Asymptotic Estimates in Information Theory with Non-Vanishing Error Probabilities

Vincent Y. F. Tan

Dept. of ECE and Dept. of Mathematics
National University of Singapore (NUS)

September 2014

Acknowledgements

This is joint work with

1. Sy-Quoc Le and Mehul Motani (NUS)

2. Jon Scarlett (Cambridge, now at EPFL)

Outline

Outline

Outline

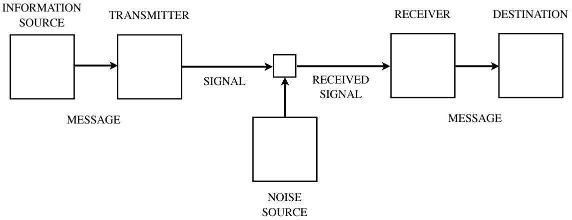
1. Motivation, Background and History
2. Gaussian Interference Channel with Very Strong Interference
3. Gaussian MAC with Degraded Message Sets

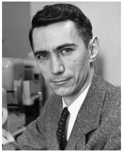
Outline

Outline

1. Motivation, Background and History
2. Gaussian Interference Channel with Very Strong Interference
3. Gaussian MAC with Degraded Message Sets
4. Conclusion

Transmission of Information

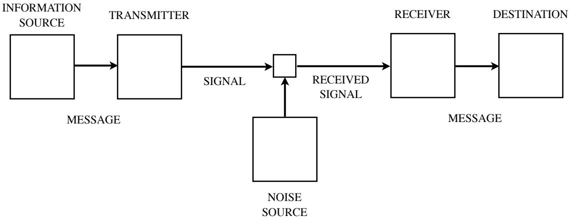


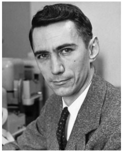


Shannon's Figure 1

- Information theory $\equiv$ Finding fundamental limits for reliable information transmission

Transmission of Information

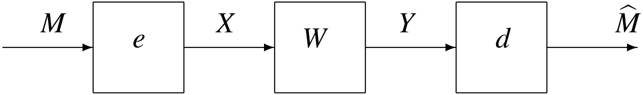




Shannon's Figure 1

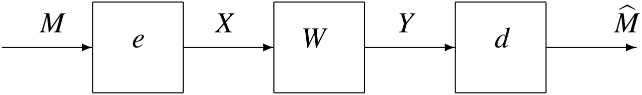
- Information theory $\equiv$ Finding fundamental limits for reliable information transmission
- Channel coding: Concerned with the maximum rate of communication in bits/channel use

Channel Coding (One-Shot)



- A code is an triple $\mathcal{C} = \{\mathcal{M}, e, d\}$ where $\mathcal{M}$ is the message set and $b(e(m)) \leq S$ for some cost function $b(\cdot)$ and cost S

Channel Coding (One-Shot)

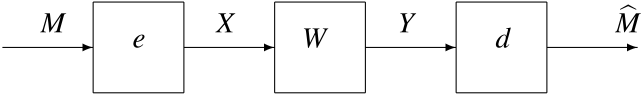


- A code is an triple $\mathcal{C} = \{\mathcal{M}, e, d\}$ where $\mathcal{M}$ is the message set and $b(e(m)) \le S$ for some cost function $b(\cdot)$ and cost S
- The average error probability $p_{\mathrm{err}}(\mathcal{C})$ is

$$p_{\mathrm{err}}(\mathcal{C}) \ := \ \Pr\left[\widehat{M} \neq M\right]$$

where M is uniform on $\mathcal{M}$

Channel Coding (One-Shot)



- A code is an triple $\mathcal{C} = \{\mathcal{M}, e, d\}$ where $\mathcal{M}$ is the message set and $b(e(m)) \le S$ for some cost function $b(\cdot)$ and cost S
- The average error probability $p_{\mathrm{err}}(\mathcal{C})$ is

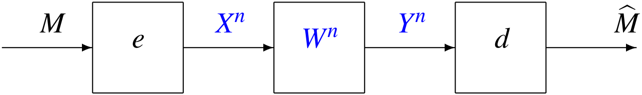
$$p_{\mathrm{err}}(\mathcal{C}) \;:=\; \Pr\left[\widehat{M} \neq M\right]$$

  where M is uniform on $\mathcal{M}$

- A non-asymptotic fundamental limit can be defined as

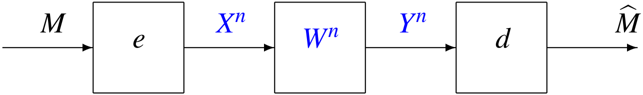
$$M^*(W, \varepsilon, S) := \sup\left\{m \in \mathbb{N} \,\middle|\, \exists\, \mathcal{C} \text{ s.t. } m = |\mathcal{M}|,\, p_{\mathrm{err}}(\mathcal{C}) \le \varepsilon\right\}$$

Channel Coding (n-Shot) for AWGN



- Consider n independent uses of an additive white Gaussian channel W^n

Channel Coding (n-Shot) for AWGN



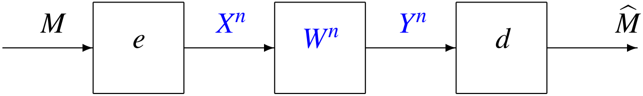
- Consider n independent uses of an additive white Gaussian channel W^n
- For vectors $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\mathbf{y} := (y_1, \ldots, y_n) \in \mathbb{R}^n$, with

$$\|\mathbf{x}\|_2^2 \le nS,$$

the channel law is

$$W^n(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_i - x_i)^2}{2}\right)$$

Channel Coding (n-Shot) for AWGN



- Consider n independent uses of an additive white Gaussian channel W^n

- For vectors $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\mathbf{y} := (y_1, \ldots, y_n) \in \mathbb{R}^n$, with

$$\|\mathbf{x}\|_2^2 \le nS,$$

  the channel law is

$$W^n(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_i - x_i)^2}{2} \right)$$

- Non-asymptotic fundamental limit for n uses of W

$$M^*(W^n, \varepsilon, S)$$

Single-User Asymptotic Evaluation

Theorem (Hayashi (2009), Polyanskiy-Poor-Verdú (2010), Tan-Tomamichel (2014))

For every $\varepsilon \in (0,1)$*, we have*

$$\log M^*(W^n, \varepsilon, S) = n\mathrm{C}(S) + \sqrt{n\mathrm{V}(S)}\Phi^{-1}(\varepsilon) + \frac{1}{2}\log n + O(1)$$

where $\mathrm{C}(S)$ *and* $\mathrm{V}(S)$ *are the capacity and dispersion defined as*

$$\mathrm{C}(S) = \frac{1}{2}\log(1+S), \qquad \mathrm{V}(S) = \frac{S(S+2)}{(S+1)^2}\log^2 e$$

Single-User Asymptotic Evaluation

Theorem (Hayashi (2009), Polyanskiy-Poor-Verdú (2010), Tan-Tomamichel (2014))

For every $\varepsilon \in (0,1)$, *we have*

$$\log M^*(W^n, \varepsilon, S) = n\mathrm{C}(S) + \sqrt{n\mathrm{V}(S)}\Phi^{-1}(\varepsilon) + \frac{1}{2}\log n + O(1)$$

where $\mathrm{C}(S)$ *and* $\mathrm{V}(S)$ *are the capacity and dispersion defined as*

$$\mathrm{C}(S) = \frac{1}{2}\log(1+S), \qquad \mathrm{V}(S) = \frac{S(S+2)}{(S+1)^2}\log^2 e$$

M. Hayashi



Polyanskiy-Poor-Verdú

M. Tomamichel

Single-User Asymptotic Evaluation: Interpretation

$$R^*(W^n, \varepsilon, S) := \frac{\log M^*(W^n, \varepsilon, S)}{n} = \underbrace{\mathrm{C}(S) + \sqrt{\frac{\mathrm{V}(S)}{n}}\Phi^{-1}(\varepsilon)}_{\text{Gaussian approximation}} + O\left(\frac{\log n}{n}\right)$$

Single-User Asymptotic Evaluation: Interpretation

$$R^*(W^n, \varepsilon, S) := \frac{\log M^*(W^n, \varepsilon, S)}{n} = \underbrace{\mathrm{C}(S) + \sqrt{\frac{\mathrm{V}(S)}{n}}\Phi^{-1}(\varepsilon)}_{\text{Gaussian approximation}} + O\left(\frac{\log n}{n}\right)$$

- Interpretation: The backoff from $\mathrm{C}(S)$ at finite blocklength n and tolerable error probability is approximately

$$\sqrt{\frac{\mathrm{V}(S)}{n}}\,\Phi^{-1}(1-\varepsilon)$$

- Small ε implies large backoff

Single-User Asymptotic Evaluation: Interpretation

$$R^*(W^n, \varepsilon, S) := \frac{\log M^*(W^n, \varepsilon, S)}{n} = \underbrace{\mathrm{C}(S) + \sqrt{\frac{\mathrm{V}(S)}{n}}\Phi^{-1}(\varepsilon)}_{\text{Gaussian approximation}} + O\Big(\frac{\log n}{n}\Big)$$

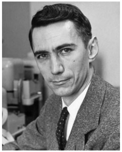
- Interpretation: The backoff from $\mathrm{C}(S)$ at finite blocklength n and tolerable error probability is approximately

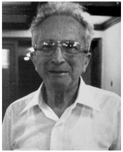
$$\sqrt{\frac{\mathrm{V}(S)}{n}}\,\Phi^{-1}(1-\varepsilon)$$

- Small ε implies large backoff
- Can compare to actual finite blocklength bounds
- Gaussian approximation is good for some n and ε

Background: Shannon's Channel Coding Theorem

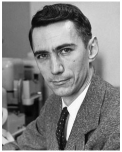
- Shannon's noisy channel coding theorem and
- Shannon's (1959), Yoshihara's (1964) and Wolfowitz's (1978) strong converse state that

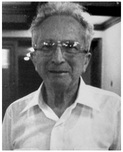




Background: Shannon's Channel Coding Theorem

- Shannon's noisy channel coding theorem and
- Shannon's (1959), Yoshihara's (1964) and Wolfowitz's (1978) strong converse state that





Theorem (Shannon (1949), Shannon (1959))

$$\lim_{n \to \infty} \frac{1}{n} \log M^*(W^n, \varepsilon, S) = \mathrm{C}(S), \qquad \forall\, \varepsilon \in (0, 1)$$

Background: Shannon's Channel Coding Theorem

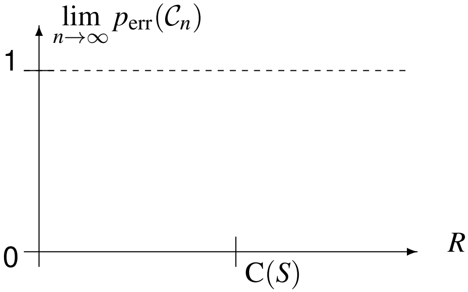
$$\lim_{n\to\infty} \frac{1}{n} \log M^*(W^n, \varepsilon, S) = \mathrm{C}(S) \quad \text{bits/channel use}$$

- Channel coding theorem for AWGN channels is independent of $\varepsilon \in (0,1)$

Background: Shannon's Channel Coding Theorem

$$\lim_{n\to\infty} \frac{1}{n} \log M^*(W^n, \varepsilon, S) = \mathrm{C}(S) \quad \text{bits/channel use}$$

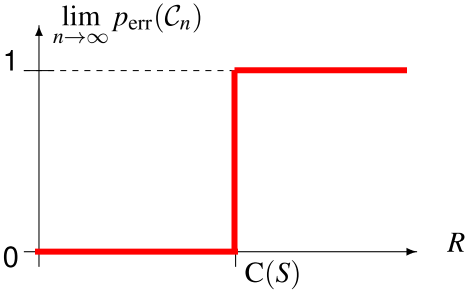
- Channel coding theorem for AWGN channels is independent of $\varepsilon \in (0,1)$



Background: Shannon's Channel Coding Theorem

$$\lim_{n\to\infty} \frac{1}{n} \log M^*(W^n, \varepsilon, S) = \mathrm{C}(S) \quad \text{bits/channel use}$$

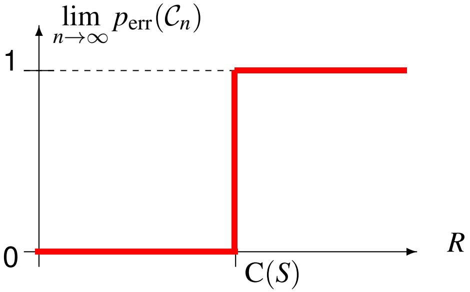
- Channel coding theorem for AWGN channels is independent of $\varepsilon \in (0, 1)$



Background: Shannon's Channel Coding Theorem

$$\lim_{n\to\infty} \frac{1}{n} \log M^*(W^n, \varepsilon, S) = \mathrm{C}(S) \quad \text{bits/channel use}$$

- Channel coding theorem for AWGN channels is independent of $\varepsilon \in (0,1)$



- Phase transition at capacity

Background: Second-Order Coding Rates

- What happens at capacity?

Background: Second-Order Coding Rates

- What happens at capacity?
- More precisely, what happens when

$$\log |\mathcal{M}_n| \approx n\mathrm{C}(S) + L\sqrt{n}$$

for some $L \in \mathbb{R}$?

Background: Second-Order Coding Rates

- What happens at capacity?
- More precisely, what happens when

$$\log |\mathcal{M}_n| \approx n\mathrm{C}(S) + L\sqrt{n}$$

for some $L \in \mathbb{R}$?

- Here L is known as the second-order coding rate of the code

Background: Second-Order Coding Rates

- What happens at capacity?
- More precisely, what happens when

$$\log |\mathcal{M}_n| \approx n\mathrm{C}(S) + L\sqrt{n}$$

  for some $L \in \mathbb{R}$?
- Here L is known as the second-order coding rate of the code
- Note that L can be negative (cf. Hayashi (2008), Hayashi (2009))

Background: Second-Order Coding Rates

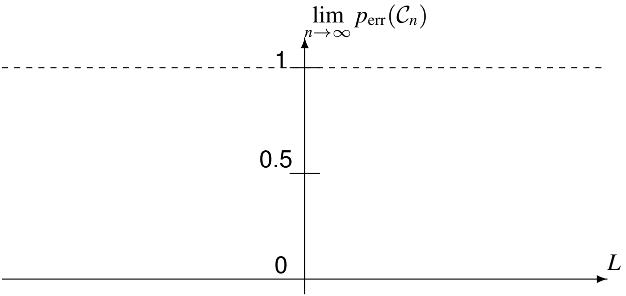
Assume rate of the code satisfies

$$\frac{1}{n}\log|\mathcal{M}_n| = \mathrm{C}(S) + \frac{L}{\sqrt{n}} + o\Big(\frac{1}{\sqrt{n}}\Big)$$

Background: Second-Order Coding Rates

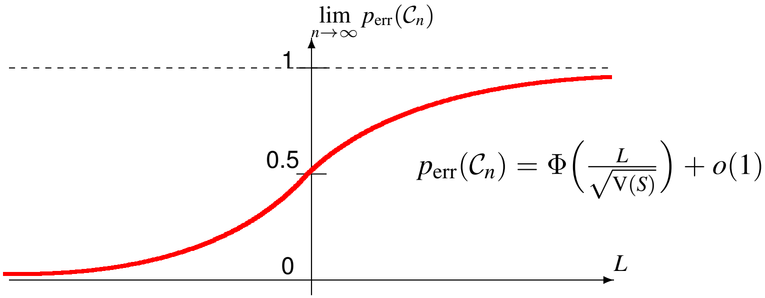
Assume rate of the code satisfies

$$\frac{1}{n}\log|\mathcal{M}_n| = \mathrm{C}(S) + \frac{L}{\sqrt{n}} + o\Big(\frac{1}{\sqrt{n}}\Big)$$



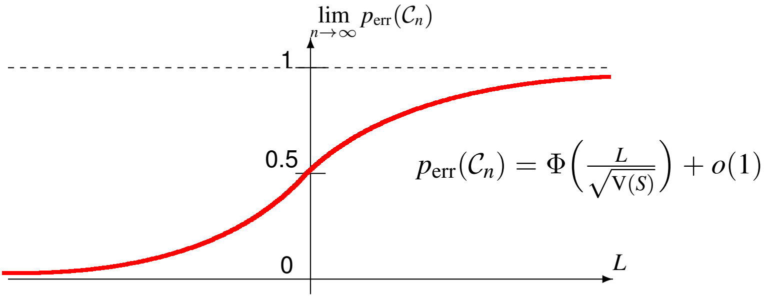
Assume rate of the code satisfies

$$\frac{1}{n}\log|\mathcal{M}_n| = \mathrm{C}(S) + \frac{L}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right)$$



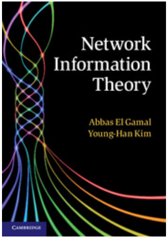
Assume rate of the code satisfies

$$\frac{1}{n}\log|\mathcal{M}_n| = \mathrm{C}(S) + \frac{L}{\sqrt{n}} + o\Big(\frac{1}{\sqrt{n}}\Big)$$



For an error probability ε, the optimum second-order coding rate is

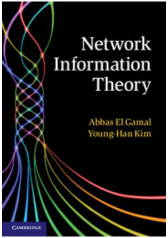
$$L^*(\varepsilon) := \sqrt{\mathrm{V}(S)}\,\Phi^{-1}(\varepsilon)$$



Network Information Theory
by A. El Gamal and Y.-H. Kim

- Many problems unsolved

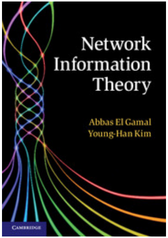
Network Information Theory



Network Information Theory
by A. El Gamal and Y.-H. Kim

- Many problems unsolved
- My agenda is to understand second-order behavior of solved NIT problems

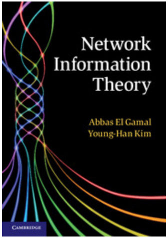
Network Information Theory



Network Information Theory
by A. El Gamal and Y.-H. Kim

- Many problems unsolved
- My agenda is to understand second-order behavior of solved NIT problems
- Gain new insights on second-order optimal coding schemes

Network Information Theory

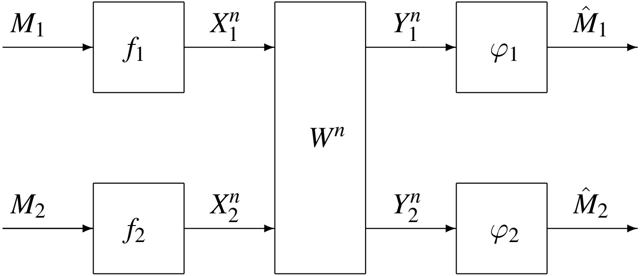


Network Information Theory
by A. El Gamal and Y.-H. Kim

- Many problems unsolved
- My agenda is to understand second-order behavior of solved NIT problems
- Gain new insights on second-order optimal coding schemes
- We study two simple NIT problems in the rest of the talk

Outline

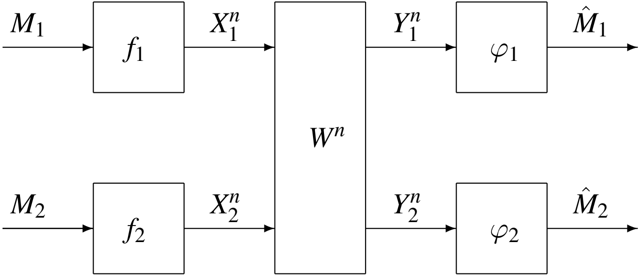
Gaussian Interference Channel



- Two-sender two-receiver GIC is given as

$$Y_{1i} = g_{11} X_{1i} + g_{12} X_{2i} + Z_{1i}, \qquad Y_{2i} = g_{21} X_{1i} + g_{22} X_{2i} + Z_{2i}$$

Gaussian Interference Channel



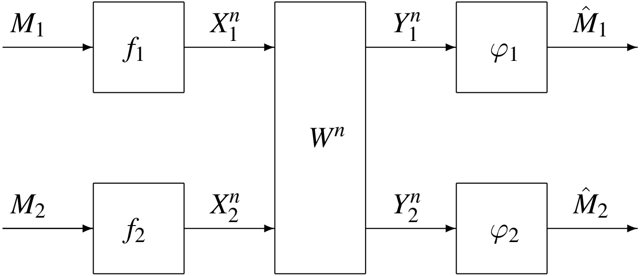
- Two-sender two-receiver GIC is given as

$$Y_{1i} = g_{11}X_{1i} + g_{12}X_{2i} + Z_{1i}, \qquad Y_{2i} = g_{21}X_{1i} + g_{22}X_{2i} + Z_{2i}$$

- Channel inputs are power limited

$$\sum_{i=1}^{n} X_{ji}^2 \le nS_j, \qquad j = 1, 2$$

Gaussian Interference Channel



- Two-sender two-receiver GIC is given as

$$Y_{1i} = g_{11}X_{1i} + g_{12}X_{2i} + Z_{1i}, \qquad Y_{2i} = g_{21}X_{1i} + g_{22}X_{2i} + Z_{2i}$$

- Channel inputs are power limited

$$\sum_{i=1}^{n} X_{ji}^2 \leq nS_j, \qquad j = 1, 2$$

- Capacity region is an open problem in NIT

Very Strong Inference

- Define the signal-to-noise ratios as

$$\mathrm{snr}_1 = g_{11}^2 S_1, \qquad \mathrm{snr}_2 = g_{22}^2 S_2$$

- Define the interference-to-noise ratios as

$$\mathrm{inr}_1 = g_{12}^2 S_2, \qquad \mathrm{inr}_2 = g_{21}^2 S_1$$

Very Strong Inference

- Define the signal-to-noise ratios as

$$\mathrm{snr}_1 = g_{11}^2 S_1, \qquad \mathrm{snr}_2 = g_{22}^2 S_2$$

- Define the interference-to-noise ratios as

$$\mathrm{inr}_1 = g_{12}^2 S_2, \qquad \mathrm{inr}_2 = g_{21}^2 S_1$$

- A GIC is said to have strictly very strong interference (SVSI) if

$$\mathrm{snr}_1 < \frac{\mathrm{inr}_2}{1+\mathrm{snr}_2}, \qquad \mathrm{snr}_2 < \frac{\mathrm{inr}_1}{1+\mathrm{snr}_1}$$

- Equivalently

$$\mathrm{C}(\mathrm{snr}_1) + \mathrm{C}(\mathrm{snr}_2) < \min\{\mathrm{C}(\mathrm{snr}_1 + \mathrm{inr}_1), \mathrm{C}(\mathrm{snr}_2 + \mathrm{inr}_2)\}$$

Very Strong Inference

- Define the signal-to-noise ratios as

$$\mathrm{snr}_1 = g_{11}^2 S_1, \qquad \mathrm{snr}_2 = g_{22}^2 S_2$$

- Define the interference-to-noise ratios as

$$\mathrm{inr}_1 = g_{12}^2 S_2, \qquad \mathrm{inr}_2 = g_{21}^2 S_1$$

- A GIC is said to have strictly very strong interference (SVSI) if

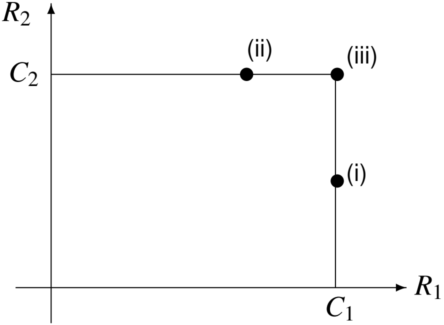
$$\mathrm{snr}_1 < \frac{\mathrm{inr}_2}{1+\mathrm{snr}_2}, \qquad \mathrm{snr}_2 < \frac{\mathrm{inr}_1}{1+\mathrm{snr}_1}$$

- Equivalently

$$\mathrm{C}(\mathrm{snr}_1) + \mathrm{C}(\mathrm{snr}_2) < \min\{\mathrm{C}(\mathrm{snr}_1 + \mathrm{inr}_1), \mathrm{C}(\mathrm{snr}_2 + \mathrm{inr}_2)\}$$

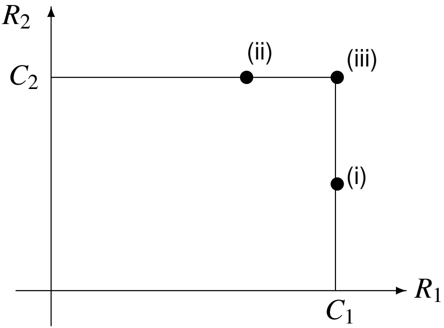
- Intuition: Receiver 1 decodes interference M_2, then uses that to decode intended message M_1 (and vice versa)

Capacity Region for GICs with SVSI



A. Carleial

Capacity Region for GICs with SVSI

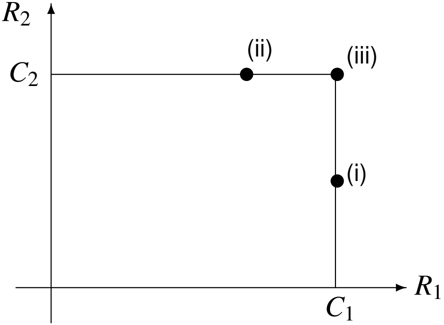


A. Carleial

- Carleial (1975) showed that the capacity region is

$$R_1 \leq C_1 := \mathrm{C}(\mathrm{snr}_1) \qquad R_2 \leq C_2 := \mathrm{C}(\mathrm{snr}_2)$$

Capacity Region for GICs with SVSI



A. Carleial

- Carleial (1975) showed that the capacity region is

$$R_1 \le C_1 := \mathrm{C}(\mathrm{snr}_1) \qquad R_2 \le C_2 := \mathrm{C}(\mathrm{snr}_2)$$

- Examine deviations of order $\frac{1}{\sqrt{n}}$ away from the boundary
- Three distinct regions.

Second-Order Coding Rate Region

- (L_1, L_2) is an $(\varepsilon, R_1^*, R_2^*)$-achievable second-order coding rate pair if there exists a sequence of $(n, M_{1n}, M_{2n}, \varepsilon_n)$-codes such that the code sizes M_{jn} satisfy

$$\liminf_{n\to\infty} \frac{1}{\sqrt{n}}(\log M_{jn} - nR_j^*) \geq L_j, \qquad j = 1, 2$$

Second-Order Coding Rate Region

- (L_1, L_2) is an $(\varepsilon, R_1^*, R_2^*)$-achievable second-order coding rate pair if there exists a sequence of $(n, M_{1n}, M_{2n}, \varepsilon_n)$-codes such that the code sizes M_{jn} satisfy

$$\liminf_{n\to\infty} \frac{1}{\sqrt{n}}(\log M_{jn} - nR_j^*) \geq L_j, \qquad j = 1, 2$$

and the average error probabilities ε_n satisfy

$$\limsup_{n\to\infty} \varepsilon_n \leq \varepsilon$$

Second-Order Coding Rate Region

- (L_1, L_2) is an $(\varepsilon, R_1^*, R_2^*)$-achievable second-order coding rate pair if there exists a sequence of $(n, M_{1n}, M_{2n}, \varepsilon_n)$-codes such that the code sizes M_{jn} satisfy

$$\liminf_{n\to\infty} \frac{1}{\sqrt{n}}(\log M_{jn} - nR_j^*) \geq L_j, \qquad j = 1, 2$$

and the average error probabilities ε_n satisfy

$$\limsup_{n\to\infty} \varepsilon_n \leq \varepsilon$$

- If (L_1, L_2) is $(\varepsilon, R_1^*, R_2^*)$-achievable, then there exists a sequence of codes such that

$$\log M_{jn} \geq nR_j^* + \sqrt{n}L_j + o(\sqrt{n})$$

and

$$\varepsilon_n \leq \varepsilon + o(1).$$

Optimum Second-Order Coding Rate Region

- The set of all $(\varepsilon, R_1^*, R_2^*)$-achievable second-order coding rate pairs is called the $(\varepsilon, R_1^*, R_2^*)$-optimum second-order coding rate region

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*)$$

Optimum Second-Order Coding Rate Region

- The set of all $(\varepsilon, R_1^*, R_2^*)$-achievable second-order coding rate pairs is called the $(\varepsilon, R_1^*, R_2^*)$-optimum second-order coding rate region

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*)$$

- Note: In the single-user case,

$$L^*(\varepsilon) := \sup \mathcal{L}(\varepsilon; R_1^* = C_1) = \sqrt{\mathrm{V}(S_1)}\Phi^{-1}(\varepsilon)$$

- This follows from Hayashi's work (2009)

Characterize $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$

- This is joint work with

S.-Q. Le (NUS)

M. Motani (NUS)

- We want to characterize $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$ for all points (R_1^*, R_2^*) on the boundary of the capacity region

Characterize $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$

- This is joint work with

S.-Q. Le (NUS)

M. Motani (NUS)

- We want to characterize $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$ for all points (R_1^*, R_2^*) on the boundary of the capacity region
- For all (R_1^*, R_2^*) in the interior of the capacity region

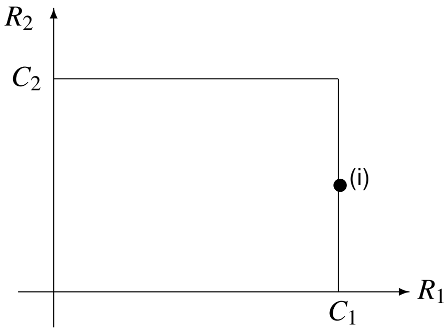
$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \mathbb{R}^2$$

- For all (R_1^*, R_2^*) in the exterior of the capacity region

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \emptyset$$

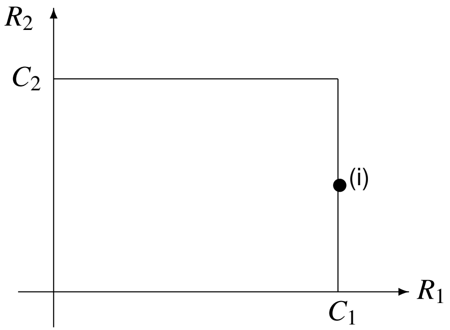
implying the strong converse

Main Result: Vertical Boundary



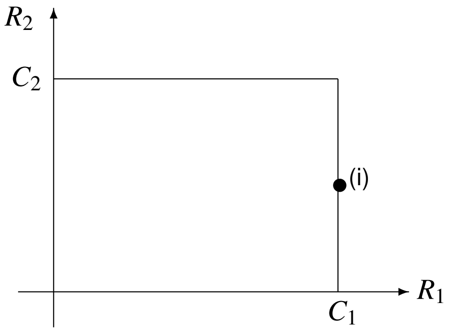
- In case (i), we are far from the horizontal boundary $R_2 < C_2$

Main Result: Vertical Boundary



- In case (i), we are far from the horizontal boundary $R_2 < C_2$
- Error event for user 2 vanishes (large deviations)

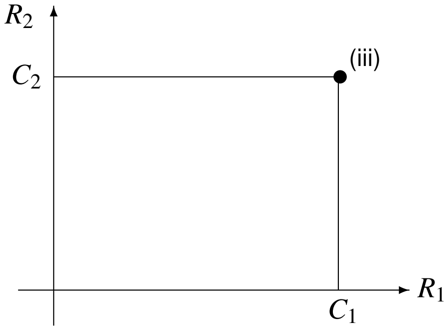
Main Result: Vertical Boundary



- In case (i), we are far from the horizontal boundary $R_2 < C_2$
- Error event for user 2 vanishes (large deviations)
- Hence second-order asymptotics only pertains to user 1

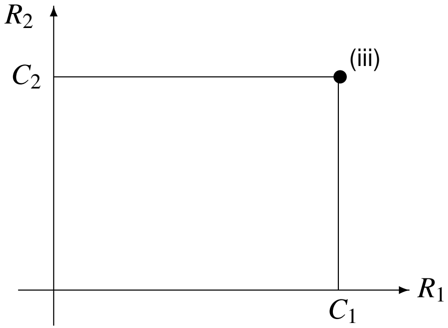
$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \left\{ (L_1, L_2) : L_1 \le \sqrt{V_1}\Phi^{-1}(\varepsilon) \right\}.$$

Main Result: Corner Point



- In case (iii), we are operating near both boundaries

Main Result: Corner Point



- In case (iii), we are operating near both boundaries
- So both constraints are active and we see L_1 and L_2 in the optimum second-order coding rate region is

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \left\{ (L_1, L_2) : \Phi\left(-\frac{L_1}{\sqrt{V_1}} \right) \Phi\left(-\frac{L_2}{\sqrt{V_2}} \right) \geq 1 - \varepsilon \right\}.$$

Heuristic Derivation of Case (iii)

- Let $\mathcal{G}_j := \{\hat{M}_j = \hat{M}\}$ be the event that message $j = 1, 2$ is decoded correctly

$$\Pr(\mathcal{G}_1 \cap \mathcal{G}_2) \geq 1 - \varepsilon$$

Heuristic Derivation of Case (iii)

- Let $\mathcal{G}_j := \{\hat{M}_j = \hat{M}\}$ be the event that message $j = 1, 2$ is decoded correctly

$$\Pr(\mathcal{G}_1 \cap \mathcal{G}_2) \ge 1 - \varepsilon$$

- Assuming independence (which does not hold generally),

$$\Pr(\mathcal{G}_1) \Pr(\mathcal{G}_2) \ge 1 - \varepsilon$$

Heuristic Derivation of Case (iii)

- Let $\mathcal{G}_j := \{\hat{M}_j = \hat{M}\}$ be the event that message $j = 1, 2$ is decoded correctly

$$\Pr(\mathcal{G}_1 \cap \mathcal{G}_2) \geq 1 - \varepsilon$$

- Assuming independence (which does not hold generally),

$$\Pr(\mathcal{G}_1) \Pr(\mathcal{G}_2) \geq 1 - \varepsilon$$

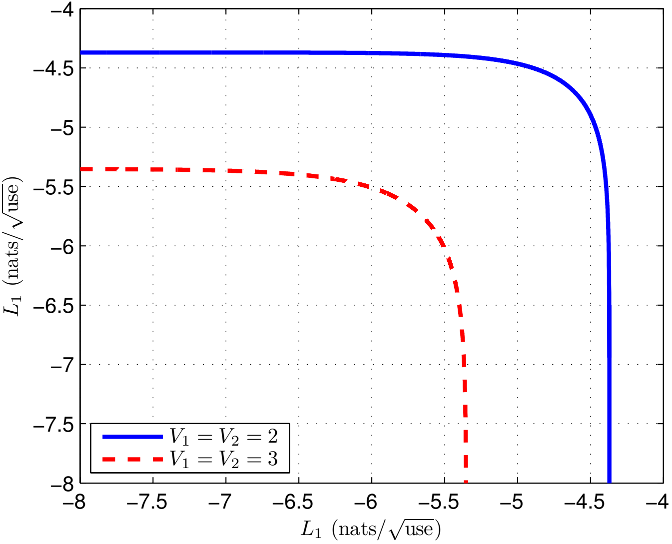
- But we know from the single-user result that the error probability

$$\Pr(\mathcal{G}_j^c) \approx \Phi\Big(\frac{L_j}{\sqrt{V_j}}\Big) \quad \Longrightarrow \quad \Pr(\mathcal{G}_j) \approx \Phi\Big(-\frac{L_j}{\sqrt{V_j}}\Big).$$

  So the set of all (ε, C_1, C_2)-achievable second-order coding rates is

$$\Phi\Big(-\frac{L_1}{\sqrt{V_1}}\Big)\Phi\Big(-\frac{L_2}{\sqrt{V_2}}\Big) \geq 1 - \varepsilon$$

Illustration of Case (iii)



Intuition Gleaned from GIC with SVSI

- Carleial (1975) mentioned that

  "Very strong interference is as innocuous as no interference at all"

Intuition Gleaned from GIC with SVSI

- Carleial (1975) mentioned that

  "Very strong interference is as innocuous as no interference at all"

- We show that SVSI is innocuous in the sense that the capacities C_j and dispersions V_j are not affected

Intuition Gleaned from GIC with SVSI

- Carleial (1975) mentioned that

  "Very strong interference is as innocuous as no interference at all"

- We show that SVSI is innocuous in the sense that the capacities C_j and dispersions V_j are not affected

- Error events are approximately independent

Intuition Gleaned from GIC with SVSI

- Carleial (1975) mentioned that

  "Very strong interference is as innocuous as no interference at all"

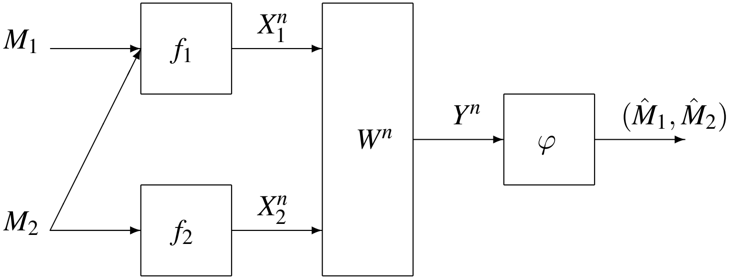
- We show that SVSI is innocuous in the sense that the capacities C_j and dispersions V_j are not affected

- Error events are approximately independent

- Intuition that the GIC with SVSI is analogous to 2 independent direct channels carries over, even in the finer second-order sense

Outline

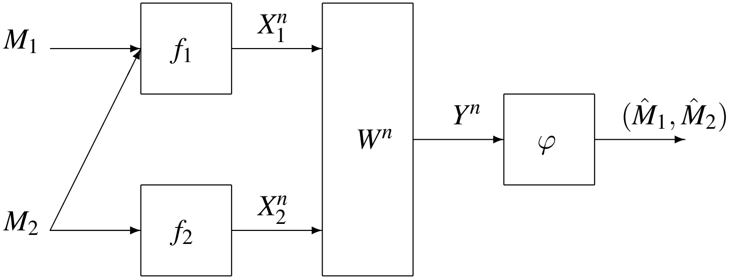
Gaussian MAC with Degraded Message Sets



- Channel law is

$$Y_i = X_{1i} + X_{2i} + Z_i$$

Gaussian MAC with Degraded Message Sets



- Channel law is

$$Y_i = X_{1i} + X_{2i} + Z_i$$

- Channel inputs are power limited

$$\sum_{i=1}^{n} X_{ji}^2 \leq nS_j, \qquad j = 1, 2$$

Capacity Region

- Capacity region is an exercise in NIT

Capacity Region

- Capacity region is an exercise in NIT
- The capacity region is the set of all (R_1, R_2) such that

$$R_1 \le \mathrm{C}(S_1(1-\rho^2))$$

$$R_1 + R_2 \le \mathrm{C}(S_1 + S_2 + 2\rho\sqrt{S_1 S_2})$$

for some $0 \le \rho \le 1$.

Capacity Region

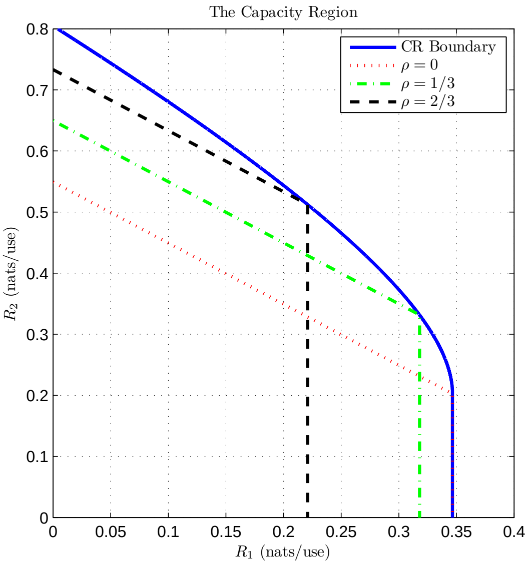
- Capacity region is an exercise in NIT
- The capacity region is the set of all (R_1, R_2) such that

$$R_1 \le \mathrm{C}(S_1(1-\rho^2))$$
$$R_1 + R_2 \le \mathrm{C}(S_1 + S_2 + 2\rho\sqrt{S_1 S_2})$$

  for some $0 \le \rho \le 1$.

- Uses superposition coding (Cover (1972))

Capacity Region



Characterize $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$

- This is joint work with

Jon Scarlett

- We want to characterize $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$ for all points (R_1^*, R_2^*) on the boundary of the capacity region

Characterize $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$

- This is joint work with

Jon Scarlett

- We want to characterize $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$ for all points (R_1^*, R_2^*) on the boundary of the capacity region
- For all (R_1^*, R_2^*) in the interior of the capacity region

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \mathbb{R}^2$$

- For all (R_1^*, R_2^*) in the exterior of the capacity region

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \emptyset$$

implying the strong converse

Some Basic Definitions

- Mutual informations

$$\mathbf{I}(\rho) := \begin{bmatrix} I_1(\rho) \\ I_{12}(\rho) \end{bmatrix} = \begin{bmatrix} \mathrm{C}\big((1-\rho^2)S_1\big) \\ \mathrm{C}\big(S_1 + S_2 + 2\rho\sqrt{S_1 S_2}\big) \end{bmatrix}$$

Some Basic Definitions

- Mutual informations

$$\mathbf{I}(\rho) := \begin{bmatrix} I_1(\rho) \\ I_{12}(\rho) \end{bmatrix} = \begin{bmatrix} \mathsf{C}\big((1-\rho^2)S_1\big) \\ \mathsf{C}\big(S_1 + S_2 + 2\rho\sqrt{S_1 S_2}\big) \end{bmatrix}$$

- Derivative of mutual informations

$$\mathbf{D}(\rho) := \frac{\partial}{\partial \rho}\mathbf{I}(\rho) = \begin{bmatrix} \frac{-S_1\rho}{1+S_1(1-\rho^2)} \\ \frac{\sqrt{S_1 S_2}}{1+S_1+S_2+2\rho\sqrt{S_1 S_2}} \end{bmatrix}$$

Some Basic Definitions

- Mutual informations

$$\mathbf{I}(\rho) := \begin{bmatrix} I_1(\rho) \\ I_{12}(\rho) \end{bmatrix} = \begin{bmatrix} \mathrm{C}\big((1-\rho^2)S_1\big) \\ \mathrm{C}\big(S_1 + S_2 + 2\rho\sqrt{S_1 S_2}\big) \end{bmatrix}$$

- Derivative of mutual informations

$$\mathbf{D}(\rho) := \frac{\partial}{\partial\rho}\mathbf{I}(\rho) = \begin{bmatrix} \frac{-S_1\rho}{1+S_1(1-\rho^2)} \\ \frac{\sqrt{S_1S_2}}{1+S_1+S_2+2\rho\sqrt{S_1S_2}} \end{bmatrix}$$

- Dispersions $\mathrm{V}(x,y) := \frac{x(y+2)}{2(x+1)(y+1)}$ and $\mathrm{V}(x) := \mathrm{V}(x,x)$

$$\mathbf{V}(\rho) := \begin{bmatrix} V_1(\rho) & V_{1,12}(\rho) \\ V_{1,12}(\rho) & V_{12,12}(\rho) \end{bmatrix}$$

where

$$V_1(\rho) := \mathrm{V}\big((1-\rho^2)S_1\big), \quad V_{12,12}(\rho) := \mathrm{V}\big(S_1 + S_2 + 2\rho\sqrt{S_1S_2}\big)$$

$$V_{1,12}(\rho) := \mathrm{V}\big((1-\rho^2)S_1, S_1 + S_2 + 2\rho\sqrt{S_1S_2}\big)$$

Generalization of Inverse CDF of a Gaussian

- For a positive semi-definite matrix $\mathbf{V}$,

$$\Psi(z_1, z_2, \mathbf{V}) = \int_{-\infty}^{z_1} \int_{-\infty}^{z_2} \mathcal{N}(\mathbf{0}, \mathbf{V}) \, \mathrm{d}\mathbf{u}$$

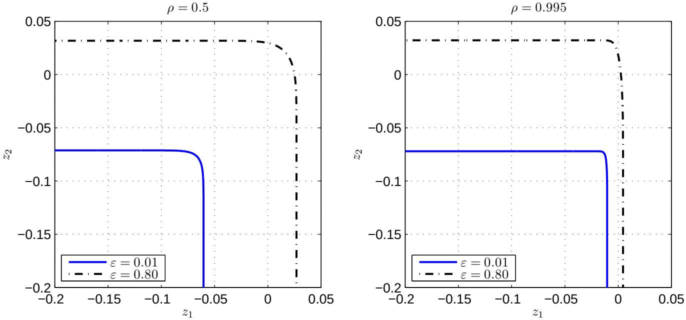
Generalization of Inverse CDF of a Gaussian

- For a positive semi-definite matrix $\mathbf{V}$,

$$\Psi(z_1, z_2, \mathbf{V}) = \int_{-\infty}^{z_1} \int_{-\infty}^{z_2} \mathcal{N}(\mathbf{0}, \mathbf{V}) \, \mathrm{d}\mathbf{u}$$

- Given $\varepsilon \in (0, 1)$,

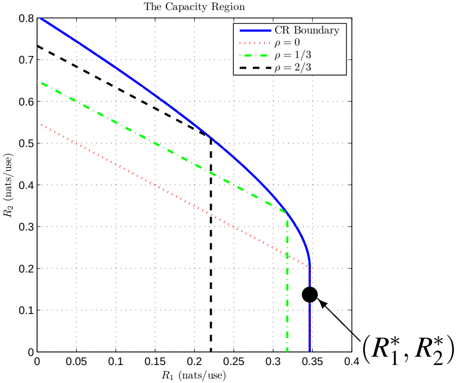
$$\Psi^{-1}(\mathbf{V}, \varepsilon) = \{(z_1, z_2) : \Psi(-z_1, -z_2, \mathbf{V}) \geq 1 - \varepsilon\} .$$



The Main Result: Vertical Boundary

Points on vertical boundary reduce to scalar dispersion as sum rate constraint is in error exponents regime

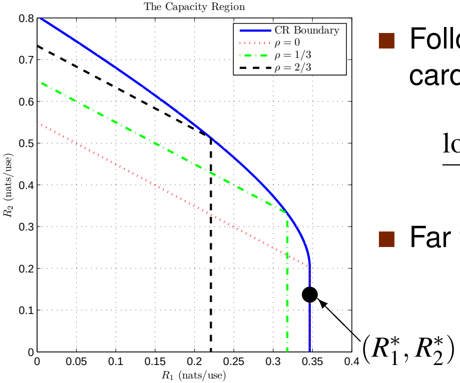
$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \{(L_1, L_2) : L_1 \le \sqrt{V_1(0)}\Phi^{-1}(\varepsilon)\}$$



The Main Result: Vertical Boundary

Points on vertical boundary reduce to scalar dispersion as sum rate constraint is in error exponents regime

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \{(L_1, L_2) : L_1 \le \sqrt{V_1(0)}\Phi^{-1}(\varepsilon)\}$$



- Following expansion holds for cardinality of first codebook

$$\frac{\log M_{1n}}{n} \approx I_1(0) + \sqrt{\frac{V_1(0)}{n}}\Phi^{-1}(\varepsilon)$$

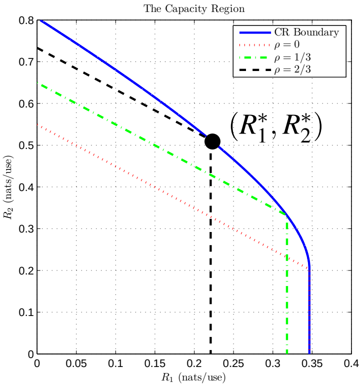
- Far from sum rate constraint

$$\lim_{n\to\infty} \frac{1}{n} \log(M_{1n}M_{2n}) < I_{12}(0)$$

The Main Result: Curved Boundary

Different behavior in the curved region

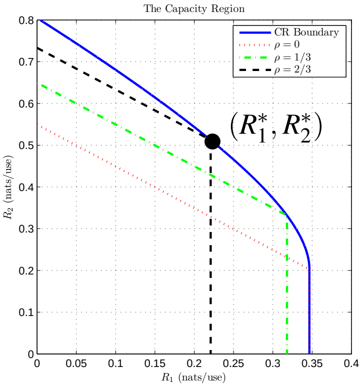
$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \left\{ (L_1, L_2) : \begin{bmatrix} L_1 \\ L_1 + L_2 \end{bmatrix} \in \bigcup_{\beta \in \mathbb{R}} \beta \mathbf{D}(\rho) + \Psi^{-1}(\mathbf{V}(\rho), \varepsilon) \right\}$$



The Main Result: Curved Boundary

Different behavior in the curved region

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \left\{ (L_1, L_2) : \begin{bmatrix} L_1 \\ L_1 + L_2 \end{bmatrix} \in \bigcup_{\beta \in \mathbb{R}} \beta \mathbf{D}(\rho) + \Psi^{-1}(\mathbf{V}(\rho), \varepsilon) \right\}$$

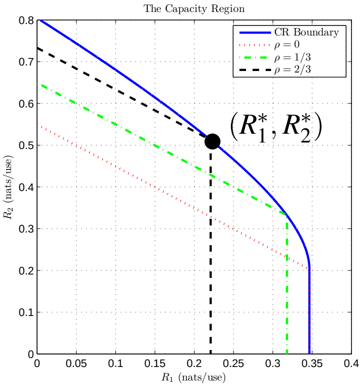


- $\mathbf{D}(\rho)$ doesn't appear usually

The Main Result: Curved Boundary

Different behavior in the curved region

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \left\{ (L_1, L_2) : \begin{bmatrix} L_1 \\ L_1 + L_2 \end{bmatrix} \in \bigcup_{\beta \in \mathbb{R}} \beta \mathbf{D}(\rho) + \Psi^{-1}(\mathbf{V}(\rho), \varepsilon) \right\}$$



- $\mathbf{D}(\rho)$ doesn't appear usually
- $\Psi^{-1}(\mathbf{V}(\rho), \varepsilon)$: corresponds using Gaussian with covariance matrix

$$\mathbf{\Sigma}(\rho) = \begin{bmatrix} S_1 & \rho\sqrt{S_1 S_2} \\ \rho\sqrt{S_1 S_2} & S_2 \end{bmatrix}$$

Achieving all Second-Order Rate Pairs

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \left\{ (L_1, L_2) : \begin{bmatrix} L_1 \\ L_1 + L_2 \end{bmatrix} \in \bigcup_{\beta \in \mathbb{R}} \beta \mathbf{D}(\rho) + \Psi^{-1}(\mathbf{V}(\rho), \varepsilon) \right\}$$

Achieving all Second-Order Rate Pairs

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \left\{ (L_1, L_2) : \begin{bmatrix} L_1 \\ L_1 + L_2 \end{bmatrix} \in \bigcup_{\beta \in \mathbb{R}} \beta \mathbf{D}(\rho) + \Psi^{-1}(\mathbf{V}(\rho), \varepsilon) \right\}$$

- Non-empty regions in CR not in trapezium achievable by

$$\mathcal{N}\left(\mathbf{0}, \begin{bmatrix} S_1 & \rho\sqrt{S_1 S_2} \\ \rho\sqrt{S_1 S_2} & S_2 \end{bmatrix} \right)$$

Achieving all Second-Order Rate Pairs

$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \left\{ (L_1, L_2) : \begin{bmatrix} L_1 \\ L_1 + L_2 \end{bmatrix} \in \bigcup_{\beta \in \mathbb{R}} \beta \mathbf{D}(\rho) + \Psi^{-1}(\mathbf{V}(\rho), \varepsilon) \right\}$$

- Non-empty regions in CR not in trapezium achievable by

$$\mathcal{N}\left(\mathbf{0}, \begin{bmatrix} S_1 & \rho\sqrt{S_1 S_2} \\ \rho\sqrt{S_1 S_2} & S_2 \end{bmatrix} \right)$$

- Use the above distribution dependent on blocklength:

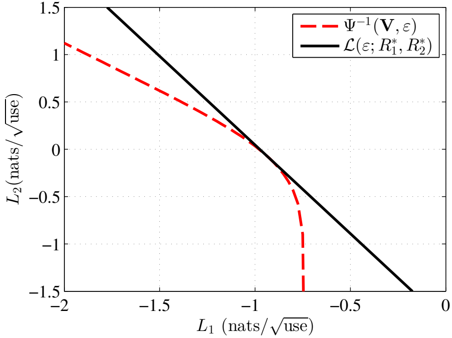
$$\rho_n = \rho + \frac{\beta}{\sqrt{n}}$$

- By a Taylor expansion,

$$\mathbf{I}(\rho_n) \approx \mathbf{I}(\rho) + (\rho_n - \rho)\mathbf{D}(\rho) = \mathbf{I}(\rho) + \frac{\beta \mathbf{D}(\rho)}{\sqrt{n}}$$

explaining the slope term.

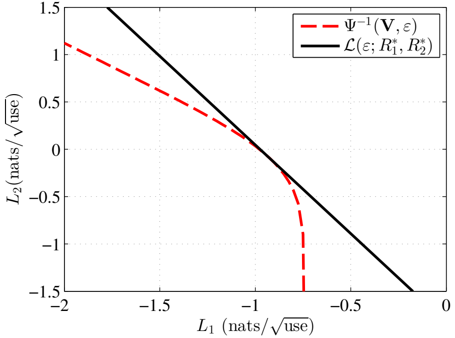
Illustration of Second-Order Coding Rates



$S_1 = S_2 = 1$ and $\rho = \frac{1}{2}$

$\mathcal{L}(\varepsilon; R_1^*, R_2^*)$ is a half-space

Illustration of Second-Order Coding Rates

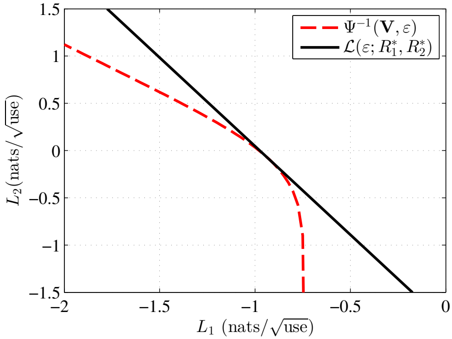


$S_1 = S_2 = 1$ and $\rho = \frac{1}{2}$

$\mathcal{L}(\varepsilon; R_1^*, R_2^*)$ is a half-space

- Second-order rates achieved using a single input distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(\rho))$ is not optimal

Illustration of Second-Order Coding Rates



$S_1 = S_2 = 1$ and $\rho = \frac{1}{2}$

$\mathcal{L}(\varepsilon; R_1^*, R_2^*)$ is a half-space

- Second-order rates achieved using a single input distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(\rho))$ is not optimal
- Need to vary input distribution with blocklength to achieve all points in $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$

Outline

Conclusion

- Asymptotic expansions for NIT problems with non-vanishing error probabilities is a very fertile area of research

Conclusion

- Asymptotic expansions for NIT problems with non-vanishing error probabilities is a very fertile area of research
- Other single-user that are solved include:
  1. Quasi-static MIMO fading channels (Yang-Durisi-Koch-Polyanskiy)
  2. Channels with discrete state (Tomamichel-Tan)
  3. Dirty-paper coding (Scarlett)

Conclusion

- Asymptotic expansions for NIT problems with non-vanishing error probabilities is a very fertile area of research
- Other single-user that are solved include:
  1. Quasi-static MIMO fading channels (Yang-Durisi-Koch-Polyanskiy)
  2. Channels with discrete state (Tomamichel-Tan)
  3. Dirty-paper coding (Scarlett)
- Shameless self-promotion:

V. Y. F. Tan
Asymptotic expansions in IT with non-vanishing error probabilities
Now Publishers
Foundations and Trends in Comms and Inf. Th.