Secret-Key Capacity of Compound Source Models with One-Way Public Communication

Nima Tavangaran[1], Holger Boche[1], and Rafael F. Schaefer[2]

[1]Technische Universität München, Germany, {nima.tavangaran, boche}@tum.de
[2]Princeton University, USA, rafaelfs@princeton.edu

***Abstract*—In the classical Secret-Key generation model, Common Randomness is generated by two terminals based on the observation of correlated components of a common source, while keeping it secret from a non-legitimate observer. It is assumed that the statistics of the source are known to all participants. In this work, the Secret-Key generation model based on a compound source is studied where the source realization is unknown. The protocol should guarantee the security and reliability of the generated Secret-Key, simultaneously for all possible realizations of the compound source. A single-letter lower-bound of the Secret-Key capacity is derived for the case where the public communication rate is limited. Furthermore, a multi-letter capacity formula is computed for the case where the public communication is unconstrained.**

I. Introduction

Current cryptographic approaches are dependent on the computational capabilities of the terminals. By increasing technological advances, the security of transmitted information can not be guaranteed for sure. In contrast, an information theoretic approach provides us with a framework for future coding schemes which are independent of computational capabilities of the eavesdroppers.

Information theoretic security was first introduced by Shannon in [1], in the so called one-time pad method where each transmitting message is encrypted by a Secret-Key (SK). Another step towards achieving communication security is to generate a shared SK based on a common source. In this model, two terminals observe correlated components of a common source and communicate over a public noiseless channel to generate a common SK, based on their knowledge. They can encrypt subsequent communication using this SK. This procedure is based on the generation of a Common Randomness (CR)-based information which was first introduced in [2] and later used by Maurer in [3], and Ahlswede and Csiszár in [4] to determine the SK capacity. The SK sharing is further studied in [5]–[7]. In practice, this kind of security can be integrated in the physical layer of wireless systems and for instance take advantage of the CR of the Ultra-Wideband (UWB) channel impulse response between two terminals to generate a SK [8].

However, in all these models which were used for SK generation, perfect knowledge of the source realization in the whole procedure was assumed. In a more general approach, the source uncertainty should be taken into account where the terminals do not have the knowledge of the actual realization of the source. While compound source coding as a related problem, was studied in [9], [10], an achievable SK rate for a compound Discrete Memoryless Multiple Source (DMMS) $\{(X, Y_s, Z_s)\}_{s\in\mathcal{S}}$ was given in [11]. In [12], the compound DMMS $\{(X_s, Y_s)\}_{s\in\mathcal{S}}$ without an eavesdropper was studied and the SK capacity was computed. Finally in [13], an achievable SK rate for a channel model with Arbitrarily Varying Channel (AVC) states at the eavesdropper was derived.

In this work, a SK generation model for a compound DMMS with one-way communication in presence of an eavesdropper is studied. The terminals observe a compound source $\mathfrak{S} := \{XYZ, s\}_{s\in\mathcal{S}} := \{(X_s, Y_s, Z_s)\}_{s\in\mathcal{S}}$ and two of them generate a shared SK by only a one-way communication over a public noiseless channel while keeping it secret from the third terminal (eavesdropper). As the source realization index $s \in \mathcal{S}$ is unknown to the terminals, an estimation method such as hypothesis testing is incorporated to find the marginal source index of the transmitter. This approach is used to generalize the model in [5], [7, Section 17.3] to the compound setup.

In Section II, the general model for SK generation is presented. Section III gives the main results followed by a short proof sketch. A single-letter lower-bound for the SK capacity is derived when the communication rate over the public channel is constrained. A multi-letter SK capacity formula is computed as well for the case where the public communication rate is unconstrained. A summary of formal proofs for the main results is given in Section IV. Finally, Section V concludes the paper.

Notation: For the typical sequences and their related sets the same definitions as in [7, Chapters 2 and 17] are taken. $\mathbb{1}_{\mathcal{A}}(\cdot)$ denotes the indicator function for a set $\mathcal{A}$. For any function f, the cardinal number of the range of the function is denoted by $\|f\|$. Random Variables (RVs) are denoted by capital letters (e.g. $X^n, U, \cdots$), their realization by small letters (e.g. $x^n, u, \cdots$), their range (alphabet) by script letters (e.g. $\mathcal{X}^n, \mathcal{U}, \cdots$), and their Probability Distribution (PD) by Roman letters (e.g. $\mathrm{P}_{X^n}, \mathrm{P}_U, \cdots$). All alphabets corresponding to RVs are supposed to be finite. $H(X)$ and $I(X;Y)$ represent the entropy of a RV X and the mutual information between X and Y respectively. $h(a)$ with $a \in [0,1]$ is the binary entropy function and is given by $h(a) := -a\log a - (1-a)\log(1-a)$. Finally, $X - Y - Z$ denotes a Markov chain for RVs X, Y, and Z.

This work of N. Tavangaran and H. Boche was supported by the German Research Foundation (DFG) under Grant BO 1734/20-1.

This work of R. F. Schaefer was supported by the German Research Foundation (DFG) under Grant WY 151/2-1.

978-1-4673-7852-9/15/$31.00 ©2015 IEEE

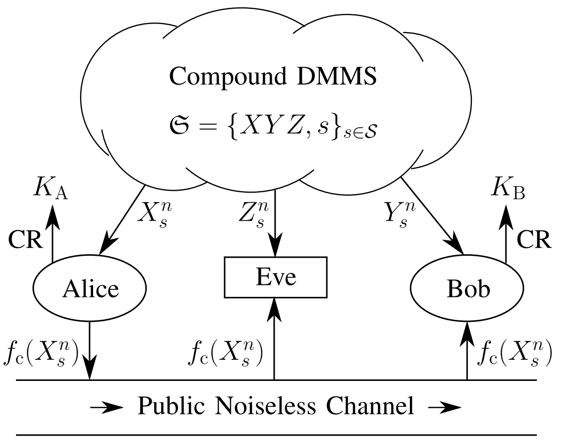

Fig. 1. SK generation protocol for compound DMMS model.

II. SK Generation Model

Figure 1 shows the SK generation model which is used throughout this work. Transmitter (Alice), receiver (Bob) and eavesdropper (Eve) observe a compound DMMS $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ for time duration $n \in \mathbb{N}$. Therefore, RVs X_s^n, Y_s^n, and Z_s^n represent their initial knowledge for the source state $s \in \mathcal{S}$. It is assumed that all terminals know the set of source states $\mathcal{S}$ as well as its statistics with PDs $\{\mathrm{P}_{XYZ,s}\}_{s\in\mathcal{S}}$. However, they do not have the knowledge of the actual realization $s \in \mathcal{S}$ of the source. The next definition describes the SK generation model which is studied through out this work.

Definition 1. *The SK generation model consists of a transmitter (Alice), a receiver (Bob), an eavesdropper (Eve), a compound DMMS which generates their initial knowledge, and a public noiseless communication channel between all terminals. The source is given for a finite set of states $\mathcal{S}$, by a sequence of generic RVs $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ taking their values in the finite set $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$.*

As RVs X_s^n and Y_s^n are correlated, Alice and Bob may generate some CR-based information by communicating over the public channel. In this work, only a one-way communication over the public channel is allowed. The following definition gives a more precise description of this procedure.

Definition 2. *A one-way SK generation protocol for the model which is given in Definition 1 with source $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ consists of the following two steps:*

- *After observing X_s^n, Alice transmits $f_{\mathrm{c}}(X_s^n)$ to Bob over the public noiseless channel. f_{c} is a deterministic function of X_s^n and is called public communication function.*
- *Next, Alice generates a SK, represented by a RV K_{A}, based on her knowledge X_s^n and Bob generates a SK, represented by a RV K_{B}, based on his knowledge $(Y_s^n, f_{\mathrm{c}}(X_s^n))$. K_{A} and K_{B} take their values in $\mathcal{K}$.*

Similarly as in [7, Problem 17.15(a)], it can be shown that a randomized f_{c} does not increase the SK rate and capacity.

As the communication over the public channel is also received by Eve, this should not reveal any information about the SK. Moreover, the generated SK should have a uniform distribution. Combining these two criteria together leads to a compact notation which was first introduced in [6] and is called security index.

Definition 3. *For RVs K_{A} and V, taking value in the sets $\mathcal{K}$ and $\mathcal{V}$ respectively, the security index is given by*

$$S(K_{\mathrm{A}}|V) := \log(|\mathcal{K}|) - H(K_{\mathrm{A}}) + I(K_{\mathrm{A}}; V).$$

K_{A} represents the SK and V Eve's knowledge. This short notation is a powerful tool which guarantees both the strong secrecy [14] and uniformity of the generated SK.

The next definition uses this concept to define an achievable SK rate and capacity. Similarly as in [4], [7], the communication rate constraint is also part of the achievability definition. This is because, in a realistic model, the information exchange rate between the terminals is restricted.

Definition 4. *A real number $R_{\mathrm{sk}} \geq 0$ is an achievable SK rate for the model in Definition 1 with source $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ and a one-way communication over the public noiseless channel with rate constraint $\Gamma \in (0, +\infty]$, if and only if, for all $\delta > 0$, and all $n \in \mathbb{N}$ large enough, there exists a SK generation protocol with public communication function f_{c}, giving rise to the RVs K_{A} and K_{B} with values in $\mathcal{K}$, for which it holds:*

- $\frac{1}{n}\log\|f_{\mathrm{c}}\| < \Gamma + \delta$,
- $R_{\mathrm{sk}} < \frac{1}{n}\log|\mathcal{K}| + \delta$,
- $\forall s \in \mathcal{S}, \quad \Pr(K_{\mathrm{A}} \neq K_{\mathrm{B}}) < \delta$,
- $\forall s \in \mathcal{S}, \quad S(K_{\mathrm{A}}|Z_s^n, f_{\mathrm{c}}(X_s^n)) < \delta$.

The SK capacity $C_{\mathrm{sk}}(\mathfrak{S}, \Gamma)$ for this model is defined to be the supremum of all achievable SK rates. If there is no communication rate constraint i.e. $\Gamma = \infty$, then the first condition in the definition is inactive and the capacity is simply denoted by $C_{\mathrm{sk}}(\mathfrak{S})$.

In the following, a subset of the compound set $\mathcal{S}$ is defined. This definition is required for stating the results in Section III.

Definition 5. *Let for the source $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$, $\hat{\mathcal{S}}$ be the set of all possible states of marginal RV X. For a given marginal state $\hat{s} \in \hat{\mathcal{S}}$, corresponding to the RV $X_{\hat{s}}$, the set of all possible joint source states is given by*

$$\mathcal{I}(\hat{s}) := \Big\{s \in \mathcal{S} : \forall x \in \mathcal{X}, \\ \sum_{y\in\mathcal{Y}}\sum_{z\in\mathcal{Z}} \mathrm{P}_{XYZ,s}(x,y,z) = \mathrm{P}_{X_{\hat{s}}}(x)\Big\}.$$

III. SK Capacity Results and Lower-Bound

In this section, a single-letter SK capacity lower-bound as well as a multi-letter SK capacity representation for the compound DMMS model is presented and a short proof sketch is provided for the first result. For Theorem 1, the capacity lower-bound is given as a function of public communication rate constraint.

Theorem 1. *For a compound DMMS model with source $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ and a one-way communication over a public noiseless channel with constraint $\Gamma \in (0,\infty]$, it holds:*

$$C_{\mathrm{sk}}(\mathfrak{S},\Gamma) \geq \min_{\hat{s}\in\hat{\mathcal{S}}} \max_{U_{\hat{s}},V_{\hat{s}}} \Big\{ \min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};Y_s|U_{\hat{s}}) - \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};Z_s|U_{\hat{s}})\Big\}, \quad (1)$$

where the outer max *is taken over all RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ such that it holds:*

$$\forall s\in\mathcal{I}(\hat{s}),\ U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}} - YZ, s \quad \textit{and}$$

$$\max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};X_{\hat{s}}|Y_s) + \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};X_{\hat{s}}|UY,s) < \Gamma.$$

Proof sketch: To achieve the SK rate in (1), Alice estimates her marginal state $\hat{s}\in\hat{\mathcal{S}}$ by hypothesis testing such that the estimation error is exponentially small [15]. Similarly as in [12], she sends it along with other information related to her observation to Bob. In Figure 1, this is denoted by $f_{\mathrm{c}}(X_s^n)$. It can be shown that the protocol guarantees all the conditions of Definition 4 for achievability.

Given an estimated marginal source state of Alice $\hat{s}\in\hat{\mathcal{S}}$, the joint source state s is not necessarily known to the terminals. However, by Definition 5, it is known that $s\in\mathcal{I}(\hat{s})$. For the correctly estimated state of Alice, say $\hat{s}\in\hat{\mathcal{S}}$, Lemma 1 from Section IV assures that Alice and Bob can generate a CR which is universal for all possible source states $s\in\mathcal{I}(\hat{s})$.

Furthermore, the coding scheme in Lemma 1, should work with respect to the given communication rate constraint $\Gamma > 0$.

Finally, as seen in Figure 1, Alice and Bob generate their SKs K_{A} and K_{B} based on this CR-based information using a SK generator. However, $f_{\mathrm{c}}(X_s^n)$ is also received by Eve. Lemma 2, again from Section IV, assures the existence of a SK generator which guarantees the strong secrecy and uniformity of the SK K_{A}, for all possible $s\in\mathcal{I}(\hat{s})$, even if the estimation RV $\hat{S}_s$ which is a part of the transmitted message $f_{\mathrm{c}}(X_s^n)$, reveals some information about the SK. □

Next, a multi-letter SK capacity formula is computed for the case where no communication rate constraint is given.

Theorem 2. *For a compound DMMS model with source $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ and a one-way communication over a public noiseless channel, it holds:*

$$C_{\mathrm{sk}}(\mathfrak{S}) = \lim_{n\to\infty} \frac{1}{n} \min_{\hat{s}\in\hat{\mathcal{S}}} \max_{U_{\hat{s}},V_{\hat{s}}} \Big\{ \min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};Y_s^n|U_{\hat{s}}) - \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};Z_s^n|U_{\hat{s}})\Big\}, \quad (2)$$

where the outer max *is taken over all RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ such that it holds:*

$$\forall s\in\mathcal{I}(\hat{s}),\ U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}}^n - Y^nZ^n, s.$$

IV. Proofs

In the following, Lemmas 1 and 2 for compound sets are presented. The proofs are similar to the non-compound versions which are available in [7, Chapter 17], [5], [16]. In Lemma 1, if the values in the equations (3)-(6) are not integers, the smallest integer which is larger than the given expression is taken.

Lemma 1. *Let $\delta > 0$ and $\sigma > \zeta > 0$ be all in $\mathbb{R}$ and sufficiently small. Let $\hat{s}\in\hat{\mathcal{S}}$ be given and for all $s\in\mathcal{I}(\hat{s})$, the Markov chains $U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}} - Y_s$ hold. Consider $N_{\hat{s},1}N_{\hat{s},2}$ sequences $u_{ij}^n(\hat{s})\in\mathcal{U}^n$, chosen independently by $\mathrm{P}_{U_{\hat{s}}^n}$ where*

$$i\in\mathcal{I} := \{1,2,\cdots,N_{\hat{s},1}\},\ j\in\mathcal{J} := \{1,2,\cdots,N_{\hat{s},2}\},$$

$$N_{\hat{s},1} := \exp\Big[n\big(\max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};X_{\hat{s}}|Y_s) + 3\delta\big)\Big], \quad (3)$$

$$N_{\hat{s},2} := \exp\Big[n\big(\min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Y_s) - 2\delta\big)\Big]. \quad (4)$$

For each $u_{ij}^n(\hat{s})$, consider $N_{\hat{s},3}N_{\hat{s},4}$ sequences ${v_{pq}^{ij}}^n(\hat{s})\in\mathcal{V}^n$, chosen conditionally independently by $\mathrm{P}_{V_{\hat{s}}|U_{\hat{s}}}^n(\cdot|u_{ij}^n(\hat{s}))$ where,

$$p\in\mathcal{P} := \{1,2,\cdots,N_{\hat{s},3}\},\ q\in\mathcal{Q} := \{1,2,\cdots,N_{\hat{s},4}\},$$

$$N_{\hat{s},3} := \exp\Big[n\big(\max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};X_{\hat{s}}|UY,s) + 3\delta\big)\Big], \quad (5)$$

$$N_{\hat{s},4} := \exp\Big[n\big(\min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};Y_s|U_{\hat{s}}) - 2\delta\big)\Big]. \quad (6)$$

Assume that all random sequences $\{u_{ij}^n(\hat{s})\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$ and $\{{v_{pq}^{ij}}^n(\hat{s})\}_{(p,q)\in\mathcal{P}\times\mathcal{Q}}$ are known to Alice and Bob. Then it holds:

a) For $n\in\mathbb{N}$ sufficiently large, there exist encoder functions $\hat{f}:\mathcal{T}\to\mathcal{I}$ and $\hat{g}:\mathcal{T}\to\mathcal{J}$, with a probability approaching 1, doubly exponentially fast where

$$\mathcal{T} := \Big\{x^n\in\mathcal{X}^n : \mathcal{T}^n_{[UX,\hat{s}]\zeta}(x^n) \neq \emptyset\Big\}, \quad (7)$$

and if $\hat{f}(x^n) = i$, $\hat{g}(x^n) = j$ then $(u_{ij}^n(\hat{s}), x^n) \in \mathcal{T}^n_{[UX,\hat{s}]\zeta}$. Alice encodes her observation $x^n\in\mathcal{T}$ by these functions to the sequence $u_{ij}^n(\hat{s})$.

Furthermore, for functions $\hat{f}$ and $\hat{g}$, extending them to $f:\mathcal{X}^n\to\mathcal{I}\cup\{0\}$ and $g:\mathcal{X}^n\to\mathcal{J}\cup\{0\}$ such that

$$\forall x^n\in\mathcal{T},\ f(x^n) = \hat{f}(x^n),\ g(x^n) = \hat{g}(x^n),$$
$$\forall x^n\notin\mathcal{T},\ f(x^n) = g(x^n) = 0,$$

there exists a decoder $\tilde{g}:\mathcal{I}\times\hat{\mathcal{S}}\times\mathcal{Y}^n\to\mathcal{J}$, with which Bob can reconstruct $g(x^n)$ from $(f(x^n),\hat{s},y^n)$, with an error probability approaching 0, exponentially fast for all $s\in\mathcal{I}(\hat{s})$.

b) For each f and g from part a), and $n\in\mathbb{N}$ sufficiently large, there exist functions $\hat{\varphi}:\mathcal{T}\to\mathcal{P}$ and $\hat{\rho}:\mathcal{T}\to\mathcal{Q}$ with a probability approaching 1, doubly exponentially fast, such that if $\hat{f}(x^n) = i$, $\hat{g}(x^n) = j$, $\hat{\varphi}(x^n) = p$, $\hat{\rho}(x^n) = q$ then $(u_{ij}^n(\hat{s}), {v_{pq}^{ij}}^n(\hat{s}), x^n) \in \mathcal{T}^n_{[UVX,\hat{s}]\sigma}$. Alice encodes her observation $x^n\in\mathcal{T}$ by these functions to the sequence ${v_{pq}^{ij}}^n(\hat{s})$.

Furthermore, for functions $\hat{\varphi}$ and $\hat{\rho}$, extending them to $\varphi:\mathcal{X}^n\to\mathcal{P}\cup\{0\}$ and $\rho:\mathcal{X}^n\to\mathcal{Q}\cup\{0\}$ such that

$$\forall x^n\in\mathcal{T},\ \varphi(x^n) = \hat{\varphi}(x^n),\ \rho(x^n) = \hat{\rho}(x^n),$$
$$\forall x^n\notin\mathcal{T},\ \varphi(x^n) = \rho(x^n) = 0,$$

there exists a decoder $\tilde{\rho}:\mathcal{I}\times\mathcal{J}\times\mathcal{P}\times\hat{\mathcal{S}}\times\mathcal{Y}^n\to\mathcal{Q}$, with which Bob can reconstruct $\rho(x^n)$ from $(f(x^n), g(x^n), \varphi(x^n), \hat{s}, y^n)$, with an error probability approaching 0, exponentially fast for all $s\in\mathcal{I}(\hat{s})$.

Lemma 2. *Let $\hat{s} \in \hat{\mathcal{S}}$ be given and C, D_s, and $\hat{S}_s$ with $s \in \mathcal{I}(\hat{s})$ be RVs taking value in $\mathcal{C}, \mathcal{D}$, and $\hat{\mathcal{S}}$ respectively. Assume $\alpha \in (0, \frac{1}{6}]$ and $\eta \in (0, \frac{1}{3}]$ with $\alpha \leq \eta$ are given and for all $s \in \mathcal{I}(\hat{s})$, there exist sets $\mathcal{B}_s \subset \mathcal{C} \times \mathcal{D}$ with*

$$\forall (c,d) \in \mathcal{B}_s,\ \mathrm{P}_{CD,s|\hat{S}_s}(c,d|\hat{s}) < \frac{1}{\alpha|\mathcal{B}_s|},$$
$$\mathrm{P}_{CD,s|\hat{S}_s}(\mathcal{B}_s|\hat{s}) \geq 1 - (\eta^2 - \alpha^2).$$

Furthermore, define the sets $\mathcal{B}_{s,d} := \big\{c \in \mathcal{C} : (c,d) \in \mathcal{B}_s\big\}$ and $\mathcal{D}_s := \big\{d \in \mathcal{D} : \mathcal{B}_{s,d} \neq \emptyset\big\}$, and assume

$$k \in \mathbb{N},\ k < \alpha^6 \min_{s \in \mathcal{I}(\hat{s}), d \in \mathcal{D}_s} |\mathcal{B}_{s,d}|\ , \quad k < \frac{\mathrm{e}^{1/\alpha}}{2|\mathcal{D}|\,|\mathcal{I}(\hat{s})|}. \tag{8}$$

Then, there exists a SK generator $\kappa : \mathcal{C} \to \{1, 2, \cdots, k\}$ such that for all $s \in \mathcal{I}(\hat{s})$,

$$S(\kappa(C)|D_s, \hat{S}_s = \hat{s}) \leq (\alpha + 2\eta)\log k + h(\alpha + \eta), \tag{9}$$

with a probability at least $1 - 2k\,|\mathcal{I}(\hat{s})|\,|\mathcal{D}|\,\mathrm{e}^{-\frac{\alpha^5 \min |\mathcal{B}_{s,d}|}{k}}$ where the min *in the exponent is taken over all $s \in \mathcal{I}(\hat{s})$ and $d \in \mathcal{D}_s$.*

In the following, short proofs of Theorems 1 and 2 are presented. Similar techniques which are used for deriving the non-compound SK capacity in [7, Section 17.3], [5] are used in the following proofs and extended to the compound setup.

Proof of Theorem 1. The proof is divided into two parts:

Part a) Assume $\min_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) > \max_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s)$ and define

$$R'_{\mathrm{sk}} := \min_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) - \max_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s).$$

We show that R'_{sk} is achievable for RV $U_{\hat{s}}$ satisfying

$$U_{\hat{s}} - X_{\hat{s}} - YZ, s \quad \text{and} \quad \max_{\hat{s} \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) < \Gamma. \tag{10}$$

As explained in Section III, Alice estimates her marginal statistic by hypothesis testing. Assume $\hat{s} \in \hat{\mathcal{S}}$ is the index corresponding to the correct decision and for all other $\tilde{s} \in \hat{\mathcal{S}} - \{\hat{s}\}$ a wrong decision is made. In the rest of the proof, it is shown that all conditions of Definition 4 are satisfied.

Let $s \in \mathcal{S}$ be given and the resulting estimated marginal state be denoted by the RV $\hat{S}_s$ taking value in $\hat{\mathcal{S}}$ and having the PD $\mathrm{P}_{\hat{S}_s}$. Therefore, it holds by [15] and [7, Problem 2.13b]

$$\mathrm{P}_{\hat{S}_s}(\hat{s}) \geq 1 - \exp(-nc_0), \tag{11}$$
$$\forall \tilde{s} \in \hat{\mathcal{S}} - \{\hat{s}\},\ \mathrm{P}_{\hat{S}_s}(\tilde{s}) \leq \exp(-nc_1), \tag{12}$$

for some $c_0, c_1 > 0$. Next, Alice sends her estimated marginal source state to Bob over the public noiseless channel.

Assume $0 < \xi < \zeta < \sigma$, and $\delta > 0$ are all in $\mathbb{R}$. For the case that hypothesis testing has led to the correct decision, it holds that $s \in \mathcal{I}(\hat{s})$. Moreover, the independently generated $N_{\hat{s},1}N_{\hat{s},2}$ sequences $u^n_{ij}(\hat{s}) \in \mathcal{U}^n$ from Lemma 1a) are known to both Alice and Bob. Based on this sequences, Lemma 1a) implies the existence of the encoder functions $f : \mathcal{X}^n \to \mathcal{I} \cup \{0\}$ and $g : \mathcal{X}^n \to \mathcal{J} \cup \{0\}$ with the given properties. Alice sends also the index $f(x^n) = i$ of $u^n_{ij}(\hat{s})$ to Bob over the public channel. Again Lemma 1a) implies the existence of a decoder $\tilde{g} : \mathcal{I} \times \hat{\mathcal{S}} \times \mathcal{Y}^n \to \mathcal{J}$, with which Bob can reconstruct $g(x^n)$ from $(f(x^n), \hat{s}, y^n)$ with an error probability approaching 0, exponentially fast for all $s \in \mathcal{I}(\hat{s})$. For all other estimation results leading to a wrong decision, the probability of happening an error is given by (12) which is also exponentially small.

The whole message which is sent over the public channel is represented by the RV $f_{\mathrm{c}}(X^n_{\hat{s}}) = (f(X^n_{\hat{s}}), \hat{S}_s)$. For the communication rate $\frac{1}{n}\log\|f_{\mathrm{c}}\|$ from Definition 4, it holds that

$$\frac{1}{n}\log\|f_{\mathrm{c}}\| = \max_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + 3\delta + \frac{1}{n}\log|\hat{\mathcal{S}}| < \Gamma + 3\delta,$$

where the equality follows by (3) and the inequality by (10).

After the index $g(x^n) = j$ is reconstructed by Bob, both Alice and Bob may generate their SK, based on this CR-based data. Thus, it remains to show that there exists a SK generator $\kappa : \mathcal{J} \to \{1, 2 \cdots, k\}$, giving rise to the RV $K_{\mathrm{A}} = \kappa(g(X^n_{\hat{s}}))$, which satisfies the security condition of Definition 4. Define

$$\mathcal{T}_s := \Big\{(x^n, z^n) : x^n \in \mathcal{T}, (u^n_{f(x^n)g(x^n)}(\hat{s}), x^n, z^n) \in \mathcal{T}^n_{[UXZ,s]\sigma}\Big\},$$
$$C := g(X^n_{\hat{s}}), \qquad D_s := \big(f(X^n_{\hat{s}}), Z^n_s, \mathbb{1}_{\mathcal{T}_s}(X^n_{\hat{s}}, Z^n_s)\big),$$
$$\mathcal{B}_s := \Big\{\big(j, (i, z^n, 1)\big) : (i,j) \in \mathcal{I} \times \mathcal{J}, z^n \in \mathcal{T}^n_{[Z_s]\xi},$$
$$\mathcal{T}^n_{[UXZ,s]\sigma}(u^n_{ij}(\hat{s}), z^n) \neq \emptyset\Big\},$$

where the set $\mathcal{T}$ was given in (7). In the following, it is shown that all conditions of Lemma 2 are satisfied.

$$\mathrm{P}_{CD,s|\hat{S}_s}(\mathcal{B}_s|\hat{s}) = \sum_{(j,(i,z^n,1)) \in \mathcal{B}_s} \mathrm{P}_{CD,s|\hat{S}_s}\big(j, (i, z^n, 1)|\hat{s}\big)$$
$$\geq 1 - \frac{1}{1 - \exp(-nc_0)}\Big[\mathrm{P}_{X^n_{\hat{s}}Z^n_s}(\mathcal{T}^{\mathrm{c}}_s) + \mathrm{P}_{Z^n_s}({\mathcal{T}^n_{[Z_s]\xi}}^{\mathrm{c}})\Big]$$
$$\geq 1 - \frac{\exp(-nc_2)}{1 - \exp(-nc_0)} \geq 1 - \exp(-nc_3),$$

for some $c_2, c_3 > 0$ and n sufficiently large. Define α and η for some arbitrary $\tau > 0$ as follows

$$\alpha := \exp(-n(\delta + 5\tau)), \quad \eta := \exp(-n\delta). \tag{13}$$

This implies for n large enough and δ and τ sufficiently small

$$\mathrm{P}_{CD,s|\hat{S}_s}(\mathcal{B}_s|\hat{s}) \geq 1 - (\eta^2 - \alpha^2).$$

Similar as in [7, Section 17.3], it can be shown that for ξ and ζ sufficiently small and n large enough,

$$|\mathcal{B}_s| \leq \exp\Big[n\big(H(Z_s) + \tau\big)\Big] \times$$
$$\exp\Big[n\big(I(U_{\hat{s}}; X_{\hat{s}}) + \delta - I(U_{\hat{s}}; Z_s) + \tau\big)\Big]. \tag{14}$$

Furthermore, for all $(j, (i, z^n, 1)) \in \mathcal{B}_s$, it holds that

$$\big(1 - \exp(-nc_0)\big) \cdot \mathrm{P}_{CD,s|\hat{S}_s}\big(j, (i, z^n, 1)|\hat{s}\big) \tag{15}$$
$$\leq \exp\Big[n\big(H(X_{\hat{s}}|UZ, s) + \tau\big)\Big]\exp\Big[-n\big(H(XZ, s) - \tau\big)\Big],$$

where the inequality follows from (11). Finally by (14), (15), and definition of α in (13), it follows that

$$|\mathcal{B}_s|\,\mathrm{P}_{CD,s|\hat{S}_s}\big(j, (i, z^n, 1)|\hat{s}\big)\,\alpha \leq \frac{\exp(-n\tau)}{1 - \exp(-nc_0)} < 1.$$

Therefore, Lemma 2 implies that there exists a SK generator function $\kappa : \mathcal{J} \to \{1, 2, \cdots, k\}$ such that (9) holds. Hence,

$$\begin{aligned} S\big(K_{\mathrm{A}} \,\big|\, f(X_{\hat{s}}^n), Z_s^n, \mathbb{1}_{\mathcal{T}_s}(X_{\hat{s}}^n, Z_s^n), \hat{S}_s = \hat{s}\big) \\ \le (\alpha + 2\eta) \log k + h(\alpha + \eta) \\ \le n \cdot \exp(-nc_4) + \exp(-nc_4) \le \exp(-nc_5), \quad (16) \end{aligned}$$

for some $c_4, c_5 > 0$ and $K_{\mathrm{A}} = \kappa(g(X_{\hat{s}}^n))$. Therefore in total, for all estimation results which may lead to a correct or incorrect decision, it holds that

$$\begin{aligned} S\big(K_{\mathrm{A}} \,\big|\, f(X_{\hat{s}}^n), Z_s^n, \mathbb{1}_{\mathcal{T}_s}(X_{\hat{s}}^n, Z_s^n), \hat{S}_s\big) \\ = \mathrm{P}_{\hat{S}_s}(\hat{s}) S\big(K_{\mathrm{A}} | Z_s^n, f(X_{\hat{s}}^n), \mathbb{1}_{\mathcal{T}_s}(X_{\hat{s}}^n, Z_s^n), \hat{S}_s = \hat{s}\big) + \\ \sum_{\tilde{s} \in \hat{\mathcal{S}} - \{\hat{s}\}} \mathrm{P}_{\hat{S}_s}(\tilde{s}) S\big(K_{\mathrm{A}} | Z_s^n, f(X_{\hat{s}}^n), \mathbb{1}_{\mathcal{T}_s}(X_{\hat{s}}^n, Z_s^n), \hat{S}_s = \tilde{s}\big) \\ \le \exp(-nc_5) + n \cdot \exp(-nc_6) \le \exp(-nc_7), \end{aligned}$$

for some $c_6, c_7 > 0$. The inequality is a result of (16) and (12).

Finally, k should also satisfy the conditions in (8). Therefore, similarly as in [7, Section 17.3], it can be shown that for $\tau > 0$ and $\zeta > 0$ both sufficiently small and n large enough,

$$\begin{aligned} \min_{s \in \mathcal{I}(\hat{s}), d \in \mathcal{D}_s} |\mathcal{B}_{s,d}| \ge \min_{s \in \mathcal{I}(\hat{s}), d \in \mathcal{D}_s} \left| \left\{ j : u_{ij}^n(\hat{s}) \in \mathcal{T}_{[UZ,s]\zeta}^n(z^n) \right\} \right| \\ \ge \exp\left[n\big(R'_{\mathrm{sk}} - 2\delta - \tau\big) \right]. \end{aligned}$$

Part b) It may be assumed that

$$\min_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) \le \max_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s). \quad (17)$$

Because otherwise, it follows that

$$\begin{aligned} R_{\mathrm{sk}} := \min_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s | U_{\hat{s}}) - \max_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s | U_{\hat{s}}) \\ \le \min_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s) - \max_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s) \\ - \Big[\min_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) - \max_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s) \Big], \end{aligned}$$

and by part a) of this lemma, R_{sk} would be achievable.

The SK rate R_{sk} can be shown to be achievable for the case when (17) holds and $R_{\mathrm{sk}} > 0$. The proof is very similar to part a) of this proof, by using Lemma 1b) and Lemma 2. □

Proof of Theorem 2. The direct part of the proof is trivial and follows using Theorem 1 and Fekete's lemma [17]. For the converse let $R_{\mathrm{sk}} > 0$ be an achievable SK rate. Next, Alice sends a message $f_{\mathrm{c}}(X_{\hat{s}}^n)$ to Bob over the public channel and generates a SK represented by K_{A}. It holds by Fano's inequality and Definitions 3 and 4 that for all $\epsilon > 0$ and n large enough,

$$\begin{aligned} \frac{1}{n} \log|\mathcal{K}| < \frac{1}{n} \Big[\min_{s \in \mathcal{I}(\hat{s})} H\big(K_{\mathrm{A}} | Z_s^n, f_{\mathrm{c}}(X_{\hat{s}}^n)\big) - \\ \max_{s \in \mathcal{I}(\hat{s})} H\big(K_{\mathrm{A}} | Y_s^n, f_{\mathrm{c}}(X_{\hat{s}}^n)\big) \Big] + \frac{1}{n} \epsilon \log|\mathcal{K}| + \frac{1}{n} + \epsilon. \end{aligned}$$

For $\delta > 0$, $\epsilon' = \epsilon/(1-\epsilon) + \delta$, n sufficiently large, and by Definition 4, it implies that

$$\begin{aligned} R_{\mathrm{sk}} < \frac{1}{n} \log|\mathcal{K}| + \delta \le \frac{1}{1-\epsilon} \cdot \frac{1}{n} \Big[\min_{s \in \mathcal{I}(\hat{s})} I\big(K_{\mathrm{A}}; Y_s^n | f_{\mathrm{c}}(X_{\hat{s}}^n)\big) \\ - \max_{s \in \mathcal{I}(\hat{s})} I\big(K_{\mathrm{A}}; Z_s^n | f_{\mathrm{c}}(X_{\hat{s}}^n)\big) \Big] + \epsilon'. \end{aligned}$$

Set RVs $U_{\hat{s}} := f_{\mathrm{c}}(X_{\hat{s}}^n)$ and $V_{\hat{s}} := (f_{\mathrm{c}}(X_{\hat{s}}^n), K_{\mathrm{A}})$. It holds by definition of $U_{\hat{s}}$ and $V_{\hat{s}}$ that $I(X_{\hat{s}}^n; U_{\hat{s}} | V_{\hat{s}}) = 0$. Furthermore, As $f_{\mathrm{c}}(X_{\hat{s}}^n)$ and K_{A} are both functions of $X_{\hat{s}}^n$, it implies that $I(Y_s^n, Z_s^n; UV, \hat{s} | X_{\hat{s}}^n) = 0$ which proves the Markov chain. Furthermore, take the maximum with respect to $U_{\hat{s}}$ and $V_{\hat{s}}$.

Finally, because $\hat{s} \in \hat{\mathcal{S}}$ was chosen arbitrarily, taking ϵ and ϵ' sufficiently small and n large enough, completes the proof. □

V. Conclusion

The SK generation protocol which was introduced in this work used a two phase approach to achieve the given SK rate. In the first step, Alice estimated her state and sent this along with other information, which was obtained from her observation, to Bob. Although, this information is also received by Eve, it was shown that the strong secrecy is still guaranteed. In the second step, Bob used this information including the estimated state of Alice to reconstruct the SK. A single-letter lower-bound for the SK capacity was derived while the one-way public communication rate between Alice and Bob was kept low by a given upper-bound. This result was further extended to a multi-letter SK capacity formula.

References

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[2] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems of Control and Information Theory*, vol. 2, pp. 149–162, 1973.

[3] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

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

[5] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 344–366, March 2000.

[6] ——, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3047–3061, December 2004.

[7] I. Csiszár and J. Körner, *Information Theory, Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University, 2011.

[8] R. Wilson, D. Tse, and R. A. Scholtz, "Channel identification: Secret sharing using reciprocity in ultrawideband channels," *IEEE T-IFS*, vol. 2, no. 3, pp. 364–375, September 2007.

[9] J.-H. Jahn, "Kodierung beliebig variierender korrelierter Quellen," Ph.D. dissertation, Universität Bielefeld, Fakultät für Mathematik, 1978.

[10] S. C. Draper and E. Martinian, "Compound conditional source coding, slepian-wolf list decoding, and applications to media coding," in *Proceedings of IEEE ISIT*, 2007, pp. 1511–1515.

[11] M. Bloch, "Channel intrinsic randomness," in *Proceedings of IEEE International Symposium on Information Theory*, 2010, pp. 2607–2611.

[12] H. Boche and R. F. Wyrembelski, "Secret key generation using compound sources - optimal key-rates and communication costs," in *Proceedings of 9th International ITG SCC*, 2013.

[13] R. A. Chou and M. R. Bloch, "Secret-key generation with arbitrarily varying eavesdroppers channel," in *Proceedings of IEEE GlobalSIP*, 2013, pp. 277–280.

[14] U. M. Maurer and S. Wolf, "Information-theorectic key agreement: From weak to strong secrecy for free," *EUROCRYPT 2000, Lecture Notes in Computer Science, Springer-Verlag*, vol. 1807, pp. 351–368, May 2000.

[15] W. Hoeffding, "Asymptotically optimal tests for multinomial distributions," *Annals of Mathematical Statistics*, vol. 36, pp. 369–401, 1965.

[16] I. Csiszár, "Almost independence and secrecy capacity," *Problems of Information Transmission*, vol. 32, no. 1, pp. 48–57, 1996.

[17] M. Fekete, "Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten," *Mathematische Zeitschrift*, vol. 17, pp. 228–249, 1923.