Finite-Blocklength Universal Coding for Multiple-Access Channels

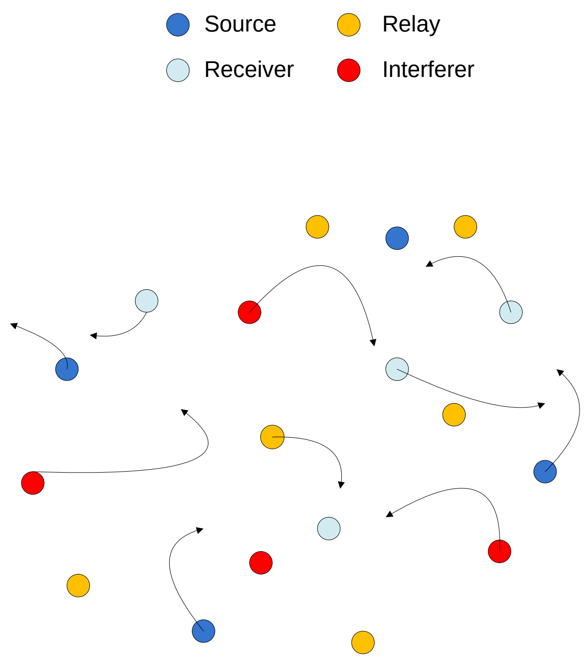
Pierre Moulin

University of Illinois at Urbana-Champaign

Electrical and Computer Engineering

DARPA ITMANET meeting
Stanford, CA
January 26, 2011

- What is the capacity region for this network?



- What is the capacity region for finite blocklengths?
- What if the link statistics are unknown?

Outline

- Capacity limits for finite blocklength:
  Strassen (1962), Polyanskiy *et al.* (2008), Hayashi (2009)
- Universal coding (unknown channel):
  Csiszár and Körner (1981)
- May'10 ITMANET meeting: achievable rates for universal decoder
- This talk:
  - Background
  - Multiple Access Channel
  - Severe Outage

PART I: Capacity for finite blocklength

- Discrete memoryless channel $\{W(y|x),\, x \in \mathcal{X},\, y \in \mathcal{Y}\}$
- Codewords $\mathbf{x}(m) \in \mathcal{X}^n$ for $m = 1, 2, \cdots, M_n$
- Code rate $R_n \triangleq \frac{1}{n} \log M_n$
- Decoding rule $\hat{m} = \phi(\mathbf{y})$
- Average error probability

$$P_e(W) = \frac{1}{M} \sum_{m=1}^{M_n} \sum_{\mathbf{y} \in \mathcal{Y}^n} W^n(\mathbf{y}|\mathbf{x}(m))\, \mathbb{1}\{\phi(\mathbf{y}) \neq m\}$$

- ϵ-capacity for blocklength n:

$$C_n(W, \epsilon) = \sup\{R_n \ :\ P_e(W) \leq \epsilon\}$$

- Shannon (1948):

$$C(W) = \lim_{\epsilon \downarrow 0} \lim_{n \to \infty} C_n(W, \epsilon) = \max_{P_X} I(P_X; W)$$

- Strassen (1962) and PPV (2008):

$$C_n(W, \epsilon) \geq C(W) - \frac{\sigma(W)}{\sqrt{n}} Q^{-1}(\epsilon) - O\left(\frac{\log n}{n}\right)$$

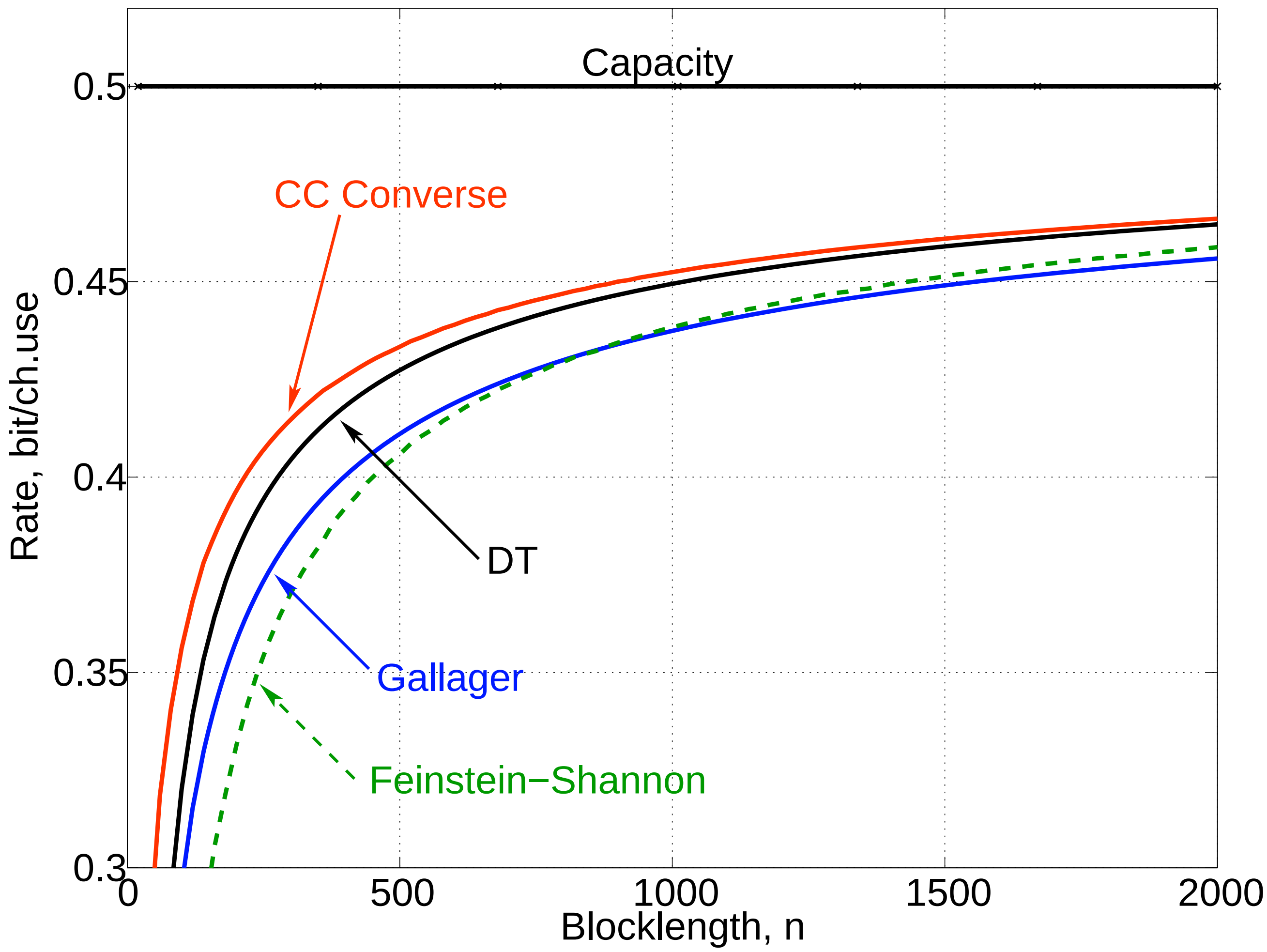
- Converse:

$$C_n(W, \epsilon) \leq C(W) - \frac{\sigma(W)}{\sqrt{n}} Q^{-1}(\epsilon) + O\left(\frac{\log n}{n}\right)$$

- The first two terms giving $C_n(W, \epsilon)$ are achieved by standard random codes and ML decoding:

$$\max_{1 \leq m \leq M_n} W^n(\mathbf{y}|\mathbf{x}(m)) \quad \Leftrightarrow \quad \max_{1 \leq m \leq M_n} i(\mathbf{x}(m); \mathbf{y})$$

where $i(\mathbf{x}(m); \mathbf{y}) = \log \frac{W^n(\mathbf{y}|\mathbf{x}(m))}{p(\mathbf{y})}$ is the information density


Capacity
CC Converse
DT
Gallager
Feinstein−Shannon
Rate, bit/ch.use
Blocklength, n
0.5
0.45
0.4
0.35
0.3
0
500
1000
1500
2000

- For iid random codes, $p(\mathbf{y})$ factors and thus

$$\frac{1}{n} i(\mathbf{x};\mathbf{y}) \quad = \quad \frac{1}{n} \sum_{i=1}^{n} \underbrace{\log \frac{W(y_i|x_i)}{P_Y(y_i)}}_{l_i = l(x_i, y_i) = \text{iid rv's}}$$

- Mutual information $I(P_X, W) = \mathbb{E}\left[l(X,Y)\right]$
- Channel dispersion $\sigma^2(P_X, W) = \text{Var}\left[l(X,Y)\right]$
- Can thus write

$$\frac{1}{n} i(\mathbf{x}(m);\mathbf{y}) = I(P_X, W) + \frac{\sigma(P_X, W)}{\sqrt{n}}\, Z_n$$

where $Z_n \stackrel{d}{\to} \mathcal{N}(0,1)$ by the Central Limit Theorem.

Refined Asymptotic Analysis

- Assume any $y \in \mathcal{Y}$ can be reached from any $x \in \mathcal{X}$
- Use Laplace's method for refined large-dev analysis, achieve

$$C(W) - \frac{\sigma(W)}{\sqrt{n}} Q^{-1}(\epsilon) + \frac{\log n}{2n} + O(1/n) \stackrel{?}{=} C_n(W, \epsilon) \quad (1)$$

- Converse? Consider CC code with type $\hat{P}_X \in \mathcal{P}_n(\mathcal{X})$

$$\Rightarrow \quad M_n \leq \max_{\hat{P}_X \in \mathcal{P}_n(\mathcal{X})} \inf_{Q \in \mathcal{P}(\mathcal{Y})} \sup_{\mathbf{x} \in T(\hat{P}_X)} \frac{1}{\beta_{1-\epsilon}(\mathbf{x}, Q)} \quad \Rightarrow (1)$$

  where $\beta_{1-\epsilon}(\mathbf{x}, Q)$ = type-II error prob of NP test between $\prod_{i=1}^n W(Y_i|x_i)$ and $\prod_{i=1}^n Q(Y_i)$ with type-I error prob $\leq \epsilon$
- Any code can be decomposed into $\leq (n+1)^{|\mathcal{X}|}$ such CC subcodes $\Rightarrow$ UB on M_n can be increased at most by $|\mathcal{X}| \frac{\log n}{n}$

Finite-Blocklength Universal Coding

- Assume d-dimensional compact family of channels W
- Use Shannon's iid random codes with $P_X = P^*$ and rate

$$R_n = I(P^*, W^*) - \frac{\sigma(P^*, W^*)}{\sqrt{n}} Q^{-1}(\epsilon) - \frac{(d-1)\log n}{2n}$$

  and a modified Maximum Mutual Info decoder
- The penalty for not knowing channel W is $\frac{d \log n}{2n}$
- Familiar?

PART II: Multiple Access Channel

- DMC $\{W(y|x_1, x_2),\ x_1 \in \mathcal{X}_1,\ x_2 \in \mathcal{X}_2,\ y \in \mathcal{Y}\}$
- Codewords $\mathbf{x}_i(m) \in \mathcal{X}_i^n$ for $i = 1, 2$ and $m = 1, 2, \cdots, M_n$
- Code rates $R_{in} \triangleq \frac{1}{n} \log M_{in}$ for $i = 1, 2$
- Decoding rule $(\widehat{m_1}, \widehat{m_2}) = \phi(\mathbf{y})$
- Average error probability

$$P_e(W) = \frac{1}{M} \sum_{m_1=1}^{M_{1n}} \sum_{m_2=1}^{M_{2n}} \sum_{\mathbf{y} \in \mathcal{Y}^n} W^n(\mathbf{y}|\mathbf{x}_1(m), \mathbf{x}_1(m))\, \mathbb{1}\{\phi(\mathbf{y}) \neq (m_1, m_2)\}$$

- Rate pair (R_1, R_2) is ϵ-achievable if $P_e(W) \leq \epsilon$
- **Problem:** characterize the **ϵ-capacity region for blocklength** n:

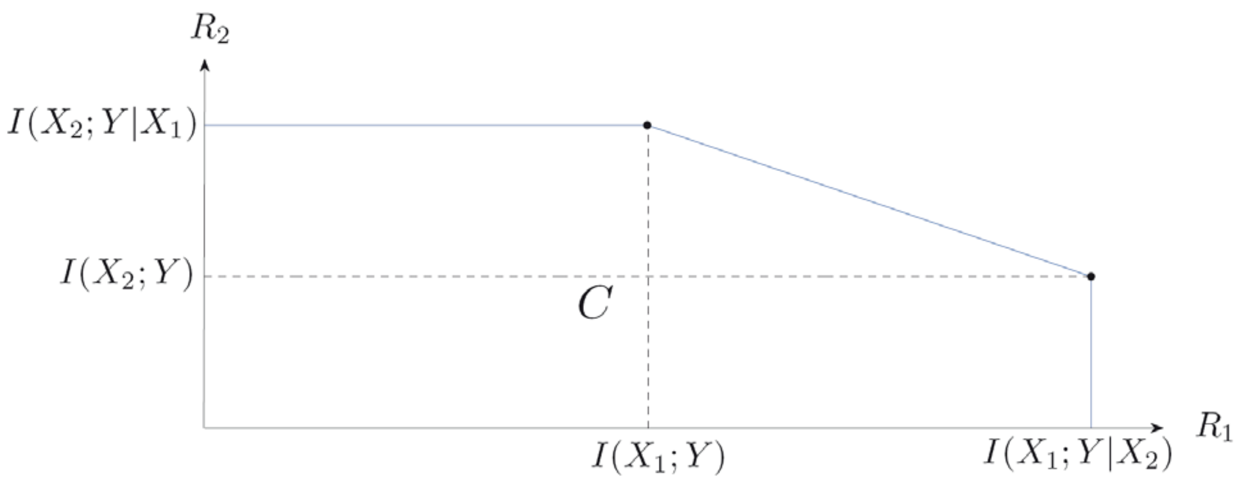
$$\mathcal{C}_n(W, \epsilon) \triangleq \mathrm{Cl}\{(R_{1n}, R_{2n})\ :\ P_e(W) \leq \epsilon\}$$

- As $\epsilon \downarrow 0$ and $n \to \infty$ this coincides with the usual MAC capacity region

$$\mathcal{C}(W) = \cup_{P_U P_{X_1|U} P_{X_2|U}} \mathcal{R}(P_U P_{X_1|U} P_{X_2|U})$$

(Ahslwede 71, Liao 72) where

$$\begin{aligned} \mathcal{R}(P_U P_{X_1|U} P_{X_2|U}) = \{(R_1, R_2) \ : \ R_1 &\leq I(X_1; Y|X_2, U) \\ R_2 &\leq I(X_2; Y|X_1, U) \\ R_1 + R_2 &\leq I(X_1, X_2; Y|U)\} \end{aligned}$$



- Rate pairs $(R_1, R_2) \in \mathcal{C}(W)$ are achieved by iid random codes and joint typicality decoding
- Fano's inequality yields weak converse

ϵ-capacity region for blocklength n

- **Theorem:**

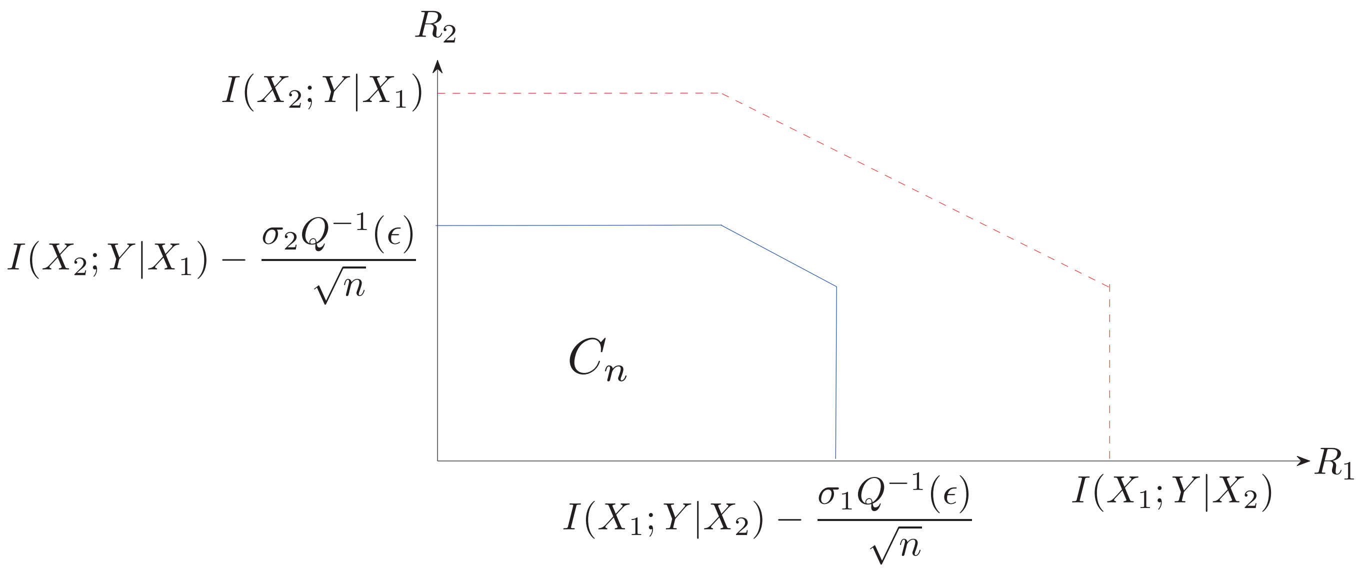
$$\mathcal{C}_n(W,\epsilon) = \cup_{P_U P_{X_1|U} P_{X_2|U}} \mathcal{R}_{n,\epsilon}(P_U P_{X_1|U} P_{X_2|U})$$

where $\mathcal{R}_{n,\epsilon}(P_U P_{X_1|U} P_{X_2|U})$ is the set of rate pairs

$$\begin{aligned}
\{(R_1, R_2) \ : \ R_1 &\le I(X_1;Y|X_2,U) - \frac{\sigma_1 Q^{-1}(\epsilon)}{\sqrt{n}} + O(\frac{\log n}{n}) \\
R_2 &\le I(X_2;Y|X_1,U) - \frac{\sigma_2 Q^{-1}(\epsilon)}{\sqrt{n}} + O(\frac{\log n}{n}) \\
R_1 + R_2 &\le I(X_1,X_2;Y|U) - \frac{\sigma_3 Q^{-1}(\epsilon)}{\sqrt{n}} + O(\frac{\log n}{n})\}
\end{aligned}$$

where $\sigma_1^2 = \mathrm{Var}\left[\log \frac{p(Y|X_1,X_2,U)}{p(Y|X_2,U)}\right]$, $\sigma_2^2 = \mathrm{Var}\left[\log \frac{p(Y|X_1,X_2,U)}{p(Y|X_1,U)}\right]$, $\sigma_3^2 = \mathrm{Var}\left[\log \frac{p(Y|X_1,X_2,U)}{p(Y|U)}\right]$ are channel dispersions

- Capacity region $\mathcal{C}_n(W, \epsilon)$:



- Achievability using iid random codes and the following threshold decoder: output all (m_1, m_2) such that

$$
\begin{aligned}
i_1(\mathbf{x}_1(m_1); \mathbf{y}|\mathbf{x}_2(m_2)) &\triangleq \log \frac{W^n(\mathbf{y}|\mathbf{x}_1(m_1), \mathbf{x}_2(m_2))}{p(\mathbf{y}|\mathbf{x}_2(m_2))} \\
&\geq I(X_1; Y|X_2, U) - \frac{\sigma_1 Q^{-1}(\epsilon)}{\sqrt{n}} \\
i_2(\mathbf{x}_2(m_2); \mathbf{y}|\mathbf{x}_1(m_1)) &\triangleq \log \frac{W^n(\mathbf{y}|\mathbf{x}_1(m_1), \mathbf{x}_2(m_2))}{p(\mathbf{y}|\mathbf{x}_1(m_1))} \\
&\geq I(X_2; Y|X_1, U) - \frac{\sigma_2 Q^{-1}(\epsilon)}{\sqrt{n}} \\
i_3(\mathbf{x}_1(m_1), \mathbf{x}_2(m_2); \mathbf{y}) &\triangleq \log \frac{W^n(\mathbf{y}|\mathbf{x}_1(m_1), \mathbf{x}_2(m_2))}{p(\mathbf{y})} \\
&\geq I(X_1, X_2; Y|U) - \frac{\sigma_3 Q^{-1}(\epsilon)}{\sqrt{n}}
\end{aligned}
$$

Strong Converse

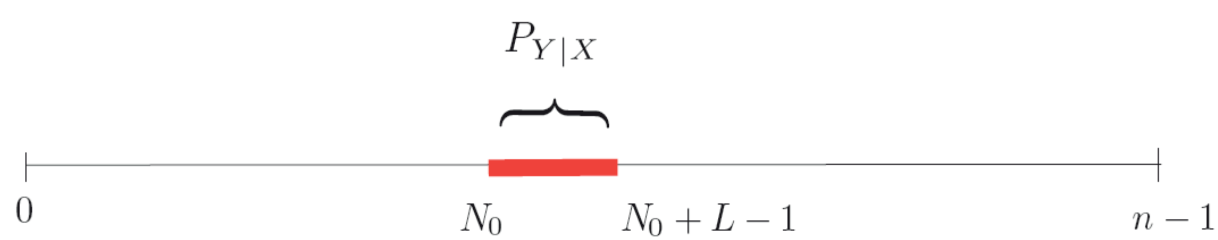
- Consider conditionally CC codes: let $\mathcal{U}$ be a finite auxiliary alphabet, and fix sequence $\mathbf{u} \in \mathcal{U}^n$ and joint type $\hat{P}_{X_1 X_2|U}$
- All codewords $(\mathbf{x}_1, \mathbf{x}_2)$ have the same conditional type $\hat{P}_{X_1 X_2|U}$
- Derive three inequalities

$$
\begin{aligned}
M_{1n} M_{2n} &\leq \inf_{Q_{Y|U}} \sup_{(\mathbf{x}_1, \mathbf{x}_2) \in T(\hat{P}_{X_1 X_2|U})} \frac{1}{\beta_{1-\epsilon}(\mathbf{x}_1, \mathbf{x}_2, Q_{Y|U})} \\
M_{1n} &\leq \inf_{Q_{Y|X_2 U}} \sup_{(\mathbf{x}_1, \mathbf{x}_2) \in T(\hat{P}_{X_1 X_2|U})} \frac{1}{\beta'_{1-\epsilon}(\mathbf{x}_1, \mathbf{x}_2, Q_{Y|X_2 U})} \\
M_{2n} &\leq \inf_{Q_{Y|X_1 U}} \sup_{(\mathbf{x}_1, \mathbf{x}_2) \in T(\hat{P}_{X_1 X_2|U})} \frac{1}{\beta''_{1-\epsilon}(\mathbf{x}_1, \mathbf{x}_2, Q_{Y|X_1 U})}
\end{aligned}
$$

- Refined large-deviations analysis $\Rightarrow \mathcal{C}_n(W, \epsilon) =$ outer region
- General codes $\Rightarrow$ add at most $|\mathcal{X}_1|\,|\mathcal{X}_2|\,|\mathcal{U}|\,\frac{\log n}{n}$ to the rates

Universal Coding for Nonstationary MAC

- Example: severe outage
  MAC $p_{Y|X_1X_2}$ is in effect between times N_0 and $N_0 + L - 1$, output is independent of input at all other times



- $N_0 \in \{1, L+1, 2L+1, \cdots\}$ is unknown to encoder & decoder
- Fix sequence $L = L(N)$ such that $1 \ll L \ll n$
- Clearly $\mathcal{C}(W) = \{(0,0)\}$
- Redefine code rates $R_{in} \triangleq \frac{1}{L} \log M_{in}$ for $i = 1, 2$

- Characterize ϵ-capacity region $\widetilde{\mathcal{C}}_{n,\epsilon}$ for blocklength n and effective transmission time L
- Use iid random codes with rate pair (R_{1n}, R_{2n})
- Use GLRT-like decoder
- **Theorem**:
  (a) If $(R_{1n}, R_{2n}) \in \widetilde{\mathcal{C}}_n(W, \epsilon)$ then $P_{e,n}(W) \leq \epsilon$
  (b) If $L \gg \log^2 n$ then $\widetilde{\mathcal{C}}_{n,\epsilon} = \mathcal{C}_{n,\epsilon}$
  (c) If $\log^2 n \gg L \gg \log n$ then $\widetilde{\mathcal{C}}_{n,\epsilon} \subset \mathcal{C}_{n,\epsilon}$
  (d) If $\frac{L}{\log n} \to R_0$ then $\widetilde{\mathcal{C}}_{n,\epsilon} = [\mathcal{C}_{n,\epsilon} - (R_0, R_0)]^+$
- Case (c) could be practical scenario (say $n = 1000$ and $L = 50$) in which case reduction of capacity region is significant

Conclusion

- Hail to Pierre-Simon Laplace
- For compound channels, the rate penalty for not knowing the channel is $\frac{d \log n}{2n}$
- Characterized capacity region for MAC with finite blocklength, up to $\frac{\log n}{n}$ terms
- Preliminary investigation of nonstationary MACs