The Wire-Tap Channel and Secrecy in Communication

Group 2
Adway Girish, Fathima Zarin Faizal

Department of Electrical Engineering
IIT Bombay

Outline

The communication system

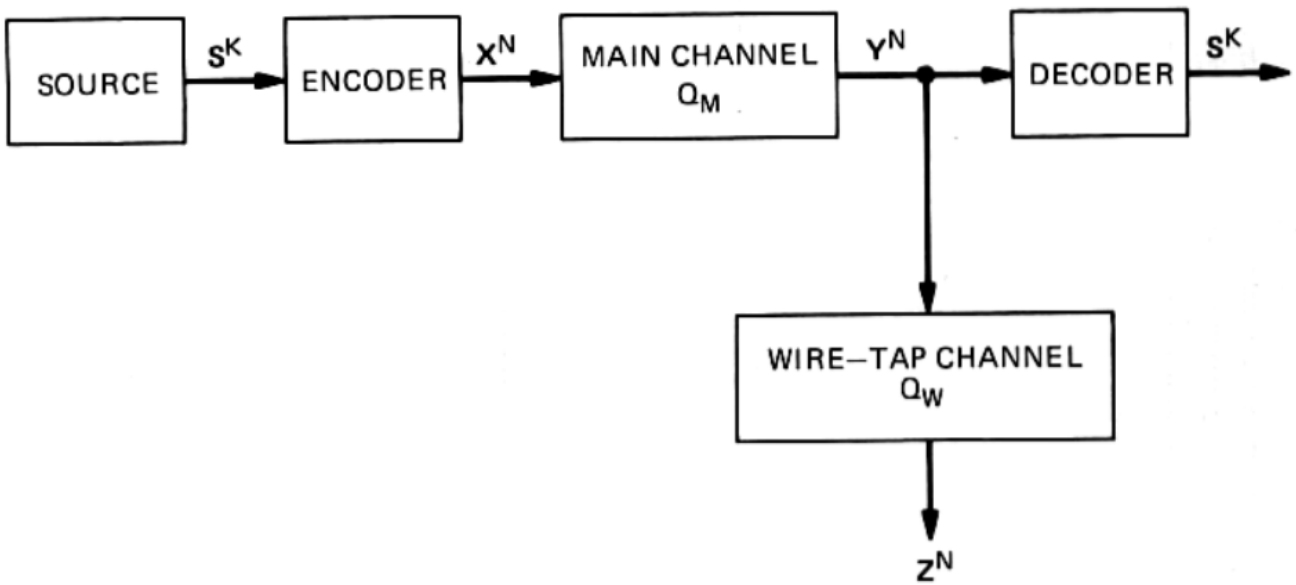

Figure: Wire-tap channel (from [1])

The communication system

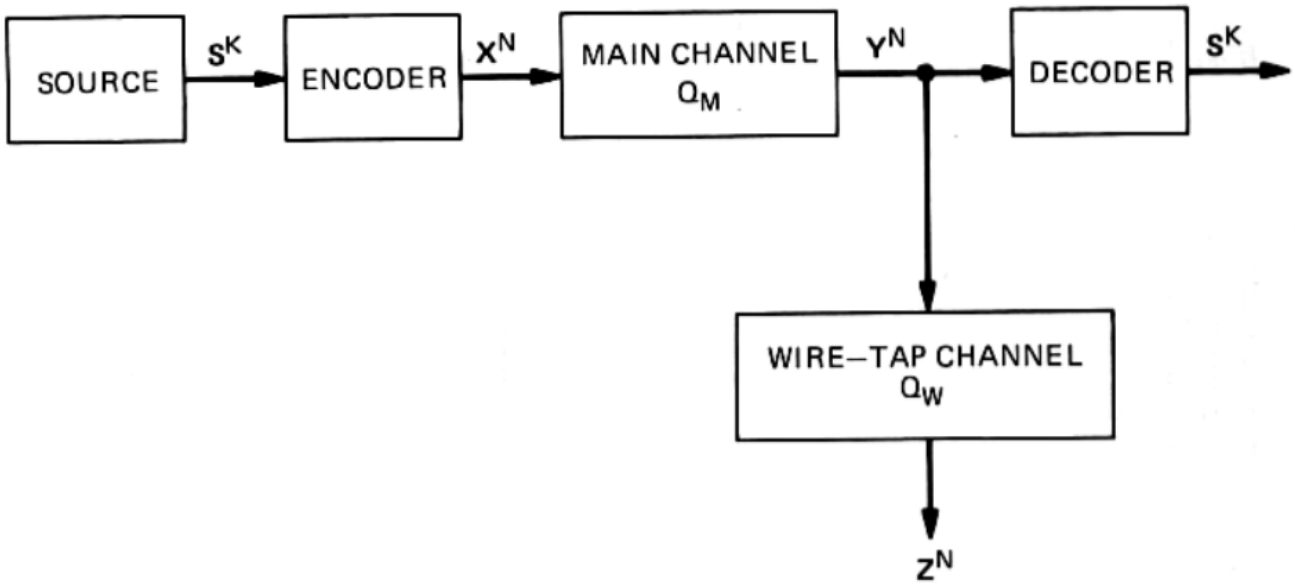

Figure: Wire-tap channel (from [1])

- **Source:**
  Defined by $\{S_k\}_1^\infty$ where S_k are iid random variables that take values in the finite set $\mathcal{S}$.

The communication system

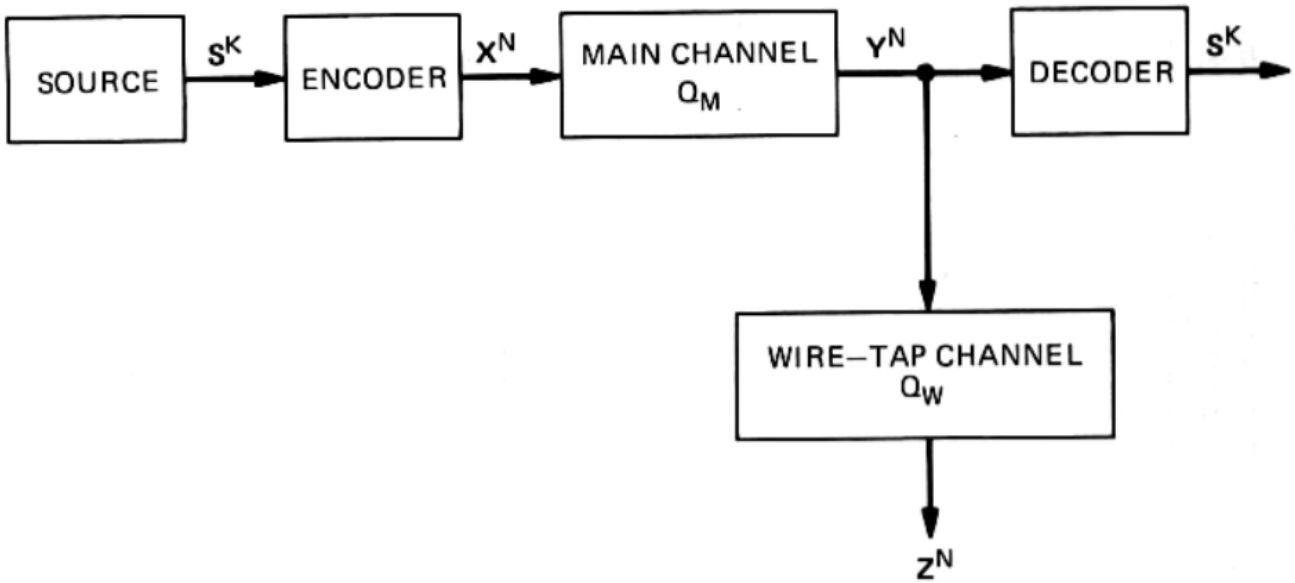

Figure: Wire-tap channel (from [1])

- **Main channel:**
  Discrete memoryless channel with finite input alphabet $\mathcal{X}$, finite output alphabet $\mathcal{Y}$ and transition probability $Q_M(y|x)$. Denote the channel capacity by C_M.

The communication system

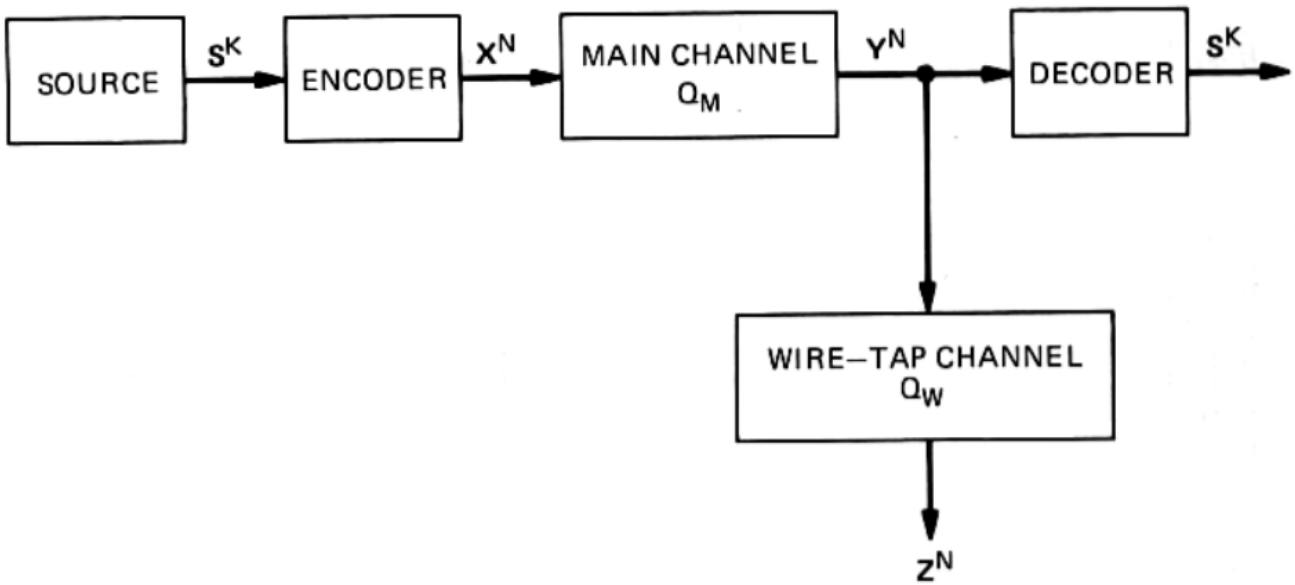

Figure: Wire-tap channel (from [1])

- **Wire-tap channel:**
  Discrete memoryless channel with input alphabet $\mathcal{Y}$, finite output alphabet $\mathcal{Z}$, and transition probability $Q_W(z|y)$.

Encoder-decoder (K, N, Δ, P_e)

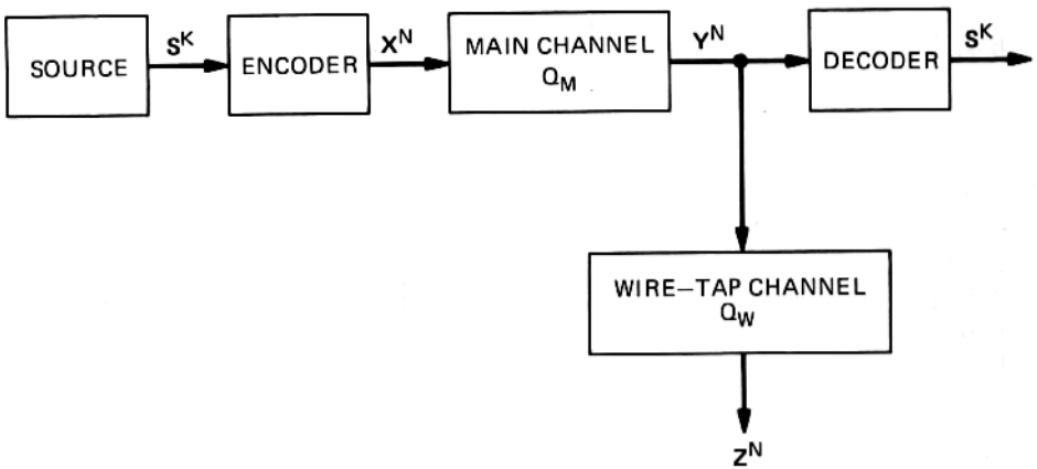

Figure: Wire-tap channel (from [1])

Encoder-decoder (K, N, Δ, P_e)

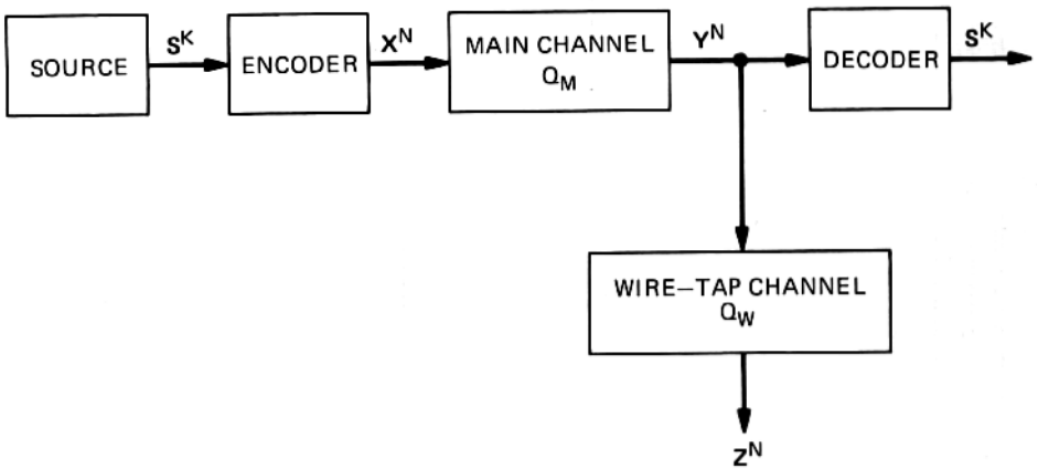

Figure: Wire-tap channel (from [1])

- **Encoder:**
  A (K, N) encoder takes in S^K and outputs the random vector X^N.

Encoder-decoder (K, N, Δ, P_e)

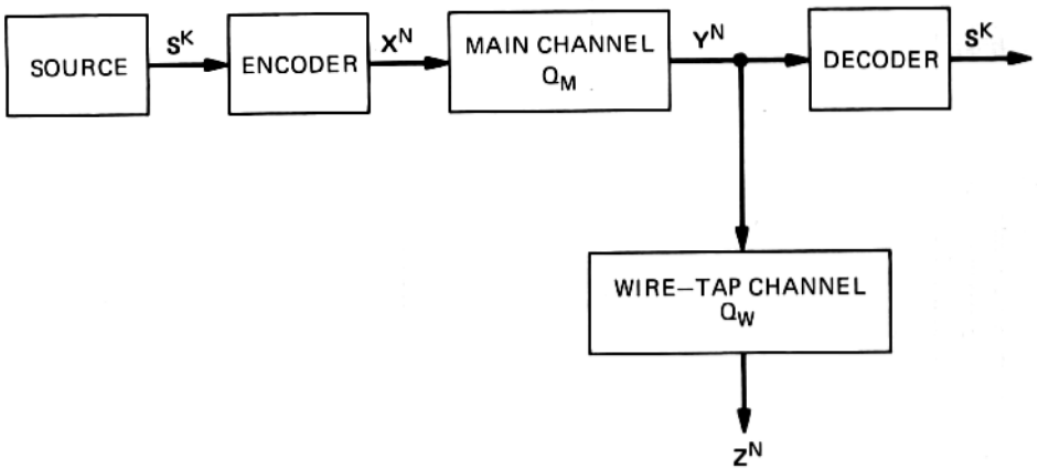

Figure: Wire-tap channel (from [1])

- **Decoder:**
  A mapping $f_D : \mathcal{Y}^N \to \mathcal{S}^K$ such that $\hat{\boldsymbol{S}} = f_D(\boldsymbol{Y})$.

Encoder-decoder (K, N, Δ, P_e)

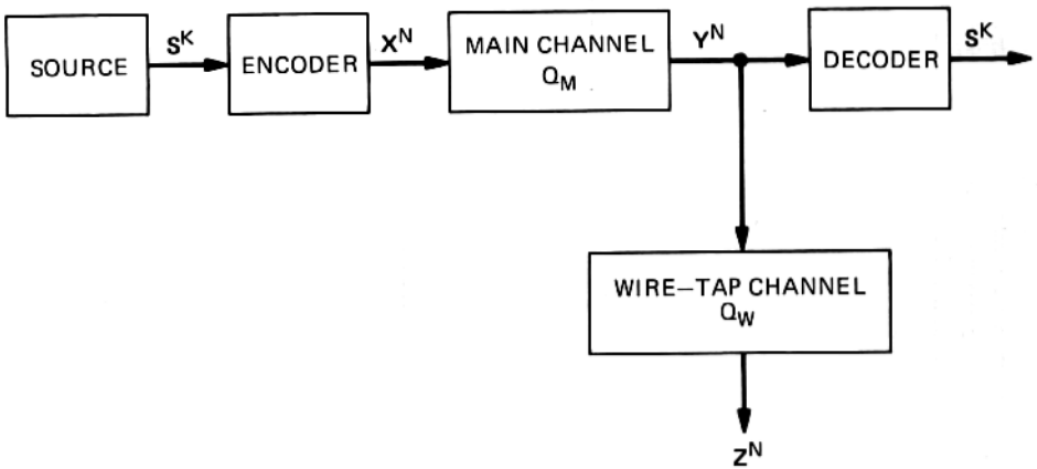

Figure: Wire-tap channel (from [1])

- For a given encoder and decoder, the **error rate** is

$$P_e = \frac{1}{K} \sum_{k=1}^{K} \mathbb{P}\{S_k \neq \hat{S}_k\}$$

Encoder-decoder (K, N, Δ, P_e)

- The **equivocation** of the source at the output of the wire-tap channel for a particular encoder is

$$\Delta \triangleq \frac{1}{K} H(\boldsymbol{S}^K | \boldsymbol{Z}^N)$$

and serves as a criterion of the wire-tapper's confusion.

Encoder-decoder (K, N, Δ, P_e)

- The **equivocation** of the source at the output of the wire-tap channel for a particular encoder is

$$\Delta \triangleq \frac{1}{K} H(\boldsymbol{S}^K | \boldsymbol{Z}^N)$$

and serves as a criterion of the wire-tapper's confusion.
Define P_{ew} as the wire-tapper's error probability. From Fano's inequality, we get

$$\Delta \leq h(P_{ew}) + P_{ew} \log |\mathcal{S}|$$

Encoder-decoder (K, N, Δ, P_e)

- The **equivocation** of the source at the output of the wire-tap channel for a particular encoder is

$$\Delta \triangleq \frac{1}{K} H(\boldsymbol{S}^K | \boldsymbol{Z}^N)$$

and serves as a criterion of the wire-tapper's confusion.
Define P_{ew} as the wire-tapper's error probability. From Fano's inequality, we get

$$\Delta \leq h(P_{ew}) + P_{ew} \log |\mathcal{S}|$$

Thus a large Δ corresponds to a large P_{ew}.

Formal problem statement

Formal problem statement

A pair (R, d) is **achievable**

Formal problem statement

A pair (R, d) is **achievable** if, $\forall \epsilon > 0$, $\exists$ an encoder-decoder (N, K, Δ, P_e) such that

$$\frac{H_S K}{N} \geq R - \epsilon \tag{1a}$$

$$\Delta \geq d - \epsilon \tag{1b}$$

$$P_e \leq \epsilon \tag{1c}$$

Formal problem statement

A pair (R, d) is **achievable** if, $\forall \epsilon > 0$, $\exists$ an encoder-decoder (N, K, Δ, P_e) such that

$$\frac{H_S K}{N} \geq R - \epsilon \tag{1a}$$

$$\Delta \geq d - \epsilon \tag{1b}$$

$$P_e \leq \epsilon \tag{1c}$$

The problem is to characterize the set $\mathcal{R}$ of achievable (R, d) pairs.

A digression

A digression

- Let $X \sim p_X(x), x \in \mathcal{X}$.
  For $R \geq 0$, let $\mathcal{P}(R) = \{p_X : I(X;Y) \geq R\}$.

A digression

- Let $X \sim p_X(x), x \in \mathcal{X}$.
  For $R \geq 0$, let $\mathcal{P}(R) = \{p_X : I(X;Y) \geq R\}$.
- For $0 \leq R \leq C_M$, define

$$\Gamma(R) \triangleq \sup_{p_X \in \mathcal{P}(R)} I(X;Y|Z) \tag{2}$$

The main theorem

The main theorem

Theorem 1

The set $\mathcal{R}$ defined above is equal to $\overline{\mathcal{R}}$, where

$$\overline{\mathcal{R}} \triangleq \{(R,d) : 0 \leq R \leq C_M, 0 \leq d \leq H_S, Rd \leq H_S\Gamma(R)\} \tag{3}$$

Setting for the special case

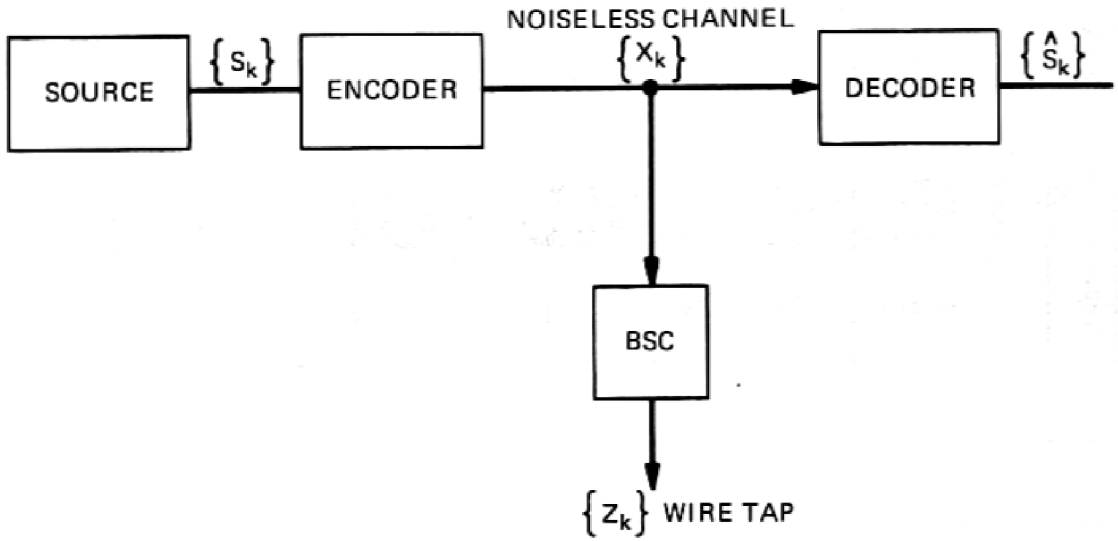

Figure: Wire-tap channel (from [1])

Setting for the special case

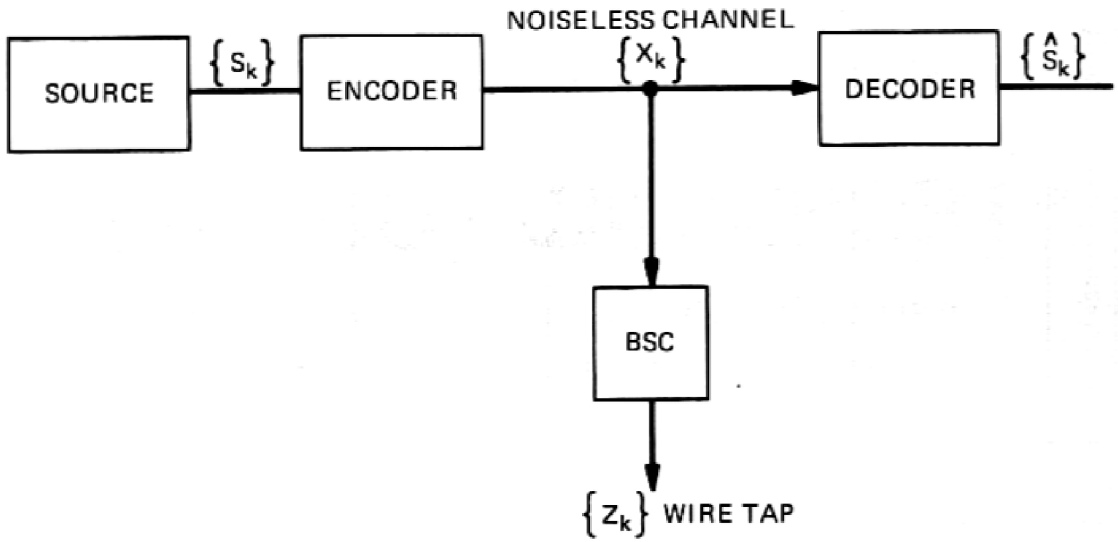

Figure: Wire-tap channel (from [1])

- **Source:**
  A simple equiprobable binary source emitting a data sequence $S_1, S_2, \ldots$, where $\Pr\{S_i = 0\} = \Pr\{S_i = 1\} = \frac{1}{2}$ for all i.

Setting for the special case

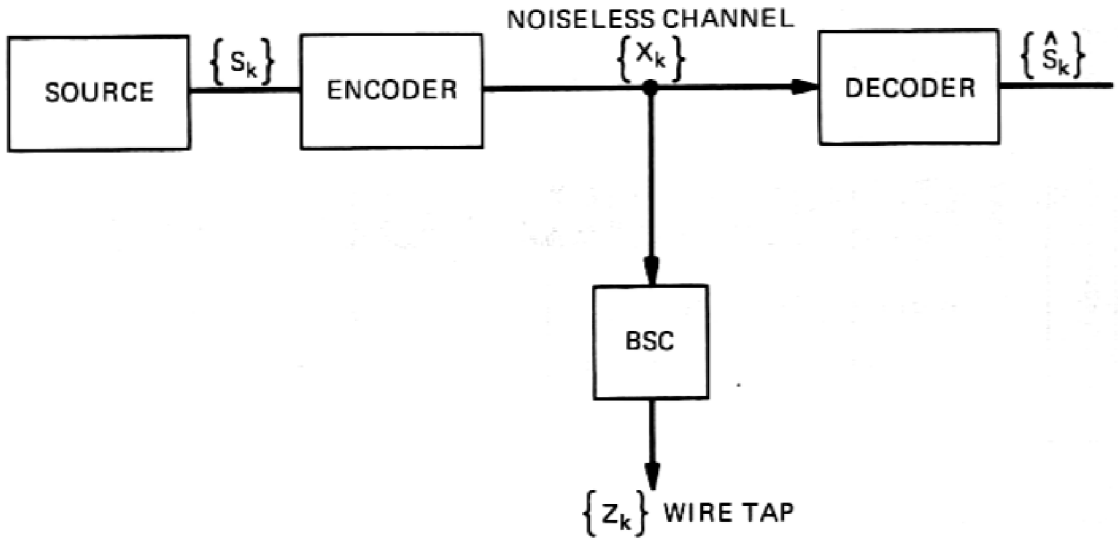

Figure: Wire-tap channel (from [1])

- **Main channel:**
  A perfectly noiseless channel, i.e.

$$Q_M(y|x) = \delta_{x,y}$$

Setting for the special case

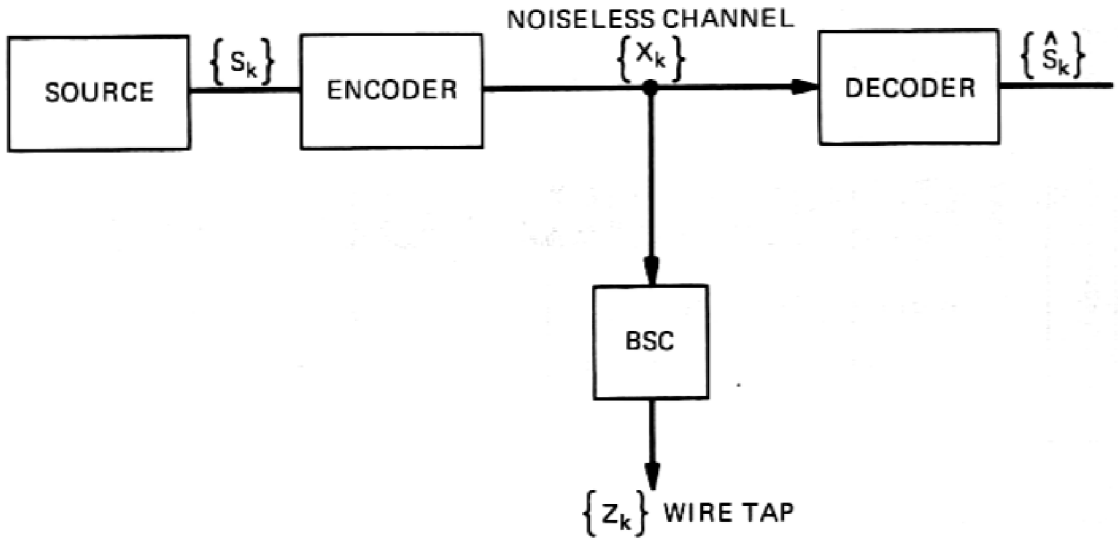

Figure: Wire-tap channel (from [1])

- **Wire-tap channel:**
  A Binary Symmetric Channel with crossover probability p_0, i.e.

$$Q_W(z|y) = (1 - p_0)\delta_{y,z} + p_0(1 - \delta_{y,z})$$

Theorem

For this example, observe that $C_M = 1$, $H_S = 1$, and $\Gamma(R) = h(p_0)$ for all $0 \leq R \leq 1$.

Theorem

For this example, observe that $C_M = 1$, $H_S = 1$, and $\Gamma(R) = h(p_0)$ for all $0 \leq R \leq 1$. Thus our theorem reduces to the following

Theorem

For this example, observe that $C_M = 1$, $H_S = 1$, and $\Gamma(R) = h(p_0)$ for all $0 \leq R \leq 1$. Thus our theorem reduces to the following

Theorem 2 (for the special case)

The pair (R, d) is achievable if and only if

$$0 \leq R \leq 1, \qquad 0 \leq d \leq 1, \qquad Rd \leq h(p_0) \tag{4}$$

Theorem

For this example, observe that $C_M = 1$, $H_S = 1$, and $\Gamma(R) = h(p_0)$ for all $0 \leq R \leq 1$. Thus our theorem reduces to the following

Theorem 2 (for the special case)

The pair (R, d) is achievable if and only if

$$0 \leq R \leq 1, \qquad 0 \leq d \leq 1, \qquad Rd \leq h(p_0) \tag{4}$$

We first prove the converse, then the direct half.

Theorem

For this example, observe that $C_M = 1$, $H_S = 1$, and $\Gamma(R) = h(p_0)$ for all $0 \leq R \leq 1$. Thus our theorem reduces to the following

Theorem 2 (for the special case)

The pair (R, d) is achievable if and only if

$$0 \leq R \leq 1, \qquad 0 \leq d \leq 1, \qquad Rd \leq h(p_0) \tag{4}$$

We first prove the converse, then the direct half.
$R \leq C_M$ being necessary is immediate from the converse to the ordinary coding theorem

Theorem

For this example, observe that $C_M = 1$, $H_S = 1$, and $\Gamma(R) = h(p_0)$ for all $0 \leq R \leq 1$. Thus our theorem reduces to the following

Theorem 2 (for the special case)

The pair (R, d) is achievable if and only if

$$0 \leq R \leq 1, \qquad 0 \leq d \leq 1, \qquad Rd \leq h(p_0) \tag{4}$$

We first prove the converse, then the direct half.
$R \leq C_M$ being necessary is immediate from the converse to the ordinary coding theorem, and $d \leq H_S$ follows directly from

$$\Delta = \frac{1}{K} H(\boldsymbol{S}^K | \boldsymbol{Z}^N) \leq \frac{1}{K} H(\boldsymbol{S}^K) = H_S$$

Closer look at $\overline{\mathcal{R}}$

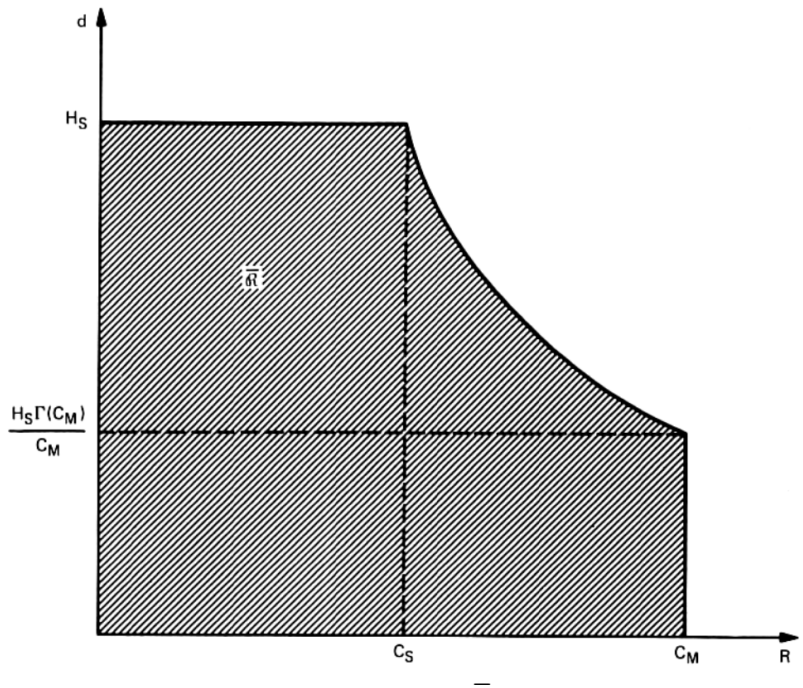

Figure: $\overline{\mathcal{R}}$ when Q_M is noiseless and Q_W a BSC (from [1])

Closer look at $\overline{\mathcal{R}}$

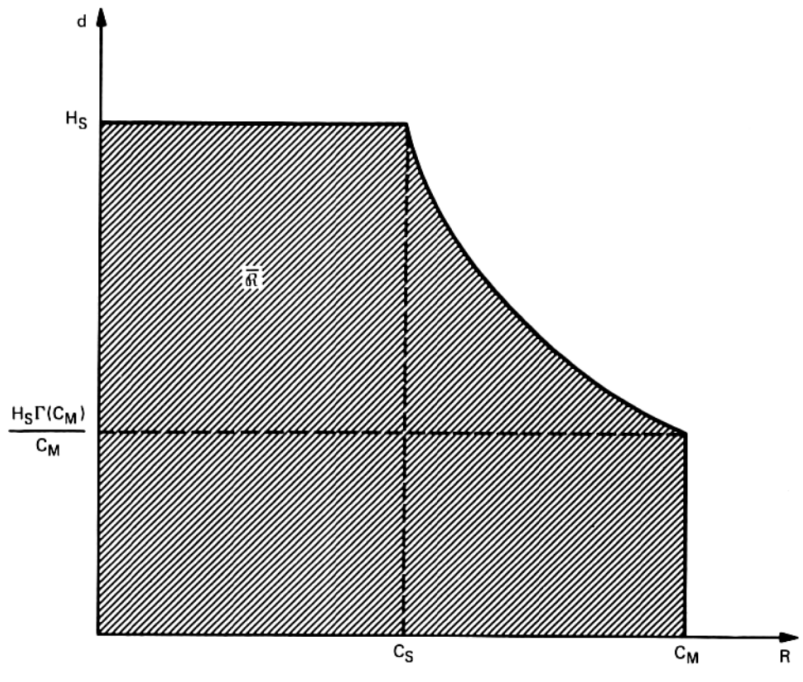

Figure: $\overline{\mathcal{R}}$ when Q_M is noiseless and Q_W a BSC (from [1])

Note that $\overline{\mathcal{R}}$ is not convex!

Converse

$$K\Delta = H(\boldsymbol{S}^K|\boldsymbol{Z}^N)$$

Converse

$$
\begin{aligned}
K\Delta &= H(\boldsymbol{S}^K|\boldsymbol{Z}^N) = H(\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{S}^K,\boldsymbol{X}^N,\boldsymbol{Z}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{Z}^N|\boldsymbol{X}^N,\boldsymbol{S}^K) + H(\boldsymbol{X}^N,\boldsymbol{S}^K) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{Z}^N|\boldsymbol{X}^N) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N)
\end{aligned}
$$

Converse

$$
\begin{aligned}
K\Delta &= H(\boldsymbol{S}^K|\boldsymbol{Z}^N) = H(\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{S}^K,\boldsymbol{X}^N,\boldsymbol{Z}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{Z}^N|\boldsymbol{X}^N,\boldsymbol{S}^K) + H(\boldsymbol{X}^N,\boldsymbol{S}^K) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= Nh(p_0) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N)
\end{aligned}
$$

Converse

$$
\begin{aligned}
K\Delta &= H(\boldsymbol{S}^K|\boldsymbol{Z}^N) = H(\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{S}^K,\boldsymbol{X}^N,\boldsymbol{Z}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{Z}^N|\boldsymbol{X}^N,\boldsymbol{S}^K) + H(\boldsymbol{X}^N,\boldsymbol{S}^K) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= Nh(p_0) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&\leq Nh(p_0) + Kh(P_e)
\end{aligned}
$$

Converse

$$
\begin{aligned}
K\Delta &= H(\boldsymbol{S}^K|\boldsymbol{Z}^N) = H(\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{S}^K,\boldsymbol{X}^N,\boldsymbol{Z}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{Z}^N|\boldsymbol{X}^N,\boldsymbol{S}^K) + H(\boldsymbol{X}^N,\boldsymbol{S}^K) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= Nh(p_0) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&\leq Nh(p_0) + Kh(P_e)
\end{aligned}
$$

$$
\Longrightarrow \quad \frac{K}{N}[\Delta - h(P_e)] \leq h(p_0)
$$

Converse

$$
\begin{aligned}
K\Delta &= H(\boldsymbol{S}^K|\boldsymbol{Z}^N) = H(\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{S}^K,\boldsymbol{X}^N,\boldsymbol{Z}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{Z}^N|\boldsymbol{X}^N,\boldsymbol{S}^K) + H(\boldsymbol{X}^N,\boldsymbol{S}^K) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= Nh(p_0) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&\leq Nh(p_0) + Kh(P_e)
\end{aligned}
$$

$$
\implies \frac{K}{N}[\Delta - h(P_e)] \leq h(p_0)
$$

Since (R, d) is achievable, this gives us

$$
(R-\epsilon)[(d-\epsilon) - h(\epsilon)] \leq h(p_0) \text{ for any } \epsilon > 0
$$

Converse

$$
\begin{aligned}
K\Delta &= H(\boldsymbol{S}^K|\boldsymbol{Z}^N) = H(\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{S}^K,\boldsymbol{X}^N,\boldsymbol{Z}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= H(\boldsymbol{Z}^N|\boldsymbol{X}^N,\boldsymbol{S}^K) + H(\boldsymbol{X}^N,\boldsymbol{S}^K) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&= Nh(p_0) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K,\boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\
&\leq Nh(p_0) + Kh(P_e)
\end{aligned}
$$

$$
\implies \frac{K}{N}[\Delta - h(P_e)] \leq h(p_0)
$$

Since (R, d) is achievable, this gives us

$$
(R-\epsilon)[(d-\epsilon) - h(\epsilon)] \leq h(p_0) \text{ for any } \epsilon > 0
$$

$$
\implies Rd \leq h(p_0) \quad \square
$$

Direct half

Now we show that any pair (R, d) satisfying

$$0 \leq R \leq 1, \qquad 0 \leq d \leq 1, \qquad Rd \leq h(p_0) \tag{5}$$

is achievable

Direct half

Now we show that any pair (R, d) satisfying

$$0 \leq R \leq 1, \qquad 0 \leq d \leq 1, \qquad Rd \leq h(p_0) \tag{5}$$

is achievable, i.e. the existence of an encoder-decoder (K, N, Δ, P_e) that gives arbitrarily low error and high equivocation.

Direct half

Let G be a group code, with block length N and $|G| = 2^{N-K}$. This has 2^K cosets, $\{C_i\}_1^{2^K}$. (with $C_1 = G$)

Direct half

Let G be a group code, with block length N and $|G| = 2^{N-K}$. This has 2^K cosets, $\{C_i\}_1^{2^K}$. (with $C_1 = G$) We use these to construct a code with rate $R = \frac{K}{N}$.

Direct half

Let G be a group code, with block length N and $|G| = 2^{N-K}$. This has 2^K cosets, $\{C_i\}_1^{2^K}$. (with $C_1 = G$) We use these to construct a code with rate $R = \frac{K}{N}$.

- **Encoder:**
  When $\boldsymbol{S}^K = i$, we let $\boldsymbol{X}^N$ be any random member of the coset C_i, i.e.

$$\Pr\{\boldsymbol{X}^N = \boldsymbol{x} | \boldsymbol{S} = i\} = \begin{cases} \frac{1}{|G|} = 2^{-(N-K)} & \boldsymbol{x} \in C_i \\ 0 & \text{else} \end{cases}$$

Direct half

Let G be a group code, with block length N and $|G| = 2^{N-K}$. This has 2^K cosets, $\{C_i\}_1^{2^K}$. (with $C_1 = G$) We use these to construct a code with rate $R = \frac{K}{N}$.

- **Encoder:**
  When $\boldsymbol{S}^K = i$, we let $\boldsymbol{X}^N$ be any random member of the coset C_i, i.e.

$$\Pr\{\boldsymbol{X}^N = \boldsymbol{x} | \boldsymbol{S} = i\} = \begin{cases} \frac{1}{|G|} = 2^{-(N-K)} & \boldsymbol{x} \in C_i \\ 0 & \text{else} \end{cases}$$

- **Decoder:**
  Simply letting $f_D(\boldsymbol{y}) = i$ when $\boldsymbol{y} \in C_i$, we have $P_e = 0$.

Direct half

We use the following equality from the converse proof,

$$K\Delta = Nh(p_0) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K, \boldsymbol{Z}^N) - H(\boldsymbol{Z}^N)$$

Direct half

We use the following equality from the converse proof,

$$\begin{aligned} K\Delta &= Nh(p_0) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K, \boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\ &\geq Nh(p_0) - h(\lambda) - \lambda(N-K) \end{aligned}$$

where λ is the error in the group code G.

Direct half

We use the following equality from the converse proof,

$$\begin{aligned} K\Delta &= Nh(p_0) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K, \boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\ &\geq Nh(p_0) - h(\lambda) - \lambda(N-K) \end{aligned}$$

where λ is the error in the group code G.

$$\implies \Delta \geq \frac{N}{K} h(p_0) - \frac{h(\lambda)}{K} - \lambda\left(\frac{N}{K} - 1\right)$$

Direct half

We use the following equality from the converse proof,

$$\begin{aligned} K\Delta &= Nh(p_0) + H(\boldsymbol{S}^K|\boldsymbol{X}^N) + H(\boldsymbol{X}^N) - H(\boldsymbol{X}^N|\boldsymbol{S}^K, \boldsymbol{Z}^N) - H(\boldsymbol{Z}^N) \\ &\geq Nh(p_0) - h(\lambda) - \lambda(N-K) \end{aligned}$$

where λ is the error in the group code G.

$$\begin{aligned} \implies \Delta &\geq \frac{N}{K}h(p_0) - \frac{h(\lambda)}{K} - \lambda\left(\frac{N}{K} - 1\right) \\ &\geq d - \left[\frac{h(\lambda)}{K} + \lambda\left(\frac{1}{R} - 1\right)\right] \end{aligned}$$

Direct half

We make use of the following lemma to see that there exists a code G with arbitrarily low λ.

Direct half

We make use of the following lemma to see that there exists a code G with arbitrarily low λ.

Lemma 3

For large enough N and $r < 1 - h(p_0)$, there exists a group code of length N with $|G| \geq 2^{Nr}$ which on passing through a BSC with crossover probability p_0 will have arbitrarily small probability of error.

Direct half

We make use of the following lemma to see that there exists a code G with arbitrarily low λ.

Lemma 3

For large enough N and $r < 1 - h(p_0)$, there exists a group code of length N with $|G| \geq 2^{Nr}$ which on passing through a BSC with crossover probability p_0 will have arbitrarily small probability of error.

Since $Rd \leq h(p_0)$ and $R = \frac{K}{N}$, we have

$$|G| = 2^{N-K} \leq 2^{N\left[1-\frac{h(p_0)}{d}\right]}$$

Direct half

We make use of the following lemma to see that there exists a code G with arbitrarily low λ.

Lemma 3

For large enough N and $r < 1 - h(p_0)$, there exists a group code of length N with $|G| \geq 2^{Nr}$ which on passing through a BSC with crossover probability p_0 will have arbitrarily small probability of error.

Since $Rd \leq h(p_0)$ and $R = \frac{K}{N}$, we have

$$|G| = 2^{N-K} \leq 2^{N\left[1-\frac{h(p_0)}{d}\right]} \leq 2^{N[1-h(p_0)]}$$

Direct half

We make use of the following lemma to see that there exists a code G with arbitrarily low λ.

Lemma 3

For large enough N and $r < 1 - h(p_0)$, there exists a group code of length N with $|G| \geq 2^{Nr}$ which on passing through a BSC with crossover probability p_0 will have arbitrarily small probability of error.

Since $Rd \leq h(p_0)$ and $R = \frac{K}{N}$, we have

$$|G| = 2^{N-K} \leq 2^{N\left[1-\frac{h(p_0)}{d}\right]} \leq 2^{N[1-h(p_0)]}$$

Thus λ can be made arbitrarily small, and we have that $\Delta \geq d - \epsilon$, and the proof is complete. □

Direct half of main theorem

Direct half of main theorem

We show that any pair (R, d) satisfying

$$0 \leq R \leq C_M \tag{6a}$$
$$0 \leq d \leq H_S \tag{6b}$$
$$Rd \leq H_S \Gamma(R) \tag{6c}$$

is achievable, i.e. $\overline{\mathcal{R}} \subseteq \mathcal{R}$.

Some preliminaries: Typical sequences

Some preliminaries: Typical sequences

Let $\mathcal{X} = \{1, 2, \cdots, A\}$. $X^* \sim p_X^*$.

Some preliminaries: Typical sequences

Let $\mathcal{X} = \{1, 2, \cdots, A\}$. $X^* \sim p_X^*$. Define

$$\#(i, \boldsymbol{x}) \triangleq \text{number of occurrences of symbol } i \text{ in } \boldsymbol{x}$$

Some preliminaries: Typical sequences

Let $\mathcal{X} = \{1, 2, \cdots, A\}$. $X^* \sim p_X^*$. Define

$$\#(i, \boldsymbol{x}) \triangleq \text{number of occurrences of symbol } i \text{ in } \boldsymbol{x}$$

Define the set of typical sequences as

$$T^*(N) = \left\{ \boldsymbol{x} \in \mathcal{X}^N : \left| \frac{\#(i, \boldsymbol{x})}{N} - p_X^*(i) \right| \leq N^{-\frac{1}{4}}, 1 \leq i \leq A \right\}$$

Some preliminaries: Typical sequences

Let $\mathcal{X} = \{1, 2, \cdots, A\}$. $X^* \sim p_X^*$. Define

$$\#(i, \boldsymbol{x}) \triangleq \text{number of occurrences of symbol } i \text{ in } \boldsymbol{x}$$

Define the set of typical sequences as

$$T^*(N) = \left\{ \boldsymbol{x} \in \mathcal{X}^N : \left| \frac{\#(i, \boldsymbol{x})}{N} - p_X^*(i) \right| \leq N^{-\frac{1}{4}}, 1 \leq i \leq A \right\}$$

Lemma 4

Let X^N, Z^N correspond to an arbitrary encoder and let X^*, Z^*, T^* correspond to an arbitrary p_X^*. Then

$$\frac{1}{N} I(\boldsymbol{X}^N; \boldsymbol{Z}^N) \leq I(\boldsymbol{X}^*; \boldsymbol{Z}^*) + (\log A)\mathbb{P}\{\boldsymbol{X}^N \notin T^*(N)\} + f_1(N)$$

where $f_1(N) \to 0$ as $N \to \infty$.

Some preliminaries: Ad-hoc encoder-decoder

We now define an ad-hoc encoder-decoder to prove the direct half.

Some preliminaries: Ad-hoc encoder-decoder

We now define an ad-hoc encoder-decoder to prove the direct half.

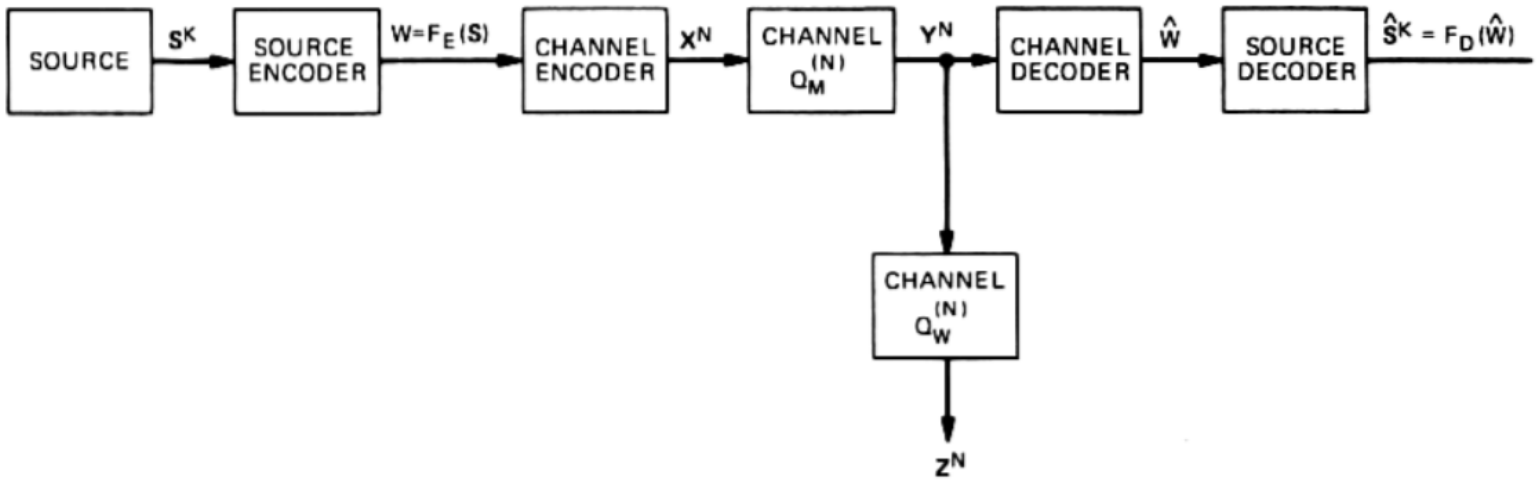

Figure: Ad-hoc encoder-decoder (from [1])

Some preliminaries: Source encoder-decoder

Some preliminaries: Source encoder-decoder

- Let the **source encoder** be the mapping $F_E : \mathcal{S}^K \to \{1, ..., M\}$ where

$$M = 2^{KH_S(1+\delta_K)}, \qquad \delta_K = K^{-\frac{1}{4}}.$$

Some preliminaries: Source encoder-decoder

- Let the **source encoder** be the mapping $F_E : \mathcal{S}^K \to \{1, ..., M\}$ where

$$M = 2^{KH_S(1+\delta_K)}, \qquad \delta_K = K^{-\frac{1}{4}}.$$

- Let the **source decoder** be the mapping $F_D : \{1, ..., M\} \to \mathcal{S}^K$. Let

$$P_{es}^{(K)} = \mathbb{P}\{F_D \circ F_E(S^K) \neq S^K\}$$

be the error probability.

Some preliminaries: Channel encoder

Let $M_1 = M_2 M$ and $\{\boldsymbol{x}_m\}_1^{M_1} \subseteq \mathcal{X}^N$.

Some preliminaries: Channel encoder

Let $M_1 = M_2 M$ and $\{\boldsymbol{x}_m\}_1^{M_1} \subseteq \mathcal{X}^N$.
$\{\boldsymbol{x}_m\}_1^{M_1}$ is partitioned into M subcodes $C_1, ..., C_M$ with cardinality M_2 each.

Some preliminaries: Channel encoder

Let $M_1 = M_2 M$ and $\{\boldsymbol{x}_m\}_1^{M_1} \subseteq \mathcal{X}^N$.
$\{\boldsymbol{x}_m\}_1^{M_1}$ is partitioned into M subcodes $C_1, ..., C_M$ with cardinality M_2 each. Assume

$$C_i = \{\boldsymbol{x}_{(i-1)M_2+1}, ..., \boldsymbol{x}_{iM_2}\}, \qquad 1 \leq i \leq M$$

Some preliminaries: Channel encoder

Let $M_1 = M_2 M$ and $\{\boldsymbol{x}_m\}_1^{M_1} \subseteq \mathcal{X}^N$.
$\{\boldsymbol{x}_m\}_1^{M_1}$ is partitioned into M subcodes $C_1, ..., C_M$ with cardinality M_2 each. Assume

$$C_i = \{\boldsymbol{x}_{(i-1)M_2+1}, ..., \boldsymbol{x}_{iM_2}\}, \qquad 1 \leq i \leq M$$

- When $W = i$, the channel encoder output X^N is a random uniformly chosen member of C_i, i.e,

$$\mathbb{P}\{\boldsymbol{X}^N = \boldsymbol{x}_{(i-1)M_2+j}\} = \frac{q_i}{M_2}, \qquad 1 \leq i \leq M, 1 \leq j \leq M_2$$

where

$$q_i \triangleq \mathbb{P}\{W = F_E(S^K) = i\}$$

Some preliminaries: Channel decoder

Some preliminaries: Channel decoder

- The mapping $G : \mathcal{Y}^N \to \{\boldsymbol{x}_m\}_1^{M_1}$ is stored by the decoder and the error probability is

$$\lambda = \mathbb{P}\{G(\boldsymbol{Y}^N) \neq \boldsymbol{X}^N\}$$

Some preliminaries: Channel decoder

- The mapping $G : \mathcal{Y}^N \to \{\boldsymbol{x}_m\}_1^{M_1}$ is stored by the decoder and the error probability is

$$\lambda = \mathbb{P}\{G(\boldsymbol{Y}^N) \neq \boldsymbol{X}^N\}$$

When the channel output $\boldsymbol{y}$ is such that $G(\boldsymbol{y}) \in C_i$, the decoder outputs $\hat{W} = i$.

Some preliminaries: Channel decoder

- The mapping $G : \mathcal{Y}^N \to \{\boldsymbol{x}_m\}_1^{M_1}$ is stored by the decoder and the error probability is

$$\lambda = \mathbb{P}\{G(\boldsymbol{Y}^N) \neq \boldsymbol{X}^N\}$$

  When the channel output $\boldsymbol{y}$ is such that $G(\boldsymbol{y}) \in C_i$, the decoder outputs $\hat{W} = i$.

- The source decoder outputs $\hat{\boldsymbol{S}}^K = F_D(\hat{\boldsymbol{W}})$ where F_D is chosen as mentioned earlier.

Some preliminaries

Observe that each C_i can be considered a code for $Q_{MW}^{(N)}$ with M_2 code words uniformly distributed.

Some preliminaries

Observe that each C_i can be considered a code for $Q_{MW}^{(N)}$ with M_2 code words uniformly distributed. If λ_i is the error probability for C_i with an optimal decoder, let

$$\overline{\lambda} = \sum_{i=1}^{M} q_i \lambda_i$$

Some preliminaries

Observe that each C_i can be considered a code for $Q_{MW}^{(N)}$ with M_2 code words uniformly distributed. If λ_i is the error probability for C_i with an optimal decoder, let

$$\overline{\lambda} = \sum_{i=1}^{M} q_i \lambda_i$$

Lemma 5

For the ad-hoc encoder-decoder defined above

$$I(\boldsymbol{X}^N; \boldsymbol{Z}^N | \boldsymbol{S}^K) \geq \log M_2 - h(\overline{\lambda}) - \overline{\lambda} \log M_2$$

Some preliminaries

Combining all the above preliminary results,

Some preliminaries

Combining all the above preliminary results,

Corollary 6

Let p_X^* be an arbitrary probability distribution on $\mathcal{X}$ and $X^*, Y^*, Z^*, T_X^*(N)$ be as defined earlier. $\boldsymbol{S}^K, \boldsymbol{X}^N, \boldsymbol{Y}^N, \boldsymbol{Z}^N$ correspond to the ad-hoc encoder-decoder. Then

$$P_e \leq P_{es}^{(K)} + \lambda \tag{7a}$$

$$\frac{K}{N}\Delta \geq \frac{K}{N}H_S + \frac{1}{N}\log M_2 - I(X^*;Z^*) - \frac{h(\overline{\lambda})}{N} - \frac{\overline{\lambda}\log M_2}{N} - (\log A)\mathbb{P}\{\boldsymbol{X}^N \notin T_X^*(N)\} - f_1(N) \tag{7b}$$

where $f_1(N) \to 0$ as $N \to \infty$.

Proof: Setting the parameters of the ad-hoc scheme

Proof: Setting the parameters of the ad-hoc scheme

Choose K, N such that

$$\frac{K}{N} = \frac{R}{H_s}$$

Proof: Setting the parameters of the ad-hoc scheme

Choose K, N such that

$$\frac{K}{N} = \frac{R}{H_s}$$

Note that this implies $\frac{H_S K}{N} \geq R - \epsilon$.

Proof: Setting the parameters of the ad-hoc scheme

Choose K, N such that

$$\frac{K}{N} = \frac{R}{H_s}$$

Note that this implies $\frac{H_S K}{N} \geq R - \epsilon$.

Let $p_X^* \in \mathcal{P}(R)$ be such that it achieves $\Gamma(R)$.

Proof: Setting the parameters of the ad-hoc scheme

Choose K, N such that

$$\frac{K}{N} = \frac{R}{H_s}$$

Note that this implies $\frac{H_S K}{N} \geq R - \epsilon$.

Let $p_X^* \in \mathcal{P}(R)$ be such that it achieves $\Gamma(R)$. We now construct our ad-hoc encoder-decoder with parameter

$$M_1 = 2^{N\left[I(\boldsymbol{X}^*;\boldsymbol{Y}^*) - \frac{\epsilon R}{2H_S}\right]}$$

Proof: Setting the parameters of the ad-hoc scheme

Choose K, N such that

$$\frac{K}{N} = \frac{R}{H_s}$$

Note that this implies $\frac{H_S K}{N} \geq R - \epsilon$.

Let $p_X^* \in \mathcal{P}(R)$ be such that it achieves $\Gamma(R)$. We now construct our ad-hoc encoder-decoder with parameter

$$M_1 = 2^{N\left[I(\boldsymbol{X}^*;\boldsymbol{Y}^*) - \frac{\epsilon R}{2H_S}\right]}$$

and M_2 follows as

$$M_2 = \frac{M_1}{M} = 2^{N\left[I(\boldsymbol{X}^*;\boldsymbol{Y}^*) - \frac{\epsilon R}{2H_S} - \frac{K}{N}H_S - \frac{K}{N}H_S\delta_K\right]}$$

Proof

Substituting for M_2 in (7b) yields

$$\frac{R\Delta}{H_S} \geq \Gamma(R) - f_2(N)$$

where

$$f_2(N) = \frac{\epsilon R}{2H_S} + \frac{h(\overline{\lambda})}{N} + \frac{\overline{\lambda} \log M_2}{N} + (\log A)\mathbb{P}\{\boldsymbol{X}^N \notin \boldsymbol{T}^*(N)\} + f_1(N)$$

Proof

Substituting for M_2 in (7b) yields

$$\frac{R\Delta}{H_S} \geq \Gamma(R) - f_2(N)$$

where

$$f_2(N) = \frac{\epsilon R}{2H_S} + \frac{h(\overline{\lambda})}{N} + \frac{\overline{\lambda} \log M_2}{N} + (\log A)\mathbb{P}\{\boldsymbol{X}^N \notin \boldsymbol{T}^*(N)\} + f_1(N)$$

Note that we have

$$Rd \leq H_S\Gamma(R)$$

Hence it is enough to bound $f_2(N)$.

Proof

Substituting for M_2 in (7b) yields

$$\frac{R\Delta}{H_S} \geq \Gamma(R) - f_2(N)$$

where

$$f_2(N) = \frac{\epsilon R}{2H_S} + \frac{h(\bar{\lambda})}{N} + \frac{\bar{\lambda}\log M_2}{N} + (\log A)\mathbb{P}\{\boldsymbol{X}^N \notin \boldsymbol{T}^*(N)\} + f_1(N)$$

Similarly, we have

$$P_e \leq P_{es}^{(K)} + \lambda$$

Hence it is enough to bound λ.

Proof: Choosing $\{\boldsymbol{x}_m\}_1^{M_1}$

The following lemma gives us the existence of a $\{\boldsymbol{x}_m\}_1^{M_1}$ such that λ and $f_2(N)$ are small.

Proof: Choosing $\{\boldsymbol{x}_m\}_1^{M_1}$

The following lemma gives us the existence of a $\{\boldsymbol{x}_m\}_1^{M_1}$ such that λ and $f_2(N)$ are small.

Lemma 7

With p_X^*, M_1, M_2 as given above, there exists for arbitrary N a set $\{\boldsymbol{x}_m\}_1^{M_1}$ such that

$$f_3(N) \geq \begin{cases} \mathbb{P}\{\boldsymbol{X}^N \notin T^*(N)\}, \\ \lambda, \\ \overline{\lambda} \end{cases}$$

where $f_3(N) \to 0$ as $N \to \infty$.

Proof

Thus, for a sufficiently large N we have $P_e \leq \epsilon$.

Proof

Thus, for a sufficiently large N we have $P_e \leq \epsilon$.
Similarly, for a sufficiently large N we have $f_2(N) \leq \frac{\epsilon R}{2H_S}$.

Proof

Thus, for a sufficiently large N we have $P_e \leq \epsilon$.
Similarly, for a sufficiently large N we have $f_2(N) \leq \frac{\epsilon R}{2H_S}$. Thus, we get

$$\Delta \geq \frac{H_S \Gamma(R)}{R} - \epsilon \geq d - \epsilon$$

Proof

Thus, for a sufficiently large N we have $P_e \leq \epsilon$.
Similarly, for a sufficiently large N we have $f_2(N) \leq \frac{\epsilon R}{2H_S}$. Thus, we get

$$\Delta \geq \frac{H_S \Gamma(R)}{R} - \epsilon \geq d - \epsilon$$

Thus, any $(R, d) \in \overline{\mathcal{R}}$ is achievable. □

Converse of main theorem

We now show that the family of achievable rates $\mathcal{R}$ is contained in $\overline{\mathcal{R}}$, i.e. every achievable (R, d) satisfies

$$0 \leq R \leq C_M, \qquad 0 \leq d \leq H_S, \qquad Rd \leq H_S\Gamma(R) \tag{8}$$

Converse of main theorem

We now show that the family of achievable rates $\mathcal{R}$ is contained in $\overline{\mathcal{R}}$, i.e. every achievable (R, d) satisfies

$$0 \leq R \leq C_M, \qquad 0 \leq d \leq H_S, \qquad Rd \leq H_S\Gamma(R) \tag{8}$$

The first two parts are immediate

Converse of main theorem

We now show that the family of achievable rates $\mathcal{R}$ is contained in $\overline{\mathcal{R}}$, i.e. every achievable (R, d) satisfies

$$0 \leq R \leq C_M, \qquad 0 \leq d \leq H_S, \qquad Rd \leq H_S \Gamma(R) \tag{8}$$

The first two parts are immediate; $0 \leq R \leq C_M$ follows from the ordinary converse to the coding theorem

Converse of main theorem

We now show that the family of achievable rates $\mathcal{R}$ is contained in $\overline{\mathcal{R}}$, i.e. every achievable (R, d) satisfies

$$0 \leq R \leq C_M, \qquad 0 \leq d \leq H_S, \qquad Rd \leq H_S \Gamma(R) \tag{8}$$

The first two parts are immediate; $0 \leq R \leq C_M$ follows from the ordinary converse to the coding theorem and $0 \leq d \leq H_S$ follows from

$$\Delta = \frac{1}{K} H(\boldsymbol{S}^K | \boldsymbol{Z}^N) \leq \frac{1}{K} H(\boldsymbol{S}^K) = H_S \tag{9}$$

Converse of main theorem

We now show that the family of achievable rates $\mathcal{R}$ is contained in $\overline{\mathcal{R}}$, i.e. every achievable (R, d) satisfies

$$0 \leq R \leq C_M, \qquad 0 \leq d \leq H_S, \qquad Rd \leq H_S \Gamma(R) \tag{8}$$

The first two parts are immediate; $0 \leq R \leq C_M$ follows from the ordinary converse to the coding theorem and $0 \leq d \leq H_S$ follows from

$$\Delta = \frac{1}{K} H(\boldsymbol{S}^K | \boldsymbol{Z}^N) \leq \frac{1}{K} H(\boldsymbol{S}^K) = H_S \tag{9}$$

We now prove the third part.

Some preliminaries

We require three properties of $\Gamma(R)$, namely that $\Gamma(R)$ is

Some preliminaries

We require three properties of $\Gamma(R)$, namely that $\Gamma(R)$ is continuous

Some preliminaries

We require three properties of $\Gamma(R)$, namely that $\Gamma(R)$ is continuous, concave

Some preliminaries

We require three properties of $\Gamma(R)$, namely that $\Gamma(R)$ is continuous, concave, and monotonically non-increasing in R.

Some preliminaries

We require three properties of $\Gamma(R)$, namely that $\Gamma(R)$ is continuous, concave, and monotonically non-increasing in R.

We also use the following notations,

- $\delta(P_e) = h(P_e) + P_e \log |\mathcal{S}|$

Some preliminaries

We require three properties of $\Gamma(R)$, namely that $\Gamma(R)$ is continuous, concave, and monotonically non-increasing in R.

We also use the following notations,

- $\delta(P_e) = h(P_e) + P_e \log |\mathcal{S}|$
- For $\boldsymbol{y} \in \mathcal{Y}^{n-1}$, $\alpha_n(\boldsymbol{y}) = I(X_n; Y_n | \boldsymbol{Y}^{n-1} = \boldsymbol{y})$

Some preliminaries

We require three properties of $\Gamma(R)$, namely that $\Gamma(R)$ is continuous, concave, and monotonically non-increasing in R.

We also use the following notations,

- $\delta(P_e) = h(P_e) + P_e \log |\mathcal{S}|$
- For $\boldsymbol{y} \in \mathcal{Y}^{n-1}$, $\alpha_n(\boldsymbol{y}) = I(X_n; Y_n | \boldsymbol{Y}^{n-1} = \boldsymbol{y})$
- For $x \in \mathcal{X}$, $p_{n,\boldsymbol{y}}(x) = \Pr\{X_n = x | \boldsymbol{Y}^{n-1} = \boldsymbol{y}\}$

Some preliminaries

We require three properties of $\Gamma(R)$, namely that $\Gamma(R)$ is continuous, concave, and monotonically non-increasing in R.

We also use the following notations,

- $\delta(P_e) = h(P_e) + P_e \log |\mathcal{S}|$
- For $\boldsymbol{y} \in \mathcal{Y}^{n-1}$, $\alpha_n(\boldsymbol{y}) = I(X_n; Y_n | \boldsymbol{Y}^{n-1} = \boldsymbol{y})$
- For $x \in \mathcal{X}$, $p_{n,\boldsymbol{y}}(x) = \Pr\{X_n = x | \boldsymbol{Y}^{n-1} = \boldsymbol{y}\}$

It is clear that $p_{n,\boldsymbol{y}} \in \mathcal{P}[\alpha_n(\boldsymbol{y})]$

Some preliminaries

We require three properties of $\Gamma(R)$, namely that $\Gamma(R)$ is continuous, concave, and monotonically non-increasing in R.

We also use the following notations,

- $\delta(P_e) = h(P_e) + P_e \log |\mathcal{S}|$
- For $\boldsymbol{y} \in \mathcal{Y}^{n-1}$, $\alpha_n(\boldsymbol{y}) = I(X_n; Y_n | \boldsymbol{Y}^{n-1} = \boldsymbol{y})$
- For $x \in \mathcal{X}$, $p_{n,\boldsymbol{y}}(x) = \Pr\{X_n = x | \boldsymbol{Y}^{n-1} = \boldsymbol{y}\}$

It is clear that $p_{n,\boldsymbol{y}} \in \mathcal{P}[\alpha_n(\boldsymbol{y})]$ and hence

$$\Gamma[\alpha_n(\boldsymbol{y})] \geq I(X_n; Y_n | Z_n, \boldsymbol{Y}^{n-1} = \boldsymbol{y})$$

Proof

We now make use of the following lemma to relate Δ with H_S.

Proof

We now make use of the following lemma to relate Δ with H_S.

Lemma 8

For any encoder-decoder (N, K, Δ, P_e),

$$\frac{K}{N}[\Delta - \delta(P_e)] \leq \frac{1}{N}\sum_{n=1}^{N} I(X_n; Y_n | Z_n, \boldsymbol{Y}^{n-1}), \tag{10a}$$

$$\frac{K}{N}[H_S - \delta(P_e)] \leq \frac{1}{N}\sum_{n=1}^{N} I(X_n; Y_n | \boldsymbol{Y}^{n-1}) \tag{10b}$$

Proof

$$\frac{K}{N}[\Delta - \delta(P_e)] \leq \frac{1}{N}\sum_{n=1}^{N} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1})$$

Proof

$$\frac{K}{N}[\Delta - \delta(P_e)] \leq \frac{1}{N}\sum_{n=1}^{N} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1})$$

$$= \frac{1}{N}\sum_{n=1}^{N}\sum_{\boldsymbol{y}\in\mathcal{Y}^{n-1}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1} = \boldsymbol{y})$$

Proof

$$\frac{K}{N}[\Delta - \delta(P_e)] \leq \frac{1}{N}\sum_{n=1}^{N} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1})$$

$$= \frac{1}{N}\sum_{n=1}^{N}\sum_{\boldsymbol{y}\in\mathcal{Y}^{n-1}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1} = \boldsymbol{y})$$

$$\leq \frac{1}{N}\sum_{n}\sum_{\boldsymbol{y}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\}\Gamma[\alpha_n(\boldsymbol{y})]$$

Proof

$$\frac{K}{N}[\Delta - \delta(P_e)] \leq \frac{1}{N}\sum_{n=1}^{N} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1})$$

$$= \frac{1}{N}\sum_{n=1}^{N}\sum_{\boldsymbol{y}\in\mathcal{Y}^{n-1}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1} = \boldsymbol{y})$$

$$\leq \frac{1}{N}\sum_{n}\sum_{\boldsymbol{y}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\}\Gamma[\alpha_n(\boldsymbol{y})]$$

$$\leq \Gamma\left[\frac{1}{N}\sum_{n}\sum_{\boldsymbol{y}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\}\alpha_n(\boldsymbol{y})\right]$$

Proof

$$\frac{K}{N}[\Delta - \delta(P_e)] \leq \frac{1}{N}\sum_{n=1}^{N} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1})$$

$$= \frac{1}{N}\sum_{n=1}^{N}\sum_{\boldsymbol{y}\in\mathcal{Y}^{n-1}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1} = \boldsymbol{y})$$

$$\leq \frac{1}{N}\sum_{n}\sum_{\boldsymbol{y}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\}\Gamma[\alpha_n(\boldsymbol{y})]$$

$$\leq \Gamma\left[\frac{1}{N}\sum_{n}\sum_{\boldsymbol{y}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\}\alpha_n(\boldsymbol{y})\right]$$

$$= \Gamma\left[\sum_{n=1}^{N} I(X_n; Y_n|\boldsymbol{Y}^{n-1})\right]$$

Proof

$$\frac{K}{N}[\Delta - \delta(P_e)] \leq \frac{1}{N}\sum_{n=1}^{N} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1})$$

$$= \frac{1}{N}\sum_{n=1}^{N}\sum_{\boldsymbol{y}\in\mathcal{Y}^{n-1}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\} I(X_n; Y_n|Z_n, \boldsymbol{Y}^{n-1} = \boldsymbol{y})$$

$$\leq \frac{1}{N}\sum_{n}\sum_{\boldsymbol{y}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\}\Gamma[\alpha_n(\boldsymbol{y})]$$

$$\leq \Gamma\left[\frac{1}{N}\sum_{n}\sum_{\boldsymbol{y}} \Pr\{\boldsymbol{Y}^{n-1} = \boldsymbol{y}\}\alpha_n(\boldsymbol{y})\right]$$

$$= \Gamma\left[\sum_{n=1}^{N} I(X_n; Y_n|\boldsymbol{Y}^{n-1})\right] \leq \Gamma\left(\frac{K}{N}[H_S - \delta(P_e)]\right)$$

Proof

Hence we have

$$\frac{H_S K}{N}[\Delta - \delta(P_e)] \leq H_S \Gamma\left(\frac{K}{N}[H_S - \delta(P_e)]\right)$$

Proof

Hence we have

$$\frac{H_S K}{N}[\Delta - \delta(P_e)] \leq H_S \Gamma\left(\frac{K}{N}[H_S - \delta(P_e)]\right)$$

Since (R, d) is achievable, we have

$$(R-\epsilon)[(d-\epsilon) - \delta(\epsilon)] \leq H_S \Gamma\left(R - \frac{K}{N}\delta(\epsilon)\right) \text{ for any } \epsilon > 0$$

Proof

Hence we have

$$\frac{H_S K}{N}[\Delta - \delta(P_e)] \leq H_S \Gamma\left(\frac{K}{N}[H_S - \delta(P_e)]\right)$$

Since (R, d) is achievable, we have

$$(R-\epsilon)[(d-\epsilon) - \delta(\epsilon)] \leq H_S \Gamma\left(R - \frac{K}{N}\delta(\epsilon)\right) \text{ for any } \epsilon > 0$$

$$\Longrightarrow \; Rd \leq H_S \Gamma(R) \quad \square$$

References

[1] A. D. Wyner.
The wire-tap channel.
Bell System Technical Journal, 54(8):1355–1387, 1975.

[2] Abbas El Gamal and Young-Han Kim.
Network information theory.
Cambridge University Press, 2018.

[3] Robert Gallager.
Information theory and reliable communication.
John Wiley, 1968.

Thank you