Error Exponents of the Degraded Broadcast Channel with Degraded Message Sets

Graduate Seminar

Yonatan Kaspi and Neri Merhav

Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa, Israel

March 19th, 2009

Outline

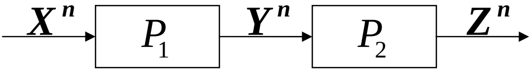
1. Introduction
   - Review of the Degraded Broadcast Channel
   - Coding Scheme
   - Previous Work

Outline

Outline

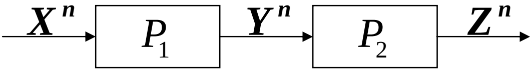
Outline

The Degraded Broadcast Channel



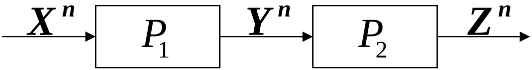
- Channel: $P(\boldsymbol{y}, \boldsymbol{z}|\boldsymbol{x}) = \prod_{t=1}^{n} P_1(y_t|x_t)P(z_t|y_t)$

The Degraded Broadcast Channel



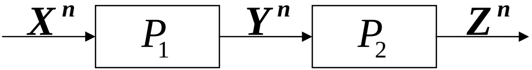
- Channel: $P(\boldsymbol{y}, \boldsymbol{z}|\boldsymbol{x}) = \prod_{t=1}^{n} P_1(y_t|x_t)P(z_t|y_t)$
- R_{yz} - The rate for both users.

The Degraded Broadcast Channel



- Channel: $P(\boldsymbol{y}, \boldsymbol{z}|\boldsymbol{x}) = \prod_{t=1}^{n} P_1(y_t|x_t)P(z_t|y_t)$
- R_{yz} - The rate for both users.
- R_y - The rate for the strong user.

The Degraded Broadcast Channel

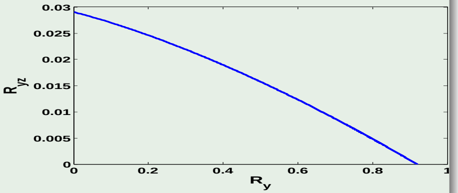


- Channel: $P(\boldsymbol{y}, \boldsymbol{z}|\boldsymbol{x}) = \prod_{t=1}^{n} P_1(y_t|x_t)P(z_t|y_t)$
- R_{yz} - The rate for both users.
- R_y - The rate for the strong user.

Capacity Region:

Convex hull of the closure of all (R_y, R_{yz}) satisfying

$$R_{yz} \leq I(Z;U)$$
$$R_y \leq I(X;Y|U)$$

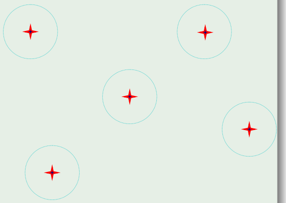


For some $P(u)P(x|u)P(y,z|x)$

Coding for the Broadcast Channel

Bergmans (73) capacity achieving scheme:

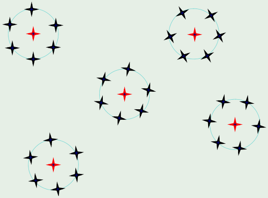
- First, draw $e^{nR_{yz}}$ "cloud centers" $\sim P(u)$.



Coding for the Broadcast Channel

Bergmans (73) capacity achieving scheme:

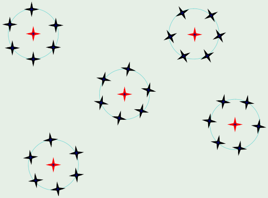
- First, draw $e^{nR_{yz}}$ "cloud centers" $\sim P(u)$.
- Draw a "cloud" of e^{nR_y} codewords $\sim P(x|u)$.



Coding for the Broadcast Channel

Bergmans (73) capacity achieving scheme:

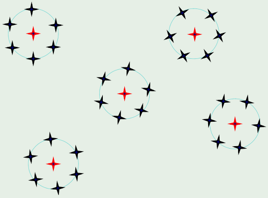
- First, draw $e^{nR_{yz}}$ "cloud centers" $\sim P(u)$.
- Draw a "cloud" of e^{nR_y} codewords $\sim P(x|u)$.
- To send message m to both users and message i to the strong user, send the i-th codeword from the m-th cloud.



Coding for the Broadcast Channel

Bergmans (73) capacity achieving scheme:

- First, draw $e^{nR_{yz}}$ "cloud centers" $\sim P(u)$.
- Draw a "cloud" of e^{nR_y} codewords $\sim P(x|u)$.
- To send message m to both users and message i to the strong user, send the i-th codeword from the m-th cloud.

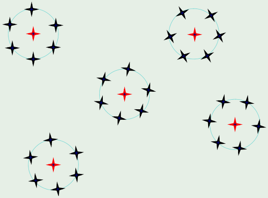


- Weak decoder determines only the cloud.

Coding for the Broadcast Channel

Bergmans (73) capacity achieving scheme:

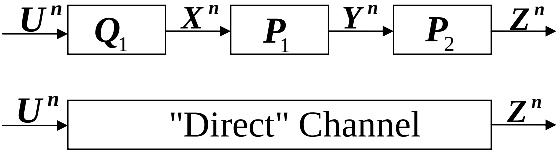
- First, draw $e^{nR_{yz}}$ "cloud centers" $\sim P(u)$.
- Draw a "cloud" of e^{nR_y} codewords $\sim P(x|u)$.
- To send message m to both users and message i to the strong user, send the i-th codeword from the m-th cloud.



- Weak decoder determines only the cloud.
- Strong decoder also determines the specific message within the cloud.

Previous Work

- Gallager, 74'. By averaging over the cloud structure, a direct channel from cloud center to the weak decoder is computed. The error exponents are given for the computed direct channel. Therefore, the error exponent for the weak decoder depends only on the common message rate.



Previous Work. Cont.

- Körner and Sgarro, 80'. Used a maximum mutual information decoder. They derived universally achievable exponents that depend on both rates.

Previous Work. Cont.

- Körner and Sgarro, 80'. Used a maximum mutual information decoder. They derived universally achievable exponents that depend on both rates.
- Sokolovski and Bross. 05'. Gave attainable error exponents for the Poisson BC with degraded message sets.

Previous Work. Cont.

- Körner and Sgarro, 80'. Used a maximum mutual information decoder. They derived universally achievable exponents that depend on both rates.
- Sokolovski and Bross. 05'. Gave attainable error exponents for the Poisson BC with degraded message sets.
- Weng et. al. 08'. Gave bounds for Gaussian broadcast and MAC channels. Defined the "Error Exponents Region".

Previous Work. Cont.

- Körner and Sgarro, 80'. Used a maximum mutual information decoder. They derived universally achievable exponents that depend on both rates.
- Sokolovski and Bross. 05'. Gave attainable error exponents for the Poisson BC with degraded message sets.
- Weng et. al. 08'. Gave bounds for Gaussian broadcast and MAC channels. Defined the "Error Exponents Region".

In this work

Our exponents pertain to optimal (ML) decoding. Namely:

- Strong decoder: $(\hat{m}(\boldsymbol{y}), \hat{i}(\boldsymbol{y})) = \arg\max_{m,i} P(\boldsymbol{y}|\boldsymbol{x}_{m,i})$.
- Weak decoder: $\tilde{m}(\boldsymbol{z}) = \arg\max_{m} \frac{1}{M_y} \sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})$.

Previous Work. Cont.

- Körner and Sgarro, 80'. Used a maximum mutual information decoder. They derived universally achievable exponents that depend on both rates.
- Sokolovski and Bross. 05'. Gave attainable error exponents for the Poisson BC with degraded message sets.
- Weng et. al. 08'. Gave bounds for Gaussian broadcast and MAC channels. Defined the "Error Exponents Region".

In this work

Our exponents pertain to <u>optimal</u> (ML) decoding. Namely:

- Strong decoder: $(\hat{m}(\boldsymbol{y}), \hat{i}(\boldsymbol{y})) = \arg\max_{m,i} P(\boldsymbol{y}|\boldsymbol{x}_{m,i})$.
- Weak decoder: $\tilde{m}(\boldsymbol{z}) = \arg\max_m \frac{1}{M_y} \sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})$.

The exponents depend on both rates.

Part I:
Gallager type bounding Method

Deriving the Exponents

The weak decoder

For the weak decoder, we start with Gallager's upper bound to the "channel" $P(\boldsymbol{z}|m) = \frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})$

$$\overline{P_{E_m}^{z}} \leq \sum_{\boldsymbol{z}} \boldsymbol{E}\left[\frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right]^{1-\rho\lambda} \cdot \boldsymbol{E}\left[\sum_{m'\neq m}\left(\frac{1}{M_y}\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho}.$$

Deriving the Exponents

The weak decoder

For the weak decoder, we start with Gallager's upper bound to the "channel" $P(\boldsymbol{z}|m) = \frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})$

$$\overline{P^z_{E_m}} \leq \sum_{\boldsymbol{z}} \boldsymbol{E}\left[\frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right]^{1-\rho\lambda} \cdot \boldsymbol{E}\left[\sum_{m'\neq m}\left(\frac{1}{M_y}\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho}.$$

We now continue with each expectation separately using Forney's method:

$$\sum_{\boldsymbol{z}} \boldsymbol{E}\left[\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right]^{1-\rho\lambda} = \sum_{\boldsymbol{z}} \boldsymbol{E}\left[\left(\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right)^{\alpha}\right]^{\frac{1-\rho\lambda}{\alpha}}$$

Deriving the Exponents

The weak decoder

For the weak decoder, we start with Gallager's upper bound to the "channel" $P(\boldsymbol{z}|m) = \frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})$

$$\overline{P_{E_m}^{z}} \leq \sum_{\boldsymbol{z}} \boldsymbol{E}\left[\frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right]^{1-\rho\lambda} \cdot \boldsymbol{E}\left[\sum_{m'\neq m}\left(\frac{1}{M_y}\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho}.$$

We now continue with each expectation separately using Forney's method:

$$\begin{aligned}\sum_{\boldsymbol{z}} \boldsymbol{E}\left[\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right]^{1-\rho\lambda} &= \sum_{\boldsymbol{z}} \boldsymbol{E}\left[\left(\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right)^{\alpha}\right]^{\frac{1-\rho\lambda}{\alpha}} \\ &\leq \sum_{\boldsymbol{z}} \boldsymbol{E}\left[\sum_{i=1}^{M_y} P^{\alpha}(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right]^{\frac{1-\rho\lambda}{\alpha}}\end{aligned}$$

We finally get:

The Weak Decoder

The weak decoder error exponent

$$E_z(R_y, R_{yz}) = \max_{0\le\rho\le 1} \max_{0\le\lambda\le 1} \max_{\lambda\le\mu\le 1} \max_{1-\rho\lambda\le\alpha\le 1} \{E_0(\rho,\lambda,\alpha,\mu) - (\alpha+\rho\mu-1)R_y - \rho R_{yz}\}$$

where,

$$E_0(\rho,\lambda,\alpha,\mu) = -\log\sum_z \left\{\sum_u Q(u)\left[\sum_x Q(x|u)P(z|x)^{1-\rho\lambda/\alpha}\right]^\alpha \times \left[\sum_{u'} Q(u')\left[\sum_x Q(x|u')P(z|x)^{\lambda/\mu}\right]^\mu\right]^\rho\right\}$$

We will look into three special cases:

The Weak Decoder

The weak decoder error exponent

$$E_z(R_y, R_{yz}) = \max_{0\leq\rho\leq 1} \max_{0\leq\lambda\leq 1} \max_{\lambda\leq\mu\leq 1} \max_{1-\rho\lambda\leq\alpha\leq 1} \{E_0(\rho,\lambda,\alpha,\mu) - (\alpha+\rho\mu-1)R_y - \rho R_{yz}\}$$

where,

$$E_0(\rho,\lambda,\alpha,\mu) = -\log \sum_z \left\{ \sum_u Q(u) \left[\sum_x Q(x|u) P(z|x)^{1-\rho\lambda/\alpha} \right]^{\alpha} \times \left[\sum_{u'} Q(u') \left[\sum_x Q(x|u') P(z|x)^{\lambda/\mu} \right]^{\mu} \right]^{\rho} \right\}$$

We will look into three special cases:

1. $\alpha = \mu$

The Weak Decoder

The weak decoder error exponent

$$E_z(R_y, R_{yz}) = \max_{0\le\rho\le1}\max_{0\le\lambda\le1}\max_{\lambda\le\mu\le1}\max_{1-\rho\lambda\le\alpha\le1}$$
$$\{E_0(\rho,\lambda,\alpha,\mu) - (\alpha+\rho\mu-1)R_y - \rho R_{yz}\}$$

where,

$$E_0(\rho,\lambda,\alpha,\mu) = -\log\sum_z\left\{\sum_u Q(u)\left[\sum_x Q(x|u)P(z|x)^{1-\rho\lambda/\alpha}\right]^\alpha \times\right.$$
$$\left.\left[\sum_{u'} Q(u')\left[\sum_x Q(x|u')P(z|x)^{\lambda/\mu}\right]^\mu\right]^\rho\right\}$$

We will look into three special cases:

1. $\alpha = \mu$
2. $\alpha = \mu = 1$
3. $\alpha = \mu = \frac{1}{1+\rho}$

The Weak Decoder, Cont.

$\alpha = \mu$

The optimal value of λ is $\frac{1}{1+\rho}$. We get an exponent with only two parameters.

$$E_z(R_y, R_{yz}) = \max_{0\leq\rho\leq1} \max_{\frac{1}{1+\rho}\leq\alpha\leq1} \left\{ E_0(\rho, \frac{1}{1+\rho}, \alpha, \alpha) - (\alpha(1+\rho)-1)R_y - \rho R_{yz} \right\}$$

The Weak Decoder, Cont.

$\alpha = \mu$

The optimal value of λ is $\frac{1}{1+\rho}$. We get an exponent with only two parameters.

$$E_z(R_y, R_{yz}) = \max_{0\le\rho\le 1} \max_{\frac{1}{1+\rho}\le\alpha\le 1} \left\{ E_0(\rho, \frac{1}{1+\rho}, \alpha, \alpha) - (\alpha(1+\rho)-1)R_y - \rho R_{yz} \right\}$$

$\alpha = \mu = 1$

In this case, there is no dependence on the coding parameter $P(x|u)$.

$$E_z(R_y, R_{yz}) = -\log\left\{ \sum_z \left[\sum_x Q(x) P(z|x)^{1/(1+\rho)} \right]^{1+\rho} \right\} - \rho(R_{yz} + R_y)$$

The Weak Decoder, Cont.

$\alpha = \mu$

The optimal value of λ is $\frac{1}{1+\rho}$. We get an exponent with only two parameters.

$$E_z(R_y, R_{yz}) = \max_{0\le\rho\le1} \max_{\frac{1}{1+\rho}\le\alpha\le1} \left\{ E_0(\rho, \frac{1}{1+\rho}, \alpha, \alpha) - (\alpha(1+\rho)-1)R_y - \rho R_{yz} \right\}$$

$\alpha = \mu = 1$

In this case, there is no dependence on the coding parameter $P(x|u)$.

$$E_z(R_y, R_{yz}) = -\log \left\{ \sum_z \left[\sum_x Q(x) P(z|x)^{1/(1+\rho)} \right]^{1+\rho} \right\} - \rho(R_{yz} + R_y)$$

$\alpha = \mu = \frac{1}{1+\rho}$

In this case, we return to the same exponent in Gallager's 74' work.

$$E_z(R_y, R_{yz}) = -\log \left\{ \sum_z \left[\sum_u Q(u) P(z|u)^{1/(1+\rho)} \right]^{1+\rho} \right\} - \rho R_{yz}$$

The Strong Decoder

The strong decoder can be wrong in two ways:

The Strong Decoder

The strong decoder can be wrong in two ways:

1. Can choose the wrong private message from the correct cloud.

The Strong Decoder

The strong decoder can be wrong in two ways:

1. Can choose the wrong private message from the correct cloud.
2. Can choose a wrong cloud.

The Strong Decoder

The strong decoder can be wrong in two ways:

1. Can choose the wrong private message from the correct cloud.
2. Can choose a wrong cloud.

The strong decoder

The Exponent is the worst exponent of the above events.

$$E_y(R_y, R_{yz}) = \min\left(\max_{0<\rho<1} E_{y1}(R_y,\rho), \max_{0<\rho<1} E_{y2}(R_y,R_{yz},\rho)\right)$$

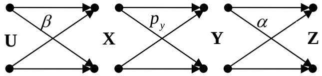
Where,

$$E_{y1}(R_y,\rho) = -\rho R_y - \log \sum_y \sum_u Q(u) \left[\sum_x Q(x|u) P(y|x)^{\frac{1}{1+\rho}}\right]^{1+\rho}$$

$$E_{y2}(R_y,R_{yz},\rho) = -\rho(R_y+R_{yz}) - \log\left\{\sum_y \left[\sum_x Q(x)P(y|x)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right\}$$

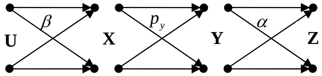
Numerical Results

Memoryless binary symmetric channel



Numerical Results

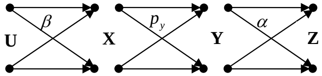
Memoryless binary symmetric channel



- Only one coding parameter (β). $Q(u) = \frac{1}{2}$ is optimal

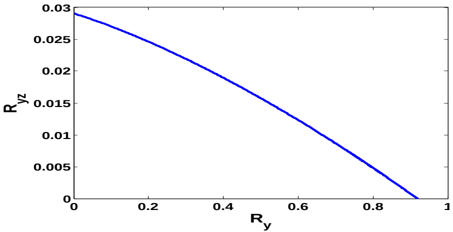
Numerical Results

Memoryless binary symmetric channel



- Only one coding parameter (β). $Q(u) = \frac{1}{2}$ is optimal
- The capacity region is:

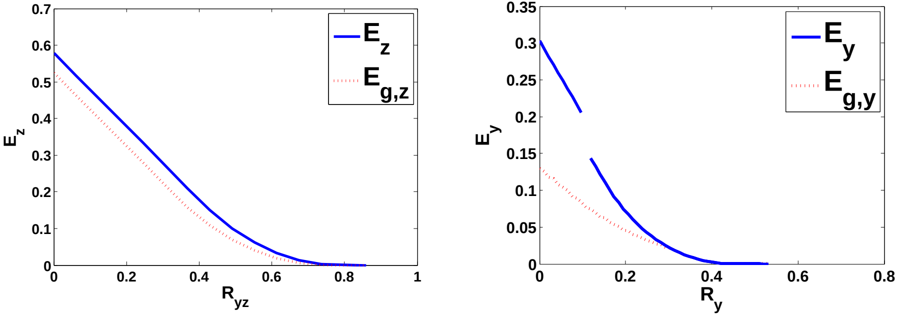
$$
\begin{aligned}
R_{yz} &\leq 1 - h(\beta * p_z) \\
R_y &\leq h(\beta * p_y) - h(p_y)
\end{aligned}
$$



- We fix one rate and plot the exponent as a function of the other rate.

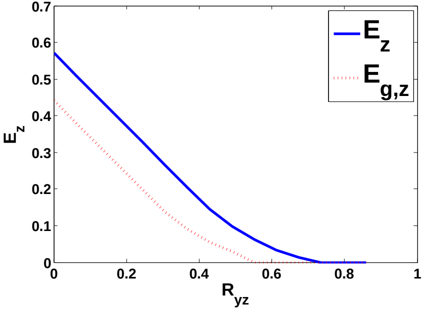
Results for the Broadcast BSC

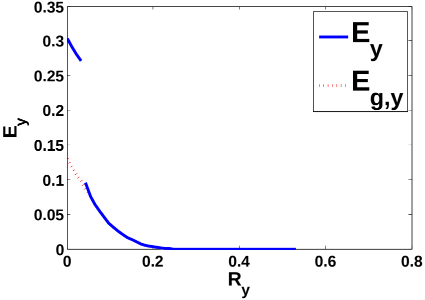
Numerical results of the exponents under the constraint that both are greater then zero. We show the best E_z and E_y (weak, strong decoder exponents respectively) while the pair (E_z, E_y) is attainable, compared to Gallager's 74' results $(E_{g,z}, E_{g,y})$.



Results for the Broadcast BSC. Cont.

Instead of requiring that the other (the one which is not drawn in each plot) exponent be positive, we can require that it will be greater then some constant value. In this case, we required that the other exponent will be greater then $\frac{1}{4}$ of its maximal attainable value.

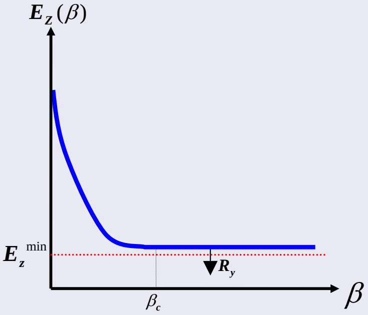




Discussion

Why is E_y discontinuous?

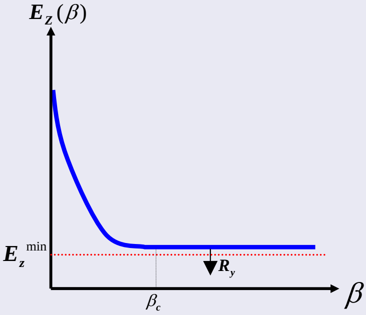
- We look for: $\max_\beta E_y$ s.t $E_z > E_z^{min}$
- If we look at E_z as a function of β :



Discussion

Why is E_y discontinuous?

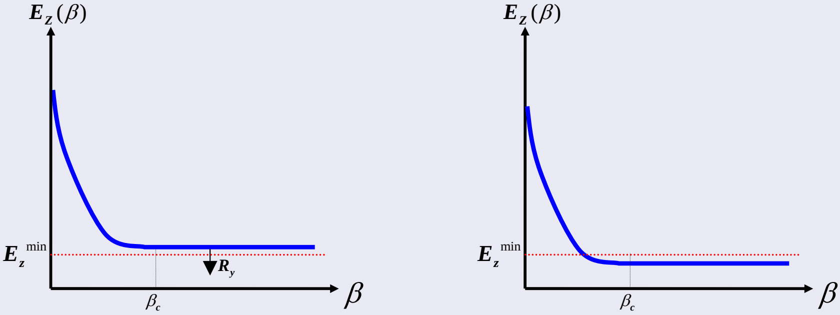
- We look for: $\max_\beta E_y$ s.t $E_z > E_z^{min}$
- If we look at E_z as a function of β :



Discussion

Why is E_y discontinuous?

- We look for: $\max_\beta E_y$ s.t $E_z > E_z^{min}$
- If we look at E_z as a function of β :



- For all $R_y > R_{y0}$ the maximization is constrained.

Part II:
Type Class Enumerator Method

Type Class Enumerator Method - Motivation

A question

In the previous section (and numerous works) Jensen's inequality is used.

$$A = \boldsymbol{E}\left[\sum_m P(\boldsymbol{y}|X_m)\right]^s \leq \left[\boldsymbol{E}\sum_m P(\boldsymbol{y}|X_m)\right]^s = B$$

Did we lose exponential tightness here? is $A \doteq B$? $(\lim_{n\to\infty} \frac{1}{n} log \frac{A}{B} = 0)$

Type Class Enumerator Method - Motivation

A question

In the previous section (and numerous works) Jensen's inequality is used.

$$A = \boldsymbol{E}\left[\sum_m P(\boldsymbol{y}|X_m)\right]^s \leq \left[\boldsymbol{E}\sum_m P(\boldsymbol{y}|X_m)\right]^s = B$$

Did we lose exponential tightness here? is $A \doteq B$? $(\lim_{n\to\infty} \frac{1}{n} log \frac{A}{B} = 0)$

Partial answer

For random coding, Gallager's exponent is tight - in his case $A \doteq B$.

- Was Gallager "lucky"?

Type Class Enumerator Method - Motivation

A question

In the previous section (and numerous works) Jensen's inequality is used.

$$A = \boldsymbol{E}\left[\sum_m P(\boldsymbol{y}|X_m)\right]^s \leq \left[\boldsymbol{E}\sum_m P(\boldsymbol{y}|X_m)\right]^s = B$$

Did we lose exponential tightness here? is $A \doteq B$? $(\lim_{n\to\infty} \frac{1}{n} log \frac{A}{B} = 0)$

Partial answer

For random coding, Gallager's exponent is tight - in his case $A \doteq B$.

- Was Gallager "lucky"?
- We will return to this question later.

A Simple Case First

BSC(p) $\beta = \log \frac{1-p}{p}$

$$\boldsymbol{E}\left[\sum_{m=1}^{e^{nR}} P(\boldsymbol{z}|\boldsymbol{X}_m)\right]^{s} = (1-p)^{ns}\boldsymbol{E}\left[\sum_{d=0}^{n} N(d)e^{-d\beta}\right]^{s}$$

- $\boldsymbol{z}$ is the "zero" word.
- $N(d)$ - # codewords around with Hamming weight d.

A Simple Case First

BSC(p) $\beta = \log \frac{1-p}{p}$

$$\boldsymbol{E}\left[\sum_{m=1}^{e^{nR}} P(\boldsymbol{z}|\boldsymbol{X}_m)\right]^{s} = (1-p)^{ns}\boldsymbol{E}\left[\sum_{d=0}^{n} N(d)e^{-d\beta}\right]^{s}$$

$$\doteq (1-p)^{ns}\boldsymbol{E}\left[\sum_{d=0}^{n} N^{s}(d)e^{-ns\beta\frac{d}{n}}\right]$$

- $\boldsymbol{z}$ is the "zero" word.
- $N(d)$ - # codewords around with Hamming weight d.

A Simple Case First

BSC(p) $\beta = \log \frac{1-p}{p}$

$$\boldsymbol{E}\left[\sum_{m=1}^{e^{nR}} P(\boldsymbol{z}|\boldsymbol{X}_m)\right]^s = (1-p)^{ns}\boldsymbol{E}\left[\sum_{d=0}^{n} N(d)e^{-d\beta}\right]^s$$

$$\doteq (1-p)^{ns}\boldsymbol{E}\left[\sum_{d=0}^{n} N^s(d)e^{-ns\beta\frac{d}{n}}\right]$$

$$= (1-p)^{ns}\sum_{d=0}^{n} \boldsymbol{E}N^s(d)e^{-ns\beta\frac{d}{n}}$$

- $\boldsymbol{z}$ is the "zero" word.
- $N(d)$ - # codewords around with Hamming weight d.

A Simple Case First

BSC(p) $\beta = \log \frac{1-p}{p}$

$$\boldsymbol{E}\left[\sum_{m=1}^{e^{nR}} P(\boldsymbol{z}|\boldsymbol{X}_m)\right]^{s} = (1-p)^{ns}\boldsymbol{E}\left[\sum_{d=0}^{n} N(d)e^{-d\beta}\right]^{s}$$

$$\doteq (1-p)^{ns}\boldsymbol{E}\left[\sum_{d=0}^{n} N^{s}(d)e^{-ns\beta\frac{d}{n}}\right]$$

$$= (1-p)^{ns}\sum_{d=0}^{n}\boldsymbol{E}N^{s}(d)e^{-ns\beta\frac{d}{n}}$$

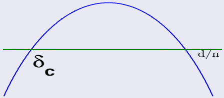
- z is the "zero" word.
- $N(d)$ - # codewords around with Hamming weight d.
- We need to evaluate the moments of $N(d)$.

Some Observations

We draw e^{nR} codewords independently and uniformly over {0,1}.
Let $N(d)$ count the number of codewords with Hamming weight d.

The Expected number of such codewords

$$\boldsymbol{E}N(d) \doteq e^{nR}e^{n(h(\frac{d}{n})-\log 2)} \triangleq e^{n\cdot g(R,d)}$$



- *sub-exponential* number of possible d's
- $g(R,d) > 0 \Rightarrow N(d)$ converges to its expectation d.e.f
- $g(R,d) < 0 \Rightarrow \Pr(N(d) = 1) = e^{n\cdot g(R,d)}$
- Moments of $N(d)$

$$\boldsymbol{E}N^{s}(d) = \begin{cases} e^{ns\cdot g(R,d)} & g(R,d) > 0 \\ e^{n\cdot g(R,d)} & g(R,d) \le 0 \end{cases}$$

Deriving the Exponent

The weak decoder

For the weak decoder, we start with Gallager's upper bound to the "channel" $P(\boldsymbol{z}|m) = \frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})$

$$\overline{P_{E_m}^{z}} \le \sum_{\boldsymbol{z}} \boldsymbol{E}\left[\frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right]^{1-\rho\lambda} \boldsymbol{E}\left[\sum_{m'\neq m}\left(\frac{1}{M_y}\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho}$$

Type class enumerators approach:

Unlike previous works that use Jensen's inequality, after this initial step, our analysis is exponentially tight.

Deriving the Exponent

The weak decoder

For the weak decoder, we start with Gallager's upper bound to the "channel" $P(\boldsymbol{z}|m) = \frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})$

$$\overline{P^{z}_{E_m}} \leq \sum_{\boldsymbol{z}} \boldsymbol{E}\left[\frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right]^{1-\rho\lambda} \boldsymbol{E}\left[\sum_{m'\neq m}\left(\frac{1}{M_y}\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho}$$

Type class enumerators approach:

Unlike previous works that use Jensen's inequality, after this initial step, our analysis is exponentially tight.

Deriving the Exponent. Cont.

The First Expectation

$$\boldsymbol{E}\left[\sum_{i=1}^{e^{nR_y}} P(z|x_{m,i})\right]^{s}$$

$$= \boldsymbol{E}_u \boldsymbol{E}_{x|u}\left[\sum_{\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}} N_{z,m}(\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}) e^{n\hat{\mathbf{E}}_{\boldsymbol{z}\boldsymbol{x}} \log P(z|x)}\right]^{s}$$

- $N_{z,m}(\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})$ - # codewords around cloud m belonging to $T_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}$. We need to evaluate its moments.
- $\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}$ plays the role of d in the binary example.

Deriving the Exponent. Cont.

The First Expectation

$$\boldsymbol{E}\left[\sum_{i=1}^{e^{nR_y}} P(z|x_{m,i})\right]^{s}$$

$$= \boldsymbol{E}_u \boldsymbol{E}_{x|u}\left[\sum_{\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}} N_{z,m}(\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}) e^{n\hat{\mathbf{E}}_{\boldsymbol{z}\boldsymbol{x}} \log P(z|x)}\right]^{s}$$

$$\doteq \boldsymbol{E}_u \boldsymbol{E}_{x|u}\left[\sum_{\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}} N^{s}_{z,m}(\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}) e^{ns\hat{\mathbf{E}}_{\boldsymbol{z}\boldsymbol{x}} \log P(z|x)}\right]$$

- $N_{z,m}(\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})$ - # codewords around cloud m belonging to $T_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}$. We need to evaluate its moments.
- $\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}$ plays the role of d in the binary example.

The Enumerator $N_{z,m}(T_{x|z})$ Properties

Given cloud $\boldsymbol{u}$, the expected number of codewords of type $T_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}$ is:

$$\begin{aligned} \boldsymbol{E}_{x|u} N_{z,m}(\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}) &\doteq e^{nR_y} \cdot e^{-n\left(\hat{\mathbf{E}}_{\boldsymbol{x}\boldsymbol{u}} \log \frac{1}{P(x|u)} - \hat{H}(\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u})\right)} \\ &\triangleq e^{n \cdot g(R_y, \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})} \end{aligned}$$

Denote

$$\mathcal{G}_{R_y}(u) = \left\{ \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}} : g(R_y, \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}) > 0 \right\}$$

and the moments of the enumerators are:

$$\boldsymbol{E}_{x|u} N^{\mathsf{s}}_{z,m}(\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}) \doteq \begin{cases} e^{n\mathsf{s} \cdot g(R_y, \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})} & \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}} \in \mathcal{G}_{R_y}(u) \\ e^{n \cdot g(R_y, \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})} & \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}} \in \mathcal{G}^c_{R_y}(u) \end{cases}$$

The Enumerator $N_{z,m}(T_{x|z})$ Properties

Given cloud $\boldsymbol{u}$, the expected number of codewords of type $T_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}$ is:

$$\begin{aligned} \boldsymbol{E}_{x|u} N_{z,m}(\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}) &\doteq e^{nR_y} \cdot e^{-n(\hat{\mathbf{E}}_{\boldsymbol{x}\boldsymbol{u}} \log \frac{1}{P(x|u)} - \hat{H}(\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}))} \\ &\triangleq e^{n \cdot g(R_y, \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})} \end{aligned}$$

Denote

$$\mathcal{G}_{R_y}(u) = \left\{ \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}} : g(R_y, \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}) > 0 \right\}$$

and the moments of the enumerators are:

$$\boldsymbol{E}_{x|u} N^{\mathbf{s}}_{z,m}(\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}) \doteq \begin{cases} e^{n\mathbf{s} \cdot g(R_y, \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})} & \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}} \in \mathcal{G}_{R_y}(u) \\ e^{n \cdot g(R_y, \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})} & \hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}} \in \mathcal{G}^c_{R_y}(u) \end{cases}$$

Now, split the sum over $\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}$ into sums over $\mathcal{G}_{R_y}$, $\mathcal{G}^c_{R_y}$

Deriving the Exponent. Cont.

$$\boldsymbol{E}\left[\sum_{i=1}^{M_y} P(z|x_{m,i})\right]^{s} \doteq \boldsymbol{E}_u\left[\sum_{\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}\in\mathcal{G}_{R_y}} e^{ns\left(g(R_y,\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})+\hat{\mathbf{E}}_{\boldsymbol{z}\boldsymbol{x}}\log P(z|x)\right)} + \sum_{\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}}\in\mathcal{G}^c_{R_y}} e^{n\left(g(R_y,\hat{Q}_{\boldsymbol{x}|\boldsymbol{z},\boldsymbol{u}})+s\hat{\mathbf{E}}_{\boldsymbol{z}\boldsymbol{x}}\log P(z|x)\right)}\right]$$

$$\doteq \boldsymbol{E}_u\left[e^{n\max_{\mathcal{G}_R} A(u,s)} + e^{n\max_{\mathcal{G}^c_R} B(u,s)}\right]$$

Observations:

- In Jensen's ineq, we take $\max A(u,s)$ $(\geq \max_{\mathcal{G}^c_R} B(u,s), 0 \leq s \leq 1)$.
- if $\max A(u,s) = \max\left\{\max_{\mathcal{G}_R} A(u,s), \max_{\mathcal{G}^c_R} B(u,s)\right\} \Rightarrow$ Jensen's ineq is tight.

Revisiting Gallager's Single User Bound

The 2nd expectation of Gallager's Bound

Our approach:

$$\boldsymbol{E}\left[\sum_{m'} P^{\frac{1}{1+\rho}}(\boldsymbol{y}|\boldsymbol{x}'_m)\right]^{\rho}$$
$$\doteq e^{n\max_{\mathcal{G}_R} A(\rho)} + e^{n\max_{\mathcal{G}_R^c} B(\rho)}$$

Jensen's inequality:

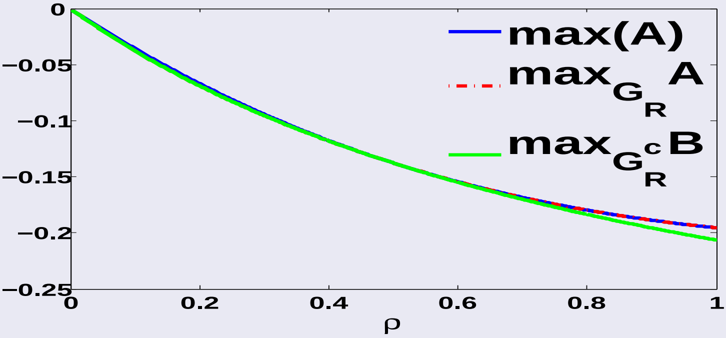
$$\boldsymbol{E}\left[\sum_{m'} P^{\frac{1}{1+\rho}}(\boldsymbol{y}|\boldsymbol{x}'_m)\right]^{\rho}$$
$$\leq e^{n\max A(\rho)}$$

- Let $\rho^*(R)$ be Gallager's optimizing ρ for each R.
- Different behavior for $R > R_c$ and $R < R_c$

$R > R_c$

$$\boldsymbol{E}\left[\sum_{m'} P^{\frac{1}{1+\rho}}(\boldsymbol{y}|\boldsymbol{x}'_m)\right]^{\rho} \doteq e^{n\max_{\mathcal{G}_R} A(\rho)} + e^{n\max_{\mathcal{G}_R^c} B(\rho)}$$

Behavior of A and B

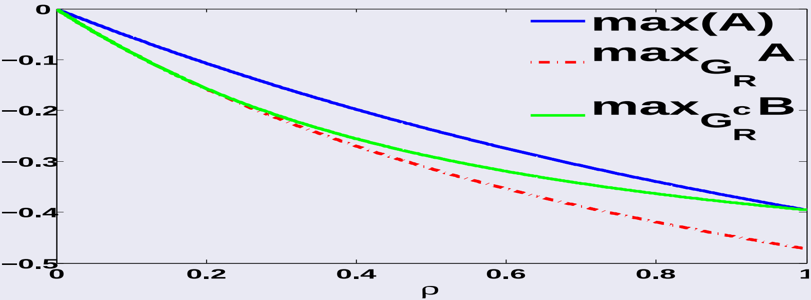


- $\max_{\mathcal{G}_R} A(\rho) \geq \max_{\mathcal{G}_R^c} B(\rho)$.
- The global maximizer of $A(\rho) \in \mathcal{G}_R$ for $\rho \geq \rho_0$.
- Gallager's bound is tight $\Rightarrow \rho^*(R) \geq \rho_0$

$R < R_c \ (\rho^*(R) = 1)$

$$\boldsymbol{E}\left[\sum_{m'} P^{\frac{1}{1+\rho}}(\boldsymbol{y}|\boldsymbol{x}'_m)\right]^{\rho} \doteq e^{n \max_{\mathcal{G}_R} A(\rho)} + e^{n \max_{\mathcal{G}_R^c} B(\rho)}$$

Behavior of A and B

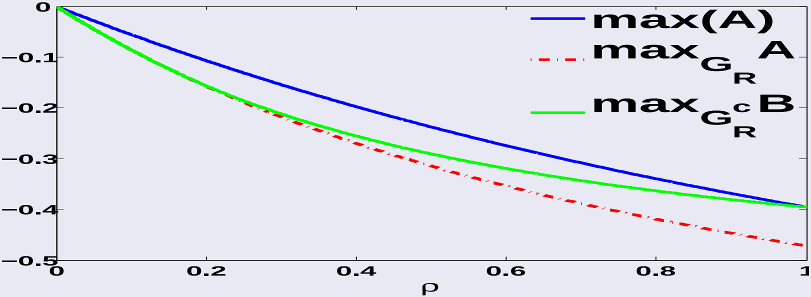


- $\max_{\mathcal{G}_R} A(\rho) \leq \max_{\mathcal{G}_R^c} B(\rho) \leq \max A(\rho)$ (?!)

$R < R_c$ $(\rho^*(R) = 1)$

$$\boldsymbol{E}\left[\sum_{m'} P^{\frac{1}{1+\rho}}(\boldsymbol{y}|\boldsymbol{x}'_m)\right]^{\rho} \doteq e^{n \max_{\mathcal{G}_R} A(\rho)} + e^{n \max_{\mathcal{G}_R^c} B(\rho)}$$

Behavior of A and B

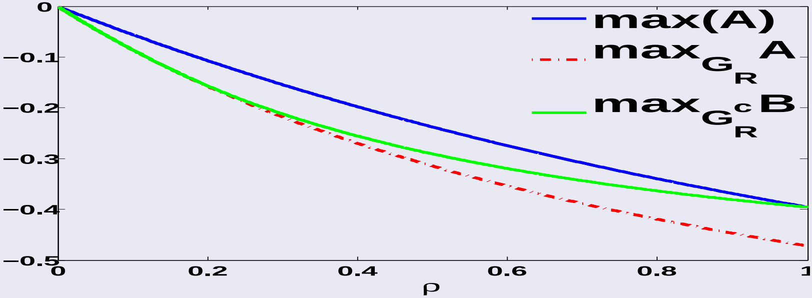


- $\max_{\mathcal{G}_R} A(\rho) \leq \max_{\mathcal{G}_R^c} B(\rho) \leq \max A(\rho)$ (?!)
- However, Gallager's bound is tight because:
  1. $\max_{\mathcal{G}_R^c} B(1) = \max A(1)$.

$R < R_c$ $(\rho^*(R) = 1)$

$$\boldsymbol{E}\left[\sum_{m'} P^{\frac{1}{1+\rho}}(\boldsymbol{y}|\boldsymbol{x}'_m)\right]^\rho \doteq e^{n \max_{\mathcal{G}_R} A(\rho)} + e^{n \max_{\mathcal{G}_R^c} B(\rho)}$$

Behavior of A and B

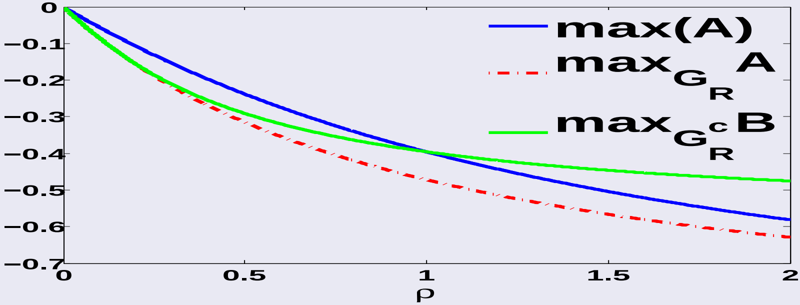


- $\max_{\mathcal{G}_R} A(\rho) \leq \max_{\mathcal{G}_R^c} B(\rho) \leq \max A(\rho)$ (?!)
- However, Gallager's bound is tight because:
  1. $\max_{\mathcal{G}_R^c} B(1) = \max A(1)$.
  2. Searching above $\rho = 1$ doesn't improve the overall bound.

$R < R_c \ (\rho^*(R) = 1)$

$$\boldsymbol{E}\left[\sum_{m'} P^{\frac{1}{1+\rho}}(\boldsymbol{y}|\boldsymbol{x}'_m)\right]^{\rho} \doteq e^{n \max_{\mathcal{G}_R} A(\rho)} + e^{n \max_{\mathcal{G}_R^c} B(\rho)}$$

Behavior of A and B



- $\max_{\mathcal{G}_R} A(\rho) \leq \max_{\mathcal{G}_R^c} B(\rho) \leq \max A(\rho)$ (?!)
- However, Gallager's bound is tight because:
  1. $\max_{\mathcal{G}_R^c} B(1) = \max A(1)$.
  2. Searching above $\rho = 1$ doesn't improve the overall bound.

Deriving the Exponent

The weak decoder

For the weak decoder, we start with Gallager's upper bound to the "channel" $P(\boldsymbol{z}|m) = \frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})$

$$\overline{P^{z}_{E_m}} \leq \sum_{\boldsymbol{z}} \boldsymbol{E}\left[\frac{1}{M_y}\sum_{i=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m,i})\right]^{1-\rho\lambda} \boldsymbol{E}\left[\sum_{m'\neq m}\left(\frac{1}{M_y}\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho}$$

Type class enumerators approach:

Unlike previous works that use Jensen's inequality, after this initial step, our analysis is exponentially tight.

The Second Expectation

The Second Expectation - Binary example $\left(\beta = \log \frac{1-p}{p}\right)$.

$$\boldsymbol{E}\left[\sum_{m'\neq m}\left(\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho} = \boldsymbol{E}\left[\sum_{m'\neq m}\left(\sum_{d=0}^{n} N_{\boldsymbol{z},m'}(d)e^{-d\beta}\right)^{\lambda}\right]^{\rho}$$

The Second Expectation

The Second Expectation - Binary example $\left(\beta = \log \frac{1-p}{p}\right)$.

$$\boldsymbol{E}\left[\sum_{m' \neq m}\left(\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho} = \boldsymbol{E}\left[\sum_{m' \neq m}\left(\sum_{d=0}^{n} N_{\boldsymbol{z},m'}(d) e^{-d\beta}\right)^{\lambda}\right]^{\rho}$$

$$\doteq \boldsymbol{E}\left[\sum_{m' \neq m}\sum_{d=0}^{n} N^{\lambda}_{\boldsymbol{z},m'}(d) e^{-d\beta\lambda}\right]^{\rho}$$

The Second Expectation

The Second Expectation - Binary example $\left(\beta = \log \frac{1-p}{p}\right)$.

$$\begin{aligned}
\boldsymbol{E}\left[\sum_{m'\neq m}\left(\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho} &= \boldsymbol{E}\left[\sum_{m'\neq m}\left(\sum_{d=0}^{n} N_{\boldsymbol{z},m'}(d)e^{-d\beta}\right)^{\lambda}\right]^{\rho} \\
&\doteq \boldsymbol{E}\left[\sum_{m'\neq m}\sum_{d=0}^{n} N_{\boldsymbol{z},m'}^{\lambda}(d)e^{-d\beta\lambda}\right]^{\rho} \\
&\doteq \sum_{d=0}^{n} e^{-d\beta\lambda\rho}\boldsymbol{E}\left[\sum_{m'\neq m} N_{\boldsymbol{z},m'}^{\lambda}(d)\right]^{\rho}
\end{aligned}$$

The Second Expectation

The Second Expectation - Binary example $\left(\beta = \log \frac{1-p}{p}\right)$.

$$\begin{aligned}
\boldsymbol{E}\left[\sum_{m' \neq m} \left(\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho} &= \boldsymbol{E}\left[\sum_{m' \neq m} \left(\sum_{d=0}^{n} N_{\boldsymbol{z},m'}(d) e^{-d\beta}\right)^{\lambda}\right]^{\rho} \\
&\doteq \boldsymbol{E}\left[\sum_{m' \neq m} \sum_{d=0}^{n} N^{\lambda}_{\boldsymbol{z},m'}(d) e^{-d\beta\lambda}\right]^{\rho} \\
&\doteq \sum_{d=0}^{n} e^{-d\beta\lambda\rho} \boldsymbol{E}\left[\sum_{m' \neq m} N^{\lambda}_{\boldsymbol{z},m'}(d)\right]^{\rho}
\end{aligned}$$

Problems:

1. There is an exponential number of m'.

The Second Expectation

The Second Expectation - Binary example $\left(\beta = \log \frac{1-p}{p}\right)$.

$$\boldsymbol{E}\left[\sum_{m' \neq m}\left(\sum_{j=1}^{M_y} P(\boldsymbol{z}|\boldsymbol{x}_{m',j})\right)^{\lambda}\right]^{\rho} = \boldsymbol{E}\left[\sum_{m' \neq m}\left(\sum_{d=0}^{n} N_{\boldsymbol{z},m'}(d) e^{-d\beta}\right)^{\lambda}\right]^{\rho}$$

$$\doteq \boldsymbol{E}\left[\sum_{m' \neq m}\sum_{d=0}^{n} N_{\boldsymbol{z},m'}^{\lambda}(d) e^{-d\beta\lambda}\right]^{\rho}$$

$$\doteq \sum_{d=0}^{n} e^{-d\beta\lambda\rho} \boldsymbol{E}\left[\sum_{m' \neq m} N_{\boldsymbol{z},m'}^{\lambda}(d)\right]^{\rho}$$

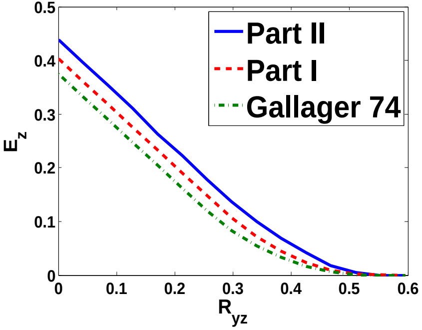
Problems:

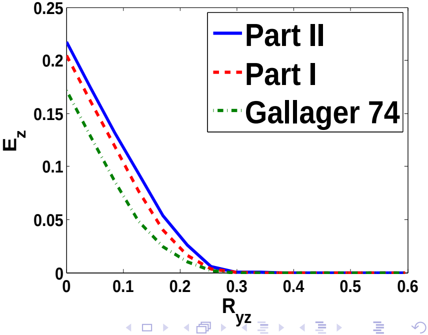
1. There is an exponential number of m'.
2. For every m', $N_{m'}(d)$ is distributed differently.

Results for the Broadcast BSC

Same setup as before

Numerical results of the weak decoder error exponent. We show the best E_z while the pair (E_z, E_y) is attainable, compared to the previous section and to Gallager 74' result. We show numerical results for two values of R_y.





Summary and Conclusions

- New and improved lower bounds to the error exponents of *optimum decoding* were given.

Summary and Conclusions

- New and improved lower bounds to the error exponents of *optimum decoding* were given.
- The dependence on both rates reveled the discontinuity we saw.

Summary and Conclusions

- New and improved lower bounds to the error exponents of *optimum decoding* were given.
- The dependence on both rates reveled the discontinuity we saw.
- An exponentially tight analysis technique was introduced.

Summary and Conclusions

- New and improved lower bounds to the error exponents of *optimum decoding* were given.
- The dependence on both rates reveled the discontinuity we saw.
- An exponentially tight analysis technique was introduced.
- The new technique gives insight into the tightness of Jensen's inequality.

Summary and Conclusions

- New and improved lower bounds to the error exponents of *optimum decoding* were given.
- The dependence on both rates reveled the discontinuity we saw.
- An exponentially tight analysis technique was introduced.
- The new technique gives insight into the tightness of Jensen's inequality.
- The new technique is applicable to other cases as well. eg. the interference channel.

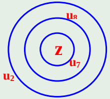
Summary and Conclusions

- New and improved lower bounds to the error exponents of *optimum decoding* were given.
- The dependence on both rates reveled the discontinuity we saw.
- An exponentially tight analysis technique was introduced.
- The new technique gives insight into the tightness of Jensen's inequality.
- The new technique is applicable to other cases as well. eg. the interference channel.

Thank you!

The Second Expectation. Cont.

Rearrange the cloud centers



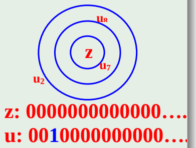
z: 0000000000000....

u: 0010000000000....

The Second Expectation. Cont.

Rearrange the cloud centers

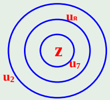
- For $\boldsymbol{u}_{m'}$ with $d_H(\boldsymbol{u}_{m'}, \boldsymbol{z}) = l_{\boldsymbol{u}\boldsymbol{z}}$, $N_{\boldsymbol{z},m'}(d)$ are i.d



The Second Expectation. Cont.

Rearrange the cloud centers

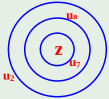
- For $\boldsymbol{u}_{m'}$ with $d_H(\boldsymbol{u}_{m'}, \boldsymbol{z}) = l_{\boldsymbol{uz}}$, $N_{\boldsymbol{z},m'}(d)$ are i.d
- $M(l_{\boldsymbol{uz}})$ - set of cloud centers distanced $l_{\boldsymbol{uz}}$ from $\boldsymbol{z}$



The Second Expectation. Cont.

Rearrange the cloud centers

- For $\boldsymbol{u}_{m'}$ with $d_H(\boldsymbol{u}_{m'}, \boldsymbol{z}) = l_{\boldsymbol{u}\boldsymbol{z}}$, $N_{\boldsymbol{z},m'}(d)$ are i.d
- $M(l_{\boldsymbol{u}\boldsymbol{z}})$ - set of cloud centers distanced $l_{\boldsymbol{u}\boldsymbol{z}}$ from $\boldsymbol{z}$



z: 0000000000000....
u: 0010000000000....

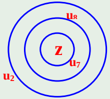
$$\boldsymbol{E}\left[\sum_{m'\neq m} N_{m'}^{\lambda}(d)\right]^{\rho} = \boldsymbol{E}\left[\sum_{l_{uz}=0}^{n} \sum_{\boldsymbol{u}_{m'}\in M(l_{uz})} N_{m'}^{\lambda}(d)\right]^{\rho}$$

$$= \sum_{l_{\boldsymbol{u}\boldsymbol{z}}=0}^{n} \boldsymbol{E}\left[\sum_{\boldsymbol{u}_{m'}\in M(l_{\boldsymbol{u}\boldsymbol{z}})} N_{m'}^{\lambda}(d)\right]^{\rho}$$

The Second Expectation. Cont.

Rearrange the cloud centers

- For $\boldsymbol{u}_{m'}$ with $d_H(\boldsymbol{u}_{m'}, \boldsymbol{z}) = l_{\boldsymbol{uz}}$, $N_{\boldsymbol{z},m'}(d)$ are i.d
- $M(l_{\boldsymbol{uz}})$ - set of cloud centers distanced $l_{\boldsymbol{uz}}$ from $\boldsymbol{z}$



z: 0000000000000....
u: 0010000000000....

$$\boldsymbol{E}\left[\sum_{m'\neq m} N_{m'}^{\lambda}(d)\right]^{\rho} = \boldsymbol{E}\left[\sum_{l_{uz}=0}^{n} \sum_{\boldsymbol{u}_{m'}\in M(l_{uz})} N_{m'}^{\lambda}(d)\right]^{\rho}$$

$$= \sum_{l_{\boldsymbol{uz}}=0}^{n} \boldsymbol{E}\left[\sum_{\boldsymbol{u}_{m'}\in M(l_{\boldsymbol{uz}})} N_{m'}^{\lambda}(d)\right]^{\rho}$$

$|M(l_{\boldsymbol{uz}})|$ is an enumerator - behaves similarly to $N(d)$.