Secrecy Generation for Multiaccess Channel Models

Imre Csiszár* and Prakash Narayan†

September 1, 2011

Abstract

Shannon theoretic secret key generation by several parties is considered for models in which a secure noisy channel with multiple input and output terminals and a public noiseless channel of unlimited capacity are available for accomplishing this goal. The secret key is generated for a set A of terminals of the noisy channel, with the remaining terminals (if any) cooperating in this task through their public communication. Single-letter lower and upper bounds for secrecy capacities are obtained when secrecy is required from an eavesdropper that observes only the public communication and perhaps also a set of terminals disjoint from A. These bounds coincide in special cases, but not in general. We also consider models in which different sets of terminals share multiple keys, one for terminals in each set with secrecy required from the eavesdropper as well as the remaining terminals in the other sets. Partial results include showing links among the associated secrecy capacity region for multiple keys, the transmission capacity region of the multiple access channel defined by the secure noisy channel, and achievable rates for a single secret key for all the terminals.

Index Terms – Multiaccess channel, multiple keys, private key, private key capacity region, secrecy capacity, secret key, source model.

*Imre Csiszár is with the A. Rényi Institute of Mathematics, Hungarian Academy of Sciences, POB 127, H-1364 Budapest, Hungary. Email: csiszar@renyi.hu.

†Prakash Narayan is with the Department of Electrical and Computer Engineering and the Institute for Systems Research, University of Maryland, College Park, MD 20742, USA. Email: prakash@umd.edu.

The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Seoul, Korea, June-July 2009.

1 Introduction

Separate terminals with the means to transact over a secure noisy channel as well as a public noiseless channel, can devise a secret key more effectively than by using the secure channel alone. A secret key, in the Shannon theoretic sense, is common randomness of near uniform distribution regarding which an eavesdropper, which observes the public communication and perhaps also possesses additional observations available or unavailable to the terminals engaged in secrecy generation, can glean only a negligible amount of information.

The first Shannon theoretic model for generating a secret key over a noisy channel was Wyner's wiretap channel [20], generalized by Csiszár and Körner [6]. This model did not allow for public communication, and secret key generation was tantamount to secure transmission over the noisy channel, when the eavesdropper had access to wiretap side information. The fact that secrecy generation could be enhanced by public communication was illustrated by Bennett, Brassard and Robert [3]. Models for secrecy generation which entailed two terminals communicating over a public noiseless channel, were examined in detail by Maurer [15] and Ahlswede and Csiszár [1]. These models involve either a discrete memoryless multiple source (DMMS) with two components accessible to one terminal each, or a discrete memoryless channel (DMC) with one input terminal and one output terminal. In both types of models, an additional "wiretapped" terminal may or may not be present. The sizable literature on such models includes Maurer [16], Bennett, Brassard, Crépeau and Maurer [4], Csiszár [5], Maurer and Wolf [17], [18], Csiszár and Narayan [8, 9], Renner and Wolf [19], Gohari and Anantharam [2], and a comprehensive treatment in Csiszár and Körner [7]. A single-letter characterization of the secrecy capacity – the largest rate at which a secret key can be generated – is known in special cases, e.g., when a wiretapped terminal is absent or when the wiretapped terminal reveals itself to the parties generating secrecy.

In our previous work, we had studied secrecy generation for a multiterminal source model where each participating terminal had access to one component of a discrete memoryless multiple source [8, 9], and for a multiterminal channel model which involved an underlying DMC with a single input and multiple outputs [9]; in both models, unrestricted and noiseless public communication between the terminals was permitted, to which the eavesdropper had full access. In this paper, which constitutes a continuation of our work in [9], [10], we examine channel models for secrecy generation which involve an underlying DMC with multiple inputs and outputs. Terminals $1, \ldots, k$ govern the inputs and terminals $k+1, \ldots, m$ observe the corresponding outputs. Following each transmission of symbols by the input terminals over the DMC, communication over a public noiseless channel of unlimited capacity is allowed between all the terminals, which may be interactive and which is observed by all the terminals[1]. The goal is to generate secret common randomness shared by a given set $A \subset \{1, \ldots, m\}$ of terminals at the largest rate possible. Thus, the resulting key must be accessible to every terminal in A. It need not be accessible to the terminals not in A, but nor is it required to be concealed from them, with the possible exception of a set D of terminals which are "wiretapped" by the eavesdropper (where $A \cap D = \emptyset$). A DMC input terminal may or may not belong to the set A or D.

We restrict ourselves to models where all the terminals cooperate, including those that are wiretapped (if $D \neq \emptyset$), in generating a secret key for the terminals in A, with secrecy being required from the eavesdropper that has access to only the public communication and the information available to the wiretapped terminals in D. Also, we assume the eavesdropper to be passive, i.e., unable to tamper with the communication of the legitimate terminals.

We do not address models with wiretap side information in which the underlying DMC also has an additional output terminal that is wiretapped by the eavesdropper and does not cooperate in secrecy generation (cf. e.g., [15, 1, 17, 19, 8, 9, 11, 12]).

The problem of secrecy generation for a general multiterminal channel model studied in this paper appears more difficult than its special case for a channel with a single input. Single-letter characterizations of secrecy capacities for the latter have been given in [9]. For the general channel model, short of providing single-letter characterizations of secrecy capacities, our main contributions are the following. One possible operational strategy in a channel model as above is source emulation which entails the channel input terminals transmitting independent sequences of random variables (rvs) over the DMC with the output terminals observing the corresponding output sequences. The emulated source model leads

[1] For ease of distinction between the use of the DMC and the use of the public channel, hereafter the former will be termed "transmission" while the latter will be referred to as "communication."

to our achievability results which furnish lower bounds for the secrecy capacities; it is likely that these bounds are not tight in general. Our converse results provide upper bounds for the secrecy capacities using familiar techniques from Shannon theory, but are difficult and rely on two entropy inequalities from our previous work [9] which may be of independent interest. Our lower and upper bounds coincide only in special cases. We also consider multiterminal channel models in which different subsets of terminals share multiple keys, one for terminals in each set with secrecy required from the eavesdropper as well as the remaining terminals in the other sets. Partial results include showing links among the associated secrecy capacity region for multiple keys, the transmission capacity region of the multiple access channel (MAC) defined by the DMC, and achievable rates for a single secret key shared by a subset of the terminals. We illustrate our results and their limitations by four examples of secrecy generation in simple multiterminal channel models.

Our problem formulations are described in Section 2. Section 3 treats secrecy generation for DMCs with a single output based on elementary tools. Our general single-letter lower and upper bounds for secrecy capacities are presented in Sections 4 and 5, respectively. Illustrative examples are given in Section 6. A closing discussion is contained in Section 7.

2 Preliminaries

All rvs are assumed to take values in finite sets, even if not stated explicitly. An rv will be denoted by an uppercase letter and its range by the corresponding script capital unless stated otherwise. The cardinality of a finite set $\mathcal{X}$ is denoted by $|\mathcal{X}|$. Logarithms are with respect to the base 2. For integers $l \leq k$, we denote $[l,k] = (l,\dots,k)$.

We consider multiterminal channel models of the following kind. Terminals $1,\dots,k$, with finite alphabets $\mathcal{X}_1,\dots,\mathcal{X}_k$, are connected to terminals $k+1,\dots,m$, with finite alphabets $\mathcal{X}_{k+1},\dots,\mathcal{X}_m$, respectively, by a DMC $W : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \mathcal{X}_{k+1} \times \ldots \times \mathcal{X}_m$. Terminals $1,\dots,k$ govern the inputs of the DMC over which they transmit *securely* sequences of length n, while terminals $k+1,\dots,m$ observe the corresponding output sequences of length n. In between consecutive symbol transmissions over the DMC (with instantaneous receptions), the terminals in $\mathcal{M} = \{1,\dots,m\}$ are allowed to communicate over a public noiseless channel of unlimited capacity. In any transmission or communication by a terminal, randomization is permitted. The public communication is observed by all the terminals in $\mathcal{M}$ as well as by an eavesdropper.

When randomization is permitted at terminal $i \in \mathcal{M}$, we shall assume that it generates at the outset a rv U_i; the rvs $U_1, \ldots, U_m$ are mutually independent. Every input terminal $i \in [1, k]$ transmits n symbols $X_{i1}, X_{i2}, \ldots, X_{in}$ over the DMC W at time instants $\tau_1 < \tau_2 < \ldots < \tau_n$, and every output terminal $i \in [k+1, m]$ observes the corresponding output symbols $X_{i1}, X_{i2}, \ldots, X_{in}$. In addition, communication among the terminals in $\mathcal{M}$ over the public channel occurs – possibly interactively and in several rounds – during the time-intervals (τ_t, τ_{t+1}), for $t = 1, \ldots, n-1$, and immediately following τ_n, which hereafter will be referred to simply as intervals $t = 1, \ldots, n$. The public communication of all the terminals in interval t is depicted collectively as F_t, and we denote $\mathbf{F} = (F_1, \ldots, F_n)$.

In general, terminal $i \in [1, k]$ determines its tth input X_{it} of the DMC W as a function of U_i for $t = 1$, and of $(U_i, F_1, \ldots, F_{t-1})$ for $t = 2, \ldots, n$. Also, the communication of terminal $i \in \mathcal{M}$ in interval t is allowed to depend on U_i, the symbols $(X_{i1}, \ldots, X_{it})$ earlier generated or observed by terminal i, and on all earlier communication $(F_1, \ldots, F_{t-1})$. While this general framework admits complex transmission and communication protocols, in our achievability proofs we shall use only simple *noninteractive communication* protocols with input terminals $i \in [1, k]$ not sending any public messages at all, and each output terminal $i \in [k+1, m]$ sending at most one public message $f_i = f_i\left(X_i^n\right)$ and that upon completion of the n transmissions over the DMC; in this case, $\mathbf{F} = F_n = \left(f_i\left(X_i^n\right),\ i \in [1, k]\right)$.

For rvs $X_i,\ i \in \mathcal{M}$, we shall use the shorthand notation $X_B = (X_i,\ i \in B)$ for sets $B \subset \mathcal{M}$, and, as a special case, $X_{[a,b]} = (X_a, \ldots, X_b)$ for $1 \leq a \leq b \leq m$. Also, we shall write $X_B^t = (X_{B1}, \ldots, X_{Bt})$ for $B \subset \mathcal{M},\ 1 \leq t \leq n$, where $X_{Bj} = \left(X_{ij}, i \in B\right),\ 1 \leq j \leq t$; in particular, $X_i^t = (X_{i1}, \ldots, X_{it})$ for $i \in \mathcal{M},\ 1 \leq t \leq n$.

The following concepts introduced in [8] will be used. Given $\epsilon > 0$, a rv U is *ϵ-recoverable* from V if $\Pr\{U \neq f(V)\} \leq \epsilon$ for some function $f(V)$ of V. For rvs K and Y, to be interpreted as representing a secret key and the eavesdropper's knowledge, respectively, the information theoretic *security index* is

$$s(K; Y) = \log |\mathcal{K}| - H(K|Y).$$

Smallness of this security index is tantamount jointly to a nearly uniform distribution for K (i.e., $\log |\mathcal{K}| - H(K)$ is small) and to the near independence of K and Y (i.e., the mutual information $I(K \wedge Y)$ is close to 0).

Definition 1: Given any set $A \subset \mathcal{M}$ of size $|A| \geq 2$, a rv K constitutes an (ϵ, δ)-*secret key* ((ϵ, δ)-SK) for the set of terminals A, achievable with n uses of the DMC W, randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F}$, if K is ϵ-recoverable from $(U_i, X_i^n, \mathbf{F})$ for each $i \in A$ and, in addition, it satisfies the secrecy condition

$$s(K; \mathbf{F}) \ \leq \ \delta. \tag{1}$$

An (ϵ, δ)-SK as above is called an (ϵ, δ)-*private key* ((ϵ, δ)-PK) for the set of terminals A, private from the set of terminals $D \subset \mathcal{M}$ with $A \cap D = \emptyset$, if it satisfies the stronger secrecy condition

$$s(K; U_D, X_D^n, \mathbf{F}) \ \leq \ \delta. \tag{2}$$

By definition, an (ϵ, δ)-SK is recoverable at the terminals in A, and is nearly uniformly distributed and effectively concealed from an eavesdropper with access to the public communication $\mathbf{F}$; it need not be concealed from the terminals in $A^c = \mathcal{M} \backslash A$. On the other hand, an (ϵ, δ)-PK for A is effectively concealed from an eavesdropper with access — in addition to the public communication $\mathbf{F}$ — also to a set $D \subset A^c$ of "wiretapped" or "compromised" terminals. This (ϵ, δ)-PK need not be concealed from the terminals in $A^c \backslash D$. Note that the compromised terminals can cooperate in the secrecy generation through their public communication. Indeed, it can be assumed without loss of generality (w.l.o.g.) that the terminals in D reveal publicly all the information in their possession (which, anyway, is accessible to the eavesdropper). This assumption will be made usually without explicit mention.

Definition 2: A number R is an achievable SK rate for a set of terminals $A \subset \mathcal{M}$ if there exist (ϵ_n, δ_n)-secret keys $K^{(n)}$ achievable for A with n uses of the DMC W, suitable randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F}^{(n)}$, such that

$$\epsilon_n \to 0, \quad \delta_n \to 0 \quad \text{and} \quad \frac{1}{n} \log |\mathcal{K}^{(n)}| \to R \quad \text{as} \quad n \to \infty.$$

The largest achievable SK rate for A is the SK capacity $C_S(A)$. Achievable PK rates and PK capacity $C_P(A|D)$ are defined similarly.

Remark: Our converse proofs stand under the requirement of "weak" secrecy, i.e., $\delta_n = o(n)$ while ϵ_n decays to 0 [15, 1]. The achievability results hold with both ϵ_n and δ_n decaying to 0 exponentially rapidly thereby affording "strong" secrecy [16, 5, 7].

In general, any number of DMC input and output terminals may be wiretapped (barring two terminals to avoid the trivial). However, it is obvious that the wiretapped input terminals (if any) can be coalesced, as can the wiretapped output terminals. The next lemma shows that attention can be restricted even to such models in which no input terminal, and at most one output terminal, is wiretapped. Nevertheless, we shall find it convenient throughout to adhere to the original model above and take recourse only occasionally to the following reduction lemma.

Lemma 1: For any DMC $W : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \mathcal{X}_{k+1} \times \ldots \times \mathcal{X}_m$ and $D \subset \mathcal{M}$ with $0 \leq |D| \leq m-2$, there exists a DMC $\tilde{W} : \tilde{\mathcal{X}}_1 \times \cdots \times \tilde{\mathcal{X}}_{\tilde{k}} \to \tilde{\mathcal{X}}_{k+1} \times \ldots \times \tilde{\mathcal{X}}_{\tilde{m}}$ with $\tilde{k}$ equal to the number of input terminals of W not in D, or $\tilde{k} = 1$ if $[1, k] \subset D$, such that for each $A \subset \mathcal{M}$ disjoint from D, the PK capacity $C_P(A|D)$ for W and D is equal to the corresponding PK capacity for $\tilde{W}$ and $\tilde{D} = \{\tilde{m}\}$.

Remark: By the Lemma, any channel model with at most one uncompromised input terminal can be reduced to a model with just one input terminal; for the latter, a single-letter solution for PK capacity is available ([9], Theorem 4.1).

Proof: By the passage preceding the Lemma, we can assume w.l.o.g. that $D = \{k\}$ or $D = \{k, m\}$. In the first case, let $\tilde{m} = m + 1$, $\tilde{\mathcal{X}}_i \triangleq \mathcal{X}_i$ for $i \in [k+1, m]$ and $\tilde{\mathcal{X}}_{\tilde{m}} \triangleq \mathcal{X}_k$, while in the second case let $\tilde{m} = m$, $\tilde{\mathcal{X}}_i = \mathcal{X}_i$ for $i \in [k+1, m-1]$ and $\tilde{\mathcal{X}}_{\tilde{m}} \triangleq \mathcal{X}_m \times \mathcal{X}_k$. Further, let $\tilde{\mathcal{X}}_1 = \mathcal{X}_1$ if $k = 1$ (when $\tilde{k} = 1$) and if $k > 1$ (when $\tilde{k} = k - 1$), let $\tilde{\mathcal{X}}_i \triangleq \mathcal{X}_i$ for $i \in [1, k-2]$ and $\tilde{\mathcal{X}}_{k-1} \triangleq \mathcal{X}_{k-1} \times \mathcal{X}_k$.

In either case, input $\tilde{k}$-tuples of $\tilde{W}$ are identified in an obvious manner with input k-tuples $\mathbf{x} = x_1, \ldots, x_k$ of W, and output $(\tilde{m} - k)$-tuples of $\tilde{W}$ are regarded as being obtained by appending an $x'_k \in \mathcal{X}_k$ to the output $(m - k)$-tuples $\mathbf{y}$ of W. The definition of $\tilde{W}$ is completed by setting

$$\tilde{W}(\mathbf{y}{x'}_k|\mathbf{x}) \triangleq \begin{cases} W(\mathbf{y}|\mathbf{x}) & \text{if } {x'}_k = x_k \\ 0 & \text{if } {x'}_k \neq x_k. \end{cases}$$

In other words, the DMC $\tilde{W}$ behaves as W but additionally transmits noiselessly the input of terminal k of W to the terminal $\tilde{m}$ of $\tilde{W}$. When $D = \{k, m\}$, the terminal $\tilde{m}$ of $\tilde{W}$ also receives the output of terminal m of W. Thus, in both the cases $D = \{k\}$ and $D = \{k, m\}$, the single wiretapped terminal $\tilde{m}$ of $\tilde{W}$ will possess the same information as the wiretapped terminal(s) of W. It follows that each protocol for W and D gives rise to a protocol for $\tilde{W}$ and $\tilde{D} = \{\tilde{m}\}$ with identical secrecy performance, and also reciprocally so. □

We shall also consider models in which different subsets of the terminals in $\mathcal{M}$ share *multiple keys*, one for the terminals in each subset with privacy from the remaining terminals in $\mathcal{M}$ that are not members of that subset.

Definition 3: Given different subsets $A_1, \ldots, A_l$ of $\mathcal{M}$, $l \geq 1$, the rvs $K_1, \ldots, K_l$ constitute (ϵ, δ)-PKs for the terminals in $A_1, \ldots, A_l$, respectively, if for each $i \in [1, l]$, K_i is an (ϵ, δ)-PK for the terminals in A_i, private from the terminals in $\mathcal{M} \backslash A_i$. The numbers $R_1, \ldots, R_l$ are achievable PK rates for the terminals in $A_1, \ldots, A_l$ if there exist (ϵ_n, δ_n)-PKs achievable for $A_1, \ldots, A_l$ with n uses of the DMC W, suitable randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F}^{(n)}$, such that

$$\epsilon_n \to 0, \quad \delta_n \to 0 \quad \text{and} \quad \frac{1}{n} \log |\mathcal{K}_i^{(n)}| \to R_i \quad \text{as} \quad n \to \infty$$

for $i = 1, \ldots, l$. The set of all achievable PK rates is the PK capacity region $\mathcal{C}_P(A_1, \ldots, A_l)$. For the case $l = 1$, $\mathcal{C}_P(A_1) = C_P(A_1|A_1^c)$ of Definition 2.

Remark: The PK capacity region is a closed convex set. The former is clear from the definition while the latter is a consequence of a standard time-sharing argument.

3 Models with Single Output

In this section, we consider DMCs with a sole output for which simple results are presented that do not require any sophisticated tools.

Let $W : \mathcal{X}_1 \times \cdots \times \mathcal{X}_{m-1} \to \mathcal{X}_m$ be a DMC with $k = m-1$ input terminals and one output terminal, and let A be any set of terminals of size $|A| \geq 2$ which contains the output terminal m. Denote by $\mathcal{C}$ the (average error) capacity region of the multiple access channel (MAC) W, and by $\mathcal{C}(A)$ its projection on the $(|A|-1)$-dimensional subspace of $\mathbb{R}^{m-1}$ spanned by the coordinate axes $\{i : i \in A\backslash\{m\}\}$. Further, consider the PK capacity region $\mathcal{C}_P(\{A_i, i \in A\backslash\{m\}\})$ for the pairs of terminals $A_i = \{i, m\},\ i \in A\backslash\{m\}$.

Theorem 2: For A and $A_i = \{i, m\},\ i \in A\backslash\{m\}$ as above, it holds that

(i) $\mathcal{C}_P(\{A_i, i \in A\backslash\{m\}\}) \supset \mathcal{C}(A)$;

(ii) Any $R > 0$ such that the $(|A|-1)$-dimensional vector $(R, \ldots, R)$ belongs to $\mathcal{C}_P(\{A_i, i \in A\backslash\{m\}\})$, is a lower bound for the PK capacity $C_P(A|A^c)$.

Corollary: It holds that

$$\begin{aligned} C_S(A) \ \geq \ C_P(A|A^c) \ &\geq \ \max\Big\{R : (R, \ldots, R) \in \mathcal{C}_P(\{A_i, i \in A\backslash\{m\}\})\Big\} \\ &\geq \ \max\Big\{R : (R, \ldots, R) \in \mathcal{C}(A)\Big\}. \end{aligned}$$

Further, any R such that $(R, \ldots, R) \in \mathcal{C}(A)$ can be achieved as an SK rate for A or PK rate for A with privacy from A^c, with no public communication by the input terminals and with only the output terminal m sending a public message.

Proof: **(i)** By definition, each $(R_i, i \in A\backslash\{m\}) \in \mathcal{C}(A)$ arises from some $(R_1, \ldots, R_{m-1}) \in \mathcal{C}$ by deleting the components R_i with $i \notin A$; it can be supposed w.l.o.g. that all these deleted components are equal to 0. It is easy to see that an achievable rate tuple $(R_1, \ldots, R_{m-1})$ for transmission over a MAC W, in which some components R_i are 0, can be achieved by codes whose message sets corresponding to the zero rates are singletons (rather than merely of subexponential size). It follows that for each $(R_i, i \in A\backslash\{m\}) \in \mathcal{C}(A)$, there exist encoders $f_i : \mathcal{K}_i \to \mathcal{X}_i^n, i \in A\backslash\{m\}$, with $|\mathcal{K}_i| = \exp(nR_i')$, with R_i' arbitrarily close to R_i, and deterministic sequences $x_i^n \in \mathcal{X}_i^n, i \in [1, m-1]\backslash A$, with the following property: If the MAC inputs are

$$X_i^n \triangleq \begin{cases} f_i(K_i), & i \in A\backslash\{m\} \\ x_i^n, & \text{otherwise,} \end{cases}$$

where the rvs K_i, uniformly distributed on $\mathcal{K}_i$, are mutually independent, then the rvs K_i are recoverable from the MAC output X_m^n with probability approaching 1 as $n \to 1$. This proves that $(R_i, i \in A\backslash\{m\})$ is an achievable PK rate-tuple for the pairs $A_i = \{i, m\},\ i \in A\backslash\{m\}$, achievable without any public communication.

(ii) Suppose that $(R, \ldots, R) \in \mathcal{C}_P(\{A_i, i \in A\backslash\{m\}\})$, and consider PKs for the pairs $A_i = \{i, m\},\ i \in A\backslash\{m\}$, represented by rvs K_i distributed on $\mathcal{K} \triangleq \{1, \ldots, \exp(nR')\}$ with R' close to R, and satisfying the secrecy condition (2) with $\mathcal{M}\backslash\{i, m\}$ in the role of D. Then, arbitrarily fixing $i_1 \in A\backslash\{m\}$, the rv K_{i_1} becomes a PK for the terminals in A, private from $D \triangleq A^c$, if terminal m broadcasts the $\mod |\mathcal{K}|$ sums $K_{i_1} + K_i,\ i \in A\backslash\{i_1, m\}$.

The Corollary is immediate. □

In Section 6, we shall give an example where the trivial inner bound for the PK capacity region $\mathcal{C}_P(\{A_i, i \in A\backslash\{m\}\})$ and the lower bounds for the SK capacity $C_S(A)$, both above, are tight. It remains open whether they are tight in general. Here we present a weaker result than the tightness of the lower bounds for $C_S(A)$ in the Corollary of Theorem 2, and which is straightforward.

Theorem 3: For any $A \ni m$, the SK capacity $C_S(A)$ is positive iff there exists $(R_1, \ldots, R_{m-1}) \in \mathcal{C}$ such that $R_i > 0$ for each $i \in A\backslash\{m\}$.

Proof: Sufficiency is obvious by the Corollary of Theorem 2. For necessity, note that if no $(R_1, \ldots, R_{m-1})$ as above exists, then – using the convexity of $\mathcal{C}$ – for some $i_1 \in A\backslash\{m\}$ we must have that $R_{i_1} = 0$ for every $(R_1, \ldots, R_{m-1}) \in \mathcal{C}$. The latter means that $W(x_m|x_1, \ldots, x_{m-1})$ does not depend on x_{i_1}, and this would imply that the SK capacity is 0 even if the terminals in $\mathcal{M}\backslash\{i_1\}$ were allowed to communicate securely among themselves. For a formal proof, note that upon regarding the terminals in $\mathcal{M}\backslash\{i_1\}$ as a consolidated party L, any use of the DMC W amounts to a randomization performed by party L (since the choice of channel input at terminal i_1 does not influence the output). However, it is well-known that two parties (here $\{i_1\}$ and L), with no resources other than the ability of randomization and of public communication, cannot generate an SK; see, for example, [15].

Remark: Theorem 3 does not extend to DMCs with two or more outputs even if there is only one input. Indeed, for a DMC with a single input and $m-1 \geq 2$ outputs, the SK capacity can be positive even if each component channel $W_i,\ i = 2, \ldots, m-1$ (where $W_i(x_i|x_1)$ equals the sum of $W(x'_2, \ldots, x'_m|x_1)$ over all $x'_2, \ldots, x'_m$ with $x'_i = x_i$) has capacity 0; see ([9], Example 1). See also Example 4 in Section 6.

4 General Lower Bounds for SK and PK Capacities

Our techniques developed in [9] will be used to derive bounds for SK and PK capacities for the general DMC model introduced in Section 2. Our results are partial; unlike in [9], the lower bounds in this section and the upper bounds in the next section agree only in special cases.

One way to generate an SK or a PK for a multiterminal channel model is by *simple source emulation*. If the input terminals in $[1, k]$ use i.i.d. repetitions of a k-tuple of rvs $X_1, \ldots, X_k$, such that the X_is assigned to the nonwiretapped terminals $i \in [1, k]\backslash D$ are conditionally independent given $X_{[1,k]\cap D}$, the DMC W will generate i.i.d. repetitions of an m-tuple of rvs $X_1, \ldots, X_m$, whose joint probability mass function (pmf) is given by

$$P_{X_{\mathcal{M}}}(x_{[1,m]}) \ = \ P_{X_{[1,k]}}(x_{[1,k]}) W(x_{[k+1,m]}|x_{[1,k]}), \quad x_{[1,m]} \in \times_{i=1}^{m} \mathcal{X}_i. \tag{3}$$

with each output terminal $i \in [k+1, \ldots, m]$ observing i.i.d. repetitions of X_i. Clearly, achievable SK rates for the source model defined by $X_{\mathcal{M}} = (X_1, \ldots, X_m)$ will be achievable for the channel model, as well.

A *general* form of source emulation entails the use of an auxiliary source. Let us consider the PK generation problem with a given set $D \subset \mathcal{M}$ of wiretapped terminals; SK generation obtains as the special case $D = \emptyset$. Let $\mathcal{V}$ be a (finite) auxiliary alphabet, and consider rvs $V, X_1, \ldots, X_k$ such that $(V, X_{[1,k]\cap D})$ has an arbitrary joint pmf, and the X_is, $i \in [1, k]\backslash D$, are conditionally independent given $(V, X_{[1,k]\cap D})$. Moreover, let $X_i,\ i \in [k+1, m]$, represent the outputs of the DMC W corresponding to input $X_{[1,k]}$, satisfying the Markov condition

$$V \multimap X_{[1,k]} \multimap X_{[k+1,m]}, \tag{4}$$

so that the pmf of $(V, X_{\mathcal{M}})$ is

$$P_{VX_{\mathcal{M}}}(v, x_{[1,m]}) \;=\; P_{VX_{[1,k]\cap D}}(\tilde{v}) \prod_{i\in[1,k]\backslash D} P_{X_i|VX_{[1,k]\cap D}}(x_i|\tilde{v}) W(x_{[k+1,m]}|x_{[1,k]}), \tag{5}$$

where $\tilde{v} = (v, \{x_i, i \in [1,k] \cap D\})$.

An associated source model is defined by assigning rvs V and $X_i,\ i \in \mathcal{M}$, with a joint pmf as above, to $m+1$ terminals $0, 1, \ldots, m$, letting the set of wiretapped terminals be $\bar{D} \triangleq D \cup \{0\}$. Clearly, this source model can be emulated by our given multiterminal channel model. First, the rvs $V, X_{[1,k]\cap D}$ with an arbitrarily specified joint pmf are generated by one of the input terminals and revealed as required by the source model (since $\{0\} \cup D \cap [1,k] \subset \bar{D}$). Then, the terminals $i \in [1,k]\backslash D$ can generate the rvs X_i conditionally independently given $V, X_{[1,k]\cap D}$, and use them as their channel inputs while the rvs $X_i,\ i \in [1,k] \cap D$, are used as channel inputs by the corresponding terminals. These inputs, in turn, give rise to the channel outputs $X_i,\ i \in [k+1,m]$.

The single-letter formulas available for the SK and PK capacities of a source model [8, 9] afford lower bounds for the corresponding capacities of the multiterminal channel model, as the suprema of SK or PK capacities of source models obtainable by simple or general source emulation as above. These lower bounds will be stated formally in Theorem 4.

As in [9], given any set $A \subset \mathcal{M}$ of size $|A| \geq 2$, we denote by $\mathcal{B}(A)$ the family of all nonempty sets $B \subset \mathcal{M}$ that do not contain A, and by $\Lambda(A)$ the set of all $|\mathcal{B}(A)|$-dimensional vectors $\lambda = \{\lambda_B : B \in \mathcal{B}(A)\}$, with $0 \leq \lambda_B \leq 1$, that satisfy

$$\sum_{B\in\mathcal{B}(A):\ B\ni i} \lambda_B \;=\; 1 \;\text{ for each }\; i \in \mathcal{M}. \tag{6}$$

Also, if a set $D \subset A^c$ is given, $\mathcal{B}(A|D)$ and $\Lambda(A|D)$ are defined analogously, restricting B to subsets of D^c and replacing (6) by

$$\sum_{B\in\mathcal{B}(A|D):\ B\ni i} \lambda_B \;=\; 1 \;\text{ for each }\; i \in D^c. \tag{7}$$

In the parlance of combinatorics, the vectors in $\Lambda(A)$ (resp. $\Lambda(A|D)$) are *fractional partitions* of $\mathcal{M}$ (resp. D^c) into members of $\mathcal{B}(A)$ (resp. $\mathcal{B}(A|D)$) (cf. e.g., [13]).

The following quantities will play an important role:

$$G_A(X_{\mathcal{M}}, V, \lambda) \;\triangleq\; H(X_{\mathcal{M}}|V) \;-\; \sum_{B\in\mathcal{B}(A)} \lambda_B H(X_B|X_{B^c}, V) \tag{8}$$

and

$$G_{A|D}(X_{\mathcal{M}}, V, \lambda) \ \triangleq \ H(X_{\mathcal{M}}|X_D, V) \ - \sum_{B\in\mathcal{B}(A|D)} \lambda_B H(X_B|X_{B^c}, V), \tag{9}$$

for rvs $X_{\mathcal{M}}, V$ (the latter with values in some finite set $\mathcal{V}$), and vectors λ in $\Lambda(A)$ (resp. $\Lambda(A|D)$). We assume throughout, without further explicit mention, that the Markov condition (4) holds and that the probability mass function (pmf) of $X_{\mathcal{M}}$ is compatible with the given DMC W, i.e.,

$$P_{VX_{\mathcal{M}}}(v, x_{[1,m]}) \ = \ P_{VX_{[1,k]}}(v, x_{[1,k]})W(x_{[k+1,m]}|x_{[1,k]}), \quad v \in \mathcal{V}, \ x_{[1,m]} \in \times_{i=1}^{m}\mathcal{X}_i. \tag{10}$$

We denote by $G_A(X_{\mathcal{M}}, \lambda)$ and $G_{A|D}(X_{\mathcal{M}}, \lambda)$ the special cases for V = constant of (8) and (9), respectively.

The quantities above are related to $G_A(Q, \lambda)$ and $G_{A|D}(Q, \lambda)$ defined in ([9], eqs.(6), (7)) for a DMC with a single input and m outputs, with Q denoting the input pmf. In order to apply the results of [9], we consider below an auxiliary channel model with underlying DMC $\bar{W} : \mathcal{V} \to \mathcal{X}_1 \times \cdots \mathcal{X}_m$ (the input alphabet $\mathcal{V}$ being any finite set), which is defined by a DMC $W_0 : \mathcal{V} \to \mathcal{X}_1 \times \cdots \mathcal{X}_k$ as

$$\bar{W}(x_{[1,m]}|v) \ \triangleq \ W_0(x_{[1,k]}|v)W(x_{[k+1,m]}|x_{[1,k]}), \quad v \in \mathcal{V}, \ x_{\mathcal{M}} \in \times_{i=1}^{m}\mathcal{X}_i. \tag{11}$$

Note that the sets $\mathcal{B}(A|D)$ and $\Lambda(A|D)$ corresponding to the original model (including when $D = \emptyset$) are the same as $\mathcal{B}(A|\bar{D})$ and $\Lambda(A|\bar{D})$ corresponding to the auxiliary model with $\bar{D} \triangleq D \cup \{0\}$, where the fictitious terminal 0 depicts the input to the DMC $\bar{W}$ (as also W_0).

The rvs V and $X_{\mathcal{M}}$ can be regarded, respectively, as the input and output of the DMC $\bar{W}$; in other words, they represent a source model that can be emulated by the auxiliary channel model iff their joint pmf is of the form (10) with

$$P_{VX_{[1,k]}}(v, x_{[1,k]}) \ = \ P_V(v)W_0(x_{[1,k]}|v), \quad v \in \mathcal{V}, \ x_{[1,k]} \in \times_{i=1}^{k}\mathcal{X}_i. \tag{12}$$

This source model can be emulated by the original channel model iff the rvs $X_i, \ i \in [1, k]\backslash D$ are conditionally independent given $V, X_{[1,k]\cap D}$. For rvs $V, X_{\mathcal{M}}$ satisfying (10), (12), the quantities in (8), (9) can be written equivalently in the notation of ([9], eqs. (6), (7)) as

$$G_A(X_{\mathcal{M}}, V, \lambda) \ = \ G_{A|\bar{D}}(P_V, \lambda), \quad \bar{D} \triangleq \{0\}, \tag{13}$$

$$G_{A|D}(X_{\mathcal{M}}, V, \lambda) \;=\; G_{A|\bar{D}}(P_V, \lambda), \quad \bar{D} \triangleq D \cup \{0\}, \tag{14}$$

the right sides meant for the underlying DMC $\bar{W}$. Hence by ([9], Theorem 4.1), the minimum of $G_A(X_{\mathcal{M}}, V, \lambda)$ with respect to $\lambda \in \Lambda(A)$ is the PK capacity of the source model defined by $V, X_{\mathcal{M}}$, with privacy from terminal 0; further, maximization over P_V yields the PK capacity of the auxiliary channel model with underlying DMC $\bar{W}$. A similar statement holds for $G_{A|D}(X_{\mathcal{M}}, V, \lambda)$, with privacy from the terminals in $D \cup \{0\}$. Furthermore, by ([9], Theorem 4.2), for the auxiliary channel model these PK rates are achievable by protocols such that terminal 0 transmits a deterministic sequence, and public communication takes place only after transmission over the DMC $\bar{W}$ has been completed and consists of public messages by the terminals in $[1, m]$ with at most one message from each terminal that is a (deterministic) function of the DMC outputs therein. Note that as a consequence of their operational meaning, the quantities in (13), (14) are nonnegative.

Theorem 4: For any V, and $X_1, \ldots, X_k$ conditionally independent given V,

$$C_S(A) \;\geq\; \min_{\lambda \in \Lambda(A)} G_A(X_{\mathcal{M}}, V, \lambda). \tag{15}$$

Similarly, for any V and $X_1, \ldots, X_k$ such that $X_i, i \in [1, k] \cap D^c$ are conditionally independent given $(V, X_{[1,k] \cap D})$,

$$C_P(A|D) \geq \min_{\lambda \in \Lambda(A|D)} G_{A|D}(X_{\mathcal{M}}, V, \lambda). \tag{16}$$

Moreover, the right sides yield the largest SK or PK rates achievable by general source emulation using a particular choice of $V, X_1, \ldots, X_k$. These rates are achievable also with the terminals in $[1, k] \cap D$ transmitting deterministic sequences over the DMC W, and with a noninteractive communication protocol.

Comments: (i) The maxima of the right sides of (15), (16) with respect to the choice of $V, X_1, \ldots, X_k$ are achieved since the cardinality of the range of V can be bounded by standard techniques.

(ii) The largest SK or PK rates achievable by simple source emulation are obtained by a similar maximization of $G_A(X_{\mathcal{M}}, \lambda)$ or $G_{A|D}(X_{\mathcal{M}}, \lambda)$.

Proof: As discussed before Theorem 4, the right side of (15) is an achievable PK rate, in the auxiliary channel model with underlying DMC $\bar{W}$, for the set of terminals $A \subset \mathcal{M}$ with privacy from the input terminal 0; moreover, it is achievable by a protocol of the mentioned special kind that, in particular, has terminal 0 transmitting a deterministic sequence $v^n = (v_1, \ldots, v_n)$. The latter circumstance can be realized in the model with DMC W with the input terminals simply transmitting mutually independent rvs $X_{[1,k]t}, t = 1, \ldots, n$, with pmfs $P_{X_{[1,k]t}} = P_{X_{[1,k]}|V=v_t}$, noting that it is at this point that the conditional independence hypothesis is used.

It follows, referring again to the preceding discussion, that the right side of (15) is an achievable SK rate for the channel model with DMC W, by means of communication protocols admitting public communication only upon completion of the DMC transmissions and with each terminal $i \in \mathcal{M}$ sending at most one public message that is a function of X_i^n alone. To complete the proof of Theorem 4 in respect of (15), it remains to show that the DMC input terminals $i \in [1, k]$ need not send public messages, to which end it may be necessary to change the pmfs of the input rvs X_i^n. This can be shown exactly as the analogous assertion of ([9], Theorem 2) was proved. Consider a "good" protocol in which the terminals $i \in [1, k]$ send public messages $f_i = f_i(X_i^n)$. Proceeding as in the cited proof (replacing f_0 there with $(f_1, \ldots, f_k)$), it follows that the protocol will remain "good" if the joint pmf of all n-length channel inputs is changed to its conditional joint pmf under the condition that the values of $f_i(X_i^n), i \in [1, k]$ are equal to suitable constants. This conditioning does not affect the independence of the inputs (although their components $X_{it}, t = 1, \ldots, n$, no longer need be independent), and it reduces the public messages $f_i(X_i^n)$ of the input terminals $i \in [1, k]$ to be constants.

The assertion concerning (16) is proved in the same manner; this time, we define an auxiliary channel model with the role of V assigned to $(V, X_{[1,k]\cap D})$. It is obvious from the definition (9) of $G_{A|D}(X_{\mathcal{M}}, V, \lambda)$ that its value remains unchanged if V is replaced by $(V, X_{[1,k]\cap D})$. □

Next, restricting attention to a MAC with a single output whose capacity region is $\mathcal{C}$, by the Corollary of Theorem 2 the condition $(R, \ldots, R) \in \mathcal{C}$ is sufficient for R to be an achievable SK rate for $A = \mathcal{M}$. While it remains unclear whether this condition is necessary, the next Proposition shows that larger SK rates cannot be achieved by means of general source emulation.

Proposition 5: For a MAC $W : \mathcal{X}_1 \times \cdots \times \mathcal{X}_{m-1} \to \mathcal{X}_m$, a necessary and sufficient condition for the achievability of SK rate R with $A = \mathcal{M}$ by general source emulation, is $(R, \ldots, R) \in \mathcal{C}$.

Comment: A similar argument shows that R is achievable as an SK rate by simple source emulation iff $(R, \ldots, R)$ belongs to a polyhedron

$$\big\{(R_1, \ldots, R_{m-1}) : \ R_i \geq 0, \ \sum_{i \in B} R_i \leq I(X_B \wedge X_m | X_{B^c \setminus \{m\}}), \ B \subset [1, m-1]\big\},$$

where $X_1, \ldots, X_{m-1}$ are i.i.d. rvs and $P_{X_m | X_{[1,m-1]}} = W$. Since the capacity region $\mathcal{C}$ equals the convex closure of the union of all such polyhedra, where the union itself may be non-convex, this shows that for some MACs general source emulation can yield larger SK rates than simple source emulation; see Example 3 in Section 6 below.

Proof: Consider general source emulation involving an auxiliary rv V and input rvs $X_1, \ldots X_{m-1}$ that are conditonally independent given V, and let X_m be the corresponding output rv. By Theorem 4, the SK rate achievable by this source emulation is $\min_{\lambda \in \Lambda(\mathcal{M})} G_{\mathcal{M}}(X_{\mathcal{M}}, V, \lambda)$.

Since $X_1, \ldots, X_{m-1}$ are conditionally independent given V, and $V \multimap X_{[1,m-1]} \multimap X_m$, the expression for $G_{\mathcal{M}}(X_{\mathcal{M}}, V, \lambda)$ in (8) simplifies. Specifically,

$$\begin{aligned} H(X_{\mathcal{M}}|V) &= H(X_{[1,m-1]}|V) + H(X_m|X_{[1,m-1]}, V) \\ &= \sum_{i=1}^{m-1} H(X_i|V) + H(X_m|X_{[1,m-1]}); \end{aligned} \tag{17}$$

for $B \ni m$,

$$\begin{aligned} H(X_B|X_{B^c}, V) &= H(X_{B \setminus \{m\}}, X_m | X_{B^c}, V) \\ &= H(X_{B \setminus \{m\}} | X_{B^c}, V) + H(X_m | X_{[1,m-1]}, V) \\ &= \sum_{i \in B \setminus \{m\}} H(X_i|V) + H(X_m|X_{[1,m-1]}); \end{aligned} \tag{18}$$

and for $B \not\ni m$,

$$\begin{aligned} H(X_B|X_{B^c}, V) &= H(X_B|X_{B^c\setminus\{m\}}, V) - I(X_B \wedge X_m|X_{B^c\setminus\{m\}}, V) \\ &= \sum_{i\in B} H(X_i|V) - I(X_B \wedge X_m|X_{B^c\setminus\{m\}}, V). \end{aligned} \tag{19}$$

Substituting (17)-(19) in (8), and using $\sum_{B:B\ni i} \lambda_B = 1$ for each $i \in \mathcal{M}$, we obtain

$$G_{\mathcal{M}}(X_{\mathcal{M}}, V, \lambda) = \sum_{B\in\Lambda(\mathcal{M}):B\not\ni m} \lambda_B I(X_B \wedge X_m|X_{B^c\setminus\{m\}}, V). \tag{20}$$

For any fixed $\tilde{B} \subset [1, m-1]$, assign $\lambda \in \Lambda(\mathcal{M})$ defined by $\lambda_B = \frac{1}{|\tilde{B}|}$ if $B = \tilde{B}$ or $B = \mathcal{M}\setminus\{i\}$ for some $i \in \tilde{B}$, and $\lambda_B = 0$ otherwise. With this λ, (20) gives

$$G_{\mathcal{M}}(X_{\mathcal{M}}, V, \lambda) = \frac{1}{|\tilde{B}|} I(X_{\tilde{B}} \wedge X_m|X_{\tilde{B}^c\setminus\{m\}}, V).$$

It follows that $R = \min_{\lambda\in\Lambda(\mathcal{M})} G_{\mathcal{M}}(X_{\mathcal{M}}, V, \lambda)$ satisfies

$$R|\tilde{B}| \le I(X_{\tilde{B}} \wedge X_m|X_{\tilde{B}^c\setminus\{m\}}, V) \tag{21}$$

for every $\tilde{B} \subset [1, m-1]$, proving the necessity part of the assertion.

For sufficiency, note that $(R, \dots, R) \in \mathcal{C}$ means that for some V and $X_1, \dots, X_{m-1}$ conditionally independent given V with $V -\!\circ\!- X_{[1,m-1]} -\!\circ\!- X_m$, the inequalities (21) are satisfied. For these rvs, (20) and (21) give

$$G_{\mathcal{M}}(X_{\mathcal{M}}, V, \lambda) \ge \sum_{B\in\Lambda(\mathcal{M}):B\not\ni m} \lambda_B R|B| \ge R\Big(\sum_{B\in\Lambda(\mathcal{M})} \lambda_B|B| - \sum_{B\in\Lambda(\mathcal{M}):B\ni m} \lambda_B(m-1)\Big).$$

Since $\sum_{B\in\Lambda(\mathcal{M})} \lambda_B|B| = \sum_{i=1}^m \sum_{B\in\Lambda(\mathcal{M}):B\ni i} \lambda_B = m$ and $\sum_{B\in\Lambda(\mathcal{M}):B\ni m} \lambda_B = 1$, this proves that $\min_{\lambda\in\Lambda(\mathcal{M})} G_{\mathcal{M}}(X_{\mathcal{M}}, V, \lambda) \ge R$. □

5 General Upper Bounds for SK and PK Capacities

In order to state our upper bounds for $C_S(A)$ and $C_P(A|D)$, we extend the notation in (8) and (9) above with a slight abuse of it. Specifically, for rvs $X_{\mathcal{M}}, V$, and for $\lambda \in \Lambda(A)$ or $\lambda \in \Lambda(A|D)$, we denote

$$G_A(X_{[1,k]}, V, \lambda) \triangleq H(X_{[1,k]}|V) - \sum_{B\in\mathcal{B}(A)} \lambda_B H(X_{[1,k]\cap B}|X_{[1,k]\cap B^c}, V), \tag{22}$$

$$G_{A|D}(X_{[1,k]}, V, \lambda) \;\triangleq\; H(X_{[1,k]}|X_D, V) \;-\; \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{[1,k]\cap B}|X_{[1,k]\cap B^c}, X_D, V), \qquad (23)$$

and denote by $G_A(X_{[1,k]}, \lambda)$ and $G_{A|D}(X_{[1,k]}, \lambda)$ the special cases for $V =$ constant of (22) and (23), respectively. As earlier, we assume that the rvs $V, X_{\mathcal{M}}$ satisfy the Markov condition $V \multimap\!\!- X_{[1,k]} \multimap\!\!- X_{[k+1,m]}$ and $P_{X_{[k+1,m]}|X_{[1,k]}} = W$. Akin to $G_A(X_{\mathcal{M}}, V, \lambda)$ and $G_{A|D}(X_{\mathcal{M}}, V, \lambda)$ in (13), (14), both $G_A(X_{[1,k]}, V, \lambda)$ and $G_{A|D}(X_{[1,k]}, V, \lambda)$, too, are nonnegative, as shown in Appendix B.

Theorem 6: The SK capacity $C_S(A)$ for a set of terminals $A \subset \mathcal{M}$ and the PK capacity $C_P(A|D)$ for A with privacy from a set of terminals $D \subset A^c$ are bounded above, respectively, as follows:

$$C_S(A) \;\le\; \sup_{P_{V,X_{[1,k]}}} \inf_{\lambda \in \Lambda(A)} \Big[G_A(X_{\mathcal{M}}, V, \lambda) \;-\; G_A(X_{[1,k]}, V, \lambda) \Big], \qquad (24)$$

and for any $i \in D^c$,

$$\begin{aligned} C_P(A|D) \;\le\; & \sup_{P_{V,X_{[1,k]}}} \inf_{\lambda \in \Lambda(A|D)} \Big[G_{A|D}(X_{\mathcal{M}}, V, \lambda) \;-\; G_{A|D}(X_{[1,k]}, V, \lambda) \\ & + \sum_{B \in \mathcal{B}(A|D):\ B \not\ni i} \lambda_B I\big(X_D \wedge X_{[1,k]\cap B}|X_{[1,k]\cap B^c}, V\big) \Big]. \end{aligned} \qquad (25)$$

Corollary: If $k = m-1$ and $D \not\ni m$, then

$$C_P(A|D) \;\le \sup_{P_{V,X_{[1,m-1]}}} \inf_{\lambda \in \Lambda(A|D)} \sum_{B \in \mathcal{B}(A|D):\ B \not\ni m} \lambda_B I\big(X_m \wedge X_B|X_{[1,m-1]\cap B^c}, V\big). \qquad (26)$$

Comments: (i) If $D \subset [1,k]$, then the last term in (25) vanishes. If $D \supset [k+1,\ldots,m]$, then the difference of the first two terms is 0.
(ii) The bound (25) is tight in the special case when no more than one DMC input is uncompromised. Indeed, then $G_{A|D}(X_{[1,k]}, V, \lambda) = 0$, and the last term in (25) also vanishes – trivially when $D \supset [1,k]$, and upon taking i to be the uncompromised input otherwise. Hence, in this case (25) gives the upper bound

$$C_P(A|D) \;\le \sup_{P_{V,X_{[1,k]}}} \inf_{\lambda \in \Lambda(A|D)} \; G_{A|D}(X_{\mathcal{M}}, V, \lambda),$$

which coincides with the lower bound in Theorem 4.
(iii) The upper bound in (24) can be weakened to $C_S(A) \;\le\; \sup_{P_{V,X_{[1,k]}}} \inf_{\lambda \in \Lambda(A)} \; G_A(X_{\mathcal{M}}, V, \lambda)$.

This weaker bound differs from the lower bound in Theorem 4 by the lack of the conditional independence of $X_1, \ldots, X_k$ given V. It remains open whether (25) can be similarly weakened.

Proof: The upper bound for PK capacity in (25) is derived first. The bound in (24) for SK capacity follows as the special case $D = \emptyset$.

The initial steps in proving (25) are identical to those in the proof of the analogous converse part in ([9], Theorem 4.1). These steps are presented first in a summarized form below, which then serve as a point of departure for the rest of the proof.

Suppose that the rv $K^{(n)}$ represents an (ϵ_n, δ_n)-PK with privacy from $D \subset A^c$, achievable with randomization $U_{\mathcal{M}}$ and public communication $\mathbf{F}^{(n)}$, where $\delta_n = o(n)$ and $\epsilon_n \to 0$; see Definition 2 and the succeeding remark. As observed in ([9], the remark preceding Definition 3), we can suppose w.l.o.g. that $\mathbf{F}^{(n)} = (U_D, X_D^n, \tilde{\mathbf{F}}^{(n)})$, where $\tilde{\mathbf{F}}^{(n)}$ consists of the communication of all the terminals in D^c. Then the secrecy condition (2) is

$$\log|\mathcal{K}^{(n)}| \; - \; H(K^{(n)}|U_D, X_D^n, \tilde{\mathbf{F}}^{(n)}) \; \le \; \delta_n.$$

Using ([9], Corollary of Lemma A.2 in Appendix A) with (U_i, X_i^n) in the role of $X_i, i \in \mathcal{M}$, and X_D^n and $\tilde{\mathbf{F}}^{(n)}$ in the roles of X_D and Y, respectively, we get as in ([9], inequality (11)) that for every $\lambda = \{\lambda_B : B \in \mathcal{B}(A|D)\} \in \Lambda(A|D)$,

$$\begin{aligned} \frac{1}{n}\log|\mathcal{K}^{(n)}| \; \le \; & \frac{\alpha_n}{n}\Big[\Big\{H(U_{\mathcal{M}}, X_{\mathcal{M}}^n|U_D, X_D^n, \tilde{\mathbf{F}}^{(n)}) \\ & - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(U_B, X_B^n|U_{B^c}, X_{B^c}^n, \tilde{\mathbf{F}}^{(n)})\Big\}\Big] \; + \; \beta_n \end{aligned} \tag{27}$$

where

$$\alpha_n \; \to \; 1 \quad \text{and} \quad \beta_n \; \to \; 0 \quad \text{as} \quad n \; \to \; \infty.$$

A main ingredient of the proof of (25) will be to show that the expression within $\Big[\cdots\Big]$ above is bounded above by

$$\begin{aligned} \textstyle\sum_{t=1}^n \Big[\; & \Big(H(X_{\mathcal{M}t}|X_{Dt}) \; - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{Bt}|X_{B^ct})\Big) \\ & - \Big(H(X_{([1,k])t}|X_{Dt}) \; - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{([1,k]\cap B)t}|X_{([1,k]\cap B^c)t}, X_{Dt})\Big) \\ & + \sum_{B \in \mathcal{B}(A|D):\; B \not\ni i} \lambda_B I\Big(X_{Dt} \wedge X_{([1,k]\cap B)t}|X_{([1,k]\cap B^c)t}\Big)\Big] \end{aligned} \tag{28}$$

for $i \in D^c$. Establishing the bound (28) entails a rather tedious manipulation of information quantities, and therefore is relegated to Appendix A.

To simplify (28), a standard technique is used: Let V be an auxiliary rv distributed uniformly on $\{1, \ldots, n\}$ and independent of $X_{\mathcal{M}}^n$, and set $\tilde{X}_i \triangleq X_{iV}$, $i \in \mathcal{M}$. Then $\sum_{t=1}^n H(X_{\mathcal{M}t}|X_{Dt}) = nH(\tilde{X}_{\mathcal{M}}|\tilde{X}_D, V)$, etc., and it holds that $V -\circ- \tilde{X}_{[1,k]} -\circ- \tilde{X}_{[k+1,m]}$ and $P_{\tilde{X}_{[k+1,m]}|\tilde{X}_{[1,k]}} = W$. Finally, omitting the tildes, we obtain for the new $X_{\mathcal{M}}$ from (27), (28) (and recalling (9), (23)) that

$$\begin{aligned}\limsup_n \frac{1}{n} \log |\mathcal{K}^{(n)}| \ &\leq \ G_{A|D}(X_{\mathcal{M}}, V, \lambda) \ - \ G_{A|D}(X_{[1,k]}, V, \lambda) \\ &\quad + \sum_{B \in \mathcal{B}(A|D):\ B \not\ni i} \lambda_B I\Big(X_D \wedge X_{[1,k]\cap B} | X_{[1,k]\cap B^c}, V\Big) \qquad (29)\end{aligned}$$

for every $\lambda = \{\lambda_B : B \in \mathcal{B}(A|D)\} \in \Lambda(A|D)$ and $i \in D^c$. The claimed upper bound for PK rates in (25) follows thereupon.

Turning to the proof of the Corollary, note first that since $D \subset [1, m-1]$, the last term in (25) vanishes. For $B \in \mathcal{B}(A|D)$, since $[1, m-1] \cap B^c \supset D$, we have in (25), by (9), (23), that

$$\begin{aligned}&G_{A|D}(X_{\mathcal{M}}, V, \lambda) - G_{A|D}(X_{[1,m-1]}, V, \lambda) \\ &= \Big(H(X_{\mathcal{M}}|X_D, V) \ - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_B|X_{B^c}, V)\Big) \\ &\quad - \Big(H(X_{[1,m-1]}|X_D, V) \ - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{[1,m-1]\cap B}|X_{[1,m-1]\cap B^c}, V)\Big) \\ &= H(X_m|X_{[1,m-1]}) \ - \sum_{B \in \mathcal{B}(A|D)} \lambda_B \Big(H(X_B|X_{B^c}, V) - H(X_{[1,m-1]\cap B}|X_{[1,m-1]\cap B^c}, V)\Big) \\ &= H(X_m|X_{[1,m-1]}) \ - \sum_{B \in \mathcal{B}(A|D)} \lambda_B \Big(H(X_{[1,m-1]\cap B}|X_{B^c}, V) + H(X_{B\setminus[1,m-1]}|X_{[1,m-1]\cap B}, X_{B^c}, V) \\ &\quad - H(X_{[1,m-1]\cap B}|X_{[1,m-1]\cap B^c}, V)\Big). \qquad (30)\end{aligned}$$

Now, for $m \in B$, in the summand in the right side of (30), the expression within $\Big(\cdots\Big)$ equals

$$H(X_{[1,m-1]\cap B}|X_{B^c}, V) + H(X_m|X_{[1,m-1]}) - H(X_{[1,m-1]\cap B}|X_{B^c}, V) = H(X_m|X_{[1,m-1]}), \quad (31)$$

while for $m \notin B$, it equals

$$H(X_{[1,m-1]\cap B}|X_{[1,m-1]\cap B^c}, X_m, V)+0-H(X_{[1,m-1]\cap B}|X_{[1,m-1]\cap B^c}, V) = I(X_m \wedge X_B|X_{[1,m-1]\cap B^c}, V). \tag{32}$$

By (31), (32) and using $\sum_{B\in\mathcal{B}(A|D):\ B\ni m} \lambda_B = 1$, it follows that the right side of (30) equals $\sum_{B\in\mathcal{B}(A|D):\ B\not\ni m} \lambda_B I(X_m \wedge X_B|X_{[1,m-1]\cap B^c}, V)$, thereby leading to (26). □

6 Examples

Example 1: Consider the DMC $W : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}_3$ with $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}_3 = \{0, 1\}$ and

$$W(x_3|x_1, x_2) = \mathbb{1}(x_3 = x_1 + x_2 \mod 2),$$

where $\mathbb{1}(\cdot)$ denotes the indicator function. First, note that its capacity region is

$$\mathcal{C} = \{(R_1, R_2) :\ 0 \le R_1 + R_2 \le 1\}.$$

Consider SK generation for $A = \mathcal{M} = \{1, 2, 3\}$. The lower bound for $C_S(\{1, 2, 3\})$ provided by Theorem 2 is equal to 0.5 since (R, R) belongs to the capacity region $\mathcal{C}$ of the MAC W iff $R \le 0.5$. This SK rate is also achieved by simple source emulation; see the Comment following Proposition 5.

The achievability of an SK rate of 0.5 can be seen also separately by the following explicit scheme which generates 1 bit of *perfect* SK (i.e., (ϵ, δ)-SK with $\epsilon = \delta = 0$) for A by means of independent transmissions over the DMC by the input terminals using $n = 2$ symbols followed by public communication *only* by the output terminal. Terminal 1 transmits $X_{11} = 0$ or 1 w.p. $(0.5, 0.5)$ and $X_{12} = 0$, while terminal 2 transmits $X_{21} = 0$ and $X_{22} = 0$ or 1 w.p. $(0.5, 0.5)$, independently of (X_{11}, X_{12}). Terminal 3 then sends the public message $f_3 = f_3(X_{31}, X_{32}) = X_{31} + X_{32} \mod 2$. Clearly, all the terminals can perfectly recover $K = X_{11}$, say, from any X_i^2 and the communication $\mathbf{F} = f_3$ while satisfying the secrecy condition

$$
\begin{aligned}
s(K;\mathbf{F}) &= 1 - H(X_{11}|X_{31}+X_{32}) \\
&= 1 - H(X_{11}|X_{11}+X_{22}) \\
&= 1 - H(X_{11}) = 0.
\end{aligned}
$$

Thus, $K = X_{11}$ is a perfect SK for $A = \{1,2,3\}$ of rate 0.5.

Next, for the upper bound, apply the Corollary of Theorem 6 with $D = \emptyset$. Then, the choice $\lambda_{\{1\}} = \lambda_{\{2\}} = \lambda_{\{3\}} = 0, \lambda_{\{12\}} = \lambda_{\{13\}} = \lambda_{\{23\}} = 0.5$ yields the sum in (26) as $0.5H(X_3|V)$. It follows that $C_S(\{1,2,3\}) \leq 0.5$. Thus, $C_S(\{1,2,3\}) = 0.5$.

Furthermore, by Theorem 2, the PK capacity region $C_P(\{1,3\},\{2,3\})$ contains $\mathcal{C}$. In fact, $C_P(\{1,3\},\{2,3\}) = \mathcal{C}$, which can be seen as follows. Suppose that $C_P(\{1,3\},\{2,3\})$ contains a rate pair outside $\mathcal{C}$. By the convexity of the PK region, it contains a rate pair (R,R) with $R > 0.5$. Then, again by Theorem 2, $R > 0.5$ would be an achievable SK rate for $A = \{1,2,3\}$, which contradicts $C_S(\{1,2,3\}) = 0.5$. □

Example 2: Consider the DMC $W : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}_3$ with $\mathcal{X}_1 = \mathcal{X}_2 = \{0,1\}$, $\mathcal{X}_3 = \{0,1,2\}$ and

$$
W(x_3|x_1,x_2) = \mathbb{1}(x_3 = x_1 + x_2).
$$

Its capacity region is $\{(R_1,R_2) : \ 0 \leq R_1, R_2 \leq 1, \ R_1 + R_2 \leq 1.5\}$. The Corollary of Theorem 2 yields that $C_S(\{1,2,3\}) \geq 0.75$. By the Comment following Proposition 5, the SK rate of 0.75 is achievable also by simple source emulation. On the other hand, the Corollary of Theorem 6 gives the upper bound $C_S(\{1,2,3\}) \leq 0.5 \log 3 = 0.78$. The exact value of the SK capacity is unknown.

Next, consider PK generation for $A = \{1,2\}$ with privacy from $D = \{3\}$. Noting in (16) that the only permissible choice of $\lambda \in \Lambda(\{1,2\}|\{3\})$ is $\lambda_{\{1\}} = \lambda_{\{2\}} = 1$, we get that the right side of (16), with X_1, X_2 conditionally independent of V, equals $H(X_1|V) + H(X_2|V) - H(X_3|V)$. Thus, by Theorem 4, it follows that

$$
\begin{aligned}
C_P(\{1,2\}|\{3\}) &\geq \max_{P_{VX_1X_2} = P_V P_{X_1|V} P_{X_2|V}} H(X_1|V) + H(X_2|V) - H(X_3|V) \\
&= \max_{P_{X_1X_2} = P_{X_1} P_{X_2}} H(X_1) + H(X_2) - H(X_3) \\
&= 0.5,
\end{aligned}
$$

with the previous maximum attained by $P_{X_1}(1) = P_{X_2}(1) = 0.5$. Thus, the largest PK rate achievable by general source emulation is 0.5. The exact value of the PK capacity remains unknown, for Theorem 6 yields only the trivial upper bound 1. □

Example 3: Consider the DMC $W : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}_3$ with

$$W(0|x_1, x_2) = W(1|x_1, x_2) = 0.5, \quad \text{if} \quad x_1 = x_2 = 1,$$

$$W(x_3|x_1, x_2) = \mathbb{1}\left(x_3 = x_1 + x_2\right) \quad \text{otherwise}.$$

Its capacity region is the same as that of the MAC in Example 1. The Corollary of Theorem 2 yields the lower bound $C_S(\{1, 2, 3\}) \geq 0.5$. The same scheme for SK generation for $A = \{1, 2, 3\}$ as in Example 1 attains an SK rate of 0.5. By Proposition 5, the SK rate of 0.5 is achievable also by general source emulation. However, by the Comment following Proposition 5, it is not achievable by simple source emulation.

By the Corollary of Theorem 4, we obtain as in Example 1 that $C_S(\{1, 2, 3\}) \leq 0.5$, so that $C_S(\{1, 2, 3\}) = 0.5$. □

Example 4: Consider the DMC $W : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \mathcal{X}_{k+1} \times \cdots \times \mathcal{X}_m$ with $\mathcal{X}_1 = \cdots = \mathcal{X}_m = \{0, 1\}$ and

$W(x_{k+1}, \ldots, x_m|x_1, \ldots, x_k)$

$$= 2^{-(m-k-1)} \, \mathbb{1}\big(x_m = \sum_{i=1}^{m-1} x_i \mod 2\big),$$

i.e., the DMC outputs at terminals $k + 1, \ldots, m - 1$ are mutually independent rvs, each distributed uniformly on $\{0, 1\}$ regardless of the inputs, and the output at terminal m is the modulo 2 sum of the inputs and the remaining outputs. For SK generation for any $A \subset \mathcal{M}$ with $|A| \geq 2$, the lower bound provided by Theorem 4 yields $C_S(A) \geq \frac{1}{|A|-1}$; this SK rate of $\frac{1}{|A|-1}$ is achieved by simple source emulation. Specifically, with the input terminals in $[1, k]$ transmitting i.i.d. repetitions of a k-tuple of mutually independent rvs $X_1, \ldots, X_k$, each distributed uniformly on $\{0, 1\}$, the DMC W generates i.i.d. repetitions of an m-tuple of rvs $X_1, \ldots, X_m$, where $X_1, \ldots, X_{m-1}$ are mutually independent with each distributed uniformly on $\{0, 1\}$ and X_m is the modulo 2 sum of $X_1, \ldots, X_{m-1}$. In this emulated source model, the largest achievable SK rate for A equals $\frac{1}{|A|-1}$; see ([8], Example 1). Furthermore, in the explicit scheme provided therein, the key generated for A satisfies the secrecy condition (2) for $D = A^c$ with $\delta = 0$, thereby constituting a perfect PK for A with privacy from $D = A^c$ in addition to being a perfect SK for A. Thus, we have $C_P(A|A^c) \geq \frac{1}{|A|-1}$, too.

For the special case $A = \mathcal{M}$, the achievability of an SK rate of $\frac{1}{m-1}$ can be seen also separately by the following explicit scheme which generates 1 bit of *perfect* SK for $\mathcal{M}$ by means of independent transmissions over the DMC by the input terminals using $n = m - 1$ symbols followed by public communication *only* by the output terminals but not by the input terminals. Specifically, the input terminal $i \in [1, k]$ transmits over the DMC a sequence $X_{i1}, \ldots, X_{in}$ with X_{ii} being $\{0, 1\}$-valued w.p. $(0.5, 0.5)$ and $X_{ij} = 0, j \neq i$; all such sequences are mutually independent. The output terminal $i \in [k + 1, m - 1]$ sends a public message $f_i = f_i(X_i^n)$ which is the block $X_i^n = (X_{i1}, \ldots, X_{in})$ excluding X_{ii}, while the output terminal m sends the public message

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

where the additions are modulo 2. It is easily seen that $K = X_{11}$, say, is perfectly recoverable from X_i^n and the public communication $\mathbf{F} = (f_{k+1}, \ldots, f_m)$. Furthermore, K satisfies the secrecy condition (1) with $\delta = 0$, and so constitutes a perfect SK of rate $\frac{1}{m-1}$.

Next, in the upper bound for $C_S(A)$ in Theorem 6, we have from (24) that for any $\lambda \in \Lambda(A)$,

$$G_A(X_{\mathcal{M}}, V, \lambda) \;-\; G_A(X_{[1,k]}, V, \lambda)$$

$$
\begin{aligned}
&= \Big[H(X_{\mathcal{M}}|V) - \sum_{B\in\mathcal{B}(A)} \lambda_B H(X_B|X_{B^c},V)\Big] \\
&\quad -\Big[H(X_{[1,k]}|V) - \sum_{B\in\mathcal{B}(A)} \lambda_B H(X_{[1,k]\cap B}|X_{[1,k]\cap B^c},V)\Big] \\
&= \Big[H(X_{\mathcal{M}}|V) - \sum_{B\in\mathcal{B}(A)} \lambda_B \big(H(X_{\mathcal{M}}|V) - H(X_{B^c}|V)\big)\Big] \\
&\quad -\Big[H(X_{[1,k]}|V) - \sum_{B\in\mathcal{B}(A)} \lambda_B \big(H(X_{[1,k]}|V) - H(X_{[1,k]\cap B^c}|V)\big)\Big] \\
&= (1 - \sum_{B\in\mathcal{B}(A)} \lambda_B)\big(H(X_{\mathcal{M}}|V) - H(X_{[1,k]}|V)\big) \\
&\quad + \sum_{B\in\mathcal{B}(A)} \lambda_B \big(H(X_{B^c}|V) - H(X_{[1,k]\cap B^c}|V)\big) \\
&= (1 - \sum_{B\in\mathcal{B}(A)} \lambda_B) H(X_{[k+1,m]}|X_{[1,k]}) + \sum_{B\in\mathcal{B}(A)} \lambda_B H(X_{B^c\setminus[1,k]}|X_{[1,k]\cap B^c},V). \qquad (33)
\end{aligned}
$$

Fix $i_0 \in A$, and consider the choice $\lambda_B = \frac{1}{|A|-1}$ for $B = A\backslash\{i_0\}$ or $B = \mathcal{M}\backslash\{i\}$ for some $i \in A\backslash\{i_0\}$, and $\lambda_B = 0$ else. For this choice of $\lambda \in \Lambda(A)$, and noting that $H(X_{[k+1,m]}|X_{[1,k]}) = m-k-1$, (33) gives

$$
\begin{aligned}
&G_A(X_{\mathcal{M}}, V, \lambda) \;-\; G_A(X_{[1,k]}, V, \lambda) \\
&= \Big(1 - \frac{|A|}{|A|-1}\Big)(m-k-1) \\
&\quad + \frac{1}{|A|-1}\Big[H(X_{(A^c\cup\{i_0\})\backslash[1,k]}|X_{[1,k]\cap(A^c\cup\{i_0\})},V) + \sum_{i\in A\backslash\{i_0\}} H(X_{\{i\}\backslash[1,k]}|X_{[1,k]\cap\{i\}},V)\Big] \\
&\le -\frac{m-k-1}{|A|-1} + \frac{1}{|A|-1}\Big[H(X_{(A^c\cup\{i_0\})\backslash[1,k]}) + \sum_{i\in A\backslash\{i_0\}} H(X_{\{i\}\backslash[1,k]})\Big] \\
&\le -\frac{m-k-1}{|A|-1} + \frac{1}{|A|-1}\sum_{i=k+1}^{m} H(X_i) \\
&\le -\frac{m-k-1}{|A|-1} + \frac{m-k}{|A|-1} \\
&= \frac{1}{|A|-1},
\end{aligned}
$$

noting in the first inequality above that the summand in the last term equals 0 if $i \notin [k+1, m]$.

Thus, $C_S(A) = \frac{1}{|A|-1}$, and, in particular, $C_S(\mathcal{M}) = \frac{1}{m-1}$. Also, $C_P(A|A^c) = \frac{1}{|A|-1}$, as a PK for A with privacy from any $D \subset A^c$ is also an SK for A. Observe that Example 1 above is a special case of the present example with $m = 3,\ k = 2$ and $A = \mathcal{M}$. □

7 Discussion

We have considered secrecy generation for multiaccess channel models whose resources consist of facilities for secure noisy channel transmission from the input to the output terminals, public noiseless communication among all the terminals, and (mutually independent) randomization at the terminals. Our main results are single-letter lower and upper bounds for SK and PK capacities which agree in special cases but not in general. The lower bounds are obtained as the largest SK or PK rates achievable by general source emulation, while the upper bounds are derived by techniques developed in our earlier work [9]. However, in general, conclusive single-letter characterizations for the secrecy capacities remain elusive. The general channel model considered here appears more defiant than its special case with a single input for which single-letter characterizations of SK and PK capacities were found in [9]. Indeed, the latter capacities were achieved by simple source emulation; the converse proofs bore the main technical difficulty.

We show for a MAC model with a single output, in which all the terminals seek to share secrecy, that a necessary and sufficient condition for R to be an achievable SK rate by general source emulation is that $(R, \ldots, R)$ must lie in the capacity region of the MAC; thus, the maximum SK rate achievable by source emulation is the largest such R in the MAC capacity region. A main open question for this special model, as well as for the general channel model, is whether secrecy rates can be achieved beyond those attainable by general source emulation, by resorting to the complex transmission and communication protocols described in Section 2. We conjecture the answer to be in the affirmative, recalling the use of only simple noninteractive protocols in our achievability proofs and noting that feedback can increase the capacity region of a MAC with the use of sophisticated transmission protocols. However, a direct connection is not apparent between secrecy rates and the feedback capacity region.

Appendix A

In order to complete the proof of the upper bound (25) in Theorem 4, we show in (27) for every $\lambda = \{\lambda_B : B \in \mathcal{B}(A|D)\} \in \Lambda(A|D)$ and $i \in D^c$ that

$$\begin{aligned}
&H(U_{\mathcal{M}}, X_{\mathcal{M}}^n | U_D, X_D^n, \tilde{\mathbf{F}}^{(n)}) \; - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(U_B, X_B^n | U_{B^c}, X_{B^c}^n, \tilde{\mathbf{F}}^{(n)}) \\
&\leq \; \sum_{t=1}^{n} \Big[\Big(H(X_{\mathcal{M}t} | X_{Dt}) \; - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{Bt} | X_{B^c t}) \Big) \\
&\quad - \Big(H(X_{([1,k])t} | X_{Dt}) \; - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{([1,k] \cap B)t} | X_{([1,k] \cap B^c)t}, X_{Dt}) \Big) \\
&\quad + \sum_{B \in \mathcal{B}(A|D):\; B \not\ni i} \lambda_B I\Big(X_{Dt} \wedge X_{([1,k] \cap D)t} | X_{([1,k] \cap B^c)t} \Big) \Big], \qquad \text{(A1)}
\end{aligned}$$

and observe that the right side above equals the expression in the claimed bound in (28).

As in ([9], Appendix B), the left side of (A1) equals

$$(1 - \lambda_{\text{sum}}) H(U_{\mathcal{M}}, X_{\mathcal{M}}^n) - H(U_D, X_D^n, \tilde{\mathbf{F}}^{(n)}) + \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(U_{B^c}, X_{B^c}^n, \tilde{\mathbf{F}}^{(n)}) \qquad \text{(A2)}$$

where

$$\lambda_{\text{sum}} = \sum_{B \in \mathcal{B}(A|D)} \lambda_B \geq 1. \qquad \text{(A3)}$$

Considering the separate terms in (A2), the counterparts of (B4), (B5) and (B7) in ([9], Appendix B) are:

$$\begin{aligned}
H(U_{\mathcal{M}}, X_{\mathcal{M}}^n) &= H(U_{\mathcal{M}}) \; + \sum_{t=1}^{n} H(X_{\mathcal{M}t} | U_{\mathcal{M}}, X_{\mathcal{M}}^{t-1}) \\
&= H(U_{\mathcal{M}}) \; + \sum_{t=1}^{n} H(X_{\mathcal{M}t} | X_{([1,k])t}) \qquad \text{(A4)}
\end{aligned}$$

since $X_{([1,k])t}$ is a function of $(U_{[1,k]}, F^{t-1}) = (U_{[1,k]}, U_D, X_D^{t-1}, \tilde{F}^{t-1})$ and so is determined by $(U_{\mathcal{M}}, X_{\mathcal{M}}^{t-1})$;

$$\begin{aligned}
H(U_D, X_D^n, \tilde{\mathbf{F}}^{(n)}) = H(U_D) \; &+ \sum_{t=1}^{n} H(X_{Dt} | U_D, X_D^{t-1}, \tilde{F}^{t-1}) \\
&+ \sum_{t=1}^{n} H(\tilde{F}_t | U_D, X_D^t, \tilde{F}^{t-1}); \qquad \text{(A5)}
\end{aligned}$$

and

$$H(U_{B^c}, X^n_{B^c}, \tilde{\mathbf{F}}^{(n)}) = H(U_{B^c}) + \sum_{t=1}^{n} H(X_{B^ct}|U_{B^c}, X^{t-1}_{B^c}, \tilde{F}^{t-1}) + \sum_{t=1}^{n} H(\tilde{F}_t|U_{B^c}, X^t_{B^c}, \tilde{F}^{t-1}). \tag{A6}$$

By (A2)-(A6), the left side of (A1) decomposes as $E_1 + E_2 + E_3$ where

$$E_1 = (1 - \lambda_{\text{sum}})H(U_{\mathcal{M}}) - H(U_D) + \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(U_{B^c})$$

and

$$E_3 = \sum_{t=1}^{n} \Big[- H(\tilde{F}_t|U_D, X^t_D, \tilde{F}^{t-1}) + \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(\tilde{F}_t|U_{B^c}, X^t_{B^c}, \tilde{F}^{t-1}) \Big]$$

are the same as in ([9], Appendix B), while

$$\begin{aligned} E_2 \;\; = \;\; & \sum_{t=1}^{n} \Big[(1 - \lambda_{\text{sum}})H(X_{\mathcal{M}t}|X_{([1,k])t}) - H(X_{Dt}|U_D, X^{t-1}_D, \tilde{F}^{t-1}) \\ & + \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{B^ct}|U_{B^c}, X^{t-1}_{B^c}, \tilde{F}^{t-1})\Big]. \end{aligned} \tag{A7}$$

By ([9], Appendix B), $E_1 = 0$ and $E_3 \leq 0$. Turning to E_2, we claim that for each $1 \leq t \leq n$, the t^{th} term of the sum in E_2, denoted by E_{2t}, satisfies

$$\begin{aligned} E_{2t} \;\; \leq \;\; & \Big[\Big(H(X_{\mathcal{M}t}|X_{Dt}) - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{Bt}|X_{B^ct})\Big) \\ & -\Big(H(X_{([1,k])t}|X_{Dt}) - \sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{([1,k]\cap B)t}|X_{([1,k])\cap B^c)t}, X_{Dt})\Big) \\ & + \sum_{B \in \mathcal{B}(A|D):\ B \not\ni i} \lambda_B I\Big(X_{Dt} \wedge X_{([1,k]\cap B)t}|X_{([1,k]\cap B^c)t}\Big)\Big], \end{aligned} \tag{A8}$$

proving which will establish (A1).

For every $i \in D^c$, the first term of E_{2t} in (A7) is

$$(1 - \lambda_{\text{sum}})H(X_{\mathcal{M}t}|X_{([1,k])t}) \;\; = \;\; - \sum_{B \in \mathcal{B}(A|D):\ B \not\ni i} \lambda_B H(X_{\mathcal{M}t}|X_{([1,k])t}) \tag{A9}$$

while the third term is

$$\sum_{B \in \mathcal{B}(A|D)} \lambda_B H(X_{B^ct}|U_{B^c}, X^{t-1}_{B^c}, \tilde{F}^{t-1})$$

$$\begin{aligned}
&= \sum_{B\in\mathcal{B}(A|D)} \lambda_B H(X_{B^ct}|X_{([1,k]\cap B^c)t}, U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}) \\
&= \sum_{B\in\mathcal{B}(A|D):\ B\not\ni i} \lambda_B H(X_{B^ct}|X_{([1,k]\cap B^c)t}, U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}) \\
&\quad + \sum_{B\in\mathcal{B}(A|D):\ B\ni i} \lambda_B H(X_{B^ct}|X_{([1,k]\cap B^c)t}, U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}) \\
&\le \sum_{B\in\mathcal{B}(A|D):\ B\not\ni i} \lambda_B H\big(X_{(B^c\setminus[1,k])t}|X_{([1,k]\cap B^c)t}\big) \\
&\quad + \sum_{B\in\mathcal{B}(A|D):\ B\ni i} \lambda_B H(X_{B^ct}|U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}), \qquad \text{(A10)}
\end{aligned}$$

where the first equality holds since $X_{([1,k]\cap B^c)t}$ is a function of $(U_{[1,k]\cap B^c}, F^{t-1}) = (U_{[1,k]\cap B^c}, U_D, X_D^{t-1}, \tilde{F}^{t-1})$ and so is determined by $(U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1})$. Further, as in ([9], p. 2451, column 2, line 18 onward), since $B \in \mathcal{B}(A|D)$ implies $B^c \supset D$,

$$H(X_{B^ct}|U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}) - H(X_{Dt}|U_D, X_D^{t-1}, \tilde{F}^{t-1})$$

$$\begin{aligned}
&\le H(X_{B^ct}|U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}) - H(X_{Dt}|U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}) \\
&= H(X_{B^ct}|X_{Dt}, U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}) \\
&= H(X_{B^ct}|X_{Dt}) - I(X_{B^ct} \wedge U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}|X_{Dt}),
\end{aligned}$$

so that in the right side of (A10),

$$\begin{aligned}
&\sum_{B\in\mathcal{B}(A|D):\ B\ni i} \lambda_B H(X_{B^ct}|U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}) \\
&\le \sum_{B\in\mathcal{B}(A|D):\ B\ni i} \lambda_B \Big[H(X_{Dt}|U_D, X_D^{t-1}, \tilde{F}^{t-1}) + H(X_{B^ct}) - H(X_{Dt}) - I(X_{B^ct}\wedge U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}|X_{Dt})\Big].
\end{aligned} \qquad \text{(A11)}$$

Combining (A9), (A10), (A11), and recalling that $\sum_{B\in\mathcal{B}(A|D):\ i\in B} \lambda_B = 1$ for $i \in D^c$, we get that E_{2t} is bounded above as

$$\begin{aligned}
E_{2t} \;\le\; &- \sum_{B\in\mathcal{B}(A|D):\ B\not\ni i} \lambda_B \Big[H(X_{\mathcal{M}t}|X_{([1,k])t}) - H\big(X_{(B^c\setminus[1,k])t}|X_{([1,k]\cap B^c)t}\big)\Big] \\
&+ \sum_{B\in\mathcal{B}(A|D):\ B\ni i} \lambda_B \Big[H(X_{B^ct}) - H(X_{Dt}) - I(X_{B^ct} \wedge U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}|X_{Dt})\Big] \text{(A12)}
\end{aligned}$$

Now, we observe in the right side of (A12) that the summands in the first and second sums, respectively, are

$$H(X_{\mathcal{M}t}) - H(X_{[1,k]t}) - H(X_{B^ct}) + H\big(X_{([1,k]\cap B^c)t}\big)$$

$$\begin{aligned}
&= H(X_{Bt}|X_{B^ct}) - H\big(X_{([1,k]\cap B)t}|X_{([1,k]\cap B^c)t}\big) \\
&= H(X_{Bt}|X_{B^ct}) - H\big(X_{([1,k]\cap B)t}|X_{([1,k]\cap B^c)t}, X_{Dt}\big) \\
&\quad -I\Big(X_{Dt} \wedge X_{([1,k]\cap B)t}|X_{([1,k]\cap B^c)t}\Big) \qquad \text{(A13)}
\end{aligned}$$

and

$$\begin{aligned}
&H(X_{B^ct}) - H(X_{Dt}) - I\Big(X_{B^ct} \wedge X_{([1,k]\cap B^c)t}, U_{B^c}, X_{B^c}^{t-1}, \tilde{F}^{t-1}|X_{Dt}\Big) \\
&\leq H(X_{B^ct}) - H(X_{Dt}) - I\Big(X_{B^ct} \wedge X_{([1,k]\cap B^c)t}|X_{Dt}\Big) \\
&= H(X_{B^ct}) - H(X_{Dt}) - H(X_{([1,k]\cap B^c)t}|X_{Dt}) \\
&= H(X_{\mathcal{M}t}) - H(X_{Bt}|X_{B^ct}) - H(X_{Dt}) - H(X_{([1,k])t}|X_{Dt}) \\
&\quad +H(X_{([1,k]\cap B)t}|X_{([1,k]\cap B^c)t}, X_{Dt}), \qquad \text{(A14)}
\end{aligned}$$

where the insertion of $X_{([1,k]\cap B^c)t}$ in the first expression in (A14) is permissible for the reason in the passage following (A10).

Finally, from (A12), (A13) and (A14), we get that for every $i \in D^c$,

$$\begin{aligned}
E_{2t} &\leq -\sum_{B\in\mathcal{B}(A|D):\ B\not\ni i} \lambda_B\Big[H(X_{Bt}|X_{B^ct}) - H(X_{([1,k]\cap B)t}|X_{([1,k])\cap B^c)t}, X_{Dt}) \\
&\quad -I\Big(X_{Dt} \wedge X_{([1,k]\cap B)t}|X_{([1,k]\cap B^c)t}\Big)\Big] \\
&\quad +\sum_{B\in\mathcal{B}(A|D):\ B\ni i} \lambda_B\Big[H(X_{\mathcal{M}t}) - H(X_{Bt}|X_{B^ct}) - H(X_{Dt}) \\
&\quad -H(X_{([1,k])t}|X_{Dt}) + H(X_{([1,k]\cap B)t}|X_{([1,k])\cap B^c)t}, X_{Dt})\Big] \\
&= \Big[\Big(H(X_{\mathcal{M}t}|X_{Dt}) - \sum_{B\in\mathcal{B}(A|D)} \lambda_B H(X_{Bt}|X_{B^ct})\Big) \\
&\quad -\Big(H(X_{([1,k])t}|X_{Dt}) - \sum_{B\in\mathcal{B}(A|D)} \lambda_B H(X_{([1,k]\cap B)t}|X_{([1,k])\cap B^c)t}, X_{Dt})\Big) \\
&\quad +\sum_{B\in\mathcal{B}(A|D):\ B\not\ni i} \lambda_B I\Big(X_{Dt} \wedge X_{([1,k]\cap B)t}|X_{([1,k]\cap B^c)t}\Big)\Big],
\end{aligned}$$

which is (A8). □

Appendix B

The proof of the nonnegativity of (23) relies on the following technical lemma; that of (22) follows with $D = \emptyset$.

Lemma B: Let $L = \{i_1, \ldots, i_l\} \subset \mathcal{M}$, $i_1 < \cdots < i_l$, and $D \subset \mathcal{M}$ be arbitrary sets with $L \backslash D \neq \emptyset$. For rvs $X_{\mathcal{M}}, Y$, and for every collection $\lambda = \{\lambda_B : B \subset D^c\}$ of weights $0 \leq \lambda_B \leq 1$ satisfying

$$\sum_{B \subset D^c:\ B \ni i} \lambda_B = 1 \ \text{ for all } \ i \in L \backslash D, \tag{B1}$$

it holds that

$$\sum_{B \subset D^c} \lambda_B H(X_{L \cap B} | X_{L \cap B^c}, Y) \leq H(X_{L \backslash D} | Y). \tag{B2}$$

Comment: Lemma B is a special case of Lemma B1 in [9] and also of Theorem 1 in [14].

Proof: We have

$$\begin{aligned}
&\sum_{B \subset D^c} \lambda_B H(X_{L \cap B} | X_{L \cap B^c}, Y) \\
&\qquad = \sum_{B \subset D^c} \lambda_B \sum_{j : i_j \in L \cap B} H(X_{i_j} | X_{\{i_1, \ldots, i_{j-1}\} \cap B}, X_{L \cap B^c}, Y) \\
&\qquad \leq \sum_{B \subset D^c} \sum_{j : i_j \in L \cap B} \lambda_B H(X_{i_j} | X_{\{i_1, \ldots, i_{j-1}\}}, Y) \\
&\qquad = \sum_{j : i_j \in L \backslash D} \sum_{B \subset D^c : B \ni i_j} \lambda_B H(X_{i_j} | X_{\{i_1, \ldots, i_{j-1}\}}, Y) \\
&\qquad = \sum_{j : i_j \in L \backslash D} H(X_{i_j} | X_{\{i_1, \ldots, i_{j-1}\}}, Y), \quad \text{by } (B1) \\
&\qquad \leq \sum_{j : i_j \in L \backslash D} H(X_{i_j} | X_{\{i_1, \ldots, i_{j-1}\} \backslash D}, Y) \\
&\qquad = H(X_{L \backslash D} | Y).
\end{aligned}$$

□

The claimed nonnegativity of (23) follows upon taking $L = X_{[1,k]}$ and $Y = (X_D, V)$ in Lemma B. This Lemma also provides a formal proof of the nonnegativity of (8), (9), with $L = X_{\mathcal{M}}$.

Acknowledgments

The work of I. Csiszár was supported by the Hungarian National Foundation for Scientific Research under Grant 76088. The work of P. Narayan was supported by the U.S. National Science Foundation under Grants CCF0830697 and CCF1117546.

References

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

[2] A. A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals I-II," *IEEE Trans. Inform. Theory*, vol. 56, pp. 3973-3996, 3997-4010, 2010.

[3] C. H. Bennett, G. Brassard and J. M. Robert," "Privacy amplification by public discussion," *Siam J. Computing*, vol. 17, no. 2, pp. 210-229, 1988.

[4] C.H. Bennett, G. Brassard, C. Crépeau and U.M. Maurer, "Generalized privacy amplification," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1915-1923, November 1995.

[5] I. Csiszár, "Almost independence and secrecy capacity," *Probl. Pered. Inform.* (Special issue devoted to M.S. Pinsker), vol. 32, no. 1, pp. 48-57, 1996.

[6] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, pp. 339-348, May 1978.

[7] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems, Second Edition*, Cambridge, 2011.

[8] I. Csiszár and P. Narayan, "The secret key capacity of multiple terminals," *IEEE Trans. Inform. Theory* vol. 50, pp. 3047-3061, Dec. 2004.

[9] I. Csiszár and P. Narayan, "Secrecy capacities for multiterminal channel models," *IEEE Trans. Inform. Theory, Special Issue on Information Theoretic Security* vol. 54, pp. 2437-2452, June 2008.

[10] I. Csiszár and P. Narayan, "Secrecy generation for multiple input multiple output channel models," *Proceedings of the IEEE International Symposium on Information Theory*, Seoul, Korea, pp. 2447-2451, June-July 2009.

[11] A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals - Part I: Source model," *Preprint*, 2008.

[12] A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals - Part II: Channel model," *Preprint*, 2008.

[13] M. Madiman and A. Barron, "Generalized entropy power inequalities and monotonicity properties of information," *IEEE Trans. Inform. Theory*, vol. 53, no. 7, pp. 2317-2329, July 2007.

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

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

[16] U. M. Maurer, "The strong secret key rate of discrete random triples," in *Communications and Cryptography: Two Sides of One Tapestry*, R. E. Blahut *et al.*, Eds., Kluwer, Norwell, MA, Ch. 26, pp. 271-285, 1994.

[17] U. M. Maurer and S. Wolf, "Unconditionally secure key agreement and the intrinsic conditional information," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 499-514, March 1999.

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

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

[20] A. D. Wyner, "The wiretap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355-1387, 1975.