Secrecy Capacities for Multiple Terminals

Imre Csiszár, *Fellow, IEEE,* and Prakash Narayan, *Fellow, IEEE*

Abstract—We derive single-letter characterizations of (strong) secrecy capacities for models with an arbitrary number of terminals, each of which observes a distinct component of a discrete memoryless multiple source, with unrestricted and interactive public communication permitted between the terminals. A subset of these terminals can serve as helpers for the remaining terminals in generating secrecy. According to the extent of an eavesdropper's knowledge, three kinds of secrecy capacity are considered: secret key (SK), private key (PK), and wiretap secret key (WSK) capacity. The characterizations of the SK and PK capacities highlight the innate connections between secrecy generation and multiterminal source coding without secrecy requirements. A general upper bound for WSK capacity is derived which is tight in the case when the eavesdropper can wiretap noisy versions of the components of the underlying multiple source, provided randomization is permitted at the terminals. These secrecy capacities are seen to be achievable with noninteractive communication between the terminals. The achievability results are also shown to be universal.

Index Terms—Common randomness, multiple source, private key, public discussion, secrecy capacity, security index, Slepian–Wolf constraints, wiretap.

I. INTRODUCTION

IT might seem surprising that separate terminals, which observe the outputs of distinct albeit correlated sources, could devise a secret key by means of public communication. In other words, these terminals are able to generate common randomness regarding which an eavesdropper, with access to this communication and perhaps also to side information comprising the outputs of other "wiretapped" sources which are correlated with the previous sources, can glean only a negligibly small amount of mutual information.

This fact was first noted, to our knowledge, by Maurer (see [8]). The case of two terminals has been extensively examined by Maurer [8], Ahlswede and Csiszár [1], and Csiszár and Narayan [6], with the last also allowing a "helper" terminal which observes the output of another source, to assist in generating secrecy. Still, a single-letter characterization of the secrecy capacity—the largest rate at which secrecy can be generated—is known only under special circumstances,

e.g., when the permissible communication is limited to just a single transmission by one of the two participating terminals (of arbitrary rate [1] or of constrained rate [6]). A simple case for which a complete and heuristically expected answer is available, entails secrecy generation by two terminals with unrestricted communication when the eavesdropper has no side information beyond this communication. The corresponding secrecy capacity is equal to the mutual information $I(Y \wedge Z)$ of the random variables (rvs) Y, Z, whose respective independent and identically distributed (i.i.d.) repetitions are observed by the two terminals [8], [1]. With the assistance of a helper, which observes i.i.d. repetitions of an rv X, the secrecy capacity is $\min\{I(XY \wedge Z), I(XZ \wedge Y)\}$ [6]; and if the helper is wiretapped, the corresponding secrecy capacity is $I(Y \wedge Z|X)$ [1], [6].

In this paper, we consider models with an arbitrary number of terminals, each of which observes a distinct component of a discrete memoryless multiple source. Unrestricted public communication is allowed between these terminals, i.e., no rate constraints are imposed on the communication, which may be interactive. It is understood that all the transmissions are observed by all the terminals. If the terminals were to communicate over a network with links of fixed connectivity and constrained rates, a more complex model would be required. Also, our models tacitly assume that all the public transmissions are impervious to any deliberate attempts at inserting corruption. In a cryptographic situation, this assumption implies, in effect, that the public transmissions are authenticated, or that the eavesdropper is passive, i.e., unable to tamper with such transmissions. (For unauthenticated public transmissions, see for instance [10].)

Our main technical contribution is the determination of the secrecy capacity when an eavesdropper observes the communication between the terminals but does not have access to any other information. Suppose that $m \geq 2$ terminals observe i.i.d. repetitions of the rvs $X_1, \ldots, X_m$, respectively. Let $A \subset \{1, \ldots, m\}$ be an arbitrary subset[1] of terminals. We show that the largest rate at which the set A of terminals can generate secrecy with the remaining terminals serving as helpers, is obtained by subtracting from the total joint entropy $H(X_1, \ldots, X_m)$ the smallest rate of communication which enables each terminal in A to reconstruct all the m components of the multiple source. The problem of determining the latter rate is, of course, one of multiterminal source coding which is *not concerned with any secrecy constraints*. Note, however, that helpers for secrecy generation do not correspond to helpers in the sense of multiterminal source coding (see [4]), and determining the rate above involves no helpers in that sense. The mentioned result highlights an *inherent connection*

[1]Throughout, the symbol $\subset$ denotes a subset which need not be a proper subset.

between secrecy generation and multiterminal source coding, and affords the following heuristically satisfying interpretation: the total rate of shared common randomness achievable by these terminals, viz. $H(X_1, \ldots, X_m)$, can be decomposed into the least rate of communication needed for this purpose and the largest secrecy rate. We provide a single-letter characterization of this smallest rate of communication, and thereby also of the secrecy capacity. The capacity result, though not in the form of an explicit formula, is remarkably simple; in particular, it does not involve any auxiliary rvs. We also show that in order to achieve the secrecy capacity, there is no need for interactive communication; a single autonomous transmission from each terminal is adequate. Furthermore, additional randomization at the terminals does not serve to enhance the secrecy capacity.

In the general case in which the eavesdropper also observes wiretapped sources in addition to the communication between the terminals, a complete solution is not to be expected since the problem remains unresolved even when $m = 2$. However, we do obtain a single-letter characterization in the special case when the wiretapped sources represent a subset of the helpers. For instance, when $A = \{1, \ldots, m\} \setminus \{j\}$ and the terminal j is wiretapped, this model depicts the situation when a centralized server j helps the other terminals generate a "private key" (PK) which is concealed even from itself. (This case, for $m = 3$, also subject to communication rate constraints, was considered in [6].) Additionally, the result for this special case provides an upper bound for the general case which is sometimes tight. The last occurs in the practically realistic situation in which the wiretapped sources constitute noisy versions of the components of the underlying multiple source.

All our capacity results mentioned above hold in a strong sense; specifically, achievability results are established for the "strong" version and converse results for the "weak" version of Definition 1 in Section II. The concept of strong secrecy capacity was introduced in [9], and our concept (used previously in [5], [6]) is even stronger. While the phenomenon of the weak and strong definitions of secrecy capacity leading to identical results has been demonstrated for various models in [9], [5], [11], [6], it is worth mentioning that our technique leads directly to strong achievability.

Section II contains the preliminaries. Our main results are stated in Section III, and the proofs are provided in Section IV. Section V addresses the computation of the secrecy capacities. In particular, we obtain explicit formulas when $X_1, \ldots, X_m$ form a Markov chain or, generally, a Markov chain on a tree. The discussion in Section VI includes the universality of our strong achievability results for the class of discrete memoryless multiple sources, and extensions of the characterizations of the various secrecy capacities to the discrete stationary ergodic multiple source.

II. PRELIMINARIES

Let $X_1, \ldots, X_m$, $m \geq 2$, be rvs with finite alphabets $\mathcal{X}_1, \ldots, \mathcal{X}_m$, respectively. For $A \subset \mathcal{M} = \{1, \ldots, m\}$, we denote $X_A = (X_i, i \in A)$ and, in particular, $X_{[a,b]} = (X_a, \ldots, X_b)$, $1 \leq a < b \leq m$. Similarly, for real numbers $R_1, \ldots, R_m$, and set $A \subset \mathcal{M}$, we shall use the notation $R_A = (R_i, i \in A)$. We

shall denote n i.i.d. repetitions of $X_{\mathcal{M}} = (X_1, \ldots, X_m)$ by $X_{\mathcal{M}}^n = (X_1^n, \ldots, X_m^n)$ and of X_A by $X_A^n = (X_i^n, i \in A)$. 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 rvs are assumed to take values in finite sets, even if not stated explicitly.

In the models considered in this paper, the components of a discrete memoryless multiple source with generic rvs $X_1, \ldots, X_m$ are, respectively, observed by m different terminals. Thus, with observation length n, the ith terminal observes X_i^n, $i \in \mathcal{M}$. Hereafter, m will always denote the number of terminals and n the observation length, unless stated otherwise. The terminals are allowed to communicate over a noiseless public channel, possibly interactively in several rounds. We assume without any loss of generality that such transmissions occur in consecutive time slots in r rounds. The communication is described in terms of mappings $f_1, \ldots, f_{rm}$, with f_ν corresponding to the transmission in slot ν by terminal i, where $i \equiv \nu \bmod m$, $1 \leq i \leq m$; in general, f_ν is allowed to yield any function of $X_i^n = x_i^n$ and of the previous transmissions described by $f_{[1,\nu-1]} = (f_1, \ldots, f_{\nu-1})$. The corresponding rvs representing the communication will be depicted by $F_1, \ldots, F_{rm}$, where $F_\nu = f_\nu(X_i^n, F_{[1,\nu-1]})$ with $F_{[1,\nu-1]} = (F_1, \ldots, F_{\nu-1})$, and we denote $\boldsymbol{F} = F_{[1,rm]}$.

For our purposes, it will transpire that noninteractive communication suffices, i.e., with $\boldsymbol{F} = (F_1, \ldots, F_m)$, where $F_i = f_i(X_i^n)$, $i \in \mathcal{M}$. Randomization at the terminals is not permitted in our basic models. However, as we show later, most of our results remain unaffected when randomization is permitted.

A function K of $X_{\mathcal{M}}^n$ will be termed ϵ-*common randomness* (ϵ-CR) for a set of terminals $A \subset \mathcal{M}$, achievable with communication $\boldsymbol{F} = F_{[1,rm]}$, if K is ϵ-recoverable from $(X_i^n, \boldsymbol{F})$ for each $i \in A$.

A function K of $X_{\mathcal{M}}^n$ constitutes an ϵ-*secret key* (ϵ-SK) for a set of terminals $A \subset \mathcal{M}$, achievable with communication $\boldsymbol{F}$, if K is ϵ-CR for A which is achievable with $\bar{\boldsymbol{F}}$, and, in addition, it satisfies the secrecy condition[2]

$$\frac{1}{n} I(K \wedge \boldsymbol{F}) \leq \epsilon \tag{1}$$

and the uniformity condition

$$\frac{1}{n} H(K) \geq \frac{1}{n} \log |\mathcal{K}| - \epsilon. \tag{2}$$

Here, $\mathcal{K}$ denotes the set of possible values of K. The terminals in A thus generate an SK, with the terminals in $A^c = \mathcal{M} \setminus A$ serving as helpers (e.g., centralized or trusted servers in a key establishment protocol) by furnishing the terminals in A with additional correlated information; the SK thus generated is effectively concealed from an eavesdropper with access to the public communication $\boldsymbol{F}$ and is nearly uniformly distributed.

A more frequently encountered practical situation arises when the eavesdropper has access to side information in addition to the public communication between the terminals. To model this situation, consider an rv Z jointly distributed with $X_{\mathcal{M}} = (X_1, \ldots, X_m)$, and suppose that the eavesdropper can wiretap the component Z^n of the i.i.d. repetitions of $(X_{\mathcal{M}}, Z)$.

---

[2]All logarithms and exponentiations are with respect to the base 2.

Then an ϵ-SK for a set of terminals $A \subset \mathcal{M}$ will be called an ϵ-*wiretap secret key* (ϵ-WSK) if it satisfies the stronger secrecy condition

$$\frac{1}{n} I(K \wedge \boldsymbol{F}, Z^n) \leq \epsilon. \tag{3}$$

A particular case, depicting the situation in which some of the helper terminals are compromised, arises when $Z = X_D$ for a set $D \subset A^c$. An ϵ-WSK for this case, i.e., when (3) is satisfied for $Z = X_D$, will be called an ϵ-*private key* (ϵ-PK) for A which is private from the terminals $D \subset A^c$. This notion is meaningful, of course, only for proper subsets $A \subset \mathcal{M}$ and for nonempty subsets $D \subset A^c$. A practical situation of this kind arises, with $D = \{j\}$, when terminal j represents a centralized server which helps the other terminals generate an SK that is concealed even from itself. For studying PK capacity (defined below), attention could be restricted to the case when D is a singleton set; the general case could be reduced to it by replacing the terminals in D with a solitary terminal which observes X_D^n. Still, we prefer to avoid this restriction as some of our results can be presented more naturally when the set D of compromised helper terminals is arbitrary.

A less special, and practically important, case of wiretap secrecy which we shall also consider arises when the eavesdropper can wiretap noisy versions of some or all of the components of the underlying multiple source. In this case, formally $Z = (Z_1, \ldots, Z_m)$ where the conditional probability mass function (pmf) of Z given $X_{\mathcal{M}}$ is of the product form

$$\Pr\{Z_1 = z_1, \ldots, Z_m = z_m | X_1 = x_1, \ldots, X_m = x_m\}$$
$$= \prod_{i=1}^{m} \Pr\{Z_i = z_i | X_i = x_i\}. \tag{4}$$

We shall consider three kinds of secrecy capacity, corresponding to the concepts above of ϵ-SK, ϵ-WSK, and ϵ-PK.

Definition 1: Given $X_{\mathcal{M}} = (X_1, \ldots, X_m)$, a number H is an achievable SK rate for a set of terminals $A \subset \mathcal{M}$ if ϵ_n-SKs $K^{(n)} = K^{(n)}(X_{\mathcal{M}}^n)$ for A are achievable with suitable communication (with the number of rounds possibly depending on n), such that

$$\epsilon_n \to 0 \quad \text{and} \quad \frac{1}{n} H(K^{(n)}) \to H. \tag{5}$$

The largest achievable SK rate for A is the SK *capacity* $C_S(A)$. Achievable WSK and PK rates are defined analogously, and so are the WSK capacity $C_W(A)$ and the PK capacity $C_P(A|D)$. An achievable SK, WSK, or PK rate will be said to be strongly achievable if ϵ_n above can be taken to vanish exponentially in n; a secrecy capacity (i.e., an SK, WSK, or PK capacity) will be termed *strong* if all smaller rates are strongly achievable.

Remarks: Clearly, the condition $\frac{1}{n} H(K^{(n)}) \to H$ could be replaced by $\frac{1}{n} \log |\mathcal{K}^{(n)}| \to H$, where $\mathcal{K}^{(n)}$ is the set of possible values of $K^{(n)}$. It is also obvious that in our definition of ϵ-CR and ϵ-SK (-WSK, -PK), we could drop the condition that K is a function of $X_{\mathcal{M}}^n$ with no change in achievable rates. It was pointed out by Maurer [9] that the secrecy and uniformity

conditions (1)–(3) were inadequate for cryptographic purposes, and should be strengthened by omission of the factor $\frac{1}{n}$. Note that our concept of strong achievability demands even more. Moreover, if K is not required to be a function of $X_{\mathcal{M}}^n$, we could strengthen the uniformity condition (2) to strict uniformity, $H(K) = \log |\mathcal{K}|$, without changing the strongly achievable rates, as in [11].

Example 1: Let $X_1, \ldots, X_m$ be binary rvs, where $X_1, \ldots, X_{m-1}$ are mutually independent with each uniformly distributed on $\{0, 1\}$ and $X_m = X_1 + \cdots + X_{m-1} \mod 2$. Note that each X_i equals the modulo 2 sum of the other X_j's, which are independent and uniformly distributed on $\{0, 1\}$. We claim that in this elementary example, 1 bit of perfect SK (i.e., ϵ-SK with $\epsilon = 0$) is achievable for any set of terminals $A \subset \mathcal{M}$, with observation length $n = |A| - 1$. First, let $A = \mathcal{M} = \{1, \ldots, m\}$, $n = m - 1$. Let each terminal i, $1 \leq i \leq m - 1$, transmit at rate $\frac{n-1}{n}$, with $F_i = f_i(X_i^n)$ being the block $X_i^n = (X_{i1}, \ldots, X_{in})$ excluding X_{ii}, and let

$$F_m = f_m(X_m^n) = (X_{m1} + X_{m2}, X_{m1} + X_{m3}, \ldots, X_{m1} + X_{mn})$$

where the additions are modulo 2. It is easily seen that $K = X_{11}$, say, is perfectly recoverable from any X_i^n and the communication $\boldsymbol{F} = (F_1, \ldots, F_m)$. Moreover, all of $X_{\mathcal{M}}^n$ is perfectly recoverable from $(X_{11}, \boldsymbol{F})$, implying that

$$H(X_{11}, \boldsymbol{F}) = H(X_{\mathcal{M}}^n) = n^2 = 1 + (n^2 - 1)$$
$$= H(X_{11}) + H(\boldsymbol{F})$$

where the last equality holds because strict inequality in $H(\boldsymbol{F}) \leq m(n - 1) = n^2 - 1$ is ruled out by

$$H(X_{11}, \boldsymbol{F}) \leq H(X_{11}) + H(\boldsymbol{F}).$$

This shows that $I(X_{11} \wedge \boldsymbol{F}) = 0$, and so $K = X_{11}$ is a perfect SK for all the terminals.

Next, let $A \subset \mathcal{M}$ be arbitrary and assume without any loss of generality that $A = \{1, \ldots, n, m\}$. In this case, let the transmissions F_i be as above but with the omission of all bits X_{ij} with $j > n$. Then $K = X_{11}$ is a perfect SK for the terminals A. Moreover, since the transmissions of the terminals i, $n + 1 \leq i < m$, equal the entire X_i^n available at those terminals, $K = X_{11}$ is also a (perfect) PK for $A = \{1, \ldots, n, m\}$, private from the set of terminals $D = A^c = \{n+1, \ldots, m-1\}$. In this example

$$C_S(\mathcal{M}) = \frac{1}{m-1}$$
$$C_S(A) = C_P(A|A^c) = \frac{1}{|A|-1}, \qquad 1 < |A| < m.$$

We have shown above that these values are achievable SK or PK rates (with $\epsilon = 0$). The results in this paper will show that they cannot be bettered (see Example 4 in Section V). □

Finally, it will be useful for us to introduce two "security indices" for the situation when an SK and an eavesdropper's knowledge are, respectively, represented by the rvs K and Y, taking values in the finite sets $\mathcal{K}$ and $\mathcal{Y}$. Ideally, K should be independent of Y and uniformly distributed on $\mathcal{K}$. The deviation

from this ideal can be measured by the information-theoretic security index

$$
\begin{aligned}
s_{\text{in}} &= \log |\mathcal{K}| - H(K) + I(K \wedge Y) \\
&= \log |\mathcal{K}| - H(K|Y).
\end{aligned} \tag{6}
$$

Note that it is the smallness of s_{in}, with an appropriate choice of Y, which is effectively called for in the definitions above of ϵ-SK, ϵ-WSK, and ϵ-PK. Simple algebra shows that s_{in} equals the information divergence between the joint pmf of K, Y, and the Cartesian product of the uniform pmf on $\mathcal{K}$ and the marginal pmf of Y. An alternative security index is the variation distance between these two pmfs on $\mathcal{K} \times \mathcal{Y}$, or equivalently, the average variation distance of the conditional pmf of K conditioned on $Y = y$ from the uniform pmf on $\mathcal{K}$, the average being taken with respect to the pmf of Y:

$$
\begin{aligned}
s_{\text{var}} &= \sum_{k \in \mathcal{K}} \sum_{y \in \mathcal{Y}} \left| \Pr\{K = k, Y = y\} - \frac{1}{|\mathcal{K}|} \Pr\{Y = y\} \right| \\
&= \sum_{y \in \mathcal{Y}} \Pr\{Y = y\} \sum_{k \in \mathcal{K}} \left| \Pr\{K = k | Y = y\} - \frac{1}{|\mathcal{K}|} \right|.
\end{aligned} \tag{7}
$$

Lemma 1: We have, with $|\mathcal{K}| \geq 4$ for the second inequality, that

$$
\frac{\log e}{2} s_{\text{var}}^2 \leq s_{\text{in}} \leq s_{\text{var}} \log \frac{|\mathcal{K}|}{s_{\text{var}}}. \tag{8}
$$

For our purposes, the second inequality will be crucial. The proof of Lemma 1 is the same as that of a similar result ([5, Lemma 1]). Still, for completeness, a proof will be provided in Appendix B.

III. STATEMENT OF RESULTS

In analogy with the SK, WSK, and PK capacities, we could also define a notion of CR capacity. However, in the context of our models, with no rate constraints imposed on the communication between the terminals, it is obvious that $X_{\mathcal{M}}^n$ is always an achievable CR, so that the CR capacity trivially equals $H(X_{\mathcal{M}})$. Interestingly enough, however, it transpires that the problem of determining the SK capacity for a set of terminals $A \subset \mathcal{M}$ is closely connected with that of determining the minimum rate of overall communication which renders $X_{\mathcal{M}}^n$ an achievable CR for these terminals, i.e., makes them "omniscient" in effect. Similarly, the PK capacity for $A \subset \mathcal{M}$, private from a set $D \subset A^c$, will be related to the minimum rate of overall communication needed to make $X_{\mathcal{M}}^n$ an achievable CR for the set of terminals A if each terminal $i \in D$ transmits all of X_i^n.

Definition 2: Given $X_{\mathcal{M}} = (X_1, \ldots, X_m)$, a number R is called an achievable rate of "communication for omniscience" (CO rate) for a set of terminals $A \subset \mathcal{M}$, if $X_{\mathcal{M}}^n$ is achievable as ϵ_n-CR for these terminals with communication $\boldsymbol{F}^{(n)} = F_{[1,rm]}^{(n)}$ (with the number r of rounds possibly depending on n), such that

$$
\epsilon_n \to 0 \quad \text{and} \quad \frac{1}{n} \log \|\boldsymbol{F}^{(n)}\| \to R \tag{9}
$$

where $\|\boldsymbol{F}^{(n)}\|$ denotes the cardinality of the range of $\boldsymbol{F}^{(n)}$. Further, given a set $D \subset A^c$, an achievable CO rate for A when each terminal $i \in D$ transmits all of X_i^n, is defined similarly but with $\boldsymbol{F}^{(n)}$ is (9) replaced by the communication of the terminals $i \in D^c$. The smallest achievable CO rate is denoted by $R_{\text{CO}}(A)$ in the former case and by $R_{\text{CO}}(A|D)$ in the latter case.

We shall relate the problem of determining the smallest achievable CO rates $R_{\text{CO}}(A)$ and $R_{\text{CO}}(A|D)$, and thereby the SK and PK capacities $C_S(A)$ and $C_P(A|D)$, to straightforward extensions of the Slepian–Wolf multiterminal source coding problem which have the achievable rate regions

$$
\mathcal{R}(A) = \left\{ R_{\mathcal{M}} : \sum_{i \in B} R_i \geq H(X_B | X_{B^c}), B \subset \mathcal{M}, \ A \not\subset B \right\} \tag{10}
$$

and

$$
\mathcal{R}(A|D) = \left\{ R_{D^c} : \sum_{i \in B} R_i \geq H(X_B | X_{B^c}), \right.
$$
$$
\left. B \subset D^c, \ A \not\subset B \right\} \tag{11}
$$

respectively. The inequalities in (10) and (11) will be referred to as the SW constraints.

Proposition 1: It holds that

$$
R_{\text{CO}}(A) = \min_{R_{\mathcal{M}} \in \mathcal{R}(A)} \sum_{i=1}^m R_i \tag{12}
$$

where $\mathcal{R}(A)$ is given by (10). Further

$$
R_{\text{CO}}(A|D) = \min_{R_{D^c} \in \mathcal{R}(A|D)} \sum_{i \in D^c} R_i \tag{13}
$$

where $\mathcal{R}(A|D)$ is given by (11). These CO rates can be achieved with noninteractive communication, and for achievable CO rates larger than $R_{\text{CO}}(A)$ (resp., $R_{\text{CO}}(A|D)$), $X_{\mathcal{M}}^n$ is achievable as ϵ_n-CR with exponentially vanishing ϵ_n.

Proposition 1 is a source coding theorem of the "Slepian–Wolf type," with the additional element that interactive communication is not *a priori* excluded, which makes the converse part nontrivial. While perhaps it has not been stated heretofore in the literature in exactly the same form, a related result appears in [14].

Theorem 1: The (strong) SK capacity $C_S(A)$ for a set of terminals $A \subset \mathcal{M}$ equals

$$
C_S(A) = H(X_{\mathcal{M}}) - R_{\text{CO}}(A) \tag{14}
$$

and can be achieved with noninteractive communication.

Remark: The proof of Theorem 1 below uses transmissions by all the terminals in $\mathcal{M}$ as well as omniscience at all the terminals in A, to show that the SK capacity can be achieved. These two features were also seen previously in Example 1. However, neither feature is necessary in order to achieve the SK capacity. For instance, in the case $m = 2$ with $\mathcal{M} = \{1, 2\}$, $A = \mathcal{M}$, the SK capacity $C_S(\mathcal{M}) = I(X_1 \wedge X_2)$ can be achieved by means of just a single transmission from either terminal 1 (at rate $H(X_1|X_2)$) or terminal 2 (at rate $H(X_2|X_1)$) (cf. [8], [1]); also, this enables omniscience at the receiving terminal but not at the

transmitting terminal. In the case $m = 3$ with $\mathcal{M} = \{1, 2, 3\}$, $A = \mathcal{M}$, and X_1, X_2, X_3 forming a Markov chain,

$$C_S(\mathcal{M}) = \min\{I(X_1 \wedge X_2), I(X_2 \wedge X_3)\}$$

(see Example 5 in Section V). This SK capacity can be achieved by a single transmission from terminal 2 (cf. [6, Theorem 2.4]); however, omniscience is not enabled at any terminal.

Theorem 2: The (strong) PK capacity $C_P(A|D)$ for a set of terminals $A \subset \mathcal{M}$, private from terminals $D \subset A^c$, equals

$$C_P(A|D) = H(X_{\mathcal{M}}|X_D) - R_{\mathrm{CO}}(A|D) \qquad (15)$$

and can be achieved with noninteractive communication.

Next, to consider the effect of randomization, assume that terminal i generates an rv M_i, $i \in \mathcal{M}$, such that $M_1, \ldots, M_m$ and $X_{\mathcal{M}}^n$ are mutually independent. The notions of SK, WSK, and PK capacities with randomization at the terminals can be formulated as above with (X_i^n, M_i) in the role of X_i^n, $i = 1, \ldots, m$.

Theorem 3: The SK capacity $C_S(A)$, $A \subset \mathcal{M}$, and the PK capacity $C_P(A|D)$, $A \subset \mathcal{M}$, $D \subset A^c$, are not increased by randomization at the terminals.

The following result concerning WSK capacity is a consequence of Theorem 2. To formulate this result, introduce a dummy terminal $m + 1$, and assign to it the rv Z whose i.i.d. repetitions are accessible to the eavesdropper. Denote by $C_P(A|Z)$ the PK capacity for the set of terminals A private from the dummy terminal, when $\mathcal{M} = \{1, \ldots, m\}$ and $X_{\mathcal{M}} = (X_1, \ldots, X_m)$ are replaced by $\{1, \ldots, m + 1\}$ and $(X_1, \ldots, X_m, Z)$; a single-letter characterization of $C_P(A|Z)$ is provided by Theorem 2.

Theorem 4: $C_P(A|Z)$ is always an upper bound for $C_W(A)$, and so is $C_P(A|V)$ for any rv V which is conditionally independent of $X_{\mathcal{M}}$ when conditioned on Z. If $Z = (Z_1, \ldots, Z_m)$, with the conditional independence property (4), then $C_W(A) = C_P(A|Z)$ provided that randomization is permitted at the terminals.

Remark: By Theorem 3, the SK capacity and the PK capacity remain unaffected upon permitting randomization at the terminals. The question as to whether the same applies to the WSK capacity, too, shall not be addressed here.

Example 2 (Special Case $m = 2$ of Theorem 4): To specialize Theorem 4 to the case $\mathcal{M} = \{1, 2\}$, $A = \mathcal{M}$, consider the augmented set of terminals, viz. $\{1, 2, 3\}$, with Z being assigned to the dummy terminal 3. Then the set in (11) with $D = \{3\}$ is determined by the two SW constraints

$$R_1 \geq H(X_1|X_2, Z), \qquad R_2 \geq H(X_2|X_1, Z)$$

whereby (15) gives

$$C_P(A|Z) = H(X_1, X_2|Z) - H(X_1|X_2, Z) - H(X_2|X_1, Z)$$
$$= I(X_1 \wedge X_2|Z).$$

Hence, Theorem 4 specialized to this case gives that the WSK capacity is bounded above by the infimum of $I(X_1 \wedge X_2|V)$ for

rvs V which are conditionally independent of (X_1, X_2) when conditioned on Z. This bound for the case $m = 2$, obtained in [1], has been the best one known until recently (see [12]). Moreover, while the WSK capacity for $m = 2$ is known to equal $I(X_1 \wedge X_2|Z)$ if X_1, X_2, Z form a Markov chain in some order (see [8], [1]), Theorem 4 implies the same equality also when Z_1, X_1, X_2, Z_2 is a Markov chain in this order, and $Z = (Z_1, Z_2)$ (provided randomization is permitted at the terminals); this latter case is apparently not covered by previous results. $\square$

Finally, while Theorems 1 and 2 give single-letter characterizations of SK and PK apacities, they may not lend themselves to a straightforward determination of whether a capacity in question is positive. The next result, whose proof is independent of Theorems 1 and 2, provides necessary and sufficient conditions for zero capacity, namely, the existence of "independent partitions."

Theorem 5:
i) For $A \subset \mathcal{M}$, the SK capacity $C_S(A) = 0$ iff $I(X_B \wedge X_{B^c}) = 0$ for some $B \subset \mathcal{M}$ with $B \cap A \neq \emptyset$, $B^c \cap A \neq \emptyset$.
ii) For $A \subset \mathcal{M}$, $D \subset A^c$, the PK capacity $C_P(A|D) = 0$ iff $I(X_B \wedge X_{B^c}|X_D) = 0$ for some $B \subset D^c$ with $B \cap A \neq \emptyset$, $B^c \cap A \neq \emptyset$.

IV. PROOFS OF THEOREMS 1–5

The main technical tools are supplied by Lemma 2 that follows and Lemma B.3 in Appendix B, which are used to prove the converse and achievability parts, respectively, of both Theorems 1 and 2. Lemma 2 is also used to prove the converse part of Proposition 1. Lemma 2 affords, in effect, a decomposition of the total joint entropy $H(X_{\mathcal{M}})$ into the normalized conditional entropy of any achievable ϵ-CR conditioned on the communication with which it is achieved, and a sum of rates which satisfy the appropriate SW conditions. Lemma B.3 guarantees, for i.i.d. repetitions (U^n, V^n) of a pair of rvs (U, V) and for a given function $g(U^n)$ of U^n, the existence of a function K of U^n with "large" range such that for this K and $Y = (g(U^n), V^n)$, the security index s_{in} defined in (6) decays to 0 exponentially rapidly as $n \to \infty$. This lemma follows from an extension of the "almost independence" result ([5, Theorem 1]), viz. Lemma B.2 in Appendix B, which, in turn, is a consequence of the "coloring lemma" in [2]. Lemma 2 is followed by the proof of the converse part of Proposition 1 (whose achievability part is proved in Appendix A), and then by the proofs of Theorems 1 and 2. The proof of Theorem 3 is then given, based on Lemma 2. Theorem 4 is proved next as a consequence of Theorem 3. Theorem 5 is established last, based on—as stated earlier—a different set of ideas from those used in the proofs of the earlier theorems.

Lemma 2: Given $X_{\mathcal{M}} = (X_1, \ldots, X_m)$, let $K = K(X_{\mathcal{M}}^n)$ be ϵ-CR (with values in a finite set $\mathcal{K}$) for a set of terminals $A \subset \mathcal{M}$, achievable with communication $\boldsymbol{F} = (F_1, \ldots, F_{rm})$. Then

$$\frac{1}{n}H(K|\boldsymbol{F}) = H(X_{\mathcal{M}}) - \sum_{i=1}^{m} R_i + \frac{m}{n}(\epsilon \log |\mathcal{K}| + 1) \quad (16)$$

for some $R_{\mathcal{M}} = (R_1, \ldots, R_m) \in \mathcal{R}(A)$ (see (10)).

Moreover, if for a set $D \subset A^c$, each terminal $i \in D$ transmits the entire X_i^n, then (16) holds with $\sum_{i=1}^m R_i$ replaced by $H(X_D) + \sum_{i \in D^c} R_i$ for a suitable $R_{D^c} \in \mathcal{R}(A|D)$ (see (11)).

Proof: Since $\mathbf{F} = (F_1, \ldots, F_{rm})$ and K are functions of $X_\mathcal{M}^n = (X_1^n, \ldots, X_m^n)$

$$H(X_\mathcal{M}^n) = H(F_1, \ldots, F_{rm}, K, X_1^n, \ldots, X_m^n)$$
$$= \sum_{\nu=1}^{rm} H(F_\nu | F_{[1, \nu-1]}) + H(K | \mathbf{F})$$
$$+ \sum_{i=1}^m H\left(X_i^n | \mathbf{F}, K, X_{[1, i-1]}^n\right).$$

Setting

$$R_i' = \frac{1}{n} \sum_{\nu: \nu \equiv i \bmod m} H(F_\nu | F_{[1, \nu-1]})$$
$$+ \frac{1}{n} H\left(X_i^n | \mathbf{F}, K, X_{[1, i-1]}^n\right) \quad (17)$$

the previous equality gives

$$\frac{1}{n} H(K | \mathbf{F}) = H(X_\mathcal{M}) - \sum_{i=1}^m R_i'. \quad (18)$$

Thus, $R_\mathcal{M}$ defined by

$$R_i = R_i' + \frac{\epsilon \log |\mathcal{K}| + 1}{n}$$

satisfies (16), and it remains to show that $R_\mathcal{M} \in \mathcal{R}(A)$. To that end

$$H(X_B^n | X_{B^c}^n) = H(F_1, \ldots, F_{rm}, K, X_B^n | X_{B^c}^n)$$
$$= \sum_{\nu=1}^{rm} H\left(F_\nu | F_{[1, \nu-1]}, X_{B^c}^n\right) + H\left(K | \mathbf{F}, X_{B^c}^n\right)$$
$$+ \sum_{i \in B} H\left(X_i^n | \mathbf{F}, K, X_{[1, i-1]}^n, X_{B^c \cap [i+1, m]}^n\right).$$

Since K is ϵ-CR for A, it is ϵ-recoverable from $\mathbf{F}$ and $X_{B^c}^n$ if $A \not\subset B$, and, hence, by Fano's inequality

$$H\left(K | \bar{\mathbf{F}}, X_{B^c}^n\right) \leq \epsilon \log |\mathcal{K}| + 1$$

for $A \not\subset B$. Further, since F_ν is a function of X_i^n and $F_{[1, \nu-1]}$ for $i \equiv \nu \bmod m$, only those terms $H\left(F_\nu | F_{[1, \nu-1]}, X_{B^c}^n\right)$ can be nonzero for which $i \equiv \nu \bmod m$, $i \in B$. Upon bounding these terms from above by omitting $X_{B^c}^n$, and similarly bounding the terms $H\left(X_i^n | \mathbf{F}, K, X_{[1, i-1]}^n, X_{B^c \cap [i+1, m]}^n\right)$ by omitting $X_{B^c \cap [i+1, m]}^n$, it follows that

$$H(X_B | X_{B^c}) \leq \sum_{i \in B} R_i' + \frac{1}{n}(\epsilon \log |\mathcal{K}| + 1) \leq \sum_{i \in B} R_i$$

whenever $A \not\subset B$. This proves that $R_\mathcal{M} \in \mathcal{R}(A)$.

In order to prove the second assertion of the lemma, assume without any loss of generality that $D = \{1, \ldots, k\}$. Then the assumption that each $i \in D$ transmits all of X_i^n means that $F_i = X_i^n$, $1 \leq i \leq k$, so that in (17), $R_i' = H(X_i | X_{[1, i-1]})$, $1 \leq i \leq k$. Then, (18) becomes

$$\frac{1}{n} H(K | \mathbf{F}) = H(X_\mathcal{M}) - H(X_{[1, k]}) - \sum_{i=k+1}^m R_i'$$
$$= H(X_\mathcal{M} | X_D) - \sum_{i \in D^c} R_i',$$

and the proof is completed as above. $\square$

Proof of Proposition 1: We prove the converse part here. The achievability part is easy to prove by standard techniques, and in Appendix A we show that it is a special case of a general theorem ([4, Theorem 3.1.14]) concerning normal source networks without helpers; this will produce a stronger result, which is also relevant for our purposes.

Suppose first that $X_\mathcal{M}^n$ is achievable as ϵ_n-CR for the terminals in $A \subset \mathcal{M}$ with (possibly interactive) communication $\mathbf{F}^{(n)}$, where $\epsilon_n \to 0$. By Lemma 2 applied to $K = X_\mathcal{M}^n$

$$\frac{1}{n} H\left(X_\mathcal{M}^n | \mathbf{F}^{(n)}\right) = H(X_\mathcal{M}) - \frac{1}{n} H(\mathbf{F}^{(n)})$$

is equal to

$$H(X_\mathcal{M}) - \sum_{i=1}^m R_i + \frac{m}{n} \left(\epsilon_n \log \prod_{i=1}^m |\mathcal{X}_i|^n + 1\right)$$

for some $R_\mathcal{M} \in \mathcal{R}(A)$. It follows that for this $R_\mathcal{M}$

$$\frac{1}{n} H(\mathbf{F}^{(n)}) = \sum_{i=1}^m R_i - m\left(\epsilon_n \sum_{i=1}^m \log |\mathcal{X}_i| + \frac{1}{n}\right). \quad (19)$$

As $\frac{1}{n} \log \|\mathbf{F}^{(n)}\| \geq \frac{1}{n} H(\mathbf{F}^{(n)})$, (19) proves the claim that

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

is a lower bound for achievable CO rates (see Definition 2).

Suppose next that $X_\mathcal{M}^n$ is achievable as ϵ_n-CR as above when each terminal $i \in D$ transmits all of X_i^n; then the total communication is

$$\mathbf{F}^{(n)} = \left(X_D^n, \widetilde{\mathbf{F}}^{(n)}\right)$$

where $\widetilde{\mathbf{F}}^{(n)}$ comprises the transmissions of the terminals $i \in D^c$. By the second part of Lemma 2, (16) holds in this case with $\sum_{i=1}^m R_i$ replaced by $H(X_D) + \sum_{i \in D^c} R_i$ for some $R_{D^c} \in \mathcal{R}(A|D)$, and, hence, (19) also holds with the same replacement. Since

$$\frac{1}{n} H(\mathbf{F}^{(n)}) \leq H(X_D) + \frac{1}{n} H(\widetilde{\mathbf{F}}^{(n)})$$
$$\leq H(X_D) + \frac{1}{n} \log \|\widetilde{\mathbf{F}}^{(n)}\|$$

it follows that

$$\frac{1}{n} \log \|\widetilde{\mathbf{F}}^{(n)}\| \geq \sum_{i \in D^c} R_i - m\left(\epsilon_n \sum_{i=1}^m \log |\mathcal{X}_i| + \frac{1}{n}\right)$$

for some $R_{D^c} \in \mathcal{R}(A|D)$. This establishes the claim that $\min_{R_{D^c} \in \mathcal{R}(A|D)} \sum_{i \in D^c} R_i$ is a lower bound for achievable CO rates when each terminal $i \in D$ transmits all of X_i^n. $\square$

Proof of Theorem 1:

Converse Part. Suppose that $K^{(n)}$ represents ϵ_n-SK for A, achievable with (possibly interactive) communication $\bar{\mathbf{F}}^{(n)}$, where $\epsilon_n \to 0$ (see Definition 1). Then (1) and Lemma 2 give that

$$\frac{1}{n} H(K^{(n)}) \leq \frac{1}{n} H(K^{(n)} | \mathbf{F}^{(n)}) + \epsilon_n$$
$$\leq H(X_\mathcal{M}) - \sum_{i=1}^m R_i$$
$$+ \frac{m}{n} \left(\epsilon_n \log |\mathcal{K}^{(n)}| + 1\right) + \epsilon_n \quad (20)$$

for some $R_\mathcal{M} \in \mathcal{R}(A)$. Here, $\sum_{i=1}^m R_i \geq R_{\mathrm{CO}}(A)$ by Proposition 1, and $\log |\mathcal{K}^{(n)}| = O(n)$ because $K^{(n)}$ is a function of $X_\mathcal{M}^n$. Hence, (20) shows that $H(X_\mathcal{M}) - R_{\mathrm{CO}}(A)$ is an upper bound for achievable SK rates, as claimed.

Achievability Part. We have to show that each $H < H(X_\mathcal{M}) - R_{\mathrm{CO}}(A)$ is a strongly achievable SK rate for the set of terminals A, achievable with noninteractive communication. To this end, fix an arbitrarily small $\delta > 0$ and note that, by Proposition 1, noninteractive communication $\boldsymbol{F}^{(n)} = \left(F_1^{(n)}, \ldots, F_m^{(n)} \right)$ with

$$\frac{1}{n} \log \|\boldsymbol{F}^{(n)}\| \leq R_{\mathrm{CO}}(A) + \delta \qquad (21)$$

suffices to achieve $X_\mathcal{M}^n$ as ϵ_n-CR for the set of terminals A, with exponentially vanishing ϵ_n. Of course, the same communication then makes any function of $X_\mathcal{M}^n$ an ϵ_n-CR. Hence, it suffices to prove the existence of a function K of $X_\mathcal{M}^n$ with range $\mathcal{K} = \{1, \ldots, \exp(nH)\}$ such that

$$s_{\mathrm{in}} = \log |\mathcal{K}| - H(K) + I(K \wedge \boldsymbol{F}^{(n)})$$

goes to 0 exponentially fast.

Recalling the bound (21) on the cardinality of the range of $\boldsymbol{F}^{(n)}$, the latter is a consequence of Lemma B.3 in Appendix B applied to $U = X_\mathcal{M}$ and $V = \mathrm{constant}$, with $g(U^n) = \boldsymbol{F}^{(n)}$, provided that $H < H(X_\mathcal{M}) - R_{\mathrm{CO}} - \delta$. Since $\delta > 0$ can be chosen arbitrarily small, this completes the proof. □

Proof of Theorem 2:

Converse Part. Let $K^{(n)}$ represent ϵ_n-PK which is private from $D \subset A^c$, achievable with (perhaps interactive) communication $\boldsymbol{F}^{(n)}$, where $\epsilon_n \to 0$. On account of the privacy condition (see (3) with $Z = X_D$), we can assume without any loss of generality that each terminal $i \in D$ transmits all of X_i^n; thus,

$$\boldsymbol{F}^{(n)} = \left(X_D^n, \widetilde{\boldsymbol{F}}^{(n)} \right)$$

where $\widetilde{\boldsymbol{F}}^{(n)}$ comprises the transmissions of the terminals $i \in D^c$. Then, by (3) and the second part of Lemma 2

$$\frac{1}{n} H(K^{(n)}) \leq \frac{1}{n} H\left(K^{(n)} | X_D^n, \widetilde{\boldsymbol{F}}^{(n)} \right) + \epsilon_n$$
$$= H(X_\mathcal{M}) - H(X_D) - \sum_{i \in D^c} R_i$$
$$+ \frac{m}{n} \left(\epsilon_n \log |\mathcal{K}^{(n)}| + 1 \right) + \epsilon_n \qquad (22)$$

for some $R_{D^c} \in \mathcal{R}(A|D)$. Here, $\sum_{i \in D^c} R_i \geq R_{\mathrm{CO}}(A|D)$ by Proposition 1, and $\log |\mathcal{K}^{(n)}| = O(n)$. Hence, it follows that $H(X_\mathcal{M}|X_D) - R_{\mathrm{CO}}(A|D)$ is an upper bound for achievable rates, as claimed.

Achievability Part. The proof is almost identical to that of the achievability part of Theorem 1, but since each terminal $i \in D$ completely reveals X_i^n, the communication that we now have to consider is

$$\widetilde{\boldsymbol{F}}^{(n)} = \left\{ F_i^{(n)}, i \in D^c \right\}.$$

Accordingly, the second part of Proposition 1 has to be invoked for the achievability of $X_\mathcal{M}^n$ as ϵ_n-CR with exponentially vanishing ϵ_n, using such communication of rate

$$\frac{1}{n} \log \|\widetilde{\boldsymbol{F}}^{(n)}\| \leq R_{\mathrm{CO}}(A|D) + \delta.$$

Moreover, to show the existence of a function K of $X_\mathcal{M}^n$ with the desired secrecy property, Lemma B.3 has to be applied to $U = X_{D^c}, V = X_D$, with $g(U^n) = \widetilde{\boldsymbol{F}}^{(n)}$ and

$$Y = (g(U^n), V^n) = \left(\widetilde{\boldsymbol{F}}^{(n)}, X_D^n \right).$$

The details are omitted. □

Proof of Theorem 3: Considering first the case of SK capacity, let $K^{(n)} = K^{(n)}(X_\mathcal{M}^n, M_\mathcal{M})$ represent ϵ_n-SK for A, achievable with communication $\boldsymbol{F} = \boldsymbol{F}^{(n)}$ when randomization is permitted, where $\epsilon_n \to 0$. Proceeding exactly as in the proof of the converse part of Theorem 1, it is seen that $K^{(n)}$ still satisfies (20) where $\sum_{i=1}^m R_i \geq R_{\mathrm{CO}}(A)$, as (16) remains unchanged in the present case. (Precisely, this can be seen from the proof of Lemma 2 by replacing the previous choice of R_i' by

$$R_i' = \frac{1}{n} \sum_{\nu : \nu \equiv i \bmod m} H(F_\nu | F_{[1,\nu-1]})$$
$$+ \frac{1}{n} H\left(X_i^n, M_i | \boldsymbol{F}, K^{(n)}, X_{[1,i-1]}^n, M_{[1,i-1]} \right) - \frac{1}{n} H(M_i)$$

in this case.) To complete the proof note that the observation following (20) concerning $|\mathcal{K}^{(n)}|$ is not valid in the present case as no prior assumption is made on the cardinality of the range of the rv $M_\mathcal{M}$. However, using (2) in lieu, we get from (20) that

$$(1 - m\epsilon_n) \frac{1}{n} H(K^{(n)})$$
$$\leq H(X_\mathcal{M}) - R_{\mathrm{CO}}(A) + m\epsilon_n^2 + \frac{m}{n} + \epsilon_n. \quad (23)$$

Hence, $H(X_\mathcal{M}) - R_{\mathrm{CO}}(A)$ remains an upper bound for all achievable SK rates with randomization at the terminals.

Turning to PK capacity, we proceed as in the proof of the converse part of Theorem 2, noting that by (3) with $Z = X_D$ and the second part of Lemma 2, (22) is satisfied where $\sum_{i \in D^c} R_i \geq R_{\mathrm{CO}}(A|D)$. The proof is then completed along the same lines as that above for SK capacity. □

Proof of Theorem 4: The upper bound $C_W(A) \leq C_P(A|Z)$ is obvious from the definitions. If V is conditionally independent of $X_\mathcal{M}$ given Z, then the secrecy condition (3) implies the same for V in the role of Z. Hence, if K is an ϵ-WSK when the eavesdropper wiretaps Z^n, it is also an ϵ-WSK when V^n is wiretapped, and hence the upper bound above holds also when Z is replaced by V.

Suppose next that $Z = Z_\mathcal{M} = (Z_1, \ldots, Z_m)$, with the conditional independence property (4). Then the terminals can simulate the wiretapped sources, provided that they are allowed to randomize. Using randomizing rvs $M_1, \ldots, M_m$ (with $M_1, \ldots, M_m, X_\mathcal{M}^n$ being mutually independent), each terminal $i \in \mathcal{M}$ can generate and reveal an rv $Y_i^n = g_i(X_i^n, M_i)$ such that the joint pmf of (X_i^n, Y_i^n) equals that of (X_i^n, Z_i^n). Then, using the assumption (4), the joint pmf of $(X_\mathcal{M}^n, Y_\mathcal{M}^n)$ is the same as that of $(X_\mathcal{M}^n, Z_\mathcal{M}^n)$. In particular, the secrecy condition (3) holds for $Z^n = Z_\mathcal{M}^n$ iff it does so for $Y_\mathcal{M}^n$ in the role of $Z_\mathcal{M}^n$.

This means that K is an ϵ-WSK for a set of terminals $A \subset \mathcal{M}$ iff in the model with $\mathcal{M}$ augmented by a dummy terminal to which $Y_{\mathcal{M}}$ is assigned (see the passage preceding Theorem 4), K is an ϵ-PK for A which is private from the dummy terminal. Since obviously $C_P(A|Y_{\mathcal{M}}) = C_P(A|Z_{\mathcal{M}})$, this completes the proof. $\square$

Proof of Theorem 5:

Sufficiency in i), ii). The sufficiency of the existence of "independent partitions" for zero capacity is easy. For any $B \subset \mathcal{M}$ with $B \cap A \neq \emptyset$, $B^c \cap A \neq \emptyset$, consider a consolidated model with two terminals which observe X_B^n and $X_{B^c}^n$, respectively. Clearly, $C_S(A)$ cannot exceed the SK capacity (without helper) for the consolidated model. The latter equals $I(X_B \wedge X_{B^c})$ (cf. [8], [1]). Hence, $C_S(A) = 0$ whenever $I(X_B \wedge X_{B^c}) = 0$ for some $B \subset \mathcal{M}$ with $B \cap A \neq \emptyset$, $B^c \cap A \neq \emptyset$.

Similarly, $C_P(A|D)$ cannot exceed the PK capacity for the consolidated model (with $B \subset D^c$) with the users maintaining additional secrecy from a wiretapped helper observing X_D^n. The PK capacity in the latter situation equals $I(X_B \wedge X_{B^c}|X_D)$ (cf. [1], [6]). The asserted sufficiency follows.

Necessity in i), ii). To prove the necessity of the existence of an independent partition for zero capacity, we give a sufficient condition for positive capacity and show that this condition, in turn, is implied by the nonexistence of independent partitions.

Consider first the case of the SK capacity. Let $\mathcal{G}$ be a graph with vertex set $\mathcal{M}$, where (i,j) forms an edge iff

$$I(X_i \wedge X_j|X_{\tilde{A}}) > 0$$

for some $\tilde{A} \subset \mathcal{M}\backslash\{i,j\}$ (possibly with $\tilde{A} = \emptyset$). For such $i,j \in \mathcal{M}$, clearly an SK can be generated with the help of $\tilde{A}$ at rate not less than $I(X_i \wedge X_j|X_{\tilde{A}})$, by interpreting $\tilde{A}$ as comprising wiretapped helpers whereby they completely reveal their respective source outputs (cf. [1], [6]). (It is immaterial here that this SK will be concealed also from $\tilde{A}$.) Then, the connectedness of the vertices in A is sufficient for $C_S(A) > 0$, which is seen as follows. If (i,j) and (i,j'), $j \neq j'$, are edges of $\mathcal{G}$ with common vertex i, then i, j and i, j' can, respectively, generate two independent SKs of 1 bit, say (possibly with help as above), in nonoverlapping time intervals, followed by i communicating to j, j' the binary sum of these SKs thereby enabling j (resp., j') to recover the SK for i, j' (resp., i, j). Any one of these two SKs constitutes a valid SK for i, j, j'. This means of SK propagation will enable all the vertices in A—by their connectedness—to share an SK, so that $C_S(A) > 0$.

It suffices to show that if the vertices in A are not connected, then an independent partition exists. Suppose that the vertex $l \in A$ is not connected to every vertex in A. Let B be the set of vertices in $\mathcal{M}$ which are connected to l. Note that $B \cap A \neq \emptyset$, $B^c \cap A \neq \emptyset$. Then X_B and X_{B^c} form an independent partition. Specifically, $I(X_B \wedge X_{B^c})$ can be represented as the sum of conditional mutual information terms of the form $I(X_{l'} \wedge X_{k'}|X_{\tilde{A}_{l'k'}})$, where $l' \in B$, $k' \in B^c$, $\tilde{A}_{l'k'} \subset \mathcal{M}\backslash\{l',k'\}$. Since (l',k') cannot be an edge of $\mathcal{G}$ if $l' \in B$, $k' \in B^c$, all these terms in the sum equal zero.

The case of the PK capacity is similar with the following modifications. We now define the graph $\mathcal{G}$ with vertex set D^c, with (i,j) forming an edge iff $I(X_i \wedge X_j|X_D, X_{\tilde{A}}) > 0$ for some

$\tilde{A} \subset D^c\backslash\{i,j\}$ (possibly with $\tilde{A} = \emptyset$). For such $i,j \in D^c$, a PK can be generated with the help of $\tilde{A}$ and private from D at rate not less than $I(X_i \wedge X_j|X_D, X_{\tilde{A}})$; this is seen by interpreting $\tilde{A}$ too as comprising wiretapped helpers (cf. [1], [6]). (The resulting additional privacy from $\tilde{A}$ as well is immaterial here.) Then the connectedness of the vertices in A is sufficient for $C_P(A|D) > 0$. The proof is completed by arguing along the same lines as for the SK capacity above; the details are omitted. $\square$

V. COMPUTATION OF SECRECY CAPACITIES

The single-letter characterizations of the SK and PK capacities in Theorems 2 and 3 reduce their computation to linear programming problems. In this context, the duality theorem of linear programming ([13, pp. 90–96]) is relevant. The version stated in this section is, in effect, the same as [13, Corollary 7.11], and is formally obtained from the latter by replacing $\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$ by $-\boldsymbol{A}$, $-\boldsymbol{b}$, $-\boldsymbol{c}$.

Duality Theorem: Let $\boldsymbol{A}$ be a $k \times m$ matrix, $\boldsymbol{b}$ a k-dimensional column vector, and $\boldsymbol{c}$ an m-dimensional row vector, such that the minimum of $\boldsymbol{cx}$ subject to $\boldsymbol{Ax} \geq \boldsymbol{b}$ is finite. Then this minimum equals the maximum of $\boldsymbol{yb}$ subject to $\boldsymbol{y} \geq 0$, $\boldsymbol{yA} = \boldsymbol{c}$, and that maximum is attained for a row vector $\boldsymbol{y}$ whose positive components correspond to linearly independent rows of $\boldsymbol{A}$.

In order to apply this theorem to compute $C_S(A)$, let $\boldsymbol{A}$ be the incidence matrix of the family of sets $B \subset \mathcal{M}$ that do not contain A, i.e., the rows of $\boldsymbol{A}$ are the incidence vectors of the sets B as above. Let the components of $\boldsymbol{b}$ be the conditional entropies $H(X_B|X_{B^c})$ for B as above, and let $\boldsymbol{c} = (1, \ldots, 1)$ (with m components). Then the duality theorem gives that $R_{\mathrm{CO}}(A)$, the minimum of $\boldsymbol{cx}$ subject to $\boldsymbol{Ax} \geq \boldsymbol{b}$, is equal to the maximum of linear combinations $\sum \lambda_B H(X_B|X_{B^c})$, with positive weights λ_B, for collections of sets B as above whose incidence vectors are linearly independent and whose linear combinations, with the weights λ_B, equal $(1, \ldots, 1)$. In particular, such a collection consists of no more than m sets. The duality theorem can be similarly applied to the computation of PK capacity, which requires a determination of $R_{\mathrm{CO}}(A|D)$; in that case, the role of m above will be played by the cardinality of D^c.

Example 3 (SK capacity for $\mathcal{M} = \{1,2,3\}$, $A = \mathcal{M}$): To determine the minimum $R_{\mathrm{CO}}(A)$ of $R_1 + R_2 + R_3$ subject to the SW constraints (cf. (10))

(i)	R_1	$\geq$	$H(X_1|X_2, X_3)$
(ii)	R_2	$\geq$	$H(X_2|X_1, X_3)$
(iii)	R_3	$\geq$	$H(X_3|X_1, X_2)$
(iv)	$R_1 + R_2$	$\geq$	$H(X_1, X_2|X_3)$
(v)	$R_1 + R_3$	$\geq$	$H(X_1, X_3|X_2)$
(vi)	$R_2 + R_3$	$\geq$	$H(X_2, X_3|X_1)$

the collections of (no more than three) subsets of $\mathcal{M}$ to be considered are: the pairs $\{B, B^c\}$ for B equal to $\{1\}$, $\{2\}$, or $\{3\}$; and the three-tuple $\{\{1,2\}, \{1,3\}, \{2,3\}\}$; the weights λ are all 1 for the first three collections and are all $\frac{1}{2}$ for the last collection. Formally, the collection $\{\{1\}, \{2\}, \{3\}\}$ also comes into account with weights 1; however, adding the inqualities (i), (ii), (iii) gives a weaker bound for $R_1 + R_2 + R_3$ than adding either

(i), (vi) or (ii), (v) or (iii), (iv). Thus, the duality theorem yields that $R_{\mathrm{CO}}(A)$ is equal to the maximum of

$$H(X_1|X_2, X_3) + H(X_2, X_3|X_1)$$
$$H(X_2|X_1, X_3) + H(X_1, X_3|X_2)$$
$$H(X_3|X_1, X_2) + H(X_1, X_2|X_3)$$

and

$$\frac{1}{2}[H(X_1, X_2|X_3) + H(X_1, X_3|X_2) + H(X_2, X_3|X_1)].$$

It follows by simple algebra that

$$C_S(\mathcal{M}) = H(X_1, X_2, X_3) - R_{\mathrm{CO}}(\mathcal{M})$$

is equal to the minimum of

$$I(X_1 \wedge X_2, X_3)$$
$$I(X_2 \wedge X_1, X_3)$$
$$I(X_3 \wedge X_1, X_2)$$

and

$$\frac{1}{2}[H(X_1) + H(X_2) + H(X_3) - H(X_1, X_2, X_3)].$$

The SK capacity $C_S(A)$ for $|A| = 2$, and the PK capacity $C_P(A|D)$ for $|A| = 2$, $D = A^c$, can also be computed similarly. (However, the latter secrecy capacities have been found earlier in [6], [1], unlike $C_S(\mathcal{M})$ above.) □

Example 4 (Upper bounds for SK capacity for arbitrary $\mathcal{M}$ and A): Consider a partition $\mathcal{B} = (B_1, \ldots, B_k)$ of $\mathcal{M} = \{1, \ldots m\}$ into (disjoint) sets, each intersecting A, where $2 \le k \le |A|$. Then the SW constraint

$$\sum_{j \in B} R_j \ge H(X_B|X_{B^c}) = H(X_{\mathcal{M}}) - H(X_{B^c}) \quad (24)$$

with the choice $B = B_i^c$ appears among those defining $\mathcal{R}(A)$ (see (10)). Substituting $B = B_i^c$ in (24) and adding these inequalities for $1 \le i \le k$ yields

$$(k-1)\sum_{j=1}^{m} R_j \ge kH(X_{\mathcal{M}}) - \sum_{i=1}^{k} H(X_{B_i})$$

as exactly one of the sets B_i and, hence, exactly $k - 1$ of the sets B_i^c contain any particular $j \in \mathcal{M}$. Therefore,

$$R_{\mathrm{CO}}(A) \ge H(X_{\mathcal{M}}) - \frac{1}{k-1}\left[\sum_{i=1}^{k} H(X_{B_i}) - H(X_{\mathcal{M}})\right].$$

Denoting by $I_k(A)$ the minimum of

$$\sum_{i=1}^{k} H(X_{B_i}) - H(X_{\mathcal{M}}) = D(P_{X_{\mathcal{M}}} \| P_{X_{B_1}} \times \cdots \times P_{X_{B_k}}) \tag{25}$$

over all partitions $\mathcal{B} = (B_1, \ldots, B_k)$ whose atoms B_i intersect A, we obtain that $C_S(A) = H(X_{\mathcal{M}}) - R_{\mathrm{CO}}(A)$ satisfies

$$C_S(A) \le \min_{2 \le k \le |A|} \frac{1}{k-1} I_k(A). \tag{26}$$

Note that the bound (26) is tight for the rvs in Example 1, where $I_{|A|}(A) = 1$, and where it is shown by an explicit construction that $C_S(A) \ge \frac{1}{|A|-1}$. Moreover, (26) is always tight for $m = 3$, $A = \mathcal{M}$; see Example 2. It is tempting to conjecture that the bound (26) is always tight. Indeed, (25) is commonly

interpreted as a measure of the mutual dependence of the rvs $X_{B_i}, 1 \le i \le k$, and hence $I_k(A)$ measures the minimum mutual dependence associated with admissible k-partitions of $X_{\mathcal{M}}$. It would be heuristically appealing if the SK capacity could always be expressed in terms of such dependence measures. Still, it is not expected that the conjecture above will hold true. By the duality theorem, $R_{\mathrm{CO}}(A)$ equals the maximum lower bound for $\sum_{i=1}^{m} R_i$ yielded by any collection of sets $B \not\supseteq A$ with linearly independent incidence vectors admitting a linear combination (with positive weights) equal to $(1, \ldots, 1)$. When m is large, there are several such collections other than those comprising the complements of atoms of a partition of $\mathcal{M}$, and no particular reason is apparent for the best lower bound for $\sum_{i=1}^{m} R_i$ to be always provided by one of the latter special collections. However, we have not delved into this issue, which remains open even for the case $m = 4$. □

We close this section with additional examples of interest in which the various secrecy capacities are computed.

Example 5 (SK capacities when $X_1, \ldots, X_m$ is a Markov Chain:
i) Consider first the case $A = \mathcal{M}$. Since

$$C_S(\mathcal{M}) \le I(X_1, \ldots, X_i \wedge X_{i+1}, \ldots, X_m)$$

for each $1 \le i \le m - 1$, and since by the assumed Markovity the mutual information term equals $I(X_i \wedge X_{i+1})$, obviously

$$C_S(\mathcal{M}) \le \min_{1 \le i \le m-1} I(X_i \wedge X_{i+1}). \tag{27}$$

We shall show, in fact, that (27) holds with equality. To this end, on account of Theorem 1, it suffices to find $(R_1, \ldots, R_m) \in \mathcal{R}(\mathcal{M})$ such that

$$H(X_{\mathcal{M}}) - \sum_{i=1}^{m} R_i = \min_{1 \le i \le m-1} I(X_i \wedge X_{i+1})$$
$$= I(X_t \wedge X_{t+1}) \tag{28}$$

say. We claim that

$$R_i = \begin{cases} H(X_i|X_{i+1}), & \text{if } i \le t \\ H(X_i|X_{i-1}), & \text{if } i > t \end{cases} \tag{29}$$

is a suitable choice. As the Markov assumption implies that

$$H(X_{\mathcal{M}}) = H(X_1, \ldots, X_t) + H(X_{t+1}, \ldots, X_m|X_t)$$
$$= H(X_1, \ldots, X_t|X_{t+1}) + I(X_t \wedge X_{t+1})$$
$$\quad + H(X_{t+1}, \ldots, X_m|X_t)$$
$$= \sum_{i=1}^{t} H(X_i|X_{i+1}) + \sum_{i=t+1}^{m} H(X_i|X_{i-1})$$
$$\quad + I(X_t \wedge X_{t+1}),$$

the choice (29) does satisfy (28). To show that (29) also satisfies the SW conditions determining the set $\mathcal{R}(\mathcal{M})$ in (10), i.e.,

$$\sum_{i \in B} R_i \ge H(X_B|X_{B^c}) \tag{30}$$

for each proper subset $B \subset \mathcal{M}$, suppose first that

$$B = \{l, l+1, \ldots, r\}.$$

Then, using the Markov property

$$H(X_B|X_{B^c}) = H(X_l, \ldots, X_r|X_{l-1}, X_{r+1}) \tag{31}$$

with the understanding that $X_0 = X_{m+1} = $ constant. As both

$$H(X_l, \ldots, X_r | X_{r+1}) = \sum_{i=l}^{r} H(X_i | X_{i+1})$$

and

$$H(X_l, \ldots, X_r | X_{l-1}) = \sum_{i=l}^{r} H(X_i | X_{i-1})$$

are upper bounds for the right-hand side of (31), it follows that (30) is certainly satisfied if $r \leq t$ or $l \geq t+1$. If $l \leq t$ and $r \geq t+1$, we bound the right-hand side of (31) by

$$
\begin{aligned}
H(X_l, &\ldots, X_r | X_{l-1}) \\
&= H(X_l, \ldots, X_t | X_{l-1}) + H(X_{t+1}, \ldots, X_r | X_t) \\
&= H(X_l, \ldots, X_t | X_{t+1}) + I(X_t \wedge X_{t+1}) \\
&\quad - I(X_{l-1} \wedge X_l) + H(X_{t+1}, \ldots, X_r | X_t)
\end{aligned}
$$

provided that $l > 1$. Then, $I(X_t \wedge X_{t+1}) \leq I(X_{l-1} \wedge X_l)$ implies that this is further bounded above by

$$\sum_{i=l}^{t} H(X_i | X_{i+1}) + \sum_{i=t+1}^{r} H(X_i | X_{i-1})$$

so that (30) is satisfied also in this case. The last argument does not work if $l = 1$, but then $r < m$ (since B is a proper subset of $\mathcal{M}$), and we can proceed similarly as above starting with the upper bound

$$
\begin{aligned}
H(X_l, \ldots, X_r | X_{r+1}) &= H(X_l, \ldots, X_t | X_{t+1}) \\
&\quad + H(X_{t+1}, \ldots, X_r | X_{r+1})
\end{aligned}
$$

for the right-hand side of (31).

Suppose next that $B \subset \mathcal{M}$ is not of the form $\{l, l+1, \ldots, r\}$. Then, it can be represented as

$$B = \bigcup_{j=1}^{s} B_j, \qquad B_j = \{l_j, l_j+1, \ldots, r_j\},$$

$$l_{j+1} \geq r_j + 2, \quad j = 1, \ldots, s-1$$

and the Markov property implies that

$$
\begin{aligned}
H(X_B | X_{B^c}) &= \sum_{j=1}^{s} H(X_{B_j} | X_{l_j-1}, X_{r_j+1}) \\
&= \sum_{j=1}^{s} H(X_{B_j} | X_{B_j^c}).
\end{aligned}
$$

Since (30) holds for each B_j, as we have just shown, it follows that it holds for $B = \bigcup_{j=1}^{s} B_j$ as well. This completes the proof of equality in (27).

ii) The result above also allows the determination of the SK capacity $C_S(A)$ for arbitrary $A \subset \mathcal{M}$. Denote the smallest and largest $i \in A$ by $i_{\min}$ and $i_{\max}$, respectively. If $i_{\min} = 1, i_{\max} = m$, then (27) holds also with $C_S(A)$ in the role of $C_S(\mathcal{M})$, for the same reason. Thus, in this case, $C_S(A) = C_S(\mathcal{M})$ (since, by definition $C_S(A) \geq C_S(\mathcal{M})$ always). In the general case, the obvious upper bound for $C_S(A)$, obtained similarly as in (27), is

$$C_S(A) \leq \min_{i_{\min} \leq i < i_{\max}} I(X_i \wedge X_{i+1}).$$

Here, again, equality must hold, since by the previous result, the right-hand side equals the SK capacity in the case where $\mathcal{M} = \{1, \ldots, m\}$ is replaced by $\{i_{\min}, i_{\min} + 1, \ldots, i_{\max}\}$; thus, it is achievable even with the helper terminals $i < i_{\min}$ and $i > i_{\max}$ not transmitting at all.

Hence, when $X_1, \ldots, X_m$ is a Markov chain, we always have that

$$C_S(A) = \min_{i_{\min} \leq i < i_{\max}} I(X_i \wedge X_{i+1}) \tag{32}$$

and only those helpers can actually help the terminals in A achieve this SK capacity whose indices fall between the smallest $i \in A$ and the largest $i \in A$. $\qquad \blacksquare$

Example 6 (PK and WSK capacities when $X_1, \ldots, X_m$ is a Markov Chain): For the PK capacity $C_P(A|D)$ to be positive, a necessary condition is that $D \cap [i_{\min}, i_{\max}] = \emptyset$, with $i_{\min}$ and $i_{\max}$ being as defined in Example 5. Indeed, if $i_{\min} \leq j \leq i_{\max}$ for some $j \in D$, then

$$C_P(A|D) \leq I\left(X_{[1,j-1] \cap D^c} \wedge X_{[j+1,m] \cap D^c} | X_D\right) = 0.$$

Subject to the necessary condition above, we have

$$C_P(A|D) = \min_{i_{\min} \leq i < i_{\max}} I(X_i \wedge X_{i+1} | X_D). \tag{33}$$

The proof is similar to that of (32), using Theorem 2 instead of Theorem 1. It suffices to consider the case when $D = [i_{\min}, i_{\max}]^c$, and then we can proceed as in Example 5, using instead of (29)

$$R_i = \begin{cases} H(X_i | X_{i+1}, X_{i_{\min}-1}), & \text{if } i \leq t \\ H(X_i | X_{i-1}, X_{i_{\max}+1}), & \text{if } i > t \end{cases}$$

for t achieving the minimum in (33). The details are omitted. We note that the PK capacity (33) depends on A and D only through $i_{\min}, i_{\max}$, and the elements of D closest to $i_{\min}$ from the left and to $i_{\max}$ from the right (if any).

The result (33) enables a determination also of the WSK capacity when the eavesdropper has access to i.i.d. repetitions of $Z = Z_1, Z_2$ such that $Z_1, X_1, \ldots, X_m, Z_2$ is a Markov chain. To see this, by Theorem 4, we consider an augmented model with $m + 2$ terminals to which are assigned the sources with generic rvs $Z_1, X_1, \ldots, X_m, Z_2$. The PK capacity $C_P(A|Z)$ for this latter model which is private from the dummy terminals to which Z_1 and Z_2 are assigned, is then equal to $C_W(A)$ (at least if this WSK capacity is defined with randomization permitted). Thus, in this case

$$C_W(A) = \min_{i_{\min} \leq i < i_{\max}} I(X_i \wedge X_{i+1} | Z_1, Z_2). \tag{34}$$

The WSK capacity (34) represents a new result even for the special case when $m = 2$ (see Example 2). $\qquad \blacksquare$

Example 7 (Markov chain on a tree): Consider a tree with vertex set $\mathcal{M} = \{1, \ldots, m\}$, i.e., a connected graph G containing no circuits. For (i, j) in the edge set $E(G)$ of G, let $B(i \leftarrow j)$ denote the set of all vertices connected with j by a path containing the edge (i, j). The rvs $X_1, \ldots, X_m$ form a Markov chain on the tree G (see [7]) if for each $(i, j) \in E(G)$, the conditional pmf of X_j given $X_{B(i \leftarrow j)}$ depends only on X_i. Note that when G is a chain, this concept reduces to that of a standard Markov chain.

The analog of (27) for this case is

$$C_S(\mathcal{M}) \leq \min_{(i,j)\in E(G)} I(X_i \wedge X_j)$$

a consequence of $C_S(\mathcal{M}) \leq I(X_{B(i\leftarrow j)}) \wedge I(X_{B(j\leftarrow i)})$ and the fact that the Markov property implies

$$I(X_{B(i\leftarrow j)}) \wedge I(X_{B(j\leftarrow i)}) = I(X_i \wedge X_j). \qquad (35)$$

Moreover, it can be seen similarly as in Example 5 that the bound above is tight. Take an edge (t, t^+) of G attaining $\min_{(i,j)\in E(G)} I(X_i \wedge X_j)$, and for $i \neq t$ let (i, i^+) be the first edge of the path in G that connects i with t. Then the analog of (28) with t^+ in the role of $t+1$ is satisfied by $R_i = H(X_i|X_{i^+})$, as can be seen from the consequence

$$H(X_{\mathcal{M}}) = H(X_{B(t\leftarrow t^+)}|X_{B(t^+\leftarrow t)})$$
$$+ H(X_{B(t^+\leftarrow t)}|X_{B(t\leftarrow t^+)}) + I(X_t \wedge X_{t^+})$$

of (35), decomposing the conditional entropies using the Markov assumption. These R_i's also satisfy the SW conditions (30) analogously as in Example 5. In that example, (30) was verified first with $B = \{l, l+1, \ldots, r\}$; now, the case to be verified first is with B being the vertex set of a subtree of G. The straightforward details are omitted.

The formula for $C_S(\mathcal{M})$ above enables us to readily obtain the following analog of the formula for $C_S(A)$ in (32):

$$C_S(A) = \min_{(i,j)\in E(G(A))} I(X_i \wedge X_j); \qquad (36)$$

here, $G(A)$ denotes the smallest subtree of G whose vertex set contains A.

The results of Example 6 also easily extend to the present case. In particular, a necessary condition for $C_P(A|D) > 0$ is that no $j \in D$ be a vertex of $G(A)$. Subject to that condition

$$C_P(A|D) = \min_{(i,j)\in E(G(A))} I(X_i \wedge X_j|X_D). \qquad (37)$$

Of particular interest is an explicit formula for the WSK capacity in the case when the eavesdropper has access to noisy versions of the sources with generic variables $X_1, \ldots, X_m$, i.e., to i.i.d. repetitions of $Z = Z_{\mathcal{M}} = (Z_1, \ldots, Z_m)$ satisfying the conditional independence assumption (4). Note in this case that the auxiliary model, with $2m$ terminals to which are assigned the rvs $X_1, \ldots, X_m, Z_1, \ldots, Z_m$, also represents a Markov chain on a tree; the underlying tree is obtained from G by adding m new edges, one from each $i \in \mathcal{M}$ to the corresponding dummy terminal to which Z_i is assigned. By Theorem 4, the WSK capacity in question is equal to the PK capacity $C_P(A|Z)$ for the auxiliary model. Hence, (37) gives

$$C_W(A) = \min_{(i,j)\in E(G(A))} I(X_i \wedge X_j|Z_{\mathcal{M}}). \qquad (38)$$

VI. DISCUSSION

We have derived single-letter characterizations of (strong) secrecy capacities for models with an arbitrary number of terminals, each of which observes a distinct component of a discrete memoryless multiple source, with unrestricted public communication allowed between these terminals. A subset of these terminals can serve as helpers which abet the remaining user terminals in generating secrecy. According to the degree of the eavesdropper's knowledge, three kinds of secrecy capacity are considered, viz. SK, PK, and WSK capacity. Our characterizations of the SK and PK capacities underscore the innate connections between secrecy generation and multiterminal source coding. In particular, the results indicate that the total joint entropy of the rvs describing the multiple source can be decomposed into the most parsimonious rate of communication among the terminals which enables the user terminals to become omniscient, i.e., reconstruct all the components of the multiple source, and the SK capacity that corresponds to (secret) CR which is concealed from an eavesdropper privy to the interterminal communication. When the eavesdropper wiretaps a subset of the helper terminals, a similar decomposition—in which the compromised helpers simply reveal their source components—yields the PK capacity. Of special practical interest is the situation in which the eavesdropper can wiretap noisy versions of the components of the underlying multiple source, in which case the previous result enables a characterization of the WSK capacity, at least when randomization is permitted at the terminals. This characterization yields a simple explicit formula in the case when the rvs of the multiple source form a Markov chain on a tree.

We have also derived a general upper bound for the WSK capacity, whose specialization to the case of two terminals coincides with the best upper bound known until recently in that case. Yet, a characterization of the WSK capacity, in general, remains unyielding; in particular, whether it has an analogous decomposition property as above is a question which merits further attention.

An aspect of our results, deliberately not emphasized in their statement for the sake of simplicity of exposition, is a form of universality, i.e., a lack of reliance on a knowledge of the joint pmf of the underlying rvs. Such an aspect in the context of CR generation without secrecy constraints, termed robust CR, has been previously studied in [2], and similar ideas apply in the present case too. Specifically, Proposition 1 and Theorems 1–3 admit the strengthening that given the rate-tuples $R_{\mathcal{M}}$ or R_{D^c}, and any $\delta > 0$, all the required mappings (in particular, the transmissions at the given rates) can be chosen in such a manner as to enable the corresponding assertions to hold regardless of the joint pmf of the underlying rvs, as long as for this joint pmf the appropriate SW conditions are satisfied even with δ added to their right sides. These strengthenings readily follow using Proposition 2 (Appendix A) and Lemma B.4 (Appendix B).

While our achievability proofs of Theorems 1–3 use transmissions by all the terminals in $\mathcal{M} = \{1, \ldots, m\}$ as well as omniscience at all the terminals in $A \subset \mathcal{M}$, neither feature is necessary for achieving the secrecy capacities. In this context, the least rate of communication between the terminals needed to achieve a secrecy capacity (perhaps without said omniscience) is of interest. This would be relevant for studying the situation in which rate constraints are imposed on permissible transmissions from the terminals, depicting bandwidth limitations associated with the use of shared public channels (cf. e.g., [6]). Also, in a cryptographic situation, it might be considered prudent to

curtail public communication to a minimum needed to generate secrecy.

Finally, the memorylessness assumed in the discrete multiple source is not essential for our results; in fact, they hold as well for the discrete stationary ergodic multiple source with only perfunctory changes. Specifically, Proposition 1 and Theorems 1–3 are now stated with all the relevant entropies and conditional entropies replaced by entropy rates and conditional entropy rates. Proposition 1 can be proved following the approach in [3], albeit without the universality and exponential decay in error probability asserted in Proposition 2 (Appendix A) under the memorylessness assumption. Theorems 1–3 also hold likewise, but with not necessarily strong secrecy capacities, and without universality.

APPENDIX A

In this appendix, we show that the achievability part of Proposition 1 is a special case of a general theorem concerning normal source networks without helpers. An important observation is that in the context of the achievability part, attention can be restricted to noninteractive communication.

In a normal source network without helpers [4], each component of a memoryless multiple source, with generic rvs $X_1, \ldots, X_m$, say, is connected to exactly one encoder, and each decoder which is required to decode certain component sources is connected to their encoders; these are the only connections in the network; in particular, interactive communication is excluded.

In the source network which underlies the problem of determining $R_{\mathrm{CO}}(A)$, the decoders are at the terminals $i \in A$, and each of them is connected to the encoders at the terminals $j \neq i$, and additionally directly to the ith source; see Fig. 1(a). While the latter is not a feature of the normal source network model, the $R_{\mathrm{CO}}(A)$ problem can be cast in the framework of a normal source network without helpers by the simple device of introducing dummy component sources with generic rvs $\tilde{X}_i = X_i$, $i \in A$; see Fig. 1(b). The dummy source $\tilde{X}_i$ is then connected to a dummy encoder whose encoding operation is the identity mapping, and this dummy encoder is connected only to the decoder at terminal i. Formally, let the component sources that the decoder at terminal $i \in A$ is required to decode be all X_j with $j \neq i$, and $\tilde{X}_i$; of course, as $\tilde{X}_i = X_i$, this is tantamount to omniscience for each $i \in A$.

By [4, Theorem 3.1.14], the achievable rate region of a normal source network without helpers consists of those m-tuples $(R_1, \ldots, R_m)$ that satisfy the SW conditions

$$\sum_{j \in L} R_j \geq H(X_L | X_{S_i \setminus L}) \tag{39}$$

for all decoders i and sets $L \subset S_i$, where S_i denotes the set of component sources which decoder i is required to decode. For the network involving the dummy sources as above, S_i, $i \in A$, consists of the indices $j \neq i$ and another index representing the
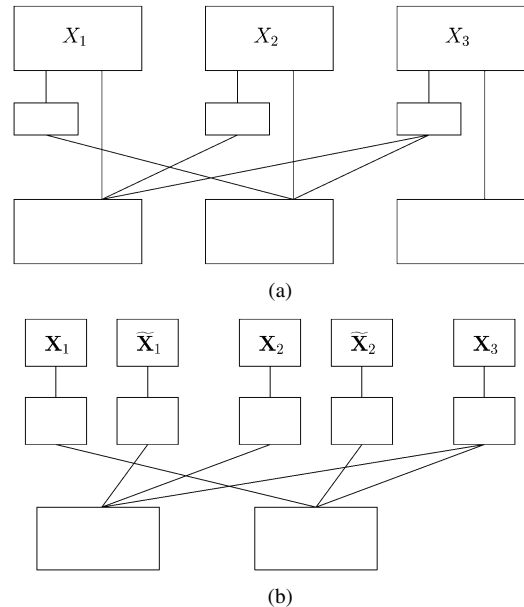

Fig. 1. (a) Source network with $m = 3$, $A = \{1, 2\}$. (b) Normal source network without helpers for Fig. 1(a).

dummy source $\tilde{X}_i$. For $L \subset S_i$ not containing the index of the dummy source $\tilde{X}_i$, the SW condition (39) reduces to

$$\sum_{j \in L} R_j \geq H(X_L | X_{\mathcal{M} \setminus L})$$

(as $\tilde{X}_i = X_i$ implies $X_{S_i \setminus L} = X_{\mathcal{M} \setminus L}$). If the previous condition is satisfied for some $L \subset S_i$, then so is the SW condition (39) for L' such that $X_{L'} = (X_L, \tilde{X}_i)$. This is obvious because

$$
\begin{aligned}
H(X_{L'} | X_{S_i \setminus L'}) &= H(\tilde{X}_i | X_{S_i \setminus L'}) + H(X_L | X_{S_i \setminus L'}, \tilde{X}_i) \\
&= H(\tilde{X}_i | X_{S_i \setminus L'}) + H(X_L | X_{\mathcal{M} \setminus L}).
\end{aligned}
$$

The source network underlying the problem of determining $R_{\mathrm{CO}}(A|D)$ is the same as in Fig. 1(a), with the understanding that at each terminal in D, the encoding operation is the identity mapping. In this case, the SW conditions need not be considered for sets B intersecting D, as their validity for sets $B \not\supset A$ contained in D^c immediately implies the same for all $B \not\supset A$.

We have thus demonstrated that the achievability part of Proposition 1 is a consequence of [4, Theorem 3.1.14]. More importantly, the strengthened assertions related to the latter concerning universally attainable error components ([4, pp. 267–268]) also apply to the omniscience problem at hand, yielding the following.

Proposition 2: Given any m-tuple $(R_1, \ldots, R_m)$, any $\delta > 0$, and sufficiently small $\eta > 0$, for sufficiently large n there exists a code for the source network underlying the $R_{\mathrm{CO}}(A)$ problem, with encoder rates not exceeding $R_j + \delta$, $j \in \mathcal{M}$, for which the error event that some of the decoders $i \in A$ do not correctly decode $X_{\mathcal{M}}^n$, has probability less than $\exp(-n\eta)$ whenever $X_{\mathcal{M}} = (X_1, \ldots, X_m)$ satisfies the SW conditions in (10). A similar statement is also true for the $R_{\mathrm{CO}}(A|D)$ problem.

Note that the code in Proposition 2 does not depend on the joint pmf of $X_{\mathcal{M}} = (X_1, \ldots, X_m)$, and is universally good

for all joint pmfs for which the SW conditions in (10) are satisfied. A similar statement holds for the code in the $R_{\mathrm{CO}}(A|D)$ problem.

APPENDIX B

In this appendix, we first prove Lemma 1, and then develop auxiliary results in a form which will directly serve to prove our achievability results.

Proof of Lemma 1

The first inequality is just the "Pinsker inequality" between variation distance and information divergence (cf. e.g., [4, p. 58]). The second inequality is proved using the bound

$$|H(P) - H(Q)| \le d(P,Q) \log \frac{s}{d(P,Q)} \qquad (40)$$

valid for arbitrary pmfs P, Q on a set of size $s \ge 4$, where $d(P,Q)$ denotes variation distance. Note that while this bound is stated in [4, p. 33], subject to $d(P,Q) \le 1/2$, the proof therein is valid whenever $d(P,Q) \le 1$, and (40) trivially holds when $d(P,Q) > 1$ since $d \log \frac{s}{d} \ge \log s$ if $s \ge 4$, $1 \le d \le 2$. Now, substituting in (40) for P the uniform pmf on $\mathcal{K}$, and for Q the conditional pmf $Q(\cdot|y)$ of K conditioned on $Y = y$, we obtain

$$\log |\mathcal{K}| - H(Q(\cdot|y)) \le \sum_{k \in \mathcal{K}} \left| Q(k|y) - \frac{1}{|\mathcal{K}|} \right|$$
$$\log \frac{|\mathcal{K}|}{\sum_{k \in \mathcal{K}} \left| Q(k|y) - \frac{1}{|\mathcal{K}|} \right|}. \qquad (41)$$

Multiplying both sides by $\Pr\{Y = y\}$ and summing over $y \in \mathcal{Y}$, the claim

$$s_{\mathrm{in}} \le s_{\mathrm{var}} \log \frac{|\mathcal{K}|}{s_{\mathrm{var}}}$$

follows by the definitions (6) and (7) of s_{in} and s_{var}, and Jensen's inequality. $\square$

The achievability proofs in this paper, similar as those in [5], [6], rely on the "coloring lemma" in [2]. Since the full force of that lemma is not needed here, we state below its assertion which we shall use.

Lemma B.1 ([2, Lemma 3.1]): Let $\mathcal{P}$ be any family of N pmfs on a finite set $\mathcal{U}$, and let $d > 0$ be such that every $P \in \mathcal{P}$ satisfies

$$P\left(\left\{u : P(u) > \frac{1}{d}\right\}\right) \le \epsilon, \qquad 0 < \epsilon < \frac{1}{9}. \qquad (42)$$

Then the probability that a randomly selected mapping $\phi : \mathcal{U} \to \{1,\ldots,k\}$ fails to satisfy

$$\sum_{i=1}^{k} \left| \sum_{u:\phi(u)=i} P(u) - \frac{1}{k} \right| < 3\epsilon \qquad (43)$$

simultaneously for each $P \in \mathcal{P}$, is less than $2Nk \exp\left(-\frac{\epsilon^2 d}{3k}\right)$.

Remark: This probability bound appears not in the statement of [2, Lemma 3.1] but in its proof, with a factor k erroneously missing. We are indebted to F. Matúš for pointing out this trivial error caused by the omission in the derivation of a multiplication by k [2 (p. 239, line 15 from below)]. Its correction does not adversely affect the applications in [2], [5], [6].

The following consequence of Lemma B.1 generalizes the "almost independence" result of [5, Theorem 1] in which the arbitrary function below did not appear. For a formal statement, we introduce for rvs U and V taking values in finite sets $\mathcal{U}$ and $\mathcal{V}$, respectively, with joint pmf P_{UV}, and for a given mapping $g : \mathcal{U} \to \{1,\ldots,l\}$, the following notation:

$$P(u|v) = \Pr\{U = u|V = v\}$$
$$= \frac{P_{UV}(u,v)}{P_V(v)} \qquad (44)$$

$$P(j,v) = \Pr\{g(U) = j, V = v\}$$
$$= \sum_{u:g(u)=j} P_{UV}(u,v) \qquad (45)$$

$$P(u|j,v) = \Pr\{U = u|g(U) = j, V = v\}$$
$$= \begin{cases} \frac{P_{UV}(u,v)}{P(j,v)}, & \text{if } j = g(u) \\ 0, & \text{otherwise.} \end{cases} \qquad (46)$$

Lemma B.2: Given U, V, and g as above, and $d > 0$ such that

$$P_{UV}\left(\left\{(u,v) : P(u|v) > \frac{1}{d}\right\}\right) \le \epsilon^2, \qquad 0 < \epsilon < 1/9 \qquad (47)$$

the probability that a randomly selected mapping $\phi : \mathcal{U} \to \{1,\ldots,k\}$ fails to satisfy

$$\sum_{j=1}^{l} \sum_{v \in \mathcal{V}} P(j,v) \sum_{i=1}^{k} \left| \sum_{u:\phi(u)=i} P(u|j,v) - \frac{1}{k} \right| < 7\epsilon \qquad (48)$$

is less than $2kl|\mathcal{V}| \exp\left(-\frac{\epsilon^3 d}{3kl}\right)$. In particular

$$kl \log(2kl|\mathcal{V}|) < \frac{1}{3}\epsilon^3 d \qquad (49)$$

is a sufficient condition for the existence of a mapping $\phi : \mathcal{U} \to \{1,\ldots,k\}$ satisfying (48).

Remark: The left-hand side of (48) is the variational security index s_{var} (see (7)), for $K = \phi(U)$, $Y = (g(U),V)$.

Proof: Using (45), (46), the hypothesis (47) can be equivalently written as

$$\sum_{j=1}^{l} \sum_{v \in \mathcal{V}} P(j,v) \sum_{u:P(u|v) > \frac{1}{d}} P(u|j,v) \le \epsilon^2 \qquad (50)$$

and, therefore, it implies, upon setting

$$D = \left\{(j,v) : \sum_{u:P(u|v) > \frac{1}{d}} P(u|j,v) \le \epsilon \right\} \qquad (51)$$

that

$$\sum_{(j,v) \in D^c} P(j,v) < \epsilon. \qquad (52)$$

Further, for (j, v) in the set

$$E = \left\{ (j, v) : P(j, v) \geq \frac{\epsilon}{l} P_V(v) \right\} \qquad (53)$$

it holds—using (46)—that $P(u|j, v) \leq \frac{l}{\epsilon} P(u|v)$, and consequently

$$\left\{ u : P(u|v) > \frac{1}{d} \right\} \supset \left\{ u : P(u|j, v) > \frac{l}{\epsilon d} \right\}.$$

This and (51) imply

$$\sum_{u : P(u|j, v) > \frac{l}{\epsilon d}} P(u|j, v) \leq \epsilon \text{ if } (j, v) \in D \cap E. \qquad (54)$$

Note also the obvious consequence of (53) that

$$\sum_{(j, v) \in E^c} P(j, v) \leq \epsilon. \qquad (55)$$

On account of (54), the family of pmfs $\{P(\cdot|j, v), (j, v) \in D \cap E\}$ satisfies the hypothesis (42) of Lemma B.1, with d replaced by $\frac{\epsilon d}{l}$. Since this family consists of at most $l|\mathcal{V}|$ pmfs, it follows by Lemma B.1 that a randomly selected $\phi : \mathcal{U} \rightarrow \{1, \ldots, k\}$ fails to make the inner sums in (48) corresponding to pairs $(j, v) \in D \cap E$ simultaneously less than 3ϵ, only with probability less than

$$2kl|\mathcal{V}| \exp \left(-\frac{\epsilon^3 d}{3kl} \right).$$

Bounding those inner sums in (48) that correspond to pairs $(j, v) \notin D \cap E$ trivially by 2, and using (52), (55), the assertion of Lemma B.2 follows. □

Lemma B.3: Given i.i.d. repetitions (U^n, V^n) of a pair of rvs (U, V), and a positive number $R < H(U|V)$, for any function $g(U^n)$ of U^n with range cardinality $\|g\| \leq \exp(nR)$, and positive number $H < H(U|V) - R$, there exists a function K of U^n with values in $\mathcal{K} = \{1, \ldots, \exp(nH)\}$ such that for this K and $Y = (g(U^n), V^n)$, the security index

$$s_{\text{in}} = \log |\mathcal{K}| - H(K) + I(K \wedge Y)$$

goes to 0 exponentially rapidly as $n \rightarrow \infty$.

Proof: Apply Lemma B.2 to (U^n, V^n) in the role of (U, V), with $d = \exp(n(H(U|V) - \delta))$, where $\delta > 0$ is sufficiently small.

With this choice, the hypothesis (47) of Lemma B.2 is obviously satisfied when n is sufficiently large, even with $\epsilon = \exp(-n\eta)$ if $\eta > 0$ is less than a threshold depending on δ. Hence, Lemma B.2 guarantees the existence of $\phi(U^n)$ with values in $\{1, \ldots, \exp(nH)\}$ such that for $K = \phi(U^n)$ and $Y = (g(U^n), V^n)$, the variational security index $s_{\text{var}} \leq 7 \exp(-n\eta)$, whenever $l = \exp(nR)$ and $k = \exp(nH)$ satisfy the condition (49) for L and ϵ as above, with $|\mathcal{V}|$ replaced by $|\mathcal{V}^n|$. Clearly,

when $H < H(U|V) - R$, this condition is satisfied for n sufficiently large, provided that $\delta > 0$ and $\eta > 0$ have been chosen sufficiently small.

Finally, the just-proved exponential convergence of s_{var} to 0 implies the same for s_{in} as well, by Lemma 1. □

In Lemma B.3, the function of U^n which serves as an appropriate choice of K, might depend not only on the given function g but also on the joint pmf of U, V. The next lemma improves upon the result by jettisoning the latter dependence.

Lemma B.4: Given the finite sets $\mathcal{U}, \mathcal{V}$, a function $g : \mathcal{U}^n \rightarrow \{1, \ldots, \exp(nR)\}$, and a positive number H, there exists a mapping $\phi : \mathcal{U}^n \rightarrow \{1, \ldots, \exp(nH)\}$ such that the security index

$$s_{\text{in}} = nH - H(\phi(U^n)) + I(\phi(U^n) \wedge g(U^n), V^n) \qquad (56)$$

goes exponentially to 0, uniformly for i.i.d. repetitions of any rvs (U, V) with

$$H(U|V) > R + H + \delta \qquad (57)$$

where $\delta > 0$ is arbitrarily small but fixed.

Proof: By Lemma 1, it suffices to prove the assertion for the security index s_{var} instead of s_{in}. The main idea of the proof is that for i.i.d. repetitions (U^n, V^n) of some pair of rvs (U, V), and given the mappings $\phi : \mathcal{U}^n \rightarrow \{1, \ldots, k\}$, $g : \mathcal{U}^n \rightarrow \{1, \ldots, l\}$, the security index

$$s_{\text{var}} = \sum_{i=1}^{k} \sum_{j=1}^{l} \sum_{\boldsymbol{v} \in \mathcal{V}^n} \Big| \Pr\{\phi(U^n) = i, g(U^n) = j, V^n = \boldsymbol{v}\}$$
$$- \frac{1}{k} \Pr\{g(U^n) = j, V^n = \boldsymbol{v}\} \Big| \qquad (58)$$

can be bounded above by an average of polynomially many terms equal to similar security indices with U^n, V^n replaced by pairs $\widetilde{U}, \widetilde{V}$ of rvs with values in $\mathcal{U}^n, \mathcal{V}^n$, uniformly distributed on joint type classes $\mathcal{T}_{\widetilde{P}}^n$, $\widetilde{P} \in \mathcal{P}_n(\mathcal{U} \times \mathcal{V})$. For notation and simple facts concerning types as used below, see, e.g., [4].

Formally, note that

$$\Big| \Pr\{\phi(U^n) = i, g(U^n) = j, V^n = \boldsymbol{v}\}$$
$$- \frac{1}{k} \Pr\{g(U^n) = j, V^n = \boldsymbol{v}\} \Big|$$
$$\leq \sum_{\widetilde{P} \in \mathcal{P}_n(\mathcal{U} \times \mathcal{V})} \Pr \left\{ (U^n, V^n) \in \mathcal{T}_{\widetilde{P}}^n \right\}$$
$$\Big| \Pr\left\{ \phi(U^n) = i, g(U^n) = j, V^n = \boldsymbol{v} | (U^n, V^n) \in \mathcal{T}_{\widetilde{P}}^n \right\}$$
$$- \frac{1}{k} \Pr \left\{ g(U^n) = j, V^n = \boldsymbol{v} | (U^n, V^n) \in \mathcal{T}_{\widetilde{P}}^n \right\} \Big|$$

and that the conditional probabilities here are the same as unconditional probabilities for $(\widetilde{U}, \widetilde{V})$, uniformly distributed on $\mathcal{T}_{\widetilde{P}}^n$, replacing (U^n, V^n). Thus, denoting by $s(\widetilde{P})$ the analog of (58)

when (U^n, V^n) is replaced by $(\widetilde{U}, \widetilde{V})$ uniformly distributed on the type class $\mathcal{T}_{\widetilde{P}}^n$, it follows that

$$s_{\text{var}} \leq \sum_{\widetilde{P} \in \mathcal{P}_n(\mathcal{U} \times \mathcal{V})} \Pr\left\{(U^n, V^n) \in \mathcal{T}_{\widetilde{P}}^n\right\} s(\widetilde{P}). \qquad (59)$$

Next, we apply Lemma B.2 to pairs $(\widetilde{U}, \widetilde{V})$ uniformly distributed on $\mathcal{T}_{\widetilde{P}}^n \subset \mathcal{U}^n \times \mathcal{V}^n$. Representing the type $\widetilde{P}$ as the joint pmf of dummy rvs X, Y (with values in $\mathcal{U}, \mathcal{V}$, respectively) with joint pmf $\widetilde{P}$, the conditional probabilities $P_{\widetilde{U}\widetilde{V}}(\boldsymbol{u}, \boldsymbol{v})/P_{\widetilde{V}}(\boldsymbol{v})$ for $(\widetilde{U}, \widetilde{V})$ uniformly distributed on $\mathcal{T}_{\widetilde{P}}^n = \mathcal{T}_{XY}^n$, constant when nonzero, are equal to $\exp(-nH(X|Y) + o(n))$. Hence, with $d = \exp(nL)$, the hypothesis (47) of Lemma B.2 is satisfied, independently of the value of ϵ, whenever

$$H(X|Y) \geq L + \frac{\delta}{4} \qquad (60)$$

say, and $n \geq n_0(\delta)$.

It follows that, given $g : \mathcal{U}^n \to \{1, \ldots, l\}$ with $l = \exp(nR)$, the probability that a randomly selected mapping $\phi : \mathcal{U}^n \to \{1, \ldots, k\}$ with $k = \exp(nH)$ fails to satisfy $s(\widetilde{P}) \leq 7\epsilon$ simultaneously for all type classes $\mathcal{T}_{\widetilde{P}}^n = \mathcal{T}_{XY}^n$ which meet (60), is less than

$$2|\mathcal{P}_n(\mathcal{U} \times \mathcal{V})|kl|\mathcal{V}^n| \exp\left(-\frac{\epsilon^3 d}{3kl}\right)$$
$$= 2|\mathcal{P}_n(\mathcal{U} \times \mathcal{V})| \exp\{n(H + R + \log|\mathcal{V}|)\}$$
$$\exp\left\{-\frac{\epsilon^3}{3} \exp[n(L - R - H)]\right\}.$$

The last bound goes (doubly exponentially) to 0 if $L = R + H + \frac{\delta}{2}$, say, even if ϵ is allowed to go to 0 exponentially with a small exponent. This proves that there exists $\phi : \mathcal{U}^n \to \{1, \ldots, \exp(nH)\}$ such that $s(\widetilde{P}) \to 0$ exponentially, simultaneously for all types $\widetilde{P} \in \mathcal{P}_n(\mathcal{U} \times \mathcal{V})$ represented by dummy

rvs X, Y satisfying $H(X|Y) \geq H + R + \frac{3\delta}{4}$. On account of (59), this completes the proof, since for (U, V) satisfying (57), the probability that (U^n, V^n) does not belong to a type class $\mathcal{T}_{\widetilde{P}}^n = \mathcal{T}_{XY}^n$ with $H(X|Y) \geq H + R + \frac{3\delta}{4}$, is exponentially vanishing uniformly for all possible (U, V). $\qquad \square$

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, July 1993.

[2] ——, "Common randomness in information theory and cryptography, part II: CR capacity," *IEEE Trans. Inform. Theory*, vol. 44, pp. 225–240, Jan. 1998.

[3] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 226–228, Mar. 1975.

[4] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1982.

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

[6] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. Inform. Theory*, vol. 46, pp. 344–366, Mar. 2000.

[7] H. O. Georgii, *Gibbs Measures and Phase Transitions*. Berlin, Germany: de Gruyter, 1988.

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

[9] ——, "The strong secret key rate of discrete random triples," in *Communications and Cryptography: Two Sides of One Tapestry*, R. E. Blahut *et al.*, Eds. Norwell, MA: Kluwer, 1994, ch. 26, pp. 271–285.

[10] U. Maurer and S. Wolf, "Privacy amplification secure against active adversaries," in *Advances in Cryptology-CRYPTO'97 (Lecture Notes in Computer Science)*, B. Kaliski, Ed. New York: Springer-Verlag, 1997, pp. 307–321.

[11] U. M. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Proc. EUROCRYPT 2000 (Lecture notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2000, pp. 352–368.

[12] R. Renner and S. Wolf, "New bounds in secret-key agreement: The gap between formation and secrecy extraction," in *Proc. EUROCRYPT 2003 (Lecture notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2003.

[13] A. Schrijver, *Theory of Linear and Integer Programming*. New York: Wiley, 1986.

[14] A. D. Wyner, J. K. Wolf, and F. M. J. Willems, "Communicating via a processing broadcast satellite," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1243–1249, June 2002.