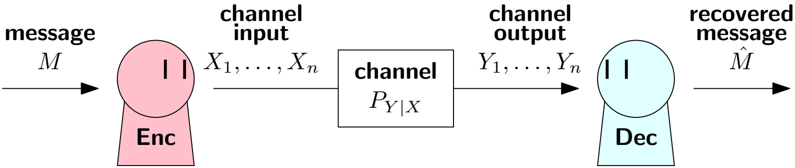
One-shot Coding using Poisson Processes

Cheuk Ting Li
Dept. of Information Engineering, Chinese University of Hong Kong
Email: ctli@ie.cuhk.edu.hk

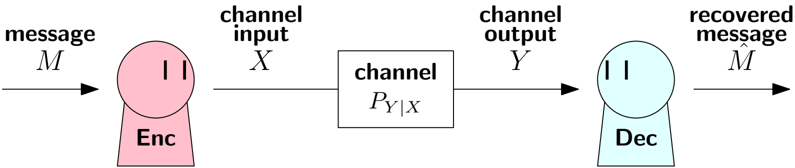
Based on joint work with Prof. Venkat Anantharam (UC Berkeley):
C. T. Li and V. Anantharam, "A Unified Framework for One-Shot Achievability via the Poisson Matching Lemma," IEEE Trans. Inf. Theory, vol. 67, no. 5, pp. 2624-2651, 2021.

Asymptotic Channel Coding



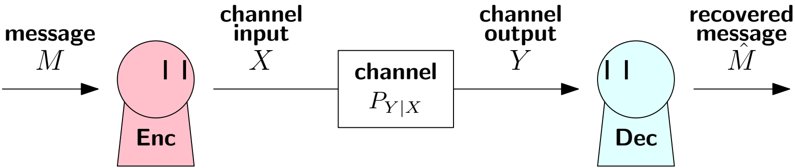
- Message $M \in \{0,1\}^{nR}$, R is the communication rate
- Encoder sees M, outputs $X_1, \ldots, X_n$, sends through noisy channel
- Decoder sees $Y_1, \ldots, Y_n$, outputs $\hat{M}$, wants $\hat{M} = M$ with high prob.
- Channel coding theorem [Shannon, 1948]: As $n \to \infty$, maximum communication rate is $C = \max_{p_X} I(X;Y)$
- But we never have $n \to \infty$ in practice!
  - Many applications (e.g. IoT) involve sending short packets
  - For large n, decoder must wait for a long delay
- Finite-blocklength analysis – study the case where n is finite

One-shot Channel Coding

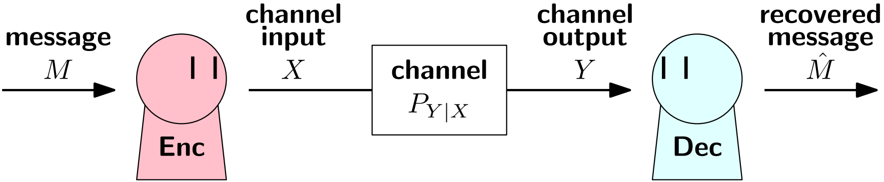


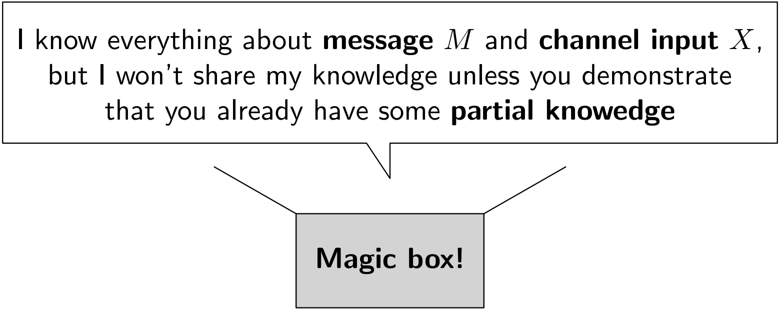
- One-shot – the extreme case $n = 1$ where only one X is sent
  - Point-to-point: [Feinstein, 1954, Shannon, 1957, Hayashi, 2009, Polyanskiy et al., 2010]
  - Multiuser: [Verdú, 2012, Yassaee et al., 2013b, Watanabe et al., 2015]
- **Misconception 1:** One-shot is too restrictive since n must be 1!
  - Wrong! One-shot is actualy the **most general** setting since we can substitute X to be a sequence, an image, a graph, or any object
  - Good one-shot results should readily subsume asymptotic (first/second-order) results
- **Misconception 2:** One-shot results are troublesome to prove!
  - Wrong! One-shot result can sometimes be just as simple (or even simpler) to prove as asymptotic results!

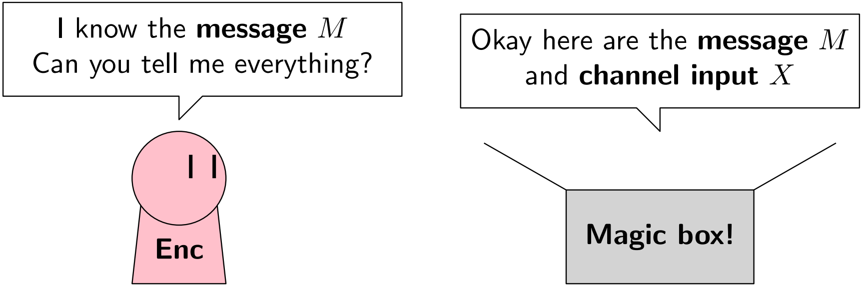
What is a Coding Scheme?

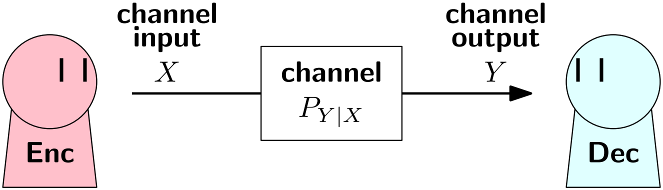


- **Conventional view:** A coding scheme is a pair of functions (f, g):
  - Encoding function f maps M to $X = f(M)$
  - Decoding function g maps Y to $\hat{M} = g(Y)$
- **Alternative view:** A coding scheme is a mechanism that can produce the pair (M, X), when provided partial knowledge about (M, X)
  - Encoder has partial knowledge on (M, X) since it knows M, and uses encoding function f to gain full knowledge on (M, X)
  - Decoder has partial knowledge on (M, X) since it knows Y (depends on X), and uses decoding function g to gain full knowledge on M, and hence X
- The alternative view provides a simple and unified way to construct coding schemes

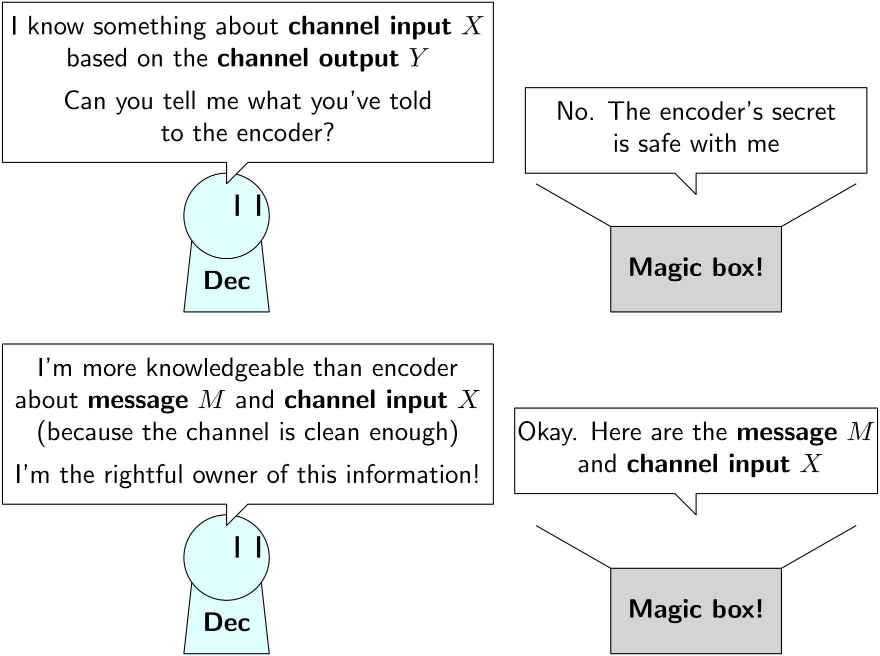

message
M
Enc
channel
input
X
channel
$P_{Y|X}$
channel
output
Y
Dec
recovered
message
$\hat{M}$


I know everything about **message** M and **channel input** X,
but I won't share my knowledge unless you demonstrate
that you already have some **partial knowedge**
Magic box!

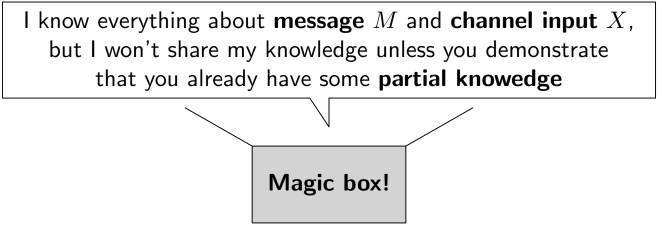

I know the **message** M
Can you tell me everything?
Enc
Okay here are the **message** M
and **channel input** X
Magic box!



channel
input
X
channel
$P_{Y|X}$
channel
output
Y
Enc
Dec

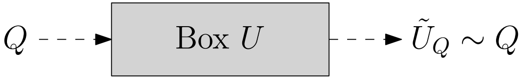

I know something about **channel input** X based on the **channel output** Y
Can you tell me what you've told to the encoder?
Dec
No. The encoder's secret is safe with me
Magic box!
I'm more knowledgeable than encoder about **message** M and **channel input** X (because the channel is clean enough)
I'm the rightful owner of this information!
Dec
Okay. Here are the **message** M and **channel input** X
Magic box!

What is this magic box? (Informal)



- **Query:** Input partial knowledge, output full knowledge
  - Encoder has partial knowledge on (M, X) since it only knows M
  - Decoder has partial knowledge on (M, X) as it knows Y (depends on X)
- **Guarantee:** If box is queried twice, two outputs will likely be the same if second partial knowledge is better than first partial knowledge
  - Encoder's partial knowledge on (M, X) is better if message M is longer
  - Decoder's partial knowledge on (M, X) is better if Y is more dependent on X, i.e., channel capacity is larger
  - Decoder's partial knowledge better than encoder's partial knowledge $\iff$ Channel capacity $>$ Length of M

What is this magic box? (Formal)



- A box about the random variable U can output samples of U
- **Query:** Input distribution Q, output a sample $\tilde{U}_Q \sim Q$
  - Q represents the party's partial knowledge on U
  - Box cannot just generate a fresh sample following Q for each query
  - Want box to have "memory"
- **Guarantee:** For distributions P, Q,

$$\mathbf{P}\left(\tilde{U}_Q \neq \tilde{U}_P \,\middle|\, \tilde{U}_P\right) \leq \frac{dP}{dQ}(\tilde{U}_P)$$

  - $P(\tilde{U}_P)$ measures how good the first partial knowdge P is about $\tilde{U}_P$
  - $\frac{dP}{dQ}(\tilde{U}_P)$ small means second partial knowledge Q better than P

What *really* is this magic box? Discrete case – Exponential Functional Representation

- Studied in Li and El Gamal [2018], Li and Anantharam [2021]
  - Related to the Gumbel-max trick [Huijben et al., 2022]
- Assume $U \in \{1, \ldots, k\}$ is discrete
- Box contains $Z_1, \ldots, Z_k \stackrel{iid}{\sim} \mathrm{Exp}(1)$
  - $Z_1, \ldots, Z_k$ generated at the time of creation of the box
- Input distribution P, output

$$\tilde{U}_P := \arg\min_u \frac{Z_u}{P(u)}$$

- Property of exponential RVs: $\tilde{U}_P \sim P$
- Poisson matching lemma [Li and Anantharam, 2021]:

$$\mathbf{P}(\tilde{U}_Q \neq \tilde{U}_P \mid \tilde{U}_P) \leq \frac{P(\tilde{U}_P)}{Q(\tilde{U}_P)}$$

Poisson Matching Lemma [Li and Anantharam, 2021]

Lemma

For finite discrete distributions P, Q,

$$\mathbf{P}(\tilde{U}_Q \neq \tilde{U}_P \mid \tilde{U}_P) \leq \frac{P(\tilde{U}_P)}{Q(\tilde{U}_P)}$$

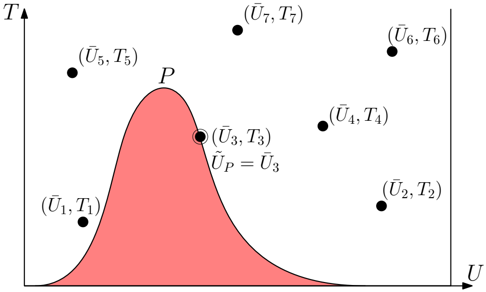
- Fact 1: If $T_i \sim \mathrm{Exp}(\gamma_i)$ indep. across $i = 1, \ldots, k$, then $\min_i T_i \sim \mathrm{Exp}(\sum_i \gamma_i)$ indep. of $\arg\min_i T_i$ with $\mathbf{P}(\arg\min_i T_i = i^*) = \frac{\gamma_{i^*}}{\sum_i \gamma_i}$
- Let $W_u := Z_u / P(u) \sim \mathrm{Exp}(P(u))$
  - $\tilde{U}_P := \arg\min_u W_u \sim P$ indep. of $W_{\tilde{U}_P} \sim \mathrm{Exp}(\sum_u P(u) = 1)$ (Fact 1)
- Conditional on the event A: "$\tilde{U}_P = u^*$, $W_{u^*} = w$":
  - Same as the event $\min_u W_u = W_{u^*} = w$
  - The conditional distribution of W_u $(u \neq u^*)$ given A is $\mathrm{Exp}(P(u)) + w$ by memoryless property of Exp since we know $W_u \geq w$
- Hence conditional on $\tilde{U}_P = u^*$, we have $W_u - W_{u^*}$ follows $\mathrm{Exp}(P(u))$ indep. across $u \neq u^*$, indep. of $W_{u^*} \sim \mathrm{Exp}(1)$

Poisson Matching Lemma [Li and Anantharam, 2021]

- Conditional on $\tilde{U}_P = u^*$, we have $W_u - W_{u^*}$ follows $\mathrm{Exp}(P(u))$ indep. across $u \neq u^*$, indep. of $W_{u^*} \sim \mathrm{Exp}(1)$

$$
\begin{aligned}
&\mathbf{P}(\tilde{U}_Q = u^* \mid \tilde{U}_P = u^*) \\
&= \mathbf{P}\left(\min_{u \neq u^*} \frac{Z_u}{Q(u)} \geq \frac{Z_{u^*}}{Q(u^*)} \,\middle|\, \tilde{U}_P = u^* \right) \\
&= \mathbf{P}\left(\min_{u \neq u^*} \frac{W_u P(u)}{Q(u)} \geq \frac{W_{u^*} P(u^*)}{Q(u^*)} \,\middle|\, \tilde{U}_P = u^* \right) \\
&\geq \mathbf{P}\left(\min_{u \neq u^*} \frac{(W_u - W_{u^*}) P(u)}{Q(u)} \geq \frac{W_{u^*} P(u^*)}{Q(u^*)} \,\middle|\, \tilde{U}_P = u^* \right) \\
&= \frac{Q(u^*)/P(u^*)}{Q(u^*)/P(u^*) + \sum_{u \neq u^*} Q(u)} \qquad \text{(Fact 1)} \\
&= \frac{Q(u^*)/P(u^*)}{Q(u^*)/P(u^*) + 1 - Q(u^*)} \\
&\geq 1 - \frac{P(u^*)}{Q(u^*)} \qquad\qquad Q.E.D.
\end{aligned}
$$

What *really* is this magic box?
General case – Poisson Functional Representation

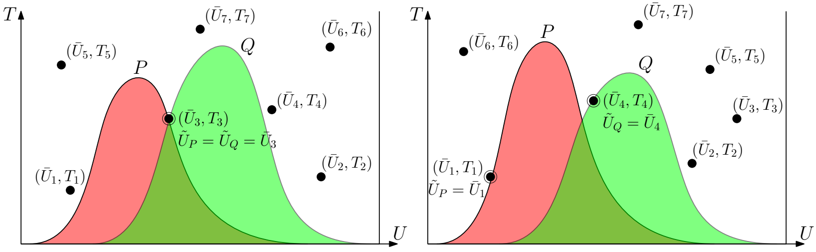


- Generalizes the exponential functional representation to general distributions (not only discrete)
  - Studied in Li and El Gamal [2018], Li and Anantharam [2021]
- Box: $\{\bar{U}_i, T_i\}_{i\in\mathbb{N}}$ points of Poisson process, intensity measure $\mu \times \lambda_{\mathbb{R}_{\geq 0}}$
- Input distribution P, output

$$K := \arg\min_i T_i \left(\frac{dP}{d\mu}(\bar{U}_i)\right)^{-1}, \ \tilde{U}_P := \bar{U}_K$$

- Mapping theorem: $\tilde{U}_P \sim P$

Guarantee – Poisson Matching Lemma



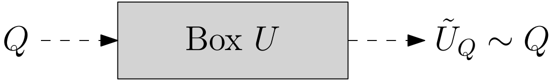
- Two distributions P, Q. Points selected are $\tilde{U}_P, \tilde{U}_Q$ resp.
- Poisson matching lemma [Li and Anantharam, 2021]:

$$\mathbf{P}\left(\tilde{U}_Q \neq \tilde{U}_P \,\middle|\, \tilde{U}_P\right) \leq \frac{dP}{dQ}(\tilde{U}_P)$$

$$= \frac{P(\tilde{U}_P)}{Q(\tilde{U}_P)} \quad \text{(for discrete } P, Q\text{)}$$

  - Refer to [Li and Anantharam, 2021] for proof

... enough peeking into the box. Now let's close it

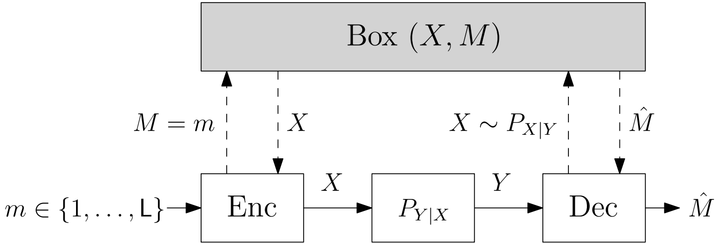


- We only use the box as a black box
  - Forget about exponential RVs, Poisson processes, etc.
- **Query:** Input distribution Q, output a sample $\tilde{U}_Q \sim Q$
  - Unifies both encoder and decoder
- **Guarantee:** For distributions P, Q,

$$\mathbf{P}\left(\tilde{U}_Q \neq \tilde{U}_P \,\middle|\, \tilde{U}_P \right) \leq \frac{dP}{dQ}(\tilde{U}_P)$$

  - Unifies both packing and covering lemmas

One-shot Channel Coding



- Box about (X, M)
- Encoder queries using $P_X \times \delta_m$ (i.e., the info $M = m$), gets X
- Decoder queries using $P_{X|Y} \times P_M$ (i.e., the info $X \sim P_{X|Y}$), gets $\hat{M}$

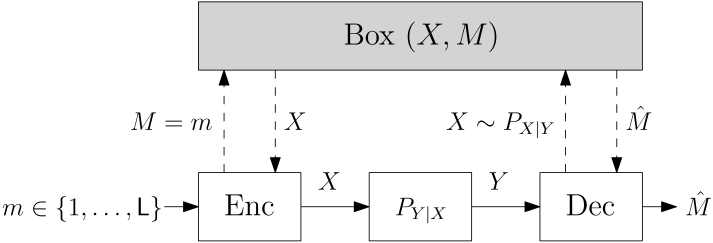
$$\mathbf{P}(M \neq \hat{M}) = \mathbf{E}\left[\mathbf{P}\left(M \neq \hat{M} \middle| M, X, Y\right)\right]$$

$$\leq \mathbf{E}\left[\min\left\{\frac{dP_X \times \delta_M}{dP_{X|Y}(\cdot|Y) \times P_M}, 1\right\}\right] = \mathbf{E}\left[\min\left\{\mathsf{L}2^{-\iota_{X;Y}(X;Y)}, 1\right\}\right]$$

where $\iota_{X;Y}$ is the information density:

$$\iota_{X;Y}(x;y) = \log\frac{dP_{X,Y}}{d(P_X \times P_Y)}(x,y) = \log\frac{P(x,y)}{P(x)P(y)} \text{ (for discrete } X, Y)$$

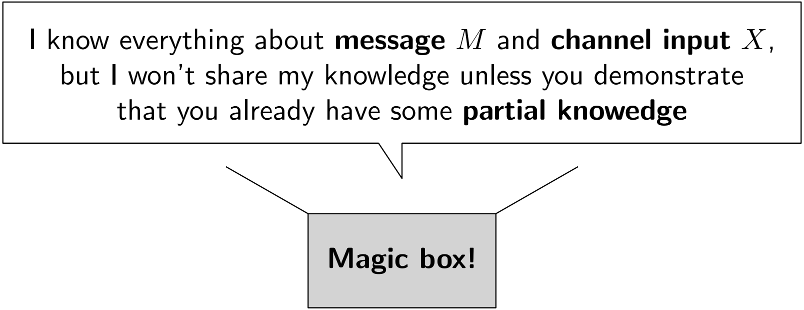
One-shot Channel Coding

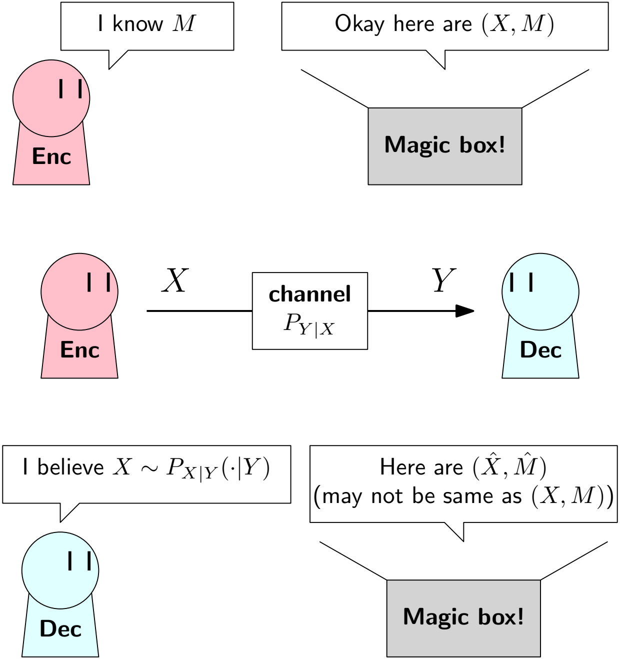


- $\mathbf{P}(M \neq \hat{M}) \leq \mathbf{E}[\min\{\mathsf{L}2^{-\iota_{X;Y}(X;Y)}, 1\}]$
- The box is random
  - Exists fixed box such that the error bound is satisfied
- Recovers (slightly weaker) dependence testing bound [Polyanskiy et al., 2010] (also see [Hayashi, 2009])
- Recovers asymptotic result: $\mathsf{L} \leftarrow 2^{nR}$, $X \leftarrow X^n = (X_1, \ldots, X_n)$, $Y \leftarrow Y^n$,

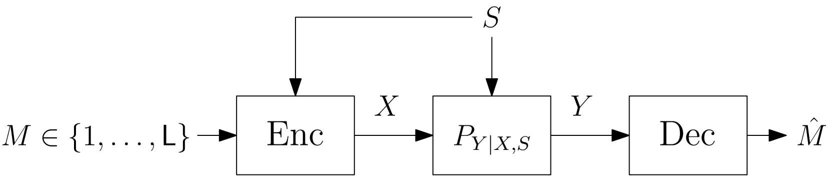
$$\iota_{X^n;Y^n}(X^n; Y^n) = \sum_{i=1}^{n} \iota_{X;Y}(X_i; Y_i) \approx nI(X; Y) \text{ (law of large numbers)}$$

- Unlike stochastic likelihood decoder [Yassaee et al., 2013a], here encoder and decoder are deterministic once the box is fixed


I know everything about **message** M and **channel input** X,
but I won't share my knowledge unless you demonstrate
that you already have some **partial knowedge**
Magic box!


I know M
Okay here are (X, M)
Enc
Magic box!
X
channel
$P_{Y|X}$
Y
Enc
Dec
I believe $X \sim P_{X|Y}(\cdot|Y)$
Here are $(\hat{X}, \hat{M})$
(may not be same as (X, M))
Dec
Magic box!

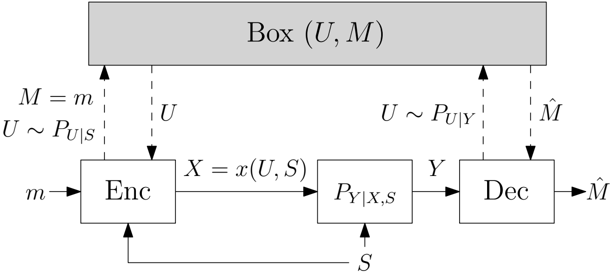
Channels with State Information at the Encoder



- Studied in [Gel'fand and Pinsker, 1980, Heegard and El Gamal, 1983, Costa, 1983]
- Message $M \sim \mathrm{Unif}[1 : \mathsf{L}]$, state $S \sim P_S$, channel $P_{Y|X,S}$
- Encoder produces X using M, S
- Decoder recovers $\hat{M}$ using Y
- Gel'fand and Pinsker [1980]: Asymptotic capacity

$$C = \sup_{P_{U|S},\, x(u,s)} \left(I(U;Y) - I(U;S) \right)$$

Channels with State Information at the Encoder



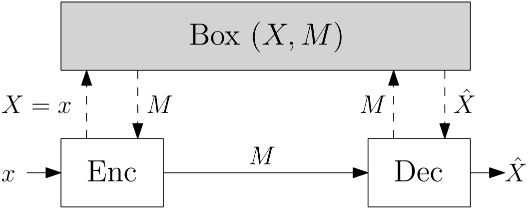
- Box about (U, M), auxiliary RV $U \sim P_{U|S}$
- Encoder queries $P_{U|S} \times \delta_m$ (i.e., the info $U \sim P_{U|S}$, $M = m$), gets U, sends $X = x(U, S)$
- Decoder queries $P_{U|Y} \times P_M$ (i.e., the info $U \sim P_{U|Y}$), gets $\hat{M}$

$$\mathbf{P}(M \neq \hat{M}) = \mathbf{E}\left[\mathbf{P}\left(M \neq \hat{M} \,\middle|\, M, S, U, Y\right)\right]$$

$$\leq \mathbf{E}\left[\min\left\{\frac{dP_{U|S}(\cdot|S) \times \delta_M}{dP_{U|Y}(\cdot|Y) \times P_M}, 1\right\}\right] = \mathbf{E}\left[\min\left\{\mathrm{L}2^{\iota_{U;S}(U;S) - \iota_{U;Y}(U;Y)}, 1\right\}\right]$$

- Implies best known second order result [Scarlett, 2015] (but much shorter proof)

Lossless Source Coding

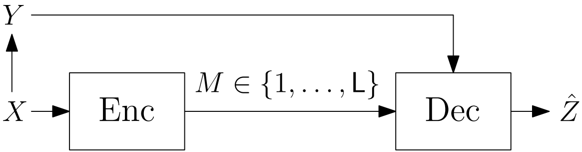


- Encode $X \sim P_X$ to $M \in \{1, \ldots, \mathsf{L}\}$. Decoder gets $\hat{X}$ [Shannon, 1948]
- Box about (X, M), let $P_M = \mathrm{Unif}\{1, \ldots, \mathsf{L}\}$
- Encoder queries using $\delta_x \times P_M$ (i.e., the info $X = x$), gets M
- Decoder queries using $P_X \times \delta_M$ (i.e., the info M), gets $\hat{X}$

$$\mathbf{P}(X \neq \hat{X}) \leq \mathbf{E}\left[\min\left\{\frac{d\delta_X \times P_M}{dP_X \times \delta_M}, 1\right\}\right]$$
$$= \mathbf{E}[\min\{\mathsf{L}^{-1} 2^{\iota_X(X)}, 1\}]$$

- Asymptotic: $X = X^n$, $\mathsf{L} = 2^{nR}$, $P_e \to 0$ if $R > H(X)$

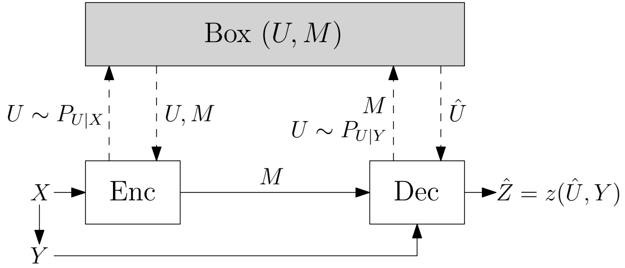
Lossy Source Coding with Side Information at the Decoder



- Studied in [Wyner and Ziv, 1976]
- Source $X \sim P_X$, message $M \in [1 : \mathsf{L}]$, side information $Y|X \sim P_{Y|X}$
- Encoder produces M using X
- Decoder recovers $\hat{Z}$ using M, Y
- Prob. of excess distortion $\mathbf{P}\{\mathsf{d}(X, \hat{Z}) > \mathsf{D}\}$
- Wyner and Ziv [1976] (asymptotic):

$$C = \inf_{P_{U|X},\, z(u,y) : \mathbf{E}[\mathsf{d}(X,Z)] \leq \mathsf{D}} (I(U;X) - I(U;Y))$$

Lossy Source Coding with Side Information at the Decoder



- Box about (U, M), auxiliary RV $U \sim P_{U|X}$
- Encoder queries $P_{U|X} \times P_M$ (i.e., the info $U \sim P_{U|X}$), gets U, M
- Decoder queries $P_{U|Y} \times \delta_M$ (i.e., the info $U \sim P_{U|Y}$ and M), gets $\hat{U}$, outputs $\hat{Z} = z(\hat{U}, Y)$

$$\begin{aligned}
\mathbf{P}(\mathsf{d}(X, \hat{Z}) > \mathsf{D}) &\leq 1 - \mathbf{P}(\mathsf{d}(X, Z) \leq \mathsf{D} \text{ and } U = \hat{U}) \quad (\text{let } Z = z(U, Y)) \\
&= \mathbf{E}\left[1 - \mathbf{1}\{\mathsf{d}(X, Z) \leq \mathsf{D}\}\mathbf{P}(U = \hat{U} \mid M, X, Y, U)\right] \\
&\leq \mathbf{E}\left[1 - \mathbf{1}\{\mathsf{d}(X, Z) \leq \mathsf{D}\} \max\left\{1 - \frac{dP_{U|X} \times P_M}{dP_{U|Y} \times \delta_M}, 0\right\}\right] \\
&= \mathbf{E}\left[1 - \mathbf{1}\{\mathsf{d}(X, Z) \leq \mathsf{D}\} \max\{1 - \mathsf{L}^{-1} 2^{\iota_{U;X}(U;X) - \iota_{U;Y}(U;Y)}, 0\}\right].
\end{aligned}$$

Asymptotic vs Conventional One-shot vs Our Approach

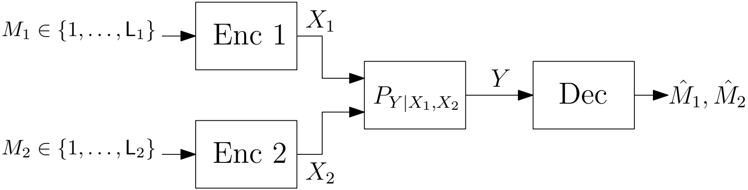
	Strong typicality	Conventional one-shot packing/covering	Poisson matching lemma
Finite blocklength results?	Maybe	Yes	Yes (sometimes stronger)
Simplicity	Simple	Complex	Simple (sometimes simpler)
General alphabet for X, Y?	Only discrete	Discrete / continuous	Discrete / continuous

Conclusion

- A new approach to proving coding theorems using Poisson functional representation and Poisson matching lemma
- Sharp one-shot / second order bounds
- Very short proofs
- Can also be applied to multiple access channels, broadcast channels, etc.
    - Refer to the paper for details

- Multiuser settings...

Multiple Access Channel

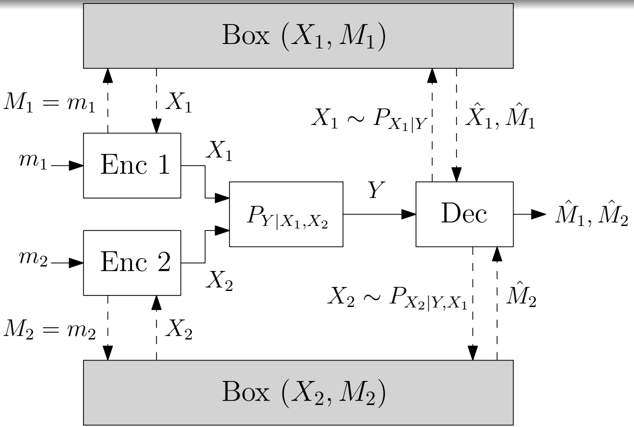


- Messages $M_1 \sim \mathrm{Unif}[1 : \mathsf{L}_1]$, $M_2 \sim \mathrm{Unif}[1 : \mathsf{L}_2]$, channel $P_{Y|X_1,X_2}$
- Encoder i produces X_i using M_i, $i = 1, 2$
- Decoder recovers $\hat{M}_1, \hat{M}_2$ using Y
- Capacity region [Ahlswede, 1971, Liao, 1972, Ahlswede, 1974] (asymptotic): Convex closure of

$$
\begin{aligned}
R_1 &< I(X_1; Y|X_2) \\
R_2 &< I(X_2; Y|X_1) \\
R_1 + R_2 &< I(X_1, X_2; Y)
\end{aligned}
$$

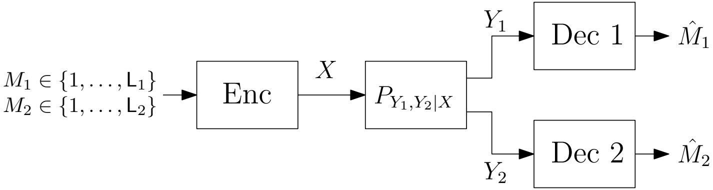
for any P_{X_1}, P_{X_2}

Multiple Access Channel



- Two boxes: (X_1, M_1) and (X_2, M_2)
- Decoder queries $P_{X_1|Y} \times P_{M_1}$ to box (X_1, M_1), gets $\hat{X}_1, \hat{M}_1$, queries $P_{X_2|Y,X_1}(\cdot|Y, \hat{X}_1) \times P_{M_2}$ to box (X_2, M_2), gets $\hat{M}_2$
- $P_e \leq \mathbf{E}[\min\{\mathsf{L}_1 2^{\iota_{X_1;Y}(X_1;Y)}, 1\}] + \mathbf{E}[\min\{\mathsf{L}_2 2^{\iota_{X_2;Y|X_1}(X_2;Y|X_1)}, 1\}]$
- Recovers corner point $R_1 = I(X_1; Y)$, $R_2 = I(X_2; Y|X_1)$ of capacity region

Broadcast Channel

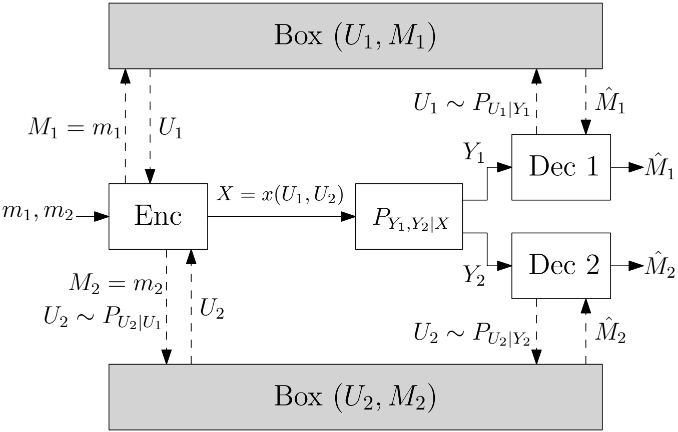


- Messages $M_1 \sim \mathrm{Unif}[1 : \mathsf{L}_1]$, $M_2 \sim \mathrm{Unif}[1 : \mathsf{L}_2]$, channel $P_{Y_1,Y_2|X}$
- Encoder produces X using M_1, M_2
- Decoder i recovers $\hat{M}_i$ using Y_i, $i = 1, 2$
- Inner bound [Marton, 1979] (asymptotic):

$$\begin{aligned} R_1 &< I(U_1; Y_1) \\ R_2 &< I(U_2; Y_2) \\ R_1 + R_2 &< I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2) \end{aligned}$$

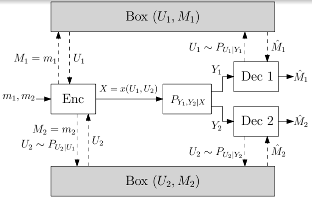
for any P_{U_1,U_2}, $x(u_1, u_2)$

Broadcast Channel



$$P_e \leq \mathbf{E}[\min\{\mathsf{L}_1 2^{\iota_{U_1;Y_1}(U_1;Y_1)}, 1\}] + \mathbf{E}[\min\{\mathsf{L}_2 2^{\iota_{U_2;Y_2}(U_2;Y_2) - \iota_{U_1;U_2}(U_1;U_2)}, 1\}]$$

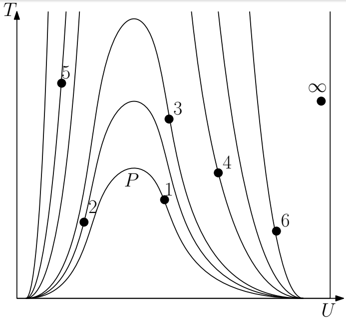
Broadcast Channel



$$P_e \leq \mathbf{E}[\min\{\mathsf{L}_1 2^{\iota_{U_1;Y_1}(U_1;Y_1)}, 1\}] \\ +\mathbf{E}[\min\{\mathsf{L}_2 2^{\iota_{U_2;Y_2}(U_2;Y_2)-\iota_{U_1;U_2}(U_1;U_2)}, 1\}]$$

- Recovers corner point $R_1 = I(U_1;Y_1)$, $R_2 = I(U_2;Y_2) - I(U_1;U_2)$ in Marton's inner bound
- To obtain the whole Marton's inner bound:
  - Time sharing – poor finite-blocklength result
  - Have Box (U_1, M_1) generate a list of U_1's instead, i.e., Box $(U_{1,1}, \ldots, U_{1,l}, M_1)$
    - Query $U_2 \sim l^{-1}\sum_i P_{U_2|U_1}(\cdot|U_{1,i})$ for Box (U_2, M_2)
  - Generalize the box to give a list of probable points
  - Modify the box to give partial information?

Generalized Poisson Matching Lemma



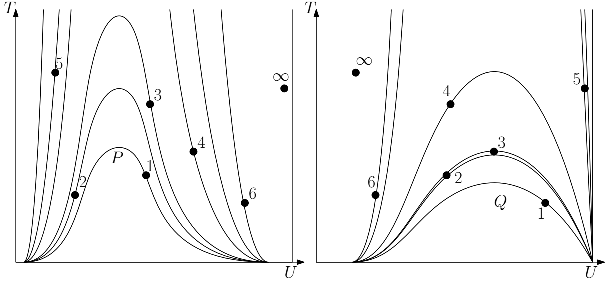
- $\{\bar{U}_i, T_i\}_i$ points of a Poisson process with intensity measure $\mu \times \lambda_{\mathbb{R}_{\geq 0}}$
- For distribution P, reorder indices $i_{P,j}$ such that

$$T_{i_{P,j}} \left(\frac{dP}{d\mu}(\bar{U}_{i_{P,j}}) \right)^{-1}$$

are sorted in ascending order, and let $\tilde{U}_P(j) := \bar{U}_{i_{P,j}}$

- $\tilde{U}_P(1), \tilde{U}_P(2), \tilde{U}_P(3) \overset{iid}{\sim} P$

Generalized Poisson Matching Lemma



$\Upsilon_{P\|Q}(1) = 2$
$\Upsilon_{P\|Q}(2) = 6$
$\Upsilon_{P\|Q}(3) = 4$
$\Upsilon_{P\|Q}(4) = 3$
$\Upsilon_{P\|Q}(5) = \infty$
$\Upsilon_{P\|Q}(6) = 1$

- For distributions P, Q, define

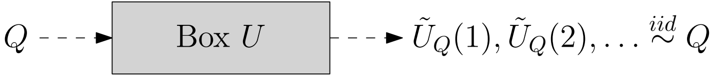
$$\Upsilon_{P\|Q}(j) := \min\{k \in \mathbb{N} : \; i_{Q,k} = i_{P,j}\}$$

  - $\tilde{U}_P = \tilde{U}_Q \Leftrightarrow \Upsilon_{P\|Q}(1) = 1$

- Generalized Poisson matching lemma [Li and Anantharam, 2021]:

$$\mathbf{E}\left[\Upsilon_{P\|Q}(j) \,\middle|\, \tilde{U}_P(j)\right] \le j \frac{dP}{dQ}(\tilde{U}_P(j)) + 1$$

The Generalized Box



- **Query:** Input distribution Q, output $\tilde{U}_Q(1), \tilde{U}_Q(2), \ldots \overset{iid}{\sim} Q$
- **Guarantee:** For distributions P, Q, $j \in \mathbb{N}$,

$$\mathbf{E}\left[\min\{k : \tilde{U}_Q(k) = \tilde{U}_P(j)\} \,\middle|\, \tilde{U}_P(j)\right] \leq j \frac{dP}{dQ}(\tilde{U}_P(j)) + 1$$

  - Implies the guarantee of single-output box for $j = 1$
- Useful for multiple access channel, broadcast channel, distributed lossy source coding, channel resolvability and channel simulation

Conclusion

- A new approach to proving coding theorems
  - Using Poisson functional representation and Poisson matching lemma
  - Sharp one-shot / second order bounds
  - Very short proofs
- Future work
  - Is there a simpler way to apply this method to multiuser settings?

Rudolf Ahlswede. Multi-way communication channels. In *2nd Int. Symp. Inform. Theory, Tsahkadsor, Armenian SSR*, pages 23–52, 1971.

Rudolf Ahlswede. The capacity region of a channel with two senders and two receivers. *The annals of probability*, 2(5):805–814, 1974.

Max H. M. Costa. Writing on dirty paper. *IEEE Trans. Inf. Theory*, 29(3):439–441, 1983.

Amiel Feinstein. A new basic theorem of information theory. *IRE Trans. Inf. Theory*, (4):2–22, 1954.

S. I. Gel'fand and M. S. Pinsker. Coding for channel with random parameters. *Probl. Contr. and Inf. Theory*, 9(1):19–31, 1980.

Masahito Hayashi. Information spectrum approach to second-order coding rate in channel coding. *IEEE Transactions on Information Theory*, 55(11):4947–4966, 2009.

Chris Heegard and A El Gamal. On the capacity of computer memory with defects. *IEEE transactions on Information theory*, 29(5):731–739, 1983.

Iris AM Huijben, Wouter Kool, Max B Paulus, and Ruud JG Van Sloun. A review of the gumbel-max trick and its extensions for discrete stochasticity in machine learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):1353–1371, 2022.

C. T. Li and A. El Gamal. Strong functional representation lemma and applications to coding theorems. *IEEE Trans. Inf. Theory*, 64(11):6967–6978, Nov 2018. ISSN 0018-9448. doi: 10.1109/TIT.2018.2865570.

Cheuk Ting Li and Venkat Anantharam. A unified framework for one-shot achievability via the Poisson matching lemma. *IEEE Transactions on Information Theory*, 67(5):2624–2651, 2021.

References II

H Liao. *Multiple Access Channels*. PhD thesis, University of Hawaii, Honolulu, HI, 1972.

K. Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inf. Theory*, 25(3):306–311, May 1979. ISSN 0018-9448. doi: 10.1109/TIT.1979.1056046.

Yury Polyanskiy, H Vincent Poor, and Sergio Verdú. Channel coding rate in the finite blocklength regime. *IEEE Trans. Inf. Theory*, 56(5):2307–2359, 2010.

Jonathan Scarlett. On the dispersions of the Gel'fand–Pinsker channel and dirty paper coding. *IEEE Trans. Inf. Theory*, 61(9):4569–4586, 2015.

Claude E Shannon. A mathematical theory of communication. *Bell system technical journal*, 27 (3):379–423, 1948.

Claude E Shannon. Certain results in coding theory for noisy channels. *Information and control*, 1 (1):6–25, 1957.

S. Verdú. Non-asymptotic achievability bounds in multiuser information theory. In *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, pages 1–8, Oct 2012. doi: 10.1109/Allerton.2012.6483192.

S. Watanabe, S. Kuzuoka, and V. Y. F. Tan. Nonasymptotic and second-order achievability bounds for coding with side-information. *IEEE Trans. Inf. Theory*, 61(4):1574–1605, April 2015. ISSN 0018-9448. doi: 10.1109/TIT.2015.2400994.

A. D. Wyner and J. Ziv. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. Inf. Theory*, 22(1):1–10, January 1976. ISSN 0018-9448. doi: 10.1109/TIT.1976.1055508.

References III

M. H. Yassaee, M. R. Aref, and A. Gohari. A technique for deriving one-shot achievability results in network information theory. In *2013 IEEE ISIT*, pages 1287–1291, July 2013a. doi: 10.1109/ISIT.2013.6620434.

Mohammad Hossein Yassaee, Mohammad Reza Aref, and Amin Gohari. A technique for deriving one-shot achievability results in network information theory. In *2013 IEEE International Symposium on Information Theory*, pages 1287–1291. IEEE, 2013b.