Rate Distortion Bounds for Binary Erasure Source Using Sparse Graph Codes

Grégory Demay[1] Vishwambhar Rathi[1,2] Lars K. Rasmussen[1,2]

[1]School of Electrical Engineering

[2]KTH Linnaeus Center ACCESS

KTH-Royal Institute of Technology, Stockholm, Sweden

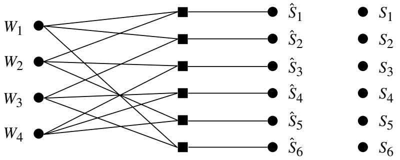
Data Compression Conference, 2010

Overview

Background and Motivation

Low-Density Generator Matrix Codes

LDGM Codes as Lossy Compressor



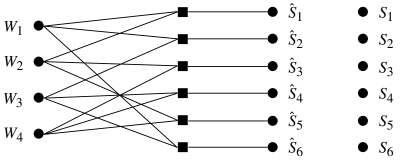
$$G = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Figure 1: LDGM code used for lossy compression $m = 6$, $R = \frac{2}{3}$, and $L(x) = x^3$.

Background and Motivation

Low-Density Generator Matrix Codes

LDGM Codes as Lossy Compressor



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

Figure 1: LDGM code used for lossy compression $m = 6$, $R = \frac{2}{3}$, and $L(x) = x^3$.

Compression/Reconstruction Process

- $s \in \mathcal{A}^m$

Background and Motivation

Low-Density Generator Matrix Codes

LDGM Codes as Lossy Compressor

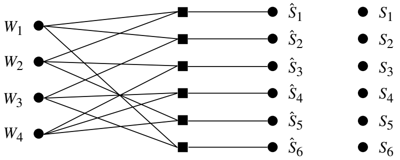


Figure 1: LDGM code used for lossy compression $m = 6$, $R = \frac{2}{3}$, and $L(x) = x^3$.

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

Compression/Reconstruction Process

- $s \in \mathcal{A}^m \xrightarrow{f} w \in \mathcal{W} = \mathbb{F}_2^{mR}$

Background and Motivation

Low-Density Generator Matrix Codes

LDGM Codes as Lossy Compressor

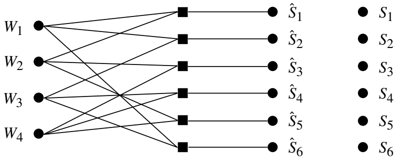


Figure 1: LDGM code used for lossy compression $m = 6$, $R = \frac{2}{3}$, and $L(x) = x^3$.

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

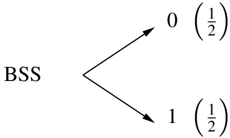
Compression/Reconstruction Process

- $s \in \mathcal{A}^m \xrightarrow{f} w \in \mathcal{W} = \mathbb{F}_2^{mR} \xrightarrow{g} \hat{s} \in \hat{\mathcal{A}}^m = \mathbb{F}_2^m$, s.t. $\hat{s} = wG$

Background and Motivation

Binary Symmetric Source

Binary Symmetric Source: BSS



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

What Has Been Done

- Performance bounds for *ensemble* of LDGM codes [1]
- Performance bounds for *individual* LDGM codes [2]

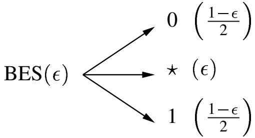
[1] A. Dimakis, M. J. Wainwright, and K. Ramchandran, "Lower bounds on the rate-distortion function of LDGM codes," in *Proc. of the IEEE Information Theory Workshop*, 2007.

[2] S. Kudekar and R. Urbanke, "Lower bounds on the rate-distortion function of individual LDGM codes," in *5th International Symposium on Turbo Codes and Related Topics*, Lausanne, Switzerland, 2008.

Background and Motivation

Binary Erasure Source

Binary Erasure Source: BES(ϵ)



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

What Has Been Done

- Duals of good LDPC codes for the BEC are good codes for the BES [3]

[3] 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 and Motivation

Distortion Measure

Distortion and Minimum Achievable Rate

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

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

Background and Motivation

Motivation

Motivation

- Past research mainly focused on BSS case
- BES(ϵ) is a generalization of a BSS
- Duals of good LDPC codes for the BEC are good codes for the BES [3]
- Better insight into the behaviour of LDGM codes

Background and Motivation

Motivation

Motivation

- Past research mainly focused on BSS case
- BES(ϵ) is a generalization of a BSS
- Duals of good LDPC codes for the BEC are good codes for the BES [3]
- Better insight into the behaviour of LDGM codes

Main Results

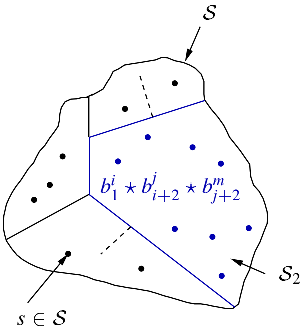
- Lower bounds on $R(D)$ for BES(ϵ) using LDGM codes via
  - counting method
  - test channel method
- Equivalence of both methods

Overview

Preliminaries

General Ideas

Partition of the Set of Source Sequences $\mathcal{S}$



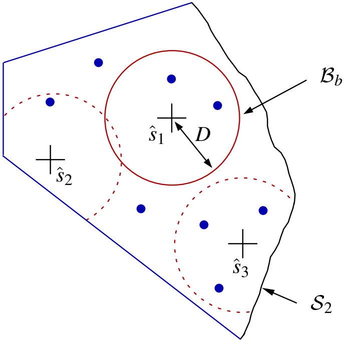
$$H_E(s) = \text{number of erasures in } s,\ \forall s \in \mathcal{S}$$

$$\mathcal{S}_b = \{s \in \mathcal{S} : H_E(s) = b\} \qquad (3)$$

Preliminaries

General Ideas

“Hamming” Balls in $\mathcal{S}_b$



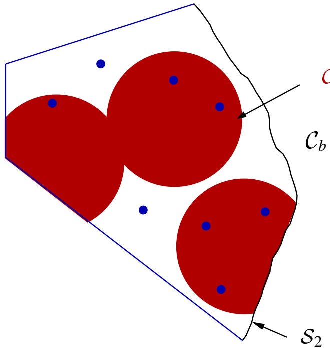
$$\forall \hat{s} \in \hat{\mathcal{S}},$$

$$\mathcal{B}_b(\hat{s}, D) = \left\{ s \in \mathcal{S}_b : \frac{1}{m} d(s, \hat{s}) \le D \right\} \tag{4}$$

Preliminaries

General Ideas

Union of “Hamming” Balls in $\mathcal{S}_b$



$$\mathcal{C}_b(D) = \bigcup_{\hat{s} \in \hat{\mathcal{S}}} \mathcal{B}_b(\hat{s}, D) = \left\{ s \in \mathcal{S}_b : \exists \hat{s} \in \hat{\mathcal{S}}, \frac{1}{m} d(s, \hat{s}) \leq D \right\} \quad (5)$$

Preliminaries

General Ideas

General Ideas - Summary

- $\mathcal{S}_b = \{s \in \mathcal{S} : \ H_E(s) = b\}$
- $\mathcal{B}_b(\hat{s}, D) = \left\{ s \in \mathcal{S}_b : \ \frac{1}{m} d(s, \hat{s}) \leq D \right\}$
- $\mathcal{C}_b(D) = \bigcup_{\hat{s} \in \hat{\mathcal{S}}} \mathcal{B}_b(\hat{s}, D)$
- Consider only $\mathcal{C}_{\epsilon m}(D)$ and **upperbound** $|\mathcal{C}_{\epsilon \mathbf{m}}(\mathbf{D})|$

Preliminaries

Simplifications

Prove the Results for

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

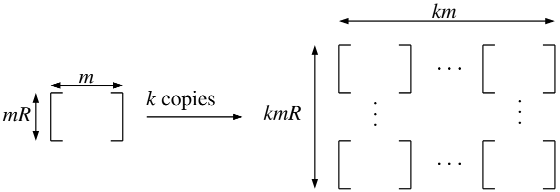


Figure 2: Construction of an arbitrarily large code with same R, D, and $L(x)$.

Preliminaries

Lower Bound on Average Distortion

Lemma I: Average Distortion and $|\mathcal{C}_{\epsilon m}(D)|$

$$\lim_{m\to\infty} \frac{1}{m} \log\left(\epsilon^{\epsilon m}\left(\frac{1-\epsilon}{2}\right)^{m-\epsilon m}|\mathcal{C}_{\epsilon m}(D)|\right) < 0, \tag{6}$$

then

$$\frac{1}{m}\mathbb{E}\left[d(S, g(f(S)))\right] \geq D(1+o(1)) \tag{7}$$

▸ proof ▸ proof Theorem I ▸ proof Theorem II

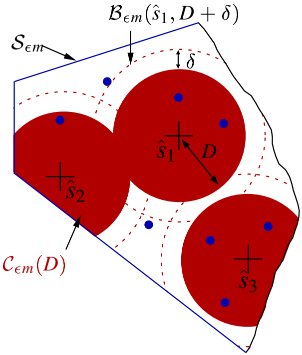
Overview

Performance of LDGM Codes as Lossy Compressors

Bounds via Counting

The Counting Method

Goal: Upper bound $|\mathcal{C}_{\epsilon m}(D)|$



- Pick $\delta \in \left[0, \frac{1-\epsilon}{2} - D\right]$ and $\delta m \in \mathbb{N}$
- $\mathcal{B}_{\epsilon m}(\hat{s}_2, D) \subset \mathcal{B}_{\epsilon m}(\hat{s}, D+\delta) \Leftrightarrow d_H(\hat{s}_2, \hat{s}_1) \leq \delta m$
- Each small $\mathcal{B}_{\epsilon m}(\hat{s}', D)$ is in $A_m(\delta m)$ big $\mathcal{B}_{\epsilon m}(\hat{s}_1, D+\delta)$

$$|\mathcal{C}_{\epsilon m}(D)| \leq \frac{1}{A_m(\delta m)} \left| \bigcup_{\hat{s} \in \hat{\mathcal{S}}} \mathcal{B}_{\epsilon m}\left(\hat{s}, D+\delta\right) \right| \tag{8}$$

- Upper bound $|\mathcal{B}_{\epsilon m}\left(\hat{s}, D+\delta\right)|$
- Lower bound $A_m(\delta m)$

Performance of LDGM Codes as Lossy Compressors

Bounds via Counting

Theorem I: Bounds via Counting

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

- *Block-length m*
- *Generator node degree distribution $L(x)$*

$$f(x) = \prod_{i=0}^{d} (1 + x^i)^{L_i}, \quad a(x) = \sum_{i=0}^{d} iL_i \frac{x^i}{1 + x^i}, \tag{9}$$

$$R(x) = (1-\epsilon)\frac{\left(1 - h\left(\frac{x}{1+x}\right)\right)}{1 - \log_2\left(\frac{f(x)}{x^{a(x)}}\right)}, \tag{10}$$

$$D(x) = \frac{x(1-\epsilon)}{1+x} - R(x)a(x). \tag{11}$$

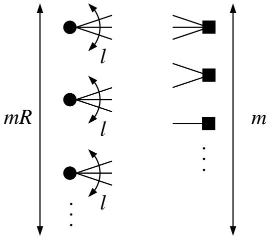
Then, the achievable rate-distortion performance of a LDGM code is lower bounded by the parametric curve $(D(x), R(x))$, $x \in [0, 1]$.

▸ Theorem II ▸ proof Theorem I

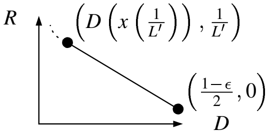
Performance of LDGM Codes as Lossy Compressors

Bound for Low Rates

Straight Line Bound



- $m \le mRl$
  $R \le \frac{1}{l} \Rightarrow m(1 - Rl)$ nodes with distortion $\frac{1-\epsilon}{2}$
- $L' = L'(1)$ average generator node degree
- $x\left(\frac{1}{L'}\right)$ unique solution of $R(x) = \frac{1}{L'}$



$$\forall R \in \left[0, \frac{1}{L'}\right], \; D = \frac{1-\epsilon}{2}\left[1 - RL'\left(1 - \frac{2}{1-\epsilon} D\left(x\left(\frac{1}{L'}\right)\right)\right)\right] \quad (12)$$

Performance of LDGM Codes as Lossy Compressors

Bounds via Counting

Regular LDGM Code

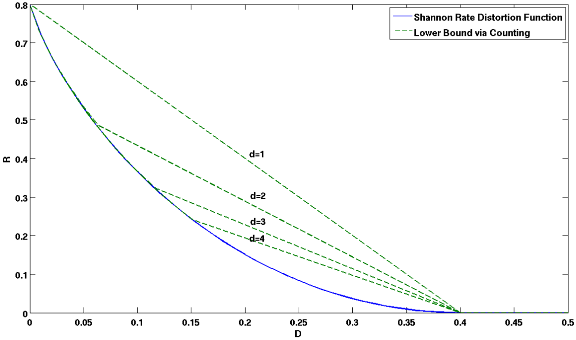


Figure 3: Rate distortion performance of generator regular LDGM code $\epsilon = 0.2$.

Performance of LDGM Codes as Lossy Compressors

Bounds via Test Channel

The Test Channel Method

Goal: Upper bound $|\mathcal{C}_{\epsilon m}(D)|$

- $\sum_{s \in \mathcal{S}_{\epsilon m}} \mathbb{P}\{S = s\} \geq \sum_{s \in \mathcal{C}_{\epsilon m}(D)} \mathbb{P}\{S = s\} \overset{?}{=} |\mathcal{C}_{\epsilon m}(D)| \, \mathbb{P}\{S = s | s \in \mathcal{C}_{\epsilon m}(D)\}$
- Lower bound $\mathbb{P}\{S = s | s \in \mathcal{C}_{\epsilon m}(D)\}$

Performance of LDGM Codes as Lossy Compressors

Bounds via Test Channel

Test Channel

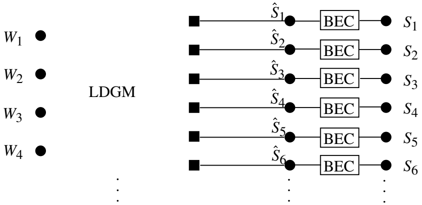


Figure 4: Test channel.

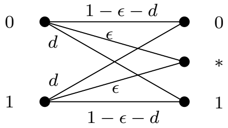


Figure 5: Binary error/erasure channel.

Performance of LDGM Codes as Lossy Compressors

Bounds via Test Channel

Theorem II: Bound via Test Channel

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

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

Then, R is lower bounded by

$$R \geq \sup_{D \leq d \leq \frac{1-\epsilon}{2}} \frac{(1-\epsilon)(1-\log_2(1-\epsilon)) + (1-D-\epsilon)\log_2(1-\epsilon-d) + D\log_2(d)}{1-\log_2\left(1+\left(\frac{d}{1-\epsilon-d}\right)^l\right)} \tag{13}$$

Theorem I
proof Theorem II

Performance of LDGM Codes as Lossy Compressors

Counting and Test Channel Methods Are Equivalent

Computation of (13) in Parametric Form

- $d \leftarrow \frac{d}{1-\epsilon-d} = x$ in (13)
- 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 m*
- *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

Conclusion

Conclusion

Key Points

- Lossy compression of a ternary source, the BES
- Derived lower bounds on $R(D)$ for the BES which are valid for
  - all LDGM codes of any length, with desired rate and generator degree distribution
  - any encoding function

Conclusion

Future Work

- Upper Bounds on the rate-distortion performance for the BES Compound construction LDGM/LDPC [4]
- Link between the test-channel model and the rate-distortion problem? [5]

Conjecture: (R, D) is achievable iff $H(S) = (1-\epsilon)m$

$$R \geq \frac{1-\epsilon+h(\epsilon)-\left(\epsilon\log_2\frac{1}{\epsilon}+D\log_2\frac{1}{D}+(1-\epsilon-D)\log_2\frac{1}{1-\epsilon-D}\right)}{1-\sum_{f=0}^{l}\binom{l}{f}\epsilon^f\sum_{w=0}^{l-f}\binom{l-f}{w}D^w(1-\epsilon-D)^{l-f-w}h\left(\frac{1}{1+\left(\frac{D}{1-\epsilon-D}\right)^{l-f-2w}}\right)}$$

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

[5] S. Kudekar and R. Urbanke, "Lower bounds on the rate-distortion function of individual LDGM codes," in *5th International Symposium on Turbo Codes and Related Topics*, Lausanne, Switzerland, 2008.

Thank you! & Questions?

Proofs

Proof of Lemma I

Lemma I

Key Steps (1/2)

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

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

Proofs

Proof of Lemma I

Lemma I

Key Steps (1/2)

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

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

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

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

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

Proofs

Proof of Lemma I

Key Steps (2/2)

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

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

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

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

Proofs

Proof of Theorem I

Theorem I

Key Steps (1/2): Upper Bound on $|\mathcal{C}_{\epsilon m}(D)|$

- $w \in \mathbb{N}$ s.t. $Dm + w \leq m\frac{(1-\epsilon)}{2}$ 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 $\omega = w/mR$.

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

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

Proofs

Proof of Theorem I

Key Steps (2/2)

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

◂ **Lemma I**

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

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

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

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

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

$$s \in \mathcal{C}_{\epsilon m}(D) \Rightarrow \exists \hat{s} \in \hat{\mathcal{S}} : d(s, \hat{s}) \leq Dm$$

Proofs

Proof of Theorem II

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

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

Proofs

Proof of Theorem II

Key Steps (3/3): Lower Bound on $|\mathcal{C}_{\epsilon m}(D)|$

$$|\mathcal{C}_{\epsilon m}(D)| \leq m2^{mR}(1+c^{l})^{-mR}(1-\epsilon-d)^{-m(1-D)}(1-\epsilon)^{m}d^{-Dm}\binom{m}{\epsilon m}\left(\frac{1-\epsilon-d}{1-\epsilon}\right)^{\epsilon m}.$$

Lemma I $\Rightarrow$

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

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

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

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

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

where

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