Achieving SK Capacity in the Source Model: When Must All Terminals Talk?

Manuj Mukherjee[†] Navin Kashyap[†] Yogesh Sankarasubramaniam[‡]

***Abstract*—In this paper, we address the problem of characterizing the instances of the multiterminal source model of Csiszár and Narayan in which communication from *all* terminals is needed for establishing a secret key of maximum rate. We give an information-theoretic sufficient condition for identifying such instances. We believe that our sufficient condition is in fact an exact characterization, but we are only able to prove this in the case of the three-terminal source model. We also give a relatively simple criterion for determining whether or not our condition holds for a given multiterminal source model.**

I. Introduction

We are concerned with the multiterminal source model of Csiszár and Narayan [3], which can be briefly described as follows. There are a certain number, $m \geq 2$, of terminals, each of which observes a distinct component of a source of correlated randomness. The terminals must agree on a shared SK by communicating over a noiseless public channel. This key must be protected from a passive eavesdropper having access to the public communication. The SK capacity, which is the supremum of the rates of SKs that can be generated, has been characterized in various ways [2], [3], [7]. What is less well-understood is the nature of public communication that is needed to achieve SK capacity in this model. In a companion paper [6], we gave a lower bound on the minimum rate of communication required to generate a maximal-rate (i.e., capacity-achieving) SK, building upon the prior work of Tyagi [9] on the two-terminal model. In this paper, we address a related question: when must all m terminals necessarily have to communicate in order to generate a maximal-rate SK?

It is well known that, in order to generate a maximal-rate SK in the two-terminal model ($m = 2$), it is sufficient for only one terminal to communicate [1], [5], [3]. All this terminal has to do is convey its local observations to the other terminal at the least possible rate of communication required to do so. Thus, when $m = 2$, it is *never necessary* for both terminals to communicate to generate a capacity-achieving SK. Even when $m > 2$, there are examples wherein not all terminals need to communicate — see remark following Theorem 1 in [3]. However, as we will show in this paper, there are plenty of other examples where all terminals *must* communicate in order to achieve SK capacity. We coin the term "omnivocality" to describe the state when all terminals communicate. The problem of interest to us is the following: *characterize the instances of the multiterminal source model in which omnivocality is necessary for maximal-rate SK generation*. In this paper, we report partial progress made towards such a characterization.

The paper is organized as follows. After establishing the required notation and background in Section II, we give, in Section III, a sufficient condition under which omnivocality is necessary for achieving SK capacity in a source model with $m \geq 3$ terminals. This condition is satisfied, for example, in the case of the complete graph pairwise independent network (PIN) model of Nitinawarat and Narayan [7]. We conjecture that our sufficient condition is also necessary, but at present, we can only prove this in the $m = 3$ case. Finally, in Section IV, we give a useful criterion for checking whether or not our condition holds for a given source model.

II. Preliminaries

Throughout, we use $\mathbb{N}$ to denote the set of positive integers. In the multiterminal source model [3], a set of $m \geq 2$ terminals, denoted by $[m] \triangleq \{1, 2, \ldots, m\}$, has access to a source $(X_1^n, X_2^n, \ldots, X_m^n)$, $n \in \mathbb{N}$, where X_i^n denotes n i.i.d. copies of a random variable (rv) X_i taking values in a finite set $\mathcal{X}_i$. The rvs $X_1, X_2, \ldots, X_m$ are in general correlated, and for each $i \in [m]$, the ith terminal observes only the component X_i^n. For any subset $A \subseteq [m]$, we will use X_A to denote the collection of rvs $(X_i : i \in A)$, and p_{X_A} to denote their joint probability mass function.

The terminals communicate through a noiseless public channel, any communication sent through which is accessible to all terminals and to potential eavesdroppers as well. The terminals communicate in a round-robin fashion, following the cyclic order $(1, 2, \ldots, m)$. Any transmission sent by the ith terminal is a deterministic function of X_i^n and all the previous communication. Formally, a *valid communication* is a finitely-supported random vector $\mathbf{F} = (F_1, F_2, \ldots, F_r)$, $r \in \mathbb{N}$, with F_j denoting a communication sent by the terminal $i \in [m]$ with $i \equiv j \pmod m$, and $H(F_j \mid F_1, \ldots, F_{j-1}, X_i^n) = 0$. The *rate* of the communication is taken to be $\frac{1}{n}\log_2|\mathcal{F}|$, where $\mathcal{F}$ is the finite set on which $\mathbf{F}$ is supported. Terminal $i \in [m]$ is said to be *silent* if $F_j = 0$ (with probability 1) for all $j \equiv i \pmod m$. An *omnivocal* communication is one in which no terminal is silent.

Given an $\epsilon > 0$, we say that an rv U is *ϵ-recoverable* from an rv V if there exists a function g of V such that $\Pr[U = g(V)] \geq 1 - \epsilon$.

[†]M. Mukherjee and N. Kashyap are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore. Email: {manuj,nkashyap}@ece.iisc.ernet.in.

[‡]Email: yogesh@gatech.edu

Definition 1. *For any $\epsilon > 0$, an* ϵ-SK *for $[m]$ is an rv $K = K^{(n)}(X_{[m]}^n)$, for some $n \in \mathbb{N}$, such that there exists a valid communication $\mathbf{F}$ with the following properties:*
(i) $I(K;\mathbf{F}) \leq \epsilon$*; and*
(ii) K *is ϵ-recoverable from $(X_i^n, \mathbf{F})$ for each $i \in [m]$.*
The rate *of this ϵ-SK is given by $\frac{1}{n}H(K)$.*

A real number $R \geq 0$ is an *achievable SK rate* if for any $\epsilon > 0$, there exists an ϵ-SK of rate greater than $R - \epsilon$. The *SK capacity* $C([m])$ is defined as the supremum of all achievable SK rates. The SK capacity can be expressed as [2, Theorem 1.1] (see also [3, Eq. (26)])

$$C([m]) = \mathbf{I}(X_{[m]}) \triangleq \min_{\mathcal{P}} \frac{1}{|\mathcal{P}|-1} D\left(p_{X_{[m]}} \parallel \prod_{A\in\mathcal{P}} p_{X_A}\right) \tag{1}$$

the minimum being taken over all partitions $\mathcal{P}$ of $[m]$, of size $|\mathcal{P}| \geq 2$. The quantity $D(\cdot\|\cdot)$ denotes relative entropy, and for a partition $\mathcal{P} = \{A_1, \ldots, A_k\}$, the notation $\prod_{A\in\mathcal{P}} p_{X_A}$ represents the product $p_{X_{A_1}} \times \cdots \times p_{X_{A_k}}$. Note that when $m = 2$, the quantity $\mathbf{I}(X_{[m]})$ defined in (1) simply reduces to the mutual information $I(X_1; X_2)$. Thus, $\mathbf{I}(X_{[m]})$ should be viewed as a multiparty extension of mutual information.

Before proceeding further, a couple of clarifications concerning Definition 1 are needed. We have adopted the notion of *strong secrecy* (property (i) in the definition), as opposed to *weak secrecy*, which only requires $\frac{1}{n}I(K;\mathbf{F}) \leq \epsilon$. All the results proved in this paper would hold just as well under either type of secrecy. In particular, our main result shows that omnivocal communication is necessary for achieving SK capacity if a certain condition on the singleton partition $\mathcal{S}$ is satisfied. Our proof of this result relies only on the expression for SK capacity given in (1), which remains the same under both forms of secrecy [3], and on a theorem of Gohari and Anantharam [4], which is stated and proved under the weak secrecy notion. Thus, our proof in fact shows that omnivocal communication is necessary even under a weak secrecy requirement on SKs.

A second clarification concerning Definition 1 is that, usually, the definition of an ϵ-SK includes an additional requirement that K be almost uniformly distributed over its alphabet $\mathcal{K}$, i.e., $H(K) \geq \log|\mathcal{K}| - \epsilon$ [3]. However, this can always be dropped without affecting SK capacity — see e.g., the discussion on p. 3976 in [4].

As mentioned above, we make use of a result of Gohari and Anantharam [4, Theorem 6] in some of our proofs. To state this result, we explicitly define a *weak ϵ-SK* for $[m]$ to be an rv K as in Definition 1, except that the strong secrecy condition (i) is replaced by the weak secrecy condition, $\frac{1}{n}I(K;\mathbf{F}) \leq \epsilon$. Then, $R \geq 0$ is an *achievable weak-SK rate* if for any $\epsilon > 0$, there exists a weak ϵ-SK of rate greater than $R - \epsilon$. It is known that the supremum of achievable weak-SK rates is the same as the SK capacity given by (1). The Gohari-Anantharam result concerns achievable weak-SK rates under the additional assumption that some fixed subset of terminals remains silent throughout. Let $T \subseteq [m]$ be such that terminals in T are allowed to communicate, while terminals in $[m] \setminus T$ must remain silent. Thus, we are restricted to valid communications $\mathbf{F}$ in which the terminals in $[m]\setminus T$ are silent, but which allow all m terminals to agree upon a weak SK. In other words, we only consider weak ϵ-SKs for $[m]$ that are obtainable through valid communications $\mathbf{F}$ in which all terminals in $[m] \setminus T$ are silent. The supremum of rates achievable by such SKs will be denoted by $C([m]\|T)$.

Theorem 1 ([4], Theorem 6)**.** $C([m]\|T) = H(X_T) - R_T^{(\min)}$*, where* $R_T^{(\min)} = \min_{\mathbf{R}\in\mathcal{R}_T} \sum_{i\in T} R_i$*, the rate region $\mathcal{R}_T$ being the set of all points $\mathbf{R} = (R_i, i \in T)$ such that*

$$\sum_{i\in A\cap T} R_i \geq H(X_{A\cap T}|X_{A^c}) \quad \forall\, A \subsetneq [m],\ A \cap T \neq \emptyset.$$

Note that if $C([m]) > C([m]\|T)$ for all $T \subset [m]$ of size $|T| = m - 1$, then omnivocality is necessary for achieving SK capacity. Thus, our approach for showing that omnivocal communication is needed in certain cases is to use Theorem 1 to prove that $C([m]) > C([m]\|T)$ for all $(m-1)$-subsets $T \subset [m]$. For this, we will need a lower bound on $R_T^{(\min)}$ when $|T| = m - 1$. To prove this bound, we use a simpler characterization (than that given in Theorem 1) of the rate region $\mathcal{R}_T$ when $|T| = m - 1$.

Lemma 2. *Let $T = [m] \setminus \{u\}$ for some $u \in [m]$. The rate region $\mathcal{R}_T$ is the set of all points $(R_i,\, i \in T)$ such that*

$$\sum_{i\in B} R_i \geq H(X_B|X_{T\setminus B}) \quad \forall\, B \subsetneq T, B \neq \emptyset, \tag{2}$$

and $\sum_{i\in T} R_i \geq H(X_T|X_u)$.

Proof: Observe that $\mathcal{R}_T$ is defined by constraints on sums of the form $\sum_{i\in B} R_i$ for non-empty subsets $B \subseteq T$. When $B = T$, the constraint is simply $\sum_{i\in T} R_i \geq H(X_T|X_u)$.

Now, consider any non-empty $B \subsetneq T$. From Theorem 1, we see that constraints on $\sum_{i\in B} R_i$ arise as constraints on $\sum_{i\in A\cap T} R_i$ in two ways: when $A = B$ and when $A = B \cup \{u\}$. Thus, we have two constraints on $\sum_{i\in B} R_i$:

$$\sum_{i\in B} R_i \geq H(X_B|X_{[m]\setminus B}),$$

obtained when $A = B$, and

$$\sum_{i\in B} R_i \geq H(X_B|X_{T\setminus B}),$$

obtained when $A = B \cup \{u\}$. The latter constraint is clearly stronger, so we can safely discard the former. ∎

We can now prove the desired lower bound on $R_T^{(\min)}$.

Lemma 3. *Let $m \geq 3$ be given. For $T \subset [m]$ with $|T| = m - 1$, we have*

$$R_T^{(\min)} \geq \frac{1}{m-2} \sum_{j\in T} H(X_{T\setminus\{j\}}|X_j).$$

Proof: Consider any $T \subset [m]$ with $|T| = m-1$. For each $j \in T$, let $B_j = T \setminus \{j\}$. Now, let $(R_i, i \in T)$ be any point in $\mathcal{R}_T$. Applying (2) with $B = B_j$, we get

$$\sum_{i \in B_j} R_i \geq H(X_{T\setminus\{j\}}|X_j),$$

for each $j \in T$. Summing over all $j \in T$, we obtain

$$\sum_{j \in T} \sum_{i \in B_j} R_i \geq \sum_{j \in T} H(X_{T\setminus\{j\}}|X_j). \tag{3}$$

Exchanging the order of summation in the double sum on the left-hand side (LHS) above, we have $\sum_{j\in T}\sum_{i\in B_j} R_i = \sum_{i\in T}\sum_{j\in B_i} R_i = \sum_{i\in T}(m-2)R_i = (m-2)\sum_{i\in T} R_i$. Putting this back into (3), we get

$$\sum_{i \in T} R_i \geq \frac{1}{m-2} \sum_{j \in T} H(X_{T\setminus\{j\}}|X_j).$$

Since this holds for any $(R_i, i \in T) \in \mathcal{R}_T$, the lemma follows. ∎

III. Omnivocal Communication

As pointed out in the Introduction, in the source model with two terminals, omnivocality is never necessary for generating a maximal-rate SK. However, the situation is different when there are three or more terminals. In this section, we give a sufficient condition for omnivocality being needed for achieving SK capacity when there are $m \geq 3$ terminals, and give an example where the sufficient condition is met. The sufficient condition also turns out to be necessary when there are exactly three terminals.

To state our results, we need a few definitions. The partition $\{\{1\}, \{2\}, \ldots, \{m\}\}$ consisting of m singleton cells will play a special role in our results; we call this the *singleton partition* and denote it by $\mathcal{S}$. For any partition $\mathcal{P}$ of $[m]$ with $|\mathcal{P}| \geq 2$, define

$$\Delta(\mathcal{P}) \triangleq \frac{1}{|\mathcal{P}|-1} \left[\sum_{A \in \mathcal{P}} H(X_A) - H(X_{[m]}) \right]. \tag{4}$$

Equivalently,

$$\Delta(\mathcal{P}) = \frac{1}{|\mathcal{P}|-1} D\left(p_{X_{[m]}} \parallel \prod_{A \in \mathcal{P}} p_{X_A} \right),$$

the notation being as in (1). Thus, $C([m]) = \mathbf{I}(X_{[m]}) = \min_{\mathcal{P}} \Delta(\mathcal{P})$. In all that follows, we say that the singleton partition $\mathcal{S}$ is a *minimizer for* $\mathbf{I}(X_{[m]})$ if $\Delta(\mathcal{S}) = \mathbf{I}(X_{[m]})$, and that $\mathcal{S}$ is the *unique minimizer for* $\mathbf{I}(X_{[m]})$ if the minimum in (1) is uniquely achieved by $\mathcal{S}$, i.e., $\Delta(\mathcal{S}) < \Delta(\mathcal{P})$ for all partitions $\mathcal{P}$ of $[m]$, $\mathcal{P} \neq \mathcal{S}$, with $|\mathcal{P}| \geq 2$.

We can now state the main result of this section.

Theorem 4. *For $m \geq 3$ terminals, if $\mathcal{S}$ is the unique minimizer for $\mathbf{I}(X_{[m]})$, then omnivocal communication is necessary for achieving the SK capacity $C([m])$.*

Before proving the theorem, we give an example where the condition of the theorem is met.

The pairwise independent network (PIN) model of Nitinawarat and Narayan [7] is defined on an underlying graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = [m]$. For $n \in \mathbb{N}$, let $\mathcal{G}^{(n)}$ be the multigraph $(\mathcal{V}, \mathcal{E}^{(n)})$, where $\mathcal{E}^{(n)}$ is the multiset of edges formed by taking n copies of each edge of $\mathcal{G}$. Associated with each edge $e \in \mathcal{E}^{(n)}$ is a Bernoulli$(1/2)$ rv ξ_e; the rvs ξ_e associated with distinct edges in $\mathcal{E}^{(n)}$ are independent. With this, the rvs X_i^n, $i \in [m]$, are defined as $X_i^n = (\xi_e : e \in \mathcal{E}^{(n)}$ and e is incident on $i)$. When $\mathcal{G} = K_m$, the complete graph on m vertices, we have the *complete graph PIN model*.

We show in the next section (Corollary 7.2) that for the complete graph PIN model, the singleton partition $\mathcal{S}$ is the unique minimizer for $\mathbf{I}(X_{[m]})$. The result below then immediately follows from Theorem 4.

Corollary 4.1. *In the PIN model defined on the complete graph K_m, $m \geq 3$, omnivocal communication is necessary for achieving $C(X_{[m]})$.*

In conjunction with Theorem 6 in [6], we now have the following picture for a capacity-achieving communication in the complete graph PIN model: the communication must be omnivocal, and if it is constrained to be a linear function of the observations $X_{[m]}^n$, then it must have rate at least $m(m-2)/2$. It should be noted that the capacity-achieving communication in the proof of [7, Theorem 1] is an omnivocal, linear communication of rate $m(m-2)/2$.

For the proof of Theorem 4, we need some convenient notation. For $T \subset [m]$, $|T| = m-1$, define $\Delta_T(\mathcal{S}) \triangleq \frac{1}{m-2}[\sum_{i\in T} H(X_i) - H(X_T)]$.

Lemma 5. *For $m \geq 3$ terminals, if the singleton partition $\mathcal{S}$ is the unique minimizer for $\mathbf{I}(X_{[m]})$, then $\Delta_T(\mathcal{S}) < \Delta(\mathcal{S})$ for all $T \subset [m]$ with $|T| = m-1$.*

Proof: For any $u \in [m]$, consider $T = [m] \setminus \{u\}$. Using $\Delta(\mathcal{S}) = \frac{1}{m-1}[\sum_{i=1}^m H(X_i) - H(X_{[m]})]$ and the definition of $\Delta_T(\mathcal{S})$ above, it is easy to verify the identity

$$\tfrac{m-1}{m-2}\Delta(\mathcal{S}) = \Delta_T(\mathcal{S}) + \tfrac{1}{m-2} I(X_u; X_T).$$

Re-arranging the above, we obtain

$$\begin{aligned} \Delta_T(\mathcal{S}) - \Delta(\mathcal{S}) &= \tfrac{1}{m-2}[\Delta(\mathcal{S}) - I(X_u; X_T)] \\ &= \tfrac{1}{m-2}[\Delta(\mathcal{S}) - \Delta(\mathcal{P})], \end{aligned} \tag{5}$$

where $\mathcal{P}$ is the 2-cell partition $\{\{u\}, T\}$ of $[m]$. By assumption, the expression in (5) is strictly negative. ∎

With this, we are ready to prove Theorem 4.

Proof of Theorem 4: We will show that $C([m]) > C([m]\|T)$ for any $T \subset [m]$ with $|T| = m-1$. First, note that since $\mathcal{S}$ is, by assumption, a minimizer for $\mathbf{I}(X_{[m]})$, we have $C([m]) = \mathbf{I}(X_{[m]}) = \Delta(\mathcal{S})$. Next, by Theorem 1 and Lemma 3, we have

$$\begin{aligned} C([m]\|T) &\leq H(X_T) - \tfrac{1}{m-2} \sum_{i \in T} H(X_{T\setminus\{i\}}|X_i), \\ &= \tfrac{1}{m-2} \left[(m-2)H(X_T) - \sum_{i \in T} [H(X_T) - H(X_i)] \right] \end{aligned}$$

$$= \Delta_T(\mathcal{S}).$$

Therefore, $C([m]\|T) \leq \Delta_T(\mathcal{S}) < \Delta(\mathcal{S}) = C([m])$, the second inequality coming from Lemma 5. ∎

For the three-terminal source model, it turns out that the unique minimizer condition in Theorem 4 is also necessary for the conclusion of the theorem to hold. Note that when $m = 3$, (1) reduces to $C(X_{[3]}) = \min\{I(X_{\{1,2\}}; X_3), I(X_{\{1,3\}}; X_2), I(X_{\{2,3\}}; X_1), \Delta(\mathcal{S})\}$; so the unique minimizer condition is equivalent to

$$\Delta(\mathcal{S}) < \min\{I(X_{\{1,2\}}; X_3), I(X_{\{1,3\}}; X_2), I(X_{\{2,3\}}; X_1)\}.$$

Theorem 6. *In the three-terminal source model, omnivocal communication is necessary for achieving SK capacity iff the singleton partition $\mathcal{S}$ is the unique minimizer for $\mathbf{I}(X_{[m]})$.*

Proof: The "if" part is by Theorem 4. For the "only if" part, suppose that $\Delta(\mathcal{S}) \geq \min\{I(X_{\{1,2\}}; X_3), I(X_{\{1,3\}}; X_2), I(X_{\{2,3\}}; X_1)\}$. Then, $\Delta(\mathcal{S})$ is either (a) greater than or equal to at least two of the three terms in the minimum, or (b) greater than or equal to exactly one term. Up to symmetry, it suffices to distinguish between two cases:

Case I: $\Delta(\mathcal{S}) \geq \max\{I(X_{\{1,2\}}; X_3), I(X_{\{1,3\}}; X_2)\}$

Case II: $\min\{I(X_{\{1,3\}}; X_2), I(X_{\{2,3\}}; X_1)\} > \Delta(\mathcal{S}) \geq I(X_{\{1,2\}}; X_3)$

In each case, we demonstrate a capacity-achieving communication in which at least one terminal remains silent.

We deal with Case I first. Observe that $\Delta(\mathcal{S}) = \frac{1}{2}\left[\sum_{i=1}^{3} H(X_i) - H(X_{[3]})\right]$ can also be written as $\frac{1}{2}[I(X_1; X_2) + I(X_{\{1,2\}}; X_3)]$. Thus, the assumption $\Delta(\mathcal{S}) \geq I(X_{\{1,2\}}; X_3)$, upon some re-organization, yields $I(X_1; X_2) \geq I(X_{\{1,2\}}; X_3)$, i.e.,

$$I(X_1; X_2) \geq I(X_1; X_3) + I(X_2; X_3|X_1). \tag{6}$$

Similarly, using the identity $\Delta(\mathcal{S}) = \frac{1}{2}[I(X_1; X_3) + I(X_{\{1,3\}}; X_2)]$ in the assumption $\Delta(\mathcal{S}) \geq I(X_{\{1,3\}}; X_2)$, we obtain $I(X_1; X_3) \geq I(X_{\{1,3\}}; X_2)$, i.e.,

$$I(X_1; X_3) \geq I(X_1; X_2) + I(X_1; X_3|X_2). \tag{7}$$

The equalities in (6) and (7) can simultaneously hold iff

$$\begin{aligned} I(X_1; X_2) = I(X_1; X_3) \quad \text{and} \\ I(X_1; X_3|X_2) = I(X_2; X_3|X_1) = 0. \end{aligned} \tag{8}$$

From (8), it is not hard to deduce that the quantities $I(X_{\{1,2\}}; X_3)$, $I(X_{\{1,3\}}; X_2)$, $I(X_{\{2,3\}}; X_1)$ and $\Delta(\mathcal{S})$ are all equal to $I(X_1; X_2)$. In particular, $C(X_{[3]}) = I(X_1; X_2)$.

From the first equality in (8), we also have $H(X_1|X_2) = H(X_1|X_3)$. Now, it can be shown by a standard random binning argument that there exists a communication from terminal 1 of rate $H(X_1|X_2) = H(X_1|X_3)$ such that X_1^n is ϵ-recoverable at both terminals 2 and 3. It then follows from the "balanced coloring lemma" [3, Lemma B.3] that an SK rate of $H(X_1) - H(X_1|X_2) = I(X_1; X_2)$ is achievable. Thus, the SK capacity, $C([3]) = I(X_1; X_2)$, is achievable by a communication in which terminals 2 and 3 are both silent.

Now, consider Case II, in which we obviously have $C([3]) = I(X_{\{1,2\}}; X_3)$. The idea here is to show that a valid communication of rate $H(X_{\{1,2\}}|X_3)$ exists in which terminal 3 is silent, and which allows ϵ-recoverability of (X_1^n, X_2^n) at all three terminals. Given this, an application of [3, Lemma B.3] shows that an SK rate of $H(X_{\{1,2\}}) - H(X_{\{1,2\}}|X_3) = I(X_{\{1,2\}}; X_3)$ is achievable. Thus, there is a $C([3])$-achieving communication in which terminal 3 is silent.

To show that the desired communication exists, we argue as follows. For $i = 1, 2$, let R_i be the rate at which terminal i communicates. A standard random binning argument shows that an achievable (R_1, R_2) region, with terminal 3 silent, for a communication intended to allow ϵ-recoverability of (X_1^n, X_2^n) at all terminals is given by

$$\begin{aligned} R_1 \geq H(X_1|X_2), \quad R_2 \geq H(X_2|X_1), \\ R_1 + R_2 \geq H(X_{\{1,2\}}|X_3). \end{aligned} \tag{9}$$

Now, using the assumption in Case II that $\Delta(\mathcal{S}) \geq I(X_{\{1,2\}}; X_3)$, we will prove that the inequality

$$H(X_1|X_2) + H(X_2|X_1) \leq H(X_{\{1,2\}}|X_3) \tag{10}$$

holds. It would then follow from (9) that there exist achievable rate pairs (R_1, R_2) with $R_1 + R_2 = H(X_{\{1,2\}}|X_3)$, thus completing the proof for Case II.

So, let us prove (10). We have $\Delta(\mathcal{S}) = \frac{1}{2}[H(X_1)+H(X_2)+H(X_3) - H(X_{[3]})]$ and $I(X_{\{1,2\}}; X_3) = H(X_{\{1,2\}}) + H(X_3) - H(X_{[3]})$. Using these expressions in the inequality $\Delta(\mathcal{S}) \geq I(X_{\{1,2\}}; X_3)$, and re-arranging terms, we obtain

$$\frac{1}{2}[H(X_1) + H(X_2) - 2H(X_{\{1,2\}})] \geq \frac{1}{2}[H(X_3) - H(X_{[3]})],$$

which is equivalent to (10). This completes the proof of the theorem. ∎

We in fact conjecture that the result of Theorem 6 should extend to more than three terminals as well.

Conjecture 1. *In the multiterminal source model with $m \geq 3$ terminals, omnivocal communication is necessary for achieving SK capacity iff the singleton partition is the unique minimizer for $\mathbf{I}(X_{[m]})$.*

At this point, we do not have a systematic approach for proving the "only if" part of the conjecture for $m \geq 4$.

IV. Singleton Partitions

The condition that the singleton partition be a unique minimizer for $\mathbf{I}(X_{[m]})$ plays a key role in the results of Section III. Thus, it would be very useful to have a way of checking whether this condition holds for a given source $X_{[m]}$, $m \geq 3$. The brute force method of comparing $\Delta(\mathcal{S})$ with $\Delta(\mathcal{P})$ for all partitions $\mathcal{P}$ with at least two parts requires an enormous amount of computation. Indeed, the number of partitions of an m-element set is the mth Bell number, B_m, an asymptotic estimate for which is $(\log w)^{1/2} w^{m-w} e^w$, where $w = \frac{m}{\log m}[1 + o(1)]$ is the solution to the equation

$m = w\log(w+1)$ [8, Example 5.4]. The proposition below brings down the number of comparisons required for verifying the unique minimizer condition to a "mere" $2^m - m - 2$.

For any non-empty subset $B = \{b_1, b_2, \ldots, b_{|B|}\}$ of $[m]$ with $|B| < m$, define $\mathcal{P}_B \triangleq \{B^c, \{b_1\}, \{b_2\}, \ldots, \{b_{|B|}\}\}$ to be the partition of $[m]$ consisting of $|B|+1$ cells, of which $|B|$ cells are singletons comprising the elements of B. Note that if $|B| = m-1$, then $\mathcal{P}_B = \mathcal{S}$.

Proposition 7. *For $m \geq 3$, let $\Omega = \{B \subset [m] : 1 \leq |B| \leq m-2\}$. The singleton partition $\mathcal{S}$ is*
(a) *a minimizer for $\mathbf{I}(X_{[m]})$ iff $\Delta(\mathcal{S}) \leq \Delta(\mathcal{P}_B)\ \forall\, B \in \Omega$;*
(b) *the unique minimizer for $\mathbf{I}(X_{[m]})$ iff $\Delta(\mathcal{S}) < \Delta(\mathcal{P}_B)$ $\forall\, B \in \Omega$.*

Proof: We prove (b); for (a), we simply have to replace the '$>$' in (11) below with a '$\geq$'.

The "only if" part is obvious. For the "if" part, suppose that $\Delta(\mathcal{S}) < \Delta(\mathcal{P}_B)$ for all $B \subset [m]$ with $1 \leq |B| \leq m-2$. Consider any partition $\mathcal{P}$ of $[m]$, $\mathcal{P} \neq \mathcal{S}$, with $|\mathcal{P}| \geq 2$. We wish to show that $\Delta(\mathcal{P}) > \Delta(\mathcal{S})$.

The following identity can be obtained from the definition in (4) by some re-grouping of terms:

$$\sum_{A\in\mathcal{P}} |A^c|\,\Delta(\mathcal{P}_{A^c}) = (|\mathcal{P}|-1)[\Delta(\mathcal{P}) + (m-1)\Delta(\mathcal{S})].$$

Thus, we have

$$\begin{aligned}\Delta(\mathcal{P}) &= \frac{1}{|\mathcal{P}|-1}\sum_{A\in\mathcal{P}} |A^c|\,\Delta(\mathcal{P}_{A^c}) - (m-1)\Delta(\mathcal{S}) \\ &> \frac{1}{|\mathcal{P}|-1}\sum_{A\in\mathcal{P}} |A^c|\,\Delta(\mathcal{S}) - (m-1)\Delta(\mathcal{S}) \qquad (11)\\ &= m\Delta(\mathcal{S}) - (m-1)\Delta(\mathcal{S}) \;=\; \Delta(\mathcal{S}). \qquad (12)\end{aligned}$$

The inequality in (11) is due to the fact that at least one $A \in \mathcal{P}$ is not a singleton cell, so that $\mathcal{P}_{A^c} \neq \mathcal{S}$, and hence, $\Delta(\mathcal{P}_{A^c}) > \Delta(\mathcal{S})$ by assumption. To verify the first equality in (12), observe that $\sum_{A\in\mathcal{P}}|A^c| = \sum_{A\in\mathcal{P}}\sum_{i\notin A} 1 = \sum_{i=1}^m \sum_{A\in\mathcal{P}: i\notin A} 1 = m(|\mathcal{P}|-1)$. ∎

Next, we apply the above proposition to some interesting special cases.

Random variables $X_1, X_2, \ldots, X_m$, $m \geq 2$, are called *isentropic* if $H(X_A) = H(X_B)$ for any pair of non-empty subsets $A, B \subseteq [m]$ having the same cardinality. Equivalently, $X_1, \ldots, X_m$ are isentropic if, for all non-empty $A \subseteq [m]$, the entropy $H(X_A)$ depends only on $|A|$. One obvious consequence of this definition is that for disjoint non-empty subsets $A, B \subset [m]$, the conditional entropy $H(X_A|X_B)$ only depends on $|A|$ and $|B|$.

Clearly, i.i.d. rvs are isentropic. More generally, exchangeable rvs are isentropic — rvs $X_1, X_2, \ldots, X_m$ are *exchangeable* if for all permutations σ of $[m]$, the joint distribution of $(X_1, X_2, \ldots, X_m)$ is the same as that of $(X_{\sigma(1)}, X_{\sigma(2)}, \ldots, X_{\sigma(m)})$. However, isentropic rvs need not be exchangeable. It may be verified that the rvs $X_1, X_2, \ldots, X_m$ in the PIN model defined on the complete graph K_m (as defined in Section III) are not exchangeable when $m \geq 3$, but they are isentropic.

Corollary 7.1. *If $X_1, X_2, \ldots, X_m$, $m \geq 3$, are isentropic rvs, then $\mathcal{S}$ is a minimizer for $\mathbf{I}(X_{[m]})$.*

Proof: For a partition $\mathcal{P}$ of $[m]$ with $|\mathcal{P}| \geq 2$, let us define

$$\delta(\mathcal{P}) \;\triangleq\; \frac{1}{|\mathcal{P}|-1}\sum_{A\in\mathcal{P}} H(X_{A^c}|X_A) \;=\; H(X_{[m]}) - \Delta(\mathcal{P}).$$

By virtue of Proposition 7(a), we need to show that $\delta(\mathcal{P}_B) \leq \delta(\mathcal{S})$ for all $B \in \Omega$.

For isentropic rvs, the quantity $H(X_B|X_{B^c})$, for any $B \subseteq [m]$, depends only on $|B|$. Hence, defining $g(k) \triangleq H(X_{[k]}|X_{[m]\setminus[k]})$ for $1 \leq k \leq m$, we can write $\delta(\mathcal{P}_B) = \frac{1}{|B|}g(|B|) + g(m-1)$ and $\delta(\mathcal{S}) = \frac{m}{m-1}g(m-1)$. Thus, we have to show that $\frac{g(|B|)}{|B|} \leq \frac{g(m-1)}{m-1}$ for all $B \in \Omega$. This follows from the fact that for isentropic rvs, the function $g(k)/k$ is non-decreasing in k — see Appendix A. ∎

Our second application of Proposition 7 is to the PIN model. Recall from Section III that this model is defined on an underlying graph $\mathcal{G} = ([m], \mathcal{E})$. From the way that the rvs X_i, $i \in [m]$, are defined, it is not difficult to verify that for any partition $\mathcal{P}$ of $[m]$ with $|\mathcal{P}| \geq 2$, we have

$$\Delta(\mathcal{P}) = \frac{|\mathcal{E}(\mathcal{P})|}{|\mathcal{P}|-1},$$

where $|\mathcal{E}(\mathcal{P})|$ denotes the number of edges $e = \{i, j\} \in \mathcal{E}$ such that i and j are in different cells of $\mathcal{P}$. This, in conjunction with Proposition 7, gives us a relatively simple criterion for verifying whether $\mathcal{S}$ is a (unique) minimizer for $\mathbf{I}(X_{[m]})$. As an illustration, we apply this to the complete graph PIN model.

Corollary 7.2. *For the PIN model on the complete graph K_m, $m \geq 3$, the singleton partition $\mathcal{S}$ is the unique minimizer for $\mathbf{I}(X_{[m]})$.*

Proof: It is easy to see that for any non-empty $B \subsetneq [m]$, $|\mathcal{E}(\mathcal{P}_B)| = \binom{m}{2} - \binom{|B^c|}{2} = \frac{1}{2}|B|(2m-|B|-1)$. Hence,

$$\Delta(\mathcal{P}_B) = \frac{|\mathcal{E}(\mathcal{P}_B)|}{|B|} = \frac{2m-|B|-1}{2} \geq \frac{m}{2} = \Delta(\mathcal{S}),$$

with equality iff $|B| = m-1$, i.e., $\mathcal{P}_B = \mathcal{S}$. The result now follows from Proposition 7(b). ∎

Appendix A

Here, we prove that for isentropic rvs $X_1, \ldots, X_m$, the function $\frac{1}{k}H(X_{[k]}|X_{[m]\setminus[k]})$, defined for $1 \leq k \leq m$, is non-decreasing in k. Define $g(k) = H(X_{[k]}|X_{[m]\setminus[k]})$. We show that the difference $kg(k+1) - (k+1)g(k)$ is always non-negative, from which the result follows.

We have $g(k+1) = H(X_{[m]}) - H(X_{\{k+2,\ldots,m\}})$ and $g(k) = H(X_{[m]}) - H(X_{\{k+1,\ldots,m\}}) = g(k+1) - H(X_{k+1}|X_{\{k+2,\ldots,m\}})$. Thus,

$$\begin{aligned}&kg(k+1) - (k+1)g(k)\\ &\qquad = (k+1)\,H(X_{k+1}|X_{\{k+2,\ldots,m\}}) - g(k+1).\end{aligned}$$

We need to show that the RHS of the equality is non-negative. This is straightforward:

$$\begin{aligned} g(k+1) &= H(X_{[k+1]}|X_{\{k+2,\ldots,m\}}) \\ &\leq \sum_{i=1}^{k+1} H(X_i|X_{\{k+2,\ldots,m\}}) \\ &= (k+1)H(X_{k+1}|X_{\{k+2,\ldots,m\}}), \end{aligned}$$

since, for $1 \leq i \leq k+1$, $H(X_i|X_{\{k+2,\ldots,m\}}) = H(X_{k+1}|X_{\{k+2,\ldots,m\}})$ by isentropy.

REFERENCES

[1] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography, part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, pp. 1121–1132, July 1993.

[2] C. Chan and L. Zheng, "Mutual dependence for secret key agreement," *Proc. 44th Annu. Conf. Inf. Sci. Syst. (CISS)*, 2010.

[3] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3047–3061, Dec. 2004.

[4] A.A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals–Part I," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3973–3996, Aug. 2010.

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

[6] M. Mukherjee and N. Kashyap, "On the communication complexity of secret key generation in the multiterminal source model." arXiv:1401.1117.

[7] S. Nitinawarat and P. Narayan, "Perfect omniscience, perfect secrecy and Steiner tree packing," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6490–6500, Dec. 2010.

[8] A.M. Odlyzko, "Asymptotic enumeration methods," in *Handbook of Combinatorics*, R.L. Graham et al., eds., 1995, pp. 1063–1229.

[9] H. Tyagi, "Common information and secret key capacity," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5627–5640, Sep. 2013.