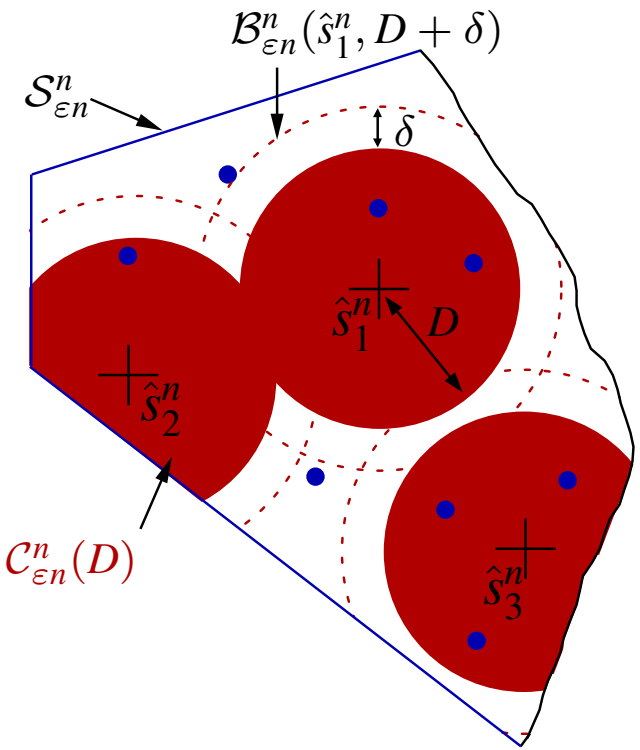
▸ **proof** ▸ **proof Theorem I** ▸ **proof Theorem II**

(Lower Bounds)

Lower Bound via Counting

The Counting Method

Goal: Upper bound $|\mathcal{C}_{\varepsilon n}^n(D)|$



- Pick $\delta \in \left[0, \frac{1-\varepsilon}{2} - D\right]$ and $\delta n \in \mathbb{N}$
- $\mathcal{B}_{\varepsilon n}^n(\hat{s}_2^n, D) \subset \mathcal{B}_{\varepsilon n}^n(\hat{s}_1^n, D+\delta) \Leftrightarrow d_H(\hat{s}_2^n, \hat{s}_1^n) \leq \delta n$
- Each small $\mathcal{B}_b^n(\hat{s}_j^n, D)$ is in $A_n(\delta n)$ big $\mathcal{B}_{\varepsilon n}^n(\hat{s}_1^n, D+\delta)$, $j \neq 1$

$$|\mathcal{C}_{\varepsilon n}^n(D)| \leq \frac{1}{A_n(\delta n)} \left| \bigcup_{\hat{s}^n \in \hat{\mathcal{S}}^n} \mathcal{B}_{\varepsilon n}^n(\hat{s}^n, D+\delta) \right| \qquad (11)$$

- Upper bound $|\mathcal{B}_{\varepsilon n}^n(\hat{s}^n, D+\delta)|$
- Lower bound $A_n(\delta n)$

(Lower Bounds)

Lower Bound via Counting

Theorem I: Bounds via Counting

Consider lossy compression of a BES(ε) using a LDGM code with

- *Blocklength n*
- *Generator node degree distribution $L(x)$*

$$f(x) = \prod_{i=0}^{d} (1 + x^i)^{L_i}, \quad a(x) = \sum_{i=0}^{d} i L_i \frac{x^i}{1 + x^i}, \tag{12}$$

$$R(x) = (1 - \varepsilon) \frac{\left(1 - h\left(\frac{x}{1+x}\right)\right)}{1 - \log_2\left(\frac{f(x)}{x^{a(x)}}\right)}, \tag{13}$$

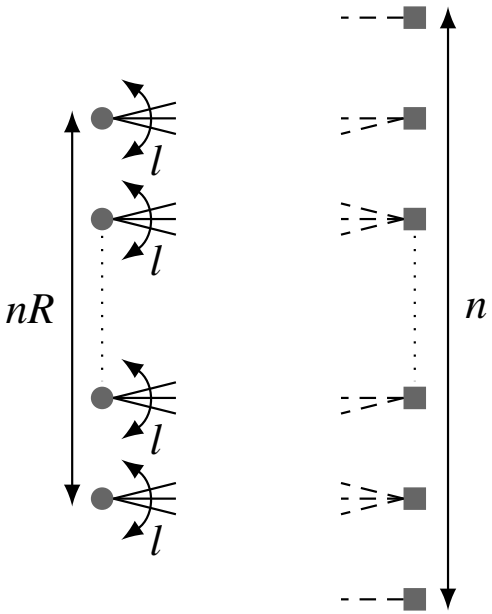
$$D(x) = \frac{x(1-\varepsilon)}{1+x} - R(x)a(x). \tag{14}$$

Then, the achievable rate-distortion performance of a LDGM code is lower bounded by the parametric curve $(D(x), R(x))$, $x \in [0, 1]$.

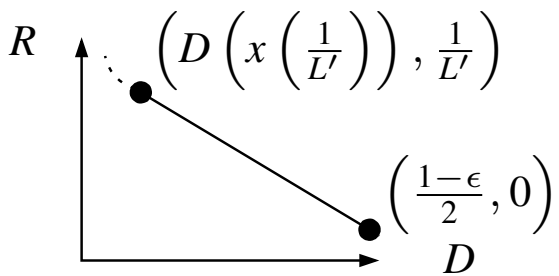
(Lower Bounds)

Lower Bound via Counting

Straight Line Bound



- $n \leq nRl$
  $R \leq \frac{1}{l} \Rightarrow n(1-Rl)$ nodes with distortion $\frac{1-\varepsilon}{2}$
- $L' = L'(1)$ average generator node degree
- $x\left(\frac{1}{L'}\right)$ unique solution of $R(x) = \frac{1}{L'}$



$$\forall R \in \left[0, \frac{1}{L'}\right], \; D = \frac{1-\varepsilon}{2}\left[1 - RL'\left(1 - \frac{2}{1-\varepsilon}D\left(x\left(\frac{1}{L'}\right)\right)\right)\right] \quad (15)$$

(Lower Bounds)

Lower Bound via Counting

Regular LDGM Codes

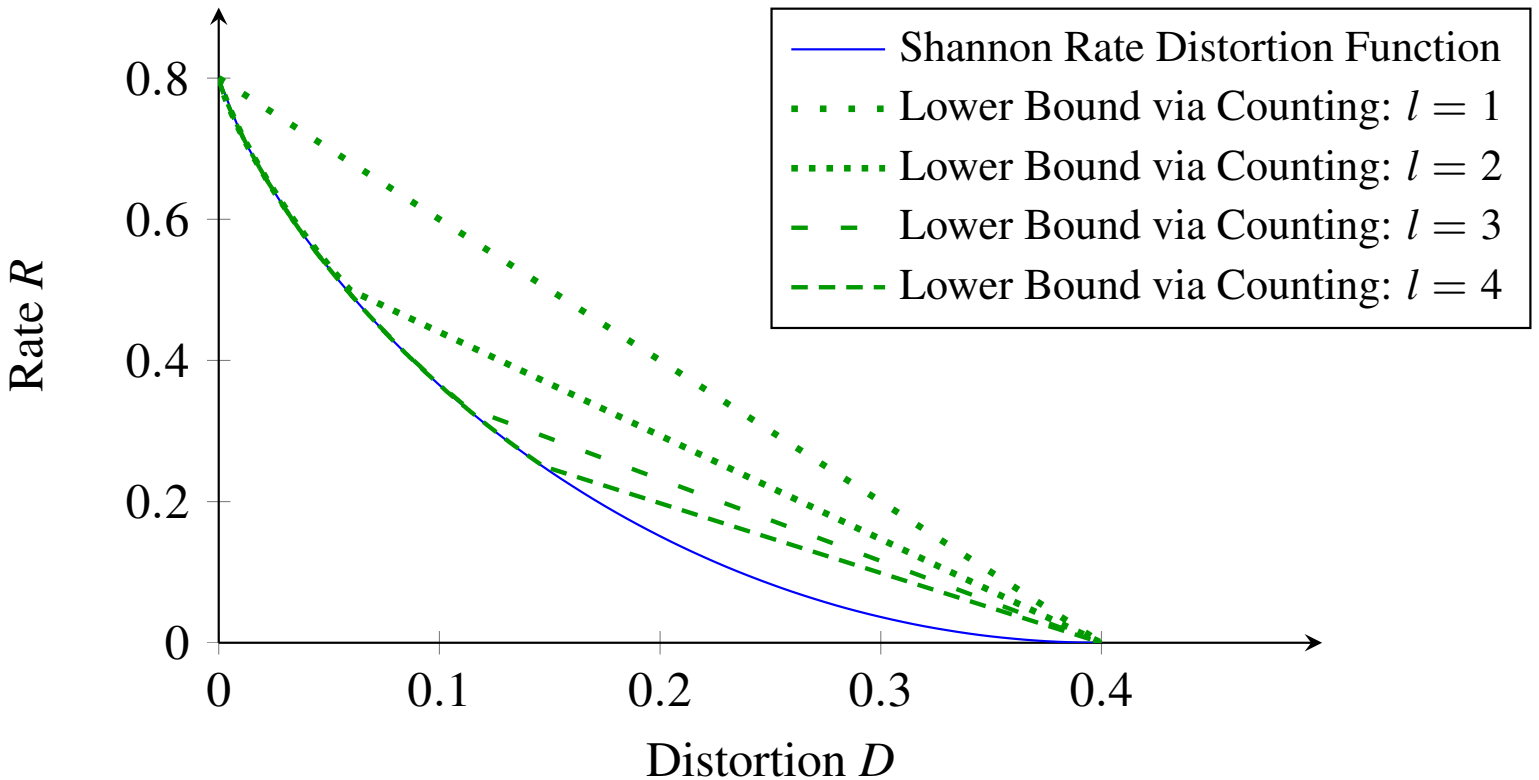


Figure 10: Rate-distortion peformances of generator regular LDGM codes for lossy compression of a BES(ε), with $\varepsilon = 0.2$.

(Lower Bounds)

Lower Bound via Test Channel

The Test Channel Method

Goal: Upper bound $|\mathcal{C}_{\varepsilon n}^n(D)|$

$$\begin{aligned}\sum_{s^n \in \mathcal{S}_{\varepsilon n}^n} \mathbb{P}\{S^n = s^n\} &\geq \sum_{s^n \in \mathcal{C}_{\varepsilon n}^n(D)} \mathbb{P}\{S^n = s^n\} \\ &\stackrel{?}{=} |\mathcal{C}_{\varepsilon n}^n(D)|\, \mathbb{P}\{S^n = s^n \mid s^n \in \mathcal{C}_{\varepsilon n}^n(D)\}\end{aligned}$$

- Lower bound $\mathbb{P}\{S^n = s^n \mid s^n \in \mathcal{C}_{\varepsilon n}^n(D)\}$

(Lower Bounds)

Lower Bound via Test Channel

Test Channel

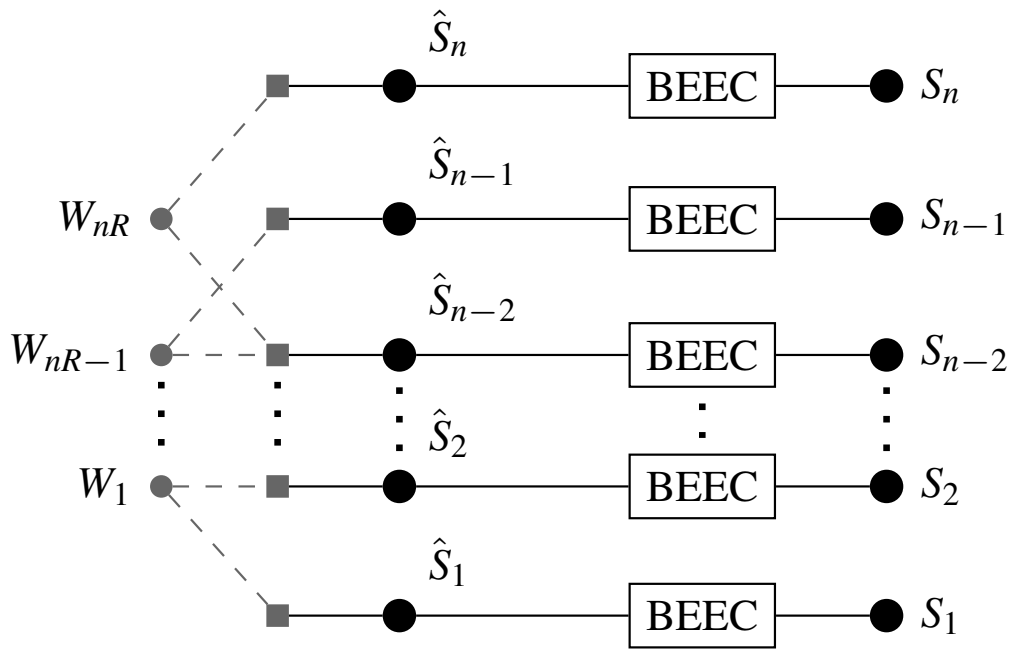


Figure 11: Test channel.

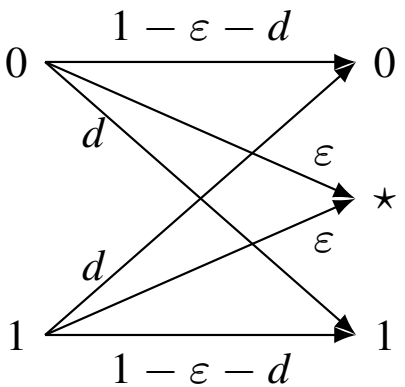


Figure 12: The binary error/erasure channel.

(Lower Bounds)

Lower Bound via Test Channel

Theorem II: Bound via Test Channel

Consider lossy compression of a BES(ε) using a LDGM code with

- *Blocklength n*
- *Generator node degree distribution $L(x) = x^l$*
- *Rate R*
- *Average normalized distortion D.*

Then, R is lower bounded by

$$R \geq \sup_{D \leq d \leq \frac{1-\varepsilon}{2}} \frac{(1-\varepsilon)\left(1-\log_2(1-\varepsilon)\right) + (1-D-\varepsilon)\log_2(1-\varepsilon-d) + D\log_2(d)}{1-\log_2\left(1+\left(\frac{d}{1-\varepsilon-d}\right)^l\right)} \tag{16}$$

◂ Theorem I ▸ proof Theorem II

(Lower Bounds)

Counting and Test Channel Methods Are Equivalent

Computation of (16) in Parametric Form

- $d \leftarrow \frac{d}{1-\varepsilon-d} = x$ in (16)
- Compute the sup in parametric form
- Identify $D(x)$ and $R(x)$

Theorem III: The bounds in Theorem I and II are equal

Consider lossy compression of a BES(ε) using a LDGM code with

- *Block-length n*
- *Generator node degree distribution $L(x) = x^l$*
- *Rate R*
- *Average normalized distortion D.*

Then the bounds on R in Theorem I and Theorem II are equal.

▸ **proof Theorem III**

Overview

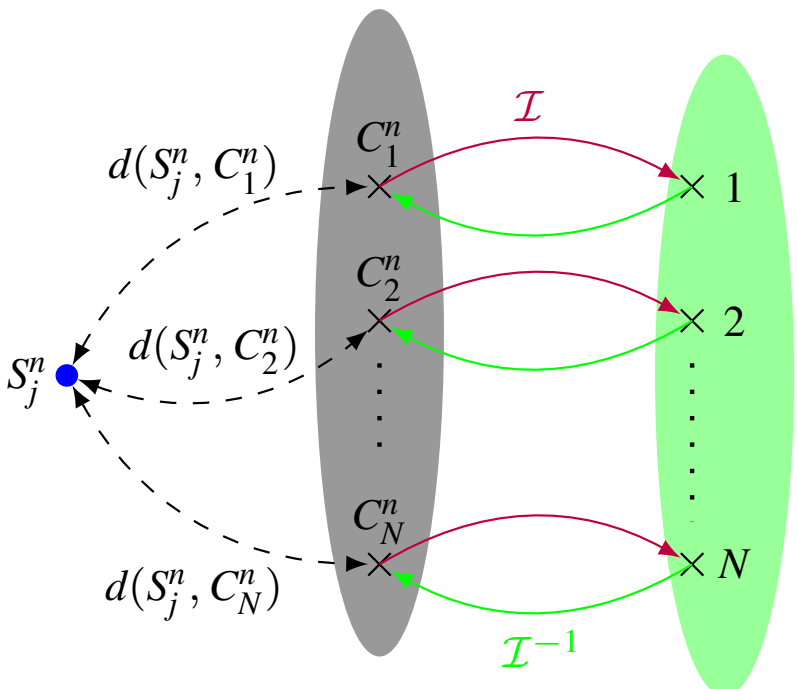
Upper Bounds

Preliminaries

Optimal Encoding

- $\mathcal{C} \in \mathfrak{L}_P\left(d_c, m, n\right) \cup \mathfrak{C}\left(d_c, d_v, d_c', m, n\right)$;

$$N = 2^{nR}, C_1^n = \underbrace{00\cdots0}_{n \text{ bits}}$$



$$f: \quad \mathcal{S}^n \rightarrow \{1, \cdots, N\}$$
$$s^n \mapsto \mathcal{I}\left(\arg\min_{c^n \in \mathcal{C}} \left\{d\left(s^n, c^n\right)\right\}\right),$$

$$g: \quad \{1, \cdots, N\} \rightarrow \hat{\mathcal{S}}^n$$
$$f(s^n) \mapsto \mathcal{I}^{-1}\left(f(s^n)\right).$$

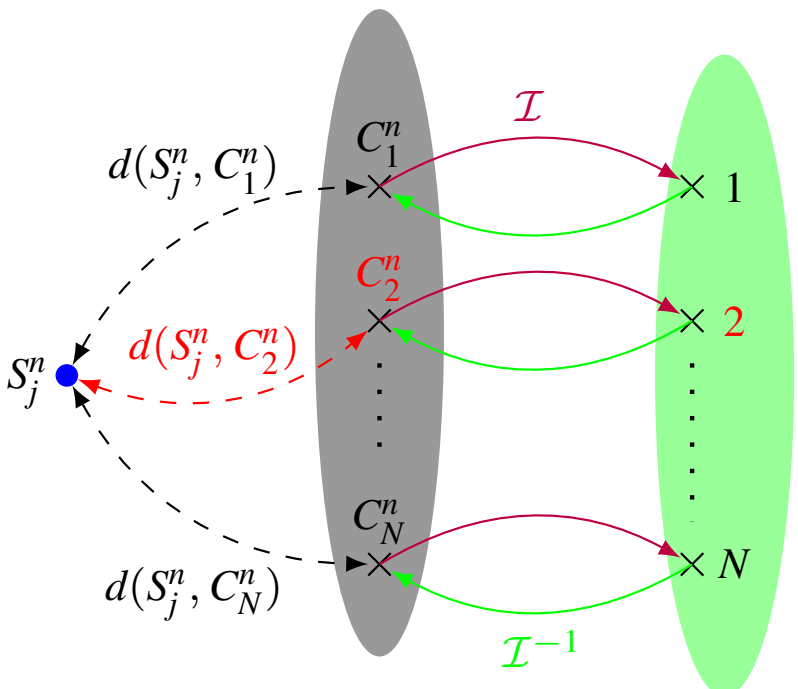
Upper Bounds

Preliminaries

Optimal Encoding

- $\mathcal{C} \in \mathfrak{L}_P\left(d_c, m, n\right) \cup \mathfrak{C}\left(d_c, d_v, d_c', m, n\right)$;

$$N = 2^{nR}, C_1^n = \underbrace{00\cdots0}_{n \text{ bits}}$$



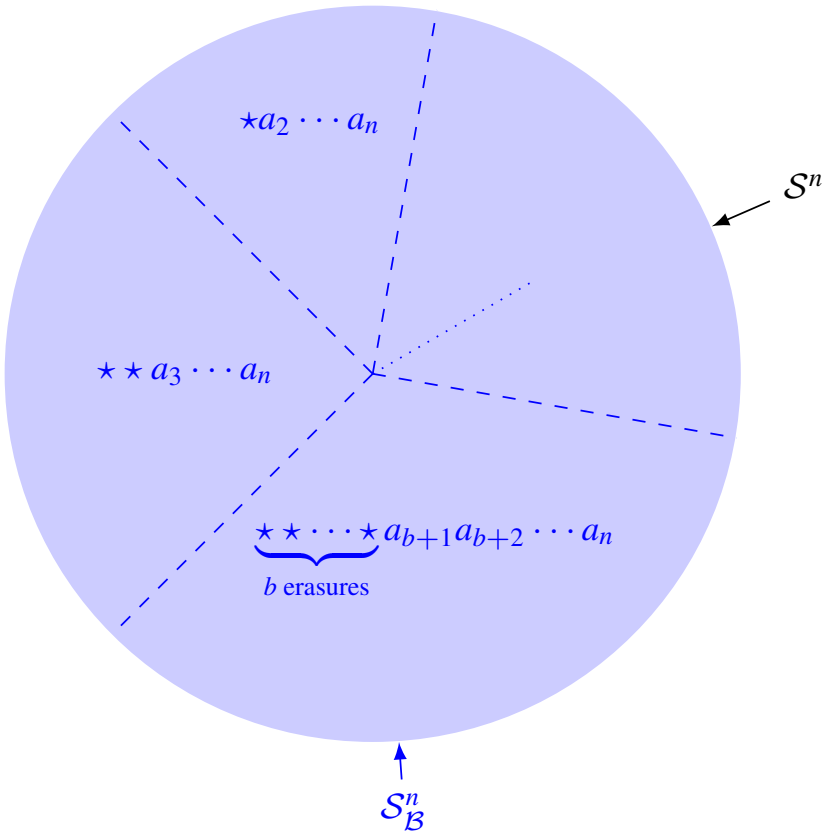
$$f: \quad \mathcal{S}^n \rightarrow \{1, \cdots, N\}$$
$$s^n \mapsto \mathcal{I}\left(\arg\min_{c^n \in \mathcal{C}} \left\{d\left(s^n, c^n\right)\right\}\right),$$

$$g: \quad \{1, \cdots, N\} \rightarrow \hat{\mathcal{S}}^n$$
$$f(s^n) \mapsto \mathcal{I}^{-1}\left(f(s^n)\right).$$

Upper Bounds

Preliminaries

General Ideas - "Partition" of the set of source sequences $\mathcal{S}^n$



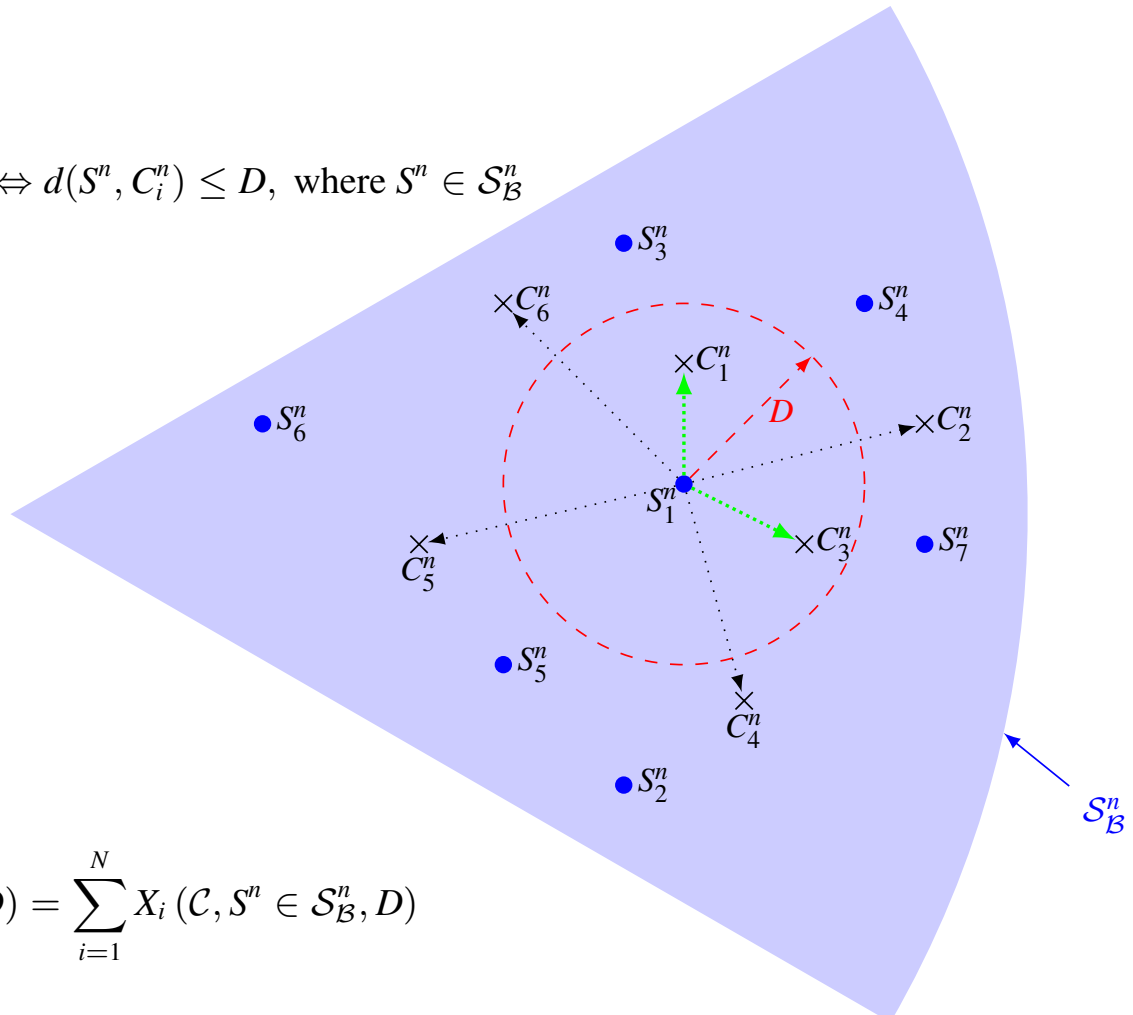
$$\mathcal{S}^n_{\mathcal{B}} = \left\{ s^n \in \mathcal{S}^n : s^n = \underbrace{\star\cdots\star}_{b \text{ erasures}} a_{b+1}\cdots a_n,\ a_i \in \{0,1\} \right\}$$

Upper Bounds

Preliminaries

General Ideas - Number of codewords within distortion D from $S^n \in \mathcal{S}_{\mathcal{B}}^n$

$X_i\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) = 1 \Leftrightarrow d(S^n, C_i^n) \leq D,$ where $S^n \in \mathcal{S}_{\mathcal{B}}^n$



$$Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) = \sum_{i=1}^{N} X_i\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)$$

Upper Bounds

Preliminaries

Lemma II: Upper Bound on the Average Distortion [1]

Assume that for a given distortion D and a given ensemble of codes, we have

$$\lim_{n\to\infty} \frac{1}{n} \log_2 \left[\sum_{b=0}^{n} \binom{n}{b} \varepsilon^b (1-\varepsilon)^{n-b} \mathbb{P}\left\{ Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) > 0 \right\} \right] \geq 0. \tag{17}$$

Then, $\forall \theta > 0$, there exists a code in the ensemble such that for sufficiently large blocklength n

$$\mathbb{E}\left[d\left(S^n, g\left(f\left(S^n\right)\right)\right) \right] < D + \theta. \tag{18}$$

[1] *M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," IEEE Trans. Inf. Theory, vol. 55, no. 3, Mar. 2009.*

Upper Bounds

Second Moment Method

The Second Moment Method [1]

Goal: Lower bound $\mathbb{P}\left\{Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) > 0\right\}$

Using Markov's inequality, we have

$$\mathbb{P}\left\{Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) > 0\right\} \geq \frac{\mathbb{E}\left[Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]^2}{\mathbb{E}\left[Z^2\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]} \tag{19}$$

[1] M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, Mar. 2009.

Upper Bounds

Second Moment Method

The Second Moment Method [1]

Goal: Lower bound $\mathbb{P}\left\{Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) > 0\right\}$

Using Markov's inequality, we have

$$\mathbb{P}\left\{Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) > 0\right\} \geq \frac{\mathbb{E}\left[Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]^2}{\mathbb{E}\left[Z^2\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]} \tag{19}$$

$\rightarrow$ lower bound 1^{st} moment $\mathbb{E}\left[Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]$;

$\rightarrow$ upper bound 2^{nd} moment $\mathbb{E}\left[Z^2\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]$.

[1] M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, Mar. 2009.

Upper Bounds

Second Moment Method

Lemma III: First Moment for any Linear Block Code

Consider any linear block code with rate R and blocklength n. The 1^{st} moment $\mathbb{E}\left[Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]$ is given by

$$\mathbb{E}\left[Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right] = 2^{nR}\mathbb{P}\left\{X_1\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) = 1\right\} \quad (20)$$

$$= 2^{nR}\sum_{j=0}^{nD}\binom{n-b}{j}\left(\frac{1}{2}\right)^{n-b} \quad (21)$$

- $\longrightarrow$ *exponential growth rate of* $\mathbb{E}\left[Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]$;

▸ **proof**

Upper Bounds

Second Moment Method

The Second Moment Method [1]

Goal: Lower bound $\mathbb{P}\left\{Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) > 0\right\}$

Using Markov's inequality, we have

$$\mathbb{P}\left\{Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) > 0\right\} \geq \frac{\mathbb{E}\left[Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]^2}{\mathbb{E}\left[Z^2\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]} \tag{22}$$

$\rightarrow$ lower bound 1^{st} moment $\mathbb{E}\left[Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]$;

$\rightarrow$ upper bound 2^{nd} moment $\mathbb{E}\left[Z^2\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right]$.

[1] M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, Mar. 2009.

Upper Bounds

Second Moment Method

Lemma IV: Second Moment for any Linear Code [1]

For any linear block code, the second moment satisfies the relation

$$\mathbb{E}\left[Z^2\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right] = \mathbb{E}\left[Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right] \times \sum_{j=1}^{N} \mathbb{P}\left\{X_j\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) = 1 \mid X_1\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) = 1\right\}. \quad (23)$$

[1] *E. Martinian and M. J. Wainwright, "Low-density codes achieve the rate-distortion bound," in Proc. of the Data Compression Conference, Snowbird, UT, Mar. 2006.*

$\longrightarrow$ **upper bound** $\sum_{j=1}^{N} \mathbb{P}\left\{X_j\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) = 1 \mid X_1\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) = 1\right\}$

Upper Bounds

Upper Bounds on the Rate-Distortion Performance

▸ CRP LDGM Ensemble ▸ Compound LDGM-LDPC Ensemble

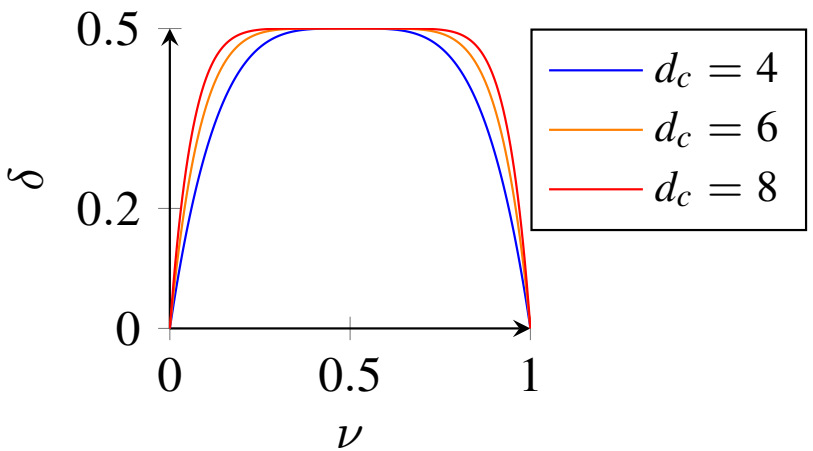
Lemma V: Distribution of random codewords generated by CRP LDGM codes

Consider the CRP LDGM ensemble $\mathfrak{L}_P(d_c, m, n)$ *and*

- *an information sequence* $w^m \in \{0,1\}^m$ *of weight* νm, $\nu \in [0,1]$;
- $\mathbb{L}(\mathbf{G})$, *selected uniformly at random in* $\mathfrak{L}_P(d_c, m, n)$;
- *a random codeword* $C^n(\nu)$ *generated by* $w^m\mathbf{G}$.

Then, $C^n(\nu)$ *is i.i.d. Bernoulli distributed with parameter*

$$\delta(\nu, d_c) \triangleq \frac{1}{2}\left[1 - (1-2\nu)^{d_c}\right]. \quad (24)$$

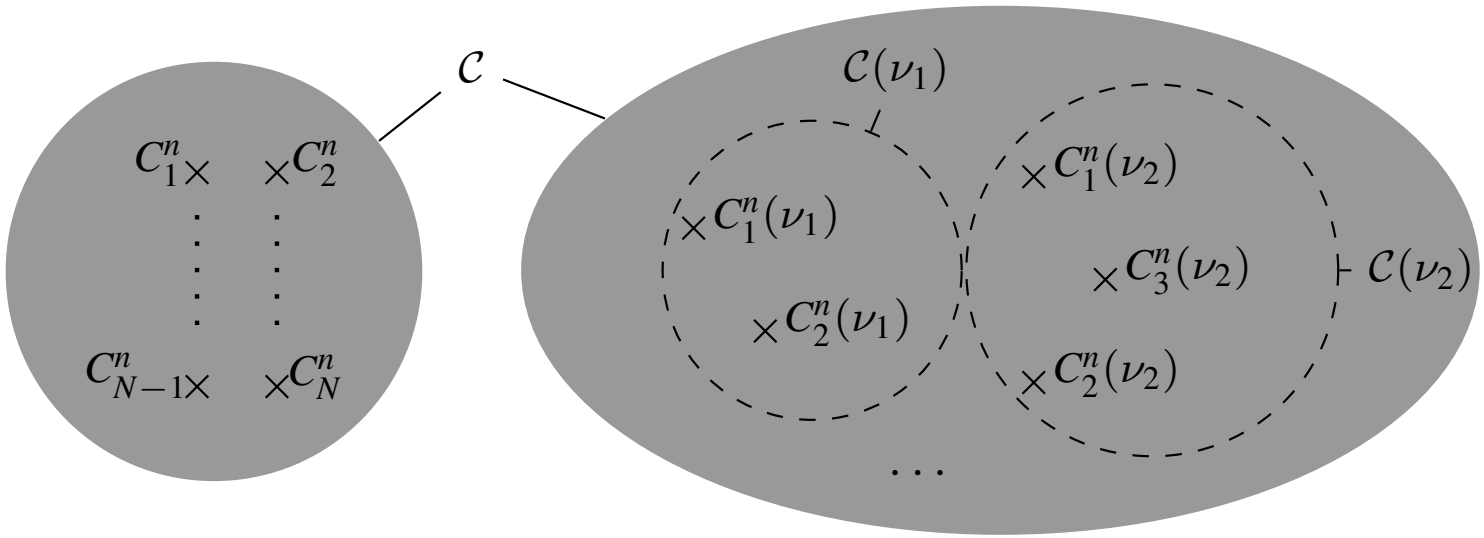


▸ proof

Upper Bounds

Upper Bounds on the Rate-Distortion Performance

Goal: Upper Bound $\sum_{j=1}^{N} \mathbb{P}\{X_j = 1 \mid X_1 = 1\}$



$$\mathbb{P}\left\{d\left(S^n, C_j^n\right) \le D \mid d\left(S^n, C_1^n\right) \le D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}$$

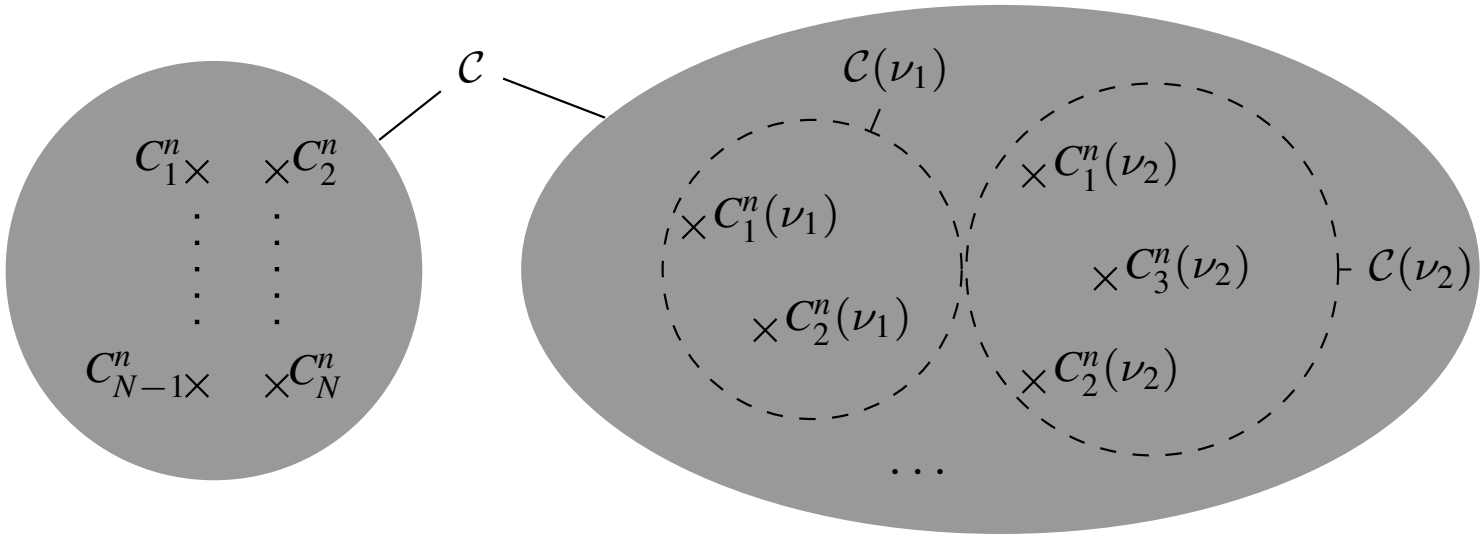
$$\updownarrow ?$$

$$\underbrace{\mathbb{P}\left\{d\left(C^n(\nu), S^n\right) \le D \mid d\left(C_1^n, S^n\right) \le D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}}_{\triangleq\, \mathcal{Q}(\nu,\beta),\ \beta = b/n}.$$

Upper Bounds

Upper Bounds on the Rate-Distortion Performance

Goal: Upper Bound $\sum_{j=1}^{N} \mathbb{P}\{X_j = 1 \mid X_1 = 1\}$



$$\sum_{j=1}^{N} \mathbb{P}\left\{d\left(S^n, C_j^n\right) \leq D \mid d\left(S^n, C_1^n\right) \leq D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}$$

$$= \sum_{\substack{\nu \in [0,1]: \\ \nu m \in \mathbb{N}}} \sum_{j=1}^{|\mathcal{C}(\nu)|} \underbrace{\mathbb{P}\left\{d\left(C^n(\nu), S^n\right) \leq D \mid d\left(C_1^n, S^n\right) \leq D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}}_{\triangleq\, \mathcal{Q}(\nu, \beta),\ \beta = b/n}.$$

Upper Bounds

Upper Bounds on the Rate-Distortion Performance

Goal: Upper Bound $\sum_{j=1}^{N} \mathbb{P}\{X_j = 1 \mid X_1 = 1\}$

$$\sum_{j=1}^{N} \mathbb{P}\left\{d\left(S^n, C_j^n\right) \leq D \mid d\left(S^n, C_1^n\right) \leq D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}$$

$$= \sum_{\substack{\nu \in [0,1]: \\ \nu m \in \mathbb{N}}} \sum_{j=1}^{|\mathcal{C}(\nu)|} \underbrace{\mathbb{P}\left\{d\left(C^n(\nu), S^n\right) \leq D \mid d\left(C_1^n, S^n\right) \leq D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}}_{\triangleq\, \mathcal{Q}(\nu,\beta),\ \beta = b/n}.$$

- For CRP LDGM codes

$$\sum_{j=1}^{N} \mathbb{P}\{X_j = 1 \mid X_1 = 1\} = \sum_{\substack{\nu \in [0,1]: \\ \nu m \in \mathbb{N}}} \binom{m}{\nu m} \mathcal{Q}(\nu, \beta).$$

Upper Bounds

Upper Bounds on the Rate-Distortion Performance

Goal: Upper Bound $\sum_{j=1}^{N} \mathbb{P}\{X_j = 1 \mid X_1 = 1\}$

$$\sum_{j=1}^{N} \mathbb{P}\left\{d\left(S^n, C_j^n\right) \leq D \mid d\left(S^n, C_1^n\right) \leq D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}$$

$$= \sum_{\substack{\nu \in [0,1]: \\ \nu m \in \mathbb{N}}} \sum_{j=1}^{|\mathcal{C}(\nu)|} \underbrace{\mathbb{P}\left\{d\left(C^n(\nu), S^n\right) \leq D \mid d\left(C_1^n, S^n\right) \leq D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}}_{\triangleq\, \mathcal{Q}(\nu,\beta),\ \beta = b/n}.$$

- For CRP LDGM codes

$$\sum_{j=1}^{N} \mathbb{P}\{X_j = 1 \mid X_1 = 1\} = \sum_{\substack{\nu \in [0,1]: \\ \nu m \in \mathbb{N}}} \binom{m}{\nu m} \mathcal{Q}(\nu, \beta).$$

$\longrightarrow$ upper bound $\mathcal{Q}(\nu, \beta)$ for $\beta = \varepsilon$

Upper Bounds

Upper Bounds on the Rate-Distortion Performance

Goal: Upper Bound $\sum_{j=1}^{N} \mathbb{P}\{X_j = 1 \mid X_1 = 1\}$

$$\sum_{j=1}^{N} \mathbb{P}\left\{d\left(S^n, C_j^n\right) \leq D \mid d\left(S^n, C_1^n\right) \leq D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}$$

$$= \sum_{\substack{\nu \in [0,1]: \\ \nu m \in \mathbb{N}}} \sum_{j=1}^{|\mathcal{C}(\nu)|} \underbrace{\mathbb{P}\left\{d\left(C^n(\nu), S^n\right) \leq D \mid d\left(C_1^n, S^n\right) \leq D, S^n \in \mathcal{S}_{\mathcal{B}}^n\right\}}_{\triangleq\, \mathcal{Q}(\nu, \beta),\ \beta = b/n}.$$

- For Compound LDGM-LDPC Codes

$$\sum_{j=1}^{N} \mathbb{P}\{X_j = 1 \mid X_1 = 1\} = \sum_{\substack{\nu \in [0,1]: \\ \nu m \in \mathbb{N}}} \mathcal{A}_m(\nu)\, \mathcal{Q}(\nu, \beta),$$

$\mathcal{A}_m(\nu)$ number of codewords of weight νm of the LDPC code.
$\longrightarrow$ upper bound $\mathcal{Q}(\nu, \beta)$ for $\beta = \varepsilon$
$\longrightarrow$ upper bound $\mathcal{A}_m(\nu)$

Upper Bounds

Upper Bounds on the Rate-Distortion Performance

Consider lossy compression of a BES(ε) using CRP LDGM codes from $\mathfrak{L}_P(d_c, m, n)$ or compound LDGM-LDPC codes from $\mathfrak{C}(d_c, d_v, d'_c, m, n)$ with

- blocklength n;
- rate R;
- distortion constraint D;
- $\frac{1}{n}\log_2 \mathcal{Q}(\nu, \beta) \leq F(\delta(\nu, d_c), \beta, D) + o(1)$.

Theorem IV: Upper Bounds for $\mathfrak{L}_P(d_c, m, n)$

There exists a code in $\mathfrak{L}_P(d_c, m, n)$ which achieves an average distortion at most D, if

$$R \geq \max_{\nu \in [0,1]} \frac{R_\varepsilon^{sh}(D) + F(\delta(\nu, d_c), \varepsilon, D)}{1 - h(\nu)}.$$

▸ **proof**

Theorem V: Upper Bounds for $\mathfrak{C}(d_c, d_v, d'_c, m, n)$

- $\frac{1}{m}\log_2 \mathcal{A}_m(\nu) \leq B(\nu) + o(1)$

There exists a code in $\mathfrak{C}(d_c, d_v, d'_c, m, n)$ which achieves an average distortion at most D, if

$$R \geq \max_{\nu \in [0,1]} \frac{R_\varepsilon^{sh}(D) + F(\delta(\nu, d_c), \varepsilon, D)}{1 - \frac{B(\nu)}{R_H}}.$$

Upper Bounds

Source Coding Optimality of the Compound LDGM-LDPC Ensemble

Rate Lower Bounds

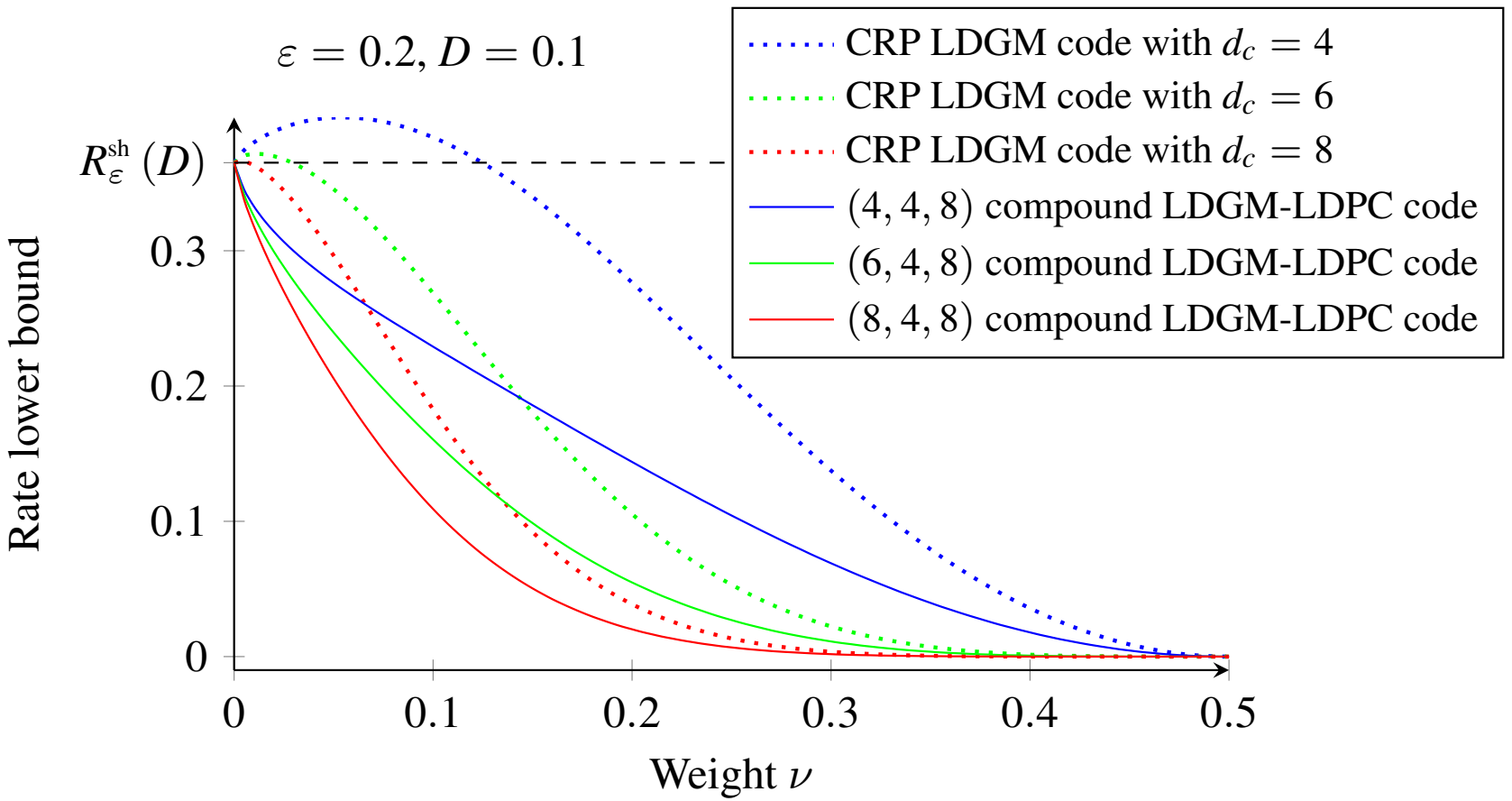


Figure 13: Rate lower bounds for the CRP LDGM and the compound LDGM-LDPC ensembles.

Upper Bounds

Source Coding Optimality of the Compound LDGM-LDPC Ensemble

Theorem VI: Source Coding Optimality of the Compound Construction

Consider lossy compression of BES(ε) using the compound construction. Given

- *a distortion D, $D \leq \frac{1-\varepsilon}{2}$;*
- *a desired compression rate R, $R < R_{\varepsilon}^{sh}(D)$,*

there exist degrees (d_c, d_v, d_c') independent of the blocklength, and a compound code $\mathbb{C}(\mathbf{G}, \mathbf{H}) \in \mathfrak{C}(d_c, d_v, d_c', m, n)$ with rate R which achieves average distortion D.

▸ **proof**

Conclusion

Conclusion

Key Points

- Lossy compression of a ternary source, the BES
- Derived upper bounds on $R(D)$ for the BES using
  - the CRP LDGM ensemble
  - the compound LDGM-LDPC ensemble
- Proved the source coding optimality of the compound LDGM-LDPC construction for lossy compression of a BES

Conclusion

Key Points

- Lossy compression of a ternary source, the BES
- Derived upper bounds on $R(D)$ for the BES using
  - the CRP LDGM ensemble
  - the compound LDGM-LDPC ensemble
- Proved the source coding optimality of the compound LDGM-LDPC construction for lossy compression of a BES

Added Value of this Internship

- 2 conference papers + 1 presentation in the U.S.A.;
- tremendous international experience;
- discover new knowledge;
- advanced technical knowledge;
- continue doing research after my master of engineering $\longrightarrow$ PhD

Thank you! & Questions?

Overview

Proofs

Proof of Lemma I

Lemma I

Key Steps (1/2)

- $|\mathcal{S}_b^n| = \binom{m}{b} 2^{m-b}$ and $\forall s \in \mathcal{S}_b^n, d\left(s, g\left(f\left(s\right)\right)\right) \geq \begin{cases} 0, & \text{if } s \in \mathcal{C}_b^n(D), \\ mD, & \text{if } s \notin \mathcal{C}_b^n(D). \end{cases}$

  $\delta(m) = \sqrt{m}(\log m)^{1/3}$

Proofs

Proof of Lemma I

Lemma I

Key Steps (1/2)

- $|\mathcal{S}_b^n| = \binom{m}{b} 2^{m-b}$ and $\forall s \in \mathcal{S}_b^n, d\left(s, g\left(f\left(s\right)\right)\right) \geq \begin{cases} 0, & \text{if } s \in \mathcal{C}_b^n(D), \\ mD, & \text{if } s \notin \mathcal{C}_b^n(D). \end{cases}$

  $\delta(m) = \sqrt{m}(\log m)^{1/3}$

- $\frac{1}{m}\mathbb{E}\left[d(S, g\left(f(S)\right))\right]$

$$= \frac{1}{m} \sum_{b=0}^{m} \mathbb{P}\left\{S \in \mathcal{S}_b^n\right\} \mathbb{E}\left[d(S, g\left(f(S)\right)) \middle| S \in \mathcal{S}_b^n\right]$$

$$\geq \frac{1}{m} \sum_{b=\epsilon m-\delta(m)}^{\epsilon m+\delta(m)} \mathbb{P}\left\{S \in \mathcal{S}_b^n\right\} \sum_{s \in \mathcal{S}_b^n} \mathbb{P}\left\{S = s | S \in \mathcal{S}_b^n\right\} d\left(s, g(f(s))\right)$$

Proofs

Proof of Lemma I

Key Steps (2/2)

- $\frac{1}{m}\mathbb{E}\left[d(S, g\left(f(S)\right))\right]$

$$\geq \frac{1}{m} \sum_{b=\epsilon m-\delta(m)}^{\epsilon m+\delta(m)} \binom{m}{b} \epsilon^b (1-\epsilon)^{m-b} \sum_{s \in \mathcal{S}_b^n \setminus \mathcal{C}_b^n(D)} \frac{Dm}{\binom{m}{b} 2^{m-b}}$$

$$\geq \underbrace{D \sum_{b=\epsilon m-\delta(m)}^{\epsilon m+\delta(m)} \binom{m}{b} \epsilon^b (1-\epsilon)^{m-b}}_{A=1+o(1)}$$

$$\underbrace{-D \sum_{b=\epsilon m-\delta(m)}^{\epsilon m+\delta(m)} \epsilon^b \left(\frac{1-\epsilon}{2}\right)^{m-b} \left|\mathcal{C}_b^n(D)\right|}_{B}.$$

Proofs

Proof of Theorem I

Theorem I

Key Steps (1/2): Upper Bound on $|\mathcal{C}^n_{\varepsilon n}(D)|$

- $w \in \mathbb{N}$ s.t. $Dm + w \leq m\frac{(1-\varepsilon)}{2}$ and $A_m(w) = \left|\left\{\hat{S} \in \hat{\mathcal{S}} :\ \mathrm{W_H}(\hat{S}) \leq w\right\}\right|$

  saddle point eq $\Rightarrow$ coef $\left(f(x)^{mR}, x^w\right) \leq \frac{f(x_\omega)^{mR}}{x_\omega^w} \leq A_m(w)$,

  $x_\omega > 0$, unique solution to $a(x) = \omega$, and $\omega = w/mR$.

$$|\mathcal{C}^n_{\varepsilon n}(D)| = \left|\bigcup_{\hat{s}\in\hat{\mathcal{S}}} \mathcal{B}_{\epsilon m}(\hat{s}, D)\right| \leq \frac{1}{A_m(w)} \sum_{\hat{s}\in\hat{\mathcal{S}}} \left|\mathcal{B}_{\epsilon m}\left(\hat{s}, D + \frac{w}{m}\right)\right|$$

$$\left|\mathcal{B}_{\epsilon m}\left(\hat{s}, D + \frac{w}{m}\right)\right| = \binom{m}{\epsilon m} \sum_{i=0}^{Dm+w} \binom{m - \epsilon m}{m - \epsilon m - i} \leq 2^{mh(\epsilon)} 2^{(m-\epsilon m)h\left(\frac{Dm+w}{m(1-\epsilon)}\right) + o(m-\epsilon m)}$$

Proofs

Proof of Theorem I

Key Steps (2/2)

$$|\mathcal{C}_{\varepsilon n}^{n}(D)| \leq 2^{m\left[-R\log_2 \frac{f(x_\omega)}{x_\omega^{a(x_\omega)}}+R+h(\epsilon)+(1-\epsilon)h\left(\frac{D+Ra(x_\omega)}{1-\epsilon}\right)+\right]+o(m-\epsilon m)}$$

Lemma I

$$\lim_{m\to\infty}\left[\frac{1}{m}\log_2\left(\varepsilon^{m\varepsilon}\left(\frac{1-\varepsilon}{2}\right)^{m(1-\varepsilon)}|\mathcal{C}_{\varepsilon n}^{n}(D)|\right)\right] \leq g(D,R),$$

$$g(D,R) = \inf_{\substack{D+a(x)R\leq\frac{1-\varepsilon}{2}\\ x\geq 0}} \underbrace{-R\log_2\frac{f(x)}{x^{a(x)}}+R+(1-\varepsilon)\left(h\left(\frac{D+Ra(x)}{1-\varepsilon}\right)-1\right)}_{h_1(x)}$$

Condition $g(D,R)=0 \Rightarrow \dfrac{D+Ra(x)}{1-\varepsilon} = \dfrac{x}{1+x}.$

Proofs

Proof of Theorem II

Theorem II

Key Steps (1/3)

- $A_m(w) = \left| \left\{ \hat{S} \in \hat{\mathcal{S}} : \ \mathrm{W_H}(\hat{S}) = w \right\} \right|$ and

$$\frac{Rc^l}{1+c^l} < \frac{1}{l} \Rightarrow \sum_{w=0}^{m} A_m(w) c^w \geq \frac{1}{m} \left(1 + c^l\right)^{mR},$$

where $c = \frac{d}{1-\varepsilon-d}$ and $D \leq d \leq \frac{1-\varepsilon}{2}$

- $$\binom{m}{\varepsilon m} \varepsilon^{\varepsilon m} (1-\varepsilon)^{m-\varepsilon m} = \sum_{s \in \mathcal{S}^n_{\varepsilon n}} \mathbb{P}\{S = s\} \geq \sum_{s \in \mathcal{C}^n_{\varepsilon n}(D)} \mathbb{P}\{S = s\}.$$

$$s \in \mathcal{C}^n_{\varepsilon n}(D) \Rightarrow \exists \hat{s} \in \hat{\mathcal{S}} : \ d(s, \hat{s}) \leq Dm$$

Proofs

Proof of Theorem II

Key Steps (2/3): Lower Bound on $\mathbb{P}\{S = s\}$

$$
\begin{aligned}
\mathbb{P}\{S=s\} &= \sum_{w=0}^{m} \sum_{\hat{s}' \in \hat{\mathcal{S}}: d(\hat{s}',\hat{s})=w} \mathbb{P}\left\{S=s, \hat{S}=\hat{s}'\right\}, \\
&= 2^{-mR} \varepsilon^{\varepsilon m} (1-\varepsilon-d)^{m-\varepsilon m} \sum_{w=0}^{m} \sum_{\hat{s}' \in \hat{\mathcal{S}}: d(\hat{s}',\hat{s})=w} c^{d(s,\hat{s}')}, \\
&\geq 2^{-mR} \varepsilon^{\varepsilon m} (1-\varepsilon-d)^{m-\varepsilon m} \sum_{w=0}^{m} \sum_{\hat{s}' \in \hat{\mathcal{S}}: d(\hat{s}',\hat{s})=w} c^{d(s,\hat{s})+d(\hat{s},\hat{s}')}, \\
&\geq 2^{-mR} e^{\varepsilon m} (1-\varepsilon-d)^{m-\varepsilon m} \sum_{w=0}^{m} A_m(w) c^{Dm+w}, \\
&\geq \frac{1}{m} 2^{-mR} \left(1+c^{l}\right)^{mR} \varepsilon^{\varepsilon m} (1-\varepsilon-d)^{m(1-D)-\varepsilon m} d^{Dm}.
\end{aligned}
$$

Proofs

Proof of Theorem II

Key Steps (3/3): Lower Bound on $|\mathcal{C}_{\varepsilon n}^{n}(D)|$

$$|\mathcal{C}_{\varepsilon n}^{n}(D)| \leq m 2^{mR}(1+c^{l})^{-mR}(1-\varepsilon-d)^{-m(1-D)}(1-\varepsilon)^{m} d^{-Dm} \binom{m}{\varepsilon m} \left(\frac{1-\varepsilon-d}{1-\varepsilon}\right)^{\varepsilon m}.$$

Lemma I $\Rightarrow$

$$R - R\log_2(1+c^{l}) + (1-\varepsilon)(-1+\log_2(1-\varepsilon)) - (1-D-\varepsilon)\log_2(1-\varepsilon-d) - D\log_2(d) < 0,$$

then the distortion is at least D.

- Prove $\frac{Rc^{l}}{1+c^{l}} < \frac{1}{l}$

Proofs

Proof of Theorem III

Theorem III

Key Steps

Theorem II

$$v(x) = \frac{(1-\varepsilon)\left(1-\log_2(1-\varepsilon)\right) + (1-D-\varepsilon)\log_2\left(\frac{1-\varepsilon}{1+x}\right) + D\log_2\left(\frac{(1-\varepsilon)x}{1+x}\right)}{1-\log_2\left(1+x^l\right)},$$

$x = \frac{d}{1-\varepsilon-d}$.

$$\frac{dv(x)}{dx} = 0 \Leftrightarrow D = (1-\varepsilon)\frac{x}{1+x} - a(x)(1-\varepsilon)\frac{\frac{x}{1+x}\log_2(x) + 1 - \log_2(1+x)}{1-\log_2\left(\frac{f(x)}{x^{a(x)}}\right)},$$

where

$$f(x) = 1+x^l, a(x) = \frac{lx^l}{1+x^l}.$$

Proofs

Proof of Lemma III

Lemma III

Key Steps

- symmetry of the code construction $\Rightarrow$

$$\begin{aligned}\mathbb{E}\left(Z\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right) &= \sum_{i=1}^{N} \mathbb{E}\left(X_i\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right)\right) \\ &= 2^{nR}\mathbb{P}\left\{X_1\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) = 1\right\}\end{aligned}$$

- $\mathbb{P}\left\{X_1\left(\mathcal{C}, S^n \in \mathcal{S}_{\mathcal{B}}^n, D\right) = 1\right\} = \sum_{j=0}^{nD} \binom{n-b}{j} \frac{1}{2^{n-b}}$
- $$\max_{j \in \{0,1,\cdots,nD\}} \binom{n-b}{j} = \begin{cases} \binom{n-b}{nD}, & \text{if } nD \leq \frac{n-b}{2}, \\ \binom{n-b}{\frac{n-b}{2}}, & \text{otherwise.} \end{cases}$$

Proofs

Proof of Lemma V

Lemma V

Key Steps

- for a code bit $C_i(\nu)$, let $V_j(i)$ be an indicator variable, which is equal to one iff the j^{th} edge starting from $C_i(\nu)$ ends up to an information bit being equal to one,

$$C_i(\nu) = V_1(i) \oplus V_2(i) \oplus \cdots \oplus V_{d_c}(i).$$

- $\mathbb{P}\{V_j(i) = 1\} = \nu$,

$$\mathbb{P}\{C_i(\nu) = 1\} = \sum_{\substack{k=0,\\ k \text{ odd}}}^{d_c} \binom{d_c}{k} \nu^k (1-\nu)^{d_c-k},$$

$$\mathbb{P}\{C_i(\nu) = 0\} = \sum_{\substack{k=0,\\ k \text{ even}}}^{d_c} \binom{d_c}{k} \nu^k (1-\nu)^{d_c-k}.$$

- $\mathbb{P}\{C_i(\nu) = 1\} = \frac{1}{2}\left[1 - \left(\mathbb{P}\{C_i(\nu) = 0\} - \mathbb{P}\{C_i(\nu) = 1\}\right)\right]$
- expand $((1-\nu) - \nu)^{d_c}$.

Proofs

Proof of Lemma VI

Lemma VI

Upper Bound on the Exponential Growth Rate of $\mathcal{Q}$

Let $\beta < 1 - 2D$. For a randomly chosen code from $\mathfrak{L}_P(d_c, m, n)$ or $\mathfrak{C}(d_c, d_v, d_c', m, n)$, the exponential growth rate of $\mathcal{Q}$ is upper bounded as

$$\frac{1}{n}\log_2 \mathcal{Q}(\nu, \beta) \leq F(\delta(\nu, d_c), \beta, D) + o(1), \tag{25}$$

where

$$F(\delta(\nu, d_c), \beta, D) = \inf_{\lambda<0} \max_{\tau \in [0,D]} \left\{ (1-\beta)\left[h\left(\frac{\tau}{1-\beta}\right) - h\left(\frac{D}{1-\beta}\right)\right] + \tau \log_2\left(f_1(\gamma, \lambda)\right) + (1-\beta-\tau)\log_2\left(f_1(1-\gamma, \lambda)\right) - \frac{\lambda D}{\log 2} \right\},$$

with $f_1(\gamma, \lambda) \triangleq (1-\gamma)e^{\lambda} + \gamma$.

Proofs

Proof of Lemma VI

Key Steps (1/2)

- let T be a RV s.t. $T = w_H(S^n)$, knowing that $S^n \in \mathcal{S}_{\mathcal{B}}^n$

$$\mathbb{P}\{T = t\} = \frac{\binom{n-b}{t}}{\sum_{i=0}^{nD}\binom{n-b}{i}}.$$

- let Y be a RV s.t. $Y = d(S^n, C^n(\nu))$, when $S^n \in \mathcal{S}_{\mathcal{B}}^n$

$$Y = \begin{cases} \sum_{j=1}^{T} U_j + \sum_{j=1}^{n-b-T} V_j, & \text{if } 1 \leq T \leq nD \\ \sum_{j=1}^{n-b-T} V_j, & \text{if } T = 0, \end{cases}$$

  where $U_j \sim \mathrm{Ber}(1-\delta)$, $V_j \sim \mathrm{Ber}(\delta)$

- $\mathcal{Q}(\nu, \beta) = \mathbb{P}\{Y \leq nD\}$.

Proofs

Proof of Lemma VI

Key Steps (1/2)

- Chernoff bound

$$\frac{1}{n}\log_2 \mathbb{P}\left\{Y \leq nD\right\} \leq \inf_{\lambda<0}\left(\frac{1}{n}\log_2 \mathbb{M}_Y\left(\lambda\right) - \frac{\lambda D}{\log 2}\right),$$

- calculate $\mathbb{M}_Y\left(\lambda\right)$
- calculate $\frac{1}{n}\log_2 \mathbb{M}_Y\left(\lambda\right)$ and use Stirling's formula.

Proofs

Proof of Theorem IV

◂ **Theorem IV**

Key Steps

- Markov's inequality+ ◂ **Lemma IV**

$$\frac{1}{n}\log_2\left(\mathbb{P}\left\{Z\left(\mathcal{C},S^n\in\mathcal{S}_{\mathcal{B}}^n,D\right)>0\right\}\right)\geq\frac{1}{n}\log_2 E\left[Z\left(\mathcal{C},S^n\in\mathcal{S}_{\mathcal{B}}^n,D\right)\right]-$$
$$\frac{1}{n}\log_2\left(1+\sum_{j\neq 1}\mathbb{P}\left\{X_j\left(\mathcal{C},S^n\in\mathcal{S}_{\mathcal{B}}^n,D\right)=1\mid X_1\left(\mathcal{C},S^n\in\mathcal{S}_{\mathcal{B}}^n,D\right)=1\right\}\right)$$

- ◂ **Lemma VI**

$$\frac{1}{n}\log_2\left(1+\sum_{j\neq 1}\mathbb{P}\left\{X_j\left(\mathcal{C},S^n\in\mathcal{S}_{\mathcal{B}}^n,D\right)=1\mid X_1\left(\mathcal{C},S^n\in\mathcal{S}_{\mathcal{B}}^n,D\right)=1\right\}\right)$$
$$=\frac{1}{n}\log_2\left(\sum_{\substack{\nu\in[0,1]:\\ \nu m\in\mathbb{N}}}\binom{m}{\nu m}\mathcal{Q}\left(\nu,\beta\right)\right)$$
$$\leq\max_{\nu\in[0,1]}\left\{Rh(\nu)+F\left(\delta(\nu,d_c),\beta,D\right)\right\}$$

- typical value $\beta=\varepsilon$, and RHS non-negative+ ◂ **Lemma II**

Proofs

Proof of Theorem VI

Theorem VI

Key Steps (1/3)

- $\frac{R_\varepsilon^{\text{sh}}(D)+F(\delta(\nu,d_c),\varepsilon,D)}{1-\frac{B(\nu)}{R_H}}$ is strictly decreasing in ν and is equal to $R_\varepsilon^{\text{sh}}(D)$ when $\nu = 0$
- need to prove degrees independent of the blocklength are sufficient

$$\frac{1}{n}\log_2 \mathbb{P}\left\{Z\left(\mathcal{C}, S^n \in \mathcal{S}^n, D\right) > 0\right\} \geq 0 \Leftrightarrow \Delta \geq \max_{\nu\in[0,1]} \left\{K(\nu, d_c)\right\},$$

where $\Delta \triangleq R - R_\varepsilon^{\text{sh}}(D)$ and $K(\nu, d_c) \triangleq \frac{R}{R_H} B(\nu, d_v, d_c') + F\left(\delta(\nu, d_c), \varepsilon, D\right)$.

1. $\exists \mu_1 > 0$, independent of d_c, such that $\forall \nu \in [0, \mu_1]$, $K(\nu, d_c) \leq \Delta$.
2. $\exists \mu_2 > 0$, independent of d_c, such that $\forall \nu \in \left[\frac{1}{2} - \mu_2, \frac{1}{2}\right]$, $K(\nu, d_c) \leq \Delta$.
3. $\exists d_c^* < \infty$ such that $\forall \nu \in [\mu_1, 1/2 - \mu_2]$, $K(\nu, d_c^*) \leq \Delta$.

Proofs

Proof of Theorem VI

Key Steps (2/3)

1) $\exists \mu_1 > 0$, independent of d_c, such that $\forall \nu \in [0, \mu_1]$, $K(\nu, d_c) \leq \Delta$.
   - ◂ **Lemma VII** last property $\Rightarrow \exists \mu_1$ independent of the LDGM part s.t. $B(\nu) \leq 0$ for all $\nu \in [0, \mu_1]$
   - $F\left(\delta(\nu, d_c), \varepsilon, D\right) \leq 0, \forall \nu$
2) $\exists \mu_2 > 0$, independent of d_c, such that $\forall \nu \in \left[\frac{1}{2} - \mu_2, \frac{1}{2}\right]$, $K(\nu, d_c) \leq \Delta$
   - $K(\nu, d_c) \leq K(\nu, 4), \forall d_c \geq 4$
   - $\exists \mu_2$, s.t. $\forall \nu \in \left[\frac{1}{2} - \mu_2, \frac{1}{2}\right]$, $\frac{\partial^2}{\partial \nu^2} K\left(\nu, d_c\right) < 0$
   - Taylor expansion of $K(\nu, d_c)$ around $\nu = \frac{1}{2}$

     $\forall \nu \in \left[\frac{1}{2} - \mu_2, \frac{1}{2}\right], \exists \tilde{\nu} \in [\nu, 1/2]$ s.t.

     $$K(\nu, 4) = \Delta + \frac{1}{2}\left(\tilde{\nu} - \frac{1}{2}\right)^2 \left.\frac{\partial^2}{\partial \nu^2} K\left(\nu, d_c\right)\right|_{\nu = \frac{1}{2}} \leq \Delta.$$

Proofs

Proof of Theorem VI

Key Steps (3/3)

3) $\exists d_c^* < \infty$ such that $\forall \nu \in [\mu_1, 1/2 - \mu_2]$, $K(\nu, d_c^*) \leq \Delta$

- **◂ Lemma VII** $\exists \sigma(\mu_2)$, $B(\nu) \leq R_H\left[1 - \sigma(\mu_2)\right]$ for all $\nu \leq \frac{1}{2} - \mu_2$
- $\lim_{d_c \to \infty} F(\delta(\mu_1, d_c), \varepsilon, D) = -R_\varepsilon^{\text{sh}}(D)$
- $F(\gamma, \varepsilon, D)$ is a decreasing function in γ, $\forall \mu_3 > 0$, $\exists d_c^* < \infty$ s.t. $F(\delta(\mu_1, d_c^*), \varepsilon, D) \leq -R_\varepsilon^{\text{sh}}(D) + \mu_3$
- Combining the results

$$\begin{aligned} K(\nu, d_c^*) &= \frac{R}{R_H} B(\nu) + F(\delta(\nu, d_c^*), \varepsilon, D) \\ &\leq R\left[1 - \sigma(\mu_2)\right] - R_\varepsilon^{\text{sh}}(D) + \mu_3 \\ &= \Delta + (\mu_3 - R\sigma(\mu_2)) \end{aligned}$$

- choose μ_3 s.t. $\mu_3 \leq R\sigma(\mu_2)$.