Proof of the Converse

Pang-Chang Lan

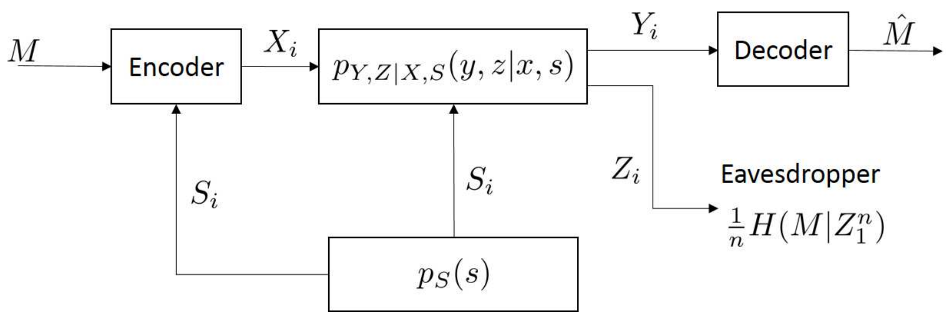

Fig. 1. Discrete memoryless wiretap channel with perfect CSIT but no CSIRE

Theorem 1. *Suppose that the main channel is less noisy than the wiretap channel, i.e., $I(U;Y) \geqslant I(U;Z)$ for all U such that $(U,S) \to (X,S) \to (Y,Z)$ forms a Markov chain. The secrecy capacity of the discrete memoryless wiretap channel with perfect CSIT but no CSIRE, as shown in Fig. 1, is given by*

$$C_s = \max_{p(t),x(t,s)} I(T;Y) - I(T;Z) \tag{1}$$

where T is the auxiliary random variables independent of S and satisfying the Markov relation $T \to (X,S) \to (Y,Z)$, the cadinality bound is $|\mathcal{T}| \leqslant \min\{(|\mathcal{X}|-1)|\mathcal{S}|+1, |\mathcal{Y}|\}$, and $x : \mathcal{T} \times \mathcal{S} \to \mathcal{X}$ is a deterministic function.

I. Achievability Proof

We first prove that the following rate is achievable in general, namely,

$$R_s = \max_{p(u)p(t|u),x(t,s)} I(U;Y) - I(U;Z) \tag{2}$$

where U and T are auxiliary random variables independent of S and satisfying the Markov relation $U \to T \to (X,S) \to (Y,Z)$, the cadinality bound is given by $|\mathcal{U}| \leqslant |\mathcal{T}|$.

Let $\mathcal{M} = \{1,2,\cdots,2^{nR}\}$. We use multicoding and a three-step randomized encoding scheme.

Codebook generation. For each message $m \in \mathcal{M}$, generate a subcodebook $\mathcal{C}(m)$ consisting of $2^{n(\tilde{R}-R)}$ sequences $u^n(l)$ for $l \in [(m-1)2^{n(\tilde{R}-R)}+1 : m2^{n(\tilde{R}-R)}]$ which is randomly and independently generated according to the distribution $\prod_{i=1}^n p_U(u_i)$. This codebook is revealed to all the nodes.

Encoding. To send message $m \in \mathcal{M}$, the encoder uniformly randomly chooses an index L from $\mathcal{C}(m)$. Then it generates $t^n(L)$ by random coding according to the distribution $\prod_{i=1}^n p_{T|U}(t_i|u_i(l))$. The channel input at time i is then given by $x_i = x(t_i(L), s_i)$ where $x(\cdot,\cdot)$ is a deterministic function.

Decoding. Given $Y^n = y^n$, find the unique $\hat{m}$ such that $(u^n(l), y^n) \in \mathcal{T}_\varepsilon^n(U,Y)$ for some $u^n(l) \in \mathcal{C}(\hat{m})$.

Analysis. By code translation, we can obtain an equivalent wiretap channel described by

$$\begin{aligned} p_{Y,Z|T}(y,z|t) &= \sum_{x\in\mathcal{X},s\in\mathcal{S}} P_{Y,Z,X,S|T}(y,z,x,s|t) \\ &= \sum_{x\in\mathcal{X},s\in\mathcal{S}} P_{Y,Z|X,T,S}(y,z|x,t,s)p_{X|T,S}(x|t,s)p_{S|T}(s|t) \\ &= \sum_{x\in\mathcal{X},s\in\mathcal{S}} P_{Y,Z|X}(y,z|x)\delta_K(x=x(t,s))p_S(s) \\ &= \sum_{s\in\mathcal{S}} P_{Y,Z|X}(y,z|x(t,s))p_S(s). \end{aligned}$$

The resulting equivalent DMC without CSI is shown in Figure 2. The rest of the analysis procedure

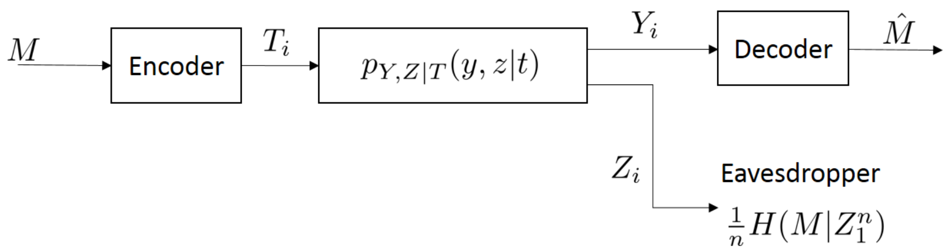

Fig. 2. Equivalent discrete memoryless wiretap channel with no CSI

follows that of the wiretap channel.

Let M be the transmitted message and L be the randomly picked index within codebook $\mathcal{C}(M)$. Let $\mathcal{E}$ be the event of error decoding. Let $\mathcal{E}_1 = \{(U^n(l), Y^n) \notin \mathcal{T}_\varepsilon^n(U,Y)$ for all $U^n(l) \in \mathcal{C}(m)\}$ and $\mathcal{E}_2 = \{(U^n(l), Y^n) \in \mathcal{T}_\varepsilon^n(U,Y)$ for some $l \neq L\}$. Then

$$P(\mathcal{E}) \leqslant P(\mathcal{E}_1 \cup \mathcal{E}_2) \leqslant P(\mathcal{E}_1) + P(\mathcal{E}_2)$$

By LLN, $P(\mathcal{E}_1) \to 0$ as $n \to \infty$. By the packing lemma, $P(\mathcal{E}_2) \to 0$ as $n \to \infty$ if $\tilde{R} < I(U;Y) - \delta(\varepsilon)$. Hence, the error probability diminishes as n goes to infinity.

By following the proof in the Network Information Theory book, the equivocation at the eavesdropper can also be shown to approach 0 as $n \to \infty$ if $\tilde{R} - R \geqslant I(U,Z)$. Hence, the achievable rate follows as (2).

Now, by the less noisy property, we have

$$\begin{aligned} R_s &= \max_{p(u)p(t|u),f(t,s)} I(U;Y) - I(U;Z) \\ &= \max_{p(u)p(t|u),f(t,s)} I(T;Y) - I(T;Z) - (I(T;Y|U) - I(T;Z|U)) \\ &= \max_{p(t),x(t,s)} I(T;Y) - I(T;Z) \end{aligned} \tag{3}$$

where the last equality follows since $I(T;Y|U) - I(T;Z|U) \geqslant 0$ and by letting $U = T$, it is made 0. Hence, the achievability follows.

II. Converse Proof

To prove the converse, we first assume that the wiretap channel is less noisy, *i.e.*, $I(U;Y) \geqslant I(U;Z)$ for all $P_U P_{X|U}$ such that $U \to X \to (Y,Z)$.

Suppose there exists a code such that $R_e \leqslant \varepsilon_n$ where $\varepsilon_n \to 0$ as $n \to \infty$. Note that by Fano's inequality, $H(M|Y^n) \leqslant n\varepsilon_n'$ where $\varepsilon_n' \to 0$ as $n \to \infty$. We have

$$\begin{aligned}
nR_s &= H(M) - H(M|Y^n) + H(M|Y^n) \\
&\leqslant I(M;Y^n) + n\varepsilon_n' \\
&= I(M;Y^n) - I(M;Z^n) + nR_e + n\varepsilon_n' \\
&\leqslant I(M;Y^n) - I(M;Z^n) + n(\varepsilon_n' + \varepsilon_n) \\
&= \sum_{i=1}^{n} \left[I(M;Y_i|Y_1^{i-1}) - I(M;Z_i|Z_{i+1}^n) \right] + n(\varepsilon_n' + \varepsilon_n) \\
&= \sum_{i=1}^{n} \left[I(M,Z_{i+1}^n;Y_i|Y_1^{i-1}) - I(Z_{i+1}^n;Y_i|Y_1^{i-1},M) - I(M,Y_1^{i-1};Z_i|Z_{i+1}^n) + I(Y_1^{i-1};Z_i|Z_{i+1}^n,M) \right] \\
&\quad + n(\varepsilon_n' + \varepsilon_n) \\
&= \sum_{i=1}^{n} \left[I(M,Z_{i+1}^n;Y_i|Y_1^{i-1}) - I(M,Y_1^{i-1};Z_i|Z_{i+1}^n) \right] + n(\varepsilon_n' + \varepsilon_n) && (4) \\
&= \sum_{i=1}^{n} \left[I(M;Y_i|Y_1^{i-1},Z_{i+1}^n) + I(Z_{i+1}^n;Y_i|Y_1^{i-1}) - I(M;Z_i|Z_{i+1}^n,Y_1^{i-1}) - I(Y_1^{i-1};Z_i|Z_{i+1}^n) \right] \\
&\quad + n(\varepsilon_n' + \varepsilon_n) \\
&= \sum_{i=1}^{n} \left[I(M;Y_i|Y_1^{i-1},Z_{i+1}^n) - I(M;Z_i|Y_1^{i-1},Z_{i+1}^n) \right] + n(\varepsilon_n' + \varepsilon_n) && (5) \\
&= \sum_{i=1}^{n} \left[I(\tilde{U}_i;Y_i|V_i) - I(\tilde{U}_i;Z_i|V_i) \right] + n(\varepsilon_n' + \varepsilon_n) && (6)
\end{aligned}$$

where (4) and (5) are due to Csiszar sum identity, and in (6) auxiliary random variables $V_i = (Y_1^{i-1}, Z_{i+1}^n)$ and $\tilde{U}_i = (M, V_i)$ are introduced such that $(V_i, S) \to (\tilde{U}_i, S) \to (X_i, S_i) \to (Y_i, Z_i)$ forms a Markov chain. Note that here V_i and $\tilde{U}_i$ are dependent on S_i.

By the less noisy property, we have

$$I(S_1^{i-1};Y_i|\tilde{U}_i) - I(S_1^{i-1};Z_i|\tilde{U}_i) \geqslant 0. \tag{7}$$

Hence,

$$\begin{aligned}
(6) &\leqslant \sum_{i=1}^{n} \left[I(\tilde{U}_i;Y_i|V_i) + I(S_1^{i-1};Y_i|\tilde{U}_i) - I(\tilde{U}_i;Z_i|V_i) - I(S_1^{i-1};Z_i|\tilde{U}_i) \right] \\
&\quad + n(\varepsilon_n' + \varepsilon_n) \\
&= \sum_{i=1}^{n} \left[I(\tilde{U}_i,S_1^{i-1};Y_i|V_i) - I(\tilde{U}_i,S_1^{i-1};Z_i|V_i) \right] + n(\varepsilon_n' + \varepsilon_n) \\
&= \sum_{i=1}^{n} \left[I(U_i;Y_i|V_i) - I(U_i;Z_i|V_i) \right] + n(\varepsilon_n' + \varepsilon_n) && (8) \\
&= \left[I(U_Q;Y_Q|V_Q,Q) - I(U_Q;Z_Q|V_Q,Q) \right] + n(\varepsilon_n' + \varepsilon_n) && (9) \\
&= \left[I(U;Y|V) - I(U;Z|V) \right] + n(\varepsilon_n' + \varepsilon_n) && (10) \\
&\leqslant \max_{v} \left[I(U;Y|V=v) - I(U;Z|V=v) \right] + n(\varepsilon_n' + \varepsilon_n) \\
&= n \max_{p(u|s)p(x|u,s)} \left[I(U;Y) - I(U;Z) \right] + n(\varepsilon_n' + \varepsilon_n) \\
&= n \max_{p(u|s)p(t|u),x(t,s)} \left[I(U;Y) - I(U;Z) \right] + n(\varepsilon_n' + \varepsilon_n) && (11)
\end{aligned}$$

where (8) follows by letting $U_i = (\tilde{U}_i, S_1^{i-1})$, in (9) a time-sharing random variable Q is introduced, (10) holds by letting $V = (V_Q, Q)$, $U = (U_Q, Q)$, $Y = Y_Q$, and $Z = Z_Q$, and (11) is because of the Markov chain $U \to (X, S) \to (Y, Z)$ and the functional representation lemma. Note that the Markov chain turns into $(U, S) \to (T, S) \to (Y, Z)$.

Now, by the less noisy property, we have

$$\begin{aligned} I(U;Y) - I(U;Z) &= I(U,T;Y) - I(U,T;Z) - (I(T;Y|U) - I(T;Z|U)) \\ &\leqslant I(T;Y) - I(T;Z) \end{aligned} \tag{12}$$

by setting $U = T$.

Lemma 1 (functional representation)**.** *Let* $(Y, S, X) \sim p(y, s, x)$. *Then* X *can be represented as a function of* (S, U) *for some random variable* U *of cardinality* $|\mathcal{U}| \leqslant |\mathcal{S}|(|\mathcal{X}| - 1) + 1$ *such that* U *is independent of* S *and* $U \to (X, S) \to Y$ *forms a Markov chain.*