Finite-blocklength schemes in information theory

Lecture 2

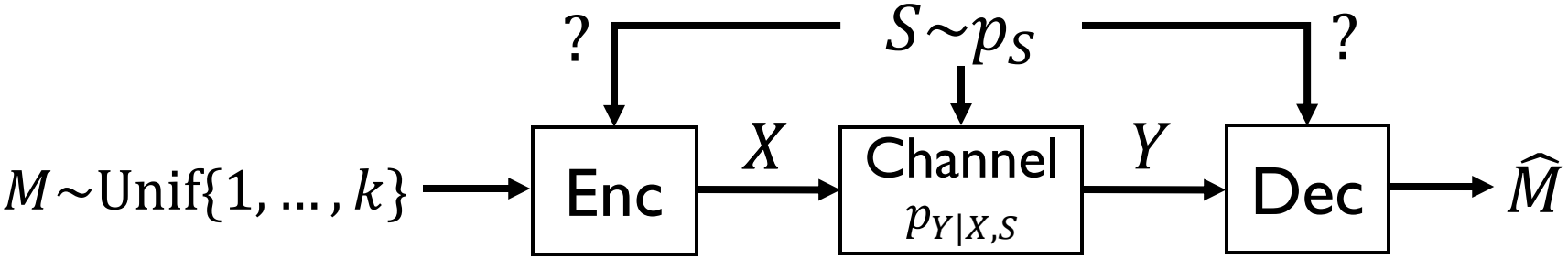
Li Cheuk Ting

Department of Information Engineering,
The Chinese University of Hong Kong

ctli@ie.cuhk.edu.hk

Part of this presentation is based on my lecture notes for Special Topics in Information Theory

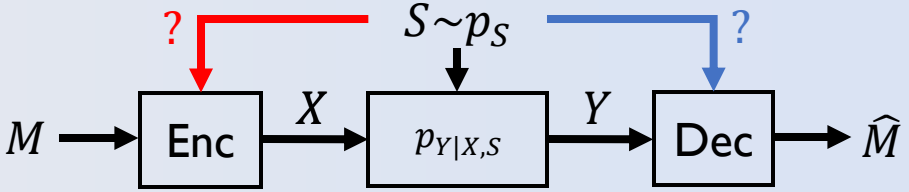
Channel with state



- Channel depends on the state $S \sim p_S$
- State may or may not be available at encoder/decoder

A. El Gamal and Y.-H. Kim, Network information theory. Cambridge University Press, 2011.

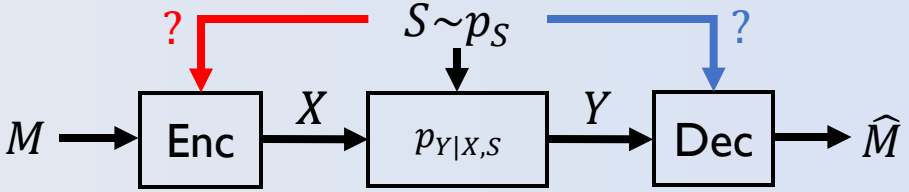
Channel w/ state



- 4 combinations: channel state available at encoder (CSIT) or not, available at decoder (CSIR) or not
- Wireless communications
  - Channel coefficient (state) is random
  - CSIR: decoder learns the state via pilot signals
  - CSIT: decoder sends the state back to encoder
- Memory with fault
  - Some cells are stuck at 0 or 1
  - CSIT: encoder examines the memory and learns the faults
- Information embedding and digital watermark
  - Hide information in a carrier signal (audio/image/video)
  - CSIT: encoder knows the state (carrier signal)

A. El Gamal and Y.-H. Kim, Network information theory. Cambridge University Press, 2011.

Channel w/ state

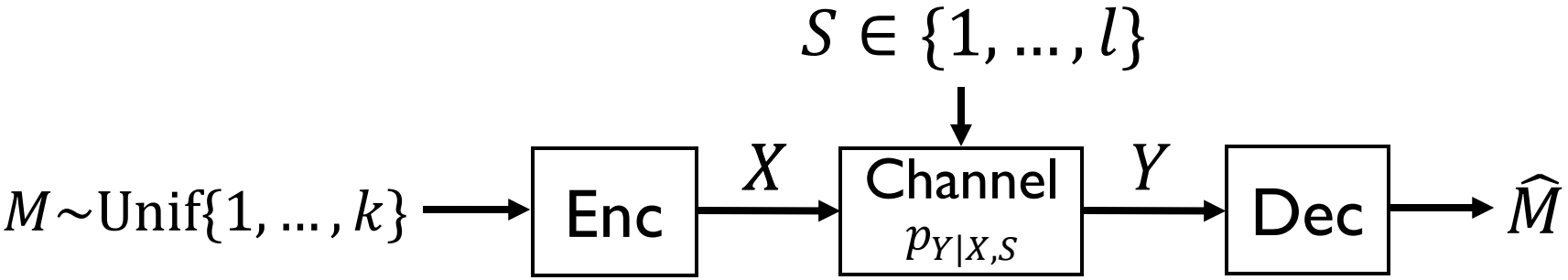


Cases:

- No CSIT/CSIR:
  - Random S: reduces to ordinary channel coding
  - Worst-case S: compound channel (next slide)
- CSIR only: treat it as a channel $X \to (Y, S)$
- CSIT and CSIR: can use a separate code for each value of S
- CSIT only: Gelfand-Pinsker (next next slide)

A. El Gamal and Y.-H. Kim, Network information theory. Cambridge University Press, 2011.

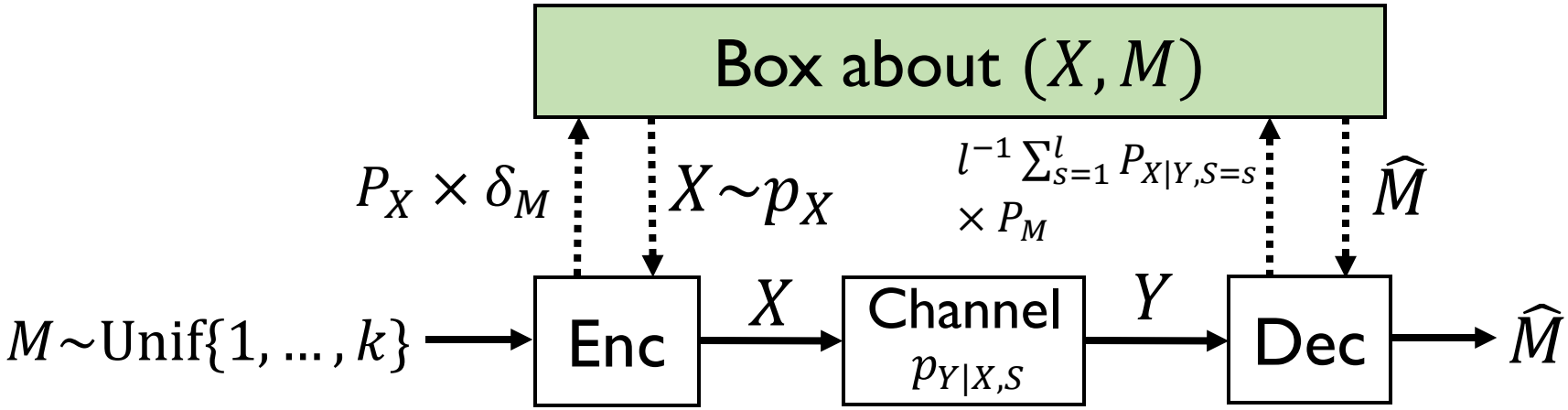
Compound channel



- State S can be any one of l choices
- The channel $p_{Y|X,S=s}$ can be any one of l choices, but the encoder and decoder do not know which one
- Consider worst-case prob. of error

$$\max_s \mathbf{P}\left(M \neq \widehat{M} \mid S = s\right)$$

A. El Gamal and Y.-H. Kim, Network information theory. Cambridge University Press, 2011.

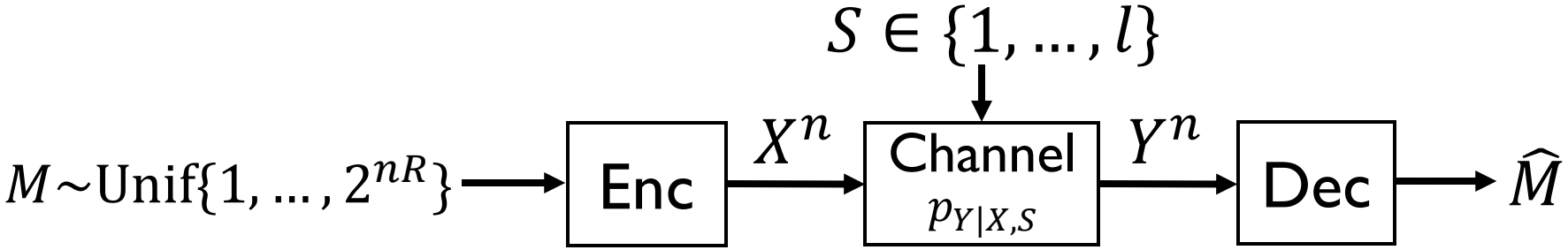


- Encoding: Query $P_X \times \delta_M$, get (X, M)
- Decoding pretends that $S \sim \mathrm{Unif}\{1, \ldots, l\}$, query $l^{-1} \sum_{s=1}^{l} P_{X|Y,S=s} \times P_M$, get $(\hat{X}, \widehat{M})$
- Poisson matching lemma:

$$\begin{aligned}
&\mathbf{P}\left(M \neq \widehat{M} | S = s\right) \\
&\leq \mathbf{E}\left[\min\left\{\frac{(P_X \times \delta_M)(X,M)}{\left(l^{-1} \sum_{s=1}^{l} P_{X|Y,S=s} \times P_M\right)(X,M)}, 1\right\} \mid S = s\right] \\
&\leq \mathbf{E}\left[\min\left\{\frac{P_X(X)}{P_{X|Y,S}(X|Y,s)/(kl)}, 1\right\} \mid S = s\right] \\
&= \mathbf{E}\left[\min\left\{kl 2^{-\imath_{X;Y|S}(X;Y|s)}, 1\right\} \mid S = s\right]
\end{aligned}$$

C. T. Li and V. Anantharam, "A Unified Framework for One-Shot Achievability via the Poisson Matching Lemma," 2021.

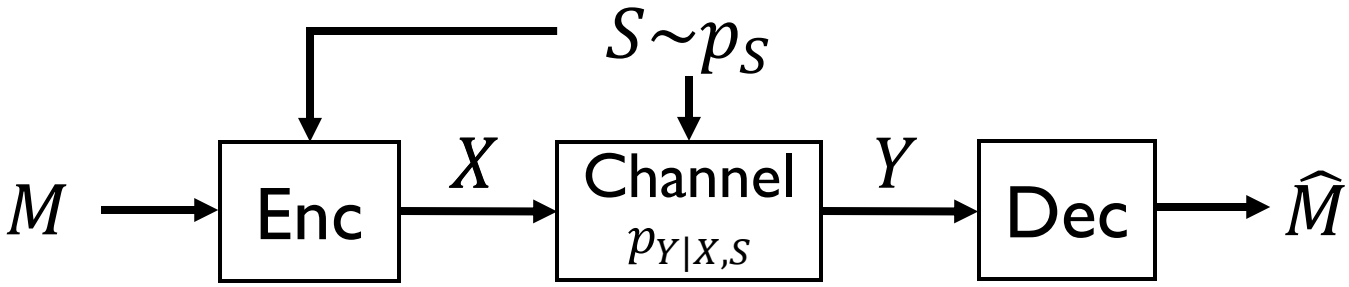
Compound channel, asymptotic



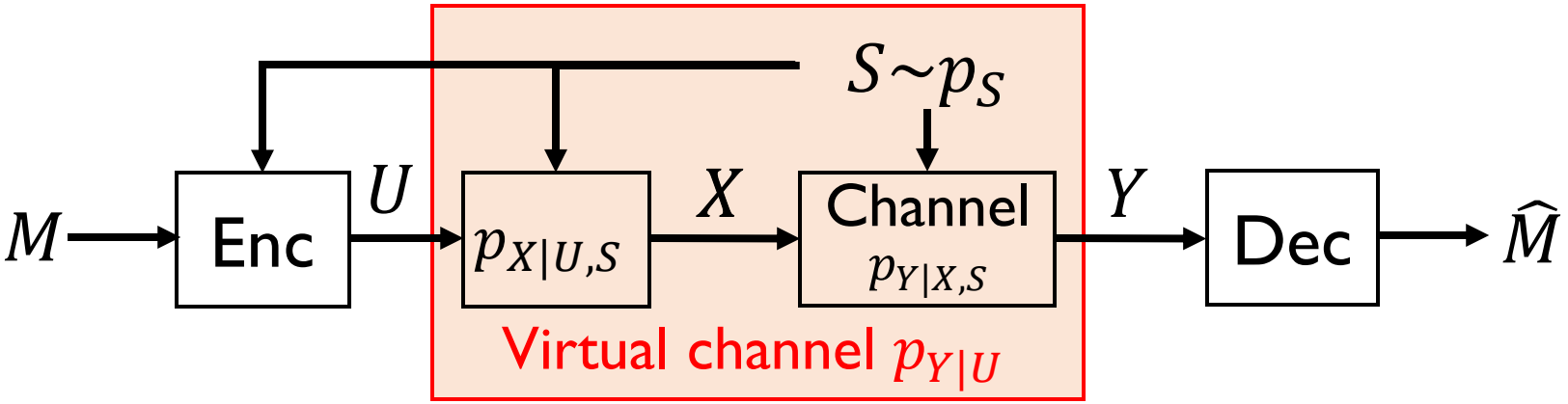
- State S is fixed throughout n channel uses
- $\mathbf{P}\left(M \neq \widehat{M} | S = s\right) \leq \mathbf{E}\left[\min\left\{l 2^{nR} 2^{-\sum_i \imath(X_i; Y_i|s)}, 1\right\} | S = s\right]$
  $\to 0$ if $R < I(X; Y|S = s)$
- Worst case $P_e \to 0$ if $R < \min_s I(X; Y|S = s)$
- Recovers (achievability part of) the capacity of compound channel
$$C = \max_{p_X} \min_s I(X; Y|S = s)$$
- C is generally smaller than $\min_s \max_{p_X} I(X; Y|S = s)$
- Arbitrarily varying channel: state can change each time

D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels," Ann. Math. Statist., vol. 30, no. 4, pp. 1229–1241, 1959

Channel with state at encoder

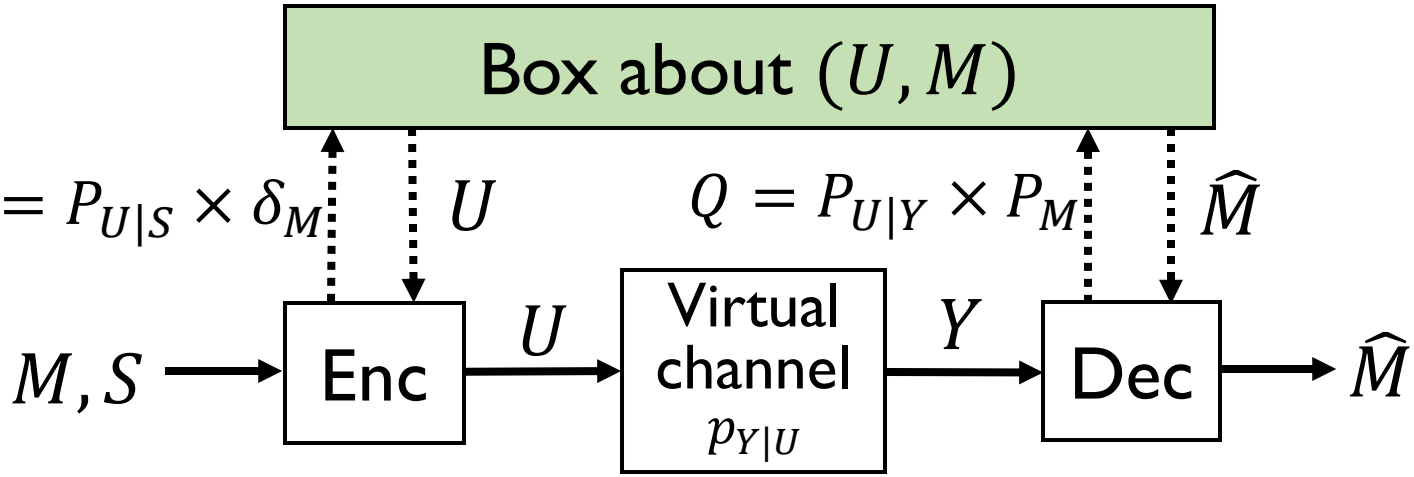


- Create a "virtual" channel input U



- We now design a code for the virtual channel $U \to Y$

A. El Gamal and Y.-H. Kim, Network information theory. Cambridge University Press, 2011.

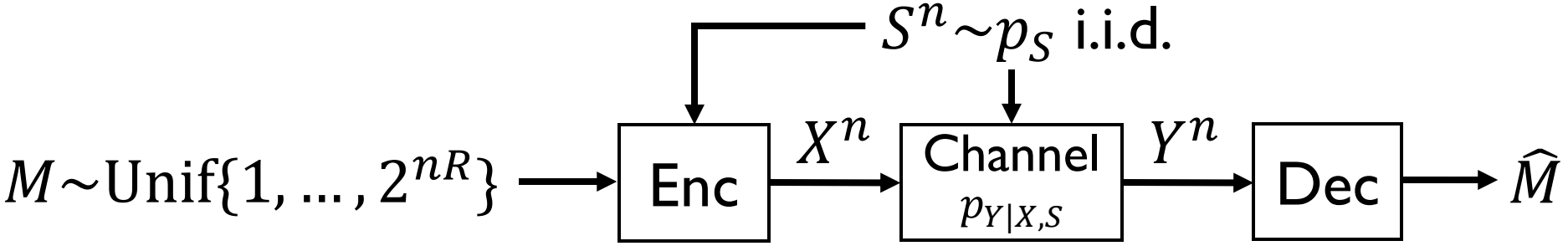


- Fix any $P_{U|S}$ and $P_{X|U,S}$
- Encoding: Query $Q = P_{U|S} \times \delta_M$, get (U, M)
- Decoding: Query $Q = P_{U|Y} \times P_M$, get $(\widehat{U}, \widehat{M})$
- Poisson matching lemma:

$$\mathbf{P}\left(M \neq \widehat{M}\right) \leq \mathbf{E}\left[\mathbf{P}\left(M \neq \widehat{M} \middle| M, X, Y\right)\right]$$
$$\leq \mathbf{E}\left[\min\left\{\frac{(P_{U|S} \times \delta_M)(U,M)}{(P_{U|Y} \times P_M)(U,M)}, 1\right\}\right]$$
$$= \mathbf{E}\left[\min\left\{\frac{P_{U|S}(U|S)}{P_{U|Y}(U|Y)/k}, 1\right\}\right]$$
$$= \mathbf{E}\left[\min\left\{k 2^{\imath_{U;S}(U;S) - \imath_{U;Y}(U;Y)}, 1\right\}\right]$$

C. T. Li and V. Anantharam, "A Unified Framework for One-Shot Achievability via the Poisson Matching Lemma," 2021.

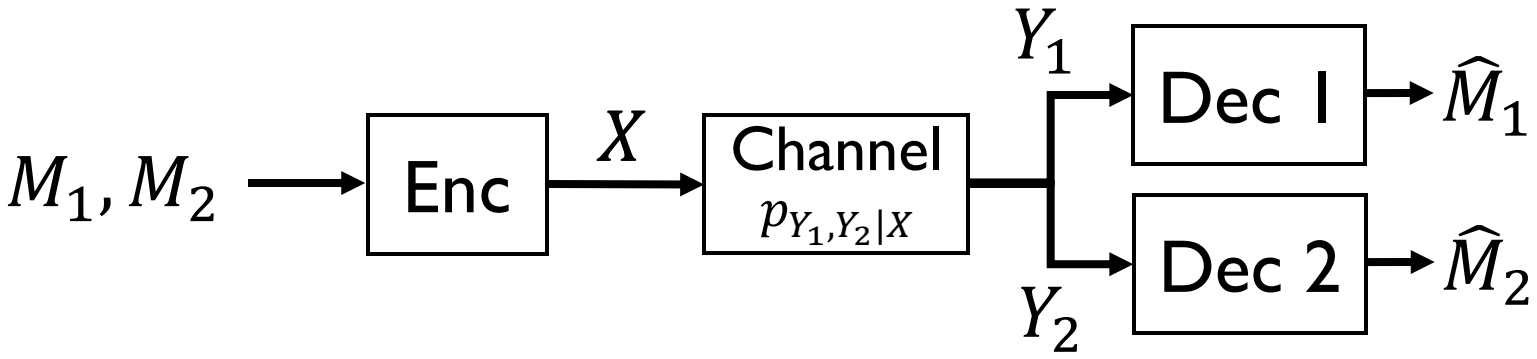
Channel with state, asymptotic noncausal



- Fix any $P_{U|S}$ and $P_{X|U,S}$
- $P_{\mathrm{e}} \leq \mathbf{E}\left[\min\left\{2^{nR}2^{\imath_{U^n;S^n}(U^n;S^n)-\imath_{U^n;Y^n}(U^n;Y^n)}, 1\right\}\right]$
- Law of large numbers:
$$\imath(U^n;S^n) - \imath(U^n;Y^n) \approx n(I(U;S) - I(U;Y))$$
- $P_{\mathrm{e}} \to 0$ if $R < I(U;Y) - I(U;S)$
- Recovers (achievability of) Gelfand-Pinsker theorem:
$$C = \sup_{P_{U|S}, P_{X|U,S}} (I(U;Y) - I(U;S))$$

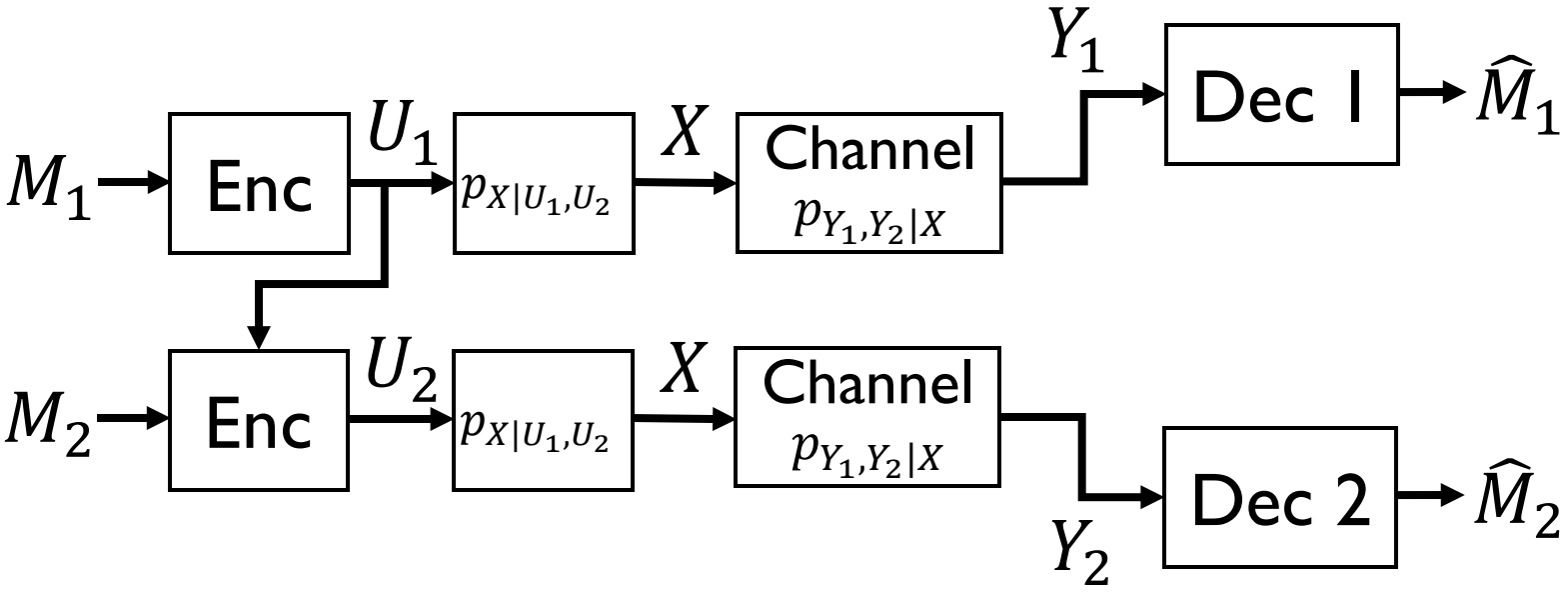
S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," Probl. Contr. and Inf. Theory, vol. 9, no. 1, pp. 19–31, 1980.

Broadcast channel



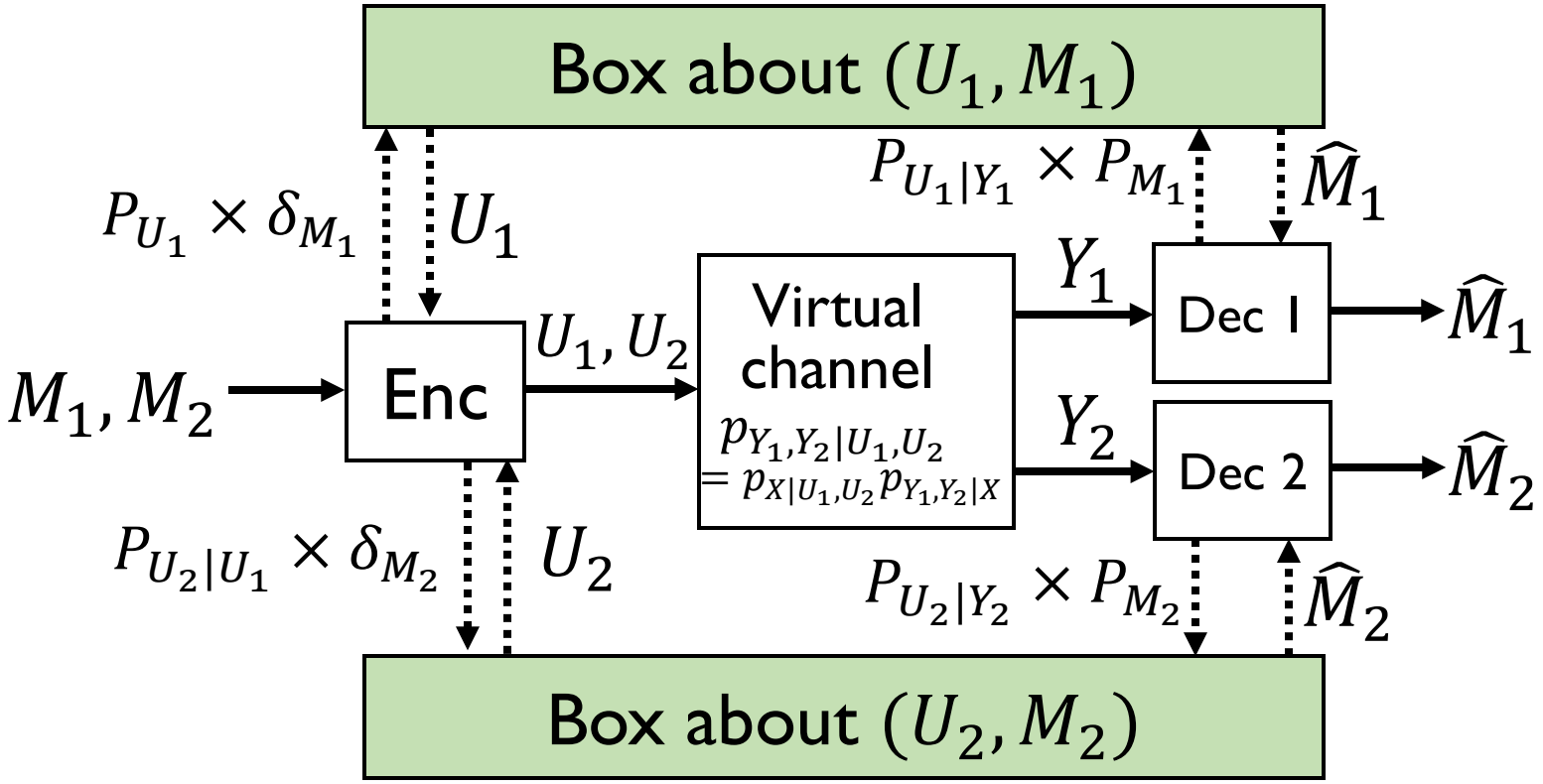
- Two independent messages $M_1 \sim \text{Unif}\{1, \ldots, k_1\}$, $M_2 \sim \text{Unif}\{1, \ldots, k_2\}$
- Encoder sends X through broadcast channel $p_{Y_1,Y_2|X}$
- Decoder i observes Y_i, wants to decode M_i

A. El Gamal and Y.-H. Kim, Network information theory. Cambridge University Press, 2011.



- Virtual inputs U_1, U_2, fix $p_{U_1,U_2}, p_{X|U_1,U_2}$
- Use two codes:
  - Transmit M_1 through the channel $U_1 \to Y_1$, treat U_2 as noise
  - Transmit M_2 through the channel $U_2 \to Y_2$ with state U_1 known at encoder

A. El Gamal and Y.-H. Kim, Network information theory. Cambridge University Press, 2011.

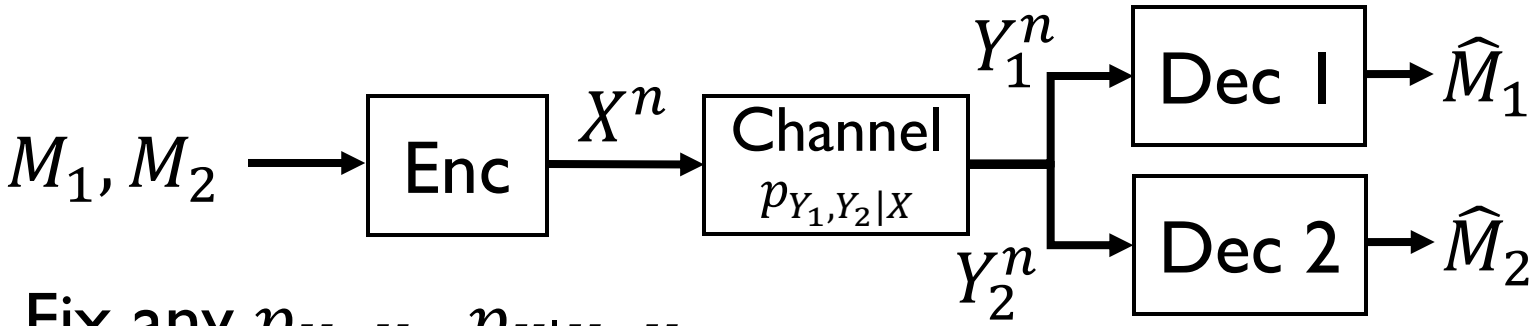


- Poisson matching lemma:

$$\mathbf{P}\big((M_1, M_2) \neq (\widehat{M}_1, \widehat{M}_2)\big)$$
$$\leq \mathbf{E}\left[\min\left\{\frac{(P_{U_1} \times \delta_{M_1})(U_1, M_1)}{(P_{U_1|Y_1} \times P_{M_1})(U_1, M_1)}, 1\right\}\right]$$
$$+\mathbf{E}\left[\min\left\{\frac{(P_{U_2|U_1} \times \delta_{M_2})(U_2, M_2)}{(P_{U_2|Y_2} \times P_{M_2})(U_2, M_2)}, 1\right\}\right]$$
$$= \mathbf{E}\big[\min\{k_1 2^{-\iota(U_1;Y_1)}, 1\}\big] + \mathbf{E}\big[\min\{k_2 2^{\iota(U_1;U_2)-\iota(U_2;Y_2)}, 1\}\big]$$

C. T. Li and V. Anantharam, "A Unified Framework for One-Shot Achievability via the Poisson Matching Lemma," 2021.

Broadcast channel, asymptotic

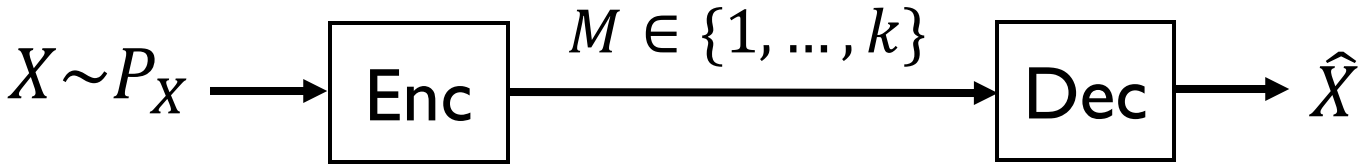


- Fix any $p_{U_1,U_2}, p_{X|U_1,U_2}$
- $P_e \leq \mathbf{E}\left[\min\left\{2^{nR_1-\iota(U_1;Y_1)}, 1\right\}\right] + \mathbf{E}\left[\min\left\{2^{nR_2+\iota(U_1;U_2)-\iota(U_2;Y_2)}, 1\right\}\right]$
- $P_e \to 0$ if $R_1 < I(U_1;Y_1), R_2 < I(U_2;Y_2) - I(U_1;U_2)$
- One corner point of Marton's inner bound

$$R_1 < I(U_1;Y_1)$$
$$R_2 < I(U_2;Y_2)$$
$$R_1 + R_2 < I(U_1;Y_1) + I(U_2;Y_2) - I(U_1;U_2)$$

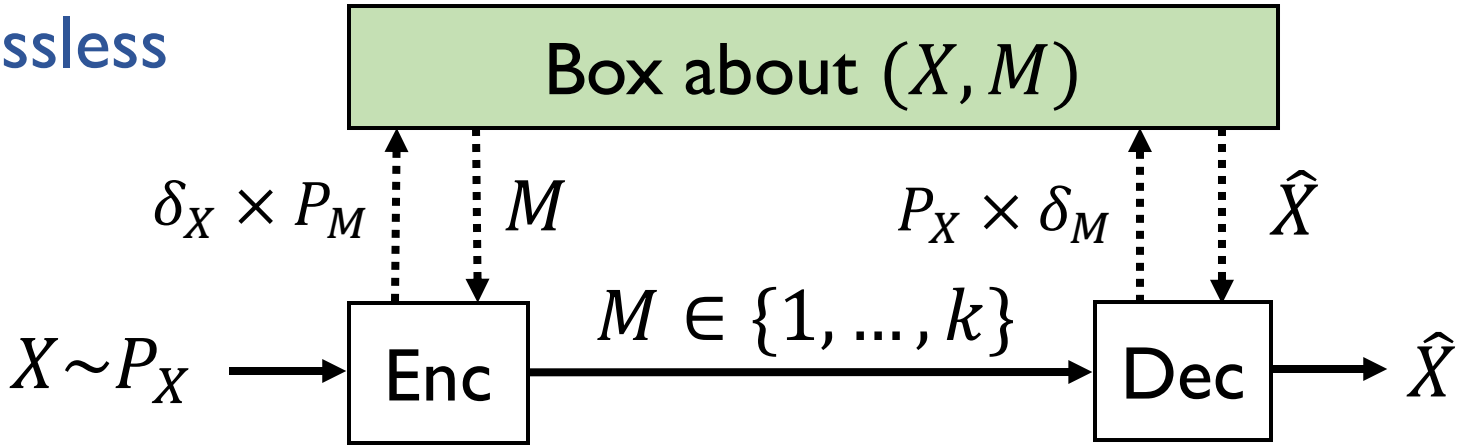
K. Marton, "A coding theorem for the discrete memoryless broadcast channel," IEEE Transactions on Information Theory, vol. 25, no. 3, pp. 306–311, May 1979.

One-shot source coding



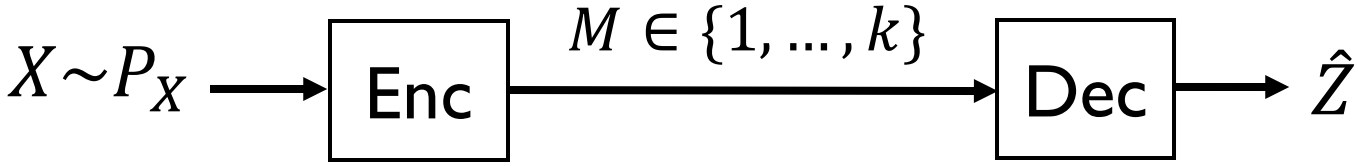
- Encoder compress source $X \sim P_X$ into $M \in \{1, \ldots, k\}$
- Decoder recovers $\hat{X}$
- Requirement:
  1. **Lossless:** $\hat{X} = X$
     - Require k to be at least the number of values of X
  2. **Almost lossless:** $\mathbf{P}(\hat{X} \neq X) < \epsilon$ (next slide)
  3. **Lossy** (next next slide)

Almost lossless



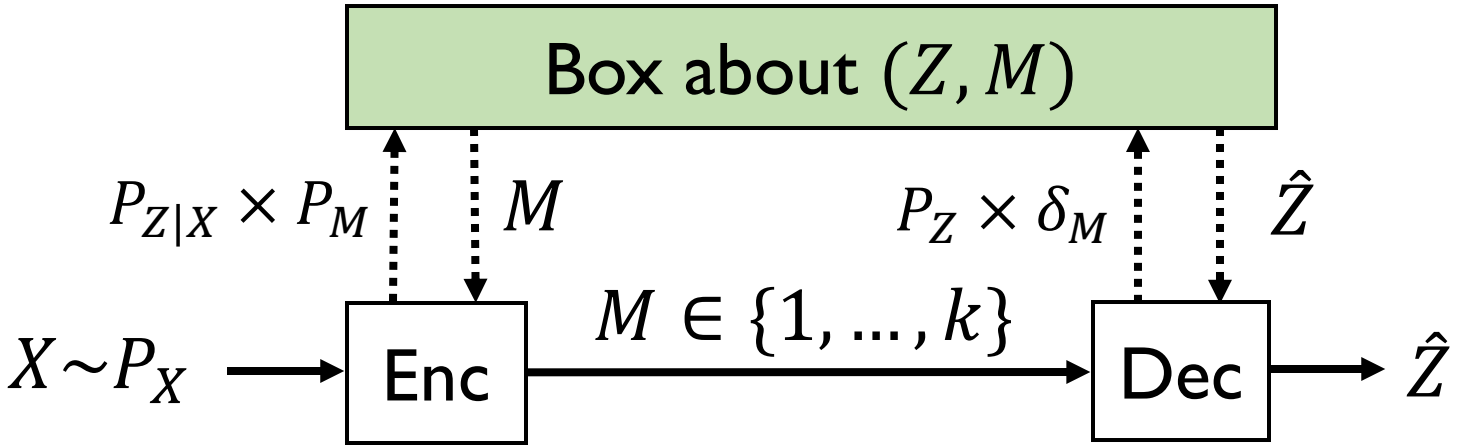
- Let P_M be $\mathrm{Unif}\{1, \ldots, k\}$
- Encoding: Query $\delta_X \times P_M$, get (X, M)
- Decoding: Query $P_X \times \delta_M$, get $(\hat{X}, \hat{M})$
- $$\mathbf{P}\left(X \neq \hat{X}\right) \leq \mathbf{E}\left[\min\left\{\frac{(\delta_X \times P_M)(X,M)}{(P_X \times \delta_M)(X,M)}, 1\right\}\right] = \mathbf{E}\left[\min\left\{\frac{1/k}{P_X(X)}, 1\right\}\right] = \mathbf{E}\left[\min\left\{k^{-1}2^{\iota_X(X)}, 1\right\}\right]$$
- Asymptotic: $k = 2^{nR}, \iota_{X^n}(X^n) \approx nH(X)$
  $P_e \to 0$ if $R > H(X)$

Lossy source coding



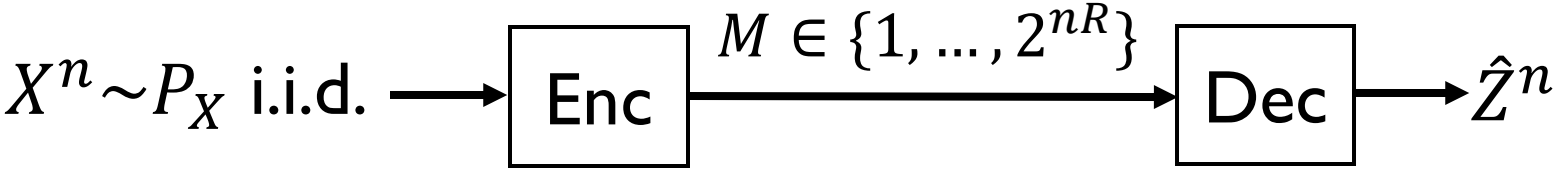
- Encoder compress source $X \sim P_X$ into $M \in \{1, \ldots, k\}$
- Decoder recovers $\hat{Z}$
- Distortion function $d(x, z) \geq 0$ measures how far the reconstruction z is from the source x
- Requires $P_e = \mathbf{P}\left(d\left(X, \hat{Z}\right) > D\right) \leq \epsilon$
- Want $\hat{Z}$ to approximately follows the conditional distribution $P_{Z|X}$ (the test channel)
  - Then we can evaluate $\mathbf{P}\left(d\left(X, \hat{Z}\right) > D\right)$ over $P_X P_{Z|X}$

Lossy



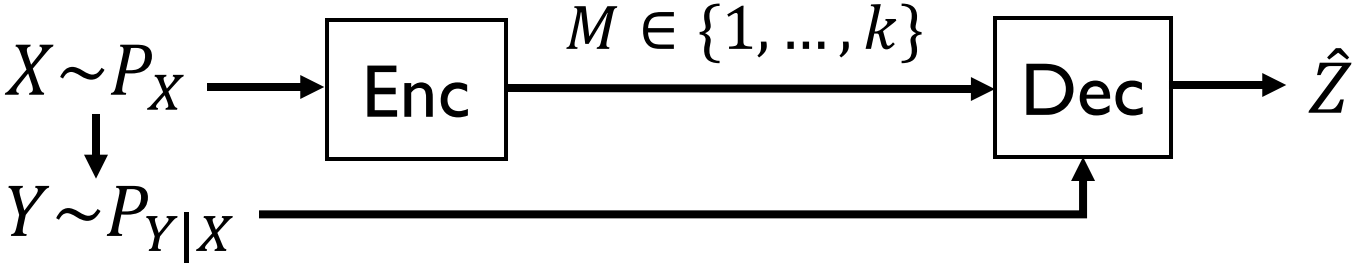
- Fix any test channel $P_{Z|X}$, box about (Z, M)
- Encoding: Query $P_{Z|X} \times P_M$, get (Z, M)
- Decoding: Query $P_Z \times \delta_M$, get $(\hat{Z}, \hat{M})$
- Note that $(X, Z) \sim P_X P_{Z|X}$
- $\mathbf{P}\left(Z \neq \hat{Z}\right)$
  $\leq \mathbf{E}\left[\min\left\{\frac{(P_{Z|X} \times P_M)(Z,M)}{(P_Z \times \delta_M)(Z,M)}, 1\right\}\right]$
  $= \mathbf{E}\left[\min\left\{k^{-1} 2^{\iota_{X;Z}(X;Z)}, 1\right\}\right]$
- $P_e = \mathbf{P}\left(d\left(X, \hat{Z}\right) > D\right)$
  $\leq \mathbf{P}(d(X, Z) > D) + \mathbf{E}\left[\min\left\{k^{-1} 2^{\iota_{X;Z}(X;Z)}, 1\right\}\right]$

Lossy source coding - asymptotic

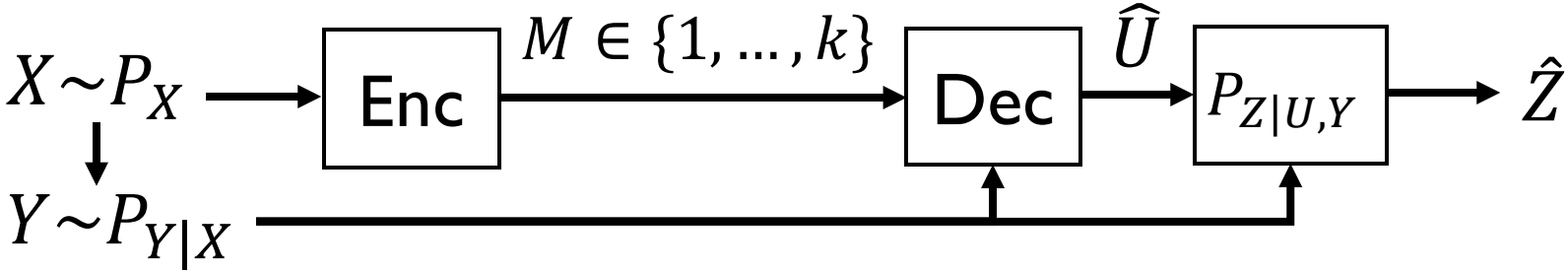


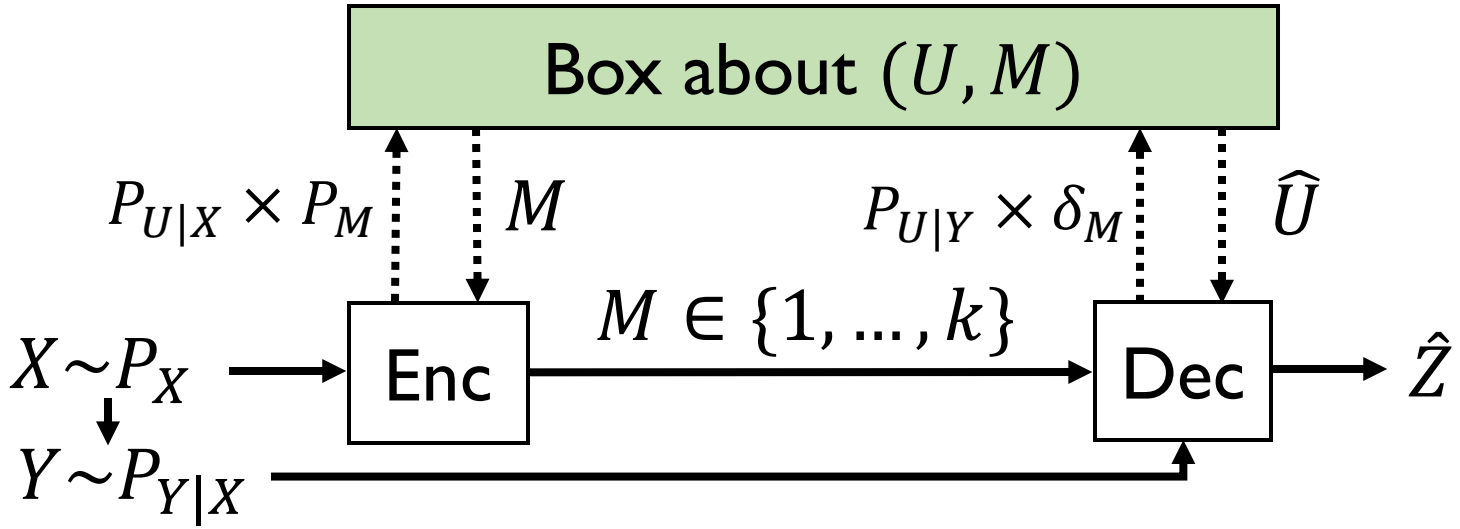
- Distortion function $d(x^n, z^n) = n^{-1} \sum_{i=1}^{n} d(x_i, z_i)$
- $P_e \leq \mathbf{P}(d(X^n, Z^n) > D) + \mathbf{E}\left[\min\left\{2^{-nR+\imath(X^n;Z^n)}, 1\right\}\right]$
- First term $\to 0$ if $\mathbf{E}[d(X, Z)] < D$
- Second term $\to 0$ if $R > I(X; Z)$
- Optimal rate is the minimum of $I(X; Z)$ over $P_{Z|X}$ subject to $\mathbf{E}[d(X, Z)] < D$
  - Rate-distortion function

Lossy source coding w/ side info available at decoder



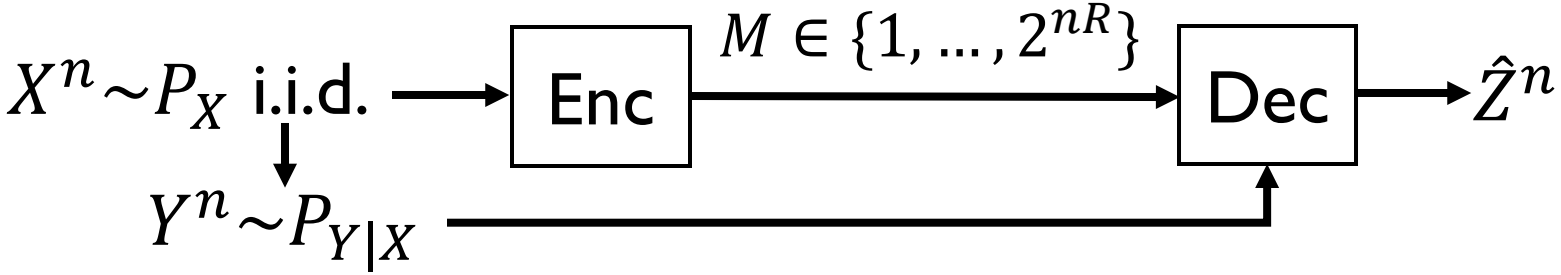
- Encoder compress source $X \sim P_X$ into $M \in \{1, \ldots, k\}$
- Decoder observes M and side info $Y \sim P_{Y|X}$, recovers $\hat{Z}$
- Requires $P_e = \mathbf{P}\left(d\left(X, \hat{Z}\right) > D\right) \leq \epsilon$
- Adds a "virtual output" $U \sim P_{U|X}$
  - Decoder recovers $\hat{U}$ and outputs $\hat{Z} \sim P_{Z|U,Y}$





- Fix any $P_{U|X}, P_{Z|U,Y}$, box about (U, M)
- Encoding: Query $P_{U|X} \times P_M$, get (U, M)
- Decoding: Query $P_{U|Y} \times \delta_M$, get $(\widehat{U}, \widehat{M})$, $\hat{Z} \sim P_{Z|U,Y}(\cdot\,|\widehat{U}, Y)$
- $(X, Y, U, Z) \sim P_X P_{Y|X} P_{U|X} P_{Z|U,Y}$
- $\mathbf{P}\left(U \neq \widehat{U}\right) \leq \mathbf{E}\left[\min\left\{\frac{\left(P_{U|X} \times P_M\right)(U,M)}{\left(P_{U|Y} \times \delta_M\right)(U,M)}, 1\right\}\right]$
$= \mathbf{E}\left[\min\left\{k^{-1} 2^{\iota(X;U) - \iota(Y;U)}, 1\right\}\right]$
- $P_e = \mathbf{P}\left(d\left(X, \hat{Z}\right) > D\right)$
$\leq \mathbf{P}(d(X, Z) > D) + \mathbf{E}\left[\min\left\{k^{-1} 2^{\iota(X;U) - \iota(Y;U)}, 1\right\}\right]$

Lossy source coding w/ side info at decoder - asymptotic



- $P_e \leq \mathbf{P}(d(X^n, Z^n) > D) + \mathbf{E}\left[\min\left\{2^{-nR+\imath(X^n;U^n)-\imath(Y^n;U^n)}, 1\right\}\right]$
- First term $\to 0$ if $\mathbf{E}[d(X, Z)] < D$
- Second term $\to 0$ if $R > I(X; U) - I(Y; U)$
- Recovers (achievability of) Wyner-Ziv theorem: Optimal rate is the minimum of $I(X; U) - I(Y; U)$ over $P_{U|X}, P_{Z|U,Y}$ subject to $\mathbf{E}[d(X, Z)] < D$

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, 1976.