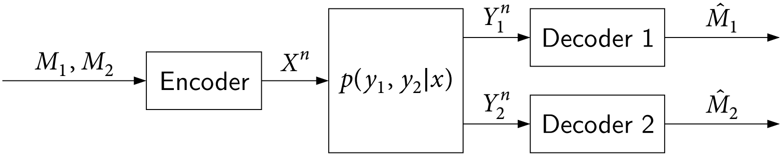
Binary Input Broadcast Channel

Abbas El Gamal

Stanford University

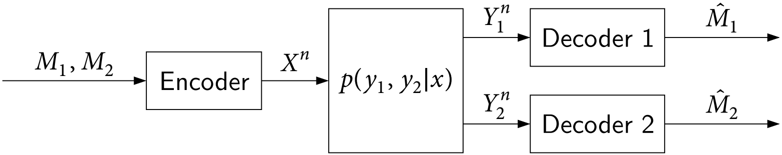
ITW, Jerusalem 2015

Broadcast channel (BC) (Cover 1972)



- Private messages: $M_1 \in [1:2^{nR_1}]$, $M_2 \in [1:2^{nR_2}]$
- Capacity region: Closure of set of achievable rate pairs (R_1, R_2)
- Capacity region is not known in general

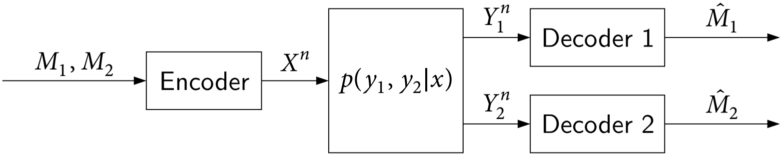
Broadcast channel (BC) (Cover 1972)



- Private messages: $M_1 \in [1 : 2^{nR_1}]$, $M_2 \in [1 : 2^{nR_2}]$
- Capacity region: Closure of set of achievable rate pairs (R_1, R_2)
- Capacity region is not known in general
- Inner and outer bounds that coincide for some special classes

  But are in general very difficult to evaluate

Broadcast channel (BC) (Cover 1972)

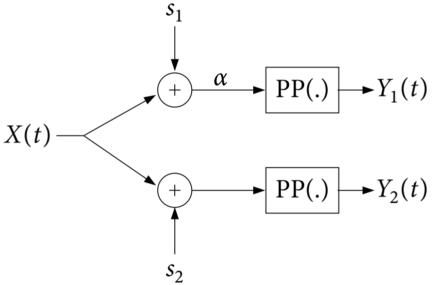


- Private messages: $M_1 \in [1 : 2^{nR_1}]$, $M_2 \in [1 : 2^{nR_2}]$
- Capacity region: Closure of set of achievable rate pairs (R_1, R_2)
- Capacity region is not known in general
- Inner and outer bounds that coincide for some special classes

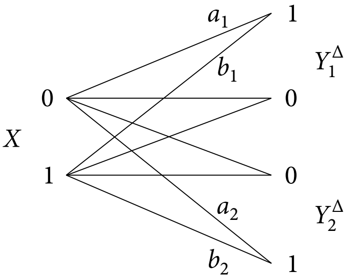
  But are in general very difficult to evaluate
- Recent progress on computing the bounds for binary input BC ($|\mathcal{X}| \le 2$)

Motivation

- Study achievable rates for Poisson BC (Kim–Nachman–EG 2015)



Continuous time Poisson BC



Binary P-BC (Wyner 1988a)

$$Y_1(t)|\{X(t) = x(t)\} \text{ is } \mathrm{PP}(\alpha(x(t) + s_1)),$$

$$Y_2(t)|\{X(t) = x(t)\} \text{ is } \mathrm{PP}(x(t) + s_2),$$

$$X(t) \in [0, 1],\ t, \alpha, s_1, s_2 \geq 0,\ s_1 \leq s_2$$

$$a_1 = \alpha s_1 \Delta + O(\Delta^2),$$

$$a_2 = s_2 \Delta + O(\Delta^2),$$

$$b_1 = \alpha(1 + s_1)\Delta + O(\Delta^2),$$

$$b_2 = (1 + s_2)\Delta + O(\Delta^2)$$

Motivation

- Study achievable rates for Poisson BC (Kim–Nachman–EG 2015)
- We used recent techniques/results on binary input BC:
  - Concave envelope method (Nair–Wang 2008)
  - Evaluation of Marton's region (Anantharam–Gohari–Nair 2013)

Motivation

- Study achievable rates for Poisson BC (Kim–Nachman–EG 2015)
- We used recent techniques/results on binary input BC:
  - Concave envelope method (Nair–Wang 2008)
  - Evaluation of Marton's region (Anantharam–Gohari–Nair 2013)
- This talk:
  - Highlight some of these techniques/results
  - Describe new results for the Poisson BC
  - Focus on computing capacity and bounds (no new coding schemes)

Outline

I. Optimality of superposition coding for BI-BC

BEC/BSC (Nair 2010) (proposed by Montanari)

- Superposition coding is always optimal
- Concave envelope method
- Degraded $\underset{\not\leftarrow}{\rightarrow}$ less noisy $\underset{\not\leftarrow}{\rightarrow}$ more capable
- Effectively less noisy

II. Poisson broadcast channel (Kim–Nachman–EG 2015):

- Superposition coding is almost always optimal

III. Evaluation of Marton region for BI-BC (Anantharam–Gohari–Nair 2013)

- Gap between Marton region and UV outer bound for Poisson BC

Outline

I. Optimality of superposition coding for BI-BC

BEC/BSC (Nair 2010) (proposed by Montanari)

- Superposition coding is always optimal
- Concave envelope method
- Degraded $\rightleftarrows$ less noisy $\rightleftarrows$ more capable
- Effectively less noisy

II. Poisson broadcast channel (Kim–Nachman–EG 2015):

- Superposition coding is almost always optimal

III. Evaluation of Marton region for BI-BC (Anantharam–Gohari–Nair 2013)

- Gap between Marton region and UV outer bound for Poisson BC

Outline

I. Optimality of superposition coding for BI-BC

BEC/BSC (Nair 2010) (proposed by Montanari)

– Superposition coding is always optimal

– Concave envelope method

– Degraded $\rightleftarrows$ less noisy $\rightleftarrows$ more capable

– Effectively less noisy

II. Poisson broadcast channel (Kim–Nachman–EG 2015):

- Superposition coding is almost always optimal

III. Evaluation of Marton region for BI-BC (Anantharam–Gohari–Nair 2013)

- Gap between Marton region and UV outer bound for Poisson BC

Superposition coding inner bound for BC

Superposition coding inner bound $\mathscr{R}_{UX}$ (Cover 1972, Bergmans 1973)

A rate pair (R_1, R_2) is achievable for the BC if

$$
\begin{aligned}
R_2 &< I(U; Y_2), \\
R_1 &< I(X; Y_1 | U), \\
R_1 + R_2 &< I(X; Y_1)
\end{aligned}
$$

for some pmf $p(u, x)$

- Y_1 recovers both M_1 and M_2
- Y_2 recovers M_2 (carried by U)
- $|\mathcal{U}| \le |\mathcal{X}| + 1$ suffices (Ahlswede–Körner 1975)

Classes of BC for which superposition coding is optimal

- Degraded (Gallager 1974): $p(y_1'|x) = p(y_1|x)$ such that $X \to Y_1' \to Y_2$
- Less noisy (Körner–Marton 1977): $I(U;Y_1) \ge I(U;Y_2)$ for every $p(u,x)$
- More capable (EG 1979): $I(X;Y_1) \ge I(X;Y_2)$ for every $p(x)$

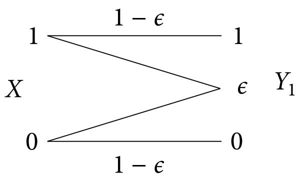
Classes of BC for which superposition coding is optimal

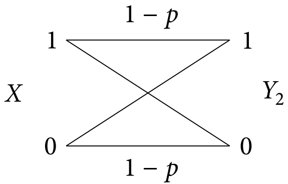
- Degraded (Gallager 1974): $p(y_1'|x) = p(y_1|x)$ such that $X \to Y_1' \to Y_2$
- Less noisy (Körner–Marton 1977): $I(U; Y_1) \geq I(U; Y_2)$ for every $p(u, x)$
- More capable (EG 1979): $I(X; Y_1) \geq I(X; Y_2)$ for every $p(x)$
- Degraded $\underset{\not\leftarrow}{\rightarrow}$ less noisy $\underset{\not\leftarrow}{\rightarrow}$ more capable (Körner–Marton 1977)

Classes of BC for which superposition coding is optimal

- Degraded (Gallager 1974): $p(y_1'|x) = p(y_1|x)$ such that $X \to Y_1' \to Y_2$
- Less noisy (Körner–Marton 1977): $I(U;Y_1) \geq I(U;Y_2)$ for every $p(u,x)$
- More capable (EG 1979): $I(X;Y_1) \geq I(X;Y_2)$ for every $p(x)$
- Degraded $\underset{\not\leftarrow}{\rightarrow}$ less noisy $\underset{\not\leftarrow}{\rightarrow}$ more capable (Körner–Marton 1977)
- Essentially less noisy and more capable (Nair 2010)
  - Less noisy $\underset{\not\leftarrow}{\rightarrow}$ essentially less noisy
  - More capable $\underset{\not\leftarrow}{\rightarrow}$ essentially more capable

BEC/BSC (Nair 2010)

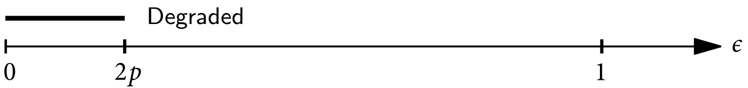




- Superposition coding is always optimal
- Concave envelope
- Degraded $\underset{\not\leftarrow}{\rightarrow}$ less noisy $\underset{\not\leftarrow}{\rightarrow}$ more capable
- Effectively less noisy BC

Y_2 degraded version of Y_1 for BEC/BSC

- Y_2 degraded version of Y_1 if $0 \le \epsilon \le 2p$



Less noisy via concave envelope (van Dijk 1997)

- Y_1 less noisy than Y_2 if: $I(U; Y_1) \geq I(U; Y_2)$ for every $p(u, x)$
- Equivalently if: $\max_{p(u,x)} \big(I(U; Y_2) - I(U; Y_1)\big) = 0,$ $|\mathcal{U}| \leq |\mathcal{X}| + 1$

Less noisy via concave envelope (van Dijk 1997)

- Y_1 less noisy than Y_2 if: $\max_{p(u,x)} \big(I(U;Y_2) - I(U;Y_1) \big) = 0,$ $|\mathcal{U}| \le |\mathcal{X}| + 1$
- Consider:

$$\max_{p(u,x)} \big(I(U;Y_2) - I(U;Y_1) \big) = \max_{p(u,x)} \big(I(X;Y_2) - I(X;Y_2|U) - I(X;Y_1) + I(X;Y_1|U) \big)$$

Less noisy via concave envelope (van Dijk 1997)

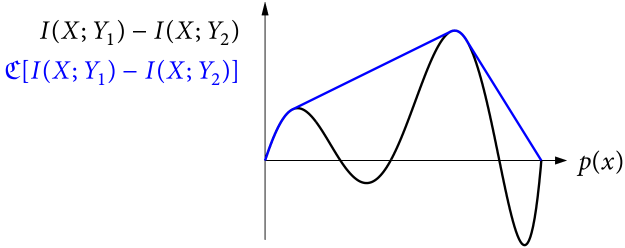
- Y_1 less noisy than Y_2 if: $\max_{p(u,x)}\big(I(U;Y_2)-I(U;Y_1)\big)=0,\ |\mathcal{U}|\le|\mathcal{X}|+1$
- Consider:

$$\begin{aligned}\max_{p(u,x)}\big(I(U;Y_2)-I(U;Y_1)\big) &= \max_{p(u,x)}\big(I(X;Y_2)-I(X;Y_2|U)-I(X;Y_1)+I(X;Y_1|U)\big)\\ &= \max_{p(x)}\Big(\max_{p(u|x):\,|\mathcal{U}|\le|\mathcal{X}|+1}\big(I(X;Y_1|U)-I(X;Y_2|U)\big)-\big(I(X;Y_1)-I(X;Y_2)\big)\Big)\end{aligned}$$

Less noisy via concave envelope (van Dijk 1997)

- Y_1 less noisy than Y_2 if: $\max_{p(u,x)} \big(I(U;Y_2) - I(U;Y_1)\big) = 0,\ |\mathcal{U}| \le |\mathcal{X}| + 1$
- Consider:

$$\begin{aligned}
\max_{p(u,x)} \big(I(U;Y_2) - I(U;Y_1)\big) &= \max_{p(u,x)} \big(I(X;Y_2) - I(X;Y_2|U) - I(X;Y_1) + I(X;Y_1|U)\big) \\
&= \max_{p(x)} \Big(\max_{p(u|x):\,|\mathcal{U}|\le|\mathcal{X}|+1} \big(I(X;Y_1|U) - I(X;Y_2|U)\big) - \big(I(X;Y_1) - I(X;Y_2)\big)\Big) \\
&= \max_{p(x)} \Big(\mathfrak{C}[I(X;Y_1) - I(X;Y_2)] - \big(I(X;Y_1) - I(X;Y_2)\big)\Big)
\end{aligned}$$



Less noisy via concave envelope (van Dijk 1997)

- Y_1 less noisy than Y_2 if: $\max_{p(u,x)} \big(I(U;Y_2) - I(U;Y_1)\big) = 0$, $|\mathcal{U}| \le |\mathcal{X}| + 1$
- Consider:

$$\begin{aligned}\max_{p(u,x)} \big(I(U;Y_2) - I(U;Y_1)\big) &= \max_{p(u,x)} \big(I(X;Y_2) - I(X;Y_2|U) - I(X;Y_1) + I(X;Y_1|U)\big)\\ &= \max_{p(x)} \Big(\max_{p(u|x):\, |\mathcal{U}|\le|\mathcal{X}|+1} \big(I(X;Y_1|U) - I(X;Y_2|U)\big) - \big(I(X;Y_1) - I(X;Y_2)\big)\Big)\\ &= \max_{p(x)} \Big(\mathfrak{C}[I(X;Y_1) - I(X;Y_2)] - \big(I(X;Y_1) - I(X;Y_2)\big)\Big)\end{aligned}$$

- Hence, Y_1 less noisy than Y_2 if:

  $I(X;Y_1) - I(X;Y_2) = \mathfrak{C}[I(X;Y_1) - I(X;Y_2)]$ for every $p(x)$

  Equivalently if, $I(X;Y_1) - I(X;Y_2)$ is concave in $p(x)$

Less noisy via concave envelope for BI-BC

- Y_1 less noisy than Y_2 if: $I(X; Y_1) - I(X; Y_2)$ concave in $p(x)$
- For BI-BC, define

$$I_j(q) = I(X; Y_j),\ \ j = 1, 2,\ \text{ where } X \sim \mathrm{Bern}(q),\ \ q \in [0, 1]$$

Less noisy via concave envelope for BI-BC

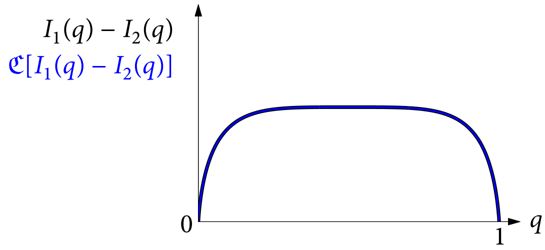
- Y_1 less noisy than Y_2 if: $I(X;Y_1) - I(X;Y_2)$ concave in $p(x)$
- For BI-BC, define

$$I_j(q) = I(X;Y_j), \ \ j = 1,2, \ \text{where } X \sim \mathrm{Bern}(q), \ \ q \in [0,1]$$

Y_1 less noisy than Y_2 if: Second derivative of $I_1(q) - I_2(q) \le 0$ for every q

Y_1 less noisy than Y_2 for BEC/BSC

- Y_1 less noisy than Y_2 if: $\epsilon \le 4p(1-p)$

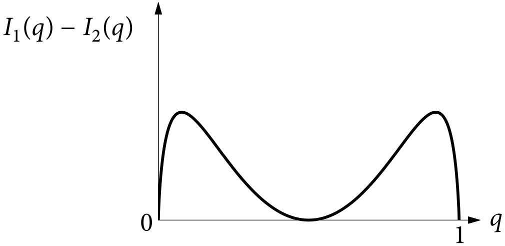


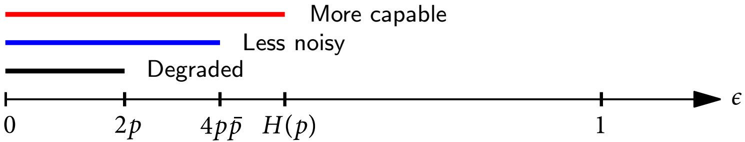
Y_1 more capable than Y_2 for BEC/BSC

- Y_1 more capable than Y_2 if: $I_1(q) - I_2(q) \geq 0$ for every $q \in [0, 1]$

Y_1 more capable than Y_2 for BEC/BSC

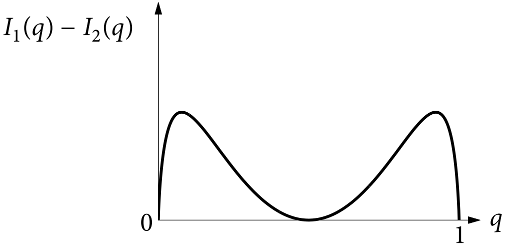
- Y_1 more capable if $\epsilon \le H(p)$

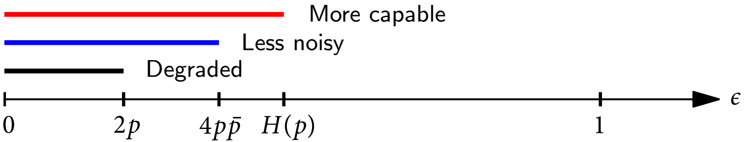




Y_1 more capable than Y_2 for BEC/BSC

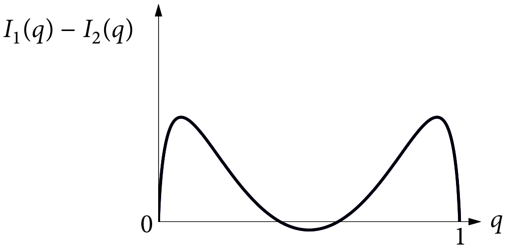
- Y_1 more capable if $\epsilon \le H(p)$
- This shows: Degraded $\overset{\longrightarrow}{\not\longleftarrow}$ less noisy $\overset{\longrightarrow}{\not\longleftarrow}$ more capable

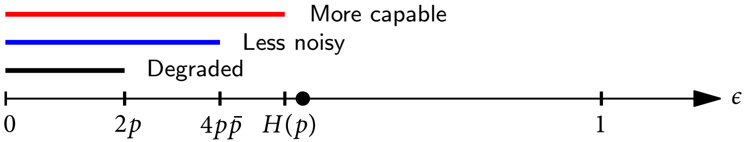




Beyond more capable

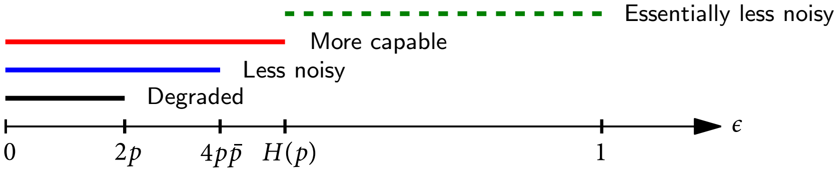
- For $H(p) < \epsilon < 1$, BEC/BSC is not more capable





Superposition coding is always optimal for BEC/BSC

- Nair (2010) showed: Superposition coding is optimal for $H(p) < \epsilon \leq 1$
  - Introduced class of essentially less noisy
  - Difficult to verify in general (when channel is not symmetric)



Superposition coding is always optimal for BEC/BSC

- Nair (2010) showed: Superposition coding is optimal for $H(p) < \epsilon \leq 1$
  - Introduced class of essentially less noisy
  - Difficult to verify in general (when channel is not symmetric)
- We show that the channel is effectively less noisy:
  - More general than essentially less noisy
  - Easier to verify using concave envelope

Effectively less noisy (Kim–Nachman–EG 2015)

- Consider the outer bound $\mathscr{R}_{\mathrm{o}}$ on BC capacity region (Marton 1979):

$$R_2 \leq I(U; Y_2),$$
$$R_1 + R_2 \leq I(U; Y_2) + I(X; Y_1 | U)$$

for some $p(u, x)$

Expressed via supporting hyperplanes:

$$\max_{(R_1, R_2) \in \mathscr{R}_{\mathrm{o}}} (\lambda R_1 + R_2) = \max_{p(u,x)} \big(\lambda I(X; Y_1 | U) + I(U; Y_2)\big) \quad \text{for } 0 \leq \lambda \leq 1$$

Effectively less noisy (Kim–Nachman–EG 2015)

- Consider the outer bound $\mathscr{R}_{\rm o}$ on BC capacity region (Marton 1979):

$$\begin{aligned} R_2 &\le I(U;Y_2), \\ R_1 + R_2 &\le I(U;Y_2) + I(X;Y_1|U) \end{aligned}$$

  for some $p(u,x)$

  Expressed via supporting hyperplanes:

$$\max_{(R_1,R_2)\in\mathscr{R}_{\rm o}} (\lambda R_1 + R_2) = \max_{p(u,x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big) \quad \text{for } 0 \le \lambda \le 1$$

- Compare to the superposition coding inner bound $\mathscr{R}_{UX}$:

$$\max_{(R_1,R_2)\in\mathscr{R}_{UX}} (\lambda R_1 + R_2) = \max_{p(u,x)} \big(\lambda \min\{I(X;Y_1|U), I(X;Y_1) - I(U;Y_2)\} + I(U;Y_2)\big)$$

  for $0 \le \lambda \le 1$

Effectively less noisy (Kim–Nachman–EG 2015)

- Consider the outer bound $\mathscr{R}_{\mathrm{o}}$ on BC capacity region (Marton 1979):

$$\begin{aligned} R_2 &\leq I(U;Y_2), \\ R_1 + R_2 &\leq I(U;Y_2) + I(X;Y_1|U) \end{aligned}$$

  for some $p(u,x)$

  Expressed via supporting hyperplanes:

$$\max_{(R_1,R_2)\in\mathscr{R}_{\mathrm{o}}} (\lambda R_1 + R_2) = \max_{p(u,x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big) \quad \text{for } 0 \leq \lambda \leq 1$$

- Compare to the superposition coding inner bound $\mathscr{R}_{UX}$:

$$\max_{(R_1,R_2)\in\mathscr{R}_{UX}} (\lambda R_1 + R_2) = \max_{p(u,x)} \big(\lambda \min\{I(X;Y_1|U), I(X;Y_1) - I(U;Y_2)\} + I(U;Y_2)\big)$$

  for $0 \leq \lambda \leq 1$

- Clearly, $\mathscr{R}_{UX} = \mathscr{R}_{\mathrm{o}}$ if $I(U;Y_1) \geq I(U;Y_2)$ for every $p(u,x)$ (less noisy)

Effectively less noisy (Kim–Nachman–EG 2015)

- Outer bound $\mathscr{R}_\text{o}$ on BC capacity region (Marton 1979):

$$\max_{(R_1,R_2)\in\mathscr{R}_\text{o}} (\lambda R_1 + R_2) = \max_{p(u,x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big) \quad \text{for } 0 \le \lambda \le 1$$

- Superposition coding inner bound $\mathscr{R}_{UX}$:

$$\max_{(R_1,R_2)\in\mathscr{R}_{UX}} (\lambda R_1 + R_2) = \max_{p(u,x)} \big(\lambda \min\{I(X;Y_1|U), I(X;Y_1) - I(U;Y_2)\} + I(U;Y_2)\big)$$

for $0 \le \lambda \le 1$

- Clearly, $\mathscr{R}_{UX} = \mathscr{R}_\text{o}$ if $I(U;Y_1) \ge I(U;Y_2)$ for every $p(u,x)$ (less noisy)

Effectively less noisy (Kim–Nachman–EG 2015)

- Outer bound $\mathscr{R}_{\rm o}$ on BC capacity region (Marton 1979):

$$\max_{(R_1,R_2)\in\mathscr{R}_{\rm o}}(\lambda R_1 + R_2) = \max_{p(u,x)}\big(\lambda I(X;Y_1|U) + I(U;Y_2)\big) \quad \text{for } 0 \le \lambda \le 1$$

- Superposition coding inner bound $\mathscr{R}_{UX}$:

$$\max_{(R_1,R_2)\in\mathscr{R}_{UX}}(\lambda R_1 + R_2) = \max_{p(u,x)}\big(\lambda \min\{I(X;Y_1|U), I(X;Y_1) - I(U;Y_2)\} + I(U;Y_2)\big)$$

for $0 \le \lambda \le 1$

- Clearly, $\mathscr{R}_{UX} = \mathscr{R}_{\rm o}$ if $I(U;Y_1) \ge I(U;Y_2)$ for every $p(u,x)$ (less noisy)

- This doesn't need to hold for every $p(u,x)$

Effectively less noisy (Kim–Nachman–EG 2015)

- Outer bound $\mathscr{R}_\mathrm{o}$ on BC capacity region (Marton 1979):

$$\max_{(R_1,R_2)\in\mathscr{R}_\mathrm{o}} (\lambda R_1 + R_2) = \max_{p(u,x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big) \quad \text{for } 0 \le \lambda \le 1$$

- Superposition coding inner bound $\mathscr{R}_{UX}$:

$$\max_{(R_1,R_2)\in\mathscr{R}_{UX}} (\lambda R_1 + R_2) = \max_{p(u,x)} \big(\lambda \min\{I(X;Y_1|U), I(X;Y_1) - I(U;Y_2)\} + I(U;Y_2)\big)$$

for $0 \le \lambda \le 1$

- Clearly, $\mathscr{R}_{UX} = \mathscr{R}_\mathrm{o}$ if $I(U;Y_1) \ge I(U;Y_2)$ for every $p(u,x)$ (less noisy)

- This doesn't need to hold for every $p(u,x)$, for example,

$\mathscr{R}_{UX} = \mathscr{R}_\mathrm{o}$ if $I(U;Y_1) \ge I(U;Y_2)$ for every $p^*(x) \in \mathscr{P}$ and every $p(u|x)$,

$$\max_{p(x)} \max_{p(u|x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big) = \max_{p^*(x)\in\mathscr{P}} \max_{p(u|x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big)$$

for $0 \le \lambda \le 1$

Effectively less noisy (Kim–Nachman–EG 2015)

Effectively less noisy BC

Y_1 is effectively less noisy than Y_2 if:

$I(U; Y_1) \geq I(U; Y_2)$ for every $p^*(x) \in \mathscr{P}$ and every $p(u|x)$ such that

$$p^*(x) = \arg\max_{p(x)} \Big(\max_{p(u|x)} \big(\lambda I(X; Y_1|U) + I(U; Y_2) \big) \Big), \ 0 \leq \lambda \leq 1$$

Effectively less noisy (Kim–Nachman–EG 2015)

Effectively less noisy BC

Y_1 is effectively less noisy than Y_2 if:

$$I(U;Y_1) \geq I(U;Y_2) \text{ for every } p^*(x) \in \mathscr{P} \text{ and every } p(u|x) \text{ such that}$$

$$p^*(x) = \arg\max_{p(x)} \Big(\max_{p(u|x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big)\Big),\ 0 \leq \lambda \leq 1$$

- In terms of concave envelope, Y_1 effectively less noisy than Y_2 if:

  $I(X;Y_1) - I(X;Y_2) = \mathfrak{C}[I(X;Y_1) - I(X;Y_2)]$ for every $p^*(x) \in \mathscr{P}$

Effectively less noisy (Kim–Nachman–EG 2015)

Effectively less noisy BC

Y_1 is effectively less noisy than Y_2 if:

$$I(U;Y_1) \geq I(U;Y_2) \text{ for every } p^*(x) \in \mathscr{P} \text{ and every } p(u|x) \text{ such that}$$

$$p^*(x) = \arg\max_{p(x)} \Big(\max_{p(u|x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big)\Big), \ 0 \leq \lambda \leq 1$$

- In terms of concave envelope, Y_1 effectively less noisy than Y_2 if:

  $I(X;Y_1) - I(X;Y_2) = \mathfrak{C}[I(X;Y_1) - I(X;Y_2)]$ for every $p^*(x) \in \mathscr{P}$

- $p^*(x)$ can also be recast in terms of concave envelope as:

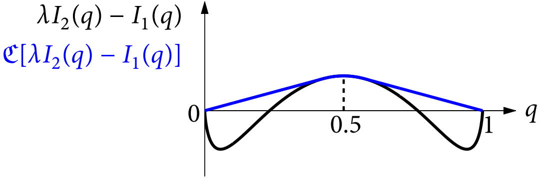
$$\begin{aligned} p^*(x) &= \arg\max_{p(x)} \Big(\max_{p(u|x)} \big(\lambda I(X;Y_2|U) + I(U;Y_1)\big)\Big) \\ &= \arg\max_{p(x)} \big(I(X;Y_1) + \mathfrak{C}[\lambda I(X;Y_2) - I(X;Y_1)]\big) \end{aligned}$$

Y_2 effectively less noisy than Y_1 for BEC/BSC

- For BEC/BSC, it can be shown that:

$$p_X^*(1) = q^* = \arg\max_q \left(I_2(q) + \mathfrak{C}[\lambda I_2(q) - I_1(q)] \right) = 0.5$$

  - $\lambda I_2(q) - I_1(q)$ is symmetric about $q = 0.5$
  - Both $I_2(q)$ and $\mathfrak{C}[\lambda I_2(q) - I_1(q)]$ are maximum at $q^* = 0.5$

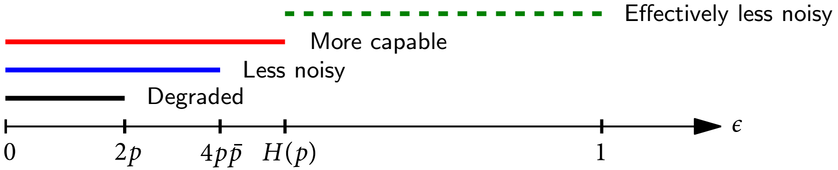


Y_2 effectively less noisy than Y_1 for BEC/BSC

- For BEC/BSC, it can be shown that:

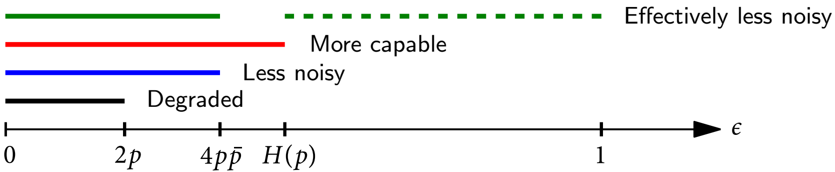
$$p_X^*(1) = q^* = \arg\max_q \big(I_2(q) + \mathfrak{C}[\lambda I_2(q) - I_1(q)]\big) = 0.5$$

- By inspection, $I_2(0.5) - I_1(0.5) = \mathfrak{C}[I_2(q) - I_1(q))]\big|_{q=0.5}$ for $H(p) < \epsilon \le 1$



BEC/BSC summary

- Superposition coding is always optimal
- Concave envelope method
- Degraded $\overset{\longrightarrow}{\not\longleftarrow}$ less noisy $\overset{\longrightarrow}{\not\longleftarrow}$ more capable
- Effectively less noisy
  - More general than essentially less noisy (Nair 2010)
  - Easier to compute in general



Outline

I. Optimality of superposition coding for BI-BC

BEC/BSC (Nair 2010)

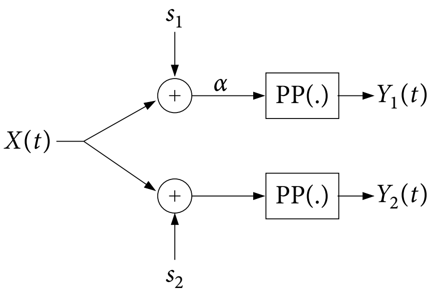
II. Poisson broadcast channel (Kim–Nachman–EG 2015):

- Superposition coding is almost always optimal

III. Evaluation of Marton region for BI-BC (Anantharam–Gohari–Nair 2013)

- Gap between Marton region and UV outer bound for Poisson BC

Poisson broadcast channel (P-BC)



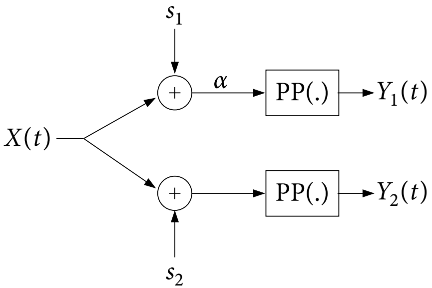
Continuous time Poisson BC

$Y_1(t)|\{X(t) = x(t)\}$ is $\mathrm{PP}(\alpha(x(t) + s_1))$,

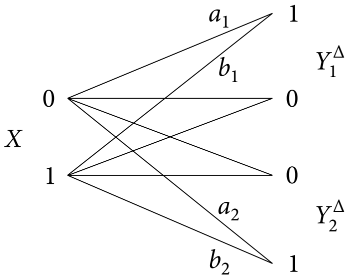
$Y_2(t)|\{X(t) = x(t)\}$ is $\mathrm{PP}(x(t) + s_2)$,

$X(t) \in [0, 1]$, $t, \alpha, s_1, s_2 \geq 0$, $s_1 \leq s_2$

Poisson broadcast channel (P-BC)



Continuous time Poisson BC

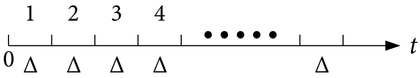


Binary P-BC (Wyner 1988a)

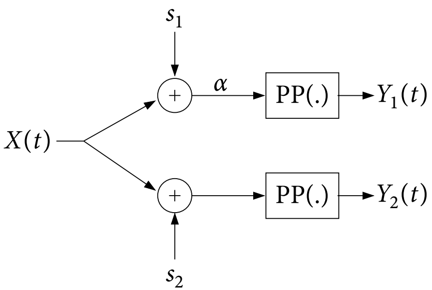
$Y_1(t)|\{X(t) = x(t)\}$ is $\mathrm{PP}(\alpha(x(t) + s_1))$,

$Y_2(t)|\{X(t) = x(t)\}$ is $\mathrm{PP}(x(t) + s_2)$,

$X(t) \in [0, 1]$, $t, \alpha, s_1, s_2 \geq 0$, $s_1 \leq s_2$



Poisson broadcast channel (P-BC)

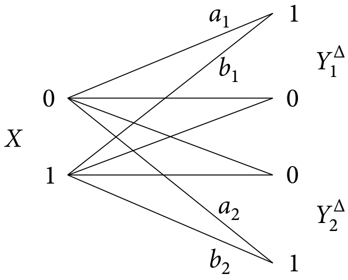


Continuous time Poisson BC

$Y_1(t)|\{X(t) = x(t)\}$ is $\mathrm{PP}(\alpha(x(t) + s_1))$,

$Y_2(t)|\{X(t) = x(t)\}$ is $\mathrm{PP}(x(t) + s_2)$,

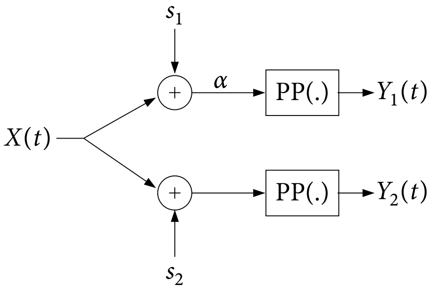
$X(t) \in [0, 1],\ t, \alpha, s_1, s_2 \geq 0,\ s_1 \leq s_2$



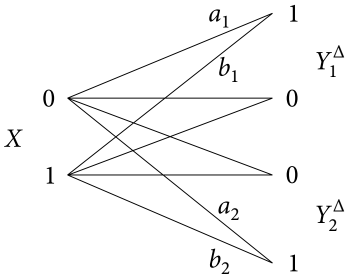
Binary P-BC (Wyner 1988a)

$$a_1 = \alpha s_1 \Delta + O(\Delta^2),$$
$$a_2 = s_2 \Delta + O(\Delta^2),$$
$$b_1 = \alpha(1 + s_1)\Delta + O(\Delta^2),$$
$$b_2 = (1 + s_2)\Delta + O(\Delta^2)$$

Poisson broadcast channel (P-BC)



Continuous time Poisson BC

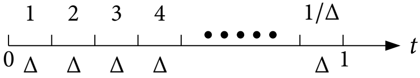


Binary P-BC (Wyner 1988a)

$Y_1(t)|\{X(t) = x(t)\}$ is $\mathrm{PP}(\alpha(x(t) + s_1))$,

$Y_2(t)|\{X(t) = x(t)\}$ is $\mathrm{PP}(x(t) + s_2)$,

$X(t) \in [0, 1],\ t, \alpha, s_1, s_2 \geq 0,\ s_1 \leq s_2$

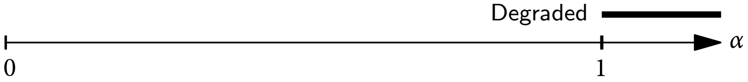


- Wyner (1988a,b), Lapidoth–Telatar–Urbanke (2003) argued that:
  Capacity of P-BC ≡ capacity of $1/\Delta$-extension binary P-BC as $\Delta \to 0$

Optimality of superposition coding for P-BC?

Optimality of superposition coding for P-BC?

- Lapidoth–Telatar–Urbanke (2003) showed:
  Y_2 degraded version of Y_1 if $\alpha \geq 1$ $(s_1 \leq s_2)$



Optimality of superposition coding for P-BC?

- Lapidoth–Telatar–Urbanke (2003) showed:
  Y_2 degraded version of Y_1 if $\alpha \geq 1$ $(s_1 \leq s_2)$
- Can define less noisy/more capable/eff. less noisy through binary P-BC:

Optimality of superposition coding for P-BC?

- Lapidoth–Telatar–Urbanke (2003) showed:
  Y_2 degraded version of Y_1 if $\alpha \geq 1$ $(s_1 \leq s_2)$

- Can define less noisy/more capable/eff. less noisy through binary P-BC:
  - Let
$$I_j(q) = \lim_{\Delta \to 0} \frac{1}{\Delta} I(X; Y_j^{\Delta}), \;\; X \sim \mathrm{Bern}(q), \;\; j = 1, 2$$

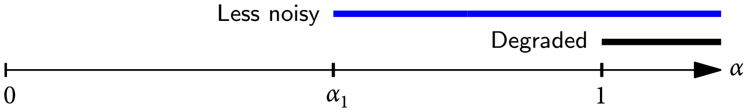
Optimality of superposition coding for P-BC?

- Lapidoth–Telatar–Urbanke (2003) showed:
  Y_2 degraded version of Y_1 if $\alpha \geq 1$ $(s_1 \leq s_2)$

- Can define less noisy/more capable/eff. less noisy through binary P-BC:
  - Let
  $$I_j(q) = \lim_{\Delta \to 0} \frac{1}{\Delta} I(X; Y_j^{\Delta}),\ X \sim \mathrm{Bern}(q),\ j = 1, 2$$
  - Less noisy: Second derivative of $I_1(q) - I_2(q) \leq 0$ for every $q \in [0, 1]$
  - More capable: $I_1(q) - I_2(q) \geq 0$ for every $q \in [0, 1]$
  - Effectively less noisy: $I_1(q^*) - I_2(q^*) = \mathfrak{C}[I_1(q) - I_2(q)]\big|_{q=q^*}$ for every $q^* \in \mathscr{P}$,
  $$q^* = \arg\max_q \big(I_2(q) + \mathfrak{C}[\lambda I_2(q) - I_1(q)]\big)$$

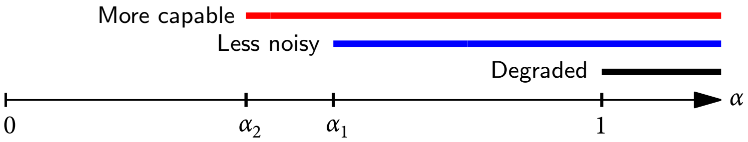
Y_1 less noisy than Y_2 for P-BC

- $s_1 \le s_2$
- $\alpha_1 = (1 + s_1)/(1 + s_2)$



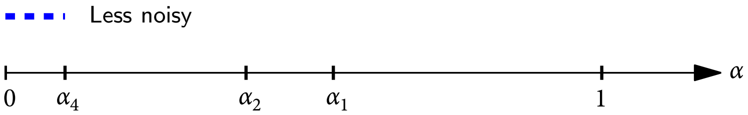
Y_1 more capable than Y_2 for P-BC

- $s_1 \leq s_2$
- $\alpha_1 = (1 + s_1)/(1 + s_2)$
- $\alpha_2 = (s_2 \log(1 + 1/s_2) - 1)/(s_1 \log(1 + 1/s_1) - 1)$



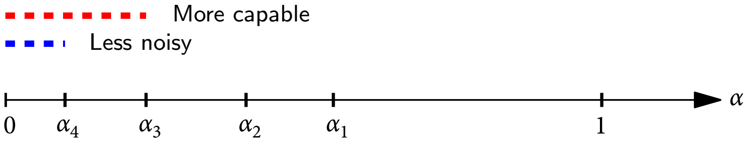
Y_2 less noisy than Y_1 for P-BC

- $s_1 \le s_2$
- $\alpha_1 = (1 + s_1)/(1 + s_2)$
- $\alpha_2 = (s_2 \log(1 + 1/s_2) - 1)/(s_1 \log(1 + 1/s_1) - 1)$
- $\alpha_4 = s_1/s_2$



Y_2 more capable than Y_1 for P-BC

- $s_1 \le s_2$
- $\alpha_1 = (1+s_1)/(1+s_2)$
- $\alpha_2 = (s_2 \log(1+1/s_2) - 1)/(s_1 \log(1+1/s_1) - 1)$
- $\alpha_4 = s_1/s_2$
- $\alpha_3 = ((1+s_2)\log(1+1/s_2) - 1)/(1+s_1)\log(1+1/s_1) - 1)$



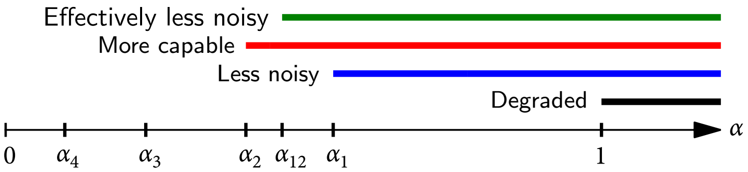
Y_1 effectively less noisy than Y_2 for P-BC

- $s_1 \le s_2$
- Y_1 effectively less noisy than Y_2 if $I_1(q^*) - I_2(q^*) = \mathfrak{C}[I_1(q) - I_2(q)]\big|_{q=q^*}$,

$$q^* \le q_2 = (1+s_2)^{1+s_2}/e s_2^{s_2} - s_2$$

- $\alpha_{12} = g(1+s_2, q_2+s_2)/g(1+s_1, q_2+s_1)$,

  where $g(x,y) = x\log(x/y) - x + y$

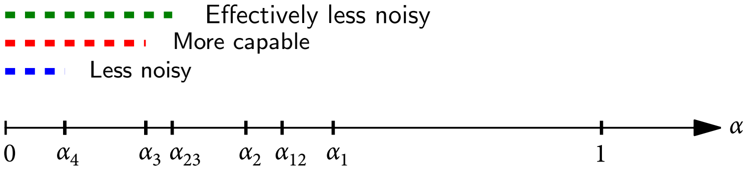


Y_2 effectively less noisy than Y_1 for P-BC

- $s_1 \leq s_2$
- Y_2 effectively less noisy than Y_1 if $I_2(q^*) - I_1(q^*) = \mathfrak{C}[I_2(q) - I_1(q)]\big|_{q=q^*}$,

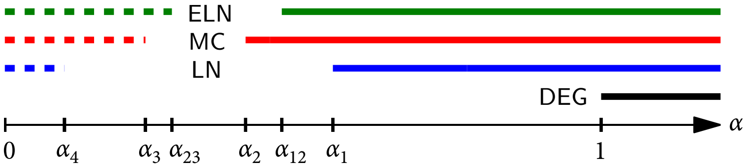
$$q^* \geq q_1 = (1 + s_1)^{1+s_1}/es_1^{s_1} - s_1$$

- $\alpha_{23} = (s_2 \log(1 + q_1/s_2) - q_1)/(s_1 \log(1 + q_1/s_1) - q_1)$



Summary

- Values of α for which superposition coding is optimal for P-BC ($s_1 \le s_2$)



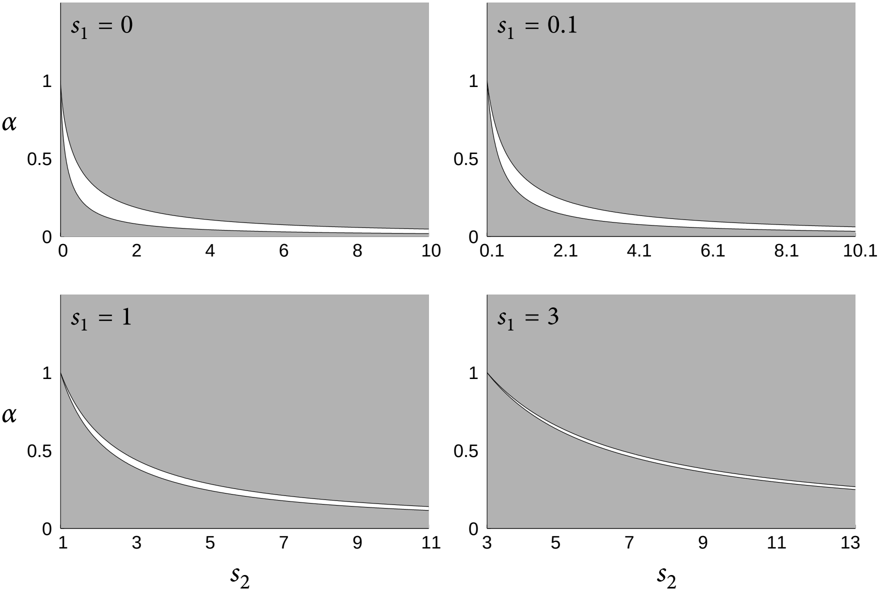
DEG: Degraded

LN: Less noisy

MC: More capable

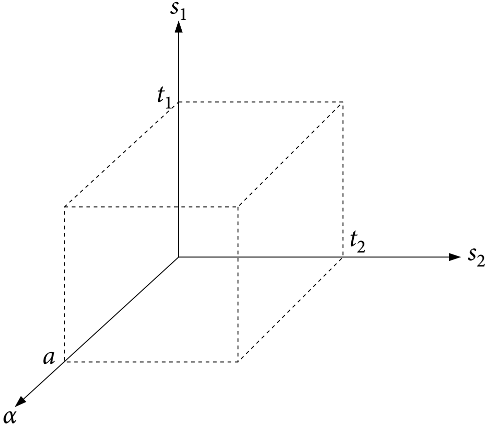
ELN: Effectively less noisy

Summary



$s_1 = 0$
$s_1 = 0.1$
$s_1 = 1$
$s_1 = 3$
α
α
1
0.5
0
0
2
4
6
8
10
0.1
2.1
4.1
6.1
8.1
10.1
1
3
5
7
9
11
3
5
7
9
11
13
s_2
s_2

Superposition coding is almost always optimal for P-BC

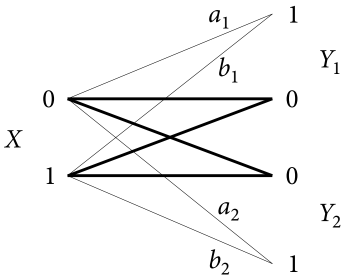


Fraction of volume s.t. superposition coding optimal $\to 1$ as $t_1, t_2, a \to \infty$

Why superposition coding almost always optimal for P-BC

- Transition probabilities for BP-BC channel have consistent magnitudes

$$a_1 = \alpha s_1 \Delta + O(\Delta^2), \qquad a_2 = s_2 \Delta + O(\Delta^2),$$
$$b_1 = \alpha(1 + s_1)\Delta + O(\Delta^2), \quad b_2 = (1 + s_2)\Delta + O(\Delta^2)$$

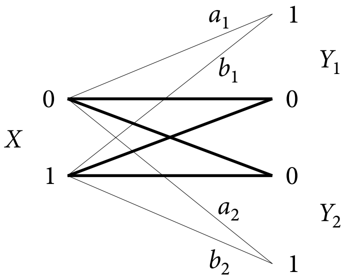


Why superposition coding almost always optimal for P-BC

- Transition probabilities for BP-BC channel have consistent magnitudes

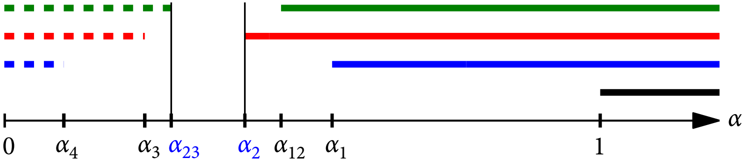
$$a_1 = \alpha s_1 \Delta + O(\Delta^2), \qquad a_2 = s_2 \Delta + O(\Delta^2),$$
$$b_1 = \alpha(1 + s_1)\Delta + O(\Delta^2), \quad b_2 = (1 + s_2)\Delta + O(\Delta^2)$$

- This is not the case for BI-BC in general



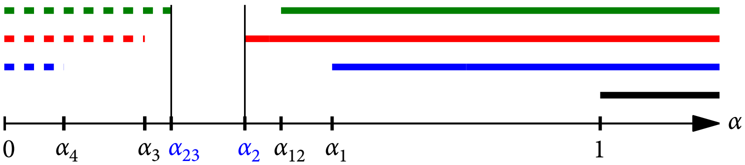
Capacity region of P-BC for remaining set of parameters

- What is the capacity region for $\alpha \in (\alpha_{23}, \alpha_2)$, $(s_1 \le s_2)$?



Capacity region of P-BC for remaining set of parameters

- What is the capacity region for $\alpha \in (\alpha_{23}, \alpha_2)$, $(s_1 \leq s_2)$?
- We compare Marton's inner bound and UV outer bound



Outline

I. Optimality of superposition coding for BI-BC

BEC/BSC (Nair 2010)

II. Poisson broadcast channel (Kim–Nachman–EG 2015):

- Superposition coding is almost always optimal

III. Evaluation of Marton region for BI-BC (Anantharam–Gohari–Nair 2013)

- Gap between Marton region and UV outer bound for Poisson BC

General inner bounds on capacity region of BC

Cover (1975), van der Meulen (1975) $\mathscr{R}_{UVW}$

A rate pair (R_1, R_2) is achievable for BC if

$$\begin{aligned} R_1 &< I(W;Y_1) + I(V;Y_1|W), \\ R_2 &< I(W;Y_2) + I(U;Y_2|W), \\ R_1 + R_2 &< \min\{I(W;Y_1), I(W;Y_2)\} + I(V;Y_1|W) + I(U;Y_2|W) \end{aligned}$$

for some pmf $p(w)p(u|w)p(v|w)$ and function $x(w,u,v)$

- Split message $M_j = (M_{j0}, M_{jj}),\ j = 1, 2$
- W carries (M_{10}, M_{20}), recovered by Y_1 and Y_2
- V carries M_{11}, recovered by Y_1; U carries M_{22}, recovered by Y_2
- Includes superposition coding inner bound ($U \leftarrow X$, $W \leftarrow U$, $V = \emptyset$)

General inner bounds on capacity region of BC

Marton (1979) $\mathscr{R}_{\mathrm{M}}$ inner bound

A rate pair (R_1, R_2) is achievable for BC if

$$\begin{aligned} R_1 &< I(W;Y_1) + I(V;Y_1|W), \\ R_2 &< I(W;Y_2) + I(U;Y_2|W), \\ R_1 + R_2 &< \min\{I(W;Y_1), I(W;Y_2)\} + I(V;Y_1|W) + I(U;Y_2|W) - I(U;V|W) \end{aligned}$$

for some pmf $p(w,u,v)$ and function $x(w,u,v)$

- U and V given W correlated (they carry independent messages!)
- Correlation results in rate penalty
- Tight for all classes of BC with known capacity regions

Marton inner bound for BI-BC ≡ randomized time sharing

Marton region for BI-BC $\mathscr{R}_{\mathrm{RTD}}$ (Anantharam–Gohari–Nair 2013)

A rate pair (R_1, R_2) is achievable for BC if

$$R_1 < I(W;Y_1) + \sum_{j=1}^{k} \alpha_j I(X;Y_1|W=j),$$

$$R_2 < I(W;Y_2) + \sum_{j=k+1}^{5} \alpha_j I(X;Y_2|W=j),$$

$$R_1 + R_2 < \min\{I(W;Y_1), I(W;Y_2)\} + \sum_{j=1}^{k} \alpha_j I(X;Y_1|W=j) + \sum_{j=k+1}^{5} \alpha_j I(X;Y_2|W=j)$$

for some $p_W(j) = \alpha_j,\ j \in [1:5]$, and $p(x|w)$

- Extends result on skewed binary BC (Hajek–Pursley 1979)

Marton inner bound for BI-BC ≡ randomized time sharing

Marton region for BI-BC $\mathscr{R}_{\mathrm{RTD}}$ (Anantharam–Gohari–Nair 2013)

A rate pair (R_1, R_2) is achievable for BC if

$$
\begin{aligned}
R_1 &< I(W;Y_1) + \sum_{j=1}^{k} \alpha_j I(X;Y_1|W=j),\\
R_2 &< I(W;Y_2) + \sum_{j=k+1}^{5} \alpha_j I(X;Y_2|W=j),\\
R_1 + R_2 &< \min\{I(W;Y_1), I(W;Y_2)\} + \sum_{j=1}^{k} \alpha_j I(X;Y_1|W=j) + \sum_{j=k+1}^{5} \alpha_j I(X;Y_2|W=j)
\end{aligned}
$$

for some $p_W(j) = \alpha_j$, $j \in [1:5]$, and $p(x|w)$

- Extends result on skewed binary BC (Hajek–Pursley 1979)
- For BI-BC: $\mathscr{R}_{\mathrm{M}} = \mathscr{R}_{UVW} = \mathscr{R}_{\mathrm{RTD}}$

  $\Rightarrow$ U and V independent given W suffices

Randomized time division (Hajek–Pursley 1979)

- Let $W \in \{1, 2\}$ and fix $p_W(1) = \alpha$, $p(x|w)$, and set

  $U = X$ if $W = 1$, $U = \epsilon$, otherwise; $V = X$ if $W = 2$, $V = \epsilon$, otherwise

Randomized time division (Hajek–Pursley 1979)

- Let $W \in \{1, 2\}$ and fix $p_W(1) = \alpha$, $p(x|w)$, and set

  $U = X$ if $W = 1$, $U = \epsilon$, otherwise; $V = X$ if $W = 2$, $V = \epsilon$, otherwise

- Generate $2^{n(R_{10}+R_{20})}$ $w^n(m_{10}, m_{20})$ sequences

$$W^n \;\Big|\; 1 \quad 2 \quad 2 \quad 1 \quad 1 \quad 1 \quad 2 \quad 1 \quad 2 \quad 2 \quad 1 \quad 2 \quad 1 \quad 1$$

Randomized time division (Hajek–Pursley 1979)

- Let $W \in \{1,2\}$ and fix $p_W(1) = \alpha$, $p(x|w)$, and set

  $U = X$ if $W = 1$, $U = \epsilon$, otherwise; $V = X$ if $W = 2$, $V = \epsilon$, otherwise

- Generate $2^{n(R_{10}+R_{20})}$ $w^n(m_{10}, m_{20})$ sequences
- For every (m_{10}, m_{20}), generate $2^{nR_{11}}$ $u^n(m_{10}, m_{20}, m_{11})$ sequences

W^n	1	2	2	1	1	1	2	1	2	2	1	2	1	1
U^n	0	ϵ	ϵ	0	0	0	ϵ	1	ϵ	ϵ	0	ϵ	0	1

Randomized time division (Hajek–Pursley 1979)

- Let $W \in \{1, 2\}$ and fix $p_W(1) = \alpha$, $p(x|w)$, and set

  $U = X$ if $W = 1$, $U = \epsilon$, otherwise; $V = X$ if $W = 2$, $V = \epsilon$, otherwise
- Generate $2^{n(R_{10}+R_{20})}$ $w^n(m_{10}, m_{20})$ sequences
- For every (m_{10}, m_{20}), generate $2^{nR_{11}}$ $u^n(m_{10}, m_{20}, m_{11})$ sequences
- For every (m_{10}, m_{20}), generate $2^{nR_{22}}$ $v^n(m_{10}, m_{20}, m_{22})$ sequences

W^n	1	2	2	1	1	1	2	1	2	2	1	2	1	1
U^n	0	ϵ	ϵ	0	0	0	ϵ	1	ϵ	ϵ	0	ϵ	0	1
V^n	ϵ	0	1	ϵ	ϵ	ϵ	1	ϵ	0	1	ϵ	1	ϵ	ϵ

Randomized time division (Hajek–Pursley 1979)

- Let $W \in \{1, 2\}$ and fix $p_W(1) = \alpha$, $p(x|w)$, and set

  $U = X$ if $W = 1$, $U = \epsilon$, otherwise; $V = X$ if $W = 2$, $V = \epsilon$, otherwise

- Generate $2^{n(R_{10}+R_{20})}$ $w^n(m_{10}, m_{20})$ sequences
- For every (m_{10}, m_{20}), generate $2^{nR_{11}}$ $u^n(m_{10}, m_{20}, m_{11})$ sequences
- For every (m_{10}, m_{20}), generate $2^{nR_{22}}$ $v^n(m_{10}, m_{20}, m_{22})$ sequences
- Set $X_i(m_1, m_2) = u_i(m_{10}, m_{20}, m_{11})$ if $u_i \neq \epsilon$, and

  $X_i(m_1, m_2) = v_i(m_{10}, m_{20}, m_{22})$ if $v_i \neq \epsilon$

W^n	1	2	2	1	1	1	2	1	2	2	1	2	1	1
U^n	0	ϵ	ϵ	0	0	0	ϵ	1	ϵ	ϵ	0	ϵ	0	1
V^n	ϵ	0	1	ϵ	ϵ	ϵ	1	ϵ	0	1	ϵ	1	ϵ	ϵ
X^n	0	0	1	0	0	0	1	1	0	1	0	1	0	1

Proving $\mathscr{R}_{\mathrm{M}} = \mathscr{R}_{\mathrm{RTD}}$ for BI-BC

- Cardinality bounds via perturbation (Gohari–Anantharam 2009):
  - $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$, $|\mathcal{W}| \le |\mathcal{X}| + 4$

Proving $\mathscr{R}_{\mathrm{M}} = \mathscr{R}_{\mathrm{RTD}}$ for BI-BC

- Cardinality bounds via perturbation (Gohari–Anantharam 2009):
  - $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$, $|\mathcal{W}| \le |\mathcal{X}| + 4$
- Sum-rate proof (Geng–Jog–Nair–Wang 2013a):
  - For BI-BC sum rate: $|\mathcal{W}| \le |\mathcal{X}| = 2$ $(|\mathcal{U}|, |\mathcal{V}| \le 2)$
  - Key inequality for BI-BC: (only few $x(u, v, w)$ need to be considered)

$$I(V;Y_1) + I(U;Y_2) - I(U;V) \le \max\{I(X;Y_1), I(X;Y_2)\}$$

Proving $\mathscr{R}_{\mathrm{M}} = \mathscr{R}_{\mathrm{RTD}}$ for BI-BC

- Cardinality bounds via perturbation (Gohari–Anantharam 2009):
  - $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$, $|\mathcal{W}| \le |\mathcal{X}| + 4$
- Sum-rate proof (Geng–Jog–Nair–Wang 2013a):
  - For BI-BC sum rate: $|\mathcal{W}| \le |\mathcal{X}| = 2$ $(|\mathcal{U}|, |\mathcal{V}| \le 2)$
  - Key inequality for BI-BC: (only few $x(u, v, w)$ need to be considered)

$$I(V;Y_1) + I(U;Y_2) - I(U;V) \le \max\{I(X;Y_1), I(X;Y_2)\}$$

- General case (Anantharam–Gohari–Nair 2013): $|\mathcal{U}| + |\mathcal{V}| \le |\mathcal{X}| + 1$
  - Concave envelope
  - Perturbation
  - A min-max theorem for rate regions

Perturbation method

- Used to bound cardinality of U, V
- Carathéodory method (Ahlswede–Körner 1975) does not work

Perturbation method

- Gohari–Anantharam (2009) used perturbation to show: $|\mathcal{U}|, |\mathcal{V}| \leq |\mathcal{X}|$

Perturbation method

- Gohari–Anantharam (2009) used perturbation to show: $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$
- Consider Marton sum rate bound $S(p(u,v|x)) = I(U;Y_1) + I(V;Y_2) - I(U;V)$
  - Want to show: $\max S(p)$ can be attained with $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$

Perturbation method

- Gohari–Anantharam (2009) used perturbation to show: $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$
- Consider Marton sum rate bound $S(p(u, v|x)) = I(U; Y_1) + I(V; Y_2) - I(U; V)$
  - Want to show: $\max S(p)$ can be attained with $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$
  - Let $p_0(u, v|x) = \arg\max S(p)$ and define perturbed pmf

$$p_\epsilon(u, v|x) = p_0(u, v|x)(1 + \epsilon L(u)), \text{ where } \min_u(1 + \epsilon L(u)) \ge 0 \text{ and}$$

$\sum_u p_\epsilon(u, v|x)L(u) = 0$ for all $x \implies$ non-zero $L(u)$ exists if $|\mathcal{U}| > |\mathcal{X}|$

Perturbation method

- Gohari–Anantharam (2009) used perturbation to show: $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$
- Consider Marton sum rate bound $S(p(u,v|x)) = I(U;Y_1) + I(V;Y_2) - I(U;V)$
  - Want to show: $\max S(p)$ can be attained with $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$
  - Let $p_0(u,v|x) = \arg\max S(p)$ and define perturbed pmf

    $p_\epsilon(u,v|x) = p_0(u,v|x)(1+\epsilon L(u))$, where $\min_u (1+\epsilon L(u)) \ge 0$ and

    $\sum_u p_\epsilon(u,v|x) L(u) = 0$ for all $x \implies$ non-zero $L(u)$ exists if $|\mathcal{U}| > |\mathcal{X}|$
  - $p_0(u,v|x)$ maximizer $\implies \frac{\partial S(p_\epsilon)}{\partial \epsilon}\Big|_{p_0} = 0, \; \frac{\partial^2 S(p_\epsilon)}{\partial \epsilon^2}\Big|_{p_0} \le 0$

    After some calculation, we find that $S(p_\epsilon)$ must be a constant!

Perturbation method

- Gohari–Anantharam (2009) used perturbation to show: $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$
- Consider Marton sum rate bound $S(p(u,v|x)) = I(U;Y_1) + I(V;Y_2) - I(U;V)$
  - Want to show: $\max S(p)$ can be attained with $|\mathcal{U}|, |\mathcal{V}| \le |\mathcal{X}|$
  - Let $p_0(u,v|x) = \arg\max S(p)$ and define perturbed pmf

    $$p_\epsilon(u,v|x) = p_0(u,v|x)(1+\epsilon L(u)), \text{ where } \min_u (1+\epsilon L(u)) \ge 0 \text{ and}$$

    $\sum_u p_\epsilon(u,v|x)L(u) = 0$ for all $x \implies$ non-zero $L(u)$ exists if $|\mathcal{U}| > |\mathcal{X}|$
  - $p_0(u,v|x)$ maximizer $\implies \left.\frac{\partial S(p_\epsilon)}{\partial \epsilon}\right|_{p_0} = 0, \left.\frac{\partial^2 S(p_\epsilon)}{\partial \epsilon^2}\right|_{p_0} \le 0$

    After some calculation, we find that $S(p_\epsilon)$ must be a constant!
  - Choose ϵ large enough: $p_\epsilon(u,v|x) = p_0(u,v|x)(1+\epsilon L(u)) = 0$ for some u
  - Repeat until $|\mathcal{U}| \le |\mathcal{X}|$

UV outer bound on capacity region of BC

UV outer bound (Nair–EG 2007)

If (R_1, R_2) is achievable, then

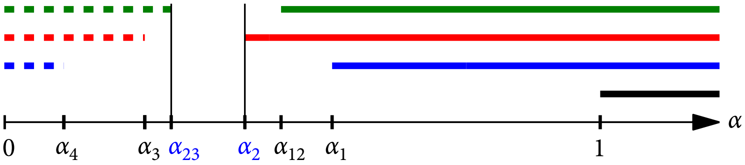
$$\begin{aligned} R_1 &\le I(V; Y_1), \\ R_2 &\le I(U; Y_2), \\ R_1 + R_2 &\le I(V; Y_1) + I(U; Y_2 | V), \\ R_1 + R_2 &\le I(V; Y_1 | U) + I(U; Y_2) \end{aligned}$$

for some $p(u, v)$ and function $x(u, v)$

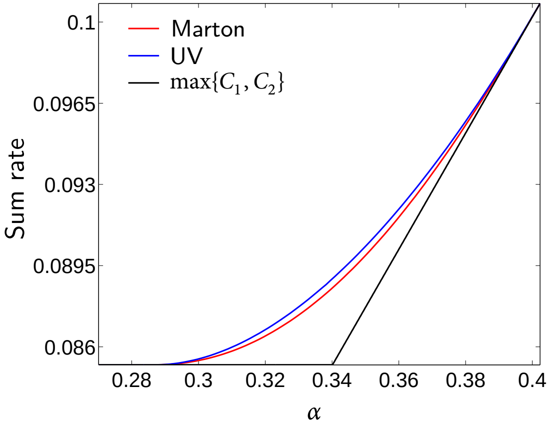
- Tight for almost all classes of BCs with known capacity regions
- Not tight in general (Geng–Gohari–Nair–Yu 2014)

Gap between Marton and UV for P-BC

- Recall, capacity region for $\alpha \in (\alpha_{23}, \alpha_2)$, $(s_1 \le s_2)$ is not known

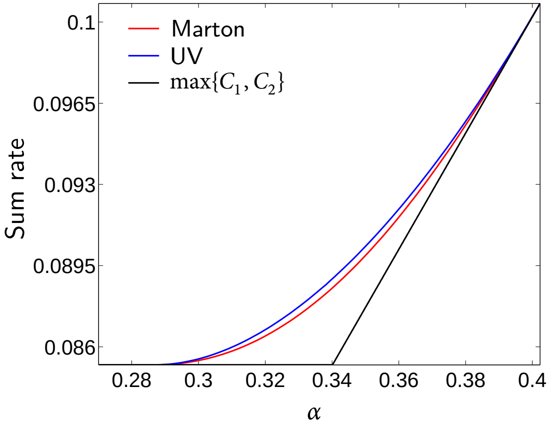


Gap between Marton and UV for P-BC



$s_1 = 0.1$ and $s_2 = 1 \implies \alpha_{23} = 0.27,\ \alpha_2 = 0.4$

Gap between Marton and UV for P-BC



$$s_1 = 0.1 \text{ and } s_2 = 1 \implies \alpha_{23} = 0.27,\ \alpha_2 = 0.4$$

- Superposition coding is optimal for $\alpha \in (0.27, 0.286)$

  No analytical expression for cutoff α

Beyond effectively less noisy

- Recall Y_1 effectively less noisy than Y_2 if:

$$I(U;Y_1) \geq I(U;Y_2) \text{ for every } p^*(x) \in \mathscr{P} \text{ and every } p(u|x), \text{ such that}$$

$$\max_{p(x)} \max_{p(u|x)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right) = \max_{p^*(x) \in \mathscr{P}} \max_{p(u|x)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right)$$

for every $0 \leq \lambda \leq 1$

Beyond effectively less noisy

- Recall Y_1 effectively less noisy than Y_2 if:

$$I(U;Y_1) \geq I(U;Y_2) \text{ for every } p^*(x) \in \mathscr{P} \text{ and every } p(u|x), \text{ such that}$$

$$\max_{p(x)} \max_{p(u|x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big) = \max_{p^*(x)\in\mathscr{P}} \max_{p(u|x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big)$$

for every $0 \leq \lambda \leq 1$

- This definition can be tightened to:

$$I(U;Y_2) \leq I(U;Y_1) \text{ for every } p^*(u,x) \in \mathscr{P}(U,X), \text{ such that}$$

$$\max_{p(u,x)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big) = \max_{p^*(u,x)\in\mathscr{P}(U,X)} \big(\lambda I(X;Y_1|U) + I(U;Y_2)\big)$$

for every $0 \leq \lambda \leq 1$

Beyond effectively less noisy

- Recall Y_1 effectively less noisy than Y_2 if:

$$I(U;Y_1) \geq I(U;Y_2) \text{ for every } p^*(x) \in \mathscr{P} \text{ and every } p(u|x), \text{ such that}$$

$$\max_{p(x)} \max_{p(u|x)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right) = \max_{p^*(x) \in \mathscr{P}} \max_{p(u|x)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right)$$

for every $0 \leq \lambda \leq 1$

- This definition can be tightened to:

$$I(U;Y_2) \leq I(U;Y_1) \text{ for every } p^*(u,x) \in \mathscr{P}(U,X), \text{ such that}$$

$$\max_{p(u,x)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right) = \max_{p^*(u,x) \in \mathscr{P}(U,X)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right)$$

for every $0 \leq \lambda \leq 1$

- This is the condition satisfied for P-BC with $\alpha \in (0.27, 0.286)$

Beyond effectively less noisy

- Recall Y_1 effectively less noisy than Y_2 if:

$$I(U;Y_1) \geq I(U;Y_2) \text{ for every } p^*(x) \in \mathscr{P} \text{ and every } p(u|x), \text{ such that}$$

$$\max_{p(x)} \max_{p(u|x)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right) = \max_{p^*(x)\in\mathscr{P}} \max_{p(u|x)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right)$$

for every $0 \leq \lambda \leq 1$

- This definition can be tightened to:

$$I(U;Y_2) \leq I(U;Y_1) \text{ for every } p^*(u,x) \in \mathscr{P}(U,X), \text{ such that}$$

$$\max_{p(u,x)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right) = \max_{p^*(u,x)\in\mathscr{P}(U,X)} \left(\lambda I(X;Y_1|U) + I(U;Y_2)\right)$$

for every $0 \leq \lambda \leq 1$

- This is the condition satisfied for P-BC with $\alpha \in (0.27, 0.286)$
- We don't have explicit analytical expression for this range in general

Summary

- Highlighted techniques for computing capacity and bounds for BI-BC
  - Concave envelope (less noisy, effectively less noisy, . . .)
  - Marton = Cover–van der Meulen = randomized time division for BI-BC

Summary

- Highlighted techniques for computing capacity and bounds for BI-BC
  - Concave envelope (less noisy, effectively less noisy, . . .)
  - Marton = Cover–van der Meulen = randomized time division for BI-BC
- Takeaway: *Only extremal distributions matter*

Summary

- Highlighted techniques for computing capacity and bounds for BI-BC
  - Concave envelope (less noisy, effectively less noisy, ...)
  - Marton = Cover–van der Meulen = randomized time division for BI-BC
- Takeaway: *Only extremal distributions matter*
- Presented new results on the Poisson BC:
  - Superposition coding inner bound is almost always tight
  - Gap between Marton and UV bounds

Related work

- Other applications of concave envelope:
  - MIMO BC with private and common messages (Geng–Nair 2014)

    Direct proof of DPC optimality (Weingarten–Steinberg–Shamai 2006)

    Extends result to setup with common message

Related work

- Other applications of concave envelope:
  - MIMO BC with private and common messages (Geng–Nair 2014)

    Direct proof of DPC optimality (Weingarten–Steinberg–Shamai 2006)

    Extends result to setup with common message
  - Product of reversely semideterministic BC (Geng–Gohari–Nair–Yu 2014)

    Shows that UV outer bound is not tight

Related work

- Other applications of concave envelope:
  - MIMO BC with private and common messages (Geng–Nair 2014)

    Direct proof of DPC optimality (Weingarten–Steinberg–Shamai 2006)

    Extends result to setup with common message
  - Product of reversely semideterministic BC (Geng–Gohari–Nair–Yu 2014)

    Shows that UV outer bound is not tight
- Binary input symmetric output BC (Geng–Nair–Shamai–Wang 2013b)
  - Either superposition coding is optimal or gap between Marton and UV

Final remarks

- Network information theory aims to:
  - Find computable (analytical and numerical) expressions for capacity
  - Coding schemes that achieve the capacity

Final remarks

- Network information theory aims to:
  - Find computable (analytical and numerical) expressions for capacity
  - Coding schemes that achieve the capacity
- We know the capacity for very few channels
- We have inner and outer bounds for other channels

Final remarks

- Network information theory aims to:
  - Find computable (analytical and numerical) expressions for capacity
  - Coding schemes that achieve the capacity
- We know the capacity for very few channels
- We have inner and outer bounds for other channels
- Computational tools play a key role in the development of this theory
  - Find explicit capacity characterizations for important channels
  - Help evaluate optimality (or lack thereof) of inner and outer bounds
  - Develop deeper understanding of the bounds

Final remarks

- Network information theory aims to:
  - Find computable (analytical and numerical) expressions for capacity
  - Coding schemes that achieve the capacity
- We know the capacity for very few channels
- We have inner and outer bounds for other channels
- Computational tools play a key role in the development of this theory
  - Find explicit capacity characterizations for important channels
  - Help evaluate optimality (or lack thereof) of inner and outer bounds
  - Develop deeper understanding of the bounds
- *Need more tools to evaluate capacity and its bounds*

Acknowledgments

- Hyeji Kim, Ben Nachman

Acknowledgments

- Hyeji Kim, Ben Nachman
- Chandra Nair
- Young-Han Kim

תודה

References

Ahlswede, R. and Körner, J. (1975). Source coding with side information and a converse for degraded broadcast channels. *IEEE Trans. Inf. Theory*, 21(6), 629–637.

Anantharam, V., Gohari, A., and Nair, C. (2013). Improved cardinality bounds on the auxiliary random variables in Marton's inner bound. In *Proc. IEEE Int. Symp. Inf. Theory*, pp. 1272–1276.

Bergmans, P. P. (1973). Random coding theorem for broadcast channels with degraded components. *IEEE Trans. Inf. Theory*, 19(2), 197–207.

Cover, T. M. (1972). Broadcast channels. *IEEE Trans. Inf. Theory*, 18(1), 2–14.

Cover, T. M. (1975). An achievable rate region for the broadcast channel. *IEEE Trans. Inf. Theory*, 21(4), 399–404.

El Gamal, A. (1979). The capacity of a class of broadcast channels. *IEEE Trans. Inf. Theory*, 25(2), 166–169.

Gallager, R. G. (1974). Capacity and coding for degraded broadcast channels. *Probl. Inf. Transm.*, 10(3), 3–14.

Geng, Y., Gohari, A., Nair, C., and Yu, Y. (2014). On marton's inner bound and its optimality for classes of product broadcast channels. *IEEE Trans. Inf. Theory*, 60(1), 22–41.

Geng, Y., Jog, V., Nair, C., and Wang, Z. (2013a). An information inequality and evaluation of marton's inner bound for binary input broadcast channels. *IEEE Trans. Inf. Theory*, 59(7), 4095–4105.

Geng, Y. and Nair, C. (2014). The capacity region of the two-receiver gaussian vector broadcast channel with private and common messages. *IEEE Trans. Inf. Theory*, 60(4), 2087–2104.

Geng, Y., Nair, C., Shamai, S., and Wang, Z. (2013b). On broadcast channels with binary inputs and symmetric outputs. *IEEE Trans. Inf. Theory*, 59(11), 6980–6989.

Gohari, A. A. and Anantharam, V. (2009). Evaluation of Marton's inner bound for the general broadcast channel. In *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, pp. 2462–2466.

Hajek, B. E. and Pursley, M. B. (1979). Evaluation of an achievable rate region for the broadcast channel. *IEEE Trans. Inf. Theory*, 25(1), 36–46.

Kim, H., Nachman, B., and El Gamal, A. (2015). Superposition coding is almost always optimal for the poisson broadcast channel. In *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong.

References (cont.)

Körner, J. and Marton, K. (1977). Comparison of two noisy channels. In I. Csiszár and P. Elias (eds.) *Topics in Information Theory (Colloquia Mathematica Societatis János Bolyai, Keszthely, Hungary, 1975)*, pp. 411–423. North-Holland, Amsterdam.

Lapidoth, A., Telatar, I., and Urbanke, R. (2003). On wide-band broadcast channels. *IEEE Trans. Inf. Theory*, 49(12), 3250–3258.

Marton, K. (1979). A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inf. Theory*, 25(3), 306–311.

Nair, C. (2010). Capacity regions of two new classes of two-receiver broadcast channels. *IEEE Trans. Inf. Theory*, 56(9), 4207–4214.

Nair, C. (2013). Upper concave envelopes and auxiliary random variables. *Int. J. Adv. Eng. Sci. Appl. Math*, 5(1), 12–20.

Nair, C. and El Gamal, A. (2007). An outer bound to the capacity region of the broadcast channel. *IEEE Trans. Inf. Theory*, 53(1), 350–355.

Nair, C. and Wang, Z. V. (2008). On the inner and outer bounds for 2-receiver discrete memoryless broadcast channels. In *Proc. UCSD Inf. Theory Appl. Workshop*, La Jolla, CA, pp. 226–229.

van der Meulen, E. C. (1975). Random coding theorems for the general discrete memoryless broadcast channel. *IEEE Trans. Inf. Theory*, 21(2), 180–190.

van Dijk, M. (1997). On a special class of broadcast channels with confidential messages. *IEEE Trans. Inf. Theory*, 43(2), 712–714.

Weingarten, H., Steinberg, Y., and Shamai, S. (2006). The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *IEEE Trans. Inf. Theory*, 52(9), 3936–3964.

Wyner, A. (1988a). Capacity and error exponent for the direct detection photon channel. i. *IEEE Trans. Inf. Theory*, 34(6), 1449–1461.

Wyner, A. (1988b). Capacity and error exponent for the direct detection photon channel. ii. *Information Theory, IEEE Transactions on*, 34(6), 1462–1471.

References (cont.)

Wyner, A. D. (1973). A theorem on the entropy of certain binary sequences and applications—II. *IEEE Trans. Inf. Theory*, 19(6), 772–777.

Wyner, A. D. and Ziv, J. (1973). A theorem on the entropy of certain binary sequences and appli-cations—I. *IEEE Trans. Inf. Theory*, 19(6), 769–772.