A short introduction to Physical Layer Security

LAURA LUZZI

SUPÉLEC
JUNE 8, 2011

Part I

Introduction

Outline

PART 1: INTRODUCTION

PART 2: WYNER'S WIRETAP CHANNEL

PART 3: SECURE COMMUNICATION OVER FADING CHANNELS

PART 4: CODING FOR SECRECY

Main references

- M. Bloch, J. Barros, "Physical-Layer Security: from Information Theory to Security Engineering", Cambridge University Press (in press, release date August 2011)

- Y. Liang, H. V. Poor, S. Shamai, "Information Theoretic Security", *Foundations and Trends in Communications and Information Theory*, vol. 5, 2008

What is Physical-Layer-Security?

- in traditional systems, reliability is guaranteed by channel coding at the physical layer, while security is ensured by encryption protocols at the upper layers

- Example of layered security architecture:

Application Layer	SSH (Secure Shell)
Transport Layer	SSL (Secure Sockets Layer)
Network Layer	IPSec (Internet Protocol Security)
Data Link Layer	WPA (Wi-Fi Protected Access)
Physical Layer	?

- physical layer security aims at exploiting the randomness inherent in noisy channels to provide an additional level of protection at the physical layer

Cryptography vs. Information-Theoretic Security

- public-key cryptography is based on the assumption that certain mathematical functions are hard to invert
  - this computational complexity assumption is mostly unproven
  - security is measured only by resistance to certain attacks, so that it is hard to compare the strengths of different cyphers
  - as the available computational power increases due to advances in technology, these methods may no longer be secure
- information-theoretic security is measured in terms of statistical independence between the attacker's observation and the message
  - even an eavesdropper with unlimited computational power cannot extract any information from the signal

Shannon's Perfect Secrecy

C. E. Shannon, "Communication Theory of Secrecy Systems", *Bell System Technical Journal*, vol. 28, p. 656, 1949.

Information-theoretic secrecy is measured by the eavesdropper's uncertainty about the message given the codeword, called the eavesdropper's equivocation.

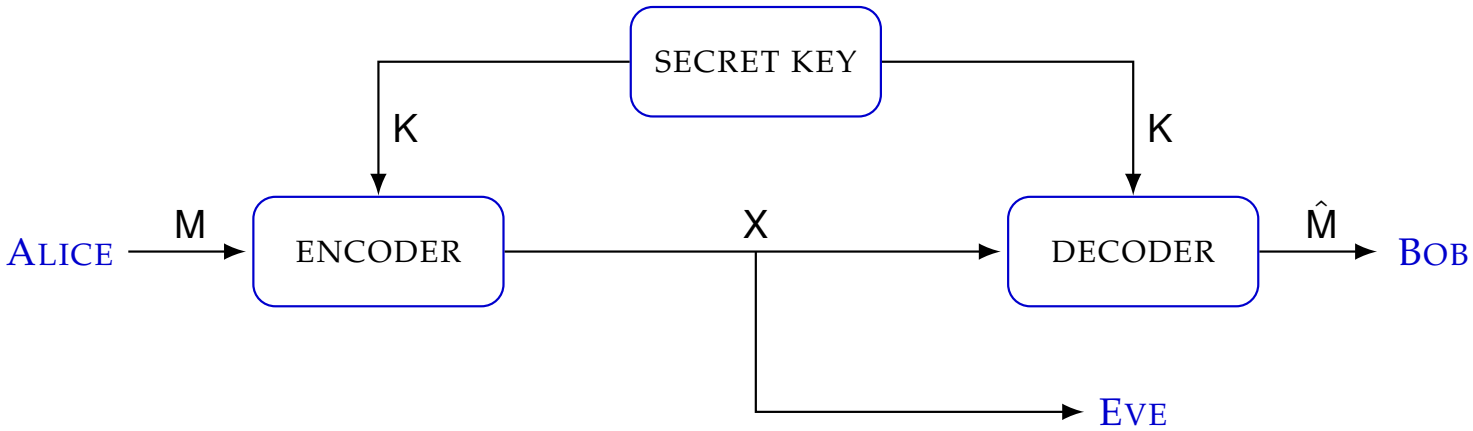
Perfect Secrecy

"After a cryptogram X is intercepted by the enemy, the *a posteriori* probabilities of the message M should be equal to the *a priori* probabilities before interception".

Perfect secrecy is equivalent to each of the following statements:

- $p_{\mathsf{MX}}(m,x) = p_{\mathsf{M}}(m)p_{\mathsf{X}}(x) \quad \forall m \in \mathcal{M}, \forall x \in \mathcal{X}$
- M and X are independent random variables.
- $\mathbb{H}(\mathsf{M}|\mathsf{X}) = \mathbb{H}(\mathsf{M})$.

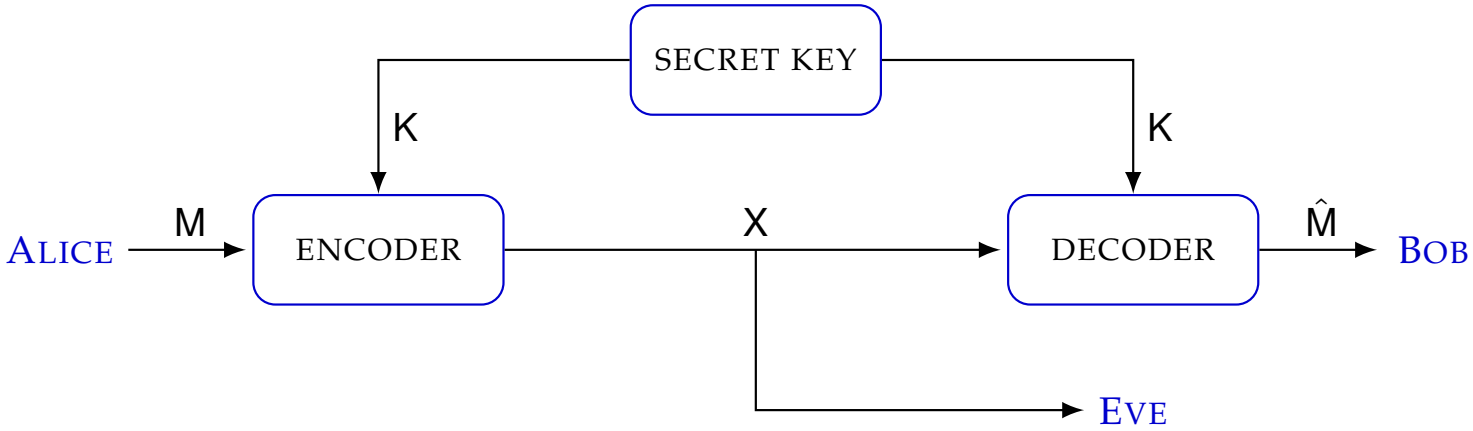
Noiseless channels: Shannon's cypher system



- M confidential message, K secret key
- $e : \mathcal{M} \times \mathcal{K} \to \mathcal{X}$ encoder, $\mathsf{X} = e(\mathsf{M}, \mathsf{K})$ transmitted codeword
- $d : \mathcal{X} \times \mathcal{K} \to \mathcal{M}$ decoder, $\hat{\mathsf{M}} = d(\mathsf{X}, \mathsf{K})$ Bob's estimate of M

- Assumption: Eve knows e and d (but not the value of K).

Noiseless channels: Shannon's cypher system



Proposition

If a coding scheme achieves perfect secrecy and Bob can recover the message without errors, then $\mathbb{H}(\mathsf{K}) \geq \mathbb{H}(\mathsf{M})$.

One-time pad: let $\mathcal{C} = \mathcal{K} = \mathcal{M}$. Let K be uniformly distributed on $\mathcal{M}$.

$$\mathsf{X} = \mathsf{M} \oplus \mathsf{K} \qquad (\text{mod } |\mathcal{M}|)$$

Proof of the Proposition.

Since $P_e = 0$, from Fano's inequality we have $\mathbb{H}(\mathsf{M}|\mathsf{X},\mathsf{K}) = 0$.

$$\mathbb{H}(\mathsf{K}) \geq \mathbb{H}(\mathsf{K}) - \mathbb{H}(\mathsf{K}|\mathsf{X},\mathsf{M}) \geq \mathbb{H}(\mathsf{K}|\mathsf{X}) - \mathbb{H}(\mathsf{K}|\mathsf{X},\mathsf{M}) = \mathbb{I}(\mathsf{K};\mathsf{M}|\mathsf{X}) =$$
$$= \mathbb{H}(\mathsf{M}|\mathsf{X}) - \mathbb{H}(\mathsf{M}|\mathsf{X},\mathsf{K}) = \mathbb{H}(\mathsf{M}|\mathsf{X}) = \mathbb{H}(\mathsf{M})$$

□

The one-time pad achieves perfect secrecy:

$$p_{\mathsf{X}}(x) = \sum_{k\in\mathcal{M}} p_{\mathsf{X}|\mathsf{K}}(x|k)p_{\mathsf{K}}(k) = \sum_{k\in\mathcal{M}} p_{\mathsf{M}}(x \ominus k)\frac{1}{|\mathcal{M}|} = \frac{1}{|\mathcal{M}|}$$

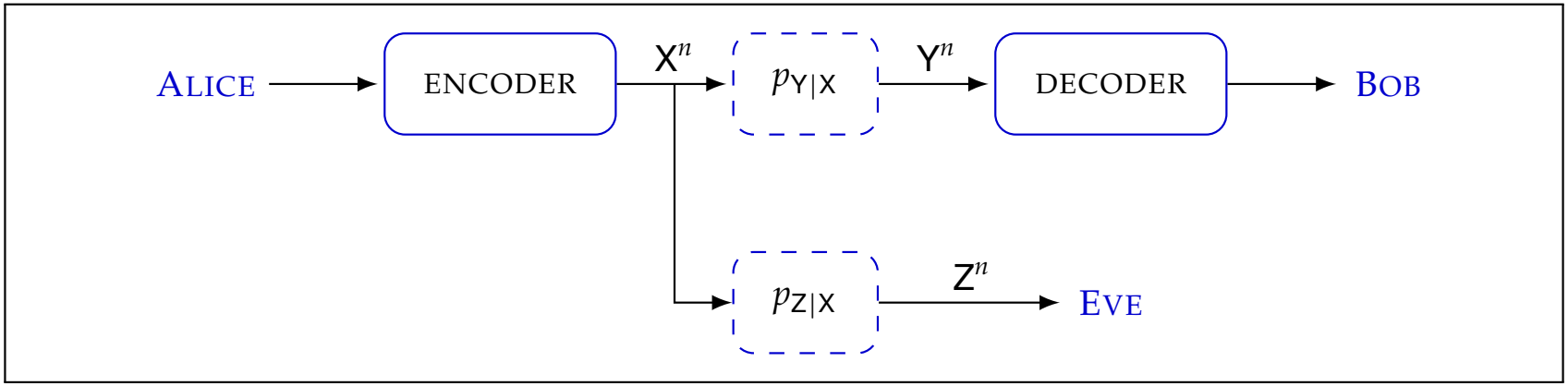
$$\mathbb{I}(\mathsf{M};\mathsf{X}) = \mathbb{H}(\mathsf{X}) - \mathbb{H}(\mathsf{X}|\mathsf{M}) = \mathbb{H}(\mathsf{X}) - \mathbb{H}(\mathsf{K}|\mathsf{M}) = \mathbb{H}(\mathsf{X}) - \mathbb{H}(\mathsf{K}) = 0$$

Conclusion

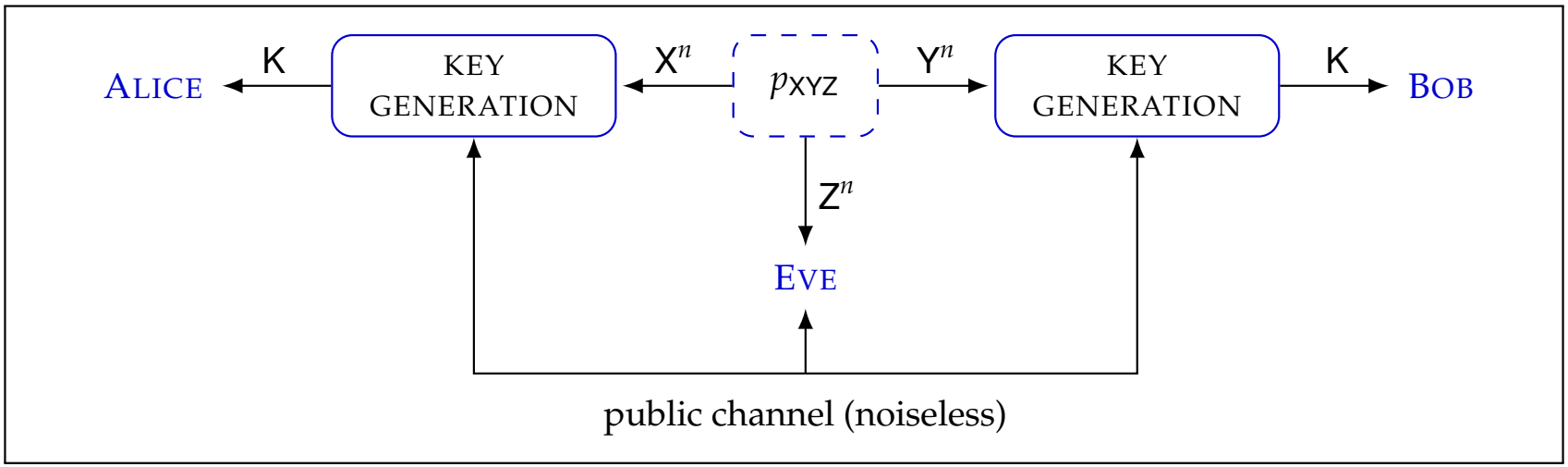
- in the case of noiseless channels, the legitimate receiver must have some advantage over the eavesdropper, otherwise the latter would be able to recover the message as well.
- the one-time pad is a rather pessimistic solution because it still requires to generate and share long keys over a secure channel.
- **Question:** Can the situation improve if we consider noisy channels?

Secure communication over noisy channels

CODING FOR SECRECY:



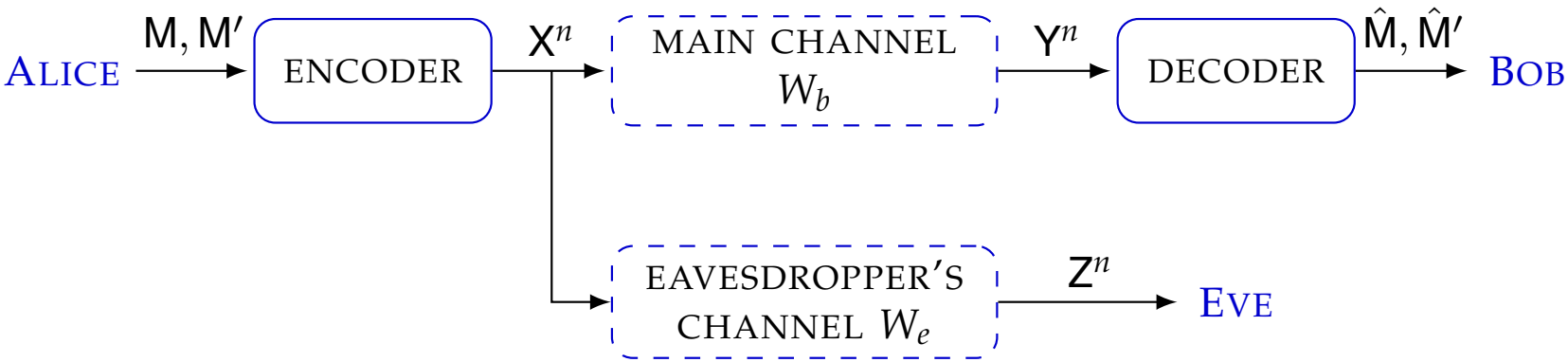
SECRET KEY AGREEMENT:



Part II

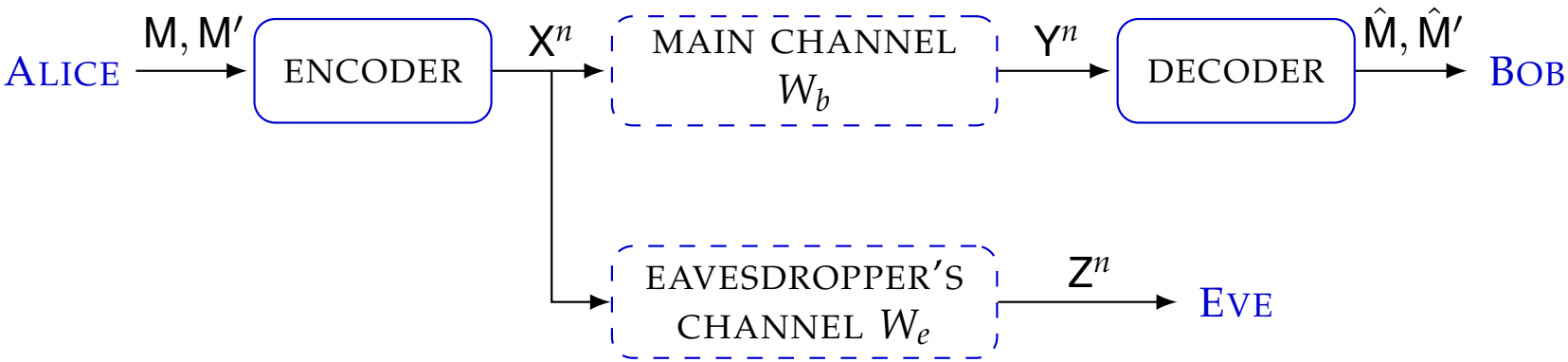
Wyner's Wiretap Channel

The wiretap channel



- M confidential message, M′ random message
- X^n codeword
- Y^n output of Bob's channel, Z^n output of Eve's channel

The wiretap channel

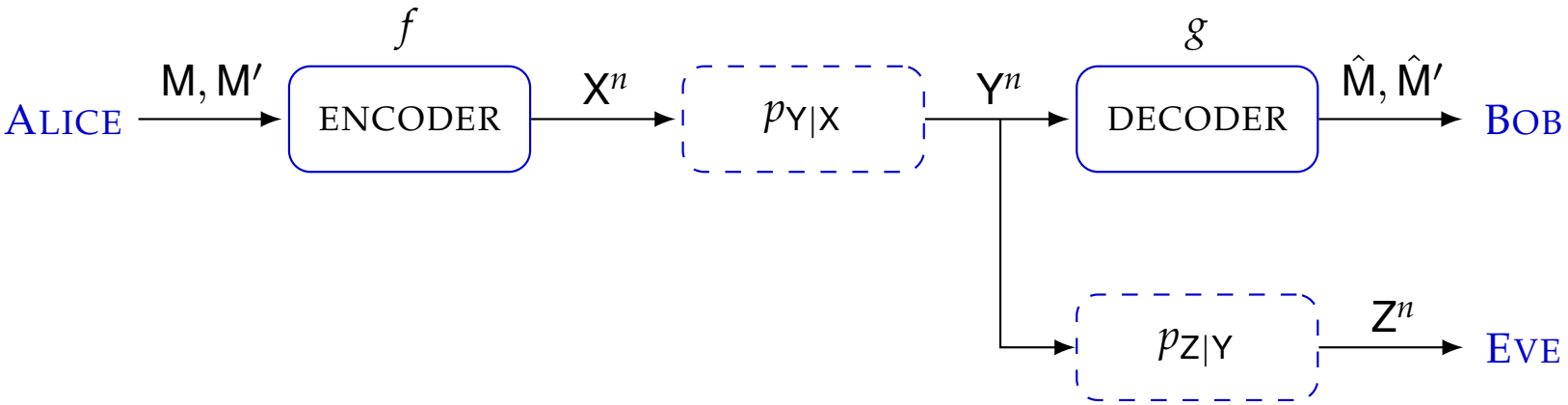


- M confidential message, M′ random message
- X^n codeword
- Y^n output of Bob's channel, Z^n output of Eve's channel

- W_b and W_e are discrete memoryless channels with capacity C_b and C_e

Wyner's model

A. D. Wyner, "The Wire-Tap Channel", *Bell System Technical Journal*, vol. 54, n. 8



Physically degraded wiretap channel

We suppose that Eve's channel $p_{Z|X}$ is physically degraded with respect to Bob's channel $p_{Y|X}$. That is, the following equivalent statements hold:

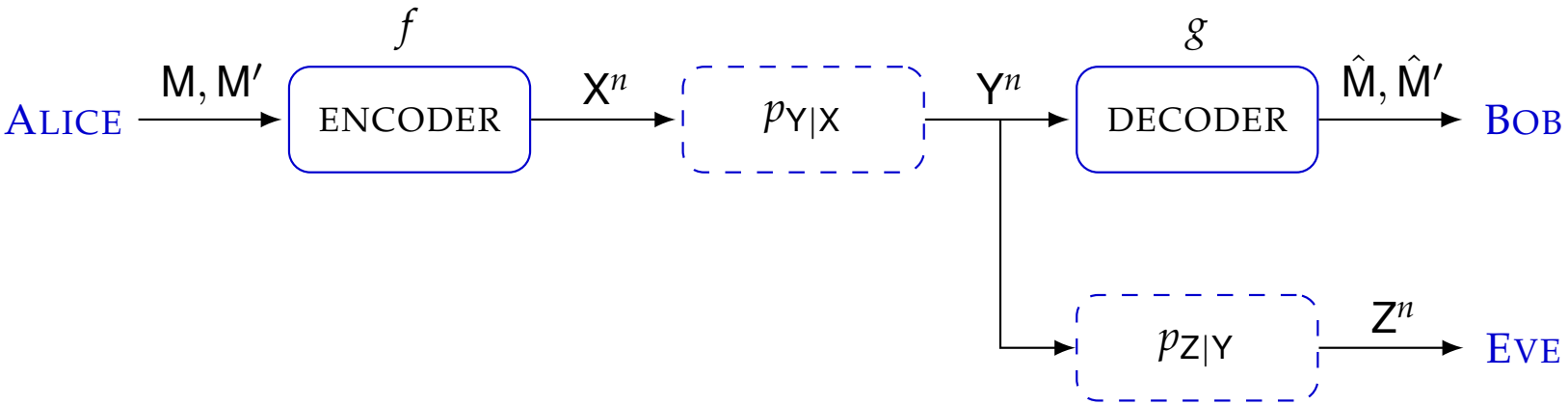
a) $X \to Y \to Z$ is a Markov chain

b) $p_{XYZ}(x,y,z) = p_X(x)p_{Y|X}(y|x)p_{Z|X}(z|x) \quad \forall x \in \mathcal{X},\ \forall y \in \mathcal{Y},\ \forall z \in \mathcal{Z}$

c) $\mathbb{I}(X;Z|Y) = 0$ (X and Z are conditionally independent given Y).

Wyner's model

A. D. Wyner, "The Wire-Tap Channel", *Bell System Technical Journal*, vol. 54, n. 8



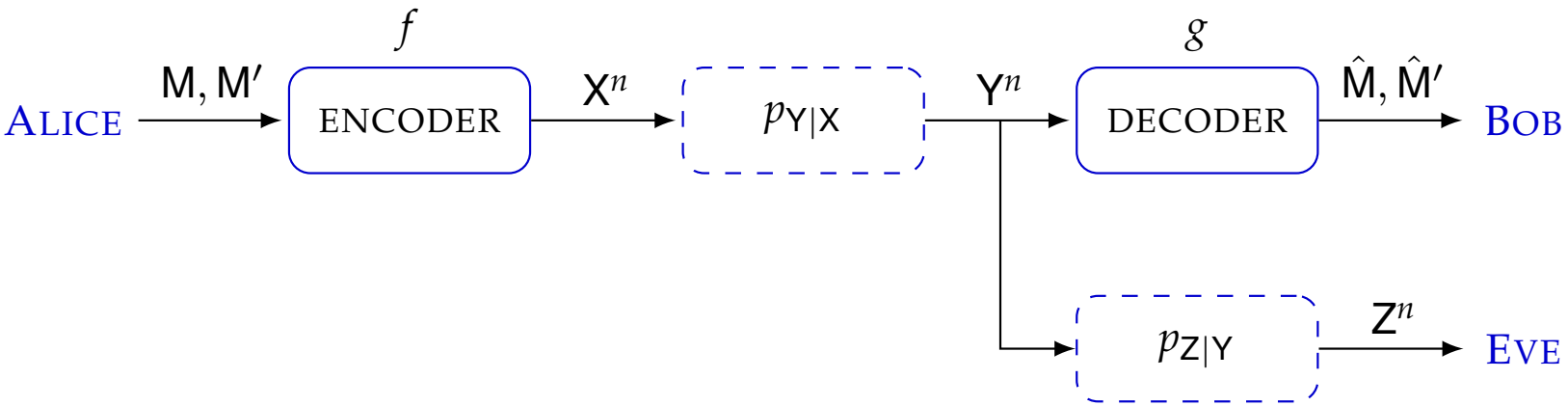
Stochastically degraded wiretap channel

Eve's channel is stochastically degraded with respect to Bob's channel if its conditional *marginal* distribution (but not the joint distribution) is the same as that of a physically degraded channel: there exists a distribution $p_{Z|Y'}$ such that

$$p_{Z|X}(z|x) = \sum_{y \in \mathcal{Y}} p_{Y|X}(y|x) p_{Z|Y'}(z|y).$$

Wyner's model

A. D. Wyner, "The Wire-Tap Channel", *Bell System Technical Journal*, vol. 54, n. 8

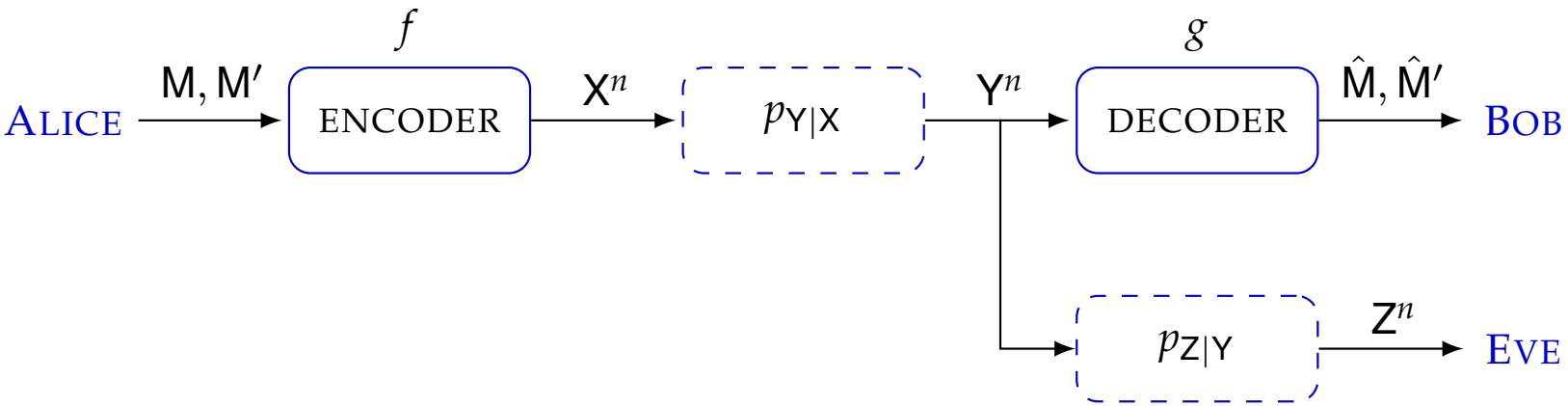


A $(2^{nR}, n)$ **wiretap code** C_n is defined by:

- $\mathcal{M}_n = [\![1, 2^{nR}]\!]$ confidential message set
- $\mathcal{M}'_n = [\![1, 2^{nR'}]\!]$ random message set
- $f : \mathcal{M}_n \times \mathcal{M}'_n \to \mathcal{X}^n$ encoding function for Alice
- $g : \mathcal{Y}^n \to \mathcal{M}_n \times \mathcal{M}'_n$ decoding function for Bob

Wyner's model

A. D. Wyner, "The Wire-Tap Channel", *Bell System Technical Journal*, vol. 54, n. 8



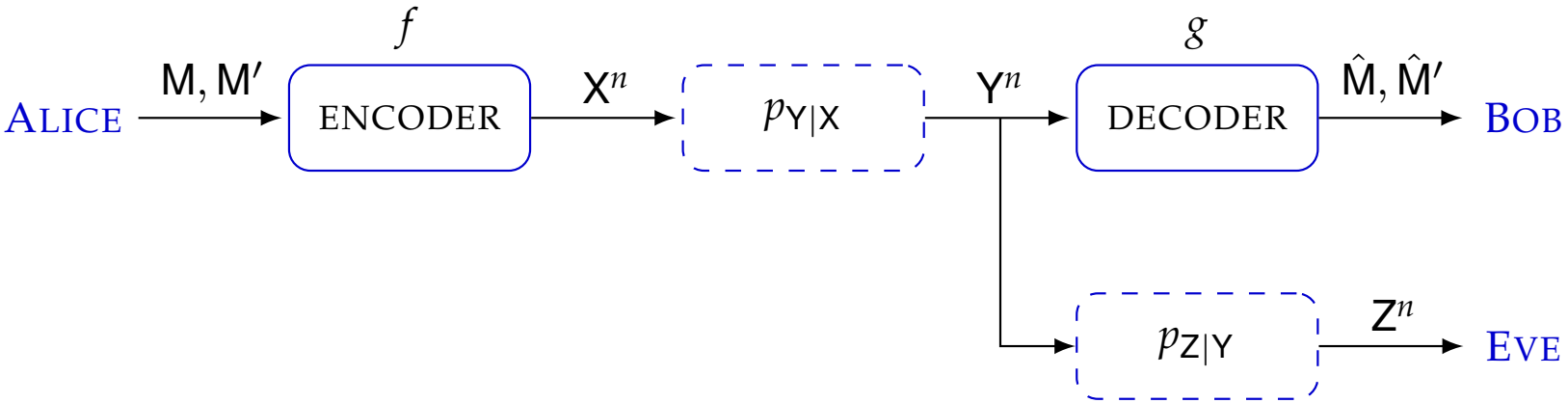
A $(2^{nR}, n)$ **wiretap code** C_n is defined by:

- $\mathcal{M}_n = [\![1, 2^{nR}]\!]$ confidential message set
- $\mathcal{M}'_n = [\![1, 2^{nR'}]\!]$ random message set
- $f : \mathcal{M}_n \times \mathcal{M}'_n \to \mathcal{X}^n$ encoding function for Alice
- $g : \mathcal{Y}^n \to \mathcal{M}_n \times \mathcal{M}'_n$ decoding function for Bob

- Assumption: the eavesdropper knows the code C_n.

Wyner's model

A. D. Wyner, "The Wire-Tap Channel", *Bell System Technical Journal*, vol. 54, n. 8



- Probability of error for Bob: $P_e(C_n) = \mathbb{P}\{\hat{M} \neq M\}$
- The equivocation of $\{C_n\}$ is Eve's uncertainty concerning the message:

$$\mathbf{E}(C_n) = \mathbb{H}(M|Z^n, C_n)$$

- The leakage of $\{C_n\}$ is the amount of information stolen by Eve:

$$\mathbf{L}(C_n) = \mathbb{I}(M; Z^n|C_n)$$

Strong and weak secrecy

- replace perfect secrecy with asymptotic secrecy when the codeword length tends to infinity
- strong secrecy: $\lim_{n\to\infty} \mathbf{L}(C_n) = \lim_{n\to\infty} \mathbb{I}(\mathsf{M}; \mathsf{Z}^n | C_n) = 0$
- weak secrecy: $\lim_{n\to\infty} \frac{1}{n}\mathbf{L}(C_n) = \lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathsf{M}; \mathsf{Z}^n | C_n) = 0$

Strong and weak secrecy

- replace perfect secrecy with asymptotic secrecy when the codeword length tends to infinity
- strong secrecy: $\lim_{n\to\infty} \mathbf{L}(C_n) = \lim_{n\to\infty} \mathbb{I}(\mathsf{M}; \mathsf{Z}^n | C_n) = 0$
- weak secrecy: $\lim_{n\to\infty} \frac{1}{n}\mathbf{L}(C_n) = \lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathsf{M}; \mathsf{Z}^n | C_n) = 0$

weak secrecy is not enough:
the amount of information stolen by Eve might still tend to infinity!

Strong and weak secrecy

- replace perfect secrecy with asymptotic secrecy when the codeword length tends to infinity
- strong secrecy: $\lim_{n\to\infty} \mathbf{L}(C_n) = \lim_{n\to\infty} \mathbb{I}(\mathsf{M}; \mathsf{Z}^n | C_n) = 0$
- weak secrecy: $\lim_{n\to\infty} \frac{1}{n}\mathbf{L}(C_n) = \lim_{n\to\infty} \frac{1}{n}\mathbb{I}(\mathsf{M}; \mathsf{Z}^n | C_n) = 0$

weak secrecy is not enough:
the amount of information stolen by Eve might still tend to infinity!

Other secrecy metrics: more generally, given a metric d on the set of joint probability distributions on $\mathcal{M}_n \times \mathcal{Z}^n$, one can require that

$$\lim_{n\to\infty} d(p_{\mathsf{M}\mathsf{Z}^n}, p_{\mathsf{M}} p_{\mathsf{Z}^n}) = 0.$$

Remark: Asymptotic weak secrecy is still much stronger than asking for example that the block error probability of the eavesdropper should be greater than some constant γ. Example: use a one-time pad to protect only one bit. The block error probability of Eve is still $\frac{1}{2}$, but on the other side, Eve knows almost all the bits in the message!
Or, I can do a one-time pad on γn bits. So the average bit error probability is $1 - \frac{\gamma}{2}$, but $\mathbb{I}(\mathsf{M}; \mathsf{Z}^n) = (1-\gamma)n$ also.

Example showing that weak secrecy is not enough:

Suppose both Bob and Eve's channels are noiseless. Alice uses a one-time-pad for the first $n - t$ bits. The last t bits are left unprotected.

$$\mathbb{I}(\mathsf{M}|\mathsf{Z}^n) = \frac{t}{n}.$$

I can for example choose $t = \lfloor \sqrt{n} \rfloor$ and I still have weak secrecy!

Simple example of wiretap code achieving strong secrecy:

- Noiseless main channel, BEC eavesdropper's channel.
- Alice associates codewords in $\{0,1\}^n$ to $\mathsf{M}=0$ or $\mathsf{M}=1$ according to parity.
- With probability $1-(1-\varepsilon)^n$, Eve loses at least 1 bit $\quad\Rightarrow\quad$ she is unable to recover the parity. Define

$$\mathsf{E}=\begin{cases}0 & \text{if } \mathsf{Z}^n \text{ contains no erasures;}\\ 1 & \text{otherwise}\end{cases}$$

- $\mathbf{E}(C_n)=\mathbb{H}(\mathsf{M}|\mathsf{Z}^n)\geq\mathbb{H}(\mathsf{M}|\mathsf{Z}^n,\mathsf{E})=\mathbb{P}(\mathsf{E}=1)\mathbb{H}(\mathsf{M}|\mathsf{Z}^n,\mathsf{E}=1)+\mathbb{P}(\mathsf{E}=0)\mathbb{H}(\mathsf{M}|\mathsf{Z}^n,\mathsf{E}=0)=\mathbb{H}(\mathsf{M})(1-(1-\varepsilon)^n)=1-(1-\varepsilon)^n$
- So we have strong secrecy. Unfortunately the rate goes to zero! Is it possible to do better?

Weak rate-equivocation region and secrecy capacity

- A weak rate-equivocation pair (R, R_e) is achievable if $\exists$ a sequence $\{C_n\}$ of $(2^{nR}, n)$ codes such that

$$\lim_{n\to\infty} P_e(C_n) = 0$$

$$\liminf_{n\to\infty} \frac{1}{n}\mathbf{E}(C_n) \geq R_e$$

- The weak rate-equivocation region of the wiretap channel is

$$\mathcal{R} = \overline{\{(R, R_e) \text{ achievable rate-equivocation pair}\}}.$$

- The weak secrecy capacity of the channel is

$$C_s = \sup\{R \mid (R, R) \in \mathcal{R}\}.$$

If M is uniformly distributed in $\mathcal{M}_n$, and (R, R) is achievable, then this implies that the weak secrecy condition is satisfied:

- The code rate is $R_n = \frac{1}{n} \log |\mathcal{M}_n| = \frac{1}{n} \mathbb{H}(\mathsf{M})$
- $\mathbf{E}(C_n) = \mathbb{H}(\mathsf{M}|\mathsf{Z}^n, C_n) \leq \mathbb{H}(\mathsf{M})$ since conditioning does not increase entropy

$$\Rightarrow \limsup_{n\to\infty} \frac{1}{n} \mathbf{L}(C_n) = \limsup_{n\to\infty} \frac{1}{n} \mathbb{I}(\mathsf{M}; \mathsf{Z}^n | C_n) =$$

$$= \limsup_{n\to\infty} \frac{1}{n} \left(\mathbb{H}(\mathsf{M}|C_n) - \mathbb{H}(\mathsf{M}|\mathsf{Z}^n, C_n) \right) =$$

$$= \limsup_{n\to\infty} \frac{1}{n} \mathbb{H}(\mathsf{M}) - \liminf_{n\to\infty} \mathbf{E}(C_n) \leq 0$$

But $\mathbf{L}(C_n) \geq 0 \;\Rightarrow$ the limits exists and is 0.

Strong rate-equivocation region and secrecy capacity

- A strong rate-equivocation pair (R, R_e) is achievable if $\exists$ a sequence $\{C_n\}$ of $(2^{nR}, n)$ codes such that

$$\lim_{n\to\infty} P_e(C_n) = 0$$
$$\liminf_{n\to\infty} (\mathbf{E}(C_n) - nR_e) \geq 0$$

- The strong rate-equivocation region of the wiretap channel is

$$\mathcal{R}^{(s)} = \overline{\{(R, R_e) \text{ achievable rate-equivocation pair}\}}.$$

- The strong secrecy capacity of the channel is

$$C_s^{(s)} = \sup\{R \mid (R, R) \in \mathcal{R}^{(s)}\}.$$

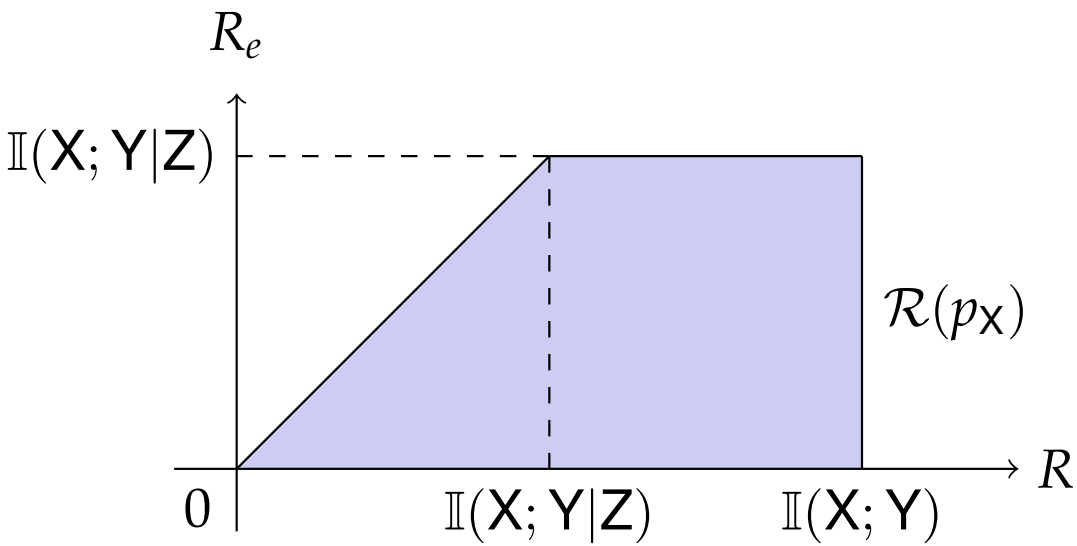
Rate-equivocation region for the wiretap channel

Theorem (Wyner)

The weak rate-equivocation region of the wiretap channel is given by

$$\mathcal{R} = \bigcup_{p_X} \mathcal{R}(p_X), \quad \textit{where}$$

$$\mathcal{R}(p_X) = \{(R, R_e) \quad | \quad 0 \leq R_e \leq R \leq \mathbb{I}(X;Y), \quad 0 \leq R_e \leq \mathbb{I}(X;Y|Z)\}$$



Rate-equivocation region for the wiretap channel

Theorem (Wyner)

The weak rate-equivocation region of the wiretap channel is given by

$$\mathcal{R} = \bigcup_{p_X} \mathcal{R}(p_X), \quad \textit{where}$$

$$\mathcal{R}(p_X) = \{(R, R_e) \quad | \quad 0 \leq R_e \leq R \leq \mathbb{I}(X;Y), \quad 0 \leq R_e \leq \mathbb{I}(X;Y|Z)\}$$

Corollary

$$C_s = \max_{p_X} \mathbb{I}(X;Y|Z) = \max_{p_X} \left(\mathbb{I}(X;Y) - \mathbb{I}(X;Z)\right)$$

$$\Rightarrow C_s \geq \max_{p_X} \mathbb{I}(X;Y) - \max_{p_X} \mathbb{I}(X;Z) \geq C_b - C_e$$

Remark: in general the inequality is strict!

Symmetric channels

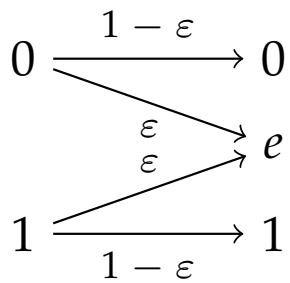
Definition (Cover and Thomas)

A DMC is CT-symmetric if:

a) every row of the transition matrix is a permutation of every other row,

b) every column is a permutation of every other column.

Problem: The BEC is not CT-symmetric:



$\Rightarrow$ transition matrix:

$$\left(\begin{array}{cc|c} 1-\varepsilon & 0 & \varepsilon \\ 0 & 1-\varepsilon & \varepsilon \end{array}\right)$$

Definition (Gallager)

A DMC is G-symmetric if the output alphabet can be partitioned into subsets such that (a) and (b) hold for all the corresponding sub-matrices.

Symmetric channels

S. K. Leung-Yan-Cheong, "On a special class of wiretap channels", *IEEE Trans. Inform. Theory*, vol. 23, n. 5, 1977.

Lemma (Gallager)

The uniform input distribution achieves the capacity of a G-symmetric channel.

Lemma

$\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$ *is a concave function of* p_{X} *for a fixed* $p_{\mathsf{YZ}|\mathsf{X}}$.

Proposition (Leung-Yan-Cheong)

If both the main channel and the eavesdropper's channel are G-symmetric, then

$$C_s = C_b - C_e.$$

Proof of Leung-Yan-Cheong's theorem.

- $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$ is a concave function of p_{X}.
- The uniform distribution $p_{\bar{\mathsf{X}}}$ maximizes both $\mathbb{I}(\mathsf{X};\mathsf{Y})$ and $\mathbb{I}(\mathsf{X};\mathsf{Z}) \Rightarrow$ in particular $p_{\bar{\mathsf{X}}}$ is a stationary point for both $\mathbb{I}(\mathsf{X};\mathsf{Y})$ and $\mathbb{I}(\mathsf{X};\mathsf{Z})$ and so it is also a stationary point for $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$.
- $p_{\bar{\mathsf{X}}}$ is an internal point of the set of discrete probability distributions on $\mathcal{X}$ which is a simplex in $\mathbb{R}^{|\mathcal{X}|}$ (indeed, it is the barycenter of the simplex).
- $\mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z})$ is also a C^1 function of p_{X} in a neighborhood of $p_{\bar{\mathsf{X}}}$. To show this, one ingredient is to recall that for G-symmetric channels, the output probabilities corresponding to a capacity-achieving input are all strictly positive (provided that each output can be reached from some input). Then you can write the expression of the mutual information and see directly that it is C^1.
- Now recall that if f concave and C^1, c internal point such that $f'(c) = 0$, then c is a global maximum.

Binning is necessary:

Notation: for the sake of simplicity, we omit the dependence on the code C_n. By contradiction, suppose that the encoder is one-to-one: $\mathbb{H}(\mathsf{X}^n|\mathsf{M}) = 0$. Bob must be able to decode, so the messages must be uniquely determined by codewords: $\mathbb{H}(\mathsf{M}|\mathsf{X}^n) = 0$. Consequently, we also have

$$\begin{aligned} \mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n|\mathsf{M}) &= \mathbb{H}(\mathsf{X}^n|\mathsf{M}) - \mathbb{H}(\mathsf{X}^n|\mathsf{Z}^n, \mathsf{M}) = 0 \\ \mathbb{I}(\mathsf{M}; \mathsf{Z}^n|\mathsf{X}^n) &= \mathbb{H}(\mathsf{M}|\mathsf{X}^n) - \mathbb{H}(\mathsf{M}|\mathsf{X}^n, \mathsf{Z}^n) = 0 \end{aligned} \qquad (*)$$

since conditioning with respect to Z^n doesn't increase entropy.

$$\begin{aligned} \mathbf{L}(C_n) = \mathbb{I}(\mathsf{M}; \mathsf{Z}^n) &\overset{(a)}{=} \mathbb{I}(\mathsf{M}, \mathsf{X}^n; \mathsf{Z}^n) - \mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n|\mathsf{M}) \overset{(a)}{=} \\ &= \mathbb{I}(\mathsf{M}; \mathsf{Z}^n|\mathsf{X}^n) + \mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n) \overset{(*)}{=} \mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n) \end{aligned}$$

For a general wiretap channel, it is not clear whether we can choose the distribution of X^n so that $\frac{1}{n}\mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n) \to 0$ but Bob obtains a non-negligible communication rate.

---

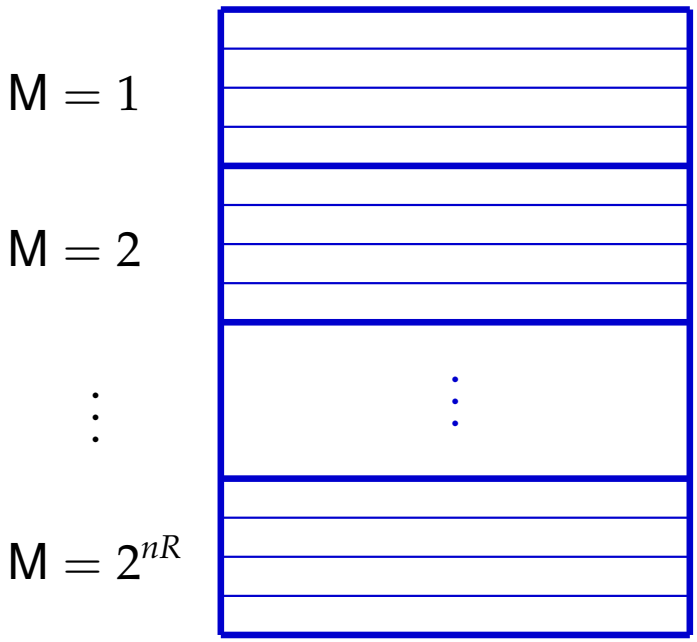
a) chain rule

Random binning for secrecy

- Random codebook generation: Choose a probability distribution p_X on $\mathcal{X}$. $\forall (m, m') \in [\![1, 2^{nR}]\!] \times [\![1, 2^{nR'}]\!]$, randomly generate the codeword $x^n(m, m') \in \mathcal{X}^n$ with i.i.d. components distributed according to p_X.

Random binning for secrecy

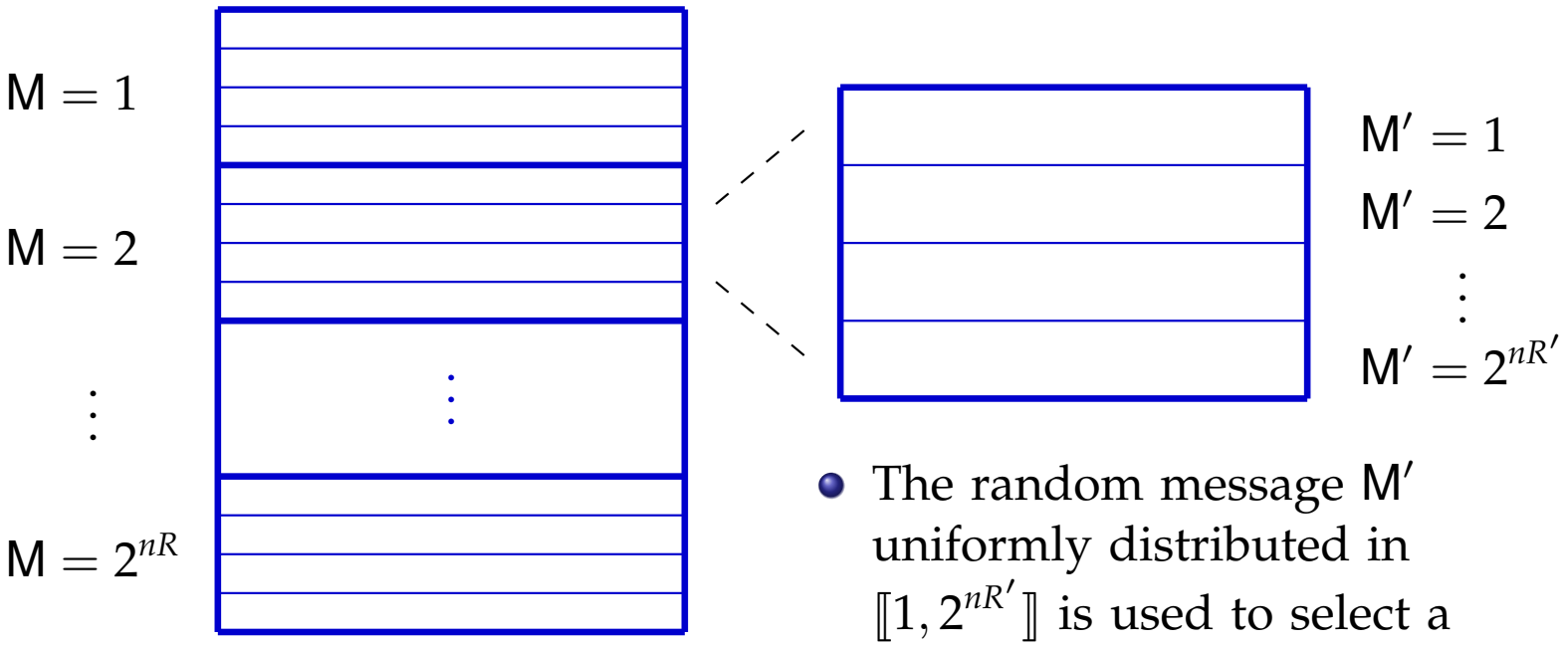
- Random codebook generation: Choose a probability distribution p_X on $\mathcal{X}$. $\forall (m, m') \in [\![1, 2^{nR}]\!] \times [\![1, 2^{nR'}]\!]$, randomly generate the codeword $x^n(m, m') \in \mathcal{X}^n$ with i.i.d. components distributed according to p_X.



  - Each message $M = m$ is associated to a subcode or bin of size $2^{nR'}$.

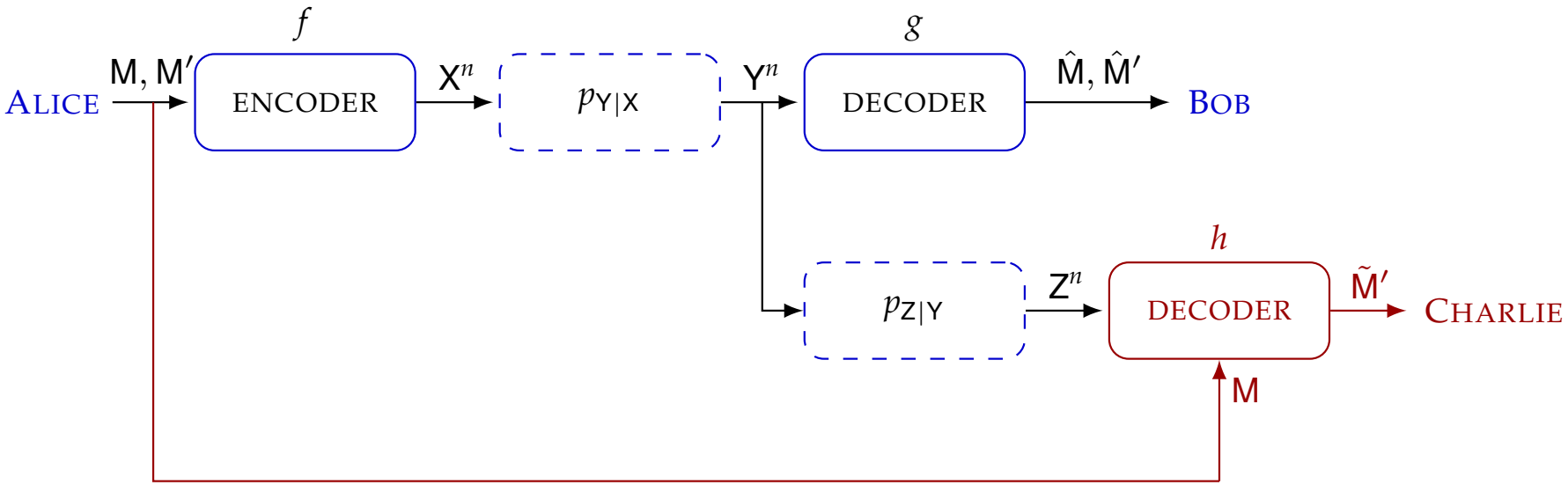
Random binning for secrecy

- Random codebook generation: Choose a probability distribution p_X on $\mathcal{X}$. $\forall (m, m') \in [\![1, 2^{nR}]\!] \times [\![1, 2^{nR'}]\!]$, randomly generate the codeword $x^n(m, m') \in \mathcal{X}^n$ with i.i.d. components distributed according to p_X.



- Each message $\mathsf{M} = m$ is associated to a subcode or bin of size $2^{nR'}$.
- The random message M' uniformly distributed in $[\![1, 2^{nR'}]\!]$ is used to select a codeword inside the bin.

Achievability proof



Introduce a new character, Charlie, who has access to Eve's observation Z^n and to the confidential message M:

- $h : \mathcal{Z}^n \times \mathcal{M}_n \to \mathcal{M}'_n$ MAP decoding function for Charlie
- $\tilde{\mathsf{M}}' = h(\mathsf{Z}^n, \mathsf{M})$ Charlie's estimate of the random message M'

Achievability proof (II)

- define the decoding function g for Bob:

$$g(y^n) = (\hat{m}, \hat{m}') \Leftrightarrow \begin{array}{l}(\hat{m}, \hat{m}') \text{ is the unique message s. t.} \\ (x^n(\hat{m}, \hat{m}'), y^n) \in \mathcal{T}_\varepsilon^n(\mathsf{X}, \mathsf{Y})\end{array}$$

- define the decoding function h for Charlie:

$$h(z^n, m) = \tilde{m}' \Leftrightarrow \begin{array}{l}\tilde{m}' \text{ is the unique message s. t.} \\ (x^n(m, \tilde{m}'), z^n) \in \mathcal{T}_\varepsilon^n(\mathsf{X}, \mathsf{Z})\end{array}$$

- $\forall$ probability distribution p_X on $\mathcal{X}$, we want to show that $\forall R < \mathbb{I}(\mathsf{X}; \mathsf{Y}|\mathsf{Z})$, the random wiretap code ensemble $\{\mathcal{C}_n\}$ generated using p_X contains at least a code $\{C_n\}$ which achieves the rate-equivocation pair (R, R).

Jointly letter-typical sets

- Empirical frequency of $(a,b) \in \mathcal{X} \times \mathcal{Y}$:

$$N(a,b|x^n,y^n) = \#\{i \in [\![1,n]\!] \,|\, (x_i,y_i) = (a,b)\}$$

- Jointly letter-typical set: $\forall \varepsilon > 0$, define

$$\mathcal{T}_\varepsilon^n(\mathsf{X},\mathsf{Y}) = \left\{ (x^n,y^n) \;:\; \forall (a,b),\; \left| \frac{1}{n} N(a,b|x^n,y^n) - p_{\mathsf{XY}}(a,b) \right| \leq \varepsilon p_{\mathsf{XY}}(a,b) \right\}$$

Theorem (Joint Asymptotic Equipartition Property)

X^n, Y^n, $\tilde{\mathsf{X}}^n$ sequences of n i.i.d. random variables with distributions p_{X}, p_{Y} and $p_{\tilde{\mathsf{X}}}$ respectively. Let $0 < \varepsilon < \min_{(x,y)\in\mathcal{X}\times\mathcal{Y}} p_{\mathsf{XY}}(x,y)$.

a) $\mathbb{P}\{(\mathsf{X}^n,\mathsf{Y}^n) \in \mathcal{T}_\varepsilon^n(\mathsf{X},\mathsf{Y})\} \geq 1 - \delta_\varepsilon(n)$

b) *If $\tilde{\mathsf{X}}$ is independent from Y,*

$$\mathbb{P}_{\tilde{\mathsf{X}}\mathsf{Y}}\{(\tilde{\mathsf{X}}^n,\mathsf{Y}^n) \in \mathcal{T}_\varepsilon^n(\mathsf{X},\mathsf{Y})\} \leq 2^{-n(\mathbb{I}(\mathsf{X};\mathsf{Y})-\delta(\varepsilon))}$$

Achievability proof (III)

- $P_e^*(C_n) = \mathbb{P}\left\{(\hat{\mathsf{M}}, \hat{\mathsf{M}}') \neq (\mathsf{M}, \mathsf{M}') \text{ or } \tilde{\mathsf{M}}' \neq \mathsf{M}'\right\} \geq P_e(C_n)$

- By symmetry of the random code ensemble $\mathcal{C}_n$,

$$\mathbb{E}[P_e^*(C_n)] = \mathbb{E}_{C_n}\left[\mathbb{P}\{(\hat{\mathsf{M}}, \hat{\mathsf{M}}') \neq (\mathsf{M}, \mathsf{M}') \text{ or } \tilde{\mathsf{M}}' \neq \mathsf{M}' \mid C_n, \mathsf{M} = 1\}\right].$$

  So we can assume that $\mathsf{M} = 1$.

- $\forall i \in [\![1, 2^{nR}]\!], \forall j \in [\![1, 2^{nR'}]\!]$, define the events

$$\mathcal{E}_{i,j} = \left\{(\mathsf{X}^n(i,j), \mathsf{Y}^n) \in \mathcal{T}_\varepsilon^n(\mathsf{X}, \mathsf{Y})\right\},$$
$$\mathcal{F}_j = \left\{(\mathsf{X}^n(1,j), \mathsf{Z}^n) \in \mathcal{T}_\varepsilon^n(\mathsf{X}, \mathsf{Z})\right\}$$

- $P_e^*(C_n) \leq \mathbb{P}(\mathcal{E}_{1,1}^c) + \sum_{(i,j)\neq(1,1)} \mathbb{P}(\mathcal{E}_{i,j}) + \mathbb{P}(\mathcal{F}_1^c) + \sum_{j\neq 1} \mathbb{P}(\mathcal{F}_j)$

Achievability proof (IV)- Reliability

- The joint AEP implies:

$$\mathbb{P}(\mathcal{E}_{1,1}^c) \leq \delta_\varepsilon(n), \quad \mathbb{P}(\mathcal{F}_1^c) \leq \delta_\varepsilon(n),$$
$$\mathbb{P}(\mathcal{E}_{i,j}) \leq 2^{-n(\mathbb{I}(\mathsf{X};\mathsf{Y})-\delta(\varepsilon))} \quad \forall (i,j) \neq (1,1),$$
$$\mathbb{P}(\mathcal{F}_j) \leq 2^{-n(\mathbb{I}(\mathsf{X};\mathsf{Z})-\delta(\varepsilon))} \quad \forall j \neq 1$$

- $P_e^*(C_n) \leq \delta_\varepsilon(n) + 2^{n(R+R'-\mathbb{I}(\mathsf{X};\mathsf{Y})+\delta(\varepsilon))} + 2^{n(R'-\mathbb{I}(\mathsf{X};\mathsf{Z})+\delta(\varepsilon))}$

- If the random ensemble is constructed by choosing

$$R + R' < \mathbb{I}(\mathsf{X};\mathsf{Y}) - \delta(\varepsilon), \quad R' < \mathbb{I}(\mathsf{X};\mathsf{Z}) - \delta(\varepsilon),$$

then $\mathbb{E}[P_e^*(C_n)] \leq \delta_\varepsilon(n)$.

Achievability proof (V)- Weak secrecy

$$\mathbf{L}(C_n) = \mathbb{I}(\mathsf{M}; \mathsf{Z}^n|C_n) = \mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n|C_n) - \mathbb{I}(\mathsf{M}'; \mathsf{Z}^n|\mathsf{M}, C_n) =$$
$$= \mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n|C_n) - \mathbb{H}(\mathsf{M}'|\mathsf{M}, C_n) + \mathbb{H}(\mathsf{M}'|\mathsf{Z}^n, \mathsf{M}, C_n) =$$
$$= \underbrace{\mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n|C_n)}_{(a)} - \underbrace{\mathbb{H}(\mathsf{M}')}_{(b)} + \underbrace{\mathbb{H}(\mathsf{M}'|\mathsf{Z}^n, \mathsf{M}, C_n)}_{(c)}$$

(a) $C_n \to \mathsf{X}^n \to \mathsf{Z}^n$ Markov chain

$$\Rightarrow \ \mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n|C_n) \leq \mathbb{I}(\mathsf{X}^n; \mathsf{Z}^n) = n\mathbb{I}(\mathsf{X}; \mathsf{Z})$$

(b) $\mathbb{H}(\mathsf{M}') = nR'$

(c) From Fano's inequality,

$$\mathbb{H}(\mathsf{M}'|\mathsf{Z}^n, \mathsf{M}) = \mathbb{H}(\mathsf{M}'|\tilde{\mathsf{M}}') \leq H_b(P_e^*) + P_e^* \log(2^{nR'} - 1)$$

So if $P_e^* \to 0$, we have $\frac{1}{n}\mathbb{H}(\mathsf{M}'|\mathsf{Z}^n, \mathsf{M}) \to 0$.

Problem with the random coding argument:

In the case of channel coding, the fact that $P_e \leq \varepsilon$ automatically implies that there is at least one code C_n such that $P_e(C_n) \leq \varepsilon$. In the case of equivocation, since conditioning might decrease entropy,

$$\frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{Z}^n) \geq \frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{Z}^n, \mathsf{C}_n) = \sum_{C_n} p_{\mathsf{C}_n}(C_n)\frac{1}{n}\mathbf{E}(C_n)$$

So we need to study $\frac{1}{n}\mathbb{H}(\mathsf{M}|\mathsf{Z}^n, \mathsf{C}_n)$ directly.

Achievability proof (VI)- Weak secrecy

- We have shown that $\frac{1}{n}\mathbf{L}(C_n) \leq \mathbb{I}(\mathsf{X};\mathsf{Z}) - R' + \delta_\varepsilon(n)$.
  So if $R' = \mathbb{I}(\mathsf{X};\mathsf{Z}) - \delta(\varepsilon)$,

$$\mathbb{E}\left[\frac{1}{n}\mathbf{L}(C_n)\right] \leq \delta(\varepsilon) + \delta_\varepsilon(n).$$

Lemma (Selection Lemma)

Let A_n be a random variable taking values in $\mathcal{A}_n$, and $\forall i \in [\![1, r]\!]$, let $f_i : \mathcal{A}_n \to \mathbb{R}^+$ be functions such that

$$\forall i \in [\![1, r]\!], \quad \mathbb{E}[f_i(\mathsf{A}_n)] \leq \delta(n).$$

Then there exists a realization A_n of A_n such that

$$\forall i \in [\![1, r]\!], \quad f_i(A_n) \leq \delta(n).$$

- Consequently, there exists a wiretap code in the ensemble which has both the reliability and the weak secrecy property.

Achievability proof (VII)

- We have shown that the rate-equivocation region

$$\mathcal{R}'(p_{\mathsf{X}}) = \{(R, R_e) \ : \ 0 \leq R < \mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}), \ 0 \leq R_e < R\}$$

  is achievable.

Achievability proof (VII)

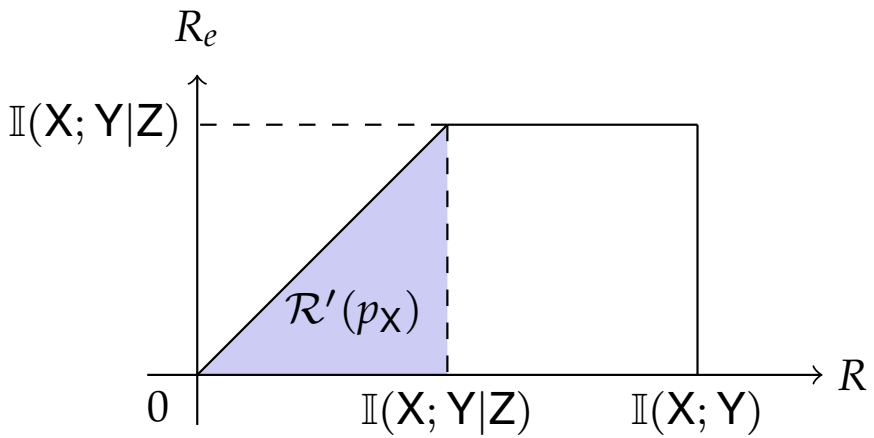
- We have shown that the rate-equivocation region

$$\mathcal{R}'(p_{\mathsf{X}}) = \{(R, R_e) \ : \ 0 \leq R < \mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}),\ 0 \leq R_e < R\}$$

  is achievable.

- For a degraded wiretap channel,

$$\mathcal{R}'(p_{\mathsf{X}}) = \{(R, R_e) \ : \ 0 \leq R < \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}),\ 0 \leq R_e < R\}.$$



Achievability proof (VII)

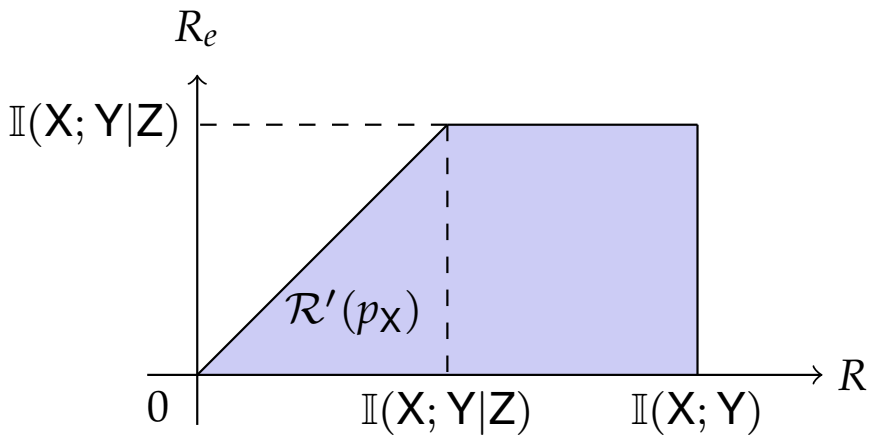
- We have shown that the rate-equivocation region

$$\mathcal{R}'(p_{\mathsf{X}}) = \{(R, R_e) \ : \ 0 \leq R < \mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}),\ 0 \leq R_e < R\}$$

  is achievable.

- For a degraded wiretap channel,

$$\mathcal{R}'(p_{\mathsf{X}}) = \{(R, R_e) \ : \ 0 \leq R < \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}),\ 0 \leq R_e < R\}.$$



Problem: How to achieve the entire region $\mathcal{R}$?

Achievability proof (VII)

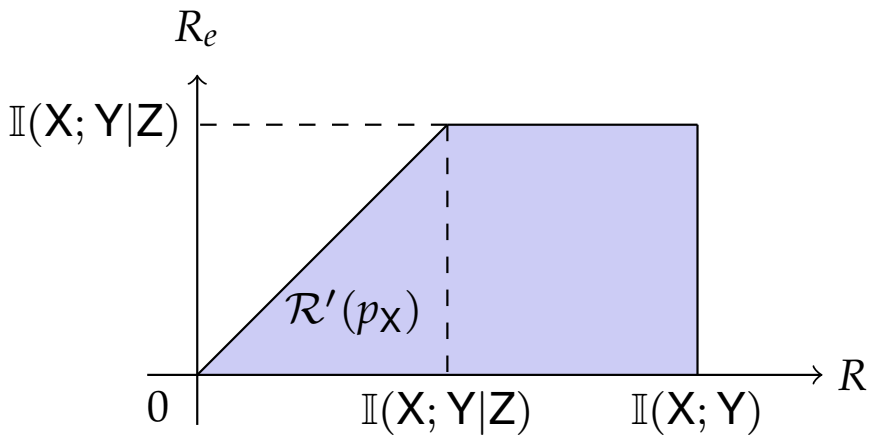
- We have shown that the rate-equivocation region

$$\mathcal{R}'(p_{\mathsf{X}}) = \{(R, R_e) \ : \ 0 \leq R < \mathbb{I}(\mathsf{X};\mathsf{Y}) - \mathbb{I}(\mathsf{X};\mathsf{Z}),\ 0 \leq R_e < R\}$$

  is achievable.

- For a degraded wiretap channel,

$$\mathcal{R}'(p_{\mathsf{X}}) = \{(R, R_e) \ : \ 0 \leq R < \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{Z}),\ 0 \leq R_e < R\}.$$



Problem: How to achieve the entire region $\mathcal{R}$?

- Idea: use the auxiliary message M' to transmit additional common information "for free".

How to achieve the entire rate-equivocation region:

We have seen that by taking $R' = \mathbb{I}(\mathsf{X};\mathsf{Z}) - \delta(\varepsilon)$, we can build an $(2^{nR}, n)$ code C_n such that $\frac{1}{n}\mathbf{L}(C_n) \leq \delta(\varepsilon) + \delta_\varepsilon(n)$.
Divide each bin of size $2^{nR'}$ in sub-bins of size 2^{nR_0} and use each sub-bin to transmit a (non secure) message M_0. The new code $\tilde{C}_n$ has rate $R + R_0$. One can show that even though the sub-bins do not have exactly the same size, the distribution of the codewords in $\tilde{C}_n$ is not much different from the distribution of the codewords in C_n.

- Reliability holds as long as $R + R_0 \leq \mathbb{I}(\mathsf{X};\mathsf{Y})$.
- As for the equivocation:

$$\mathbf{E}(\tilde{C}_n) = \mathbb{H}(\mathsf{M},\mathsf{M}_0|\mathsf{Z}^n,\tilde{C}_n) \geq \mathbb{H}(\mathsf{M}|\mathsf{Z}^n,\tilde{C}_n) = \mathbb{H}(\mathsf{M}|\tilde{C}_n) - \mathbb{I}(\mathsf{M};\mathsf{Z}^n|\tilde{C}_n)$$
$$\mathbb{H}(\mathsf{M}|\tilde{C}_n) = \mathbb{H}(\mathsf{M}|C_n) = nR$$
$$\mathbb{I}(\mathsf{M};\mathsf{Z}^n|\tilde{C}_n) \leq \mathbb{I}(\mathsf{M};\mathsf{Z}^n|C_n) + \delta(n)$$

since $\mathbb{I}(\mathsf{M};\mathsf{Z}^n|\tilde{C}_n)$ is a continous function of p_{X^n}. So

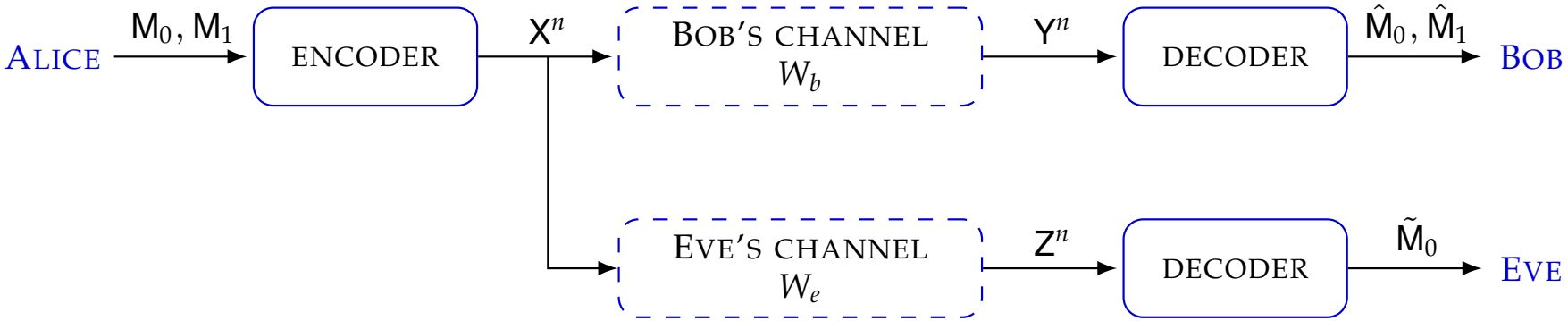
$$\frac{1}{n}\left(\mathbb{H}(\mathsf{M}|\tilde{C}_n) - \mathbb{I}(\mathsf{M};\mathsf{Z}^n|\tilde{C}_n)\right) \geq R - \delta(\varepsilon) + \frac{\mathbf{L}(C_n)}{n} - \frac{\delta(n)}{n}$$

Remark: this works only because the messages M are uniformly distributed!

- One can show that the rate-equivocation region $\mathcal{R}$ is convex.
- We still need the proof of converse: given an achievable rate-equivocation pair (R, R_e), show that it belongs to $\mathcal{R}$. We will skip this proof for lack of time.
- Now let's consider a more general scenario in which Alice also wants to reliably send some message to Eve, while keeping some other message secret: the broadcast channel with confidential messages.
- It is more general also because we don't assume that Eve's channel is degraded.

Broadcast channel with confidential messages

I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages", *IEEE Trans. Inform. Theory*, vol. 24, n. 3, 1978.



- Alice sends a common message M_0 intended for both Bob and Eve, and a confidential message M_1 for Bob treating Eve as an eavesdropper.

- $P_e(C_n) = \mathbb{P}\{(\hat{M}_0, \hat{M}_1) \neq (M_0, M_1) \text{ or } \tilde{M}_0 \neq M_0\}$ error probability
- $\mathbf{E}(C_n) = \mathbb{H}(M_1|Z^n)$ equivocation
- $\mathcal{R} = \overline{\{(R_0, R_1, R_e) \text{ achievable rate}\}}$ rate-equivocation region
- $C_s = \overline{\{(R_0, R_1) \mid (R_0, R_1, R_1) \in \mathcal{R}\}}$ secrecy capacity region

Broadcast channel with confidential messages (II)

Theorem (Csiszár and Körner)

The rate-equivocation region of the broadcast channel with confidential messages is given by

$$\mathcal{R} = \bigcup_{p_{\mathsf{UVX}} \in \mathcal{P}} \mathcal{R}(p_{\mathsf{UVX}}),$$

where $\mathsf{U} \to \mathsf{V} \to \mathsf{X}$ *is a Markov chain, and*

$$\mathcal{R}(p_{\mathsf{UVX}}) = \left\{ (R_0, R_1, R_e) : \begin{array}{l} 0 \leq R_e \leq R_1 \\ 0 \leq R_e \leq \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}) \\ 0 \leq R_0 \leq \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \\ 0 \leq R_1 + R_0 \leq \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \end{array} \right\}$$

Corollary

$$C_s = \bigcup_{p_{\mathsf{UVX}} \in \mathcal{P}} \left\{ (R_0, R_1) : \begin{array}{l} 0 \leq R_0 \leq \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \\ 0 \leq R_1 \leq \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}) \end{array} \right\}$$

Meaning of auxiliary random variables:

- Roughly, U represents the common message.
- V represents the codeword used by the encoder to transmit the confidential message.

Superposition coding technique for broadcast channels:

- Generate a random codebook $\{u^n(m_0)\}$ of size 2^{nR_0} with i.i.d. entries distributed according to p_{U}.
- Generate a second codebook $\{x^n(m_0, m_1, m')\}$ of size $2^{n(R_1+R')}$ where each codeword $x^n(m_0, m_1, m')$ has i.i.d. entries distributed according to $p_{\mathsf{X}|\mathsf{U}=u_i(m_0)}$.

1. $0 \leq R_e \leq R_1$
   Obviously, the secure rate cannot be higher than the rate of the confidential message.
2. $0 \leq R_e \leq \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U})$
   This comes from the wiretap channel, assuming that both Bob and Eve know the common message.
3. $0 \leq R_0 \leq \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z}))$
   This comes from the channel coding theorem.
4. $0 \leq R_1 + R_0 \leq \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) + \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z}))$
   This comes from the achievable region for the broadcast channel (without secrecy constraints): suppose I want to send M_0 and M_1 to the first receiver, and just M_0 to the second receiver. Then the achievable rate region is the convex hull

$$\mathcal{R} = \text{conv}\left(\bigcup_{p_{\mathsf{UX}}} \mathcal{R}(p_{\mathsf{UX}})\right), \quad \text{where}$$

$$\mathcal{R}(p_{\mathsf{UX}}) = \left\{ (R_0, R_1) \mid \begin{array}{l} 0 \leq R_0 \leq \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \\ 0 \leq R_1 \leq \mathbb{I}(\mathsf{X};\mathsf{Y}|\mathsf{U}) \end{array} \right\}$$

(*Reference:* P. Bergmans, *IEEE Trans. Inform. Theory* vol 19 n.2, 1973)

Channel comparison

Corollary (Csiszár and Körner)

The secrecy capacity of the general (non-degraded) wiretap channel is

$$C_s = \max_{p_{VX}} \left(\mathbb{I}(V;Y) - \mathbb{I}(V;Z) \right)$$

Lemma (Liang *et al.*)

The secrecy capacity of a wiretap channel depends on the transition probability $p_{YZ|X}$ only through the marginal transition probabilities $p_{Y|X}$ and $p_{Z|X}$.

So there is no difference between physically degraded channels and stochastically degraded channels from the point of view of capacity.

Proposition

Consider two eavesdroppers channels W_e, $\tilde{W}_e$ such that W_e is stochastically degraded with respect to $\tilde{W}_e$. Then a wiretap code ensuring equivocation R_e for $\tilde{W}_e$ guarantees the same equivocation for W_e.

Conclusion

- we have proved the existence of wiretap codes that achieve reliability and security at the same time, and characterized the achievable secrecy capacity

- **Problem:** Wyner's model deals with discrete channels. Can we extend this result to Gaussian and fading channels? (Part III)

- **Problem:** how to design explicit codes for a given wiretap channel? (Part IV)

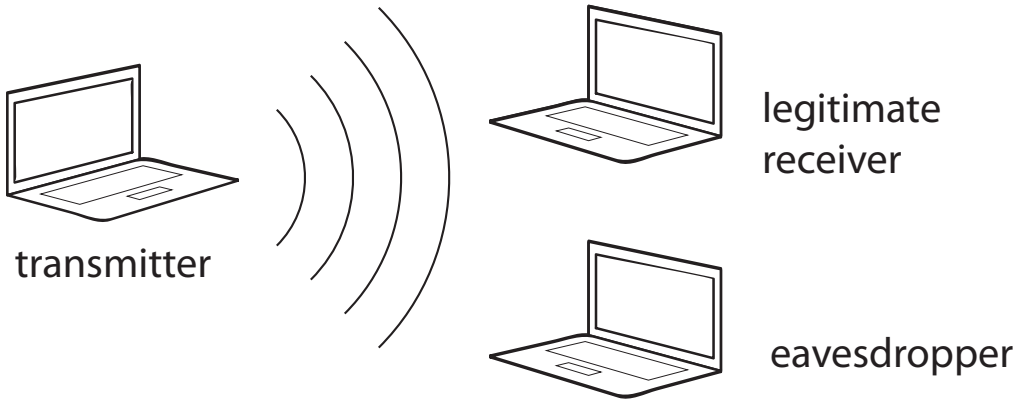
References for Part II

- I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages", *IEEE Trans. Inform. Theory*, vol. 24, n. 3, 1978.
- S. K. Leung-Yan-Cheong, "On a special class of wiretap channels", *IEEE Trans. Inform. Theory*, vol. 23, n. 5, 1977.
- A. D. Wyner, "The Wire-Tap Channel", *Bell System Technical Journal*, vol. 54, n. 8, 1975.

Part III

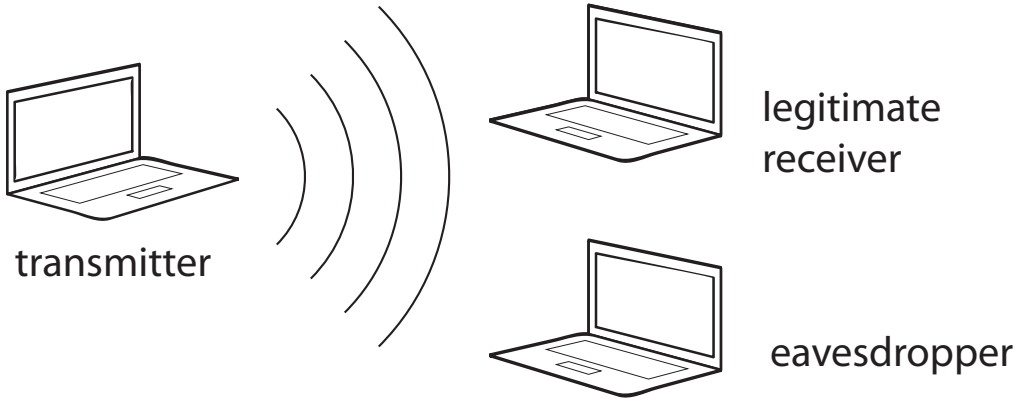
Secure Communication over Fading Channels

Introduction



- more and more often, users rely on wireless devices to transmit sensitive data, such as banking information, administrative documents or medical records
- wireless networks are particularly vulnerable to attacks, since every node in the network is a potential eavesdropper
- however, wireless channels are also a source of randomness that can be harnessed to provide security

Introduction



Classical cryptography techniques might not be the best choice for heterogeneous and decentralized wireless networks:

- symmetric algorithms rely on secure distribution of secret keys though a noiseless link, which is possible only for highly centralized systems
- asymmetric algorithms are too complex for mobile phones with limited battery and computational power

The Gaussian broadcast channel with confidential messages

- System model:

$$\begin{cases} Y_i = X_i + N_{b,i} & N_{b,i} \sim \mathcal{N}(0, \sigma_b^2), \\ Z_i = X_i + N_{e,i} & N_{e,i} \sim \mathcal{N}(0, \sigma_e^2) \end{cases}$$

- Power constraint: $\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i^2] \leq P$.
- In the case of the Gaussian BCC, either Eve's channel is stochastically degraded with respect to Bob's channel or Bob's channel is stochastically degraded with respect to Eve's channel.
- For example, if $\sigma_e^2 \geq \sigma_b^2$, the Gaussian BCC is equivalent to

$$\begin{cases} Y_i = X_i + N_{b,i} & N_{b,i} \sim \mathcal{N}(0, \sigma_b^2), \\ Z_i = Y_i + N'_{e,i} & N'_{e,i} \sim \mathcal{N}(0, \sigma_e^2 - \sigma_b^2) \end{cases}$$

The Gaussian broadcast channel with confidential messages

Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure Communication Over Fading Channels", *IEEE Trans. Inform. Theory*, vol. 54, n. 6, 2008.

Theorem (Liang, Poor and Shamai)

$$C_s = \bigcup_{\beta \in [0,1]} \left\{ \begin{array}{l} R_0 \leq \min \left(\frac{1}{2} \log \left(1 + \frac{(1-\beta)P}{\sigma_b^2 + \beta P} \right), \frac{1}{2} \log \left(1 + \frac{(1-\beta)P}{\sigma_e^2 + \beta P} \right) \right) \\ R_1 \leq \left(\frac{1}{2} \log \left(1 + \frac{\beta P}{\sigma_b^2} \right) - \frac{1}{2} \log \left(1 + \frac{\beta P}{\sigma_e^2} \right) \right)^+ \end{array} \right\}$$

Sketch of the achievability proof.

Achievability follows from the general theorem for the BCC: let $\mathsf{U} \sim \mathcal{N}(0, (1-\beta)P)$ common message, $\mathsf{X}' \sim \mathcal{N}(0, \beta P)$ confidential message, $\mathsf{V} = \mathsf{X} = \mathsf{U} + \mathsf{X}'$. □

Actually we have to pay attention when passing from the discrete to the continuous case:

- Letter-typical sequences have to be replaced by entropy-typical sequences (see Cover and Thomas)

$$A_\varepsilon^n = \left\{ (x_1, \ldots, x_n) \ : \ \left| -\frac{1}{n} \log_2 p_{\mathsf{X}^n}(x_1, \ldots, x_n) - \mathbb{H}(\mathsf{X}) \right| \leq \varepsilon \right\}$$

  A continuous version of the AEP still holds.

- In the random coding argument, we need to keep into account the power constraint, and declare an error event whenever it is not satisfied. But this event is negligible.

Remark: the assumption that Alice knows Eve's channel is realistic only for the BCC, not for the wiretap channel.

To obtain the special case of the Gaussian wiretap channel, just take $\beta = 1$ (no common message).

Achievability proof- continued:

The system becomes

$$\begin{cases} \mathsf{Y} = \mathsf{U} + \mathsf{X}' + \mathsf{N}_b & \mathsf{X}' + \mathsf{N}_b \sim \mathcal{N}_{\mathbb{C}}(0, \beta P + \sigma_b^2), \\ \mathsf{Z} = \mathsf{U} + \mathsf{X}' + \mathsf{N}_e & \mathsf{X}' + \mathsf{N}_e \sim \mathcal{N}_{\mathbb{C}}(0, \beta P + \sigma_e^2) \end{cases}$$

Replacing in the expression

$$C = \bigcup_{p_{\mathsf{UVX}} \in \mathcal{P}} \left\{ (R_0, R_1) : \begin{array}{l} 0 \le R_0 \le \min(\mathbb{I}(\mathsf{U};\mathsf{Y}), \mathbb{I}(\mathsf{U};\mathsf{Z})) \\ 0 \le R_1 \le \mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U}) \end{array} \right\}$$

we find:

$$\mathbb{I}(\mathsf{U};\mathsf{Y}) = \frac{1}{2}\log\left(1 + \frac{(1-\beta)P}{\beta P + \sigma_b^2}\right), \quad \mathbb{I}(\mathsf{U};\mathsf{Z}) = \frac{1}{2}\log\left(1 + \frac{(1-\beta)P}{\beta P + \sigma_e^2}\right)$$

$$\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) = \mathbb{I}(\mathsf{U},\mathsf{V};\mathsf{Y}) - \mathbb{I}(\mathsf{U};\mathsf{Y}) = \mathbb{I}(\mathsf{V};\mathsf{Y}) + \underbrace{\mathbb{I}(\mathsf{U},\mathsf{Y}|\mathsf{V})}_{=0} - \mathbb{I}(\mathsf{U};\mathsf{Y}) =$$

$$= \frac{1}{2}\log\left(1 + \frac{P}{\sigma_b^2}\right) - \frac{1}{2}\log\left(1 + \frac{(1-\beta)P}{\beta P + \sigma_b^2}\right) = \frac{1}{2}\log\left(1 + \frac{\beta P}{\sigma_b^2}\right)$$

The Gaussian wiretap channel

S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian Wire-Tap Channel", *IEEE Trans. Inform. Theory*, vol. 24, n. 4, 1978.

Corollary (Leung-Yan-Cheong, Hellman)

$$C_s = \left(\frac{1}{2}\log\left(1+\frac{P}{\sigma_b^2}\right) - \frac{1}{2}\log\left(1+\frac{P}{\sigma_e^2}\right)\right)^+ = (C_b - C_e)^+$$

- unlike the capacity, the secrecy capacity is not unbounded when the power $P \to \infty$:

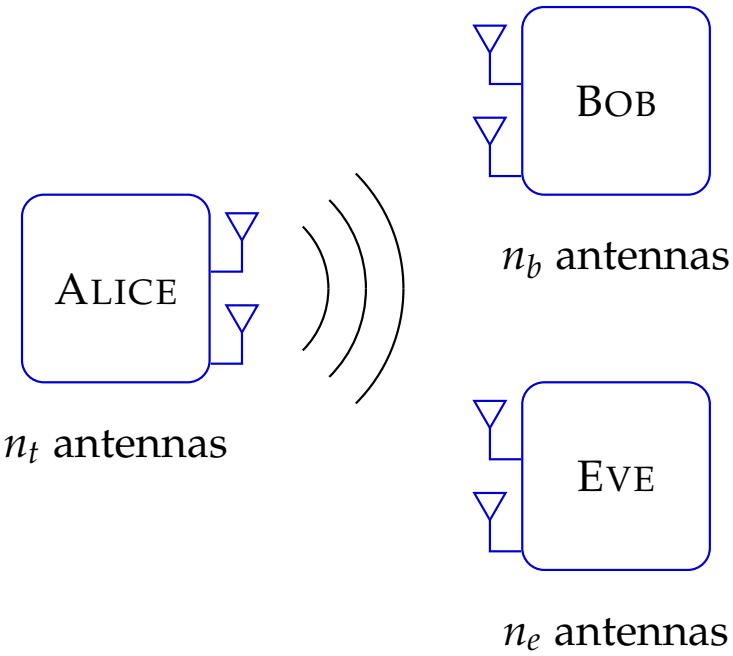
$$\lim_{P\to\infty} C_s = \left(\frac{1}{2}\log\frac{\sigma_e^2}{\sigma_b^2}\right)^+$$

- secure communication is possible only if Bob has a better SNR than Eve.

- **Question:** does the situation improve if we use multiple antennas?

MIMO Gaussian wiretap channel

A. Khisti, G. Wornell,"Secure Transmission with Multiple Antennas-II: The MIMOME Wiretap Channel", *IEEE Trans. Inform. Theory*, vol 56 n.11, 2010



$$\begin{cases} \mathsf{Y}_i^{n_b} = \mathbf{H}_b \mathsf{X}_i^{n_t} + \mathsf{N}_{b,i}^{n_b}, \\ \mathsf{Z}_i^{n_e} = \mathbf{H}_e \mathsf{X}_i^{n_t} + \mathsf{N}_{e,i}^{n_e} \end{cases}$$

- unlike the scalar case, in general Eve's channel is not degraded with respect to Bob's channel

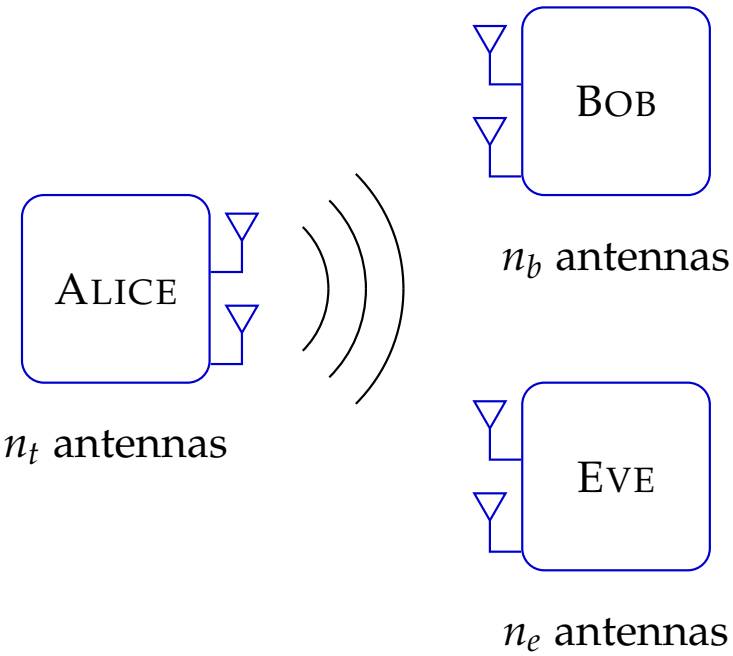
Theorem (Khisti and Wornell, Oggier and Hassibi, Liu and Shamai)

$$C_s = \max_{\mathbf{Q}_\mathsf{X}} \left(\log \left| \mathbf{I}_{n_r} + \tfrac{1}{\sigma_b^2} \mathbf{H}_b \mathbf{Q}_\mathsf{X} \mathbf{H}_b^H \right| - \log \left| \mathbf{I}_{n_e} + \tfrac{1}{\sigma_e^2} \mathbf{H}_e \mathbf{Q}_\mathsf{X} \mathbf{H}_e^H \right| \right),$$

over all the covariance matrices $\mathbf{Q}_\mathsf{X}$ *which satisfy the power constraint* $\mathrm{Tr}(\mathbf{Q}_\mathsf{X}) \leq P$.

MIMO Gaussian wiretap channel

A. Khisti, G. Wornell,"Secure Transmission with Multiple Antennas-II: The MIMOME Wiretap Channel", *IEEE Trans. Inform. Theory*, vol 56 n.11, 2010



$$\begin{cases} \mathsf{Y}_i^{n_b} = \mathbf{H}_b \mathsf{X}_i^{n_t} + \mathsf{N}_{b,i}^{n_b}, \\ \mathsf{Z}_i^{n_e} = \mathbf{H}_e \mathsf{X}_i^{n_t} + \mathsf{N}_{e,i}^{n_e} \end{cases}$$

- unlike the scalar case, in general Eve's channel is not degraded with respect to Bob's channel

Proposition (Khisti and Wornell)

$$C_s = 0 \quad \Leftrightarrow \quad \sup_{\mathbf{v}} \frac{\sigma_e \left\| \mathbf{H}_b \mathbf{v} \right\|}{\sigma_b \left\| \mathbf{H}_e \mathbf{v} \right\|} \leq 1.$$

Sketch of the proof of the Proposition.

- If $\mathrm{Ker}(\mathbf{H}_b) \nsubseteq \mathrm{Ker}(\mathbf{H}_e)$, there exists $\mathbf{v} \in \mathrm{Ker}(\mathbf{H}_e)$ such that $\|\mathbf{H}_b\mathbf{v}\| > 0$. In this case, Alice can communicate securely by communicating her signal in the direction $\mathbf{v}$, and no wiretap coding is needed. (Beamforming alone is enough).
- If $\mathrm{Ker}(\mathbf{H}_b) \subseteq \mathrm{Ker}(\mathbf{H}_e)$, consider $\lambda = \frac{\sigma_e \|\mathbf{H}_b\mathbf{v}\|}{\sigma_b \|\mathbf{H}_e\mathbf{v}\|}$. If $\lambda > 1$, there exists $\mathbf{v}$ such that $\frac{\|\mathbf{H}_b\mathbf{v}\|}{\sigma_b} > \frac{\|\mathbf{H}_e\mathbf{v}\|}{\sigma_e}$, so there is one direction in which Bob has a better SNR and we are in the situation of the SISO Gaussian channel. (Wiretap coding and beamforming is more powerful than beamforming alone). □

Wireless wiretap channel with fading

- System model:

$$\begin{cases} Y_i = H_{b,i} X_i + N_{b,i} & N_{b,i} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_b^2), \\ Z_i = H_{e,i} X_i + N_{e,i} & N_{e,i} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_e^2) \end{cases}$$

$$G_{b,i} = |H_{b,i}|^2, \quad G_{e,i} = |H_{e,i}|^2 \quad \text{fading gains}$$

- Power constraint: $\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i^2] \leq P$.
- We suppose that Alice, Bob and Eve have full knowledge of *all the fading channels*
- Three fading models according to the coherence time: the channel remains constant during the transmission of
  - one symbol $\Rightarrow$ ergodic fading
  - several symbols $\Rightarrow$ block fading
  - a whole codeword $\Rightarrow$ quasi-static fading

Ergodic case

Y. Liang, H. V. Poor, and S. Shamai, "Secure Communication Over Fading Channels", *IEEE Trans. Inform. Theory*, vol. 54, n. 6, 2008.

- **Idea:** decompose the fading wiretap channel into k^2 time-invariant Gaussian wiretap channels by partitioning the range of G_b and G_e into k intervals $\{I_b^{(l)}\}_{l=1,\ldots,k}$, $\{I_e^{(m)}\}_{m=1,\ldots,k}$ respectively.
- $\forall (l,m)$, Alice chooses a power allocation $\gamma_{l,m}$ such that the average satisfies the power constraint: $\mathbb{E}[\gamma(\mathsf{G}_b, \mathsf{G}_e)] \leq P$.
  Characterize the optimal γ in the limit for $k \to \infty$.
- If Eve has a better instantaneous SNR, Alice allocates zero power.
- Achievability follows from the corresponding result for the Gaussian wiretap channel.

Theorem (Liang *et al.*)

$$C_s = \max_{\gamma} \mathbb{E}_{\mathsf{G}_b,\mathsf{G}_e}\left[\log\left(1+\frac{\gamma(\mathsf{G}_b,\mathsf{G}_e)\mathsf{G}_b}{\sigma_b^2}\right) - \log\left(1+\frac{\gamma(\mathsf{G}_b,\mathsf{G}_e)\mathsf{G}_e}{\sigma_e^2}\right)\right]$$

Ergodic case

P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels", *IEEE Trans. Inform. Theory*, vol. 54, n. 10, 2008.

- Fading is beneficial for secrecy: C_s is always strictly positive provided that $\mathbb{P}\left\{\frac{G_b}{\sigma_b^2} > \frac{G_e}{\sigma_e^2}\right\} > 0$. In the fading case, Alice can decide to transmit only when Bob has a better channel than Eve.

- However, the assumption that Alice knows Eve's channel is completely unrealistic. Without this assumption, $C_s = 0$ in the ergodic fading case.

- However, for some block-fading models, $C_s > 0$ even without CSI concerning Eve's channel [Gopala, Lai and El Gamal 2008].

Quasi-static fading case

J. Barros and M. R. D. Rodrigues, "Secrecy Capacity of Wireless Channels", *Proc. ISIT 2006*

- With full channel state information, the secrecy capacity is the average of the instantaneous capacity: $C_s^{\text{av}} = \mathbb{E}_{\mathsf{G}_b,\mathsf{G}_e}[C_s(\mathsf{G}_b, \mathsf{G}_e)]$
- Without knowledge of Eve's channel at the transmitter, $C_s^{\text{av}} = 0$: during the transmission of one codeword, the probability that Eve has a better channel than Bob is always strictly positive.
- In this case, the outage probability of the secrecy capacity can be defined instead.
- If Alice has no knowledge of Bob's channel either, Bob's outage probability must also be taken into account.

Conclusions

- if the transmitter has channel state information, SNR fluctuations are beneficial for secrecy
- positive secure communication rates are achievable even if Eve has a better average SNR
- for some fading models, this is true even without channel state information at the transmitter

References for Part III

- J. Barros and M. R. D. Rodrigues, "Secrecy Capacity of Wireless Channels", *Proc. ISIT 2006*
- P. K. Gopala, L. Lai, and H. El Gamal, "On the Secrecy Capacity of Fading Channels", *IEEE Trans. Inform. Theory*, vol. 54, n. 10, 2008.
- A. Khisti, G. Wornell,"Secure Transmission with Multiple Antennas-II: The MIMOME Wiretap Channel", *IEEE Trans. Inform. Theory*, vol 56 n.11, 2010
- S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian Wire-Tap Channel", *IEEE Trans. Inform. Theory*, vol. 24, n. 4, 1978.
- Y. Liang, H. V. Poor, and S. Shamai, "Secure Communication Over Fading Channels", *IEEE Trans. Inform. Theory*, vol. 54, n. 6, 2008.
- T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel", *IEEE Trans. Inform. Theory*, vol. 55 n.6, 2009
- F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channels", *Proc. ISIT* 2008

Part IV

Coding for Secrecy

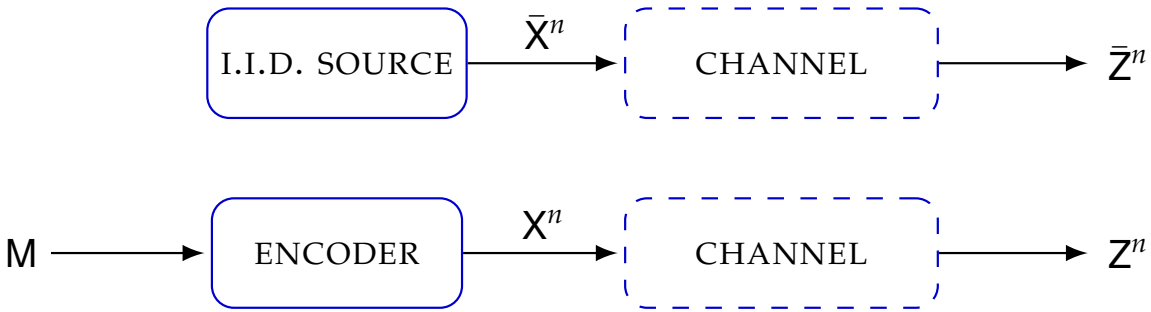
Introduction

- the secrecy capacity of wiretap channels has been determined in many scenarios using non-constructive random coding arguments
- however, the design of practical codes remains mostly an open problem even for very simple channels
- two different approaches to achieve secrecy are compared: capacity and resolvability
- resolvability is a more promising tool to obtain strong secrecy, while capacity-based constructions fail to achieve it

Outline

Coding for resolvability

T. Han and S. Verdù, "Approximation Theory of Output Statistics", *IEEE Trans. Inform. Theory*, vol. 39, n. 3, 1993



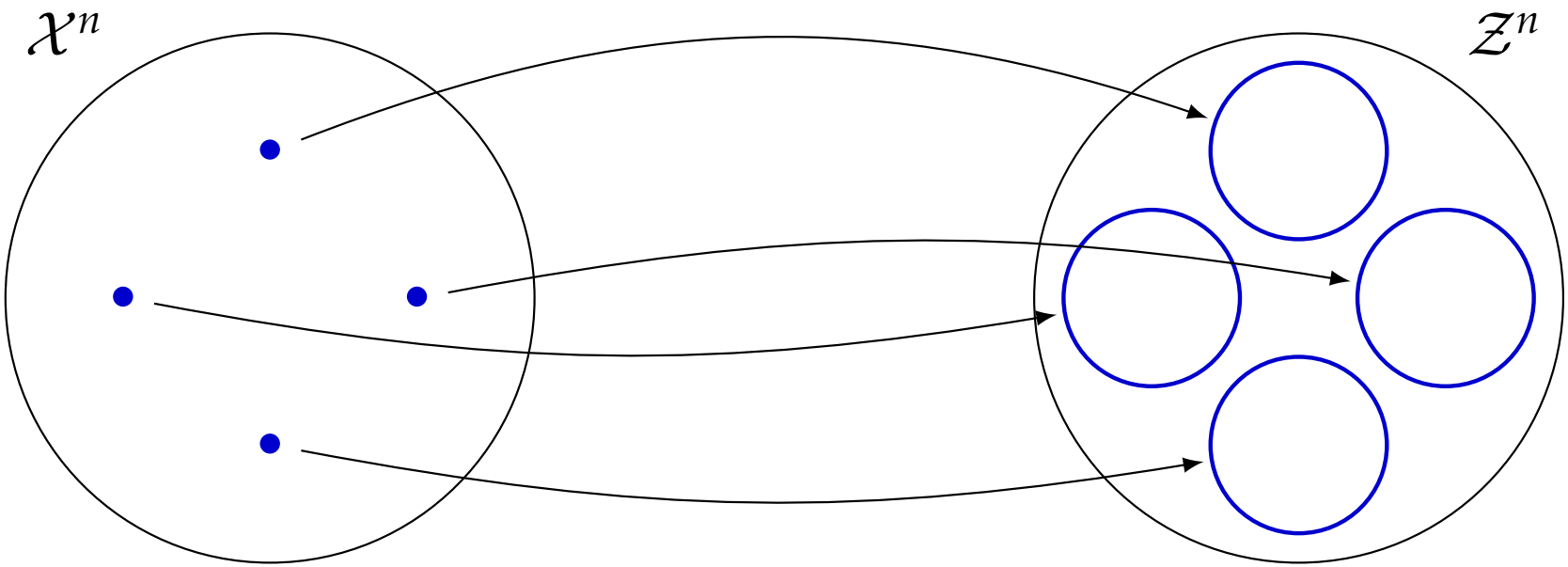
Resolvability codes

Given $p_{\bar{X}}$ input probability distribution, construct a sequence of codes such that the output Z^n of the code approaches the output $\bar{Z}^n$ of an i.i.d. source with distribution $p_{\bar{X}}$ (in variational distance):

$$\lim_{n\to\infty} \mathbb{V}(p_{Z^n}, p_{\bar{Z}^n}) = 0$$

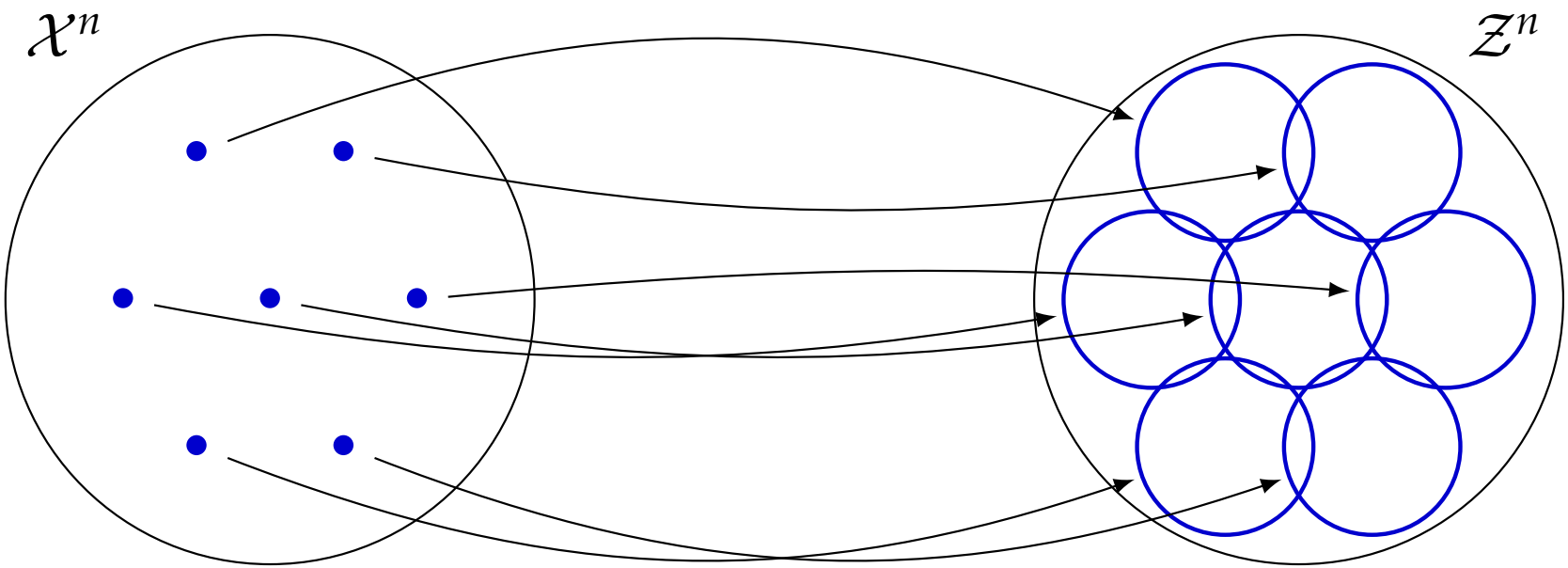
- the channel resolvability is the minimum rate such that resolvability codes exist for any distribution $p_{\bar{X}}$
- in the case of discrete memoryless channels, resolvability = capacity

Coding for reliability



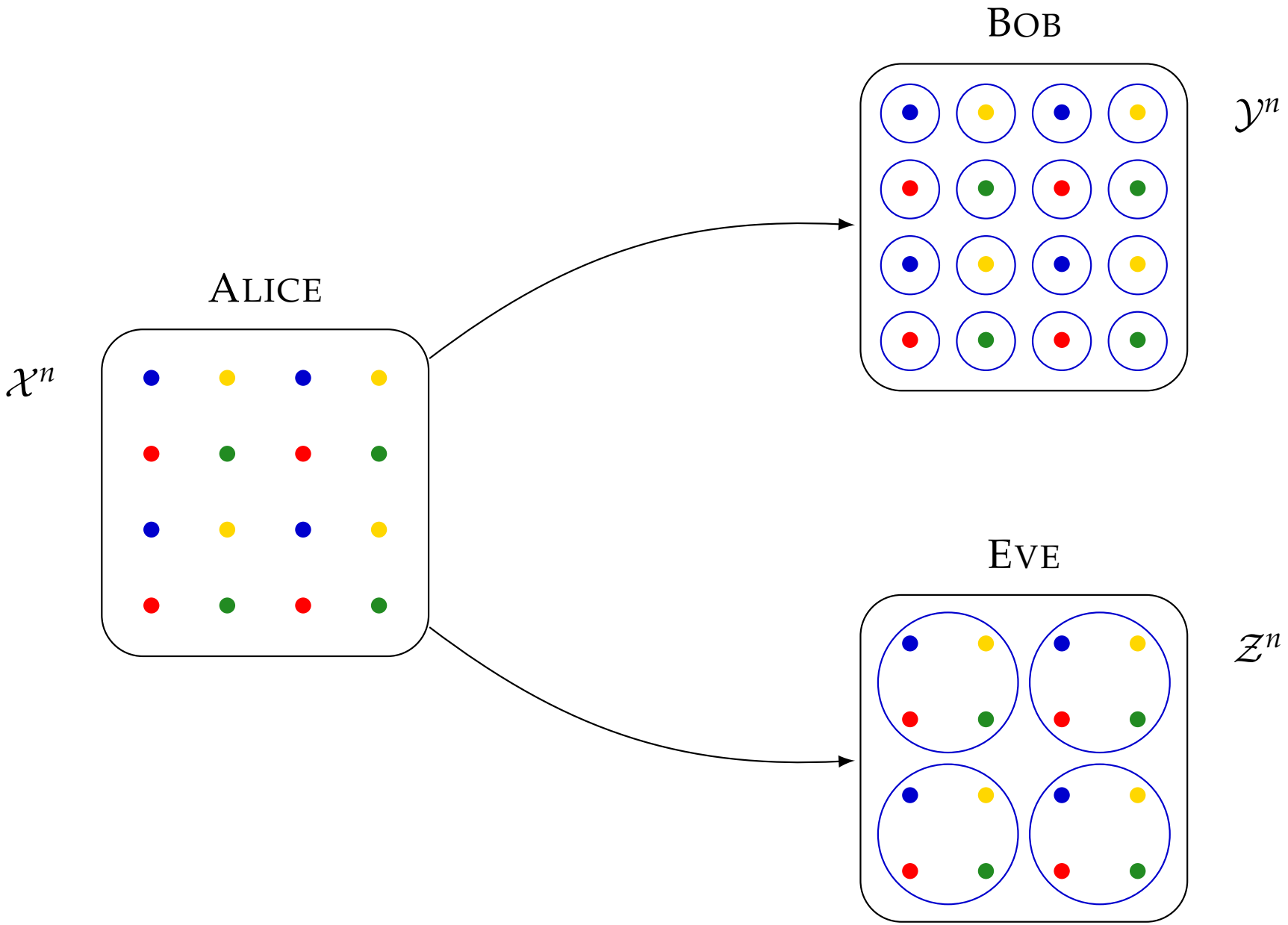
$\Rightarrow$ rate < channel capacity

Coding for resolvability

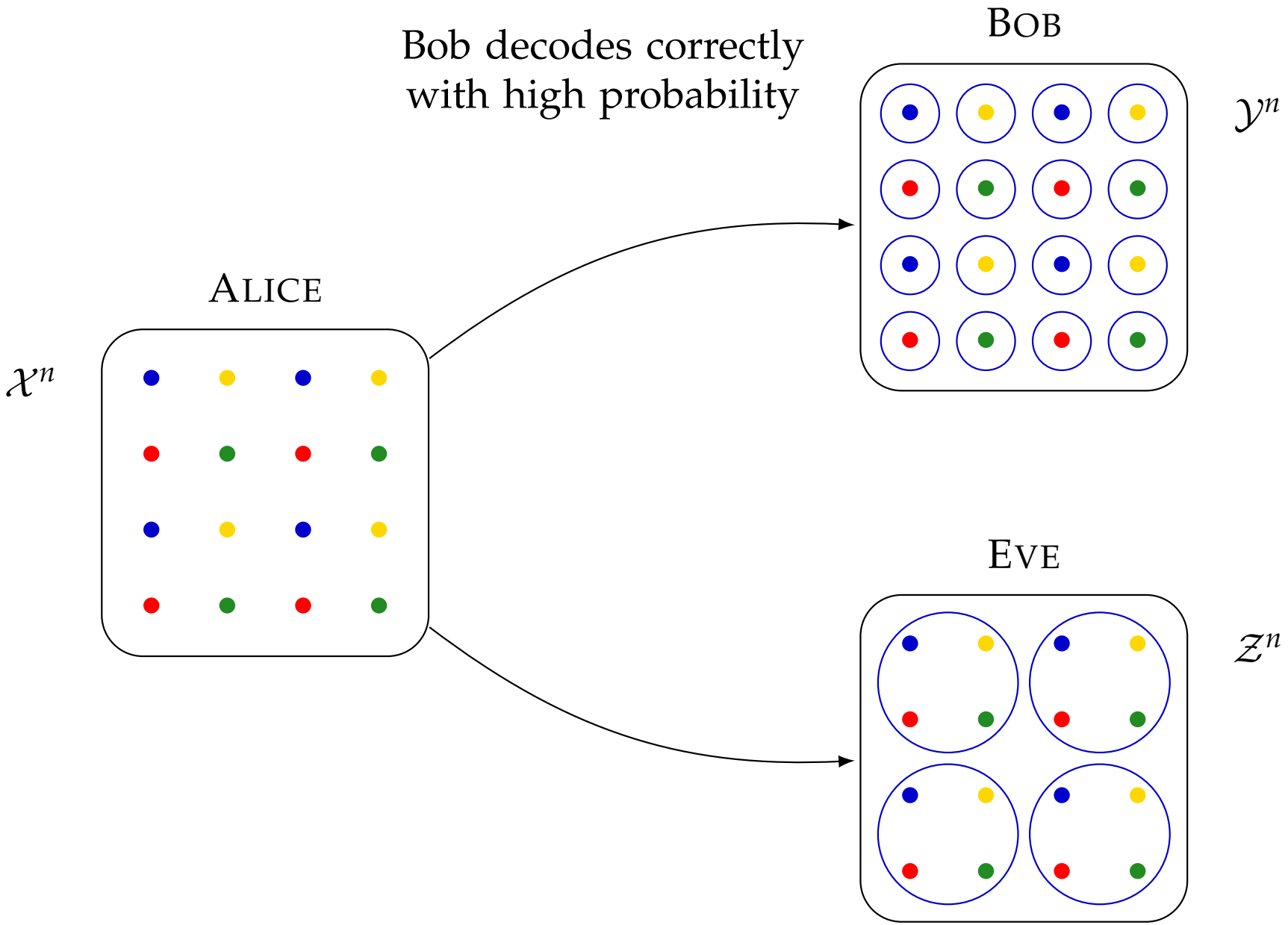


rate > channel capacity

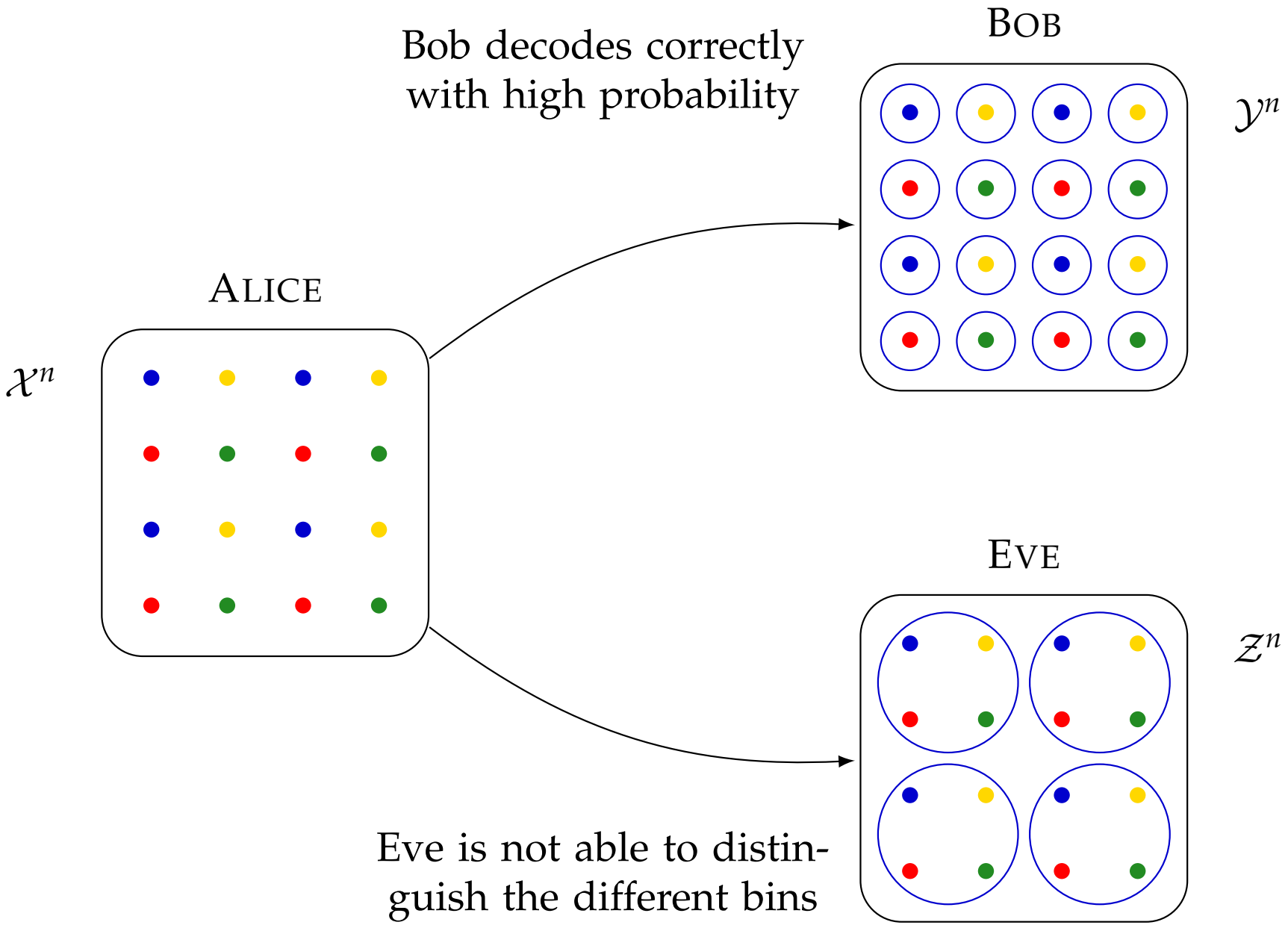
Random binning for secrecy



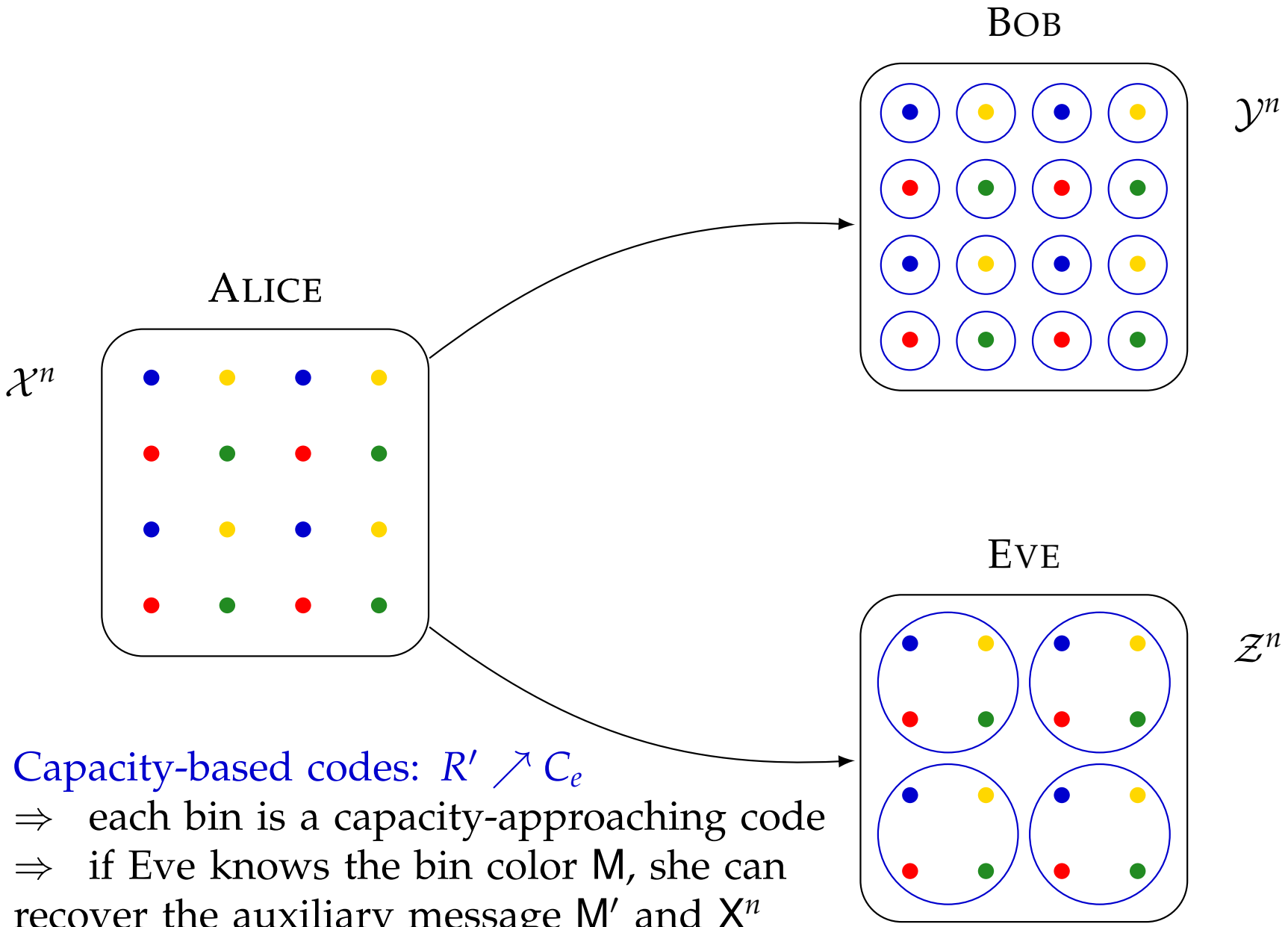
Random binning for secrecy



Random binning for secrecy



Capacity-based codes vs. resolvability-based codes

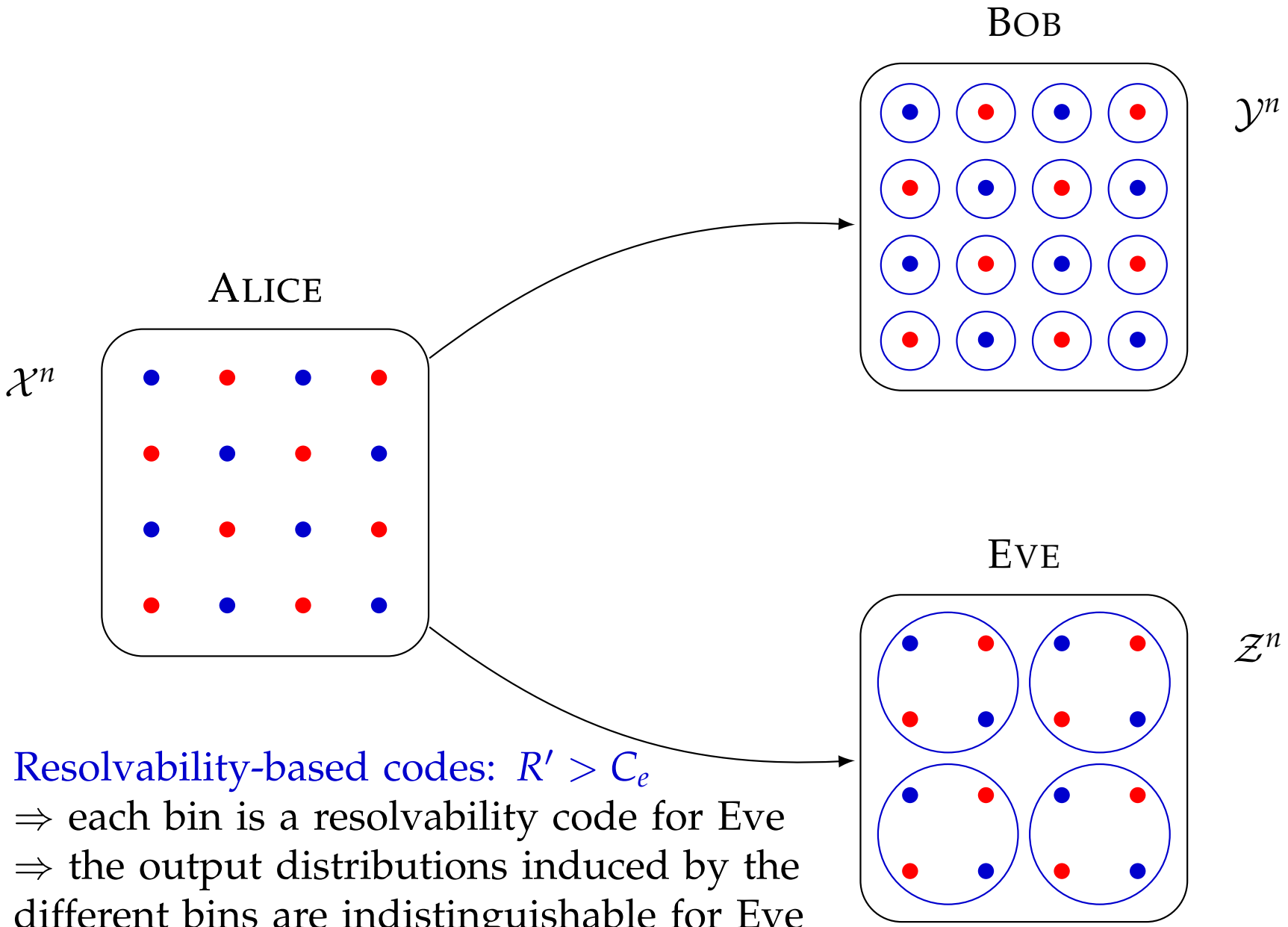


Capacity-based codes: $R' \nearrow C_e$

$\Rightarrow$ each bin is a capacity-approaching code

$\Rightarrow$ if Eve knows the bin color M, she can recover the auxiliary message M′ and X^n

Capacity-based codes vs. resolvability-based codes



Resolvability-based codes: $R' > C_e$

$\Rightarrow$ each bin is a resolvability code for Eve

$\Rightarrow$ the output distributions induced by the different bins are indistinguishable for Eve

Resolvability-based codes achieve strong secrecy

M. Bloch, N. Laneman, "Secrecy from resolvability", submitted, http://arxiv.org/abs/1105.5419

Useful fact: for bin rates above channel capacity, the output of one random bin resembles the output of a uniform source.

- $p_{Z^n|M=m}$ output distribution of the bin corresponding to the message m
- $p_{\bar{Z}^n}$ output distribution of a uniform source

Lemma (Cloud mixing)

If the bin rate $R' > C_e$, then with high probability, $\mathbb{V}(p_{Z^n|M=m}, p_{\bar{Z}^n}) \to 0$.

Lemma (Strong secrecy from resolvability)

If $R + R' < C_b$ and $R' > C_e$, then $\mathbb{I}(M; Z^n) \to 0$ with high probability.

Problem: can capacity-based codes also achieve strong secrecy?

Capacity-based codes cannot achieve strong secrecy

Unfortunately, a capacity-based code picked at random in the code ensemble will probably be bad:

Proposition

If $R + R' < C_b$ *and* $R' \nearrow C_e$, *then* $\exists \eta > 0$ *such that* $\forall \varepsilon > 0$,

$$\mathbb{I}(\mathsf{M}; \mathsf{Z}^n) > \eta - \varepsilon - f_\varepsilon(n)$$

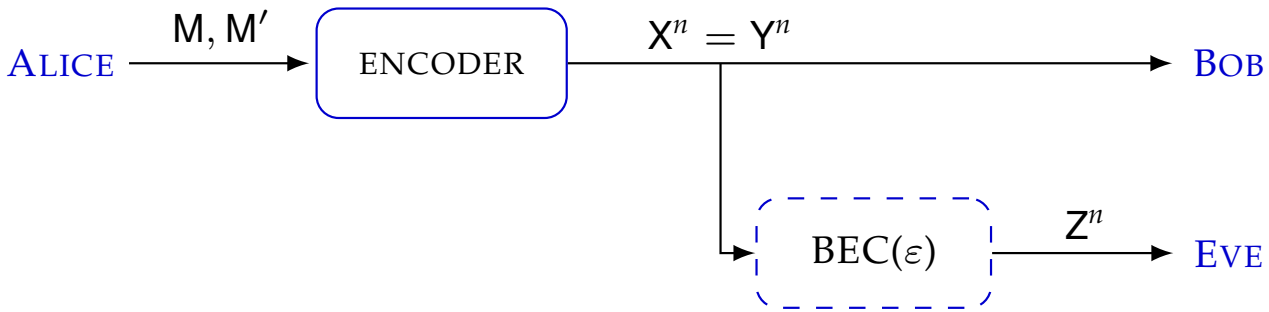
with (exponentially) high probability, and $\lim_{n\to\infty} f_\varepsilon(n) = 0$.

- the proof is based on a result on source coding with side information [Watanabe *et al.* 2009]

Outline

LDPC codes for the binary erasure wiretap channel

A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla,
"Applications of LDPC Codes to the Wiretap Channels", *IEEE Trans. Inform. Theory*, vol. 53, no. 8, 2007.



- **Idea**: use cosets of binary linear codes as "bins". [Thangaraj *et al.* 2007]
- capacity-based: use the cosets of a good code
  - achieves weak secrecy
  - capacity-achieving codes are hard to design
- resolvability-based: use the cosets of the dual of a good code
  - achieves strong secrecy in some special cases [Suresh *et al* 2010]
  - does not work if main channel is noisy (the dual of a good code is not a good code in general)

Dual code construction

- $C \sim (n, k)$ binary linear code (LDPC code),
  $H : \mathbb{F}_2^n \to \mathbb{F}_2^{n-k}$ parity-check matrix of C

- $C_\perp \sim (n, n-k)$ dual code of C,
  $G_\perp = H^T : \mathbb{F}_2^{n-k} \to \mathbb{F}_2^n$ generator matrix of $C_\perp$

- Choose $G' : \mathbb{F}_2^k \to \mathbb{F}_2^n$ such that the union of the columns of $G_\perp$ and G' is a basis of $\mathbb{F}_2^n$.

- **Encoder:** $\mathbf{m} \in \mathbb{F}_2^k$ confidential message, $\mathbf{m}' \in \mathbb{F}_2^{n-k}$ random message.

$$\mathbf{x} = G_\perp \mathbf{m}' + G'\mathbf{m} \ \in \mathbb{F}_2^n \quad \text{transmitted codeword}$$

- The bin of possible codewords $\mathbf{x}$ associated to the message $\mathbf{m}$ is a coset

$$\{G'\mathbf{m} + C_\perp\}$$

Dual code construction

- $\mathbf{z} \in \{0, 1, e\}^n$ output of Eve's channel.
- If a coset of $C_\perp$ contains at least one vector that coincides with $\mathbf{z}$ in all the unerased positions, we say that the coset is consistent with $\mathbf{z}$.
  We say that $\mathbf{z}$ is secured by $C_\perp$ if all cosets are consistent with $\mathbf{z}$.

Theorem (Ozarow and Wyner)

Let $G_{n-|\mathcal{E}|}$ be the submatrix of $G_\perp$ corresponding to the unerased bits of $\mathbf{z}$. Then

$$\mathbf{z} \text{ is secured by } C_\perp \quad \Leftrightarrow \quad G_{n-|\mathcal{E}|} \text{ has rank } n - |\mathcal{E}| \,.$$

Threshold property of LDPC codes

An ensemble of LDPC codes has block error probability threshold α if the block error probability on a BEC(ε) under belief propagation decoding tends to 0 as $n \to \infty$ if and only if $\varepsilon < \alpha$.

The dual code construction achieves weak secrecy

Remark

Let $\mathbf{H}$ be the parity-check matrix of an LDPC code C selected at random in an ensemble whose block error probability threshold is α, and construct a submatrix $\mathbf{H}'$ by selecting each *column* of $\mathbf{H}$ with probability $\varepsilon < \alpha$. Then with probability $1 - \delta(n)$, $\mathbf{H}'$ is full rank.

Proof.

We know that $\mathbf{Hz} = 0$ has a unique solution $\mathbf{z} \in C$ with probability $1 - \delta(n)$. So the matrix $\mathbf{H}'$ corresponding to the *erased* bits must be full rank. □

Proposition

The dual code construction achieves weak secrecy over any BEC(ε) with $1 - \varepsilon < \alpha$.

Proof.

With probability $1-\delta(n)$, the submatrix of $G_\perp = H^T$ formed by the *rows* corresponding to *unerased bits* is full rank and so $\mathbf{z}$ is secured by $C_\perp$. Let

$$\mathsf{Q} = \begin{cases} 1 & \text{if } \mathsf{Z}^n \text{ is secured by } C_\perp, \\ 0 & \text{otherwise.} \end{cases}$$

Then we have:

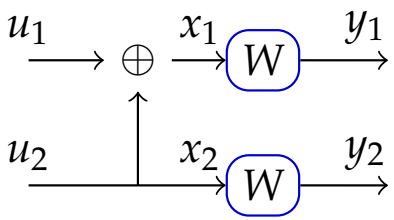
$$\begin{aligned} &\mathbb{I}(\mathsf{M};\mathsf{Z}^n) = \mathbb{H}(\mathsf{M}) - \mathbb{H}(\mathsf{M}|\mathsf{Z}^n) \leq \mathbb{H}(\mathsf{M}) - \mathbb{H}(\mathsf{M}|\mathsf{Z}^n,\mathsf{Q}) = \\ &= \mathbb{H}(\mathsf{M}) - \mathbb{H}(\mathsf{M}|\mathsf{Z}^n,\mathsf{Q}=1)\mathbb{P}(\mathsf{Q}=1) - \mathbb{H}(\mathsf{M}|\mathsf{Z}^n,\mathsf{Q}=0)\mathbb{P}(\mathsf{Q}=0) \leq \\ &\leq \mathbb{H}(\mathsf{M}) - \mathbb{H}(\mathsf{M})(1-\delta(n)) = \mathbb{H}(\mathsf{M})\delta(n) \end{aligned}$$

Outline

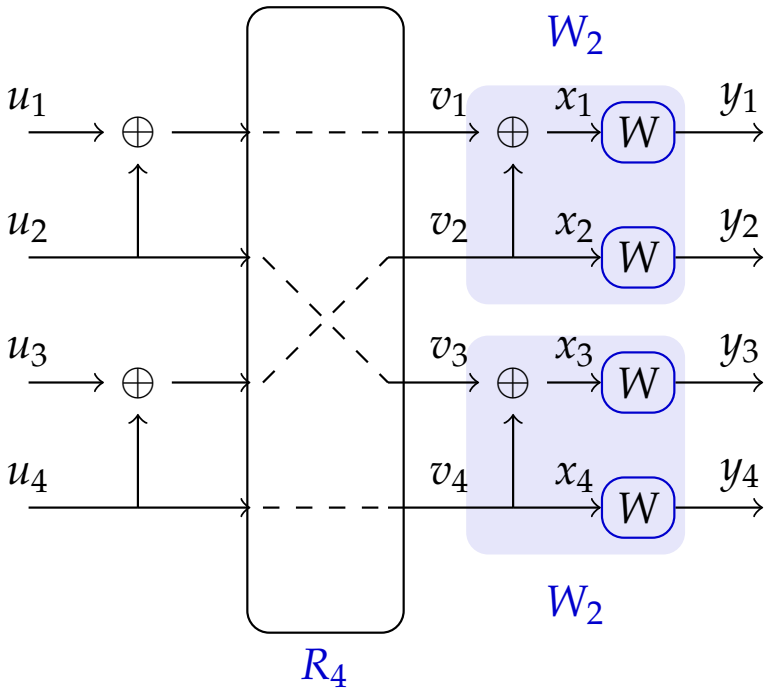
Polar Codes

E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels", *IEEE Trans. Inform. Theory*, vol. 55, n.7, 2009.

- $n = 2$



- $n = 4$



- $W : \{0,1\} \to \mathcal{Y}$ binary-input symmetric DMC
- Channel combining: for $n = 2^k$, recursively define a vector channel $W_n : \{0,1\}^n \to \mathcal{Y}^n$

$$W_n(\mathbf{y}|\mathbf{u}) = W^n(\mathbf{y}|\mathbf{x}), \quad \mathbf{x} = G_n\mathbf{u}$$

- Channel splitting: define the bit channels $W_n^{(i)} : \{0,1\} \to \mathcal{Y}^n \times \{0,1\}^{i-1}$

$$W_n^{(i)}(\mathbf{y}, (u_1, \ldots, u_{i-1})|u_i) =$$
$$= \sum_{u_{i+1},\ldots,u_n} \frac{1}{2^{n-1}} W_n(\mathbf{y}|\mathbf{u})$$

Channel polarization

Bhattacharyya parameter

For a binary-input channel $W : \{0,1\} \to \mathcal{Y}$, $Z(W) = \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)}$

- $C(W) \geq \log \frac{2}{1+Z(W)}$, $C(W) \leq \sqrt{1 - Z(W)^2}$.
- $Z(W)$ close to 1 $\Rightarrow$ very noisy channel
- $Z(W)$ close to 0 $\Rightarrow$ almost noiseless channel

Consider the bit-channels for $n = 2$: $W' = W_2^{(1)}$, $W'' = W_2^{(2)}$

$$W'(y_1, y_2|u_1) = \sum_{u_1'} \frac{1}{2} W(y_1|u_1 \oplus u_2') W(y_2|u_2')$$

$$W''(y_1, y_2|u_1, u_2) = \frac{1}{2} W(y_1|u_1 \oplus u_2) W(y_2|u_2)$$

The transformation $(W, W) \mapsto (W', W'')$ has the following properties:

$$C(W') + C(W'') = 2C(W), \quad C(W') \leq C(W'')$$

$$Z(W'') = Z(W)^2, \quad Z(W') \leq 2Z(W) - Z(W)^2, \quad Z(W') \leq Z(W) \leq Z(W'')$$

Channel polarization

For $0 < \beta < \frac{1}{2}$, define

$$\mathcal{G}_n(W,\beta) = \left\{ i \in [\![1,n]\!] \mid Z(W_n^{(i)}) < \frac{2^{-n^\beta}}{n} \right\} \quad \text{good bit channels}$$

$$\mathcal{B}_n(W,\beta) = \left\{ i \in [\![1,n]\!] \mid Z(W_n^{(i)}) \geq \frac{2^{-n^\beta}}{n} \right\} \quad \text{bad bit channels}$$

Theorem (Channel polarization)

When $n \to \infty$, the fraction $\frac{|\mathcal{G}_n(W,\beta)|}{n}$ of good bit channels tends to $C(W)$ and the fraction $\frac{|\mathcal{B}_n(W,\beta)|}{n}$ of bad bit channels tends to $1 - C(W)$.

Polar Codes for channel coding

Theorem (Channel polarization)

When $n \to \infty$, the fraction $\frac{|\mathcal{G}_n(W,\beta)|}{n}$ of good bit channels tends to $C(W)$ and the fraction $\frac{|\mathcal{B}_n(W,\beta)|}{n}$ of bad bit channels tends to $1 - C(W)$.

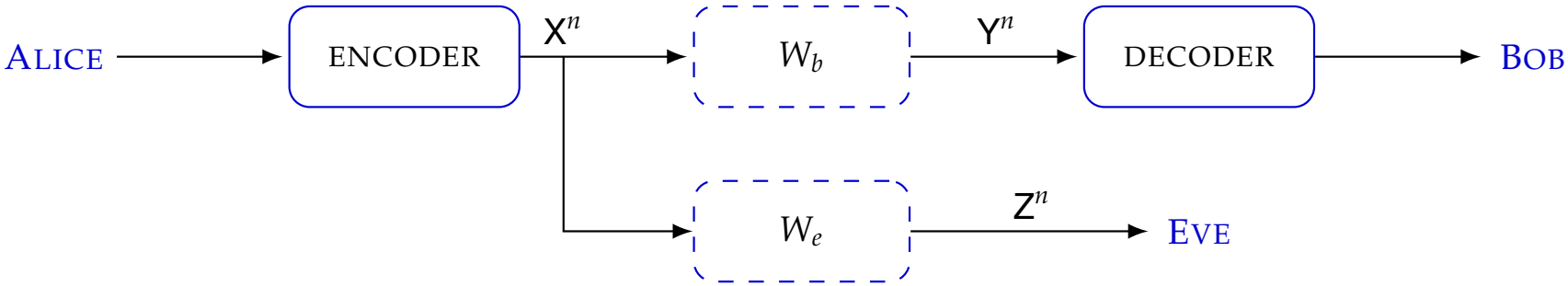
	BIT CHANNELS:	INPUT:
$\mathcal{A} = \mathcal{G}_n(W, \beta)$	good bit channels	information bits
$\mathcal{B} = \mathcal{B}_n(W, \beta)$	bad bit channels	zeros

Theorem (Arikan and Telatar)

$\forall \beta < \frac{1}{2}$, the polar coding scheme $C_n(\mathcal{A})$ has error probability
$P_e \leq \sum_{i \in \mathcal{A}} Z(W_n^{(i)}) \leq 2^{-n^\beta}$.

Polar coding for the binary symmetric wiretap channel: weak secrecy scheme

H. Mahdavifar and A. Vardy, "Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes", *ISIT*, 2010.



- W_b, W_e binary-input symmetric channels such that W_e is degraded with respect to W_b.

	BIT CHANNELS:	INPUT:
$\mathcal{R} = \mathcal{G}_n(W_e, \beta)$	good for both Bob and Eve	random bits
$\mathcal{A} = \mathcal{G}_n(W_b, \beta) \setminus \mathcal{G}_n(W_e, \beta)$	good for Bob but bad for Eve	information bits
$\mathcal{B} = \mathcal{B}_n(W_b, \beta)$	bad for both Bob and Eve	zeros

Polar Codes achieve weak secrecy

Let $r = |\mathcal{R}|, k = |\mathcal{A}|$.

Lemma (Mahdavifar and Vardy)

$$\mathbb{H}(\mathsf{V}_{\mathcal{R}}|\mathsf{Z}^n, \mathsf{V}_{\mathcal{A}}) \leq H_b(2^{-n^\beta}) + r2^{-n^\beta}$$

Lemma (Mahdavifar and Vardy)

Let $\varepsilon_n = C(W) - \frac{r}{n}$. *Then*

$$\mathbb{I}(\mathsf{U}; \mathsf{Z}^n) \leq n\varepsilon_n + H_b(2^{-n^\beta}) + (n-k)2^{-n^\beta}$$

Polar Codes achieve the weak secrecy capacity:

$$R_n = \frac{|\mathcal{A}|}{n} = \frac{|\mathcal{G}_n(W_b, \beta)|}{n} - \frac{|\mathcal{G}_n(W_e, \beta)|}{n}$$
$$\Rightarrow \lim_{n\to\infty} R_n = C_b - C_e$$

Proof of the first Lemma.

From Fano's inequality,

$$\mathbb{H}(\mathsf{V}_{\mathcal{R}}|\mathsf{Z}^n,\mathsf{V}_{\mathcal{A}}) \leq H_b(\lambda) + r\lambda \leq H_b(2^{-n^\beta}) + r2^{-n^\beta},$$

where $\lambda = \mathbb{P}\{\hat{\mathsf{V}}_{\mathcal{R}} \neq \mathsf{V}_{\mathcal{R}}\} \leq 2^{-n^\beta}$ because of Arikan and Telatar's theorem. □

Proof of the second Lemma.

$$\mathbb{I}(\mathsf{U};\mathsf{Z}^n) = \mathbb{I}(\mathsf{V}_{\mathcal{A}\cup\mathcal{B}};\mathsf{Z}^n) = \mathbb{I}(\mathsf{V};\mathsf{Z}^n) - \mathbb{I}(\mathsf{V}_{\mathcal{R}};\mathsf{Z}^n|\mathsf{V}_{\mathcal{A}\cup\mathcal{B}}) =$$
$$\mathbb{I}(\mathsf{V};\mathsf{Z}^n) - \mathbb{I}(\mathsf{V}_{\mathcal{R}};\mathsf{Z}^n|\mathsf{V}_{\mathcal{A}}) = \mathbb{I}(\mathsf{V};\mathsf{Z}^n) - \mathbb{H}(\mathsf{V}_{\mathcal{R}}|\mathsf{V}_{\mathcal{A}}) + \mathbb{H}(\mathsf{V}_{\mathcal{R}}|\mathsf{Z}^n,\mathsf{V}_{\mathcal{A}}) \leq$$
$$\leq nC_e - r + \mathbb{H}(\mathsf{V}_{\mathcal{R}}|\mathsf{Z}^n,\mathsf{V}_{\mathcal{A}}) = n\varepsilon_n + \mathbb{H}(\mathsf{V}_{\mathcal{R}}|\mathsf{Z}^n,\mathsf{V}_{\mathcal{A}}) \leq n\varepsilon_n + H_b(2^{-n^\beta}) + r2^{-n^\beta}$$

□

Strong secrecy scheme

H. Mahdavifar and A. Vardy, "Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes", submitted to *IEEE Trans. Inform. Theory*, 2010.

Simply by modifying the definition of good and bad bit channels, we can transform the previous capacity-based wiretap scheme into a resolvability-based one.

$$\mathcal{P}_n(W, \delta_n) = \left\{ i \in [\![1, n]\!] \mid C(W_n^{(i)}) < \delta_n \right\} \qquad \text{"poor" bit channels}$$

	BIT CHANNELS:	INPUT:
$\mathcal{R} = [\![1, n]\!] \setminus \mathcal{P}_n(W_e, \delta_n)$	good for Eve	random bits
$\mathcal{A} = \mathcal{G}_n(W_b, \beta) \cap \mathcal{P}_n(W_e, \delta_n)$	poor for Eve, good for Bob	information bits
$\mathcal{B} = \mathcal{P}_n(W_e, \delta_n) \setminus \mathcal{G}_n(W_b, \beta)$	poor for Eve, bad for Bob	zeros

$$\frac{r}{n} = \frac{|\mathcal{R}|}{n} \searrow C(W_e)$$

Strong secrecy scheme

- Define a new channel $Q_n(W_e, \mathcal{R})$ which includes the polar code and the random bits:

$$Q_n(\mathbf{z}|\mathbf{x}) = \frac{1}{2^r} \sum_{\mathbf{e}\in\{0,1\}^r} W_e^n \left(\mathbf{z} \middle| G_n \begin{pmatrix} \mathbf{x} \\ \mathbf{e} \end{pmatrix} \right)$$

- $Q_n(W_e, \mathcal{R})$ is a symmetric channel. Therefore

$$C(Q_n(W_e, \mathcal{R})) = I(\mathsf{V}_{\mathcal{R}^c}; \mathsf{Z}^n) = \sum_{i\in\mathcal{R}^c} C(W_{e,n}^{(i)}) \leq \delta_n \, |\mathcal{P}_n(W_e, \delta_n)|$$

Proposition (Mahdavifar and Vardy)

Regardless of the distribution of the input U, $\mathbb{I}(\mathsf{U}; \mathsf{Z}^n) \leq \delta_n \, |\mathcal{P}_n(W_e, \delta_n)|$.
Moreover, we can choose $\delta_n = o\left(\frac{1}{n}\right)$ *so that the scheme guarantees strong secrecy.*

Proof.

$$\mathbb{I}(\mathsf{U}; \mathsf{Z}^n) = \mathbb{I}(\mathsf{V}_{\mathcal{A}}; \mathsf{Z}^n) = \mathbb{I}(\mathsf{V}_{\mathcal{A}\cup\mathcal{B}}; \mathsf{Z}^n) \leq C(Q_n(W_e, \mathcal{R}))$$ □

References for Part IV

- A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC Codes to the Wiretap Channel", *IEEE Trans. Inform. Theory*, vol. 53, n. 8, 2007.
- H. Mahdavifar and A. Vardy, "Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes", *Proc. ISIT 2010*
- H. Mahdavifar and A. Vardy, "Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes", submitted to *IEEE Trans. Inform. Theory*, 2010.
- E. Hof and S. Shamai, "Secrecy-achieving polar-coding", *Proc. ITW 2010*.
- A. T. Suresh, A. Subramanian, A. Thangaraj, M. Bloch, and S. McLaughlin, "Strong Secrecy for Erasure Wiretap Channels", *Proc. ITW 2010*.