Common Information and Secret Key Capacity

Himanshu Tyagi†

***Abstract*—We study the generation of a secret key of maximum rate by a pair of terminals observing correlated sources and with the means to communicate over a noiseless public communication channel. Our main result establishes a structural equivalence between the generation of a maximum rate secret key and the generation of a common randomness that renders the observations of the two terminals conditionally independent. The minimum rate of such common randomness, termed interactive common information, is related to Wyner's notion of common information, and serves to characterize the minimum rate of interactive public communication required to generate an optimum rate secret key. This characterization yields a single-letter expression for the aforementioned communication rate when the number of rounds of interaction are bounded. An application of our results shows that interaction does not reduce this rate for binary symmetric sources. Further, we provide an example for which interaction does reduce the minimum rate of communication. Also, certain invariance properties of common information quantities are established that may be of independent interest.**

***Index Terms*—Common information, common randomness, interactive communication, interactive common information, secret key capacity.**

I. Introduction

Consider secret key (SK) generation by a pair of terminals that observe independent and identically distributed (i.i.d.) repetitions of two discrete, finite-valued random variables (rvs) of known joint probability mass function. The terminals communicate over a noiseless public channel of unlimited capacity, interactively in multiple rounds, to agree upon the value of the key. The key is required to be (almost) independent of the public communication. The maximum rate of such an SK, termed the secret key capacity, was characterized in [13], [1].

In the works of Maurer and Ahlswede-Csiszár [13], [1], SK generation of maximum rate entailed both the terminals recovering the observations of one of the terminals, using the least rate of communication required to do so. Later, it was shown by Csiszár-Narayan [5] that a maximum rate SK can be generated also by the terminals recovering the observations of both the terminals. Clearly, the latter scheme requires more communication than the former. In this paper, we address the following question, which was raised in [5, Section VI]:
What is the minimum overall rate of interactive communication R_{SK} required to establish a maximum rate SK?

This work was supported by the U.S. National Science Foundation under Grants CCF0830697 and CCF1117546.

†Department of Electrical and Computer Engineering, and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA. Email: tyagi@umd.edu.

A preliminary version of this paper was presented at the IEEE International Symposium on Information Theory, St. Petersburg, Russia, July 31 - August 5, 2011.

We answer this question by characterizing the form of common randomness (CR) (i.e., shared bits, see [2]) that the terminals must establish in order to generate a maximum rate SK; two examples of such common randomness are the observations of any one terminal [13], [1] and of both terminals [5]. While our main result does not yield a single-letter characterization, it nonetheless reveals a central link between secrecy generation and Wyner's notion of common information (CI) between two dependent rvs X and Y [16]. Wyner defined CI as the minimum rate of a function of i.i.d. repetitions of two correlated random variables X and Y that facilitated a certain distributed source coding task. Alternatively, it can be defined as the minimum rate of a function of i.i.d. repetitions of X and Y such that, conditioned on this function, the i.i.d. sequences are (almost) independent; this definition, though not stated explicitly in [16], follows from the analysis therein. We introduce a variant of this notion of CI called the *interactive CI* where we seek the minimum rate of CR that renders the mentioned sequences conditionally independent. Clearly, interactive CI cannot be smaller than Wyner's CI, and can exceed it. Our main contribution is to show a one-to-one correspondence between such CR and the CR established for generating an optimum rate SK. This correspondence is used to characterize the minimum rate of communication R_{SK} required for generating a maximum rate SK. In fact, it is shown that R_{SK} is simply interactive CI minus the secret key capacity.

When the number of rounds of interaction are bounded, this characterization yields a single-letter expression for R_{SK}. Using this expression we show that an interactive communication scheme can have less rate than a noninteractive one, in general. However, interaction offers no advantage for binary symmetric sources. This expression also illustrates the role of sufficient statistics in SK generation. We further dwell on this relationship and show that many CI quantities of interest remain unchanged if the sources are replaced by their corresponding sufficient statistics (with respect to each other). Interestingly, the effect of substitution by sufficient statistics has been studied in the context of the rate-distortion problem for a remote source in [7, Lemma 2], and recently, for the lossy and lossless distributed source coding problems in [17]. Here, in effect, we study this substitution for the distributed source coding problems underlying the CI quantities.

The basic notions of CR and SK are explained in the next section. The definition of interactive CI and the heuristics underlying our approach are given in Section III. Our main results are provided in Section IV, followed by illustrative examples in the subsequent section. Section VI explores the connection between sufficient statistics and common information quantities. A discussion of our results and possible

extensions is given in the final section.

Notation. The rvs X and Y take values in finite sets $\mathcal{X}$ and $\mathcal{Y}$, respectively. Let $X^n = (X_1, ..., X_n)$ and $Y^n = (Y_1, ..., Y_n)$ denote n i.i.d. repetitions of X and Y, respectively. For a collection of rvs $U_1, ..., U_r$, for $i \leq j$ let U_i^j denote $U_i, U_{i+1} ..., U_j$; when $i = 1$, we use $U^j = U_1, ..., U_j$. For rvs U, V, and $0 < \epsilon < 1$, we say U is ϵ-recoverable from V if there is a function g of V such that

$$\mathbb{P}\left(U = g(V)\right) \geq 1 - \epsilon.$$

Denote the cardinality of the range space of a mapping f by $\|f\|$, and similarly, with a slight abuse of notation, the (fixed) range space of a random mapping $\mathbf{F}$ by $\|\mathbf{F}\|$.

II. Interactive Communication, Common Randomness and Secret Keys

Terminals $\mathcal{X}$ and $\mathcal{Y}$ (with a slight abuse of notation) communicate interactively, with, say, terminal $\mathcal{X}$ transmitting first. Each terminal then communicates alternately for r rounds. Specifically, an *r-interactive communication* $\mathbf{f} = (f_1, f_2, ..., f_r)$ is a sequence of finite-valued mappings with

$$f_{2i+1} : \mathcal{X}^n \times \mathcal{F}^{2i} \to \mathcal{F}_{2i+1}, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor,$$
$$f_{2i} : \mathcal{Y}^n \times \mathcal{F}^{2i-1} \to \mathcal{F}_{2i}, \quad 1 \leq i \leq \lfloor r/2 \rfloor,$$

where $\{\mathcal{F}_i\}_{i=1}^r$ are finite sets and $\mathcal{F}_0 = \emptyset$. This set-up subsumes protocols where terminal $\mathcal{Y}$ initiates the communication upon choosing $f_1 =$ constant. Let $\mathbf{F} = \mathbf{f}\left(X^n, Y^n\right)$ describe collectively the corresponding rv. The rate of this communication is given by

$$\frac{1}{n} \log \|\mathbf{F}\|.$$

We assume that the communication from each terminal is a (deterministic) function of its knowledge. In particular, randomization is not allowed. This is not a limiting assumption; see Section VII-A.

Definition 1. Given interactive communication $\mathbf{F}$ as above, a function L of (X^n, Y^n) is *ϵ-common randomness* (ϵ-CR) recoverable from[1] $\mathbf{F}$ if there exist mappings $L_1 = L_1^{(n)}(X^n, \mathbf{F})$ and $L_2 = L_2^{(n)}(Y^n, \mathbf{F})$ such that

$$\mathbb{P}\left(L = L_1 = L_2\right) \geq 1 - \epsilon.$$

Definition 2. A function K of (X^n, Y^n), with values in a set $\mathcal{K}$, forms an *ϵ-secret key for X and Y* (ϵ-SK) if K is ϵ-CR recoverable from X^n or Y^n and (interactive public communication) $\mathbf{F}$, and

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

For convenience, simplistically, the ϵ-SK K is said to be recoverable from $\mathbf{F}$. A rate $R > 0$ is an achievable SK rate if for every $0 < \epsilon < 1$ there exists, for some[2] $n \geq 1$, an ϵ-SK $K = K^{(n)}$ with $(1/n)H(K) \geq R - \epsilon$. The supremum of all achievable SK rates is denoted by C, and is called the SK capacity.

The following result[3] is well known.

Theorem 1. [13], [1] *The SK capacity for X and Y is given by*

$$C = I(X \wedge Y). \tag{2}$$

III. Relation between Secret Key and Wyner's Common Information

We interpret Wyner's CI for a pair of rvs (X, Y) as the minimum rate of a function of their i.i.d. repetitions (X^n, Y^n) that renders X^n and Y^n conditionally independent. Formally,

Definition 3. $R \geq 0$ is an achievable CI rate if for every $0 < \epsilon < 1$ there exists an $n \geq 1$ and a (finite-valued) rv $L = L\left(X^n, Y^n\right)$ of rate $(1/n)H(L) \leq R + \epsilon$ that satisfies the property:

$$\frac{1}{n} I\left(X^n \wedge Y^n \mid L\right) \leq \epsilon. \tag{3}$$

Obvious examples of such an rv L are $L = (X^n, Y^n)$ or X^n or Y^n. The infimum of all achievable CI rates, denoted $CI_W(X \wedge Y)$, is called the CI of X and Y. This definition of CI, though not stated explicitly in [16], follows from the analysis therein. The following theorem characterizes $CI_W(X \wedge Y)$.

Theorem 2. [16] *The CI of the rvs X, Y is*

$$CI_W(X \wedge Y) = \min_{W} I(X, Y \wedge W), \tag{4}$$

where the rv W takes values in a (finite) set $\mathcal{W}$ with $|\mathcal{W}| \leq |\mathcal{X}||\mathcal{Y}|$ and satisfies the Markov condition $X \multimap W \multimap Y$.

The direct part follows from [16, equation (5.12)]. The proof of the converse is straightforward. Further, it is a simple exercise to infer from (4) that $CI_W(X \wedge Y) \geq I(X \wedge Y)$.

Definition 4. An achievable *r-interactive* CI rate is defined in a manner analogous to the achievable CI rate, but with the restriction that the rvs L in (3) be ϵ-CR, i.e., $L = (J, \mathbf{F})$, where $\mathbf{F}$ is an r-interactive communication and J is ϵ-recoverable from $\mathbf{F}$. The infimum of all achievable r-interactive CI rates, denoted $CI_i^r(X; Y)$, is called the r-interactive CI of the rvs X and Y. By definition, the nonnegative sequence $\{CI_i^r(X; Y)\}_{r=1}^{\infty}$ is nonincreasing in r and is bounded below by $CI_W(X \wedge Y)$. Define

$$CI_i(X \wedge Y) = \lim_{r \to \infty} CI_i^r(X; Y).$$

Then $CI_i(X \wedge Y) \geq CI_W(X \wedge Y) \geq 0$. Note that $CI_i^r(X; Y)$ may not be symmetric in X and Y since the communication is initiated at terminal $\mathcal{X}$. However, since

$$CI_i^{r+1}(X; Y) \leq CI_i^r(Y; X) \leq CI_i^{r-1}(X; Y),$$

[1]The rv L is ϵ-recoverable from $(X^n, \mathbf{F})$ or $(Y^n, \mathbf{F})$ but not necessarily from $\mathbf{F}$ alone. The deliberate misuse of the terminology "recoverable from $\mathbf{F}$" simplifies presentation.

[2]Our results hold even if the phrase "for some $n \geq 1$" is replaced by "for all n sufficiently large;" the former has been chosen here for convenience.

[3]It is shown in [14], [3] that SK capacity remains unchanged even if the notion of "weak secrecy" of K in (1) is tightened to "strong secrecy" by omitting the normalization with respect to n, and an additional uniformity constraint $H(K) \geq \log |\mathcal{K}| - \epsilon$ is imposed.

clearly,

$$\begin{aligned} CI_i(X \wedge Y) &= \lim_{r\to\infty} CI_i^r(X;Y) \\ &= \lim_{r\to\infty} CI_i^r(Y;X) \\ &= CI_i(Y \wedge X). \end{aligned} \tag{5}$$

Further, for all $0 < \epsilon < 1$, $J = X^n$ is ϵ-recoverable from Y^n and a communication (of a Slepian-Wolf codeword) $F = F(X^n)$, and $L = (J, F)$ satisfies (3). Hence, $CI_i(X \wedge Y) \leq H(X)$; similarly, $CI_i(X \wedge Y) \leq H(Y)$. To summarize, we have

$$0 \leq CI_W(X \wedge Y) \leq CI_i(X \wedge Y) \leq \min\{H(X), H(Y)\}, \tag{6}$$

where the first and the last inequalities can be strict. In Section V-A we show that the second inequality is strict for binary symmetric rvs X, Y.

The r-interactive CI plays a pivotal role in optimum rate SK generation. Loosely speaking, our main result asserts the following. *A CR that satisfies (3) can be used to generate an optimum rate SK and conversely, an optimum rate SK yields a CR satisfying (3). In fact, such a CR of rate R can be recovered from an interactive communication of rate $R - C$, where C is the SK capacity for X and Y. Therefore, to find the minimum rate of interactive communication needed to generate an optimum rate SK, it is sufficient to characterize $CI_i(X \wedge Y)$.*

IV. Main Results

Definition 5. A rate $R' \geq 0$ is an achievable r-interactive communication rate for CI_i^r if, for all $0 < \epsilon < 1$, there exists, for some $n \geq 1$, an r-interactive communication $\mathbf{F}$ of rate $(1/n)\log\|\mathbf{F}\| \leq R' + \epsilon$, and an ϵ-CR J recoverable from $\mathbf{F}$, with $L = (J, \mathbf{F})$ satisfying (3). Let R_{CI}^r denote the infimum of all achievable r-interactive communication rates for CI_i^r. Similarly, $R'' \geq 0$ is an achievable r-interactive communication rate for SK capacity if, for all $0 < \epsilon < 1$, there exists, for some $n \geq 1$, an r-interactive communication $\mathbf{F}$ of rate $(1/n)\log\|\mathbf{F}\| \leq R'' + \epsilon$, and an ϵ-SK K, recoverable from $\mathbf{F}$, of rate $(1/n)H(K) \geq I(X \wedge Y) - \epsilon$; R_{SK}^r denotes the infimum of all achievable r-interactive communication rates for SK capacity. Note that by their definitions, both R_{CI}^r and R_{SK}^r are nonincreasing with increasing r, and are bounded below by zero. Define

$$R_{CI} = \lim_{r\to\infty} R_{CI}^r, \quad R_{SK} = \lim_{r\to\infty} R_{SK}^r.$$

Although $R_{CI}^r(X;Y)$ and $R_{SK}^r(X;Y)$ are not equal to $R_{CI}^r(Y;X)$ and $R_{SK}^r(Y;X)$, respectively, the quantities R_{CI} and R_{SK} are symmetric in X and Y using an argument similar to the one leading to (5).

Theorem 3. *For every $r \geq 1$,*

$$R_{SK}^r = R_{CI}^r = CI_i^r(X;Y) - I(X \wedge Y). \tag{7}$$

Corollary. *It holds that*

$$R_{SK} = R_{CI} = CI_i(X \wedge Y) - I(X \wedge Y). \tag{8}$$

Remark. The relation (8) can be interpreted as follows. Any CR J recoverable from (interactive communication) $\mathbf{F}$, with $L = (J, \mathbf{F})$ satisfying (3), can be decomposed into two mutually independent parts: An SK K of maximum rate and the interactive communication $\mathbf{F}$. It follows upon rewriting (8) as $CI_i(X \wedge Y) = I(X \wedge Y) + R_{CI}$ that the communication $\mathbf{F}$ is (approximately) of rate R_{CI}. Furthermore, R_{CI} is the same as R_{SK}.

A computable characterization of the operational term $CI_i(X \wedge Y)$ is not known. However, the next result gives a single-letter characterization of $CI_i^r(X;Y)$.

Theorem 4. *Given rvs X, Y and $r \geq 1$, we have*

$$CI_i^r(X;Y) = \min_{U_1,\ldots,U_r} I(X, Y \wedge U_1, ..., U_r), \tag{9}$$

where the minimum is taken over rvs $U_1, ..., U_r$ taking values in finite sets $\mathcal{U}_1, ..., \mathcal{U}_r$, respectively, that satisfy the following conditions

$$\begin{aligned} &(P1)\ U_{2i+1} \multimap X, U^{2i} \multimap Y, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor, \\ &\qquad U_{2i} \multimap Y, U^{2i-1} \multimap X, \quad 1 \leq i \leq \lfloor r/2 \rfloor, \\ &(P2)\ X \multimap U^r \multimap Y, \\ &(P3)\ |\mathcal{U}_{2i+1}| \leq |\mathcal{X}| \prod_{j=1}^{2i} |\mathcal{U}_j| + 1, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor, \\ &\qquad |\mathcal{U}_{2i}| \leq |\mathcal{Y}| \prod_{j=1}^{2i-1} |\mathcal{U}_j| + 1, \quad 1 \leq i \leq \lfloor r/2 \rfloor, \end{aligned}$$

with $\mathcal{U}_0 = \emptyset$ and $U_0 =$ constant.

Remark. Note that (9) has the same form as the expression for $CI_W(X \wedge Y)$ in (4) with W replaced by $(U_1, ..., U_r)$ satisfying the conditions above.

Before presenting the proof of our main Theorems 3 and 4, we give some technical results that will constitute central tools for the proofs.

Lemma 5. *For an interactive communication $\mathbf{F}$ it holds that*

$$H(\mathbf{F} \mid X^n) + H(\mathbf{F} \mid Y^n) \leq H(\mathbf{F}). \tag{10}$$

Lemma 6. *For an r-interactive communication $\mathbf{F}$, define*

$$\mathbf{F}_i = \mathbf{F}\left(X_{n(i-1)+1}^{ni}, Y_{n(i-1)+1}^{ni}\right), \quad 1 \leq i \leq k.$$

Then, for all $k \geq k_0(n, \epsilon, |\mathcal{X}|, |\mathcal{Y}|)$ there exists an r-interactive communication $\mathbf{F}' = \mathbf{F}'\left(X^{nk}, Y^{nk}\right)$ of rate

$$\frac{1}{nk}\log\|\mathbf{F}'\| \leq \frac{1}{n}\left[H\left(\mathbf{F}|X^n\right) + H\left(\mathbf{F}|Y^n\right)\right] + \epsilon, \tag{11}$$

such that $\mathbf{F}^k$ is an ϵ-CR recoverable from $\mathbf{F}'$.

Remark. Lemma 6 says that, in essence, for an optimum rate communication $\mathbf{F}$,

$$\frac{1}{n}\log\|\mathbf{F}\| \approx \frac{1}{n}\left[H\left(\mathbf{F}|X^n\right) + H\left(\mathbf{F}|Y^n\right)\right].$$

Lemma 7. *(A General Decomposition) For a CR J recoverable from an interactive communication $\mathbf{F}$ we have*

$$\begin{aligned} &nI(X \wedge Y) \\ &= I\left(X^n \wedge Y^n \mid J, \mathbf{F}\right) + H(J, \mathbf{F}) - H\left(\mathbf{F} \mid X^n\right) \\ &\quad - H\left(\mathbf{F} \mid Y^n\right) - H\left(J \mid X^n, \mathbf{F}\right) - H\left(J \mid Y^n, \mathbf{F}\right). \end{aligned} \tag{12}$$

Lemma 5 is a special case of [6, Lemma B.1] (also, see [12]). The proofs of Lemma 6 and Lemma 7 are given in the Appendix.

Note that a simplification of (12) gives

$$\begin{aligned} I(X \wedge Y) \leq \frac{1}{n}\Big[& I\left(X^n \wedge Y^n \mid J, \mathbf{F}\right) + \\ & H(J, \mathbf{F}) - H\left(\mathbf{F} \mid X^n\right) - H\left(\mathbf{F} \mid Y^n\right)\Big]. \end{aligned} \tag{13}$$

If J is an ϵ-CR recoverable from $\mathbf{F}$, Fano's inequality implies

$$\begin{aligned} \frac{1}{n}\big[H(J \mid X^n, \mathbf{F}) + H(J \mid Y^n, \mathbf{F})\big] &\leq 2\epsilon \log|\mathcal{X}||\mathcal{Y}| + 2h(\epsilon) \\ &= \delta(\epsilon), \text{ say,} \end{aligned} \tag{14}$$

where $h(\epsilon) = -\epsilon\log\epsilon - (1-\epsilon)\log(1-\epsilon)$, and $\delta(\epsilon) \to 0$ as $\epsilon \to 0$. Combining (12) and (14) we get

$$\begin{aligned} I(X \wedge Y) \geq \frac{1}{n}\Big[& I\left(X^n \wedge Y^n \mid J, \mathbf{F}\right) + H(J, \mathbf{F}) \\ & - H\left(\mathbf{F} \mid X^n\right) - H\left(\mathbf{F} \mid Y^n\right)\Big] - \delta(\epsilon), \end{aligned} \tag{15}$$

and further, by (10),

$$\begin{aligned} I(X \wedge Y) \geq \frac{1}{n}\left[I\left(X^n \wedge Y^n \mid J, \mathbf{F}\right) + H(J, \mathbf{F}) - H(\mathbf{F})\right] \\ - \delta(\epsilon). \end{aligned} \tag{16}$$

A. Proof of Theorem 3

In this section we give a proof for (7). The proof of (8) then follows upon taking limit $r \to \infty$ on both sides of (7). The proof of (7) follows from claims 1-3 below. In particular, the proofs of claims 1-3 establish a structural equivalence between a maximum rate SK and an SK of rate $\approx \frac{1}{n}H(J \mid \mathbf{F})$ extracted from a CR J recoverable from $\mathbf{F}$ such that $L = (J, \mathbf{F})$ satisfies (3).

Claim 1: $R^r_{CI} \geq CI^r_i(X;Y) - I(X \wedge Y)$.

Proof. By the definition of R^r_{CI}, for every $0 < \epsilon < 1$ there exists, for some $n \geq 1$, an r-interactive communication $\mathbf{F}$ of rate

$$\frac{1}{n}\log\|\mathbf{F}\| \leq R^r_{CI} + \epsilon, \tag{17}$$

and J, an ϵ-CR recoverable from $\mathbf{F}$, such that $L = (J, \mathbf{F})$ satisfies (3). It follows upon rearranging the terms in (16) that

$$\frac{1}{n}H(J, \mathbf{F}) \leq I(X \wedge Y) + \frac{1}{n}H(\mathbf{F}) + \delta(\epsilon),$$

which with (17) gives

$$\frac{1}{n}H(J, \mathbf{F}) \leq I(X \wedge Y) + R^r_{CI} + \epsilon + \delta(\epsilon). \tag{18}$$

Since $(J, \mathbf{F})$ satisfies

$$\frac{1}{n}I\left(X^n \wedge Y^n \mid J, \mathbf{F}\right) \leq \epsilon \leq \epsilon + \delta(\epsilon),$$

the inequality (18), along with the fact that $(\epsilon + \delta(\epsilon)) \to 0$ as $\epsilon \to 0$, implies that $I(X \wedge Y) + R^r_{CI}$ is an achievable r-interactive CI rate; hence, $CI^r_i(X;Y) \leq I(X \wedge Y) + R^r_{CI}$.

Claim 2: $R^r_{SK} \geq R^r_{CI}$.

Proof. Using the definition of R^r_{SK}, for $0 < \epsilon < 1$ there exists, for some $n \geq 1$, an r-interactive communication $\mathbf{F}$ of rate $\frac{1}{n}\log\|\mathbf{F}\| \leq R^r_{SK} + \epsilon$, and an ϵ-SK K recoverable from $\mathbf{F}$ of rate

$$\frac{1}{n}H(K) \geq I(X \wedge Y) - \epsilon. \tag{19}$$

By choosing $J = K$ in (16) and rearranging the terms we get,

$$\frac{1}{n}I\left(X^n \wedge Y^n \mid K, \mathbf{F}\right) \leq I(X \wedge Y) - \frac{1}{n}H(K \mid \mathbf{F}) + \delta(\epsilon).$$

Next, from $(1/n)I(K \wedge \mathbf{F}) < \epsilon$, we have

$$\begin{aligned} \frac{1}{n}I\left(X^n \wedge Y^n \mid K, \mathbf{F}\right) &\leq I(X \wedge Y) - \frac{1}{n}H(K) + \epsilon + \delta(\epsilon) \\ &\leq 2\epsilon + \delta(\epsilon), \end{aligned}$$

where the last inequality follows from (19). Since $(2\epsilon + \delta(\epsilon)) \to 0$ as $\epsilon \to 0$, R^r_{SK} is an achievable r-interactive communication rate for CI^r_i, and thus, $R^r_{SK} \geq R^r_{CI}$.

Claim 3: $R^r_{SK} \leq CI^r_i(X;Y) - I(X \wedge Y)$.

Proof. For $0 < \epsilon < 1$, let J be an ϵ-CR recoverable from an r-interactive communication $\mathbf{F}$, with

$$\frac{1}{n}H(J, \mathbf{F}) \leq CI^r_i(X;Y) + \epsilon, \tag{20}$$

such that $L = (J, \mathbf{F})$ satisfies (3), and so, by (13),

$$\begin{aligned} &\frac{1}{n}\left[H(\mathbf{F} \mid X^n) + H(\mathbf{F} \mid Y^n)\right] \\ &\leq \frac{1}{n}H(J, \mathbf{F}) - I(X \wedge Y) + \epsilon \\ &\leq CI^r_i(X;Y) - I(X \wedge Y) + 2\epsilon. \end{aligned} \tag{21}$$

To prove the assertion in claim 3, we show that for some $N \geq 1$ there exists $\Delta(\epsilon)$-SK $K = K(X^N, Y^N)$ of rate

$$\frac{1}{n}\log\|K\| \geq I(X \wedge Y) - \Delta(\epsilon)$$

recoverable from an r-interactive communication $\mathbf{F}'' = \mathbf{F}''(X^N, Y^N)$ of rate

$$\frac{1}{N}\log\|\mathbf{F}''\| \leq \frac{1}{n}\left[H(\mathbf{F} \mid X^n) + H(\mathbf{F} \mid Y^n)\right] + \Delta(\epsilon) - 2\epsilon, \tag{22}$$

where $\Delta(\epsilon) \to 0$ as $\epsilon \to 0$. Then (22), along with (21), would yield

$$\frac{1}{N}\log\|\mathbf{F}''\| \leq CI^r_i(X;Y) - I(X \wedge Y) + \Delta(\epsilon), \tag{23}$$

so that $CI^r_i(X;Y) - I(X \wedge Y)$ is an achievable r-interactive communication rate for SK capacity, thereby establishing the claim.

It remains to find K and $\mathbf{F}''$ as above. To that end, let J be recovered as $J_1 = J_1(X^n, \mathbf{F})$ and $J_2 = J_2(Y^n, \mathbf{F})$ by terminals $\mathcal{X}$ and $\mathcal{Y}$, respectively, i.e.,

$$\mathbb{P}\left(J = J_1 = J_2\right) \geq 1 - \epsilon.$$

Further, for $k \geq 1$, let

$$J_{1i} = J_1\left(X_{n(i-1)+1}^{ni}, \mathbf{F}_i\right),$$
$$J_{2i} = J_2\left(Y_{n(i-1)+1}^{ni}, \mathbf{F}_i\right), \quad 1 \leq i \leq k,$$

where $\mathbf{F}_i = \mathbf{F}\left(X_{n(i-1)+1}^{ni}, Y_{n(i-1)+1}^{ni}\right)$. For odd r, we find an r-interactive communication $\mathbf{F}''$ such that $\left(J_1^k, \mathbf{F}^k\right)$ is a ϵ-CR recoverable from $\mathbf{F}''$, for all k sufficiently large; the the SK K will be chosen to be a function of $\left(J_1^k, \mathbf{F}^k\right)$ of appropriate rate. The proof for even r is similar and is obtained by interchanging the roles of J_1 and J_2. In particular, by Lemma 6, for all k sufficiently large there exists an r-interactive communication $\mathbf{F}'$ such that $\mathbf{F}^k$ is ϵ-CR recoverable from $\mathbf{F}'$ of rate given by (11). Next, from Fano's inequality

$$\frac{1}{n}\max\{H(J \mid J_1); H(J_1 \mid J_2)\} \leq \epsilon \log|\mathcal{X}||\mathcal{Y}| + h(\epsilon). \quad (24)$$

By the Slepian-Wolf theorem [15] there exists a mapping f of J_1^k of rate

$$\frac{1}{k}\log\|f\| \leq H(J_1 \mid J_2) + n\epsilon, \quad (25)$$

such that

$$J_1^k \text{ is } \epsilon\text{-recoverable from } \left(f\left(J_1^k\right), J_2^k\right), \quad (26)$$

for all k sufficiently large. It follows from (24), (25) that

$$\frac{1}{nk}\log\|f\| \leq \epsilon + \epsilon\log|\mathcal{X}||\mathcal{Y}| + h(\epsilon). \quad (27)$$

For $N = nk$, we define the r-interactive communication $\mathbf{F}'' = \mathbf{F}''\left(X^N, Y^N\right)$ as

$$F_i'' = F_i', \quad 1 \leq i \leq r-1,$$
$$F_k'' = F_r', f(J_1^k), \quad i = r,$$

Thus, $\left(J_1^k, \mathbf{F}^k\right)$ is 2ϵ-CR recoverable from $\mathbf{F}''$, where, by (11) and (27), the rate of communication $\mathbf{F}''$ is bounded by

$$\frac{1}{nk}\log\|\mathbf{F}''\| \leq \frac{1}{n}\left[H\left(\mathbf{F}|X^n\right) + H\left(\mathbf{F}|Y^n\right)\right] + 2\epsilon + \epsilon\log|\mathcal{X}||\mathcal{Y}| + h(\epsilon). \quad (28)$$

Finally, to construct the SK $K = K\left(J_1^k, \mathbf{F}^k\right)$, using the corollary of Balanced Coloring Lemma in [5, Lemma B.3], with

$$U = (J_1, \mathbf{F}),\ V = \phi,\ n = k,\ g = \mathbf{F}',$$

we get from (28) that there exists a function K of $J_1^k, \mathbf{F}^k$ such that

$$\frac{1}{k}\log\|K\| \geq H(U) - \frac{1}{k}\log\|\mathbf{F}''\| \geq H(J_1, \mathbf{F}) - H(\mathbf{F} \mid X^n) - H(\mathbf{F} \mid Y^n) - n(2\epsilon + \epsilon\log|\mathcal{X}||\mathcal{Y}| + h(\epsilon)), \quad (29)$$

and

$$I(K \wedge \mathbf{F}') \leq \exp(-ck),$$

where $c > 0$, for all sufficiently large k. We get from (29) and (13) that the rate of K is bounded below as follows:

$$\frac{1}{nk}\log\|K\| \geq I(X \wedge Y) - \frac{1}{n}I\left(X^n \wedge Y^n \mid J_1, \mathbf{F}\right) - 2\epsilon - \epsilon\log|\mathcal{X}||\mathcal{Y}| - h(\epsilon). \quad (30)$$

Observe that

$$\begin{aligned} I(X^n \wedge Y^n \mid J, \mathbf{F}) &= I(J_1, X^n \wedge Y^n \mid J, \mathbf{F}) \\ &\geq I(X^n \wedge Y^n \mid J, J_1, \mathbf{F}) \\ &\geq I(X^n \wedge Y^n \mid J_1, \mathbf{F}) - H(J \mid J_1), \end{aligned}$$

which along with (24), and the fact that $L = (J, \mathbf{F})$ satisfies (3), yields

$$\frac{1}{n}I(X^n \wedge Y^n \mid J_1, \mathbf{F}) \leq \epsilon + \epsilon\log|\mathcal{X}||\mathcal{Y}| + h(\epsilon). \quad (31)$$

Upon combining (30) and (31) we get,

$$\frac{1}{nk}\log\|K\| \geq I(X \wedge Y) - 3\epsilon - 2\epsilon\log|\mathcal{X}||\mathcal{Y}| - 2h(\epsilon).$$

Thus, for $\Delta(\epsilon) = 4\epsilon + 2\epsilon\log|\mathcal{X}||\mathcal{Y}| + 2h(\epsilon)$ K is a $\Delta(\epsilon)$-SK of rate $(1/nk)\log\|K\| \geq I(X \wedge Y) - \Delta(\epsilon)$, recoverable from r-interactive communication $\mathbf{F}''$, which with (28), completes the proof. □

B. Proof of Theorem 4

Achievability. Consider rvs $U_1, ..., U_r$ satisfying conditions (P1)-(P3) in the statement of Theorem 4. It suffices to show for every $0 < \epsilon < 1$, for some $n \geq 1$, there exists an r-interactive communication $\mathbf{F}$, and ϵ-CR J recoverable from $\mathbf{F}$, such that

$$I(X, Y \wedge U^r) - \epsilon \leq \frac{1}{n}H(J, \mathbf{F}) \leq I(X, Y \wedge U^r) + \epsilon, \quad (32)$$

and

$$\frac{1}{n}H(\mathbf{F}) \leq I(X, Y \wedge U^r) - I(X \wedge Y) + \epsilon, \quad (33)$$

since from (16), (32) and (33), we have

$$\begin{aligned} &\frac{1}{n}I\left(X^n \wedge Y^n \mid J, \mathbf{F}\right) \\ &\leq \frac{1}{n}H(\mathbf{F}) - \frac{1}{n}H(J, \mathbf{F}) + I(X \wedge Y) + \delta(\epsilon) \\ &\leq 2\epsilon + \delta(\epsilon). \end{aligned}$$

We show below that

$$\begin{aligned} &I(X, Y \wedge U^r) - I(X \wedge Y) \quad (34) \\ &= \sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(X \wedge U_{2i+1} \mid Y, U^{2i}) \\ &\quad + \sum_{i=1}^{\lfloor r/2 \rfloor} I(Y \wedge U_{2i} \mid X, U^{2i-1}). \quad (35) \end{aligned}$$

Thus, the proof will be completed upon showing that there exists an ϵ-CR J, recoverable from $\mathbf{F}$ of rate

$$\frac{1}{n}H(\mathbf{F}) \leq \sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(X \wedge U_{2i+1} \mid Y, U^{2i}) + \sum_{i=1}^{\lfloor r/2 \rfloor} I(Y \wedge U_{2i} \mid X, U^{2i-1}) + \epsilon, \tag{36}$$

such that $(J, \mathbf{F})$ satisfies (32). For $r = 2$, such a construction was given by Ahlswede-Csiszár [2, Theorem 4.4]. (In their construction, $\mathbf{F}$ was additionally a function of J.) The extension of their construction to a general r is straightforward, and is relegated to the appendix.

It remains to prove (35). Note

$$\begin{aligned} & I(X, Y \wedge U^r) - \sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(X \wedge U_{2i+1} \mid Y, U^{2i}) \\ & \qquad - \sum_{i=1}^{\lfloor r/2 \rfloor} I(Y \wedge U_{2i} \mid X, U^{2i-1}) \\ & = \sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(Y \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=1}^{\lfloor r/2 \rfloor} I(X \wedge U_{2i} \mid U^{2i-1}). \end{aligned} \tag{37}$$

Further, from conditions (P1)-(P3) it follows that

$$\begin{aligned} & \sum_{i=0}^{\lfloor (r-1)/2 \rfloor} I(Y \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=1}^{\lfloor r/2 \rfloor} I(X \wedge U_{2i} \mid U^{2i-1}) \\ & \qquad - I(Y \wedge X) \\ & = \sum_{i=1}^{\lfloor (r-1)/2 \rfloor} I(Y \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=2}^{\lfloor r/2 \rfloor} I(X \wedge U_{2i} \mid U^{2i-1}) \\ & \quad + I(X \wedge U_2 \mid U_1) + I(Y \wedge U_1) - I(Y \wedge X) \\ & = \sum_{i=1}^{\lfloor (r-1)/2 \rfloor} I(Y \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=2}^{\lfloor r/2 \rfloor} I(X \wedge U_{2i} \mid U^{2i-1}) \\ & \quad + I(X \wedge U_2 \mid U_1) - I(X \wedge Y \mid U_1) \\ & = \sum_{i=1}^{\lfloor (r-1)/2 \rfloor} I(Y \wedge U_{2i+1} \mid U^{2i}) + \sum_{i=2}^{\lfloor r/2 \rfloor} I(X \wedge U_{2i} \mid U^{2i-1}) \\ & \qquad - I(X \wedge Y \mid U_1, U_2) \\ & = \ldots = -I(X \wedge Y \mid U^r) = 0. \end{aligned} \tag{38}$$

Combining (37) and (38) we get (35).

Converse. Let $R \geq 0$ be an achievable r-interactive CI rate. Then, for all $0 < \epsilon < 1$, for some $n \geq 1$, there exists an r-interactive communication $\mathbf{F}$, and ϵ-CR J recoverable from $\mathbf{F}$, such that $(1/n)H(J, \mathbf{F}) \leq R + \epsilon$ and $L = (J, \mathbf{F})$ satisfies (3). Let J be recovered as $J_1 = J_1(X^n, \mathbf{F})$ and $J_2 = J_2(Y^n, \mathbf{F})$ by terminals $\mathcal{X}$ and $\mathcal{Y}$, respectively, i.e., $\mathbb{P}(J = J_1 = J_2) \geq 1 - \epsilon$. Further, let rv T be distributed uniformly over the set $\{1, ..., n\}$. Define rvs U^r as follows:

$$\begin{aligned} U_1 &= F_1, X^{T-1}, Y_{T+1}^n, T, \\ U_i &= F_i, \qquad 2 \leq i < r, \\ U_r &= \begin{cases} (F_r, J_1), & r \text{ odd}, \\ (F_r, J_2), & r \text{ even}. \end{cases} \end{aligned}$$

We complete the proof for odd r; the proof for even r can be completed similarly. It was shown by Kaspi [10, equations (3.10)-(3.13)] that

$$\begin{aligned} & U_{2i+1} \multimap X_T, U^{2i} \multimap Y_T, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor, \\ & U_{2i} \multimap Y_T, U^{2i-1} \multimap X_T, \quad 1 \leq i \leq \lfloor r/2 \rfloor. \end{aligned}$$

Next, note from (31) that

$$\begin{aligned} & \epsilon + \epsilon \log |\mathcal{X}||\mathcal{Y}| + h(\epsilon) \\ & \geq \frac{1}{n} I(X^n \wedge Y^n \mid J_1, \mathbf{F}) \\ & \geq \frac{1}{n} \sum_{i=1}^{n} I(X_i \wedge Y^n \mid X^{i-1}, J_1, \mathbf{F}) \\ & \geq \frac{1}{n} \sum_{i=1}^{n} I(X_i \wedge Y_i \mid X^{i-1}, Y_{i+1}^n, J_1, \mathbf{F}) \\ & = I(X_T \wedge Y_T \mid U^r). \end{aligned} \tag{39}$$

Similarly, it holds that

$$\epsilon + \epsilon \log |\mathcal{X}||\mathcal{Y}| + h(\epsilon) \geq I(X_T \wedge Y_{T+1}^n \mid X^{T-1}, J_1, \mathbf{F}, T). \tag{40}$$

The entropy rate of $(J, \mathbf{F})$ is now bounded as

$$\begin{aligned} & \frac{1}{n} H(J, \mathbf{F}) \\ & \geq \frac{1}{n} H(J_1, \mathbf{F}) - \frac{1}{n} H(J_1 \mid J) \\ & \geq \frac{1}{n} H(J_1, \mathbf{F}) - \epsilon \log |\mathcal{X}||\mathcal{Y}| - h(\epsilon) \\ & = \frac{1}{n} I(X^n, Y^n \wedge J_1, \mathbf{F}) - \epsilon \log |\mathcal{X}||\mathcal{Y}| - h(\epsilon) \\ & = H(X_T, Y_T) - \frac{1}{n} H(X^n \mid J_1, \mathbf{F}) \\ & \qquad - \frac{1}{n} H(Y^n \mid X^n, J_1, \mathbf{F}) - \epsilon \log |\mathcal{X}||\mathcal{Y}| - h(\epsilon) \\ & = H(X_T, Y_T) - H(X_T \mid X^{T-1}, J_1, \mathbf{F}, T) \\ & \qquad - H(Y_T \mid X^{T-1}, Y_{T+1}^n, X_T, X_{T+1}^n, J_1, \mathbf{F}, T) \\ & \qquad - \epsilon \log |\mathcal{X}||\mathcal{Y}| - h(\epsilon) \\ & \geq I(X_T, Y_T \wedge U^r) - \epsilon - 2\epsilon \log |\mathcal{X}||\mathcal{Y}| - 2h(\epsilon), \end{aligned}$$

where the second inequality follows from Fano's inequality, and the last inequality follows from (40). Consequently,

$$\begin{aligned} R & \geq \frac{1}{n} H(J, \mathbf{F}) - \epsilon \\ & \geq I(X_T, Y_T \wedge U^r) - 2(\epsilon + \epsilon \log |\mathcal{X}||\mathcal{Y}| + h(\epsilon)). \end{aligned} \tag{41}$$

We now replace the rvs $U_1, ..., U_r$ with those taking values in finites sets $\mathcal{U}_1, ..., \mathcal{U}_r$, respectively, with $\mathcal{U}_1, ..., \mathcal{U}_r$ satisfying the cardinality bounds in condition (iii). Similar bounds were derived in the context of interactive function computation in

[11]. For $1 \leq l \leq r$, assume that rvs $\mathcal{U}_1, ..., \mathcal{U}_{l-1}$ satisfy the cardinality bounds. We consider odd l; the steps for even l are similar. If the rv U_l does not satisfy the cardinality bound, from the Support Lemma [4, Lemma 15.4], we can replace it with another rv $\tilde{U}_l$ that takes less than or equal to $|\mathcal{X}| \prod_{i=1}^{l-1} |\mathcal{U}_i| + 1$ values, while keeping the following quantities unchanged:

$$\mathtt{P}_{X_T U^{l-1}}, \ I(X_T \wedge Y_T \mid U^r), \text{ and } I(X_T, Y_T \wedge U^r).$$

Note that we have only altered $\mathtt{P}_{U_l}$ in the joint pmf $\mathtt{P}_{X_T Y_T U^r} = \mathtt{P}_{U_l} \mathtt{P}_{X_T U^{l-1}|U_l} \mathtt{P}_{Y_T|XU^{l-1}}$. Hence, the Markov relations in (P1) remain unaltered. Furthermore, $\mathtt{P}_{X_T Y_T} = \mathtt{P}_{XY}$. Finally, since the set of pmfs on a finite alphabet is compact, and the choice of ϵ above was arbitrary, it follows upon taking $\epsilon \to 0$ in (39) and (41) that there exists U_1^r satisfying (P1)-(P3) such that

$$R \geq I(X, Y \wedge U^r),$$

which completes the proof. □

V. Can interaction reduce the communication rate?

It is well known that the SK capacity can be attained by using a simple one-way communication from terminal $\mathcal{X}$ to terminal $\mathcal{Y}$ (or from $\mathcal{Y}$ to $\mathcal{X}$). Here we derive the minimum rate R_{NI} of such noninteractive communication using the expression for $CI_i^r(X;Y)$ in (9). Since this expression has a *double Markov structure*, it can be simplified by the following observation (see [4, Problem 16.25]): If rvs U, X, Y satisfy

$$U \multimap X \multimap Y, \quad X \multimap U \multimap Y, \tag{42}$$

then there exist functions $f = f(U)$ and $g = g(X)$ such that

(i) $\mathbb{P}(f(U) = g(X)) = 1$;
(ii) $X \multimap g(X) \multimap Y$.

In particular, for rvs U, X, Y that satisfy (42), it follows from (i) above that

$$I(X, Y \wedge U) = I(X \wedge U) \geq I(g(X) \wedge f(U)) = H(g(X)).$$

Turning to (9), for rvs U^r with r odd, the observations above applied to the rvs X and Y conditioned on each realization $U^{r-1} = u^{r-1}$ implies that there exists a function $g_1 = g_1\left(X, U^{r-1}\right)$ such that

$$X \multimap g\left(X, U^{r-1}\right), U^{r-1} \multimap Y, \tag{43}$$

and

$$I\left(X, Y \wedge U^r\right) \geq I\left(X, Y \wedge U^{r-1}\right) + H\left(g\left(X, U^{r-1}\right) \mid U^{r-1}\right),$$

where rv U^{r-1} satisfies (P1), (P3). Similar observations hold for even r. Thus, for the minimization in (9), conditioned on arbitrarily chosen rvs U^{r-1} satisfying (P1), (P3), the rv U_r is selected as a *sufficient statistic for Y given the observation X* (sufficient statistic for X given the observation Y) when r is odd (r is even). Specifically, for $r = 1$, we have

$$CI_i^1(X;Y) = \min_{X \multimap g_1(X) \multimap Y} H\left(g_1(X)\right), \tag{44}$$

and

$$CI_i^1(Y;X) = \min_{Y \multimap g_2(Y) \multimap X} H\left(g_2(Y)\right). \tag{45}$$

The answer to the optimization problems in (44) and (45) can be given explicitly. In fact, we specify next a minimal sufficient statistic for Y on the basis of X. Define an equivalence relation on $\mathcal{X}$ as follows:

$$x \sim x' \Leftrightarrow \mathtt{P}_{Y|X}(y \mid x) = \mathtt{P}_{Y|X}(y \mid x'), \quad y \in \mathcal{Y}. \tag{46}$$

Let g_1^* be the function corresponding to the equivalence classes of $\sim$. We claim that g_1^* is a minimal sufficient statistic for Y on the basis of X. This expression for the minimal sufficient statistic was also given in [9, Lemma 3.5(4)]. Specifically, $X \multimap g_1^*(X) \multimap Y$ since with $g_1^*(X) = c$, say, we have

$$\begin{aligned}
&\mathtt{P}_{Y|g_1^*(X)}(y \mid c) \\
&= \sum_{x \in \mathcal{X}} \mathtt{P}_{Y,X|g_1^*(X)}(y, x \mid c) \\
&= \sum_{x: g_1^*(x) = c} \mathtt{P}_{X|g_1^*(X)}(x \mid c)\, \mathtt{P}_{Y|X,g_1^*(X)}(y \mid x, c) \\
&= \mathtt{P}_{Y|X,g_1^*(X)}(y \mid x, c), \quad \forall\, x \text{ with } g_1^*(x) = c.
\end{aligned}$$

Also, if $g_1(X)$ satisfies $X \multimap g_1(X) \multimap Y$ then g_1^* is a function of g_1. To see this, let $g_1(x) = g_1(x') = c$ for some $x, x' \in \mathcal{X}$. Then,

$$\mathtt{P}_{Y|g_1(X)}(y \mid c) = \mathtt{P}_{Y|X}(y \mid x) = \mathtt{P}_{Y|X}(y \mid x'), \quad y \in \mathcal{Y},$$

so that $g_1^*(x) = g_1^*(x')$. Since g_1^* is a minimal sufficient statistic for Y on the basis of X, it follows from (44) that

$$CI_i^1(X;Y) = H\left(g_1^*(X)\right),$$

and similarly, with $g_2^*(Y)$ defined analogously,

$$CI_i^1(Y;X) = H\left(g_2^*(Y)\right).$$

Therefore, from (7), the minimum rate R_{NI} of a noninteractive communication for generating a maximum rate SK is given by

$$R_{NI} = \min\left\{H\left(g_1^*(X)\right), H\left(g_1^*(X)\right)\right\} - I(X \wedge Y). \tag{47}$$

From the expression for R_{NI}, it is clear that the rate of noninteractive communication can be reduced by replacing X and Y with their respective minimal sufficient statistics $g_1^*(X)$ and $g_2^*(Y)$. Can the rate of communication required for generating an optimum rate SK be reduced by resorting to complex interactive communication protocols defined in Section II? To answer this question we must compare the expression for R_{NI} with R_{SK}. Specifically, from Theorem 3 and the Corollary following it, interaction reduces the rate of communication iff, for some $r > 1$,

$$CI_i^r(X;Y) < \min\left\{H\left(g_1^*(X)\right), H\left(g_1^*(X)\right)\right\}, \tag{48}$$

where g_1^* and g_2^* are as in (47); interaction does not help iff

$$CI_i(X \wedge Y) = \min\left\{H\left(g_1^*(X)\right), H\left(g_1^*(X)\right)\right\}.$$

Note that instead of comparing with $CI_i^r(X;Y)$ in (48), we can also compare with $CI_i^r(Y;X)$.

We shall explore this question here, and give an example where the answer is in the affirmative. In fact, we first show that interaction does not help in the case of binary symmetric sources. Then we give an example where interaction does help.

A. Binary Symmetric Sources

For binary rvs X and Y, we note a property of rvs U^r that satisfy the conditions (P1)-(P3) in Theorem 4.

Lemma 8. *Let X and Y be $\{0,1\}$ valued rvs with $I(X\wedge Y)\neq 0$. Then, for rvs $U_1,...,U_r$ that satisfy the conditions (P1)-(P3) in Theorem 4, for every realization $u_1,...,u_r$ of $U_1,...,U_r$, one of the following holds:*

$$H(X\mid U^r=u^r)=0,\quad \textit{or}\quad H(Y\mid U^r=u^r)=0. \tag{49}$$

Proof. Given a sequence u^r, assume that

$$H(X\mid U^r=u^r)>0\ \text{ and }\ H(Y\mid U^r=u^r)>0,$$

which is equivalent to

$$\begin{aligned}&\mathtt{P}_{X|U^r}\left(1\mid u^r\right)\mathtt{P}_{X|U^r}\left(0\mid u^r\right)>0\ \text{ and}\\ &\mathtt{P}_{Y|U^r}\left(1\mid u^r\right)\mathtt{P}_{Y|U^r}\left(0\mid u^r\right)>0.\end{aligned} \tag{50}$$

We consider the case when r is even; the case of odd r is handled similarly. From the Markov conditions $X\multimap U^r\multimap Y$ and $X\multimap Y,U^{r-1}\multimap U_r$, we have

$$\begin{aligned}&\mathtt{P}_{X,Y|U^r}\left(x,y\mid u^r\right)\\ &=\mathtt{P}_{X|U^r}\left(x\mid u^r\right)\mathtt{P}_{Y|U^r}\left(y\mid u^r\right)\\ &=\mathtt{P}_{X|Y,U^{r-1}}\left(x\mid y,u^{r-1}\right)\mathtt{P}_{Y|U^r}\left(y\mid u^r\right),\quad x,y\in\{0,1\}.\end{aligned}$$

Since $\mathtt{P}_{Y|U^r}\left(y\mid u^r\right)>0$ from (50), we have

$$\mathtt{P}_{X|U^r}\left(x\mid u^r\right)=\mathtt{P}_{X|Y,U^{r-1}}\left(x\mid y,u^{r-1}\right),\quad x,y\in\{0,1\},$$

which further implies

$$\mathtt{P}_{X|Y,U^{r-1}}\left(x\mid 1,u^{r-1}\right)=\mathtt{P}_{X|Y,U^{r-1}}\left(x\mid 0,u^{r-1}\right),\quad x\in\{0,1\}.$$

Hence, $I(X\wedge Y\mid U^{r-1}=u^{r-1})=0$. Noting from (50) that

$$\mathtt{P}_{X|U^{r-1}}\left(1\mid u^{r-1}\right)\mathtt{P}_{X|U^{r-1}}\left(0\mid u^{r-1}\right)>0,$$

we can do the same analysis as above, again for $r-1$. Upon repeating this process r times we get $I(X\wedge Y)=0$, which is a contradiction. Therefore, either $H(X\mid U^r=u^r)=0$ or $H(Y\mid U^r=u^r)=0$ holds. □

Note that

$$CI_i^r(X;Y)=H(X,Y)-\max_{U^r}H(X,Y\mid U^r),$$

where the max is taken over rvs U^r as in Theorem 4. If $H(X\mid U^i=u^i)=0$, it follows that

$$\begin{aligned}I\left(X\wedge Y\mid U^i=u^i\right)&=0,\ \text{and}\\ H(X,Y\mid U^i=u^i,U_{i+1}^r)&=H(Y\mid U^i=u^i,U_{i+1}^r)\\ &\le H(Y\mid U^i=u^i).\end{aligned} \tag{51}$$

Similarly, $H(Y\mid U^i=u^i)=0$ implies

$$\begin{aligned}I\left(X\wedge Y\mid U^i=u^i\right)&=0,\ \text{and}\\ H(X,Y\mid U^i=u^i,U_{i+1}^r)&\le H(X\mid U^i=u^i).\end{aligned} \tag{52}$$

For a sequence u^r with $\mathtt{P}_{U^r}\left(u^r\right)>0$, let $\tau(u^r)$ be the minimum value of i such that

$$H(X\mid U^i=u^i)=0\text{ or }H(Y\mid U^i=u^i)=0;$$

if X and Y are independent, $\tau(u^r)=0$. Note that τ is a stopping-time adapted to $U_1,...,U_r$. Then, from (51), (52), $CI_i^r(X;Y)$ remains unchanged if we restrict the support of U^r to sequences u^r with $u_i=\phi$ for all $i>\tau(u^r)$. Furthermore, the Markov condition (P1) implies that if for a sequence u^r, $\tau=\tau(u^r)$ is odd then

$$\mathtt{P}_{Y|X,U^\tau}\left(y\mid x,u^\tau\right)=\mathtt{P}_{Y|X,U^{\tau-1}}\left(y\mid x,u^{\tau-1}\right),$$

and so if

$$\mathtt{P}_{X|U^\tau}\left(1\mid u^\tau\right)\mathtt{P}_{X|U^\tau}\left(0\mid u^\tau\right)>0,$$

it holds from the definition of τ that

$$\mathtt{P}_{Y|U^\tau}\left(1\mid u^\tau\right)\mathtt{P}_{Y|U^\tau}\left(0\mid u^\tau\right)>0,$$

which is a contradiction. Therefore, we have $H(X\mid U^\tau=u^\tau)=0$. Similarly, $H(Y\mid U^\tau=u^\tau)=0$ holds for even τ. To summarize,

$$CI_i^r(X;Y)=\min_{U^\tau}I\left(X,Y\wedge U^\tau\right), \tag{53}$$

where U^r are rvs satisfying (P1)-(P3), and τ is the stopping-time defined above.

We show next that for binary symmetric sources, interaction can never reduce the rate of communication for optimum rate SK generation. In fact, *we conjecture that for any binary rvs X,Y, $R_{NI}=R_{SK}$.*

Theorem 9. *Let X and Y be $\{0,1\}$-valued rvs, with*

$$\begin{aligned}&\mathbb{P}\left(X=0,Y=0\right)=\mathbb{P}\left(X=1,Y=1\right)=\frac{1}{2}(1-\delta),\\ &\mathbb{P}\left(X=0,Y=1\right)=\mathbb{P}\left(X=1,Y=0\right)=\frac{1}{2}\delta,\quad 0<\delta<\frac{1}{2}.\end{aligned} \tag{54}$$

Then,

$$CI_i(X\wedge Y)=\min\{H(X);H(Y)\},$$

i.e., interaction does not help to reduce the communication required for optimum rate SK generation.

Remark. As a consequence of Theorem 9, for sources with joint distribution as in (54), the second inequality in (6) can be strict. Specifically, it was noted by Wyner (see the discussion following equation (1.19) in [16]) that for binary symmetric sources, $CI_W(X\wedge Y)<1$. From Theorem 9, we have

$$CI_i(X\wedge Y)=\min\{H(X);H(Y)\}=1.$$

Thus, for such sources, $CI_W(X\wedge Y)<CI_i(X\wedge Y)$.

Proof. Denote by $\mathcal{U}_0^r$ the following set of stopped sequences in $\mathcal{U}^r$:
For $i\le r$, for a sequence $u^r\in\mathcal{U}^r$ the stopped sequence $u^i\in\mathcal{U}_0^r$ if:

$$H\left(X\mid U^j=u^j\right)>0, H\left(Y\mid U^j=u^j\right)>0,\quad\forall\, j<i,\ \text{and}$$
$$H\left(X\mid U^i=u^i\right)=0\text{ or }H\left(Y\mid U^i=u^i\right)=0.$$

For $i\in\{0,1\}$, define the following subsets of $\mathcal{U}_0^r$:

$$\begin{aligned}\mathcal{U}_i^X&=\left\{u^\tau\in\mathcal{U}_0^r:\tau\text{ is odd},\mathtt{P}_{X|U^\tau}\left(i\mid u^\tau\right)=1\right\},\\ \mathcal{U}_i^Y&=\left\{u^\tau\in\mathcal{U}_0^r:\tau\text{ is even},\mathtt{P}_{Y|U^\tau}\left(i\mid u^\tau\right)=1\right\}.\end{aligned}$$

By their definition the sets $\mathcal{U}_0^X, \mathcal{U}_1^X, \mathcal{U}_0^Y$, and $\mathcal{U}_1^Y$ are disjoint, whereby we have

$$\begin{aligned}\mathtt{P}_{U^\tau}\left(\mathcal{U}_0^r\right) &= \mathtt{P}_{U^r}\left(\mathcal{U}_0^X \bigcup \mathcal{U}_1^X \bigcup \mathcal{U}_0^Y \bigcup \mathcal{U}_1^Y\right) \\ &= \sum_{i=0}^{1}\left[\mathtt{P}_{U^r}\left(\mathcal{U}_i^X\right) + \mathtt{P}_{U^r}\left(\mathcal{U}_i^Y\right)\right] = 1. \end{aligned} \tag{55}$$

For $u^\tau \in \mathcal{U}_0^r$, denote by $p(u^\tau)$ the probability $\mathtt{P}_{U^\tau}(u^\tau)$. Further, for $u^\tau \in \mathcal{U}_0^X \bigcup \mathcal{U}_1^X$, denote by $W^{u^\tau}: \mathcal{X} \to \mathcal{Y}$ the stochastic matrix corresponding to $\mathtt{P}_{Y|X,U^\tau}(\cdot \mid \cdot, u^\tau)$, and for $u^\tau \in \mathcal{U}_0^Y \bigcup \mathcal{U}_1^Y$, denote by $T^{u^\tau}: \mathcal{Y} \to \mathcal{X}$ the stochastic matrix corresponding to $\mathtt{P}_{X|Y,U^\tau}(\cdot \mid \cdot, u^\tau)$. With this notation, the following holds:

$$\begin{aligned}&\frac{1}{2}(1-\delta) \\ &= \mathtt{P}_{X,Y}(i,i) \\ &= \sum_{u^\tau \in \mathcal{U}_i^X} p(u^\tau) W^{u^\tau}(i \mid i) + \sum_{u^\tau \in \mathcal{U}_i^Y} p(u^\tau) T^{u^\tau}(i \mid i), \\ &\qquad\qquad i \in \{0,1\}, \end{aligned} \tag{56}$$

since the sets $\mathcal{U}_0^X, \mathcal{U}_1^X, \mathcal{U}_0^Y, \mathcal{U}_1^Y$ are disjoint. Upon adding (56) for $i = 0, 1$, we get

$$\begin{aligned}&\sum_{i=0}^{1}\left[\sum_{u^\tau \in \mathcal{U}_i^X} p(u^\tau) W^{u^\tau}(i \mid i) + \sum_{u^\tau \in \mathcal{U}_i^Y} p(u^\tau) T^{u^\tau}(i \mid i)\right] \\ &= (1-\delta).\end{aligned}$$

Furthermore, from (55) we get

$$1 = \sum_{i=0}^{1}\sum_{u^\tau \in \mathcal{U}_i^X} p(u^\tau) + \sum_{u^\tau \in \mathcal{U}_i^Y} p(u^\tau).$$

Therefore, since the function $g(z) = -z\log z$ is concave for $0 < z < 1$, the Jensen's inequality yields

$$\begin{aligned}g(1-\delta) \geq \sum_{i=0}^{1}\sum_{u^\tau \in \mathcal{U}_i^X} p(u^\tau) g\left(W^{u^\tau}(i \mid i)\right) + \\ \sum_{u^\tau \in \mathcal{U}_i^Y} p(u^\tau) g\left(T^{u^\tau}(i \mid i)\right)\end{aligned} \tag{57}$$

Similarly, using

$$\begin{aligned}\frac{1}{2}\delta &= \mathtt{P}_{X,Y}(i,j) \\ &= \sum_{u^\tau \in \mathcal{U}_i^X} p(u^\tau)\left(1 - W^{u^\tau}(i \mid i)\right) + \\ &\quad \sum_{u^\tau \in \mathcal{U}_j^Y} p(u^\tau)\left(1 - T^{u^\tau}(j \mid j)\right), \quad i \neq j, i, j \in \{0,1\},\end{aligned}$$

we get

$$\begin{aligned}g(\delta) \geq \sum_{i=0}^{1}\sum_{u^\tau \in \mathcal{U}_i^X} p(u^\tau) g\left(1 - W^{u^\tau}(i \mid i)\right) + \\ \sum_{u^\tau \in \mathcal{U}_i^Y} p(u^\tau) g\left(1 - T^{u^\tau}(i \mid i)\right)\end{aligned} \tag{58}$$

On adding (57) and (58) we get

$$\begin{aligned}h(\delta) &= g(\delta) + g(1-\delta) \\ &\geq \sum_{i=0}^{1}\sum_{u^\tau \in \mathcal{U}_i^X} p(u^\tau) h\left(W^{u^\tau}(i \mid i)\right) + \\ &\qquad \sum_{u^\tau \in \mathcal{U}_i^Y} p(u^\tau) h\left(T^{u^\tau}(i \mid i)\right),\end{aligned}$$

where h is the binary entropy function. Note that the right side above equals $H(X, Y \mid U^\tau)$, which yields

$$h(\delta) = \max\{H(X \mid Y); H(Y \mid X)\} \geq H(X, Y \mid U^\tau).$$

Since rvs U^r above were arbitrary, we have from (53),

$$\begin{aligned}CI_i^r(X;Y) &\geq H(X,Y) - \max\{H(X \mid Y); H(Y \mid X\}) \\ &= \min\{H(X); H(Y)\}.\end{aligned}$$

Combining this with (6), we obtain

$$CI_i^r(X;Y) = \min\{H(X); H(Y)\}.$$

□

B. An example where interaction does help

Consider rvs X and Y with $\mathcal{X} = \mathcal{Y} = \{0,1,2\}$, and with joint pmf:

$$\begin{bmatrix} a & a & a \\ b & a & a \\ a & c & a \end{bmatrix},$$

where a, b, c are nonnegative, $7a+b+c = 1$, and $c \neq a$, which holds iff $b \neq 1 - 8a$. Assume that

$$2a > b > a. \tag{59}$$

From (48), to show that interaction helps, it suffices to find rvs $U_1, ..., U_r$ satisfying (P1)-(P3) such that

$$I\left(X, Y \wedge U_1, ..., U_r\right) < \min\left\{H\left(g_1^*(X)\right), H\left(g_2^*(Y)\right)\right\}, \tag{60}$$

where g_1^* and g_2^* are as in (47). From (46), $g_1^*(x) = g_1^*(x')$ iff

$$\frac{\mathtt{P}_{Y,X}(y,x)}{\mathtt{P}_{Y,X}(y,x')} = \frac{\mathtt{P}_X(x)}{\mathtt{P}_X(x')}, \quad y \in \mathcal{Y}, \tag{61}$$

i.e., the ratio $\frac{\mathtt{P}_{Y,X}(y,x)}{\mathtt{P}_{Y,X}(y,x')}$ does not depend on y. Therefore, for the pmf above, $g_1^*(X)$ and $g_2^*(Y)$ are equivalent to X and Y, respectively. Thus,

$$\min\left\{H\left(g_1^*(X)\right), H\left(g_2^*(Y)\right)\right\} = \min\{H(X); H(Y)\},$$

where $H(X) = H(Y)$ for the given pmf.

Next, let $U_1 = f_1(X)$, $U_2 = f_2(Y, f_1(X))$, where f_1 and f_2 are given below:

$$f_1(x) = \begin{cases} 1, & x = 2, \\ 2, & x = 0, 1, \end{cases}$$

$f_2(y,1) = 0, \forall\, y \in \{0,1,2\}$, and $f_2(y,2) = \begin{cases} 1, & y = 0, \\ 2, & y = 1, 2. \end{cases}$

Clearly, U_1 and U_2 satisfy (P1) and (P2). For (P3), note that if $(U_1, U_2) = (1, 0)$, then $X = 2$, and if $(U_1, U_2) = (2, 1)$, then $Y = 0$. Finally, if $(U_1, U_2) = (2, 2)$, then $X \in \{0, 1\}$ and $Y \in \{1, 2\}$, implying

$$\begin{aligned}\mathtt{P}_{X,Y|U_1,U_2}(x, y \mid 2, 2) &= \frac{\mathtt{P}_{X,Y}(x, y)}{4a} \\ &= \frac{1}{4}, \quad \forall\ (x, y) \in \{0, 1\} \times \{1, 2\}.\end{aligned}$$

Therefore, $I(X \wedge Y \mid U_1, U_2) = 0$, and so U_1, U_2 satisfy (P3). We show that (60) holds for this choice of U_1, U_2. Specifically, $I(X, Y \wedge U_1, U_2) = H(U_1, U_2)$, and the following holds:

$$\begin{aligned}&H(Y) - H(U_1, U_2) \\ &= H(X) - H(U_1, U_2) \\ &= H(X|U_1) - H(U_2|U_1) \\ &= \mathbb{P}(f_1(X) = 2) \Big[H(X|f_1(X) = 2) \\ &\qquad\qquad - H(f_2(2, Y)|f_1(X) = 2)\Big] \\ &= (5a + b)\left[h\left(\mathtt{P}_{X|f_1(X)}(0|2)\right) - h\left(\mathtt{P}_{Y|f_1(X)}(0|2)\right)\right] \\ &= (5a + b)\left[h\left(\frac{3a}{5a + b}\right) - h\left(\frac{a + b}{5a + b}\right)\right].\end{aligned}$$

Then, from (59),

$$\frac{a + b}{5a + b} < \frac{3a}{5a + b} < \frac{1}{2},$$

which implies (60) for U_1, U_2.

VI. Sufficient statistics and common information quantities

In this work we encountered three CI quantities: Shannon's mutual information $I(X \wedge Y)$, Wyner's CI $CI_W(X \wedge Y)$, and interactive CI $CI_i(X \wedge Y)$. In fact, the first notion of CI was given by Gács and Körner in the seminal work [8]. In particular, they specified the maximal common function of X and Y, denoted here as $\mathtt{mcf}(X, Y)$, such that any other common function of X and Y is a function of $\mathtt{mcf}(X, Y)$; the Gács-Körner CI is given by $H(\mathtt{mcf}(X, Y))$. The following inequality ensues (see [8], [16], and inequality (6)):

$$H(\mathtt{mcf}(X, Y)) \le I(X \wedge Y) \le CI_W(X \wedge Y) \le CI_i(X \wedge Y).$$

Since any good notion of CI between rvs X and Y measures the correlation between X and Y, it is reasonable to expect the CI to remain unchanged if X and Y are replaced by their respective sufficient statistics. The following theorem establishes this for the quantities $H(\mathtt{mcf}(X, Y)), I(X \wedge Y), CI_W(X \wedge Y)$, and $H(\mathtt{mcf}(X, Y))$.

Theorem 10. *For rvs X and Y, let functions g_1 of X and g_2 of Y be such that $X \multimap g_1(X) \multimap Y$ and $X \multimap g_2(Y) \multimap Y$. Then the following relations hold:*

$$\begin{aligned}H(\mathtt{mcf}(X, Y)) &= H(\mathtt{mcf}(g_1(X), g_2(Y))), \\ I(X \wedge Y) &= I(g_1(X) \wedge g_2(Y)), \\ CI_W(X \wedge Y) &= CI(g_1(X) \wedge g_2(Y)), \\ CI_i^r(X; Y) &= CI_i^r(g_1(X); g_2(Y)), \quad r \ge 1, \\ CI_i(X \wedge Y) &= CI_i(g_1(X) \wedge g_2(Y)).\end{aligned}$$

Remark. (i) Theorem 10 implies that the minimum rate of communication for generating a maximum rate secret key remains unchanged if X and Y are replaced by $g_1(X)$ and $g_2(Y)$ as above, respectively.
(ii) Note that $g_1(X)$ and $g_2(Y)$ above are, respectively, functions of $g_1^*(X)$ and $g_2^*(Y)$ defined through (46).

Proof. First note that

$$I(X \wedge Y) = I(g_1(X) \wedge Y) = I(g_1(X) \wedge g_2(Y)). \tag{62}$$

Next, consider the interactive CI. From (62), any protocol that generates an optimum rate SK for the sources $g_1(X)$ and $g_2(Y)$ also generates an optimum rate SK for the sources X and Y. Thus, the minimum communication rate for prior protocols is bounded below by the minimum communication rate for the latter protocols, so that by Theorem 3,

$$\begin{aligned}&CI_i^r(g_1(X); g_2(Y)) - I(g_1(X) \wedge g_2(Y)) \\ &\ge CI_i^r(X; Y) - I(X \wedge Y),\end{aligned}$$

which, by (62), is

$$CI_i^r(g_1(X); g_2(Y)) \ge CI_i^r(X; Y). \tag{63}$$

In fact, (63) holds with equality: We claim that any choice of rvs U^r that satisfy (P1)-(P3) also satisfy the following Markov relations:

$$\begin{aligned}&U_{2i+1} \multimap g_1(X), U^{2i} \multimap g_2(Y), \quad 0 \le i \le \lfloor (r-1)/2 \rfloor, \\ &U_{2i} \multimap g_2(Y), U^{2i-1} \multimap g_1(X), \quad 1 \le i \le \lfloor r/2 \rfloor, \\ &g_1(X) \multimap U^r \multimap g_2(Y).\end{aligned} \tag{64}$$

It follows that

$$\begin{aligned}CI_i^r(g_1(X); g_2(Y)) &\le I(g_1(X), g_2(Y) \wedge U^r) \\ &\le I(X, Y \wedge U^r),\end{aligned}$$

and consequently,

$$CI_i^r(g_1(X); g_2(Y)) \le CI_i^r(X; Y).$$

Thus, by (63),

$$CI_i^r(g_1(X); g_2(Y)) = CI_i^r(X; Y). \tag{65}$$

Taking the limit $r \to \infty$ we get

$$CI_i(g_1(X) \wedge g_2(Y)) = CI_i(X \wedge Y).$$

It remains to establish (64); instead, using induction we establish the following stronger Markov relations: For $1 \le i \le r$,

$$\begin{aligned}&U_i \multimap g_1(X), U^{i-1} \multimap Y, \quad i \text{ odd}, \\ &U_i \multimap g_2(Y), U^{i-1} \multimap X, \quad i \text{ even}, \\ &X \multimap g_1(X), U^i \multimap Y \text{ and } X \multimap g_2(Y), U^i \multimap Y.\end{aligned} \tag{66}$$

Clearly, (66) implies the first two Markov relations in (64). The last Markov chain in (64) follows upon observing

$$0 = I(X \wedge Y \mid U^r) \ge I(g_1(X) \wedge g_2(Y) \mid U^r).$$

To see that (66) holds for $i = 1$ note that

$$\begin{aligned}&I(X \wedge Y \mid g_1(X), U_1) \\ &\le I(X \wedge Y \mid g_1(X)) + I(U_1 \wedge Y \mid g_1(X), X) = 0,\end{aligned}$$

and

$$I\left(X \wedge Y \mid g_2(Y), U_1\right) \leq I\left(X \wedge Y \mid g_2(Y)\right) + I\left(U_1 \wedge Y, g_2(Y) \mid X\right) = 0.$$

Next, assume that (66) holds for an even i. Then, from (P1) we get:

$$\begin{aligned} & I\left(Y \wedge U_{i+1} \mid X, U^i\right) = 0 \\ \Leftrightarrow & I\left(Y \wedge U_{i+1} \mid X, g_1(X), U^i\right) = 0 \\ \Leftrightarrow & I\left(Y \wedge X, U_{i+1} \mid g_1(X), U^i\right) = I\left(Y \wedge X \mid g_1(X), U^i\right) = 0, \end{aligned}$$

where the last equality follows from (66). From the last inequality above we have

$$U_{i+1} \multimap g_1(X), U^i \multimap Y \quad \text{and} \quad X \multimap g_1(X), U^{i+1} \multimap Y.$$

Furthermore, it also follows from (66) that

$$\begin{aligned} I\left(X \wedge Y \mid g_2(Y), U^{i+1}\right) &\leq I\left(X, U_{i+1} \wedge Y \mid g_2(Y), U^i\right) \\ &= I\left(U_{i+1} \wedge Y \mid g_2(Y), X, U^i\right) \\ &\leq I\left(U_{i+1} \wedge Y \mid X, U^i\right) = 0, \end{aligned}$$

where the last equality follows from (P1). Thus, we have

$$X \multimap g_2(Y), U^{i+1} \multimap Y,$$

establishing the validity of (66) for $i+1$. The proof of (64) can be completed by induction by using a similar argument for odd i.

Next, we consider the Gács-Körner CI. Note that any common function of $g_1(X)$ and $g_2(Y)$ is also a common function of X and Y. Consequently,

$$H(\mathtt{mcf}(X,Y)) \geq H(\mathtt{mcf}(g_1(X), g_2(Y))). \tag{67}$$

For the reverse inequality, observe that for an rv U such that $H(U|Y) = H(U|X) = 0$ we have

$$U \multimap X \multimap g_1(X) \multimap Y.$$

Thus, $H\left(U|g_1(X)\right) \leq H(U|Y) = 0$, and similarly, $H\left(U|g_2(Y)\right) = 0$. In particular, it holds that

$$H\left(\mathtt{mcf}(X,Y)|g_1(X)\right) = H\left(\mathtt{mcf}(X,Y)|g_2(Y)\right) = 0,$$

and so,

$$H(\mathtt{mcf}(X,Y)) \leq H(\mathtt{mcf}(g_1(X), g_2(Y))),$$

which along with (67) yields

$$H(\mathtt{mcf}(X,Y)) = H(\mathtt{mcf}(g_1(X), g_2(Y))).$$

Finally, we consider Wyner's CI and claim that this, too, remains unchanged upon replacing the sources with their respective sufficient statistics (for the other source). It suffices to show that

$$CI_W(X \wedge Y) = CI_W(g(X) \wedge Y),$$

for a function g such that $X \multimap g(X) \multimap Y$. Consider an rv W for which $X \multimap W \multimap Y$ is satisfied. We have

$$0 = I(X \wedge Y \mid W) \geq I\left(g(X) \wedge Y \mid W\right).$$

It follows from (4) that

$$CI_W(X \wedge Y) \geq CI_W\left(g(X) \wedge Y\right). \tag{68}$$

On the other hand, for an rv $L = L\left(g^n\left(X^n\right), Y^n\right)$ we have

$$\frac{1}{n} I\left(X^n \wedge Y^n \mid L\right) = \frac{1}{n} I\left(g^n\left(X^n\right) \wedge Y^n \mid L\right),$$

since

$$\begin{aligned} I\left(X^n \wedge Y^n \mid L, g^n\left(X^n\right)\right) &\leq I\left(X^n \wedge Y^n, L \mid g^n\left(X^n\right)\right) \\ &= I\left(X^n \wedge Y^n \mid g^n\left(X^n\right)\right) = 0. \end{aligned}$$

Thus, from the definition of $CI_W(g(X) \wedge Y)$ we get

$$CI_W(X \wedge Y) \leq CI_W(g(X) \wedge Y),$$

so that, by (68),

$$CI_W(X \wedge Y) = CI_W(g(X) \wedge Y).$$

□

VII. Discussion

A. Local Randomization

Although independent local randomization was not allowed in our formulation, our main result characterizing R_{SK} holds even when such randomization is available. Consider a model where terminals $\mathcal{X}$ and $\mathcal{Y}$, in additional to their respective observations X^n and Y^n, have access to finite-valued[4] rvs T_1 and T_2, respectively. The rvs T_1, T_2, and (X^n, Y^n) are mutually independent. The SK capacity is defined as before, with X^n and Y^n now replaced by (X^n, T_1) and (Y^n, T_2), respectively. It is known [13], [2] that even with randomization the SK capacity equals $I(X \wedge Y)$. For this model, denote the minimum rate of r-interactive communication required to generate an SK of rate $I(X \wedge Y)$ by $\tilde{R}^r_{SK}$.

Lemma 11. *For $r \geq 1$,*

$$\tilde{R}^r_{SK} = R^r_{SK}.$$

To see this, we define quantities $\tilde{R}^r_{CI}$ and $\tilde{CI}^r_i$ analogously to R^r_{SK} and CI^r_i, with X^n and Y^n replaced by (X^n, T_1) and (Y^n, T_2), respectively. Note that this substitution is made even in condition (3), i.e., the CR J and the communication $\mathbf{F}$ now are required to satisfy:

$$\frac{1}{n} I\left(X^n, T_1 \wedge Y^n, T_2 \mid J, \mathbf{F}\right) \leq \epsilon. \tag{69}$$

We observe that (12) still holds, with (X^n, T_1) and (Y^n, T_2) replacing, respectively, X^n and Y^n on the right-side. Therefore, the proof of Theorem 3 is valid, and we get:

$$\tilde{R}^r_{CI} = \tilde{R}^r_{SK} = \tilde{CI}^r_i - I(X \wedge Y). \tag{70}$$

By its definition $\tilde{R}^r_{CI} \leq R^r_{CI}$, since $L = (J, \mathbf{F}) = L(X^n, Y^n)$ satisfying (3) will meet (69) as well. We claim that $\tilde{R}^r_{CI} \geq R^r_{CI}$, which by (70) and Theorem 3 implies Lemma 11.

[4]The cardinalities of the range spaces of T_1 and T_2 are allowed to be at most exponential in n.

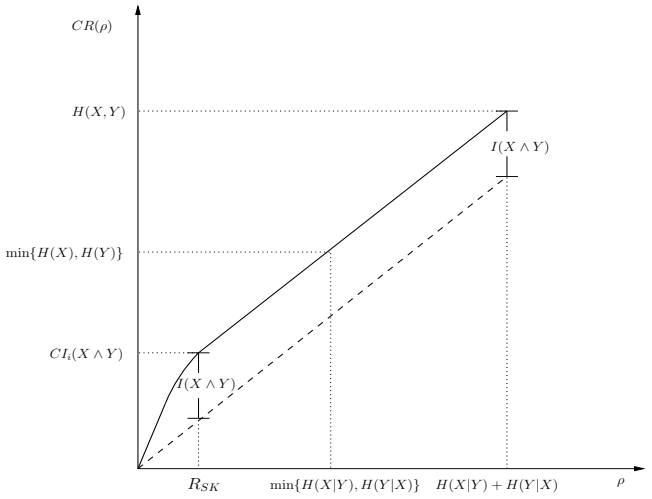

Fig. 1. Minimum rate of communication R_{SK} for optimum rate SK generation

Indeed, consider CR J recoverable from $\mathbf{F}$ such that $(J, \mathbf{F})$ attain $\tilde{R}^r_{CI}$. Then, the condition (69) gives

$$\frac{1}{n} I\left(X^n \wedge Y^n \mid J, \mathbf{F}, T_1, T_2\right) \approx 0.$$

So, there exist t_1, t_2 such that conditioned on $T_1 = t_1, T_2 = t_2$ the CR J is still recoverable from $\mathbf{F}$, and

$$\frac{1}{n} I\left(X^n \wedge Y^n \mid J, \mathbf{F}, T_1 = t_1, T_2 = t_2\right) \approx 0.$$

Thus, with $T_1 = t_1, T_2 = t_2$ fixed, $(J, \mathbf{F})$ constitutes a feasible choice in the definition of R^r_{CI}. Since the number of values taken by $\mathbf{F}$ can only decrease upon fixing $T_1 = t_1, T_2 = t_2$, we get $\tilde{R}^r_{CI} \geq R^r_{CI}$. Therefore, the availability of local randomization does not decrease the rate of communication required for generating an optimum rate SK.

B. Less-than-optimum rate SKs

SK generation is linked intrinsically to the efficient generation of CR. For $\rho \geq 0$, a rate $R \geq 0$ is an achievable CR rate for ρ if for every $0 < \epsilon < 1$ there exists, for some $n \geq 1$, an ϵ-CR L with

$$\frac{1}{n} H(L) \geq R - \epsilon,$$

recoverable from an r-interactive communication $\mathbf{F}$, for arbitrary r, of rate

$$\frac{1}{n} H(\mathbf{F}) \leq \rho + \epsilon;$$

the maximum achievable CR rate for ρ is denoted by $CR(\rho)$. Similarly, denote by $C(\rho)$ the maximum rate of an SK that can be generated using a communication as above. It can be shown in a straightforward manner that

$$C(\rho) = CR(\rho) - \rho. \tag{71}$$

The graph of CR as a function of ρ is plotted in Fig. 1. $CR(\rho)$ is an increasing and a concave function of ρ, as seen from a simple time-sharing argument. Since R_{SK} is the minimum rate of communication required to generate a maximum rate SK, $CR(\rho) - \rho = I(X \wedge Y)$ for $\rho \geq R_{SK}$. Thus, our results characterize the graph of $CR(\rho)$ for all $\rho \geq R_{SK}$. The quantity R_{SK} is the minimum value of ρ for which the slope of $CR(\rho)$ is 1; $CR\left(R_{SK}\right)$ is equal to the interactive common information $CI_i(X \wedge Y)$. Furthermore, from the proof of Theorem 3, a CR L that satisfies (3) must yield an optimum rate SK. Thus, any CR recoverable from a communication of rate less than R_{SK} cannot satisfy (3). A characterization of $CR(\rho)$ for $\rho < R_{SK}$ is central to the characterization of $C(\rho)$, and this, along with a single-letter characterization of R_{SK}, remains an interesting open problem.

APPENDIX

Proof of Lemma 6:

From the Slepian-Wolf theorem [15], there exist mappings $f_1, ..., f_r$ of $F_1^k, ..., F_r^k$, respectively, of rates

$$\frac{1}{k} \log \|f_{2i+1}\| \leq H(F_{2i+1} \mid Y^n, F_1, ..., F_{2i}) + \frac{n\epsilon}{2r}, \quad 0 \leq i \leq \lfloor (r-1)/2 \rfloor,$$

$$\frac{1}{k} \log \|f_{2i}\| \leq H(F_{2i} \mid X^n, F_1, ..., F_{2i-1}) + \frac{n\epsilon}{2r}, \quad 1 \leq i \leq \lfloor r/2 \rfloor,$$

such that

F_{2i+1}^k is $\frac{\epsilon}{2r}$-recoverable from $\left(f_{2i+1}(F_{2i+1}^k), Y^N, F_1^k, ..., F_{2i}^k\right), 0 \leq i \leq \lfloor (r-1)/2 \rfloor$,

F_{2i}^k is $\frac{\epsilon}{2r}$-recoverable from $\left(f_{2i}(F_{2i}^k), X^N, F_1^k, ..., F_{2i-1}^k\right), \quad 1 \leq i \leq \lfloor r/2 \rfloor$,

for all k sufficiently large. Thus, the communication $\mathbf{F}'$ given by $F_i' = f_i\left(F_i^k\right)$, $1 \leq i \leq r$ constitutes the required communication of rate

$$\frac{1}{nk} \log \|\mathbf{F}'\| \leq \frac{1}{n} \left[H\left(\mathbf{F}|X^n\right) + H\left(\mathbf{F}|Y^n\right)\right] + \epsilon.$$

□

Proof of Lemma 7:

For $T = T\left(X^n, Y^n\right)$ we have,

$$\begin{aligned}
&nI(X \wedge Y) \\
&= H\left(X^n, Y^n\right) - H\left(X^n \mid Y^n\right) - H\left(Y^n \mid X^n\right) \\
&= H\left(X^n, Y^n \mid T\right) - H\left(X^n \mid Y^n, T\right) - H\left(Y^n \mid X^n, T\right) \\
&\quad + H(T) - H\left(T \mid X^n\right) - H\left(T \mid Y^n\right) \\
&= I\left(X^n \wedge Y^n \mid T\right) + H(T) - H\left(T \mid X^n\right) - H\left(T \mid Y^n\right).
\end{aligned}$$

Lemma 7 follows upon choosing $T = J, \mathbf{F}$. □

Proof of (32) and (36):

It remains to prove that there exists ϵ-CR J, recoverable from $\mathbf{F}$ such that $J, \mathbf{F}$ satisfy (32) and (36). We provide a CR generation scheme with r stages. For $1 \leq k \leq r$, denote by $\mathcal{E}_k$ the error event in the kth stage (defined below recursively in terms of $\mathcal{E}_{k-1}$), and by $\mathcal{E}_0$ the negligible probability event corresponding to X^n, Y^n not being P_{XY}-typical.

Consider $1 \leq k \leq r$, k odd. For brevity, denote by V the rvs U^{k-1} and by U the rv U_k; for $k = 1$, V is taken to

be a constant. Suppose that conditioned on $\mathcal{E}_{k-1}^c$ terminals $\mathcal{X}$ and $\mathcal{Y}$ observe, respectively, sequences $\mathbf{x} \in \mathcal{X}^n$ and $\mathbf{y} \in \mathcal{Y}^n$, as well as a common sequence $\mathbf{v} \in \mathcal{V}^n$ such that $(\mathbf{v}, \mathbf{x}, \mathbf{y})$ are jointly P_{VXY}-typical. For $\delta > 0$, generate at random $\exp\left[n(I(X, Y \wedge U \mid V) + \delta)\right]$ sequences $\mathbf{u} \in \mathcal{U}^n$ that are jointly P_{UV}-typical with $\mathbf{v}$, denoted by $\mathbf{u}_{ij}$, $1 \leq i \leq N_1$, $1 \leq j \leq N_2$, where

$$N_1 = \exp\left[n\left(I(X \wedge U \mid Y, V) + 3\delta\right)\right],$$
$$N_2 = \exp\left[n\left(I(Y \wedge U \mid V) - 2\delta\right)\right].$$

The sequences $\mathbf{u}_{ij}$ are generated independently for different indices ij. Denote by $L^{(k)}(\mathbf{v}, \mathbf{x})$ a sequence $\mathbf{u}_{ij}$, $1 \leq i \leq N_1$, $1 \leq j \leq N_2$, that is jointly P_{UVX}-typical with $(\mathbf{v}, \mathbf{x})$ (if there exist more than one such sequences, choose any of them). The error event when no such sequence is found is denoted by $\mathcal{E}_{k1}$; this happens with probability vanishing to 0 doubly exponentially in n. The communication $F_k(\mathbf{v}, \mathbf{x})$ is defined to equal the first index i of $\mathbf{u}_{ij} = L^{(k)}(\mathbf{v}, \mathbf{x})$. Upon observing $F_k(\mathbf{v}, \mathbf{x}) = i$, the terminal $\mathcal{Y}$ computes $L_2^{(k)}(\mathbf{v}, \mathbf{y}, i)$ as the unique sequence in $\{\mathbf{u}_{ij}, 1 \leq j \leq N_2\}$, that is jointly typical with $(\mathbf{v}, \mathbf{y})$. If no such sequence is found or if several such sequences are found an error event $\mathcal{E}_{k2}$ occurs. Clearly, the rate of communication F_k is bounded above by

$$\begin{aligned}\frac{1}{n}\log N_1 &= I(X \wedge U \mid Y, V) + 3\delta \\ &= I(X \wedge U_k \mid Y, U^{k-1}) + 3\delta, \end{aligned} \tag{A1}$$

and also, for large n,

$$\begin{aligned}\frac{1}{n}H(L^{(k)}) &\leq \frac{1}{n}\log(1 + N_1 N_2) \leq I(X, Y \wedge U \mid V) + 2\delta \\ &= I(X, Y \wedge U_k \mid Y, U^{k-1}) + 2\delta. \end{aligned} \tag{A2}$$

Denote by $\mathcal{E}_{k3}$ the event $\left(L^{(k)}(\mathbf{v}, \mathbf{x}), \mathbf{v}, \mathbf{x}, \mathbf{y}\right)$ not being jointly P_{UVXY}-typical. The error event $\mathcal{E}_k$ is defined as $\mathcal{E}_k = \mathcal{E}_{k-1} \cup \mathcal{E}_{k1} \cup \mathcal{E}_{k2} \cup \mathcal{E}_{k3}$. Then, conditioned on $\mathcal{E}_k^c$ the terminals share sequences $(\mathbf{u}_{ij}, \mathbf{v})$ that are jointly typical with $(\mathbf{x}, \mathbf{y})$. In the next stage $k+1$, the sequence $(\mathbf{u}_{ij}, \mathbf{v})$ plays the role of the sequence $\mathbf{v}$. The scheme for stages with even k is defined analogously with roles of $\mathcal{X}$ and $\mathcal{Y}$ interchanged. We claim that $L^{(1)}, ..., L^{(r)}$ constitutes the required CR along with the communication $\mathbf{F} = F_1, ..., F_k$. Then, (36) follows from (A1), and the second inequality in (32) follows from (A2). Moreover, for every realization $\mathbf{u}_1, ..., \mathbf{u}_r$ of $L^{(1)}, ..., L^{(r)}$, with $E = \mathbf{1}_{\mathcal{E}_r}$ we have,

$$\begin{aligned}&\mathbb{P}\left(L^{(1)}, ..., L^{(r)} = \mathbf{u}_1, ..., \mathbf{u}_r \mid E = 0\right) \\ &\leq \mathbb{P}\left(\{(\mathbf{x}, \mathbf{y}) : (\mathbf{u}_1, ..., \mathbf{u}_r, \mathbf{x}, \mathbf{y}) \text{ are jointly } P_{U^r XY} \text{ typical}\}\right) \\ &\leq \exp\left[-n(I(X, Y \wedge U^r) - \delta)\right],\end{aligned}$$

for n large, which further yields

$$\frac{1}{n}H(L^{(1)}...L^{(r)} \mid E = 0) \geq I(X, Y \wedge U^r) - \delta.$$

Therefore,

$$\begin{aligned}&\frac{1}{n}H(L^{(1)}...L^{(r)}) \\ &\geq \frac{1}{n}H(L^{(1)}...L^{(r)} \mid E = 0) - \mathbb{P}\left(\mathcal{E}_r\right)\log|\mathcal{X}||\mathcal{Y}| \\ &\geq I(X, Y \wedge U^r) - \delta - \mathbb{P}\left(\mathcal{E}_r\right)\log|\mathcal{X}||\mathcal{Y}|.\end{aligned}$$

Thus, the claim will follow upon showing that $\mathbb{P}\left(\mathcal{E}_r\right) \to 0$ as $n \to \infty$. In particular, it remains to show that $\mathbb{P}\left(\mathcal{E}_{k2}\right) \to 0$ and $\mathbb{P}\left(\mathcal{E}_{k3}\right) \to 0$, $k = 1, ..., r$, as $n \to \infty$. As before, we show this for odd k and the proof for even k follows *mutatis mutandis*. To that end, note first that for any jointly P_{UVX}-typical $(\mathbf{u}, \mathbf{v}, \mathbf{x})$, the set of $\mathbf{y} \in \mathcal{Y}^n$ such that $(\mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y})$ are jointly typical with $(\mathbf{u}, \mathbf{v}, \mathbf{x})$ has conditional probability close to 1 conditioned on $U^n = \mathbf{u}, V^n = \mathbf{v}, X^n = \mathbf{x}$, and so by the Markov relation $Y \multimap V, X \multimap U$, also conditioned on $V^n = \mathbf{v}, X^n = \mathbf{x}$. Upon choosing $\mathbf{u} = L^{(k)}(\mathbf{v}, \mathbf{x})$ in the argument above, we get $\mathbb{P}\left(\mathcal{E}_{k2}\right) \to 0$. Finally, we show that $\mathbb{P}\left(\mathcal{E}_{k3}\right)$ will be small, for large probability choices of the random codebook $\{\mathbf{u}_{ij}\}$. Specifically, for fixed typical sequences $(\mathbf{v}, \mathbf{x}, \mathbf{y})$, the probability $\mathbb{P}\left(\mathcal{E}_{k3} \mid V^n = \mathbf{v}, X^n = \mathbf{x}, Y^n = \mathbf{y}\right)$ is bounded above exactly as in [2, equation (4.15)]:

$$\begin{aligned}&\mathbb{P}\left(\mathcal{E}_{k3} \mid V^n = \mathbf{v}, X^n = \mathbf{x}, Y^n = \mathbf{y}\right) \\ &\leq \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\sum_{l=1, l\neq j}^{N_2} \mathbb{P}\Big((\mathbf{u}_{ij}, \mathbf{v}, \mathbf{x}) \text{ jointly } P_{UVX}\text{-typical}, \\ &\qquad\qquad (\mathbf{u}_{il}, \mathbf{v}, \mathbf{u}_{il}) \text{ jointly } P_{UVY}\text{-typical}\Big) \\ &\leq N_1 N_2^2 . \exp[-n(I(X \wedge U \mid V) + I(X \wedge U \mid V) + o(n))] \\ &\leq \exp[-n\delta + o(n)],\end{aligned}$$

for all n sufficiently large. Note that the probability distribution in the calculation above comes from codebook generation, and in particular, the second inequality above uses the fact that $\mathbf{u}_{il}$ and $\mathbf{u}_{ij}$ are independently selected for $l \neq j$. Thus, $\mathbb{P}\left(\mathcal{E}_{k3} \mid \mathcal{E}_{k2}\right) \to 0$ for an appropriately chosen codebook, which completes the proof. □

ACKNOWLEDGEMENTS

The ideas presented in this work are based on heuristics for the interplay between CR generation and SK generation, developed, over the years, jointly with Prof. Prakash Narayan. Further, his comments on an earlier version of this manuscript have improved the presentation, especially that of Section VI, where he also simplified the proofs.

REFERENCES

[1] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography–part i: Secret sharing," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.

[2] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography–part ii: CR capacity," *IEEE Trans. Inform. Theory*, vol. 44, no. 1, pp. 225–240, January 1998.

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

[4] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless channels*. 2nd Edition. Cambridge University Press, 2011.

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

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

[7] K. Eswaran and M. Gastpar, "Rate loss in the ceo problem," *Proc. Conference on Information Sciences and Systems*, March 2005.

[8] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems of Control and Information Theory*, vol. 2, no. 2, pp. 149–162, 1973.

[9] S. Kamath and V. Ananthram, "A new dual to the Gács-Körner common information defined via the Gray-Wyner system," *Proc. Conference on Communication, Control, and Computing (Allerton)*, Oct 2010.
[10] A. H. Kaspi, "Two-way source coding with a fidelity criterion," *IEEE Trans. Inform. Theory*, vol. 31, no. 6, pp. 735–740, November 1985.
[11] N. Ma and P. Ishwar, "Some results on distributed source coding for interactive function computation.," *IEEE Trans. Inform. Theory*, vol. 57, no. 9, pp. 6180–6195, September 2011.
[12] M. Madiman and P. Tetali, "Information inequalities for joint distributions, with interpretations and applications," *IEEE Trans. Inform. Theory*, vol. 56, no. 6, pp. 2699–2713, June 2010.
[13] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
[14] U. M. Maurer, "The strong secret key rate of discrete random triples," *Communications and Cryptography: Two sides of One Tapestry*, pp. 271–285, 1994.
[15] D. Slepian and J. Wolf, "Noiseless coding of correlated information source," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.
[16] A. D. Wyner, "The common information of two dependent random variables," *IEEE Trans. Inform. Theory*, vol. 21, no. 2, pp. 163–179, March 1975.
[17] G. Xu and B. Chen, "The sufficiency principle for decentralized data reduction," *Proc. IEEE Int. Symp. Inform. Theory*, July 2012.

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