Second-Order Asymptotics for the Gaussian MAC with Degraded Message Sets

Jonathan Scarlett[†] and **Vincent Y. F. Tan**[‡]

[†] Department of Engineering, University of Cambridge

[‡] Electrical and Computer Engineering, National University of Singapore

Information Theory and Applications Workshop (Feb 2013)

Full version: `http://arxiv.org/abs/1310.1197`

Second-Order Asymptotics for Point-to-Point Channels

- Second-order asymptotics for the DMC and the AWGN channel are well-understood

Second-Order Asymptotics for Point-to-Point Channels

- Second-order asymptotics for the DMC and the AWGN channel are well-understood
- $M^*(n, \varepsilon, S)$: maximum codebook size for n uses of the AWGN channel with SNR S and average error probability ε

Second-Order Asymptotics for Point-to-Point Channels

- Second-order asymptotics for the DMC and the AWGN channel are well-understood
- $M^*(n, \varepsilon, S)$: maximum codebook size for n uses of the AWGN channel with SNR S and average error probability ε
- Hayashi (2009) and Polyanskiy-Poor-Verdú (2010) showed that

$$\log M^*(n, \varepsilon, S) = n\mathsf{C}(S) + \sqrt{n\mathsf{V}(S)}\Phi^{-1}(\varepsilon) + o(\sqrt{n}) \quad \text{nats},$$

where the Gaussian capacity and Gaussian dispersion functions are defined as

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

Second-Order Asymptotics for Point-to-Point Channels

- Second-order asymptotics for the DMC and the AWGN channel are well-understood
- $M^*(n, \varepsilon, S)$: maximum codebook size for n uses of the AWGN channel with SNR S and average error probability ε
- Hayashi (2009) and Polyanskiy-Poor-Verdú (2010) showed that

$$\log M^*(n, \varepsilon, S) = n\mathsf{C}(S) + \sqrt{n\mathsf{V}(S)}\Phi^{-1}(\varepsilon) + o(\sqrt{n}) \quad \text{nats},$$

where the Gaussian capacity and Gaussian dispersion functions are defined as

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

- Second-order coding rate $=$ Largest coefficient of $\sqrt{n}$ term $= \sqrt{\mathsf{V}(S)}\Phi^{-1}(\varepsilon)$.

Second-Order Asymptotics in Networks

- Second-order asymptotics in networks has been more modest

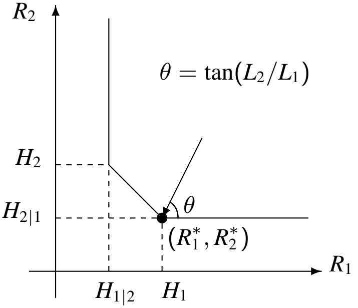
Second-Order Asymptotics in Networks

- Second-order asymptotics in networks has been more modest
- Tan-Kosut (2014) and Nomura-Han (2012)–Slepian-Wolf problem

Second-Order Asymptotics in Networks

- Second-order asymptotics in networks has been more modest
- Tan-Kosut (2014) and Nomura-Han (2012)–Slepian-Wolf problem
- Let $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$ be the set of all $(L_1, L_2) \in \mathbb{R}^2$ such that there exists length-n codes of sizes $(M_{1,n}, M_{2,n})$ and errors ε_n such that

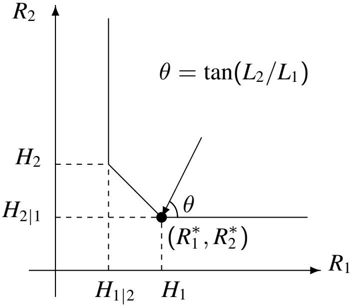
$$\limsup_{n\to\infty} \frac{1}{\sqrt{n}}(\log M_{j,n} - nR_j^*) \le L_j, \quad j = 1, 2, \quad \limsup_{n\to\infty} \varepsilon_n \le \varepsilon.$$



Second-Order Asymptotics in Networks

- Second-order asymptotics in networks has been more modest
- Tan-Kosut (2014) and Nomura-Han (2012)–Slepian-Wolf problem
- Let $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$ be the set of all $(L_1, L_2) \in \mathbb{R}^2$ such that there exists length-n codes of sizes $(M_{1,n}, M_{2,n})$ and errors ε_n such that

$$\limsup_{n\to\infty} \frac{1}{\sqrt{n}}(\log M_{j,n} - nR_j^*) \le L_j, \quad j = 1,2, \quad \limsup_{n\to\infty} \varepsilon_n \le \varepsilon.$$

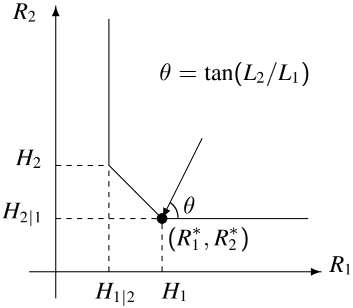


$$(L_1, L_2) \in \mathcal{L}(\varepsilon; R_1^*, R_2^*)$$

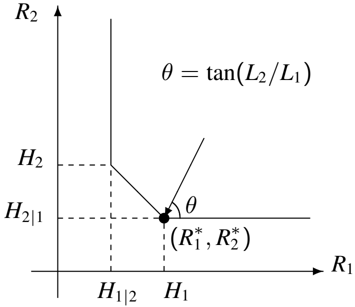
implies exists ε-reliable codes with

$$\log M_{j,n} \le nR_j^* + \sqrt{n}\, L_j + o(\sqrt{n})$$

Second-Order Asymptotics in Networks

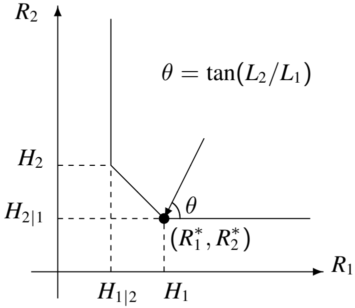


Second-Order Asymptotics in Networks



$$\mathcal{L}(\varepsilon; H_1, H_{2|1}) = \{(L_1, L_2) : \Psi(L_2, L_1 + L_2; \mathbf{V}_{2,12}) \geq 1 - \varepsilon\} .$$

where



$$\mathcal{L}(\varepsilon; H_1, H_{2|1}) = \{(L_1, L_2) : \Psi(L_2, L_1 + L_2; \mathbf{V}_{2,12}) \geq 1 - \varepsilon\} .$$

where

$$\Psi(z_2, z_3; \mathbf{V}) := \int_{-\infty}^{z_2} \int_{-\infty}^{z_3} \mathcal{N}(\mathbf{0}, \mathbf{V}) \, \mathrm{d}\mathbf{u}, \quad \text{and}$$

$$\mathbf{V}_{2,12} := \mathrm{Cov}\left(\left[-\log p_{X_2|X_1} \quad -\log p_{X_1 X_2} \right]' \right)$$

Second-Order Asymptotics in the MAC

- The “next-easiest” network problem is the MAC

Second-Order Asymptotics in the MAC

- The "next-easiest" network problem is the MAC
- Numerous attempts to characterize second-order asymptotics for the MAC including Tan-Kosut (2014), Huang-Moulin (2012), MolavianJazi-Laneman (2012), Haim-Erez-Kochman (2012), Scarlett-Martinez-Guillén i Fàbregas (2013) etc.

Second-Order Asymptotics in the MAC

- The “next-easiest” network problem is the MAC
- Numerous attempts to characterize second-order asymptotics for the MAC including Tan-Kosut (2014), Huang-Moulin (2012), MolavianJazi-Laneman (2012), Haim-Erez-Kochman (2012), Scarlett-Martinez-Guillén i Fàbregas (2013) etc.
- No tight local-dispersion characterization yet for the DM-MAC or Gaussian MAC

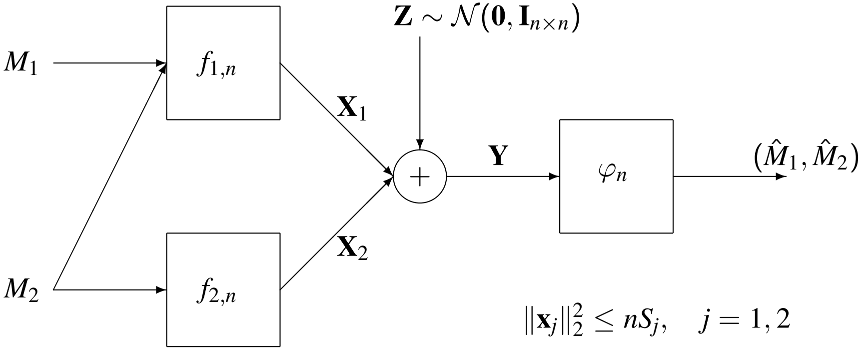
Second-Order Asymptotics in the MAC

- The “next-easiest” network problem is the MAC
- Numerous attempts to characterize second-order asymptotics for the MAC including Tan-Kosut (2014), Huang-Moulin (2012), MolavianJazi-Laneman (2012), Haim-Erez-Kochman (2012), Scarlett-Martinez-Guillén i Fàbregas (2013) etc.
- No tight local-dispersion characterization yet for the DM-MAC or Gaussian MAC
- Some problems (primarily the converse):
  1. Difficulty in characterizing the boundary of the CR
  2. Independent input distributions

Second-Order Asymptotics in the MAC

- The “next-easiest” network problem is the MAC
- Numerous attempts to characterize second-order asymptotics for the MAC including Tan-Kosut (2014), Huang-Moulin (2012), MolavianJazi-Laneman (2012), Haim-Erez-Kochman (2012), Scarlett-Martinez-Guillén i Fàbregas (2013) etc.
- No tight local-dispersion characterization yet for the DM-MAC or Gaussian MAC
- Some problems (primarily the converse):
  1. Difficulty in characterizing the boundary of the CR
  2. Independent input distributions
- We consider a MAC-like model that retains the main characteristics of MAC but skirts the problems above

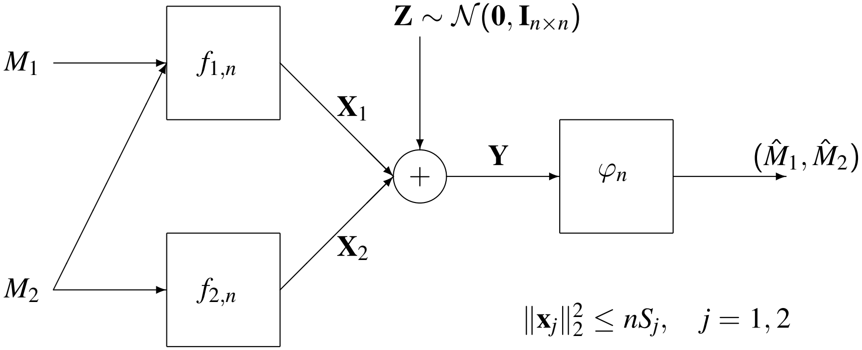
Gaussian MAC with Degraded Message Sets



$$\|\mathbf{x}_j\|_2^2 \le nS_j, \quad j = 1, 2$$

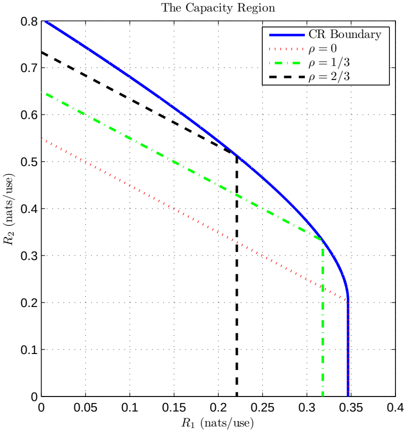
- Encoder 1 has access to both messages

Gaussian MAC with Degraded Message Sets



- Encoder 1 has access to both messages
- Capacity region is well known [Exercise 5.18(b), El Gamal and Kim (2012)]; achieved using superposition coding

Gaussian MAC with Degraded Message Sets



$$R_1 \leq \mathsf{C}\big((1-\rho^2)S_1\big)$$

$$R_1 + R_2 \leq \mathsf{C}\big(S_1 + S_2 + 2\rho\sqrt{S_1 S_2}\big)$$

$$\mathsf{C}(x) = \frac{1}{2}\log(1+x)$$

$\rho \in [0,1]$ parametrizes curved boundary and indicates the amount of correlation between users' codewords.

Second-Order Asymptotics for Gaussian MAC w/ DMS

- (R_1^*, R_2^*): arbitrary point on the boundary of the capacity region

Second-Order Asymptotics for Gaussian MAC w/ DMS

- (R_1^*, R_2^*): arbitrary point on the boundary of the capacity region
- $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$: set of all $(L_1, L_2) \in \mathbb{R}^2$ such that there exists length-n codes of sizes $(M_{1,n}, M_{2,n})$ and average errors ε_n such that

$$\liminf_{n\to\infty} \frac{1}{\sqrt{n}}(\log M_{j,n} - nR_j^*) \geq L_j, \quad j = 1, 2, \quad \limsup_{n\to\infty} \varepsilon_n \leq \varepsilon$$

and $\|\mathbf{x}_j\|_2^2 \leq nS_j$, for $j = 1, 2$

Second-Order Asymptotics for Gaussian MAC w/ DMS

- (R_1^*, R_2^*): arbitrary point on the boundary of the capacity region
- $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$: set of all $(L_1, L_2) \in \mathbb{R}^2$ such that there exists length-n codes of sizes $(M_{1,n}, M_{2,n})$ and average errors ε_n such that

$$\liminf_{n\to\infty} \frac{1}{\sqrt{n}}(\log M_{j,n} - nR_j^*) \geq L_j, \quad j = 1, 2, \quad \limsup_{n\to\infty} \varepsilon_n \leq \varepsilon$$

  and $\|\mathbf{x}_j\|_2^2 \leq nS_j$, for $j = 1, 2$

- If $(L_1, L_2) \in \mathcal{L}(\varepsilon; R_1^*, R_2^*)$, there exists ε-reliable codes s.t.

$$\log M_{j,n} \geq nR_j^* + \sqrt{n}\, L_j + o(\sqrt{n}), \quad j = 1, 2.$$

Second-Order Asymptotics for Gaussian MAC w/ DMS

- (R_1^*, R_2^*): arbitrary point on the boundary of the capacity region
- $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$: set of all $(L_1, L_2) \in \mathbb{R}^2$ such that there exists length-n codes of sizes $(M_{1,n}, M_{2,n})$ and average errors ε_n such that

$$\liminf_{n\to\infty} \frac{1}{\sqrt{n}}(\log M_{j,n} - nR_j^*) \geq L_j, \quad j = 1,2, \quad \limsup_{n\to\infty} \varepsilon_n \leq \varepsilon$$

and $\|\mathbf{x}_j\|_2^2 \leq nS_j$, for $j = 1,2$

- If $(L_1, L_2) \in \mathcal{L}(\varepsilon; R_1^*, R_2^*)$, there exists ε-reliable codes s.t.

$$\log M_{j,n} \geq nR_j^* + \sqrt{n}\, L_j + o(\sqrt{n}), \quad j = 1,2.$$

- Main Contribution: A complete characterization of $\mathcal{L}(\varepsilon; R_1^*, R_2^*)$
- First complete characterization of second-order asymptotics for a channel-type network information theory problem

Some Basic Definitions

- Mutual informations

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

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

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

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

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

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

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

where

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

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

Generalization of Inverse CDF of a Gaussian

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

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

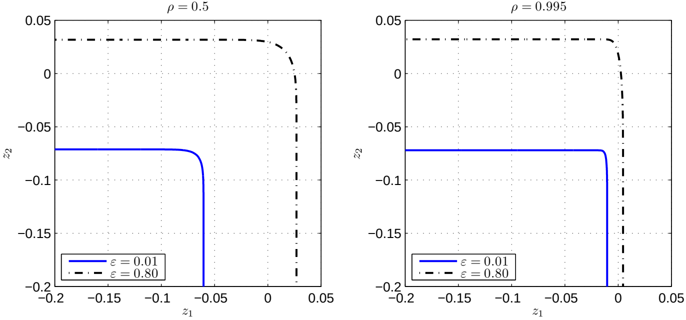
Generalization of Inverse CDF of a Gaussian

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

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

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

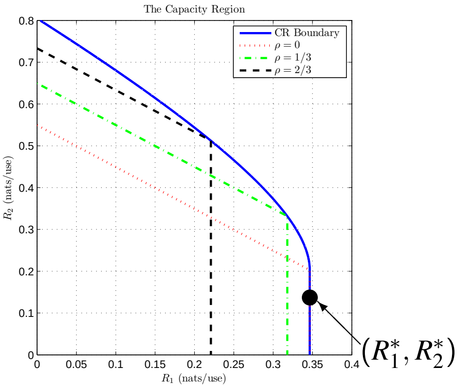
$$\Psi^{-1}(\mathbf{V}, \varepsilon) = \{(z_1, z_2) : \Psi(-z_1, -z_2, \mathbf{V}) \geq 1 - \varepsilon\} .$$



The Main Result: Vertical Boundary

Points on vertical boundary reduce to scalar dispersion as sum rate constraint is in error exponents regime [Haim-Erez-Kochman (2012)]

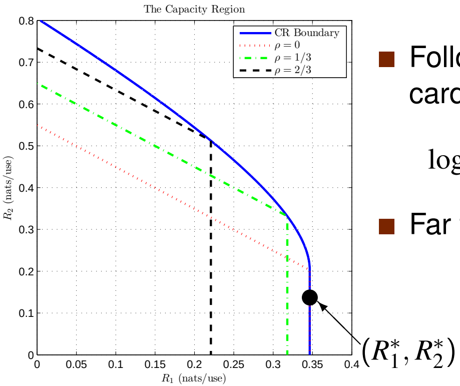
$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \{(L_1, L_2) : L_1 \le \sqrt{V_1(0)}\Phi^{-1}(\varepsilon)\}$$



The Main Result: Vertical Boundary

Points on vertical boundary reduce to scalar dispersion as sum rate constraint is in error exponents regime [Haim-Erez-Kochman (2012)]

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



- Following expansion holds for cardinality of first codebook

$$\log M_{1,n} \approx nI_1(0) + \sqrt{nV_1(0)}\Phi^{-1}(\varepsilon)$$

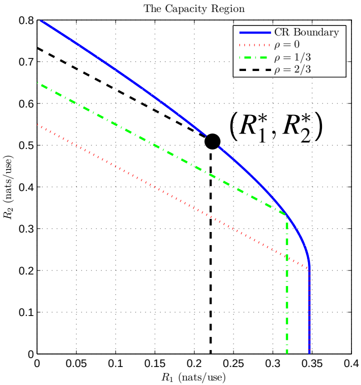
- Far from sum rate constraint

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

The Main Result: Curved Boundary

Radically different behavior in the curved region

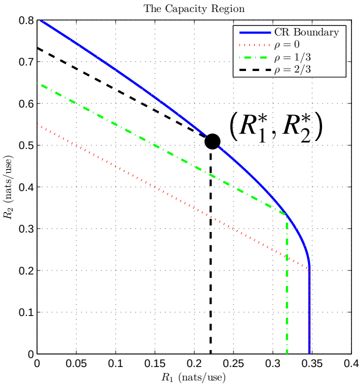
$$\mathcal{L}(\varepsilon; R_1^*, R_2^*) = \left\{ (L_1, L_2) : \begin{bmatrix} L_1 \\ L_1 + L_2 \end{bmatrix} \in \bigcup_{\beta \in \mathbb{R}} \beta \mathbf{D}(\rho) + \Psi^{-1}(\mathbf{V}(\rho), \varepsilon) \right\}$$



The Main Result: Curved Boundary

Radically different behavior in the curved region

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

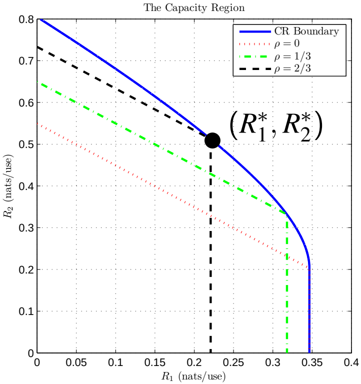


- $\mathbf{D}(\rho)$ doesn't appear for SW

The Main Result: Curved Boundary

Radically different behavior in the curved region

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



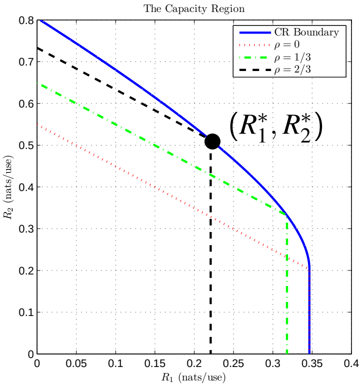
- $\mathbf{D}(\rho)$ doesn't appear for SW
- $\Psi^{-1}(\mathbf{V}(\rho), \varepsilon)$: corresponds to only using $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(\rho))$ where

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

The Main Result: Curved Boundary

Radically different behavior in the curved region

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



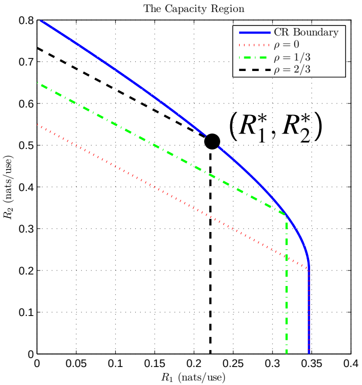
- $\mathbf{D}(\rho)$ doesn't appear for SW
- $\Psi^{-1}(\mathbf{V}(\rho), \varepsilon)$: corresponds to only using $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(\rho))$ where

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

- Non-empty regions in CR not in trapezium achievable by $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(\rho))$

Radically different behavior in the curved region

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

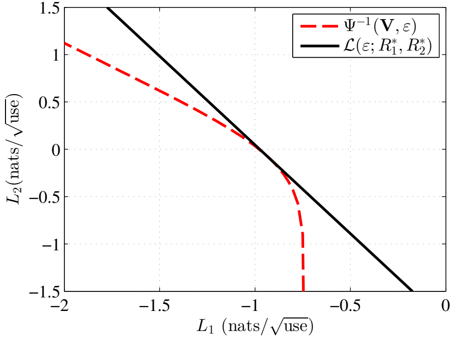


- $\mathbf{D}(\rho)$ doesn't appear for SW
- $\Psi^{-1}(\mathbf{V}(\rho), \varepsilon)$: corresponds to only using $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(\rho))$ where

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

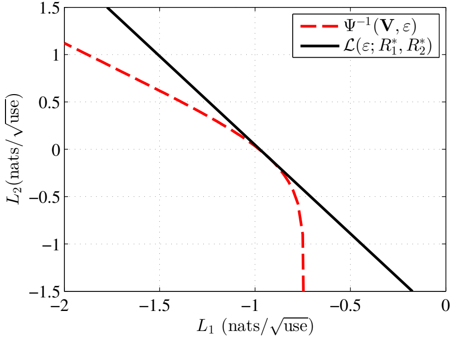
- Non-empty regions in CR not in trapezium achievable by $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(\rho))$
- Use $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(\rho_n))$ and $\rho_n = \rho + \beta/\sqrt{n}$

Illustration of Second-Order Coding Rates



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

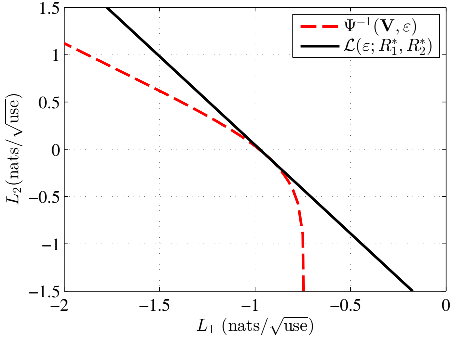
Illustration of Second-Order Coding Rates



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

- 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 $\rho = \frac{1}{2}$

- Second-order rates achieved using a single input distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}(\rho))$ is not optimal
- The optimal second-order coding rate region is a half-space

Main Ideas in Converse Proof: Part I

- By a standard $n \to n+1$ argument [Shannon (1959)], we may consider codes with equal power constraints

$$\|\mathbf{x}_j\|_2^2 = nS_j, \qquad j = 1, 2.$$

Main Ideas in Converse Proof: Part I

- By a standard $n \to n+1$ argument [Shannon (1959)], we may consider codes with equal power constraints

$$\|\mathbf{x}_j\|_2^2 = nS_j, \qquad j = 1, 2.$$

- Reduction from average to maximal error probability via expurgation of polynomially many codeword pairs

  –Not possible for standard MAC [Dueck (1978)]

Main Ideas in Converse Proof: Part I

- By a standard $n \to n+1$ argument [Shannon (1959)], we may consider codes with equal power constraints

$$\|\mathbf{x}_j\|_2^2 = nS_j, \qquad j = 1, 2.$$

- Reduction from average to maximal error probability via expurgation of polynomially many codeword pairs

  –Not possible for standard MAC [Dueck (1978)]

- Reduction to constant correlation type classes

$$\mathcal{T}_n(k) = \left\{ (\mathbf{x}_1, \mathbf{x}_2) : \frac{\langle \mathbf{x}_1, \mathbf{x}_2 \rangle}{\|\mathbf{x}_1\|_2 \|\mathbf{x}_2\|_2} \in \left(\frac{k-1}{n}, \frac{k}{n} \right] \right\}, \quad k = 1, 2, \ldots, n.$$

  Without too much loss in rate

Main Ideas in Converse Proof: Part II

- Verdú-Han-type converse: For any $\gamma > 0$ and any $(Q_{\mathbf{Y}|\mathbf{X}_2}, Q_{\mathbf{Y}})$, have the following non-asymptotic converse bound

$$\varepsilon_n \geq \Pr\Bigg(j_1(\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}) \leq \frac{1}{n}\log M_{1,n} - \gamma \quad \text{or} \quad j_{12}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}) \leq \frac{1}{n}\log(M_{1,n}M_{2,n}) - \gamma \Bigg) - 2e^{-n\gamma}$$

where $j_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \frac{1}{n}\log\frac{W^n(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)}{Q_{\mathbf{Y}|\mathbf{X}_2}(\mathbf{y}|\mathbf{x}_2)}$ and $j_{12}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \frac{1}{n}\log\frac{W^n(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)}{Q_{\mathbf{Y}}(\mathbf{y})}$.

Main Ideas in Converse Proof: Part II

- Verdú-Han-type converse: For any $\gamma > 0$ and any $(Q_{\mathbf{Y}|\mathbf{X}_2}, Q_{\mathbf{Y}})$, have the following non-asymptotic converse bound

$$\varepsilon_n \geq \Pr\Bigg(j_1(\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}) \leq \frac{1}{n} \log M_{1,n} - \gamma \quad \text{or} \quad j_{12}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}) \leq \frac{1}{n} \log(M_{1,n} M_{2,n}) - \gamma \Bigg) - 2e^{-n\gamma}$$

where $j_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \frac{1}{n} \log \frac{W^n(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2)}{Q_{\mathbf{Y}|\mathbf{X}_2}(\mathbf{y}|\mathbf{x}_2)}$ and $j_{12}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \frac{1}{n} \log \frac{W^n(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2)}{Q_{\mathbf{Y}}(\mathbf{y})}$.

- Let $\mathbf{j} = [j_1, j_{12}]^T$. Choose $(Q_{\mathbf{Y}|\mathbf{X}_2}, Q_{\mathbf{Y}})$ and for $(\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{T}_n(k)$,

$$\mathbb{E}\big[\mathbf{j}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{Y})\big] \approx \mathbf{I}(\rho) \quad \text{and} \quad \mathrm{Cov}\big[\sqrt{n}\,\mathbf{j}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{Y})\big] \approx \mathbf{V}(\rho)$$

where

$$\frac{k-1}{n} \leq \rho \leq \frac{k}{n}$$

Main Ideas in Converse Proof: Part III

- By evaluating Verdú-Han using multivariate Berry-Esseen, there exists a sequence $\{\rho_n\}_{n\geq 1} \subset [0,1]$ satisfying

$$\rho_n = \rho + O\left(\frac{1}{\sqrt{n}}\right)$$

such that

$$\frac{1}{n}\begin{bmatrix} \log M_{1,n} \\ \log(M_{1,n}M_{2,n}) \end{bmatrix} \in \mathbf{I}(\rho_n) + \frac{\Psi^{-1}(\mathbf{V}(\rho_n),\varepsilon)}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right)\mathbf{1}$$

Main Ideas in Converse Proof: Part III

- By evaluating Verdú-Han using multivariate Berry-Esseen, there exists a sequence $\{\rho_n\}_{n\geq 1} \subset [0,1]$ satisfying

$$\rho_n = \rho + O\left(\frac{1}{\sqrt{n}}\right)$$

such that

$$\frac{1}{n}\begin{bmatrix} \log M_{1,n} \\ \log(M_{1,n}M_{2,n}) \end{bmatrix} \in \mathbf{I}(\rho_n) + \frac{\Psi^{-1}(\mathbf{V}(\rho_n),\varepsilon)}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right)\mathbf{1}$$

- By a Taylor expansion of $\mathbf{I}(\rho_n)$ around $\mathbf{I}(\rho)$

$$\mathbf{I}(\rho_n) \approx \mathbf{I}(\rho) + (\rho_n - \rho)\mathbf{D}(\rho)$$

and the Bolzano-Weierstrass theorem, we establish the converse for a point on the curved boundary

Conclusion and Perspectives

- Why is this possible vis-à-vis second-order analysis for DM-MAC?

Conclusion and Perspectives

- Why is this possible vis-à-vis second-order analysis for DM-MAC?
- Gaussianity: Allows us to identify the boundary of the CR and time-sharing not necessary

Conclusion and Perspectives

- Why is this possible vis-à-vis second-order analysis for DM-MAC?
- Gaussianity: Allows us to identify the boundary of the CR and time-sharing not necessary
- Degraded Message Sets: Codeword pairs do not have to be independent for the strong converse

Conclusion and Perspectives

- Why is this possible vis-à-vis second-order analysis for DM-MAC?
- Gaussianity: Allows us to identify the boundary of the CR and time-sharing not necessary
- Degraded Message Sets: Codeword pairs do not have to be independent for the strong converse
- Key distinction relative to existing second-order works:

$$\mathbf{D}(\rho) = \frac{\partial}{\partial \rho}\mathbf{I}(\rho)$$

- Gaussian broadcast channel is similar to this problem

  – Techniques may carry over

Conclusion and Perspectives

- Why is this possible vis-à-vis second-order analysis for DM-MAC?
- Gaussianity: Allows us to identify the boundary of the CR and time-sharing not necessary
- Degraded Message Sets: Codeword pairs do not have to be independent for the strong converse
- Key distinction relative to existing second-order works:

$$\mathbf{D}(\rho) = \frac{\partial}{\partial \rho} \mathbf{I}(\rho)$$

- Gaussian broadcast channel is similar to this problem

  –Techniques may carry over
- Full version: `http://arxiv.org/abs/1310.1197`