An "Umbrella" Bound of the Lovász-Gallager Type

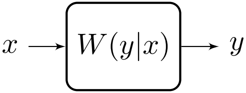
Marco Dalai

Department of Information Engineering
University of Brescia - Italy

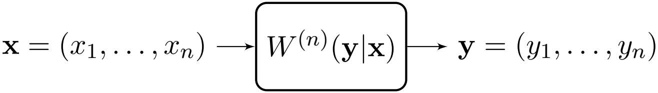
International Symposium on Information Theory, 2013

Classical DMCs

- **Discrete channel** $\mathcal{W} : \mathcal{X} \to \mathcal{Y}$, ($\mathcal{X}, \mathcal{Y}$ finite)



- **Memoryless extension** $\mathcal{W}^{(n)} : \mathcal{X}^n \to \mathcal{Y}^n$



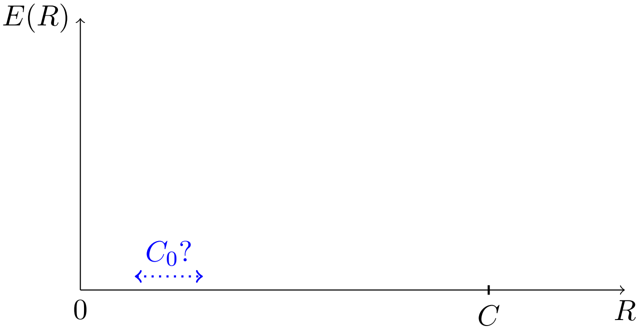
- **Code**: M codewords $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M\} \subset \mathcal{X}^n$
- **Decoder**: M disjoint decision regions $\{Y_1, \ldots, Y_M\} \subseteq \mathcal{Y}^n$
- **Probability of error**: $P_{e|m} = W^{(n)}(\overline{Y_m}|\mathbf{x}_m)$

Probability of error and reliability function

$$P_e \approx e^{-nE(R)}$$

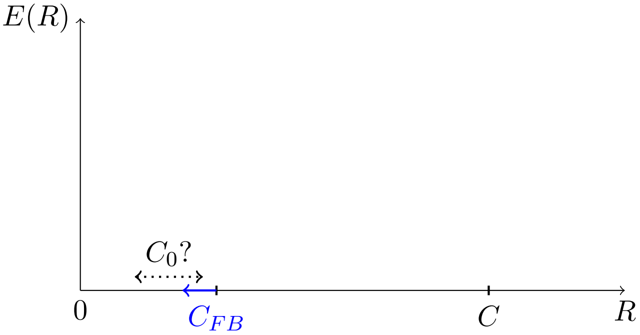
$$R < C_0 \implies E(R) \text{ infinite}$$
$$R > C_0 \implies E(R) \text{ finite}$$

Introduction



Shannon, 1956

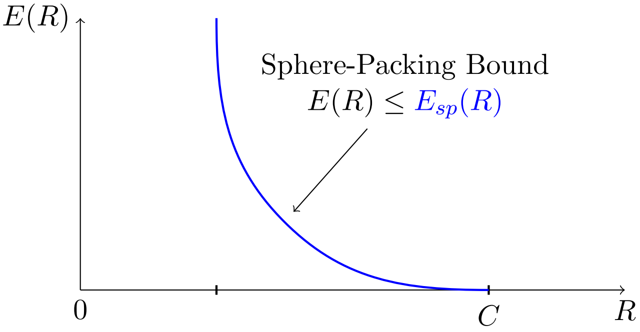
Introduction



Shannon, 1956

Upper bounded by the zero-error capacity with feedback

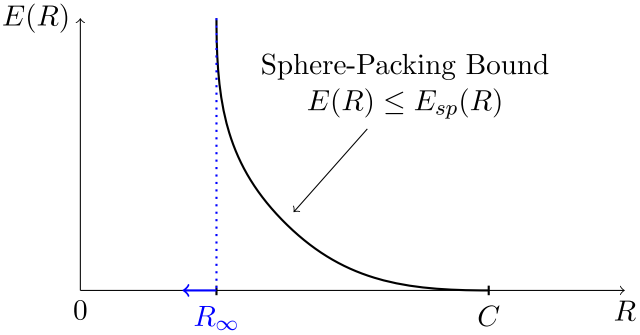
$$C_{FB} = \max_P \left[-\log \max_y \sum_{x:W(y|x)>0} P(x) \right]$$



Fano, 1961 - Shannon, Gallager and Berlekamp, 1967

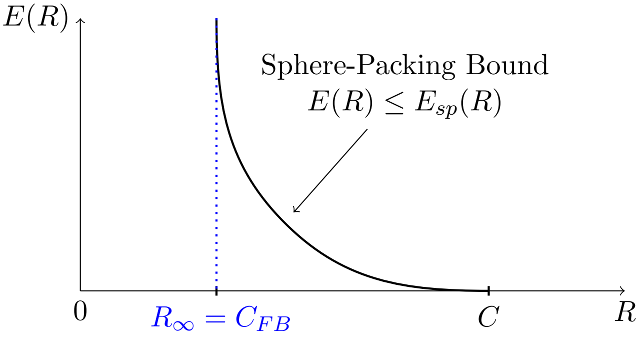
$$E_{sp}(R) = \sup_{\rho \geq 0} \max_{P} \left[-\log \sum_{y} \left(\sum_{x} P(x) W(y|x)^{1/(1+\rho)} \right)^{1+\rho} - \rho R \right]$$

Introduction



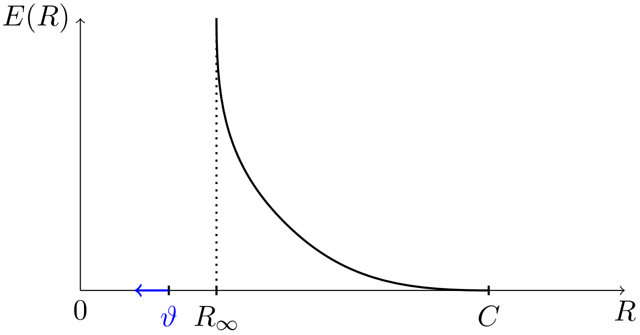
- $E_{sp}(R)$ is finite for $R > R_\infty$
- Hence, $C_0 \leq R_\infty$

Introduction



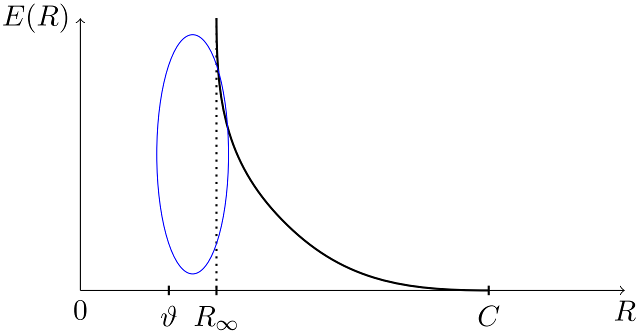
- $E_{sp}(R)$ is finite for $R > R_\infty$
- Hence, $C_0 \leq R_\infty$
- It turns out that $R_\infty = C_{FB}$ (whenever $C_0 > 0$)
- Coherent bounds to C_0 and to $E(R)$

Introduction



Lovász, 1979

- New bound: $C_0 \leq \vartheta$
- Using *geometric representations of graphs*
- Combinatorial: bound to C_0, not to P_e



Lovász, 1979

- New bound: $C_0 \leq \vartheta$
- Using *geometric representations of graphs*
- Combinatorial: bound to C_0, not to P_e
- **Objective**: bound $E(R)$ for $R > \vartheta$

- **Orthonormal Representation**:
  A set of unit norm vectors $\{\boldsymbol{u}_x\}$, $x \in \mathcal{X}$

$$x, x' \text{ not confusable} \quad \Longrightarrow \quad \boldsymbol{u}_x^\dagger \boldsymbol{u}_{x'} = 0$$

  No constraints if x, x' are confusable

- **Value**:

$$V(\{u_x\}) = \min_{\boldsymbol{c}} \max_x \log \frac{1}{|\boldsymbol{u}_x^\dagger \boldsymbol{c}|^2} \qquad (\|\boldsymbol{c}\| = 1)$$

- **The bound**:

$$C_0 \leq V(\{\boldsymbol{u}_x\})$$

- **Theta function**:

$$\vartheta = \min_{\{\boldsymbol{u}_x\}} V(\{\boldsymbol{u}_x\})$$

Lovász' Bound

- **Orthonormal Representation**:
  A set of unit norm vectors $\{\boldsymbol{u}_x\}$, $x \in \mathcal{X}$

$$x, x' \text{ not confusable} \quad \Longrightarrow \quad \boldsymbol{u}_x^\dagger \boldsymbol{u}_{x'} = 0$$

  No constraints if x, x' are confusable

- **Value**:

$$V(\{u_x\}) = \min_{\boldsymbol{c}} \max_x \log \frac{1}{|\boldsymbol{u}_x^\dagger \boldsymbol{c}|^2} \qquad (\|\boldsymbol{c}\| = 1)$$

- **The bound**:

$$C_0 \leq V(\{\boldsymbol{u}_x\})$$

- **Theta function**:

$$\vartheta = \min_{\{\boldsymbol{u}_x\}} V(\{\boldsymbol{u}_x\})$$

- **Orthonormal Representation**:
  A set of unit norm vectors $\{\boldsymbol{u}_x\}$, $x \in \mathcal{X}$

$$x, x' \text{ not confusable} \quad \Longrightarrow \quad \boldsymbol{u}_x^\dagger \boldsymbol{u}_{x'} = 0$$

  No constraints if x, x' are confusable

- **Value**:

$$V(\{u_x\}) = \min_{\boldsymbol{c}} \max_x \log \frac{1}{|\boldsymbol{u}_x^\dagger \boldsymbol{c}|^2} \qquad (\|\boldsymbol{c}\| = 1)$$

- **The bound**:

$$C_0 \leq V(\{\boldsymbol{u}_x\})$$

- **Theta function**:

$$\vartheta = \min_{\{\boldsymbol{u}_x\}} V(\{\boldsymbol{u}_x\})$$

Lovász' Bound

- **Orthonormal Representation**:
  A set of unit norm vectors $\{\boldsymbol{u}_x\}$, $x \in \mathcal{X}$

$$x, x' \text{ not confusable} \quad \Longrightarrow \quad \boldsymbol{u}_x^\dagger \boldsymbol{u}_{x'} = 0$$

  No constraints if x, x' are confusable

- **Value**:

$$V(\{u_x\}) = \min_{\boldsymbol{c}} \max_x \log \frac{1}{|\boldsymbol{u}_x^\dagger \boldsymbol{c}|^2} \qquad (\|\boldsymbol{c}\| = 1)$$

- **The bound**:

$$C_0 \leq V(\{\boldsymbol{u}_x\})$$

- **Theta function**:

$$\vartheta = \min_{\{\boldsymbol{u}_x\}} V(\{\boldsymbol{u}_x\})$$

Channels State Vectors

- Given a channel W define the *state vectors*

$$\boldsymbol{\psi}_x = \sqrt{W(\cdot|x)}$$

- Similarly, for $\boldsymbol{x} = (x_1, \ldots, x_n)$

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \sqrt{W^{(n)}(\cdot|\boldsymbol{x})}$$

- Then

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \boldsymbol{\psi}_{x_1} \otimes \boldsymbol{\psi}_{x_2} \otimes \cdots \otimes \boldsymbol{\psi}_{x_n}$$

- In a binary hypothesis test between two codewords $\boldsymbol{x}_m, \boldsymbol{x}_{m'}$

$$-\frac{1}{n} \log P_e \leq -\frac{2}{n} \log \left(\boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger} \boldsymbol{\Psi}_{\boldsymbol{x}_{m'}} \right) + o(1)$$

- So, an important quantity is

$$\max_{m \neq m'} \boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger} \boldsymbol{\Psi}_{\boldsymbol{x}'_m}$$

Channels State Vectors

- Given a channel W define the *state vectors*

$$\boldsymbol{\psi}_x = \sqrt{W(\cdot|x)}$$

- Similarly, for $\boldsymbol{x} = (x_1, \ldots, x_n)$

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \sqrt{W^{(n)}(\cdot|\boldsymbol{x})}$$

- Then

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \boldsymbol{\psi}_{x_1} \otimes \boldsymbol{\psi}_{x_2} \otimes \cdots \otimes \boldsymbol{\psi}_{x_n}$$

- In a binary hypothesis test between two codewords $\boldsymbol{x}_m, \boldsymbol{x}_{m'}$

$$-\frac{1}{n} \log P_e \leq -\frac{2}{n} \log \left(\boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger} \boldsymbol{\Psi}_{\boldsymbol{x}_{m'}} \right) + o(1)$$

- So, an important quantity is

$$\max_{m \neq m'} \boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger} \boldsymbol{\Psi}_{\boldsymbol{x}'_m}$$

Channels State Vectors

- Given a channel W define the *state vectors*

$$\boldsymbol{\psi}_x = \sqrt{W(\cdot|x)}$$

- Similarly, for $\boldsymbol{x} = (x_1, \ldots, x_n)$

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \sqrt{W^{(n)}(\cdot|\boldsymbol{x})}$$

- Then

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \boldsymbol{\psi}_{x_1} \otimes \boldsymbol{\psi}_{x_2} \otimes \cdots \otimes \boldsymbol{\psi}_{x_n}$$

- In a binary hypothesis test between two codewords $\boldsymbol{x}_m, \boldsymbol{x}_{m'}$

$$-\frac{1}{n} \log P_e \leq -\frac{2}{n} \log \left(\boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger} \boldsymbol{\Psi}_{\boldsymbol{x}_{m'}} \right) + o(1)$$

- So, an important quantity is

$$\max_{m \neq m'} \boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger} \boldsymbol{\Psi}_{\boldsymbol{x}'_m}$$

Channels State Vectors

- Given a channel W define the *state vectors*

$$\boldsymbol{\psi}_x = \sqrt{W(\cdot|x)}$$

- Similarly, for $\boldsymbol{x} = (x_1, \ldots, x_n)$

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \sqrt{W^{(n)}(\cdot|\boldsymbol{x})}$$

- Then

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \boldsymbol{\psi}_{x_1} \otimes \boldsymbol{\psi}_{x_2} \otimes \cdots \otimes \boldsymbol{\psi}_{x_n}$$

- In a binary hypothesis test between two codewords $\boldsymbol{x}_m, \boldsymbol{x}_{m'}$

$$-\frac{1}{n}\log P_e \leq -\frac{2}{n}\log\left(\boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger}\boldsymbol{\Psi}_{\boldsymbol{x}_{m'}}\right) + o(1)$$

- So, an important quantity is

$$\max_{m\neq m'} \boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger}\boldsymbol{\Psi}_{\boldsymbol{x}'_m}$$

Channels State Vectors

- Given a channel W define the *state vectors*

$$\boldsymbol{\psi}_x = \sqrt{W(\cdot|x)}$$

- Similarly, for $\boldsymbol{x} = (x_1, \ldots, x_n)$

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \sqrt{W^{(n)}(\cdot|\boldsymbol{x})}$$

- Then

$$\boldsymbol{\Psi}_{\boldsymbol{x}} = \boldsymbol{\psi}_{x_1} \otimes \boldsymbol{\psi}_{x_2} \otimes \cdots \otimes \boldsymbol{\psi}_{x_n}$$

- In a binary hypothesis test between two codewords $\boldsymbol{x}_m, \boldsymbol{x}_{m'}$

$$-\frac{1}{n} \log P_e \leq -\frac{2}{n} \log \left(\boldsymbol{\Psi}^\dagger_{\boldsymbol{x}_m} \boldsymbol{\Psi}_{\boldsymbol{x}_{m'}} \right) + o(1)$$

- So, an important quantity is

$$\max_{m \neq m'} \boldsymbol{\Psi}^\dagger_{\boldsymbol{x}_m} \boldsymbol{\Psi}_{\boldsymbol{x}'_m}$$

This Paper: Tilted State Vectors

- **Orthonormal representation of degree** $\rho \geq 1$:
  A set of "tilted" state vectors $\{\tilde{\boldsymbol{\psi}}_x\}$ such that

$$|\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'}| \leq (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$$

- **Comment**: $|\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^\dagger \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}'}| \leq (\boldsymbol{\Psi}_{\boldsymbol{x}}^\dagger \boldsymbol{\Psi}_{\boldsymbol{x}'})^{1/\rho}$
- **Value**:

$$V(\{u_x\}) = \min_{\boldsymbol{f}} \max_x \log \frac{1}{|\tilde{\boldsymbol{\psi}}_x^\dagger \boldsymbol{f}|^2} \qquad (\|\boldsymbol{f}\| = 1)$$

- **Theta function of degree** ρ

$$\vartheta(\rho) = \min_{\{\tilde{\boldsymbol{\psi}}_x\}} V(\{\tilde{\boldsymbol{\psi}}_x\})$$

  where the minimum is over representations of degree ρ

This Paper: Tilted State Vectors

- **Orthonormal representation of degree** $\rho \geq 1$:
  A set of "tilted" state vectors $\{\tilde{\boldsymbol{\psi}}_x\}$ such that

$$|\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'}| \leq (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$$

- **Comment**: $|\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^\dagger \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}'}| \leq (\boldsymbol{\Psi}_{\boldsymbol{x}}^\dagger \boldsymbol{\Psi}_{\boldsymbol{x}'})^{1/\rho}$
- **Value**:

$$V(\{u_x\}) = \min_{\boldsymbol{f}} \max_x \log \frac{1}{|\tilde{\psi}_x^\dagger \boldsymbol{f}|^2} \qquad (\|\boldsymbol{f}\| = 1)$$

- **Theta function of degree** ρ

$$\vartheta(\rho) = \min_{\{\tilde{\psi}_x\}} V(\{\tilde{\psi}_x\})$$

where the minimum is over representations of degree ρ

This Paper: Tilted State Vectors

- **Orthonormal representation of degree $\rho \geq 1$:**
  A set of "tilted" state vectors $\{\tilde{\boldsymbol{\psi}}_x\}$ such that

$$|\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'}| \leq (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$$

- **Comment**: $|\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^\dagger \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}'}| \leq (\boldsymbol{\Psi}_{\boldsymbol{x}}^\dagger \boldsymbol{\Psi}_{\boldsymbol{x}'})^{1/\rho}$
- **Value**:

$$V(\{u_x\}) = \min_{\boldsymbol{f}} \max_x \log \frac{1}{|\tilde{\boldsymbol{\psi}}_x^\dagger \boldsymbol{f}|^2} \qquad (\|\boldsymbol{f}\| = 1)$$

- **Theta function of degree ρ**

$$\vartheta(\rho) = \min_{\{\tilde{\psi}_x\}} V(\{\tilde{\psi}_x\})$$

where the minimum is over representations of degree ρ

This Paper: Tilted State Vectors

- **Orthonormal representation of degree** $\rho \geq 1$:
  A set of "tilted" state vectors $\{\tilde{\boldsymbol{\psi}}_x\}$ such that

$$|\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'}| \leq (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$$

- **Comment**: $|\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^\dagger \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}'}| \leq (\boldsymbol{\Psi}_{\boldsymbol{x}}^\dagger \boldsymbol{\Psi}_{\boldsymbol{x}'})^{1/\rho}$
- **Value**:

$$V(\{u_x\}) = \min_{\boldsymbol{f}} \max_x \log \frac{1}{|\tilde{\boldsymbol{\psi}}_x^\dagger \boldsymbol{f}|^2} \qquad (\|\boldsymbol{f}\| = 1)$$

- **Theta function of degree** ρ

$$\vartheta(\rho) = \min_{\{\tilde{\boldsymbol{\psi}}_x\}} V(\{\tilde{\boldsymbol{\psi}}_x\})$$

where the minimum is over representations of degree ρ

Bounds

Theorem

For any code of block-length n with M codewords and any $\rho \geq 1$

$$\max_m \sum_{m' \neq m} (\boldsymbol{\Psi}_m^\dagger \boldsymbol{\Psi}_{m'}) \geq \frac{\left(M e^{-n\vartheta(\rho)} - 1\right)^\rho}{(M-1)^{\rho-1}}$$

Corollary

For any $R > \vartheta(\rho)$,

$$E(R) \leq 2\rho\,\vartheta(\rho)$$

where '2' can be removed if the channel is pairwise reversible.

Proof

- Take the handle $\boldsymbol{f}$ and consider $\boldsymbol{F} = \boldsymbol{f}^{\otimes n}$. For any $\boldsymbol{x}$

$$\begin{aligned} |\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^{\dagger}\boldsymbol{F}|^2 &= \prod_{i=1}^{n} |\tilde{\boldsymbol{\psi}}_{x_i}^{\dagger}\boldsymbol{f}|^2 \\ &\geq e^{-n\vartheta(\rho)} \end{aligned}$$

- Rewrite as

$$\boldsymbol{F}^{\dagger}\left(\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^{\dagger}\right)\boldsymbol{F} \geq e^{-n\vartheta(\rho)}$$

- Averaging over $\boldsymbol{x} = \boldsymbol{x}_1, \ldots, \boldsymbol{x}_M$

$$\boldsymbol{F}^{\dagger}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\dagger}\boldsymbol{F} \geq e^{-n\vartheta(\rho)}.$$

  where $\boldsymbol{\Phi} = \left(\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}_1}, \ldots, \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}_M}\right)/\sqrt{M}$

- So, $\lambda_{\max}\left(\boldsymbol{\Phi}\boldsymbol{\Phi}^{\dagger}\right) \geq e^{-n\vartheta(\rho)}$ and hence $\lambda_{\max}\left(\boldsymbol{\Phi}^{\dagger}\boldsymbol{\Phi}\right) \geq e^{-n\vartheta(\rho)}$

Proof

- Take the handle $\boldsymbol{f}$ and consider $\boldsymbol{F} = \boldsymbol{f}^{\otimes n}$. For any $\boldsymbol{x}$

$$
\begin{aligned}
|\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^{\dagger} \boldsymbol{F}|^2 &= \prod_{i=1}^{n} |\tilde{\psi}_{x_i}^{\dagger} \boldsymbol{f}|^2 \\
&\geq e^{-n\vartheta(\rho)}
\end{aligned}
$$

- Rewrite as

$$
\boldsymbol{F}^{\dagger} \left(\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}} \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^{\dagger} \right) \boldsymbol{F} \geq e^{-n\vartheta(\rho)}
$$

- Averaging over $\boldsymbol{x} = \boldsymbol{x}_1, \ldots, \boldsymbol{x}_M$

$$
\boldsymbol{F}^{\dagger} \boldsymbol{\Phi} \boldsymbol{\Phi}^{\dagger} \boldsymbol{F} \geq e^{-n\vartheta(\rho)}.
$$

where $\boldsymbol{\Phi} = \left(\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}_1}, \ldots, \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}_M} \right) / \sqrt{M}$

- So, $\lambda_{\max} \left(\boldsymbol{\Phi} \boldsymbol{\Phi}^{\dagger} \right) \geq e^{-n\vartheta(\rho)}$ and hence $\lambda_{\max} \left(\boldsymbol{\Phi}^{\dagger} \boldsymbol{\Phi} \right) \geq e^{-n\vartheta(\rho)}$

Proof

- Take the handle $\boldsymbol{f}$ and consider $\boldsymbol{F} = \boldsymbol{f}^{\otimes n}$. For any $\boldsymbol{x}$

$$\begin{aligned} |\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^{\dagger} \boldsymbol{F}|^2 &= \prod_{i=1}^{n} |\tilde{\psi}_{x_i}^{\dagger} \boldsymbol{f}|^2 \\ &\geq e^{-n\vartheta(\rho)} \end{aligned}$$

- Rewrite as

$$\boldsymbol{F}^{\dagger} \left(\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}} \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^{\dagger} \right) \boldsymbol{F} \geq e^{-n\vartheta(\rho)}$$

- Averaging over $\boldsymbol{x} = \boldsymbol{x}_1, \ldots, \boldsymbol{x}_M$

$$\boldsymbol{F}^{\dagger} \boldsymbol{\Phi} \boldsymbol{\Phi}^{\dagger} \boldsymbol{F} \geq e^{-n\vartheta(\rho)}.$$

where $\boldsymbol{\Phi} = \left(\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}_1}, \ldots, \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}_M} \right) / \sqrt{M}$

- So, $\lambda_{\max} \left(\boldsymbol{\Phi} \boldsymbol{\Phi}^{\dagger} \right) \geq e^{-n\vartheta(\rho)}$ and hence $\lambda_{\max} \left(\boldsymbol{\Phi}^{\dagger} \boldsymbol{\Phi} \right) \geq e^{-n\vartheta(\rho)}$

Proof

- Take the handle $\boldsymbol{f}$ and consider $\boldsymbol{F} = \boldsymbol{f}^{\otimes n}$. For any $\boldsymbol{x}$

$$\begin{aligned} |\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^{\dagger}\boldsymbol{F}|^2 &= \prod_{i=1}^{n} |\tilde{\psi}_{x_i}^{\dagger}\boldsymbol{f}|^2 \\ &\geq e^{-n\vartheta(\rho)} \end{aligned}$$

- Rewrite as

$$\boldsymbol{F}^{\dagger}\left(\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}}^{\dagger}\right)\boldsymbol{F} \geq e^{-n\vartheta(\rho)}$$

- Averaging over $\boldsymbol{x} = \boldsymbol{x}_1, \ldots, \boldsymbol{x}_M$

$$\boldsymbol{F}^{\dagger}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\dagger}\boldsymbol{F} \geq e^{-n\vartheta(\rho)}.$$

  where $\boldsymbol{\Phi} = \left(\tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}_1}, \ldots, \tilde{\boldsymbol{\Psi}}_{\boldsymbol{x}_M}\right)/\sqrt{M}$

- So, $\lambda_{\max}\left(\boldsymbol{\Phi}\boldsymbol{\Phi}^{\dagger}\right) \geq e^{-n\vartheta(\rho)}$ and hence $\lambda_{\max}\left(\boldsymbol{\Phi}^{\dagger}\boldsymbol{\Phi}\right) \geq e^{-n\vartheta(\rho)}$

Proof

- For any matrix A

$$\lambda_{\max}(A) \leq \max_i \sum_j |A_{i,j}|.$$

- If $A = \mathbf{\Phi}^\dagger \mathbf{\Phi}$ then $A_{i,j} = \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_i}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_j}/M$
- So

$$e^{-n\vartheta(\rho)} \leq \lambda_{\max}\left(\mathbf{\Phi}^\dagger \mathbf{\Phi}\right) \leq \max_i \sum_j \frac{|\tilde{\mathbf{\Psi}}_{\boldsymbol{x}_i}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_j}|}{M}$$

- Using some algebra and $|\tilde{\mathbf{\Psi}}_{\boldsymbol{x}}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}'}| \leq (\mathbf{\Psi}_{\boldsymbol{x}}^\dagger \mathbf{\Psi}_{\boldsymbol{x}'})^{1/\rho}$ we get the theorem

$$\max_m \sum_{m' \neq m} (\mathbf{\Psi}_{\boldsymbol{x}_m}^\dagger \mathbf{\Psi}_{\boldsymbol{x}_{m'}}) \geq \frac{\left(Me^{-n\vartheta(\rho)} - 1\right)^\rho}{(M-1)^{\rho-1}}$$

Proof

- For any matrix A

$$\lambda_{\max}(A) \leq \max_i \sum_j |A_{i,j}|.$$

- If $A = \mathbf{\Phi}^\dagger \mathbf{\Phi}$ then $A_{i,j} = \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_i}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_j} / M$
- So

$$e^{-n\vartheta(\rho)} \leq \lambda_{\max}\left(\mathbf{\Phi}^\dagger \mathbf{\Phi}\right) \leq \max_i \sum_j \frac{|\tilde{\mathbf{\Psi}}_{\boldsymbol{x}_i}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_j}|}{M}$$

- Using some algebra and $|\tilde{\mathbf{\Psi}}_{\boldsymbol{x}}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}'}| \leq (\mathbf{\Psi}_{\boldsymbol{x}}^\dagger \mathbf{\Psi}_{\boldsymbol{x}'})^{1/\rho}$ we get the theorem

$$\max_m \sum_{m' \neq m} (\mathbf{\Psi}_{\boldsymbol{x}_m}^\dagger \mathbf{\Psi}_{\boldsymbol{x}_{m'}}) \geq \frac{\left(M e^{-n\vartheta(\rho)} - 1\right)^\rho}{(M-1)^{\rho-1}}$$

Proof

- For any matrix A

$$\lambda_{\max}(A) \leq \max_i \sum_j |A_{i,j}|.$$

- If $A = \mathbf{\Phi}^\dagger \mathbf{\Phi}$ then $A_{i,j} = \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_i}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_j} / M$
- So

$$e^{-n\vartheta(\rho)} \leq \lambda_{\max}\left(\mathbf{\Phi}^\dagger \mathbf{\Phi}\right) \leq \max_i \sum_j \frac{|\tilde{\mathbf{\Psi}}_{\boldsymbol{x}_i}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_j}|}{M}$$

- Using some algebra and $|\tilde{\mathbf{\Psi}}_{\boldsymbol{x}}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}'}| \leq (\mathbf{\Psi}_{\boldsymbol{x}}^\dagger \mathbf{\Psi}_{\boldsymbol{x}'})^{1/\rho}$ we get the theorem

$$\max_m \sum_{m' \neq m} (\mathbf{\Psi}_{\boldsymbol{x}_m}^\dagger \mathbf{\Psi}_{\boldsymbol{x}_{m'}}) \geq \frac{\left(Me^{-n\vartheta(\rho)} - 1\right)^\rho}{(M-1)^{\rho-1}}$$

Proof

- For any matrix A

$$\lambda_{\max}(A) \leq \max_i \sum_j |A_{i,j}|.$$

- If $A = \mathbf{\Phi}^\dagger \mathbf{\Phi}$ then $A_{i,j} = \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_i}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_j} / M$
- So

$$e^{-n\vartheta(\rho)} \leq \lambda_{\max}\left(\mathbf{\Phi}^\dagger \mathbf{\Phi}\right) \leq \max_i \sum_j \frac{|\tilde{\mathbf{\Psi}}_{\boldsymbol{x}_i}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}_j}|}{M}$$

- Using some algebra and $|\tilde{\mathbf{\Psi}}_{\boldsymbol{x}}^\dagger \tilde{\mathbf{\Psi}}_{\boldsymbol{x}'}| \leq (\mathbf{\Psi}_{\boldsymbol{x}}^\dagger \mathbf{\Psi}_{\boldsymbol{x}'})^{1/\rho}$ we get the theorem

$$\max_m \sum_{m' \neq m} (\mathbf{\Psi}_{\boldsymbol{x}_m}^\dagger \mathbf{\Psi}_{\boldsymbol{x}_{m'}}) \geq \frac{\left(M e^{-n\vartheta(\rho)} - 1\right)^\rho}{(M-1)^{\rho-1}}$$

Proof

- **Corollary**: if $R > \vartheta(\rho)$

$$
\begin{aligned}
\max_{m \neq m'} \boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger} \boldsymbol{\Psi}_{\boldsymbol{x}_{m'}} &\geq \frac{1}{M-1} \max_m \sum_{m' \neq m} \left(\boldsymbol{\Psi}_{\boldsymbol{x}_m}^{\dagger} \boldsymbol{\Psi}_{\boldsymbol{x}_{m'}} \right) \\
&\geq \left(\frac{M e^{-n\vartheta(\rho)} - 1}{M-1} \right)^{\rho} \\
&\geq \left(e^{-n\vartheta(\rho)} - 1/M \right)^{\rho} \\
&\approx e^{-n\rho\,\vartheta(\rho)}
\end{aligned}
$$

Computation

- Set $K = |\mathcal{X}|$ and consider the $(K+1)\times(K+1)$ Gram matrix

$$G = [\tilde{\boldsymbol{\psi}}_1, \ldots, \tilde{\boldsymbol{\psi}}_K, \boldsymbol{f}]^\dagger [\tilde{\boldsymbol{\psi}}_1, \ldots, \tilde{\boldsymbol{\psi}}_K, \boldsymbol{f}]$$

- We can find the optimal representation solving the problem

$$\begin{array}{rrcl} & \max & t & \\ \text{s.t.} & G(k, K+1) & \geq & t, \quad \forall k \leq K \\ & G(k,k) & = & 1, \quad \forall k \\ & G(k,i) & \leq & \boldsymbol{\psi}_k^\dagger \boldsymbol{\psi}_i^{1/\rho} \\ & & & 1 \leq k < K,\, k < i \leq K \\ & G & \text{is} & \text{positive semidefinite} \end{array}$$

- The solution is $e^{-\vartheta(\rho)}$

Properties of $\vartheta(\rho)$

When $\rho \to \infty$

- The condition
$$|\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'}| \leq (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$$
reduces to Lovász' contraint
$$x, x' \text{ not confusable} \quad \Longrightarrow \quad \tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'} = 0$$
- Hence
$$\lim_{\rho\to\infty} \vartheta(\rho) = \vartheta$$
- We deduce
$$R > \vartheta \implies E(R) \text{ finite}$$
- **Comments**:
  - $\vartheta(\rho)$ can be constant for $\rho > \rho_{\max}$
  - The bound is finite but not really good

Properties of $\vartheta(\rho)$

When $\rho = 1$

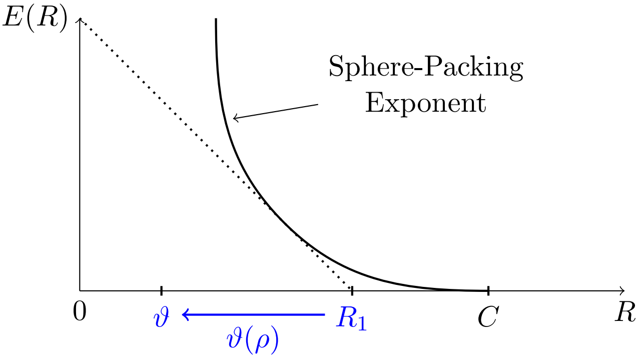
- The condition becomes $|\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'}| \leq (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})$, optimally satisfied by the trivial representation $\tilde{\boldsymbol{\psi}}_x = \boldsymbol{\psi}_x$.
- So,

$$\begin{aligned} \vartheta(1) &= \min_{\boldsymbol{f}} \max_x \log \frac{1}{|\boldsymbol{\psi}_x^\dagger \boldsymbol{f}|^2} \\ &= \min_Q \max_x \left(-2 \log \sum_y \sqrt{W(y|x)Q(y)} \right) \end{aligned}$$

- As proved by Csiszár (1995), this is the cutoff rate of the channel

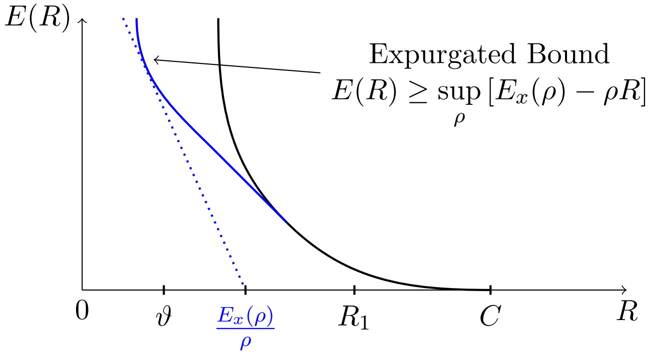
$$R_1 = \max_P \left[-\log \sum_{x,x'} P(x)P(x') \left(\sum_y \sqrt{W(y|x)W(y|x')} \right) \right]$$

Properties of $\vartheta(\rho)$: Range



When ρ varies from 1 to ∞, $\vartheta(\rho)$ moves from R_1 to ϑ

Properties of $\vartheta(\rho)$: Connections with the Expurgated Bound



In general, $\vartheta(\rho)$ is related to Gallager's expurgated coefficient

$$E_x(\rho) = \max_P \left[-\rho \log \sum_{x,x'} P(x)P(x')(\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho} \right].$$

Properties of $\vartheta(\rho)$

Connections with the Expurgated Bound

- Consider Jelinek's *non-negative definite channels*, for which the matrix with elements $(\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$ is positive semidefinite for all $\rho \geq 1$.

- For these channels the expurgated coefficient $E_x^{(n)}(\rho)$ computed on the n-fold extension equals $E_x(\rho)$
- And
$$C_0 = \lim_{\rho\to\infty} \frac{E_x(\rho)}{\rho} = \vartheta$$
- Note that, for all $\rho \geq 1$, there is a representation that satisfies the constraints with equality
$$\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'} = (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$$

Properties of $\vartheta(\rho)$

Connections with the Expurgated Bound

- Consider Jelinek's *non-negative definite channels*, for which the matrix with elements $(\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$ is positive semidefinite for all $\rho \geq 1$.
- For these channels the expurgated coefficient $E_x^{(n)}(\rho)$ computed on the n-fold extension equals $E_x(\rho)$
- And

$$C_0 = \lim_{\rho \to \infty} \frac{E_x(\rho)}{\rho} = \vartheta$$

- Note that, for all $\rho \geq 1$, there is a representation that satisfies the constraints with equality

$$\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'} = (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$$

Connections with the Expurgated Bound

- Consider Jelinek's *non-negative definite channels*, for which the matrix with elements $(\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$ is positive semidefinite for all $\rho \geq 1$.

- For these channels the expurgated coefficient $E_x^{(n)}(\rho)$ computed on the n-fold extension equals $E_x(\rho)$
- And

$$C_0 = \lim_{\rho \to \infty} \frac{E_x(\rho)}{\rho} = \vartheta$$

- Note that, for all $\rho \geq 1$, there is a representation that satisfies the constraints with equality

$$\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'} = (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$$

Connections with the Expurgated Bound

- Consider Jelinek's *non-negative definite channels*, for which the matrix with elements $(\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$ is positive semidefinite for all $\rho \geq 1$.

- For these channels the expurgated coefficient $E_x^{(n)}(\rho)$ computed on the n-fold extension equals $E_x(\rho)$
- And

$$C_0 = \lim_{\rho \to \infty} \frac{E_x(\rho)}{\rho} = \vartheta$$

- Note that, for all $\rho \geq 1$, there is a representation that satisfies the constraints with equality

$$\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'} = (\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$$

Connections with the Expurgated Bound

- Then it can be proved that

$$\begin{aligned}\vartheta(\rho) &= \min_{\boldsymbol{f}} \max_{x} \log \frac{1}{|\tilde{\boldsymbol{\psi}}_x^\dagger \boldsymbol{f}|^2} \\ &= \max_P \left[-\log \sum_{x,x'} P(x)P(x')\tilde{\boldsymbol{\psi}}_x^\dagger \tilde{\boldsymbol{\psi}}_{x'} \right] \\ &= \max_P \left[-\log \sum_{x,x'} P(x)P(x')(\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho} \right] \\ &= \frac{E_x(\rho)}{\rho}\end{aligned}$$

Connections with the Expurgated Bound

- For general channels, the matrix with elements $(\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$ is not positive semidefinite for $\rho > \bar{\rho}$
- It can be proved (not in this paper) that

$$\frac{E_x(\rho)}{\rho} \leq \frac{E_x^{(n)}(\rho)}{\rho} \leq \vartheta(\rho), \quad \forall n$$

- This generalizes the following property of ϑ

$$C_0 = \sup_n \left(\lim_{\rho \to \infty} \frac{E_x^{(n)}(\rho)}{\rho} \right) \leq \vartheta$$

Connections with the Expurgated Bound

- For general channels, the matrix with elements $(\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$ is not positive semidefinite for $\rho > \bar{\rho}$
- It can be proved (not in this paper) that

$$\frac{E_x(\rho)}{\rho} \leq \frac{E_x^{(n)}(\rho)}{\rho} \leq \vartheta(\rho), \quad \forall n$$

- This generalizes the following property of ϑ

$$C_0 = \sup_n \left(\lim_{\rho\to\infty} \frac{E_x^{(n)}(\rho)}{\rho} \right) \leq \vartheta$$

Properties of $\vartheta(\rho)$

Connections with the Expurgated Bound

- For general channels, the matrix with elements $(\boldsymbol{\psi}_x^\dagger \boldsymbol{\psi}_{x'})^{1/\rho}$ is not positive semidefinite for $\rho > \bar{\rho}$
- It can be proved (not in this paper) that

$$\frac{E_x(\rho)}{\rho} \leq \frac{E_x^{(n)}(\rho)}{\rho} \leq \vartheta(\rho), \quad \forall n$$

- This generalizes the following property of ϑ

$$C_0 = \sup_n \left(\lim_{\rho\to\infty} \frac{E_x^{(n)}(\rho)}{\rho} \right) \leq \vartheta$$

Properties of $\vartheta(\rho)$: Example

Typewriter channel with 5 inputs and crossover probability 1/2