Technische Universität München
Fakultät für Elektrotechnik und Informationstechnik
Lehrstuhl für Theoretische Informationstechnik

Robust Secret-Key Generation under Source Uncertainty and Communication Rate Constraint

Nima Tavangaran

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr.-Ing. Georg Sigl

Prüfer der Dissertation:

1. Prof. Dr.-Ing. Dr. rer. nat. Holger Boche
2. Prof. H. Vincent Poor, Ph.D. (schriftliche Beurteilung)
   Prof. Dr.-Ing. Antonia Wachter-Zeh (mündliche Prüfung)
3. Jun.-Prof. Dr.-Ing. Rafael F. Schaefer

Die Dissertation wurde am 20.09.2017 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 26.02.2018 angenommen.

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. A key is then generated using this common randomness while keeping it secret from a non-legitimate observer. It is assumed that the probability distribution of the source is known to all participants. On the other hand, for secret-key generation by using a compound source, the actual source and thus its probability distribution are unknown to the participants. It is assumed that the actual source belongs to a compound set which is known to the participants and does not change during the observation of the source. The secret-key generation protocol in this case, should guarantee the reliability and security of the generated secret-key simultaneously for all elements of the compound set.

In this work, secret-key generation based on a three-party compound source is studied where eavesdropper's side information is also taken into account and strong secrecy is guaranteed. At the same time, the public communication rate constraint between the legitimate users is part of the secret-key generation protocol. To this end, in the first step finite compound sources are applied in the model. A single-letter lower bound of the secret-key capacity is derived as a function of the public communication rate constraint. The multi-letter secret-key capacity formula is computed as well, as a function of the public communication rate constraint. In the second step the result for single-letter lower bound is then reformulated as a region of secret-key rate versus communication rate constraint pairs. It is shown that this region is convex, even if the compound set is infinite. Based on this fact and by using approximation techniques, the secret-key capacity results are then extended to arbitrary (possibly infinite) compound sources with finite marginal sets. Furthermore, the single-letter secret-key capacity of a degraded compound source is derived where the compound set is again arbitrary (possibly infinite) and the set of marginals is finite.

This setup is also applied to two-party biometric authentication systems. In such systems, the authenticity of each user is determined by comparing two secret-keys which are generated based on the user's biometric enrollment and authentication sequences. These sequences are generated, based on the correlated components of a common source whose probability distribution is unknown to the users and belongs to a compound set. In this case, the protocol should guarantee the security and reliability of the authentication procedure, simultaneously for all realizations of the compound source. In this work, a single-letter secret-key rate versus privacy leakage rate capacity region of a compound biometric authentication system is given for the case where strong secrecy is also guaranteed. To this end, a special case of the results for three-party secret-key compound model is applied to the biometric authentication system.

Acknowledgements

First and foremost, I would like to thank my advisor Prof. Holger Boche for giving me the opportunity to work with him at Chair of Theoretical Information Technology, Technische Universität München, and for all his advice and motivation. I express my sincere gratitude to Prof. H. Vincent Poor and Prof. Rafael F. Schaefer, who agreed to be the second and third referees of my dissertation. Further, I thank Prof. Georg Sigl for serving as the chairman of the dissertation committee and Prof. Antonia Wachter-Zeh for acting as the additional referee. I also thank all my colleagues at Technische Universität München, who were always a constant source of inspiration. Many thanks to Dr. Harout Aydinian for sharing the office with me during a large part of my doctoral studies. At last but not least, my special thanks and gratitude go to my parents for their encouragement and endless motivation.

Contents

Notation

$:=$	To be defined as
iff	If and only if
$\mathbb{N}$	Set of natural numbers
$\mathbb{R}$	Set of real numbers
$\mathbb{R}^{+}$	Set of all $x \in \mathbb{R}$ such that $x \geq 0$
$\mathbb{R}^{++}$	Set of all $x \in \mathbb{R}$ such that $x > 0$
$\mathcal{A}, \mathcal{B}, \cdots$	Sets
$\emptyset$	Empty set
$\mathcal{A} \subset \mathcal{B}$	$\mathcal{A}$ is a subset of $\mathcal{B}$
$\mathcal{A} \supset \mathcal{B}$	$\mathcal{A}$ is a superset of $\mathcal{B}$
$\mathcal{A}^{\mathrm{c}}$	Complement of a set $\mathcal{A}$
$\mathcal{A} - \mathcal{B}$	The same as $\mathcal{A} \cap \mathcal{B}^{\mathrm{c}}$
$2^{\mathcal{A}}$	Set of all subsets of a given set $\mathcal{A}$
$\mathscr{A}, \mathscr{G}, \cdots$	σ-Algebras, class of sets
$(\mathcal{A}_i)_{i \in \mathbb{N}} \in \mathscr{A}$	Sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \cdots$ all as elements of $\mathscr{A}$
$(\mathcal{A}_i)_{i \in \mathbb{N}} \subset \mathscr{A}$	Sequence of sets $\mathcal{A}_1, \mathcal{A}_2, \cdots$ all as subsets of $\mathscr{A}$
$\mathfrak{S}, \hat{\mathfrak{S}}, \cdots$	Compound source sets
$X, Y, Z, U, V, \hat{S}, \cdots$	Random variables
$\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{U}, \mathcal{V}, \hat{\mathcal{S}}, \cdots$	Alphabets (ranges) of random variables
$P_X, P_Y, P_Z, P_U, P_V, P_{\hat{S}}, \cdots$	Probability distributions of random variables
$\mathcal{P}(\mathcal{A})$	Family of all probability distributions with range $\mathcal{A}$
$P \ll Q$	P is absolutely continuous with respect to Q
$\|P - Q\|$	Norm-1 distance between $P, Q \in \mathcal{P}(\mathcal{X})$
$D(P\|Q)$	Kullback-Leibler divergence between $P, Q \in \mathcal{P}(\mathcal{X})$
$\mathrm{supp}(\cdot)$	The support set of a given probability distribution
$f : \mathcal{X} \to \mathcal{Y}$	Function with domain $\mathcal{X}$ and image $f(\mathcal{X}) \subset \mathcal{Y}$
$\|f\|$	Cardinal number of the image of the function f
$W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$	Stochastic matrix
$\times$	Cartesian product
$\otimes$	Tensor product (product σ-algebras or measures)
X^n	Random variable $(X_1, X_2, \cdots X_n)$
x^n	Realization $(x_1, x_2, \cdots x_n)$ of random variable X^n
$X^n Y^n Z^n$	Random variable $((X_1, Y_1, Z_1), (X_2, Y_2, Z_2), \cdots (X_n, Y_n, Z_n))$
$x^n y^n z^n$	Realization $((x_1, y_1, z_1), (x_2, y_2, z_2), \cdots (x_n, y_n, z_n))$ of random variable $X^n Y^n Z^n$
$\mathbb{E}[\cdot]$	Mathematical expectation

$U - X - YZ$	Markov chain
$\mathbb{1}_{\mathcal{A}}(\cdot)$	Indicator function
$\log(\cdot)$	Logarithm function to the base 2
$\exp(\cdot)$	Exponential function with base 2
e	Euler's number
ln	Natural logarithm function with base e
$H(\cdot)$	Entropy of a given random variable
$h(\cdot)$	Binary entropy function
$I(\cdot\,;\cdot)$	Mutual information between two random variables
PW	For $P \in \mathcal{P}(\mathcal{X})$, $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ $PW \in \mathcal{P}(\mathcal{Y})$ and $\forall y \in \mathcal{Y},\ PW(y) := \sum_{x \in \mathcal{X}} P(x)W(y\|x)$

Abbreviations

AVC	Arbitrarily Varying Channel
CR	Common Randomness
DMC	Discrete Memoryless Channel
DMS	Discrete Memoryless Source
DMMS	Discrete Memoryless Multiple Source
iid	independently and identically distributed
PD	Probability Distribution
RV	Random Variable
SK	Secret-Key

1 Introduction

Current cryptographic approaches are dependent on the computational capabilities of the non-legitimate terminals. By increasing technological advances and computational capabilities of adversaries, the security of transmitted information between legitimate users can not be guaranteed for sure. In contrast, an information theoretic [1, 2] approach provides us with a framework for future coding schemes that guarantee security independent of computational capabilities of the eavesdroppers.

Cryptographic approaches are not only dependent on the computational capabilities of non-legitimate terminals, but also vulnerable to unknown attacks and potential security gaps. Information theoretic security in contrast, gives an absolute guarantee on the security of the implemented system based on a mathematical model [3].

The next generation of wireless communication systems should support applications with very low communication latency, availability, high reliability and security. Tactile Internet is an example of such systems which has many applications in industry, robotics, virtual and augmented reality, and health care. To address the security aspects of such systems, a low latency security protocol is required. The classical encryption algorithms are implemented in the higher protocol layers of the communication systems, causing extra delays. In contrast, information theoretic security works directly in the physical layer and does not generate much latency. This leads to an absolute security in the sense of secure communication and authentication of participants with very low latency. To achieve this kind of security, new system infrastructures should be implemented to satisfy the mentioned requirements [4, 5]. As an example, this kind of security can be integrated in the physical layer of wireless systems [6].

The information theoretic security is currently developing very fast such that leading network operators and institutes for secure public communication have made this field as their central research area [7, 8]. The information theoretic source model which is investigated in this work addresses both Secret-Key (SK) generation and authentication in such systems.

Information theoretic security was first introduced by Shannon in [2] and later by Wyner in [9]. SK sharing by using a common source is an example of a model which is based on information theoretic security. In this model, two terminals (Alice and Bob) observe correlated components of a common source and communicate over a public noiseless channel to generate a common SK. Afterwards, they can encrypt subsequent communication using this SK. This procedure relies on the generation of a Common Randomness (CR) which was introduced in [10] by Gács and Körner. This concept was further used by Maurer in [11] and Ahlswede and Csiszár in [12] to determine the SK capacity. In [13] the CR capacity was derived.

Among the available SK sharing source models, some setups are based on three-party

sources where the knowledge of transmitter (Alice), receiver (Bob), and eavesdropper (Eve) is represented by a three-party joint Random Variable (RV). In such models, Eve's side information (Eve's knowledge) is also taken into account as an RV. The SK generation protocol in this case, should guarantee the security of the generated SK also based on this parameter. On the other hand, if no Eve's side information is considered, the model is based on a two-party joint RV representing only the knowledge of Alice and Bob.

To guarantee the security of the generated SK, different kinds of secrecy conditions are applied. In weak secrecy, only the rate of information which is leaked from the generated SK to eavesdropper is verified. On the contrary, in strong secrecy [14], the whole information which is leaked is considered. This concept is explained in more details, where the SK sharing and biometric authentication models are introduced in Chapters 3 and 6 respectively.

The communication between legitimate users over the public channel can be a one-way (forward) communication initiated by Alice or it can be a multi-way communication. In this work, only the forward communication models are studied. The rate of forward communication which is required to generate a shared SK can be interpreted as part of the cost in a real system. Furthermore, the transmitted message over the public channel is a function of marginal source representing Alice's knowledge. Therefore, the public communication rate can also be interpreted as a measure of information about the source itself, which is revealed to the eavesdroppers (privacy leakage) [15, 16]. This concept is explained more where the biometric authentication systems are introduced as a closely related problem to the SK sharing.

In [11], a three-party SK sharing source model was introduced and SK capacity bounds were determined. In [12], a two-party SK sharing source model was introduced and then extended to a three-party system where Eve's side information is also taken into account. The SK capacity was determined in this case. In both [11] and [12], only one-way communication between legitimate users (Alice and Bob) is allowed and no constraint is put on the public communication rate between them. Furthermore, weak secrecy is guaranteed.

In [17], Csiszár and Narayan introduced a three-party SK sharing source model where Alice and Bob generate a shared SK based on their observations and communication over a public noiseless channel. Furthermore, they keep the generated key secret from Eve. In the protocol design, the communication rate constraint between Alice and Bob is taken into account and an exponentially strong secrecy is used in the definition of security. Both forward and multi-way communication scenarios are investigated and SK capacities are derived. The forward SK capacity is derived in this case as a function of communication rate upper bound (constraint). This work is also available in [18, Theorem 17.21]. The SK sharing is further studied in [18–25].

As a closely related problem, biometric authentication systems can also be studied by information theoretic models [15, 26]. The biometric systems are very attractive to be used in authentication systems due to their uniqueness and almost time-invariable characteristics. Classical non-biometric authentication systems usually work with secret passwords or physical tokens to guarantee the legitimacy of a user. Biometric authentica-

tion systems in contrast, use the physical characteristics and biometric measurements of the authenticating person to guarantee the legitimacy of the person.

In [15], both two-party and three-party biometric authentication systems are studied and capacity regions are derived. In [16], a two-party biometric authentication system is introduced and the capacity regions are given as well, where weak secrecy is guaranteed.

Biometric authentication systems which are given in [16] consist of two phases, namely the enrollment and authentication. In the enrollment phase, the biometric measurements of each person are gathered and based on that an SK is generated. As the captured biometric measurements might be affected by noise, some helper data must be used to deal with this noisy data. Next, this helper data together with an encrypted version of the SK is stored in a central database. In the authentication phase, the user generates again an SK, based on his newly generated biometric sequence and the stored helper data from the database. He encrypts the SK with the same method which was used in the enrollment phase. The authenticity is approved if both encrypted SKs from enrollment and authentication phases are identical. As the database is publicly available, the challenge in this scheme is to prevent the helper data to reveal asymptotically any information about the SK (secrecy leakage) and as low as possible information about the biometric sequence (privacy leakage).

The privacy leakage limitation [27] is a crucial condition in biometric authentication systems. This is because the biometric data of each person is limited and can not be renewed in case it is stolen by an eavesdropper. A stronger condition which also guarantees the privacy leakage rate limitation in two-party systems is the communication rate constraint from the model in [17].

In all these models for SK generation, perfect knowledge of the source statistics is 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. However, they know that the actual source belongs to a known uncertainty (compound) set and that it remains constant during the entire observation of the source. The SK generation protocol in this case should guarantee the security and reliability of the generated SK simultaneously for all elements of the compound set.

In [28], a three-party finite compound source is studied where Alice's marginal source is fixed. A lower bound for the SK capacity is given in this case, where no communication rate constraint is considered. In [29], a two-party finite compound source is studied. The SK capacity is computed in this case, where again no communication rate constraint is considered. In [30] a two-party finite compound source in a biometric authentication system is studied and the capacity region is given. In [31], a three-party compound source is studied in the quantum regime where the compound source is assumed to be regular and arbitrary large (possibly infinite). The multi-letter SK capacity formula is derived in this case, where no communication rate constraint is considered. As other related problems, compound source coding, SK generation with Arbitrarily Varying Channels (AVCs), compound wiretap channels, and AVC capacity are studied in [32–41].

In this work, the non-compound source model in [17], [18, Section 17.3] is extended to a compound source system. To this end, an SK generation model for a three-party

system (Alice, Bob, and Eve) using an arbitrary (possibly infinite) compound source is introduced where the class of marginal sources is assumed to be finite. In this case, Eve's side information is also taken into account. The terminals observe the compound source and two of them (Alice and Bob) generate a shared SK by communicating over a public noise channel. The communication is assumed to be one-way and initiated by Alice. Furthermore, the information which is transmitted over the public channel should not reveal asymptotically any information about the generated SK to eavesdropper. The public communication rate constraint between Alice and Bob is further considered in the protocol design. As the model is based on a compound source, the actual source realization is unknown to the terminals. In this case, an estimation method such as hypothesis testing is incorporated to find the Alice's source (marginal).

Chapter 2 gives a short review on basics of measure theory, Lebesgue integrals, probability theory, and Shannon theory. In Section 2.1, the fundamental definitions like measure functions and probability measures are explained. Based on these definitions, the Lebesgue integrals are defined in Section 2.2. This enables us to state Lemma 2.3 (Support lemma [18]) which is required in Chapters 4 and 6. In Section 2.3, some important results which are required during this work are given. Lemma 2.6 (Blackwell et al. [42]) is used in Chapter 5 to extend the SK capacity results based on finite compound sources to infinite one. Lemma 2.7 (Extractor [18]) is used in Chapter 3 for showing the existence of an SK generator. In Section 2.4, Lemma 2.9 (continuity of mutual information [43]) is given which is required in Chapter 5 again for extension from finite to infinite case. Finally, typical sequences are reviewed and some important propositions from [18, Chapter 17] which are required in this work are given.

In Chapter 3, Section 3.1, a short review on SK sharing using non-compound sources and channels is given and some known results are presented. In Section 3.2, the model for SK generation under source uncertainty is presented. To this end, the SK generation model for [17, Theorem 2.6], [18, Theorem 17.21] is extended to the compound setup and the SK protocol for the compound source is introduced. In Section 3.3, the concept of hypothesis testing as part of SK generation protocol is explained. Finally in Sections 3.4 and 3.5, Lemma 3.2 (random coding) and Lemma 3.3 (SK generator) are given. These two lemmas are required in Chapter 4 in the proof of Theorem 4.1.

Chapter 4 gives the SK capacity results for finite compound sources. In Theorem 4.1, a single-letter lower bound for the SK capacity of a finite compound source is derived as a function of the communication rate constraint over the public channel. In Theorem 4.2, the multi-letter SK capacity formula of a finite compound source is computed again as a function of the communication rate constraint. By using Lemma 2.3 (Support lemma [18]), some upper bounds for the alphabet size of auxiliary RVs are derived in Section 4.3.

In Chapter 5, the SK capacity results for arbitrary (possibly infinite) compound sources are given. In Theorem 5.1, a single-letter SK capacity formula for a degraded compound source is derived, where the compound set may be infinite and the set of marginals is finite (cf. [31, 36]). Compared with previous theorems, this result is more practical in the sense that the SK capacity is single-letter and also valid for infinite compound sets.

In Section 5.2, the result from Theorem 4.1 is reformulated as an SK rate versus communication rate region. It is shown that such regions are convex for arbitrary (possibly infinite) compound sources. Based on this fact, the continuity of the SK rate as a function of communication rate constraint is given. From a practical point of view, this means that the variations of SK rate are not very large if the variations of communication rate constraint is small enough. Moreover in Section 5.3, the continuity is required in the proof of Theorem 5.3.

In Theorem 5.3, a single-letter lower bound for the SK capacity of an arbitrary (possibly infinite) compound source is derived as a function of the communication rate constraint. In Theorem 5.4, the multi-letter SK capacity formula of an arbitrary (possibly infinite) compound source is computed as well, as a function of the communication rate constraint.

Compared with previous results, the generalization in Theorems 5.3 and 5.4 makes them more practical in the sense that the protocol guarantees the security and reliability of the generated SK simultaneously for infinitely many statistical states that Bob and Eve might observe. This level of generality is particularly crucial in the sense that all possible statistical states of the eavesdropper (possibly infinite) are taken into account.

Extending the results to an arbitrary (possibly infinite) compound source is a non-trivial task. For this, the infinite compound source is first approximated to a sequence of finite compound sources (approximating compound sources). Afterwards, it is shown that the SK protocols which are used for the approximating sources, also guarantee the achievability of the given SK rate for the infinite compound source. To this end, in extending the results of Chapters 3 and 4 to the infinite case, the constants which are used in the definition of the SK protocol are slightly adjusted to be universal. Furthermore, choosing the size of the approximating source is a crucial step. If it is too small, then the approximation error will lead to a zero SK rate and if it is too large, then the SK generation protocol of the approximating source fails.

Finally, in Chapter 6, the compound biometric authentication models are studied and a single-letter capacity region for a finite compound source is derived where strong secrecy is also guaranteed. To this end, a special case of the model which is given in Chapter 3, is incorporated into the biometric security setup to guarantee the strong secrecy of the achievability part. Furthermore, an upper bound for the auxiliary RV is derived by using Lemma 2.3. The upper bound depends on the size of the compound set for each given marginal state.

The author's publications and preprints which are used in this dissertation and appeared in conference proceedings and journals are [44–48], [60].

2 Fundamentals

In this chapter, some fundamental definitions and results in measure, integration, and probability theory are shortly reviewed. Shannon's information quantities like mutual information and entropy are introduced. Finally, typical sequences as well as their properties are introduced. For a complete review of these concepts, readers are referred to [18, 42, 43, 49–51].

2.1 Some Definitions in Measure Theory

In the following, some fundamental definitions in measure theory are given. These notions are required to have a better understanding of probability theory and thus information theoretical concepts. Moreover, to state Lemma 2.3, some of these definitions are needed. The first definition introduces the σ-Algebra on a given set.

Definition 2.1. *A σ-Algebra on a set $\mathcal{X}$ is a set $\mathscr{A} \subset 2^{\mathcal{X}}$ with the following properties*

$$
\begin{aligned}
&i) \quad \mathcal{X} \in \mathscr{A},\\
&ii) \quad \forall \mathcal{A},\ \mathcal{A} \in \mathscr{A} \ \Rightarrow\ \mathcal{A}^{\mathrm{c}} \in \mathscr{A},\\
&iii) \quad \text{for any sequence } (\mathcal{A}_n)_{n\in\mathbb{N}} \text{ with } \mathcal{A}_n \in \mathscr{A},\\
&\qquad \text{it holds } \bigcup_{n\in\mathbb{N}} \mathcal{A}_n \in \mathscr{A}.
\end{aligned}
$$

Let $\mathscr{G} \subset 2^{\mathcal{X}}$ be given. The smallest σ-Algebra on the set $\mathcal{X}$ which is a superset of $\mathscr{G}$ is denoted by $\sigma(\mathscr{G})$ and is called the σ-Algebra generated by $\mathscr{G}$. Based on this concept, measure functions and measure spaces are defined.

Definition 2.2. *A measure μ on the σ-Algebra $\mathscr{A} \subset 2^{\mathcal{X}}$, is a mapping $\mu : \mathscr{A} \to [0, \infty]$ such that*

$$
\begin{aligned}
&i) \quad \mu(\emptyset) = 0,\\
&ii) \quad \text{for any sequence } (\mathcal{A}_n)_{n\in\mathbb{N}} \text{ with distinct elements } \mathcal{A}_n \in \mathscr{A},\\
&\qquad \text{it holds } \mu\Big(\bigcup_{n\in\mathbb{N}} \mathcal{A}_n\Big) = \sum_{n\in\mathbb{N}} \mu(\mathcal{A}_n).
\end{aligned}
$$

The 3-tuple $(\mathcal{X}, \mathscr{A}, \mu)$ is called a measure space. If $\mu(\mathcal{X}) = 1$ then $(\mathcal{X}, \mathscr{A}, \mu)$ is called a probability space and μ is called a probability measure in this space.

A measure μ is called σ-finite iff

$$\exists(\mathcal{A}_i)_{i\in\mathbb{N}} \in \mathscr{A}, \quad \bigcup_{i=1}^{\infty} \mathcal{A}_i = \mathcal{X} \quad \text{and} \quad \mu(\mathcal{A}_n) < \infty.$$

In the next definition, topological spaces are introduced. We need this concept to define the Borel σ-Algebra.

Definition 2.3. *Let $\mathscr{G} \subset 2^{\mathcal{X}}$ be given. A topological space is an ordered pair $(\mathcal{X}, \mathscr{G})$ such that it holds*

i) $\emptyset \in \mathscr{G}$ *and* $\mathcal{X} \in \mathscr{G}$,

ii) *any finite intersection of members of $\mathscr{G}$ belongs to $\mathscr{G}$,*

iii) *any arbitrary union of members of $\mathscr{G}$ belongs to $\mathscr{G}$.*

Members of the set $\mathscr{G}$ are called open sets.

Let the topological space $(\mathcal{X}, \mathscr{G})$ be given. The σ-Algebra generated by $\mathscr{G}$ is called the Borel σ-Algebra on the set $\mathcal{X}$ and is denoted by $\mathscr{B}(\mathcal{X}) := \sigma(\mathscr{G})$. The Borel σ-Algebra on the set of real numbers $\mathbb{R}$ is simply denoted by $\mathscr{B}$. In the following, some more definitions from measure theory are given:

- Let $\mathscr{A}$ and $\mathscr{A}'$ be σ-Algebras on the sets $\mathcal{X}$ and $\mathcal{X}'$ respectively. A function $f : \mathcal{X} \to \mathcal{X}'$ is called $\mathscr{A}/\mathscr{A}'$-measurable iff

  $$\forall \mathcal{A}' \in \mathscr{A}', \quad f^{-1}(\mathcal{A}') \in \mathscr{A}.$$

  A $\mathscr{A}/\mathscr{A}'$-measurable function f is also denoted as

  $$f : (\mathcal{X}, \mathscr{A}) \to (\mathcal{X}', \mathscr{A}').$$

  Real valued measurable functions $f : (\mathcal{X}, \mathscr{A}) \to (\mathbb{R}, \mathscr{B})$ are simply called $\mathscr{A}$-measurable (instead of $\mathscr{A}/\mathscr{B}$-measurable).

- Let the measure space $(\mathcal{X}, \mathscr{A}, \mu)$ and the $\mathscr{A}/\mathscr{A}'$-measurable function f be given. Measure μ' which is defined in the following, is called the image measure of μ under function f:

  $$\forall \mathcal{A}' \in \mathscr{A}', \quad \mu'(\mathcal{A}') := \mu(f^{-1}(\mathcal{A}')).$$

- Let the measure spaces $(\Omega, \mathscr{A}, \mu)$ and $(\Omega, \mathscr{A}, \nu)$ be given. Then, ν is said to be absolutely continuous with respect to μ and denoted by $\nu \ll \mu$, iff

  $$\forall \mathcal{A} \in \mathscr{A}, \quad \mu(\mathcal{A}) = 0 \Rightarrow \nu(\mathcal{A}) = 0.$$

  This notion is required to define Kullback-Leibler divergence.

- Let measure spaces

$$(\Omega_1, \mathscr{A}_1, \mu_1), (\Omega_2, \mathscr{A}_2, \mu_2), \cdots (\Omega_n, \mathscr{A}_n, \mu_n)$$

be given and $\mu_1, \mu_2, \cdots \mu_n$ be σ-finite. Define the set

$$\mathscr{G} := \Big\{ \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \mathcal{A}_n : \ \mathcal{A}_1 \in \mathscr{A}_1, \mathcal{A}_2 \in \mathscr{A}_2, \cdots, \mathcal{A}_n \in \mathscr{A}_n \Big\}.$$

The generated σ-Algebra $\sigma(\mathscr{G})$ is called the product σ-Algebra of $\mathscr{A}_1, \mathscr{A}_2, \cdots \mathscr{A}_n$ and is denoted by $\otimes_{i=1}^n \mathscr{A}_i = \mathscr{A}_1 \otimes \mathscr{A}_2, \otimes \cdots \mathscr{A}_n$.

It can be shown [49] that there exists exactly one measure μ which is σ-finite with measure space $(\times_{i=1}^n \Omega_i, \otimes_{i=1}^n \mathscr{A}_i, \mu)$ such that

$$\forall \mathcal{A}_1 \in \mathscr{A}_1, \forall \mathcal{A}_2 \in \mathscr{A}_2, \cdots \forall \mathcal{A}_n \in \mathscr{A}_n, \quad \mu(\mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \mathcal{A}_n) = \mu_1(\mathcal{A}_1)\mu_2(\mathcal{A}_2) \cdots \mu_n(\mathcal{A}_n).$$

The measure μ is called the product measure of $\mu_1, \mu_2, \cdots \mu_n$ and is denoted by

$$\otimes_{i=1}^n \mu_i = \mu_1 \otimes \mu_2 \otimes \cdots \mu_n.$$

2.2 Lebesgue Integrals and Support Lemma

Let $\mathscr{A}$ be a σ-Algebra on the set $\mathcal{X}$. For any $\mathcal{A} \in \mathscr{A}$, the indicator function $\mathbb{1}_{\mathcal{A}} : \mathcal{X} \to \{0, 1\}$ is given by

$$\mathbb{1}_{\mathcal{A}}(x) := \begin{cases} 1 & \text{if } x \in \mathcal{A}, \\ 0 & \text{otherwise.} \end{cases}$$

For finite sequences $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_n \in \mathscr{A}$ and $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{R}$, the following function $S : \mathcal{X} \to \mathbb{R}$ is called a simple function:

$$S(x) := \sum_{i=1}^{n} \alpha_i \mathbb{1}_{\mathcal{A}_i}(x), \quad \text{for all } x \in \mathcal{X}. \tag{2.1}$$

The following lemma is required to define the Lebesgue integrals.

Lemma 2.1 ([49])**.** *Let the real valued measurable function $f : (\mathcal{X}, \mathscr{A}) \to (\mathbb{R}, \mathscr{B})$ be given such that for all $x \in \mathcal{X}$, $f(x) \geq 0$. Then, there exists a sequence of simple functions $(S_n)_{n \in \mathbb{N}}$ with $S_n : \mathcal{X} \to \mathbb{R}^+$ such that for all $x \in \mathcal{X}$,*

$$\begin{aligned} &i) \quad S_1(x) \leq S_2(x) \leq \cdots \leq f(x), \\ &ii) \quad \lim_{n \to \infty} S_n(x) = f(x). \end{aligned}$$

The Lebesgue integrals are defined in the following based on the concept of simple functions. Let the measure space $(\mathcal{X}, \mathscr{A}, \mu)$ be given. The Lebesgue integral is defined in three steps:

1) The Lebesgue integral of a simple function $S : (\mathcal{X}, \mathscr{A}) \to (\mathbb{R}, \mathscr{B})$ as given in equation (2.1) is defined by

$$\int S \mathrm{d}\mu := \sum_{i=1}^{n} \alpha_i \mu(\mathcal{A}_i).$$

2) Let $f : (\mathcal{X}, \mathscr{A}) \to (\mathbb{R}, \mathscr{B})$ be given where for all $x \in \mathcal{X},\ f(x) \geq 0$. It follows by using Lemma 2.1 that there exists a sequence of simple functions $(S_n)_{n\in\mathbb{N}}$ with the given properties. The Lebesgue integral of f is defined as

$$\int f \mathrm{d}\mu := \lim_{n\to\infty} \int S_n \mathrm{d}\mu.$$

3) Finally, to define the Lebesgue integral of a function $f : (\mathcal{X}, \mathscr{A}) \to (\mathbb{R}, \mathscr{B})$ in general, define the following functions:

$$f^+(x) := \max\{f(x), 0\} \quad \text{and} \quad f^-(x) := -\min\{f(x), 0\}.$$

It holds that $f = f^+ - f^-,\ f^+(x) \geq 0,\ f^-(x) \geq 0$ and that the functions f^+ and f^- are $\mathscr{A}$-measurable. If $\int f^+ \mathrm{d}\mu < \infty$ and $\int f^- \mathrm{d}\mu < \infty$, the Lebesgue integral of f is defined by

$$\int f \mathrm{d}\mu := \int f^+ \mathrm{d}\mu - \int f^- \mathrm{d}\mu.$$

Let the measure space $(\mathcal{X}, \mathscr{A}, \mu)$ and the function $f : \mathcal{X} \to \mathbb{R}$ be given. The function f is called integrable with respect to the measure μ iff

$$\begin{aligned} &i) \quad f \text{ is } \mathscr{A} - \text{measurable}, \\ &ii) \quad \int f^+ \mathrm{d}\mu < \infty \quad \text{and} \int f^- \mathrm{d}\mu < \infty. \end{aligned}$$

The set of all such functions is denoted by $\mathcal{L}_1(\mathcal{X}, \mathscr{A}, \mu)$.

As the first practical property of Lebesgue integrals, the Fubini's lemma is given in the following which is useful in calculations.

Lemma 2.2 (Fubini [49])**.** *Let the measure spaces*

$$(\Omega_1, \mathscr{A}_1, \mu_1), (\Omega_2, \mathscr{A}_2, \mu_2), \cdots, (\Omega_n, \mathscr{A}_n, \mu_n)$$

and the function

$$f \in \mathcal{L}_1(\times_{i=1}^n \Omega_i,\ \otimes_{i=1}^n \mathscr{A}_i,\ \otimes_{i=1}^n \mu_i)$$

be given. Let also the measures $\mu_1, \mu_2, \cdots, \mu_n$ be σ-finite. Then, it holds that

$$\int f \,\mathrm{d} \otimes_{i=1}^n \mu_i = \int \int \cdots \int f \,\mathrm{d}\mu_1 \mathrm{d}\mu_2 \cdots \mathrm{d}\mu_n.$$

Furthermore, the sequence of integration can be arbitrarily changed.

Finally, in the following, the Support lemma [18, Lemma 15.4] is stated. This lemma is required in Chapters 4 and 6 to specify the upper bounds for auxiliary RVs which are used in characterization of the SK capacities.

Lemma 2.3 (Support [18])**.** *Let continuous functions $f_j : \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ with $j = 1, 2, \cdots, k$ be given. Furthermore, let μ be a probability measure defined on the Borel σ-Algebra of $\mathcal{P}(\mathcal{X})$. Then, there exist k probability measures $P_i \in \mathcal{P}(\mathcal{X})$ and constants $\alpha_i \in \mathbb{R}^+$ with $i = 1, 2, \cdots, k$ and $\sum_{i=1}^{k} \alpha_i = 1$ such that*

$$\forall j \in \{1, 2, \cdots, k\}, \qquad \int f_j \mathrm{d}\mu = \sum_{i=1}^{k} \alpha_i f_j(P_i).$$

2.3 Probability Measures and Stochastic Matrices

Probability spaces as special cases of measure spaces were defined in Section 2.1. RVs are measurable functions mapping a given underlying probability space to another probability space. Similarly as in [18], the underlying probability space in this work is denoted by $(\Omega, \mathscr{F}, \mu)$ and is assumed to be rich enough. All RVs are assumed to be measurable mappings from Ω to some finite alphabet (range).

As an example, consider an RV $X : \Omega \to \mathcal{X}$ which maps the underlying probability space $(\Omega, \mathscr{F}, \mu)$ to $(\mathcal{X}, \mathscr{A}, P_X)$ where $\mathscr{A}$ and P_X are the corresponding σ-Algebra and image measure of μ respectively. P_X is called the Probability Distribution (PD) of the RV X. Similarly as in [18], for a given $\mathcal{A} \in \mathscr{A}$, the probability Pr is defined as follows

$$\Pr(X \in \mathcal{A}) := \mu(\{\omega \in \Omega : X(\omega) \in \mathcal{A}\}) = P_X(\mathcal{A}).$$

Furthermore, for singletons of σ-Algebras e.g. $\mathcal{A} = \{x\}$, the following notations are equivalent:

$$\Pr(X = x) := \mu(\{\omega \in \Omega : X(\omega) = x\}) = P_X(x).$$

In this work, joint PDs are always denoted in terms of the corresponding multivariate RVs. For example, the joint PD of the RVs X, Y, and Z is denoted by P_{XYZ} where its marginals are P_X, P_Y, and P_Z.

Consider the underlying probability space $(\Omega, \mathscr{F}, \mu)$. The sets $\mathcal{E}_1, \mathcal{E}_2, \cdots \mathcal{E}_n \in \mathscr{F}$ (also called events) are called to be independent with respect to μ iff

$$\forall \mathcal{I} \subset \{1, 2, \cdots n\}, \quad \mu\big(\bigcap_{i \in \mathcal{I}} \mathcal{E}_i\big) = \prod_{i \in \mathcal{I}} \mu(\mathcal{E}_i).$$

Similarly, consider the probability spaces $(\mathcal{X}_i, \mathscr{A}_i, P_{X_i})$ with $i = 1, 2, \cdots n$. The RVs $X_1, X_2, \cdots X_n$ are called to be independent iff

$$\forall \mathcal{A}_i \in \mathscr{A}_i, \forall i \in \{1, 2, \cdots n\}, \quad \mu\big(\bigcap_{i=1}^{n} X_i^{-1}(\mathcal{A}_i)\big) = \prod_{i=1}^{n} \mu\big(X_i^{-1}(\mathcal{A}_i)\big).$$

Stochastic matrices $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ are functions which map each $x \in \mathcal{X}$ to a PD $W(\cdot|x) \in \mathcal{P}(\mathcal{Y})$. Furthermore, let RVs X and Y be given such that for all $x \in \mathcal{X}$, it holds $P_X(x) > 0$. The stochastic matrix $P_{Y|X} : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ is defined for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ by

$$P_{Y|X}(y|x) := \frac{P_{XY}(x,y)}{P_X(x)}.$$

Stochastic matrices W and $P_{Y|X}$ are also called discrete channels with input set $\mathcal{X}$ and output set $\mathcal{Y}$.

Let $X_1, X_2, \cdots X_n$ be a sequence of RVs with range $\mathcal{X}$ which are mutually independent and have the same PD P_X. This sequence is an independently and identically distributed (iid) sequence of RVs and is called a Discrete Memoryless Source (DMS). In this case, it holds for all $x^n \in \mathcal{X}^n$ that

$$P_{X^n}(x^n) = P_X^n(x^n) := \prod_{i=1}^{n} P_X(x_i).$$

Furthermore, an iid sequence of RVs $(X_1, Y_1, Z_1), (X_2, Y_2, Z_3), \cdots (X_n, Y_n, Z_n)$ is called a Discrete Memoryless Multiple Source (DMMS). In this case, it holds for all $(x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$ that

$$P_{X^nY^nZ^n}(x^n, y^n, z^n) = P_{XYZ}^n(x^n, y^n, z^n) := \prod_{i=1}^{n} P_{XYZ}(x_i, y_i, z_i).$$

In these definitions, DMSs or DMMSs are used (observed) n times independently and with the same PD.

Similarly, let a stochastic matrix (discrete channel) $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ be used n times independently. Define the sequence of channels $\left(W^n : \mathcal{X}^n \to \mathcal{P}(\mathcal{Y}^n)\right)_{n\in\mathbb{N}}$ as:

$$W^n(y^n|x^n) := \prod_{i=1}^{n} W(y_i|x_i),$$

where $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$. The sequence $\left(W^n : \mathcal{X}^n \to \mathcal{P}(\mathcal{Y}^n)\right)_{n\in\mathbb{N}}$ is called a Discrete Memoryless Channel (DMC).

To determine the difference between different PDs, the following quantities are introduced. The first one is the norm-1 distance. It is defined between two PDs $P, Q \in \mathcal{P}(\mathcal{X})$ as follows

$$\|P - Q\| := \sum_{x\in\mathcal{X}} |P(x) - Q(x)|.$$

The second quantity is the Kullback-Leibler divergence. By using the convention $0 \log 0 = 0 \log 0/0 = 0$, it is defined between two PDs $P, Q \in \mathcal{P}(\mathcal{X})$ as follows

$$D(P\|Q) := \begin{cases} \sum_{x\in\mathcal{X}} P(x) \log \frac{P(x)}{Q(x)} & \text{if} \quad P \ll Q, \\ \infty & \text{otherwise.} \end{cases}$$

It can be shown [18] that the Kullback-Leibler divergence is always non-negative $D(P\|Q) \geq 0$. In contrary to norm-1 distance, the Kullback-Leibler divergence is not commutative.

In the following, two practical inequalities are given. The first lemma is required in the calculations. The second one gives a better understanding of the relation between the norm-1 distance and Kullback-Leibler divergence.

Lemma 2.4 (Markov inequality [50]). *Let X be an RV and the function $f:[0,\infty) \to [0,\infty)$ be monotonically increasing. Then, it holds for all $\epsilon > 0$ with $f(\epsilon) > 0$ that*

$$\Pr(|X| \geq \epsilon) \;\leq\; \frac{\mathbb{E}\big[f(|X|)\big]}{f(\epsilon)}.$$

Lemma 2.5 (Pinsker inequality [18]). *Let $p, q \in \mathcal{P}(\mathcal{A})$ be given. Then, it holds*

$$\frac{1}{2\ln 2} \parallel p - q \parallel^2 \leq D(p \parallel q).$$

Next, a useful lemma [42, Lemma 4] is given in the following for approximating an infinite class of stochastic matrices by a finite one. This lemma is used in Chapter 5 when the SK capacity results for infinite compound sources are derived.

Lemma 2.6 (Blackwell [42]). *Let $\mathcal{S}$ be an arbitrary set and possibly infinite and $\{W_s : \mathcal{X} \to \mathcal{P}(\mathcal{U})\}_{s\in\mathcal{S}}$ be a family of stochastic matrices. Then, for every $l \in \mathbb{N}$ with $l \geq 2|\mathcal{U}|^2$, there exists a family of stochastic matrices $\{W_{s'} : \mathcal{X} \to \mathcal{P}(\mathcal{U})\}_{s'\in\mathcal{S}'}$ with a finite set $\mathcal{S}'$, such that $|\mathcal{S}'| \leq (l+1)^{|\mathcal{X}\times\mathcal{U}|}$, where the following properties hold:*

$$\begin{aligned} \forall s \in \mathcal{S},\ \exists s' \in \mathcal{S}', \forall x \in \mathcal{X}, \forall u \in \mathcal{U},& \\ |W_s(u|x) - W_{s'}(u|x)| &\leq \frac{1}{l}|\mathcal{U}|, \\ W_s(u|x) &\leq \mathrm{e}^{2|\mathcal{U}|^2/l} W_{s'}(u|x). \end{aligned}$$

Next lemma [18, Lemma 17.3] provides a tool for proving the existence of an SK generator. This lemma is used in Chapter 3 to show that the generated SK guarantees all security criteria.

Lemma 2.7 (Extractor [13, 18]). *Let $\epsilon, \eta, \lambda \in \mathbb{R}^{++}$ and $k \in \mathbb{N}$ be given and U be an RV taking its values in $\mathcal{U}$. If $P_U(\{u \in \mathcal{U} : P_U(u) \leq \frac{1}{\lambda}\}) \geq 1 - \eta$, then for a randomly selected function $\kappa : \mathcal{U} \to \{1, 2, \cdots, k\}$, it holds that*

$$\Pr\Big(\big\|\kappa(P_U) - P_0\big\| > \epsilon + 2\eta\Big) \leq 2k\mathrm{e}^{-\frac{\lambda\epsilon^2(1-\eta)}{2k(1+\epsilon)}},$$

where $P_0(i) = 1/k$ for all $i = 1, 2, \cdots k$. Random selection means that the $\kappa(u), u \in \mathcal{U}$ are chosen iid uniformly.

2.4 Shannon's Information Quantities

The Shannon's information quantities such as entropy and mutual information are defined in the following. Again in all definitions which follow, we use the convention that $0 \log 0 = 0 \log 0/0 = 0$. For RV X with range $\mathcal{X}$, the entropy is defined as

$$H(X) := -\sum_{x \in \mathcal{X}} P_X(x) \log P_X(x).$$

If the alphabet is binary e.g. $\mathcal{X} = \{0, 1\}$ with $P_X(0) = p$ and $P_X(1) = 1 - p$, the entropy is given as a function of p and is called the binary entropy function. The binary entropy function is denoted by $h(p)$, where

$$h(p) := -p \log p - (1 - p) \log(1 - p).$$

Next, for RVs X and Y with ranges $\mathcal{X}$ and $\mathcal{Y}$ respectively, the conditional entropy is defined by

$$H(Y|X = x) := -\sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \log P_{Y|X}(y|x),$$
$$H(Y|X) := \sum_{x \in \mathcal{X}} P_X(x) H(Y|X = x).$$

The mutual information is defined as

$$I(X;Y) := \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_{XY}(x, y) \log \frac{P_{XY}(x, y)}{P_X(x) P_Y(y)}.$$

By definition, mutual information can be expressed as Kullback-Leibler divergence between P_{XY} and $P_X \otimes P_Y$:

$$I(X;Y) = D(P_{XY} \| P_X \otimes P_Y).$$

By this interpretation, mutual information is a quantity which measures the dependency between RVs X and Y. If X and Y are independent, then $I(X;Y) = 0$.

Similarly, for RVs X, Y, and Z with ranges $\mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$ respectively, the conditional mutual information is defined as

$$I(X;Y|Z = z) := \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_{XY|Z}(x, y|z) \log \frac{P_{XY|Z}(x, y|z)}{P_{X|Z}(x|z) P_{Y|Z}(y|z)},$$
$$I(X;Y|Z) := \sum_{z \in \mathcal{Z}} P_Z(z) I(X;Y|Z = z).$$

Remark. *The following relations hold*

$$H(X, Y) = H(X) + H(Y|X) \text{ (Chain Rule)},$$
$$I(X;Y, Z) = I(X;Y) + I(X;Z|Y) \text{ (Chain Rule)},$$
$$I(X;Y|Z) = H(X|Z) - H(X|Y, Z),$$
$$I(X;Y) \geq 0, \quad H(Y) \geq H(Y|X), \quad H(X) = I(X;X) \geq 0.$$

An upper bound for conditional entropy is a useful tool which is required in Chapters 4 and 6 for the proof of converse. The Fano's inequality gives such an upper bound.

Lemma 2.8 (Fano's inequality [18]). *Let RVs X and $\hat{X}$ with range $\mathcal{X}$ be given. Let also $p_{\mathrm{e}} := Pr\{X \neq \hat{X}\}$ give the error probability. Then, it holds that*

$$H(X|\hat{X}) \leq h(p_{\mathrm{e}}) + p_{\mathrm{e}} \log(|\mathcal{X}| - 1).$$

Finally, in the following lemma, the continuity of mutual information as a function of the joint PD P_{XY} is given. In this kind of continuity which is also a uniform continuity [52], large variations of $I(X;Y)$ is to be excluded if the joint PD P_{XY} only variates minimally.

Lemma 2.9 ([43]). *Let (X,Y) and (X',Y') be two pairs of RVs taking values in $\mathcal{X} \times \mathcal{Y}$ with PDs P_{XY} and $P_{X'Y'}$ respectively. Furthermore, let $\gamma := \frac{1}{2}\|P_{XY} - P_{X'Y'}\|$ and $\gamma \leq 1 - \frac{1}{|\mathcal{X}\times\mathcal{Y}|}$. Then, it holds that*

$$|I(X;Y) - I(X';Y')| \leq 3\gamma \log(|\mathcal{X} \times \mathcal{Y}| - 1) + 3h(\gamma).$$

This lemma is an alternative to Alicki-Fannes inequality in quantum regime [53].

2.5 Typical Sequences

For the typical sequences and their related sets, the same definitions as in [18, Chapters 2 and 17] are taken. Let $N(x|x^n)$ give the number of repetitions of an element x in the sequence x^n and $N(x,y|x^n,y^n)$ the number of repetitions of the pair (x,y) in the pair sequence (x^n,y^n). For RV X and $\epsilon > 0$, the set of typical sequences is given by

$$\begin{aligned}\mathcal{T}^n_{[X]_\epsilon} := \Big\{ & x^n \in \mathcal{X}^n : \forall x \in \mathcal{X}, \\ & \big|P_X(x) - \frac{1}{n}N(x|x^n)\big| \leq \epsilon \quad \wedge \quad \big(P_X(x) = 0 \Rightarrow N(x|x^n) = 0\big)\Big\}.\end{aligned}$$

Furthermore, for two RVs X and Y, stochastic matrix $P_{Y|X} : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, and $\epsilon > 0$, define

$$\begin{aligned}\mathcal{T}^n_{[XY]_\epsilon} := \Big\{ & (x^n,y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : \quad \forall (x,y) \in \mathcal{X} \times \mathcal{Y}, \\ & \big|P_{XY}(x,y) - \frac{1}{n}N(x,y|x^n,y^n)\big| \leq \epsilon \\ & \wedge \quad \big(P_{XY}(x,y) = 0 \Rightarrow N(x,y|x^n,y^n) = 0\big)\Big\}, \\ \mathcal{T}^n_{[Y|X]_\epsilon}(x^n) := \Big\{ & y^n \in \mathcal{Y}^n : \quad \forall (x,y) \in \mathcal{X} \times \mathcal{Y}, \\ & \big|\frac{1}{n}N(x|x^n)P_{Y|X}(y|x) - \frac{1}{n}N(x,y|x^n,y^n)\big| \leq \epsilon \\ & \wedge \quad \big(P_{Y|X}(y|x) = 0 \Rightarrow N(x,y|x^n,y^n) = 0\big)\Big\}, \\ \mathcal{T}^n_{[XY]_\epsilon}(x^n) := \Big\{ & y^n \in \mathcal{Y}^n : \quad (x^n,y^n) \in \mathcal{T}^n_{[XY]_\epsilon}\Big\}.\end{aligned}$$

In the following, some propositions from [18, Lemma 17.8 (Typicality)] are given which will be used through out this work in the proofs.

Lemma 2.10 ([18]). *Let U, X, and Y be RVs taking their values in $\mathcal{U}, \mathcal{X}$, and $\mathcal{Y}$ respectively. Assume $0 < \xi < \zeta < \sigma$, and $\tau > 0$ are all in $\mathbb{R}$ and $n \in \mathbb{N}$. Then, it holds that*

1. $\forall x^n \in \mathcal{T}^n_{[X]\xi}, \quad \mathcal{T}^n_{[XY]\zeta}(x^n) \supset \mathcal{T}^n_{[Y|X]\zeta-\xi}(x^n)$.
2. $P^n_X(\mathcal{T}^n_{[X]\xi}) \geq 1 - 2|\mathcal{X}|\mathrm{e}^{-2\xi^2 n}$.
3. *If $u^n \in \mathcal{T}^n_{[U]\xi}$ then* $P^n_{X|U}(\mathcal{T}^n_{[UX]\zeta}(u^n)|u^n) \geq 1 - 2|\mathcal{U}||\mathcal{X}|\mathrm{e}^{-2(\zeta-\xi)^2 n}$.
4. $\forall \tau > 0$, $\forall \xi > 0$ *sufficiently small, and* $\forall x^n \in \mathcal{T}^n_{[X]\xi}$, *it holds*
$$\big| -\frac{1}{n} \log P^n_X(x^n) - H(X) \big| < \tau.$$
5. $\forall \tau > 0$, $\forall \xi > 0$ *sufficiently small, and* $\forall (u^n, x^n) \in \mathcal{T}^n_{[UX]\xi}$, *it holds*
$$\big| -\frac{1}{n} \log P^n_{X|U}(x^n|u^n) - H(X|U) \big| < \tau.$$
6. $\forall \tau > 0$, $\forall \xi > 0$ *sufficiently small, and* $\forall n \in \mathbb{N}$ *sufficiently large, it holds*
$$\big| \frac{1}{n} \log |\mathcal{T}^n_{[X]\xi}| - H(X) \big| < \tau.$$
7. $\forall \tau > 0$, $\forall \zeta > 0$ *sufficiently small,* $\forall n \in \mathbb{N}$ *sufficiently large, and* $\forall u^n$ *with* $\mathcal{T}^n_{[UX]\zeta}(u^n) \neq \emptyset$ *it holds*
$$\big| \frac{1}{n} \log |\mathcal{T}^n_{[UX]\zeta}(u^n)| - H(X|U) \big| < \tau.$$
8. *For Markov chain $U - X - Y$ and* $\forall (u^n, x^n) \in \mathcal{T}^n_{[UX]\xi}$ *it holds*
$$P^n_{Y|X}\big(\mathcal{T}^n_{[UXY]\sigma}(u^n, x^n)|x^n\big) \geq 1 - 2|\mathcal{U}||\mathcal{X}||\mathcal{Y}|\mathrm{e}^{-2(\sigma-\xi)^2 n}.$$
9. $\forall \tau > 0$, $\forall \zeta > 0$ *sufficiently small,* $\forall n \in \mathbb{N}$ *sufficiently large, and* $\forall y^n \in \mathcal{Y}^n$, *if* $\mathcal{T}^n_{[XY]\zeta}(y^n) \neq \emptyset$, *then it holds*
$$\Big| -\frac{1}{n} \log P^n_X\big(\mathcal{T}^n_{[XY]\zeta}(y^n)\big) - I(X;Y) \Big| < \tau.$$
10. $\forall \tau > 0$, $\forall \sigma > 0$ *sufficiently small,* $\forall n \in \mathbb{N}$ *sufficiently large, and* $\forall (u^n, y^n) \in \mathcal{U}^n \times \mathcal{Y}^n$, *if* $\mathcal{T}^n_{[UXY]\sigma}(u^n, y^n) \neq \emptyset$, *then it holds*
$$\big| -\frac{1}{n} \log P^n_{X|U}\big(\mathcal{T}^n_{[UXY]\sigma}(u^n, y^n)|u^n\big) - I(X;Y|U) \big| < \tau.$$

Remark. *Some of the propositions which are stated in 2.10, [18, Lemma 17.8 (Typicality)] do not specify directly the exact size of constants $0 < \xi < \zeta < \sigma$, and $\tau > 0$ or the length of sequences $n \in \mathbb{N}$. The inequalities are mostly stated for $0 < \xi < \zeta < \sigma$ sufficiently small and/or for n large enough. In Chapters 3 and 4, the SK generation protocol is designed by taking these values into account. The size of these constants are particularly essential, when the results are extended in Chapter 5 to the case where the compound source is infinite.*

In the following, the size of constant $\tau > 0$ is given as a function of ξ, ζ, σ and/or sequence length n together with other constant quantities.

a) In Lemma 2.10.4:

$$\tau = \frac{\xi}{\mu_X} H(X), \quad \text{where} \quad \mu_X = \min_{x \in \mathrm{supp}(P_X)} P_X(x). \tag{2.2}$$

b) In Lemma 2.10.5:

$$\tau = \frac{\xi}{\mu_{UX}} H(X|U), \quad \text{where} \quad \mu_{UX} = \min_{(u,x) \in \mathrm{supp}(P_{UX})} P_{UX}(u,x). \tag{2.3}$$

c) In Lemma 2.10.6:

$$\tau = \frac{|\mathcal{X}|}{n} \log(n+1) - |\mathcal{X}| \xi \log \xi. \tag{2.4}$$

d) In Lemma 2.10.7:

$$\tau = \frac{|\mathcal{U}||\mathcal{X}|}{n} \log(n+1) - |\mathcal{U}||\mathcal{X}| \zeta \log \frac{4\zeta^2}{|\mathcal{X}|}. \tag{2.5}$$

e) In Lemma 2.10.9:

$$\tau = \frac{\zeta |\mathcal{Y}|}{\mu_X} H(X) + \frac{|\mathcal{X}||\mathcal{Y}|}{n} \log(n+1) - |\mathcal{X}||\mathcal{Y}| \zeta \log \frac{4\zeta^2}{|\mathcal{X}|}, \quad \text{where} \quad \mu_X = \min_{x \in \mathrm{supp}(P_X)} P_X(x). \tag{2.6}$$

f) In Lemma 2.10.10:

$$\tau = \frac{\sigma |\mathcal{Y}|}{\mu_{UX}} H(X|U) + \frac{|\mathcal{U}||\mathcal{X}||\mathcal{Y}|}{n} \log(n+1) - |\mathcal{U}||\mathcal{X}||\mathcal{Y}| \sigma \log \frac{4\sigma^2}{|\mathcal{X}|}, \quad \text{where} \quad \mu_{UX} = \min_{(u,x) \in \mathrm{supp}(P_{UX})} P_{UX}(u,x). \tag{2.7}$$

Finally, the next two lemmas from [18, Corollaries 17.9A and 17.9B] are required to show that the encoders in SK protocol which is given in Chapter 3 exist.

Lemma 2.11 ([18])**.** *Let RVs U and X take their values in $\mathcal{U}$ and $\mathcal{X}$ respectively. Consider $N = \exp(nR)$ sequences u_l^n with $l \in \mathcal{L} := \{1, 2 \cdots N\}$, which are independently drawn by a PD P_U^n, such that it holds*

$$I(U;X) < R.$$

Then for all $\tau \in (0, R - I(U;X))$, all $\zeta > 0$ sufficiently small, all $n \in \mathbb{N}$ sufficiently large, and all x^n such that $\mathcal{T}^n_{[UX]\zeta}(x^n) \neq \emptyset$, it holds that

$$\left| \frac{1}{n} \log \left| \left\{ l \in \mathcal{L} : u_l^n \in \mathcal{T}^n_{[UX]\zeta}(x^n) \right\} \right| - \big(R - I(U;X)\big) \right| < \tau,$$

with a probability approaching one, doubly exponentially fast.

Lemma 2.12 ([18])**.** *Let RVs U, V, and X take their values in $\mathcal{U}, \mathcal{V}$, and $\mathcal{X}$ respectively. Consider $N = \exp(nR)$ sequences v_l^n with $l \in \mathcal{L} := \{1, 2 \cdots N\}$, which are conditionally independently drawn by a PD $P^n_{V|U}(.|u^n)$ for a given $u^n \in \mathcal{U}^n$, such that it holds*

$$I(V;X|U) < R.$$

Then for all $\tau \in (0, R - I(V;X|U))$, all $\sigma > 0$ sufficiently small, all $n \in \mathbb{N}$ sufficiently large, and all x^n such that $\mathcal{T}^n_{[UVX]\sigma}(u^n, x^n) \neq \emptyset$, it holds that

$$\left| \frac{1}{n} \log \left| \left\{ l \in \mathcal{L} : v_l^n \in \mathcal{T}^n_{[UVX]\sigma}(u^n, x^n) \right\} \right| - \big(R - I(V;X|U)\big) \right| < \tau,$$

with a probability approaching one, doubly exponentially fast. This holds uniformly for all given $u^n \in \mathcal{U}^n$.

Similarly as in Lemma 2.10, the constant τ and sequence length n are not explicitly specified in Lemmas 2.11 and 2.12 from [18, Corollaries 17.9A and 17.9B]. However, by using equations (2.6) and (2.7) in the proofs of these two lemmas, the constant τ is computable as a function of ζ, σ, the sequence length n, and other constant quantities as follows.

g) In Lemma 2.11:

$$\tau = \frac{\zeta|\mathcal{X}|}{\mu_U} H(U) + \frac{|\mathcal{U}||\mathcal{X}|}{n} \log(n+1) - |\mathcal{U}||\mathcal{X}|\zeta \log \frac{4\zeta^2}{|\mathcal{U}|}, \tag{2.8}$$
$$\text{where} \quad \mu_U = \min_{u \in \text{supp}(P_U)} P_U(u).$$

h) In Lemma 2.12:

$$\tau = \frac{\sigma|\mathcal{X}|}{\mu_{UV}} H(V|U) + \frac{|\mathcal{U}||\mathcal{V}||\mathcal{X}|}{n} \log(n+1) - |\mathcal{U}||\mathcal{V}||\mathcal{X}|\sigma \log \frac{4\sigma^2}{|\mathcal{V}|}, \tag{2.9}$$

$$\text{where} \quad \mu_{UV} = \min_{(u,v)\in \text{supp}(P_{UV})} P_{UV}(u,v).$$

The size of constant τ is again required in Chapters 3 and 4 in the protocol design and also in Chapter 5 for extending the results to the case where the compound set is infinite.

3 Secret-Key Generation and Compound Sources

In this chapter, a mathematical model for SK generation using compound sources is presented. The hypothesis testing is further applied to this SK model for the case when the set of marginals is finite. Next, a coding scheme is introduced to generate the CR between Alice and Bob. Finally, the existence of an SK generator is shown which works for all elements of the compound source.

3.1 Secret-Key Generation with Perfect Knowledge of Source or Channel

In this section, a very short review on SK generation using non-compound sources (or channels) is given [18, Section 17.3], [17]. In this case the source (or channel) is perfectly known to all participants. Two main approaches are source and channel models.

In the source model, Alice, Bob, and Eve observe a DMMS $XYZ \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ for a time duration n so that their knowledge is given by X^n, Y^n, and Z^n respectively. Alice and Bob generate an SK based on their knowledge, using randomization, and communication over a public noiseless channel.

In the channel model, Alice selects the inputs represented by RV X. Based on a DMC $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y} \times \mathcal{Z})$, Bob observes the Y-output and Eve observes the Z-output. Similarly as in source model, Alice and Bob generate an SK based on their knowledge, randomization, and communication over a public noiseless channel.

In both models (source and channel), Eve can observe all public communications between Alice and Bob, however she is not allowed to make any changes in the transmitted information.

There are in general two kinds of SK generation protocols for both channel and source models. In the first approach, the public communication between Alice and Bob is unrestricted and might be interactive (multi-way). In this case both Alice and Bob are allowed to transmit messages over the public noiseless channel in several rounds. In the second approach, only forward (one-way) communication initiated by Alice is allowed. More details and exact definitions of SK protocols, rate achievability and SK capacity in each case are available in [18, Section 17.3].

In the following two bounds for source and channel models are given where public communication is unrestricted [18, Theorem 17.17]. In this case Alice and Bob might communicate interactively as many times as they wish.

Theorem 3.1 ([18]). *The SK capacity of an unrestricted source model with DMMS XYZ is bounded by:*

$$I(X;Y) - \min\big\{I(X;Z), I(Y;Z)\big\} \leq C_{\mathrm{sk}} \leq I(X;Y|Z). \tag{3.1}$$

For an unrestricted channel model with DMC $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y} \times \mathcal{Y})$ the bounds for the SK capacity are given as follow:

$$\max\Big\{I(X;Y) - \min\big\{I(X;Z), I(Y;Z)\big\}\Big\} \leq C_{\mathrm{sk}} \leq \max\Big\{I(X;Y|Z)\Big\},$$

where the max *is taken over all RVs X, Y, and Z such that $P_{YZ|X} = W$.*

The following result [18, Theorem 17.21], [17] is given for a source model based on a forward public communication initiated by Alice.

Theorem 3.2 ([18]). *Consider the SK source model with DMMS XYZ and a one-way communication over a public noiseless channel with communication rate upper bound $\Gamma \in (0, \infty]$. Then, it holds that*

$$C_{\mathrm{sk}}(\Gamma) = \max_{U,V} \big\{I(V;Y|U) - I(V;Z|U)\big\},$$

where the max *is taken over all RVs U and V such that it holds:*

$$U - V - X - YZ \qquad \text{and} \qquad I(V;X|Y) \leq \Gamma.$$

Furthermore, it may be assumed that $V = (U, V')$ where both U and V' have alphabet size of at most $|\mathcal{X}| + 1$

In this dissertation only source models with forward public communication are considered. We will extend the result from Theorem 3.2 to compound sources in Chapters 4 and 5. In this case, the actual statistics of the source is unknown to the participants.

3.2 Mathematical Model with Compound Sources

In this section, the SK generation model based on a compound source is introduced. To this end, the non-compound source model with forward public communication from [12, 17, 18] which was shortly introduced in Section 3.1 is extended to the compound setup.

Let an arbitrary set (possibly infinite) of 3-party DMMSs be given by a sequence of generic RVs $\mathfrak{S} := \{XYZ, s\}_{s \in \mathcal{S}}$ where $\mathcal{S}$ is the set of source indices. Let also all sources in $\mathfrak{S}$ take their values in the finite set $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. The set of all marginal RVs $X_{\hat{s}}$ corresponding to $\mathfrak{S}$ whose PDs for some $s \in \mathcal{S}$ are given by

$$P_{X_{\hat{s}}}(\cdot) = \sum_{y \in \mathcal{Y}, z \in \mathcal{Z}} P_{XYZ,s}(\cdot, y, z),$$

is denoted by $\hat{\mathfrak{S}} := \{X_{\hat{s}}\}_{\hat{s} \in \hat{\mathcal{S}}}$. The sets $\hat{\mathfrak{S}}$ and $\hat{\mathcal{S}}$ are called the set of marginals and marginal indices respectively and are assumed to be always finite in this work.

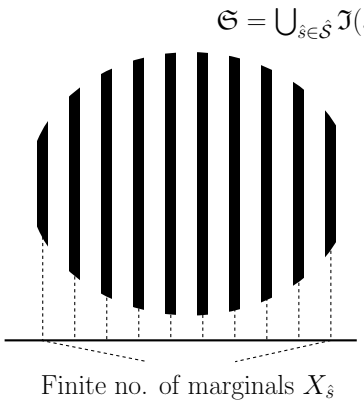

Figure 3.1: Informal illustration of an example of a set of 3-party discrete memoryless multiple sources

Similarly as in [29], for each $X_{\hat{s}} \in \hat{\mathfrak{S}}$, the set of all DMMSs $XYZ, s \in \mathfrak{S}$ whose marginal is $X_{\hat{s}}$, is denoted by $\mathfrak{I}(\hat{s}) = \{XYZ, s\}_{s\in\mathcal{I}(\hat{s})}$ and is given by

$$\mathfrak{I}(\hat{s}) := \Big\{XYZ, s \in \mathfrak{S} :\ \forall x \in \mathcal{X},\ \sum_{y\in\mathcal{Y}}\sum_{z\in\mathcal{Z}} P_{XYZ,s}(x,y,z) = P_{X_{\hat{s}}}(x)\Big\}. \tag{3.2}$$

The set $\mathcal{I}(\hat{s})$ is the index set of $\mathfrak{I}(\hat{s})$. It holds by definition that for all $\hat{s} \in \hat{\mathcal{S}}$,

$$\mathfrak{I}(\hat{s}) \subset \mathfrak{S}.$$

Figure 3.1 depicts an informal illustration of an example of such an arbitrary (possibly infinite) set $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ of 3-party DMMSs with a finite set $\hat{\mathfrak{S}} := \{X_{\hat{s}}\}_{\hat{s}\in\hat{\mathcal{S}}}$ of marginals. The horizontal axis represents the marginal RVs $X_{\hat{s}} \in \hat{\mathfrak{S}}$. The vertical axis represents the corresponding DMMSs XYZ, s for each marginal $X_{\hat{s}}$. Therefore, each bar in figure 3.1 represents a set $\mathfrak{I}(\hat{s}) \subset \mathfrak{S}$. The set $\mathfrak{S}$ can be described as the union of all $\mathfrak{I}(\hat{s})$.

Next, we use this set $\mathfrak{S}$ of 3-party DMMSs to define the compound sources and SK generation procedure. Figure 3.2 shows the SK generation model which is used throughout this work. Transmitter (Alice), receiver (Bob) and eavesdropper (Eve) observe a DMMS which belongs to the set $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ for time duration $n \in \mathbb{N}$. Therefore, RVs $X_{\hat{s}}^n, Y_s^n$ and Z_s^n represent their initial knowledge respectively for the marginal index $\hat{s} \in \hat{\mathcal{S}}$ and source index $s \in \mathcal{I}(\hat{s})$.

It is assumed that all terminals know the set $\mathfrak{S}$ as well as its statistics i.e. PDs $\{P_{XYZ,s}\}_{s\in\mathcal{S}}$. However, they do not know which element of $\mathfrak{S}$ is the actual realization. They only know that the actual PD of the source belongs to the set $\{P_{XYZ,s}\}_{s\in\mathcal{S}}$. This set $\mathfrak{S}$ of sources with the mentioned properties is called a compound source.

Furthermore, as part of the model, a noiseless public channel between all participants is given. Communication over such channels is noiseless in the sense that the transmitted

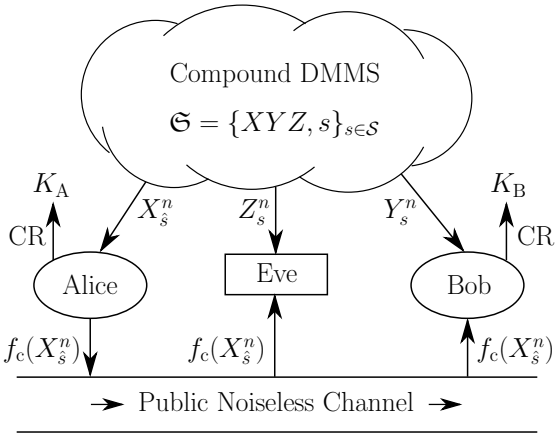

Figure 3.2: Secret-key generation protocol for compound DMMS model

and received messages are identical. Furthermore, as the channel is public, Eve receives the transmitted messages too. In this work, only a one-way communication over the public noiseless channel, initiated by Alice, is allowed. Furthermore, similarly as in [17] and [18, Section 17.3], the communication rate is upper bounded by some parameter $\Gamma > 0$ and in all SK generation protocols in this model, the public communication rate should be lower than Γ. To this end, the coding scheme which is given in Section 3.4 works with respect to this parameter. This problem for the case where no communication constraint is given i.e. $\Gamma = \infty$ is a special case of this model and is easier to solve. A non-compound version is stated in [18, Problem 17.15b].

As RVs $X_{\hat{s}}^n$ and Y_s^n are correlated, Alice and Bob may generate some CR by using their initial knowledge (i.e. $X_{\hat{s}}^n$ and Y_s^n respectively) and communicating over this noiseless public channel. The message which is transmitted over the public channel is a function of Alice initial knowledge and is denoted in Figure 3.2 by $f_{\mathrm{c}}(X_{\hat{s}}^n)$. Eventually, after generating the CR, Alice and Bob have the goal to generate a shared SK $K_{\mathrm{A}} = K_{\mathrm{B}}$ based on the CR while keeping it secret from Eve. The following definition gives a more precise description of this procedure.

Definition 3.1. *A one-way SK generation protocol with compound source set* $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ *consists of the following two steps:*

- *After observing* $X_{\hat{s}}^n$*, Alice transmits a message* $f_{\mathrm{c}}(X_{\hat{s}}^n)$ *to Bob over the public noiseless channel.* f_{c} *is called public communication function.*[1]

- *Next, Alice generates an SK, represented by an RV* K_{A}*, based on her knowledge* $X_{\hat{s}}^n$ *and Bob generates an SK, represented by an RV* K_{B}*, based on his knowledge* $(Y_s^n, f_{\mathrm{c}}(X_{\hat{s}}^n))$*.* K_{A} *and* K_{B} *take their values in* $\mathcal{K}$*.*

[1] Similarly as in [18, Problem 17.15(a)], it can be shown that a randomized f_{c} in the one-way SK generation protocol does not increase the SK capacity. Therefore, the communication function f_{c} is assumed to be a deterministic function of $X_{\hat{s}}^n$ and no randomization is considered here.

As the communication over the public channel is overheard by Eve, this should not reveal any information about the SK (secrecy leakage). Moreover, the generated SK should have a uniform distribution. Combining these two criteria together leads to a compact notion, called security index, which was first introduced in [20] and is given in the following.

Definition 3.2. *For RVs K_{A} and V, taking values in the sets $\mathcal{K}$ and $\mathcal{V}$ respectively, the security index is given by*

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

In the context of Definition 3.2, K_{A} represents the SK and V Eve's knowledge. This short notion is a powerful tool which can be used to describe both strong secrecy [14] and the uniformity of the generated SK. The next definition, uses this concept to define an achievable SK rate and capacity of a compound source.

Definition 3.3. *A number $R_{\mathrm{sk}} \in \mathbb{R}^+$ is an achievable SK rate using the compound source $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$ and a one-way communication over the public noiseless channel with rate constraint $\Gamma \in (0, +\infty]$, iff for all $\delta > 0$, there exists an $n_0 > 0$ such that for all $n \in \mathbb{N}, n \geq n_0$, there exists an 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 for all $s \in \mathcal{S}$:*

$$\frac{1}{n}\log\|f_{\mathrm{c}}\| < \Gamma + \delta, \tag{3.3}$$

$$R_{\mathrm{sk}} < \frac{1}{n}\log|\mathcal{K}| + \delta, \tag{3.4}$$

$$\Pr(K_{\mathrm{A}} \neq K_{\mathrm{B}}) < \delta, \tag{3.5}$$

$$S(K_{\mathrm{A}}|Z_s^n, f_{\mathrm{c}}(X_{\hat{s}}^n)) < \delta. \tag{3.6}$$

The SK capacity $C_{\mathrm{sk}}(\mathfrak{S}, \Gamma)$ is defined to be the supremum of all achievable SK rates. If there is no communication rate constraint, i.e. $\Gamma = \infty$, then condition (3.3) *in the definition is inactive and the capacity is denoted by $C_{\mathrm{sk}}(\mathfrak{S})$.*

Remark. *Similarly as in [17, 18], the communication rate constraint is also part of Definition 3.3. This is because, in a realistic model where the communication cost is an important parameter, the information exchange rate between the terminals is restricted.*

Furthermore, Condition (3.6) *of Definition 3.3 guarantees the strong secrecy. This is because, in Definition 3.2, the whole leaked information to eavesdropper i.e. $I(K_{\mathrm{A}};V)$ is considered. Meanwhile, in models with weak secrecy, only the rate of leaked information i.e. $n^{-1}I(K_{\mathrm{A}};V)$ is taken into account.*

In the next sections of this chapter, we introduce the hypothesis testing and give some technical lemmas for random coding and SK generation. The SK capacity results are given in Chapters 4 and 5. In Theorem 4.1, in Chapter 4, a lower bound for SK capacity of a finite compound source is given as a function of the communication rate constraint. In the proof of that theorem, the techniques and lemmas from Sections 3.3, 3.4, and 3.5 are required.

3.3 Hypothesis Testing and Marginal Estimation

As mentioned in Section 3.2, in a compound source model, Alice, Bob, and Eve know that the actual source belongs to a known set $\mathfrak{S}$. As given in Definition 3.3, any SK generation protocol, which is to achieve a given SK rate $R > 0$, should work for all elements of the compound source $\mathfrak{S}$. In this section, we incorporate the hypothesis testing technique to this setup for the case when the set of marginals $\hat{\mathfrak{S}}$ is finite.

Let $X_{\hat{s}} \in \hat{\mathfrak{S}}$ be the actual marginal RV of the compound source and all other $X_{\tilde{s}} \in \hat{\mathfrak{S}} - \{X_{\hat{s}}\}$ be the wrong marginals. Alice estimates the marginal PD $P_{X_{\hat{s}}}$ by observing a sample x^n and using hypothesis testing. Similarly as in [29], she transmits then the marginal index $\hat{s}$ along with other information related to her observation over the public channel to Bob. By doing this, Alice and Bob would know the actual PD of the marginal RV with a probability exponentially close to one [54]. They may then simplify the design of the encoder and decoder of the SK generation protocol. In this case, it is sufficient that the coding scheme works only for the estimated marginal RV $X_{\hat{s}}$ instead of all elements of the set of marginals $\hat{\mathfrak{S}}$.

By knowing the actual marginal source $X_{\hat{s}} \in \hat{\mathfrak{S}}$, the actual source $XYZ_s \in \mathfrak{S}$ itself is not necessarily known to the terminals. However, by definition in (3.2), it is known that $XYZ_s \in \mathfrak{I}(\hat{s}) \subset \mathfrak{S}$. Therefore, it is sufficient that the coding scheme works only for all $XYZ_s \in \mathfrak{I}(\hat{s})$ instead of all elements of the bigger set $\mathfrak{S}$.

The following lemma from [18, Problem 2.13(b)] shows how to define a hypothesis test and distinguish the actual marginal $X_{\hat{s}} \in \hat{\mathfrak{S}}$ from other marginals $X_{\tilde{s}} \in \hat{\mathfrak{S}} - \{X_{\hat{s}}\}$.

Lemma 3.1 ([18])**.** *Let $\epsilon > 0$ and RV $X_{\hat{s}} \in \hat{\mathfrak{S}}$ with range $\mathcal{X}$ be given. Let also $X_{\tilde{s}}$ be any RV in $\hat{\mathfrak{S}} - \{X_{\hat{s}}\}$ with the same range $\mathcal{X}$. Then, there exists a sequence of sets $(\mathcal{A}_n)_{n\in\mathbb{N}} \subset \mathcal{X}^n$ such that*

$$\lim_{n\to\infty} \frac{1}{n} \log\left(1 - P^n_{X_{\hat{s}}}(\mathcal{A}_n)\right) \leq -\epsilon,$$
$$\frac{1}{n} \log P^n_{X_{\tilde{s}}}(\mathcal{A}_n) \leq - \min_{P: D(P\|P_{X_{\hat{s}}}) \leq \epsilon} D(P\|P_{X_{\tilde{s}}}).$$

Lemma 2.5 gives the relationship between Kullback-Leibler divergence D and norm-1 distance between PDs.

Let RV $\hat{S}$ with range $\hat{\mathcal{S}}$ (cf. Section 3.2, set of marginal indices) represent the outcome of the hypothesis test. Based on the test which is given in Lemma 3.1, the following bounds for PD $P_{\hat{S}}$ are given:

$$P_{\hat{S}}(\hat{s}) \geq 1 - \exp(-nc_0), \tag{3.7}$$
$$\forall \tilde{s} \in \hat{\mathcal{S}} - \{\hat{s}\},\ P_{\hat{S}}(\tilde{s}) \leq \exp(-nc_1), \tag{3.8}$$

where $c_0, c_1 > 0$ are constants.

3.4 Random Coding Scheme

As seen in Section 3.3, the actual marginal source PD $P_{X_{\hat{s}}}$ can be estimated by Alice with a probability close to one. Therefore, in the following lemma the marginal source $X_{\hat{s}}$ is assumed to be fixed. The encoder and decoder schemes which are introduced should work simultaneously for all elements of the compound set $\mathfrak{I}(\hat{s})$. In this section, the compound set $\mathfrak{I}(\hat{s})$ is assumed to be finite. This coding scheme is required in the proof of Theorem 4.1, in Chapter 4.

Similar techniques which are used in the non-compound versions in [18, Lemma 17.22], [17], are used in the proofs and extended to the compound setup. For completeness, all proofs in detail are presented. Assume in Lemma 3.2, if the values in the equations (3.9), (3.10), (3.11), and (3.12) are not integer numbers, then the smallest integer which is larger than the given expression is taken.

Lemma 3.2. *Let $\delta > 0$ be given. Let Alice's marginal index $\hat{s} \in \hat{\mathcal{S}}$ be given and RVs $X_{\hat{s}}$ and Y_s with $s \in \mathcal{I}(\hat{s})$, take their values in $\mathcal{X}$ and $\mathcal{Y}$ respectively. Let also RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ with ranges $\mathcal{U}_{\hat{s}}$ and $\mathcal{V}_{\hat{s}}$ respectively be given such that for all $s \in \mathcal{I}(\hat{s})$ the Markov chains $U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}} - Y_s$ hold.*

Assume $N_{\hat{s},1}N_{\hat{s},2}$ random sequences $u_{ij}^n(\hat{s}) \in \mathcal{U}_{\hat{s}}^n$, chosen independently according to PD $P_{U_{\hat{s}}}^n$, are given and known to Alice and Bob where

$$
\begin{aligned}
i \in \mathcal{I} &:= \big\{1, 2, \cdots, N_{\hat{s},1}\big\},\\
j \in \mathcal{J} &:= \big\{1, 2, \cdots, N_{\hat{s},2}\big\},\\
N_{\hat{s},1} &:= \exp\Big[n\big(\max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + 3\delta\big)\Big], \qquad (3.9)\\
N_{\hat{s},2} &:= \exp\Big[n\big(\min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) - 2\delta\big)\Big]. \qquad (3.10)
\end{aligned}
$$

Moreover, let for each $u_{ij}^n(\hat{s})$, $N_{\hat{s},3}N_{\hat{s},4}$ random sequences $v_{pq}^{ij\,n}(\hat{s}) \in \mathcal{V}_{\hat{s}}^n$, chosen conditionally independently according to $P_{V_{\hat{s}}|U_{\hat{s}}}^n(\cdot|u_{ij}^n(\hat{s}))$, be given and known to Alice and Bob where

$$
\begin{aligned}
p \in \mathcal{P} &:= \big\{1, 2, \cdots, N_{\hat{s},3}\big\},\\
q \in \mathcal{Q} &:= \big\{1, 2, \cdots, N_{\hat{s},4}\big\},\\
N_{\hat{s},3} &:= \exp\Big[n\big(\max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_s) + 3\delta\big)\Big], \qquad (3.11)\\
N_{\hat{s},4} &:= \exp\Big[n\big(\min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s|U_{\hat{s}}) - 2\delta\big)\Big]. \qquad (3.12)
\end{aligned}
$$

Then CR can be generated between Alice and Bob in two ways:

a) For all $\zeta > 0$ small enough and $n \in \mathbb{N}$ sufficiently large, there exist encoder functions $f : \mathcal{T} \to \mathcal{I}$ and $g : \mathcal{T} \to \mathcal{J}$, with a probability approaching 1 doubly exponentially fast, such that

$$
\text{if} \quad f(x^n) = i,\ g(x^n) = j \quad \text{then} \quad (u_{ij}^n(\hat{s}), x^n) \in \mathcal{T}_{[UX,\hat{s}]\zeta}^n, \qquad (3.13)
$$

and the set $\mathcal{T}$ is given by

$$\mathcal{T} := \Big\{ x^n \in \mathcal{X}^n : \mathcal{T}^n_{[UX,\hat{s}]\zeta}(x^n) \neq \emptyset \Big\}. \tag{3.14}$$

Alice encodes her observation $x^n \in \mathcal{T}$ by these functions to the sequence $u^n_{ij}(\hat{s})$ where j is the CR.

Next, extend the domain of functions f and g to the set $\mathcal{X}^n$ by defining for all $x^n \notin \mathcal{T}$, $f(x^n) = g(x^n) = 0$. Then, for all $\sigma > \zeta$ small enough and $n \in \mathbb{N}$ sufficiently large, there exists a decoder function $\tilde{g} : \mathcal{I} \times \hat{\mathcal{S}} \times \mathcal{Y}^n \to \mathcal{J}$ (depending on σ) such that for all $s \in \mathcal{I}(\hat{s})$,

$$\Pr\Big\{ g(X^n_{\hat{s}}) \neq \tilde{g}\big(f(X^n_{\hat{s}}), \hat{s}, Y^n_s\big) \Big\} < \exp(-n\delta_0), \tag{3.15}$$

for some $\delta_0 > 0$. The constant δ_0 can be made larger by making σ smaller and n larger. On the other hand, δ_0 becomes smaller by making δ smaller. Thus, Bob can reconstruct $g(x^n) = j$ from $(f(x^n), \hat{s}, y^n)$ by using this decoder for given realizations $X^n_{\hat{s}} = x^n$ and $Y^n_s = y^n$ with the error probability in (3.15)*.*

b) Let the encoder functions f and g from part a) be given such that $f(x^n) = i$ and $g(x^n) = j$. Then, for all $\sigma > 0$ small enough and $n \in \mathbb{N}$ sufficiently large, there exist encoder functions $\varphi : \mathcal{T} \to \mathcal{P}$ and $\rho : \mathcal{T} \to \mathcal{Q}$ with a probability approaching 1 doubly exponentially fast, such that

$$\text{if} \quad \varphi(x^n) = p,\ \rho(x^n) = q \quad \text{then} \quad (u^n_{ij}(\hat{s}), v^{ij\,n}_{pq}(\hat{s}), x^n) \in \mathcal{T}^n_{[UVX,\hat{s}]\sigma}. \tag{3.16}$$

Alice encodes her observation $x^n \in \mathcal{T}$ by these functions to the sequence $v^{ij\,n}_{pq}(\hat{s})$ where q is the CR.

Next, extend the domain of functions φ and ρ to the set $\mathcal{X}^n$ by defining for all $x^n \notin \mathcal{T}$, $\varphi(x^n) = \rho(x^n) = 0$. Then, for all $\vartheta > \sigma$ small enough and $n \in \mathbb{N}$ sufficiently large, there exists a decoder function $\tilde{\rho} : \mathcal{I} \times \mathcal{J} \times \mathcal{P} \times \hat{\mathcal{S}} \times \mathcal{Y}^n \to \mathcal{Q}$ (depending on ϑ) such that for all $s \in \mathcal{I}(\hat{s})$,

$$\Pr\Big\{ \rho(X^n_{\hat{s}}) \neq \tilde{\rho}\big(f(X^n_{\hat{s}}), g(X^n_{\hat{s}}), \varphi(X^n_{\hat{s}}), \hat{s}, Y^n_s\big) \Big\} < \exp(-n\delta_0'), \tag{3.17}$$

for some $\delta_0' > 0$. The constant δ_0' can be made larger by making ϑ smaller and n larger. On the other hand, δ_0' becomes smaller by making δ smaller. Thus, Bob can reconstruct $\rho(x^n) = q$ from $(f(x^n), g(x^n), \varphi(x^n), \hat{s}, y^n)$ by using this decoder for given realizations $X^n_{\hat{s}} = x^n$ and $Y^n_s = y^n$ with error probability in (3.17)*.*

Proof. a) Let R be the rate of choosing the sequences $\{u^n_{ij}(\hat{s})\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$ which implies that $N_{\hat{s},1}N_{\hat{s},2} = \exp(nR)$. Therefore, by (3.9) and (3.10) and properties of the Markov chain, it follows that

$$\begin{aligned} R &= \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) + \delta \\ &= I(U_{\hat{s}}; X_{\hat{s}}) + \max_{s\in\mathcal{I}(\hat{s})} [I(U_{\hat{s}}; Y_s|X_{\hat{s}}) - I(U_{\hat{s}}; Y_s)] + \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) + \delta \\ &= I(U_{\hat{s}}; X_{\hat{s}}) + \delta. \end{aligned} \tag{3.18}$$

Similarly as in [18, Lemma 17.22], for all $x^n \in \mathcal{T}$, it holds by using the definition in (3.14) that $\mathcal{T}^n_{[UX,\hat{s}]\zeta}(x^n) \neq \emptyset$. Thus, Lemma 2.11 together with (3.18) implies for all $\tau \in (0, R - I(U_{\hat{s}}; X_{\hat{s}}))$, all $\zeta > 0$ sufficiently small, all $n \in \mathbb{N}$ sufficiently large, and all $x^n \in \mathcal{T}$ that

$$\frac{1}{n} \log \Big| \Big\{ (i,j) \in \mathcal{I} \times \mathcal{J} : u^n_{ij}(\hat{s}) \in \mathcal{T}^n_{[UX,\hat{s}]\zeta}(x^n) \Big\} \Big| \geq R - I(U_{\hat{s}}; X_{\hat{s}}) - \tau,$$

with a probability approaching one, doubly exponentially fast. The size of τ as a function of ζ, n, and other constant quantities is given by (2.8), where τ can be made arbitrarily small, if ζ is made small enough and n sufficiently large. The constants are universal for all sets $\mathcal{I}(\hat{s})$. Therefore, for each $x^n \in \mathcal{T}$, the number of chosen sequences $u^n_{ij}(\hat{s})$ which are in $\mathcal{T}^n_{[UX,\hat{s}]\zeta}(x^n)$ is non-zero. As a result, the encoder functions f and g as mentioned in (3.13) do exist with $f(x^n) = i$ and $g(x^n) = j$.

Next, Let $\sigma > \zeta$ be given. Define for all $i \in \mathcal{I}$ and $y^n \in \mathcal{Y}^n$, the decoder function as follows:

$$\tilde{g}(i, \hat{s}, y^n) := \begin{cases} j & \text{if } j \in \mathcal{J},\ u^n_{ij}(\hat{s}) \in \bigcup\limits_{s \in \mathcal{I}(\hat{s})} \mathcal{T}^n_{[UY,s]\sigma|\mathcal{X}|}(y^n) \\ & \quad \text{and } \ \forall m \in \mathcal{J}, m \neq j \Rightarrow u^n_{im}(\hat{s}) \notin \bigcup\limits_{s \in \mathcal{I}(\hat{s})} \mathcal{T}^n_{[UY,s]\sigma|\mathcal{X}|}(y^n), \\ 0 & \text{otherwise.} \end{cases} \tag{3.19}$$

Define the set

$$\mathcal{T}_0 := \Big\{ (x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : \quad x^n \in \mathcal{T} \ \wedge \ (u^n_{ij}(\hat{s}), x^n, y^n) \in \bigcup_{s \in \mathcal{I}(\hat{s})} \mathcal{T}^n_{[UXY,s]\sigma} \Big\}.$$

In the following, it is shown that Alice and Bob's observation $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ is in the set $\mathcal{T}_0$ with a probability exponentially close to one. It holds that

$$\begin{aligned} P^n_{XY,s}(\mathcal{T}_0{}^{\mathrm{c}}) &= \sum_{x^n \in \mathcal{T}^{\mathrm{c}}, y^n \in \mathcal{Y}^n} P^n_{XY,s}(x^n, y^n) + \sum_{x^n \in \mathcal{T}, (x^n,y^n) \notin \mathcal{T}_0} P^n_{XY,s}(x^n, y^n) \\ &= P^n_{X_{\hat{s}}}(\mathcal{T}^{\mathrm{c}}) + \sum_{x^n \in \mathcal{T}} P^n_{X_{\hat{s}}}(x^n)\ P^n_{Y_s|X_{\hat{s}}}\Big(\bigcap_{s \in \mathcal{I}(\hat{s})} \mathcal{T}^n_{[UXY,s]\sigma}{}^{\mathrm{c}}(u^n_{ij}(\hat{s}), x^n) \big| x^n \Big). \end{aligned} \tag{3.20}$$

Furthermore, Lemma 2.10.1 implies for every $x^n \in \mathcal{T}^n_{[X_{\hat{s}}]\xi}$ with constant $\xi \in (0, \zeta)$ that $\mathcal{T}^n_{[UX,\hat{s}]\zeta}(x^n) \neq \emptyset$. Therefore, by using the definition in (3.14), it follows that $\mathcal{T} \supset \mathcal{T}^n_{[X_{\hat{s}}]\xi}$ and thus Lemma 2.10.2 implies that

$$P^n_{X_{\hat{s}}}(\mathcal{T}^{\mathrm{c}}) < 2|\mathcal{X}| \mathrm{e}^{-2\xi^2 n}. \tag{3.21}$$

On the other hand, it holds that $(u^n_{ij}(\hat{s}), x^n) \in \mathcal{T}^n_{[UX,\hat{s}]\zeta}$. Since for all $s \in \mathcal{I}(\hat{s})$, the Markov chains $U_{\hat{s}} - X_{\hat{s}} - Y_s$ hold, Lemma 2.10.8 implies that

$$P^n_{Y_s|X_{\hat{s}}}\big(\mathcal{T}^n_{[UXY,s]\sigma}{}^{\mathrm{c}}(u^n_{ij}(\hat{s}), x^n) | x^n\big) < 2|\mathcal{U}_{\hat{s}}||\mathcal{X}||\mathcal{Y}| \mathrm{e}^{-2(\sigma-\zeta)^2 n}.$$

This inequality together with (3.20) and (3.21) gives

$$P_{XY,s}^n(\mathcal{T}_0^{\mathrm{c}}) < 2|\mathcal{X}|\mathrm{e}^{-2\xi^2 n} + 2|\mathcal{U}_{\hat{s}}||\mathcal{X}||\mathcal{Y}|\mathrm{e}^{-2(\sigma-\zeta)^2 n}, \tag{3.22}$$

where the upper bound goes to zero exponentially fast as n goes to infinity. The constants are universal for all sets $\mathcal{I}(\hat{s})$.

Therefore, to compute the upper bound of the probability in (3.15), we may concentrate just on all $(x^n, y^n) \in \mathcal{T}_0$ with

$$\tilde{g}(i, \hat{s}, y^n) \neq g(x^n). \tag{3.23}$$

It holds for all $(x^n, y^n) \in \mathcal{T}_0$ that $x^n \in \mathcal{T}$. Thus, for $f(x^n) = i$ and $g(x^n) = j$, it follows by using (3.13) that

$$(u_{ij}^n(\hat{s}), x^n) \in \mathcal{T}_{[UX,\hat{s}]\zeta}^n. \tag{3.24}$$

Furthermore, a necessary condition for $(x^n, y^n) \in \mathcal{T}_0$ is given by

$$(u_{ij}^n(\hat{s}), y^n) \in \bigcup_{s \in \mathcal{I}(\hat{s})} \mathcal{T}_{[UY,s]\sigma|\mathcal{X}|}^n,$$

which together with (3.23) and (3.19) implies that

$$\exists m \neq j,\ (u_{im}^n(\hat{s}), y^n) \in \bigcup_{s \in \mathcal{I}(\hat{s})} \mathcal{T}_{[UY,s]\sigma|\mathcal{X}|}^n. \tag{3.25}$$

Define the RV

$$\tilde{U}_{\hat{s}} := \{U_{ij,\hat{s}}^n\}_{(i,j)\in\mathcal{I}\times\mathcal{J}} = (U_{11,\hat{s}}^n, U_{12,\hat{s}}^n, \cdots, U_{|\mathcal{I}||\mathcal{J}|,\hat{s}}^n)$$

and let $\tilde{u}(\hat{s}) := \{u_{ij}^n(\hat{s})\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$ be an arbitrary realization. Define for all $(x^n, y^n) \in \mathcal{T}_0$

$$e\big(\tilde{u}(\hat{s})\big) := \sum_{\substack{i\in\mathcal{I}, j\in\mathcal{J} \\ m\in\mathcal{J}-\{j\}}} \mathbb{1}_{\mathcal{T}_{[UX,\hat{s}]\zeta}^n}(u_{ij}^n(\hat{s}), x^n)\mathbb{1}_{\bigcup_{s\in\mathcal{I}(\hat{s})} \mathcal{T}_{[UY,s]\sigma|\mathcal{X}|}^n}(u_{im}^n(\hat{s}), y^n). \tag{3.26}$$

The relations (3.24), (3.25), and (3.26) give the following upper bound for the error probability in (3.15) when $s \in \mathcal{I}(\hat{s})$:

$$\sum_{(x^n,y^n)\in\mathcal{T}_0} P_{XY,s}^n(x^n, y^n) e\big(\tilde{u}(\hat{s})\big). \tag{3.27}$$

To find an upper bound of $e(\tilde{u}(\hat{s}))$ for a given $(x^n, y^n) \in \mathcal{T}_0$, its expectation is calculated in the following with respect to $\tilde{U}_{\hat{s}}$:

$$\mathbb{E}_{\tilde{U}_{\hat{s}}}\Big[e(\tilde{U}_{\hat{s}})\Big] = \sum_{\substack{i\in\mathcal{I}, j\in\mathcal{J} \\ m\in\mathcal{J}-\{j\}}} \mathbb{E}_{(U_{ij,\hat{s}}^n, U_{im,\hat{s}}^n)}\Big[\mathbb{1}_{\mathcal{T}_{[UX,\hat{s}]\zeta}^n}(U_{ij,\hat{s}}^n, x^n)\mathbb{1}_{\bigcup_{s\in\mathcal{I}(\hat{s})} \mathcal{T}_{[UY,s]\sigma|\mathcal{X}|}^n}(U_{im,\hat{s}}^n, y^n)\Big]. \tag{3.28}$$

The equality (3.28) follows by using Lemma 2.2 and the fact that RVs $\{U^n_{ij,\hat{s}}\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$ are independent. In the following, an upper bound for (3.28) is derived, which automatically gives a universal upper bound for the error upper bound in (3.27) for all $s \in \mathcal{I}(\hat{s})$.

Let $i \in \mathcal{I}, j \in \mathcal{J}, m \in \mathcal{J} - \{j\}$ and $\tau, \tau' > 0$ be given such that $\delta > \tau + \tau'$. It holds for all $(x^n, y^n) \in \mathcal{T}_0$, all $\sigma > \zeta$ sufficiently small, and n sufficiently large that

$$\begin{aligned}
&\mathbb{E}_{(U^n_{ij,\hat{s}}, U^n_{im,\hat{s}})}\Big[\mathbb{1}_{\mathcal{T}^n_{[UX,\hat{s}]\zeta}}(U^n_{ij,\hat{s}}, x^n)\mathbb{1}_{\bigcup\limits_{s\in\mathcal{I}(\hat{s})}\mathcal{T}^n_{[UY,s]\sigma|\mathcal{X}|}}(U^n_{im,\hat{s}}, y^n)\Big] \\
&\qquad = \Pr\Big(U^n_{ij,\hat{s}} \in \mathcal{T}^n_{[UX,\hat{s}]\zeta}(x^n)\Big)\Pr\Big(U^n_{im,\hat{s}} \in \bigcup_{s\in\mathcal{I}(\hat{s})}\mathcal{T}^n_{[UY,s]\sigma|\mathcal{X}|}(y^n)\Big) \\
&\qquad \le |\mathcal{I}(\hat{s})|\exp\Big[-n\big(I(U_{\hat{s}};X_{\hat{s}}) + \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Y_s) - \tau - \tau'\big)\Big], \qquad (3.29)
\end{aligned}$$

where the equality follows again by the fact that RVs $\{U^n_{ij,\hat{s}}\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$ are independent. The inequality is a result of Lemma 2.10.9. The size of $\tau + \tau'$ as a function of ζ, σ, n, and other constant quantities is given by using (2.6). Based on (2.6), $\tau + \tau'$ can be made arbitrarily small, if σ is made small enough and n sufficiently large. The constants are universal for all sets $\mathcal{I}(\hat{s})$.
Moreover, the definitions in (3.9) and (3.10) imply that

$$\begin{aligned}
N_{\hat{s},1}N_{\hat{s},2}(N_{\hat{s},2} - 1) &\le N_{\hat{s},1}N_{\hat{s},2}N_{\hat{s},2} \\
&= \exp\Big[\, n\big(I(U_{\hat{s}};X_{\hat{s}}) + \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Y_s) - \delta\big)\,\Big]. \qquad (3.30)
\end{aligned}$$

By using (3.28), (3.29) and (3.30), it follows that

$$\mathbb{E}_{\tilde{U}_{\hat{s}}}\Big[e(\tilde{U}_{\hat{s}})\Big] \le |\mathcal{I}(\hat{s})|\exp\big[-n(\delta - \tau - \tau')\big]. \qquad (3.31)$$

Let δ_1 be sufficiently small such that $0 < \delta_1 < \delta - \tau - \tau'$. Lemma 2.4 (Markov inequality) implies that

$$\Pr\Big(e(\tilde{U}_{\hat{s}}) \ge \exp(-n\delta_1)\Big) \le \frac{\mathbb{E}_{\tilde{U}_{\hat{s}}}\Big[e(\tilde{U}_{\hat{s}})\Big]}{\exp(-n\delta_1)}. \qquad (3.32)$$

Therefore, (3.31) and (3.32) imply for all ζ and σ sufficiently small, and all n large enough that

$$\Pr\Big(e(\tilde{U}_{\hat{s}}) < \exp(-n\delta_1)\Big) \ge 1 - |\mathcal{I}(\hat{s})|\exp(-n(\delta - \tau - \tau' - \delta_1)). \qquad (3.33)$$

The inequality (3.33) implies that there exists a realization $\tilde{u}(\hat{s}) = \{u^n_{ij}(\hat{s})\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$ of the RV $\tilde{U}_{\hat{s}}$, such that for all $s \in \mathcal{I}(\hat{s})$ and $(x^n, y^n) \in \mathcal{T}_0$, the upper bound of the decoding error probability in (3.27) is given by $\exp(-n\delta_1)$.

This implies by using (3.22) that the total error probability in (3.15) is upper bounded by $\exp(-n\delta_0)$ with some $\delta_0 > 0$ such that $\delta_0 < \delta_1 < \delta - \tau - \tau'$. The size of $\tau + \tau'$ is given

by using (2.6). Thus, the constant δ_0 can be made larger by making σ smaller and n larger. On the other hand, δ_0 becomes smaller by making δ smaller.

b) The proof of the second part is very similar to the first part. For a given $(i,j) \in \mathcal{I} \times \mathcal{J}$ with $f(x^n) = i$ and $g(x^n) = j$, the number of chosen random sequences $v_{pq}^{ij\,n}$ is $N_{\hat{s},3} N_{\hat{s},4} = \exp(nR')$, where R' is the rate of choosing the random sequences. Similar to (3.18), conditions (3.11) and (3.12) imply that $R' = I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}) + \delta$. Furthermore, as a result of Lemma 2.10.1, from $(u_{ij}^n(\hat{s}), x^n) \in \mathcal{T}^n_{[UX,\hat{s}]\zeta}$ follows $\mathcal{T}^n_{[UVX,\hat{s}]\sigma}(u_{ij}^n(\hat{s}), x^n) \neq \emptyset$. Therefore, Lemma 2.12 implies for $\tau \in (0, R - I(V;X|U))$, σ small enough and n sufficiently large that the encoder functions φ and ρ as mentioned in (3.16) do exist with $\phi(x^n) = p$, and $\rho(x^n) = q$. The size of τ as a function of σ, n, and other constant quantities is given by (2.9), where τ can be made arbitrarily small, if σ is made small enough and n sufficiently large. Similarly as in part a), the constants are universal for all sets $\mathcal{I}(\hat{s})$.

According to part a) of this lemma, Bob is able to reconstruct $g(x^n) = j$, by knowing $f(x^n), \hat{s}$ and y^n. Therefore, he knows also $u_{ij}^n(\hat{s})$. Let $\vartheta \in \mathbb{R}$ be given such that $\vartheta > \sigma$. Similar to (3.19), the decoder is defined as

$$\tilde{\rho}(i,j,p,\hat{s},y^n) := \begin{cases} q & \text{if } q \in \mathcal{Q}, \quad v_{pq}^{ij\,n}(\hat{s}) \in \bigcup\limits_{s \in \mathcal{I}(\hat{s})} \mathcal{T}^n_{[UVY,s]\vartheta|\mathcal{X}|}(u_{ij}^n(\hat{s}), y^n) \\ & \quad \text{and} \quad \forall r \in \mathcal{Q}, r \neq q \Rightarrow v_{pr}^{ij\,n}(\hat{s}) \notin \bigcup\limits_{s \in \mathcal{I}(\hat{s})} \mathcal{T}^n_{[UVY,s]\vartheta|\mathcal{X}|}(u_{ij}^n(\hat{s}), y^n), \\ 0 & \text{otherwise.} \end{cases} \tag{3.34}$$

Define the set $\mathcal{T}_0'$ as follows:

$$\mathcal{T}_0' := \Big\{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : \quad x^n \in \mathcal{T} \wedge (u_{ij}^n(\hat{s}), v_{pq}^{ij\,n}(\hat{s}), x^n, y^n) \in \bigcup_{s \in \mathcal{I}(\hat{s})} \mathcal{T}^n_{[UVXY,s]\vartheta}\Big\}.$$

Similar to (3.22), the probability $P^n_{XY,s}(\mathcal{T}_0'^{\mathrm{c}})$ goes to zero exponentially fast as n goes to infinity. Therefore, to find the error upper bound, we may concentrate just on all $(x^n, y^n) \in \mathcal{T}_0'$ with

$$\tilde{\rho}(i,j,p,\hat{s},y^n) \neq \rho(x^n). \tag{3.35}$$

For a given $(i,j) \in \mathcal{I} \times \mathcal{J}$ define the RV

$$\tilde{V}_{ij,\hat{s}} := \{V_{pq,\hat{s}}^{ij\,n}\}_{(p,q) \in \mathcal{P} \times \mathcal{Q}} ,$$

and let $\tilde{v}_{ij}(\hat{s}) := \{v_{pq}^{ij\,n}(\hat{s})\}_{(p,q) \in \mathcal{P} \times \mathcal{Q}}$ be an arbitrary realization. Define for $(x^n, y^n) \in \mathcal{T}_0'$

$$e\big(\tilde{v}_{ij}(\hat{s})\big) := \sum_{\substack{p \in \mathcal{P},\, q \in \mathcal{Q} \\ r \in \mathcal{Q} - \{q\}}} \mathbb{1}_{\bigcup\limits_{s \in \mathcal{I}(\hat{s})} \mathcal{T}^n_{[UVXY,s]\vartheta}}(u_{ij}^n(\hat{s}), v_{pq}^{ij\,n}(\hat{s}), x^n, y^n) \cdot \mathbb{1}_{\bigcup\limits_{s \in \mathcal{I}(\hat{s})} \mathcal{T}^n_{[UVY,s]\vartheta|\mathcal{X}|}}(u_{ij}^n(\hat{s}), v_{pr}^{ij\,n}(\hat{s}), y^n). \tag{3.36}$$

Similarly as in part a), it can be shown by using (3.34) and (3.35) that the error probability in (3.17) is upper bounded for all $(x^n, y^n) \in \mathcal{T}_0'$ by

$$\sum_{i\in\mathcal{I},j\in\mathcal{J}} \sum_{\substack{(x^n,y^n)\in\mathcal{T}_0' \\ f(x^n)=i,\, g(x^n)=j}} P_{XY,s}^n(x^n,y^n)\, e\big(\tilde{v}_{ij}(\hat{s})\big). \tag{3.37}$$

For $0 < \tau + \tau' < \delta$, Lemma 2.10.10 implies that

$$\begin{aligned}\mathbb{E}_{(V_{pq,\hat{s}}^{ij\,n}, V_{pr,\hat{s}}^{ij\,n})}\Big[&\mathbb{1}_{\bigcup_{s\in\mathcal{I}(\hat{s})} \mathcal{T}^n_{[UVXY,s]\vartheta}}(u_{ij}^n(\hat{s}), V_{pq,\hat{s}}^{ij\,n}, x^n, y^n) \\ &\mathbb{1}_{\bigcup_{s\in\mathcal{I}(\hat{s})} \mathcal{T}^n_{[UVY,s]\vartheta|\mathcal{X}|}}(u_{ij}^n(\hat{s}), V_{pr,\hat{s}}^{ij\,n}, y^n) \;\Big|\; U_{ij,\hat{s}}^n = u_{ij}^n(\hat{s})\Big] \\ &\le \big|\mathcal{I}(\hat{s})\big|^2 \cdot \exp\Big[-n\big(I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}) + \min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s|U_{\hat{s}}) - \tau - \tau' \big)\Big]. \end{aligned} \tag{3.38}$$

The size of $\tau + \tau'$ as a function of ϑ, n, and other constant quantities is given by using (2.7). Based on (2.7), $\tau + \tau'$ can be made arbitrarily small, if ϑ is made small enough and n sufficiently large. Similarly as in part a), the constants are universal for all sets $\mathcal{I}(\hat{s})$. Moreover, by using (3.11) and (3.12), it follows that

$$\begin{aligned} N_{\hat{s},3}N_{\hat{s},4}(N_{\hat{s},4}-1) &\le N_{\hat{s},3}N_{\hat{s},4}N_{\hat{s},4} \\ &= \exp\Big[\, n\big(I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}) + \min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s|U_{\hat{s}}) - \delta\big)\,\Big]. \end{aligned} \tag{3.39}$$

Thus, (3.36), (3.38) and (3.39) imply that

$$\mathbb{E}_{\tilde{V}_{ij,\hat{s}}}\Big[e(\tilde{V}_{ij,\hat{s}})\big|U_{ij,\hat{s}}^n = u_{ij}^n(\hat{s})\Big] \le \big|\mathcal{I}(\hat{s})\big|^2 \exp\Big[-n(\delta - \tau - \tau')\Big].$$

Therefore, it follows by using Lemma 2.4 that for all $0 < \delta_1' < \delta - \tau - \tau'$,

$$\begin{aligned}\Pr\Big(e(\tilde{V}_{ij,\hat{s}}) < \exp(-n\delta_1') \;\Big|\; U_{ij,\hat{s}}^n = u_{ij}^n(\hat{s})\Big) \\ \ge 1 - \big|\mathcal{I}(\hat{s})\big|^2 \exp\big(-n(\delta - \tau - \tau' - \delta_1')\big).\end{aligned}$$

This implies for a given (i,j), that there exists a realization $\tilde{v}_{ij}(\hat{s})$ of the RV $\tilde{V}_{ij,\hat{s}}$, such that for all $s \in \mathcal{I}(\hat{s})$ and $(x^n, y^n) \in \mathcal{T}_0'$, the upper bound of the decoding error in (3.37) is given by $\exp(-n\delta_1')$.

The probability $P_{XY,s}^n(\mathcal{T}_0'^c)$ goes to zero exponentially fast as n goes to infinity. Therefore, the total decoding error probability in (3.17) is upper bounded by $\exp(-n\delta_0')$ with some $\delta_0' > 0$ such that $\delta_0' < \delta_1' < \delta - \tau - \tau'$. The size of $\tau + \tau'$ is given by using (2.7). Thus, the constant δ_0' can be made larger by making ϑ smaller and n larger. On the other hand, δ_0' becomes smaller by making δ smaller. □

3.5 Secret-Key Generator

Similarly as in Section 3.4, the actual marginal source $X_{\hat{s}}$ is taken fixed in the following lemma. It is shown that there exists an SK generator function κ which maps the generated CR between Alice and Bob (cf. Section 3.4) to an SK. The SK should guarantee the security criteria which is given in (3.6), simultaneously for all elements of the compound set $\mathfrak{I}(\hat{s})$. In this section, the compound set $\mathfrak{I}(\hat{s})$ is assumed to be finite. This lemma is required in the proof of Theorem 4.1, in Chapter 4.

Similar techniques which are used in the non-compound version in [18, Lemma 17.5], [55], are used in the proofs and extended to the compound setup. For completeness, all proofs in detail are presented.

Lemma 3.3. *Let Alice's marginal index $\hat{s} \in \hat{\mathcal{S}}$ be given. Let RVs C, D_s with $s \in \mathcal{I}(\hat{s})$, and $\hat{S}$ take their values in $\mathcal{C}, \mathcal{D}$, and $\hat{\mathcal{S}}$ respectively. RVs C and D_s denote the CR and part of Eve's knowledge respectively. RV $\hat{S}$ denotes the result of the hypothesis test by Alice. Assume that $\alpha \in (0, \frac{1}{6}]$ 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*

$$\forall (c,d) \in \mathcal{B}_s,\ P_{CD,s|\hat{S}}(c,d|\hat{s}) < (\alpha|\mathcal{B}_s|)^{-1}, \tag{3.40}$$

$$P_{CD,s|\hat{S}}(\mathcal{B}_s|\hat{s}) \geq 1 - (\eta^2 - \alpha^2). \tag{3.41}$$

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 let $k \in \mathbb{N}$ be given such that

$$k < \min\Big\{ \alpha^6 \min_{s \in \mathcal{I}(\hat{s}), d \in \mathcal{D}_s} |\mathcal{B}_{s,d}|, \quad \mathrm{e}^{1/\alpha} (2|\mathcal{D}|\,|\mathcal{I}(\hat{s})|)^{-1} \Big\}. \tag{3.42}$$

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

$$S\big(\kappa(C)|D_s, \hat{S} = \hat{s}\big) \leq (\alpha + 2\eta) \log k + h(\alpha + \eta), \tag{3.43}$$

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 $d \in \mathcal{D}_s$.*

Proof. Let $s \in \mathcal{I}(\hat{s})$ be given. Define:

$$\lambda := \alpha^3 \min_{s \in \mathcal{I}(\hat{s}), d \in \mathcal{D}_s} |\mathcal{B}_{s,d}|, \tag{3.44}$$

$$\mathcal{D}'_s := \Big\{ d \in \mathcal{D} : P_{D_s|\hat{S}}(d|\hat{s}) \geq \frac{\alpha^2 |\mathcal{B}_{s,d}|}{|\mathcal{B}_s|} \Big\}, \tag{3.45}$$

$$\mathcal{B}'_s := \mathcal{B}_s \cap (\mathcal{C} \times \mathcal{D}'_s), \tag{3.46}$$

$$\mathcal{G}_s := \Big\{ (c,d) \in \mathcal{C} \times \mathcal{D} : P_{C|D\hat{S},s}(c|d,\hat{s}) \leq \frac{1}{\lambda} \Big\}, \tag{3.47}$$

$$\mathcal{G}_{s,d} := \Big\{ c \in \mathcal{C} : (c,d) \in \mathcal{G}_s \Big\}, \tag{3.48}$$

$$\mathcal{E}_s := \Big\{ d \in \mathcal{D} : P_{C|D\hat{S},s}(\mathcal{G}_{s,d}|d,\hat{s}) < 1 - \eta \Big\}. \tag{3.49}$$

Similarly as in [18, Lemma 17.5], we show in the first step that the following inequality is true:

$$P_{D_s|\hat{S}}(\mathcal{E}_s|\hat{s}) < \eta. \tag{3.50}$$

This inequality is required later to show that (3.43) holds. For this, let $(c,d) \in \mathcal{B}'_s$ be given and $s \in \mathcal{I}(\hat{s})$. It follows that

$$\begin{aligned} P_{C|D\hat{S},s}(c|d,\hat{s}) &= \frac{P_{CD,s|\hat{S}}(c,d|\hat{s})}{P_{D_s|\hat{S}}(d|\hat{s})} \\ &\leq \frac{(\alpha|\mathcal{B}_s|)^{-1}}{\alpha^2|\mathcal{B}_{s,d}||\mathcal{B}_s|^{-1}} \\ &\leq \frac{1}{\lambda}, \end{aligned}$$

where the first inequality follows by (3.40), (3.45), and (3.46) and the last one by (3.44). This implies by the definition in (3.47) that

$$\mathcal{B}'_s \subset \mathcal{G}_s\,. \tag{3.51}$$

Moreover, by using (3.45) it holds for all $s \in \mathcal{I}(\hat{s})$ that

$$\begin{aligned} P_{D_s|\hat{S}}(\mathcal{D}'^c_s|\hat{s}) &= \sum_{d\in\mathcal{D}'^c_s} P_{D_s|\hat{S}}(d|\hat{s}) \\ &< \sum_{d\in\mathcal{D}} \frac{\alpha^2|\mathcal{B}_{s,d}|}{|\mathcal{B}_s|} \\ &= \alpha^2. \end{aligned} \tag{3.52}$$

On the other hand, the following equation is a result of (3.46):

$$\begin{aligned} P_{CD,s|\hat{S}}(\mathcal{B}_s \cup (\mathcal{C}\times\mathcal{D}'_s)|\hat{s}) = P_{CD,s|\hat{S}}(\mathcal{B}_s|\hat{s}) \\ + P_{D_s|\hat{S}}(\mathcal{D}'_s|\hat{s}) - P_{CD,s|\hat{S}}(\mathcal{B}'_s|\hat{s}). \end{aligned} \tag{3.53}$$

The relations (3.52) and (3.53) together with the assumption (3.41) of the lemma imply that

$$\begin{aligned} P_{CD,s|\hat{S}}(\mathcal{B}'_s|\hat{s}) &\geq P_{CD,s|\hat{S}}(\mathcal{B}_s|\hat{s}) - P_{D_s|\hat{S}}(\mathcal{D}'^c_s|\hat{s}) \\ &\geq 1-(\eta^2-\alpha^2)-\alpha^2 \\ &= 1-\eta^2. \end{aligned} \tag{3.54}$$

By using (3.51) and (3.54), it follows that

$$P_{CD,s|\hat{S}}(\mathcal{G}_s|\ \hat{s}) \geq 1-\eta^2. \tag{3.55}$$

The lower bound (3.55) implies by the definition in (3.48) that for all $s \in \mathcal{I}(\hat{s})$

$$\begin{aligned} 1-\eta^2 &\leq \sum_{d\in\mathcal{D}} P_{D_s|\hat{S}}(d|\hat{s}) P_{C|D\hat{S},s}(\mathcal{G}_{s,d}|d,\hat{s}) \\ &< \sum_{d\in\mathcal{E}_s} P_{D_s|\hat{S}}(d|\hat{s})(1-\eta) + \sum_{d\notin\mathcal{E}_s} P_{D_s|\hat{S}}(d|\hat{s}) \\ &= -\eta \sum_{d\in\mathcal{E}_s} P_{D_s|\hat{S}}(d|\hat{s}) + 1, \end{aligned}$$

where the second inequality follows by (3.49). The desired relation (3.50) follows by simplifying this inequality.

In the second step, we show that an SK generator κ, satisfying (3.43), exists. For this, consider each member of the family of PDs $P_{C|D\hat{S},s}(.|d,\hat{s})$ with $d \notin \mathcal{E}_s$ and $s \in \mathcal{I}(\hat{s})$. Lemma 2.7 implies for a randomly selected SK generator κ that

$$\Pr\Big(\big\|\kappa(P_{C|D\hat{S},s}(.|d,\hat{s})) - P_0\big\| > 2(\alpha+\eta)\Big) \leq 2k\mathrm{e}^{-\frac{\lambda\alpha^2}{k}}, \tag{3.56}$$

where $P_0(i) = 1/k$ for all $i = 1, 2\cdots k$. The universal upper bound in (3.56) was calculated by taking $\epsilon = 2\alpha$ in Lemma 2.7 and by the inequalities $\alpha \leq 1/6$ and $\eta \leq 1/3$ from the assumption of the lemma. Therefore, for the following events

$$\mathcal{A}_{s,d} := \Big\{\big\|\kappa(P_{C|D\hat{S},s}(.|d,\hat{s})) - P_0\big\| \leq 2(\alpha+\eta)\Big\},$$

it follows by (3.44) and (3.56) that

$$\begin{aligned} \Pr\Big(\bigcap_{s\in\mathcal{I}(\hat{s}),d\notin\mathcal{E}_s} \mathcal{A}_{s,d}\Big) &\geq 1 - \sum_{s\in\mathcal{I}(\hat{s}),d\notin\mathcal{E}_s} \Pr\big(\mathcal{A}^{\mathrm{c}}_{s,d}\big) \\ &\geq 1 - 2k\,|\mathcal{D}|\,|\mathcal{I}(\hat{s})|\mathrm{e}^{-\frac{\alpha^5 \min|\mathcal{B}_{s,d}|}{k}}, \end{aligned} \tag{3.57}$$

where the min in (3.57) is taken over all $s \in \mathcal{I}(\hat{s})$ and $d \in \mathcal{D}_s$.

This means that an SK generator κ satisfies the relation $\|\kappa(P_{C|D\hat{S},s}(.|d,\hat{s})) - P_0\| \leq 2(\alpha+\eta)$ simultaneously for all $d \notin \mathcal{E}_s$ and $s \in \mathcal{I}(\hat{s})$ with the probability stated in (3.57). Therefore, it holds by the same probability that

$$\begin{aligned} S\big(\kappa(C)|D_s,\hat{S}=\hat{s}\big) &= \sum_{d\in\mathcal{D}} P_{D_s|\hat{S}}(d|\hat{s})\Big[\log k - H\big(\kappa(C)|D_s=d,\hat{S}=\hat{s}\big)\Big] \\ &\leq \sum_{d\notin\mathcal{E}_s} P_{D_s|\hat{S}}(d|\hat{s})\big[(\alpha+\eta)\log k + h(\alpha+\eta)\big] \\ &\qquad + \sum_{d\in\mathcal{E}_s} P_{D_s|\hat{S}}(d|\hat{s})\log k \\ &\leq (\alpha+2\eta)\log k + h(\alpha+\eta), \end{aligned}$$

which gives the desired relation in (3.43). The equality is the result of Definition 3.2 and the first inequality follows from the uniform continuity of entropy [43], [18, Problem 3.10].

The last step is a result of the inequality (3.50). Moreover, assumption (3.42) implies that

$$k \ln \left(2k|\mathcal{D}| \cdot |\mathcal{I}(\hat{s})| \right) < \alpha^5 \cdot \min_{s \in \mathcal{I}(\hat{s}), d \in \mathcal{D}_s} |\mathcal{B}_{s,d}|,$$

and consequently, the probability in (3.57) is non-zero. □

4 Secret-Key Capacity of Finite Compound Sources

In this chapter, SK capacity results for the SK model from Chapter 3, Section 3.2 are presented. It is assumed that the compound set is finite. Furthermore, some upper bounds for the range size of the auxiliary RVs are derived.

4.1 Single-Letter Secret-Key Capacity Lower Bound

In the following, Theorem 4.1 gives a single-letter lower bound for the SK capacity as a function of the compound DMMS and public communication rate upper bound.

Similar techniques which are used for deriving the non-compound SK capacity results in [18, Theorem 17.21], [17], are used and extended to the finite compound setup. For completeness, all proofs are presented in detail.

Theorem 4.1. *Consider the SK model from Section 3.2 with a finite compound DMMS* $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$*, set of marginals* $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s}\in\hat{\mathcal{S}}}$*, and a one-way communication over a public noiseless channel with communication rate upper bound* $\Gamma \in (0, \infty]$*. Then, it holds that*

$$C_{\mathrm{sk}}(\mathfrak{S}, \Gamma) \geq \min_{\hat{s}\in\hat{\mathcal{S}}} \max_{U_{\hat{s}}, V_{\hat{s}}} \Big\{ \min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s|U_{\hat{s}}) - \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s|U_{\hat{s}}) \Big\}, \tag{4.1}$$

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

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}} - Y_s Z_s\,, \tag{4.2}$$

$$\max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_s) \leq \Gamma. \tag{4.3}$$

Proof. Let $\hat{\mathcal{S}} = \{\hat{s}_1, \dots \hat{s}_m\}$ be the set of marginal indices and $\hat{s} \in \hat{\mathcal{S}}$ be given. Let $U_{\hat{s}}$ and $V_{\hat{s}}$ satisfy for all $s \in \mathcal{I}(\hat{s})$ the Markov chains $U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}} - Y_s Z_s$.

Consider $N_{\hat{s},1}N_{\hat{s},2}$ sequences $u_{ij}^n(\hat{s}) \in \mathcal{U}_{\hat{s}}^n$ as rows of Table 4.1 for each $\hat{s} \in \hat{\mathcal{S}}$, which are chosen independently by PD $P_{U_{\hat{s}}}^n$ with

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

and $N_{\hat{s},1}$ and $N_{\hat{s},2}$ satisfying (3.9) and (3.10) from Lemma 3.2. Moreover, for every element $u_{ij}^n(\hat{s})$ of Table 4.1, consider $N_{\hat{s},3}N_{\hat{s},4}$ sequences $v_{pq}^{ij\,n}(\hat{s}) \in \mathcal{V}_{\hat{s}}^n$, which are chosen conditionally independently by PD $P_{V_{\hat{s}}|U_{\hat{s}}}^n(.|u_{ij}^n(\hat{s}))$ with

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

$\hat{s}_1$	$u_{11}^n(\hat{s}_1)$	$\cdots$	$u_{ij}^n(\hat{s}_1)$	$\cdots$	$u_{N_{\hat{s}_1,1}N_{\hat{s}_1,2}}^n(\hat{s}_1)$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
$\hat{s}_m$	$u_{11}^n(\hat{s}_m)$	$\cdots$	$u_{ij}^n(\hat{s}_m)$	$\cdots$	$u_{N_{\hat{s}_m,1}N_{\hat{s}_m,2}}^n(\hat{s}_m)$

Table 4.1: Random sequences for a DMMS with $|\hat{\mathcal{S}}| = m$

and $N_{\hat{s},3}$ and $N_{\hat{s},4}$ satisfying (3.11) and (3.12) from Lemma 3.2. Assume that the random sequences $u_{ij}^n(\hat{s})$ in Table 4.1 and their corresponding sequences $\{v_{pq}^{ij\,n}(\hat{s})\}_{(p,q)\in\mathcal{P}\times\mathcal{Q}}$ are known to Alice and Bob.

To show the achievability of (4.1), the proof is divided into two parts. In part a), the following rate is shown to be achievable:

$$R'_{\mathrm{sk}} := \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Y_s) - \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Z_s), \tag{4.4}$$

when R'_{sk} is positive and RV $U_{\hat{s}}$ satisfies for all $s \in \mathcal{I}(\hat{s})$

$$U_{\hat{s}} - X_{\hat{s}} - Y_s Z_s \quad \text{and} \quad \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};X_{\hat{s}}|Y_s) \le \Gamma. \tag{4.5}$$

This gives a special case of (4.1), (4.2) and (4.3). In part b), the achievability of the SK rate in (4.1) is shown, when it is positive.

Part a) Let $\delta > 0$ be given. Assume that R'_{sk} from (4.4) is positive i.e.

$$\min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Y_s) > \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};Z_s). \tag{4.6}$$

As explained in Section 3.3, Alice estimates her marginal statistic by hypothesis testing. Assume that $\hat{s} \in \hat{\mathcal{S}}$ is the index corresponding to the correct decision and all other $\tilde{s} \in \hat{\mathcal{S}} - \{\hat{s}\}$ correspond to a wrong decision. The PD $P_{\hat{S}}$ is given by (3.7) and (3.8).

Next, Alice sends her estimated marginal index to Bob over the public noiseless channel. Assume that the hypothesis testing has led to the correct decision $\hat{s}$. Alice and Bob find the corresponding family of sequences $\{u_{ij}^n(\hat{s})\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$ from Table 4.1 by knowing $\hat{s}$. Lemma 3.2a implies for all $\zeta > 0$ small enough and $n \in \mathbb{N}$ sufficiently large, the existence of some encoder functions $f : \mathcal{X}^n \to \mathcal{I} \cup \{0\}$ and $g : \mathcal{X}^n \to \mathcal{J} \cup \{0\}$. Alice uses these encoders to find the indices $f(x^n) = i$ and $g(x^n) = j$ of the sequence $u_{ij}^n(\hat{s})$ to be chosen from the family of sequences $\{u_{ij}^n(\hat{s})\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$.

As shown in Figure 4.1, in addition to the transmitted $\hat{s}$, Alice sends further the index $f(x^n) = i$ to Bob over the public channel. Lemma 3.2a implies for all $\sigma > \zeta$ small enough and $n \in \mathbb{N}$ sufficiently large, the existence of a decoder function $\tilde{g} : \mathcal{I} \times \hat{\mathcal{S}} \times \mathcal{Y}^n \to \mathcal{J}$, with which, Bob can reconstruct the index $g(x^n) = j$. This j is the CR between Alice and Bob.

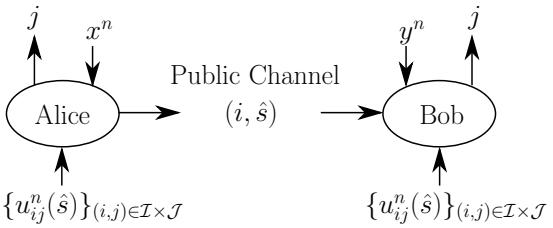

Figure 4.1: Generating the CR j

In total, for all Alice's estimation results which may lead to a correct or incorrect decision, the error probability upper bound for all $s \in \mathcal{I}(\hat{s})$ is given by

$$\begin{aligned}\Pr\Big\{g(X_{\hat{s}}^n) \neq \tilde{g}\big(f(X_{\hat{s}}^n), \hat{S}, Y_s^n\big)\Big\} &= \Pr\Big\{g(X_{\hat{s}}^n) \neq \tilde{g}\big(f(X_{\hat{s}}^n), \hat{s}, Y_s^n\big) \wedge \hat{S} = \hat{s}\Big\} \\ &\quad + \sum_{\tilde{s}\in\hat{\mathcal{S}}-\{\hat{s}\}} P_{\hat{S}}(\tilde{s}) \Pr\Big\{g(X_{\hat{s}}^n) \neq \tilde{g}\big(f(X_{\hat{s}}^n), \hat{S}, Y_s^n\big) \,\Big|\, \hat{S} = \tilde{s}\Big\} \\ &\leq \exp(-n\delta_0) + \exp(-nc_1) \cdot |\hat{\mathcal{S}}|, \end{aligned} \tag{4.7}$$

where the inequality is a result of (3.8) and (3.15) for some $c_1, \delta_0 > 0$. Thus, condition (3.5) of Definition 3.3 is satisfied.

The whole message which is sent over the public channel is represented by RV $f_{\mathrm{c}}(X_{\hat{s}}^n) = (f(X_{\hat{s}}^n), \hat{S})$ having the range size $\|f\| \cdot |\hat{\mathcal{S}}|$. As shown in the following, the communication rate satisfies condition (3.3) of Definition 3.3:

$$\begin{aligned}\frac{1}{n}\log\|f_{\mathrm{c}}\| &= \frac{1}{n}\log(\|f\| \cdot |\hat{\mathcal{S}}|) \\ &= \frac{1}{n}\log(N_{\hat{s},1} \cdot |\hat{\mathcal{S}}|) \\ &= \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + 3\delta + \frac{1}{n}\log|\hat{\mathcal{S}}| \\ &\leq \Gamma + 4\delta, \end{aligned}$$

where the last equality follows by (3.9) and the inequality is a result of (4.5) and is valid for all n sufficiently large.

After the index $g(x^n) = j$ is reconstructed by Bob, both Alice and Bob may generate their SK, based on this CR. Thus, it remains to show that there exists an SK generator $\kappa : \mathcal{J} \to \{1, 2 \cdots, k\}$, giving rise to the RV $K_{\mathrm{A}} = \kappa(g(X_{\hat{s}}^n))$, which satisfies condition (3.6) of Definition 3.3.

Again the condition is verified for both estimation results. Assume hypothesis testing has led to the correct decision and $\hat{s}$ is sent to Bob over the public channel. Define for all $s \in \mathcal{I}(\hat{s})$,

$$\mathcal{T}_s := \Big\{(x^n, z^n) \in \mathcal{X}^n \times \mathcal{Z}^n : \; x^n \in \mathcal{T} \;\wedge\; (u_{ij}^n(\hat{s}), x^n, z^n) \in \mathcal{T}_{[UXZ,s]\sigma}^n\Big\},$$

where $\mathcal{T}$ is given in (3.14). A similar discussion as for (3.22) in the proof of Lemma 3.2, implies that

$$\begin{aligned} P^n_{XZ,s}(\mathcal{T}_s^{\mathrm{c}}) &= \sum_{x^n\in\mathcal{T}^{\mathrm{c}},z^n\in\mathcal{Z}^n} P^n_{XZ,s}(x^n,z^n) + \sum_{x^n\in\mathcal{T},(x^n,z^n)\notin\mathcal{T}_s} P^n_{XZ,s}(x^n,z^n) \\ &= P^n_{X_{\hat{s}}}(\mathcal{T}^{\mathrm{c}}) + \sum_{x^n\in\mathcal{T}} P^n_{X_{\hat{s}}}(x^n)\; P^n_{Z_s|X_{\hat{s}}}\Big(\mathcal{T}^{n\;\mathrm{c}}_{[UXZ,s]\sigma}(u^n_{ij}(\hat{s}),x^n)\big|x^n\Big) \\ &\leq 2|\mathcal{X}|\mathrm{e}^{-2\xi^2 n} + 2|\mathcal{U}_{\hat{s}}||\mathcal{X}||\mathcal{Z}|\mathrm{e}^{-2(\sigma-\zeta)^2 n}, \end{aligned} \tag{4.8}$$

where the upper bound goes to zero exponentially fast as n goes to infinity. The constant ξ was given in deriving (3.22) such that $\xi\in(0,\zeta)$.

Similarly as in [18, Theorem 17.21] for the non-compound version, define for all $s\in\mathcal{I}(\hat{s})$, the RVs C and D_s and the set $\mathcal{B}_s$ to be used in Lemma 3.3, as follows

$$\begin{aligned} C &:= g(X^n_{\hat{s}}), \\ D_s &:= \big(f(X^n_{\hat{s}}), Z^n_s, \mathbb{1}_{\mathcal{T}_s}(X^n_{\hat{s}},Z^n_s)\big), \\ \mathcal{B}_s &:= \Big\{\big(j,(i,z^n,1)\big) : (i,j)\in\mathcal{I}\times\mathcal{J}, z^n\in\mathcal{T}^n_{[Z_s]\xi}, \mathcal{T}^n_{[UXZ,s]\sigma}(u^n_{ij}(\hat{s}),z^n)\neq\emptyset\Big\}. \end{aligned} \tag{4.9}$$

Assume that RVs C and D_s take their values in the sets $\mathcal{C}$ and $\mathcal{D}$ respectively. Moreover, let the sets $\mathcal{D}_s$ and $\mathcal{B}_{s,d}$ be defined as in Lemma 3.3. In the following, it is shown that all conditions of Lemma 3.3 are satisfied. It holds that

$$\begin{aligned} P_{CD,s|\hat{S}}(\mathcal{B}_s|\hat{s}) &= \sum_{(j,(i,z^n,1))\in\mathcal{B}_s} P_{CD,s|\hat{S}}\big(j,(i,z^n,1)|\hat{s}\big) \\ &= \sum_{(j,(i,z^n,1))\in\mathcal{B}_s} \sum_{\substack{x^n: f(x^n)=i, g(x^n)=j,\\ \mathbb{1}_{\mathcal{T}_s}(x^n,z^n)=1}} P_{X^n_{\hat{s}}Z^n_s|\hat{S}}(x^n,z^n|\hat{s}) \\ &= P_{X^n_{\hat{s}}Z^n_s|\hat{S}}\Big(\mathcal{T}_s\cap\Big\{(x^n,z^n)\in\mathcal{X}^n\times\mathcal{Z}^n : z^n\in\mathcal{T}^n_{[Z_s]\xi}\Big\}\;\Big|\;\hat{s}\Big) \\ &\geq 1-\Big[P_{X^n_{\hat{s}}Z^n_s|\hat{S}}(\mathcal{T}_s^{\mathrm{c}}|\hat{s}) + P_{Z^n_s|\hat{S}}(\mathcal{T}^{n\;\mathrm{c}}_{[Z_s]\xi}|\hat{s})\Big] \\ &\geq 1-\frac{P_{X^n_{\hat{s}}Z^n_s}(\mathcal{T}_s^{\mathrm{c}}) + P_{Z^n_s}(\mathcal{T}^{n\;\mathrm{c}}_{[Z_s]\xi})}{1-\exp(-nc_0)}, \end{aligned} \tag{4.10}$$

where the last inequality follows by using (3.7). Lemma 2.10.2 and (4.8) imply that the lower bound in (4.10) approaches one as n goes to infinity. The constants are universal for all sets $\mathcal{I}(\hat{s})$. Define the parameters α and η for some arbitrary $\tau>0$ and $\nu\in(0,\xi)$ as follows

$$\alpha := \exp(-n(\delta+5\tau)), \quad \eta := \exp(-n\nu). \tag{4.11}$$

It follows by using (4.8) and (4.10) that for n sufficiently large and ν small enough,

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

This guarantees condition (3.41) of Lemma 3.3. Moreover, the conditions $\alpha \in (0, 1/6]$ and $\eta \in (0, 1/3]$ of Lemma 3.3 are satisfied, if n is sufficiently large.

To check condition (3.40) of Lemma 3.3, we find in the first step an upper bound for $|\mathcal{B}_s|$. The non-emptiness constraint $\mathcal{T}^n_{[UXZ,s]\sigma}(u^n_{ij}(\hat{s}), z^n) \neq \emptyset$ from the definition of the set $\mathcal{B}_s$ in (4.9) is a sufficient condition for $u^n_{ij}(\hat{s}) \in \mathcal{T}^n_{[UZ,s]\sigma|\mathcal{X}|}(z^n)$ and thus

$$|\mathcal{B}_s| \leq \sum_{z^n \in \mathcal{T}^n_{[Z_s]\xi}} \Big| \Big\{ (i,j) \in \mathcal{I} \times \mathcal{J} : u^n_{ij}(\hat{s}) \in \mathcal{T}^n_{[UZ,s]\sigma|\mathcal{X}|}(z^n) \Big\} \Big|. \tag{4.12}$$

Furthermore, for R being the rate of choosing the random sequences $\{u^n_{ij}(\hat{s})\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$, it holds by using (3.9) and (3.10) that

$$\begin{aligned} R &= \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) + \delta \\ &> \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s), \end{aligned} \tag{4.13}$$

where the inequality is the result of assumption (4.6). Moreover, for all $(j, (i, z^n, 1)) \in \mathcal{B}_s$ it holds that $z^n \in \mathcal{T}^n_{[Z_s]\xi}$, which implies by using Lemma 2.10.1 that $\mathcal{T}^n_{[UZ,s]\zeta}(z^n) \neq \emptyset$. Assume that τ is small enough such that

$$\tau < R - \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s). \tag{4.14}$$

It follows by using (4.12), (4.13), (4.14), and Lemma 2.11 that for $\sigma > \zeta > \xi > 0$ sufficiently small and n large enough,

$$\begin{aligned} |\mathcal{B}_s| &\leq \big|\mathcal{T}^n_{[Z_s]\xi}\big| \exp\Big[n\big(R - I(U_{\hat{s}}; Z_s) + \tau\big)\Big] \\ &\leq \exp\Big[n\big(H(Z_s) + \tau\big)\Big] \exp\Big[n\big(I(U_{\hat{s}}; X_{\hat{s}}) + \delta - I(U_{\hat{s}}; Z_s) + \tau\big)\Big], \end{aligned} \tag{4.15}$$

where the last inequality follows by using (3.18) and Lemma 2.10.6. The size of τ as a function of ξ, σ, n, and other constant quantities is given by using (2.4) and (2.8). The constants are universal for all sets $\mathcal{I}(\hat{s})$.

In the second step of verifying (3.40), an upper bound for $P_{CD,s|\hat{S}}\big(j, (i, z^n, 1)|\hat{s}\big)$ is derived. For all $(j, (i, z^n, 1)) \in \mathcal{B}_s$, it holds by using (3.7) and (4.9) that

$$\begin{aligned} \big(1 - \exp(-nc_0)\big) P_{CD,s|\hat{S}}\big(j, (i, z^n, 1)|\hat{s}\big) &\leq P_{CD,s}\big(j, (i, z^n, 1)\big) \\ &\leq \sum_{x^n \in \mathcal{T}^n_{[UXZ,s]\sigma}(u^n_{ij}(\hat{s}), z^n)} P^n_{XZ,s}(x^n, z^n) \\ &\leq \big|\mathcal{T}^n_{[UXZ,s]\sigma}(u^n_{ij}(\hat{s}), z^n)\big| \exp\Big[-n\big(H(X_{\hat{s}}Z_s) - \tau\big)\Big] \\ &\leq \exp\Big[n\big(H(X_{\hat{s}}|U_{\hat{s}}Z_s) - H(X_{\hat{s}}Z_s) + 2\tau\big)\Big], \end{aligned} \tag{4.16}$$

where the second inequality follows by the fact that $\mathbb{1}_{\mathcal{T}_s}(X_{\hat{s}}^n, Z_s^n) = 1$. The third inequality follows by using Lemma 2.10.4 and the last one by Lemma 2.10.7 for σ sufficiently small and n large enough.

The size of τ as a function of σ, n, and other constant quantities is given by using (2.2) and (2.5). In contrast with previous cases, for a given τ, the constant σ depends on $\mu_{X_{\hat{s}}Z_s}$ as given by (2.2) and possibly different for each $s \in \mathcal{I}(\hat{s})$. To make it universal for all elements of the compound set, we take the minimum of all σ for each $s \in \mathcal{I}(\hat{s})$. As in this chapter, the set $\mathcal{I}(\hat{s})$ is assumed to be constant and finite, this minimum exists. For the case where the compound set is infinite, the finite approximating compound sets $\mathcal{I}_n(\hat{s})$ are applied. In this case the sets $\mathcal{I}_n(\hat{s})$ are not constant and change by increasing code length n. This problem will be explained in more details in Chapter 5.
By using Markov chains in (4.5), it holds that

$$H(X_{\hat{s}}|U_{\hat{s}}Z_s) - H(X_{\hat{s}}Z_s) = -I(U_{\hat{s}}; X_{\hat{s}}) - H(Z_s) + I(U_{\hat{s}}; Z_s). \tag{4.17}$$

The relations (4.15), (4.16), and (4.17) together with the definition of α in (4.11) imply for n large enough that

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

Thus, condition (3.40) of Lemma 3.3 is also satisfied.

Therefore, Lemma 3.3 implies that there exists an SK generator $\kappa : \mathcal{J} \to \{1, 2, \cdots, k\}$ with k satisfying (3.42), such that the relation (3.43) holds. By using the definitions in (4.11), it follows that both α and η approach zero exponentially fast. Moreover, by using (3.42) and (4.15), it follows that k does not increase faster than exponentially. Thus, the following upper bound which is given by using (3.43) for $K_{\mathrm{A}} = \kappa(g(X_{\hat{s}}^n))$ approaches zero exponentially fast, as n goes to infinity:

$$S\big(K_{\mathrm{A}} \,\big|\, f(X_{\hat{s}}^n), Z_s^n, \mathbb{1}_{\mathcal{T}_s}(X_{\hat{s}}^n, Z_s^n), \hat{S} = \hat{s}\big) \leq (\alpha + 2\eta)\log k + h(\alpha + \eta). \tag{4.18}$$

In total, for all estimation results which may lead to a correct or incorrect decision in hypothesis testing by Alice, the security index is given by

$$\begin{aligned} &S\big(K_{\mathrm{A}} \,\big|\, f(X_{\hat{s}}^n), Z_s^n, \mathbb{1}_{\mathcal{T}_s}(X_{\hat{s}}^n, Z_s^n), \hat{S}\big) \\ &\quad = P_{\hat{S}}(\hat{s}) S\big(K_{\mathrm{A}}|Z_s^n, f(X_{\hat{s}}^n), \mathbb{1}_{\mathcal{T}_s}(X_{\hat{s}}^n, Z_s^n), \hat{S} = \hat{s}\big) \\ &\qquad + \sum_{\tilde{s} \in \hat{\mathcal{S}} - \{\hat{s}\}} P_{\hat{S}}(\tilde{s}) S\big(K_{\mathrm{A}}|Z_s^n, f(X_{\hat{s}}^n), \mathbb{1}_{\mathcal{T}_s}(X_{\hat{s}}^n, Z_s^n), \hat{S} = \tilde{s}\big). \end{aligned}$$

By using (3.7), (3.8) and (4.18), it follows that the security index also goes to zero exponentially fast. Therefore, condition (3.6) of Definition 3.3 is satisfied.

In the rest of this part, we show that R'_{sk} which was defined in (4.4), satisfies condition (3.4) of Definition 3.3. By using the definition of α in (4.11) and the set $\mathcal{D}$ being the alphabet of RV D_s, it follows that the expression $\mathrm{e}^{1/\alpha}(2|\mathcal{D}|\,|\mathcal{I}(\hat{s})|)^{-1}$ from (3.42) increases doubly exponentially fast. On the other hand, for $\mathcal{B}_{s,d}$ from Lemma 3.3 with

$d := (i, z^n, 1) \in \mathcal{D}_s$, it follows by using (4.15) that $|\mathcal{B}_{s,d}|$ does not increase faster than exponentially. Therefore, for n large enough, it holds that

$$\alpha^6 \min_{s\in\mathcal{I}(\hat{s}),d\in\mathcal{D}_s} \big|\mathcal{B}_{s,d}\big| < \mathrm{e}^{1/\alpha}(2|\mathcal{D}|\,|\mathcal{I}(\hat{s})|)^{-1}. \tag{4.19}$$

Thus, to guarantee condition (3.42) of Lemma 3.3 it is necessary that k be lower than the left hand side of (4.19). For this, a lower bound for $\min_{s\in\mathcal{I}(\hat{s}),d\in\mathcal{D}_s} |\mathcal{B}_{s,d}|$ should first be determined.

To this end, we apply Lemma 2.10.1 which implies that $u_{ij}^n(\hat{s}) \in \mathcal{T}^n_{[UZ,s]\zeta}(z^n)$ is a sufficient condition for $\mathcal{T}^n_{[UXZ,s]\sigma}(u_{ij}^n(\hat{s}), z^n) \neq \emptyset$. Therefore for all $s \in \mathcal{I}(\hat{s})$ and $d = (i, z^n, 1) \in \mathcal{D}_s$, it holds that

$$\big|\mathcal{B}_{s,d}\big| \geq \Big|\Big\{j \in \mathcal{J} : u_{ij}^n(\hat{s}) \in \mathcal{T}^n_{[UZ,s]\zeta}(z^n)\Big\}\Big|. \tag{4.20}$$

Furthermore by using (3.10), the rate of choosing the random sequences $u_{ij}^n(\hat{s})$ for a fixed index i, is given by

$$R_2 = \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) - 2\delta. \tag{4.21}$$

Based on assumption (4.6), assume that

$$0 < 2\delta < \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) - \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s). \tag{4.22}$$

This implies by using (4.21) that

$$R_2 > \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s). \tag{4.23}$$

Assume that τ is small enough such that

$$\tau < R_2 - \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s). \tag{4.24}$$

Moreover, in the proof of (4.15), it was shown that $\mathcal{T}^n_{[UZ,s]\zeta}(z^n) \neq \emptyset$. Thus, Lemma 2.11 together with (4.20), (4.21), (4.23), and (4.24) implies that

$$\begin{aligned}
\alpha^6 \min_{s\in\mathcal{I}(\hat{s}),d\in\mathcal{D}_s} |\mathcal{B}_{s,d}| &\geq \exp(-6n(\delta+5\tau)) \min_{s\in\mathcal{I}(\hat{s}),d\in\mathcal{D}_s} \Big|\Big\{j : u_{ij}^n(\hat{s}) \in \mathcal{T}^n_{[UZ,s]\zeta}(z^n)\Big\}\Big| \\
&> \exp(-6n(\delta+5\tau)) \exp\Big[n\big(R_2 - \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s) - \tau\big)\Big] \\
&= \exp\Big[n\big(\min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) - \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s) - 8\delta - 31\tau\big)\Big].
\end{aligned} \tag{4.25}$$

The size of τ as a function of ζ, n, and other constant quantities is given by using (2.8). The constants in this case are universal for all sets $\mathcal{I}(\hat{s})$.

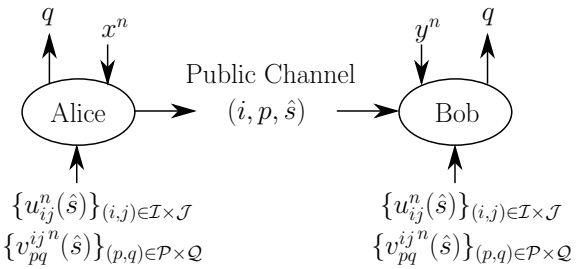

Figure 4.2: Generating the CR q

By keeping $|\mathcal{K}| = k$ lower than the left hand side of (4.19), condition (3.42) of Lemma 3.3 is guaranteed. Therefore, it follows by using (4.25) and the definition of R'_{sk} from (4.4) that for $\tau > 0$ small enough

$$\frac{1}{n}\log|\mathcal{K}| > R'_{\mathrm{sk}} - 8\delta.$$

This gives condition (3.4) of Definition 3.3.

Part b) Let $\delta > 0$ be given. In this part, it is shown that if

$$R_{\mathrm{sk}} := \min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s|U_{\hat{s}}) - \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s|U_{\hat{s}}) > 0, \tag{4.26}$$

then R_{sk} is achievable for all RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ satisfying (4.2) and (4.3). It may be assumed that

$$\min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) \le \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s). \tag{4.27}$$

Because otherwise, it follows by using the Markov chains in (4.2) that

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

which implies by using part a) that R_{sk} is achievable.

Similarly as in part a), Alice sends her estimated marginal index to Bob over the public noiseless channel. For the case that hypothesis testing has led to the correct decision $\hat{s}$, both Alice and Bob find the corresponding family of sequences $\{u_{ij}^n(\hat{s})\}_{(i,j)\in\mathcal{I}\times\mathcal{J}}$ from Table 4.1 and also the related families of sequences $\{v_{pq}^{ij\,n}(\hat{s})\}_{(p,q)\in\mathcal{P}\times\mathcal{Q}}$ for each member of the chosen row of the table. Lemma 3.2b implies for all $\sigma > 0$ small enough and $n \in \mathbb{N}$ sufficiently large, the existence of some encoder functions $\varphi : \mathcal{X}^n \to \mathcal{P}\cup\{0\}$ and $\rho : \mathcal{X}^n \to \mathcal{Q}\cup\{0\}$. These encoders give the indices $\varphi(x^n) = p$ and $\rho(x^n) = q$ of the

sequence $v_{pq}^{ij\,n}(\hat{s})$ to be chosen from the family of sequences $\{v_{pq}^{ij\,n}(\hat{s})\}_{(p,q)\in\mathcal{P}\times\mathcal{Q}}$ for the given indices i and j.

As shown in Figure 4.2, in addition to the transmitted $\hat{s}$, Alice sends further the indices $f(x^n) = i$ and $\varphi(x^n) = p$ to Bob over the public channel. By part a) of this theorem, $g(x^n) = j$ is known to Bob by a probability close to one. Lemma 3.2b implies for all $\vartheta > \sigma$ small enough and n sufficiently large, the existence of a decoder function $\tilde{\rho} : \mathcal{I} \times \mathcal{J} \times \mathcal{P} \times \hat{\mathcal{S}} \times \mathcal{Y}^n \to \mathcal{Q}$, with which Bob can reconstruct the index q to be used as the CR. The upper bound of the error probability was given in (3.17) in Lemma 3.2. Due to (4.27), the index $g(x^n) = j$ can not be used as the CR any more. This is because in this case assumption (4.22) can not be made.

Similarly as in (4.7) in part a), for all estimation results which may lead to a correct or incorrect decision, the error probability upper bound for all $s \in \mathcal{I}(\hat{s})$ is exponentially small. Therefore, condition (3.5) of Definition 3.3 is satisfied.
The public communication function is represented by the RV

$$f_{\mathrm{c}}(X_{\hat{s}}^n) = (f(X_{\hat{s}}^n), \varphi(X_{\hat{s}}^n), \hat{S}).$$

As shown in the following, the communication rate satisfies condition (3.3) of Definition 3.3:

$$\begin{aligned}\frac{1}{n}\log\left(\|f\| \cdot \|\varphi\| \cdot |\hat{\mathcal{S}}|\right) &= \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_s) + 6\delta + \frac{1}{n}\log|\hat{\mathcal{S}}| \\ &\leq \Gamma + 7\delta,\end{aligned}$$

for n sufficiently large. The equality is a result of (3.9) and (3.11) from Lemma 3.2 and the inequality follows by (4.3).

For showing that conditions (3.4) and (3.6) of Definition 3.3 also hold, the RVs C, D_s, and the set $\mathcal{B}_s$ are defined similarly as in part a), but this time for the coding scheme from Lemma 3.2b). To this end, an analogue to the set $\mathcal{T}_s$ from part a) is first defined. Similarly as in [18, Theorem 17.21] for the non-compound version, let $\vartheta > \sigma|\mathcal{X}|$ be given and define

$$\mathcal{T}_s' := \left\{(x^n, z^n) \in \mathcal{X}^n \times \mathcal{Z}^n : x^n \in \mathcal{T} \ \wedge\ (u_{ij}^n(\hat{s}), v_{pq}^{ij\,n}(\hat{s}), x^n, z^n) \in \mathcal{T}_{[UVXZ,s]\vartheta}^n\right\}.$$

Similarly as in (4.8), it can be shown that $P_{XZ,s}^n(\mathcal{T}_s'^{\mathrm{c}})$ goes to zero exponentially fast as n goes to infinity.

Define for all $s \in \mathcal{I}(\hat{s})$ the RVs C and D_s and the set $\mathcal{B}_s$ to be used in Lemma 3.3 as follows

$$\begin{aligned}C &:= \rho(X_{\hat{s}}^n), \\ D_s &:= \left(f(X_{\hat{s}}^n), g(X_{\hat{s}}^n), \varphi(X_{\hat{s}}^n), Z_s^n, \mathbb{1}_{\mathcal{T}_s'}(X_{\hat{s}}^n, Z_s^n)\right), \\ \mathcal{B}_s &:= \Big\{\big(q, (i, j, p, z^n, 1)\big) : (i, j, p, q) \in \mathcal{I} \times \mathcal{J} \times \mathcal{P} \times \mathcal{Q}, \\ &\qquad z^n \in \mathcal{T}_{[Z_s]\xi}^n \ \wedge\ \mathcal{T}_{[UXZ,s]\sigma}^n(u_{ij}^n(\hat{s}), z^n) \neq \emptyset \ \wedge\ \mathcal{T}_{[UVXZ,s]\vartheta}^n(u_{ij}^n(\hat{s}), v_{pq}^{ij\,n}(\hat{s}), z^n) \neq \emptyset\Big\}.\end{aligned}$$

Assume that RVs C and D_s take their values in the sets $\mathcal{C}$ and $\mathcal{D}$ respectively. Moreover, let the sets $\mathcal{D}_s$ and $\mathcal{B}_{s,d}$ be defined as in Lemma 3.3.

Define α and η as in (4.11). Similarly as in part a), it can be shown that all conditions of Lemma 3.3 including conditions (3.40) and (3.41) are satisfied. Thus, there exists an SK generator $\kappa : \mathcal{Q} \to \{1, 2 \cdots, k\}$, giving rise to the RV $K = \kappa(\rho(X_{\hat{s}}^n))$, which satisfies condition (3.6) of Definition 3.3.

The rate of choosing sequences $v_{pq}^{ij\,n}(\hat{s})$ for fixed indices i, j, and p is given by using (3.12) as follows.

$$R_3 = \min_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s | U_{\hat{s}}) - 2\delta.$$

Based on assumption (4.26), assume that

$$0 < 2\delta < \min_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s | U_{\hat{s}}) - \max_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s | U_{\hat{s}}).$$

This implies that

$$R_3 > \max_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s | U_{\hat{s}}).$$

Therefore, Lemma 2.12 with this rate and some θ satisfying $\sigma|\mathcal{X}| < \theta < \vartheta$ implies that

$$\begin{aligned} \alpha^6 \min_{s \in \mathcal{I}(\hat{s}), d \in \mathcal{D}_s} |\mathcal{B}_{s,d}| &\geq \exp(-6n(\delta + 5\tau)) \min_{s \in \mathcal{I}(\hat{s}), d \in \mathcal{D}_s} \left| \left\{ q : v_{pq}^{ij\,n}(\hat{s}) \in \mathcal{T}^n_{[UVZ,s]\theta}(u_{ij}^n(\hat{s}), z^n) \right\} \right| \\ &> \exp\Big[n\big(\min_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s | U_{\hat{s}}) - \max_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s | U_{\hat{s}}) - 8\delta - 31\tau \big) \Big]. \end{aligned} \tag{4.28}$$

The size of τ as a function of θ, n, and other constant quantities is given by using (2.9). Similarly as in part a), the constants in this case are universal for all sets $\mathcal{I}(\hat{s})$.

By keeping k lower than $\alpha^6 \min_{s \in \mathcal{I}(\hat{s}), d \in \mathcal{D}_s} |\mathcal{B}_{s,d}|$, condition (3.42) of Lemma 3.3 is guaranteed. Therefore, it follows by using (4.28) and the definition of R_{sk} from (4.26) that

$$\frac{1}{n} \log |\mathcal{K}| > R_{\mathrm{sk}} - 8\delta.$$

This satisfies condition (3.4) of Definition 3.3, which completes the proof. □

4.2 Multi-Letter Secret-Key Capacity Formula

Next, a multi-letter SK capacity formula as a function of a finite compound DMMS and public communication rate upper bound is derived.

Similar techniques which are used for deriving the non-compound SK capacity results in [18, Theorem 17.21], [17], are used and extended to the finite compound setup.

Theorem 4.2. *Consider the SK model from Section 3.2 with a finite compound DMMS* $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$*, set of marginals* $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s}\in\hat{\mathcal{S}}}$*, and a one-way communication over a public noiseless channel with communication rate upper bound* $\Gamma \in (0, \infty]$*. Then, the SK capacity is given by*

$$C_{\mathrm{sk}}(\mathfrak{S}, \Gamma) = \lim_{n\to\infty} \frac{1}{n} \min_{\hat{s}\in\hat{\mathcal{S}}} \max_{U_{\hat{s}}, V_{\hat{s}}} \Big\{ \min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s^n|U_{\hat{s}}) - \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s^n|U_{\hat{s}}) \Big\}, \tag{4.29}$$

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

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}}^n - Y_s^n Z_s^n, \tag{4.30}$$

$$\frac{1}{n} \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}^n|Y_s^n) + \frac{1}{n} \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}^n|U_{\hat{s}}Y_s^n) \le \Gamma. \tag{4.31}$$

Proof. The achievability of the SK rate in (4.29) follows by a code-blocking argument. For this, we apply the result of Theorem 4.1 to an n-fold product of the source. This gives the achievability of the following SK rate:

$$\frac{1}{n} \min_{\hat{s}\in\hat{\mathcal{S}}} \max_{U_{\hat{s}}, V_{\hat{s}}} \Big\{ \min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s^n|U_{\hat{s}}) - \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s^n|U_{\hat{s}}) \Big\}, \tag{4.32}$$

where $n \in \mathbb{N}$ and the outer max is taken over all RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ satisfying conditions (4.30) and (4.31).

To show the achievability of (4.29), we have to show that the limit exists. Similarly as in [36], the proof follows by using the Fekete's lemma [56] which states that if a sequence a_n is superadditive i.e. $a_{n+m} \ge a_n + a_m$, then $\lim_{n\to\infty} a_n/n$ exists.
To this end, define the sequences

$$\begin{aligned} a_n(\hat{s}) &:= \max_{U_{\hat{s}}, V_{\hat{s}}} \Big\{ \min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s^n|U_{\hat{s}}) - \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s^n|U_{\hat{s}}) \Big\}, \\ a_n &:= \min_{\hat{s}\in\hat{\mathcal{S}}} a_n(\hat{s}). \end{aligned} \tag{4.33}$$

Take two arbitrary independent Markov chains

$$U_{\hat{s},1} - V_{\hat{s},1} - X_{\hat{s},1}^n - Y_{s,1}^n Z_{s,1}^n\,, \tag{4.34}$$

$$U_{\hat{s},2} - V_{\hat{s},2} - X_{\hat{s},2}^m - Y_{s,2}^m Z_{s,2}^m\,, \tag{4.35}$$

such that $m, n \in \mathbb{N}$ and it holds:

$$\begin{aligned} \hat{U}_{\hat{s}} &:= (U_{\hat{s},1}, U_{\hat{s},2}), \\ \hat{V}_{\hat{s}} &:= (V_{\hat{s},1}, V_{\hat{s},2}), \\ X_{\hat{s}}^{n+m} &:= (X_{\hat{s},1}^n, X_{\hat{s},2}^m), \\ Y_s^{n+m} &:= (Y_{s,1}^n, Y_{s,2}^m), \\ Z_s^{n+m} &:= (Z_{s,1}^n, Z_{s,2}^m). \end{aligned}$$

As the two Markov chains are independent, any RV from (4.34) is independent of all RVs in (4.35) and vice versa. Therefore, the Markov chain $\hat{U}_{\hat{s}} - \hat{V}_{\hat{s}} - X_{\hat{s}}^{n+m} - Y_s^{n+m} Z_s^{n+m}$ holds and

$$\begin{aligned} a_{n+m}(\hat{s}) &= \max_{\hat{U}_{\hat{s}}, \hat{V}_{\hat{s}}} \Big\{ \min_{s \in \mathcal{I}(\hat{s})} I(\hat{V}_{\hat{s}}; Y_s^{n+m} | \hat{U}_{\hat{s}}) - \max_{s \in \mathcal{I}(\hat{s})} I(\hat{V}_{\hat{s}}; Z_s^{n+m} | \hat{U}_{\hat{s}}) \Big\} \\ &\geq \max_{\hat{U}_{\hat{s}}, \hat{V}_{\hat{s}}} \Big\{ \min_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s},1}; Y_{s,1}^n | U_{\hat{s},1}) - \max_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s},1}; Z_{s,1}^n | U_{\hat{s},1}) \\ &\qquad + \min_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s},2}; Y_{s,2}^m | U_{\hat{s},2}) - \max_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s},2}; Z_{s,2}^m | U_{\hat{s},2}) \Big\} \\ &= a_n(\hat{s}) + a_m(\hat{s}). \end{aligned}$$

Thus, by taking the minimum from both sides of this inequality and using the definition in (4.33), it follows that

$$\begin{aligned} a_{n+m} &\geq \min_{\hat{s} \in \hat{\mathcal{S}}} \{a_n(\hat{s}) + a_m(\hat{s})\} \\ &\geq \min_{\hat{s} \in \hat{\mathcal{S}}} \{a_n(\hat{s})\} + \min_{\hat{s} \in \hat{\mathcal{S}}} \{a_m(\hat{s})\} \\ &= a_n + a_m. \end{aligned}$$

Therefore, the limit in (4.29) exists and is achievable.

For the converse proof, let $R_{\mathrm{sk}} > 0$ be an achievable SK rate and $s \in \mathcal{S}$ be given. Assume that $\hat{s} \in \hat{\mathcal{S}}$ is the marginal index and thus $s \in \mathcal{I}(\hat{s})$. Alice sends a message $f_{\mathrm{c}}(X_{\hat{s}}^n)$ to Bob over the public channel and generates an SK represented by RV $K = K_{\mathrm{A}}$.

It holds by using Definition 3.2 and condition (3.6) of Definition 3.3 that for all $\delta > 0$ and n sufficiently large

$$\frac{1}{n} \log |\mathcal{K}| - \frac{1}{n} \min_{s \in \mathcal{I}(\hat{s})} H\big(K_{\mathrm{A}} | Z_s^n, f_{\mathrm{c}}(X_{\hat{s}}^n)\big) < \delta. \tag{4.36}$$

Moreover, by using Lemma 2.8 (Fano's inequality), it holds that

$$\frac{1}{n} \max_{s \in \mathcal{I}(\hat{s})} H\big(K_{\mathrm{A}} | Y_s^n, f_{\mathrm{c}}(X_{\hat{s}}^n)\big) < \frac{1}{n} \delta \log |\mathcal{K}| + \frac{1}{n}. \tag{4.37}$$

By adding (4.36) and (4.37), it follows that

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

Thus, by using condition (3.4) of Definition 3.3 and taking

$$\epsilon := \delta/(1-\delta) + \frac{1}{n(1-\delta)} + \delta,$$

it follows that for n sufficiently large

$$\begin{aligned} R_{\mathrm{sk}} &< \frac{1}{n}\log|\mathcal{K}| + \delta \\ &\leq \frac{1}{1-\delta}\cdot\frac{1}{n}\Big[\min_{s\in\mathcal{I}(\hat{s})} I\big(K_{\mathrm{A}};Y_s^n|f_{\mathrm{c}}(X_{\hat{s}}^n)\big) - \max_{s\in\mathcal{I}(\hat{s})} I\big(K_{\mathrm{A}};Z_s^n|f_{\mathrm{c}}(X_{\hat{s}}^n)\big)\Big] + \epsilon. \end{aligned} \tag{4.38}$$

The last inequality follows by adding and subtracting the term $H(K_{\mathrm{A}}|f_{\mathrm{c}}(X_{\hat{s}}^n))$. Next, define RVs

$$\begin{aligned} U_{\hat{s}} &:= f_{\mathrm{c}}(X_{\hat{s}}^n) \\ V_{\hat{s}} &:= (f_{\mathrm{c}}(X_{\hat{s}}^n), K_{\mathrm{A}}). \end{aligned} \tag{4.39}$$

It holds 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 holds that

$$I(Y_s^n Z_s^n; U_{\hat{s}}V_{\hat{s}}|X_{\hat{s}}^n) = 0.$$

This proves that the Markov chains in (4.30) are valid.

Next, we show that the RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ satisfy also the communication rate condition (4.31). It holds that

$$\frac{1}{n}\max_{s\in\mathcal{I}(\hat{s})} H(f_{\mathrm{c}}(X_{\hat{s}}^n)|Y_s^n) \leq \frac{1}{n}\log\|f_{\mathrm{c}}\|. \tag{4.40}$$

By adding (4.37) and (4.40), it follows that

$$\begin{aligned} \frac{1}{n}\max_{s\in\mathcal{I}(\hat{s})} H(f_{\mathrm{c}}(X_{\hat{s}}^n)|Y_s^n) + \frac{1}{n}\max_{s\in\mathcal{I}(\hat{s})} H\big(K_{\mathrm{A}}|Y_s^n, f_{\mathrm{c}}(X_{\hat{s}}^n)\big) \\ < \frac{1}{n}\log\|f_{\mathrm{c}}\| + \frac{1}{n}\delta\log|\mathcal{K}| + \frac{1}{n}. \end{aligned} \tag{4.41}$$

Furthermore, it holds trivially that

$$\begin{aligned} H(f_{\mathrm{c}}(X_{\hat{s}}^n)|X_{\hat{s}}^n, Y_s^n) &= H\big(K_{\mathrm{A}}, f_{\mathrm{c}}(X_{\hat{s}}^n)|X_{\hat{s}}^n, Y_s^n, f_{\mathrm{c}}(X_{\hat{s}}^n)\big) \\ &= 0. \end{aligned}$$

By subtracting these terms from the values inside the maximums in (4.41), it follows that

$$\begin{aligned} \frac{1}{n}\max_{s\in\mathcal{I}(\hat{s})} I(X_{\hat{s}}^n; f_{\mathrm{c}}(X_{\hat{s}}^n)|Y_s^n) + \frac{1}{n}\max_{s\in\mathcal{I}(\hat{s})} I\big(X_{\hat{s}}^n; K_{\mathrm{A}}, f_{\mathrm{c}}(X_{\hat{s}}^n)|Y_s^n, f_{\mathrm{c}}(X_{\hat{s}}^n)\big) \\ < \frac{1}{n}\log\|f_{\mathrm{c}}\| + \frac{1}{n}\delta\log|\mathcal{K}| + \frac{1}{n}. \end{aligned} \tag{4.42}$$

By using condition (3.3) and inserting the values of $U_{\hat{s}}$ and $V_{\hat{s}}$ from (4.39) into (4.42), it follows for δ sufficiently small and n large enough that condition (4.31) also holds.

Finally, by taking the maximum with respect to $U_{\hat{s}}$ and $V_{\hat{s}}$ and the minimum with respect to $\hat{s} \in \hat{\mathcal{S}}$ in (4.38), it follows for $\delta > 0$ sufficiently small and n large enough that

$$R_{\mathrm{sk}} \leq \frac{1}{n} \min_{\hat{s}\in\hat{S}} \max_{U_{\hat{s}},V_{\hat{s}}} \Big[\min_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s^n|U_{\hat{s}}) - \max_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s^n|U_{\hat{s}}) \Big],$$

which completes the proof. □

4.3 Alphabet Size of Auxiliary Random Variables

In this section, some upper bounds for the range size of the auxiliary RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ in Theorem 4.1 are derived. The upper bounds in this case are given for each $\hat{s} \in \hat{\mathcal{S}}$ by:

$$|\mathcal{U}_{\hat{s}}| \leq |\mathcal{X}| + 4|\mathcal{I}(\hat{s})| - 1, \tag{4.43}$$

$$|\mathcal{V}'_{\hat{s}}| \leq |\mathcal{X}| + 2|\mathcal{I}(\hat{s})|, \tag{4.44}$$

with $V_{\hat{s}} := (U_{\hat{s}}, V'_{\hat{s}})$, and $\mathcal{U}_{\hat{s}}$ and $\mathcal{V}'_{\hat{s}}$ being the alphabets of $U_{\hat{s}}$ and $V'_{\hat{s}}$ respectively. To achieve the SK rate in Theorem 4.1 for a finite $\mathfrak{S}$, it suffices that the alphabet sizes of $U_{\hat{s}}$ and $V'_{\hat{s}}$ are at most the upper bounds in (4.43) and (4.44) respectively. Although these upper bounds are not necessarily optimal, but they show how the size of compound sources might influence the alphabet size of auxiliary RVs.

To derive (4.43) and (4.44), similarly as in [18, Theorem 17.13], we show that any $U_{\hat{s}}$ and $V_{\hat{s}}$ which satisfy the Markov conditions in (4.2) but not conditions (4.43) and (4.44), can be replaced by other RVs $\bar{U}_{\hat{s}}$ and $\bar{V}_{\hat{s}}$ satisfying these conditions. We also show that this replacement does not affect the following quantities and thus the SK capacity remains unchanged:

$$\begin{aligned} &I(U_{\hat{s}}; X_{\hat{s}}|Y_s), \qquad I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_s), \\ &I(V_{\hat{s}}; Y_s|U_{\hat{s}}), \qquad I(V_{\hat{s}}; Z_s|U_{\hat{s}}). \end{aligned} \tag{4.45}$$

To this end, define the following $|\mathcal{X}| + 4|\mathcal{I}(\hat{s})| - 1$ real valued continuous functions for all $P \in \mathcal{P}(\mathcal{V}_{\hat{s}})$ as follows:
a) For every $x \in \mathcal{X}$ but one,

$$f_x(P) := \sum_{v\in\mathcal{V}_{\hat{s}}} P(v)W(x|v), \tag{4.46}$$

where $W : \mathcal{V}_{\hat{s}} \to \mathcal{P}(\mathcal{X})$ is a stochastic matrix with $W := P_{X_{\hat{s}}|V_{\hat{s}}}$.
b) For all $s \in \mathcal{I}(\hat{s})$,

$$\begin{aligned} f_s^{(1)}(P) &:= H(PW_s^{(1)}) - H(PW), \\ f_s^{(2)}(P) &:= \sum_{v\in\mathcal{V}_{\hat{s}}} P(v)\Big[H\big(W_s^{(1)}(\cdot|v)\big) - H\big(W(\cdot|v)\big)\Big], \\ f_s^{(3)}(P) &:= H(PW_s^{(1)}) - \sum_{v\in\mathcal{V}_{\hat{s}}} P(v)H\big(W_s^{(1)}(\cdot|v)\big), \\ f_s^{(4)}(P) &:= H(PW_s^{(2)}) - \sum_{v\in\mathcal{V}_{\hat{s}}} P(v)H\big(W_s^{(2)}(\cdot|v)\big), \end{aligned}$$

where $W_s^{(1)}$ and $W_s^{(2)}$ are stochastic matrices as follows

$$W_s^{(1)} : \mathcal{V}_{\hat{s}} \to \mathcal{P}(\mathcal{Y}), \quad W_s^{(1)} := P_{Y_s|V_{\hat{s}}},$$
$$W_s^{(2)} : \mathcal{V}_{\hat{s}} \to \mathcal{P}(\mathcal{Z}), \quad W_s^{(2)} := P_{Z_s|V_{\hat{s}}}.$$

The definitions in a) and b) imply that

$$P_{X_{\hat{s}}}(x) = \sum_{u \in \mathcal{U}_{\hat{s}}} P_{U_{\hat{s}}}(u) f_x(P_{V_{\hat{s}}|U_{\hat{s}}}(\cdot|u)), \tag{4.47}$$

$$I(U_{\hat{s}}; X_{\hat{s}}|Y_s) = H(X_{\hat{s}}) - H(Y_s) + \sum_{u \in \mathcal{U}_{\hat{s}}} P_{U_{\hat{s}}}(u) f_s^{(1)}(P_{V_{\hat{s}}|U_{\hat{s}}}(\cdot|u)), \tag{4.48}$$

$$I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_s) = \sum_{u \in \mathcal{U}_{\hat{s}}} P_{U_{\hat{s}}}(u) \big[f_s^{(2)}(P_{V_{\hat{s}}|U_{\hat{s}}}(\cdot|u)) - f_s^{(1)}(P_{V_{\hat{s}}|U_{\hat{s}}}(\cdot|u)) \big], \tag{4.49}$$

$$I(V_{\hat{s}}; Y_s|U_{\hat{s}}) = \sum_{u \in \mathcal{U}_{\hat{s}}} P_{U_{\hat{s}}}(u) f_s^{(3)}(P_{V_{\hat{s}}|U_{\hat{s}}}(\cdot|u)), \tag{4.50}$$

$$I(V_{\hat{s}}; Z_s|U_{\hat{s}}) = \sum_{u \in \mathcal{U}_{\hat{s}}} P_{U_{\hat{s}}}(u) f_s^{(4)}(P_{V_{\hat{s}}|U_{\hat{s}}}(\cdot|u)). \tag{4.51}$$

Next, by applying Lemma 2.3 (Support) to (4.47)-(4.51), it follows that there exist PDs $P_{\bar{u}} \in \mathcal{P}(\mathcal{V}_{\hat{s}}), \bar{u} \in \bar{\mathcal{U}}_{\hat{s}}$ such that

$$P_{X_{\hat{s}}}(x) = \sum_{\bar{u} \in \bar{\mathcal{U}}_{\hat{s}}} \alpha_{\bar{u}} f_x(P_{\bar{u}}), \tag{4.52}$$

$$I(U_{\hat{s}}; X_{\hat{s}}|Y_s) = H(X_{\hat{s}}) - H(Y_s) + \sum_{\bar{u} \in \bar{\mathcal{U}}_{\hat{s}}} \alpha_{\bar{u}} f_s^{(1)}(P_{\bar{u}}), \tag{4.53}$$

$$I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_s) = \sum_{\bar{u} \in \bar{\mathcal{U}}_{\hat{s}}} \alpha_{\bar{u}} \big[f_s^{(2)}(P_{\bar{u}}) - f_s^{(1)}(P_{\bar{u}}) \big], \tag{4.54}$$

$$I(V_{\hat{s}}; Y_s|U_{\hat{s}}) = \sum_{\bar{u} \in \bar{\mathcal{U}}_{\hat{s}}} \alpha_{\bar{u}} f_s^{(3)}(P_{\bar{u}}), \tag{4.55}$$

$$I(V_{\hat{s}}; Z_s|U_{\hat{s}}) = \sum_{\bar{u} \in \bar{\mathcal{U}}_{\hat{s}}} \alpha_{\bar{u}} f_s^{(4)}(P_{\bar{u}}), \tag{4.56}$$

where $\sum_{\bar{u} \in \bar{\mathcal{U}}_{\hat{s}}} \alpha_{\bar{u}} = 1$, $\alpha_{\bar{u}} \geq 0$, and

$$\bar{\mathcal{U}}_{\hat{s}} := \{1, 2, \cdots, |\mathcal{X}| + 4|\mathcal{I}(\hat{s})| - 1\}. \tag{4.57}$$

By using condition (4.52), we may define the RV $\bar{U}_{\hat{s}}$ with values in $\bar{\mathcal{U}}_{\hat{s}}$, and RV $\tilde{V}_{\hat{s}}$ with values in $\mathcal{V}_{\hat{s}}$ such that

$$P_{\bar{U}\tilde{V}XYZ,s}(\bar{u}, v, x, y, z) = \alpha_{\bar{u}} P_{\bar{u}}(v) W(x|v) P_{YZ,s|X_{\hat{s}}}(y, z|x),$$

where $(\bar{u}, v, x, y, z) \in \bar{\mathcal{U}}_{\hat{s}} \times \mathcal{V}_{\hat{s}} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. This definition implies that the Markov chains

$$\bar{U}_{\hat{s}} - \tilde{V}_{\hat{s}} - X_{\hat{s}} - Y_s Z_s$$

hold. The equations (4.53)-(4.56) imply that the values in (4.45) do not change by replacing $U_{\hat{s}}$ by $\bar{U}_{\hat{s}}$, and $V_{\hat{s}}$ by $\tilde{V}_{\hat{s}}$.

Again, Similarly as in [18, Theorem 17.13], we fix $\bar{U}_{\hat{s}}$ and define in the following an RV $V'_{\hat{s}}$. We replace then $\tilde{V}_{\hat{s}}$ by RV $\bar{V}_{\hat{s}} := (\bar{U}_{\hat{s}}, V'_{\hat{s}})$. Finally, we show that this replacement leaves the values in (4.45) unchanged.

For this, define the following $|\mathcal{X}| + 2|\mathcal{I}(\hat{s})|$ real valued continuous functions for all $P \in \mathcal{P}(\mathcal{X})$

$$g_x(P) := P(x), \quad \text{for every } x \in \mathcal{X} \text{ but one}, \tag{4.58}$$

$$g(P) := H(P), \tag{4.59}$$

$$g_s^{(1)}(P) := H(P\hat{W}_s^{(1)}), \quad g_s^{(2)}(P) := H(P\hat{W}_s^{(2)}), \tag{4.60}$$

where $\hat{W}_s^{(1)}$ and $\hat{W}_s^{(2)}$ are stochastic matrices as follows

$$\hat{W}_s^{(1)} : \mathcal{X} \to \mathcal{P}(\mathcal{Y}), \quad \hat{W}_s^{(1)} := P_{Y_s|X_{\hat{s}}},$$
$$\hat{W}_s^{(2)} : \mathcal{X} \to \mathcal{P}(\mathcal{Z}), \quad \hat{W}_s^{(2)} := P_{Z_s|X_{\hat{s}}}.$$

Let $\bar{U}_{\hat{s}} = \bar{u}$ be given. Similarly as in (4.47)-(4.51), the quantities

$$P_{X_{\hat{s}}|\bar{U}_{\hat{s}}}(x|\bar{u}), \quad H(X_{\hat{s}}|\bar{U}_{\hat{s}} = \bar{u}, \tilde{V}_{\hat{s}}),$$
$$H(Y_s|\bar{U}_{\hat{s}} = \bar{u}, \tilde{V}_{\hat{s}}), \quad \text{and} \quad H(Z_s|\bar{U}_{\hat{s}} = \bar{u}, \tilde{V}_{\hat{s}})$$

can be expressed as average values of the functions in (4.58)-(4.60), at points $P_{X_{\hat{s}}|V_{\hat{s}}}(\cdot|v)$ with weights $P_{\bar{u}}(v)$. Again applying Lemma 2.3 to these values implies that there exist PDs $P_{\bar{u}v'} \in \mathcal{P}(\mathcal{X})$, with $v' \in \mathcal{V}'_{\hat{s}}$ such that

$$P_{X_{\hat{s}}|\bar{U}_{\hat{s}}}(x|\bar{u}) = \sum_{v' \in \mathcal{V}'_{\hat{s}}} \beta_{\bar{u}v'} g_x(P_{\bar{u}v'}), \tag{4.61}$$

$$H(X_{\hat{s}}|\bar{U}_{\hat{s}} = \bar{u}, \tilde{V}_{\hat{s}}) = \sum_{v' \in \mathcal{V}'_{\hat{s}}} \beta_{\bar{u}v'} g(P_{\bar{u}v'}), \tag{4.62}$$

$$H(Y_s|\bar{U}_{\hat{s}} = \bar{u}, \tilde{V}_{\hat{s}}) = \sum_{v' \in \mathcal{V}'_{\hat{s}}} \beta_{\bar{u}v'} g_s^{(1)}(P_{\bar{u}v'}), \tag{4.63}$$

$$H(Z_s|\bar{U}_{\hat{s}} = \bar{u}, \tilde{V}_{\hat{s}}) = \sum_{v' \in \mathcal{V}'_{\hat{s}}} \beta_{\bar{u}v'} g_s^{(2)}(P_{\bar{u}v'}), \tag{4.64}$$

where $\sum_{v' \in \mathcal{V}'_{\hat{s}}} \beta_{\bar{u}v'} = 1$, $\beta_{\bar{u}v'} \geq 0$, and

$$\mathcal{V}'_{\hat{s}} := \{1, 2, \cdots, |\mathcal{X}| + 2|\mathcal{I}(\hat{s})|\}. \tag{4.65}$$

By using condition (4.61), we may define the RV $V'_{\hat{s}}$ such that

$$P_{V'XYZ,s|\bar{U}_{\hat{s}}}(v', x, y, z|\bar{u}) = \beta_{\bar{u}v'} P_{\bar{u}v'}(x) P_{YZ,s|X_{\hat{s}}}(y, z|x),$$

where $(\bar{u}, v', x, y, z) \in \bar{\mathcal{U}}_{\hat{s}} \times \mathcal{V}'_{\hat{s}} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. This definition implies that the Markov chains

$$\bar{U}_{\hat{s}} - \bar{U}_{\hat{s}} V'_{\hat{s}} - X_{\hat{s}} - Y_s Z_s$$

hold. Moreover, the equations (4.62)-(4.64) imply that the values in (4.45) whose RVs were replaced by $\bar{U}_{\hat{s}}$ and $\tilde{V}_{\hat{s}}$ remain again unchanged if $\tilde{V}_{\hat{s}}$ is replaced by $\bar{V}_{\hat{s}} = (\bar{U}_{\hat{s}}, V'_{\hat{s}})$. The upper bounds in (4.43) and (4.44) follow by using (4.57) and (4.65).

5 Secret-Key Capacity of Infinite Compound Sources

In this chapter, the SK capacity results for arbitrary (possibly infinite) compound sources are presented, where the SK model from Chapter 3, Section 3.2 is used. It is assumed that the set of marginals is finite.

5.1 Single-Letter Secret-Key Capacity Formula (Degraded Compound Sources)

In the following, a single-letter SK capacity formula is given for a degraded compound source with an arbitrary set of source states $\mathfrak{S}$ which might be infinite. The set of marginals $\hat{\mathfrak{S}}$ is assumed to be finite.

Theorem 5.1. *Consider the SK model from Section 3.2 with an arbitrary (possibly infinite) compound DMMS $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$, a finite set of marginals $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s}\in\hat{\mathcal{S}}}$, and a one-way communication over a public noiseless channel. If the following Markov chains are satisfied (i.e. the compound source is degraded),*

$$\forall \hat{s} \in \hat{\mathcal{S}},\ \forall r,t \in \mathcal{I}(\hat{s}), \quad X_{\hat{s}} - Y_r - Z_t, \tag{5.1}$$

then, the SK capacity is given by

$$C_{\mathrm{sk}}(\mathfrak{S}) = \min_{\hat{s}\in\hat{\mathcal{S}}} \Big\{ \inf_{r\in\mathcal{I}(\hat{s})} I(X_{\hat{s}};Y_r) - \sup_{t\in\mathcal{I}(\hat{s})} I(X_{\hat{s}};Z_t) \Big\}. \tag{5.2}$$

Proof. The achievability of the SK rate in (5.2) for a finite source follows directly from the special case of Theorem 4.1. By taking RV $U_{\hat{s}}$ to be $X_{\hat{s}}$ in part a) of the proof of Theorem 4.1 in (4.4), it follows that the given rate in (5.2) for a finite $\mathfrak{S}$ is achievable.

To show the achievability of the SK rate in (5.2) for an infinite $\mathfrak{S}$ and a finite set of marginals $\hat{\mathfrak{S}}$, let $\hat{s} \in \hat{\mathcal{S}}$ be given and consider the following infinite class of stochastic matrices:

$$\{P_{YZ,s|X_{\hat{s}}} : \mathcal{X} \to \mathcal{P}(\mathcal{Y}\times\mathcal{Z})\}_{s\in\mathcal{I}(\hat{s})}. \tag{5.3}$$

The approximation follows Similarly as in [38]. Lemma 2.6 implies by making l dependent on codeword length n and taking $l = n^3$ that for any $n^3 > 2|\mathcal{Y}\times\mathcal{Z}|^2$, there exists a finite set of stochastic matrices

$$\{W_{s'} : \mathcal{X} \to \mathcal{P}(\mathcal{Y}\times\mathcal{Z})\}_{s'\in\mathcal{I}_n(\hat{s})}, \quad \text{with} \quad |\mathcal{I}_n(\hat{s})| \le (n^3+1)^{|\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}|}, \tag{5.4}$$

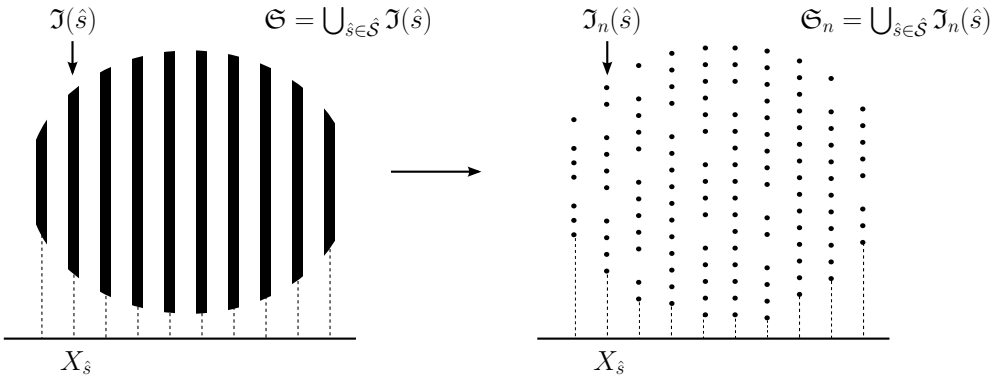

Figure 5.1: Informal illustration of an example of a compound source approximation

which approximates the one in (5.3) such that

$$\forall s \in \mathcal{I}(\hat{s}),\ \ \exists s' \in \mathcal{I}_n(\hat{s}),\ \ \forall (x,y,z) \in \mathcal{X}\times\mathcal{Y}\times\mathcal{Z},$$

$$\left|P_{YZ,s|X_{\hat{s}}}(y,z|x) - W_{s'}(y,z|x)\right| \le \frac{1}{n^3}|\mathcal{Y}\times\mathcal{Z}|, \tag{5.5}$$

$$P_{YZ,s|X_{\hat{s}}}(y,z|x) \le \mathrm{e}^{\frac{2|\mathcal{Y}\times\mathcal{Z}|^2}{n^3}} W_{s'}(y,z|x). \tag{5.6}$$

As $\hat{s} \in \hat{\mathcal{S}}$ was chosen arbitrarily, we may repeat this procedure for all $\hat{s} \in \hat{\mathcal{S}}$ and define the following finite source

$$\mathfrak{S}_n := \bigcup_{\hat{s}\in\hat{\mathcal{S}}} \mathfrak{I}_n(\hat{s}) = \bigcup_{\hat{s}\in\hat{\mathcal{S}}} \left\{XYZ, s'\right\}_{s'\in\mathcal{I}_n(\hat{s})}, \tag{5.7}$$

where for all $\hat{s} \in \hat{\mathcal{S}}, s' \in \mathcal{I}_n(\hat{s})$, and $(x,y,z) \in \mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$, the PDs of RVs in $\mathfrak{S}_n$ are given by

$$P_{XYZ,s'}(x,y,z) := P_{X_{\hat{s}}}(x) W_{s'}(y,z|x).$$

An example of this procedure is depicted informally in Figure 5.1. The left diagram in the figure represents an example of an infinite compound source $\mathfrak{S}$ which was also given in Chapter 3, Figure 3.1. The right diagram in Figure 5.1 represents an example of a finite compound source $\mathfrak{S}_n$ which approximates $\mathfrak{S}$.

The assumption $l = n^3$ in Lemma 2.6 is necessary as the approximation should become more precise by increasing the codeword length (number of source uses) n. Define for a marginal index $\hat{s}$ and indices $r, t \in \mathcal{I}(\hat{s})$,

$$f(\hat{s}, r, t) := I(X_{\hat{s}}; Y_r) - I(X_{\hat{s}}; Z_t). \tag{5.8}$$

Since the approximating compound sources $\mathfrak{S}_n$ given by (5.7) are finite, Theorem 4.1 implies the achievability of the following SK rates for each $\mathfrak{S}_n$ as a fixed source:

$$\min_{\hat{s}\in\hat{\mathcal{S}}} \min_{r',t'\in\mathcal{I}_n(\hat{s})} f(\hat{s}, r', t'). \tag{5.9}$$

On the other hand, the sets $\mathfrak{S}_n$ and $\mathfrak{I}_n(\hat{s})$ are dependent on the codeword length n. This means that by increasing n, both $\mathfrak{S}_n$ and $\mathfrak{I}_n(\hat{s})$ change too. Therefore, the proof of Theorem 4.1 as well as its corresponding lemmas should be slightly adjusted such that the constants are always universal for all $\mathfrak{I}_n(\hat{s})$ and n.

To this end, first of all, the constant τ which is defined in the proof of Theorem 4.1, should be chosen such that it remains fixed by increasing n. However, if $\mathfrak{S}_n$ is applied in Theorem 4.1, then the size of τ is dependent on the compound sources $\mathfrak{I}_n(\hat{s})$ and thus on n in assumptions (4.14) and (4.24). To find a universal τ, we must show that the term

$$\max_{s' \in \mathcal{I}_n(\hat{s})} I(U_{\hat{s}}; Z_{s'}),$$

which appears on the right hand sides of (4.14) and (4.24) approaches

$$\sup_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Z_s), \tag{5.10}$$

as n goes to infinity. The proof follows similarly as in (5.46) by using (5.5). Therefore, as the infinite compound source $\mathfrak{I}(\hat{s})$ in (5.10) is fixed, τ can be chosen independent of n. Similar argument holds for choosing δ in (4.22).

Moreover, in deriving the upper bound (4.16) in the proof of Theorem 4.1, the size of constant σ was given by (2.2) for a given τ and each element of the compound source. However, if $\mathfrak{S}_n$ is applied in Theorem 4.1, then the size of σ is dependent on the compound sources $\mathfrak{I}_n(\hat{s})$ and thus on n. To find a universal σ, we must first show that the term

$$\min_{s' \in \mathcal{I}_n(\hat{s})} \mu_{X_{\hat{s}} Z_{s'}}$$

approaches

$$\inf_{s \in \mathcal{I}(\hat{s})} \mu_{X_{\hat{s}} Z_s},$$

as n goes to infinity. The proof follows by using (5.12). Therefore, σ can be made sufficiently small such that

$$0 < \sigma < \inf_{s \in \mathcal{I}(\hat{s})} \mu_{X_{\hat{s}} Z_s}. \tag{5.11}$$

Finally, as $\mathfrak{I}(\hat{s})$ in (5.11) is fixed, σ would be independent of n. The upper bound in (5.11) is non-zero, because the marginal $X_{\hat{s}}$ in $\mu_{X_{\hat{s}} Z_s}$ is fixed.

In the following, it is shown that for some universal $n_0 > 0$ and all $n \geq n_0$, the SK generation protocols which guarantee the achievability of the SK rates in (5.9) for finite compound sources $\mathfrak{S}_n$, also guarantee the achievability of the SK rate given in (5.2) for the infinite compound source $\mathfrak{S}$. To this end, all conditions of Definition 3.3 will be verified in the following. For each $\mathfrak{S}_n$, the constant n_0 from Definition 3.3 depends only linearly on $\log |\mathcal{I}_n(\hat{s})|$ (cf. proof of Theorem 4.1). Furthermore, we showed in (5.10) and (5.11) that the constants which are used in the SK generation protocols are universal for all $\mathfrak{I}_n(\hat{s})$. Therefore, a universal n_0 can be determined which works for all $\mathfrak{S}_n$ with $n \geq n_0$. This will be explained in more detail in the rest of this proof.

Condition (3.4) *(SK Rate):* It follows by using (5.5) that for all $r, t \in \mathcal{I}(\hat{s})$ and their corresponding indices $r', t' \in \mathcal{I}_n(\hat{s})$,

$$\begin{aligned}\gamma_1 &:= \frac{1}{2}\|P_{XY,r} - P_{XY,r'}\| \leq \frac{1}{2n^3}|\mathcal{Y} \times \mathcal{Z}|^2,\\ \gamma_2 &:= \frac{1}{2}\|P_{XZ,t} - P_{XZ,t'}\| \leq \frac{1}{2n^3}|\mathcal{Y} \times \mathcal{Z}|^2.\end{aligned} \tag{5.12}$$

Since $\gamma_1 \leq 1 - (|\mathcal{X}||\mathcal{Y}|)^{-1}$ and $\gamma_2 \leq 1 - (|\mathcal{X}||\mathcal{Z}|)^{-1}$ hold for n large enough, it follows by Lemma 2.9 that

$$|I(X_{\hat{s}}; Y_r) - I(X_{\hat{s}}; Y_{r'})| < 3\gamma_1 \log(|\mathcal{X} \times \mathcal{Y}|) + 3h(\gamma_1), \tag{5.13}$$

$$|I(X_{\hat{s}}; Z_t) - I(X_{\hat{s}}; Z_{t'})| < 3\gamma_2 \log(|\mathcal{X} \times \mathcal{Z}|) + 3h(\gamma_2). \tag{5.14}$$

By using (5.8), (5.13), and (5.14), it holds that

$$\begin{aligned}|f(\hat{s}, r, t) - f(\hat{s}, r', t')| &\leq |I(X_{\hat{s}}; Y_r) - I(X_{\hat{s}}; Y_{r'})| + |I(X_{\hat{s}}; Z_t) - I(X_{\hat{s}}; Z_{t'})|\\ &< \frac{3|\mathcal{Y} \times \mathcal{Z}|^2}{2n^3} \log(|\mathcal{X}^2 \times \mathcal{Y} \times \mathcal{Z}|) + 6h\Big(\frac{|\mathcal{Y} \times \mathcal{Z}|^2}{2n^3}\Big).\end{aligned} \tag{5.15}$$

Let $\epsilon > 0$ be given. To show that the SK rate of this protocol is close to the one which is given in (5.2), it is sufficient to show that

$$\Big|\min_{\hat{s} \in \hat{\mathcal{S}}} \inf_{r,t \in \mathcal{I}(\hat{s})} f(\hat{s}, r, t) - \min_{\hat{s} \in \hat{\mathcal{S}}} \min_{r', t' \in \mathcal{I}_n(\hat{s})} f(\hat{s}, r', t')\Big| < \epsilon. \tag{5.16}$$

Similarly as in [57, 58], let $\tau > 0$ be given. There exist then $r_0, t_0 \in \mathcal{I}(\hat{s})$ such that

$$\begin{aligned}\inf_{r,t \in \mathcal{I}(\hat{s})} f(\hat{s}, r, t) &> f(\hat{s}, r_0, t_0) - \tau\\ &> \min_{r', t' \in \mathcal{I}_n(\hat{s})} f(\hat{s}, r', t') - \epsilon,\end{aligned}$$

where the second inequality holds for τ small enough and is a result of (5.15) and the concavity of the binary entropy function. Similarly, there exist $r_0', t_0' \in \mathcal{I}_n(\hat{s})$ such that

$$\begin{aligned}\min_{r', t' \in \mathcal{I}_n(\hat{s})} f(\hat{s}, r', t') &> f(\hat{s}, r_0', t_0') - \tau\\ &> \inf_{r,t \in \mathcal{I}(\hat{s})} f\big(\hat{s}, r, t\big) - \epsilon.\end{aligned}$$

Therefore, as the index $\hat{s} \in \hat{\mathcal{S}}$ was taken arbitrarily, the relation (5.16) follows directly. This implies that condition (3.4) of Definition 3.3 is satisfied for the infinite source $\mathfrak{S}$.

Conditions (3.5) *and* (3.6) *(Reliability and Security):* For verifying condition (3.5) of Definition 3.3, the following inequality is required:

$$\begin{aligned}P_{Y_s|X_{\hat{s}}}^n(y^n|x^n) &= \prod_{i=1}^n P_{Y_s|X_{\hat{s}}}(y_i|x_i)\\ &\leq \mathrm{e}^{\frac{2|\mathcal{Y} \times \mathcal{Z}|^2}{n^2}} V_{s'}^n(y^n|x^n),\end{aligned} \tag{5.17}$$

where $V_{s'}(y|x) := \sum_{z\in\mathcal{Z}} W_{s'}(y,z|x)$ for any $x \in \mathcal{X}, y \in \mathcal{Y}$. The second inequality in (5.17) is a result of (5.6).

In order to use the finite protocol for the infinite source $\mathfrak{S}$, Alice estimates her marginal PD by hypothesis testing. Let the correct estimation decision be given by index $\hat{s}$. The result of estimation for the source $\mathfrak{S}$ is identical with the one from the source $\mathfrak{S}_n$. Moreover, assume that the encoder g and decoder $\tilde{g}$ which are given by Lemma 3.2a), are used to generate the CR. The error probability upper bound for all $s \in \mathcal{I}(\hat{s})$ is given by

$$\begin{aligned}
&\Pr\Big\{g(X_{\hat{s}}^n) \neq \tilde{g}\big(f(X_{\hat{s}}^n), \hat{S}, Y_s^n\big)\Big\} \\
&\qquad \leq P_{XY,s}^n\big(\big\{(x^n,y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : g(x^n) \neq \tilde{g}(f(x^n),\hat{s},y^n)\big\}\big) \\
&\qquad\quad + \sum_{\tilde{s}\in\hat{\mathcal{S}}-\{\hat{s}\}} P_{\hat{S}}(\tilde{s}) \Pr\Big\{g(X_{\hat{s}}^n) \neq \tilde{g}\big(f(X_{\hat{s}}^n), \hat{S}, Y_s^n\big) \big| \hat{S} = \tilde{s}\Big\} \\
&\qquad \leq \mathrm{e}^{\frac{2|\mathcal{Y}\times\mathcal{Z}|^2}{n^2}} \sum_{\substack{(x^n,y^n)\in\mathcal{X}^n\times\mathcal{Y}^n: \\ g(x^n)\neq\tilde{g}(f(x^n),\hat{s},y^n)}} P_{X_{\hat{s}}}^n(x^n) V_{s'}^n(y^n|x^n) + \exp(-nc_1)\cdot|\hat{\mathcal{S}}| \\
&\qquad \leq \mathrm{e}^{\frac{2|\mathcal{Y}\times\mathcal{Z}|^2}{n^2}} P_{XY,s'}^n\big(\big\{(x^n,y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : \\
&\qquad\qquad\qquad g(x^n) \neq \tilde{g}(f(x^n),\hat{s},y^n)\big\}\big) + \exp(-nc_1)\cdot|\hat{\mathcal{S}}| \\
&\qquad \leq \mathrm{e}^{\frac{2|\mathcal{Y}\times\mathcal{Z}|^2}{n^2}} \exp(-n\delta_0) + \exp(-nc_1)\cdot|\hat{\mathcal{S}}|,
\end{aligned}$$

for some $c_1, \delta_0 > 0$. The second inequality follows by (5.17) and (3.8). The last inequality is a result of (3.15) from Lemma 3.2a). Therefore, the error probability is exponentially small.

The probability that such a decoder exists for finite $\mathfrak{I}_n(\hat{s})$ was given partly in (3.33). As the constants τ, τ', and δ_1 in (3.33) are universal for all $\mathfrak{I}_n(\hat{s})$, it follows by using the inequality in (5.4) that a universal n_0 exists such that for all $n \geq n_0$, the probability in (3.33) is non-zero and thus such a coding scheme for finite sources $\mathfrak{S}_n$ exists. This is because the upper bound $(n^3+1)^{|\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}|}$ in (5.4) increases only polynomially with respect to n. On the other hand, the term $\exp(-n(\delta-\tau-\tau'-\delta_1))$ in (3.33) goes to zero exponentially fast, as n goes to infinity.

To show that condition (3.6) of Definition 3.3 is also guaranteed, assume again that the encoder g and decoder $\tilde{g}$ from Lemma 3.2a) are used and the correct estimation decision is $\hat{s}$. Let $\epsilon > 0$ be given. As the protocol guarantees condition (3.6) of Definition 3.3 for finite sources $\mathfrak{S}_n$, it holds by using Definitions 3.2 that for all $s' \in \mathcal{I}_n(\hat{s})$

$$\begin{aligned}
S(K_{\mathrm{A}}|Z_{s'}^n, f(X_{\hat{s}}^n), \hat{S} = \hat{s}) = \log|\mathcal{K}| - H(\kappa(g(X_{\hat{s}}^n))|\hat{S} = \hat{s}) \\
+ I(\kappa(g(X_{\hat{s}}^n)); Z_{s'}^n, f(X_{\hat{s}}^n)|\hat{S} = \hat{s}) < \epsilon, \qquad (5.18)
\end{aligned}$$

where $K_{\mathrm{A}}, \mathcal{K}$, and κ are given by Definition 3.1 and Lemma 3.3. An upper bound for (5.18) is given by (4.18) in the proof of Theorem 4.1.

Furthermore, it holds that

$$\begin{aligned}\gamma_3 :&= \frac{1}{2}\big\|P_{\kappa(g(X_{\hat{s}}^n))Z_s^n f(X_{\hat{s}}^n)|\hat{S}}(\cdot|\hat{s}) - P_{\kappa(g(X_{\hat{s}}^n))Z_{s'}^n f(X_{\hat{s}}^n)|\hat{S}}(\cdot|\hat{s})\big\| \\ &\leq \frac{1}{2-2\exp(-nc_0)}\big\|P_{\kappa(g(X_{\hat{s}}^n))Z_s^n f_{\mathrm{c}}(X_{\hat{s}}^n)} - P_{\kappa(g(X_{\hat{s}}^n))Z_{s'}^n f_{\mathrm{c}}(X_{\hat{s}}^n)}\big\| \\ &\leq \frac{n\big\|P_{XZ,s} - P_{XZ,s'}\big\|}{2-2\exp(-nc_0)} \\ &\leq \frac{|\mathcal{Y}\times\mathcal{Z}|^2}{2n^2(1-\exp(-nc_0))}\,, \end{aligned} \tag{5.19}$$

where RVs $f(X_{\hat{s}})$ with alphabet $\mathcal{I}$ and $f_{\mathrm{c}}(X_{\hat{s}}^n) = (f(X_{\hat{s}}^n), \hat{S})$ are given by Lemma 3.2a) and in the proof of Theorem 4.1. The first inequality follows by (3.7) and the second one by the fact that no mapping of the PDs increases their 1-norm distance. The last inequality is a result of (5.5). Since $\gamma_3 \leq 1 - \frac{1}{|\mathcal{K}\times\mathcal{Z}^n\times\mathcal{I}|}$, then Lemma 2.9 implies by using (5.19) that

$$\begin{aligned}\big|I(\kappa(g(X_{\hat{s}}^n)); Z_s^n, f(X_{\hat{s}}^n)|\hat{S}=\hat{s}) - I(\kappa(g(X_{\hat{s}}^n)); Z_{s'}^n, f(X_{\hat{s}}^n)|\hat{S}=\hat{s})\big| \\ &< 3\gamma_3\log(|\mathcal{K}\times\mathcal{Z}^n\times\mathcal{I}|) + 3h(\gamma_3) \\ &\leq \frac{3|\mathcal{Y}\times\mathcal{Z}|^2c_2}{2n(1-\exp(-nc_0))} + 3h\Big(\frac{|\mathcal{Y}\times\mathcal{Z}|^2}{2n^2(1-\exp(-nc_0))}\Big), \end{aligned} \tag{5.20}$$

for some $c_2 > 0$ and n sufficiently large. The second inequality follows by the fact that the argument of the log function does not increase faster than exponentially. Therefore, it follows by using (3.7), (3.8), (5.18), and (5.20) that for all $s \in \mathcal{I}(\hat{s})$

$$\begin{aligned} S(K_{\mathrm{A}}|Z_s^n, f(X_{\hat{s}}^n), \hat{S}) &= P_{\hat{S}}(\hat{s})S(K_{\mathrm{A}}|Z_s^n, f(X_{\hat{s}}^n), \hat{S}=\hat{s}) \\ &\quad + \sum_{\tilde{s}\in\hat{\mathcal{S}}-\{\hat{s}\}} P_{\hat{S}}(\tilde{s})S\big(K_{\mathrm{A}}|Z_s^n, f(X_{\hat{s}}^n), \hat{S}=\tilde{s}\big) \\ &< \epsilon + \frac{3|\mathcal{Y}\times\mathcal{Z}|^2c_2}{2n(1-\exp(-nc_0))} + 3h\Big(\frac{|\mathcal{Y}\times\mathcal{Z}|^2}{2n^2(1-\exp(-nc_0))}\Big) \\ &\quad + n\exp(-nc_1)|\hat{\mathcal{S}}|c_2, \end{aligned}$$

for some $c_2 > 0$. Therefore, the security index goes to zero as n goes to infinity.

The existence of the SK generator κ for finite sources $\mathfrak{S}_n$ follows again by applying (5.4) to (3.42). As a result, the inequality in (4.19) always hold because the upper bound $(n^3+1)^{\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}}$ in (5.4) increases only polynomially with respect to n. Therefore, the probability in (3.57) is non-zero.

For the converse proof, assume that R_{sk} is achievable. By using the upper bound in (3.1) from Theorem 3.1 or [12, Theorem 1], it follows that

$$R_{\mathrm{sk}} \leq \min_{\hat{s}\in\hat{\mathcal{S}}} \inf_{r,t\in\mathcal{I}(\hat{s})} I(X_{\hat{s}}; Y_r|Z_t). \tag{5.21}$$

Furthermore, for any given $\hat{s} \in \hat{\mathcal{S}}$ and all $r, t \in \mathcal{I}(\hat{s})$, it holds by Markov conditions in (5.1) that

$$I(X_{\hat{s}}; Y_r) - I(X_{\hat{s}}; Z_t) = I(X_{\hat{s}}; Y_r | Z_t).$$

This identity together with (5.21) completes the proof. □

5.2 Convexity of the Rate Region

In this section, the SK capacity result which was derived in Chapter 4, Theorem 4.1 for finite compound sources, is reformulated as a set (region) of SK rate versus communication rate constraint pairs i.e. $(R_{\mathrm{sk}}, \Gamma)$. It is shown that this region is in general convex for compound sources. Moreover, the convexity holds if the compound set is infinite. The proof of convexity follows by a similar approach as in non-compound two-party model in [16, Section 3.3.5] but this time for a compound three-party source model.

Based on this result, it follows that the SK upper bound which is used in the definition of this region is a continuous function with respect to the communication rate upper bound Γ. This property is used in Section 5.3 to extend the SK capacity result of Theorem 4.1 from finite to an arbitrary (possibly infinite) compound source with a finite set of marginals.

Let an arbitrary (possibly infinite) compound source $\mathfrak{S} = \{XYZ, s\}_{s \in \mathcal{S}}$ with set of marginals $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s} \in \hat{\mathcal{S}}}$ be given. Define the rate region as follows:

$$\begin{aligned}\mathcal{R}(\mathfrak{S}) := \bigcap_{\hat{s} \in \hat{\mathcal{S}}} \bigcup_{U_{\hat{s}}, V_{\hat{s}}} \Big\{ (R_{\mathrm{sk}}, \Gamma) \in \mathbb{R}^+ \times \mathbb{R}^{++} : \\ R_{\mathrm{sk}} \leq \inf_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s | U_{\hat{s}}) - \sup_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s | U_{\hat{s}}) \\ \wedge \sup_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}} | Y_s) + \sup_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}} | U_{\hat{s}} Y_s) \leq \Gamma \Big\}, \quad (5.22)\end{aligned}$$

where the union is taken over all RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ such that

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}} - Y_s Z_s\,. \quad (5.23)$$

In the following theorem, the convexity of this set is claimed.

Theorem 5.2. *For an arbitrary (possibly infinite) DMMS $\mathfrak{S} = \{XYZ, s\}_{s \in \mathcal{S}}$ the rate region $\mathcal{R}(\mathfrak{S})$ is convex.*

Proof. Let $\hat{s} \in \hat{\mathcal{S}}$ be given. Let $(R_1, \Gamma_1), (R_2, \Gamma_2) \in \mathcal{R}(\mathfrak{S})$ be given. It follows by using (5.22) and (5.23) that there exist $P^{(1)}_{UV_{\hat{s}}|X_{\hat{s}}}$ and $P^{(2)}_{UV_{\hat{s}}|X_{\hat{s}}}$ with

$$\begin{aligned}R_1 &\leq \inf_{s \in \mathcal{I}(\hat{s})} I(V^{(1)}_{\hat{s}}; Y_s | U^{(1)}_{\hat{s}}) - \sup_{s \in \mathcal{I}(\hat{s})} I(V^{(1)}_{\hat{s}}; Z_s | U^{(1)}_{\hat{s}}), \\ R_2 &\leq \inf_{s \in \mathcal{I}(\hat{s})} I(V^{(2)}_{\hat{s}}; Y_s | U^{(2)}_{\hat{s}}) - \sup_{s \in \mathcal{I}(\hat{s})} I(V^{(2)}_{\hat{s}}; Z_s | U^{(2)}_{\hat{s}}),\end{aligned}$$

such that

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}}^{(1)} - V_{\hat{s}}^{(1)} - X_{\hat{s}} - Y_s Z_s\,, \tag{5.24}$$

$$\sup_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}^{(1)}; X_{\hat{s}}|Y_s) + \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}^{(1)}; X_{\hat{s}}|U_{\hat{s}}^{(1)}Y_s) \leq \Gamma, \tag{5.25}$$

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}}^{(2)} - V_{\hat{s}}^{(2)} - X_{\hat{s}} - Y_s Z_s\,, \tag{5.26}$$

$$\sup_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}^{(2)}; X_{\hat{s}}|Y_s) + \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}^{(2)}; X_{\hat{s}}|U_{\hat{s}}^{(2)}Y_s) \leq \Gamma. \tag{5.27}$$

Define a time sharing RV T with range $\{1,2\}$ being independent of all other RVs as follows

$$P_T(1) := \alpha, \qquad P_T(2) := \bar{\alpha} = 1-\alpha, \qquad \text{where } \alpha \in [0,1]. \tag{5.28}$$

Assume that RVs $U_{\hat{s}}^{(1)}$ and $U_{\hat{s}}^{(2)}$ have the same alphabet $\mathcal{U}_{\hat{s}}$. Similarly, let $V_{\hat{s}}^{(1)}$ and $V_{\hat{s}}^{(2)}$ have the same alphabet $\mathcal{V}_{\hat{s}}$. Construct the auxiliary RV $U_{\hat{s}}$ with range $\bar{\mathcal{U}}_{\hat{s}}$, and $V_{\hat{s}}$ with range $\bar{\mathcal{V}}_{\hat{s}} = \mathcal{V}_{\hat{s}} \times \{1,2\}$ such that for all $(\bar{u},\bar{v},x,y,z) = (u,t,v,t,x,y,z) \in \bar{\mathcal{U}}_{\hat{s}} \times \bar{\mathcal{V}}_{\hat{s}} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ their joint PD with $X_{\hat{s}}Y_sZ_s$ is given by:

$$\begin{aligned} P_{U_{\hat{s}}V_{\hat{s}}X_{\hat{s}}Y_sZ_s}(\bar{u},\bar{v},x,y,z) &:= P_{U_{\hat{s}}^{(T)}V_{\hat{s}}^{(T)}TX_{\hat{s}}Y_sZ_s}(u,v,t,x,y,z) \\ &:= P_T(t) P_{U_{\hat{s}}^{(t)}V_{\hat{s}}^{(t)}X_{\hat{s}}Y_sZ_s}(u,v,x,y,z). \end{aligned} \tag{5.29}$$

Next, to prove the convexity of the set $\mathcal{R}(\mathfrak{S})$, we show that the following Markov chains hold

$$\forall s \in \mathcal{I}(\hat{s}), \quad U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}} - Y_s Z_s. \tag{5.30}$$

Furthermore, we show that

$$\alpha R_1 + \bar{\alpha} R_2 \leq \inf_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Y_s|U_{\hat{s}}) - \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; Z_s|U_{\hat{s}}), \tag{5.31}$$

$$\alpha\Gamma_1 + \bar{\alpha}\Gamma_2 \geq \sup_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_s). \tag{5.32}$$

It holds for all $\bar{u} \in \bar{\mathcal{U}}_{\hat{s}}, \bar{v} = (v,t) \in \bar{\mathcal{V}}_{\hat{s}}, x \in \mathcal{X}$, and $s \in \mathcal{I}(\hat{s})$ that

$$\begin{aligned} P_{X_{\hat{s}}|U_{\hat{s}}V_{\hat{s}}}(x|\bar{u},\bar{v}) &= \frac{P_{U_{\hat{s}}V_{\hat{s}}X_{\hat{s}}}(\bar{u},\bar{v},x)}{P_{U_{\hat{s}}V_{\hat{s}}}(\bar{u},\bar{v})} \\ &= \frac{P_T(t) P_{U_{\hat{s}}^{(t)}V_{\hat{s}}^{(t)}X_{\hat{s}}}(u,v,x)}{P_T(t) P_{U_{\hat{s}}^{(t)}V_{\hat{s}}^{(t)}}(u,v)} \\ &= P_{X_{\hat{s}}|V_{\hat{s}}^{(t)}}(x|v) \\ &= \frac{P_T(t) P_{V_{\hat{s}}^{(t)}X_{\hat{s}}}(v,x)}{P_T(t) P_{V_{\hat{s}}^{(t)}}(v)} \\ &= P_{X_{\hat{s}}|V_{\hat{s}}}(x|\bar{v}), \end{aligned}$$

where the second equality follows by (5.29). The third equality is a result of (5.24) and (5.26).
Similarly, it follows for all $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$ that

$$\begin{aligned} P_{Y_s Z_s|U_{\hat{s}} V_{\hat{s}} X_{\hat{s}}}(y,z|\bar{u},\bar{v},x) &= \frac{P_{U_{\hat{s}} V_{\hat{s}} X_{\hat{s}} Y_s Z_s}(\bar{u},\bar{v},x,y,z)}{P_{U_{\hat{s}} V_{\hat{s}} X_{\hat{s}}}(\bar{u},\bar{v},x)} \\ &= \frac{P_T(t) P_{U_{\hat{s}}^{(t)} V_{\hat{s}}^{(t)} X_{\hat{s}} Y_s Z_s}(u,v,x,y,z)}{P_T(t) P_{U_{\hat{s}}^{(t)} V_{\hat{s}}^{(t)} X_{\hat{s}}}(u,v,x)} \\ &= P_{Y_s Z_s|U_{\hat{s}}^{(t)} V_{\hat{s}}^{(t)} X_{\hat{s}}}(y,z|u,v,x) \\ &= P_{Y_s Z_s|X_{\hat{s}}}(y,z|x). \end{aligned}$$

Thus, the Markov chains in (5.30) hold. Next, the inequality (5.31) is shown. It holds by using (5.28) and (5.29) that

$$\begin{aligned} \alpha R_1 + \bar{\alpha} R_2 &\leq \alpha \inf_{r,s \in \mathcal{I}(\hat{s})} \big\{ I(V_{\hat{s}}^{(1)}; Y_r|U_{\hat{s}}^{(1)}) - I(V_{\hat{s}}^{(1)}; Z_s|U_{\hat{s}}^{(1)}) \big\} \\ &\qquad + \bar{\alpha} \inf_{r,s \in \mathcal{I}(\hat{s})} \big\{ I(V_{\hat{s}}^{(2)}; Y_r|U_{\hat{s}}^{(2)}) - I(V_{\hat{s}}^{(2)}; Z_s|U_{\hat{s}}^{(2)}) \big\} \\ &\leq \inf_{r,s \in \mathcal{I}(\hat{s})} \big\{ \alpha I(V_{\hat{s}}^{(1)}; Y_r|U_{\hat{s}}^{(1)}) \\ &\qquad - \alpha I(V_{\hat{s}}^{(1)}; Z_s|U_{\hat{s}}^{(1)}) \\ &\qquad + \bar{\alpha} I(V_{\hat{s}}^{(2)}; Y_r|U_{\hat{s}}^{(2)}) \\ &\qquad - \bar{\alpha} I(V_{\hat{s}}^{(2)}; Z_s|U_{\hat{s}}^{(2)}) \big\} \\ &= \inf_{r,s \in \mathcal{I}(\hat{s})} \Big\{ \sum_{t \in \{1,2\}} P_T(t) H(Y_r|U_{\hat{s}}^{(t)}, T=t) \\ &\qquad - \sum_{t \in \{1,2\}} P_T(t) H(Y_r|U_{\hat{s}}^{(t)} V_{\hat{s}}^{(t)}, T=t) \\ &\qquad - \sum_{t \in \{1,2\}} P_T(t) H(Z_s|U_{\hat{s}}^{(t)}, T=t) \\ &\qquad + \sum_{t \in \{1,2\}} P_T(t) H(Z_s|U_{\hat{s}}^{(t)} V_{\hat{s}}^{(t)}, T=t) \Big\} \\ &= \inf_{r,s \in \mathcal{I}(\hat{s})} \big\{ I(V_{\hat{s}}; Y_r|U_{\hat{s}}) - I(V_{\hat{s}}; Z_s|U_{\hat{s}}) \big\}. \end{aligned}$$

This gives the inequality (5.31).

Next, we show in the following that (5.32) holds. It follows by using (5.25), (5.27), (5.28), and (5.29) that

$$
\begin{aligned}
\alpha\Gamma_1+\bar{\alpha}\Gamma_2 &\geq \alpha \sup_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}^{(1)};X_{\hat{s}}|Y_s) + \alpha \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}^{(1)};X_{\hat{s}}|U_{\hat{s}}^{(1)}Y_s) \\
&\quad + \bar{\alpha} \sup_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}^{(2)};X_{\hat{s}}|Y_s) + \bar{\alpha} \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}}^{(2)};X_{\hat{s}}|U_{\hat{s}}^{(2)}Y_s) \\
&\geq \sup_{s\in\mathcal{I}(\hat{s})} \left\{\alpha I(U_{\hat{s}}^{(1)};X_{\hat{s}}|Y_s) + \bar{\alpha} I(U_{\hat{s}}^{(2)};X_{\hat{s}}|Y_s)\right\} \\
&\quad + \sup_{s\in\mathcal{I}(\hat{s})} \left\{\alpha I(V_{\hat{s}}^{(1)};X_{\hat{s}}|U_{\hat{s}}^{(1)}Y_s) + \bar{\alpha} I(V_{\hat{s}}^{(2)};X_{\hat{s}}|U_{\hat{s}}^{(2)}Y_s)\right\} \\
&= \sup_{s\in\mathcal{I}(\hat{s})} \Big\{H(X_{\hat{s}}|Y_s) - \sum_{t\in\{1,2\}} P_T(t) H(X_{\hat{s}}|Y_s U_{\hat{s}}^{(t)}, T=t)\Big\} \\
&+ \sup_{s\in\mathcal{I}(\hat{s})} \Big\{ \sum_{t\in\{1,2\}} P_T(t) H(X_{\hat{s}}|Y_s U_{\hat{s}}^{(t)}, T=t) - \sum_{t\in\{1,2\}} P_T(t) H(X_{\hat{s}}|Y_s U_{\hat{s}}^{(t)} V_{\hat{s}}^{(t)}, T=t)\Big\} \\
&= \sup_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};X_{\hat{s}}|Y_s) + \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};X_{\hat{s}}|U_{\hat{s}}Y_s).
\end{aligned}
$$

This gives the inequality (5.32). It follows then by using (5.30), (5.31), and (5.32) that

$$\alpha(R_1,\Gamma_1) + \bar{\alpha}(R_2,\Gamma_2) \in \mathcal{R}(\mathfrak{S}).$$

This completes the proof of convexity. □

5.3 Single-Letter Secret-Key Capacity Lower Bound

In the following, Theorem 5.3 gives a single-letter lower bound for the SK capacity as a function of the compound DMMS and public communication rate upper bound. It is assumed that the compound set $\mathfrak{S}$ is arbitrary (possibly infinite) and the set of marginals $\hat{\mathfrak{S}}$ is finite.

Theorem 5.3. *Consider the SK model from Section 3.2 with an arbitrary (possibly infinite) compound DMMS $\mathfrak{S} = \{XYZ, s\}_{s\in\mathcal{S}}$, a finite set of marginals $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s}\in\hat{\mathcal{S}}}$, and a one-way communication over a public noiseless channel with communication rate upper bound $\Gamma \in (0,\infty]$. Then, it holds that*

$$C_{\mathrm{sk}}(\mathfrak{S},\Gamma) \geq \min_{\hat{s}\in\hat{\mathcal{S}}} \sup_{U_{\hat{s}},V_{\hat{s}}} \Big\{ \inf_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};Y_s|U_{\hat{s}}) - \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};Z_s|U_{\hat{s}})\Big\}, \tag{5.33}$$

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

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}} - Y_s Z_s\,, \tag{5.34}$$

$$\sup_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};X_{\hat{s}}|Y_s) + \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};X_{\hat{s}}|U_{\hat{s}}Y_s) \leq \Gamma. \tag{5.35}$$

Proof. Let the compound source $\mathfrak{S}$ be infinite and the set of marginals $\hat{\mathfrak{S}}$ be finite. Let $\hat{s} \in \hat{\mathcal{S}}$ be given and consider the infinite class of stochastic matrices as follows:

$$\{P_{YZ,s|X_{\hat{s}}} : \mathcal{X} \to \mathcal{P}(\mathcal{Y} \times \mathcal{Z})\}_{s \in \mathcal{I}(\hat{s})}. \tag{5.36}$$

In the first step, similarly as in Theorem 5.1, we approximate the class of stochastic matrices given in (5.36). Lemma 2.6 implies by taking $l = n^3$ that for any $n^3 > 2|\mathcal{Y} \times \mathcal{Z}|^2$, there exists a finite set of stochastic matrices

$$\{W_{s'} : \mathcal{X} \to \mathcal{P}(\mathcal{Y} \times \mathcal{Z})\}_{s' \in \mathcal{I}_n(\hat{s})} \qquad \text{with} \qquad |\mathcal{I}_n(\hat{s})| \leq (n^3+1)^{|\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}|}, \tag{5.37}$$

which approximates the one in (5.36) such that conditions (5.5), (5.6), and (5.7) hold.

The proof is done first for a special case in part a). The proof of general case follows in part b).

Part a) Define for all $\hat{s} \in \hat{\mathcal{S}}$ and $r, t \in \mathcal{I}(\hat{s})$

$$\bar{f}(\hat{s}, r, t) := I(U_{\hat{s}}; Y_r) - I(U_{\hat{s}}; Z_t). \tag{5.38}$$

We show that the following SK rate is achievable:

$$\bar{R}_{\mathrm{sk}}(\mathfrak{S}, \Gamma) := \min_{\hat{s} \in \hat{\mathcal{S}}} \sup_{U_{\hat{s}}} \inf_{r,t \in \mathcal{I}(\hat{s})} \bar{f}(\hat{s}, r, t), \tag{5.39}$$

where the sup is taken over all RVs $U_{\hat{s}}$ such that

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - X_{\hat{s}} - Y_s Z_s\,, \quad \text{and} \quad \sup_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}} | Y_s) \leq \Gamma. \tag{5.40}$$

By using (5.38) and Theorem 4.1 for finite compound sources $\mathfrak{S}_n$, it follows that the following SK rates are achievable:

$$\bar{R}_{\mathrm{sk}}(\mathfrak{S}_n, \Gamma) := \min_{\hat{s} \in \hat{\mathcal{S}}} \max_{U_{\hat{s}}} \min_{r', t' \in \mathcal{I}_n(\hat{s})} \bar{f}(\hat{s}, r', t'), \tag{5.41}$$

where the max is taken over all $U_{\hat{s}}$ with alphabet $\mathcal{U}_{\hat{s}}$ such that

$$\forall s' \in \mathcal{I}_n(\hat{s}),\ U_{\hat{s}} - X_{\hat{s}} - Y_{s'} Z_{s'}, \quad \max_{s' \in \mathcal{I}_n(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}} | Y_{s'}) \leq \Gamma. \tag{5.42}$$

The argument regarding universal constants which was given in the proof of Theorem 5.1 holds also here. These constants should be chosen universally for all sets $\mathcal{I}_n(\hat{s})$ as explained there.

Similarly as in the proof of Theorem 5.1, we show next for some $n_0 > 0$ and all $n \geq n_0$ that the SK generation protocols which guarantee the achievability of the SK rates in (5.41) for finite compound sources $\mathfrak{S}_n$, also guarantee the achievability of the SK rate in (5.39) for the infinite compound source $\mathfrak{S}$. To this end, we verify in the following, all conditions of Definition 3.3. A universal n_0 can be determined, because it depends only linearly on $\log |\mathcal{I}_n(\hat{s})|$ (cf. proof of Theorem 4.1).

Condition (3.3) *(Communication Rate Constraint):* For $\mathfrak{S}_n$, the communication rate between Alice and Bob is given in the proof of Theorem 4.1, part a) as

$$\max_{s' \in \mathcal{I}_n(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_{s'}) + 3\delta + \frac{1}{n} \log |\hat{\mathcal{S}}|. \tag{5.43}$$

Next, we show that by applying the infinite source $\mathfrak{S}$, the communication rate $\frac{1}{n} \log \|f_c\|$ still satisfies condition (3.3) of Definition 3.3.

The Markov chain conditions in (5.40) imply that for all $s \in \mathcal{I}(\hat{s})$ and $(u, x, y, z) \in \mathcal{U}_{\hat{s}} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$,

$$P_{UXYZ,s}(u, x, y, z) = P_{U_{\hat{s}}|X_{\hat{s}}}(u|x) P_{XYZ,s}(x, y, z).$$

Take the same stochastic matrix $P_{U_{\hat{s}}|X_{\hat{s}}}$ as above and define

$$P_{UXYZ,s'}(u, x, y, z) := P_{U_{\hat{s}}|X_{\hat{s}}}(u|x) P_{XYZ,s'}(x, y, z),$$

where $s' \in \mathcal{I}_n(\hat{s})$. Thus $U_{\hat{s}}$ satisfies the Markov chains in (5.42).

Furthermore, it follows by using (5.5) and the Markov chain conditions in (5.40) and (5.42) that

$$\gamma_n^{(1)} := \frac{1}{2} \big\| P_{UY,s} - P_{UY,s'} \big\| \leq \frac{1}{2n^3} |\mathcal{Y} \times \mathcal{Z}|^2. \tag{5.44}$$

By using Lemma 2.9 together with (5.44), it follows that

$$\begin{aligned} &|I(U_{\hat{s}}; Y_s) - I(U_{\hat{s}}; Y_{s'})| < \epsilon_n^{(1)}, \\ &\text{where} \quad \epsilon_n^{(1)} := 3\gamma_n^{(1)} \log(|\mathcal{U}_{\hat{s}} \times \mathcal{Y}|) + 3h(\gamma_n^{(1)}). \end{aligned} \tag{5.45}$$

By adding and subtracting the term $I(U_{\hat{s}}; X_{\hat{s}})$ inside the absolute value in (5.45) and using the Markov chain properties, and also using the infimum and minimum properties as in the proof of (5.16), it follows that

$$\big| \sup_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) - \max_{s' \in \mathcal{I}_n(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_{s'}) \big| \leq \epsilon_n^{(1)}. \tag{5.46}$$

The sequence $\gamma_n^{(1)}$ and thus $\epsilon_n^{(1)}$ approach zero as n goes to infinity. Thus, condition (3.3) follows by using (5.43) and (5.46).

Condition (3.4) *(SK Rate):* Similarly as in (5.44), it can be shown by using (5.5) and the Markov chain conditions in (5.40) and (5.42) that for all $t \in \mathcal{I}(\hat{s})$ and its corresponding $t' \in \mathcal{I}_n(\hat{s})$ from the approximating source,

$$\gamma_n^{(2)} := \frac{1}{2} \big\| P_{UZ,t} - P_{UZ,t'} \big\| \leq \frac{1}{2n^3} |\mathcal{Y} \times \mathcal{Z}|^2. \tag{5.47}$$

Again, by using Lemma 2.9 together with (5.47), it follows that

$$|I(U_{\hat{s}}; Z_t) - I(U_{\hat{s}}; Z_{t'})| < \epsilon_n^{(2)}, \tag{5.48}$$

$$\text{where} \qquad \epsilon_n^{(2)} := 3\gamma_n^{(2)} \log(|\mathcal{U}_{\hat{s}} \times \mathcal{Z}|) + 3h(\gamma_n^{(2)}).$$

Similarly as in the proof of (5.16), the inequalities (5.45) and (5.48) together with the definition in (5.38), triangle inequality, and infimum properties imply that

$$\inf_{s,t \in \mathcal{I}(\hat{s})} \bar{f}(\hat{s}, s, t) \leq \min_{s', t' \in \mathcal{I}_n(\hat{s})} \bar{f}(\hat{s}, s', t') + \epsilon_n^{(1)} + \epsilon_n^{(2)}. \tag{5.49}$$

Next, we take maximum of both sides in (5.49) over all $U_{\hat{s}}$ which satisfy

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - X_{\hat{s}} - Y_s Z_s, \quad \sup_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}} | Y_s) \leq \Gamma - 1/n. \tag{5.50}$$

The term $\epsilon_n^{(1)}$ in (5.46) approaches zero faster than $1/n$. Thus, in the maximization process in (5.49), each $U_{\hat{s}}$ that satisfies (5.50), also satisfies $\max_{s' \in \mathcal{I}_n(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}} | Y_{s'}) \leq \Gamma$. By using (5.39), (5.41), (5.49) and minimizing over $\hat{s} \in \hat{\mathcal{S}}$, it follows that

$$\bar{R}_{\text{sk}}(\mathfrak{S}, \Gamma - 1/n) \leq \bar{R}_{\text{sk}}(\mathfrak{S}_n, \Gamma) + \epsilon_n^{(1)} + \epsilon_n^{(2)}. \tag{5.51}$$

On the other hand, $\bar{R}_{\text{sk}}(\mathfrak{S}, \cdot)$ is a continuous function with respect to Γ and it holds for all $n \in \mathbb{N}$ that

$$\bar{R}_{\text{sk}}(\mathfrak{S}, \Gamma) \leq \bar{R}_{\text{sk}}(\mathfrak{S}, \Gamma - 1/n) + \epsilon_n^{(3)}, \tag{5.52}$$

where $\epsilon_n^{(3)}$ approaches zero as n goes to infinity. The continuity of $\bar{R}_{\text{sk}}(\mathfrak{S}, \cdot)$ follows from the fact that the set of all $\big(\bar{R}_{\text{sk}}(\mathfrak{S}, \Gamma), \Gamma\big)$ pairs forms a convex set. The convexity was discussed in Section 5.2, Theorem 5.2.

By using (5.51), (5.52), and the fact that $\bar{R}_{\text{sk}}(\mathfrak{S}_n, \Gamma)$ are achievable, it follows that the SK rate $\bar{R}_{\text{sk}}(\mathfrak{S}, \Gamma)$ for the infinite source $\mathfrak{S}$ also satisfies condition (3.4).

Conditions (3.5) *and* (3.6) *(Reliability and Security):* These two steps follow similarly as in the proof of Theorem 5.1 by using (5.5) and (5.6). There, it was shown that the decoding error probability and security index approach zero for all $s \in \mathcal{I}(\hat{s})$.

Part b) In this part, the proof of general case is given. Define for all $\hat{s} \in \hat{\mathcal{S}}$ and $r, t \in \mathcal{I}(\hat{s})$

$$f(\hat{s}, r, t) := I(V_{\hat{s}}; Y_r | U_{\hat{s}}) - I(V_{\hat{s}}; Z_t | U_{\hat{s}}). \tag{5.53}$$

We show that the following SK rate is achievable:

$$R_{\text{sk}}(\mathfrak{S}, \Gamma) := \min_{\hat{s} \in \hat{\mathcal{S}}} \sup_{U_{\hat{s}}, V_{\hat{s}}} \inf_{r, t \in \mathcal{I}(\hat{s})} f(\hat{s}, r, t), \tag{5.54}$$

where $R_{\text{sk}}(\mathfrak{S}, \Gamma)$ is the lower bound in (5.33) and the supremum is taken over all RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ satisfying (5.34) and (5.35).

By using (5.53) and Theorem 4.1 for finite compound sources $\mathfrak{S}_n$, it follows that the following SK rates are achievable:

$$R_{\text{sk}}(\mathfrak{S}_n, \Gamma) := \min_{\hat{s} \in \hat{\mathcal{S}}} \max_{U_{\hat{s}}, V_{\hat{s}}} \min_{r', t' \in \mathcal{I}_n(\hat{s})} f(\hat{s}, r', t'), \tag{5.55}$$

where the max is taken over all RVs $U_{\hat{s}}$ and $V_{\hat{s}}$ such that

$$\forall s' \in \mathcal{I}_n(\hat{s}),\ U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}} - Y_{s'} Z_{s'}\,, \tag{5.56}$$

$$\max_{s' \in \mathcal{I}_n(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_{s'}) + \max_{s' \in \mathcal{I}_n(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_{s'}) \leq \Gamma. \tag{5.57}$$

Similarly as in part a), all conditions of Definition 3.3 are verified in the following.

Condition (3.3) *(Communication Rate Constraint):* For $\mathfrak{S}_n$, the communication rate between Alice and Bob is given in the proof of Theorem 4.1, part b) as

$$\max_{s' \in \mathcal{I}_n(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_{s'}) + \max_{s' \in \mathcal{I}_n(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_{s'}) + 6\delta + \frac{1}{n} \log |\hat{\mathcal{S}}|. \tag{5.58}$$

Next, we show that by applying the infinite source $\mathfrak{S}$, the communication rate $\frac{1}{n} \log \|f_{\mathrm{c}}\|$ still satisfies condition (3.3).

The Markov chain conditions in (5.34) imply that for all $s \in \mathcal{I}(\hat{s})$ and $(u, v, x, y, z) \in \mathcal{U}_{\hat{s}} \times \mathcal{V}_{\hat{s}} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$,

$$P_{UVXYZ,s}(u, x, y, z) = P_{U_{\hat{s}}V_{\hat{s}}|X_{\hat{s}}}(u, v|x) P_{XYZ,s}(x, y, z).$$

Take the same stochastic matrix $P_{U_{\hat{s}}V_{\hat{s}}|X_{\hat{s}}}$ as above and define

$$P_{UVXYZ,s'}(u, x, y, z) := P_{U_{\hat{s}}V_{\hat{s}}|X_{\hat{s}}}(u, v|x) P_{XYZ,s'}(x, y, z),$$

where $s' \in \mathcal{I}_n(\hat{s})$. Thus, $U_{\hat{s}}$ and $V_{\hat{s}}$ satisfy the Markov chains in (5.56). Similarly as in (5.46) and by using Lemma 2.9, it can be shown that

$$\begin{aligned} \Big| \sup_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + \sup_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_s) \\ - \max_{s' \in \mathcal{I}_n(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_{s'}) - \max_{s' \in \mathcal{I}_n(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_{s'}) \Big| \leq \epsilon_n, \end{aligned} \tag{5.59}$$

where ϵ_n approaches zero as n goes to infinity. This implies by using (5.58) that condition (3.3) of Definition 3.3 is satisfied.

Condition (3.4) *(SK Rate):* Similarly as in (5.49), it can be shown by using Lemma 2.9 and (5.53) that

$$\inf_{s,t \in \mathcal{I}(\hat{s})} f(\hat{s}, s, t) \leq \min_{s', t' \in \mathcal{I}_n(\hat{s})} f(\hat{s}, s', t') + \epsilon'_n, \tag{5.60}$$

where ϵ'_n approaches zero as n goes to infinity. We take the maximum of both sides in (5.60) over all $U_{\hat{s}}$ and $V_{\hat{s}}$ which satisfy (5.34) and

$$\sup_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) + \sup_{s \in \mathcal{I}(\hat{s})} I(V_{\hat{s}}; X_{\hat{s}}|U_{\hat{s}}Y_s) \leq \Gamma - 1/n. \tag{5.61}$$

Similarly as in (5.46), the term ϵ_n in (5.59) approaches zero faster than $1/n$. Thus, in the maximization process in (5.60), each $U_{\hat{s}}$ and $V_{\hat{s}}$ that satisfy (5.61), satisfies also (5.57). Therefore, by using (5.54), (5.55), (5.60) and minimizing over $\hat{s} \in \hat{\mathcal{S}}$, it follows that

$$R_{\mathrm{sk}}(\mathfrak{S}, \Gamma - 1/n) \leq R_{\mathrm{sk}}(\mathfrak{S}_n, \Gamma) + \epsilon'_n. \tag{5.62}$$

On the other hand, $R_{\mathrm{sk}}(\mathfrak{S},\cdot)$ is a continuous function with respect to Γ and it holds for all $n \in \mathbb{N}$ that

$$R_{\mathrm{sk}}(\mathfrak{S},\Gamma) \leq R_{\mathrm{sk}}(\mathfrak{S},\Gamma - 1/n) + \epsilon_n'', \tag{5.63}$$

where ϵ_n'' approaches zero as n goes to infinity. Again, the continuity of $R_{\mathrm{sk}}(\mathfrak{S},\cdot)$ follows from the fact that the set of all $\big(R_{\mathrm{sk}}(\mathfrak{S},\Gamma),\Gamma\big)$ pairs forms a convex set and the convexity is a result of Theorem 5.2.

By using (5.62), (5.63), and the fact that $R_{\mathrm{sk}}(\mathfrak{S}_n,\Gamma)$ is achievable, it follows that the SK rate $R_{\mathrm{sk}}(\mathfrak{S},\Gamma)$ for the infinite source $\mathfrak{S}$ also satisfies condition (3.4).

Conditions (3.5) *and* (3.6) *(Reliability and Security):* These two steps follow again similarly as in the proof of Theorem 5.1 by using (5.5) and (5.6). This completes the proof. □

5.4 Multi-Letter Secret-Key Capacity Formula

Next, a multi-letter SK capacity formula as a function of a compound DMMS and public communication rate upper bound is given. The compound source is again assumed to be arbitrary (possibly infinite). The set of marginals is finite.

Theorem 5.4. *Consider the SK model from Section 3.2 with an arbitrary (possibly infinite) compound DMMS* $\mathfrak{S} = \{XYZ,s\}_{s\in\mathcal{S}}$*, a finite set of marginals* $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s}\in\hat{\mathcal{S}}}$*, and a one-way communication over a public noiseless channel with communication rate upper bound* $\Gamma \in (0,\infty]$*. Then, the SK capacity is given by*

$$C_{\mathrm{sk}}(\mathfrak{S},\Gamma) = \lim_{n\to\infty}\frac{1}{n}\min_{\hat{s}\in\hat{\mathcal{S}}}\sup_{U_{\hat{s}},V_{\hat{s}}}\Big\{\inf_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};Y_s^n|U_{\hat{s}}) - \sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};Z_s^n|U_{\hat{s}})\Big\}, \tag{5.64}$$

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

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - V_{\hat{s}} - X_{\hat{s}}^n - Y_s^n Z_s^n,$$
$$\frac{1}{n}\sup_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}};X_{\hat{s}}^n|Y_s^n) + \frac{1}{n}\sup_{s\in\mathcal{I}(\hat{s})} I(V_{\hat{s}};X_{\hat{s}}^n|U_{\hat{s}}Y_s^n) \leq \Gamma.$$

Proof. The achievability proof of the SK rate in (5.64) follows by a code-blocking argument. For this, we apply the result of Theorem 5.3 to an n-fold product of the source. Similarly as in the proof of Theorem 4.2, the existence of the limit in (5.64) is shown by using the Fekete's lemma [56]. The argument for the converse proof is similar as in the proof of Theorem 4.2, but this time for infinite compound sources. □

6 Compound Biometric Authentication Systems with Strong Secrecy

In this chapter, the SK model and capacity region of the compound biometric authentication systems are given. To this end, a special case of the SK capacity results from Chapter 4 is used in the biometric model.

6.1 Mathematical Model with Compound Sources

Figure 6.1 shows the biometric authentication model using a 2-party finite compound DMMS $\mathfrak{S} = \{XY, s\}_{s\in\mathcal{S}}$. The source takes its values in the finite set $\mathcal{X} \times \mathcal{Y}$.

Similarly as in [16], the model consists of two steps. In the enrollment phase, the user's biometric information is captured and based on that the biometric enrollment sequence $X_{\hat{s}}^n$ is generated in time duration $n \in \mathbb{N}$. Next, in the authentication phase, a biometric authentication sequence Y_s^n is generated based on the newly captured biometric information of the same user. Both sequences are given by observing the correlated components of the compound DMMS $\mathfrak{S}$.

As seen in Figure 6.1, the enrollment sequence $X_{\hat{s}}^n$ is fed to an enrollment terminal which has access to a public database. It may store information there which is publicly available. The authentication sequence Y_s^n is similarly fed to an authentication terminal which has access to the public database as well. As the database is publicly available, an adversary has also access to the database.

It is assumed that the set $\mathfrak{S}$ as well as its statistics with PDs $\{P_{XY,s}\}_{s\in\mathcal{S}}$ are known to the enrollment and authentication terminals. However, they do not have the knowledge of the actual realization of the source.

As RVs $X_{\hat{s}}^n$ and Y_s^n are correlated, enrollment and authentication terminals may generate a shared SK by communicating over the public database. Only a one-way communication over the public database is allowed. The following definition gives a more precise description of this procedure which is similar to the non-compound protocol [16].

Definition 6.1. *A biometric authentication protocol with compound source set* $\mathfrak{S} = \{XY, s\}_{s\in\mathcal{S}}$ *consists of the following steps:*

- *After observing* $X_{\hat{s}}^n$*, the enrollment terminal stores some helper data* M *in the public database. Assume that* M *takes its values in the set* $\mathcal{M}$*.*
- *Furthermore, the enrollment terminal generates an SK, represented by an RV* K*, based on its knowledge* $X_{\hat{s}}^n$*.* K *takes its values in* $\mathcal{K}$*. The enrollment terminal stores*

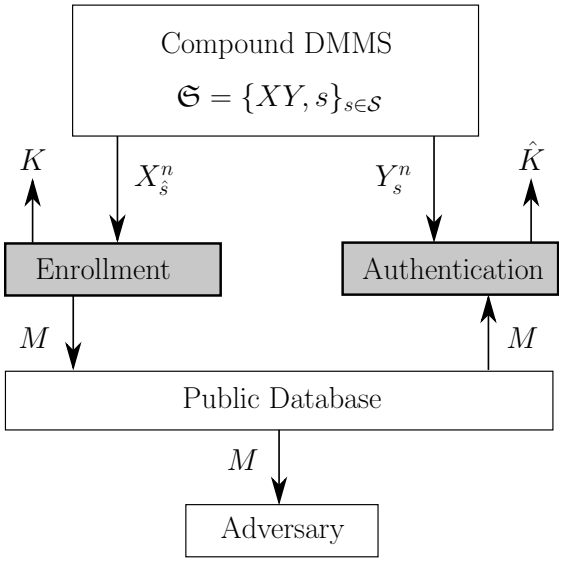

Figure 6.1: Compound biometric authentication system.

an encrypted version of K in the database. The encryption is performed by using a one-way function.

- *Next, the authentication terminal reads the helper data M from the database. It then generates an SK, represented by an RV $\hat{K}$, based on its knowledge (Y_s^n, M). $\hat{K}$ takes its values in $\mathcal{K}$ as well.*

- *Finally, the authentication terminal generates the encrypted version of $\hat{K}$ by the same one-way function which was used by enrollment terminal. It then compares it with the encrypted version of K which was stored in the public database. If they are identical, then the authentication is approved and otherwise failed.*

As the helper data is stored in the public database and consequently available to Eve, it should not reveal any information about the SK. Moreover, the generated SK should have a uniform distribution.

Another critical constraint is related to the privacy leakage [27]. The privacy leakage is defined to be the amount of information which the helper data M reveals about the biometric enrollment sequence $X_{\hat{s}}^n$ [16]. The helper data should not reveal much information about the user's biometric information. In the following, the definition of an achievable SK rate R_{sk} versus privacy leakage rate constraint L for a compound source is given where the strong secrecy [14] is guaranteed.

Definition 6.2. *A pair of numbers $(R_{\text{sk}}, L) \in \mathbb{R}^+ \times \mathbb{R}^{++}$ is achievable for the compound source $\mathfrak{S} = \{XY, s\}_{s\in\mathcal{S}}$, iff for all $\delta > 0$, and all $n \in \mathbb{N}$ large enough, there exists a biometric authentication protocol with helper data M, giving rise to the RVs K and $\hat{K}$*

with values in $\mathcal{K}$, for which it holds for all $s \in \mathcal{S}$ that

$$R_{\mathrm{sk}} < \frac{1}{n} \log |\mathcal{K}| + \delta, \tag{6.1}$$

$$\Pr(K \neq \hat{K}) < \delta, \tag{6.2}$$

$$I(K; M) < \delta, \tag{6.3}$$

$$\log |\mathcal{K}| < H(K) + \delta, \tag{6.4}$$

$$\frac{1}{n} I(X_{\hat{s}}^n; M) < L + \delta. \tag{6.5}$$

$\mathcal{R}(\mathfrak{S})$ is defined to be the SK capacity region containing all pairs (R_{sk}, L) which are achievable.

Remark. *In the protocol from Definition 6.1, no randomization was assumed. Similarly as in [18, Problem 17.15(a)], it can be shown that each rate pair (R_{sk}, L) that is achievable by using a randomized M, can also be achieved without randomization.*

Condition (6.1) from Definition 6.2 gives the SK rate condition. Condition (6.2) ensures that the generated SKs K and $\hat{K}$ are reliably identical and thus the authentication procedure is done with a sufficiently low error probability.

Condition (6.3) guarantees the strong secrecy. Unlike the weak secrecy where only the rate of mutual information should be small, the strong secrecy ensures that the helper data M reveals asymptotically no information about the generated SK K. Condition (6.4) of Definition 6.2 guarantees the uniformity of the generated SK.

Finally, condition (6.5) ensures that the information rate which the helper data M reveals about the biometric enrollment sequence $X_{\hat{s}}^n$ is limited by constraint L. Each of these conditions should be satisfied simultaneously for all realizations of the compound source with indices $s \in \mathcal{S}$.

Remark. *To understand the importance of strong secrecy in condition* (6.3) *of Definition 6.2, the readers are recommended to refer to the examples in [59, Section 3]. Weak secrecy which is given by the condition $n^{-1} I(K; M) < \delta$, does not always guarantee an operational secrecy. This is because, some protocols in information theoretic security are vulnerable to serious attacks and yet they might satisfy the weak secrecy condition.*

In the following, similarly as in Section 3.2 and [29, 30], a subset of the compound set $\mathfrak{S}$ is defined. This definition is required for stating the results and defining the rate regions.

Definition 6.3. *Let compound source $\mathfrak{S} = \{XY, s\}_{s \in \mathcal{S}}$ be given and $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s} \in \hat{\mathcal{S}}}$ be its set of marginals. For a given marginal $X_{\hat{s}} \in \hat{\mathfrak{S}}$, the set of all possible joint sources is denoted by $\mathfrak{I}(\hat{s}) = \{XY, s\}_{s \in \mathcal{I}(\hat{s})}$ and is given by*

$$\mathfrak{I}(\hat{s}) := \Big\{ XY, s \in \mathfrak{S} : \quad \forall x \in \mathcal{X}, \sum_{y \in \mathcal{Y}} P_{XY,s}(x, y) = P_{X_{\hat{s}}}(x) \Big\}.$$

Finally, the rate region $\mathcal{R}^*(\mathfrak{S})$ is defined in the following. This region is required for stating the main result in Section 6.2.

Definition 6.4. *Let compound source $\mathfrak{S} = \{XY, s\}_{s\in\mathcal{S}}$ be given and $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s}\in\hat{\mathcal{S}}}$ be its set of marginals. Then*

$$\mathcal{R}^*(\mathfrak{S}) := \bigcap_{\hat{s}\in\hat{\mathcal{S}}} \bigcup_{U_{\hat{s}}} \Big\{(R_{\mathrm{sk}}, L) \in \mathbb{R}^+ \times \mathbb{R}^{++} : R_{\mathrm{sk}} \leq \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) \\ \wedge \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) \leq L\Big\},$$

where the union is taken over all RVs $U_{\hat{s}}$ such that it holds

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - X_{\hat{s}} - Y_s.$$

6.2 Secret-Key Capacity Region

To show the achievability of a given pair (R_{sk}, L) from Section 6.1, we use the SK model from Section 3.2 and the results from Section 4.1. In that model three terminals observe the correlated components of a compound source

$$\mathfrak{S} := \{XYZ, s\}_{s\in\mathcal{S}}$$

and two of them (Alice and Bob) generate a shared SK by only a one-way communication over a public noiseless channel while keeping it secret from the third terminal (Eve). Moreover, the communication rate between Alice and Bob is constrained.

In Definition 3.3 of achievability for the mentioned model, the security and uniformity conditions of biometric model which were given in conditions (6.3) and (6.4) of Definition 6.2 are combined and replaced by

$$\forall s \in \mathcal{S}, \quad S(K|Z_s^n, M) < \delta. \tag{6.6}$$

This condition guarantees the strong secrecy and uniformity of the generated SK K simultaneously.

Furthermore, in Definition 3.3 instead of the privacy leakage rate constraint (6.5) of Definition 6.2, the following stronger condition is required:

$$\frac{1}{n}\log|\mathcal{M}| < L + \delta, \tag{6.7}$$

where L is interpreted as the public communication rate constraint between Alice and Bob.

In the following a single-letter SK capacity lower bound as a function of public communication rate upper bound L is given. This result is a special case of Theorem 4.1 and follows by taking Z_s = constant. We use this Lemma to show the achievability of the SK rate for compound biometric authentication systems and find the capacity regions. We note that in [30] the capacity regions are derived by using a different approach for the achievability proof which is based on the setup in [16].

Lemma 6.1. *For a finite compound DMMS $\mathfrak{S} = \{XY, s\}_{s\in\mathcal{S}}$, set of marginals $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s}\in\hat{\mathcal{S}}}$, and a one-way communication over a public noiseless channel with constraint $L \in (0, \infty]$, the following SK rate is achievable:*

$$\min_{\hat{s}\in\hat{\mathcal{S}}} \max_{U_{\hat{s}}} \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s), \tag{6.8}$$

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

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - X_{\hat{s}} - Y_s \ \text{ and } \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) \leq L.$$

Corollary 6.1. *By taking $U_{\hat{s}} = X_{\hat{s}}$ in Lemma 6.1, the privacy leakage rate condition reduces to $\max_{s\in\mathcal{I}(\hat{s})} H(X_{\hat{s}}|Y_s) \leq L$ and the* max *in* (6.8) *becomes obsolete. The achievable rate* (6.8) *reduces then to the special case $\min_{s\in\mathcal{S}} I(X_s; Y_s)$ which is given in [29] and the non-compound version in [11, 12].*

Finally, the single-letter SK capacity region of a finite compound biometric authentication model from Section 6.1, is given in the following. In this approach, strong secrecy is guaranteed. The extension from finite to infinite compound sources (with a finite set of marginals) follows similarly as in Chapter 5.

Theorem 6.1. *Consider the biometric authentication model from Section 6.1 with a finite compound DMMS $\mathfrak{S} = \{XY, s\}_{s\in\mathcal{S}}$, set of marginals $\hat{\mathfrak{S}} = \{X_{\hat{s}}\}_{\hat{s}\in\hat{\mathcal{S}}}$, and the corresponding regions $\mathcal{R}(\mathfrak{S})$ and $\mathcal{R}^*(\mathfrak{S})$ from Definitions 6.2 and 6.4. Then, it holds that*

$$\mathcal{R}(\mathfrak{S}) = \mathcal{R}^*(\mathfrak{S}).$$

Proof. To prove the achievability part, we assume that $(R_{\mathrm{sk}}, L) \in \mathcal{R}^*(\mathfrak{S})$ and show that $(R_{\mathrm{sk}}, L) \in \mathcal{R}(\mathfrak{S})$. Definition 6.4 implies for all $\hat{s} \in \hat{\mathcal{S}}$ that there exists some $U_{\hat{s}}$ with

$$\forall s \in \mathcal{I}(\hat{s}),\ U_{\hat{s}} - X_{\hat{s}} - Y_s,$$

such that it holds:

$$0 \leq R_{\mathrm{sk}} \leq \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s) \ \wedge \ \max_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) \leq L.$$

As $\hat{s}$ was taken arbitrarily, this implies that

$$0 \leq R_{\mathrm{sk}} \leq \min_{\hat{s}\in\hat{\mathcal{S}}} \max_{U_{\hat{s}}} \min_{s\in\mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s). \tag{6.9}$$

On the other hand, by using Lemma 6.1, the upper bound in (6.9) is an achievable SK rate with respect to the conditions (6.1), (6.2), (6.6), and (6.7). This implies that the upper bound is also achievable with respect to Definition 6.2. This is because, for a given $\delta > 0$ and $s \in \mathcal{S}$, it follows [16] that

$$\begin{aligned} \frac{1}{n} I(X_{\hat{s}}^n; M) &= \frac{1}{n} H(M) \\ &\leq \frac{1}{n} \log |\mathcal{M}| \\ &\leq L + \delta, \end{aligned}$$

where the last inequality is a result of (6.7). The strong secrecy and uniformity conditions (6.3) and (6.4) from Definition 6.2 follow as well by using condition (6.6). Therefore, it holds that $(R_{\text{sk}}, L) \in \mathcal{R}(\mathfrak{S})$ and the achievability proof is complete.

The converse proof follows similarly as in [16, 30]. Assume that $(R_{\text{sk}}, L) \in \mathcal{R}(\mathfrak{S})$. Let $\delta > 0$ be given and $\hat{s} \in \hat{\mathcal{S}}$ be the marginal index. Furthermore, let $s \in \mathcal{I}(\hat{s})$ be given. Define the RVs

$$\begin{aligned} U_{\hat{s},i} &:= KMX_{\hat{s}}^{i-1} \quad \text{with} \qquad i \in \{1, 2, \cdots n\}, \\ U_{\hat{s}} &:= U_{\hat{s},T}T, \qquad X_{\hat{s}} := X_{\hat{s},T}, \qquad Y_s := Y_{s,T}, \end{aligned} \tag{6.10}$$

where T is considered to be a uniform RV on the set $\{1, 2, \cdots n\}$ and independent of all other RVs. Furthermore, $X_{\hat{s},i}$ is the i'th element in the sequence $X_{\hat{s}}^n$, and $Y_{s,i}$ is the i'th element in the sequence Y_s^n. It holds that

$$\begin{aligned} H(K) &= I(K; MY_s^n) + H(K|MY_s^n\hat{K}) \\ &\leq I(K; MY_s^n) + H(K|\hat{K}) \\ &\leq I(K; M) + I(KM; Y_s^n) + 1 + Pr(K \neq \hat{K}) \log |\mathcal{K}| \\ &\leq I(K; M) + \sum_{i=1}^{n} I(KMY_s^{i-1}; Y_{s,i}) + 1 + \delta \log |\mathcal{K}| \\ &\leq I(K; M) + \sum_{i=1}^{n} I(KMX_{\hat{s}}^{i-1}; Y_{s,i}) + 1 + \delta \log |\mathcal{K}| \\ &= I(K; M) + \sum_{i=1}^{n} I(U_{\hat{s},i}; Y_{s,i}) + 1 + \delta \log |\mathcal{K}| \\ &= I(K; M) + nI(U_{\hat{s},T}; Y_{s,T}|T) + 1 + \delta \log |\mathcal{K}| \\ &= I(K; M) + nI(U_{\hat{s}}; Y_s) + 1 + \delta \log |\mathcal{K}|, \end{aligned} \tag{6.11}$$

where the second inequality follows by Lemma 2.8 and the third one by condition (6.2) of Definition 6.2.

By using the inequality (6.11) and conditions (6.1), (6.3) and (6.4) of Definition 6.2 and the fact that $|\mathcal{K}| \leq |\mathcal{X}|^n$, it follows that

$$\begin{aligned} R_{\text{sk}} &\leq \frac{1}{n} H(K) + 2\delta \\ &\leq \frac{1}{n} I(K; M) + I(U_{\hat{s}}; Y_s) + \frac{1}{n} + \frac{\delta}{n} \log |\mathcal{X}|^n + 2\delta. \end{aligned}$$

Therefore, as $s \in \mathcal{I}(\hat{s})$ was chosen arbitrarily, it implies for δ sufficiently small and n large enough that

$$0 \leq R_{\text{sk}} \leq \min_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; Y_s). \tag{6.12}$$

Similarly as in [16, 30], it can be shown for all $i \in \{1, \cdots n\}$ and $(k, m, x^{i-1}, x_i, y_i) \in \mathcal{K} \times \mathcal{M} \times \mathcal{X}^{i-1} \times \mathcal{X} \times \mathcal{Y}$ that

$$\begin{aligned} &P_{KMX_{\hat{s}}^{i-1}X_{\hat{s},i}Y_{s,i}}(k, m, x^{i-1}, x_i, y_i) \\ &\qquad = P_{X_{\hat{s},i}}(x_i) P_{Y_{s,i}|X_{\hat{s},i}}(y_i|x_i) P_{KMX_{\hat{s}}^{i-1}|X_{\hat{s},i}}(k, m, x^{i-1}|x_i). \end{aligned}$$

This implies for all $s \in \mathcal{I}(\hat{s})$ that the Markov chains $U_{\hat{s},i} - X_{\hat{s},i} - Y_{s,i}$ hold. Therefore, it follows by using (6.10) that the following Markov chains also hold

$$U_{\hat{s}} - X_{\hat{s}} - Y_s. \tag{6.13}$$

In the last step, it holds by using the definitions in (6.10) and Lemma 2.8 that

$$\begin{aligned} I(X_{\hat{s}}^n; M) &= H(M) - H(M|X_{\hat{s}}^n) \\ &\geq H(M|Y_s^n) - H(KM|X_{\hat{s}}^n) \\ &\geq H(KM|Y_s^n) - H(K|\hat{K}) - H(KM|X_{\hat{s}}^n) \\ &\geq I(KM; X_{\hat{s}}^n) - I(KM; Y_s^n) - 1 - \delta \log |\mathcal{K}| \\ &\geq nI(U_{\hat{s}}; X_{\hat{s}}) - nI(U_{\hat{s}}; Y_s) - 1 - \delta \log |\mathcal{K}|. \end{aligned}$$

Thus, by using condition (6.5) of Definition 6.2, it follows that

$$L > I(U_{\hat{s}}; X_{\hat{s}}) - I(U_{\hat{s}}; Y_s) - \frac{1}{n} - \delta \log |\mathcal{X}| - \delta.$$

As $s \in \mathcal{I}(\hat{s})$ was taken arbitrarily, it follows for δ sufficiently small and n large enough that

$$\max_{s \in \mathcal{I}(\hat{s})} I(U_{\hat{s}}; X_{\hat{s}}|Y_s) \leq L. \tag{6.14}$$

The inequalities (6.12) and (6.14) together with Markov conditions (6.13) and the fact that $\hat{s} \in \hat{\mathcal{S}}$ was taken arbitrarily, imply that $(R_{\text{sk}}, L) \in \mathcal{R}^*(\mathfrak{S})$. This completes the proof of converse. □

6.3 Alphabet Size of Auxiliary Random Variables

In this section, we use Lemma 2.3 to determine an upper bound for the alphabet size of the auxiliary RV $U_{\hat{s}}$ with range $\mathcal{U}_{\hat{s}}$. In this case, the upper bound is given by

$$|\mathcal{U}_{\hat{s}}| \leq |\mathcal{X}| + |\mathcal{I}(\hat{s})|. \tag{6.15}$$

Remark. *As seen in* (6.15)*, for each marginal index $\hat{s}$, the upper bound for the size of $\mathcal{U}_{\hat{s}}$ depends on the size of compound index set $\mathcal{I}(\hat{s})$. In the special case of non-compound sources, this upper bound reduces to $|\mathcal{X}| + 1$ as given in [16].*

To prove (6.15), let $U_{\hat{s}}$ fulfill the Markov chains $U_{\hat{s}} - X_{\hat{s}} - Y_s$ for all $s \in \mathcal{I}(\hat{s})$. We show that there is an RV $\bar{U}_{\hat{s}}$, also satisfying the Markov conditions and the following equations

$$I(\bar{U}_{\hat{s}}; Y_s) = I(U_{\hat{s}}; Y_s), \tag{6.16}$$
$$I(\bar{U}_{\hat{s}}; X_{\hat{s}}) - I(\bar{U}_{\hat{s}}; Y_s) = I(U_{\hat{s}}; X_{\hat{s}}) - I(U_{\hat{s}}; Y_s), \tag{6.17}$$

for all $s \in \mathcal{I}(\hat{s})$ and range size $|\mathcal{X}| + |\mathcal{I}(\hat{s})|$. Using this RV instead of $U_{\hat{s}}$ will not affect the capacity results as it holds by using the property of Markov chain conditions that

$$I(U_{\hat{s}}; X_{\hat{s}}|Y_s) = I(U_{\hat{s}}; X_{\hat{s}}) - I(U_{\hat{s}}; Y_s).$$

Define the stochastic matrices $W_s \colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ with $W_s := P_{Y_s|X_{\hat{s}}}$. Define also the following $|\mathcal{X}| + |\mathcal{I}(\hat{s})|$ real valued continuous functions on $\mathcal{P}(\mathcal{X})$

$$f_x(P_{X_{\hat{s}}}) := P_{X_{\hat{s}}}(x) \quad \text{for every } x \in \mathcal{X} \text{ but one}, \tag{6.18}$$
$$f(P_{X_{\hat{s}}}) := H(P_{X_{\hat{s}}}), \tag{6.19}$$
$$g_s(P_{X_{\hat{s}}}) := H(P_{X_{\hat{s}}} W_s) \quad \text{for every } s \in \mathcal{I}(\hat{s}). \tag{6.20}$$

Furthermore, the following equations hold

$$\sum_{u \in \mathcal{U}_{\hat{s}}} P_{U_{\hat{s}}}(u) f_x(P_{X_{\hat{s}}|U_{\hat{s}}}(\cdot|u)) = P_{X_{\hat{s}}}(x) \qquad \text{for every } x \in \mathcal{X},$$
$$\sum_{u \in \mathcal{U}_{\hat{s}}} P_{U_{\hat{s}}}(u) f(P_{X_{\hat{s}}|U_{\hat{s}}}(\cdot|u)) = H(X_{\hat{s}}|U_{\hat{s}}),$$
$$\sum_{u \in \mathcal{U}_{\hat{s}}} P_{U_{\hat{s}}}(u) g_s(P_{X_{\hat{s}}|U_{\hat{s}}}(\cdot|u)) = H(Y_s|U_{\hat{s}}) \quad \text{for every } s \in \mathcal{I}(\hat{s}).$$

Similarly as in [18, Lemma 15.5], by using these equations and Lemma 2.3, it follows that for all functions in (6.18)-(6.20), there are $|\mathcal{X}| + |\mathcal{I}(\hat{s})|$ elements $P_u \in \mathcal{P}(\mathcal{X})$ with

$$u \in \bar{\mathcal{U}}_{\hat{s}} = \{1, \cdots, |\mathcal{X}| + |\mathcal{I}(\hat{s})|\}, \tag{6.21}$$

and $\alpha_u \geq 0$ with $\sum_{u=1}^{|\mathcal{X}|+|\mathcal{I}(\hat{s})|} \alpha_u = 1$, such that it holds

$$P_{X_{\hat{s}}}(x) = \sum_{u=1}^{|\mathcal{X}|+|\mathcal{I}(\hat{s})|} \alpha_u f_x(P_u), \tag{6.22}$$
$$H(X_{\hat{s}}|U_{\hat{s}}) = \sum_{u=1}^{|\mathcal{X}|+|\mathcal{I}(\hat{s})|} \alpha_u f(P_u), \tag{6.23}$$
$$H(Y_s|U_{\hat{s}}) = \sum_{u=1}^{|\mathcal{X}|+|\mathcal{I}(\hat{s})|} \alpha_u g_s(P_u). \tag{6.24}$$

By using condition (6.22), we may define the RV $\bar{U}_{\hat{s}}$ such that

$$P_{\bar{U}XY,s}(u, x, y) := \alpha_u P_u(x) P_{Y_s|X_{\hat{s}}}(y|x),$$

where $(u, x, y) \in \bar{\mathcal{U}}_{\hat{s}} \times \mathcal{X} \times Y$. This definition implies that the Markov chains $\bar{U}_{\hat{s}} - X_{\hat{s}} - Y_s$ hold. Furthermore, conditions (6.23) and (6.24) imply that the conditions in (6.16) and (6.17) also hold. Therefore, RV $\bar{U}_{\hat{s}}$ fulfills all the requirements and its range size is given by (6.21) and upper bounded by $|\mathcal{X}| + |\mathcal{I}(\hat{s})|$.

7 Conclusion

The SK capacity results which are given in this dissertation are first expressed for finite compound sources and then extended to arbitrary (possibly infinite) ones where the set of marginals is assumed to be finite.

The SK generation protocol for the finite case, uses a two phase approach to achieve the given SK rate. In the first step, Alice estimates her marginal $X_{\hat{s}}$ and sends $\hat{s}$ along with other information which is obtained from her observation to Bob. Although, this information is also received by Eve, it is shown that the strong secrecy and uniformity of the generated SK is still guaranteed. In the second step, Bob uses this information including the estimated marginal of Alice to generate an SK which satisfies the achievability conditions for all elements of $\mathfrak{I}(\hat{s})$. The SK capacity results are then given as a function of communication rate constraint. This model is also applied to the compound biometric authentication systems to determine the SK region when strong secrecy is guaranteed.

For extending the results to arbitrary (possibly infinite) compound sources, it is shown that for any infinite compound source, with a finite set of marginals, there exists a sequence of approximating finite compound sources whose SK generation protocols guarantee also the achievability of the given SK rate for the infinite source. The SK capacity results are given here again as a function of communication rate constraint. In case of degraded compound sources, the single-letter SK capacity is derived.

If the set of marginals $\hat{\mathfrak{S}}$ is also infinite, the SK rate might be smaller. This is because in this case Alice might not always be able to estimate her exact marginal $X_{\hat{s}}$. Even small estimation errors might lead to completely different compound sets $\mathfrak{I}(\hat{s})$, and thus different decoding results. An example of an infinite compound source with infinite set of marginals is given in [31]. The SK capacities for that source are different in case when $X_{\hat{s}}$ is exactly known to Alice and when this is not the case. In that work, an extra condition called regularity is put on the infinite compound sources (quantum) to derive the multi-letter SK capacity for the case when no communication rate constraint is taken into account.

As future work, this extra regularity condition can be applied to infinite compound sources (classical or quantum) to characterize the multi-letter SK capacity, where the communication rate parameter is also taken into account. In this case, the set of marginals would be also arbitrary and possibly infinite. This problem is still open and will be the subject of future research.

Author's Publication List

This list collects the author's publications which are used in this dissertation and appeared in conference proceedings or journals. They are also included in the Bibliography.

- N. Tavangaran, H. Boche, and R. F. Schaefer, "Secret-key generation using compound sources and one-way public communication," IEEE Trans. Inf. Forensics Security, vol. 12, no. 1, pp. 227-241, Jan. 2017.
- N. Tavangaran, H. Boche, and R. F. Schaefer, "Secret-key capacity of compound source models with one-way public communication," in Proc. IEEE Inf. Theory Workshop (ITW), Oct. 2015, pp. 252-256.
- N. Tavangaran, S. Baur, A. Grigorescu, and H. Boche, "Compound biometric authentication systems with strong secrecy," in Proc. Int. ITG Conf. Syst., Commun. Coding (SCC), Feb. 2017, pp. 1-5.
- N. Tavangaran, H. Boche, and R. F. Schaefer, "Secret-key capacity of infinite compound sources with communication rate constraint," in Proc. IEEE Int. Conf. Commun. (ICC), May 2017.
- N. Tavangaran, R. F. Schaefer, H. V. Poor, and H. Boche, "Secret-key generation and convexity of the rate region using infinite compound sources," IEEE Trans. Inf. Forensics Security, vol. 13, no. 8, pp. 2075–2086, Aug. 2018.

Preprint:

- N. Tavangaran, H. Boche, and R. F. Schaefer, "Secret-key generation using compound sources and one-way public communication," arXiv:1601.07513 [cs.IT], Jan. 2016.

Bibliography

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423,623–656, Oct. 1948.

[2] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[3] O. Goldreich, *Foundations of Cryptography, Basic Tools.* Cambridge University Press, 2004.

[4] G. Fettweis, H. Boche, T. Wiegand *et al.*, "The tactile internet," ITU-T Technology Watch Report, Aug. 2014. [Online]. Available: http://www.itu.int/oth/T2301000023/en

[5] R. F. Schaefer and H. Boche, "Physical layer service integration in wireless networks – signal processing challenges," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 147–156, May 2014.

[6] R. Wilson, D. Tse, and R. A. Scholtz, "Channel identification: Secret sharing using reciprocity in ultrawideband channels," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 364–375, Sept. 2007.

[7] Deutsche Telekom AG Laboratories, "Next generation mobile networks: (R)evolution in mobile communications," Technology Radar Edition III, Feature Paper, 2010. [Online]. Available: http://www.lti.ei.tum.de/en/people/boche

[8] U. Helmbrecht and R. Plaga, "New challenges for IT-security research in ICT," in *World Federation of Scientists, International Seminars on Planetary Emergencies, Erice, Italy*, Aug. 2008, pp. 1–6.

[9] A. D. Wyner, "The wire-tap channel," *Bell Sys. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[10] 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.

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

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

[13] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography - part II: CR capacity," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 225–240, Jan. 1998.

[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] L. Lai, S.-W. Ho, and H. V. Poor, "Privacy-security trade-offs in biometric security systems-part I: Single use case," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 122–139, Mar. 2011.

[16] T. Ignatenko and F. M. Willems, "Biometric security from an information-theoretical perspective," *Foundations and Trends in Communication and Information Theory*, vol. 7, no. 2, pp. 135–316, 2012.

[17] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 344–366, Mar. 2000.

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

[19] H. V. Poor and R. F. Schaefer, "Wireless physical layer security," in *Proc. Natl. Acad. Sci. U.S.A.*, vol. 114, no. 1, Jan. 2017, pp. 19–26.

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

[21] I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states," in *Proc. R. Soc. A 461*, 2005, pp. 207–235.

[22] P. Narayan and H. Tyagi, "Multiterminal secrecy by public discussion," *Foundations and Trends in Communications and Information Theory*, vol. 13, no. 2, pp. 129–275, 2016.

[23] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4, pp. 355–580, 2009.

[24] R. F. Schaefer, H. Boche, A. Khisti, and H. V. Poor, *Information Theoretic Security and Privacy of Information Systems*. Cambridge, UK: Cambridge University Press, 2017.

[25] S. Watanabe and Y. Oohama, "Secret key agreement from vector Gaussian sources by rate limited public communication," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 541–550, Sept. 2011.

[26] T. Ignatenko and F. M. Willems, "Biometric security: Privacy and secrecy aspects," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 4, pp. 956–973, 2009.

[27] T. Ignatenko and F. M. Willems, "Privacy leakage in biometric secrecy systems," in *Proc. 46th Annual Allerton Conference on Communication, Control, and Computing*, Sept. 2008, pp. 850–857.

[28] M. Bloch, "Channel intrinsic randomness," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2010, pp. 2607–2611.

[29] H. Boche and R. F. Wyrembelski, "Secret key generation using compound sources - optimal key-rates and communication costs," in *Proc. 9th Int. ITG Conf. Syst., Commun. Coding (SCC)*, Jan. 2013, pp. 1–6.

[30] A. Grigorescu, H. Boche, and R. F. Schaefer, "Robust biometric authentication from an information theoretic perspective," *Entropy*, vol. 19, no. 9, Sept. 2017.

[31] H. Boche and G. Janßen, "Distillation of secret-key from a class of compound memoryless quantum sources," *J. Math. Phys. 57, 082201*, Aug. 2016.

[32] R. F. Schaefer, H. Boche, and H. V. Poor, "Secure communication under channel uncertainty and adversarial attacks," in *Proc. IEEE*, vol. 103, no. 10, Oct. 2015, pp. 1796–1813.

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

[34] S. C. Draper and E. Martinian, "Compound conditional source coding, Slepian-Wolf list decoding, and applications to media coding," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2007, pp. 1511–1515.

[35] R. A. Chou and M. R. Bloch, "Secret-key generation with arbitrarily varying eavesdroppers channel," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2013, pp. 277–280.

[36] I. Bjelaković, H. Boche, and J. Sommerfeld, "Secrecy results for compound wiretap channels," *Problems Inf. Transmiss.*, vol. 49, no. 1, pp. 73–98, 2013.

[37] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wiretap channels," *EURASIP Journal on Wireless Communications and Networking*, no. 142374, 2009.

[38] R. F. Schaefer and S. Loyka, "The secrecy capacity of compound Gaussian MIMO wiretap channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5535–5552, July 2015.

[39] H. Boche, R. Schaefer, and H. V. Poor, "On the continuity of the secrecy capacity of compound and arbitrarily varying wiretap channels," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 12, pp. 2531–2546, Dec. 2015.

[40] R. F. Schaefer, A. Grigorescu, H. Boche, and H. V. Poor, "Broadcast channels with confidential messages: Channel uncertainty, robustness, and continuity," *in Physical and Data-Link Security Techniques for Future Communication Systems, Springer*, vol. 358, pp. 69–91, 2016.

[41] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 44, no. 2, pp. 159–175, 1978.

[42] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels," *Ann. Math. Stat.*, vol. 30, no. 4, pp. 1229–1241, Dec. 1959.

[43] Z. Zhang, "Estimating mutual information via Kolmogorov distance," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3280–3282, Sept. 2007.

[44] N. Tavangaran, H. Boche, and R. F. Schaefer, "Secret-key generation using compound sources and one-way public communication," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 227–241, Jan. 2017.

[45] N. Tavangaran, H. Boche, and R. F. Schaefer, "Secret-key generation using compound sources and one-way public communication," *arXiv:1601.07513 [cs.IT]*, Jan. 2016.

[46] N. Tavangaran, H. Boche, and R. F. Schaefer, "Secret-key capacity of compound source models with one-way public communication," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Oct. 2015, pp. 252–256.

[47] N. Tavangaran, S. Baur, A. Grigorescu, and H. Boche, "Compound biometric authentication systems with strong secrecy," in *Proc. Int. ITG Conf. Syst., Commun. Coding (SCC)*, Feb. 2017, pp. 1–5.

[48] N. Tavangaran, H. Boche, and R. F. Schaefer, "Secret-key capacity of infinite compound sources with communication rate constraint," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017.

[49] H. Bauer, *Maß- und Integrationstheorie*, 2nd ed. De Gruyter, 1992.

[50] A. Klenke, *Wahrscheinlichkeitstheorie*, 3rd ed. Springer Spektrum, 2013.

[51] J. R. Munkres, *Topology*, 2nd ed. Pearson, 2000.

[52] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. McGraw-Hill, 1976.

[53] R. Alicki and M. Fannes, "Continuity of quantum conditional information," *J. Phys A*, vol. 37, no. 5, pp. L55–L57, 2004.

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

[55] I. Csiszár, "Almost independence and secrecy capacity," *Problems Inf. Transmiss.*, vol. 32, no. 1, pp. 48–57, 1996.

[56] M. Fekete, "Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten," *Mathematische Zeitschrift*, vol. 17, pp. 228–249, 1923.

[57] I. Bjelaković, H. Boche, and J. Nötzel, "On quantum capacity of compound channels," *Physical Review A*, vol. 78, no. 4, p. 042331, 2008.

[58] I. Bjelaković, H. Boche, and G. Janßen, "Universal quantum state merging," *J. Math. Phys. 54, 032204*, Mar. 2013.

[59] M. Bloch and J. Barros, *Physical-Layer Security, From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[60] N. Tavangaran, R. F. Schaefer, H. V. Poor, and H. Boche, "Secret-key generation and convexity of the rate region using infinite compound sources," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 8, pp. 2075–2086, Aug. 2018.