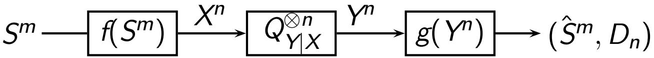
A Lower Bound on the Expected Distortion of Joint Source-Channel Coding

Or Ordentlich (Hebrew University of Jerusalem)
Joint work with Yuval Kochman (HUJI) and Yury Polyanskiy (MIT)

ISIT,
Paris,
July 9, 2019

Joint Source-Channel Coding

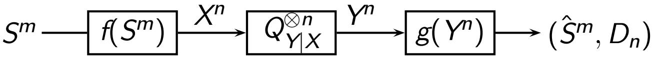
$$n = \rho m$$



- $S^m \sim P_S^{\otimes m}$ is memoryless source
- $Q_{Y|X}^{\otimes n}(y^n|x^n) = \prod_{i=1}^{n} Q_{Y|X}(y_i|x_i)$ is a memoryless channel
- $d(s^m, \hat{s}^m) = \frac{1}{m} \sum_{i=1}^{m} d(s_i, \hat{s}_i)$ is a decomposable distortion measure

Joint Source-Channel Coding

$$n = \rho m$$



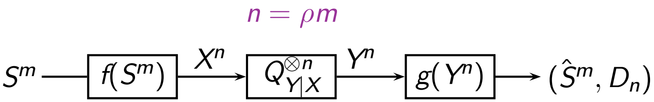
- $S^m \sim P_S^{\otimes m}$ is memoryless source
- $Q_{Y|X}^{\otimes n}(y^n|x^n) = \prod_{i=1}^n Q_{Y|X}(y_i|x_i)$ is a memoryless channel
- $d(s^m, \hat{s}^m) = \frac{1}{m}\sum_{i=1}^m d(s_i, \hat{s}_i)$ is a decomposable distortion measure

Define

$$D_n^* = D^*(n, \rho, P_S, Q_{Y|X}) \triangleq \min_{\substack{f: \mathcal{S}^m \mapsto \mathcal{X}^n \\ g: \mathcal{Y}^n \mapsto \hat{\mathcal{S}}^m}} \frac{1}{m} \mathbb{E} d(S^m, g(Y^n)),$$

where $m = \frac{n}{\rho}$ and Y^n is the output of $Q_{Y|X}^{\otimes n}$ applied on $X^n = f(S^m)$.

Joint Source-Channel Coding



Source-Channel Separation Theorem

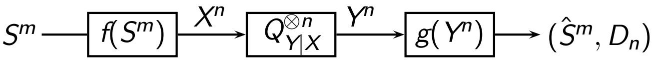
$$\lim_{n \to \infty} D_n^* = D(\rho C_Q)$$

where

- $D(R)$ is the distortion-rate function of P_S
- C_Q is the capacity of the channel $Q_{Y|X}$

Joint Source-Channel Coding

$$n = \rho m$$



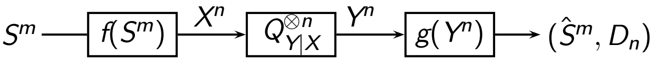
Source-Channel Separation Theorem

$$\lim_{n \to \infty} D_n^* = D(\rho C_Q)$$

Define $\Delta_n = \Delta(n, \rho, P_S, Q_{Y|X}) = D_n^* - D(\rho C_Q)$

Joint Source-Channel Coding

$n = \rho m$



Source-Channel Separation Theorem

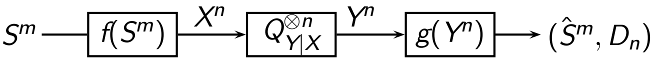
$$\lim_{n \to \infty} D_n^* = D(\rho C_Q)$$

Define $\Delta_n = \Delta(n, \rho, P_S, Q_{Y|X}) = D_n^* - D(\rho C_Q)$

How does Δ_n scale with n?

Joint Source-Channel Coding

$$n = \rho m$$



Source-Channel Separation Theorem

$$\lim_{n\to\infty} D_n^* = D(\rho C_Q)$$

Define $\Delta_n = \Delta(n, \rho, P_S, Q_{Y|X}) = D_n^* - D(\rho C_Q)$

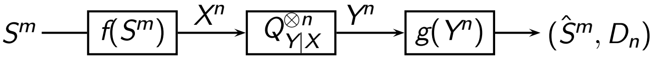
How does Δ_n scale with n?

A lot is known about FBL JSCC under **excess distortion** criterion

Practically nothing is known under **mean distortion** criterion

Joint Source-Channel Coding

$n = \rho m$



Source-Channel Separation Theorem

$$\lim_{n\to\infty} D_n^* = D(\rho C_Q)$$

Define $\Delta_n = \Delta(n, \rho, P_S, Q_{Y|X}) = D_n^* - D(\rho C_Q)$

How does Δ_n scale with n?

A lot is known about FBL JSCC under **excess distortion** criterion

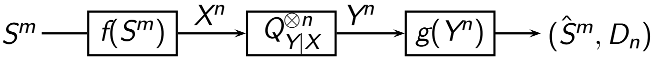
Practically nothing is known under **mean distortion** criterion

Maybe for a good reason...

$\Delta_n = 0$ when "Gastpar's conditions" satisfied

Joint Source-Channel Coding

$n = \rho m$

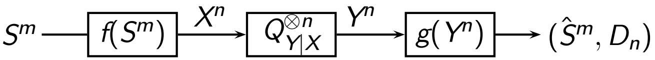


Recall:

- For channel coding $C - R_n^* = \sqrt{\frac{V}{n}} Q^{-1}(\epsilon) + o(\frac{1}{\sqrt{n}})$
- For source coding $D_n^*(R) - D(R) = -D'(R)\frac{\log n}{2n} + o(\frac{\log n}{n})$ [Zhang-Yang-Wei'97]

Joint Source-Channel Coding

$n = \rho m$



Recall:

- For channel coding $C - R_n^* = \sqrt{\frac{V}{n}} Q^{-1}(\epsilon) + o(\frac{1}{\sqrt{n}})$
- For source coding $D_n^*(R) - D(R) = -D'(R)\frac{\log n}{2n} + o(\frac{\log n}{n})$

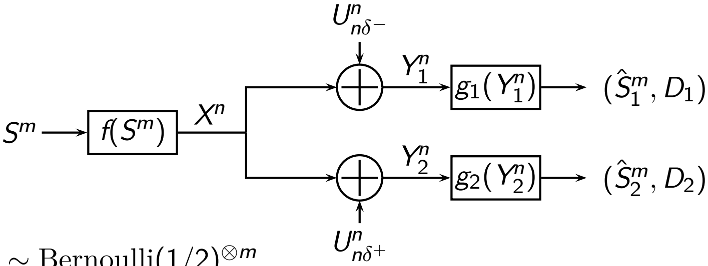
Our main result

For a transmission of a symmetric binary source over a binary symmetric channel with crossover probability δ

$$\Delta_n > \frac{c(\rho,\delta) \cdot D(\rho C_\delta)}{\sqrt{n}} + \mathcal{O}(n^{-3/4}\log n),$$

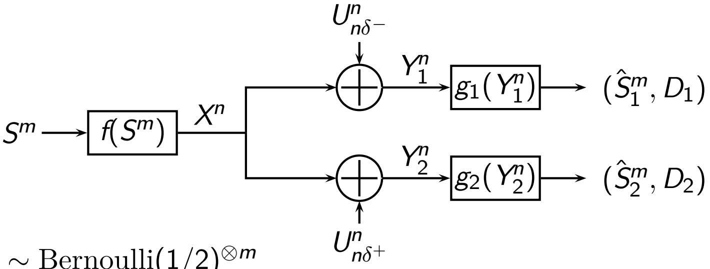
where $c(\rho,\delta) > 0$ for $\rho > 1$

Main Ingredient in the Proof



- $S^m \sim \mathrm{Bernoulli}(1/2)^{\otimes m}$
- $U^n_{n\delta}$ is uniform on $n\delta$-Hamming sphere ("type")
- $\delta^- = \delta - \frac{a}{\sqrt{n}}$, $\delta^+ = \delta + \frac{a}{\sqrt{n}}$ for some $a > 0$

Main Ingredient in the Proof

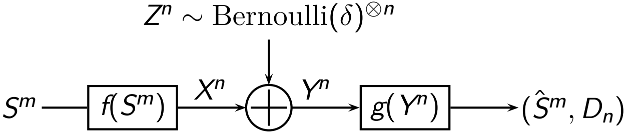


- $S^m \sim \text{Bernoulli}(1/2)^{\otimes m}$
- $U^n_{n\delta}$ is uniform on $n\delta$-Hamming sphere ("type")
- $\delta^- = \delta - \frac{a}{\sqrt{n}}$, $\delta^+ = \delta + \frac{a}{\sqrt{n}}$ for some $a > 0$

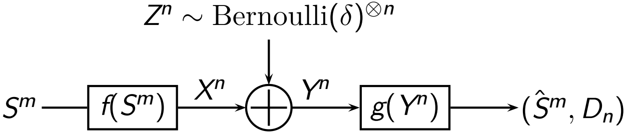
Main Technical Contribution

For any f, g_1, g_2 :
$\frac{D_1+D_2}{2} \geq D(\rho C_\delta) + \eta(\rho, \delta) D(\rho C_\delta) \frac{a}{2\sqrt{n}} + \mathcal{O}(n^{-3/4} \log n)$,
where $\eta(\rho, \delta) > 0$ for $\rho > 1$

Application of Theorem

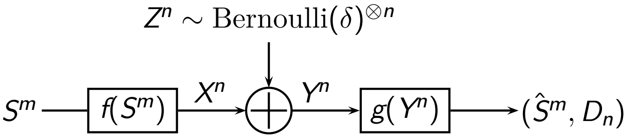


Application of Theorem



Let $A = \frac{w_H(Z^n) - n\delta}{\sqrt{n}}$
By CLT: "$A \sim \mathcal{N}(0, \delta(1-\delta))$"

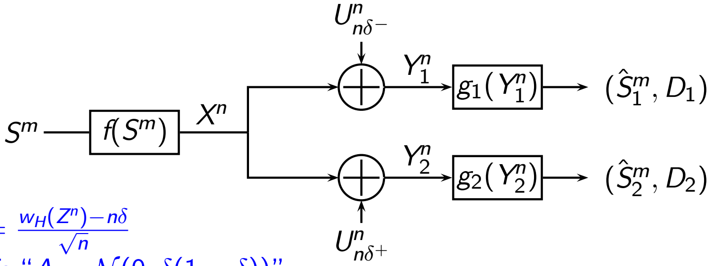
Application of Theorem



Let $A = \frac{w_H(Z^n) - n\delta}{\sqrt{n}}$
By CLT: "$A \sim \mathcal{N}(0, \delta(1-\delta))$"

$$D_n = \mathbb{E}[D_n| \ |A|]$$

Application of Theorem



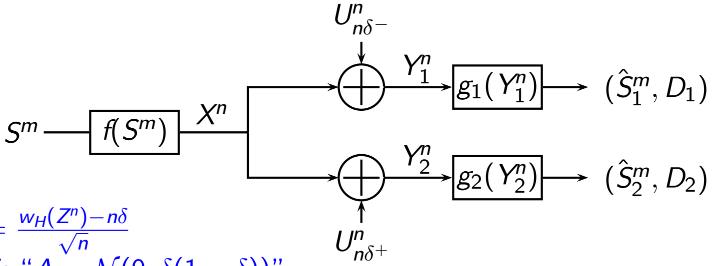
Let $A = \frac{w_H(Z^n) - n\delta}{\sqrt{n}}$

By CLT: "$A \sim \mathcal{N}(0, \delta(1-\delta))$"

Recall: $\delta^- = \delta - \frac{|A|}{\sqrt{n}}$, $\delta^+ = \delta + \frac{|A|}{\sqrt{n}}$

$$D_n = \mathbb{E}[D_n|\ |A|] \geq \frac{\mathbb{E}\left[D_1(|A|) + D_2(|A|)\right]}{2}$$

Application of Theorem



Let $A = \frac{w_H(Z^n) - n\delta}{\sqrt{n}}$

By CLT: "$A \sim \mathcal{N}(0, \delta(1-\delta))$"

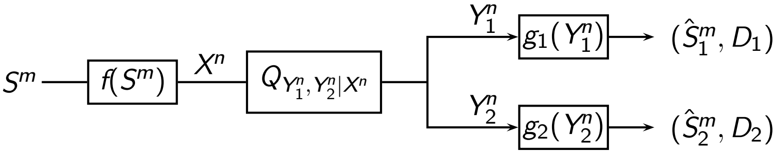
Recall: $\delta^- = \delta - \frac{|A|}{\sqrt{n}}$, $\delta^+ = \delta + \frac{|A|}{\sqrt{n}}$

$$D_n = \mathbb{E}[D_n |\ |A|] \geq \frac{\mathbb{E}\left[D_1(|A|) + D_2(|A|)\right]}{2}$$

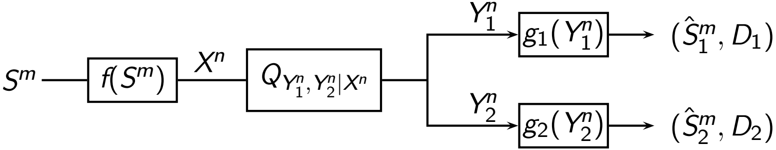
$$\geq D(\rho C_\delta)\left(1 + \frac{\mathbb{E}(|A|)}{2\sqrt{n}}\eta(\rho, \delta)\right) + \mathcal{O}(n^{-3/4}\log n)$$

Due to our Theorem

JSCC Over Broadcast

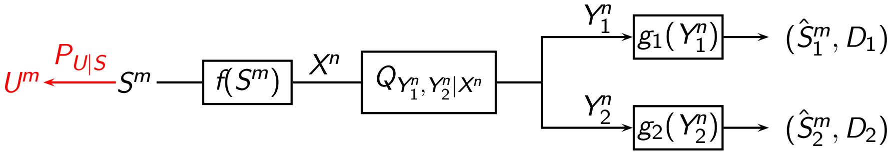


JSCC Over Broadcast



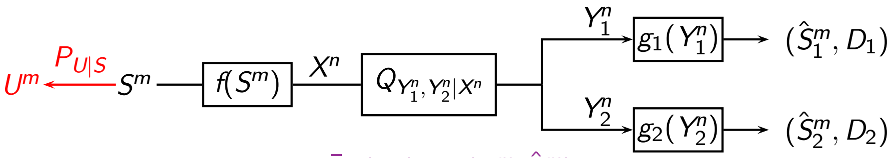
Nontrivial tradeoff between (D_1, D_2)
How to capture this?

JSCC Over Broadcast



Ozarow/Reznic-Feder-Zamir trick for converse:
Bound (D_1, D_2) tradeoff w.r.t. an auxiliary U^m generated by the DMC $P_{U|S}$

JSCC Over Broadcast



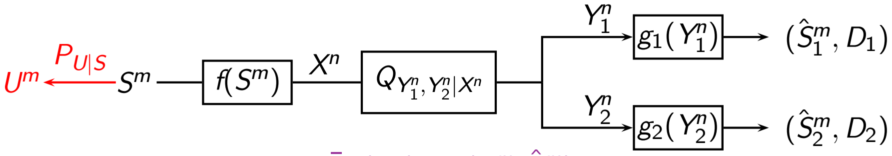
$$m \cdot \bar{R}_P(D_2) \leq I(U^m; \hat{S}_2^m)$$

$$\bar{R}_P(D) \triangleq \min_{\substack{\hat{S}\,:\; U-S-\hat{S} \\ \mathbb{E}d(S,\hat{S}) \leq D}} I(U; \hat{S})$$

Lemma

Let $\hat{S}^m$ be a random vector satisfying the Markov chain $U^m - S^m - \hat{S}^m$ and $\mathbb{E}d(S^m, \hat{S}^m) \leq D$, then $I(U^m; \hat{S}^m) \geq m\bar{R}_P(D)$

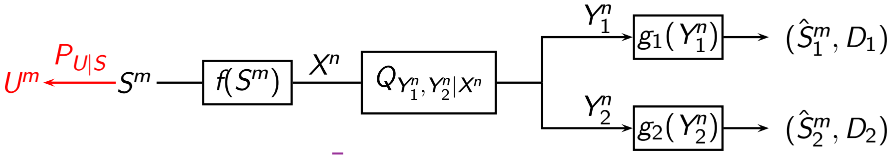
JSCC Over Broadcast



$$
\begin{aligned}
m \cdot \bar{R}_P(D_2) &\leq I(U^m; \hat{S}_2^m) \\
&\leq I(U^m; Y_2^n)
\end{aligned}
$$

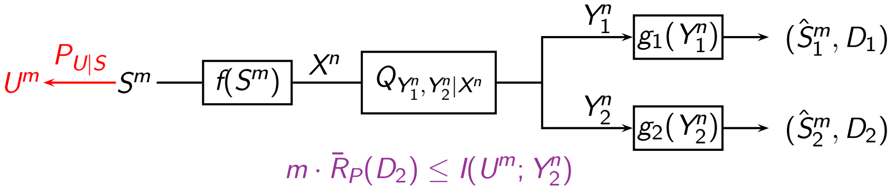
DPI

JSCC Over Broadcast



$$m \cdot \bar{R}_P(D_2) \leq I(U^m; Y_2^n)$$

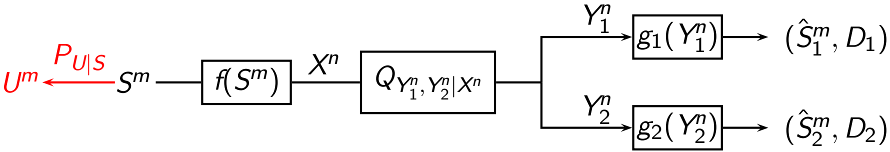
JSCC Over Broadcast



Let $Q^n = Q^n_{Y_1^n, Y_2^n | X^n}$.

$$G_{Q^n}(t) \triangleq \max_{\substack{W, X^n \,:\, W - X^n - Y_1^n - Y_2^n \\ I(X^n; Y_1^n | W) \geq t}} I(Y_2^n; W)$$

JSCC Over Broadcast

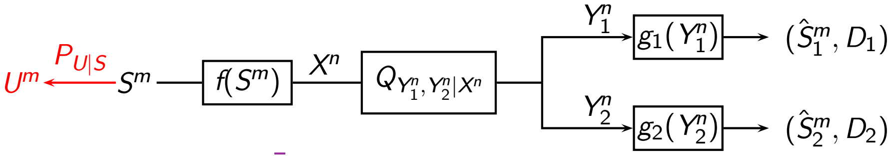


$$\begin{aligned} m \cdot \bar{R}_P(D_2) &\leq I(U^m; Y_2^n) \\ &\leq G_{Q^n}\left(I(X^n; Y_1^n | U^m)\right) \end{aligned}$$

Let $Q^n = Q^n_{Y_1^n, Y_2^n | X^n}$.

$$G_{Q^n}(t) \triangleq \max_{\substack{W, X^n \,:\, W - X^n - Y_1^n - Y_2^n \\ I(X^n; Y_1^n | W) \geq t}} I(Y_2^n; W)$$

JSCC Over Broadcast

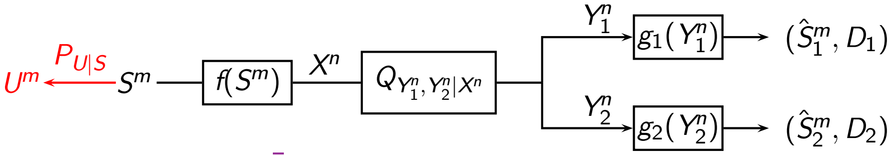


$$m \cdot \bar{R}_P(D_2) \leq G_{Q^n}\left(I(X^n; Y_1^n | U^m)\right)$$

Let $Q^n = Q^n_{Y_1^n, Y_2^n | X^n}$.

$$G_{Q^n}(t) \triangleq \max_{\substack{W, X^n \,:\, W - X^n - Y_1^n - Y_2^n \\ I(X^n; Y_1^n | W) \geq t}} I(Y_2^n; W)$$

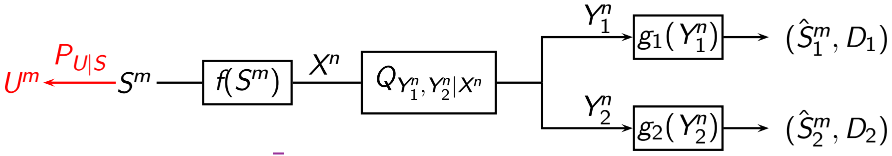
JSCC Over Broadcast



$$\begin{aligned} m \cdot \bar{R}_P(D_2) &\leq G_{Q^n}\left(I(X^n; Y_1^n | U^m)\right) \\ &\leq G_{Q^n}\left(I(S^m; Y_1^n | U^m)\right) \end{aligned}$$

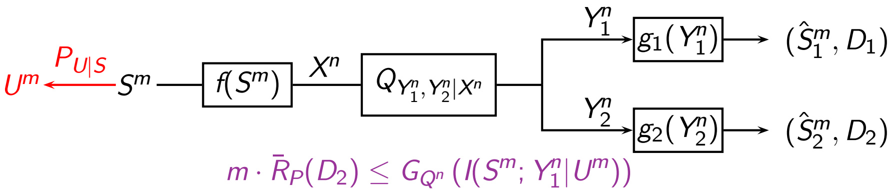
DPI

JSCC Over Broadcast



$$m \cdot \bar{R}_P(D_2) \leq G_{Q^n}\left(I(S^m; Y_1^n | U^m)\right)$$

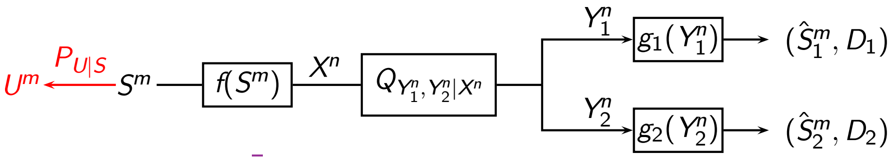
JSCC Over Broadcast



Let $P = P_{SU} = P_S P_{U|S}$ and define

$$F_P(t) \triangleq \min_{\substack{V \,:\, U - S - V \\ I(S;V) \geq t}} I(S; V | U)$$

JSCC Over Broadcast

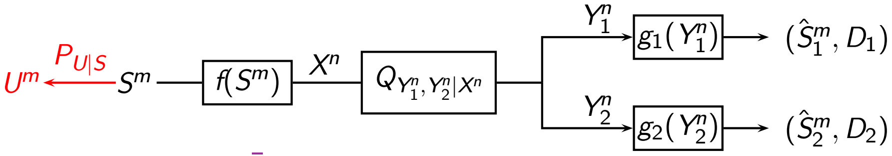


$$\begin{aligned} m \cdot \bar{R}_P(D_2) &\leq G_{Q^n}\left(I(S^m; Y_1^n | U^m)\right) \\ &\leq G_{Q^n}\left(F_{P^{\otimes m}}\left(I(S^m; Y_1^n)\right)\right) \end{aligned}$$

Let $P = P_{SU} = P_S P_{U|S}$ and define

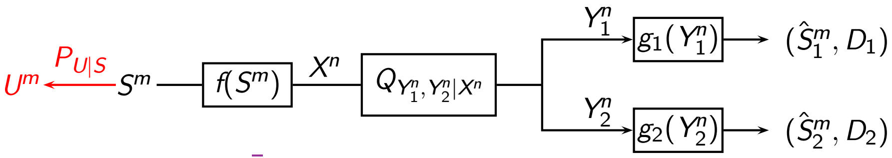
$$F_P(t) \triangleq \min_{\substack{V \,:\, U - S - V \\ I(S;V) \geq t}} I(S; V | U)$$

JSCC Over Broadcast



$$m \cdot \bar{R}_P(D_2) \leq G_{Q^n}\left(F_{P^{\otimes m}}\left(I(S^m; Y_1^n)\right)\right)$$

JSCC Over Broadcast

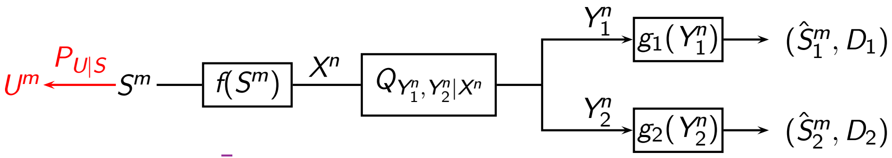


$$m \cdot \bar{R}_P(D_2) \leq G_{Q^n}\left(F_{P^{\otimes m}}\left(I(S^m; Y_1^n)\right)\right)$$

Lemma

The function $F_P(t)$ is monotone non-decreasing, convex, and tensorizes, i.e., $F_{P^m}(mt) = m \cdot F_P(t)$.

JSCC Over Broadcast

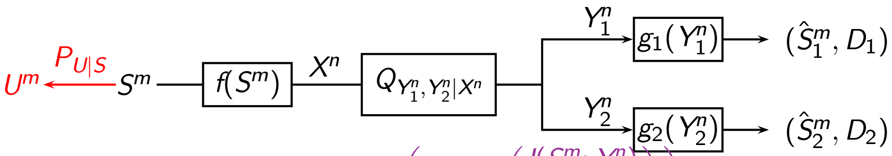


$$m \cdot \bar{R}_P(D_2) \le G_{Q^n}\left(F_{P^{\otimes m}}\left(I(S^m; Y_1^n)\right)\right)$$
$$\le G_{Q^n}\left(m \cdot F_P\left(\frac{I(S^m; Y_1^n)}{m}\right)\right)$$

Lemma

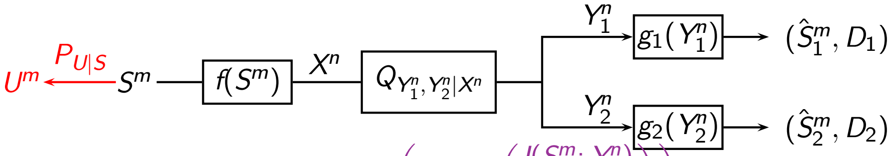
The function $F_P(t)$ is monotone non-decreasing, convex, and tensorizes, i.e., $F_{P^m}(mt) = m \cdot F_P(t)$.

JSCC Over Broadcast



$$m \cdot \bar{R}_P(D_2) \leq G_{Q^n}\left(m \cdot F_P\left(\frac{I(S^m; Y_1^n)}{m}\right)\right)$$

JSCC Over Broadcast

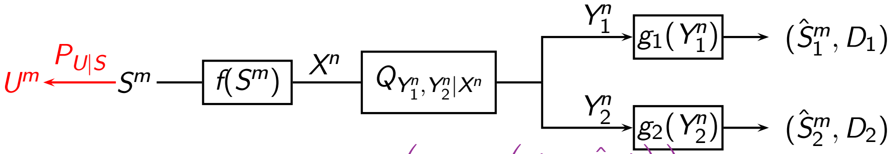


$$m \cdot \bar{R}_P(D_2) \leq G_{Q^n}\left(m \cdot F_P\left(\frac{I(S^m; Y_1^n)}{m}\right)\right)$$

$$\leq G_{Q^n}\left(m \cdot F_P\left(\frac{I(S^m; \hat{S}_1^m)}{m}\right)\right)$$

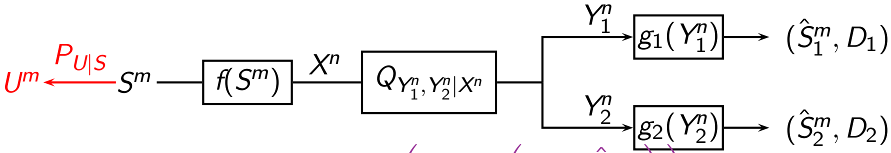
DPI

JSCC Over Broadcast



$$m \cdot \bar{R}_P(D_2) \leq G_{Q^n}\left(m \cdot F_P\left(\frac{I(S^m; \hat{S}_1^m)}{m}\right)\right)$$

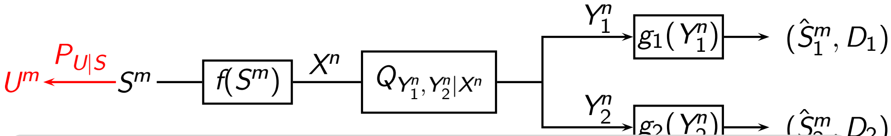
JSCC Over Broadcast



$$
\begin{aligned}
m \cdot \bar{R}_P(D_2) &\leq G_{Q^n}\left(m \cdot F_P\left(\frac{I(S^m; \hat{S}_1^m)}{m}\right)\right) \\
&\leq G_{Q^n}\left(m \cdot F_P\left(R(D_1)\right)\right)
\end{aligned}
$$

RDF

JSCC Over Broadcast

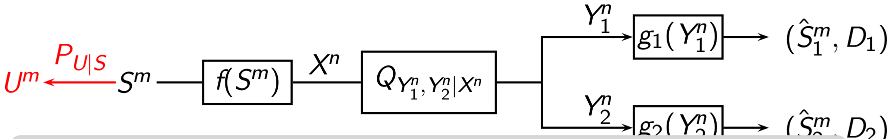


Theorem

If (D_1, D_2) are achievable over the broadcast Q^n, then for any choice of $P_{U|S}$ we have

$$m\bar{R}_P(D_2) \leq G_{Q^n}\left(m \cdot F_P\left(R(D_1)\right)\right)$$

JSCC Over Broadcast



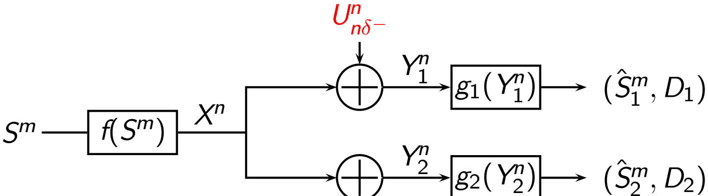
Theorem

If (D_1, D_2) are achievable over the degraded memoryless broadcast $Q^{\otimes n}$, then for any choice of $P_{U|S}$ we have

$$\bar{R}_P(D_2) \leq \rho G_Q\left(\frac{F_P(R(D_1))}{\rho}\right)$$

- If Q^n is degraded memoryless BC, then $G_{Q^{\otimes n}}(nt) = nG_Q(t)$
- For Quadratic Gaussian over GBC, recovers [Reznic-Feder-Zamir, IT'06]
- But is implied by [Khezeli-Chen IT'16]

JSCC Over Spherical Noise Broadcast

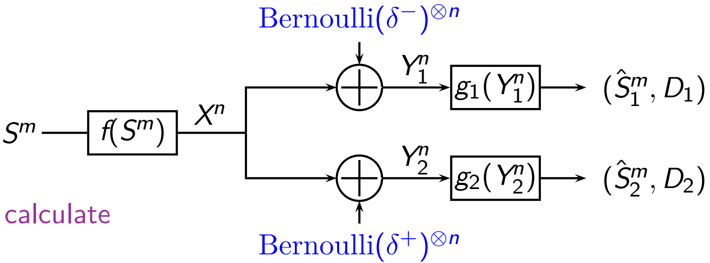


How to calculate

$$G_{Q^n}(t) \triangleq \max_{\substack{W, X^n \,:\, W - X^n - Y^n_1 - Y^n_2 \\ I(X^n; Y^n_1 | W) \geq t}} I(Y^n_2; W)$$

for this particular Q^n?

JSCC Over Spherical Noise Broadcast



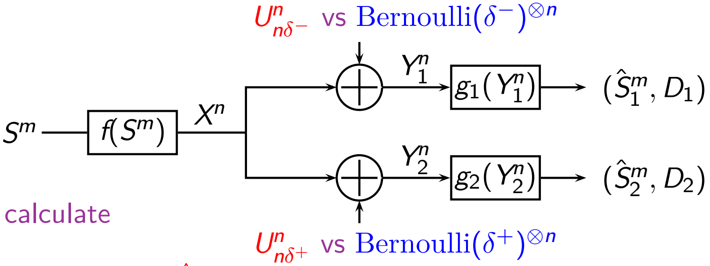
How to calculate

$$G_{Q^n}(t) \triangleq \max_{\substack{W,X^n \,:\, W - X^n - Y_1^n - Y_2^n \\ I(X^n;Y_1^n|W)\geq t}} I(Y_2^n; W)$$

for this particular Q^n?

If spherical noise replaced with Bernoulli noise, BC Q^n becomes degraded memoryless BC $\tilde{Q}^{\otimes n}$,and $G_{\tilde{Q}^{\otimes n}}(t)$ tensorizes

JSCC Over Spherical Noise Broadcast



How to calculate

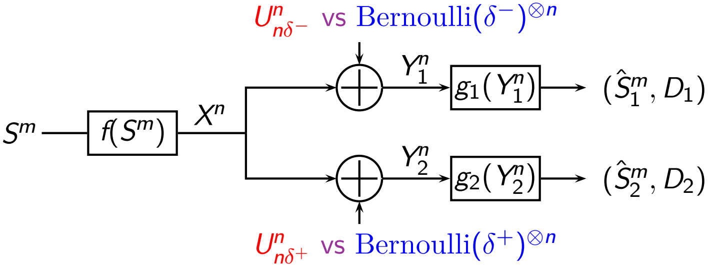
$$G_{Q^n}(t) \triangleq \max_{\substack{W, X^n \,:\, W - X^n - Y_1^n - Y_2^n \\ I(X^n; Y_1^n | W) \geq t}} I(Y_2^n; W)$$

for this particular Q^n?

If spherical noise replaced with Bernoulli noise, BC Q^n becomes degraded memoryless BC $\tilde{Q}^{\otimes n}$,and $G_{\tilde{Q}^{\otimes n}}(t)$ tensorizes

How different can $G_{Q^n}(t)$ and $G_{\tilde{Q}^{\otimes n}}(t)$be?

JSCC Over Spherical Noise Broadcast



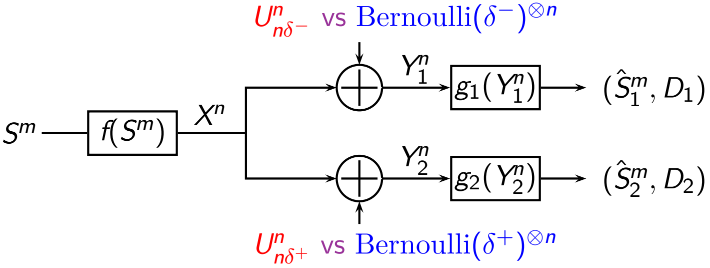
Theorem

$$G_{Q^n}(nt) \leq G_{\tilde{Q}^{\otimes n}}(nt) + n\epsilon_n = n\left(G_{\tilde{Q}}(t) + \epsilon_n\right),$$

where

$$\epsilon_n = \mathcal{O}(n^{-3/4}\log n) = o(n^{-1/2})$$

JSCC Over Spherical Noise Broadcast

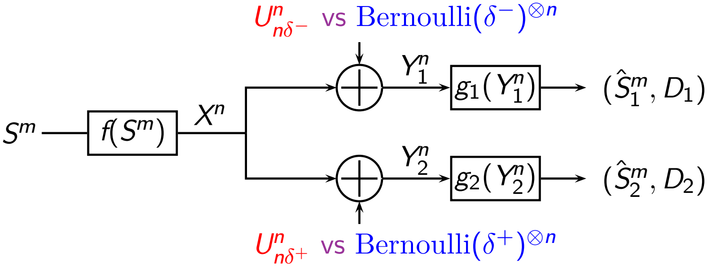


Main ingredient in the proof:

Let $Z^n \sim \mathrm{Bernoulli}(\mu)^{\otimes n}$, s.t. $\mu * \delta^- = \delta^+$. For any W, X^n

$$|H(X^n + U^n_{n\delta^+}|W) - H(X^n + U^n_{n\delta^-} + Z^n|W)| \leq \mathcal{O}(n^{1/4}\log n)$$

JSCC Over Spherical Noise Broadcast



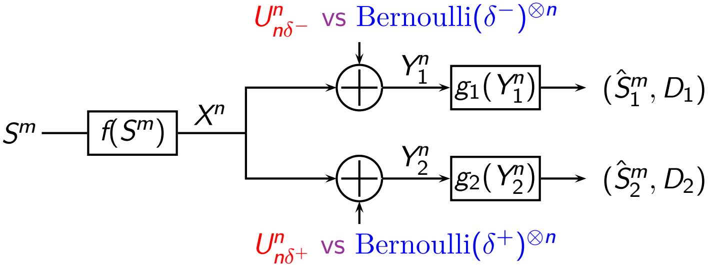
Main ingredient in the proof:

Let $Z^n \sim \text{Bernoulli}(\mu)^{\otimes n}$, s.t. $\mu * \delta^- = \delta^+$. For any W, X^n

$$|H(X^n + U^n_{n\delta^+}|W) - H(X^n + U^n_{n\delta^-} + Z^n|W)| \leq \mathcal{O}(n^{1/4} \log n)$$

Established via coupling technique of [Polyanskiy-Wu IT'16]

JSCC Over Spherical Noise Broadcast



Main ingredient in the proof:

Let $Z^n \sim \text{Bernoulli}(\mu)^{\otimes n}$, s.t. $\mu * \delta^- = \delta^+$. For any W, X^n

$$|H(X^n + U^n_{n\delta^+}|W) - H(X^n + U^n_{n\delta^-} + Z^n|W)| \leq \mathcal{O}(n^{1/4}\log n)$$

$\Rightarrow$ Q^n can be replaced with degraded BC, up to negligible terms

Summary

- D_n^* does not decay faster than $\frac{1}{\sqrt{n}}$ in general
- Technique goes through a reduction to transmitting a source over a broadcast channel
- Seems possible to extend to Quadratic Gaussian JSCC problem
- What about achievability? Separation only gives $\Delta_n < O\left(\sqrt{\frac{\log n}{n}}\right)$