Information-Theoretic Key Agreement of Multiple Terminals - Part II: Channel Model

Amin Aminzadeh Gohari and Venkat Anantharam[1]

[1] Department of Electrical Engineering and Computer Science

University of California, Berkeley

{aminzade,ananth}@eecs.berkeley.edu

Abstract

This is the second part of a two-part paper on information-theoretically secure secret key agreement. This paper focuses on the secret key rate problem under the *channel model*. In the channel model, a set of two or more terminals wish to create a shared secret key that is information-theoretically secure from an eavesdropper. The first terminal can choose a sequence of inputs to a discrete memoryless broadcast channel, which has outputs at the other terminals and at the eavesdropper. After each channel use, the terminals can engage in arbitrarily many rounds of interactive authenticated communication over a public channel; thus, each input by the first terminal can depend on the previous inputs and the public communication so far. At the end of the process each terminal should be able to generate the key. We introduce a technique for proving that a given expression bounds the secrecy rate from above. Using this technique, a new upper bound on the secrecy rate in the general multi-terminal case is proposed that strictly improves the currently best known upper bound. We also derive a new lower bound on the secrecy rate and prove that it strictly improves what is essentially the best known lower bound.

Keywords: Secret key agreement, unconditional security, secrecy capacity, common randomness, public discussion, channel model, security.

I. Introduction

In this paper, we study the problem of determining the maximum information-theoretically secure secret key rate against a passive eavesdropper in a well-known setting in the information-theoretic security literature, called the *channel model*.

The history of this model dates back to an early work by Wyner [10], who studied what may be called a “degraded broadcast scenario”. In this setting Alice is connected to Bob by a discrete memoryless

channel. The eavesdropper, Eve, receives a noisy version of the output at Bob's end. In a subsequent work, Csiszár and Körner [3] generalized Wyner's model by assuming that Alice is connected to Bob and Eve through a broadcast channel. The channel from Alice to Eve in this model is not necessarily a degraded version of the channel between Alice and Bob. In this scenario, the secrecy rate, as one might expect, would be zero if the channel from Alice to Eve is stronger than the channel from Alice to Bob.

The scenario considered by Csiszár and Körner was further generalized by Maurer [7], who made the interesting observation that even if the channel from Alice to Eve is stronger than the channel from Alice to Bob, Alice and Bob may still be able to generate a common secret key if we allow Bob to send authenticated but public messages to Alice. This observation led to the formulation of the two main models in this area, introduced by the works of Ahlswede and Csiszár [1], Csiszár and Narayan [4] and Maurer [7], the *source model* and *channel model*. In this paper, we focus on the channel model. There are m terminals interested in secret key generation against an adversary Eve. A discrete memoryless broadcast channel exists from the first terminal to all other terminals, and to Eve. The input to the channel is governed by the first terminal while the other terminals, as well as Eve, observe the outputs of the broadcast channel at their end. In what is traditionally called the channel model, after each use of the channel by the first terminal, all the m terminals are allowed to engage in arbitrary many rounds of interactive authenticated communication over a public channel. We consider a generalization of this where only the first u terminals ($1 \leq u \leq m$) are allowed such communication; terminals $u+1 \leq i \leq m$ listen and must participate in secret key generation, but cannot talk. This generalization is motivated by the desire to put one-way capacity and interactive capacity on the same footing, and fits naturally with the corresponding generalization that we made in the source model [6]. Note that we assume, mostly for notational convenience, that terminal 1 is allowed to participate in the interactive authenticated public communication.

Note that each input to the broadcast channel by the first terminal is allowed to depend on the past inputs and on the public communication so far. At the end of the entire process, i.e. of the n inputs and of the interactive public communication after each input, each terminal $1 \leq i \leq m$ generates random variable S_i as its secret key. All S_i's should with high probability be equal to each other and they should be approximately independent of Eve's whole information after the communication, i.e. the n outputs at Eve's end of the broadcast channel, and the entire public discussion. The achieved secret key rate would then be roughly $\frac{1}{n}H(S_1)$. The highest achievable secret key rate, asymptotically in n, is called the secrecy capacity. For a precise formulation see section 2.

Calculation of the exact secrecy capacity remains an unsolved problem, although some lower and

upper bounds on this quantity are known. For the case of $m = 2$, the best known upper bound explicitly mentioned in the literature, as far as we are aware, is $\min[\sup_{p(x)} I(X;Y), \sup_{p(x)} I(X;Y|Z))]$, which was proposed by Maurer [7]. This can however be easily generalized to $\inf_{\overline{Z}-Z-XY}[\sup_{p(x)} I(X;Y|\overline{Z})]$. The best know lower bound, as far as we are aware, is

$\sup_{p(x)} \max\{\sup_{V-U-X-YZ}[I(U;Y|V) - I(U;Z|V)], \sup_{V-U-Y-XZ}[I(U;X|V) - I(U;Z|V)]\}$, which one can find in [4], [7].

In this paper, we improve the above mentioned upper bounds on the secret key rate. Our proof technique is similar to the one for proving upper bounds in the first part of this paper [6], but this paper can be read independently of [6]. The idea is to define a potential function and show that for any valid secret key generating protocol, the potential function starts from the upper bound and decreases as we move along the protocol, and eventually becomes equal to the achieved secret key rate of the protocol.

We also derive lower bounds on the secrecy rate by exploiting our new lower bound on the secrecy rate in the source model, which was proved in [6]. An example is provided to show that the new bound is strictly better than $\inf_{\overline{Z}-Z-XY}[\sup_{p(x)} I(X;Y|\overline{Z})]$.

The outline of this paper is as follows. In section 2, we introduce the basic notations and definitions used in this paper. Section 3 contains the main results of this paper. This is followed by section 4 and two appendices which give proofs for the results.

II. Definitions and Notation

Throughout this paper we assume X_1, X_2, ..., X_m and Z are $m + 1$ possibly dependent random variables each taking values from a finite set.

Our multi-terminal channel model is basically the same as the one studied in the literature, see e.g. [8]. We however relax the uniformity condition on the generated secret key. Maurer in [7] argued that the assumption of uniformity could always be added without loss of generality. We study the weak notion of secrecy throughout this paper and assume that all m terminals are interested in secret key generation. It is known that the weak and strong secret key rates are equal [8]. Our model is however somewhat more general in the sense that it assumes that only some of the terminals are able to participate in the public discussion. Throughout this paper, we assume that terminals $1, 2, ..., u$ ($1 \leq u \leq m$) are allowed to talk while terminals $u + 1, u + 2, ..., m$ are silent.

Given an ordered sequence of n random variables taking values from some finite set we denote the i^{th} of these by notation such as $X(i)$. We write $X^{1:i}$ for $(X(1), X(2), ..., X(i))$. For $X^{1:n}$ we will often instead write X^n.

Definition 1. Let $\mathrm{q}(\mathrm{x}_2, \mathrm{x}_3, ..., \mathrm{x}_\mathrm{m}, \mathrm{z}|\mathrm{x}_1)$ be a conditional distribution, n be a natural number, ϵ be a positive real number, $\overrightarrow{C} = (\overrightarrow{C}_1, \overrightarrow{C}_2, ..., \overrightarrow{C}_n)$ be a collection of n finite sets of discrete random variables $\overrightarrow{C}_i : i = 1, 2, ..., n$. Each $\overrightarrow{C}_i$ is a finite set of discrete random variables: $\overrightarrow{C}_i = (\overrightarrow{C}_i(1), \overrightarrow{C}_i(2), ..., \overrightarrow{C}_i(r(i)))$. Let M_1, M_2, ..., M_u, X_1^n, X_2^n, ..., X_m^n, Z^n and S_1, S_2, ..., S_m be $u + (m+1)n + m$ discrete random variables.

Consider the following conditions:

1) For $i = 1, 2, ..., n$:

   $p(X_2(i) = x_2(i), ..., X_m(i) = x_m(i), Z(i) = z(i)|X_1^{1:i} = x_1^{1:i}, X_2^{1:i-1} = x_2^{1:i-1}, ..., X_m^{1:i-1} = x_m^{1:i-1}, Z^{1:i-1} = z^{1:i-1}, M_1 = m_1, ..., M_u = m_u) =$

   $\mathrm{q}(\mathrm{x}_2(\mathrm{i}), ..., \mathrm{x}_\mathrm{m}(\mathrm{i}), \mathrm{z}(\mathrm{i})|\mathrm{x}_1(\mathrm{i}))$;

2) For $i = 1, ..., n$:

   $H(X_1(i)|\overrightarrow{C}_1, \overrightarrow{C}_2, ..., \overrightarrow{C}_{i-1}, M_1, X_1^{1:i-1}) = 0$;

3) $p(M_1...M_u X_1(1), X_2(1), ..., X_m(1), Z(1)) = p(M_1)...p(M_u)p(X_1(1), X_2(1), ..., X_m(1), Z(1))$;

4) $H(\overrightarrow{C}_i(j)|\overrightarrow{C}_1, \overrightarrow{C}_2, ..., \overrightarrow{C}_{i-1}\overrightarrow{C}_i^{1:j-1}X_s^{1:i}M_s) = 0 \quad \forall s : 1 \leq s \leq u, \ s = j$ modulo m. This means that the indexing of the communications is done in round robin order and each communication is adapted to the available information of the communicator.

   Furthermore, $\overrightarrow{C}_i(j) = 0 \ \forall i, j, s : j = s$ modulo m and $s > u$. This means that s-th terminal is not allowed to participate in the communication;

5) $H(S_i|\overrightarrow{C}, X_i^n M_i) = 0$ for $1 \leq i \leq u$

   $H(S_i|\overrightarrow{C}, X_i^n) = 0$ for $u + 1 \leq i \leq m$.

   This means that S_i is created by i-th terminal at the end of the entire process;

6) $P(S_1 = S_2 = S_3 = ... = S_m) > 1 - \epsilon$. This ensures the reliability of the generated keys;

7) $\frac{1}{n} I(S_1; Z^n, \overrightarrow{C}) < \epsilon$. This ensures that the generated key is almost hidden from the eavesdropper.

Intuitively, n represents the number of communication rounds; $\overrightarrow{C}_i$ represents communications at the i-th stage; M_1, M_2, ..., M_u represents external randomness provided to the first u terminals.

The data typing condition SK_C(n, ϵ, S_1, S_2, S_3,..., S_m, $\overrightarrow{C}$, $M_1, M_2, ..., M_u$, $X_1^n, X_2^n, ..., X_m^n, Z^n$) is said to hold iff all above-mentioned conditions are satisfied. To any SK_C data type, we assign a number called the *gain* of the SK_C data type which is defined as $\frac{1}{n} H(S_1)$. •

Definition 2: $C_{CH}^{\epsilon}(u, \mathrm{q}(\mathrm{x}_2\mathrm{x}_3...\mathrm{x}_\mathrm{m}\mathrm{z}|\mathrm{x}_1))$, the ϵ secret key rate, is defined as:

$$\limsup_{n \to \infty} \sup_{SK_C(n, \epsilon, S_1, S_2, S_3, ..., S_m, \overrightarrow{C}, M_1, ..., M_u, X_1^n, X_2^n, ..., X_m^n, Z^n)} Gain(SK)$$

•

Definition 3: $C_{CH}(u, \mathrm{q}(\mathrm{x_2x_3...x_mz|x_1}))$, the channel model secret key rate, is defined as:

$$\lim_{\epsilon \to 0} C^{\epsilon}_{CH}(u, \mathrm{q}(\mathrm{x_2x_3...x_mz|x_1}))$$

●

Note that we have allowed the first user to participate in the public discussion and to randomize. Further, all the terminals who participate in the public discussion, i.e. terminals $1 \leq i \leq u$, are allowed to randomize. The assumption on the participation of the first terminal in the public discussion can be removed but this terminal must be allowed to randomize. Otherwise, the inputs to the broadcast channel will be always a deterministic function of the public communication and thus known to the eavesdropper, resulting in zero secret key rate. It is legitimate to differentiate between the ability to randomize and the ability to participate in the public discussion as long as the first user is concerned. For the sake of notational simplicity, however, we allow the first user to participate in the public discussion.

III. Statement of the results

In this section, the main results of the paper are formally presented as theorems 1 through 4. Following each result there is an informal discussion of it in order to give an intuitive feeling for the result.

Theorem 1. For each $j \geq 1$, let $\varphi_j(p(x_1, x_2, ..., x_m, z))$ be a real-valued function from the set of *all* probability distributions defined on a product of *any* $m+1$ finite sets. We sometimes use the notation $\varphi_j(X_1; X_2; X_3; ...; X_m \| Z)$ to refer to $\varphi_j(p(x_1, x_2, ..., x_m, z))$ when $(X_1, X_2, ..., X_m, Z)$ has the law $p(x_1, ..., x_m, z)$. For any conditional distribution $\mathrm{q}(\mathrm{x_2, x_3, ..., x_m, z|x_1})$,

$$\phi(\mathrm{q}(\mathrm{x_2, x_3, ..., x_m, z|x_1})) = \sup_{q(x_1)} \varphi_1(q(x_1).\mathrm{q}(\mathrm{x_2, x_3, ..., x_m, z|x_1}))$$

would be an upper bound on $C_{CH}(u, \mathrm{q}(\mathrm{x_2, x_3, ...x_m, z|x_1}))$, the channel model secrecy rate assuming that only the first u terminals are permitted to talk, if $\varphi_j (j = 1, 2, ...)$ satisfy the following for all $p(x_1, x_2, ..., x_m, z)$:

1) Whenever

   $H(X'_1|X_1) = 0$ and

   $X_1X_2...X_mZ - X_1 - X'_1 - X'_1X'_2...X'_mZ'$ and

   $p(x'_2, x'_3, ..., x'_m, z'|x'_1) = \mathrm{q}(\mathrm{x_2, x_3, ..., x_m, z|x_1})$

   are true, we have:

   $\varphi_{j+1}(X_1X'_1; X_2X'_2; ...; X_mX'_m \| ZZ') \leq \varphi_j(X_1; X_2; ...; X_m \| Z) + \phi(\mathrm{q}(\mathrm{x_2, x_3, ..., x_m, z|x_1}))$;

2) For any random variable F such that $\exists i \leq u : H(F|X_i) = 0$, we have:

   $\varphi_j(X_1; X_2; ...; X_m \| Z) \geq \varphi_j(X_1F; X_2F; ...; X_mF \| ZF)$;

3) For any random variables $X_1', X_2', ..., X_m'$ such that $\forall i : H(X_i'|X_i) = 0$, we have:

$\varphi_j(X_1; X_2; ...; X_m \| Z) \geq \varphi_j(X_1'; X_2'; ...; X_m' \| Z)$;

4) $\varphi_j(X_1; X_2; ...; X_m \| Z) \geq H(X_1|Z) - \sum_{i=2}^{m} H(X_1|X_i)$;

5) Whenever for random variables M_1, M_2, ..., M_u

$p(M_1, M_2, ..., M_u, X_1, X_2, X_3, ..., X_m, Z) = p(M_1)p(M_2)...p(M_u)p(X_1, X_2, X_3, ..., X_m, Z)$

is true, we have:

$\varphi_j(X_1; X_2; ...; X_m \| Z) \geq \varphi_j(M_1X_1; M_2X_2; ...; M_uX_u; X_{u+1}; ...; X_m \| Z)$.

Discussion: For each $j \geq 1$, the quantity $\varphi_j(p(x_1, x_2, ..., x_m, z))$ can be intuitively understood as representing the secret key rate per channel use that is possible if we insist on first using j channel uses to create the distribution $p(x_1, x_2, ..., x_m, z)$ and then work with this distribution as the "raw" joint distribution across a new discrete memoryless channel. With this rough picture in mind, condition 1 can be understood as saying that having already insisted on working with a j-channel use $p(x_1, x_2, ..., x_m, z)$, one more use of the channel can at most buy us the channel capacity on a per use basis. Condition 2 says that further insistence on working with a distribution that results from a particular kind of use of the authenticated public channel by any terminal $1 \leq i \leq u$ cannot increase the per channel use secrecy rate. Condition 3 has a similar interpretation. Each of these conditions has been stated only for the case where the corresponding maps are deterministic; this is sufficient because the possibility of randomization by any of the first u terminals is covered by condition 5. The right hand side of condition 4 is a convenient expression that is easily seen to be a lower bound on the corresponding secrecy rate; other such expressions would have worked as well. Finally, condition 5 would apply if independent randomization was freely available to the terminals who can talk, i.e. terminals $1 \leq i \leq u$. ●

Theorem 2. Let $[m]$ and $[u]$ respectively denote the sets $\{1, 2, ..., m\}$, $\{1, 2, ..., u\}$. For every $\Lambda = (\lambda_B, B \subseteq [m])$ such that for each u-tuple $(R_1, R_2, ..., R_u)$ of real numbers we have

$\sum_{B:B\subset[m], B\bigcap[u]\neq\emptyset, B\neq[m]} \lambda_B \sum_{j\in B\bigcap[u]} R_j = \sum_{j=1}^{u} R_j$,

the following inequality holds:

$$C_{CH}(u, \mathrm{q}(\mathrm{x_2}, \mathrm{x_3}, ...\mathrm{x_m}, \mathrm{z}|\mathrm{x_1})) \leq$$
$$\sup_{p(x_1)}\{\inf_{p(J|X_1,...,X_m,Z)} \big([H(X_1...X_u|J) - \tau^{\Lambda}(X_1, X_2, ..., X_u, X_{u+1}^{(s)}, ..., X_m^{(s)} \| J) + I(X_1X_2...X_m; J|Z)]\big)\}.$$

In this expression $(X_1, X_2, ..., X_m, J, Z)$ have the law $p(x_1)\mathrm{q}(\mathrm{x_2}, ..., \mathrm{x_m}, \mathrm{z}|\mathrm{x_1})p(j|x_1, ..., x_m, z)$, Λ is the mnemonic for $(\lambda_B, B \subseteq [m])$, and $\tau^{\Lambda}(X_1, X_2, ..., X_u, X_{u+1}^{(s)}, ..., X_m^{(s)} \| J)$ is defined as

$\sum_{B:B\subset[m], B\bigcap[u]\neq\emptyset, B\neq[m]} \lambda_B H(X_{B\bigcap[u]}|X_{B^c}J)$.

Discussion: The above upper bound can be written as the infimum over the set of all valid Λ of

$$\sup_{p(x_1)}\{\inf_{p(J|X_1,...,X_m,Z)}\big([H(X_1...X_u|J)-\tau^{\Lambda}(X_1,X_2,...,X_u,X_{u+1}^{(s)},...,X_m^{(s)}\|J)+ I(X_1X_2...X_m;J|Z)]\big)\}.$$

If the infimum over Λ is swapped with the supremum over $p(x_1)$, one gets the following lower bound on our upper bound by applying theorem 6 of the first part of this paper:

$$\sup_{p(x_1)}\{\inf_{p(J|X_1,...,X_m,Z)}\big([S(X_1J;X_2J;...;X_uJ;(X_{u+1}J)^{(s)}...;(X_mJ)^{(s)}\|J)+ I(X_1X_2...X_m;J|Z)]\big)\}.$$

For the notation see [6]. We were not able to prove that this smaller expression is an upper bound on $C_{CH}(u,\mathrm{q}(\mathrm{x}_2,\mathrm{x}_3,...\mathrm{x}_\mathrm{m},\mathrm{z}|\mathrm{x}_1))$. •

Theorem 3. In the case of $m=2$, the new upper bound on $C_{CH}(2,\mathrm{q}(\mathrm{y},\mathrm{z}|\mathrm{x}))$ equals

$$\sup_{p(x)}\inf_J[I(X;Y|J)+I(XY;J|Z)].$$

This bound strictly improves the $\min[\sup_{p(x)} I(X;Y),\sup_{p(x)} I(X;Y\|Z))]$ bound proposed by Maurer [7]. It also improves the stronger upper bound $\inf_{\overline{Z}-Z-XY}\sup_{p(x)} I(X;Y|\overline{Z})$ mentioned in the introduction.

Discussion: $\inf_{\overline{Z}-Z-XY}\sup_{p(x)} I(X;Y|\overline{Z})$ is an upper bound on $C_{CH}(2,\mathrm{q}(\mathrm{y},\mathrm{z}|\mathrm{x}))$ because for every choice of $p(\overline{z}|z)$ the channel model secrecy rate is no bigger than $\sup_{p(x)} I(X;Y|\overline{Z})$. We will in fact prove that the new bound is strictly smaller than $\sup_{p(x)}\inf_{\overline{Z}-Z-XY} I(X;Y|\overline{Z})$, which in turn is no bigger than $\inf_{\overline{Z}-Z-XY}\sup_{p(x)} I(X;Y|\overline{Z})$. •

Theorem 4. Assume that $q\leq p$ are two arbitrary natural numbers and $(U_1,U_2,...,U_p)$ are arbitrary random variables satisfying the following properties:

- U_i $(i=1,2,...,p)$ takes values from a finite set;
- $p(U_1,U_2,...,U_p|X_1,X_2,X_3,...,X_m,Z)=\prod_{i=1}^{p} p(U_i|U_{1:i-1}X_{i \bmod m})$;
- For all $r>u$, we have $U_i=0\ \forall i: i-r\equiv^m 0$.

$C_{CH}(u,\mathrm{q}(\mathrm{x}_2,\mathrm{x}_3,...\mathrm{x}_\mathrm{m},\mathrm{z}|\mathrm{x}_1))$ is bounded from below by

$$\sup_{p(x_1)}\sum_{j=q}^{p}[\min_{1\leq r\leq m} I(U_j;X_r|U_{1:j-1})-I(U_j;Z|U_{1:j-1})]$$

where $(X_1,X_2,...,X_m,Z,U_1,...,U_p)$ inside the supremum has joint distribution

$$p(X_1)\mathrm{q}(\mathrm{X}_2,\mathrm{X}_3,...\mathrm{X}_\mathrm{m},\mathrm{z}|\mathrm{X}_1)p(U_1,U_2,...,U_p|X_1,X_2,X_3,...,X_m,Z).$$

In the case of $m=2$, the new lower bound on $C_{CH}(2,\mathrm{q}(\mathrm{y},\mathrm{z}|\mathrm{x}))$ derived by taking supremum over all valid $(q,p,U_1,U_2,...,U_p)$ strictly improves the $\sup_{p(x)}[\max(S(X;Y^{(s)}\|Z),S(X^{(s)};Y\|Z))]$ lower bound, where in this expression, $S(X;Y^{(s)}\|Z)$ is the source-model one way secrecy rate from X to Y in the presence of Z, see [6].

In this Theorem, q (written in italics) is a non-negative integer and should not be confused with the conditional distribution $\mathrm{q}(\mathrm{x}_2,\mathrm{x}_3,...\mathrm{x}_\mathrm{m},\mathrm{z}|\mathrm{x}_1)$.

Discussion: $C_{CH}(u, \mathrm{q}(\mathrm{x}_2, \mathrm{x}_3, ...\mathrm{x}_\mathrm{m}, \mathrm{z}|\mathrm{x}_1))$ is bounded from below by

$\sup_{p(x_1)} S(X_1; X_2; ...; X_u; X_{u+1}^{(s)}; ...; X_m^{(s)} \| Z)$ (see [6] for the notation)

because the first terminal can always insert i.i.d. repetitions of any $p(x_1)$ at the input of the broadcast channel[7]. We then apply theorem 7 of the first part of this paper [6] to bound from below $S(X_1; X_2; ...; X_u; X_{u+1}^{(s)}; ...; X_m^{(s)} \| Z)$ by

$\sum_{j=q}^{p} [\min_{1 \le r \le m} I(U_j; X_r | U_{1:j-1}) - I(U_j; Z | U_{1:j-1})]$.

The proof for this theorem mainly involves proving that in the case of $m = 2$, the new lower bound strictly improves the $\sup_{p(x)} [\max(S(X; Y^{(s)} \| Z), S(X^{(s)}; Y \| Z))]$ lower bound. •

IV. Proofs of Theorems 1-4

Proof of Theorem 1. Fix a probability distribution $\mathrm{q}(\mathrm{x}_2, \mathrm{x}_3, ..., \mathrm{x}_\mathrm{m}, \mathrm{z}|\mathrm{x}_1)$ and assume that $X_1, X_2, ..., X_m$ and Z take values from the discrete finite sets Δ_i, $i = 1...m+1$. For every $\delta > 0$ and $\epsilon > 0$, one can find a valid data type $\mathrm{SK}_C(n, \epsilon, S_1, S_2, S_3, ..., S_m, \overrightarrow{C}, M_1, M_2, ..., M_u, X_1^n, X_2^n, ..., X_m^n, Z^n)$ whose gain is within δ of $C_{CH}^{\epsilon}(u, \mathrm{q}(\mathrm{x}_2, ..., \mathrm{x}_\mathrm{m}, \mathrm{z}|\mathrm{x}_1))$.

We have:

$n\phi(q(x_2, x_3, ..., x_m, z|x_1)) \ge$

$(n-1)\phi(q(x_2, x_3, ..., x_m, z|x_1)) + \varphi_1(X_1^1; X_2^1; ...; X_m^1 \| Z^1) \ge$

$(n-1)\phi(q(x_2, x_3, ..., x_m, z|x_1)) + \varphi_1(M_1 X_1^1; M_2 X_2^1; ...; M_u X_u^1; X_{u+1}^1 ... X_m^1 \| Z^1) \ge$

$(n-1)\phi(q(x_2, x_3, ..., x_m, z|x_1))$
$\quad + \varphi_1(M_1 X_1^1 \overrightarrow{C}_1; M_2 X_2^1 \overrightarrow{C}_1; ...; M_u X_u^1 \overrightarrow{C}_1; X_{u+1}^1 \overrightarrow{C}_1 ... X_m^1 \overrightarrow{C}_1 \| Z^1 \overrightarrow{C}_1) \ge^i$

$(n-2)\phi(q(x_2, x_3, ..., x_m, z|x_1))$
$\quad + \varphi_2(M_1 X_1^{1:2} \overrightarrow{C}_1; M_2 X_2^{1:2} \overrightarrow{C}_1; ...; M_u X_u^{1:2} \overrightarrow{C}_1; X_{u+1}^{1:2} \overrightarrow{C}_1 ... X_m^{1:2} \overrightarrow{C}_1 \| Z^{1:2} \overrightarrow{C}_1) \ge$

$(n-2)\phi(q(x_2, x_3, ..., x_m, z|x_1))$
$\quad + \varphi_2(M_1 X_1^{1:2} \overrightarrow{C}_{1:2}; M_2 X_2^{1:2} \overrightarrow{C}_{1:2}; ...; M_u X_u^{1:2} \overrightarrow{C}_{1:2}; X_{u+1}^{1:2} \overrightarrow{C}_{1:2} ... X_m^{1:2} \overrightarrow{C}_{1:2} \| Z^{1:2} \overrightarrow{C}_{1:2}) \ge$

$(n-3)\phi(q(x_2, x_3, ..., x_m, z|x_1))$
$\quad + \varphi_3(M_1 X_1^{1:3} \overrightarrow{C}_{1:2}; M_2 X_2^{1:3} \overrightarrow{C}_{1:2}; ...; M_u X_u^{1:3} \overrightarrow{C}_{1:2}; X_{u+1}^{1:3} \overrightarrow{C}_{1:2} ... X_m^{1:3} \overrightarrow{C}_{1:2} \| Z^{1:3} \overrightarrow{C}_{1:2}) \ge$

...

$\varphi_n(M_1 X_1^{1:n} \overrightarrow{C}_{1:n}; M_2 X_2^{1:n} \overrightarrow{C}_{1:n}; ...; M_u X_u^{1:n} \overrightarrow{C}_{1:n}; X_{u+1}^{1:n} \overrightarrow{C}_{1:n} ... X_m^{1:n} \overrightarrow{C}_{1:n} \| Z^{1:n} \overrightarrow{C}_{1:n}) \ge$

$\varphi_n(M_1 X_1^{1:n} \overrightarrow{C}_{1:n}; M_2 X_2^{1:n} \overrightarrow{C}_{1:n}; ...; M_u X_u^{1:n} \overrightarrow{C}_{1:n}; X_{u+1}^{1:n} \overrightarrow{C}_{1:n} ... X_m^{1:n} \overrightarrow{C}_{1:n} \| Z^{1:n} \overrightarrow{C}_{1:n}) \ge$

$\varphi_n(S_1; S_2; ...; S_m \| Z^{1:n} \overrightarrow{C}_{1:n}) \ge$

$H(S_1 | Z^{1:n} \overrightarrow{C}_{1:n}) - \sum_{j=2}^{m} H(S_1|S_j) \ge$

$nC_{CH}^{\epsilon}(u, \mathrm{q}(\mathrm{x}_2, \mathrm{x}_3, ...\mathrm{x}_\mathrm{m}, \mathrm{z}|\mathrm{x}_1))) - n\delta - (m-1)[h(\epsilon) + \epsilon . n \log \prod_{i=1}^{m} |\Delta_i|]$

The last inequality in this chain is a consequence of Fano's inequality. The inequality i and its analogs are valid because φ satisfies property number 1 of theorem 1. All the other inequalities are consequence of other properties required in theorem 1, in a straightforward way.

The above inequalities show that

$$\phi(X_1; X_2; X_3; ...; X_m \| Z) \geq C_{CH}^{\epsilon}(u, \mathrm{q}(\mathrm{x}_2, \mathrm{x}_3, ...\mathrm{x}_\mathrm{m}, \mathrm{z}|\mathrm{x}_1))) - \delta - \tfrac{m-1}{n}[h(\epsilon) + \epsilon.n \log \textstyle\prod_{i=1}^{m} |\Delta_i|].$$

The first part of the theorem is proved by taking the limit as ϵ and δ go to zero. ●

Proof of Theorem 2. Fix a $\Lambda = (\lambda_B, B \subseteq [m])$ satisfying the conditions of the theorem. In order to prove this theorem, it is enough to verify the five conditions of theorem 1 when for all $j \geq 1$ we set:

$$\begin{aligned} \varphi_j(X_1; X_2; X_3; ...; X_m \| Z) &:= \varphi_1(X_1; X_2; X_3; ...; X_m \| Z) \\ &= \inf_J \big(H(X_1 ... X_u | J) - \tau^{\Lambda}(X_1, X_2, ..., X_u, X_{u+1}^{(s)}, ..., X_m^{(s)} \| J) \\ &\qquad + I(X_1 X_2 ... X_m; J \| Z) \big) \ , \end{aligned} \tag{1}$$

where $\tau^{\Lambda}(X_1, X_2, ..., X_u, X_{u+1}^{(s)}, ..., X_m^{(s)} \| J)$ is as in the statement of the Theorem. In appendix I, we show that this choice satisfies the five conditions of Theorem 1, thus completing the proof. ●

Proof of Theorem 3. The only possible value for $\lambda_{\{1\}}$ and $\lambda_{\{2\}}$ in the case of $m = u = 2$ is one. The upper bound, therefore reduces to $\sup_{p(x)} \inf_J [I(X; Y|J) + I(XY; J|Z)]$. In order to prove that this bound strictly improves $\sup_{p(x)} \inf_{\overline{Z}-Z-XY} I(X; Y|\overline{Z})$ we use the example of Renner and Wolf in [9]. X and Y take values from the set $\{0, 1, 2, 3\}$. Assuming that $P(X = i) = p_i$, Table (I) characterizes the conditional probability distribution of Y given X. The conditional distribution of Z given X and Y is specified by the following equation:

$$Z = \begin{cases} (X \ + \ Y) \mod 2 & \text{if } X \in \{0, 1\} \\ X \qquad\quad \mod 2 & \text{if } X \in \{2, 3\} \end{cases}$$

Renner and Wolf proved that for the choice of $p_i = \frac{1}{4}$ for $i = 0, 1, 2, 3$ and $U = \lfloor \frac{X}{2} \rfloor$, one has

$$I(X; Y \downarrow Z) = \frac{3}{2} \qquad I(X; Y \downarrow ZU) = 0$$

where $I(X; Y \downarrow Z)$, known as the intrinsic information is defined as $\inf_{\overline{Z}-Z-XY} I(X; Y|\overline{Z})$ [9].

Therefore

$$\sup_{p(x)} [I(X; Y \downarrow Z)] \geq \tfrac{3}{2}$$

The proof will be completed if we can show that $\sup_{p(x)} \inf_J [I(X; Y|J) + I(XY; J|Z)] \leq 1$.

TABLE I

JOINT PROBABILITY DISTRIBUTION OF X AND Y

Y \ X	0	1	2	3
0	$\frac{1}{2}p_0$	$\frac{1}{2}p_1$	0	0
1	$\frac{1}{2}p_0$	$\frac{1}{2}p_1$	0	0
2	0	0	p_2	0
3	0	0	0	p_3

Let

$$J_0 = \begin{cases} U & \text{if U=0} \\ UZ & \text{if U=1} \end{cases}$$

We can upper bound $\sup_{p(x)} \inf_J [I(X;Y|J) + I(XY;J|Z)]$ by $\sup_{p(x)} [I(X;Y|J_0) + I(XY;J_0|Z)]$.

Since $I(X;Y|J_0) = 0$ and $I(XY;J_0|Z) \leq 1$ for all $p(x)$, $\sup_{p(x)} \inf_J [I(X;Y|J) + I(XY;J|Z)]$ is less than or equal to one. ●

Proof of Theorem 4. $C_{CH}(u, \mathrm{q}(\mathrm{x}_2, \mathrm{x}_3, ...\mathrm{x}_\mathrm{m}, \mathrm{z}|\mathrm{x}_1))$ is bounded from below by

$\sup_{p(x_1)} S(X_1; X_2; ...; X_u; X_{u+1}^{(s)}; ...; X_m^{(s)} \| Z)$ (see [6] for the notation)

because the first terminal can always insert i.i.d. repetitions of any $p(x_1)$ at the input of the broadcast channel[7]. We apply theorem 7 of the first part of this paper to bound from below

$S(X_1; X_2; ...; X_u; X_{u+1}^{(s)}; ...; X_m^{(s)} \| Z)$ by

$\sum_{j=q}^{p} [\min_{1 \leq r \leq m} I(U_j; X_r | U_{1:j-1}) - I(U_j; Z | U_{1:j-1})]$.

For the case of $m = 2$, we first prove that the new lower bound on $C_{CH}(2, \mathrm{q}(\mathrm{y}, \mathrm{z}|\mathrm{x}))$ is not worse than $\sup_{p(x)} [\max(S(X; Y^{(s)} \| Z), S(X^{(s)}; Y \| Z))]$.

Take some particular $p(x)$, and consider random variables X, Y and Z with the joint distribution $p(x)\mathrm{q}(\mathrm{y}, \mathrm{z}|\mathrm{x})$. Take arbitrary random variables V_1 and V_2 satisfying the Markov chain $V_2 - V_1 - X - YZ$. Let $p = q = 3$ and $U_1 = V_2$, $U_2 = 0$, $U_3 = V_1$. The lower bound achieved by this choice of $p(x)$, p, q and (U_1, U_2, U_3) is $I(V_2; Y|V_1) - I(V_2; Z|V_1)$. Therefore the new lower bound is no worse than $\sup_{p(x)} S(X; Y^{(s)} \| Z)$. It can be similarly proved that the new lower bound is no worse than $\sup_{p(x)} S(X; Y^{(s)} \| Z)$.

Now we construct an example in which the new lower bound strictly improves

$\sup_{p(x)}[\max(S(X;Y^{(s)}\|Z), S(X^{(s)};Y\|Z))]$.

Assume that $X=(X_1,X_2)$, $Y=(Y_1,Y_2)$, $Z=(Z_1,Z_2)$. The conditional distribution of (Y_1,Y_2,Z_1,Z_2) given X_1 and X_2 is defined in figure 1. We prove that the upper bound $\sup_{p(X)} I(X;Y|Z)$ is equal to the new lower bound but is not strictly greater than the previous lower bound.

In appendix II, with reference to figure 1 with $X=(X_1,X_2)$, $Y=(Y_1,Y_2)$ and $Z=(Z_1,Z_2)$, it is shown that for any $0<\epsilon<1$, $I(X;Y|Z)$ strictly increases when

- X_1 and X_2 are not independent and we replace $p(X_1,X_2)p(Y,Z|X)$ with $p(X_1)p(X_2)p(Y,Z|X)$;
- we change the distribution of X_1 to a uniform distribution if X_1 and X_2 are independent but X_1 is not uniform;
- we change the distribution of X_2 to a uniform distribution if X_1 and X_2 are independent but X_2 is not uniform.

But when X_1 and X_2 are independent, the pairs (X_1,Y_1,Z_1) and (X_2,Y_2,Z_2) will become independent and the upper bound $I(X;Y|Z)=I(X_1;Y_1|Z_1)+I(X_2;Y_2|Z_2)=I(X_1;Y_1)-I(X_1;Z_1)+I(Y_2;X_2)-I(Y_2;Z_2)$ will become achievable by the choice of $U_1=X_1$ and $U_2=Y_2$.

Now, we will prove that

$\sup_{p(x)}[\max(S(X;Y^{(s)}\|Z), S(X^{(s)};Y\|Z))]$

is strictly less than

$\sup_{p(X)} I(X;Y|Z)$.

Assume that this is not the case. Since for every choice of $p(x)$, $I(X;Y|Z)$ is as big as

$\max(S(X;Y^{(s)}\|Z), S(X^{(s)};Y\|Z))$

the supremum of the maximum of the two one way rates must happen when X_1 and X_2 are independent and have a uniform distribution. But in the proof of theorem 7 of the first part of this paper, it is shown that under these circumstances $I(X;Y|Z)$ strictly exceeds $\max(S(X;Y^{(s)}\|Z), S(X^{(s)};Y\|Z))$.

●

V. Discussion

We have derived a new lower bound and upper bound on the secrecy rate under the channel model. The latter was proved using a general technique for proving that a certain expression bounds the secrecy rate from above, while the former was proved using the fact that $C_{CH}(u, \mathrm{q}(\mathrm{x}_2,\mathrm{x}_3,...\mathrm{x}_\mathrm{m},\mathrm{z}|\mathrm{x}_1))$ is bounded from below by $\sup_{p(x_1)} S(X_1;X_2;...;X_u;X_{u+1}^{(s)};...;X_m^{(s)}\|Z)$.

The exact relation of the secrecy rate under the channel model and source model remains an open problem. Both the new lower bound and the new upper bound have the generic form of

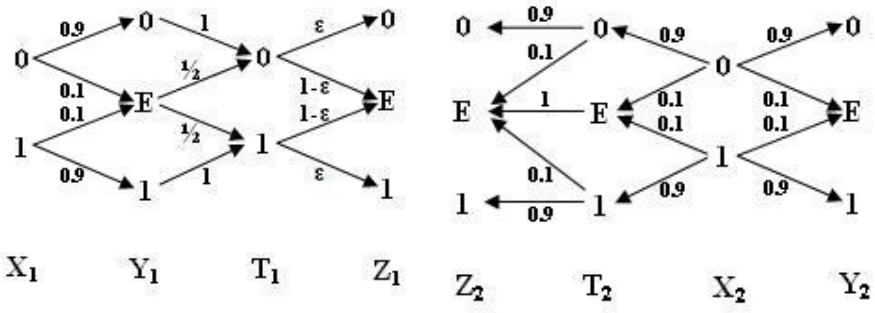

Fig. 1. The conditional distribution of (Y_1, Y_2, Z_1, Z_2) given X_1 and X_2.

$\sup_{p(x_1)} F(p(x_1)\mathrm{q}(\mathrm{x}_2, \mathrm{x}_3, ...\mathrm{x}_\mathrm{m}, \mathrm{z}|\mathrm{x}_1))$.

One can then conjecture that $C_{CH}(u, \mathrm{q}(\mathrm{x}_2, \mathrm{x}_3, ...\mathrm{x}_\mathrm{m}, \mathrm{z}|\mathrm{x}_1))$ equals

$\sup_{p(x_1)} S(X_1; X_2; ...; X_u; X_{u+1}^{(s)}; ...; X_m^{(s)} \| Z)$.

If true, using the theorem 5 of the first part of this paper for $m = 2$, one can bound $C_{CH}(2, \mathrm{q}(\mathrm{y}, \mathrm{z}|\mathrm{x}))$ from above by

$$\sup_{p(x_1)} \inf_J f^{-1}\{f(S(X_1; X_2; ...; X_u; (X_{u+1})^{(s)}; ...; (X_m)^{(s)} \| J)) + S_{f-one-way}(X_1 X_2 ... X_m; J^{(s)} \| Z)\}$$

$f : \mathbb{R}_{\geq 0} \mapsto \mathbb{R}_{\geq 0}$ is an arbitrary strictly increasing convex function and *f-one-way secrecy rate* is defined as

$S_{f-one-way}(X; Y^{(s)} \| Z) = \sup_{V-U-X-YZ}[f(H(U|ZV)) - f(H(U|YV))]$.

We do not know if this expression actually serves as an upper bound on $C_{CH}(2, \mathrm{q}(\mathrm{y}, \mathrm{z}|\mathrm{x}))$ for all appropriate choices of f, or less ambitiously for the particular choice of $f(x) = x$. If it does, it may represent an strict improvement over previous bounds. Otherwise, it will be evidence against the original conjecture.

APPENDIX I

In this appendix, we prove that the φ_j, $j \geq 1$ proposed in eqn. (1) satisfy the five properties of Theorem 1. Recall that $\Lambda = (\lambda_B, B \subseteq [m])$ is assumed to verify the conditions in the statement of theorem 2.

Let

$\theta^\Lambda(X_1; X_2; X_3; ...; X_m; J \| Z) = H(X_1 ... X_u | J) -$

$\tau^\Lambda(X_1, X_2, ..., X_u, X_{u+1}^{(s)}, ..., X_m^{(s)} \| J) + I(X_1 X_2 ... X_m; J \| Z)$

where $\tau^\Lambda(X_1, X_2, ..., X_u, X_{u+1}^{(s)}, ..., X_m^{(s)} \| J)$ is as in the statement of theorem 2.

In this appendix, for any subset B of $[m] = \{1, 2, 3, ..., m\}$, we use the notation X_B in reference to the set of random variables $(X_k, k \in B)$.

Property 1.

It is required to verify that:

$\inf_{\widetilde{J}}(\theta^{\Lambda}(X_1X_1'; X_2X_2'; X_3X_3'; ...; X_mX_m'; \widetilde{J}\|ZZ')) \leq$

$\inf_{\widetilde{J}'}(\theta^{\Lambda}(X_1; X_2; X_3; ...; X_m; \widetilde{J}'\|Z)) + \phi(q(x_2, x_3, ..., x_m, z\|x_1))$.

From the continuity of the relevant information theoretic functions, for any $\epsilon > 0$ one can bound from below

$\phi(q(x_2, x_3, ..., x_m, z\|x_1))$ by

$\theta(X_1'; X_2'; X_3'; ...; X_m'; J''\|Z') - \epsilon$ for some J''.

We will prove that the above inequality holds when we replace $\phi(q(x_2, x_3, ..., x_m, z\|x_1))$ by this lower bound. Without loss of generality, we can further assume that

$\widetilde{J}' - X_1X_2...X_mZ - X_1 - X_1' - X_1'X_2'...X_m'Z' - J''$

because in the corresponding optimization problems depend only

on $p(\widetilde{J}'|X_1X_2...X_mZ)$ and $p(J''|X_1'X_2'...X_m'Z')$.

In order to prove that

$\inf_{\widetilde{J}}\theta^{\Lambda}(X_1X_1'; X_2X_2'; X_3X_3'; ...; X_mX_m'; \widetilde{J}\|ZZ') \leq$

$\inf_{\widetilde{J}'}\theta^{\Lambda}(X_1; X_2; X_3; ...; X_m; \widetilde{J}'\|Z) + \theta^{\Lambda}(X_1'; X_2'; X_3'; ...; X_m'; J''\|Z')$

it would be enough to show that for any arbitrary J' satisfying

$J' - X_1X_2...X_mZ - X_1 - X_1' - X_1'X_2'...X_m'Z' - J''$,

the following inequality holds:

$\theta^{\Lambda}(X_1X_1'; X_2X_2'; X_3X_3'; ...; X_mX_m'; J'J''\|ZZ') \leq$

$\theta^{\Lambda}(X_1; X_2; X_3; ...; X_m; J'\|Z) + \theta^{\Lambda}(X_1'; X_2'; X_3'; ...; X_m'; J''\|Z')$.

We claim that the following two inequalities hold:

$H(X_1...X_uX_1'...X_u'|J', J'') - \tau^{\Lambda}(X_1X_1', X_2X_2', ..., X_uX_u', (X_{u+1}X_{u+1}')^{(s)}, ..., (X_mX_m')^{(s)}\|J'J'')$

$\leq H(X_1...X_u|J') - \tau^{\Lambda}(X_1, X_2, ..., X_u, X_{u+1}^{(s)}, ..., X_m^{(s)}\|J') +$

$H(X_1'...X_u'|J'') - \tau^{\Lambda}(X_1', X_2', ..., X_u', X_{u+1}^{'(s)}, ..., X_m^{'(s)}\|J'')$;

and

$I(X_1X_2...X_mX_1'X_2'...X_m'; J'J''|ZZ') \leq$

$I(X_1X_2...X_m; J'|Z) + I(X_1'X_2'...X_m'; J''|Z')$.

Starting from the last inequality:

$I(X_1X_2...X_mX_1'X_2'...X_m'; J'J''|ZZ') =$

$$H(J'J''|ZZ') - H(J'J''|ZZ'X_1X_2...X_mX_1'X_2'...X_m') \leq$$
$$H(J'|ZZ') + H(J''|ZZ') -$$
$$H(J'|ZZ'X_1X_2...X_mX_1'X_2'...X_m') - H(J''|J'ZZ'X_1X_2...X_mX_1'X_2'...X_m') \leq^i$$
$$H(J'|Z) + H(J''|Z') - H(J'|ZX_1X_2...X_m) - H(J''|Z'X_1'X_2'...X_m') =$$
$$I(X_1X_2...X_m; J'|Z) + I(X_1'X_2'...X_m'; J''|Z')$$

In step i, we have used the Markov property

$$J' - X_1X_2...X_mZ - X_1 - X_1' - X_1'X_2'...X_m'Z' - J''.$$

It remains to prove the other inequality. We first prove that for every set $B \subseteq [m]$:

$$H(X_{B\bigcap[u]}X'_{B\bigcap[u]}|X_{B^c}X'_{B^c}J'J'') - H(X_1|X_{B^c}X'_{B^c}J'J'') =$$
$$H(X_{B\bigcap[u]}|X_{B^c}J') - H(X_1|X_{B^c}J') + H(X'_{B\bigcap[u]}|X'_{B^c}J'') - H(X_1'|X'_{B^c}J'')$$

This equality is true because

$$H(X_{B\bigcap[u]}X'_{B\bigcap[u]}|X_{B^c}X'_{B^c}J'J'') =$$
$$H(X_{B\bigcap[u]}X'_{B\bigcap[u]}X_1|X_{B^c}X'_{B^c}J'J'') =$$
$$H(X_1|X_{B^c}X'_{B^c}J'J'') + H(X_{B\bigcap[u]}X'_{B\bigcap[u]}|X_1X_{B^c}X'_{B^c}J'J'') =^i$$
$$H(X_1|X_{B^c}X'_{B^c}J'J'') + H(X_{B\bigcap[u]}|X_1X_{B^c}X'_{B^c}J'J'') +$$
$$H(X'_{B\bigcap[u]}|X_1X_1'X_{B\bigcap[u]}X_{B^c}X'_{B^c}J'J'') =^{ii}$$
$$H(X_1|X_{B^c}X'_{B^c}J'J'') + H(X_{B\bigcap[u]}|X_1X_{B^c}J') + H(X'_{B\bigcap[u]}|X_1'X'_{B^c}J'') =$$
$$H(X_1|X_{B^c}X'_{B^c}J'J'') + H(X_{B\bigcap[u]}|X_{B^c}J') - H(X_1|X_{B^c}J') + H(X'_{B\bigcap[u]}|X'_{B^c}J'') - H(X_1'|X'_{B^c}J'').$$

In step i, we have used the fact that $H(X_1'|X_1) = 0$ and in step ii, we have used the Markov property

$$J' - X_1X_2...X_mZ - X_1 - X_1' - X_1'X_2'...X_m'Z' - J''.$$

This property lets us to rewrite the inequality we would like to prove in a new form:

$$H(X_1|J', J'') -$$
$$\sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m]} \lambda_B H(X_1|X_{B^c}X'_{B^c}J', J'') \leq$$
$$H(X_1|J') -$$
$$\sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m]} \lambda_B H(X_1|X_{B^c}J') +$$
$$H(X_1'|J'') -$$
$$\sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m]} \lambda_B H(X_1'|X'_{B^c}J'')$$

Further, we can restrict the summation on those sets B such that $1 \in B$ (otherwise the term in question would be zero).

From the definition of Λ, we can write:

$\sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m],1\in B} \lambda_B = 1$ (this could be proved by setting $R_1 = 1$, and $R_i = 0$ for $1 < i \leq u$).

Therefore

$H(X_1|J',J'') - \sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m]} \lambda_B H(X_1|X_{B^c}X'_{B^c}J'J'') =$
$\sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m]} \lambda_B [H(X_1|J',J'') - H(X_1|X_{B^c}X'_{B^c}J'J'')] =$
$\sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m]} \lambda_B I(X_1; X_{B^c}X'_{B^c}|J'J'').$

Similarly we can rewrite the two other expressions. It would be then enough to prove that

$I(X_1; X_{B^c}X'_{B^c}|J'J'') \leq I(X_1; X_{B^c}|J') + I(X_1; X'_{B^c}|J'')$

for all $B \subseteq [m]$ such that $B \neq [m]$ and $1 \in B$.

We have:

$I(X_1; X_{B^c}X'_{B^c}|J'J'') = H(X_{B^c}X'_{B^c}|J'J'') - H(X_{B^c}X'_{B^c}|J'J''X_1) \leq$
$H(X_{B^c}|J') + H(X'_{B^c}|J'') - H(X_{B^c}X'_{B^c}|J'J''X_1) =^i$
$H(X_{B^c}|J') + H(X'_{B^c}|J'') - H(X_{B^c}|J'X_1) - H(X'_{B^c}|J''X'_1) =$
$I(X_1; X_{B^c}|J') + I(X_1; X'_{B^c}|J'').$

In step i, we have used $H(X'_1|X_1) = 0$ and the Markov property

$J' - X_1X_2...X_mZ - X_1 - X'_1 - X'_1X'_2...X'_mZ' - J''.$ •

Property 2.

Let $1 \leq i \leq u$ and let $H(F|X_i) = 0$. We need to prove that:

$\inf_{\widetilde{J}}(\theta^\Lambda(X_1; X_2; X_3; ...; X_m; \widetilde{J}\|Z)) \geq$
$\inf_{\widetilde{J'}}(\theta^\Lambda(X_1F; X_2F; X_3F; ...; X_mF; \widetilde{J'}\|ZF))$

It is enough to prove that for any J, there is a J' such that:

$\theta^\Lambda(X_1; X_2; X_3; ...; X_m; J\|Z) \geq \theta^\Lambda(X_1F; X_2F; X_3F; ...; X_mF; J'\|ZF)$

Let $J' = JF$. Since $I(F; J \mid Z) \geq 0$, one can show that the above inequality would hold if:

$H(F|J) - \sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m]} \lambda_B H(F|X_{B^c}J) \geq 0.$

Since $H(F|X_i) = 0$, we can rewrite the above inequality as follows:

$H(F|J) - \sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m],i\in B} \lambda_B H(F|X_{B^c}J) \geq 0.$

$H(F|X_{B^c}J)$ is bounded from above by $H(F|J)$ hence

$H(F|J) - \sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m],i\in B} \lambda_B H(F|X_{B^c}J) \geq$
$H(F|J).(1 - \sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m],i\in B} \lambda_B).$

But

$1 - \sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m],i\in B} \lambda_B = 0$

This could be proved by setting $R_i = 1$, and $R_j = 0$ otherwise in the inequality involving R_j's. •

Property 3.

We need to prove that:

$\inf_{\widetilde{J}}(\theta(X_1;X_2;X_3;...;X_m;\widetilde{J}\|Z)) \geq$

$\inf_{\widetilde{J'}}(\theta(X'_1;X'_2;X'_3;...;X'_m;\widetilde{J'}\|Z))$

It is enough to prove that for any J:

$\theta(X_1;X_2;X_3;...;X_m;J\|Z) \geq \theta(X'_1;X'_2;X'_3;...;X'_m;J\|Z)$

It is clear that

$I(X_1X_2...X_m;J|Z) \geq I(X'_1X'_2...X'_m;J|Z)$.

It remains to show that the first two terms of the expression, that is

$H(X_1...X_u|J) - \tau^{\Lambda}(X_1,X_2,...,X_u,X^{(s)}_{u+1},...,X^{(s)}_m\|J)$,

does not increase when we replace $(X_1,X_2,...,X_m,Z,J)$ with $(X'_1,X'_2,...,X'_m,Z,J)$.

Since we can replace $(X_1,X_2,...,X_m)$s with $(X'_1,X'_2,...,X'_m)$ one at a time, it is enough to consider the case that we only change one component, that is we replace $(X_1,X_2,...,X_m)$ by $(X_1,X_2,...,X_{j-1},X'_j,X_{j+1},...,X'_m)$.

The proof can be completed by considering the two cases of $j>u$ and $j\leq u$ separately. In the case $j>u$, we note that $\tau^{\Lambda}(X_1,X_2,...,X_u,X^{(s)}_{u+1},...,X^{(s)}_m\|J)$ increases term by term while $H(X_1X_2...X_u|J)$ remains constant. In case $j\leq u$, we note that for every set B that does not contain j, the term $-\lambda_B H(X_{B\bigcap[u]}|X_{B^c}J)$ decreases as we replace X_j by X'_j. If the set B includes j, we have:

$H(X_{B\bigcap[u]}|X_{B^c}J) = H(X_{(B\bigcap[u])-\{j\}}X_j|X_{B^c}J) =$

$H(X_{(B\bigcap[u])-\{j\}}X_jX'_j|X_{B^c}J) =$

$H(X_{(B\bigcap[u])-\{j\}}X'_j|X_{B^c}J) + H(X_j|X'_jX_{B^c}X_{(B\bigcap[u])-\{j\}}J) \leq$

$H(X_{(B\bigcap[u])-\{j\}}X'_j|X_{B^c}J) + H(X_j|X'_jX_{[u]-\{j\}}J)$

So, in order to prove the inequality, it would be enough to prove that

$H(X_j|X'_jX_{[u]-\{j\}}J) - \sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m],j\in B}\lambda_B H(X_j|X'_jX_{[u]-\{j\}}J) \geq 0$.

But the left hand side is zero since $\sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m],j\in B}\lambda_B = 1$. ●

Property 4.

$\varphi_j(X_1;X_2;X_3;...;X_m\|Z)$ can be lower bounded as follows (in the following formula a Λ is called valid if it verifies the conditions in the statement of Theorem 2):

$$\varphi_j(X_1;X_2;X_3;...;X_m\|Z) \geq \inf_{valid\ \Lambda}\{\inf_J\big(H(X_1...X_u|J) - \tau^{\Lambda}(X_1,X_2,...,X_u,X^{(s)}_{u+1},...,X^{(s)}_m\|J) + I(X_1X_2...X_m;J\|Z)\big)\} =$$

$$\inf_J\{\inf_{valid\ \Lambda}\big(H(X_1...X_u|J) - \tau^{\Lambda}(X_1,X_2,...,X_u,X^{(s)}_{u+1},...,X^{(s)}_m\|J) + I(X_1X_2...X_m;J\|Z)\big)\}$$

By applying Theorem 6 of the first part of this paper [6] and the duality theory, one gets the following lower bound on $\varphi_j(X_1;X_2;X_3;...;X_m\|Z)$:

$\varphi_1(X_1;X_2;X_3;...;X_m\|Z) \geq$

$\inf_J \big(S(X_1J;X_2J;...;X_uJ;(X_{u+1}J)^{(s)}...;(X_mJ)^{(s)}\|J) + I(X_1X_2...X_m;J^{(s)}\|Z)\big)$.

According to Theorem 5 of the [6],

$\inf_J \big(S(X_1J;X_2J;...;X_uJ;(X_{u+1}J)^{(s)}...;(X_mJ)^{(s)}\|J) + I(X_1X_2...X_m;J^{(s)}\|Z)\big)$

is an upper bound on

$S(X_1;X_2;...;X_u;X_{u+1}^{(s)}...;X_m^{(s)}\|Z)$

which is in turn bounded from below by

$H(X_1|Z) - \sum_{i=2}^m H(X_1|X_i)$.

Therefore $\varphi(X_1;X_2;X_3;...;X_m\|Z) \geq H(X_1|Z) - \sum_{i=2}^m H(X_1|X_i)$. ●

Property 5.

We need to prove that:

$\inf_{\widetilde{J}}(\theta^\Lambda(X_1;X_2;X_3;...;X_m;\widetilde{J}\|Z)) \geq$

$\inf_{\widetilde{J}'}(\theta^\Lambda(X_1M_1;X_2M_2;...;X_uM_u;X_{u+1}...;X_m;\widetilde{J}'\|Z))$

It is enough to prove that for any J, there is a J' such that:

$\theta^\Lambda(X_1;X_2;X_3;...;X_m;J\|Z) \geq \theta^\Lambda(X_1M_1;X_2M_2;...;X_uM_u;X_{u+1}...;X_m;J'\|Z)$

We define J' in a way that it has the same joint distribution with $(X_1,X_2,...,X_m,Z)$ as J has but at the same be independent of $M_1M_2...M_u$. One can then prove that:

$H(X_1M_1...X_uM_u|J') - \tau^\Lambda(X_1M_1,X_2M_2,...,X_uM_u,X_{u+1}^{(s)},...,X_m^{(s)}\|J') +$

$I(X_1X_2...X_mM_1...M_u;J'|Z) =$

$H(X_1...X_u|J) - \tau^\Lambda(X_1,X_2,...,X_u,X_{u+1}^{(s)},...,X_m^{(s)}\|J) + I(X_1X_2...X_m;J|Z) +$

$H(M_1) + ... + H(M_u) -$

$\sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m]} \lambda_B \sum_{i\in B\bigcap[u]} H(M_i)$

But

$H(M_1) + H(M_2) + ... + H(M_u) -$

$\sum_{B:B\subset[m],B\bigcap[u]\neq\emptyset,B\neq[m]} \lambda_B \sum_{i\in B\bigcap[u]} H(M_i)$

is zero (this could be proved using the definition of Λ and by setting $R_i = H(M_i)$ for $1 \leq i \leq u$). ●

APPENDIX II

In this appendix, we will prove that for any ϵ in the interval $(0,1)$, $I(X;Y|Z)$ strictly increases in the following three cases:

- X_1 and X_2 are not independent and we replace $p(X_1,X_2)p(Y,Z|X)$ with $p(X_1)p(X_2)p(Y,Z|X)$.
- we change the distribution of X_1 to a uniform distribution if X_1 and X_2 are independent but X_1 is not uniform.
- we change the distribution of X_2 to a uniform distribution if X_1 and X_2 are independent but X_2 is not uniform.

Case 1:

$I(X;Y|Z) = I(X_1X_2;Y_1Y_2|Z_1Z_2) = H(Y_1Y_2|Z_1Z_2) - H(Y_1Y_2|Z_1Z_2X_1X_2)$.

Since $Y_1Z_1 - X_1 - X_2 - Y_2Z_2$, we can work out the second term

$H(Y_1Y_2|Z_1Z_2X_1X_2) = H(Y_1|Z_1Z_2X_1X_2) + H(Y_2|Z_1Z_2X_1X_2Y_1) = H(Y_1|Z_1X_1) + H(Y_2|X_2Z_2)$.

The first term can be bounded from above as follows:

$H(Y_1Y_2|Z_1Z_2) = H(Y_2|Z_1Z_2) + H(Y_1|Z_1Z_2Y_2) \leq H(Y_2|Z_2) + H(Y_1|Z_1)$.

Therefore $I(X;Y|Z) \leq I(X_1;Y_1|Z_1) + I(X_2;Y_2|Z_2)$. This would mean that if we replace $p(X_1,X_2)p(Y,Z|X)$ with $p(X_1)p(X_2)p(Y,Z|X)$, $I(X;Y|Z)$ does not decrease.

We prove that $I(X;Y|Z)$ strictly increases by contradiction. Assume $I(X;Y|Z)$ does not increase. In this case, $H(Y_1|Z_1Z_2Y_2)$ must be equal to $H(Y_1|Z_1)$ implying that $I(Y_1;Y_2|Z_1) = 0$. Since $Z_1 - Y_1 - Y_2$ form a Markov chain, the $I(Y_1;Y_2|Z_1) = 0$ constraint implies that $I(Y_2;Z_1) = I(Y_2;Y_1)$. But since

$I(Y_2;Y_1) \geq I(Y_2;T_1) \geq I(Y_2;Z_1)$,

we get $I(Y_2;T_1) = I(Y_2;Z_1)$.

$I(Y_2;Z_1) = I(Y_2;Z_1,\mathbb{1}[Z_1 = E]) =$

$I(Y_2;\mathbb{1}[Z_1 = E]) + I(Y_2;Z_1|\mathbb{1}[Z_1 = E]) = 0 + \epsilon.I(Y_2;T_1)$.

Since $\epsilon < 1$, $I(Y_2;T_1) = I(Y_2;Z_1)$ can hold only when $I(Y_2;T_1) = I(Y_2;Z_1) = I(Y_2;Y_1) = 0$.

$0 = I(Y_2;Y_1) = I(Y_2,\mathbb{1}[Y_2 = E];Y_1,\mathbb{1}[Y_1 = E]) \geq$

$I(Y_2;Y_1|\mathbb{1}[Y_2 = E],\mathbb{1}[Y_1 = E]) \geq$

$p(Y_2 \neq E).p(Y_1 \neq E).I(Y_2;Y_1|Y_2 \neq E, Y_1 \neq E) = 0.81I(X_1;X_2)$.

Therefore $I(X_1;X_2) = 0$ meaning that X_1 and X_2 are independent. This is a contradiction. ●

Case 2:

$I(X_1;Y_1|Z_1) = I(X_1;Y_1) - I(X_1;Z_1) = H(Y_1) - H(Y_1|X_1) - H(Z_1) + H(Z_1|X_1)$ can be thought of as a function of $p(X_1 = 0) = a$. $H(Y_1|X_1)$ and $H(Z_1|X_1)$ are constant not depending on a. The

marginal distribution of Z_1 equals $(\epsilon.(0.9a+0.05), 1-\epsilon, \epsilon.(-0.9a+0.95))$, and the marginal distribution of Y_1 equals $(0.9a, 0.1, 0.9-0.9a))$. Therefore it is enough to show that $H(Y_1) - H(Z_1)$ reaches its maximum at and only at $a = 0.5$. This can be seen by noting that the derivative of $\frac{1}{0.9}(H(Y_1) - H(Z_1))$ with respect to a equals: $\log \frac{0.5-(a-0.5)}{0.5+(a-0.5)} - \epsilon \log \frac{0.5-0.9(a-0.5)}{0.5+0.9(a-0.5)}$ which is zero only at $a = 0.5$. ●

Case 3:

$I(X_2; Y_2|Z_2) = I(X_2; (Y_2, \mathbb{1}[Y_2 = E])|Z_2) = I(X_2; \mathbb{1}[Y_2 = E]|Z_2) + I(X_2; Y_2|\mathbb{1}[Y_2 = E], Z_2) =$
$0 + P(Y_2 = E).0 + P(Y_2 \neq E).H(X_2|Z_2) = 0.9H(X_2|Z_2)$.

But $H(X_2|Z_2) = P(Z_2 = 0).0 + P(Z_2 = 1).0 + P(Z_2 = E).H(X_2)$. Therefore

$I(X_2; Y_2|Z_2) = 0.9 * 0.19H(X_2)$.

We are done by noting that $H(X_2)$ strictly increases when the distribution of X_2 is changed to uniform.

ACKNOWLEDGMENT

The authors would like to thank TRUST (The Team for Research in Ubiquitous Secure Technology), which receives support from the National Science Foundation (NSF award number CCF-0424422) and the following organizations: Cisco, ESCHER, HP, IBM, Intel, Microsoft, ORNL, Pirelli, Qualcomm, Sun, Symantec, Telecom Italia and United Technologies, for their support of this work. The research was also partially supported by NSF grant numbers CCF-0500023, CCF-0635372 and CNS-0627161.

REFERENCES

[1] R. Ahlswede and I. Csiszár, "Common Randomness in Information Theory and Cryptography. Part I: Secret sharing", *IEEE Trans. Inform. Theory*, Vol. 39, No. 4, July 1993, pp. 1121 -1132.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.

[3] I. Csiszár and J. Korner, "Broadcast Channels with Confidential Messages", *IEEE Trans. Inform. Theory*, Vol. 24, No. 3, May 1978, pp. 339-348.

[4] I. Csiszár and P. Narayan, "Secrecy Capacities for Multiple Terminals", *IEEE Trans. Inform. Theory*, Vol. 50, No. 12, Dec 2004, pp. 3047-3061.

[5] A. A. Gohari and V. Anantharam, "Communication for Omniscience by a Neutral Observer and Information-Theoretic Key Agreement of Multiple Terminals", *Proceedings of International Symposium on Information Theory (ISIT)*, 2007, pp. 2056-2060.

[6] A. A. Gohari and V. Anantharam, "Information-Theoretic Key Agreement of Multiple Terminals - Part I: Source Model", *Preprint submitted for publication*, 2008.

[7] U. M. Maurer, "Secret Key Agreement by Public Discussion From Common Information", *IEEE Trans. Inform. Theory*, Vol. 39, No.3, May 1993,pp. 733-742.

[8] U. M. Maurer and S. Wolf, "From Weak to Strong Information-Theoretic Key Agreement", *Proceedings of International Symposium on Information Theory (ISIT)*, 2000, p.18.

[9] R. Renner and S. Wolf, "New Bounds in Secret-Key Agreement: The Gap Between Formation and Secrecy Extraction", *Proceedings of EUROCRYPT 2003*, LNCS, Springer-Verlag, Vol. 2656, May 2003, pp.562577.

[10] A. D. Wyner, “The Wiretap Channel”, *Bell System Technical Journal*, Vol. 54, No. 8, Oct. 1975, pp. 1355-1387.