

When is a Function Securely Computable?

Himanshu Tyagi*, Prakash Narayan* and Piyush Gupta†

*Dept. of Electrical and Computer Engineering
and Institute for Systems Research
University of Maryland
College Park, MD 20742, USA
Email: {tyagi, prakash}@umd.edu

†Bell Labs, Alcatel-Lucent
Murray Hill, NJ 07974, USA
Email: pgupta@research.bell-labs.com

Abstract—A subset of a set of terminals that observe correlated signals seek to compute a given function of the signals using public communication. It is required that the value of the function be kept secret from an eavesdropper with access to the communication. We show that the function is securely computable if and only if its entropy is less than the “aided secret key” capacity of an associated secrecy generation model, for which a single-letter characterization is provided.

I. INTRODUCTION

Suppose that the terminals in $\mathcal{M} = \{1, \dots, m\}$ observe correlated signals, and that a subset $\mathcal{A} = \{1, \dots, a\}$ of them are required to compute “securely” a given (single-letter) function g of all the signals. To this end, following their observations, all the terminals are allowed to communicate interactively over a public noiseless channel of unlimited capacity, with all such communication being observed by all the terminals. The terminals in \mathcal{A} seek to compute g in such a manner as to keep its value information theoretically secret from an eavesdropper with access to the public interterminal communication. A typical application arises in a wireless network of colocated sensors which seek to compute a given function of their correlated measurements using public communication that does not give away the value of the function.

Our goal is to characterize necessary and sufficient conditions under which such secure computation is feasible. We formulate a new Shannon theoretic multiterminal source model that addresses the elemental question: *When can a function g be computed so that its value is independent of the public communication used in its computation?*

In [9], we had addressed the special case $\mathcal{A} = \mathcal{M}$ and shown that g is securely computable if and only if its entropy does not exceed the secret key capacity of a standard secrecy generation model [7], [1], [5], [6]. In the general case considered here, a natural extension of this result does not hold, as seen from the Example 1 given in Appendix A, which partly motivates the present work.

We establish that the answer to the question above is innately connected to a *new* problem of secret key (SK) generation in which all the terminals in \mathcal{M} seek to generate “secret

common randomness” at the largest rate possible, when the terminals in $\mathcal{A}^c = \mathcal{M}/\mathcal{A}$ are provided with side information for limited use, by means of public communication from which an eavesdropper can glean only a negligible amount of information about the SK. The public communication from a terminal can be any function of its own observed signal and of all previous communication. Side information is provided to the terminals in \mathcal{A}^c in the form of the value of g , and can be used only for recovering the key. Such a key, termed an aided secret key (ASK), constitutes a modification of the original notion of a SK in [7], [1], [5], [6]. The largest rate of such an ASK is the ASK capacity C . Since a securely computable function g for \mathcal{A} will yield an ASK (for \mathcal{M}) of rate equal to its entropy H , it is clear that g necessarily must satisfy $H \leq C$. We show that surprisingly, $H < C$ is a sufficient condition for the existence of a protocol for the secure computation of g for \mathcal{A} . When all the terminals in \mathcal{M} seek to compute g securely, the corresponding ASK capacity reduces to the standard SK capacity for \mathcal{M} [5], [6]. Furthermore, we show that a function that is securely computed by \mathcal{A} can be augmented by residual secret common randomness to yield a SK for \mathcal{A} of optimum rate, bringing out an operational decomposition of the entropy in the model.

Preliminaries and the problem formulation are contained in section II. The characterization of the secure computability of g is provided in Section III and a decomposition result for the total entropy of the model is provided in Section IV. The proof of the characterization of the secure computability of g is given in Section V followed by the concluding remarks in Section VI. A full-length version of this submission is currently under review [10].

II. PRELIMINARIES

Let X_1, \dots, X_m , $m \geq 2$, be rvs with finite alphabets $\mathcal{X}_1, \dots, \mathcal{X}_m$, respectively. For any nonempty set $A \subseteq \mathcal{M} = \{1, \dots, m\}$, we denote $X_A = (X_i, i \in A)$. Similarly, for real numbers R_1, \dots, R_m and $A \subseteq \mathcal{M}$, we denote $R_A = (R_i, i \in A)$. Let A^c be the set $\mathcal{M} \setminus A$. We denote n i.i.d. repetitions of $X_{\mathcal{M}} = (X_1, \dots, X_m)$ with values in

$\mathcal{X}_{\mathcal{M}} = \mathcal{X}_1 \times \dots \times \mathcal{X}_m$ by $X_{\mathcal{M}}^n = (X_1^n, \dots, X_m^n)$ with values in $\mathcal{X}_{\mathcal{M}}^n = \mathcal{X}_1^n \times \dots \times \mathcal{X}_m^n$. Following [5], given $\epsilon > 0$, for rvs U, V , we say that U is ϵ -recoverable from V if $\Pr(U \neq f(V)) \leq \epsilon$ for some function $f(V)$ of V . All logarithms and exponentials are with respect to the base 2.

We consider a multiterminal source model for secure computation with public communication; this basic model was introduced in [5] in the context of SK generation with public transaction. Terminals $1, \dots, m$ observe, respectively, the sequences X_1^n, \dots, X_m^n , of length n . Let $g : \mathcal{X}_{\mathcal{M}} \rightarrow \mathcal{Y}$ be a given mapping, where \mathcal{Y} is a finite alphabet. For $n \geq 1$, the (single-letter) mapping $g^n : \mathcal{X}_{\mathcal{M}}^n \rightarrow \mathcal{Y}^n$ is defined by

$$g^n(x_{\mathcal{M}}^n) = (g(x_{11}, \dots, x_{m1}), \dots, g(x_{1n}, \dots, x_{mn})),$$

$$x_{\mathcal{M}}^n = (x_1^n, \dots, x_m^n) \in \mathcal{X}_{\mathcal{M}}^n.$$

For convenience, we shall denote the rv $g^n(X_{\mathcal{M}}^n)$ by G^n , $n \geq 1$, and, in particular, $G^1 = g(X_{\mathcal{M}})$ simply by G . The terminals in a given set $\mathcal{A} \subseteq \mathcal{M}$ wish to “compute securely” the function $g^n(x_{\mathcal{M}}^n)$, $x_{\mathcal{M}}^n \in \mathcal{X}_{\mathcal{M}}^n$. To this end, the terminals are allowed to communicate over a noiseless public channel, possibly interactively in several rounds. Randomization at the terminals is permitted; we assume that terminal i generates a rv U_i , $i \in \mathcal{M}$, such that U_1, \dots, U_m and $X_{\mathcal{M}}^n$ are mutually independent. While the cardinalities of range spaces of $U_i, i \in \mathcal{M}$, are unrestricted, we assume that $H(U_{\mathcal{M}}) < \infty$.

Definition 1. Assume without any loss of generality that the communication of the terminals in \mathcal{M} occurs in consecutive time slots in r rounds; such communication is described in terms of the mappings

$$f_{11}, \dots, f_{1m}, f_{21}, \dots, f_{2m}, \dots, f_{r1}, \dots, f_{rm},$$

with f_{ji} corresponding to a message in time slot j by terminal i , $1 \leq j \leq r$, $1 \leq i \leq m$; in general, f_{ji} is allowed to yield any function of (U_i, X_i^n) and of previous communication described in terms of $\{f_{kl} : k < j, l \in \mathcal{M} \text{ or } k = j, l < i\}$. The corresponding rvs representing the communication will be depicted collectively as

$$\mathbf{F} = \{F_{11}, \dots, F_{1m}, F_{21}, \dots, F_{2m}, \dots, F_{r1}, \dots, F_{rm}\},$$

where $\mathbf{F} = \mathbf{F}^{(n)}(U_{\mathcal{M}}, X_{\mathcal{M}}^n)$. A special form of such communication will be termed *noninteractive communication* if $\mathbf{F} = (F_1, \dots, F_m)$, where $F_i = f_i(U_i, X_i^n)$, $i \in \mathcal{M}$.

Definition 2. For $\epsilon_n > 0, n \geq 1$, we say that g is ϵ_n -securely computable (ϵ_n -SC) by (the terminals in) a given set $\mathcal{A} \subseteq \mathcal{M}$ with $|\mathcal{A}| \geq 1$ from observations of length n , randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F} = \mathbf{F}^{(n)}$, if

(i) g^n is ϵ_n -recoverable from (U_i, X_i^n, \mathbf{F}) for every $i \in \mathcal{A}$, i.e., there exists $\hat{g}_i^{(n)}$ satisfying

$$\Pr(\hat{g}_i^{(n)}(U_i, X_i^n, \mathbf{F}) \neq G^n) \leq \epsilon_n, \quad i \in \mathcal{A}, \quad (1)$$

and

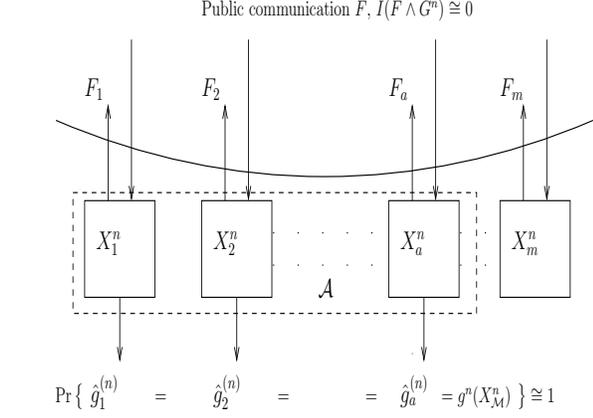


Fig. 1. Secure computation of g

(ii) g^n satisfies the “strong” secrecy condition¹

$$I(G^n \wedge \mathbf{F}) \leq \epsilon_n. \quad (2)$$

By definition, an ϵ_n -SC function g is recoverable (as g^n) at the terminals in \mathcal{A} and is effectively concealed from an eavesdropper with access to the public communication \mathbf{F} .

Definition 3. We say that g is *securely computable* by \mathcal{A} if g is ϵ_n -SC by \mathcal{A} from observations of length n , suitable randomization $U_{\mathcal{M}}$ and public communication \mathbf{F} , such that $\lim_n \epsilon_n = 0$.

Figure 1 shows the setup for secure computing.

III. WHEN IS g SECURELY COMPUTABLE?

A characterization of securely computable functions will be seen to be linked inherently to a new SK generation problem that is formulated next.

We consider an extension of the SK generation problem [7], [1], [5], [6] which now involves additional side information G^n that is provided to the terminals not in \mathcal{A} for use in *only the recovery stage* of SK generation.

Definition 4. For $\epsilon_n > 0, n \geq 1$, a function K of $X_{\mathcal{M}}^n$ is an ϵ_n -aided secret key (ϵ_n -ASK) for (the terminals in) \mathcal{M} , achievable from observations of length n , randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F} = \mathbf{F}^{(n)}(U_{\mathcal{M}}, X_{\mathcal{M}}^n)$ as above, if

(i) K is ϵ_n -recoverable from (U_i, X_i^n, \mathbf{F}) for every $i \in \mathcal{A}$ and from $(U_i, X_i^n, G^n, \mathbf{F})$ for every $i \in \mathcal{A}^c$;

(ii) K satisfies the “strong” secrecy condition

$$\log |\mathcal{K}| - H(K | \mathbf{F}) = \log |\mathcal{K}| - H(K) + I(K \wedge \mathbf{F}) \leq \epsilon_n, \quad (3)$$

where $\mathcal{K} = \mathcal{K}^{(n)}$ denotes the set of possible values of K . The ASK capacity $C_{ASK}(\mathcal{M}) = C_{ASK}(\mathcal{M}; g, \mathcal{A})$ is the largest rate $\lim_n (1/n)H(K)$ of ϵ_n -ASKs such that $\lim_n \epsilon_n = 0$.

¹The notion of strong secrecy for SK generation was introduced in [8], and developed further in [2], [4].

Remark. The secrecy condition (3) is tantamount jointly to a nearly uniform distribution for K (i.e., $\log |\mathcal{K}| - H(K)$ is small) and to the near independence of K and \mathbf{F} (i.e., $I(K \wedge \mathbf{F})$ is small).

A single-letter characterization of the ASK capacity $C_{ASK}(\mathcal{M})$ is given by the following theorem.

Theorem 1. *The ASK capacity $C_{ASK}(\mathcal{M})$ equals*

$$C_{ASK}(\mathcal{M}) = H(X_{\mathcal{M}}) - R_g(\mathcal{A}) \quad (4)$$

where

$$R_g(\mathcal{A}) = \min_{R_{\mathcal{M}} \in \mathcal{R}_g(\mathcal{A})} \sum_{i \in \mathcal{M}} R_i \quad (5)$$

with

$$\begin{aligned} \mathcal{R}_g(\mathcal{A}) = \{ & R_{\mathcal{M}} : \forall B \subsetneq \mathcal{M}, \\ & R_B \geq H(X_B | X_{B^c}), \text{ if } \mathcal{A} \not\subseteq B \\ & R_B \geq H(X_B | X_{B^c}, G), \text{ if } \mathcal{A} \subseteq B \}. \end{aligned}$$

Furthermore, the ASK capacity can be achieved with noninteractive communication and without recourse to randomization at the terminals in \mathcal{M} .

The proof of Theorem 1 can be given along the lines of [5, Theorem 1] or as a consequence of a general result in our extended paper [10, Theorem 4].

Remark. It is seen that the ASK capacity $C_{ASK}(\mathcal{M})$ is not increased if the secrecy condition (3) is replaced by the following weaker requirement:

$$\frac{1}{n} I(K \wedge \mathbf{F}) \leq \epsilon_n. \quad (6)$$

A comparison of the conditions in (2) and (6) that must be met by a securely computable g and an ASK K , respectively, shows for a given g to be securely computable, it is necessary that

$$H(G) \leq C_{ASK}(\mathcal{M}). \quad (7)$$

Our main result says that the necessary condition (7) is tight.

Theorem 2. *A function g is securely computable by $\mathcal{A} \subseteq \mathcal{M}$ if*

$$H(G) < C_{ASK}(\mathcal{M}). \quad (8)$$

Furthermore, under the condition above, g is securely computable with noninteractive communication and without recourse to randomization at the terminals in \mathcal{M} .

Conversely, if g is securely computable by $\mathcal{A} \subseteq \mathcal{M}$, then $H(G) \leq C_{ASK}(\mathcal{M})$.

IV. DECOMPOSITION RESULTS

We recall from [5] the definition of a (standard) SK for the terminals in \mathcal{A} as a rv $K = K(X_{\mathcal{M}}^n)$ that satisfies the conditions of Definition 4 but without the recoverability requirement for the terminals in \mathcal{A}^c in (i). The corresponding

SK capacity for \mathcal{A} , denoted by $C_{SK}(\mathcal{A})$, equals [5], [6]

$$C_{SK}(\mathcal{A}) = H(X_{\mathcal{M}}) - R_{CO}(\mathcal{A}), \quad (9)$$

where

$$R_{CO}(\mathcal{A}) = \min_{R_{\mathcal{M}} \in \mathcal{R}(\mathcal{A})} \sum_{i=1}^m R_i$$

with

$$\mathcal{R}(\mathcal{A}) = \left\{ R_{\mathcal{M}} : B \subsetneq \mathcal{M}, \mathcal{A}' \not\subseteq B, \right. \\ \left. R_B \geq H(X_B | X_{B^c}) \right\}.$$

The sufficiency condition (8) prompts the following two natural questions: Does the difference $C_{ASK}(\mathcal{M}) - H(G)$ possess an operational significance? If g is securely computable by terminals in \mathcal{A} , clearly G^n forms a SK for \mathcal{A} . Can G^n be augmented suitably to form a SK (see [5]) for \mathcal{A} of maximum achievable rate?

The answers to both these questions are in the affirmative. In particular, our approach to the second question involves a characterization of the minimum rate of communication for omniscience for \mathcal{A} , under the additional requirement that this communication be independent of G^n . Specifically, we show that for a securely computable function g , this minimum rate remains $R_{CO}(\mathcal{A})$, i.e., the minimum rate of communication for omniscience for \mathcal{A} in absence of additional secrecy constraints on the communication.

Addressing the first question, we introduce a rv $K_g = K_g^{(n)}$ such that $K = (K_g, G^n)$ constitutes an ϵ_n -ASK satisfying the additional requirement

$$I(K_g \wedge G^n) \leq \epsilon_n. \quad (10)$$

Let the largest rate $\lim_n (1/n) H(K_g)$ of such an ASK be $C^g(\mathcal{A})$. Observe that since K is required to be nearly independent of \mathbf{F} , where \mathbf{F} is the public communication involved in its formation, it follows by (10) that K_g is nearly independent of (G^n, \mathbf{F}) .

Turning to the second question, in the same vein let K'_g be a rv such that $K' = (K'_g, G^n)$ constitutes an ϵ_n -SK for $\mathcal{A} \subseteq \mathcal{M}$ and satisfying (10). Let $C^{g'}(\mathcal{A})$ denote the largest rate of K'_g . As noted above, K'_g will be nearly independent of (G^n, \mathbf{F}') , where \mathbf{F}' is the public communication involved in the formation of K' .

Proposition 3. *For $\mathcal{A} \subseteq \mathcal{M}$, it holds that*

- (i) $C^g(\mathcal{A}) = C_{ASK}(\mathcal{M}) - H(G)$,
- (ii) $C^{g'}(\mathcal{A}) = C_{SK}(\mathcal{A}) - H(G)$.

Remarks. (i) For the case $\mathcal{A} = \mathcal{M}$, both (i) and (ii) above reduce to $C^g(\mathcal{M}) = C_{SK}(\mathcal{M}) - H(G)$.

(ii) Proposition 3 (ii) and (9) lead to the observation

$$H(X_{\mathcal{M}}) = R_{CO}(\mathcal{A}) + H(G) + C^{g'}(\mathcal{A}),$$

which admits the following heuristic interpretation. The “total

randomness” $X_{\mathcal{M}}^n$ that corresponds to omniscience decomposes into three “nearly mutually independent” components: a minimum-sized communication for omniscience for \mathcal{A} and the independent parts of an optimum-rate SK for \mathcal{A} composed of G^n and K'_g .

V. OUTLINE OF THE PROOF OF THEOREM 2

The necessity of (7) follows by the comments preceding Theorem 2.

The sufficiency of (8) will be established by showing the existence of *noninteractive* public communication comprising source codes that enable omniscience corresponding to $X_{\mathcal{M}}^n$ at the terminals in \mathcal{A} , and thereby the computation of g . Furthermore, the corresponding codewords are selected so as to be simultaneously independent of G^n , thus assuring security.

First, from (8) and (4), there exists $\delta > 0$ such that $R_g(\mathcal{A}) + \delta < H(X_{\mathcal{M}}|G)$, using $G = g(X_{\mathcal{M}})$. For each i and $R_i \geq 0$, consider a (map-valued) rv J_i that is uniformly distributed on the family \mathcal{J}_i of all mappings $\mathcal{X}_i^n \rightarrow \{1, \dots, \lceil \exp(nR_i) \rceil\}$, $i \in \mathcal{M}$. The rvs $J_1, \dots, J_m, X_{\mathcal{M}}^n$ are taken to be mutually independent. Denote $Z_{\mathcal{M}} = Z_{\mathcal{M}}(\mathcal{A}) = \{Z_i\}_{i \in \mathcal{M}}$ with

$$Z_i = \begin{cases} 0, & i \in \mathcal{A} \\ G, & i \in \mathcal{A}^c. \end{cases} \quad (11)$$

Fix ϵ, ϵ' , with $\epsilon' > m\epsilon$ and $\epsilon + \epsilon' < 1$. It follows from the proof of the general source network coding theorem [3, Lemma 3.1.13 and Theorem 3.1.14] that for all sufficiently large n ,

$$\Pr\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : X_{\mathcal{M}}^n \text{ is } \epsilon_n\text{-recoverable from } \left(X_i^n, j_{\mathcal{M} \setminus \{i\}} \left(X_{\mathcal{M} \setminus \{i\}}^n\right), Z_i^n\right), i \in \mathcal{M}\right\}\right) \geq 1 - \epsilon, \quad (12)$$

provided $R_{\mathcal{M}} = (R_1, \dots, R_m) \in \mathcal{R}_g(\mathcal{A})$, where ϵ_n vanishes exponentially rapidly in n . This assertion follows exactly as in the proof of [5, Proposition 1, with $A = \mathcal{M}$] but with \tilde{X}_i there equal to (X_i, Z_i) rather than X_i , $i \in \mathcal{M}$. In particular, we shall choose $R_{\mathcal{M}} \in \mathcal{R}_g(\mathcal{A})$ such that

$$\sum_{i=1}^m R_i \leq R_g(\mathcal{A}) + \frac{\delta}{2}. \quad (13)$$

Below we shall establish that

$$\Pr(\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : I(j_{\mathcal{M}}(X_{\mathcal{M}}^n) \wedge G^n) \geq \epsilon_n\}) \leq \epsilon', \quad (14)$$

for all n sufficiently large, to which end it suffices to show that

$$\Pr\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : I\left(j_i(X_i^n) \wedge G^n, j_{\mathcal{M} \setminus \{i\}} \left(X_{\mathcal{M} \setminus \{i\}}^n\right)\right) \geq \frac{\epsilon_n}{m}\right\}\right) \leq \frac{\epsilon'}{m}, \quad i \in \mathcal{M}, \quad (15)$$

since

$$\begin{aligned} & I(j_{\mathcal{M}}(X_{\mathcal{M}}^n) \wedge G^n) \\ & \leq \sum_{i=1}^m I\left(j_i(X_i^n) \wedge G^n, j_{\mathcal{M} \setminus \{i\}} \left(X_{\mathcal{M} \setminus \{i\}}^n\right)\right). \end{aligned}$$

Then it would follow from (12), (14) and definition of $Z_{\mathcal{M}}$ in (11) that

$$\Pr\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : G^n \text{ is } \epsilon_n\text{-recoverable from } \left(X_i^n, j_{\mathcal{M} \setminus \{i\}} \left(X_{\mathcal{M} \setminus \{i\}}^n\right)\right), i \in \mathcal{A}, \text{ and } I(j_{\mathcal{M}}(X_{\mathcal{M}}^n) \wedge G^n) < \epsilon_n\right\}\right) \geq 1 - \epsilon - \epsilon'.$$

This shows the existence of a particular realization $j_{\mathcal{M}}$ of $J_{\mathcal{M}}$ such that G^n is ϵ_n -SC from

$$\left(X_i^n, j_{\mathcal{M} \setminus \{i\}} \left(X_{\mathcal{M} \setminus \{i\}}^n\right)\right) \text{ for each } i \in \mathcal{A}.$$

It now remains to prove (15) leading to the existence of communication $j_{\mathcal{M}} = \{j_1, \dots, j_m\}$ that achieves omniscience while being independent simultaneously of G^n . This step of the proof is performed relying on our generalized version of the “balanced coloring Lemma” [5, Lemma B.2]. The generalization is stated without proof in the Appendix B. The details of this step are omitted. \square

VI. DISCUSSION

We obtain simple necessary and sufficient conditions for secure computability involving function entropy and ASK capacity. The latter is the largest rate of a SK for a new model in which side information is provided for use in only the recovery stage of SK generation. This model could be of independent interest. In particular, a function is securely computable if its entropy is less than ASK capacity of an associated secrecy model. The difference is shown to correspond to the maximum achievable rate of an ASK which is independent of the securely computed function and, together with it, forms an ASK of optimum rate. Also, a function that is securely computed by \mathcal{A} can be augmented to form a SK for \mathcal{A} of maximum rate.

Our results extend to functions defined on a block of symbols of *fixed* length in an obvious manner by considering larger alphabets composed of supersymbols of such length. However, they do not cover functions of symbols of increasing length (in n), e.g., a running average (in n).

In our proof of Theorem 2, g was securely computed from omniscience at all the terminals in $\mathcal{A} \subseteq \mathcal{M}$ that was attained using noninteractive public communication. However omniscience is not necessary for the secure computation of g , and it is possible to make do with communication of rate less than $R_{CO}(\mathcal{A})$ using an interactive protocol. A related unresolved question is: What is the minimum rate of public communication for secure computation?

A natural generalization of the conditions for secure computability of g by $\mathcal{A} \subseteq \mathcal{M}$ given here entails a characterization

of conditions for the secure computability of multiple functions g_1, \dots, g_k by subsets $\mathcal{A}_1, \dots, \mathcal{A}_k$ of \mathcal{M} , respectively. This unsolved problem, in general, will not permit omniscience for any $\mathcal{A}_i, i = 1, \dots, k$. For instance with $m = 2$, $\mathcal{A}_1 = \{1\}$, $\mathcal{A}_2 = \{2\}$, and X_1 and X_2 being independent, the functions $g_i(x_i) = x_i, i = 1, 2$, are securely computable trivially, but not through omniscience since, in this example, public communication is forbidden for the secure computation of g_1, g_2 .

VII. APPENDIX A

Example 1. Let $m = 3$, $A = \{1, 2\}$ and consider rvs X_1, X_2, X_3 with $X_1 = X_2$, where X_1 is independent of X_3 and $H(X_3) < H(X_1)$. Let g be defined by $g(x_1, x_2, x_3) = x_3, x_i \in \mathcal{X}_i, 1 \leq i \leq 3$. Clearly, $C_{SK}(\{1, 2\}) = H(X_1)$. Therefore, $H(G) = H(X_3) < C_{SK}(\{1, 2\})$. However, for g to be computed by the terminals 1 and 2, its value must be conveyed to them necessarily by public communication from terminal 3. Thus, g is not securely computable.

APPENDIX B

Our proof of Theorem 2 calls for a balanced coloring of a set corresponding to a rv that differs from another rv for which probability bounds are used. However, both rvs agree with high probability when conditioned on a set of interest.

Consider rvs U, U', V with values in finite sets $\mathcal{U}, \mathcal{U}', \mathcal{V}$, respectively, where U' is a function of U , and a mapping $h : \mathcal{U} \rightarrow \{1, \dots, r'\}$. For $\lambda > 0$, let \mathcal{U}_0 be a subset of \mathcal{U} such that

- (i) $\Pr(U \in \mathcal{U}_0) > 1 - \lambda^2$;
- (ii) given $U \in \mathcal{U}_0, h(U) = j, U' = u', V = v$, there exists $u = u(u') \in \mathcal{U}_0$ satisfying

$$\begin{aligned} \Pr(U = u \mid h(U) = j, V = v, U \in \mathcal{U}_0) \\ = \Pr(U' = u' \mid h(U) = j, V = v, U \in \mathcal{U}_0), \\ 1 \leq j \leq r', v \in \mathcal{V}. \end{aligned}$$

Then the following holds.

Lemma B1. *Let the rvs U, U', V and the set \mathcal{U}_0 be as above. Further, assume that*

$$P_{UV} \left(\left\{ (u, v) : \Pr(U = u \mid V = v) > \frac{1}{d} \right\} \right) \leq \lambda^2.$$

Then, a randomly selected mapping $\phi : \mathcal{U}' \rightarrow \{1, \dots, r\}$ fails to satisfy

$$\begin{aligned} \sum_{j=1}^{r'} \sum_{v \in \mathcal{V}} \Pr(h(U) = j, V = v) \times \\ \sum_{i=1}^r \left| \sum_{u' \in \mathcal{U}': \phi(u')=i} \Pr(U' = u' \mid h(U) = j, V = v) - \frac{1}{r} \right| \\ < 14\lambda, \end{aligned}$$

with probability less than $2rr'|\mathcal{V}| \exp\left(-\frac{c\lambda^3 d}{rr'}\right)$ for a constant $c > 0$.

ACKNOWLEDGEMENTS

The work of H. Tyagi and P. Narayan was supported by the U.S. National Science Foundation under Grants CCF0635271 and CCF0830697. P. Gupta acknowledges support from NSF Grant CNS-519535.

REFERENCES

- [1] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography—part i: Secret sharing," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1121–1132, 1993.
- [2] I. Csiszár, "Almost independence and secrecy capacity," *Prob. Pered. Inform.*, vol. 32, no. 1, pp. 48–57, 1996.
- [3] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless channels*. Academic Press, 1981.
- [4] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. Inform. Theory*, vol. 46, pp. 344–366, March 2000.
- [5] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3047–3061, 2004.
- [6] —, "Secrecy capacities for multiterminal channel models," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2437–2452, 2008.
- [7] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inform. Theory*, vol. 39, pp. 733–742, May 1993.
- [8] U. M. Maurer, *Communications and Cryptography: Two sides of One Tapestry*, R.E. Blahut et al., Eds. ed. Norwell, MA: Kluwer, 1994, ch. 26, pp. 271–285.
- [9] H. Tyagi, P. Narayan, and P. Gupta, "Secure computing," *Proc. Int. Symp. Inform. Theory*, pp. 2612 – 2616, June 2010.
- [10] H. Tyagi, P. Narayan, and P. Gupta, "When is a function securely computable?" *IEEE Trans. Inform. Theory*, submitted, 2010; [arXiv:1007.2945v1 \[cs.IT\]](https://arxiv.org/abs/1007.2945v1).