II

Transmission, Identification and Common Randomness Capacities for Wire-Tape Channels with Secure Feedback from the Decoder

R. Ahlswede and N. Cai

Abstract. We analyze wire-tape channels with secure feedback from the legitimate receiver. We present a lower bound on the transmission capacity (Theorem 1), which we conjecture to be tight and which is proved to be tight (Corollary 1) for Wyner's original (degraded) wire-tape channel and also for the reversely degraded wire-tape channel for which the legitimate receiver gets a degraded version from the enemy (Corollary 2).

Somewhat surprisingly we completely determine the capacities of secure common randomness (Theorem 2) and secure identification (Theorem 3 and Corollary 3). Unlike for the DMC, these quantities are different here, because identification is linked to non-secure common randomness.

1 Introduction

The main results are mentioned in the abstract.

After having given standard concepts in Section 2 and known results and techniques for the wire-tape channel in Section 3, we state and prove Theorem 1 in Section 4. Our code construction relies upon a lemma for balanced coloring from [2], which has proved already useful for secrecy problems in [3].

The transmission capacities for the two kinds of degraded wire-tape channels are derived in Section 5. Particularly interesting is an example of a reversely degraded channel, where the channel $W_1' : \mathcal{X} \to \mathcal{Z}$ for the wiretapper is noiseless (for instance with binary alphabets) and the channel $W_2' : \mathcal{Z} \to \mathcal{Y}$ for the legal receiver is a noisy binary symmetric channel with crossover probability $p \in (0, 1/2)$. Here the wiretapper is in a better position than the legal user and therefore the capacity is zero, if there is no feedback. However, by our Corollary the capacity is positive, because the feedback serves as a secure key shared by sender and receiver.

In Section 6 a discussion based on the construction for transmission in Section 4 and known results and constructions for identification [8], [9], [15], and common randomness [9], [7], and all other references builds up the intuition for our solutions of the capacity problems for common randomness and identification in Section 7 and 8.

2 Notation and Definitions

Throughout the paper $\mathcal{U}$, $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ are finite sets and their elements are written as corresponding lower letters e.g. u, x, y, and z. The letters U, X, Y,

R. Ahlswede et al. (Eds.): Information Transfer and Combinatorics, LNCS 4123, pp. 258–275, 2006.

Z etc. will be used for random variables with values in the corresponding sets, $\mathcal{U}, \ldots$. $\mathcal{T}^n_X$, $\mathcal{T}^n_{Y|X}(x^n)$, $\mathcal{T}^n_{XYZ}$, etc. are sets of n-typical, conditional typical and joint typical sequences, and sets of δ-typical, conditional typical and joint typical sequences are written as $\mathcal{T}^n_{X,\delta}$, $\mathcal{T}^n_{Y|X,\delta}(x^n)$, $\mathcal{T}^n_{XYZ,\delta}$, etc.

Then a (discrete memoryless) wire-tape channel is specified by a stochastic matrix $W : \mathcal{X} \to \mathcal{Y} \times \mathcal{Z}$, where $\mathcal{X}$ serves as input alphabet, $\mathcal{Y}$ as output alphabet of the legal receiver and $\mathcal{Z}$ as output alphabet of a wiretapper. The channel works as follows: the legal receiver receives an output sequence y^n and the wiretapper receives an output sequence z^n with probability

$$W^n(y^n z^n|x^n) = \prod_{t=1}^{n} W(y_t z_t|x_t).$$

In the case of transmission the sender's goal is to send to the receiver a message U uniformly distributed on an as large as possible set of messages with vanishing probability of error such that the wiretapper almost knows nothing about the message. Randomization at the sender side is allowed. The wiretapper, who knows the coding scheme, but not the message, tries to learn about the message as much as possible.

For given $\lambda, \mu > 0$, a (λ, μ)-code of length n with a set of messages $\mathcal{M}$ is a system $\{(Q_m : \mathcal{D}_m) : m \in \mathcal{M}\}$, where the Q_m's for $m \in \mathcal{M}$ are probability distributions on $\mathcal{X}^n$, and the $\mathcal{D}_m$'s are pairwise disjoint subsets of $\mathcal{Y}^n$, such that

$$|\mathcal{M}|^{-1} \sum_{m \in \mathcal{M}} \sum_{x^n \in \mathcal{X}^n} Q_m(x^n) \sum_{z^n \in \mathcal{Z}^n} W^n(\mathcal{D}_m, z^n|x^n) > 1 - \lambda, \tag{2.1}$$

and

$$\frac{1}{n} I(U; Z^n) < \mu, \tag{2.2}$$

if Z^n is the random output sequence generated by the message U through the channel. The transmission capacity of the wire-tape channel is the maximal non-negative number C_{wt} such that for $\mathcal{M}$, $\lambda, \mu, \varepsilon > 0$ and all sufficiently large length n, there exists a (λ, μ)-code with rate $\frac{1}{n} \log |\mathcal{M}| > C_{wt} - \varepsilon$. The security criterion (2.2) is strengthened in [11] to

$$I(U; Z) < \mu. \tag{2.3}$$

In the current paper we assume the output y_t at time t is completely and immediately feedback to the sender via a secure noiseless channel such that the wiretapper has no knowledge about the feedback (except his own output z^n). Then for $\lambda, \mu > 0$, a (λ, μ)-code of length n for the wire-tape channel with secure feedback is a system $\{(Q, \mathcal{D}_m) : m \in \mathcal{M}\}$ where $\mathcal{D}_m$, $m \in \mathcal{M}$, are pairwise disjoint subsets of $\mathcal{Y}^n$ as before and Q is a stochastic matrix $Q : \mathcal{M} \times \mathcal{Y}^{n-1} \to \mathcal{X}^n$ with

$$Q(x^n|m, y^{n-1}) = \prod_{t=1}^{n} Q(x_t|m, y^{t-1})$$

for $x^n \in \mathcal{X}$, $y^{n-1} \in \mathcal{Y}^{n-1}$, and $m \in \mathcal{M}$, such that

$$|\mathcal{M}|^{-1} \sum_{m\in\mathcal{M}} \sum_{x^n\in\mathcal{X}} \sum_{z^n\in\mathcal{Z}^n} \sum_{y^n\in\mathcal{D}_m} Q(x^n|m, y^{n-1})W^n(y^n, z^n|x^n) > 1-\lambda \quad (2.4)$$

and (2.2) holds. The transmission capacity is defined analogously and denoted by C_{wtf}. In Theorem 1 in Section 4 we shall prove our (direct) coding theorem with the stronger security criterion (2.3).

3 Previous and Auxiliary Results

Our code construction is based on a coding lemma and a code for wire-tape channel without feedback. A balanced coloring lemma originally was introduced by R. Ahlswede [2] and we need its following variation.

Lemma 1. *For all $\delta, \eta > 0$, sufficiently large n, all n-type P_{XY} and all $x^n \in \mathcal{T}^n_X$, there exists a γ-coloring $c : \mathcal{T}^n_{Y|X}(x^n) \to \{0, 1, 2, \ldots, \gamma-1\}$ of $\mathcal{T}^n_{Y|X}(x^n)$ such that for all joint n-type P_{XYZ} with marginal distribution P_{XY} and $\gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n, z^n)| > 2n\eta$, $x^n, z^n \in \mathcal{T}^n_{XZ}$,*

$$|c^{-1}(k)| \le \gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n, z^n)|(1+\delta), \quad (3.1)$$

for $k = 0, 1, \ldots, \gamma-1$, where c^{-1} is the inverse image of c.

Proof: Let us randomly and independently color $y^n \in \mathcal{T}^n_{Y|X}(x^n)$ with γ colors and uniform distribution over $\mathcal{T}^n_{Y|X}(x^n)$. Let for $k = 0, 1, \ldots, \gamma-1$

$$S_k(y^n) = \begin{cases} 1 & \text{if } y^n \text{ is colored by } k \\ 0 & \text{else.} \end{cases} \quad (3.2)$$

Then for a joint type P_{XZY} and $z^n \in \mathcal{T}^n_{Z|X}(x^n)$, by Chernoff bound,

$$\begin{aligned}
&Pr\left\{\sum_{y^n\in\mathcal{T}^n_{Y|XZ}(x^n,z^n)} S_k(y^n) > \gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n, y^n)|(1+\delta)\right\} \\
&\le e^{-\frac{\delta}{2}\gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n,z^n)|(1+\delta)} \prod_{y^n\in\mathcal{T}^n_{Y|XZ}(x^n,z^n)} E\, e^{\frac{\delta}{2}S_k(y^n)} \\
&= c^{-\frac{\delta}{2}\gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n,z^n)|(1+\delta)} \left[(1-\gamma^{-1}) + \gamma^{-1}e^{\frac{\delta}{2}}\right]^{|\mathcal{T}^n_{Y|XZ}(x^n,z^n)|} \\
&= e^{-\frac{\delta}{2}\gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n,z^n)|(1+\delta)} \left[1 + (e^{\frac{\delta}{2}} - 1)\gamma^{-1}\right]^{|\mathcal{T}^n_{Y|XZ}(x^n,z^n)|} \\
&\le e^{-\frac{\delta}{2}\gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n,z^n)|(1+\delta)} \left[1 + \gamma^{-1}(\frac{\delta}{2} + \frac{\delta^2}{8}e)\right]^{|\mathcal{T}^n_{Y|XZ}(x^n,z^n)|}
\end{aligned}$$

$$\leq \exp_e \left\{ -\frac{\delta}{2}\gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n,z^n)|(1+\delta) + \gamma^{-1}\left(\frac{\delta}{2}+\frac{\delta^2}{8}e\right)|\mathcal{T}^n_{Y|XZ}(x^n,z^n)| \right\}$$

$$= \exp_e \left\{ -\frac{\delta}{2}\gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n,z^n)|\left(1-\frac{e}{4}\right)\delta \right\}$$

$$\leq e^{-\frac{e\delta^2}{24}\gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n,z^n))|}$$

$$\leq e^{-\frac{e\delta^2}{24}2^{n\eta}}, \tag{3.3}$$

if $\gamma^{-1}|\mathcal{T}^n_{Y|XZ}(x^n,z^n)| > 2^{n\eta}$ and $\frac{\delta}{2} \leq 1$.
Here, to obtain the 2nd and 3rd inequalities, we use for $x \in [0,1]$ the inequalities $e^x \leq 1 + x + \frac{e}{2}x^2$ and $1 + x \leq e^x$ respectively.

(3.1) follows from (3.3) because the numbers of sequences z^n and n-types increase exponentially and polynomially respectively as the length increases. □

To prove (the direct part of) the coding theorem for the wire-tape channel (without feedback) [11] Csiszár and Körner used a special code, Ahlswede's partition of typical input sequences into sets of code words, obtained by iterative maximal coding [1]. An easier proof appears in [2], part II, as consequence of the "link". We shall use its following improvement obtained with a Balanced Coloring Lemma of [2] and presented in [10].

For a given wire-tape channel such that for an input random variable X and its output random variables Y and Z for the legal user and wiretapper respectively

$$I(X;Y) - I(X;Z) > 0 \tag{3.4}$$

all $\lambda', \mu' > 0$ $0 < \varepsilon' < I(X;Y) - I(X;Z)$ and sufficiently large n, there exists a set of codewords

$$\{u_{m,\ell} : m = 0,1,2,\ldots,M-1,\ \ell = 0,1,2,\ldots,L-1\}$$

in $\mathcal{T}^n_X$ having the following properties.

$$I(X;Y) - I(X;Z) - \varepsilon' < \frac{1}{n}\log M \leq I(X;Y) - I(X;Z) - \frac{\varepsilon'}{2} \tag{3.5}$$

$$I(X;Z) + \frac{\varepsilon'}{8} \leq \frac{1}{n}\log L < I(X;Z) + \frac{\varepsilon'}{4}. \tag{3.6}$$

For a set of properly chosen decoding sets $\{\mathcal{D}_{m,\ell}\}$,

$$\{(u_{m,\ell}, \mathcal{D}_{m,\ell}) : m = 0,1,2,\ldots,M-1, \ell = 0,1,2,\ldots,L-1\}$$

is a λ-code for the legal user.

Let $V, \tilde{Z}$ be random variables taking values in $\mathcal{M} \times \mathcal{Z}^n$, where $\mathcal{M} = \{0,1,\ldots,M-1\}$, with probability for $(m,z^n) \in \mathcal{M} \times \mathcal{Z}^n$

$$Pr\{V,\tilde{Z}) = (m,z^n)) = \sum_{\ell=0}^{L-1} L^{-1} P^n_{Z|X}(z^n|u_{m,\ell}).$$

Then

$$I(V;\tilde{Z}) < \mu'. \tag{3.7}$$

4 The Coding Theorem for Transmission and Its Proof

Let $\mathcal{Q}$ be the set of quadruples of random variables (U, X, Y, Z) taking values in $\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ for a finite set $\mathcal{U}$ with probability

$$Pr((U, X, Y, Z) = (u, x, y, z)) = P_{UX}(ux)W(yz|x) \tag{4.1}$$

for $(u, x, y, z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$.

Then

Theorem 1. *The capacity of a wire-tape channel with feedback satisfies*

$$C_{wtf} \geq \max_{(U,X,Y,Z)\in\mathcal{Q}} \min[|I(U;Y) - I(U;Z)|^+ + H(Y|U,Z), I(U;Y)]. \tag{4.2}$$

Proof: For a $(U, X, Y, Z) \in \mathcal{Q}$, to show the achievability, one may introduce an auxiliary channel $P_{X|U}$ and construct a code for the channel

$$W'(y, z|u) = \sum_x P_{X|U}(x|u)W(y, z|x).$$

Then it is sufficient to show that $|I(X;Y) - I(X;Z)|^+ + H(Y|XZ)$ is achievable. Let us fix $\lambda, \mu, \varepsilon > 0$ and construct a (λ, μ)-code with rate

$$|I(X;Y) - I(X,Z)|^+ + H(Y|XZ) - \varepsilon. \tag{4.3}$$

To this end, let $\lambda', \mu', \varepsilon'$ be positive small real numbers specified later.

Let $\mathcal{U} = \{u_{m,\ell} : m = 0, 1, 2, \ldots, M-1, \ell = 0, 1, 2, \ldots, L-1\}$ be the codebook if in the previous section for a sufficiently large n (3.4) holds i.e., $I(X;Y) - I(X;Z) > 0$.

In the case that (3.4) does not hold we choose $M = 1$ and take a codebook of an arbitrary λ'-code for the legal user, with rate $I(X;Y) - \varepsilon' < R \triangleq \frac{1}{n}\log L \leq I(X;Y) - \frac{\varepsilon'}{2}$ as our codebook:

$$\mathcal{U} = \{u_{0,\ell} : \ell = 0, 1, 2, \ldots, L-1\}.$$

Our code consists of N blocks of length n and sends a message $(U'_1, U'_2U''_2, \ldots, U'_NU''_N)$ uniformly distributed on $\mathcal{M}' \times (\mathcal{M}' \times \mathcal{M}'')^{N-1}$, where

$$\mathcal{M}' = \{0, 1, 2, \ldots, M-1\}, \mathcal{M}'' = \{0, 1, \ldots, L''-1\}, \tag{4.4}$$

and $L'' = \min\{L, 2^{n(H(Y|XZ)-\frac{3}{4})}\}$.

In particular $M = 1$, $\mathcal{M}'$ is a dummy message set. Then the rate of the messages is

$$R^* = \frac{1}{n}\log M + \frac{1}{n}\log L'' - \frac{1}{nN}\log L'' \geq \frac{1}{n}\log M + \frac{1}{n}\log L'' - \frac{1}{N}\log|\mathcal{Y}|.$$

That is by (3.5), (3.6)

$$R^* \geq \begin{cases} I(X;Y)-I(X;Z)-\varepsilon'+\min\left[I(X;Z)+\frac{\varepsilon'}{8}, H(Y|XZ)-\frac{\varepsilon}{4}\right]-\frac{1}{N}\log|\mathcal{Y}| \\ \qquad \text{if } I(X;Y)-I(X;Z(>0 \\ \min\left[I(X;Y)-\frac{\varepsilon'}{2}, H(Y|XZ)-\frac{\varepsilon}{4}\right]-\frac{1}{N}\log|\mathcal{Y}| \\ \qquad \text{else.} \end{cases} \tag{4.5}$$

By choosing $\varepsilon' < \frac{\varepsilon}{2}$ and $N > 2\varepsilon^{-1}\log|\mathcal{Y}|$ in (4.5) we have

$$R^* > \min[|I(X;Y)-I(X;Z)|^+ + H(Y|XZ), I(X;Y)] - \varepsilon \tag{4.6}$$

our desired rate.

In each block, we use a codebook

$$\mathcal{U} = \{u_{m,\ell} : m = 0,1,\ldots,M-1, \ell = 0,1,2,\ldots,L-1\}$$

defined as above. Suppose the sender wants to send a message $(m'_1, m'_2m''_2, \ldots, m'_Nm''_N)$ to the receiver. Then our code consists of the following components.

1. In the first block the sender randomly chooses a $u_{m'_1,\ell}$ from the codebook with uniform distribution on $\{u_{m'_1}, j : j = 0,1,\ldots,L-1\}$ and sends the codeword to the receiver. Then by choosing a proper decoder the receiver can decode $u_{m'_1,\ell}$ and therefore m'_1 correctly with probability $1-\lambda'$.
2. From the first to the $N-1$st blocks, for all $u_{m,\ell} \in \mathcal{U}$, color all $\mathcal{T}^n_{\bar{Y}|\bar{X}}(u_{m,\ell}) \subset \mathcal{T}^n_{Y|X,\delta_1}(u_{m,\ell})$ with L'' colors such that for a suitably small $\delta_2 > 0$ all n-joint type $P_{\bar{X}\bar{Y}Z}$ with $P_{\bar{X}} = P_X$ and
$$\sum_{yz} |P_{\bar{Y}\bar{Z}\bar{X}}(y,z|x) - P_{YZ|X}(yz|x)| < \delta_2. \tag{4.7}$$
$\mathcal{T}^n_{\bar{Y}|\bar{X}\bar{Z}}(u_{m,\ell}, z^n)$ is properly colored in the sense of Lemma 1.
3. For $j = 1,2,\ldots,N-1$ after the sender receives output y^n of the jth block, he gives up if $y^n \notin \mathcal{T}^n_{Y|X,\delta_1}(u(j))$, where $u(j)$ is the input sequence in $\mathcal{X}^n$ sent by the sender in the jth block. Then the probability for giving up at the jth block is exponentially small in n. In the case $y^n \in \mathcal{T}^n_{Y|X,\delta_1}(u(j))$, y^n receives a coloring $c_{u(j)}(y^n) \in \{0,1,\ldots,L''-1\}$ in the coloring for $\mathcal{T}^n_{\bar{Y}|\bar{X}}(u(j))$, where $P_{\bar{X}\bar{Y}}$ is the joint type of $(u(j), Y^n)$.
   3.1. In the case $L \leq 2^{[H(Y|XZ)-\frac{3}{4}]}$ i.e., $L'' = L$, the sender sends $U_{m'_{j+1}m''_{j+1}} \oplus c_{m(j)}(y^n) \triangleq u(j+1)$ in the codebook $\mathcal{U}$ in the $j+1$st block, where $\oplus$ is the addition modulo L''.
   3.2. In the case $L > 2^{n[H(Y|XZ)-\frac{3}{4}]}$, without loss of generality, we assume $L''|L$. Then the sender partitions $\{0,1,\ldots,L-1\}$ into L'' segments of equal size. He randomly chooses an ℓ''_{j+1} in the $m''_{j+1} \oplus c_{u(j)}(y^n)$ segment with equal probability and sends $u_{m'_{j+1},\ell''_{j+1}}$ in the codebook in the $j+1$st block.

4. For $j = 1, 2, \ldots, N$ in the jth block the receiver decode separately by using a proper decoder and obtains a $\bar{u}(j)$ in the jth block. Thus $\bar{u}(j) = u(j)$ with probability λ' for a given j. Let $\lambda' < M^{-1}\lambda$, then $\bar{u}(j) = u(j)$ with probability larger than $1 - \lambda$ for all j. The receiver declares $m'_1 = \bar{m}'_1$ if $\bar{u}(1) = u_{\bar{m}'_1,\ell}$ for some ℓ. The receiver declares $m'_j m''_j = \bar{m}'_j \bar{m}''_j$ for $\bar{m}''_j = \ell_j \ominus c_{\bar{u}(j-1)}(y^n)$ if in the $j-1$st block he receives y^n and $\bar{u}(j) = u_{\bar{m}'_j \ell_j}$ in the case $L'' = L$ and $\bar{u}(j) = u_{\bar{m}'_j,\ell'_j}$ for an ℓ'_j in the ℓ_jth segment in the case $L'' < L$, for $j = 2, 3, \ldots, N$. Obviously

$$(\bar{m}'_1, \bar{m}'_2\bar{m}''_2, \ldots, \bar{m}'_N\bar{m}''_N) = (m'_1 m m'_2 m''_2, \ldots, m'_N m''_N)$$

if $\bar{u}(j) = u(j)$ for all j.

We have seen that the probability of error is smaller than λ and it is sufficient for us to verify the security criterion.

Denote by $\tilde{X}_j$, $\tilde{Y}_j$ and $\tilde{Z}_j$, the random input and outputs in the jth block generated by the code and the random message, $(U'_1, U'_2U''_2, \ldots, U'_N U''_N)$ respectively, for $j = 1, 2, \ldots, N$. Notice here $\tilde{X}_j$, $\tilde{Y}_j$, and $\tilde{Z}_j$ are random sequences of length n. Let K_j be the coloring of the random output sequences of the legal receiver in the jth block. Write $U'^N = (U'_1, U'_2, \ldots, U'_N)$, $U''^N = (U''_1, U''_2, \ldots, U''_N)$ (where U'_1 is a dummy constant), $\tilde{X}^N = (\tilde{X}_1, \ldots, \tilde{X}_N)$, $\tilde{Y}^N = (\tilde{Y}_1, \ldots, \tilde{Y}_N)$ and $\tilde{Z}^N = (\tilde{Z}_1, \ldots, \tilde{Z}_N)$. Then we are concerned about an upper bound to $I(U'^N U''^N; \tilde{Z}^N)$.

At first we bound $I(U'^N; \tilde{Z}^N)$ with (3.7). Denote $\tilde{Z}^{\bar{j}} = (\tilde{Z}_1, \tilde{Z}_2, \ldots, \tilde{Z}_{j-1}, \tilde{Z}_{j+1}, \ldots, \tilde{Z}_N)$.

Then by symmetry, independent of $\tilde{Z}^{\bar{j}}$ and U'^{j-1}, given $U'_j = m$, the input of the channel in the jth block is uniformly distributed on the sub-codebook

$$\{u_{m,\ell} : \ell = 0, 1, \ldots, L-1\}.$$

For $j = 1$ it immediately follows from the step 1 of the coding scheme. For $j > 1$, it is sufficient for us to show that $P_{U''_j \oplus K_{j-1}|U'^{j-1}\tilde{Z}^{\bar{j}}}$ is uniform. Indeed, for all ℓ, u'^{j-1}, and $z^{\bar{j}}$

$$\begin{aligned} &Pr\{U''_j \oplus K_{j-1} = \ell | U'^{j-1} = u'^{j-1}, \tilde{Z}^{\bar{j}} = z^{\bar{j}}\} \\ &\quad = \sum_{m''=0}^{L''-1} L''^{-1} Pr\{K_{j-1} = \ell \ominus m'' | U^{j-1} = u'^{j-1}; Z^{\bar{j}} = Z^{\bar{j}}\} = L''^{-1}. \end{aligned}$$

This means that for all j and $(V, \tilde{Z})$ in (3.7) we have

$$H(U'_j | U'^{j-1}\tilde{Z}^N) = H(U'_j | \tilde{Z}_j, U'^{j-1}Z^{\bar{j}}) = H(U|\tilde{Z})$$

and therefore by (3.7)

$$I(U'_j; U'^{j-1}\tilde{Z}^N) < \mu'$$

since U'_j and V have the same distribution.

Consequently

$$I(U'^N; Z^N) = \sum_{j=1}^{N} I(U_j; Z^N|U^{j-1}) \leq \sum_{j=1}^{N} I(U_j; U'^{j-1} Z^N) \leq N\mu'. \tag{4.8}$$

Next we bound $I(U''_j; \tilde{Z}^N|U'^N U''^{j-1})$. At first we observe that by our coding scheme U''_j is independent of $U'^N U''^{j-1} \tilde{Z}^i$ for all $i < j$ and therefore

$$\begin{aligned}
&I(U''_j; \tilde{Z}_i|U'^N U''^{j-1} \tilde{Z}^{i-1}) = 0, \text{ or} \\
&I(U''_j; \tilde{Z}^N|U'^N U''^{j-1}) = \sum_{i=1}^{j-1} I(U''_j; \tilde{Z}_i|U'^N U''^{j-1} \tilde{Z}^{i-1}) \\
&+ I(U''_j; \tilde{Z}_j|U'^N U''^{j-1} \tilde{Z}^{j-1}) + I(U''_j; \tilde{Z}^N_{j+1}|U'^N U''^{j-1} \tilde{Z}^j) \\
&= I(U''_j; \tilde{Z}_j|U'^N U''^{j-1} \tilde{Z}^{j-1}) + I(U''_j; \tilde{Z}^N_{j+1}|U'^N U''^{j-1} \tilde{Z}^j),
\end{aligned} \tag{4.9}$$

where $\tilde{Z}^N_{j+1} = (\tilde{Z}_{j+1}, \ldots, \tilde{Z}_N)$.

Moreover by our coding scheme under the condition given $U'^N U''^{j-1} \tilde{Z}^{j-1}$

$$U''_j \Leftrightarrow U''_j \oplus K_{j-1} \Leftrightarrow \tilde{Z}_j$$

form a Markov chain i.e., by the data processing inequality.

$$\begin{aligned}
&I(U''_j; \tilde{Z}_j|U''^N U''^{j-1} Z^{j-1}) \leq I(U''_j; U''_j \oplus K_{j-1}|U'^N U''^{j-1} Z^{j-1}) \\
&= I(U''_j; K_{j-1}|U'^N U''^{j-1} Z^{j-1}) \leq I(U'^N U''^j Z^{j-1}; K_{j-1}).
\end{aligned} \tag{4.10}$$

However, because $U'^N U''^j \tilde{Z}^{j-1} \Leftrightarrow \tilde{X}_{j-1} \tilde{Z}_{j-1} \Leftrightarrow K_{j-1}$ forms a Markov chain, (4.10) implies

$$I(U''_j; \tilde{Z}_j|U'^N U''^j Z^{j-1}) \leq I(\tilde{X}_{j-1} \tilde{Z}_{j-1}; K_{j-1}). \tag{4.11}$$

For $j-1$

$$W_{j-1} = \begin{cases} 0 & \text{if } \tilde{Y}_{j-1} \in \mathcal{T}^n_{Y|X,\delta_1}(\tilde{X}_{j-1}) \\ 1 & \text{else,} \end{cases}$$

then recalling that the output of legal user is colored by Lemma 1 in the $j-1$st block, by AEP we have

$$Pr\{K_{j-1} = k|\tilde{X}_{j-1} = x^n, \tilde{Z}_{j-1} = j^n W_{j-1} = 0\} \leq L''^{-1}(1+\delta).$$

Thus

$$\begin{aligned}
H(K_{j-1}|\tilde{X}_{j-1} \tilde{Z}_{j-1}) &\geq (1 - 2^{-n\theta}) H(K_{j-1}|\tilde{X}_{j-1} \tilde{Z}_{j-1} W_{j-1} = 0) \\
&\geq (1 - 2^{-n\theta})[\log L'' - \log(1+\delta)],
\end{aligned}$$

for a $\theta > 0$ as $Pr(W_j = 0) > 1 - 2^{-n\theta}$. Thus for a $\mu'' > 0$ with $\mu'' \to 0$ as $\delta \to 0$,

$$I(\tilde{X}_{j-1}\tilde{Z}_{j-1}; K_{j-1}) = H(K_{j-1}) - \log L'' + \mu'' \leq \mu'', \tag{4.12}$$

for sufficiently large n. Similarly by the coding scheme under the condition given U'^N

$$U''^j Z^j \Leftrightarrow K_j \Leftrightarrow Z_{j+1}^N$$

forms a Markov chain and therefore

$$I(U_j''; Z_{j+1}^N | U''^N U''^{j-1}) \leq I(U''^j Z^j; \tilde{Z}_{j+1}^N | U'^N) \leq I(U''^j Z^j; K_j | U'^N) \leq I(U'^N U''^j \tilde{Z}^j; K_j). \tag{4.13}$$

However, by the coding scheme $U'^N U''^j \tilde{Z}^j \Leftrightarrow \tilde{X}_j \tilde{Z}_j \Leftrightarrow K_j$ forms a Markov chain and so we can continue to bound (4.13) as

$$I(U_j''; Z_{j+1}^N | U'^N U''^{j-1} Z^j) \leq I(\tilde{X}_j \tilde{Z}_j; K_j). \tag{4.14}$$

By replacing $j-1$ by j in (4.12) and applying it to (4.14) we have

$$I(U_j''; Z_{j+1}^N | U'^N U''^{j-1} Z^j) \leq \mu''. \tag{4.15}$$

Finally, we combine (4.8), (4.9), (4.10), (4.11), and (4.15), to obtain

$$\begin{aligned}
&I(U'^N U''^N; \tilde{Z}^N) \\
&= I(U'^N; \tilde{Z}^N) + I(U''^N; \tilde{Z}^N | U'^N) \\
&\leq N\mu' + \sum_{j=2}^{N} I(U_j''; \tilde{Z}^N | U'^N U''^{j-1}) \\
&= N\mu' + \sum_{j=2}^{N} [I(U_j''; \tilde{Z}_j | U'^N U''^{j-1} \tilde{Z}^{j-1}) + I(U_j''; \tilde{Z}_{j+1}^N | U'^N U''^{j-1} \tilde{Z}_j)] \\
&\leq N\mu' + \sum_{j=2}^{N} [I(\tilde{X}_{j-1}\tilde{Z}_{j-1}; K_{j-1}) + I(U_j''; \tilde{Z}_{j+1}^N | U'^N U''^{j-1} \tilde{Z}_j)] \\
&\leq N\mu' + 2(N-1)\mu'' < \mu,
\end{aligned}$$

for sufficiently small μ' and μ''. This completes our proof.

5 Capacity of Two Special Families of Wire-Tape Channel

In this section we apply Theorem 1 to show the following upper bound of capacity, which is believed not to be tight in general, but is tight for wire-tape channels with certain Markovities.

Let $\mathcal{Q}'$ be the set of triples of random variables (X, Y, Z) with joint distribution $P_{XYZ}(x, y, z) = P_X(x)W(y, z|x)$ for $x \in \mathcal{X}$, $y \in \mathcal{Y}$, and $z \in \mathcal{Z}$.

Then

Lemma 2. *For all wire-tape channels*

$$C_{wtf} \leq \max_{(X,Y,Z)\in\mathcal{Q}'} \min[H(Y|Z), I(X;Y)]. \tag{5.1}$$

Proof: For a given (λ, μ)-code for the wire-tape channel, let X^n, Y^n, Z^n be the input and outputs generated by uniformly distributed messages U through the code. Then in the same way to show the converse coding theorem of a (two terminal) noisy channel with feedback, one obtains that

$$C_{wtf} \leq \frac{1}{n}\sum_{t=1}^{n} I(X_t; Y_t) + \varepsilon' \tag{5.2}$$

where $\varepsilon' \to 0$ as $\lambda \to 0$.

On the other hand, by the security condition and Fano's inequality we have

$$\begin{aligned}
C_{wtf} &= \frac{1}{n}H(U) \leq \frac{1}{n}H(U|Z^n) + \mu \\
&\leq \frac{1}{n}H(U|Z^n) - \frac{1}{n}H(H|Y^n) + \lambda \log|\mathcal{X}| + \frac{1}{n}h(\lambda) + \mu \\
&\leq \frac{1}{n}H(U|Z^n) - \frac{1}{n}H(U|Y^nZ^n) + \lambda \log|\mathcal{X}| + \frac{1}{n}h(\lambda) + \mu \\
&= \frac{1}{n}I(U; Y^n|Z^n) + \varepsilon'' \leq \frac{1}{n}H(Y^n|Z^n) + \varepsilon'' \\
&= \frac{1}{n}\sum_{t=1}^{n} H(Y_t|Z^nY^{t-1}) + \varepsilon'' \leq \frac{1}{n}\sum_{t=1}^{n} H(Y_t|Z_t) + \varepsilon'',
\end{aligned} \tag{5.3}$$

where $h(\lambda) = -\lambda\log\lambda - (1-\lambda)\log(1-\lambda)$ and $\varepsilon'' = \lambda\log|\mathcal{X}| + \frac{1}{n}h(\lambda) + \mu \to 0$ as $\lambda, \mu \to 0$.

Let $(UXYZ)$ be a quadruple of random variables with distribution

$$P_{UXYZ}(t, z, y, z) = \frac{1}{n}\sum_{t=1}^{n} P_{X_tY_tZ_t}(x, y, z)$$

for $t \in \{1, 2, \ldots, n\}$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $z \in \mathcal{Z}$. Then $(XYZ) \in \mathcal{Q}'$ and by (5.2) and (5.3) for $\varepsilon = \max(\varepsilon', \varepsilon'')$

$$C_{wtf} \leq \min[H(Y|ZU), I(X;Y|U)] + \varepsilon \leq \min[H(Y|Z), I(X;Y)] + \varepsilon,$$

where $\varepsilon \to 0$ as $\lambda, \mu \to 0$. That is, (5.1).

Corollary 1. *For a wire-tape channel W such that there exist $W_1 : \mathcal{X} \to \mathcal{Y}$, and $W_2 : \mathcal{Y} \to \mathcal{Z}$ with*

$$W(y, z|x) = W_1(y|x)W_2(z|y), \tag{5.4}$$

for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$, *and* $z \in \mathcal{Z}$ $C_{wtf} = \max_{(X,Y,Z)\in\mathcal{Q}'}, \min[H(Y|Z), I(X;Y)]$.

Proof: By Markov condition (5.4), we have that for all $(X, Y, Z) \in \mathcal{Q}'$

$$I(X;Y) - I(X,Z) \geq 0 \tag{5.5}$$

and

$$I(X;Z|Y) = 0. \tag{5.6}$$

Thus

$$\begin{aligned} |I(X;Y) - I(X;Z)|^+ + H(Y|XZ) &= H(X|Z) - H(X|Y) + H(Y|XZ) \\ &= H(XY|Z) - H(X|Y) \\ &= H(Y|Z) + H(X|YZ) - H(X|Y) \\ &= H(Y|Z) + I(X;Z|Y) \\ &= H(Y|Z). \end{aligned}$$

Then corollary follows from Theorem 1 and Lemma 2.

Corollary 2. *For a wire-tape channel such that there exist* $W_1' : \mathcal{X} \to \mathcal{Z}$ *and* $W_2' : \mathcal{Z} \to \mathcal{Y}$ *with*

$$W(y,z|x) = W_1'(z|x)W_2'(y|z) \tag{5.7}$$

for $x \in \mathcal{X}$, $y \in \mathcal{Y}$, *and* $z \in \mathcal{Z}$

$$C_{wtf} = \max_{(X,Y,Z)\in\mathcal{Q}'}, \min[H(Y|Z), I(X;Y)].$$

Proof: The Markov condition (5.7) implies that

$$I(X;Y) - I(X;Z) \leq 0 \tag{5.8}$$

and

$$H(Y|XZ) = H(Y|Z), \tag{5.9}$$

which yield

$$|I(X;Y) - I(X;Z)|^+ + H(Y|XZ) = H(Y|XZ) = H(Y|Z). \tag{5.10}$$

Thus the corollary follows from Theorem 1 and Lemma 2.

Example: An interesting example is a special channel for which W_1' is a noiseless channel and W_2' is a noisy channel in Corollary 2 e.g., W_1 is a noiseless binary channel, W_2'' is a binary symmetric channel with crossover probability $p \in \left(0, \frac{1}{2}\right)$. For this channel the wiretapper is in a better position than the legal user. So the capacity is zero without feedback. The feedback makes the capacity positive by our Corollary 2 as it serves as a secure key shared by sender and receiver.

6 Discussion: Transmission, Building Common Randomness and Identification

As goals of communications are considered transmission i.e., sending a given message from a set of messages, building common randomness i.e., to provide a random resource shared by users, and identification i.e., identifying whether an event of interest to a particular user occurs ([3], [4], [5], [13]).

Roughly saying in a given communication system, the capacity of transmission is upper bounded by the capacity of common randomness, since common randomness shared by a sender and receiver can be built by transmission whereas the capacity of identification is lower bounded by capacity of common randomness, if the former is positive, which is shown by a scheme in [5] to build identification codes by common randomness. That is,

$$\begin{aligned} \text{capacity of transmission} \ &\leq \ \text{capacity of common randomness} \\ &\leq \ \text{capacity of identification.} \end{aligned} \tag{6.1}$$

However, in different communication systems equalities in (6.1) may or may not hold. In this section we illustrate the variety in two-terminal channels and wiretape channels. More examples in more complicated communication systems can be found e.g. in [3], [12], [15].

First of all, obviously the first inequality in (6.1) is always an equality for a two terminal channel without feedback, because all information obtained by the receiver is from the transmission via the channel. Moreover, it has been shown in [4] that the second inequality is an equality and therefore the three quantities in (6.1) are actually the same if the channel is discrete memoryless. A channel with rapidly increasing alphabet (as the length of codes grows) for which the capacity of identification is strictly larger than capacity of common randomness was described in [6]. It was shown in [8] that under a certain condition the capacity of common randomness (which is equal to the capacity of transmission) for Gaussian channels is finite whereas the capacity of identification is infinite in the same communication system. We notice that Gaussian channels have continuous, or infinite alphabets. It is natural to expect that for a discrete channel whose input alphabet "reasonably" increases the last two quantities, or consequently the three quantities in (6.1) are equal. This was shown in [14] for all channels whose input alphabets exponentially increase as the lengths of codes linearly increase.

The situation of two terminal channels is different when feedback is present. In this case the capacity of identification, which is equal to the capacity of common randomness, is strictly larger than the capacity of transmission for simplest channels, namely discrete memoryless channels [5]. The reason is clear. On one hand, it is well known, feedback does not increase the capacity of transmission for discrete memoryless channels. On the other hand, the feedback provides a random resource, shared by sender and receiver, the random output, whose rate, roughly speaking, is input entropy. Obviously it increases common randomness between sender and receiver and therefore capacity of identification.

Next we turn to wire-tape channels without feedback. More precisely, we mean secure common randomness shared by sender and receiver, about which the wiretapper has (almost) no knowledge. By the same reason as for two terminal channels without feedback, the capacity of (secure) common randomness is not larger than the capacity of transmission over the wire-tape channel. In fact it is shown in [3], that it may not be larger than the capacity of transmission even in the case where a public forward channel with unbounded capacity is available to the sender and receiver. This intuitively is not surprising. R. Ahlswede and Z. Zhang observed in [7] that to keep the message to be identified in secret a secure common randomness with positive rate is sufficient and the *major part of common randomness between the legitimate communicator applied in the identification code in* [5] *can be publically sent.*

Based on this observation they show that the capacity of identification is strictly larger than the capacity of secure common randomness. A more detailed analysis in [9] shows that the amount of secure common randomness needed only depends on the probability of second error and security criterion and is independent of the rate of messages. For fixed criterion of error and security, a constant amount – or zero-rate – of secure common randomness is sufficient, if provided with sufficiently large public common randomness.

Let us return to our main topic wire-tape channels with secure feedback and investigate (6.1) in this communication system. We immediately find that the observation about wire-tape channels without feedback is still valid when feedback is present, because there is nothing in the observation which links to the existence of feedback. This means that the capacity of identification must be the capacity of "public" common randomness between sender and receiver i.e., the maximum rate of common randomness shared by the sender and the receiver, neglecting whether or how much the wiretapper knows about it once a positive amount of secure common randomness is provided. But now the public common randomness is the maximum output entropy for the channel $W_1 : \mathcal{X} \to \mathcal{Y}$ defined by

$$W_1(y|x) = \sum_{z \in \mathcal{Z}} W(y, z|x) \text{ for all } x \in \mathcal{X}, y \in \mathcal{Y}, \tag{6.2}$$

or in other words $\max\limits_{(X,Y,Z)\in\mathcal{Q}'} H(Y)$, for $\mathcal{Q}'$ as defined in Section 5. So we conclude that in this case the capacity of identification is either zero or $\max\limits_{(X,Y,Z)\in\mathcal{Q}'}, H(Y)$. The only problem left is to find suitable conditions for the positivity of the capacity. We shall discuss this later.

To see the relation of the first pair of quantities in (6.1), we take a look at our main result

Theorem 1. *The information theoretical meaning of mutual information in (4.2) is obvious. The capacity of transmission with security criterion can not exceed that without it. So we expect this term could be removed in the formula of capacity of common randomness. To investigate the remaining term in (4.2), let us recall our coding scheme in Section 4.*

From the first block to the second last block, the transmission in each block has two tasks, sending a secret message m'_j (in the jth block) with a rate $\sim |I(U;Y) - I(U;Z)|$; and generating a secure common randomness with a rate $\sim H(Y|UZ)$, which will be used as a private key to send message m''_{j+1} in the next block. This gives us a secure common randomness with rate $\sim H(Y|UZ)$. The reason for the fact that U occurs in the "condition" is that the key for the $j+1$st block has to be independent of the message sent in the jth block. For secure common randomness itself this is not necessary. So we expect that the capacity of common randomness is $\max\limits_{(X,Y,Z)\in\mathcal{Q}'} H(Y|Z)$, which actually is shown in the next section.

But before this we have a remaining problem, namely the positivity of the capacity of identification, which should be discussed. First we notice that to have positive capacity of identification, the capacity of the channel W_1 in (6.2), where we do not count wiretapper's role, has to be positive. By counting wiretapper's role, we look for an input random variable X, the conditional entropy $H(Y|Z)$ for output random variable Y and Z has to be positive, because otherwise the wiretapper would know everything known by the legal receiver. *We shall show that the two necessary conditions together are sufficient for the positivity.*

7 The Secure Common Randomness Capacity in the Presence of Secure Feedback

Let $\mathcal{J}_n = \{0, 1, \ldots, J_n - 1\}$ be a finite set (whose size depends on n), $\lambda, \mu > 0$. An (n, J_n, λ, μ)-common randomness for the wire-tape channel with secure feedback is a pair of random variables (K_n, L_n) defined on the same domain $\mathcal{J}_n$ with the following properties.

There exists a random variable U taking value in a finite set $\mathcal{U}$ and three functions $\theta^n : \mathcal{U} \times \mathcal{Y}^{n-1} \to \mathcal{X}^n$, $\varphi : \mathcal{U} \times \mathcal{Y}^n \to \mathcal{J}_n$, and $\Psi : \mathcal{Y}^n \to \mathcal{J}_n$ such that for all $u \in \mathcal{U}$ and $y^{n-1} \in \mathcal{Y}^{n-1}$

$$\theta^n(u, y^{n-1}) = (\theta_1(u), \theta_2(u, y_1), \ldots, \theta_n(u, y^{n-1})), \tag{7.1}$$

$$K_n = \varphi(U, Y^n) \tag{7.2}$$

$$\text{and} \qquad L_n = \Psi(Y^n), \tag{7.3}$$

where Y^n and Z^n are output random variables for the legal receiver and the wiretapper, respectively, generated by random variable U, encoding function θ^n, and the channel W.

I.e.

$$Pr((Y^n, Z^n) = (y^n, z^n)) = \sum_{u\in\mathcal{U}} Pr(U = u) W(y_1, z_1|\theta_1(u)) \prod_{t=2}^{n} W(y_t, z_t|\theta_t(u, y^{t-1})). \tag{7.4}$$

$$Pr(K_n \neq L_n) < \lambda, \tag{7.5}$$

$$\frac{1}{n}H(K_n|Z^n) > \frac{1}{n}\log J_n - \mu. \tag{7.6}$$

$\frac{1}{n}\log J_n$ is called rate of the code and the capacity of the (secure) common randomness, denoted by C_{wtf-cr}, is defined as the maximum achievable rate in the standard way.

Theorem 2

$$C_{wtf-cr} = \max_{(X,Y,Z)\in\mathcal{Q}'} H(Y|Z), \tag{7.7}$$

in particular, the RHS of (7.7) is achievable if (7.6) is replaced by a stronger condition

$$H(K_n|Z^n) > \log J_n - \mu. \tag{7.8}$$

Proof: The proofs to both, direct and converse parts, are straightforward. They immediately follow from the proofs for Theorem 1 and Lemma 2, respectively.

Let $(X', Y, Z) \in \mathcal{Q}'$ achieve the maximum at RHS (7.7). Apply Lemma 1 to color sets of typical remaining sequences $\mathcal{T}^n_{Y'} \subset \mathcal{T}^n_{Y,\delta}$ [1], then it follows from the proof of Theorem 1 (the part to show (4.11)) that for any fixed $\mu > 0$ and sufficiently large n

$$H(\tilde{K}|Z^n) > \log J_n - \mu,$$

where $\tilde{K}$ is the random J_n-coloring obtained from Lemma 1.
Choose $K_n = L_n = \tilde{K}$, then the proof of the direct part is done. To show the converse part we apply Fano's inequality to (7.5). Then

$$\begin{aligned}
\frac{1}{n}\log J_n &\le \frac{1}{n}H(K_n|Z^n) + \mu \\
&\le \frac{1}{n}H(K_n|Z^n) - \frac{1}{n}H(K_n|Y^n) + \mu + \frac{1}{n}\lambda\log J_n + \frac{1}{n}h(\lambda) \\
&\le \frac{1}{n}H(K_n|Z^n) - \frac{1}{n}H(K_n|Y^n, Z^n) + \mu + \frac{1}{n}\lambda\log J_n + \frac{1}{n}h(\lambda) \\
&\le \frac{1}{n}I(K_n; Y^n|Z^n) + \mu + \frac{1}{n}\lambda\log J_n + \frac{1}{n}h(\lambda) \\
&\le \frac{1}{n}H(Y^n|Z^n) + \mu + \frac{1}{n}\lambda\log J_n + \frac{1}{n}h(\lambda).
\end{aligned}$$

Now the converse follows as in the proof for Lemma 2.

8 The Secure Identification Capacity in the Presence of Secure Feedback

In this section let us take a look at the coding theorem for identification codes. First we have to formally define the codes and capacity. An $(n, |\mathcal{M}|, \lambda_1, \lambda_2, \mu)$-

[1] More precisely, let $\mathcal{X}_0 = \{x_0\}$, $x^n = (x_0, x_0, \ldots, x_0)$, and (X, X', Y, Z) be random variables with joint distribution $Pr((X, X', Y, Z) = (x^n, x'^n, y^n, z^n)) = P_{X'YZ}(x'^n, y^n, z^n)$ for all x'^n, y^n, z^n and coloring for the "conditional" typical sequences $\mathcal{T}^n_{Y|X}(x^n) = \mathcal{T}^n_Y$.

identification code for a wire-tape channel with secure feedback is a system $\{Q, \mathcal{D}_m : m \in \mathcal{M}\}$ such that $Q : \mathcal{M} \times Y^{n-1} \to \mathcal{X}^n$ is a stochastic matrix with

$$Q(x^n|m, y^{n-1}) = Q_1(x_1|m) \prod_{t=2}^{n} Q_t(x_t|m, y^{t-1})$$

for $m \in \mathcal{M}$, $y^{n-1} \in \mathcal{Y}^{n-1}$, for all $m \in \mathcal{M}$

$$\sum_{x^n \in \mathcal{X}^n} \sum_{y^n \in \mathcal{D}_m} Q_n(x_1|m) \prod_{t=2}^{n} Q_t(x_t|m, y^{t-1}) W_1(y_t|x_t) > 1 - \lambda_1,$$

for $m, m' \in \mathcal{M}$ with $m \neq m'$

$$\sum_{x^n \in \mathcal{X}^n} \sum_{y^n \in \mathcal{D}'_m} Q_1(x_1|m) \prod_{t=2}^{n} Q_t(x_t|m, y^{t-1}) W_1(y_t|x_t) < \lambda_2,$$

and for all $m, m' \in \mathcal{M}$, $m \neq m'$ and $\mathcal{V} \subset Z^n$

$$\sum_{x^n \in \mathcal{X}^n} \sum_{y^n \in \mathcal{Y}^n} Q_1(x_1|m') \prod_{t=2}^{n} Q_t(x_t|m', y^{t-1}) W(y^n, \mathcal{V}|x^n)$$

$$+ \sum_{x^n \in \mathcal{X}^n} \sum_{y^n \in \mathcal{Y}^n} Q_1(x_1|m) \prod_{t=2}^{n} Q_t(x_t|m, y^{t-1}) W(y^n, \mathcal{V}^c|x^n) > 1 - \mu.$$

Then capacity of identification is defined in the standard way and denoted by C_{wtf-id}.

C_{wtf-id} is upper bounded by the RHS of (8.1), follows from the converse of the coding theorem of identification with feedback for channel W_1 [5]. In the case that II holds, one can construct a code achieving $H(Y)$ asymptotically from the code in [7] by replacing the ordinary code for W_1 by a uniform partition of output sequences for the legal receiver and a code for the wire-tape channel without feedback by a code for the same channel but with feedback.

Furthermore the two conditions in III

Theorem 3. *The following statements are equivalent.*

I $C_{wtf-id} = \max_{(X,Y,Z) \in \mathcal{Q}'} H(Y)$ (8.1)

II $C_{wtf} > 0$

III There exists an $(X, Y, Z) \in \mathcal{Q}'$ such that

$$H(Y|Z) > 0$$

and the channel W_1 has positive capacity.

Proof: The converse of the coding theorem i.e., C_{wtf-id} is upper bounded by the right hand side of (8.1) follows from the converse of coding theorem of

identification with feedback for channel W_1 [4]. In the case that II holds, one can construct a code achieving $H(Y)$ asymptotically from the code in [6] by replacing the ordinary code for W_1 by a uniform partition of output sequences for the legal receiver and a code for the wiretape channel without feedback by a code for the same channel but with feedback.

Furthermore the two conditions in III obviously are necessary for positivity of C_{wtf-id}. The only thing left to be proved is that III implies II. Let $(X_i, Y_i, Z_i) \in \mathcal{Q}'$ for $i = 0, 1$ such that $H(Y_0|Z_0) > 0$ and $I(X_1, Y_1) > 0$. By Theorem 1, it is sufficient for us to find $(U, X, Y, Z) \in \mathcal{Q}$ such that $I(U;Y) > 0$ and $H(Y|UZ) > 0$. Obviously we are done, if $I(X_0; Y_0) > 0$ or $H(Y_1|U_1, Z_1) > 0$. Otherwise we have to construct a quadruple of random variables $(U, X, Y, Z) \in \mathcal{Q}$ from (X_0, Y_0, Z_0) and (X_1, Y_1, Z_1) such that $H(Y|UZ) > 0$ and $I(U;Y) > 0$. To this end, let $\mathcal{U} = \mathcal{X} \cup \{u_0\}$, (where u_0 is a special letter not in $\mathcal{X}$), and for all $u \in \mathcal{U}$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$, let (U, X, Y, Z) be a quadruple of random variables such that

$$P_{UXYZ}(u,x,y,z) = \begin{cases} \frac{1}{2} P_{X_0Y_0Z_0}(x,y,z) & \text{if } u = u_0 \\ \frac{1}{2} P_{X_1Y_1Z_1}(x,y,z) & \text{if } u \in \mathcal{X} \text{ and } u = x \\ 0 & \text{otherwise.} \end{cases}$$

Then $(U, X, Y, Z) \in \mathcal{Q}$, $P_{YZ|U}(y|u_0) = P_{Y_0Z_0}(yz)$ for all $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$. $P_0(u_0) = \frac{1}{2}$ and therefore

$$H(Y|UZ) = \sum_{u \in \mathcal{U}} P_U(u) H(Y|U = uZ) \geq \frac{1}{2} H(Y|U = u_0 Z) = \frac{1}{2} H(Y_0|Z_0) > 0.$$

On the other hand for

$$S = \begin{cases} 0 & \text{if } U = u_0 \\ 1 & \text{otherwise,} \end{cases}$$

for all $u \in \mathcal{X}, y \in \mathcal{Y}$

$$P_{UY|S}(u, y|S = 1) = P_{X_1Y_1}(u, y)$$

and $P_s(1) = \frac{1}{2}$ and consequently

$$I(U;Y) = I(US;Y) \geq I(U;Y|S) \geq P_s(1) I(U;Y|S = 1) = \frac{1}{2} I(X_1; Y_1) > 0.$$

That is, (U, X, Y, Z) is as desired. We conclude with the

Corollary 3

$$C_{wtf-id} = \begin{cases} \max\limits_{(X,Y,Z) \in \mathcal{Q}'} H(Y|Z) \\ 0 \end{cases}$$

and $C_{wtf-id} = 0$ iff for all $(X, Y, Z) \in \mathcal{Q}'$ $H(Y|Z) = 0$ or the capacity of W_1 is zero.

Proof: That for all $(X, Y, Z) \in \mathcal{Q}'$, $H(Y|Z) = 0$ implies that the wiretapper knows what the receiver receives with probability one no matter how the sender chooses the input and that the capacity of W_1 is zero means the sender may not change the output distributions at the terminal for the legal receiver. So in both cases $C_{wtf-id} = 0$. Thus the corollary follows from Theorem 3.

References

1. R. Ahlswede, Universal coding, Paper presented at the 7th Hawai International Conference on System Science, Jan. 1974, Published in [A21].
2. R. Ahlswede, Coloring hypergraphs: A new approach to multi-user source coding, Part I, J. Comb. Inform. Syst. Sci., Vol. 4, No. 1, 76-115, 1979; Part II, Vol. 5, No. 3, 220-268, 1980.
3. R. Ahlswede and I. Csiszár, Common randomness in information theory and cryptography, Part I: Secret sharing, IEEE Trans. Inf. Theory, Vol. 39, No. 4, 1121-1132, 1993; Part II: CR capacity, Vol. 44, No. 1, 55-62, 1998.
4. R. Ahlswede and G. Dueck, Identification via channels, IEEE Trans. Inform. Theory, Vol. 35, No. 1, 15-29, 1989.
5. R. Ahlswede and G. Dueck, Identification in the presence of feedback – a discovery of new capacity formulas, IEEE Trans. Inform. Theory, Vol. 35, No. 1, 30-39, 1989.
6. R. Ahlswede, General theory of information transfer, Preprint 97–118, SFB 343 "Diskrete Strukturen in der Mathematik", Universität Bielefeld, 1997; General theory of information transfer:updated, General Theory of Information Transfer and Combinatorics, a Special Issue of Discrete Applied Mathematics, to appear.
7. R. Ahlswede and Z. Zhang, New directions in the theory of identification via channels, IEEE Trans. Inform. Theory, Vol. 41, No. 4, 1040-1050, 1995.
8. M. Burnashev, On identification capacity of infinite alphabets or continuous time, IEEE Trans. Inform. Theory, Vol. 46, 2407-2414, 2000.
9. N. Cai and K.-Y. Lam, On identification secret sharing scheme, Inform. and Comp., 184, 298-310, 2002.
10. I. Csiszár, Almost independence and secrecy capacity, Probl. Inform. Trans., Vol. 32, 40-47, 1996.
11. I. Csiszár and J. Körner, Broadcast channel with confidential message, IEEE Trans. Inform. Theory, Vol. 24, 339-348, 1978.
12. I. Csiszár and P. Narayan, Common randomness and secret key generation with a helper, IEEE Trans. Inform. Theory, Vol. 46, No. 2, 344–366, 2000.
13. C.E. Shannon, A mathematical theory of communication, Bell. Sys. Tech. J., 27, 379-423, 1948.
14. Y. Steinberg, New converses in the theory of identification via channels, IEEE Trans. Inform. Theory, Vol. 44, 984-998, 1998.
15. S. Venkatesh and V. Anantharam, The common randomness capacity of a network of discrete memoryless channels, IEEE Trans. Inform. Theory, Vol. 46, 367-387, 2000.
16. A.D. Wyner, The wiretape channel, Bell. Sys. Tech. J., Vol. 54, 1355-1387, 1975.