Optimal Finite-Length Relay Matrix

Shenghao Yang

School of Science and Engineering
The Chinese University of Hong Kong, Shenzhen

Open Problem Session, BIID, 2026

Problem Formulation

- Mutual information of two random variables X and Y with distribution $p(x,y) = p_X(x)Q_{Y|X}(y|x)$:

$$I(X;Y) = I(p_X, Q_{Y|X})$$

- Our problem: for given stochastic matrices A and B,

$$\max_{p,\Phi} I(p, A\Phi B).$$

Problem Formulation

- Mutual information of two random variables X and Y with distribution $p(x,y) = p_X(x)Q_{Y|X}(y|x)$:

$$I(X;Y) = I(p_X, Q_{Y|X})$$

  - $I(X;Y)$ is a concave $\cap$ function of p_X for given $Q_{Y|X}$.
  - $I(X;Y)$ is a convex $\cup$ function of $Q_{Y|X}$ for given p_X.
- Our problem: for given stochastic matrices A and B,

$$\max_{p,\Phi} I(p, A\Phi B).$$

  - Optimized by deterministic Φ.

Optimal Finite-Length Relay

$$a \xrightarrow{Q} r \xrightarrow{Q} b$$

- Let Q be the stochastic matrix of a discrete memoryless channel (DMC).
- Suppose each channel is used n times. Then the stochastic matrix of the channel is $Q^{\oplus n} = Q \otimes Q \otimes \cdots \otimes Q$ (n times).
- If node r performs the relay operation Φ, the end-to-end channel from a to b is $Q^{\oplus n} \Phi Q^{\oplus n}$.
- The maximum reliable communication rate from a to b is

$$C_n = \frac{1}{n} \max_{p,\Phi} I(p, Q^{\oplus n} \Phi Q^{\oplus n}).$$

  - C_n is a non-decreasing function of n.
  - $C := \cancel{\lim_{n\to\infty} C_n = \max}_p\, I(p, Q)$, the min-cut.

A Special Case

Consider

$$C_n = \frac{1}{n} \max_{p,\Phi} I(p, Q^{\oplus n} \Phi Q^{\oplus n})$$

for the binary symmetric channel (BSC) with crossover probability ϵ, i.e.,

$$Q = \begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix}.$$

- $C = \lim_{n\to\infty} C_n = 1 - h(\epsilon)$, achieved by decode-and-forward.
- When $n = 1, 2, 3$, $C_n = 1 - h(2\epsilon(1-\epsilon))$, achieved by identity Φ.
- There must exist a value n_0 (depending on ϵ) such that for all $n \geq n_0$, identity Φ is NOT optimal.

A Special Case

Consider

$$C_n = \frac{1}{n} \max_{p,\Phi} I(p, Q^{\oplus n} \Phi Q^{\oplus n})$$

for the binary symmetric channel (BSC) with crossover probability ϵ, i.e.,

$$Q = \begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix}.$$

- $C = \lim_{n\to\infty} C_n = 1 - h(\epsilon)$, achieved by decode-and-forward.
  - For an (n, k) code, $\Phi = DE$, where D is a $2^n \times 2^k$ matrix and E is a $2^k \times 2^n$ matrix.
- When $n = 1, 2, 3$, $C_n = 1 - h(2\epsilon(1-\epsilon))$, achieved by identity Φ.
- There must exist a value n_0 (depending on ϵ) such that for all $n \geq n_0$, identity Φ is ~~NOT optimal~~.

Multihop Network

$$a \xrightarrow{Q} r_1 \xrightarrow{Q} r_2 \xrightarrow{Q} r_3 \xrightarrow{Q} b$$

- The achievable rate is $\frac{1}{n} I(p, Q^{\oplus n} \Phi_1 Q^{\oplus n} \cdots Q^{\oplus n} \Phi_\ell Q^{\oplus n})$, where Φ_i is the relay matrix for the relay node r_i.
- We know the asymptotic behavior:
  - When $n \to \infty$, the optimal rate is the capacity of the channel Q.
  - When $\ell \to \infty$, the optimal rate is the *zero-error capacity* of Q.
- Little is known about the finite-length behavior.
  - Are identical relay matrices optimal for the multihop network?
  - When identity relay matrices are optimal?

Lanni

$\rho^{\otimes n}$ vs $\sigma^{\otimes n}$: $\mathrm{Stein}(\rho^{iid} \| \sigma^{iid}) = D(\rho\|\sigma)$
$= \mathrm{Tr}\,\rho(\log\rho - \log\sigma)$

$\{\rho_1^{\otimes n}, \rho_2^{\otimes 2}\}$ vs $\sigma^{\otimes n}$: $\mathrm{Stein}(\{\rho_1,\rho_2\}^{iid} \| \sigma^{iid}) = \min_{i=1,2} D(\rho_i\|\sigma)$

① $\rho^{\otimes n}$ vs $\{\sigma_1^{\otimes n}, \sigma_2^{\otimes n}\}$: $\mathrm{Stein}(\rho^{iid} \| \{\sigma_1,\sigma_2\}^{iid}) = ?$

② Thm 3, 2510.06340

$$\boxed{= \lim_{n\to\infty} \frac{1}{n} D\left(\rho^{\otimes n} \,\Big\|\, \frac{\sigma_1^{\otimes n}+\sigma_2^{\otimes n}}{2}\right) = ?}$$

Naive : $\cancel{\overset{?}{=}}$ $<$ $\min_{i=1,2} D(\rho\|\sigma_i)$

2011.04645 :
$\exists\ \sigma_1,\sigma_2,\rho$:
$\mathrm{Stein}(\rho^{iid} \| \{\sigma_1,\sigma_2\}^{iid})$
$\overset{?}{=}$ $\leq$ $\min_{t\in[0,1]} D(\rho \| \underline{\sigma_1 \#_t \sigma_2})$
$\underline{<}$ $\min_i D(\rho\|\sigma_i)$

$$\sigma_1 \#_t \sigma_2 := \sigma_1^{1/2}(\sigma_1^{-1/2}\sigma_2\sigma_1^{-1/2})^t \sigma_1^{1/2}$$

Conjecture:

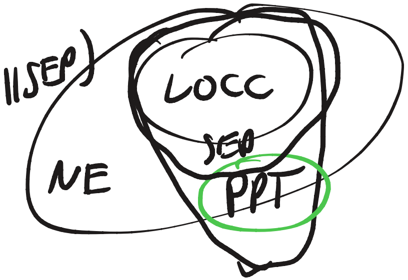
$$\lim_{n\to\infty} \frac{1}{n} D\left(\rho^{\otimes n} \,\Big\|\, \frac{\sigma_1^{\otimes n}+\sigma_2^{\otimes n}}{2}\right) \overset{?}{=} \min_t D(\rho\|\sigma_1 \#_t \sigma_2)$$

$\boxed{\text{Regula}}$

$\rho^{\otimes n} \xrightarrow{\varepsilon_n} \Psi_+^{\otimes R n}$ $\qquad \varepsilon_n \to 0$

iid $\qquad$ non-iid

$R(\rho \xrightarrow[NE]{} \Psi_+) = \mathrm{Stein}(\rho \| SEP) \underset{GQSL}{=} D^\infty(\rho \| SEP)$



$R(\rho \xrightarrow[PPT]{} \Psi_+) = \mathrm{Stein}(\rho \| R) = ?$

$\{ R = R^\dagger : \| R^\Gamma \|_1 \le 1 \}$

conjecture: $D^\infty(\rho \| R_+)$ $\leftarrow$ efficiently computable

$R_0(\rho \to \Psi_+)$

$D_H^\varepsilon(\rho \| R)$

$\varepsilon = 0$: $\quad D_H^0(\rho \| R) = -\log \min \{ \| Q^\Gamma \|_\infty : \Pi_\rho \le Q \le \mathbb{1} \}$

$\lim_{n\to\infty} \frac{1}{n} D_H^0(\rho^{\otimes n} \| R) = ?$

conjecture: $\quad -\log \min \{ \| Q^\Gamma \|_\infty : \Pi_\rho \le Q \}$

$\boxed{\text{Fang}}$

$\alpha(A_n, M_n) = \text{Type I}.$

$\beta(B_n, M_n) = \text{Type II}$

$$\beta_\varepsilon(A_n \| B_n) = \inf_{0 \le M_n \le I} \{\beta(B_n, M_n),\ \alpha(A_n, M_n) \le \varepsilon\}$$

GQSL: $\displaystyle \lim_{n\to\infty} -\frac{1}{n} \log \beta_\varepsilon(A_n \| B_n) = \boxed{D^\infty(A \| B)}$ ✓.

$$\boxed{A_n = \{\rho^{\otimes n}\}. \quad B_n = SEP_n}$$

$\forall\, \varepsilon, \alpha \in (0,1)$ $\quad D_H^\varepsilon(\rho \| \sigma) \ge D_{p,\alpha}(\rho \| \sigma) + \frac{\alpha}{1-\alpha} \log \frac{1}{\varepsilon}$ ✓

$$D_H^\varepsilon(A_n \| B_n) \ge D_{p,\alpha}(A_n \| B_n) + \frac{\alpha}{1-\alpha} \log \frac{1}{\varepsilon}$$

$$\lim_{n\to\infty} \frac{1}{n} D_H^\varepsilon(A_n \| B_n) \ge \lim_{n\to\infty} \frac{1}{n} D_{p,\alpha}(A_n \| B_n)$$

$$:= D^\infty_{p,\alpha}(A \| B) \quad \forall\, \alpha \in (0,1)$$

$A_n = \{\rho^{\otimes n}\}$, $B_n = \mathrm{SEP}$ $\longrightarrow$ $\geq \boxed{\sup_{\alpha \in (0,1)} D^{\infty}_{p.\alpha}(A\|B)}$

☆ **Conjecture:** $\sup_{\alpha \in (0,1)} D^{\infty}_{p.\alpha}(A\|B) \overset{?}{=} D^{\infty}(A\|B)$

① A_n, B_n convex compact

② $A_n \otimes A_m \subseteq A_{n+m}$.