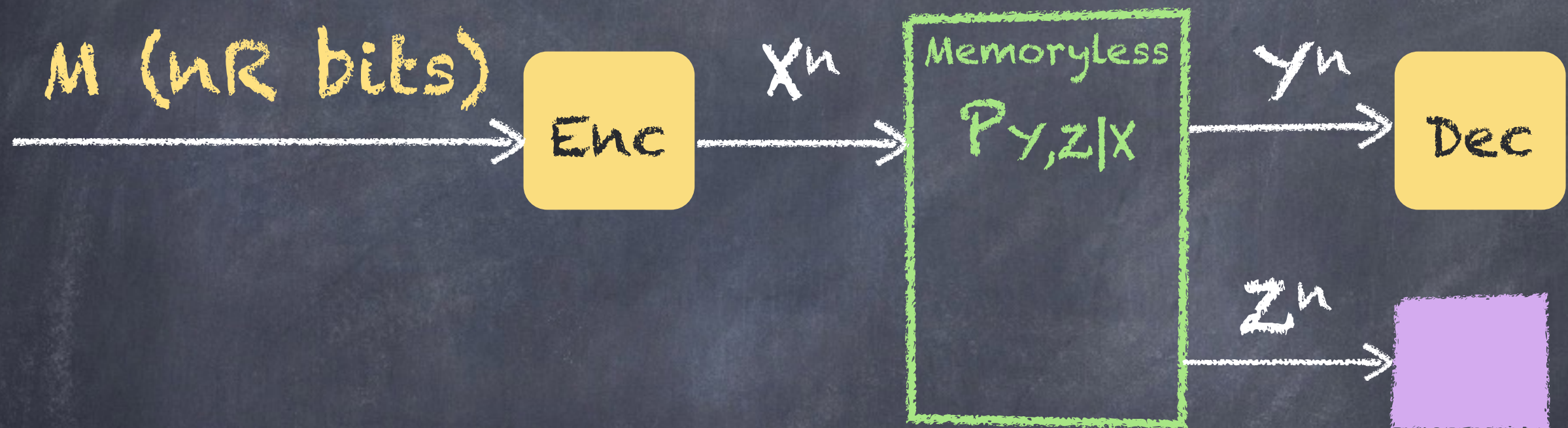


Wiretap Channels with Random States

Paul Cuff (Princeton University),
Ziv Gelfeld, Haim Permuter



Wiretap Channel



Secrecy Capacity:

- Reliable communication
- Z^n contains no information about M

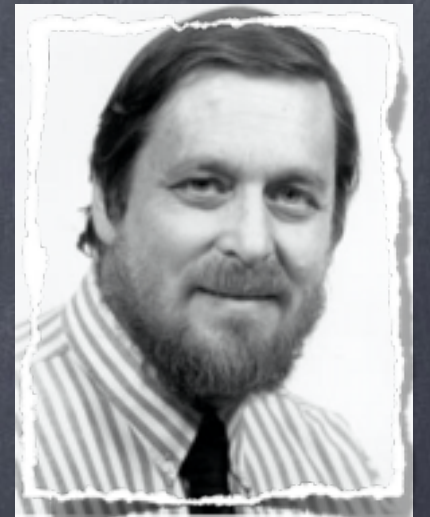
Solution

$$C_s = \max_{P_{X,U}} I(U; Y) - I(U; Z)$$

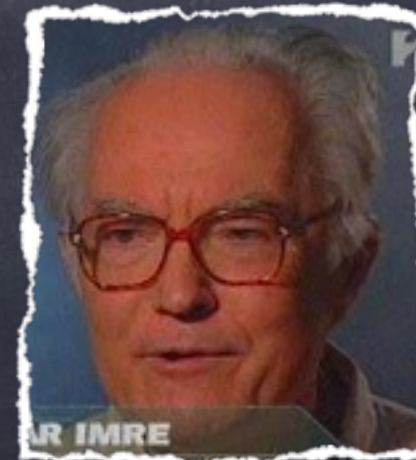


Solutions

- 1975: Wyner introduced the problem and gave solution for degraded channels ($U=X$ is sufficient)

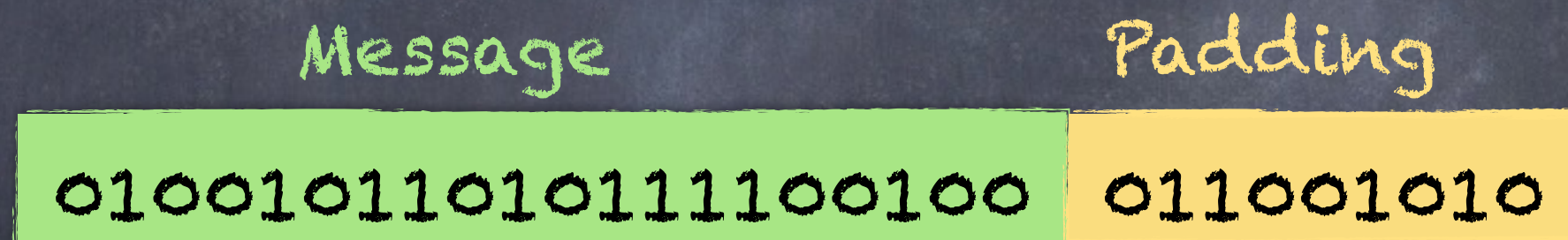


- 1978: Csiszár and Körner gave solution for all channels



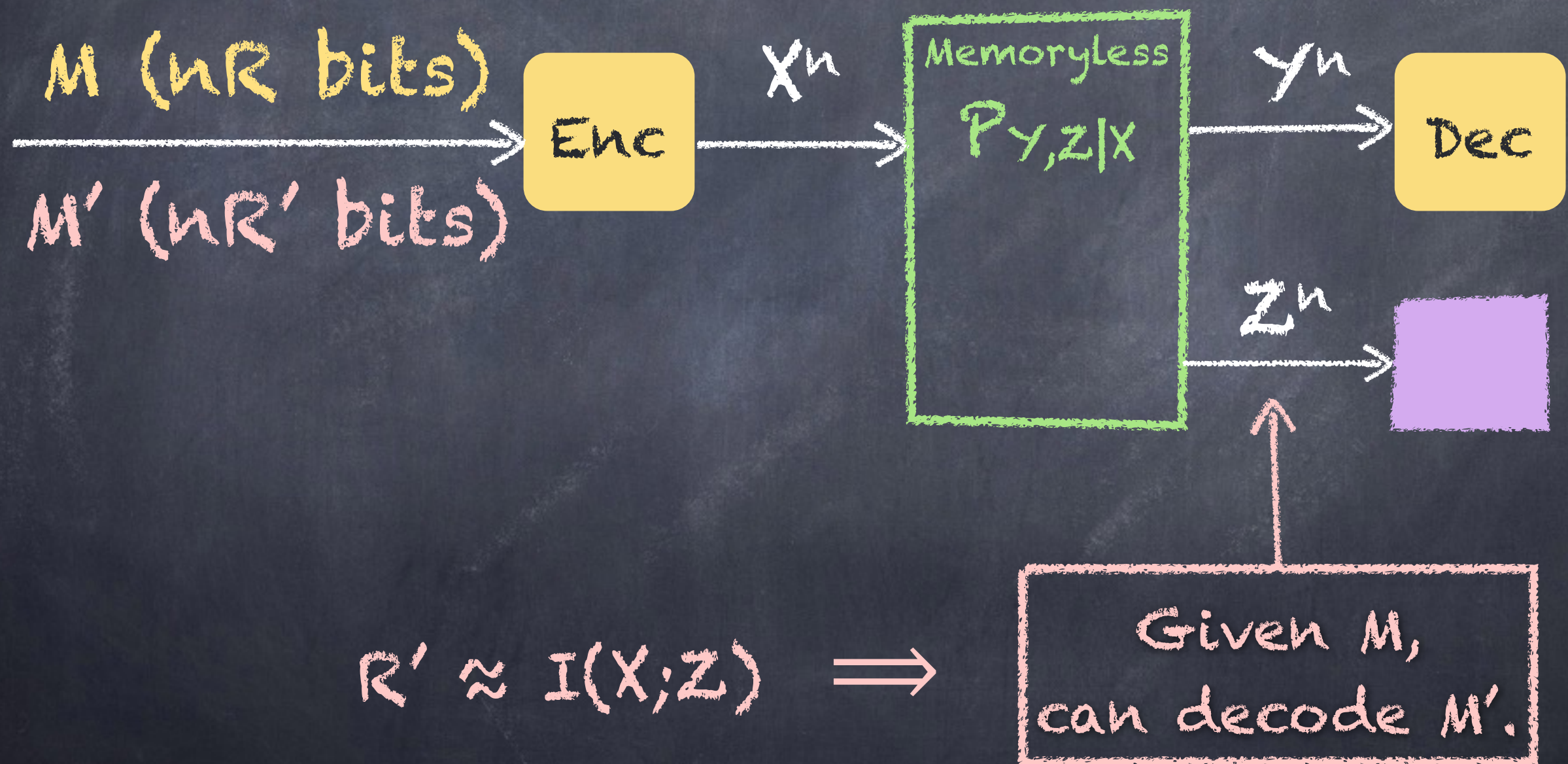
Encoding

- Random Codebook
- Pad with random garbage bits



Transmitted together in one block

Encoding Diagram



Wyner's security argument

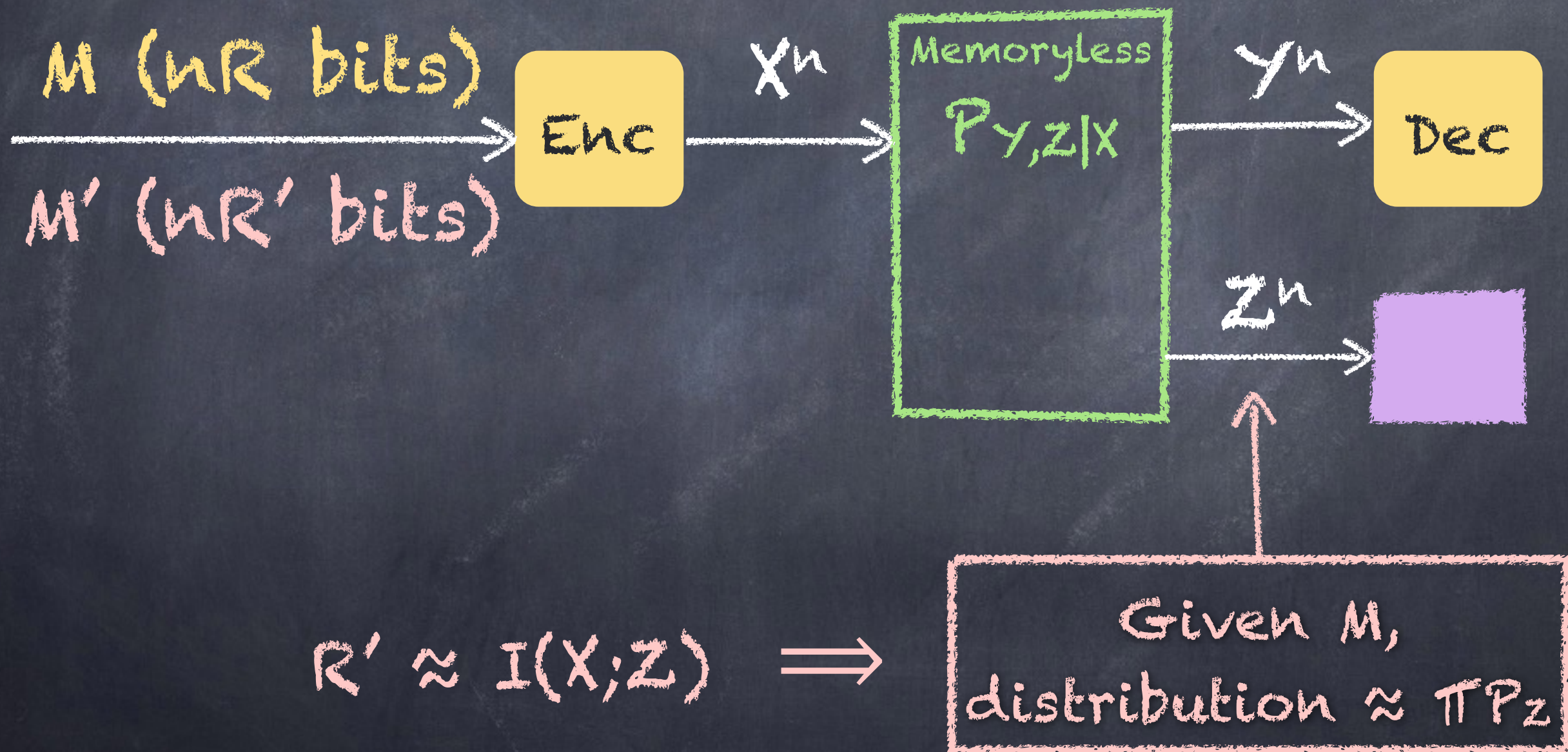
$$I(M, M'; Z^n) = I(M; Z^n) + I(M'; Z^n | M)$$

$$\uparrow$$
$$I(X^n; Z^n) \approx n I(X; Z)$$

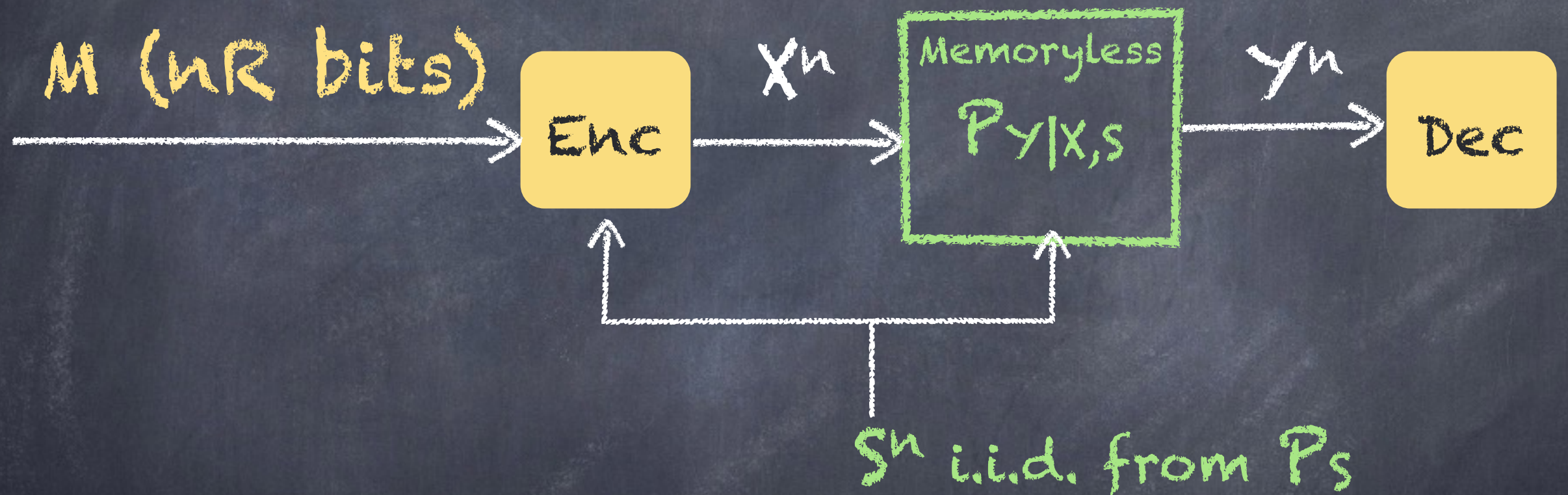
$$\uparrow$$
$$H(M') = n R'$$

Decodable if
 $R' < I(X; Z)$

Encoding Concept



Gelfand-Pinsker (state known to encoder)

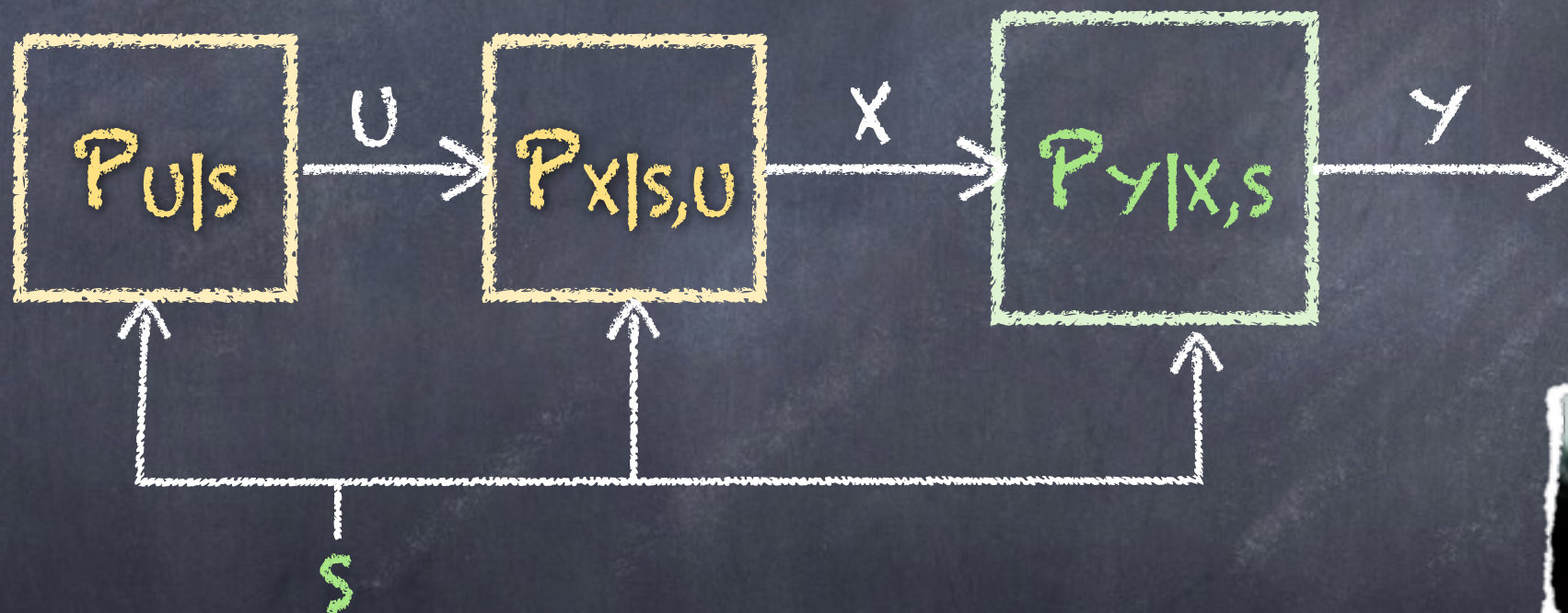


Capacity:

- Reliable communication

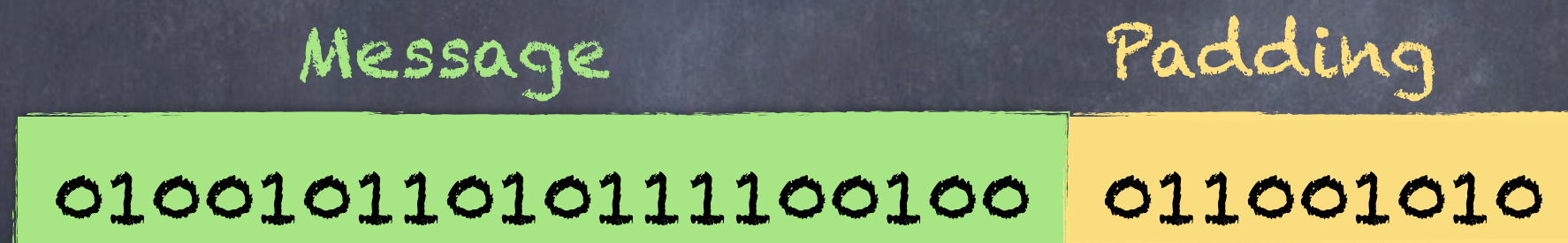
Solution (1980)

$$C = \max_{P_{X,U|S}} I(U; Y) - I(U; S)$$



Encoding

- Random Codebook
- Pad with skillfully chosen bits



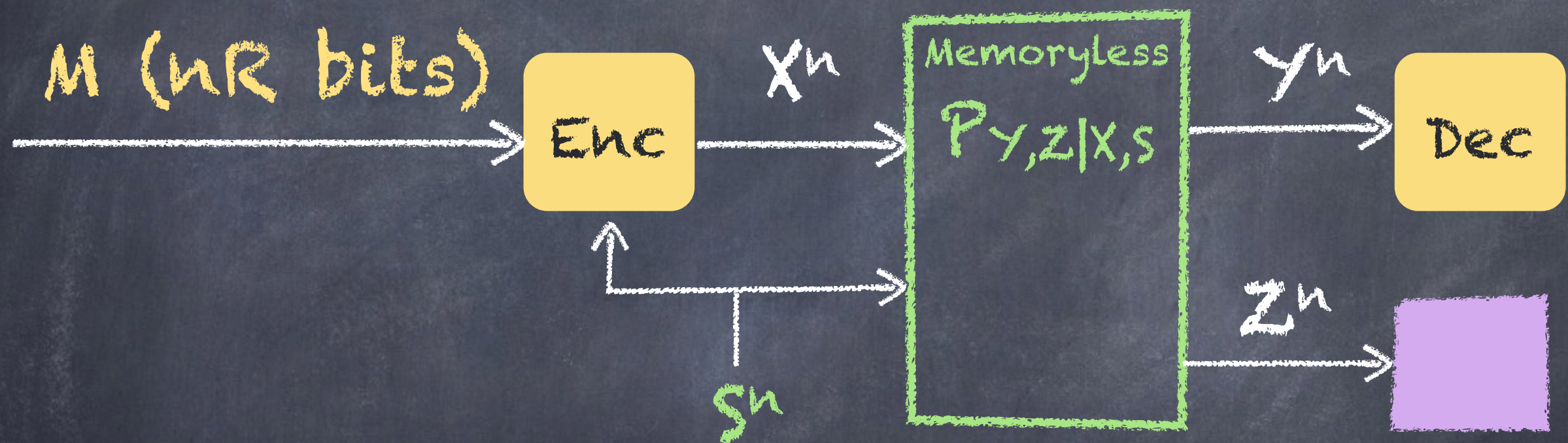
Transmitted together in one block

Two red arrows originate from the text 'Transmitted together in one block' and point upwards to the 'Message' and 'Padding' columns of the table above, indicating that both are part of a single transmission block.

Similarities

- Virtually the same
 - Same encoding
 - Same converse (except, iid S^n allows a skipped step)
 - Same problem statement:
 - Wiretap: M independent of Z^n
 - Gelfand-Pinsker: M independent of S^n

Wiretap Channel with State

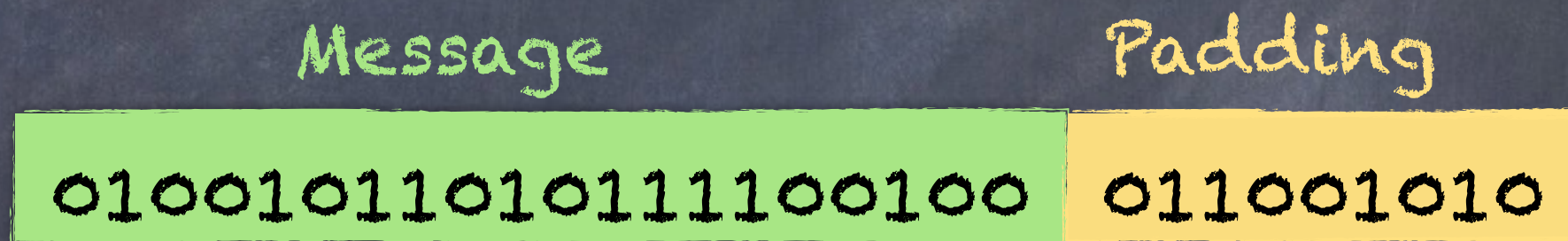


Secrecy Capacity:

- Reliable communication
- Z^n contains no information about M

Same Encoding

$$C_s \geq \max_{P_{X,U|S}} I(U; Y) - \max \left\{ \begin{array}{l} I(U; Z), \\ I(U; S) \end{array} \right\}$$



Transmitted together in one block

Extract Key

Assume S is known to the intended receiver as well:

$$C_s \geq \max_{P_{X,U|S}} \min \left\{ \begin{array}{l} I(U; Y|S), \\ H(S|Z, U) \end{array} \right\}$$

Better in some cases!

Chia and El Gamal, 2012

Note: They consider causal state information.

This region is adapted to take advantage of non-causal state information.

Combined

Assume S is known to the intended receiver as well:

$$C_s \geq \max_{P_{X,U|S}} \min \left\{ \begin{array}{l} I(U; Y|S), \\ H(S|Z, U) + [I(U; Y, S) - I(U; Z)]_+ \end{array} \right\}$$

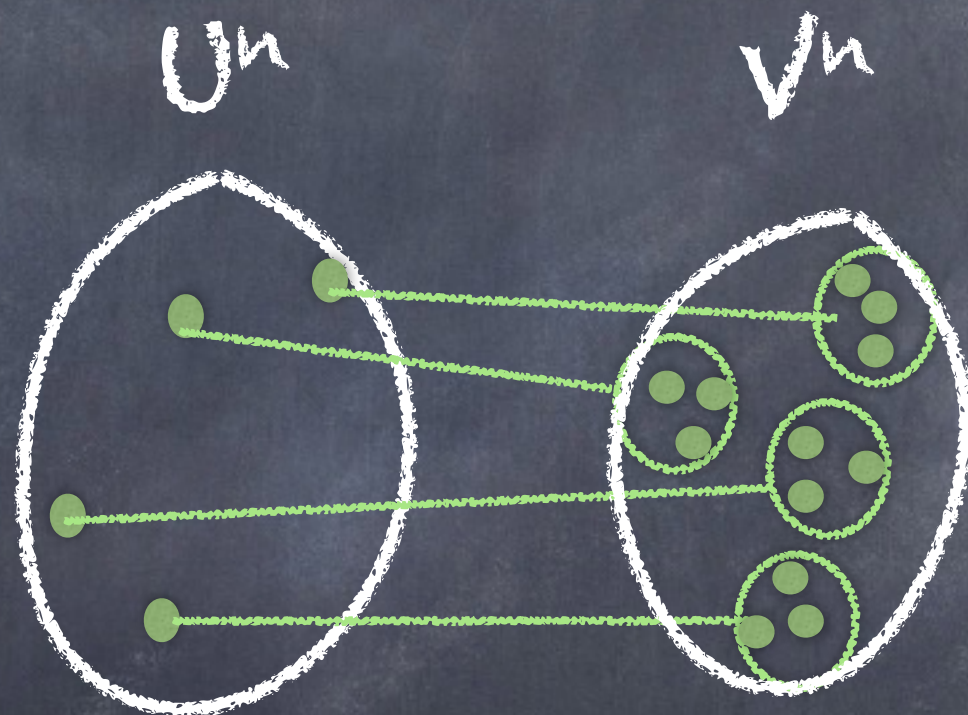
Chia and El Gamal, 2012

Note: They consider causal state information.

This region is adapted to take advantage of non-causal state information.

Our Scheme

Superposition code



U^n index is padding only

V^n index is message and padding

All secrecy comes from V

U^n is decoded by the eavesdropper

Our Scheme

$$C_s \geq \max_{P_{X,U,V|S} : I(U;Y) \geq I(U;S)} \min \left\{ \begin{array}{l} I(U,V;Y) - I(U,V;S), \\ I(V;Y|U) - I(V;Z|U) \end{array} \right\}$$

Can mimic Chia and El Gamal's key
extraction by setting $V=S$

Beats previous regions

Other Related Work

- Prabhakaran, Eswaran, and Ramchandran, 2012:
 - Same superposition code but require $U-V-(S,X)$ and $U \perp S$.
- Bassi, Bunin, Piantanida, and Shamai, 2016 (several papers):
 - Key generation and secure communication
 - Sources independent of channel
 - Generalized feedback

Simple Special case

- Unlimited public noise-free channel
- "Key Capacity with one-way communication"

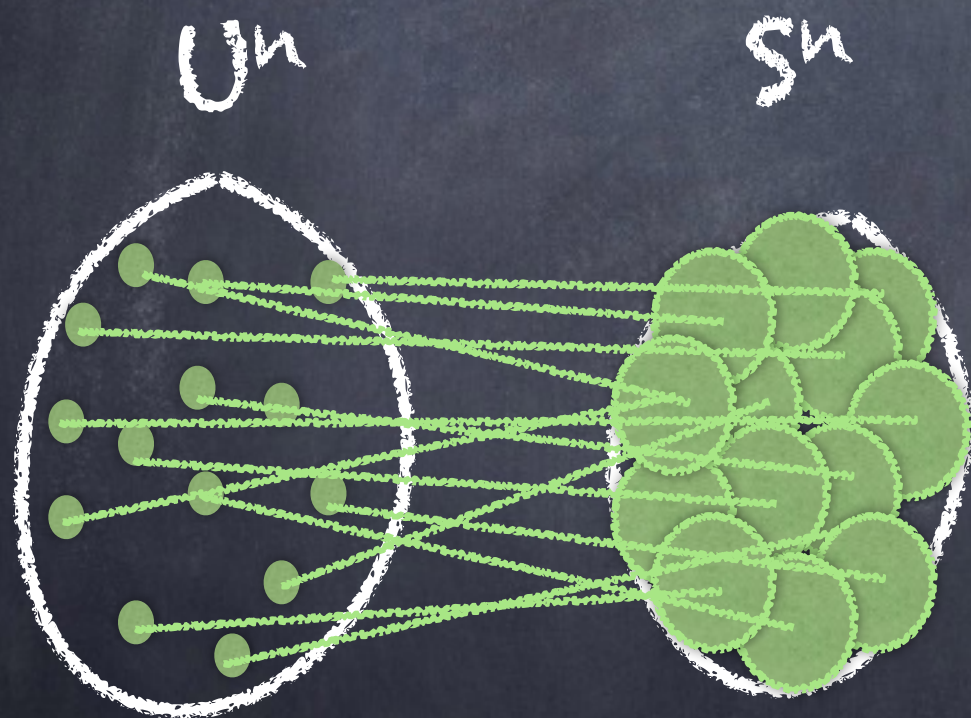
$$C_s = \max_{P_{U,V|S_x}} I(V; S_y | U) - I(V; S_z | U)$$

Achieved by our scheme

Apparently not by any others

Analysis Trick

- Likelihood encoder to choose padding



Soft Covering

Distribution 1:

- Choose U^n codeword uniformly at random
- Generate S^n memorylessly from U^n

Distribution 2:

- Use the above conditional distribution of U^n given V^n (this is the likelihood encoder)
- Let S^n be iid

Observation

- Lower layer of superposition code intended to be decoded by eavesdropper (i.e. "decoy")
- Villard-Piantanida secure source coding
- Key agreement with one way communication

Differential Privacy as a Mutual Information Constraint

Paul Cuff and Langqing Yu



Database Privacy

- Let $X_1, X_2, \dots, X_n$ be entries in a database
 - E.g. X_i is personal information about person i
- Let Y be the response to a query
 - We can denote Y_q to indicate that it depends on the query
- The job of the information provider is to answer queries and protect individual privacy

Differential Privacy

- ϵ -DP:

- Let x and x' differ in only one entry (i.e. $x_i = x'_i$ for all but one i)
- $p(y|x) \leq e^\epsilon p(y|x')$
- Why x and x' differ in only one spot?
 - Convince someone to put their data in your database
- Why multiplicative constraint?
 - Posterior update is small

A Common Technique

- Add Laplacean noise

Weaker DP

- (ϵ, δ) -DP:
 - $P(Y \in A | x) \leq e^\epsilon P(Y \in A | x') + \delta$
- Additive Gaussian noise often provides privacy

Mutual Information Differential Privacy

ϵ -MI-DP:

$$\max_{i, P_{X^n}} I(X_i; Y | X^{i-1}, X_{i+1}^n) < \epsilon$$

Claim

$$\epsilon\text{-DP} > \text{MI-DP} > (\epsilon, \delta)\text{-DP}$$

Furthermore, if input or output alphabet is finite,

$$\text{MI-DP} = (\epsilon, \delta)\text{-DP}$$

Privacy Ordering

- α -DP $>$ β -DP if for all $\beta > 0$ there exists α such that α -DP $\Rightarrow$ β -DP.

Subadditivity of DP

- Multiple queries:
 - If k queries $Y_{q1}, Y_{q2}, \dots, Y_{qk}$ each have differential privacy ϵ and are conditionally independent, the combined they have $k\epsilon$ privacy.

Simple MI-DP Proof:

$$\begin{aligned} I(X; Y_1, Y_2) &= I(X; Y_1) + I(X; Y_2 | Y_1) \\ &\leq I(X; Y_1) + I(X; Y_2) \end{aligned}$$

For clarity, conditioned database variables are omitted.

Common complaint

- Differentially privacy doesn't not mean that you can't learn about X_i .
- Consider a database with correlated entries.

Simple MI-DP Explanation:

$$I(X_i; Y) \not\leq I(X_i; Y | X^{i-1}, X_{i+1}^n)$$

Precise Bounds

(ϵ, δ) -closeness

$$P \stackrel{(\epsilon, \delta)}{\approx} Q$$

if

$$P(A) \leq e^\epsilon Q(A) + \delta, \quad \forall A \in \mathcal{F},$$

$$Q(A) \leq e^\epsilon P(A) + \delta, \quad \forall A \in \mathcal{F}.$$

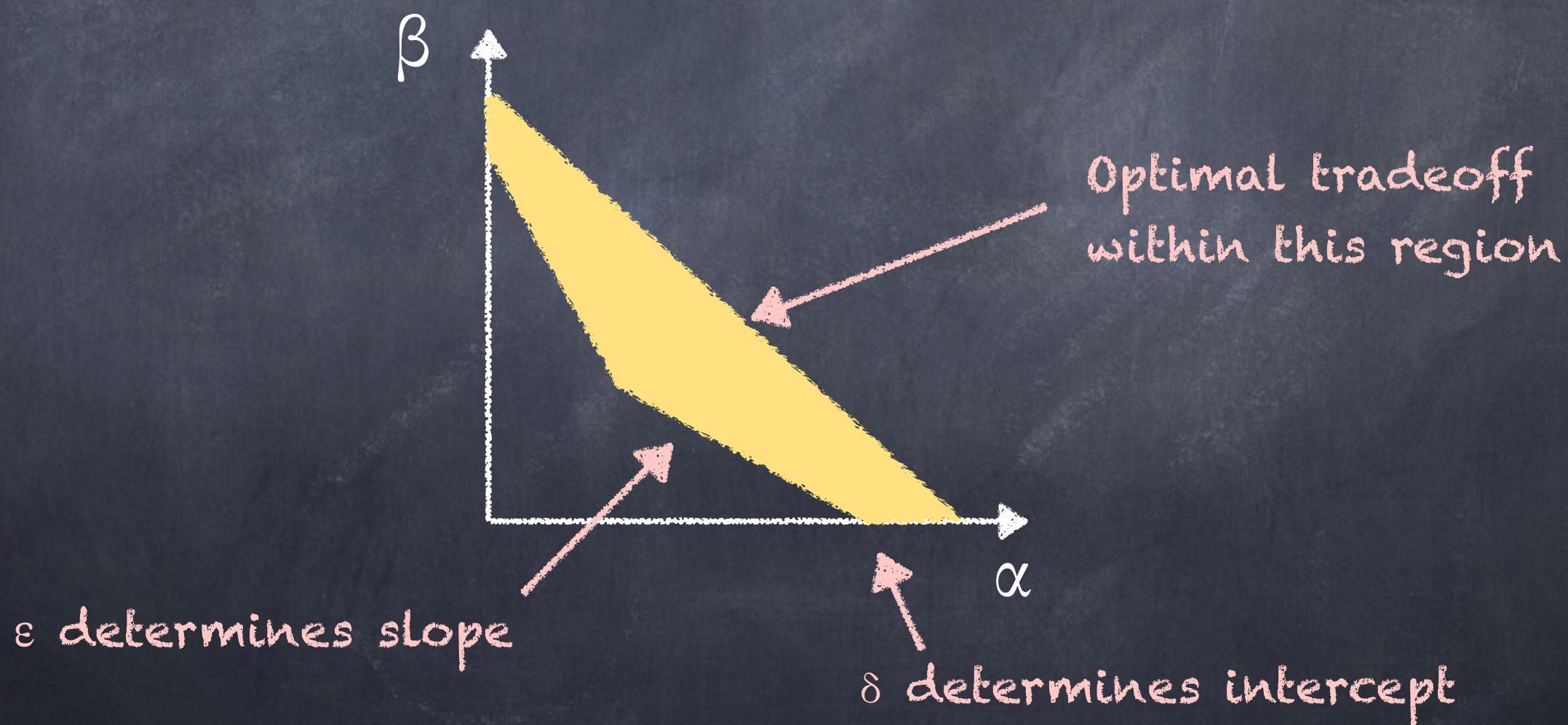
Special Cases

$$P \stackrel{(\epsilon, 0)}{\approx} Q \iff \left| \ln \frac{dP}{dQ}(a) \right| \leq \epsilon \quad \forall a \in \Omega.$$

$$P \stackrel{(0, \delta)}{\approx} Q \iff \|P - Q\|_{TV} \leq \delta.$$

Relation to Detection Theory

$$P \stackrel{(\epsilon, \delta)}{\approx} Q$$



Tightest (ϵ, δ) Conversion

$$P \stackrel{(\epsilon, \delta)}{\approx} Q \implies P \stackrel{(\epsilon', \delta')}{\approx} Q.$$

for

$$\epsilon' \leq \epsilon$$

$$\delta' = 1 - \frac{(e^{\epsilon'} + 1)(1 - \delta)}{e^{\epsilon} + 1}$$

Simple Claim

$$P \stackrel{(\epsilon, 0)}{\approx} Q \implies \begin{aligned} D(P \| Q) &\leq \epsilon \text{ nats}, \\ D(Q \| P) &\leq \epsilon \text{ nats}. \end{aligned}$$

Tightest Claim

$$P \stackrel{(\epsilon, 0)}{\approx} Q \implies \begin{aligned} D(P \| Q) &\leq \epsilon \frac{(e^\epsilon - 1)(1 - e^{-\epsilon})}{(e^\epsilon - 1) + (1 - e^{-\epsilon})} \text{ nats}, \\ D(Q \| P) &\leq \epsilon \frac{(e^\epsilon - 1)(1 - e^{-\epsilon})}{(e^\epsilon - 1) + (1 - e^{-\epsilon})} \text{ nats}. \end{aligned}$$

Pinsker

$$D(P\|Q) \leq \epsilon \text{ nats} \implies P \stackrel{(0, \sqrt{\epsilon/2})}{\approx} Q.$$

Relative entropy to Mutual Information

If

$$D(P_{Y|X=x_1} \| P_{Y|X=x_2}) \leq \epsilon \quad \forall x_1, x_2 \in \mathcal{X}$$

then

$$I(X; Y) \leq \epsilon$$

Hint: Radius of information ball

Mutual Information to Total Variation

$$I(X; Y) \leq \epsilon \quad \implies \quad \left\| P_{Y|X=x_1} - P_{Y|X=x_2} \right\|_{TV} \leq \delta' \\ \forall x_1, x_2 \in \mathcal{X}$$

$$\delta' = 1 - 2h^{-1}(\ln 2 - \epsilon) \\ \leq \sqrt{2\epsilon}$$

Tightest bound,
achieved with
binary channel



Finite Alphabet

$$\left\| P_{Y|X=x_1} - P_{Y|X=x_2} \right\|_{TV} \leq \delta \quad \forall x_1, x_2 \in \mathcal{X} \quad \implies \quad I(X; Y) \leq \epsilon'$$

$$\epsilon' = 2h(\delta) + 2\delta \ln \left(\min \left\{ |\mathcal{Y}|, \max_i |\mathcal{X}_i| + 1 \right\} \right)$$

Continuity of entropy

Continuity of conditional entropy

inspired by Alicki and Fannes, 2004

Observation

$$\max_{P_{X^n}} I(X_i; Y | X^{i-1}, X_{i+1}^n) = \max_{\prod_{t=1}^n P_{X_t}} I(X_i; Y) \quad \forall i$$

Either could be used for definition of MI-DP