

Communication For Omniscience by a Neutral Observer and Information-Theoretic Key Agreement of Multiple Terminals

Amin Aminzadeh Gohari
EECS Department
University of California Berkeley
Berkeley, CA 94720
aminzade@eecs.berkeley.edu

Venkat Anantharam
EECS Department
University of California Berkeley
Berkeley, CA 94720
ananth@eecs.berkeley.edu

Abstract—We derive a new upper bound on the secrecy capacity in the source model with eavesdropper which strictly improves the currently best upper bound, i.e. the double intrinsic information bound of Renner and Wolf [2]. Furthermore, unlike that bound, which is defined only in the case of two terminals, the new upper bound is not specific to the two terminals case. We define a problem of communication for omniscience by a neutral observer and establish the equivalence between this new problem and the problem of secret key agreement.

I. INTRODUCTION

Complete secrecy is the most desirable form of security measure as it does not make any assumptions on the computational power of the adversary. Shannon was the first who precisely formulated the problem of information theoretic secret key generation [9]. Since then, the work of Shannon has been much developed and modified (for example see [7], [5], [4]). In this paper we deal with an important model, known as the “source model”, introduced by the works of Ahlswede and Csiszár [4] and Csiszár and Narayan [1]. In this model, there are m parties interested in secret key generation against an adversary Eve. The m parties and Eve have access to i.i.d. repetition of jointly correlated random variables X_i ($i = 1, \dots, m$), and Z . We assume that all the m parties desire to agree on a random variable which is hidden from Eve. The parties are permitted to have an interactive authenticated public communication after observing, say n i.i.d repetition of their random variables. Following the communication, the parties generate random variables S_i 's as the secret key ($i = 1, 2, \dots, m$). 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 (that is the n i.i.d repetitions of Z and the public discussion). The achieved secret key rate would then be $\frac{1}{n}H(S_1)$. The highest achievable secret key rate is called the secrecy capacity. Calculation of the exact secrecy capacity remains an unsolved problem, although some lower and upper bounds have been proved. For the case of $m = 2$, the best know upper bound is that of Renner and Wolf [2]. This bound is known as the “double intrinsic information” bound and is equal to

$\inf_U [H(U) + I(X_1; X_2 \downarrow ZU)]$ where $I(X; Y \downarrow Z)$ is defined as $\inf_{X'Y-Z-\bar{Z}} I(X; Y|\bar{Z})$ and is called the “intrinsic information” [3]. The best known lower bound, proved using random binning arguments, is due to Ahlswede and Csiszár [4]: the maximum of $\sup_{V-U-X-YZ} (I(U; Y|V) - I(U; Z|V))$ and $\sup_{V-U-Y-XZ} (I(U; X|V) - I(U; Z|V))$. Under some special circumstances, Csiszár and Narayan [1] derived a single-letter characterization of the secrecy capacity, notably when Z is independent of (X_1, X_2, \dots, X_m) . In [1] the connection between a problem of communication for omniscience by the terminals and the secret key generation problem is introduced. In the former problem the required condition is that at the end of the communication all the terminals become omniscient about each other's random variables, and the goal is to minimize the communication rate required to achieve this.

In this paper, we improve the above mentioned results relating to multi-terminal secret key rate calculations [1] and the state of art upper bound on secret key rate of [2].

The outline of this paper is as follows. In section II, we introduce or review the definitions and the notations which are frequently used in this paper. Among the definitions discussed are those relating to the Communication for Secret Key generation in the presence of an eavesdropper (SK), and the Communication For Omniscience by a neutral observer (CFO). Section III contains the main results of this paper followed by section IV which gives proof sketches for the results. Appendix I contains an example showing that our upper bound for secret key rate is strictly better than the currently best know upper bound from [2].

II. DEFINITIONS AND NOTATIONS

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

We use $C.H.\{\Phi(p(x_1, x_2, \dots, x_m, z))\}$ to refer to the concave hull of the function Φ defined on the set of probability distributions on $(X_1, X_2, \dots, X_m, Z)$.

We basically use the same multi-terminal model defined by Csiszár and Narayan in [1]. We however relax the uniformity

condition on the generated secret key i.e. equation (2) in [1] (Maurer in [5] 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 parties are interested in secret key generation.

We proceed with the following definitions:

Definition 1: The pair (n, \vec{C}) , where n is a natural number and $\vec{C} = (C_1, C_2, \dots, C_r)$ is a finite set of discrete random variables is considered a “valid communication” if:

- $H(C_i | C_{[1:i-1]}, X_j^{1:n}) = 0 \forall j : 1 \leq j \leq m, i - j \equiv m \pmod{0}$

Please note that if (n, \vec{C}) is valid, then one has $H(\vec{C} | X_1^n, X_2^n, \dots, X_m^n) = 0$.

Definition 2. Let n be a natural number, ϵ be a positive real number, $\vec{C} = (C_1, C_2, \dots, C_r)$ be a finite set of discrete random variables and S_1, \dots, S_m be m discrete random variables. Consider the following conditions.

- 1) the pair (n, \vec{C}) is a valid communication.
- 2) $H(S_i | C_{[1:r]}, X_i^{1:n}) = 0$ for all $1 \leq i \leq m$
- 3) $P(S_1 = S_2 = S_3 = \dots = S_m) > 1 - \epsilon$
- 4) $\frac{1}{n} I(S_1; Z^{1:n}, C_{[1:r]}) < \epsilon$
- 5) $\frac{1}{n} H(X_1^{1:n}, X_2^{1:n}, \dots, X_m^{1:n} | Z^{1:n}, S_1, S_2, \dots, S_m) < \epsilon$

The data typing condition $SK(n, \epsilon, S_1, S_2, S_3, \dots, S_m, \vec{C})$ is said to hold iff conditions 1, 2, 3 and 4 are satisfied. To any SK data type, we assign a number called the “gain” of the SK data type which is defined as $\frac{1}{n} H(S_1)$.

The data typing condition $CFO(n, \epsilon, S_1, S_2, S_3, \dots, S_m, \vec{C})$ is said to hold iff conditions 1, 2, 3 and 5 are satisfied. To any CFO data type, we assign a number called the “cost” of the CFO data type which is defined as $\frac{1}{n} H(\vec{C} | Z^n)$.

A valid communication (n, \vec{C}) for which, for some $\epsilon > 0$ and some (S_1, S_2, \dots, S_m) the data typing condition $SK(n, \epsilon, S_1, S_2, S_3, \dots, S_m, \vec{C})$ holds is called a Communication for Secret Key generation in the presence of an eavesdropper. The intuitive reason for this terminology should be clear from the definition.

A valid communication (n, \vec{C}) for which, for some $\epsilon > 0$ and some (T_1, T_2, \dots, T_m) the data typing condition $CFO(n, \epsilon, T_1, T_2, T_3, \dots, T_m, \vec{C})$ holds is called a Communication For Omniscience by a neutral observer (CFO). Intuitively speaking, a Communication For Omniscience protocol works as follows. The terminals will conduct a public discussion in order to agree (with high probability) on a common randomness, but there is no secrecy constraint. We can assume that there is a neutral party, say Charles, who receives Z^n from Eve and the common randomness T_i 's obtained by the terminals. Charles is required to become omniscient about $X_1^n, X_2^n, \dots, X_m^n$. The “cost” of the communication would be the entropy of the overall communication conditioned on Z^n .

If Z is independent of (X_1, \dots, X_m) , then Charles will not learn anything about $X_1^n, X_2^n, \dots, X_m^n$ from Z^n and thus each T_i should be approximately equal to $X_1^n, X_2^n, \dots, X_m^n$ meaning that each terminal has learned the random variables of all other terminals. The Communication For Omniscience by a neutral observer would be transformed to a simple Communication

For Omniscience studied by Csiszár and Narayan [1]. The cost of communication in this case, as expected from the work of Csiszár and Narayan, is equal to the sum of the rates of the communications. Therefore the Communication For Omniscience by a neutral observer is a generalization of the Communication For Omniscience of [1].

Definition 3: $S_{no-r}^\epsilon(X_1; X_2; X_3; \dots; X_m \| Z)$, the ϵ secret key rate when the terminals cannot randomize, is defined as: $\limsup_{n \rightarrow \infty} \sup_{SK(n, \epsilon, S_1, S_2, S_3, \dots, S_m, \vec{C})} Gain(SK)$. Similarly, $T^\epsilon(X_1; X_2; X_3; \dots; X_m \| Z)$ is defined: $\liminf_{n \rightarrow \infty} \inf_{CFO(n, \epsilon, T_1, T_2, T_3, \dots, T_m, \vec{C})} Cost(CFO)$.

$S_{no-r}(X_1; X_2; X_3; \dots; X_m \| Z)$, the secret key rate when the terminals cannot randomize, is defined as the limit of $S_{no-r}^\epsilon(X_1; X_2; X_3; \dots; X_m \| Z)$ as ϵ goes to zero. $T(X_1; X_2; X_3; \dots; X_m \| Z)$ is defined similarly.

$S(X_1; X_2; X_3; \dots; X_m \| Z)$, the secret key rate when the terminals can randomize, is defined as supremum over all (M_1, M_2, \dots, M_m) satisfying: $p(M_1, \dots, M_m, X_1, \dots, X_m, Z) = p(M_1) \cdot p(M_2) \cdot \dots \cdot p(M_m) \cdot p(X_1, \dots, X_m, Z)$ of $S_{no-r}(X_1 M_1; X_2 M_2; X_3 M_3; \dots; X_m M_m \| Z)$.

III. STATEMENT OF THE RESULTS

Theorem 1. Let $\varphi(X_1; X_2; X_3; \dots; X_m \| Z)$ be a function from the set of all probability distributions defined on $(X_1, X_2, X_3, \dots, X_m, Z)$ (where X_i 's and Z are taking values from discrete sets), to real numbers. $\varphi(X_1; X_2; X_3; \dots; X_m \| Z)$ is an upper bound on $S_{no-r}(X_1; X_2; X_3; \dots; X_m \| Z)$ if it satisfies the following properties:

- 1) For any natural number n :
 $n\varphi(X_1; X_2; \dots; X_m \| Z) \geq \varphi(X_1^n; X_2^n; \dots; X_m^n \| Z^n)$
- 2) For any random variable F such that $\exists i : H(F | X_i) = 0$, we have:
 $\varphi(X_1; X_2; \dots; X_m \| Z) \geq \varphi(X_1 F; X_2 F; \dots; X_m F \| Z F)$
- 3) For any random variables X'_1, X'_2, \dots, X'_m such that $\forall i : H(X'_i | X_i) = 0$, we have:
 $\varphi(X_1; X_2; \dots; X_m \| Z) \geq \varphi(X'_1; X'_2; \dots; X'_m \| Z)$.
- 4) $\varphi(X_1; X_2; \dots; X_m \| Z) \geq H(X_1 | Z) - \sum_{i=2}^m H(X_1 | X_i)$

Further $S_{no-r}(X_1; X_2; X_3; \dots; X_m \| Z)$ satisfies these properties.

Theorem 2. Let $\psi(X_1; X_2; X_3; \dots; X_m \| Z)$ be a function from the set of all probability distributions defined on $(X_1, X_2, X_3, \dots, X_m, Z)$ (where X_i 's and Z are taking values from discrete sets), to real numbers. $\psi(X_1; X_2; X_3; \dots; X_m \| Z)$ is a lower bound on $T(X_1; X_2; X_3; \dots; X_m \| Z)$ if it satisfies the following properties:

- 1) For any natural number n :
 $n\psi(X_1; X_2; \dots; X_m \| Z) \leq \psi(X_1^n; X_2^n; \dots; X_m^n \| Z^n)$
- 2) For any random variable F such that $\exists i : H(F | X_i) = 0$, we have:

$$\psi(X_1; X_2; \dots; X_m \| Z) \leq \psi(X_1 F; X_2 F; \dots; X_m F \| Z F) + H(F | Z)$$

3) For any random variables X'_1, X'_2, \dots, X'_m such that $\forall i: H(X'_i|X_i) = 0$, we have:

$$\psi(X_1; X_2; \dots; X_m \| Z) \leq \psi(X'_1; X'_2; \dots; X'_m \| Z) + H(X_1 \dots X_m | X'_1 \dots X'_m Z).$$

4) $\psi(X_1; X_2; \dots; X_m \| Z) \leq H(X_2 \dots X_m | X_1 Z) + \sum_{i=2}^m H(X_1 | X_i)$

Further $T(X_1; X_2; X_3; \dots; X_m \| Z)$ satisfies these properties.

Theorem 3. For any joint distribution $p(x_1, x_2, \dots, x_m, z)$, we have: $S_{no-r}(X_1; X_2; X_3; \dots; X_m \| Z) + T(X_1; X_2; X_3; \dots; X_m \| Z) = H(X_1, X_2, \dots, X_m | Z)$.

Theorem 4. If $\psi(X_1; X_2; X_3; \dots; X_m \| Z)$ is a lower bound on $T(X_1; X_2; X_3; \dots; X_m \| Z)$ satisfying the four properties of Thm. 2, so would be $\psi'(X_1; X_2; X_3; \dots; X_m \| Z)$ which is defined in the following way:

$$\psi'(X_1; X_2; X_3; \dots; X_m \| Z) := \inf_r \frac{1}{r} C.H.\{\psi\}(X_1^r; X_2^r; \dots; X_m^r \| Z^r)$$

Theorem 5. For the special case of $n = 2$, we get the upper bound $\inf_{J,J'} (I(X; Y|J) + I(XY; J|Z))$ on $S_{no-r}(X; Y \| Z)$ where the infimum is taken over all random variables J . This expression is also an upper bound on $S(X; Y \| Z)$ and further the infimum could be replaced by minimum. This bound strictly improves the Renner-Wolf double intrinsic information upper bound.

IV. PROOFS OF THEOREMS 1-5

Proof of Theorem 1. Fix a probability distribution $p(x_1, x_2, \dots, x_m, z)$ on $(X_1, X_2, \dots, X_m, Z)$ and assume that X_1, X_2, \dots, X_m, Z take values in the discrete finite sets $\Delta_i, i = 1 \dots m + 1$. For every $\delta > 0$ and $\epsilon > 0$, one can find data type $SK(n, \epsilon, S_1, S_2, S_3, \dots, S_m, \vec{C})$ whose Gain is within δ of $S_{no-r}^\epsilon(X_1; X_2; \dots; X_m \| Z)$. We have:

$$\begin{aligned} n\varphi(X_1; X_2; X_3; \dots; X_m \| Z) &\geq^i \\ \varphi(X_1^n; X_2^n; X_3^n; \dots; X_m^n \| Z^n) &\geq^{ii} \\ \varphi(X_1^n C_1; X_2^n C_1; \dots; X_m^n C_1 \| Z^n C_1) &\geq^{iii} \\ \varphi(X_1^n C_1 C_2; X_2^n C_1 C_2; \dots; X_m^n C_1 C_2 \| Z^n C_1 C_2) \dots &\geq^{iv} \\ \varphi(X_1^n \vec{C}; X_2^n \vec{C}; \dots; X_m^n \vec{C} \| Z^n \vec{C}) &\geq^v \\ \varphi(S_1; S_2; \dots; S_m \| Z^n \vec{C}) &\geq^{vi} H(S_1 | Z^n \vec{C}) - \\ \sum_{j=2}^m H(S_1 | S_j) &\geq^{vii} nS_{no-r}^\epsilon(X_1; X_2; \dots; X_m \| Z) \} - \\ n\delta - (m-1)[h(\epsilon) + \epsilon.n \log \prod_{i=1}^m |\Delta_i|] & \end{aligned}$$

Inequalities i, ii, iii, iv, v, vi are true respectively because of the properties 1,2,2,2,3,4. Inequality vii is true because of the Fano inequality, and the fact that the gain of $SK(n, \epsilon, S_1, S_2, S_3, \dots, S_m, \vec{C})$ is within δ of $S_{no-r}^\epsilon(X_1; X_2; \dots; X_m \| Z)$.

Therefore we get $\varphi(X_1; X_2; X_3; \dots; X_m \| Z) \geq S_{no-r}^\epsilon(X_1; X_2; \dots; X_m \| Z) - \delta - \frac{m-1}{n}[h(\epsilon) + \epsilon.n \log \prod_{i=1}^m |\Delta_i|]$. The theorem is proved by taking the limit as ϵ and δ go to zero.

$S_{no-r}(X_1; X_2; \dots; X_m \| Z)$ itself satisfies the four properties (the details are suppressed). It satisfies the last property since $S_{no-r}(X_1; X_2; \dots; X_m \| Z)$ is greater than or equal to the one way secret key rate from X_1 to X_2, \dots, X_m in the presence of Z which in turn is greater than or equal to the stated lower bound (the details are suppressed).

Proof of Theorem 2. Fix the probability distribution $p(x_1, x_2, \dots, x_m, z)$ on $(X_1, X_2, \dots, X_m, Z)$ and assume that $(X_1, X_2, \dots, X_m, Z)$ take values in the discrete finite sets $\Delta_i,$

$i = 1 \dots m + 1$. For every $\delta > 0$ and $\epsilon > 0$, one can find data type CFO($n, \epsilon, S_1, S_2, S_3, \dots, S_m, \vec{C}$) whose Cost is within δ of $T^\epsilon(X_1; X_2; \dots; X_m \| Z)$. We have:

$$\begin{aligned} n\psi(X_1; X_2; X_3; \dots; X_m \| Z) &\leq^i \\ \psi(X_1^n; X_2^n; X_3^n; \dots; X_m^n \| Z^n) &\leq^{ii} \\ \psi(X_1^n C_1; X_2^n C_1; \dots; X_m^n C_1 \| Z^n C_1) + H(C_1 | Z^n) &\leq^{iii} \\ \psi(X_1^n C_1 C_2; X_2^n C_1 C_2; \dots; X_m^n C_1 C_2 \| Z^n C_1 C_2) &+ \\ H(C_1 C_2 | Z^n) \dots &\leq^{iv} \psi(X_1^n \vec{C}; X_2^n \vec{C}; \dots; X_m^n \vec{C} \| Z^n \vec{C}) + \\ H(\vec{C} | Z^n) &\leq^v \psi(S_1; S_2; \dots; S_m \| Z^n \vec{C}) + \\ H(X_1^n X_2^n \dots X_m^n | S_1 S_2 \dots S_m Z^n) &+ H(\vec{C} | Z^n) \leq^{vi} \\ H(S_2 S_2 \dots S_m | S_1 Z^n \vec{C}) &+ \sum_{j=2}^m H(S_1 | S_j) + \\ H(X_1^n X_2^n \dots X_m^n | S_1 S_2 \dots S_m Z^n) &+ H(\vec{C} | Z^n) \leq^{vii} \\ h(\epsilon) + \epsilon.n \log \prod_{i=1}^m |\Delta_i| + (m-1)[h(\epsilon) + \epsilon.n \log \prod_{i=1}^m |\Delta_i|] &+ \\ n\epsilon + nT^\epsilon(X_1; X_2; \dots; X_m \| Z) \} &+ n\delta \end{aligned}$$

Inequalities i, ii, iii, iv, v, vi are true respectively because of the properties 1,2,2,2,3,4. Inequality vii is true due to the Fano inequality, and the fact that the cost of CFO($n, \epsilon, S_1, S_2, S_3, \dots, S_m, \vec{C}$) is within δ of $T^\epsilon(X_1; X_2; \dots; X_m \| Z)$.

Therefore we get $\psi(X_1; X_2; X_3; \dots; X_m \| Z) \leq T^\epsilon(X_1; X_2; \dots; X_m \| Z) + \delta + \frac{m}{n}[h(\epsilon) + \epsilon.n \log \prod_{i=1}^m |\Delta_i|] + \epsilon$. The theorem is proved by taking the limit as ϵ and δ go to zero.

$T(X_1; X_2; \dots; X_m \| Z)$ itself satisfies the four properties. For property number 3, intuitively, one possible communication for omniscience for $(X_1; X_2; \dots; X_m; Z)$ is to first conduct a communication for omniscience for $(X'_1; X'_2; \dots; X'_m; Z)$. The party who wants to become omniscient, Charles, would be able to approximately learn $(X'_1; X'_2; \dots; X'_m; Z)$ with the cost of $T(X'_1; X'_2; \dots; X'_m \| Z)$. If Charles exactly knew $(X'_1; X'_2; \dots; X'_m; Z)$, the m parties could have revealed their extra knowledge by revealing $H(X_1 \dots X_m | X'_1 X'_2 \dots X'_m Z)$ bits using a Slepian-Wolf type communication scheme. Even though Charles does not exactly know $(X'_1; X'_2; \dots; X'_m; Z)$, the Slepian-Wolf algorithm still works; the details are suppressed.

Proof of Theorem 3. It can be easily shown that $\psi(X_1; X_2; X_3; \dots; X_m \| Z)$ satisfies the four properties of Theorem 2 if and only if $H(X_1 X_2 \dots X_m | Z) - \psi(X_1; X_2; X_3; \dots; X_m \| Z)$ satisfies the four properties of Theorem 1. $T(X_1; X_2; \dots; X_m \| Z)$ itself satisfies the four properties of Theorem 2. Hence $H(X_1 X_2 \dots X_m | Z) - T(X_1; X_2; \dots; X_m \| Z) \geq S_{no-r}(X_1; X_2; \dots; X_m \| Z)$. Further since $S_{no-r}(X_1; X_2; \dots; X_m \| Z)$ itself satisfies the four properties of Theorem 1, we get $H(X_1 X_2 \dots X_m | Z) - S_{no-r}(X_1; X_2; \dots; X_m \| Z) \leq T(X_1; X_2; \dots; X_m \| Z)$. Therefore $H(X_1 X_2 \dots X_m | Z) = S_{no-r}(X_1; X_2; \dots; X_m \| Z) + T(X_1; X_2; \dots; X_m \| Z)$.

Proof of Theorem 4. We prove that $\psi'(X_1; X_2; X_3; \dots; X_m \| Z)$ satisfies the four properties of Theorem 2. We prove this in two stages: First we assume that C.H. $\{\psi\}$ satisfies the last three properties and prove that $\psi'(X_1; X_2; X_3; \dots; X_m \| Z)$ satisfies all four properties; and then we prove that C.H. $\{\psi\}$ satisfies the last three properties.

Property number 1: $\psi'(X_1; X_2; X_3; \dots; X_m \| Z) = \inf_{r=1,2,\dots} \frac{1}{r} C.H.\{\psi\}(X_1^r; X_2^r; \dots; X_m^r \| Z^r) \leq$

$$\begin{aligned} \inf_{r=n,2n,\dots} \frac{1}{r} \text{C.H.}\{\psi\}(X_1^r; X_2^r; \dots; X_m^r \| Z^r) &= \\ \inf_{r=1,2,\dots} \frac{1}{nr} \text{C.H.}\{\psi\}((X_1^n)^r; (X_2^n)^r; \dots; (X_m^n)^r \| (Z^n)^r) &= \\ \frac{1}{n} \psi'(X_1^n; X_2^n; X_3^n; \dots; X_m^n \| Z^n). & \end{aligned}$$

Property number 2: Assuming that $\text{C.H.}\{\psi\}$ satisfies property number (2), we have $\psi'(X_1; X_2; X_3; \dots; X_m \| Z) = \inf_r \frac{1}{r} \text{C.H.}\{\psi\}(X_1^r; X_2^r; \dots; X_m^r \| Z^r) \leq \inf_r \frac{1}{r} [\text{C.H.}\{\psi\}(X_1^r F^r; X_2^r F^r; \dots; X_m^r F^r \| Z^r F^r) + H(F^r | Z^r)] = \inf_r \frac{1}{r} \text{C.H.}\{\psi\}(X_1^r F^r; X_2^r F^r; \dots; X_m^r F^r \| Z^r F^r) + H(F | Z) = \psi'(X_1 F; X_2 F; X_3 F; \dots; X_m F \| Z F) + H(F | Z).$

The two other properties can be proved similarly (the details are suppressed).

Now, we prove that $\text{C.H.}\{\psi\}$ satisfies properties (2-4). $\text{C.H.}\{\psi\} = \sup_J \sum_j P(J = j) \cdot \psi(X_1^j; X_2^j; X_3^j; \dots; X_m^j \| Z^j)$ where J is a jointly distributed random variable with $(X_1; X_2; X_3; \dots; X_m; Z)$, and $(X_1^j; X_2^j; X_3^j; \dots; X_m^j; Z^j)$ is distributed according to the probability distribution of $p(x_1; x_2; x_3; \dots; x_m; z | J = j)$. We have also assumed that ψ satisfies properties (1-4).

Property number 2: $\text{C.H.}\{\psi\}(X_1; X_2; X_3; \dots; X_m \| Z) = \sup_J \sum_j P(J = j) \cdot \psi(X_1^j; X_2^j; X_3^j; \dots; X_m^j \| Z^j) \leq \sup_J [\sum_j P(J = j) \cdot \psi(X_1^j F^j; X_2^j F^j; \dots; X_m^j F^j \| Z^j F^j) + H(F | Z, J = j)] = \sup_J [\sum_j P(J = j) \cdot \psi(X_1^j F^j; X_2^j F^j; \dots; X_m^j F^j \| Z^j F^j) + H(F | Z J)] \leq \text{C.H.}\{\psi\}(X_1 F; X_2 F; \dots; X_m F \| Z F) + H(F | Z)$

The other properties can be proved similarly (the details are suppressed).

Proof of Theorem 5. $\inf_{XY-Z-\bar{Z}} I(X; Y | \bar{Z})$ is an upper bound on $S_{no-r}(X; Y | Z)$ that satisfies properties (1-4) of Theorem 1. Theorem 3 implies that $\psi(X; Y | Z) = H(XY | Z) - \inf_{XY-Z-\bar{Z}} I(X; Y | \bar{Z})$ is a lower bound on $T(X; Y | Z)$. According to Theorem 4, $\psi'(X; Y | Z) = \inf_r \frac{1}{r} \text{C.H.}\{\psi\}(X^r; Y^r \| Z^r)$ would also be a lower bound on $T(X; Y | Z)$. Since $\psi(X; Y | Z) \geq H(XY | Z) - I(X; Y) := \tilde{\psi}(X; Y | Z)$, we have: $\psi'(X; Y | Z) \geq \inf_r \frac{1}{r} \text{C.H.}\{\tilde{\psi}\}(X^r; Y^r \| Z^r) := \tilde{\psi}'(X; Y | Z)$. It can be proved that $\tilde{\psi}'(X; Y | Z) = \text{C.H.}\{\tilde{\psi}\}(X; Y | Z)$. Therefore $\sup_J (H(XY | Z J) - I(X; Y | J))$ would be a lower bound on $T(X; Y | Z)$ and hence $\inf_J I(X; Y | J) + I(XY; J | Z)$ would be an upper bound on $S_{no-r}(X; Y | Z)$. It turns out that $\inf_J [I(X; Y | J) + I(XY; J | Z)]$ satisfies the properties (1-4) of Theorem 1, so nothing can be gained by using it instead of $\inf_{XY-Z-\bar{Z}} I(X; Y | \bar{Z})$ to initiate the proof. It can be proved that $\inf_J [I(X; Y | J) + I(XY; J | Z)]$ is also an upper bound on $S(X; Y | Z)$. The proof uses the fact that for every pair (M_1, M_2) satisfying: $p(M_1, M_2, X, Y, Z) = p(M_1) \cdot p(M_2) \cdot p(X, Y, Z)$ we have: $\inf_J [I(X; Y | J) + I(XY; J | Z)] = \inf_J [I(X M_1; Y M_2 | J) + I(X M_1 Y M_2; J | Z)]$. In order to prove that the new bound is not worse than the double intrinsic information bound, it is sufficient to prove that for any random variable U , there is a random variable J such that $I(X; Y | J) + I(XY; J | Z) \leq [H(U) + \min_{\bar{Z}: X-Y-ZU-\bar{Z}} I(X; Y | \bar{Z})]$. Choosing $J = \bar{Z}$, we will have $I(X; Y | J) = I(X; Y | \bar{Z})$ and also $I(XY; J | Z) = I(XY; U | Z) - I(XY; U | Z J) \leq I(XY; U | Z) \leq H(U)$.

Therefore the new bound is no worse than the double intrinsic information bound. Appendix I contains an example for which the new bound is strictly better than the double intrinsic information bound. Infimum over J could be replaced by minimum over J using Carathodory's theorem (details are suppressed).

V. DISCUSSION

The process of finding new upper bounds in Thm. 5 could be generalized to the case of m users. We can prove similar results on the secret key rate when randomization is allowed at the terminals by adding the following condition to the four conditions of Theorem 1: For any set of random variables M_1, M_2, \dots, M_m being jointly independent of each other and of the pair (X_1, X_2, \dots, X_m) , $\varphi(X_1; X_2; \dots; X_m \| Z) = \varphi(X_1 M_1; X_2 M_2; \dots; X_m M_m \| Z F)$.

Other upper bounds could be derived by noting that $\inf_{J_1, J_2, \dots, J_t} [\max_i (S_{no-r}(X_1; X_2; \dots; X_m \| J_i)) + S_{no-r}(X_1; X_2; \dots; X_m; J_1^{(s)}; J_2^{(s)}; \dots; J_t^{(s)} \| Z)]$ satisfies the four properties of Thm. 1 and therefore is an upper bound on $S_{no-r}(X_1; X_2; \dots; X_m \| Z)$. In this formulation $S(X_1; X_2; \dots; X_m; J_1^{(s)}; J_2^{(s)}; \dots; J_t^{(s)} \| Z)$ refers to the secret key rate for the problem in which the $m + t$ parties $X_1; X_2; \dots; X_m; J_1; J_2; \dots; J_t$ are desiring to create a common secret key hidden from Z but that the last t parties $J_1; J_2; \dots; J_t$ are silent and do not participate in the communications. By taking $t = 1$ and $n = 2$, we get $S(X; Y | Z) \leq \inf_J [S(X; Y | J) + S(X; Y; J^s | Z)] \leq \inf_J [I(X; Y | J) + S(XY; J^s | Z)]$. The value of $S(XY; J^s | Z)$, the one way secret key rate from XY to J in the presence of Z , is known. The $\inf_J [I(X; Y | J) + S(XY; J^s | Z)]$ upper bound is always less than or equal to the $\inf_J [I(X; Y | J) + I(XY; J | Z)]$ upper bound and may improve it.

We have also studied the secret key generation in the channel model case studied by I. Csiszár and P. Narayan [8]. Similar techniques could be carried over to the channel model case and for example it can be proved that $C_s(P(y, z|x)) \leq \sup_{p(x)} [\inf_J I(X; Y | J) + I(XY; J | Z)]$. The details will be provided in another paper.

VI. CONCLUSION

We have derived a new upper bound on the secret key rate which generalizes and improves the double intrinsic information bound of [2] to the multi-terminal case. We have also strengthened the results of [1] via a newly formulated problem of communication for omniscience by a neutral observer.

APPENDIX I

In this Appendix we prove existence of a joint probability distribution on X, Y, Z for which the new bound is strictly better than the double intrinsic information bound. We need the following Lemmas which we will prove at the end of this Appendix:

Lemma A1.1 Assume that $\inf_U [H(U) + I(X; Y \downarrow ZU)] = \min_J [I(X; Y | J) + I(XY; J | Z)]$, then there is a sequence of

TABLE I
JOINT PROBABILITY DISTRIBUTION OF X AND Y

Y	X			
	0	1	2	3
0	$\frac{1}{8}$	$\frac{1}{8}$	0	0
1	$\frac{1}{8}$	$\frac{1}{8}$	0	0
2	0	0	$\frac{1}{4}$	0
3	0	0	0	$\frac{1}{4}$

random variables U_i , $i = 1, 2, \dots$ taking values in finite sets Ω_i , and a sequence of positive real numbers δ_i converging to zero, such that:

- 1) $H(U_i) + I(X; Y \downarrow ZU_i) \rightarrow \inf_U [H(U) + I(X; Y \downarrow ZU)]$ as $i \rightarrow \infty$
- 2) $H(U_i | XYZ) \rightarrow 0$ as $i \rightarrow \infty$
- 3) $I(U_i; Z) \rightarrow 0$ as $i \rightarrow \infty$
- 4) $|p(U_i = u_j | X = x, Y = y, Z = z) - \frac{1}{2}| \geq \frac{1}{2} - \delta_i \quad \forall u_j \in \Omega_i, (x, y, z) : p(x, y, z) > 0$
- 5) The variational distance $d(\mathfrak{S}(U_i | Z = z_i), \mathfrak{S}(U_i | Z = z_j)) \rightarrow 0$ as $i \rightarrow \infty \quad \forall z_i, z_j : p(Z = z_i) > 0, p(Z = z_j) > 0$

Lemma A1.2 Continuity of $I(X; Y \downarrow Z)$: $\forall \xi > 0, \exists \delta > 0$ such that for all random variables T having entropy less than δ , we have $|I(X; Y \downarrow ZT) - I(X; Y \downarrow Z)| < \xi$. \square

We will perturb the example that Renner and Wolf provided in [2] in order to prove that the double intrinsic information bound can be strictly better than the intrinsic information bound of Maurer [3]. Table (I) shows the joint probability distribution between X and Y in that example. Z is defined as:

$$Z = \begin{cases} X + Y \pmod{2} & \text{if } X, Y \in \{0, 1\} \\ X \pmod{2} & \text{if } X \in \{2, 3\} \end{cases}$$

Renner and Wolf proved that for the choice of $U = \lfloor \frac{X}{2} \rfloor$, one has

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

And therefore their bound would be less than or equal to $H(U) + I(X; Y \downarrow ZU) = 1$, while $I(X; Y \downarrow Z) = \frac{3}{2} > 1$.

Take a binary random variable V satisfying the $\tilde{V} - U - XYZ$ Markov property and the following property $\{0, p(U = 0 | V = 0), 1 - p(U = 0 | V = 0), \frac{1}{2}p(U = 0 | V = 0), 1 - \frac{1}{2}p(U = 0 | V = 0), 1\} \cap \{0, p(U = 0 | V = 1), 1 - p(U = 0 | V = 1), \frac{1}{2}p(U = 0 | V = 1), 1 - \frac{1}{2}p(U = 0 | V = 1), 1\} = \{0, 1\}$.

Let $\tilde{X} = X$, $\tilde{Y} = Y$, $\tilde{Z} = (Z, V)$. We would like to prove that the new bound is strictly better than the double intrinsic information bound for the triple $(\tilde{X}, \tilde{Y}, \tilde{Z})$.

Assuming that the new bound is not better than the double intrinsic information bound, we can apply Lemma A1.1 to get a sequence U_i having the five properties given in Lemma A1.1. Using the properties number 2,3, and 4 and the fact that $h(x) = x \log(\frac{1}{x})$ is monotonic for all $x < \frac{1}{2}$ and for all $x > \frac{1}{2}$, it can be proved that $H(U_i) \rightarrow 0$ as $i \rightarrow \infty$ (the details are suppressed).

Hence, the limit of $H(U_i) + I(\tilde{X}; \tilde{Y} \downarrow \tilde{Z}U_i)$ is the same as that of $I(\tilde{X}; \tilde{Y} \downarrow \tilde{Z}U_i)$. The property number 1 of Lemma A1.1 states that the series converges to the double intrinsic information upper bound which is assumed to be equal to $\min_J [I(\tilde{X}; \tilde{Y} | J) + I(\tilde{X}\tilde{Y}; J | \tilde{Z})]$.

Evaluating the expression at $J = \tilde{Z}U$, gives us $0 + I(XY; UZV | ZV) = I(XY; U | ZV) \leq 1$

Therefore we should have: $\lim_{i \rightarrow \infty} I(X; Y \downarrow ZVU_i) \leq 1$. On the other hand, Renner and Wolf have shown that $I(X; Y \downarrow Z) = \frac{3}{2}$. Letting $H(V) \rightarrow 0$, this would be in contradiction with Lemma A1.2 noting that $H(U_i) \rightarrow 0$ as $i \rightarrow \infty$.

Now, we prove the Lemmas mentioned at the beginning of this Appendix:

Proof of Lemma A1.1 Take a sequence U_1, U_2, \dots such that $H(U_i) + I(X; Y \downarrow ZU_i) \rightarrow \inf_U [H(U) + I(X; Y \downarrow ZU)]$.

For every U_i , there exists J_i such that $I(X; Y \downarrow ZU_i) = I(X; Y | J_i)$, and also $XY - ZU_i - J_i$ forming a Markov chain.

We have: $I(XY; J_i | Z) = I(XY; U_i | Z) - I(XY; U_i | ZJ_i) \leq I(XY; U_i | Z) = H(U_i | Z) - H(U_i | XYZ) = H(U_i) - I(U_i; Z) - H(U_i | XYZ)$. Hence $H(U_i) + I(X; Y \downarrow ZU_i) \geq [I(U_i; Z) + H(U_i | XYZ)] + [I(X; Y | J_i) + I(XY; J_i | Z)] \geq [I(U_i; Z) + H(U_i | XYZ)] + \min_J [I(X; Y | J) + I(XY; J | Z)] = [I(U_i; Z) + H(U_i | XYZ)] + \inf_U [H(U) + I(X; Y \downarrow ZU)]$.

Taking the limit as $i \rightarrow \infty$, we conclude that $[I(U_i; Z) + H(U_i | XYZ)] \rightarrow 0$ as $i \rightarrow \infty$. Properties 1-4 could be proved using this statement (the details are suppressed).

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 number CCF-0500023.

REFERENCES

- [1] I. Csiszár and P. Narayan, *Secrecy Capacities for Multiple Terminals*, IEEE Trans. Inform. Theory, Vol. 50, No. 12, Dec 2004.
- [2] 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, 2003
- [3] U. Maurer and S. Wolf, *Unconditionally Secure Key Agreement and the Intrinsic Mutual Information*, IEEE Trans. Inform. Theory, Vol. 45, No. 2, pp. 499 -514, 1999.
- [4] R. Ahlswede and I. Csiszár, *Common randomness in Information Theory and Cryptography. Part I: Secret sharing*, IEEE Trans. Inform. Theory, Vol. 39, No. 4, pp. 1121 -1132, 1993.
- [5] Ueli M. Maurer, *Secret Key Agreement by Public Discussion From Common Information*, IEEE Trans. Inform. Theory, Vol. 39, p. 733-742, 1993.
- [6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.
- [7] I. Csiszár and J. Körner, *Broadcast Channels with Confidential Messages*, IEEE Trans. Inform. Theory, Vol. 24, No. 3, pp. 339-348, 1978.
- [8] I. Csiszár and P. Narayan, *Secrecy Capacities for Multiterminal Channel Models*, IEEE International Symposium on Information Theory, pp.2138 - 2141, 2005.
- [9] C.E. Shannon, *Communication Theory of Secrecy Systems*, Bell System Technical Journal, Vol. 28, p. 656-715, Oct 1949.