

One-Shot Coding over General Noisy Networks

ISIT 2024

Yanxiao Liu and Cheuk Ting Li

The Chinese University of Hong Kong



Background

Some key questions/challenges in network information theory

- 1 **Blocklength** in information transmission: asymptoticity & finite blocklength & one-shot achievability?
- 2 **Noisy network coding**: capacity of noisy networks & coding schemes?
- 3 **Unified Coding Scheme**: channel coding & source coding & coding for computing?

Our contributions

- 1 **One-shot Coding**: A unified **one-shot** coding framework for communication and compression of messages among multiple nodes across a **general acyclic noisy network**.
 - Proof technique: exponential process refinement lemma
- 2 **Achievability Results** :
 - Novel one-shot achievability results discovered.
 - Existing one-shot & asymptotic results recovered.



Background: One-shot Information Theory

One-shot information theory

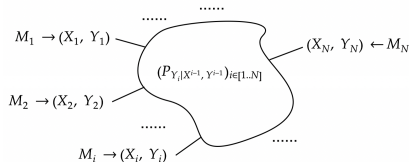
- ① **Conventional Shannon theory:** **asymptotic** information transmission rates based on the law of large numbers.
- ② **Finite-blocklength regime^a:** maximal information transmission rate at a given blocklength and error probability?
- ③ **One-shot achievability^b:** What if the blocklength can be as short as 1 (each source and channel is only used once)?
 - ① Sources and channels can be **arbitrary**: no need to be memoryless or ergodic.
 - ② Goal: can recover existing (first-order and second-order) asymptotic results when applied to memoryless sources and channels.

^aPolyanskiy, Yury, H. Vincent Poor, and Sergio Verdú. "Channel coding rate in the finite blocklength regime." IEEE Transactions on Information Theory 56, no. 5 (2010): 2307-2359.

^bLi, Cheuk Ting, and Venkat Anantharam. "A unified framework for one-shot achievability via the Poisson matching lemma." IEEE Transactions on Information Theory 67, no. 5 (2021): 2624-2651.



Background: Noisy Network Coding



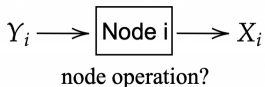
Noisy Network Coding

- **Noisy network coding**^a: communicating messages between **multiple** sources and destinations over a general **noisy** network.
- Generalizing:
 - 1 Noiseless network coding by Ahlswede, Cai, Li and Yeung.
 - 2 Compress-forward coding for relay channels by Cover and El Gamal.
 - 3 Coding for relay networks, coding for erasure networks, etc.

^aLim, Sung Hoon, Young-Han Kim, Abbas El Gamal, and Sae-Young Chung. "Noisy network coding." IEEE Transactions on Information Theory 57, no. 5 (2011): 3132-3152.



Background: A Unified Random Coding Bound



A Unified random coding bound

- ① Unified random coding bound^a: work for **any** combination of channel coding and source coding problems.
- ② Unifying and generalizing known relaying strategies; can yield bounds without error analysis.

^aLee, Si-Hyeon, and Sae-Young Chung. "A unified random coding bound." IEEE Transactions on Information Theory 64, no. 10 (2018): 6779-6802.

Overview: Our Contributions



Our contributions

- 1 A unified one-shot coding scheme
- 2 over general noisy acyclic discrete networks (ADN)
- 3 that is applicable to any combination of source coding, channel coding and coding for computing problems.

Special cases

- Novel one-shot achievability results for:
 - 1 One-shot relay channels
 - 2 One-shot primitive relay channels
 - Compress-and-forward bound
 - Partial-decode-and-forward bound
- Recovered one-shot & asymptotic results on:
 - 1 Source and channel coding
 - 2 Gelfand-Pinsker, Wyner-Ziv and coding for computing
 - 3 Multiple access channels
 - 4 Broadcast channels



Techniques

Poisson functional representation

- For a finite set \mathcal{U} , let $\mathbf{U} := (Z_u)_{u \in \mathcal{U}}$ be i.i.d. $\text{Exp}(1)$ random variables.
- Given a distribution P over \mathcal{U} , **Poisson functional representation**^a:

$$\mathbf{U}_P := \operatorname{argmin}_u \frac{Z_u}{P(u)} \quad (1)$$

- We have $\mathbf{U}_P \sim P$.
- Various applications: minimax learning, neural network compression, differential privacy, etc.

^aLi, Cheuk Ting, and Abbas El Gamal. "Strong functional representation lemma and applications to coding theorems." *IEEE Transactions on Information Theory* 64, no. 11 (2018): 6967-6978.

Techniques



Poisson functional representation

- Given a distribution P over \mathcal{U} , **Poisson functional representation**:

$$\mathbf{U}_P := \operatorname{argmin}_u \frac{Z_u}{P(u)}$$

Generalized Poisson matching lemma

- Let $\mathbf{U}_P(1), \dots, \mathbf{U}_P(|\mathcal{U}|) \in \mathcal{U}$ be the elements of \mathcal{U} sorted in ascending order of $Z_u/P(u)$, let $\mathbf{U}_P^{-1} : \mathcal{U} \rightarrow [|\mathcal{U}|]$ for the inverse function of $i \mapsto \mathbf{U}_P(i)$.
- Generalized Poisson matching lemma**^a: For distributions P, Q over \mathcal{U} , we have the following almost surely:

$$\mathbf{E} \left[\mathbf{U}_Q^{-1}(\mathbf{U}_P) \mid \mathbf{U}_P \right] \leq \frac{P(\mathbf{U}_P)}{Q(\mathbf{U}_P)} + 1.$$

^aLi, Cheuk Ting, and Venkat Anantharam. "A unified framework for one-shot achievability via the Poisson matching lemma." IEEE Transactions on Information Theory 67, no. 5 (2021): 2624-2651.

Techniques



Refining a distribution by an exponential process

- For a joint distribution $Q_{V,U}$ over $\mathcal{V} \times \mathcal{U}$, the **refinement** of $Q_{V,U}$ by \mathbf{U} , denoted as $Q_{V,U}^{\mathbf{U}}$, is a joint distribution

$$Q_{V,U}^{\mathbf{U}}(v, u) := \frac{Q_V(v)}{\left(\mathbf{U}_{Q_{U|V}^{-1}(\cdot|v)}(u) \sum_{i=1}^{|\mathcal{U}|} i^{-1} \right)} \quad (2)$$

for all (v, u) in the support of $Q_{V,U}$, where Q_V is the V -marginal of $Q_{V,U}$ and $Q_{U|V}$ is the conditional distribution of U given V .

- The refinement $Q_U^{\mathbf{U}}(u)$ is for the **soft decoding** of U , which gives a distribution over U , with \mathbf{U}_{Q_U} having the largest probability.
- Useful in non-unique decoding.
- If the distribution $Q_{V,U}$ represents our “**prior distribution**” of (V, U) , then the refinement $Q_{V,U}^{\mathbf{U}}$ is our updated “**posterior distribution**” after taking the exponential process \mathbf{U} into account.



Techniques

Exponential Process Refinement Lemma

- For a distribution P over \mathcal{U} and a joint distribution $Q_{V,U}$ over a finite $\mathcal{V} \times \mathcal{U}$, for every $v \in \mathcal{V}$, we have, almost surely,

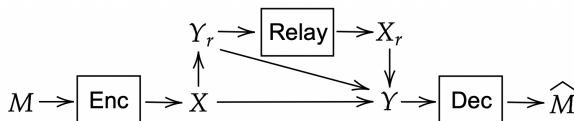
$$\mathbf{E} \left[\frac{1}{Q_{V,U}^{\mathbf{U}}(v, \mathbf{U}_P)} \middle| \mathbf{U}_P \right] \leq \frac{\ln |\mathcal{U}| + 1}{Q_V(v)} \left(\frac{P(\mathbf{U}_P)}{Q_{U|V}(\mathbf{U}_P|v)} + 1 \right). \quad (3)$$

Purpose

It keeps track of the evolution of the “posterior probability” of the correct values of a large number of random variables through the refinement process.



One-Shot Relay Channel

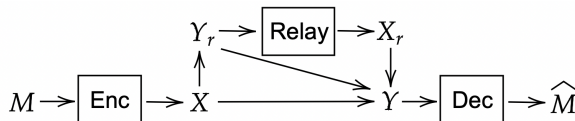


One-Shot Relay Channel

- Inherent limitation of one-shot setting: unable to model cyclic networks, e.g., two-way communication channels, conventional relay channels that depend on its **past**.
- One-shot version of **relay-with-unlimited-look-ahead**^a.
- “Best one-shot approximation” of the conventional relay channel.

^aEl Gamal, Abbas, Navid Hassanpour, and James Mammen. "Relay networks with delays." IEEE Transactions on Information Theory 53, no. 10 (2007): 3413-3431.

One-Shot Relay Channel

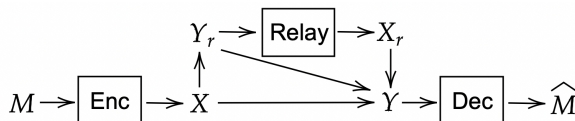


One-Shot Relay Channel

- ① Encoder observes $M \sim \text{Unif}[L]$ and outputs X , which is passed through the channel $P_{Y_r|X}$.
- ② Relay observes Y_r and outputs X_r .
- ③ (X, X_r, Y_r) is passed through the channel $P_{Y|X, X_r, Y_r}$.
 - Y depends on all of X, X_r, Y_r and X_r may interfere with (X, Y_r) .
- ④ Decoder observes Y and recovers \hat{M} .

Practical to scenarios where the relay outputs X_r instantaneously or the channel has a long memory, or it is a storage device.

One-Shot Relay Channel



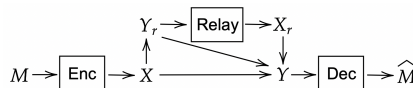
Corollary ((One-Shot Achievable Bound))

For any P_X , $P_{U|Y_r}$, function $x_r(y_r, u)$, there is a coding scheme for the one-shot relay channel such that the error probability satisfies

$$P_e \leq \mathbf{E} \left[\min \left\{ \gamma L 2^{-\iota(X;U,Y)} \left(2^{-\iota(U;Y) + \iota(U;Y_r)} + 1 \right), 1 \right\} \right],$$

where $(X, Y_r, U, X_r, Y) \sim P_X P_{Y_r|X} P_{U|Y_r} \delta_{x_r(Y_r, U)} P_{Y|X, Y_r, X_r}$, and $\gamma := \ln |\mathcal{U}| + 1$.

One-Shot Relay Channel



Proof

- ① “Random codebooks” $\mathbf{U}_1, \mathbf{U}_2$: independent exponential processes.
- ② Encoder: $U_1 = (\mathbf{U}_1)_{P_{U_1} \times \delta_M}$, then outputs X -component of U_1 .
- ③ Relay: $U_2 = (\mathbf{U}_2)_{P_{U_2|Y_r}(\cdot|Y_r)}$, then outputs $X_r = x_r(Y_r, U_2)$.
- ④ Decoder observes Y , and:
 - Refine $P_{U_2|Y}(\cdot|Y)$ to $Q_{U_2} := P_{U_2|Y}^{U_2}$. By Exponential Process Refinement Lemma:

$$\mathbb{E} \left[\frac{1}{Q_{U_2}(U_2)} \middle| U_2, Y, Y_r \right] \leq (\ln |\mathcal{U}_2| + 1) \left(\frac{P_{U_2|Y_r}(U_2)}{P_{U_2|Y}(U_2)} + 1 \right).$$

- Compute $Q_{U_2} P_{U_1|U_2, Y}$ over $\mathcal{U}_1 \times \mathcal{U}_2$, and let its U_1 -marginal be \tilde{Q}_{U_1} .
- Let $\tilde{U}_1 = (\mathbf{U}_1)_{\tilde{Q}_{U_1} \times P_M}$, and output its M -component.



One-Shot Relay Channel

Proof

$$\begin{aligned}
 & \mathbf{P}(\tilde{U}_1 \neq U_1 \mid X, Y_r, U_2, X_r, Y, M) \\
 & \stackrel{(a)}{\leq} \mathbf{E} \left[\min \left\{ \frac{P_{U_1}(U_1) \delta_M(M)}{P_{U_1|U_2,Y}(U_1|U_2, Y) Q_{U_2}(U_2) P_M(M)}, 1 \right\} \mid X, Y_r, U_2, X_r, Y, M \right] \\
 & \stackrel{(b)}{=} \mathbf{E} \left[\min \left\{ L \frac{P_{U_1}(U_1)}{P_{U_1|U_2,Y}(U_1|U_2, Y) Q_{U_2}(U_2)}, 1 \right\} \mid X, Y_r, U_2, X_r, Y, M \right] \\
 & \stackrel{(c)}{\leq} \min \left\{ L \frac{P_{U_1}(U_1)}{P_{U_1|U_2,Y}(U_1|U_2, Y)} (\ln |U_2| + 1) \left(\frac{P_{U_2|Y_r}(U_2)}{P_{U_2|Y}(U_2)} + 1 \right), 1 \right\} \\
 & = \min \left\{ (\ln |U_2| + 1) L 2^{-\iota(X; U_2, Y)} (2^{-\iota(U_2; Y) + \iota(U_2; Y_r)} + 1), 1 \right\}.
 \end{aligned}$$

(a) is by the Poisson matching lemma; (b) is by $\delta_M(M) = 1$, $P_M(M) = 1/L$; (c) is by the refinement step (previous page) and Jensen's inequality.

For some $P_{U|Y_r}$ and function $x_r(y_r, u_2)$, it yields the asymptotic achievable rate:

$$R \leq I(X; U, Y) - \max \{ I(U; Y_r) - I(U; Y), 0 \}.$$



One-Shot Primitive Relay Channel

$Y = (Y', Y'')$ and the channel $P_{Y|X, X_r, Y_r} = P_{Y'|X, Y_r} P_{Y''|X_r}$ can be decomposed into two orthogonal components.

Corollary

For any $P_X, P_{X_r}, P_{U'|Y_r}$, there is a coding scheme for the one-shot primitive relay channel with $M \sim \text{Unif}[\mathcal{L}]$ such that

$$P_e \leq \mathbf{E} \left[\min \left\{ (\ln(|\mathcal{U}'||\mathcal{X}_r|) + 1) L 2^{-\iota(X; U', Y')} (2^{-\iota(X_r; Y'') + \iota(U'; Y_r | Y')} + 1), 1 \right\} \right],$$

$(X, Y_r, U', Y') \sim P_X P_{Y_r|X} P_{U'|Y_r} P_{Y'|X, Y_r}$ independent of $(X_r, Y'') \sim P_{X_r} P_{Y''|X_r}$.

- $C_r = \max_{P_{X_r}} I(X_r; Y'')$ is the capacity of channel $P_{Y''|X_r}$, then the asymptotic achievable rate is $R \leq I(X; U', Y') - \max\{I(U'; Y_r | Y') - C_r, 0\}$.
- It recovers the **compress-and-forward bound**^a.

^aKim, Young-Han. "Coding techniques for primitive relay channels." In Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput, p. 2007. 2007.



One-Shot Relay Channel

Corollary (Partial-Decode-and-Forward Bound)

Fix any $P_{X,V}$, $P_{U|Y_r,V}$, function $x_r(y_r, u, v)$, and J which is a factor of L . There exists a deterministic coding scheme for the one-shot relay channel with

$$P_e \leq \mathbf{E} \left[\min \left\{ J 2^{-\iota(V; Y_r)} + (\ln(J|\mathcal{U}|) + 1)(\ln(J|\mathcal{V}|) + 1) L J^{-1} 2^{-\iota(X; U, Y|V)} \right. \right. \\ \left. \left. \cdot (2^{-\iota(U; V, Y) + \iota(U; V, Y_r)} + 1) (J 2^{-\iota(V; Y)} + 1), 1 \right\} \right],$$

where $(X, V, Y_r, U, X_r, Y) \sim P_{X,V} P_{Y_r|X,V} P_{U|Y_r,V} \delta_{x_r(Y_r, U, V)} P_{Y|X, Y_r, X_r}$.

Split the message: $M \sim \text{Unif}[L] \Rightarrow M_1 \sim \text{Unif}[J], M_2 \sim \text{Unif}[L/J]$.

- ① Node 1 sees $Y_1 = M_1$, outputs $X_1 = V$, and has an auxiliary $U_1 = (M_1, V)$
- ② Node 2 sees $Y_2 = (M_1, M_2, V)$, outputs $X_2 = X$, and $U_2 = (M_1, M_2, X)$.
- ③ Relay sees $Y_3 = Y_r$, decodes U_1 , outputs $X_3 = X_r$, and $U_3 = (M_1, U)$.
- ④ Decoder sees $Y_4 = Y$ and uses the decoding order “ $U_2, U_3?, U_1?$ ”.

It recovers existing asymptotic partial-decode-and-forward Bounds.

General Acyclic Discrete Network



Acyclic discrete network (ADN)

- Nodes are labelled by $1, \dots, N$; node i sees $Y_i \in \mathcal{Y}_i$ and produces $X_i \in \mathcal{X}_i$.
- Y_i depends on all previous inputs and outputs X^{i-1}, Y^{i-1} .
- ADN**: a collection of channels $(P_{Y_i|X^{i-1}, Y^{i-1}})_{i \in [N]}$, where $P_{Y_i|X^{i-1}, Y^{i-1}}$ is a conditional distribution from $\prod_{j=1}^{i-1} \mathcal{X}_j \times \prod_{j=1}^{i-1} \mathcal{Y}_j$ to \mathcal{Y}_i .

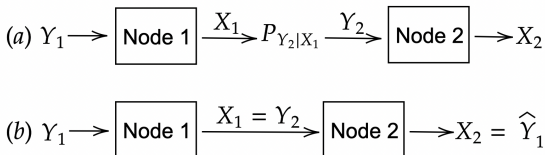


Figura 1: (a) Channel coding. (b) Source coding.



Coding scheme

Public-randomness coding scheme

A pair $(P_W, (f_i)_{i \in [N]})$, where P_W is the distribution of the public randomness $W \in \mathcal{W}$ available to all nodes and $f_i : \mathcal{Y}_i \times \mathcal{W} \rightarrow \mathcal{X}_i$ is the encoding function of node i mapping its observation Y_i and the public randomness W to its output X_i . The operations are as follows. First, generate $W \sim P_W$. For $i = 1, \dots, N$, generate \tilde{Y}_i conditional on $\tilde{X}^{i-1}, \tilde{Y}^{i-1}$ according to $P_{Y_i|X^{i-1}, Y^{i-1}}$, and take $\tilde{X}_i = f_i(\tilde{Y}_i, W)$.

Deterministic coding scheme

A sequence of encoding functions $(f_i)_{i \in [N]}$, where $f_i : \mathcal{Y}_i \rightarrow \mathcal{X}_i$. For $i = 1, \dots, N$, the following operations are performed:

- **Noisy channel.** The output \tilde{Y}_i is generated conditional on $\tilde{X}^{i-1}, \tilde{Y}^{i-1}$ according to $P_{Y_i|X^{i-1}, Y^{i-1}}$. For $i = 1$, $\tilde{Y}_1 \sim P_{Y_1}$ can be regarded as a source or a channel state.
- **Node operation.** Node i observes \tilde{Y}_i and outputs $\tilde{X}_i = f_i(\tilde{Y}_i)$.

Coding scheme



Achievability

- \tilde{X}_i, \tilde{Y}_i denote the **actual** random variables from the coding scheme.
- X_i, Y_i denote the random variables following an **ideal** distribution.
- The goal (the “**achievability**”) is to make the actual joint distribution $P_{\tilde{X}^N, \tilde{Y}^N}$ “approximately as good as” the ideal joint distribution P_{X^N, Y^N} . For an “error set” $\mathcal{E} \subseteq (\prod_{i=1}^N \mathcal{X}_i) \times (\prod_{i=1}^N \mathcal{Y}_i)$ that we do not want $(\tilde{X}^N, \tilde{Y}^N)$ to fall into, we want

$$\mathbf{P}((\tilde{X}^N, \tilde{Y}^N) \in \mathcal{E}) \lesssim \mathbf{P}((X^N, Y^N) \in \mathcal{E}), \quad (4)$$

which can be guaranteed by $P_{\tilde{X}^N, \tilde{Y}^N}$ being close to P_{X^N, Y^N} in TV distance.



Main Theorem

Theorem

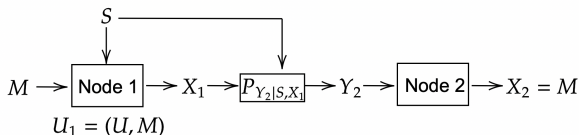
Fix any ADN $(P_{Y_i|X^{i-1}, Y^{i-1}})_{i \in [M]}$. For any collection of indices $(a_{i,j})_{i \in [M], j \in [d_i]}$ where $(a_{i,j})_{j \in [d_i]}$ is a sequence of distinct indices in $[i-1]$ for each i , any sequence $(d'_i)_{i \in [M]}$ with $0 \leq d'_i \leq d_i$ and any collection of conditional distributions $(P_{U_i|Y_i, \bar{U}'_i}, P_{X_i|Y_i, U_i, \bar{U}'_i})_{i \in [M]}$ (where $\bar{U}_{i,S} := (U_{a_{i,j}})_{j \in S}$ for $S \subseteq [d_i]$ and $\bar{U}'_i := \bar{U}_{i,[d'_i]}$), which induces the joint distribution of X^N, Y^N, U^N (the “ideal distribution”), there exists a public-randomness coding scheme $(P_W, (f_i)_{i \in [M]})$ such that the joint distribution of \tilde{X}^N, \tilde{Y}^N induced by the scheme satisfies

$$\delta_{\text{TV}}(P_{X^N, Y^N}, P_{\tilde{X}^N, \tilde{Y}^N}) \leq \mathbf{E} \left[\min \left\{ \sum_{i=1}^N \sum_{j=1}^{d'_i} B_{i,j}, 1 \right\} \right],$$

where $\gamma_{i,j} := \prod_{k=j+1}^{d_i} (\ln |\mathcal{U}_{a_{i,k}}| + 1)$ and

$$B_{i,j} := \gamma_{i,j} \prod_{k=j}^{d_i} (2^{-\ell(\bar{U}_{i,k}; \bar{U}_{i,[d_i] \setminus [j..k]}, Y_i) + \ell(\bar{U}_{i,k}; \bar{U}'_{a_{i,k}}, Y_{a_{i,k}})} + \mathbf{1}\{k > j\}).$$

ADN: Gelfand-Pinsker Problem



Gelfand-Pinsker Problem

- **ADN**: $Y_1 := (M, S)$, $Y_2 := Y$, $P_{Y_2|Y_1,X_1}$ be $P_{Y|S,X}$, and $X_2 := M$.
- **Auxiliary** on node 1: $U_1 = (U, M)$ for some U following $P_{U|S}$ given S .
- **Decoding order**: on node 2 “ U_1 ” (i.e., it only wants U_1).

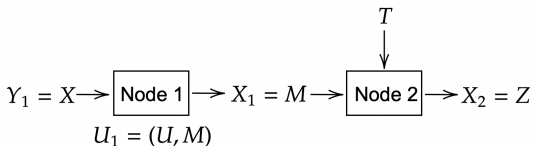
Corollary (Gelfand-Pinsker)

Fix $P_{U|S}$ and function $x : \mathcal{U} \times \mathcal{S} \rightarrow \mathcal{X}$. There exists a coding scheme for the channel $P_{Y|X,S}$ with $S \sim P_S$, $M \sim \text{Unif}[L]$ such that

$$P_e \leq \mathbf{E} \left[\min \left\{ L 2^{-\iota(U;Y) + \iota(U;S)}, 1 \right\} \right],$$

where $S, U, X, Y \sim P_S P_{U|S} \delta_{x(U,S)} P_{Y|X,S}$.

ADN: Wyner-Ziv Problem and Coding for Computing



Corollary (Wyner-Ziv)

Fix $P_{U|X}$ and function $z : \mathcal{U} \times \mathcal{Y} \rightarrow \mathcal{Z}$. There exists a coding scheme s.t.

$$P_e \leq \mathbf{E} \left[\min \left\{ \mathbf{1}\{d(X, Z) > D\} + L^{-1} 2^{-\iota(U; T) + \iota(U; X)}, 1 \right\} \right],$$

where $X, Y, U, Z \sim P_X P_{Y|X} P_{U|X} \delta_{Z(U, Y)}$.

Coding for Computing

- Coding for computing: node 2 recovers a function $f(X, T)$,
 $P_e \leq \mathbf{E}[\min\{\mathbf{1}\{d(f(X, T), Z) > D\} + L^{-1} 2^{-\iota(U; T) + \iota(U; X)}, 1\}].$



ADN: Multiple Access Channel

Multiple Access Channel

- ADN: $Y_1 := M_1$, $Y_2 := M_2$, $Y_3 := Y$ and $X_3 := (M_1, M_2)$.
- Auxiliaries: $U_1 := (X_1, M_1)$ and $U_2 := (X_2, M_2)$.
- Decoding order of node 3: “ U_2, U_1 ” (i.e., decode U_1 (soft), and then U_2 (unique), and then U_1 (unique)).

Corollary

Fix P_{X_1}, P_{X_2} . There exists a coding scheme for the multiple access channel $P_{Y|X_1, X_2}$ with

$$P_e \leq \mathbf{E} \left[\min \left\{ \gamma L_1 L_2 2^{-\iota(X_1, X_2; Y)} + \gamma L_2 2^{-\iota(X_2; Y|X_1)} + L_1 2^{-\iota(X_1; Y|X_2)}, 1 \right\} \right],$$

where $\gamma := \ln(L_1 | \mathcal{X}_1 |) + 1$, $(X_1, X_2, Y) \sim P_{X_1} P_{X_2} P_{Y|X_1, X_2}$.

Asymptotic region: $R_1 < I(X_1; Y|X_2)$, $R_2 < I(X_2; Y|X_1)$, $R_1 + R_2 < I(X_1, X_2; Y)$.



Summary

Summary

- We provide a **unified one-shot coding framework** for communication and compression over general noisy networks.
- We design a proof technique “**exponential process refinement lemma**” that can keep track of a large number of auxiliary random variables.
- We provide **novel one-shot results** for various multi-hop settings.
- We recover most of the best-known one-shot results of different settings.

Future Directions

- Continuous case of the current framework.
- One-shot versions of other asymptotic bounds for relay channels.
- A unified coding scheme is useful to design automated theorem proving tools, e.g., PSITIP^a. Extensions to one-shot results is left for future study.

^aLi, Cheuk Ting. "An automated theorem proving framework for information-theoretic results." IEEE Transactions on Information Theory (2023).