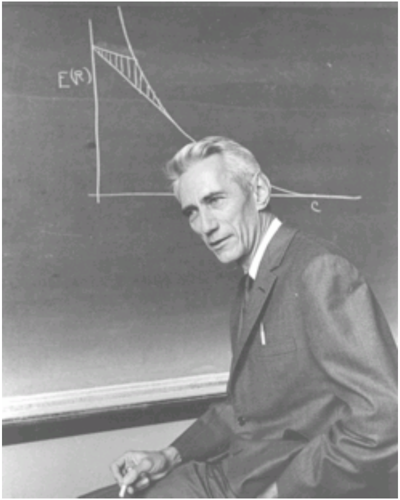
15.094J: Robust Modeling, Optimization, Computation

Lecture 23-24: Network Information Theory via Robust Optimization

Outline

1. Origins of Information Theory
2. Motivation
3. Single Use Gaussian Channel
4. Exponential single user channel
5. Two-User Gaussian Interference Channel
6. Effect of Noise
7. Conclusions

Claude Shannon (1916-2001)



- BS, University of Michigan, 1936.
- SM, MIT, 1937.
- PhD, MIT, 1940.
- Institute of Advanced Study, Princeton, 1941.
- Bell Labs, 1942-1956.
- MIT, professor, 1956-1983.
- MIT, emeritus, 1983-2001.

Perspectives

- Shannon-1948 "A Mathematical Theory of Communication", started the field of Information theory.
- The paper is described as the Magna Carta of the modern information age.
- Like Einstein in 1906, 1915, Shannon was asking questions nobody else was asking at the time.

Citations in the Mathematical/Economics Sciences in the 20th century

- **Shannon-1948: 75,000 citations**.
- Keynes -1937: 25,000 citations.
- Kahneman-Tversky -1979: 35,000 citations.
- Metropolis -1953: 30,000 citations.
- von Neumann-Morgensten -1944: 25,000 citations.
- Kalman -1953: 20,000 citations.
- RSA -1978: 15,000 citations
- Karp -1972 : 9,000 citations
- Dantzig -1947: 8,000 citations
- Nash -1950: 5,000 citations

Information theory

Successes

- Characterization of the capacity of a single user channel, Shannon-1948.
- Closed form formula for the capacity of a variety of single user channels (example: additive Gaussian channel).
- Random coding with minimum-distance decoding achieves this capacity.

Challenges

- Network Information theory.
- Multi sender, multi receiver channels with interference.
- Understanding the limitations of wireless networks.

Research Objectives

- Use RO to attack stochastic systems in high dimensions.
- Connect RO and Information Theory to achieve tractability and address network information theory.
- Understand the tractability of the approach.
- Understand the limitations of the approach.

Robustness and Information Theory

- Typical sets introduced by Shannon can be viewed as uncertainty sets in RO.
- Decoding is a robustness property.

Typical Sets

Incorporating Distributional Information

- Shannon (1948) introduced the idea of Typical Sets:
- Property (a): A typical set has probability nearly 1.
- Property (b): All elements of typical set are nearly equiprobable.
- Given pdf $f(\cdot)$,

$$\mathcal{U}^{f-\text{Typical}} = \left\{ (z_1, \ldots, z_n) \left| -\Gamma \leq \frac{\sum_{i=1}^{n} \log f(z_i) - n \cdot \mu_{\log f}}{\sigma_{\log f} \cdot \sqrt{n}} \leq \Gamma. \right. \right\},$$

$$\mu_{\log f} = \int_{-\infty}^{\infty} f(x) \log f(x) \, \mathrm{d}x,$$

$$\sigma_{\log f_j} = \int_{-\infty}^{\infty} f(x) \left(\log f(x) - \mu_{\log f} \right)^2 \mathrm{d}x.$$

Examples of Typical Sets

- $\tilde{z}_i \sim N(0, \sigma)$

$$\mathcal{U}_\epsilon^{\mathrm{G}} = \left\{ \mathbf{z} \,\middle|\, -\Gamma_\epsilon^{\mathrm{G}} \leq \|\mathbf{z}\|^2 - n\sigma^2 \leq \Gamma_\epsilon^{\mathrm{G}} \right\}.$$

- $\tilde{z}_i \sim Exp(\lambda)$

$$\mathcal{U}_\epsilon^{\mathrm{E}} = \left\{ \mathbf{z} \,\middle|\, \frac{n}{\lambda} - \frac{\sqrt{n}}{\lambda} \cdot \Gamma_\epsilon^{\mathrm{E}} \leq \sum_{j=1}^{n} z_j \leq \frac{n}{\lambda} + \frac{\sqrt{n}}{\lambda} \cdot \Gamma_\epsilon^{\mathrm{E}},\ \mathbf{z} \geq \mathbf{0} \right\}.$$

- $\tilde{z}_i \sim U[a, b]$

$$\mathcal{U}_\epsilon^{U} = \left\{ \mathbf{z} \,\middle|\, \begin{array}{c} n\dfrac{a+b}{2} - \Gamma_\epsilon^{U}\sqrt{n} \leq \displaystyle\sum_{j=1}^{n} z_j \leq n\dfrac{a+b}{2} + \Gamma_\epsilon^{U}\sqrt{n}, \\ a \leq z_j \leq b,\ j = 1, \ldots, n, \end{array} \right\}..$$

- $\tilde{z}_i \sim \mathrm{Bin}(p)$

$$\mathcal{U}_\epsilon^{B} = \left\{ \mathbf{z} \,\middle|\, \begin{array}{c} np - \Gamma_\epsilon^{B}\sqrt{n} \leq \displaystyle\sum_{j=1}^{n} z_j \leq np + \Gamma_\epsilon^{B}\sqrt{n}, \\ z_j \in \{0, 1\},\ j = 1, \ldots, n, \end{array} \right\}.$$

Single Use Gaussian Channel

- A transmitter wants to send M messages index by $i \in \mathcal{M}$.
- He codes the i^{th} message as a vector $\mathbf{x}_i \in \mathbb{R}^n$, and transmits it.
- The receiver, receives a vector $\mathbf{y}_i \in \mathbb{R}^n$ given by

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{z}_i,$$

  $\mathbf{z}_i$ is noise.
- Noise could be different than additive.

The Key Problem

- *Coding Problem:* Given power P, select $\mathbf{x}_i$ to represent the i^{th} message such that $||\mathbf{x}_i||^2 \leq nP$.
- *Decoding Problem:* Find a decoding function $g(\mathbf{y}_i)$ that maps $\mathbf{y}_i$ to one of the code words:

$$\frac{1}{M}\sum_{i=1}^{M}\mathbb{P}\left[g(\mathbf{y}_i) \neq i\right] \leq \epsilon_n,$$

  with $\epsilon_n \to 0$, as $n \to \infty$.
- *Can we send M messages of length n with arbitrary small error ϵ_n?*

Key Insights

- Decoder
$$g\left(\mathbf{y}\right) = \arg\min_{i \in \mathcal{M}} \left\|\mathbf{y} - \mathbf{x}_i\right\|.$$
- Minimum distance decoding is a *robustness property*
$$\left\|\mathbf{x}_i + \mathbf{z} - \mathbf{x}_{i'}\right\| \geq \left\|\mathbf{z}\right\| \ \ \forall \mathbf{z} \in \mathcal{U}_\epsilon^{\mathrm{G}}, \forall i, i' \neq i$$
$$\mathcal{U}_\epsilon^{\mathrm{G}} = \left\{\mathbf{z} \,\middle|\, -\Gamma_\epsilon^{\mathrm{G}} \leq \left\|\mathbf{z}\right\|^2 - n\sigma^2 \leq \Gamma_\epsilon^{\mathrm{G}}\right\}.$$
- Capacity characterization
$$\begin{aligned} M_n^*(\epsilon) = \max \quad & M \\ \text{s.t.} \quad & \left\|\mathbf{x}_i + \mathbf{z} - \mathbf{x}_{i'}\right\| \geq \left\|\mathbf{z}\right\|, \quad \forall \mathbf{z} \in \mathcal{U}_\epsilon^{\mathrm{G}}, \forall i, i' \in \mathcal{M}, i' \neq i, \\ & \left\|\mathbf{x}_i\right\|^2 \leq nP, \qquad \forall i \in \mathcal{M}. \end{aligned}$$
- Theorem (Shannon (1948))
$$\lim_{n\to\infty,\ \epsilon\to 0} \frac{\log_2 M_n^*(\epsilon)}{n} = \frac{1}{2} \cdot \log\left(1 + \frac{P}{\sigma^2}\right).$$

Optimal Coding

- *Inputs: R, P, n, σ, ϵ, ν*.
- Select γ_ϵ, $\mathbb{P}\left[\|\tilde{\mathbf{z}}_G\| \leq \gamma_\epsilon\right] \geq 1-\epsilon$, $\tilde{\mathbf{z}}_G \sim N\left(\mathbf{0}, \sigma \cdot I\right)$;
- $T = \left(\frac{1+\nu}{\zeta\nu} \cdot \frac{\gamma_\epsilon}{\sqrt{n}}\right)^n$, with $\zeta = \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}\left(1 - O(\epsilon)\right)$,
- $M_0 = (1+\nu) \cdot \gamma_\epsilon.$;
- Let $\mathcal{Z} = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_T\}$ uniformly distributed on $\mathcal{S}_n(M_0) = \left\{\mathbf{z} \in \mathbb{R}^n \middle| \|\mathbf{z}\| = M_0\right\}$.
- Wyner (1967) developed methods to construct deterministic sequences to model uniformly distributed points.

Optimal Coding

- To count fraction of correct decoding:

$$v_{it} = \begin{cases} 1, & \text{if } \left\|\mathbf{x}_i + \mathbf{z}_t - \mathbf{x}_{i'}\right\| \geq \left\|\mathbf{z}_t\right\|, \forall i' \in \mathcal{M}, \\ 0, & \text{otherwise.} \end{cases}$$

- Encoding Algorithm

$$\begin{aligned}
&\left\|\mathbf{x}_i\right\|^2 \leq nP, && \forall i \in \mathcal{M}, \\
&\left\|\mathbf{x}_i - \mathbf{x}_k + \mathbf{z}_t\right\| + (1 - v_{it}) M_0 \geq \left\|\mathbf{z}_t\right\|, && \forall t \in \mathcal{T}, \forall i, k \in \mathcal{M},\ k \neq i, \\
&\sum_{t=1}^{T} v_{it} \geq (1 - \epsilon) T, && \forall i \in \mathcal{M}, \\
&v_{it} \in \{0, 1\}, && \forall i \in \mathcal{M}, t \in \mathcal{T},
\end{aligned}$$

Reformulation

The set of quadratic, possibly non-convex, constraints

$$f_k(\mathbf{y}) = \mathbf{y}'\mathbf{A}_k\mathbf{y} + \mathbf{b}_k'\mathbf{y} + c_k \leq 0, \;\; \forall k \in \mathcal{K}.$$

is equivalent to the semidefinite optimization problem

$$\begin{gathered} \tilde{\mathbf{A}}_k \bullet \mathbf{Y} \leq 0, \;\; \forall k \in \mathcal{K}, \\ Y_{11} = 1, \; \mathbf{Y} \succeq \mathbf{0}, \; \mathrm{rank}\,(\mathbf{Y}) = 1, \end{gathered}$$

where

$$\mathbf{Y} = \begin{pmatrix} 1 \\ \mathbf{y} \end{pmatrix} \left(1, \mathbf{y}'\right), \;\; \tilde{\mathbf{A}}_k = \begin{pmatrix} c_k & \mathbf{b}_k \\ \mathbf{b}_k & \mathbf{A}_k \end{pmatrix}.$$

The overall optimization problem

$$
\begin{array}{lll}
\min & \operatorname{rank}(\mathbf{Y}) & \\
\text{s.t.} & \mathbf{A}_i \bullet \mathbf{Y} \le 0, & \forall i \in \mathcal{M}, \\
& \mathbf{B}_{ikt} \bullet \mathbf{Y} \le 0, & \forall t \in \mathcal{T}, \forall i, k \in \mathcal{M},\ k \neq i, \\
& \mathbf{C}_i \bullet \mathbf{Y} \le 0, & \forall i \in \mathcal{M}, \\
& \mathbf{D}_{it} \bullet \mathbf{Y} = 0, & \forall i \in \mathcal{M}, t \in \mathcal{T}, \\
& \mathbf{Y} \succeq \mathbf{0}, &
\end{array}
$$

Algorithm

- **Input :** R, P, σ, n, ν, ϵ.
- Solve the rank minimization SOP to compute r^*, codewords$\{\mathbf{x}_i\}_{i \in \mathcal{M}}$.
- If $r^* = 1$, then $R \in \mathcal{R}_n\left[P, \sigma, 2\epsilon\right]$.
- If $r^* \geq 2$, then $R \notin \mathcal{R}_n\left[P, (1 + 3\nu)\, \sigma, O(\epsilon)\right]$.
- As $n \to \infty$, $\epsilon \to 0$, $\nu \to 0$, the characterization of *the asymptotic capacity of the channel is tight.*

Remarks

- By solving a rank minimization SOP, we find the asymptotic capacity and the matching optimal code.
- *Similar RO problems for a variety of other type of channels:*
  - Binary symmetric channel,
  - Binary erasure channel,
  - Additive uniform noise channel.
  - Additive exponential channel.

Algorithm from Fazell, Hindi, Boyd –2003

- Solve the convex optimization problem

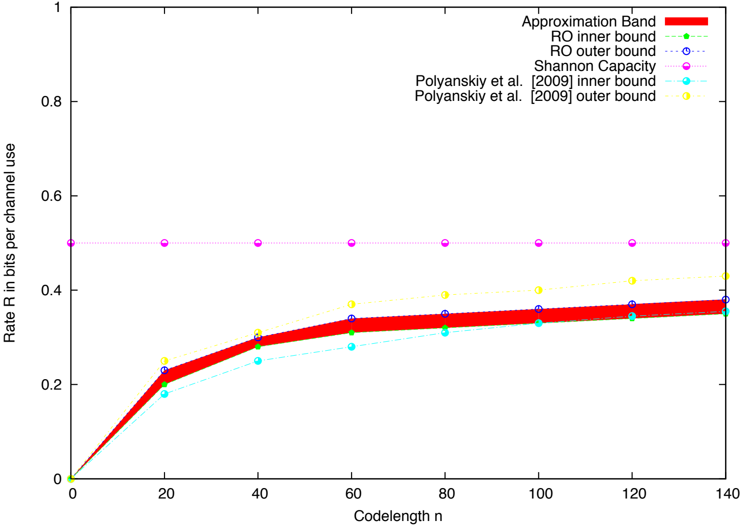
$$\begin{aligned} \min \quad & \mathrm{Tr}\,(\mathbf{X}) \\ \text{s.t.} \quad & \tilde{\mathbf{A}} \bullet \mathbf{X} \le 0, \\ & \mathbf{X} \succeq 0, \end{aligned}$$

  and let $\mathbf{X}^0$ denote the optimal solution.

- For each iteration $k = 1, \ldots, K$, solve the optimization problem

$$\begin{aligned} \min \quad & \mathrm{Tr}\left(\left(\mathbf{X}^{k-1} + \delta I\right)^{-1} \mathbf{X}\right) \\ \text{s.t.} \quad & \tilde{\mathbf{A}} \bullet \mathbf{X} \le 0, \\ & \mathbf{X} \succeq 0. \end{aligned}$$

Single User Gaussian Channel



Exponential single user channel

- Recall Typical set

$$\mathcal{U}_\epsilon(\lambda) = \left\{ \mathbf{z} \,\middle|\, \frac{n}{\lambda} - \frac{\sqrt{n}}{\lambda} \cdot \Gamma_\epsilon^{\mathrm{E}} \le \sum_{j=1}^{n} z_j \le \frac{n}{\lambda} + \frac{\sqrt{n}}{\lambda} \cdot \Gamma_\epsilon^{\mathrm{E}}, \ \mathbf{z} \ge \mathbf{0} \right\}.$$

- MLE Decoder

$$\arg\min_{i \in \mathcal{B}(\mathbf{y})} \sum_{j=1}^{n} (y_j - x_{ij}),$$

where $\mathcal{B}(\mathbf{y}) = \left\{ i \in \mathcal{B} \,\middle|\, y_j \ge x_{ij}, \ \forall j = 1, \ldots, n \right\}$.

- Generate $\mathbf{z}_t$ uniformly on $\mathcal{U}_\epsilon\left(\frac{\lambda}{1+2\nu}\right)$ to help us count the error probability.

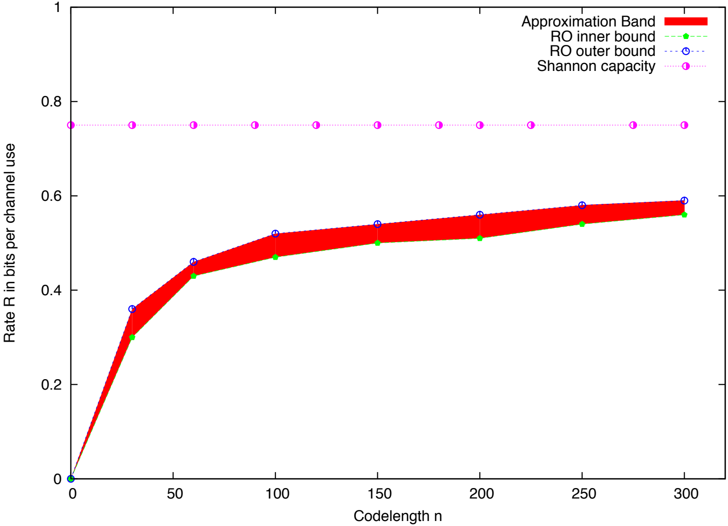
Algorithm

- Solve

$$
\begin{aligned}
\max \quad & \sum_{i,k,t} v_{ikt} & \\
& \sum_{j=1}^{n} x_{ij} \leq nP, & \forall i = 1, \ldots, 2^{nR}, \\
& \sum_{j=1}^{n} x_{ij} + (2 - v_{it} - v_{ikt}) M_0 \geq \sum_{j=1}^{n} x_{kj}, & \forall t, \forall i, k \neq i, \\
& x_{ij} + z_{tj} \geq x_{kj} - M_0 (1 - v_{ikt}), & \forall i, k, j, t, \\
& \sum_{t=1}^{T} v_{it} \geq (1 - \epsilon) T, & \forall i, \\
& v_{it}, v_{ikt} \in \{0, 1\}, & \forall i, k, t,
\end{aligned}
$$

- If feasible, then $R \in \mathcal{R}_n^E [P, \lambda, 2\epsilon]$. Otherwise, then $R \notin \mathcal{R}_n^E \left[P, \frac{\lambda}{1+2\nu}, \epsilon\right]$.

Single User Exponential Channel



Two-User Gaussian Interference Channel

- Two transmitters, two receivers.
- User 1 selects a message i and transmits $\mathbf{x}_i^1$.
- User 2 selects a message k and transmits $\mathbf{x}_k^2$.
- The signal vectors $\mathbf{x}_i^1$ and $\mathbf{x}_k^2$ are power constrained, that is,

$$\left\|\mathbf{x}_i^1\right\|^2 \leq nP_1, \quad \left\|\mathbf{x}_i^2\right\|^2 \leq nP_2.$$

- The received signals $\mathbf{y}^1$, $\mathbf{y}^2$ are

$$\begin{aligned} \mathbf{y}^1 &= \mathbf{x}_i^1 + h_{12}\mathbf{x}_k^2 + \tilde{\mathbf{z}}_1, \\ \mathbf{y}^2 &= \mathbf{x}_k^2 + h_{21}\mathbf{x}_i^1 + \tilde{\mathbf{z}}_2, \end{aligned}$$

where h_{12}, h_{21} are interference parameters, and $\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2 \sim N\left(\mathbf{0}, \sigma \cdot I\right)$.

Decoding

Suppose the noise $\tilde{\mathbf{z}}$ is distributed **uniformly** in $\mathcal{B}_n(r) = \{\mathbf{z} \in \mathbb{R}^n \,|\, \|\mathbf{z}\| \leq r\}$. The maximum likelihood decoder for this channel is given by

$$i_1^* = \arg\max_{i \in \mathcal{M}^1} \left|\mathcal{B}_i^1\right|, \text{ where } \mathcal{B}_i^1 = \left\{k \in \mathcal{M}^2 \,:\, \left\|\mathbf{y}^1 - \left(\mathbf{x}_i^1 + h_{12}\mathbf{x}_k^2\right)\right\| \leq r\right\}$$

$$i_2^* = \arg\max_{i \in \mathcal{M}^2} \left|\mathcal{B}_i^2\right|, \text{ where } \mathcal{B}_i^2 = \left\{k \in \mathcal{M}^1 \,:\, \left\|\mathbf{y}^2 - \left(\mathbf{x}_i^2 + h_{21}\mathbf{x}_k^1\right)\right\| \leq r\right\}.$$

Parameters

- Parameter γ_ϵ: $\mathbb{P}\left[\|\tilde{\mathbf{z}}_G\| \leq \gamma_\epsilon\right] \geq 1-\epsilon$, $\tilde{\mathbf{z}}_G \sim N\left(\mathbf{0}, \sigma \cdot I\right)$.
- $T = \left(\frac{1+\nu}{\eta\nu} \cdot \frac{\gamma_\epsilon}{\sqrt{n}}\right)^n$, with $\eta = \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}\left(1-\epsilon^{1/4}\right)$,
- $M_0 = (1+\nu) \cdot \gamma_\epsilon$.
- $\mathcal{Z} = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_T\}$ as before, $\|\mathbf{z}_t\| = M_0$, $t = 1, \ldots, T$.
- "Cushion" parameters: δ_1, δ_2, α_1, α_2 with the property that as $\nu \to 0$ and $n \to \infty$, $\delta_1, \delta_2 \to 0$ and $\alpha_1, \alpha_2 \to 1$.

Decoding

- For correct decoding we would require

$$\left|\left\{k' \in \mathcal{B}^2 \mid \left\|\mathbf{y}^1 - \mathbf{x}^1_{i_1^*} - h_{12}\mathbf{x}^2_k\right\|^2 \leq M_0^2\right\}\right| \geq$$

$$\left|\left\{k' \in \mathcal{B}^2 \mid \left\|\mathbf{y}^1 - \mathbf{x}^1_i - h_{12}\mathbf{x}^2_k\right\|^2 \leq M_0^2\right\}\right|, \forall i \in \mathcal{M}^1.$$

- We create a "cushion" δ_1, and instead require

$$\left|\left\{k' \in \mathcal{B}^2 \, : \, \left\|\mathbf{y}^1 - \mathbf{x}^1_{i_1^*} - h_{12}\mathbf{x}^2_k\right\|^2 \leq M_0^2 + \delta_1\right\}\right| \geq$$

$$\left|\left\{k' \in \mathcal{B}^2 \, : \, \left\|\mathbf{y}^1 - \mathbf{x}^1_i - h_{12}\mathbf{x}^2_k\right\|^2 \leq M_0^2 - \delta_1\right\}\right|, \forall i \in \mathcal{M}^1.$$

- Using this decoder, we want to ensure that the average probability of error is at most $\epsilon^{1/4}$, that is,

$$\frac{1}{M_2} \sum_{k \in \mathcal{M}^2} \mathbb{P}\left[g^1\left(\mathbf{y}^1\right) \neq i \,\middle|\, m^1 = i, m^2 = k\right] \leq \epsilon^{1/4}.$$

Variables

- "Counting" variables $\left\{v_i^1, v_{ik}^1, v_{ikt}^1\right\}_{i\in\mathcal{M}^1, k\in\mathcal{M}^2, t\in\mathcal{T}}$:

$$v_{ikt}^1 = \begin{cases} 1, & \text{if } \left|\mathcal{B}_{ikt,i'}^1\right| \leq \left|\mathcal{B}_{ikt,i}^1\right|, \ \forall i' \in \mathcal{M}^1, \\ 0, & \text{otherwise}, \end{cases}$$

$$v_{ik}^1 = \begin{cases} 1, & \text{if } \sum_{t\in\mathcal{T}} v_{ikt}^1 \geq \left(1-\epsilon^{1/4}\right)\cdot T, \\ 0, & \text{otherwise}, \end{cases}$$

$$v_i^1 = \begin{cases} 1, & \text{if } \sum_{k\in\mathcal{M}^2} v_{ik}^1 \geq \left(1-\epsilon^{1/4}\right)\cdot M_2, \\ 0, & \text{otherwise}, \end{cases}$$

where

$$\mathcal{B}_{ikt,i'}^1 = \left\{k' \in \mathcal{B}^2 \ : \ \left\|\mathbf{x}_i^1 - \mathbf{x}_{i'}^1 + h_{12}\left(\mathbf{x}_k^2 - \mathbf{x}_{k'}^2\right) + \mathbf{z}_t\right\|^2 \leq M_0^2 - \delta_1\right\},$$

$$\mathcal{B}_{ikt,i}^1 = \left\{k' \in \mathcal{B}^2 \ : \ \left\|h_{12}\left(\mathbf{x}_k^2 - \mathbf{x}_{k'}^2\right) + \mathbf{z}_t\right\|^2 \leq M_0^2 + \delta_1\right\}.$$

More Variables

- $$v^1_{ii'kk't} = \begin{cases} 1, & \text{if } \left\|\mathbf{x}^1_i - \mathbf{x}^1_{i'} + h_{12}\left(\mathbf{x}^2_k - \mathbf{x}^2_{k'}\right) + \mathbf{z}_t\right\|^2 \leq M_0^2 - \delta_1, \\ 0, & \text{otherwise,} \end{cases}$$
  $$v^1_{iikk't} = \begin{cases} 1, & \text{if } \left\|h_{12}\left(\mathbf{x}^2_k - \mathbf{x}^2_{k'}\right) + \mathbf{z}_t\right\|^2 \leq M_0^2 + \delta_1, \\ 0, & \text{otherwise.} \end{cases}$$
- The variables $\left\{v^2_k, v^2_{ki}, v^2_{kit}, v^2_{kk'ii't}\right\}$ corresponding to User 2 are defined in a similar manner.

Algorithm: **Input** : n, R_1, R_2, σ, P_1, P_2, ϵ, ν.

$$\left\|\mathbf{x}_i^1\right\|^2 \leq nP_1,$$

$$\left\|\mathbf{x}_i^1 - \mathbf{x}_{i'}^1 + h_{12}\left(\mathbf{x}_k^2 - \mathbf{x}_{k'}^2\right) + \mathbf{z}_t\right\|^2 \leq M_0^2 - \delta_1 + \left(1 - v_{ii'kk't}^1\right) M_0^2,$$

$$\left\|h_{12}\left(\mathbf{x}_k^2 - \mathbf{x}_{k'}^2\right) + \mathbf{z}_t\right\|^2 \leq M_0^2 + \delta_1 + \left(1 - v_{iikk't}^1\right) M_0^2,$$

$$\sum_{k'=1}^{M_2} v_{iikk't}^1 \geq \sum_{k'=1}^{M_2} v_{ii'kk't}^1,$$

$$v_{ii'kk't}^1 \leq v_{ikt}^1, \qquad v_{ikt}^1 \leq v_{ik}^1, \qquad v_{ik}^1 \leq v_i^1,$$

$$\sum_{t=1}^{T} v_{ikt}^1 \geq \left(1 - \epsilon^{1/4}\right) \cdot T \cdot v_{ik}^1,$$

$$\sum_{k=1}^{M_2} v_{ik}^1 \geq \left(1 - \epsilon^{1/4}\right) \cdot M_2 \cdot v_i^1,$$

$$\sum_{i=1}^{M_1} v_i^1 \geq \left(1 - \epsilon^{1/4}\right) \cdot M_1,$$

Algorithm Continued

$$\left\|\mathbf{x}_k^2\right\|^2 \leq nP_2,$$

$$\left\|\mathbf{x}_k^2 - \mathbf{x}_{k'}^2 + h_{21}\left(\mathbf{x}_i^1 - \mathbf{x}_{i'}^1\right) + \mathbf{z}_t\right\|^2 \leq M_0^2 - \delta_2 + \left(1 - v_{kk'ii't}^2\right) M_0^2,$$

$$\left\|h_{21}\left(\mathbf{x}_i^1 - \mathbf{x}_{i'}^1\right) + \mathbf{z}_t\right\|^2 \leq M_0^2 + \delta_2 + \left(1 - v_{kkii't}^2\right) M_0^2,$$

$$\sum_{i'=1}^{M_1} v_{kkii't}^2 \geq \sum_{i'=1}^{M_1} v_{kk'ii't}^2,$$

$$v_{kk'ii't}^2 \leq v_{kit}^2, \quad v_{kit}^2 \leq v_{ki}^2, \quad v_{ki}^2 \leq v_k^2,$$

$$\sum_{t=1}^{T} v_{kit}^2 \geq \left(1 - \epsilon^{1/4}\right) \cdot T \cdot v_{ki}^2,$$

$$\sum_{i=1}^{M_1} v_{ki}^2 \geq \left(1 - \epsilon^{1/4}\right) \cdot M_1 \cdot v_k^2,$$

$$\sum_{k=1}^{M_2} v_k^2 \geq \left(1 - \epsilon^{1/4}\right) \cdot M_2,$$

Central Result

Capacity Region

- Reformulate as a rank minimization SOP

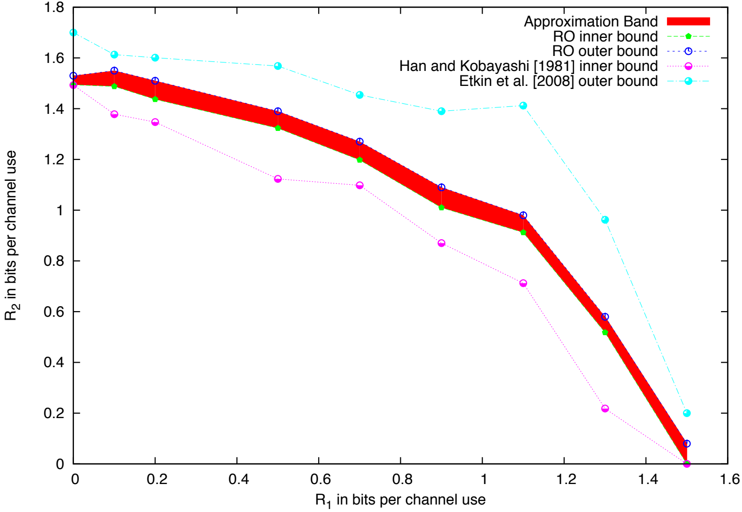
$$\begin{aligned} r^* = \min \quad & \operatorname{rank}(\mathbf{Y}) \\ \text{s.t.} \quad & \mathbf{A}_i \bullet \mathbf{Y} \leq 0, \\ & \mathbf{B}_i \bullet \mathbf{Y} = 0, \\ & \mathbf{Y} \succeq \mathbf{0}. \end{aligned}$$

- If $r^* = 1$, then $(R_1, R_2) \in \mathcal{R}_n^{\mathrm{IC}} \left[\alpha_1 P_1, \alpha_2 P_2, \frac{h_{12}}{\alpha_2}, \frac{h_{21}}{\alpha_1}, \sigma, O(\epsilon^{1/4}) \right]$,
- If $r^* \geq 2$, then $(R_1, R_2) \notin \mathcal{R}_n^{\mathrm{IC}} \left[P_1, P_2, h_{12}, h_{21}, (1 + 3\nu)\,\sigma, O(\epsilon) \right]$,
- Note that as $n \to \infty, \epsilon,\ \nu \to 0,\ \alpha_1, \alpha_2 \to 1$ and *the characterization of the asymptotic capacity is tight.*

Optimization problem as a function of the noise

Noise	Typical Set	Optimization Problem
Gaussian (independent)	Ball	*Rank minimization* with semidefinite constraints
Gaussian (correlated)	Ellipsoid	*Rank minimization* with semidefinite constraints
Exponential	Polyhedron	*Binary mixed linear* optimization problem
Uniform	Polyhedron	*Binary mixed linear* optimization problem
Binary symmetric noise	Polyhedron	*Binary optimization* problem

Two-user Gaussian Channel, $n = 60$



Extensions

- Multi-Cast channel.
- Multi-Access channel.
- Many transmitters, many receivers.

Conclusions

- The reason size becomes an issue is because of error probability guarantees.
- RO brings the power of *optimization* to the analysis of information theory.
- If underlying problem mixed binary, we can solve $n = 300 - 500$ and $M = 200,000$.
- If underlying problem SOP with rank constraints, we can solve $n = 100 - 150$ and $M = 100,000$.