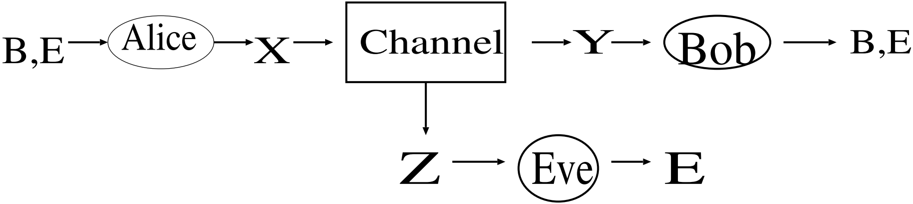
Strong security and separated code constructions for the broadcast channels with confidential messages

Ryutaroh Matsumoto[1]
Masahito Hayashi[2]

[1]Tokyo Institute of Technology, Japan
[2]Tohoku University, Japan/National University of Singapore

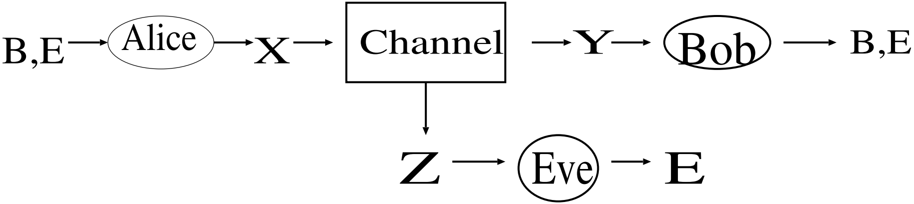
Sept. 29, 2010
at Institute of Network Coding
Chinese University of Hong Kong

Broadcast channel with degraded message sets



- Common message E is delivered to both Bob and Eve.
- Private message B is delivered to Bob.
- Eve is allowed to know B.

Broadcast channel with confidential messages



- Common message *E* is delivered to both Bob and Eve.
- Secret message *B* is delivered only to Bob.
- Eve is **NOT** allowed to know *B*.

Q: How can we measure Eve's lack of knowledge on *B*??

Weak&strong security criteria

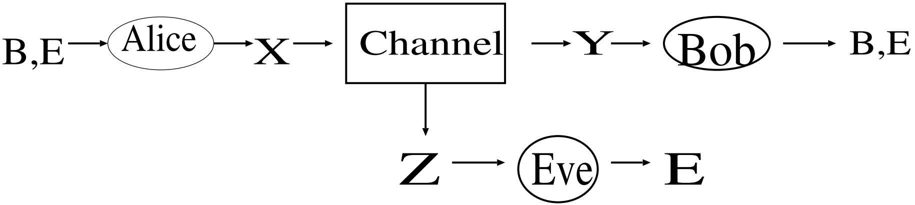
n: the number of channel use
The weak security criterion requires

$$\lim_{n\to\infty} \frac{I(B_n; Z^n)}{n} = 0.$$

Is it really ensure the secrecy?? Suppose that the first $\sqrt{n}$ symbols in B_n and Z^n are always the same and the rest are statistically independent. Although Eve knows infinitely much information on B_n, it is judged secure!!!

The strong security [Maurer 1994] requires $I(B_n; Z^n) \to 0$.

Capacity region of BCC I



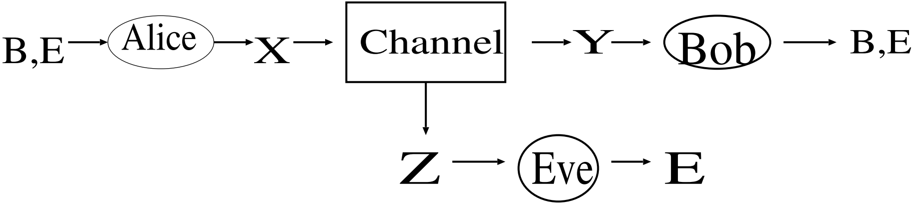
Which rate pairs for *B* and *E* are achievable?

The capacity region under the weak security was found by Csiszár and Körner.

That under the strong security remains unknown. I will show that it is the same as the weak security.

Proof of the converse part is unnecessary. I will show the direct part.

Capacity region of BCC II



R_0: rate of E (common message)
R_1: rate of B (secret message)

(R_0, R_1) is achievable if
$U \to V \to X \to YZ$
$R_0 < \min\{I(U;Y), I(U;Z)\}$
$R_1 < I(V;Y|U) - I(V;Z|U)$

An interpretation of the capacity region of BCC I

$U \to V \to X \to YZ$
$R_0 < \min\{I(U;Y), I(U;Z)\}$
$R_1 < I(V;Y|U) - I(V;Z|U)$

Interpretation:
U: common message
V: common message + private (not secret) message
$V \to X$: artificial noise increasing Bob's advantage over Eve

The common message rate is the minimum of channel capacities to Bob and Eve.

The secret message rate is the capacity to Bob minus that to Eve.

An interpretation of the capacity region of BCC II

$U \to V \to X \to YZ$
$R_0 \le \min\{I(U;Y), I(U;Z)\}$
$R_1 \le I(V;Y|U) - I(V;Z|U)$

If

- we set $X = V$, and
- delete the term $-I(V;Z|U)$,

then the region is almost the same as that of BC with degraded message sets.

This suggests

- coding for BC with degraded message sets can be used for that for BC with confidentiality, and
- replacing $nI(V;Z|U)$ symbols in the private message to Bob with random garbage makes the rest of the private message secret to Eve.

I will formalize the above by the inverse hashing construction introduced last week.

Review of random coding for BC with degraded message sets

R_0: rate of the common message
R'_1: rate of the private (not secret) message to Bob
Given $U \to V \to X \to YZ$

1. Generate $\exp(nR_0)$ codewords of length n according to P_U.
2. For each codeword in the above step, generate $\exp(nR'_1)$ codewords according to $P_{V|U}$.

Adding secrecy to a code for BC with degraded message sets I

E_n: common message
B_n: private (not secret) message in the degraded message sets
S_n: secret message
F_n: hash function from $\mathcal{B}_n$ to $\mathcal{S}_n$

Structure of the transmitter

1. given S_n, uniformly randomly choose B_n from $\{b \in \mathcal{B}_n \mid F_n(b) = S_n\}$,
2. encode B_n and E_n by the encoder for BC with degraded message sets,
3. Apply the artificial noise $P_{X|V}$.

Adding secrecy to a code for BC with degraded message sets II

E_n: common message
B_n: private (not secret) message in the degraded message sets
S_n: secret message
F_n: hash function from $\mathcal{B}_n$ to $\mathcal{S}_n$
Y^n: Bob's received signal

Structure of Bob's receiver

1. decode B_n from Y^n
2. apply F_n to B_n to get S_n

The transmitter and the receiver have to agree on the choice of F_n in advance.
I will discuss this in the next slide.

Pre-agreement on the choice of hash functions

Since $I(S_n, Z^n | F_n) = \mathbf{E}_f I(S_n, Z^n | F_n = f)$ will be shown to be small, almost every choice of hash function f makes $I(S_n, Z^n | F_n = f)$ small. The transmitter and the receiver can agree on the choice of f in advance, and the receiver can compute S_n by applying f to decoded B_n.

Calculation of the mutual information – Setup

We don't have to care about the decoding error probability, because we just use the code for BC with degraded message sets without change.

We need to evaluate $I(S_n;Z^n)$. I fix notations:
F_n: hash function (to be elaborated in the next slide)
B_n: private message (not secret) to Bob
E_n: common message
Λ: RV representing the selection of random codebook
All RVs below incorporates Λ's effect.
U^n: codeword for the common message
V^n: codeword for the common+private message
Z^n: Eve's received signal

Assumptions

- S_n is uniformly distributed (can be relaxed by the last week's argument).
- The distribution of E_n can be arbitrary.
- B_n and E_n are independent.
- F_n is from a family of two-universal hash functions.
- Each function F_n is surjective.
- $\{b \in \mathcal{B}_n \mid F_n(b) = s\}$ has the constant number of elements for every pair of F_n and s.

Family of two-universal hash functions

We will propose a construction attaching a hash function at the source node to an existing code for a BC with degraded message sets.

Definition

Let $\mathcal{F}$ be a set of functions from a set $\mathcal{S}_1$ to $\mathcal{S}_2$, F a (uniform) RV on $\mathcal{F}$. If for all $x_1 \neq x_2 \in \mathcal{S}_1$ we have

$$\Pr[F(x_1) = F(x_2)] \leq \frac{1}{|\mathcal{S}_2|},$$

then $\mathcal{F}$ is said to be a family of two-universal hash functions.

Privacy amplification theorem

(X, Z): a pair of discrete RVs
$\mathcal{F}$: a family two-universal hash functions from $\mathcal{X}$ to $\mathcal{S}$
F: an RV on $\mathcal{F}$ statistically independent of (X, Z).

$$
\begin{aligned}
I(F(X); Z|F) &\leq \frac{|\mathcal{S}|^\rho \mathbf{E}[P_{X|Z}(X|Z)^\rho]}{\rho} \\
&= \frac{|\mathcal{S}|^\rho}{\rho} \sum_{x,z} P_{XZ}(x,z) P_{X|Z}(x|z)^\rho \\
&= \frac{|\mathcal{S}|^\rho}{\rho} \sum_{x,z} P_{XZ}(x,z)^{1+\rho} P_Z(z)^{-\rho}
\end{aligned}
$$

for all $0 < \rho \leq 1$.

All the logarithms, including ones in H and I, have to be the natural ones.

$\rho = 1$: Bennett et al. (1995).
$0 < \rho \leq 1$: Hayashi (2009).
$\rho \to 0$ gives the best result for our application.

Calculation of the mutual information – Outline

We first fix realizations of E_n (common message) and Λ (selection of the random codebook)

Goal: Derive an upper bound that is concave with respect to the input distribution.

By the concavity, we become able to include averaging by the random coding Λ into the upper bound.

$$
\begin{aligned}
& I(F_n(B_n);Z^n|F_n,\Lambda=\lambda)\\
\leq\ & I(F_n(B_n);Z^n,E_n|F_n,\Lambda=\lambda)
\end{aligned}
$$

Giving the common message E_n does not increase I much.

$$
\begin{aligned}
=\ & \underbrace{I(F_n(B_n);E_n|F_n,\Lambda=\lambda)}_{=0}+I(F_n(B_n);Z^n|F_n,E_n,\Lambda=\lambda)\\
=\ & \sum_e P_{E_n}(e)I(F_n(B_n);Z^n|F_n,E_n=e,\Lambda=\lambda)\\
\leq\ & \sum_e P_{E_n}(e)\frac{\exp(n\rho R_1)}{\rho}\times\sum_{b,z}P_{B_n,Z^n|E_n=e,\Lambda=\lambda}(b,z)^{1+\rho}P_{Z^n|E_n=e,\Lambda=\lambda}(z)^{-\rho}\\
=\ & \sum_e P_{E_n}(e)\frac{\exp(n\rho R_1)}{\rho}\sum_{b,z}P_{B_n|\Lambda=\lambda}(b)^{1+{\color{red}\rho}}P_{Z^n|B_n,E_n=e,\Lambda=\lambda}(z|b)^{1+\rho}\\
& P_{Z^n|E_n=e,\Lambda=\lambda}(z)^{-\rho}\\
=\ & \sum_e P_{E_n}(e)\frac{\exp(n\rho R_1)}{\rho{\color{red}\exp(n\rho R_1')}}\sum_{b,z}P_{B_n|\Lambda=\lambda}(b)P_{Z^n|B_n,E_n=e,\Lambda=\lambda}(z|b)^{1+\rho}P_{Z^n|E_n=e,\Lambda=\lambda}(z)^{-\rho}
\end{aligned}
$$

by the uniformity of B_n. This makes the desired concavity.

Rewriting the upper bound by U and V

$$
\begin{aligned}
& \sum_e P_{E_n}(e)\frac{\exp(n\rho R_1)}{\rho\exp(n\rho R'_1)}\sum_{b,z} P_{B_n|\Lambda=\lambda}(b)P_{Z^n|B_n,E_n=e,\Lambda=\lambda}(z|b)^{1+\rho} \\
& P_{Z^n|E_n=e,\Lambda=\lambda}(z)^{-\rho} \\
= & \sum_e P_{E_n}(e)\frac{\exp(n\rho R_1)}{\rho\exp(n\rho R'_1)}\sum_{b,z} P_{B_n|\Lambda=\lambda}(b)[P_{Z^n|B_n,E_n,\Lambda=\lambda}(z|b,e)]^{1+\rho} \\
& P_{Z^n|E_n=e,\Lambda=\lambda}(z)^{-\rho} \\
= & \sum_e P_{E_n}(e)\frac{\exp(n\rho R_1)}{\rho\exp(n\rho(R'_1))}\sum_{v,z} \underbrace{\sum_{b:\lambda(b,e)=v} P_{B_n|\Lambda=\lambda}(b)}_{=P_{V^n|E_n=e,\Lambda=\lambda}(v)} \underbrace{P_{Z^n|B_n,E_n,\Lambda=\lambda}(z|b,e)^{1+\rho}}_{=P_{Z^n|V^n,\Lambda=\lambda}(z|v)^{1+\rho}} \\
& P_{Z^n|E_n=e,\Lambda=\lambda}(z)^{-\rho}
\end{aligned}
$$

Rewriting with a concave function ψ

$$
\begin{aligned}
&= \sum_e P_{E_n}(e) \frac{\exp(n\rho R_1)}{\rho \exp(n\rho(R'_1))} \sum_{v,z} P_{V^n|E_n=e,\Lambda=\lambda}(v) P_{Z^n|V^n,\Lambda=\lambda}(z|v)^{1+\rho} \\
&\quad P_{Z^n|E_n=e,\Lambda=\lambda}(z)^{-\rho} \\
&= \sum_e P_{E_n}(e) \frac{\exp(n\rho R_1 + \psi(\rho, P_{Z^n|V^n,\Lambda=\lambda}, P_{V^n|E_n=e,\Lambda=\lambda}))}{\rho \exp(n\rho(R'_1))} \\
&= \sum_e P_{E_n}(e) \frac{\exp(n\rho R_1 + \psi(\rho, P_{Z^n|V^n}, P_{V^n|E_n=e,\Lambda=\lambda}))}{\rho \exp(n\rho(R'_1))} \\
&= \sum_e P_{E_n}(e) \frac{\exp(n\rho(R_1 - R'_1) + \psi(\rho, P_{Z^n|V^n}, P_{V^n|E_n=e,\Lambda=\lambda}))}{\rho}
\end{aligned}
$$

$$\psi(\rho, P_{Z|L}, P_L) = \ln \sum_z \frac{\sum_\ell P_L(\ell)(P_{Z|L}(z|\ell))^{1+\rho}}{P_Z(z)^\rho}$$ concave w.r.t. P_L with $P_{Z|L}$ fixed.

Averaging over the random codebooks

$$
\begin{aligned}
& \sum_{e} P_{E_n}(e) I(F_n(B_n); Z^n | F_n, \Lambda, E_n = e) \\
= \quad & \sum_{\lambda} P_{\Lambda}(\lambda) \sum_{e} P_{E_n}(e) \frac{\exp(n\rho(R_1 - R_1') + \psi(\rho, P_{Z^n|V^n}, P_{V^n|E_n=e,\Lambda=\lambda}))}{\rho} \\
= \quad & \sum_{\lambda} P_{\Lambda}(\lambda) \sum_{e} P_{E_n}(e) \frac{\exp(n\rho(R_1 - R_1') + \psi(\rho, P_{Z^n|V^n}, P_{V^n|U^n=\lambda(e),\Lambda=\lambda}))}{\rho} \\
= \quad & \sum_{\lambda} P_{\Lambda}(\lambda) \sum_{u} P_{U^n|\Lambda=\lambda}(u) \frac{\exp(n\rho(R_1 - R_1') + \psi(\rho, P_{Z^n|V^n}, P_{V^n|U^n=u,\Lambda=\lambda}))}{\rho} \\
= \quad & \sum_{u} P_{U^n}(u) \sum_{\lambda} P_{\Lambda|U^n=u}(\lambda) \frac{\exp(n\rho(R_1 - R_1') + \psi(\rho, P_{Z^n|V^n}, P_{V^n|U^n=u,\Lambda=\lambda}))}{\rho} \\
\leq \quad & \sum_{u} P_{U^n}(u) \frac{\exp(n\rho(R_1 - R_1') + \psi(\rho, P_{Z^n|V^n}, \sum_{\lambda} P_{\Lambda|U^n=u}(\lambda) P_{V^n|U^n=u,\Lambda=\lambda}))}{\rho}
\end{aligned}
$$

(concavity of $\exp(\psi)$ is used)

Single-letterize the formula

$$\begin{aligned}
& \sum_{u} P_{U^n}(u) \frac{\exp(n\rho(R_1 - R_1') + \psi(\rho, P_{Z^n|V^n}, \sum_{\lambda} P_{\Lambda|U^n=u}(\lambda) P_{V^n|U^n=u,\Lambda=\lambda}))}{\rho} \\
= & \frac{1}{\rho} \sum_{u^n \in \mathcal{U}^n} P_{U^n}(u^n) \exp(n\rho(R_1 - R_1') + \psi(\rho, P_{Z^n|V^n}, P_{V^n|U^n=u^n})) \\
= & \frac{1}{\rho} \sum_{u^n \in \mathcal{U}^n} \prod_{i=1}^{n} P_U(u_i) \exp(\rho(R_1 - R_1') + \psi(\rho, P_{Z|V}, P_{V|U=u_i})) \\
= & \frac{1}{\rho} \prod_{i=1}^{n} \sum_{u_i \in \mathcal{U}} P_U(u_i) \exp(\rho(R_1 - R_1') + \psi(\rho, P_{Z|V}, P_{V|U=u_i})) \\
= & \frac{1}{\rho} \left[\exp(\rho(R_1 - R_1')) \left(\sum_{u \in \mathcal{U}} P_U(u) \exp(\psi(\rho, P_{Z|V}, P_{V|U=u})) \right) \right]^n
\end{aligned}$$

Under what condition the upper bound goes to zero?

Taking the logarithm of the upper bound we have

$$-\log\rho + n\rho\left[R_1 - R_1' + \underbrace{\frac{1}{\rho}\log\left(\sum_{u\in\mathcal{U}} P_U(u)\exp(\psi(\rho, P_{Z|V}, P_{V|U=u}))\right)}_{(*)}\right]$$

I will show that (*) → $I(V;Z|U)$ as $\rho \to 0$.
This shows that the amount of random garbage required to make $S_n = F_n(B_n)$ secret from Eve is $I(V;Z|U)$ per channel use. By choosing $R_0 = \min\{I(U,Y), I(U,Z)\} - \delta$ and $R_1' = I(V;Y|U) - \delta$, we have completed the direct part proof.

The upper bound goes to $I(V;Z|U)$

I will use l'Hôpital's rule to find the limit of (*). Switching to the Cover-Thomas notation.

$$\begin{aligned}\text{numerator of } (*) &= \log\left[\sum_{u\in\mathcal{U}} P_U(u)\exp(\psi(\rho, P_{Z|V}, P_{V|U=u}))\right] \\ &= \log\left[\sum_{u,v,z} p(u,v,z)p(z|v)^{\rho}p(z|u)^{-\rho}\right]\end{aligned}$$

Derivating the numerator w.r.t. ρ and substituting $\rho = 0$, we have

$$\begin{aligned}\sum_{u,v,z} p(u,v,z)\log\frac{p(z|v)}{p(z|u)} &= \sum_{u,v,z} p(u,v,z)\log\frac{p(z|v)p(v|u)}{p(z|u)p(v|u)} \\ &= \sum_{u,v,z} p(u,v,z)\log\frac{p(z,v|u)}{p(z|u)p(v|u)} \\ &= I(V;Z|U)\end{aligned}$$

Discussion

The evaluation of $I(S_n; Z^n|F_n)$ is similar to Hayashi (2009) studying the wiretap channel and the secret key agreement. Major differences are

- We evaluated $I(S_n; Z^n, E_n|F_n)$ instead of $I(S_n; Z^n|F_n)$. Giving the common message E_n to Eve does not worsen our evaluation.
- We do not move averaging over U into ψ while moving that over Λ into ψ. Otherwise we get much worse upper bound.

Overall, this presentation is not so technically novel relative to Hayashi (2009).

The evaluation of mutual information is independent of that of decoding error probability. We can combine our argument with the best research on the decoding error probability for BC with degraded message sets, provided that the type of random coding is the same. Provision of secrecy is separated from error correction.

The proposed method is universal in the sense that it makes the mutual information small as far as $R_1' - R_1 > I(V; Z|U)$.

Practical construction of secrecy codes

Suppose that we are given single pair of encoder and decoder for a broadcast channel with degraded message sets. We want to construct a code for BC with confidential messages. If we do this, then the practical study of codes for BC with confidential messages becomes unnecessary.

The size of secret message set $\mathcal{S}$ must satisfy

$$\min_{0<\rho\leq 1} \frac{|\mathcal{S}|^{\rho}\mathbf{E}[P_{X|Z}(X|Z)^{\rho}]}{\rho} \leq \text{acceptable value},$$

where X is the uniform distribution on the codebook.
When the number of codewords is, say 2^{1000}, evaluation of the left hand side is practically impossible.

I introduce another form of the privacy amplification theorem so that we can compute a suitable size of $\mathcal{S}$. What follows is an extension on our result on the wiretap channel (ISIT 2010).

Another privacy amplification theorem

Assume that the given family of two-universal hash function F from $\mathcal{L}$ to $\mathcal{M}$ satisfies that

$$|F^{-1}(m)| = \frac{|\mathcal{L}|}{|\mathcal{M}|}, \quad \forall m,$$

that the statistically independent random variable K and L obey the uniform distributions on $\mathcal{K}$ and $\mathcal{L}$, respectively, and that a fixed conditional probability $Q_{Z|K,L}$ is given. We also assume that F is statistically independent of K and L. Then,

$$I(F(L);Z|F) = \mathbf{E}_F I(F(L);Z) \leq \frac{|\mathcal{M}|^\rho \exp(\bar{\phi}(\rho, Q_{Z|K,L}))}{(|\mathcal{K}| \times |\mathcal{L}|)^\rho \rho}, \tag{1}$$

for $0 < \rho \leq 1/2$, where $\mathbf{E}_F$ expresses the expectation concerning the random variable F,

$$\bar{\phi}(\rho, Q_{Z|K,L}) = \log \int_{\mathcal{Z}} \left(\mathbf{E}_{KL}(Q_{Z|K,L}(z|K,L)^{1/(1-\rho)}) \right)^{1-\rho} dz$$

and dz is an arbitrary measure.

Proof is omitted

Evaluation of the mutual information

$\mathcal{B}_n$: set of the private messages (not secret)
$\mathcal{E}_n$: set of the common messages
$e_n : \mathcal{B}_n \times \mathcal{E}_n \to \mathcal{V}^n$: encoder for BC with degraded message sets
B_n: uniform RV on $\mathcal{B}_n$
E_n: uniform RV on $\mathcal{E}_n$, independent of B_n
F_n: hash function
Z^n: Eve's received signal

By another privacy amplification theorem we have

$$I(F_n(B_n); Z_n|F_n) \leq \frac{|\mathcal{S}_n|^\rho \exp(\phi(\rho, P^n_{Z|V}, P_{e_n(B_n,E_n)}))}{|\mathcal{B}_n \times \mathcal{E}_n|^\rho \rho},$$

where

$$\phi(\rho, Q_{Z|V}, P_V) = \ln \sum_z \left(\sum_{v \in \mathcal{V}} P_V(v)(Q_{Z|V}(z|v))^{1/(1-\rho)} \right)^{1-\rho}.$$

This is essentially Gallager's function E_0.

Numerical evaluation of ϕ

$$
\begin{aligned}
& \exp(\phi(\rho, P^n_{Z|V}, P_{e_n(B_n,E_n)})) \\
\leq \; & \max_{P_n \text{ on } \mathcal{V}^n} \exp(\phi(\rho, P^n_{Z|V}, P_{e_n(B_n,E_n)})) \\
= \; & \max_{P_1 \text{ on } \mathcal{V}^1} \underbrace{\exp(n\phi(\rho, P^n_{Z|V}, P_{e_n(B_n,E_n)}))}_{(**)} \text{ by Arimoto (1973)}
\end{aligned}
$$

- (**) is concave w.r.t. P_1, so its maximization can be computed in practice.
- $\min_{P_1}$(**) is convex w.r.t. ρ, so its minimization w.r.t. ρ can also be computed.

Under which condition the second construction achieves a rate pair?

Since ϕ is Gallager's function, $\lim_{\rho \to 0} \phi(\rho, P_{Z|V}, P_V)/\rho = I(V;Z)$. The mutual information $I(F_n(B_n); Z^n | F_n)$ goes to zero if $R_0 + R_1' - R_1 > \max I(V;Z)$.

Sufficient condition to achieve (R_0, R_1)

- (R_0, R_1') is achievable in the BC $(P_{YZ|V})$ with degraded message sets with some artificial noise $P_{X|V}$, and
- $R_1 < R_0 + R_1' - \max_{P_V} I(V;Z)$.

What happens if we use the second PA theorem in place of the first

The second PA theorem ensures the mutual information to be small if $R_0 + R_1' - R_1 > I(V;Z)$. When $R_0 \simeq I(U,Z)$ and $R_1' = I(V;Y|U)$, by noting $I(V;Z) = I(U,V;Z) = I(U;Z) + I(V;Z|U)$, we have

$$\begin{aligned} & R_0 + R_1' - R_1 = I(V;Z) \\ \Longleftrightarrow\quad & I(U;Z) + I(V;Y|U) - R_1 = I(V;Z) \\ \Longleftrightarrow\quad & R_1 = I(V;Y|U) - I(V;Z|U). \end{aligned}$$

This means that if $I(U;Y) < I(U;Z)$ then we cannot achieve that rate pairs.

Some References

Arimoto 1973 S. Arimoto, “On the converse to the coding theorem for discrete memoryless channels,” *IEEE Trans. Inform. Theory*, vol. 19, no. 3, pp. 357–359, May 1973.

Hayashi 2009 M. Hayashi, “Exponential decreasing rate of leaked information in universal random privacy amplification,” to be published in IEEE Trans. Inform. Theory, arXiv:0904.0308.

Maurer 1994 U. M. Maurer, The Strong Secret Key Rate of Discrete Random Triples, Communication and Cryptography — Two Sides of One Tapestry, (R. Blahut et al., Eds.), Kluwer Academic Publishers, pp. 271-285, 1994.