Distributed Function Computation with Confidentiality

Himanshu Tyagi[†]

Abstract—A set of terminals observe correlated data and seek to compute functions of the data using interactive public communication. At the same time, it is required that the value of a private function of the data remains concealed from an eavesdropper observing this communication. In general, the private function and the functions computed by the nodes can be all different. We show that a class of functions are securely computable if and only if the conditional entropy of data given the value of private function is greater than the least rate of interactive communication required for a related multiterminal source-coding task. A single-letter formula is provided for this rate in special cases.

Index Terms—Balanced coloring lemma, distributed computing, function computation, omniscience, secure computation.

I. INTRODUCTION

We consider the following distributed function computation problem with a confidentiality requirement. The terminals in a set $\mathcal{M} = \{1, ..., m\}$ observe correlated data, and wish to compute functions $g_1, ..., g_m$, respectively, of their collective data. To this end, they communicate interactively over a noiseless channel of unlimited capacity. It is required that this communication must not reveal the value of a specified private function g_0 of the data. If such a communication protocol exists, the functions $g_0, g_1, ..., g_m$ are said to be *securely computable*. We formulate a Shannon theoretic multiterminal source model that addresses the basic question: *When are the functions $g_0, g_1, ..., g_m$ securely computable?*

Applications of this formulation include distributed computing over public communication networks and function computation over sensor networks in hostile environments. In contrast to the classic notion of secure computing in cryptography [21], we assume that the nodes are trustworthy but their public communication network can be accessed by an eavesdropper. We examine the feasibility of certain distributed computing tasks without revealing a critical portion of the data to the eavesdropper; the function g_i, $i = 1, ..., m$, denotes the computation requirements of the ith terminal, while the critical data is represented by the value of private function g_0. As an example, consider a data download problem in a sensor network. The central server terminal 1 downloads binary data from terminals $2, ..., m$, while the latter terminals compute the

symbolwise maxima. An observer of the communication must not learn of the data of terminals $2, ..., m$.

The answer to the general question above remains open. The simplest case of interest when the terminals in a subset $\mathcal{A}$ of $\mathcal{M}$ compute only the private function g_0 and those not in $\mathcal{A}$ perform no computation was introduced in [19]. The data download problem, upon dropping the computation requirements for terminals $2, ..., m$, reduces to this setting. It was shown that if g_0 is securely computable (by the terminals in $\mathcal{A}$), then

$$H(X_{\mathcal{M}}|G_0) = H(X_{\mathcal{M}}) - H(G_0) \geq R^*, \qquad (1)$$

and g_0 is securely computable if

$$H(X_{\mathcal{M}}|G_0) > R^*, \qquad (2)$$

where R^* has the operational significance of being the minimum overall rate of communication needed for a specific multiterminal source-coding task that necessitates the recovery of entire data at all the terminals in $\mathcal{A}$; this task does not involve any security constraint (see Section II for a detailed discussion). Loosely speaking, denoting the collective data of the terminals by the random variable (rv) $X_{\mathcal{M}}$ and the random value of the function g_0 by the rv G_0, the maximum rate of randomness (in the data) that is independent of G_0 is $H(X_{\mathcal{M}}|G_0)$. The conditions above imply, in effect, that g_0 is securely computable if and only if this residual randomness of rate $H(X_{\mathcal{M}}|G_0)$ contains an interactive communication, of rate R^*, for the mentioned source-coding task.

In this paper, for a broad class of settings involving the secure computation of multiple functions, we establish necessary and sufficient conditions for secure computation of the same form as (1) and (2), respectively. The rate R^* now corresponds to, roughly, the minimum overall rate of communication that allows each terminal to:

(i) accomplish its required computation task, and,
(ii) along with the private function value, recover the entire data.

This characterization of secure computability is obtained via a general heuristic principle that leads to new results and further explains the results of [19] in a broader context.

Using the sufficient condition (2), we present a specific secure computing protocol in Section IV with a communication of rate R^*. Under (2), the secure computing scheme in [19] recovered the entire data, i.e., the collective observations of all the terminals, at the (function seeking) terminals in $\mathcal{A}$ using communication that is independent of G_0. In fact, we observe that this is a special case of the following more general

principle: a terminal that computes the private function g_0, may recover the entire data without affecting the conditions for secure computability.

Unlike [19], we do not provide a single-letter formula for the quantity R^*, in general; nevertheless, conditions (1) and (2) provide a structural characterization of securely computable functions in a broader setting, generalizing the results in [19]. A general recipe for single-letter characterization is presented which, in Example 1 and Corollary 4 below, yields single-letter results that are new and cannot be obtained from the analysis in [19]. To the best of our knowledge, the general analysis presented here is the only known method to prove the necessity of the single-letter conditions for secure computability in these special cases. Furthermore, for the cases with single-letter characterizations, the aforementioned heuristic interpretation of R^* is made precise (see the remark following Lemma 2 below).

The algorithms for exact function computation by multiple parties, without secrecy requirements, were first considered in [20], and have since been studied extensively (cf. e.g., [8], [9], [10]). An information-theoretic version with asymptotically accurate (in observation length) function computation was considered in [16], [11]. The first instance of the exact function computation problem with secrecy appears in [15]. A basic version of the secure computation problem studied here was introduced in [18], [19]; [3] gives an alternative proof of the results in [18], [19].

The problem of secure computing for multiple functions is formulated in the next section, followed by our results in section III. The proofs are given in sections IV and V. The final section discusses alternative forms of the necessary conditions.

Notation. The set $\{1, ..., m\}$ is denoted by $\mathcal{M}$. For $i < j$, denote by $[i, j]$ the set $\{i, ..., j\}$. Let $X_1, ..., X_m, m \geq 2$, be rvs taking values in finite sets $\mathcal{X}_1, ..., \mathcal{X}_m$, respectively, and with a known probability mass function. Denote by $X_{\mathcal{M}}$ the collection of rvs $(X_1, ..., X_m)$, and by $X_{\mathcal{M}}^n = (X_{\mathcal{M},1}, ..., X_{\mathcal{M},n})$ the n independent and identically distributed (i.i.d.) repetitions of the rv $X_{\mathcal{M}}$. For a subset $\mathcal{A}$ of $\mathcal{M}$, denote by $X_{\mathcal{A}}$ the rvs $(X_i, i \in \mathcal{A})$. Given $R_i \geq 0$, $1 \leq i \leq m$, let $R_{\mathcal{A}}$ denote the sum $\sum_{i \in \mathcal{A}} R_i$. Denote the cardinality of the range-space of an rv U by $\|U\|$.

Finally, for $0 < \epsilon < 1$, an rv U is ϵ-recoverable from an rv V if there exists a function g of V such that $\Pr(U = g(V)) \geq 1 - \epsilon$.

II. PROBLEM FORMULATION

We consider a multiterminal source model for function computation using public communication, with a confidentiality requirement. This basic model was introduced in [6] in a separate context of SK generation with public transaction. Terminals $1, \ldots, m$ observe, respectively, the sequences $X_1^n, \ldots, X_m^n$ of length n. For $0 \leq i \leq m$, let $g_i : \mathcal{X}_{\mathcal{M}} \to \mathcal{Y}_i$ be given mappings, where the sets $\mathcal{Y}_i$ are finite. Further, for $0 \leq i \leq m$ and $n \geq 1$, the (single-letter) mapping

$g_i^n : \mathcal{X}_{\mathcal{M}}^n \to \mathcal{Y}_i^n$ is defined by

$$g_i^n(x_{\mathcal{M}}^n) = (g_i(x_{11}, \ldots, x_{m1}), \ldots, g_i(x_{1n}, \ldots, x_{mn})),$$
$$x_{\mathcal{M}}^n = (x_1^n, \ldots, x_m^n) \in \mathcal{X}_{\mathcal{M}}^n.$$

For convenience, we shall denote the rv $g_i^n(X_{\mathcal{M}}^n)$ by $G_i^n, n \geq 1$, and, in particular, $G_i^1 = g_i(X_{\mathcal{M}})$ simply by G_i.

Each terminal $i \in \mathcal{M}$ wishes to compute the function $g_i^n(x_{\mathcal{M}}^n)$, without revealing $g_0^n(x_{\mathcal{M}}^n)$, $x_{\mathcal{M}}^n \in \mathcal{X}_{\mathcal{M}}^n$. To this end, the terminals are allowed to communicate over a noiseless public channel, possibly interactively in several rounds.

Definition 1. An *r-rounds interactive communication protocol* consists of mappings

$$f_{11}, ..., f_{1m}, ..., f_{r1}, ..., f_{rm},$$

where f_{ij} denotes the communication sent by the jth node in the ith round of the protocol; specifically, f_{ij} is a function of X_j^n and the communication sent in the previous rounds $\{f_{kl} : 1 \leq k \leq i - 1, l \in \mathcal{M}\}$. Denote the rv corresponding to the communication by

$$\mathbf{F} = F_{11}, ..., F_{1m}, ..., F_{r1}, ..., F_{rm},$$

noting that $\mathbf{F} = \mathbf{F}^{(n)}(X_{\mathcal{M}}^n)$. The rate[1] of $\mathbf{F}$ is $\frac{1}{n} \log \|\mathbf{F}\|$.

Definition 2. For $\epsilon_n > 0$, $n \geq 1$, we say that functions[2] $g_{\mathcal{M}} = (g_0, g_1, ..., g_m)$, with private function g_0, are ϵ_n-*securely computable* (ϵ_n- SC) from observations of length n, and public communication $\mathbf{F} = \mathbf{F}^{(n)}$, if

(i) G_i^n is ϵ_n- recoverable from $(X_i^n, \mathbf{F})$ for every $i \in \mathcal{M}$, and

(ii) $\mathbf{F}$ satisfies the secrecy condition

$$\frac{1}{n} I(G_0^n \wedge \mathbf{F}) \leq \epsilon_n.$$

Remark. The definition of secrecy here corresponds to "weak secrecy" [1], [13]. When our results have a single-letter form, our achievability schemes for secure computing attain "strong secrecy" in the sense of [14], [4], [6]. In fact, when we have a single-letter form, our proof in section IV yields "strong secrecy" upon minor modification.

By definition, for ϵ_n-SC functions $g_{\mathcal{M}}$, the private function G_0 is effectively concealed from an eavesdropper with access to the public communication $\mathbf{F}$.

Definition 3. For private function g_0, we say that functions $g_{\mathcal{M}}$ are *securely computable* if $g_{\mathcal{M}}$ are ϵ_n- SC from observations of length n and public communication $\mathbf{F} = \mathbf{F}^{(n)}$, such that $\lim_n \epsilon_n = 0$. Figure 1 shows the setup for secure computing.

In this paper, we give necessary and sufficient conditions for the secure computability of certain classes of functions $g_{\mathcal{M}} = (g_0, g_1, ..., g_m)$. The formulation in [19], in which the terminals in a given subset $\mathcal{A}$ of $\mathcal{M}$ are required to compute

---

[1]All logarithms are with respect to the base 2.
[2]The abuse of notation $g_{\mathcal{M}} = (g_0, g_1, ..., g_m)$ simplifies our presentation.
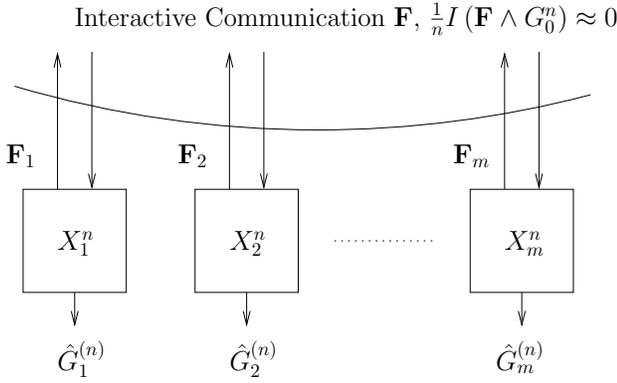
Fig. 1. Secure computation of $g_1, ..., g_m$ with private function g_0

(only) g_0 securely, is a special case with

$$g_i = \begin{cases} g_0, & i \in \mathcal{A}, \\ \text{constant}, & \text{otherwise}. \end{cases} \quad (3)$$

It was shown in [19] that (1) and (2) constitute, respectively, necessary and sufficient conditions for the functions above to be securely computable, with R^* being the minimum rate of interactive communication $\mathbf{F}$ that enables all the terminals in $\mathcal{M}$ to attain *omniscience* (see [6]), i.e., recover *all* the data $X_{\mathcal{M}}^n$, using $\mathbf{F}$ and the *decoder side information* G_0^n given to the terminals in $\mathcal{M} \setminus \mathcal{A}$. In fact, it was shown that when condition (2) holds, it is possible to recover $X_{\mathcal{M}}^n$ using communication that is independent of G_0^n.

The guiding heuristic in this work is the following general principle, which is also consistent with the results of [19]:

Conditions (1) and (2) constitute, respectively, the necessary and sufficient conditions for functions $g_{\mathcal{M}} = (g_0, g_1, ..., g_m)$ to be securely computable, where R^* is the infimum of the rates of interactive communication $\mathbf{F}'$ such that, for each $1 \leq i \leq m$, the following hold simultaneously:

(P1) G_i^n *is ϵ_n-recoverable from $(X_i^n, \mathbf{F}')$, and*
(P2) $X_{\mathcal{M}}^n$ *is ϵ_n-recoverable from $(X_i^n, G_0^n, \mathbf{F}')$, i.e., terminals attain omniscience, with G_0^n as side information that is used only for decoding (but is not used for the communication $\mathbf{F}'$),*

where $\epsilon_n \to 0$ as $n \to \infty$.

Thus, (P1) and (P2) require any terminal computing g_0 to become omniscient, an observation that was also made for the special case in [19]. The first condition (P1) above is straightforward and ensures the computability of the functions $g_1, ..., g_m$, by the terminals $1, ..., m$, respectively. The omniscience condition (P2) facilitates the decomposition of total entropy into mutually independent components that include the random values of the private function G_0^n and the communication $\mathbf{F}'$. For the specific case in (3), R^* above has a single-letter formula. In general, a single-letter expression for R^* is not known.

Our results, described in section III, are obtained by simple adaptations of this principle. Unlike [19], our conditions, in general, are not of a single-letter form. Nevertheless, they provide a structural characterization of secure computability.

As an application, our results provide simple conditions for secure computability in the following illustrative example.

Example 1. We consider the case of $m = 2$ terminals that observe binary symmetric sources (BSS) with underlying rvs X_1, X_2 with joint pmf given by

$$\Pr(X_1 = 0, X_2 = 0) = \Pr(X_1 = 1, X_2 = 1) = \frac{1-\delta}{2},$$

$$\Pr(X_1 = 0, X_2 = 1) = \Pr(X_1 = 1, X_2 = 0) = \frac{\delta}{2},$$

where $0 < \delta < 1/2$. The results of this paper will allow us to provide conditions for the secure computability of the four choices of g_0, g_1, g_2 below; it will follow by Theorem 1 that functions g_0, g_1, g_2 are securely computable if

$$h(\delta) < \tau,$$

and conversely, if the functions above are securely computable, then

$$h(\delta) \leq \tau,$$

where $h(\tau) = -\tau \log \tau - (1-\tau) \log(1-\tau)$, and the constant $\tau = \tau(\delta)$ depends on the choice of the function. These characterizations are summarized in the next table. Denote the AND and the OR of two random bits X_1 and X_2 by $X_1.X_2$ and $X_1 \oplus X_2$, respectively.

g_0	g_1	g_2	τ
$X_1 \oplus X_2$	$X_1 \oplus X_2$	$X_1 \oplus X_2$	$1/2$
$X_1 \oplus X_2$	$X_1 \oplus X_2$	ϕ	1
$X_1 \oplus X_2, \ X_1.X_2$	$X_1 \oplus X_2, \ X_1.X_2$	$X_1.X_2$	$2\delta/3$
$X_1 \oplus X_2$	$X_1 \oplus X_2$	$X_1.X_2$	$2/3$

The results for the first two settings follow from [19]. The third and fourth results are new. In these settings, terminal 1 is required to recover the private function; our results below show that the conditions for the secure computability in these cases remain unchanged even if this terminal is required to attain omniscience. Note that since $h(\delta) < 1$ for all $0 < \delta < 1/2$, there exists a communication protocol for securely computing the functions in the second setting. By contrast, a secure computing protocol for the functions in the third setting does not exist for any $0 < \delta < 1/2$, since $h(\delta) > 2\delta/3$. □

III. CHARACTERIZATION OF SECURELY COMPUTABLE FUNCTIONS

In this section, we characterize securely computable functions for three settings. Our necessary and sufficient conditions entail the comparison of $H(X_{\mathcal{M}}|G_0)$ with a rate R^*; the specific choice of R^* depends on the functions $g_{\mathcal{M}}$. Below we consider three different classes of functions $g_{\mathcal{M}}$. Although the first class is a special case of the second, the two are handled separately as the more restrictive case is amenable to simpler analysis. Furthermore, for $m = 2$, the obtained necessary and sufficient conditions for secure computability take a single-letter form in the first case (see Corollary 4).

(1) In the first class we consider, values of all the functions

$g_1, ..., g_m$ must be kept secret. In addition, at least one of the terminals must compute all the functions $g_1, ..., g_m$. This case arises in distributed function computation over a network where all the computed values are collated at a single sink node, and we are interested in securing the collated function values. Alternatively, denoting the function computed at the sink node by the private function g_0, the computed functions $g_1, ..., g_m$ can be restricted to be functions of g_0. Specifically, for $0 < m_0 < m$, and for private function g_0, let

$$g_i = \begin{cases} g_0, & i \in [1, m_0], \\ g_i(g_0), & i \in [m_0 + 1, m]. \end{cases} \quad (4)$$

(2) The next case is a relaxation of the previous model in that the restriction $g_i = g_i(g_0)$ for $i \in [m_0 + 1, m]$ is dropped. For this general case, our analysis below implies roughly that requiring the terminals $[1, m_0]$ that compute the private function g_0 to recover the entire data $X_\mathcal{M}^n$ does not change the conditions for secure computability, which is a key observation of this paper.

(3) The last class of problems we study is an instance of *secure multiterminal source coding*, which arises in the data download problems in sensor networks where each node is interested in downloading the data observed by a subset of nodes. Specifically, we consider the situation where each terminal wishes to recover some subset $X_{\mathcal{M}_i}^n$ of the sources where $\mathcal{M}_i \subseteq \mathcal{M} \setminus \{i\}$, i.e.,

$$g_i(X_\mathcal{M}) = X_{\mathcal{M}_i}, \quad i \in \mathcal{M}. \quad (5)$$

This last case appears to be disconnected from the previous two cases a priori. However, our characterizations of secure computability below have the same form for all cases above. Moreover, the same heuristic principle, highlighted in (P1) and (P2), leads to a characterization of secure computability in all three cases.

The necessary and sufficient conditions for secure computability are stated in terms of quantities $R_i^*(g_\mathcal{M})$, $i = 1, 2, 3$, which are defined next. The subscript i corresponds to case (i) above. In particular, the quantity R^* corresponds to the minimum rate of communication needed for an appropriate modification of the source-coding task in (P1), (P2). Below we give specific expressions for R_i^*, $i = 1, 2, 3$, along with their operational roles (for a complete description of this role see the sufficiency proof in Section IV).

Denote by $\mathcal{R}_1^*(g_\mathcal{M})$ the closure of the (nonempty) set of pairs[3]

$$\left(R_\mathbf{F}^{(1)}, \frac{1}{n} I(G_0^n \wedge \mathbf{F}) \right),$$

for all $n \geq 1$ and interactive communication $\mathbf{F}$, where

$$R_\mathbf{F}^{(1)} = \frac{1}{n} H(\mathbf{F}) + \frac{1}{n} \sum_{i=m_0+1}^{m} H(G_i^n | X_i^n, \mathbf{F}) + \inf R_\mathcal{M}, \quad (6)$$

with the infimum taken over the rates $R_1, ..., R_m$ satisfying the following constraints:

---

[3]The first term accounts for the rate of the communication and the second term tracks the information about G_0^n leaked by $\mathbf{F}$ (see (11)) below

(1a) $\forall \mathcal{L} \subsetneq \mathcal{M}, [1, m_0] \nsubseteq \mathcal{L}$,

$$R_\mathcal{L} \geq \frac{1}{n} H\left(X_\mathcal{L}^n | X_{\mathcal{M} \setminus \mathcal{L}}^n, \mathbf{F} \right);$$

(1b) $\forall \mathcal{L} \subsetneq \mathcal{M}, [1, m_0] \subseteq \mathcal{L}$,

$$R_\mathcal{L} \geq \frac{1}{n} H\left(X_\mathcal{L}^n | X_{\mathcal{M} \setminus \mathcal{L}}^n, G_0^n, \mathbf{F} \right).$$

The quantity $\inf_{n, \mathbf{F}} R_\mathbf{F}^{(1)}$ corresponds to the solution of a multiterminal source coding problem. Specifically, it is the infimum of the rates of interactive communication that satisfy (P1) and (P2) above (see [5, Theorem 13.15], [6]).

Next, let $\mathcal{R}_2^*(g_\mathcal{M})$ denote the closure of the set of pairs

$$\left(R_\mathbf{F}^{(2)}, \frac{1}{n} I(G_0^n \wedge \mathbf{F}) \right),$$

for all $n \geq 1$ and interactive communication $\mathbf{F}$, where

$$R_\mathbf{F}^{(2)} = \frac{1}{n} H(\mathbf{F}) + \inf \left[R'_{[m_0+1, m]} + R_\mathcal{M} \right], \quad (7)$$

with the infimum taken over the rates $R_1, ..., R_m$ and $R'_{m_0+1}, ..., R'_m$ satisfying the following constraints:

(2a) $\forall \mathcal{L} \subsetneq \mathcal{M}, [1, m_0] \nsubseteq \mathcal{L}$,

$$R_\mathcal{L} \geq \frac{1}{n} H\left(X_\mathcal{L}^n | X_{\mathcal{M} \setminus \mathcal{L}}^n, \mathbf{F} \right);$$

(2b) for $m_0 < j \leq m$,

$$R'_j \geq \frac{1}{n} H\left(G_j^n | X_j^n, \mathbf{F} \right);$$

(2c) $\forall \mathcal{L} \subseteq \mathcal{M}, [1, m_0] \subseteq \mathcal{L}$, and $\mathcal{L}' \subseteq [m_0 + 1, m]$ with either $\mathcal{L} \neq \mathcal{M}$ or $\mathcal{L}' \neq [m_0 + 1, m]$,

$$R'_{\mathcal{L}'} + R_\mathcal{L} \geq \frac{1}{n} H\left(G_{\mathcal{L}'}^n, X_\mathcal{L}^n | G_{[m_0+1, m] \setminus \mathcal{L}'}^n, X_{\mathcal{M} \setminus \mathcal{L}}^n, G_0^n, \mathbf{F} \right).$$

The quantity $\inf_{n, \mathbf{F}} R_\mathbf{F}^{(2)}$ corresponds to the solution of a multiterminal source coding problem, and is the infimum of the rates of interactive communication $\mathbf{F}'$ that satisfy (P1) and (P2) above, and additionally satisfies:

(P3) $X_\mathcal{M}^n$ *is ϵ_n-recoverable from* $\left(G_j^n, G_0^n, \mathbf{F}' \right)$, $m_0 < j \leq m$.

This modification corresponds to the introduction of $m - m_0$ dummy terminals, with the jth dummy terminal observing G_j^n, $m_0 < j \leq m$ (see section VI); the dummy terminals can be realized by a terminal i in $[1, ..., m_0]$ that recovers $X_\mathcal{M}^n$ from $(X_i^n, \mathbf{F})$. The conditions (P2) and (P3) above correspond to the omniscience at the terminals in the extended model, with G_0^n provided as side information only for decoding.

Finally, denote by $\mathcal{R}_3^*(g_\mathcal{M})$ the closure of the set of pairs

$$\left(R_\mathbf{F}^{(3)}, \frac{1}{n} I(G_0^n \wedge \mathbf{F}) \right),$$

for all interactive communication $\mathbf{F}$, where

$$R_\mathbf{F}^{(3)} = \frac{1}{n} H(\mathbf{F}) + \inf R_\mathcal{M}, \quad (8)$$

with rates $R_1, ..., R_m$ satisfying the following constraints:

(3a) For $1 \leq i \leq m, \forall \mathcal{L} \subseteq \mathcal{M}_i \subseteq \mathcal{M} \setminus \{i\}$,

$$R_\mathcal{L} \geq \frac{1}{n} H\left(X_\mathcal{L}^n | X_{\mathcal{M}_i \setminus \mathcal{L}}^n, X_i^n, \mathbf{F} \right);$$

(3b) $\forall \mathcal{L} \subsetneq \mathcal{M}$,

$$R_{\mathcal{L}} \geq \frac{1}{n} H\left(X_{\mathcal{L}}^n | X_{\mathcal{M} \setminus \mathcal{L}}^n, G_0^n, \mathbf{F}\right).$$

As before, the quantity $\inf_{n, \mathbf{F}} R_{\mathbf{F}}^{(3)}$ corresponds to the infimum of the rates of interactive communication that satisfy (P1) and (P2) above.

Our main result below characterizes securely computable functions for the three settings above.

Theorem 1. *For $i = 1, 2, 3$, with functions $g_0, g_1, ..., g_m$ as in the case (i) above, the functions $g_{\mathcal{M}}$ are securely computable if the following condition holds:*

$$H\left(X_{\mathcal{M}} | G_0\right) > R_i^*\left(g_{\mathcal{M}}\right). \tag{9}$$

Conversely, if the functions above are securely computable, then

$$H\left(X_{\mathcal{M}} | G_0\right) \geq R_i^*\left(g_{\mathcal{M}}\right), \tag{10}$$

where

$$R_i^*\left(g_{\mathcal{M}}\right) = \inf_{(x, 0) \in \mathcal{R}_i^*(g_{\mathcal{M}})} x, \quad i = 1, 2, 3. \tag{11}$$

Remark. Although the first setting above is a special case of the second, it is unclear if for $g_{\mathcal{M}}$ in (4) the quantities $R_1^*(g_{\mathcal{M}})$ and $R_2^*(g_{\mathcal{M}})$ are identical (also, see Section VI). In general, the multi-letter characterizations of secure computability of $g_{\mathcal{M}}$ above can have different forms. For case (1) with $m = 2$, Corollary 4 below provides a single-letter formula for $R_1^*(g_{\mathcal{M}})$. However, a similar single-letter formula for $R_2^*(g_{\mathcal{M}})$ is not known.

Theorem 1 affords the following heuristic interpretation. The quantity $H\left(X_{\mathcal{M}} | G_0\right)$ represents the maximum rate of randomness in $X_{\mathcal{M}}^n$ that is (nearly) independent of G_0^n. On the other hand, $R_i^*\left(g_{\mathcal{M}}\right)$ is an appropriate rate of communication for the computation of $g_{\mathcal{M}}$; we show that latter being less than $H\left(X_{\mathcal{M}} | G_0\right)$ guarantees the secure computability of $g_{\mathcal{M}}$.

Although the characterization in Theorem 1 is not of a single-letter form, the following result provides a sufficient condition for obtaining such forms. Denote by $R_{\text{constant}}^{(i)}$, $i = 1, 2, 3$, the quantity $R_{\mathbf{F}}^{(i)}$ for $\mathbf{F} = \text{constant}$.

Lemma 2. *For case (i), $i = 1, 2, 3$, if for all $n \geq 1$ and interactive communication $\mathbf{F}$*

$$R_{\mathbf{F}}^{(i)} \geq R_{\text{constant}}^{(i)}, \tag{12}$$

then $R_i^*\left(g_{\mathcal{M}}\right) = R_{\text{constant}}^{(i)} = \inf_{n, \mathbf{F}} R_{\mathbf{F}}^{(i)}$.

The proof is a simple consequence of the definition of $R_i^*\left(g_{\mathcal{M}}\right)$ in (11). Note that $R_{\text{constant}}^{(i)}$ has a single-letter form.

Remark. As mentioned before, the quantity $\inf_{n, \mathbf{F}} R_{\mathbf{F}}^{(i)}$ is the infimum of the rates of interactive communication that satisfies (P1), (P2) for $i = 1, 3$, and satisfies (P1)-(P3) for $i = 2$. Thus, when the conditions of Lemma 2 hold, we have from Theorem 1 that $g_{\mathcal{M}}$ are securely computable if

$$H\left(X_{\mathcal{M}} | G_0\right) > R_{\text{constant}}^{(i)},$$

and if $g_{\mathcal{M}}$ are securely computable then

$$H\left(X_{\mathcal{M}} | G_0\right) \geq R_{\text{constant}}^{(i)},$$

where $R_{\text{constant}}^{(i)}$ is the minimum rate of communication that satisfies (P1), (P2) for $i = 1, 3$, and satisfies (P1)-(P3) for $i = 2$.

As a consequence of Lemma 2, we obtain below a single-letter characterization of securely computable functions, with $m = 2$, in a special case; the following lemma, which is a special case of [7, Lemma B.1] (see also [12, Theorem 1]), is instrumental to our proof.

Lemma 3. *Let $m = 2$. For an interactive communication $\mathbf{F}$, we have*

$$H(\mathbf{F}) \geq H\left(\mathbf{F} | X_1^n\right) + H\left(\mathbf{F} | X_2^n\right).$$

We next consider case (1) for two terminals.

Corollary 4. *For $m = 2$, for functions g_0, g_1, g_2 with $g_1 = g_0$ and $g_2 = g_2\left(g_0\right)$, we have*

$$R_1^*\left(g_{\mathcal{M}}\right) = H\left(X_2 | X_1\right) + H\left(G_2 | X_2\right) + H\left(X_1 | X_2, G_0\right). \tag{13}$$

Proof: The constraints (1a) and (1b) satisfied by rates R_1, R_2 in the definition of $R_{\mathbf{F}}^{(1)}$ are

$$R_2 \geq \frac{1}{n} H\left(X_2^n | X_1^n, \mathbf{F}\right),$$

$$R_1 \geq \frac{1}{n} H\left(X_1^n | X_2^n, G_0^n, \mathbf{F}\right),$$

which further yields

$$R_{\mathbf{F}}^{(1)} = \frac{1}{n} \left[H\left(\mathbf{F}\right) + H\left(G_2^n | X_2^n, \mathbf{F}\right) \right.$$
$$\left. + H\left(X_2^n | X_1^n, \mathbf{F}\right) + H\left(X_1^n | X_2^n, G_0^n, \mathbf{F}\right) \right]. \tag{14}$$

Thus, $R_{\text{constant}}^{(1)}$ equals the term on the right side of (13). Upon manipulating the expression for $R_{\mathbf{F}}^{(1)}$ above, we get

$$R_{\mathbf{F}}^{(1)} = \frac{1}{n} \left[H(\mathbf{F}) - H\left(\mathbf{F} | X_1^n\right) - H\left(\mathbf{F} | X_2^n, G_0^n\right) \right.$$
$$\left. - I\left(G_2^n \wedge \mathbf{F} | X_2^n\right) \right] + R_{\text{constant}}^{(1)}. \tag{15}$$

Further, since $H\left(G_2 | G_0\right) = 0$, it holds that

$$I\left(G_2^n \wedge \mathbf{F} | X_2^n\right) \leq I\left(G_0^n \wedge \mathbf{F} | X_2^n\right),$$

which along with (15) yields

$$R_{\mathbf{F}}^{(1)} \geq \frac{1}{n} \left[H(\mathbf{F}) - H\left(\mathbf{F} | X_1^n\right) - H\left(\mathbf{F} | X_2^n\right) \right] + R_{\text{constant}}^{(1)}$$

$$\geq R_{\text{constant}}^{(1)},$$

where the last inequality follows from Lemma 3. The result then follows from Lemma 2. $\qquad \square$

We next derive simple conditions for secure computability for the BSS in Example 1

Example 2. Consider the setup of Example 1, with $g_0 = g_1 = X_1 \oplus X_2, X_1.X_2$ and $g_2 = X_1.X_2$. By Corollary 4 and the observation $H\left(G_2 | X_2\right) = h(\delta)/2$, we get $R_1^*\left(g_{\mathcal{M}}\right) = 3h(\delta)/2$. Since $H\left(X_1, X_2 \mid G_0\right) = H\left(X_1, X_2 \mid X_1 \oplus X_2\right) - H\left(X_1.X_2 \mid X_1 \oplus X_2\right) = \delta$, the characterization of secure

computability claimed in Example 1 follows from Theorem 1. $\qquad\square$

Example 3. In the setup of Example 1, consider $g_0 = g_1 = X_1 \oplus X_2$ and $g_2 = X_1.X_2$. This choice of g_0, g_1, g_2 is an instance of case (2) above. For an interactive communication $\mathbf{F}$, the constraints (2a), (2b), (2c) in the definition of $R_{\mathbf{F}}^{(2)}$, upon simplification, reduce to

$$R_1 \geq \frac{1}{n} H\left(X_1^n | X_2^n, G_0^n, G_2^n, \mathbf{F}\right),$$

$$R_2 \geq \frac{1}{n} H\left(X_2^n | X_1^n, \mathbf{F}\right),$$

$$R_1 + R_2 \geq \frac{1}{n} H\left(X_1^n, X_2^n | G_0^n, G_2^n, \mathbf{F}\right),$$

$$R_2' \geq \frac{1}{n} H\left(G_2^n | X_2^n, \mathbf{F}\right).$$

Therefore, $\inf\left[R_1 + R_2 + R_2'\right]$ with R_1, R_2, R_2' satisfying (2a), (2b), (2c), is given by

$$\frac{1}{n}\left[H\left(X_1^n | X_2^n, G_0^n, G_2^n, \mathbf{F}\right)\right.$$
$$+ \max\left\{H\left(X_2^n | G_0^n, G_2^n, \mathbf{F}\right), H\left(X_2^n | X_1^n, \mathbf{F}\right)\right\}$$
$$\left.+ H\left(G_2^n | X_2^n, \mathbf{F}\right)\right],$$

which further gives

$$R_{\mathbf{F}}^{(2)} = \frac{1}{n}\left[H(\mathbf{F}) + H\left(X_1^n | X_2^n, G_0^n, G_2^n, \mathbf{F}\right)\right.$$
$$+ \max\left\{H\left(X_2^n | G_0^n, G_2^n, \mathbf{F}\right), H\left(X_2^n | X_1^n, \mathbf{F}\right)\right\}$$
$$\left.+ H\left(G_2^n | X_2^n, \mathbf{F}\right)\right]. \tag{16}$$

It follows from $H\left(X_1^n | X_2^n, G_0^n, G_2^n, \mathbf{F}\right) = 0$ that

$$R_{constant}^{(2)} = H\left(G_2 | X_2\right)$$
$$+ \max\left\{H\left(X_2 | G_0, G_2\right), H\left(X_2 | X_1\right)\right\}$$
$$= \frac{h(\delta)}{2} + \max\left\{\delta, h(\delta)\right\} = \frac{3}{2}h(\delta), \tag{17}$$

as $h(\delta) > \delta$ for $0 < \delta < 1/2$.

Next, note from (16) that for any interactive communication $\mathbf{F}$

$$R_{\mathbf{F}}^{(2)} \geq \frac{1}{n}\left[H(\mathbf{F}) + H\left(X_2^n | X_1^n, \mathbf{F}\right) + H\left(G_2^n | X_2^n, \mathbf{F}\right)\right]$$
$$= \frac{1}{n}\left[H(\mathbf{F}) + H\left(X_2^n | X_1^n\right)\right.$$
$$- H\left(\mathbf{F} | X_1^n\right) + H\left(G_2^n, \mathbf{F} | X_2^n\right) - H\left(\mathbf{F} | X_2^n\right)\right]$$
$$\geq \frac{1}{n}\left[H(\mathbf{F}) - H\left(\mathbf{F} | X_1^n\right) - H\left(\mathbf{F} | X_2^n\right)\right]$$
$$+ H\left(G_2 | X_2\right) + H\left(X_2 | X_1\right)$$
$$\geq H\left(G_2 | X_2\right) + H\left(X_2 | X_1\right) = \frac{3}{2}h(\delta), \tag{18}$$

where the last inequality above follows from Lemma 3. The characterization in Example 1 follows from (17), (18), and $H\left(X_1, X_2 | G_0\right) = 1$, using Lemma 2 and Theorem 1. $\qquad\square$

IV. PROOF OF SUFFICIENCY IN THEOREM 1

Sufficiency of (9) for $i = 1$: We propose a two step protocol

for securely computing $g_0, g_1, ..., g_m$. In the first step, for sufficient large N, the terminals $[1, m_0]$ (g_0-seeking terminals) attain omniscience, using an interactive communication $\mathbf{F}'' = \mathbf{F}''\left(X_{\mathcal{M}}^N\right)$ that satisfies

$$\frac{1}{N} I\left(G_0^N \wedge \mathbf{F}''\right) \leq \epsilon, \tag{19}$$

where $\epsilon > 0$ is sufficiently small. Next, upon attaining omniscience, one of the terminals in $[1, m_0]$ computes the following for $m_0 < j \leq m$:

(i) Slepian-Wolf codewords $\hat{F}_j = \hat{F}_j\left(G_j^N\right)$ of appropriate rates R_j' for a recovery of G_j^N by a decoder with the knowledge of X_j^N and previous communication $\mathbf{F}''$, and

(ii) the rvs $K_j = K_j\left(X_j^N\right)$ of rates R_j' that satisfy:

$$\left|\frac{1}{N} H\left(K_j\right) - R_j'\right| \leq \epsilon, \tag{20}$$

$$\frac{1}{N} I\left(K_j \wedge G_0^N, \mathbf{F}'', \left\{K_l \oplus \hat{F}_l\right\}_{m_0 < l \leq j-1}\right) \leq \epsilon. \tag{21}$$

Note that $K_j \oplus \hat{F}_j$ denotes the encrypted version of the Slepian-Wolf code $\hat{F}_j$, encrypted with a one-time pad using the secret key (SK) K_j. Thus, terminal j, with the knowledge of K_j, can recover $\hat{F}_j$ from $K_j \oplus \hat{F}_j$, and hence can recover G_j^N. The operation $K_j \oplus \hat{F}_j$ is valid since the SK K_j has size greater than $\|\hat{F}_j\|$. Furthermore, we have from (19) and (21) that

$$\frac{1}{N} I\left(G_0^N \wedge \mathbf{F}'', \left\{K_j \oplus \hat{F}_j\right\}_{m_0 < j \leq m}\right)$$
$$\leq \frac{1}{N} I\left(G_0^N \wedge \left\{K_j \oplus \hat{F}_j\right\}_{m_0 < j \leq m} \mid \mathbf{F}''\right) + \epsilon$$
$$\leq \sum_{j=m_0+1}^{m} \frac{1}{N}\left[\log\|K_j \oplus \hat{F}_j\|\right.$$
$$\left.- H\left(K_j \oplus \hat{F}_j \mid \mathbf{F}'', \left\{K_i \oplus \hat{F}_i\right\}_{m_0 < i \leq j-1}, G_0^N\right)\right] + \epsilon$$
$$\leq \sum_{j=m_0+1}^{m} \frac{1}{N}\left[H\left(K_j\right)\right.$$
$$\left.- H\left(K_j \oplus \hat{F}_j \mid \mathbf{F}'', \left\{K_i \oplus \hat{F}_i\right\}_{m_0 < i \leq j-1}, G_0^N\right)\right] + 2\epsilon$$
$$= \sum_{j=m_0+1}^{m} \frac{1}{N}\left[H\left(K_j\right)\right.$$
$$\left.- H\left(K_j \mid \mathbf{F}'', \left\{K_i \oplus \hat{F}_i\right\}_{m_0 < i \leq j-1}, G_0^N\right)\right] + 2\epsilon \tag{22}$$
$$\leq 3m\epsilon,$$

where the third inequality above uses (20) and the last inequality follows from (21). The equality in (22) follows from the fact that $\hat{F}_j = \hat{F}_j\left(G_j^N\right)$ is a function of G_0^N, since G_j is a function of G_0. We note that this is the only place in the proof where the functional relation between G_j and G_0 is used.

Thus, the communication $\left(\mathbf{F}'', K_j \oplus \hat{F}_j, m_0 < j \leq m\right)$ constitutes the required secure computing protocol for $g_{\mathcal{M}}$. It remains to show the existence of $\mathbf{F}''$ and $K_j, m_0 < j \leq m$ that satisfy (19)-(21).

Specifically, when (9) holds for $i = 1$, we have from

the definition of $R_1^*(g_{\mathcal{M}})$ in (11) that for all $0 < \epsilon \leq \epsilon_0$ (ϵ_0 to be specified later), there exists $n \geq 1$ and interactive communication $\mathbf{F} = \mathbf{F}(X_{\mathcal{M}}^n)$ such that

$$\frac{1}{n} I(G_0^n \wedge \mathbf{F}) < \epsilon, \qquad (23)$$

and

$$R_{\mathbf{F}}^{(1)} \leq R_1^*(g_{\mathcal{M}}) + \frac{\epsilon}{2},$$

where $R_{\mathbf{F}}^{(1)}$ is as in (6). This further implies that there exist $R_1, ..., R_m$ satisfying (1a) and (1b) (for $\mathbf{F}$) such that

$$\frac{1}{n} H(\mathbf{F}) + \frac{1}{n} \sum_{i=m_0+1}^{m} H(G_j^n \mid X_j^n, \mathbf{F}) + R_{\mathcal{M}} \leq R_1^*(g_{\mathcal{M}}) + \epsilon. \qquad (24)$$

Choosing

$$\epsilon_0 < H(X_{\mathcal{M}} \mid G_0) - R_1^*(g_{\mathcal{M}}) - \delta,$$

for some $\delta < H(X_{\mathcal{M}} \mid G_0) - R_1^*(g_{\mathcal{M}})$, we get from (23) and (24) upon simplification:

$$\frac{1}{n} \sum_{i=m_0+1}^{m} H(G_j^n \mid X_j^n, \mathbf{F}) + R_{\mathcal{M}} + \delta < \frac{1}{n} H(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}). \qquad (25)$$

Next, for $k \geq 1$, denote by $\mathbf{F}^k = (\mathbf{F}_1, ..., \mathbf{F}_k)$ the i.i.d. rvs $\mathbf{F}_i = \mathbf{F}(X_{\mathcal{M},n(i-1)+1}, ..., X_{\mathcal{M},ni})$, $1 \leq i \leq k$. Further, let $N = nk$. In Appendix A, we follow the approach in the proof of [19, Theorem 5] and use (25) to show that for sufficiently large k there exists an interactive communication $\mathbf{F}' = \mathbf{f}'(X_{\mathcal{M}}^{nk})$ of overall rate $R_{\mathcal{M}} + \delta/2$ that satisfies the following:

$$X_{\mathcal{M}}^{nk} \text{ is } \epsilon\text{-recoverable from } (X_i^N, \mathbf{F}^k, \mathbf{F}') \text{ for } 1 \leq i \leq m_0,$$
$$\text{and from } (X_i^N, \mathbf{F}^k, G_0^N, \mathbf{F}') \text{ for } m_0 < i \leq m, \qquad (26)$$

and further,

$$\frac{1}{N} I(G_0^N, \mathbf{F}^k \wedge \mathbf{F}') < \epsilon. \qquad (27)$$

The proposed communication $\mathbf{F}''$ comprises $\mathbf{F}', \mathbf{F}^k$, and condition (19) follows from (23) and (27). Finally, we show the existence of $\hat{F}_j$ and K_j, $m_0 < j \leq m$, as above. From the Slepian-Wolf theorem [17], there exist rvs $\hat{F}_j = \hat{F}_j(G_j^N)$ of rates

$$R_j' \leq \frac{1}{N} H(G_j^N \mid X_j^N, \mathbf{F}^k) + \frac{\delta}{2m}, \qquad (28)$$

such that G_j^N is ϵ-recoverable from $(X_j^N, \mathbf{F}^k, \hat{F}_j)$, $m_0 < j \leq m$, for k sufficiently large. Suppose the rvs $K_{m_0+1}, K_{m_0+2}, ..., K_j$ of rates $R_{m_0+1}', R_{m_0+2}', ..., R_j'$, respectively, satisfy (20) and (21) for some $j \leq m - 1$. Denote by $\mathbf{F}'(j)$ the communication $(\mathbf{F}', K_i \oplus \hat{F}_i, m_0 < i \leq j)$ of rate $R^{(j)}$ that satisfies

$$R^{(j)} \leq R_{\mathcal{M}} + \frac{1}{N} \sum_{i=m_0+1}^{j} H(G_i^N \mid X_i^N, \mathbf{F}^k) + \delta \qquad (29)$$

We have from (25)-(29) that

$$R_{j+1}' < \frac{1}{N} H(X_{\mathcal{M}}^N \mid G_0^N, \mathbf{F}^k) - R^{(j)}. \qquad (30)$$

Heuristically, since $X_{\mathcal{M}}^N$ is recoverable from $(X_{j+1}^N, \mathbf{F}^k, \mathbf{F}')$, (30) gives

$$\frac{1}{N} H(X_{j+1}^N \mid G_0^N, \mathbf{F}^k, \mathbf{F}'(j))$$
$$\approx \frac{1}{N} H(X_{\mathcal{M}}^N \mid G_0^N, \mathbf{F}^k) - \frac{1}{N} H(\mathbf{F}'(j) \mid G_0^N, \mathbf{F}^k)$$
$$\geq \frac{1}{N} H(X_{\mathcal{M}}^N \mid G_0^N, \mathbf{F}^k) - R^{(j)}$$
$$> R_{j+1}'.$$

Thus, a randomly chosen mapping $K_{j+1} = K_{j+1}(X_{j+1}^N)$ of rate R_{j+1}' is almost jointly-independent of $G_0^N, \mathbf{F}^k, \mathbf{F}'(j)$ (see [4]). This argument is made rigorous using a version of the "balanced coloring lemma" (see [2], [6]) given in Appendix B. Specifically, in Lemma B1, set $U = X_{\mathcal{M}}^N$, $U' = X_{j+1}^N$, $V = G_0^N, \mathbf{F}^k$, $h = \mathbf{F}'(j)$, and

$$\mathcal{U}_0 = \Big\{ x_{\mathcal{M}}^N \in \mathcal{X}_{\mathcal{M}}^N :$$
$$x_{\mathcal{M}}^N = \psi_{j+1}(x_{j+1}^N, f'(x_{\mathcal{M}}^N), \mathbf{F}^k, g_0^n(x_{\mathcal{M}}^N)) \Big\},$$

for some mapping ψ_{j+1}, where $f'(X_{\mathcal{M}}^N) = \mathbf{F}'$ is as in (26). By the definition of $\mathbf{F}'$,

$$\Pr(U \in \mathcal{U}_0) \geq 1 - \epsilon,$$

so that condition (B1)(i) preceding Lemma B1 is met. Condition (B1)(ii), too, is met from the definition of $\mathcal{U}_0, h$ and V.

Upon choosing

$$d = \exp\left[k\left(H(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}) - \frac{n\delta}{2m}\right)\right],$$

in (B2), the hypotheses of Lemma B1 are satisfied for appropriately chosen λ, and for sufficiently large k. Then, by Lemma B1, with

$$r = \lceil \exp(NR_{j+1}') \rceil, \quad r' = \lceil \exp(NR^{(j)}) \rceil,$$

and with K_{j+1} in the role of ϕ, it follows from (B4) that there exists rv $K_{j+1} = K_{j+1}(X_{j+1}^N)$ that satisfies (20) and (21), for k sufficiently large. The proof is completed upon repeating this argument for $m_0 < j < m$. $\square$

Sufficiency of (9) for $i = 2$: The secure computing protocol for this case also consists of two stages. In the first stage, as before, the terminals $[1, m_0]$ (g_0-seeking terminals) attain omniscience, using an interactive communication $\mathbf{F}'' = \mathbf{F}''(X_{\mathcal{M}}^N)$. The second stage, too, is similar to the previous case and involves one of the omniscience-attaining terminals in $[1, m_0]$ transmitting communication $\hat{F}_j = \hat{F}_j(G_j^N)$ to the terminals j, for $m_0 < j \leq m$. However, the encryption-based scheme of the previous case is not applicable here; in particular, (22) no longer holds. Instead, the communication $\hat{F}_j$ now consists of the Slepian-Wolf codewords for G_j^N given X_j^N, and previous communication $\mathbf{F}''$. We show below that if (9) holds, then

there exist communication $\mathbf{F}''$ and $\hat{F}_j$, $m_0 < j \leq m$, of appropriate rate such that the following holds:

$$\frac{1}{N} I\left(G_0^N \wedge \mathbf{F}'', \hat{F}_{m_0+1}, ..., \hat{F}_m\right) < \epsilon,$$

for sufficiently large N.

Specifically, when (9) holds for $i = 2$, using similar manipulations as in the previous case we get that for all $0 < \epsilon < \epsilon_0$, there exist interactive communication $\mathbf{F} = \mathbf{F}\left(X_{\mathcal{M}}^n\right)$, and rates $R_1, ..., R_m, R'_{m_0+1}, ..., R'_m$ satisfying (2a)-(2c) (for $\mathbf{F}$) such that

$$\frac{1}{n} I\left(G_0^n \wedge \mathbf{F}\right) < \frac{\epsilon}{2},$$

and

$$R_{\mathcal{M}} + R'_{[m_0+1,m]} + \delta < \frac{1}{n} H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right), \qquad (31)$$

with $\delta < H\left(X_{\mathcal{M}} \mid G_0\right) - R_2^*\left(g_{\mathcal{M}}\right) - \epsilon_0$; (31) replaces (25) in the previous case.

Next, for $N = nk$ consider $2m - m_0$ correlated sources X_j^N, $1 \leq j \leq m$, and G_j^N, $m_0 < j \leq m$. Since $R_1, ..., R_m, R'_{m_0+1}, ..., R'_m$ satisfy (2a)-(2c), random mappings $F'_j = F'_j\left(X_j^N\right)$ of rates R_j, $1 \leq j \leq m$, and $F'_{j+m-m_0} = F'_{j+m-m_0}\left(G_j^N\right)$ of rates R'_j, $m_0 < j \leq m$ satisfy the following with high probability, for k sufficiently large (see [5, Lemma 13.13 and Theorem 13.14]):

(i) for $1 \leq i \leq m$, $X_{\mathcal{M}}^{nk}$ is ϵ-recoverable from $\left(F'_1, ..., F'_m, \mathbf{F}^k, X_i^{nk}\right)$;

(ii) for $m_0 < j \leq m$, G_j^{nk} is ϵ-recoverable from $\left(F'_{j+m-m_0}, \mathbf{F}^k, X_j^{nk}\right)$;

(iii) for $m_0 < j \leq m$, $X_{\mathcal{M}}^{nk}$ is ϵ-recoverable from $\left(\mathbf{F}', \mathbf{F}^k, X_j^{nk}, G_0^{nk}\right)$ and from $\left(\mathbf{F}', \mathbf{F}^k, G_j^{nk}, G_0^{nk}\right)$,

where $\mathbf{F}^k = \left(\mathbf{F}_1, ..., \mathbf{F}_k\right)$ are i.i.d. rvs $\mathbf{F}_i = \mathbf{F}\left(X_{\mathcal{M}, n(i-1)+1}, ..., X_{\mathcal{M}, ni}\right)$, $1 \leq i \leq k$. It follows from (31) in a manner similar to the proof in Appendix A that there exist communication F'_j, $1 \leq j \leq 2m - m_0$ as above such that

$$\frac{1}{nk} I\left(G_0^{nk} \wedge \mathbf{F}', \mathbf{F}^k\right) < \epsilon,$$

for sufficiently large k.

The first stage of the protocol entails transmission of $\mathbf{F}^k$, followed by the transmission of $F'_1, ..., F'_m$, i.e., $\mathbf{F}'' = \left(\mathbf{F}^k, F'_1, ..., F'_m\right)$. The second stage of communication $\hat{F}_j$ is given by F'_{j+m-m_0}, for $m_0 < j \leq m$. $\square$

Sufficiency of (9) for $i = 3$: Using the definition of $R_3^*\left(g_{\mathcal{M}}\right)$ and the manipulations above, the sufficiency condition (9) implies that for all $0 < \epsilon < \epsilon_0$, there exist interactive communication $\mathbf{F} = \mathbf{F}\left(X_{\mathcal{M}}^n\right)$, and rates $R_1, ..., R_m$ satisfying (3a), (3b) (for $\mathbf{F}$) such that

$$\frac{1}{n} I\left(G_0^n \wedge \mathbf{F}\right) < \frac{\epsilon}{2},$$

and

$$R_{\mathcal{M}} + \delta < \frac{1}{n} H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right), \qquad (32)$$

for $\delta < H\left(X_{\mathcal{M}} \mid G_0\right) - R_3^*\left(g_{\mathcal{M}}\right) - \epsilon_0$. Denoting by $\mathbf{F}^k =$

$\left(\mathbf{F}_1, ..., \mathbf{F}_k\right)$ the i.i.d. rvs $\mathbf{F}_i = \mathbf{F}\left(X_{\mathcal{M}, n(i-1)+1}^{ni}\right)$, $1 \leq i \leq k$, it follows from (3a) and (3b) that for $N = nk$ the random mappings $F'_i = F'_i\left(X_i^{nk}\right)$ of rates R_i, $1 \leq i \leq m$, satisfy the following with high probability, for k sufficiently large (see [5, Lemma 13.13 and Theorem 13.14]):

(i) for $i \in \mathcal{M}$, $X_{\mathcal{M}_i}^{nk}$ is ϵ-recoverable from $\left(\mathbf{F}', \mathbf{F}^k, X_i^{nk}\right)$;

(ii) for $i \in \mathcal{M}$, $X_{\mathcal{M}}^{nk}$ is ϵ-recoverable from $\left(\mathbf{F}', \mathbf{F}^k, X_i^{nk}, G_0^{nk}\right)$.

From (32), the approach of Appendix A implies that there exist F'_i, $i \in \mathcal{M}$, as above such that

$$\frac{1}{nk} I\left(G_0^{nk} \wedge \mathbf{F}', \mathbf{F}^k\right) < \epsilon,$$

for sufficiently large k. The interactive communication $\left(\mathbf{F}', \mathbf{F}^k\right)$ constitutes the protocol for securely computing $g_{\mathcal{M}}$, where $g_i\left(X_{\mathcal{M}}\right) = X_{\mathcal{M}_i}$, $i \in \mathcal{M}$. $\square$

V. Proof of Necessity in Theorem 1

Necessity of (10) for $i = 1$: If functions $g_{\mathcal{M}}$ are securely computable then there exists an interactive communication $\mathbf{F}$ such that G_i^n is ϵ_n-recoverable from $\left(X_i^n, \mathbf{F}\right)$, $i \in \mathcal{M}$, and

$$\frac{1}{n} I\left(G_0^n \wedge \mathbf{F}\right) < \epsilon_n, \qquad (33)$$

where $\epsilon_n \to 0$ as $n \to \infty$. It follows from the Fano's inequality that[4]

$$\frac{1}{n} H\left(G_i^n \mid X_i^n, \mathbf{F}\right) < c_1 \epsilon_n, \qquad i \in \mathcal{M}. \qquad (34)$$

Using an approach similar to that in [6], we have from (33):

$$\frac{1}{n} H\left(X_{\mathcal{M}}^n\right)$$
$$= \frac{1}{n} H\left(G_0^n, \mathbf{F}\right) + \frac{1}{n} H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right)$$
$$\geq \frac{1}{n} H\left(G_0^n\right) + \frac{1}{n} H\left(\mathbf{F}\right) + \frac{1}{n} H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right) - \epsilon_n, \qquad (35)$$
$$= \frac{1}{n} H\left(G_0^n\right) + \frac{1}{n} H\left(\mathbf{F}\right) + \frac{1}{n} \sum_{i=1}^m H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right)$$
$$\qquad\qquad - \epsilon_n. \qquad (36)$$

Next, for $\mathcal{L} \subsetneq \mathcal{M}$, with $[1, m_0] \nsubseteq \mathcal{L}$, we have

$$\frac{1}{n} H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M}\setminus\mathcal{L}}^n, \mathbf{F}\right)$$
$$= \frac{1}{n} H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M}\setminus\mathcal{L}}^n, G_0^n, \mathbf{F}\right) + \frac{1}{n} H\left(G_0^n \mid X_{\mathcal{M}\setminus\mathcal{L}}^n, \mathbf{F}\right)$$
$$\leq \frac{1}{n} H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M}\setminus\mathcal{L}}^n, G_0^n, \mathbf{F}\right) + c_1 \epsilon_n,$$

where the last step follows from (34) and the assumption that $g_i = g_0$ for $i \in [1, m_0]$. Continuing with the inequality above, we get

$$\frac{1}{n} H\left(X_{\mathcal{L}}^n \mid X_{\mathcal{M}\setminus\mathcal{L}}^n, \mathbf{F}\right)$$
$$\leq \frac{1}{n} \sum_{i \in \mathcal{L}} \left[H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1 \epsilon_n\right], \qquad (37)$$

[4]The constants c_1, c_2, c_3, c_4 depend only on $\log\|\mathcal{X}_{\mathcal{M}}\|$, m, m_0 (and not on n).

Letting

$$R_i = \frac{1}{n} H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n, \quad i \in \mathcal{M},$$

by (37) $R_1, ..., R_m$ satisfy (1a) and (1b) for $\mathbf{F}$, whereby it follows from (34) and (36) that

$$H\left(X_\mathcal{M} \mid G_0\right)$$
$$\geq \frac{1}{n}H(\mathbf{F}) + \frac{1}{n}\sum_{i=m_0+1}^{m} H\left(G_i^n \mid X_i^n, \mathbf{F}\right) + R_\mathcal{M} - c_2\epsilon_n$$
$$\geq R_\mathbf{F}^{(1)} - c_2\epsilon_n,$$

where $\mathbf{F}$ satisfies (33). Taking the limit $n \to \infty$, and using the definition of $R_1^*(g_\mathcal{M})$ we get $H\left(X_\mathcal{M} \mid G_0\right) \geq R_1^*\left(g_\mathcal{M}\right).$ $\square$

Necessity of (10) for $i = 2$: If $g_\mathcal{M}$ are securely computable, the approach above implies that there exists an interactive communication $\mathbf{F}$ satisfying (33) and (34) such that, with

$$R_i = \begin{cases} \frac{1}{n} H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n, & 1 \leq i \leq m_0, \\ \\ \frac{1}{n} H\left(X_i^n \mid X_{[1,i-1]}^n, G_{[m_0+1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n, \\ & m_0 < i \leq m, \end{cases}$$

$$R_j' = c_1\epsilon_n, \quad m_0 < j \leq m,$$

we have by (35),

$$H\left(X_\mathcal{M} \mid G_0\right)$$
$$\geq \frac{1}{n}H(\mathbf{F}) + \frac{1}{n}H\left(X_\mathcal{M}^n \mid G_0^n, \mathbf{F}\right) - \epsilon_n$$
$$\geq \frac{1}{n}H(\mathbf{F}) + \frac{1}{n}\sum_{i=1}^{m_0} H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right)$$
$$+ \frac{1}{n}\sum_{i=m_0+1}^{m} H\left(X_i^n \mid X_{[1,i-1]}^n, G_{[m_0+1,i-1]}^n, G_0^n, \mathbf{F}\right) - \epsilon_n$$
$$\geq \frac{1}{n}H(\mathbf{F}) + R_\mathcal{M} + R_{[m_0+1,m]}' - c_3\epsilon_n. \tag{38}$$

Furthermore, (34) and the assumption $g_i = g_0$, $1 \leq i \leq m_0$, yield for $[1, m_0] \not\subseteq \mathcal{L} \subsetneq \mathcal{M}$ that

$$\frac{1}{n}H\left(X_\mathcal{L}^n \mid X_{\mathcal{M}\setminus\mathcal{L}}^n, \mathbf{F}\right)$$
$$\leq \frac{1}{n}H\left(X_\mathcal{L}^n \mid X_{\mathcal{M}\setminus\mathcal{L}}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n$$
$$\leq \sum_{i\in\mathcal{L}, i\leq m_0}\left[\frac{1}{n}H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n\right] +$$
$$\sum_{i\in\mathcal{L}, i>m_0}\left[\frac{1}{n}H\left(X_i^n \mid X_{[1,i-1]}^n, G_{[m_0+1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n\right]$$
$$= R_\mathcal{L}, \tag{39}$$

and similarly, for $[1, m_0] \subseteq \mathcal{L} \subseteq \mathcal{M}$, $\mathcal{L}' \subseteq [m_0 + 1, m]$, with either $\mathcal{L} \neq \mathcal{M}$ or $\mathcal{L}' \neq [m_0 + 1, m]$ that

$$\frac{1}{n}H\left(G_{\mathcal{L}'}^n, X_\mathcal{L}^n \mid G_{[m_0+1,m]\setminus\mathcal{L}'}^n, X_{\mathcal{M}\setminus\mathcal{L}}^n, G_0^n, \mathbf{F}\right)$$
$$= \frac{1}{n}H\left(X_\mathcal{L}^n \mid G_{[m_0+1,m]\setminus\mathcal{L}'}^n, X_{\mathcal{M}\setminus\mathcal{L}}^n, G_0^n, \mathbf{F}\right)$$

$$\leq \frac{1}{n}H\left(X_\mathcal{L}^n \mid X_{\mathcal{M}\setminus\mathcal{L}}^n, G_0^n, \mathbf{F}\right)$$
$$\leq R_\mathcal{L} + R_{\mathcal{L}'}', \tag{40}$$

Therefore, (39), (34) and (40) imply that $R_1, ..., R_m$, $R_{m_0}', ..., R_m'$ satisfy (2a)-(2c) for $\mathbf{F}$, which along with (38) yields

$$H\left(X_\mathcal{M} \mid G_0\right) \geq R_\mathbf{F}^{(2)} - c_3\epsilon_n,$$

where $R_\mathbf{F}^{(2)}$ is as in (7), and $\mathbf{F}$ satisfies (33), which completes the proof of necessity (10) for $i = 2$ upon taking the limit $n \to \infty$. $\square$

Necessity of (10) for $i = 3$: If the functions $g_\mathcal{M}$ in (5) are securely computable then, as above, there exists an interactive communication $\mathbf{F}$ that satisfies (33) and (34). Defining

$$R_i = \frac{1}{n} H\left(X_i^n \mid X_{[1,i-1]}^n, G_0^n, \mathbf{F}\right) + c_1\epsilon_n, \quad i \in \mathcal{M},$$

similar manipulations as above yield

$$H\left(X_\mathcal{M} \mid G_0\right) \geq \frac{1}{n}H(\mathbf{F}) + R_\mathcal{M} - c_4\epsilon_n. \tag{41}$$

Further, from (34) we get that $R_1, ..., R_m$ satisfy (3a) and (3b) for $\mathbf{F}$. It follows from (41) that

$$H\left(X_\mathcal{M} \mid G_0\right) \geq R_\mathbf{F}^{(3)} - c_4\epsilon_n,$$

where $R_\mathbf{F}^{(2)}$ is as in (8), and $\mathbf{F}$ satisfies (33), which completes the proof of necessity (10) for $i = 3$ as above. $\square$

VI. DISCUSSION: ALTERNATIVE NECESSARY CONDITIONS FOR SECURE COMPUTABILITY

The necessary condition (10) for secure computing given in section III is in terms of quantities $R_\mathbf{F}^{(i)}$, $i = 1, 2, 3$, defined in (6), (7), (8), respectively. As remarked before, for $i = 1, 3$, the quantity $\inf_\mathbf{F} R_\mathbf{F}^{(i)}$ is the infimum over the rates of interactive communication that satisfy conditions (P1) and (P2). However, this is not true for $i = 2$. Furthermore, although $i = 1$ is special case of $i = 2$, it is not clear if the necessary condition (10) for $i = 2$ reduces to that for $i = 1$ upon imposing the restriction in (4). In this section, we shed some light on this baffling observation.

First, consider the functions $g_\mathcal{M}$ in (3). For this choice of functions, denoting by R_0^* the minimum rate of interactive communication that satisfies (P1) and (P2), the results in [19] imply that (1) constitutes a necessary condition for secure computability, with $R^* = R_0^*$.

Next, consider an augmented model obtained by introducing a new terminal $m + 1$ that observes rv $X_{m+1} = \tilde{g}(X_\mathcal{M})$ and seeks to compute $g_{m+1} = \emptyset$. Further, the terminal does not communicate, i.e., observation X_{m+1}^n is available only for decoding. Clearly, secure computability in the original model implies secure computability in the new model. It follows from the approach of [19] that for the new model also, (1) constitutes a necessary condition for secure computability, with R^* now being the minimum rate of interactive communication that satisfies (P1) and (P2) when terminal $m + 1$ does not communicate; this R^* is given by

$$\max\{H\left(X_\mathcal{M} \mid \tilde{g}(X_\mathcal{M}), G_0\right), R_0^*\}.$$

Note that the new necessary condition (1) is

$$H\left(X_{\mathcal{M}} \mid G_0\right) \geq R_0^* = \max\{H\left(X_{\mathcal{M}} \mid \tilde{g}(X_{\mathcal{M}}), G_0\right), R_0^*\},$$

which is, surprisingly, same as the original condition

$$H\left(X_{\mathcal{M}} \mid G_0\right) \geq R_0^*.$$

Our necessary condition (10) for $i = 2$ is based on a similar augmentation that entails introduction of $m - m_0$ new terminals observing $g_{m_0+1}\left(X_{\mathcal{M}}\right), ..., g_m\left(X_{\mathcal{M}}\right)$ (to be used only for decoding). Now, however, this modification may result in a different necessary condition.

APPENDIX A

From (25), we have

$$nR_{\mathcal{M}} + \frac{\delta}{2} < H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right),$$

where $R_1, ..., R_m$ satisfy conditions (1a) and (1b). 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 $\mathcal{X}_i^{nk} \rightarrow \{1, ..., \lceil \exp(knR_i)\rceil\}$, $i \in \mathcal{M}$. The rvs $J_1, ..., J_m, X_{\mathcal{M}}^{nk}$ are taken to be mutually independent.

Fix ϵ, ϵ', with $\epsilon' > m\epsilon$ and $\epsilon + \epsilon' < 1$. It follows from the proof of the general source network coding theorem [5, Lemma 13.13 and Theorem 13.14] that for all sufficiently large k,

$$\Pr\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : X_{\mathcal{M}}^{nk} \text{ is } \epsilon\text{-recoverable from}\right.\right.$$
$$\left.\left. \left(X_i^{nk}, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right), Z_i^k\right), i \in \mathcal{M}\right\}\right) \geq 1 - \epsilon,$$
$$\text{(A1)}$$

where, for $i \in \mathcal{M}$,

$$Z_i^k = \begin{cases} \mathbf{F}^k, & j \in [1, m_0], \\ \left(\mathbf{F}^k, G_0^{nk}\right), & m_0 < j \leq m. \end{cases}$$

Below we shall establish that

$$\Pr\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : \frac{1}{nk} I\left(j_{\mathcal{M}}(X_{\mathcal{M}}^{nk}) \wedge G_0^{nk}, \mathbf{F}^k\right) \geq \epsilon\right\}\right) \leq \epsilon',$$
$$\text{(A2)}$$

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

$$\Pr\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : \right.\right.$$
$$\left.\left. \frac{1}{nk} I\left(j_i(X_i^{nk}) \wedge G_0^{nk}, \mathbf{F}^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right)\right) \geq \frac{\epsilon}{m}\right\}\right)$$
$$\leq \frac{\epsilon'}{m}, \quad i \in \mathcal{M}, \tag{A3}$$

since

$$I\left(j_{\mathcal{M}}\left(X_{\mathcal{M}}^{nk}\right) \wedge G_0^{nk}, \mathbf{F}^k\right)$$
$$= \sum_{i=1}^{m} I\left(j_i\left(X_i^{nk}\right) \wedge G_0^{nk}, \mathbf{F}^k \mid j_1\left(X_1^{nk}\right), ..., j_{i-1}\left(X_{i-1}^{nk}\right)\right)$$
$$\leq \sum_{i=1}^{m} I\left(j_i\left(X_i^{nk}\right) \wedge G_0^{nk}, \mathbf{F}^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right)\right).$$

Then it would follow from (A1), (A2), and definition of $Z_{\mathcal{M}}$ that

$$\Pr\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : X_{\mathcal{M}}^{nk} \text{ is } \epsilon\text{-recoverable from}\right.\right.$$
$$\left(X_i^{nk}, Z_i^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right)\right), i \in \mathcal{M}, \text{ and}$$
$$\left.\left. \frac{1}{nk} I\left(j_{\mathcal{M}}(X_{\mathcal{M}}^{nk}) \wedge G_0^{nk}, \mathbf{F}^k\right) < \epsilon\right\}\right) \geq 1 - \epsilon - \epsilon'.$$

This shows the existence of a particular realization $\mathbf{F}'$ of $J_{\mathcal{M}}$ that satisfies (26) and (27).

It now remains to prove (A3). Defining

$$\tilde{\mathcal{J}}_i = \left\{j_{\mathcal{M}\setminus\{i\}} \in \mathcal{J}_{\mathcal{M}\setminus\{i\}} : X_{\mathcal{M}}^{nk} \text{ is } \epsilon\text{-recoverable from}\right.$$
$$\left.\left(X_i^{nk}, Z_i^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right),\right)\right\},$$

we have by (A1) that $\Pr\left(J_{\mathcal{M}\setminus\{i\}} \in \tilde{\mathcal{J}}_i\right) \geq 1 - \epsilon$. It follows that

$$\Pr\left(\left\{j_{\mathcal{M}} \in \mathcal{J}_{\mathcal{M}} : \right.\right.$$
$$\left.\left. \frac{1}{nk} I\left(j_i(X_i^{nk}) \wedge G_0^{nk}, \mathbf{F}^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right)\right) \geq \frac{\epsilon}{m}\right\}\right)$$
$$\leq \epsilon + \sum_{j_{\mathcal{M}\setminus\{i\}} \in \tilde{\mathcal{J}}_i} \Pr\left(J_{\mathcal{M}\setminus\{i\}} = j_{\mathcal{M}\setminus\{i\}}\right) p\left(j_{\mathcal{M}\setminus\{i\}}\right),$$

since J_i is independent of $J_{\mathcal{M}\setminus\{i\}}$, where $p\left(j_{\mathcal{M}\setminus\{i\}}\right)$ is defined as

$$\Pr\left(\left\{j_i \in \mathcal{J}_i : \right.\right.$$
$$\left.\left. \frac{1}{nk} I\left(j_i(X_i^{nk}) \wedge G_0^{nk}, \mathbf{F}^k, j_{\mathcal{M}\setminus\{i\}}\left(X_{\mathcal{M}\setminus\{i\}}^{nk}\right)\right) \geq \frac{\epsilon}{m}\right\}\right).$$

Thus, (A3) will follow upon showing that

$$p\left(j_{\mathcal{M}\setminus\{i\}}\right) \leq \frac{\epsilon'}{m} - \epsilon, \quad j_{\mathcal{M}\setminus\{i\}} \in \tilde{\mathcal{J}}_i, \tag{A4}$$

for all k sufficiently large. Fix $j_{\mathcal{M}\setminus\{i\}} \in \tilde{\mathcal{J}}_i$. We take recourse to Lemma B1 in Appendix B, and set $U = X_{\mathcal{M}}^{nk}$, $U' = X_i^{nk}, V = \left(G_0^{nk}, \mathbf{F}^k\right), h = j_{\mathcal{M}\setminus\{i\}}$, and

$$\mathcal{U}_0 = \left\{x_{\mathcal{M}}^{nk} \in \mathcal{X}_{\mathcal{M}}^{nk} : x_{\mathcal{M}}^{nk} = \psi_i\left(x_i^{nk}, j_{\mathcal{M}\setminus\{i\}}\left(x_{\mathcal{M}\setminus\{i\}}^{nk}\right),\right.\right.$$
$$\left.\left. \mathbf{F}^k\left(x_{\mathcal{M}}^{nk}\right), g_0^n\left(x_{\mathcal{M}}^n\right) \mathbf{1}\left(m_0 < i \leq m\right)\right)\right\}$$

for some mapping ψ_i. By the definition of $\tilde{\mathcal{J}}_i$,

$$\Pr\left(U \in \mathcal{U}_0\right) \geq 1 - \epsilon,$$

so that condition (B1)(i) preceding Lemma B1 is met. Condition (B1)(ii), too, is met from the definition of $\mathcal{U}_0, h$ and V.

Upon choosing

$$d = \exp\left[k\left(H\left(X_{\mathcal{M}}^n \mid G_0^n, \mathbf{F}\right) - \frac{\delta}{2}\right)\right],$$

in (B2), the hypotheses of Lemma B1 are satisfied, for

appropriately chosen λ, and for sufficiently large k. Then, by Lemma B1, with

$$r = \lceil \exp(knR_i) \rceil, \quad r' = \lceil \exp(knR_{\mathcal{M}\setminus i}) \rceil,$$

and with J_i in the role of ϕ, (A4) follows from (B3) and (B4). $\qquad\square$

APPENDIX B

Our proof of sufficiency in Theorem 1 requires random mappings to satisfy certain "almost independence" and "almost uniformity" properties. The following version of the "balanced coloring lemma" given in [19] constitutes the key step in the derivation of these properties.

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 $h : \mathcal{U} \to \{1, \ldots, r'\}$. For $0 < \lambda < 1$, let $\mathcal{U}_0$ be a subset of $\mathcal{U}$ such that
(i) $\Pr(U \in \mathcal{U}_0) > 1 - \lambda^2$;
(ii) given the event $\{U \in \mathcal{U}_0, h(U) = j, U' = u', V = v\}$, there exists $u = u(u') \in \mathcal{U}_0$ satisfying

$$\Pr(U' = u' \mid h(U) = j, V = v, U \in \mathcal{U}_0)$$
$$= \Pr(U = u \mid h(U) = j, V = v, U \in \mathcal{U}_0), \quad \text{(B1)}$$

for $1 \leq j \leq r'$ and $v \in \mathcal{V}$. Then the following holds.

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

$$P_{UV}\left(\left\{(u, v) : \Pr(U = u \mid V = v) > \frac{1}{d}\right\}\right) \leq \lambda^2. \quad \text{(B2)}$$

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

$$\sum_{j=1}^{r'} \sum_{v \in \mathcal{V}} \Pr(h(U) = j, V = v) \times$$

$$\sum_{i=1}^{r} \left| \sum_{\substack{u' \in \mathcal{U}': \\ \phi(u') = i}} \Pr(U' = u' \mid h(U) = j, V = v) - \frac{1}{r} \right| < 14\lambda,$$
$$\text{(B3)}$$

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

Remark. Denoting by s_{var} the left side of (B3), it follows from [6, Lemma 1] that

$$\log r - H(\phi(U)) + I(\phi(U) \wedge h(U), V) \leq s_{var} \log \frac{r}{s_{var}}.$$

Since the function $f(x) = x \log(r/x)$ is increasing for $0 < x < re$, it follows from (B3) that

$$\log r - H(\phi(U)) + I(\phi(U) \wedge h(U), V) \leq 14\lambda \log \frac{|\mathcal{U}|}{14\lambda}. \quad \text{(B4)}$$

ACKNOWLEDGEMENTS

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] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography–part ii: CR capacity," *IEEE Trans. Inform. Theory*, vol. 44, pp. 225–240, 1998.

[3] C. Chan, "Multiterminal secure source coding for a common secret source," *in Proceedings of 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 188 - 195, 2011.

[4] I. Csiszár, "Almost independence and secrecy capacity," *Prob. Pered. Inform.*, vol. 32, no. 1, pp. 48–57, 1996.

[5] I. Csiszár and J. Körner, *Information theory: Coding Theorems for Discrete Memoryless Channels*. 2nd Edition. Cambridge University Press, 2011.

[6] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3047–3061, 2004.

[7] I. Csiszár and P. Narayan, "Secrecy capacities for multiterminal channel models," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2437–2452, 2008.

[8] R. G. Gallager, "Finding parity in a simple broadcast nework," *IEEE Trans. Inform. Theory*, vol. 34, no. 2, pp. 176–180, 1988.

[9] A. Giridhar and P. Kumar, "Computing and communicating functions over sensor networks," *IEEE Journ. on Select. Areas in Commun.*, vol. 23, no. 4, pp. 755–764, 2005.

[10] E. Kushilevitz and N. Nisan, *Communication complexity*. Cambridge University Press, 1997.

[11] N. Ma, P. Ishwar, and P. Gupta, "Information-theoretic bounds for multi-round function computation in collocated networks," *IEEE International Symposium on Information Theory (ISIT)*, pp. 2306–2310, 2009.

[12] M. Madiman and P. Tetali, "Information inequalities for joint distributions, with interpretations and applications," *IEEE Trans. Inform. Theory*, vol. 56, pp. 2699–2713, 2010.

[13] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inform. Theory*, vol. 39, pp. 733–742, 1993.

[14] ——, *Communications and Cryptography: Two sides of One Tapestry*, R.E. Blahut et al., Eds. ed. Norwell, MA: Kluwer, ch. 26, pp. 271–285, 1994.

[15] A. Orlitsky and A. E. Gamal, "Communication with secrecy constraints," *Proc. 16th Ann. Symp. on Theory of Computing*, pp. 217–224, 1984.

[16] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 903–917, 2001.

[17] D. Slepian and J. Wolf, "Noiseless coding of correlated information source," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, 1973.

[18] H. Tyagi, P. Narayan, and P. Gupta, "Secure computing," *Proc. Int. Symp. Inform. Theory*, pp. 2612 – 2616, June 2010.

[19] H. Tyagi, P. Narayan, and P. Gupta, "When is a function securely computable?" *IEEE Trans. Inform. Theory*, vol. 57, no. 10, 2011.

[20] A. C. Yao, "Some complexity questions related to distributive computing," *Proc. 11th Ann. Symp. on Theory of Computing*, pp. 209–213, 1979.

[21] A. C. Yao, "Protocols for secure computations," *Proc. 23rd Ann. Symp. on Foundations of Computer Science*, pp. 160–164, 1982.

Himanshu Tyagi received the Bachelor of Technology degree in electrical engineering and the Master of Technology degree in communication and information technology, both from the Indian Institute of Technology, Delhi, India in 2007. He is currently a Ph.D. candidate at the University of Maryland, College Park, USA.