Fundamental Limits of Biometric Identification Systems with Noisy Enrollment

Vamoua Yachongka

Department of Computer and Network Engineering
The University of Electro-Communications

This dissertation is submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

February 4, 2021

Supervisory Committees

Chairperson: Associate Professor Hideki Yagi

1. Member: Professor Tsutomu Kawabata

2. Member: Professor Yasutada Oohama

3. Member: Associate Professor Tomohiro Ogawa

4. Member: Associate Professor Mitsugu Iwamoto

Date of the final defense: February 4, 2021.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor, Associate Professor Hideki Yagi, for his consistent support. His kind guidance is very helpful for me to build up the basic skills needed for doing research, and without that this dissertation will not work out the same. To be honest, as I am the first ever Ph.D. candidate in his Lab, there are few people that I can talk or consult about the difficulties regarding academic and personal endeavours that I have gone through during these intense academic years. He has provided great care for me, which makes my life here in Japan easier and I would be able to fully focus on conducting my research.

I am grateful to all members in the joint research group (information theory laboratory) where I belong to. The meetings and conversations with them are inspiring me to learn from different perspectives and think outside the box.

Moreover, I want to thank Professor Tsutomu Kawabata, Professor Yasutada Oohama, Associate Professor Tomohoro Ogawa, and Associate Professor Mitsugu Iwamoto, for accepting to be a part of supervisory committees for this dissertation.

Finally, I cannot forget to thank my family, relatives, and friends who have been warmly watching and encouraging me.

Abstract

Biometric identification is a procedure of matching the digital files of users in the system database with the presented biological data. In this thesis, we treat biometric identification systems as a mathematical model and analyze its fundamental performances by information theoretic methods. Two models of biometric identification systems are being investigated. The first scenario is the biometric identification system for secret key-based identification and authentication, and the second one is the system dealing with both chosen and generated secrecy. In the latter scenario, we concentrate on two settings: discrete memoryless and Gaussian sources.

The first model can be categorized further into two submodels: generated-secret and chosen-secret biometric identification system models. Ignatenko and Willems (2015) investigated these models with visible sources, where the enrollment channel is noiseless, and they clarified the fundamental trade-off between identification and secrecy rates under a privacy constraint. However, when the biometric data is scanned for enrolling, it is likely that noise is added to the extracted sequence, and so as to reduce the cost of hardware architecture, it is important to constrain the size of the storage. We improve the models by assuming hidden sources, where the enrollment channel is noisy, and adding the constraint on the size of the storage. We derive the optimal trade-off between identification and secrecy rates of the two models under both privacy and storage constraints. As special cases, our results agree with the ones given by Ignatenko and Willems (2015), and coincide with the results derived by Günlü and Kramer (2018) in which there is only one individual.

In the second model, two secret keys are used together. That is, in the enrollment phase, the encoder encodes biometric data sequence with a secret key (chosen-secret key), chosen independently, to generate helper data and another secret key (generated-secret key). In the identification phase, the decoder should estimate the identified user and her two secret keys reliably. Here, we allow the two keys to be correlated. We characterize the capacity region among identification, chosen- and generated- secrecy, template, and privacy-leakage rates of the system for discrete memoryless sources. As a result, a larger sum of identification, chosen- and generated-secrecy rates is achieved due to permitting the correlation, and when the sum of the identification and chosen-secrecy rates increases, a larger storage space for storing the templates is required, but the generated-secrecy rate does not affect the memory space. Furthermore, only the changes of the identification rate directly affect the minimum value of the privacy-leakage. Later, we extend the model with both chosen and generated secrecy to Gaussian sources. We introduce a technique for deriving the capacity region of these rates by converting the system to one where the data flow is in one-way direction. Also, we provide

numerical calculations of three different examples, and as a result, it seems difficult to achieve both high generated secret key rate and small privacy-leakage at the same time.

Table of contents

List of figures

List of tables

Nomenclature

Roman Symbols

$d(\cdot)$	System decoder
$e(\cdot)$	System encoder
$H(\cdot)$	The Shannon entropy of discrete random variable
$h(\cdot)$	Differential entropy
$H_b(\cdot)$	Binary entropy
i	The index of user i
$j(i)$	The template of user i
k	Variable
M_I	The number of users in the BIS
M_J	The number of templates (codewords) in the database
M_S	The number of secret keys
M_C	The number of chosen-secret keys
M_G	The number of generated-secret keys
$I(\cdot;\cdot)$	Mutual information
n	Bloch length
$\Pr[\mathcal{E}]$	The probability of the even $\mathcal{E}$
R_C	The chosen-secrecy rate
R_G	The generated-secrecy rate
R_I	The identification rate

R_J	The template rate
R_L	The privacy-leakage rate
R_S	The secrecy rate for the model where the chosen- and generated-secret keys are treated separately
$s(i)$	The secret key of user i
$\mathcal{T}_{\varepsilon}^{(n)}(\cdot)$	The strongly ε-typical set
$\mathcal{B}_{\varepsilon}^{(n)}(\cdot)$	A modified ε-typical set
$\mathcal{A}_{\varepsilon}^{(n)}(\cdot)$	The weakly ε-typical set
X	A random variable
$\mathcal{X}$	A set and its size can be either finite or infinite
x^n	The biometric data sequence
y^n	The biometric data sequence observed via the enrollment channel
z^n	The biometric data sequence observed via the identification channel

Greek Symbols

α	A positive value lies in the range of $(0,1]$
Γ	A positive value
δ	A small enough positive value
δ_n,δ'_n,ε_n	Positive values converge to zero when $n \to \infty$ and $\delta \downarrow 0$
ε	A small enough positive value
η	A positive value lies in the range of $[1, \frac{1}{1-\rho_1^2\rho_2^2})$
Γ	A positive value
γ	A positive value lies in the range of $[0,0.5]$
ρ_1,ρ_2	The Pearson's correlation coefficients

Acronyms / Abbreviations

AEP	Asymptotic equipartition property
auth.	Authentication

Bio-data	Biological data or biometric data
BIS	Biometric identification system
DMC	Discrete memoryless channel
DMS	Discrete memoryless source
EPI	Entropy power inequality
HSM	Hidden source model
id.	Identification
i.i.d.	Independent and identically distributed
MGL	Mrs. Gerber's lemma
PDF	Probability density function
RV	Random variable
s. t.	Such that
VSM	Visible source model

Chapter 1

Introduction

In this chapter, we go through some common knowledge of biometric identification, introduce previous works of this study, state our goals and motivations, and finally brief organizations of the present thesis.

1.1 Background

When we hear the term biometrics, our imagination often links to the captured pictures such as fingerprints, faces, and irises. In fact, the term relates to a science of identifying users using physical characteristics, or *biological data* (bio-data) [47] where a certain part of human body is transformed into a matrix form so that the processor is able to compute or recognize it. Biometric identification indicates an automated process of recognizing individual by matching the individual's biometrics with

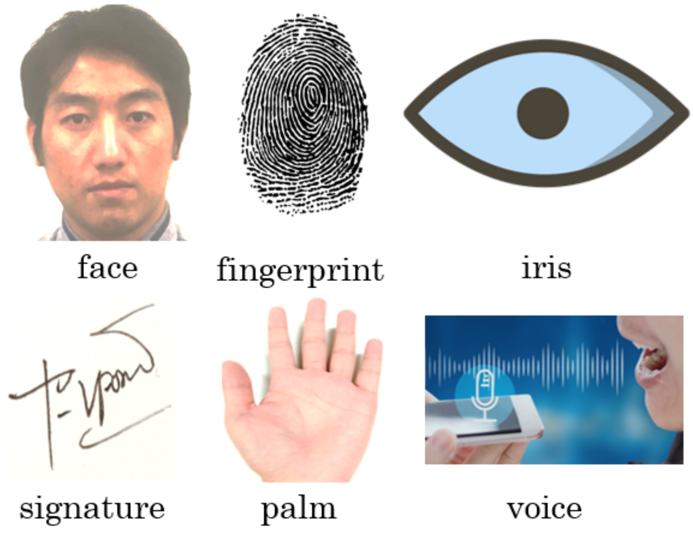

Fig. 1.1 Biometric modalities; Some examples of body traits that can be used for biometric recognition.

the references in the system database [48]. Basically, any bio-data can be used for biometric identification, but some common ones include fingerprint, iris, face, voice, palm, and so on [36] (cf. Fig. 1.1). The history of the *biometric identification systems* (BISs) is not very new. It is believed that its technologies have been initially used in the late age of the 19th century [1] as a purpose for criminal detection. At that time, fingerprints were a major resource to identify the suspects due to the limited technologies available [37], but since then, the researches related to biometric identifications have been rapidly developed. Nowadays we could enjoy many convenient applications utilizing biometric technologies based on different modalities. Examples include, but not limited to, online mobile payment using fingerprints, face scan at airport, sound recognition for granting access into a data center, handwritten signature for digital documents, and for more details, the reader should refer to [34].

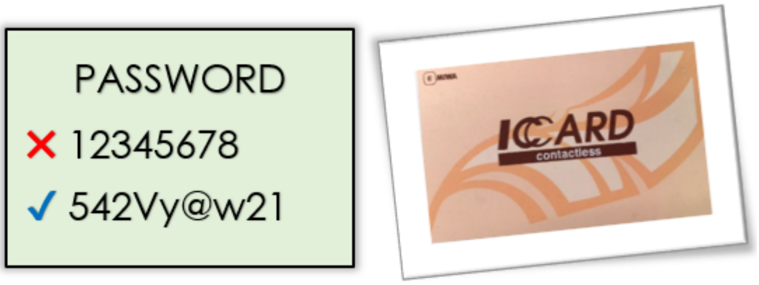

Fig. 1.2 Traditional methods for identification; Password and IC-card or smart-card based identification are well-known conventional methods for user's identification.

Traditional methods, as shown in Fig. 1.2, for identifying individuals are based on information that we know (passwords) or something that we own (smart cards or tokens) [47], [36]. In these methods, the disadvantages are that the passwords can be predicted by the intruder since users trend to use easy ones, and in case the identified user forgets her password, it is impossible to gain system access. Moreover, smart cards or tokens can be stolen, and thus the abused issue becomes a great concern [13], [64]. On the other hand, biometrics based identification provides a promising solution to the issues mentioned above since biometric traits are unique to each individual [13], [35], [64], and the samplings (scanned version) of the bio-data basically is in structure of a large dimensional matrix, which is hard for imposter to guess. Even the imposter tries to deceive, the machine can easily detect such cheats [3], [46]. Although biometric identification can provide a greater confidence in terms of security concerns, of course, dark sides for the method exist as well. Unlike the passwords or smart-cards, which can be changed at any time, the bio-data has no or every few alternatives [52], [64]. Bio-data is not easily revoked and it might end up in redesigning the feature extractor of the system [53]. Therefore, privacy protection of the users is crucial for designing the BIS [57], [89]. Indeed, it is required to ensure that user's privacy is well protected or leaked only negligible amount even the database is hacked. Lastly, one more important indicator is the constraint on the storage device. It should be minimized to save the memory space and reduce the cost of hardware architectures, especially, when a large number of users is using the system [16], [21].

1.2 Related Works

Generally speaking, cancelable biometrics and biometric secrecy systems are two main approaches in researches regarding BISs in recent years. The former approach focuses on transforming the bio-data in the feature domain, and matched in the transformed domain directly. In this fashion, the requirement of storage (database) is avoidable. However, the weak point of this method is that it is always hard to construct transformations with explicitly provable security. For this approach, we do not go into deep details. For those who are interested, a understandable explanation of cancelable biometrics can be seen in Jain et al. [35] or Ratha et al. [58]. On the other hand, the latter one is based on the concept of information-theoretic method in which Shannon's entropy and mutual information [65] become important indicators for evaluating security of the system. In this thesis, we deploys the latter approach to clarify the fundamental performances of the BIS.

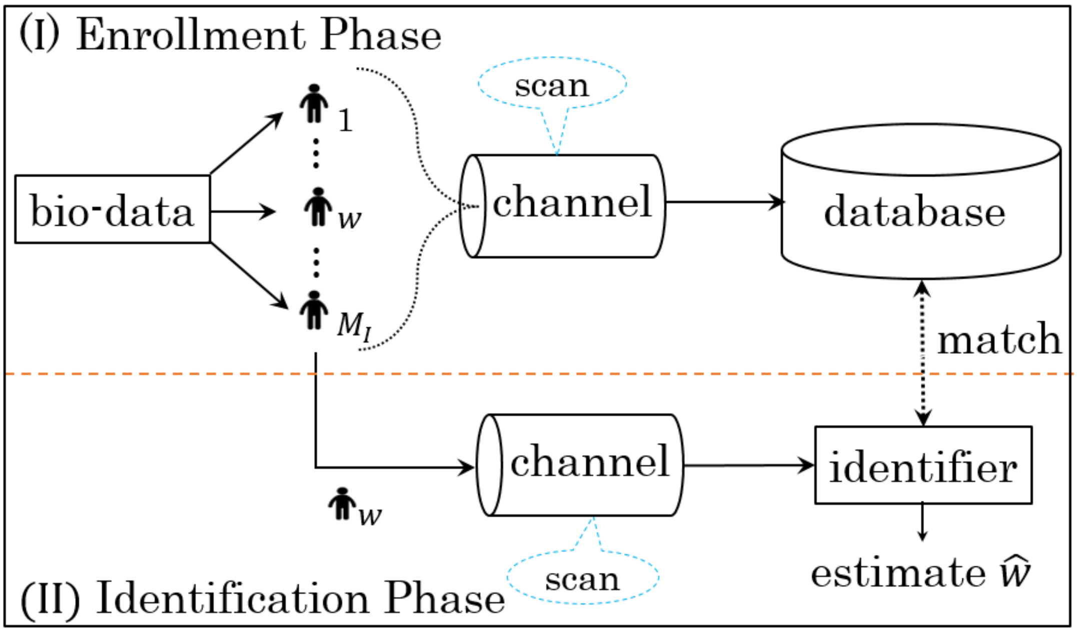

Fig. 1.3 A simple BIS model; A BIS consists of two parts: Enrollment and Identification Phases.

A simple BIS model is illustrated in Fig. 1.3. A detailed explanation of this model is provided in Section 2.3. Here, only a short and simple description is given, which it can help the readers to grasp the scenario of this part. In general, the BIS consists of two phrases: Enrollment and Identification Phases. Assume that there are M_I users utilizing the system. In Enrollment Phase, bio-data of these users are enrolled into the system database via a scanner. In Identification Phase, a user w presents his/her identify to the identifier and it estimates the user based on the information inside the database. By focusing on only user w, the data flow of the user is shown in Fig. 1.4. In this figure, we suppose that the biometric source is a thumb. The original shape of the thumb is represented by x_w^n, generated from source P_X. It is scanned through a scanner, modeled as $P_{Y|X}$, in Enrollment Phase and output as y_w^n. The sequence y_w^n is stored in the database for future identification. In Identification Phase,

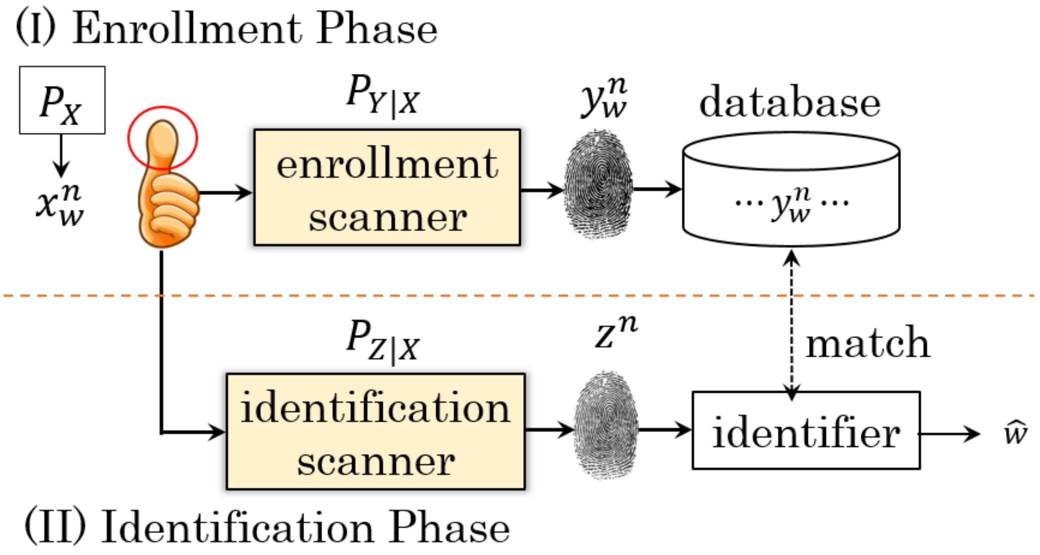

Fig. 1.4 Modeling of user w: This is an illustration of how user w is modeled within information-theoretic framework.

the same thumb is again scanned via a scanner, modeled as $P_{Z|X}$, and the output sequence is z^n. The identifier sees z^n, and finally predicts who is being identified.

O'Sullivan and Schmid [56] and Willems et al. [76] initially investigated the fundamental performances of the BIS from information-theoretic point of views. In [56], they analyzed the error exponent of the BIS and showed the maximum identification rate that guarantees the error probability converges to zero. On the other hands, Willems et al. [76] took a different approach and used arguments of channel coding, e.g., the asymptotic equipartition property (AEP) and Fano's inequality, to characterize the identification capacity of the BIS. The identification capacity means the maximum achievable rate of the number of individuals as the error probability converges to zero when the length of bio-data sequences goes to infinity. They proved that the error probability of the BIS trends to zero if and only if (iff) the identification rate is below the identification capacity $I(Z;Y)$ (cf. Theorem 2.1). In [56] and [76], however, it is assumed that bio-data sequences are stored in the system database without encoded, so some critical problems like enormous storage consuming and magnificent privacy-leakage could happen. Tuncel [69], [70] extended the model in [76] by incorporating compression of bio-data sequences before stored in the system database. More specifically, he considered an encoder in the enrollment phase to generate helper data (in this entire thesis, the helper data and its rate are called *template* and *template rate*, respectively) for bio-data sequences and the helper data or templates corresponding to each bio-data sequence are saved in the system database. The fundamental trade-off between identification and compression rates was characterized in [69] for single stage and [70] for both single and multiple stages. Some extended works of [69], [70] can be found in [71] to recover noisy reconstruction (rate-distortion) and in order to speed up the search complexity of the BIS, hierarchical identification with a pre-processing at the decoder is considered in [75]. Error exponents of the BIS is examined in [80] from Arimoto's arguments [5], and [91] and [92] from information spectrum perspectives [27].

For the purpose of security, the BIS incorporating the generation of both a secret key and a helper data for each user in the enrollment phase has been found in many studies, e.g., [31], [32], [33], [40], and [81]. In this system, not only the index, but also the secret key of the identified user is estimated in the identification phase. In general, there are two types of models: *generated-secret* and *chosen-secret* BIS models are frequently discussed. In the generated-secret BIS model, the secret key is extracted from bio-data sequences, while in the chosen-secret BIS model, the secret key is chosen independently of the bio-data sequences. These names came from the literature [21], and we also use the terms to call the models for the sake of the readers. Note that in the field of cryptography, the meaning of the word "chosen" is closer to "arbitrary", for instances, chosen-ciphertext attack, chosen-plaintext attack, etc. Here, we simply use it to indicate the operation that the secret key is chosen uniformly and independently of other RVs from a set. These models can be viewed as the BIS supporting authentication since the estimated index informs whom the identified user is and the estimated secret key may be used as a tool to claim that it is actually the same user trying to access the system. Ignatenko and Willems [32], [33] have characterized the capacity regions of identification, secrecy, and privacy-leakage rates in the two models. Here, privacy-leakage is defined as the amount of information that leaks from a template to its original bio-data sequence. Fundamentally, the privacy-leakage is unavoidable and it is hard to make this value negligibly small [66]. The reason is because the templates are generated from the bio-data sequences and their correlations remain at certain level. In order to achieve negligible privacy-leakage, an private key available to both the encoder and decoder was introduced in [29], [31, Section 3.5], [90]. Instead of analyzing the privacy-leakage rate, a model considering the template rate in the generated-secret BIS model can be found in [81]. In [81], the authors came to a conclusion that the template and privacy-leakage rates are equivalent. Furthermore, Kittichokechai and Caire [40] applied the concept of broadcast channel to assume the presence of an adversary in the generated-secret BIS model. In [40], a constraint on the storage was also taken into account. However, a common assumption in above studies is that the identified user is uniformly distributed and the BIS is analysed under *visible source model* (VSM). In the VSM, it is assumed that the enrollment channel of the BIS is noiseless, and thus the encoder is able to observe the biometric source sequences directly.

Another stream of studies deals with only the estimation of one user's secret key. This type of the BIS can be viewed as key-agreement model between two terminals. Ahlswede and Csiszár[2], Csiszár and Narayan [16], and Maurer [49] analyzed the model without imposing privacy constraint. The privacy constraint was taken into consideration in the literature, e.g., Ignatenko and Willems [29], Lai et al. [43], [44], Willems and Ignatenko [77], Koide and Yamamoto [42], and Günlü and Kramer [21]. More specifically, Ignatenko and Willems [29] and Lai et al. [43], [44] investigated the fundamental trade-off of secrecy and privacy-leakage rates for *dicreste memoryless soruces* (DMSs), and Willems and Ignatenko [77] provided the trade-off of these rates for *Gaussian sources*. Koide and Yamamoto [42] constrained the template to the model considered in [29], and loosened the secrecy-leakage condition. Günlü and Kramer [21] made a successful attempt for characterizing the capacity region of the *hidden source model* (HSM), where the enrollment channel is noisy and the encoder can only see the noisy sequences of biometric source. Furthermore, fuzzy commitment scheme, allowing noisy data to be used as input for generating a secret key, and in the phase of construction, the secret key is estimated from close values, but not necessarily identical to the original one without sacrificing the security requirement, is an authentication method that is quite similar to the concept of the BIS with single user (with no estimation of the index). The scheme was proposed by Juels and Martin [39], and developed in [18], [66] with adopting error correcting codes. These studies designed codes based on Hamming distance, edit distance, and set difference to asses the maximum secrecy rate and the minimum entropy loss, corresponding to the maximum privacy-leakage. However, to reduce the risk that original bio-data sequence of a user is being hacked by the attackers, the amount of the privacy-leakage should be minimized [29]. Here, instead of constructing certain codes, we focus on deriving the fundamental performances of the BIS by applying information theoretic approaches.

1.3 Goals of This Thesis

The goals of this thesis are listed as follows:

- First, we extend the system models considered in [33] to incorporate the HSM and add a constraint on the storage. Also, we assume that the prior distribution of the identified user is unknown. We want to find the optimal trade-off between identification and secrecy rates for both generated- and chosen-secret BIS models under privacy and storage constraints.
- Second, we analyze a model in which the chosen- and generated-secret keys are used together, and aim to characterize the fundamental limits among identification, chosen- and generated-secrecy, template, and privacy-leakage rates of the BIS for DMSs.
- Lastly, we further extend the model of the second goal to Gaussian sources and channels.

1.4 Motivations

The motivations for each goal of this thesis are summarized as follows:

For our first goal, this extension is considered to be more natural from the following reasons.

- When users enroll their bio-data to the database for identification, the bio-data must be captured via a scanner, and through this process, there is high possibility that noise is added to the captured version. Hence, the encoder possibly cannot access to the perfect source sequences, and assuming the HSM for evaluating the fundamental performances of the BIS is more natural setting.
- In order to reduce the cost of hardware architecture, it is important to constrain the template rate.

- In real-life situation, the frequencies of coming to use the system for each user are unlikely identical, so it would be more realistic to analyze the BIS under the condition that the prior distribution of the identified user is uniform.

For the second goal, in the previous studies, e.g., [21], [29], [33], [85], the chosen- and generated-secret keys are assumed in the separate models, namely, chosen- and generated-secret BIS models, respectively. However, an interesting question is when the two keys are used in the same system, how the chosen- and generated-secrecy rates affect the fundamental performances of the BIS. The answer to this question has not yet been known, and it is not trivial from the results of the previous studies. Yet, we allows chosen- and generated-secret keys to be correlated at some level. The reason behind the scene of this is that we wish to achieve a larger sum rate of identification, chosen- and generated-secrecy rates. Another motivation is that we want to figure out the fundamental trade-off of the BIS supporting two-factor authentication, where in this case, chosen- and generated-secret keys can be used for the first and second stages of authentications.

For the analysis of Gaussian sources, we are motivated by the fact that the signal of bio-data is basically represented by vectors with continuous elements in real applications, and most communication links can be modeled as Gaussian channels. Moreover, in [77], the fundamental trade-off of secrecy and privacy-leakage rates in the BIS with the VSM for Gaussian case was clarified. However, when the model is shifted from the VSM to the HSM, the problem becomes more challenging and many techniques used for deriving the capacity region of the VSM is not directly applicable [21]. Thus, analysing the BIS for Gaussian sources is of both theoretical and practical interest.

1.5 Organizations of This Thesis

This thesis is organized as follows:

In Chapter 2, we define the notation used in the subsequent chapters, recap some well-known results in information theory, and introduce the simple model analyzed in [76].

Chapter 3 focuses on characterizing the capacity regions among identification, secrecy, template, and privacy-leakage rates for both generated- and chosen-secret BIS models. We extend the model considered in [32] and [33] to include the noisy enrollment and constrain the storage. We derive the regions via two auxiliary random variables (RVs). The obtained results show that in the generated-secret BIS model, like the results derived in [33], [82], the minimum required amounts of the privacy-leakage and template rates vary based on the number of users. However, different to a conclusion in [81] (the result of the VSM), in which the minimum amount of the privacy-leakage rate is smaller than the template rate, the two rates are bounded In the chosen-secret BIS, we need to store the templates at full rate in order to reconstruct the secret key reliably. We also simplify the derived regions of the generated-secret BIS model for binary hidden sources by applying Mrs. Gerber's lemma (MGL), and give numerical results of an example concerning the simplified rate region.

Chapter 4 addresses on the BIS with both chosen and generated secrecy. We provide the optimal trade-off of identification, chosen- and generated-secrecy rates in the BIS under privacy and storage constraints. Additionally, we allow correlation between the two secret keys. As a result, it turns out that the identification, chosen- and generated-secrecy rates are in a trade-off relation, and a bigger sum of these rates is achievable compared to the result in [86]. The minimum amount of the template rate belongs to both identification and chosen-secrecy rates, but that of the privacy-leakage rate is only affected by the identification rate, and has nothing to do with the chosen-secrecy rate. Like in Chapter 3, we simplify the derived region for binary hidden source and give numerical calculations for an example.

In Chapter 5, we extend the model considered in Chapter 4 to Gaussian sources and channels, and characterize the capacity region of identification, chosen- and generated-secrecy, template, and privacy-leakage rates of the BIS. We introduce an idea of converting the system to another one where the data flow of each user is in the same direction, which enables us to characterize the capacity region. More specifically, in establishing the outer bound of the region, the converted system allows us to use the well-known *entropy power inequality* (EPI) [65] twice in two opposite directions, and also its property facilitates the derivation of the inner bound. We also provide numerical computations of three different examples. From the results of these examples, we may conclude that it is difficult to achieve both high secrecy and small privacy-leakage rates together. To achieve a small privacy-leakage rate, the gain of the secrecy rate is scarified somehow.

In Chapter 6, we provide some concluding remarks and future directions for this thesis

This thesis is written based on the author's joint works which are partially published in [83], [84], [86], and [87].

Chapter 2

Preliminaries

In this chapter, we first introduce and define notation that is used in this paper. After that, we summarize some basic results in information theory, which are useful in the arguments of upcoming chapters, and introduce the result of pioneer research [76].

2.1 Notation and its Definition

Calligraphic letter $\mathcal{A}$ stands for a finite set and its cardinality is written as $|\mathcal{A}|$. Upper-case such that A and lower-case $a \in \mathcal{A}$ denote a RV and its realization, respectively. $A^n = (A_1, A_2, \cdots, A_n)$ represents a string of RVs, taking values in $\mathcal{A}^n$, and subscripts represent the position of a RV in the string. $P_A(a) = \Pr[A = a]$ represents the probability distribution on $\mathcal{A}$ and P_{A^n} represents the probability distribution of RV $A^n \in \mathcal{A}^n$. In case A is continuous RV, the probability density function (PDF) of A is denoted by f_A. $P_{A^nB^n}$ represents the joint probability distribution of a pair of RVs (A^n, B^n) and its conditional probability distribution $P_{A^n|B^n}$ is defined as

$$P_{A^n|B^n}(a^n|b^n) = \frac{P_{A^nB^n}(a^n, b^n)}{P_{B^n}(b^n)} \tag{2.1}$$

for any $a^n \in \mathcal{A}^n$, $b^n \in \mathcal{B}^n$ such that $P_{B^n}(b^n) > 0$. Integers a and b such that $a < b$, $[a:b]$ denotes the set $\{a, a+1, \cdots, b\}$. A partial sequence of a sequence c^n from the first symbol to the tth symbol $(c_1, c_2, \cdots, c_t)$ is represented by c^t. For $x > 0$, $\log x$ and $\ln x$ stand for the base of two and natural logarithm, respectively.

A standard information measure in information theory is entropy and mutual information. The entropy tells us about the average value of information or uncertainty inherent in a RV. Its concept was introduced by Shannon in his pioneering paper [65] in 1948. There are also other types of entropy such as Hartley or max-entropy [26], collision or quadratic entropy, min-entropy, and Rényi entropy [59]. The Rényi entropy generalizes the Hartley entropy, the Shannon entropy, the quadratic entropy, and the min-entropy, corresponding the cases in which The Rényi order is equal to $0, 1, 2$, and ∞, respectively. In this thesis, however, we use only the Shannon's entropy [65]. $H(\cdot)$ and $h(\cdot)$ denote the entropy of discrete and continuous RV, respectively. On the other hand, mutual information is a quantity that how much one RV tells about another. The mutual information between RVs A and B is denoted by $I(A;B)$. $H_b(a) = a\log\frac{1}{a} + (1-a)\log\frac{1}{(1-a)}$ denotes the binary entropy for $0 \leq a \leq 1$ and $H_b(0) = H_b(1) = 0$, and $H_b^{-1}(\cdot)$ is inverse function of $H_b(\cdot)$. Overall, we basically follow the standard notation in Cover and Thomas [14], and El Gamal and Kim [19].

2.2 Basic Results in Information Theory

In this section, we will review many classical results in information theory. First, we begin with the definition of weakly ε-typical set. The definition holds for both discrete and continuous RV, but here we provide only the continuous version. It is formally defined below, see also in [14, Chapter 9].

Definition 2.1. *(Weakly ε-typical set for continuous RVs [14, Chapter 9])*

Assume $(X_1, X_2, \cdots, X_k)$ be a finite collection of continuous RVs with joint probability density function (PDF) $f_{X_1X_2\cdots X_k}(x_1, x_2, \cdots, x_k)$ and differential entropy $h(X_1, X_2, \cdots, X_k)$. $f_V(v)$ is marginal PDF of the joint PDF $f_{X_1X_2\cdots X_k}(x_1, x_2, \cdots, x_k)$ with differential entropies $h(V)$, where RV $V \subseteq \{X_1, X_2, \cdots, X_k\}$. The jointly ε-typical set, denoted by $\mathcal{A}_\varepsilon^{(n)}(X_1X_2\cdots X_k)$, is the set of sequences $(x_1^n, x_2^n, \cdots, x_k^n) \in \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{k}$ satisfying:

$$\mathcal{A}_\varepsilon^{(n)}(X_1X_2\cdots X_k) = \left\{ (x_1^n, x_2^n, \cdots, X_k^n) : \left| -\frac{1}{n}\log f_{V^n}(v^n) - h(V) \right| \leq \varepsilon \right\}, \tag{2.2}$$

where $v^n \subseteq \{x_1^n, x_2^n, \cdots, x_k^n\}$ corresponding to RV V and $f_{V^n}(v^n) = \prod_{k=1}^n f_{V_k}(v_k)$. Moreover, the conditional typicality is defined as

$$\mathcal{A}_\varepsilon^{(n)}(X_k|x_2^n, \cdots, x_{k-1}^n) = \left\{ X_k^n : (x_1^n, x_2^n, \cdots, X_k^n) \in \mathcal{A}_\varepsilon^{(n)}(X_1X_2\cdots X_k) \right\}. \tag{2.3}$$

Next, we provide several properties regarding the weakly ε-typical set.

Lemma 2.1. *(Some properties of weakly ε-typical set [14], and [31, Lemma A.1])*

Let $\varepsilon > 0$. We have that

1) From (2.2) and $\forall V \subseteq \{X_1, X_2, \cdots, X_k\}$, it follows that

$$2^{-n(h(V)+\varepsilon)} \leq f_{V^n}(v^n) \leq 2^{-n(h(V)-\varepsilon)}. \tag{2.4}$$

2) For $\forall V \subseteq \{X_1, X_2, \cdots, X_k\}$ and large enough n

$$\Pr\{v^n \in \mathcal{A}_\varepsilon^{(n)}(V)\} \geq 1 - \varepsilon. \tag{2.5}$$

3) For $\forall V \subseteq \{X_1, X_2, \cdots, X_k\}$, it holds that

$$(1-\varepsilon)2^{n(h(V)-\varepsilon)} \leq |\mathcal{A}_\varepsilon^{(n)}(V)| \leq 2^{n(h(V)+\varepsilon)}. \tag{2.6}$$

4) For $\forall V, W \subseteq \{X_1, X_2, \cdots, X_k\}$, we have that

$$(1-\varepsilon)2^{n(h(V|W)-2\varepsilon)} \leq |\mathcal{A}_\varepsilon^{(n)}(V|w^n)| \leq 2^{n(h(V|W)+2\varepsilon)}. \tag{2.7}$$

5) Fix $k = 2$. If $(\tilde{X}_1^n, \tilde{X}_2^n)$ are independent sequence with the same marginals as $f_{X_1^n X_2^n}(x_1^n, x_2^n)$, then

$$\Pr\{(\tilde{X}_1^n, \tilde{X}_2^n) \in \mathcal{A}_\varepsilon^{(n)}(X_1X_2)\} \leq 2^{-n(I(X_1;X_2)-2\varepsilon)}. \tag{2.8}$$

Moreover, for n large enough,

$$\Pr\{(\tilde{X}_1^n, \tilde{X}_2^n) \in \mathcal{A}_\varepsilon^{(n)}(X_1X_2)\} \geq (1-\varepsilon)2^{-n(I(X_1;X_2)+2\varepsilon)}. \tag{2.9}$$

See [14, Section 15.2] for detailed proofs of the above properties.

These properties will be used in the derivation of the capacity region of the BIS under Gaussian source in Section 5. For the weakly ε-typical set for discrete RVs, it can be defined similarly to the above definition and it also holds for all properties in Lemma 2.1 by replacing differential entropy $h(\cdot)$ with the entropy of discrete RVs $H(\cdot)$.

Next, we define the strongly ε-typical set for discrete RVs. The strong typicality had a long history. It was first studied by Wolfowitz [78] and developed in Berger [6] and Csiszár and Korner [15]. A more comprehensible version of the strong typicality can be found in [14] and [19].

Definition 2.2. *(Strongly ε-typical set [14], and [19])*

Let $N(a_1, a_2 \cdots, a_k | a_1^n, a_2^n, \cdots, a_k^n)$ be the number of occurrences of $(a_1, a_2 \cdots, a_k)$ in $(a_1^n, \cdots, a_k^n)$. The strongly ε-typical set with respect to a distribution $P_{A_1^n A_2^n \cdots A_k^n}(a_1^n, a_2^n \cdots, a_k^n)$ on $\mathcal{A}_1^n \times \mathcal{A}_2^n \times \cdots \times \mathcal{A}_k^n$, denoted by $\mathcal{T}_\varepsilon^{(n)}(A_1A_2 \cdots A_k)$, is the set of sequence $(a_1^n, a_2^n \cdots, a_k^n) \in \mathcal{A}_1^n \times \mathcal{A}_2^n \times \cdots \times \mathcal{A}_k^n$ satisfying:

1) $P_{AB}(a_1, a_2 \cdots, a_k) = 0$ implies $\frac{1}{n}N(a_1, a_2 \cdots, a_k | a_1^n, a_1^n, \cdots, a_k^n) = 0$
for all $(a_1, a_2 \cdots, a_k) \in \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_k$,

2) $|\frac{1}{n}N(a_1, a_2 \cdots, a_k | a_1^n, a_1^n, \cdots, a_k^n) - P_{A_1A_2 \cdots A_k}(a_1, a_2 \cdots, a_k)| \leq \frac{\varepsilon}{|\mathcal{A}_1||\mathcal{A}_2| \cdots |\mathcal{A}_k|}$
if $P_{A_1A_2 \cdots A_k}(a_1, a_2 \cdots, a_k) > 0$ for all $(a_1, a_2 \cdots, a_k) \in \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_k$.

Note that a well-known relation of the two sets is that strong typicality implies weak typicality, but the converse claim is not guaranteed. In general, strong typicality is more powerful and flexible than weak typicality as a tool for proving the achievability (direct part) in many memoryless problems. However, unlike the weak typicality, which can be extended to cover the continuous RVs, the strong typicality is applicable only for RVs with finite alphabets.

Here, we provide the definition of the modified ε-typical set [33, Appendix A-A], and [29, Appendix C-A]. The modified set gives the so-called Markov lemma for weak typicality, and two properties of this set enable us to establish the inner bound of the capacity region on the BIS for Gaussian sources in Chapter 5.

Definition 2.3. *(Modified ε-typical set [33, Appendix A])*

Consider that (X,Y,U) forms a Markov chain $X-Y-U$, i.e., $f_{XYU}(x,y,u) = f_{XY}(x,y)f_{U|Y}(u|y)$. The modified ε-typical set $\mathcal{B}_\varepsilon^{(n)}(YU)$ is defined as

$$\mathcal{B}_\varepsilon^{(n)}(YU) = \Big\{(y^n,u^n) : \Pr\{X^n \in \mathcal{A}_\varepsilon^{(n)}(X|y^n,u^n)|(Y^n,U^n) = (y^n,u^n)\} \geq 1-\varepsilon\Big\}, \tag{2.10}$$

where ε is small enough positive, and X^n is drawn i.i.d. from the transition probability $\prod_{k=1}^n f_{X|Y}(X_k|y_k)$. In addition, define $\mathcal{B}_\varepsilon^{(n)}(U|y^n) = \{u^n : (u^n,y^n) \in \mathcal{B}_\varepsilon^{(n)}(YU)\}$ for all y^n, and $\mathcal{B}_\varepsilon^{(n)}(U|y^n)^c$ denotes the complementary set of $\mathcal{B}_\varepsilon^{(n)}(U|y^n)$.

Property 2.1. *There are two useful properties regarding the modified typical set*

(1) If $(y^n,u^n) \in \mathcal{B}_\varepsilon^{(n)}(YU)$, then $(y^n,u^n) \in \mathcal{A}_\varepsilon^{(n)}(YU)$.

(2) For large enough n, it holds that

$$\iint_{\mathcal{B}_\varepsilon^{(n)}(YU)} f_{Y^nU^n}(y^n,u^n)d(y^n,u^n) \geq 1-\varepsilon. \tag{2.11}$$

Proof: The proofs of both properties are given in [33, Appendix C]. □

The following lemma is often used in evaluating the lower bounds of uniformity of the secret keys, secrecy- and privacy-leakage for discrete sources. In Chapter 5,

Lemma 2.2. *(Kittichokechai et al. [41])*

Assume that (X^n,Y^n,U^n) are jointly typical with high probability[1] for a given codebook $\mathcal{C}_n$. Then, it holds that

$$\frac{1}{n}H(Y^n|U^n,\mathcal{C}_n) \leq H(Y|U) + \delta_n, \tag{2.12}$$

$$\frac{1}{n}H(Y^n|X^n,U^n,\mathcal{C}_n) \leq H(Y|X,U) + \delta_n, \tag{2.13}$$

where $\delta_n \downarrow 0$ as $n \to \infty$.

Proof: The proof can be found in [41, Appendix C]. □

Another important tool for deriving our results is the selection lemma. The lemma was proposed in [10, Lemma 2.2], and it is mainly used to assert the existence of a good code for the achievability proofs in the succeeding chapters.

[1] It means that $\Pr\{(X^n,Y^n,U^n) \in \mathcal{T}_\varepsilon^{(n)}(XYU)\} \to 1$ as $n \to \infty$, where $\mathcal{T}_\varepsilon^{(n)}(XYU)$ denotes the set of strongly ε-typical sequences for RVs X,Y, and U.

Lemma 2.3. *(Selection lemma [10, Lemma 2.2])*

Let $X^n \in \mathcal{X}^n$ be a random variable and $\mathcal{F}$ be a finite set of functions $f : \mathcal{X}^n \longrightarrow \mathbb{R}^+$ such that $|\mathcal{F}|$ does not depend on n and

$$\forall f \in \mathcal{F}, \quad \mathbb{E}_{X^n}[f(X^n)] \leq \delta(n). \tag{9a}$$

Then, there exists a specific realization x^n of X^n such that

$$\forall f \in \mathcal{F}, \quad f(x^n) \leq (|\mathcal{F}|+1)\delta(n). \tag{9b}$$

(Proof): See the proof of [14, Lemma 2.2]. □

Lastly, we introduce Shannon's EPI, which will be used in the proofs of Section 5. Let RVs $A \sim f_A$ and $B \sim f_B$ be independent continuous RVs. The EPI tells us that a lower bound of the differential entropy of the sum of RVs A and B is given by

$$e^{2h(A+B)} \geq e^{2h(A)} + e^{2h(B)} \tag{2.14}$$

with equality if both RVs A and B are Gaussian RVs. The EPI was first proposed in Shannon [65] without a rigorous proof. Later, a complete proof of the EPI was given by Stam [68] and Blachman [9], based on Fisher information inequality. The proof is simplified by Dembo et al. [17]. Yet, other much simpler proofs can also be found in [25], [72] using minimum mean-square error and Rioul [60]–[62] via only the properties of mutual information, avoiding both Fisher information inequality and minimum mean-square error. The conditional version of the EPI is shown in [7, Lemma II]. In the converse proof of Gaussian sources (Chapter 5), the conditional version of the EPI plays important role in deriving the outer bound of the capacity region.

2.3 The Primitive BIS

In this section, we review a classical model of the pioneering work. We explain the system model and introduce the main result given by Willems et al. [76].

2.3.1 System Model

The model is shown in Figure 2.1. Basically, a BIS consists of two big phases; (I) *Enrollment Phase* and (II) *Identification Phase*. In this subsection, the details of each phase and the result of this model are provided within information theoretic framework.

(I) *Enrollment Phase*

We assume that there are M_I individuals in the BIS. Each user is assigned by an index from the set $\mathcal{I} = [1 : M_I]$. The raw or original bio-data sequence of user i, $x_i^n = (x_{i1}, x_{i2}, \cdots, x_{in}) \in \mathcal{X}^n$, with symbols $x_{ik}(i \in \mathcal{I}, k \in [1,n])$ takes a value in a finite alphabet $\mathcal{X}$. We also assume that all of these

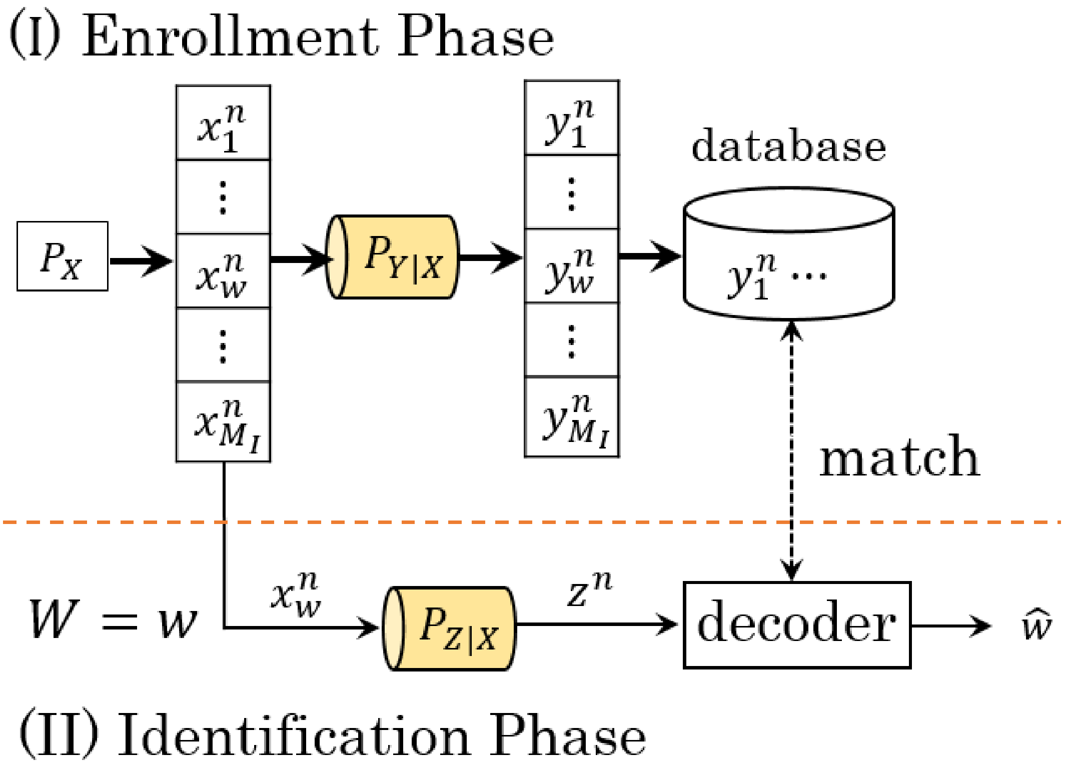

Fig. 2.1 Primitive BIS model; This is the system model considered in Willems et al. [76], where the captured biological data are stored in the system database in the plain forms. This type of model is called noisy BIS, also known as HSM, because the noise in the enrollment phase is taken into account.

bio-data sequences are generated independently and identically distributed (i.i.d.) from a stationary memoryless source P_X. For all $i \in \mathcal{I}$, the generating probability for each sequence $x_i^n \in \mathcal{X}^n$ is given by

$$P_{X^n}(x_i^n) \;=\; \Pr[X^n = x_i^n] \;=\; \prod_{k=1}^{n} P_X(x_{ik}). \tag{2.15}$$

All bio-data sequences x_i^n $(i \in \mathcal{I})$ are observed via a *discrete memoryless channel* (DMC) $\{\mathcal{Y}, P_{Y|X}, \mathcal{X}\}$, called the enrollment channel, where $\mathcal{Y}$ is a finite output-alphabet of $P_{Y|X}$. Therefore, the corresponding probability that a bio-data sequence $x_i^n \in \mathcal{X}^n$ is observed as $y_i^n = (y_{i1}, y_{i2}, \cdots, y_{in}) \in \mathcal{Y}^n$ via the DMC $P_{Y|X} : \mathcal{X} \to \mathcal{Y}$ is

$$P_{Y_i^n|X_i^n}(y_i^n|x_i^n) \;=\; \Pr[Y_i^n = y_i^n|X_i^n = x_i^n] \;=\; \prod_{k=1}^{n} P_{Y|X}(y_{ik}|x_{ik}) \tag{2.16}$$

for all $i \in \mathcal{I}$. Here, y_i^n is the output sequence of individual i via $P_{Y|X}$. All $\{y_1^n, y_2^n, \cdots, y_{M_I}^n\}$ are saved into the system database, which can be accessed by a decoder in the identification phase. For simplicity reason, we denote the system database as $\boldsymbol{J} = \{y_1^n, y_2^n, \cdots, y_{M_I}^n\}$.

(II) Identification Phase

In this phase, an unknown user w, who has already gone through the Enrollment Phase, presents his bio-data sequence to the system and the sequence is observed via a DMC $\{\mathcal{Z}, P_{Z|X}, \mathcal{X}\}$, called the identification channel, where $\mathcal{Z}$ is a finite output-alphabet via $P_{Z|X} : \mathcal{X} \to \mathcal{Z}$. Therefore, the corresponding probability that a bio-data sequence $x_w^n \in \mathcal{X}^n$ is observed as $z^n = (z_1, z_2, \cdots, z_n) \in \mathcal{Z}^n$ via $P_{Z|X} : \mathcal{X} \to \mathcal{Z}$ is

$$P_{Z^n|X_w^n}(z^n|x_w^n) \ = \ \Pr[Z^n = z^n|X_w^n = x_w^n] \ = \ \prod_{k=1}^{n} P_{Z|X}(z_k|x_{wk}). \tag{2.17}$$

z^n is passed to the decoder $d : \mathcal{Z}^n \times \boldsymbol{J} \longrightarrow \mathcal{I}$ and it compares z^n with all sequences y_i^n $(i \in \mathcal{I})$ in the database $\boldsymbol{J}$ and outputs an estimate of the unknown user's index

$$\widehat{w} = d(z^n, \boldsymbol{J}). \tag{2.18}$$

2.3.2 Identification Capacity

Willems et al. [76] applied information theoretic methods to investigate the maximum achievable rate of the individuals with vanishing decoding error probability. We first note that RVs X_i^n, Y_i^n, and Z^n form a Markov chain $Y_i^n - X_i^n - Z^n$ [14], and thus the joint probability distribution among them can be written as

$$\begin{aligned} P_{X_i^n Y_i^n Z^n}(x_i^n, y_i^n, z^n) &= P_{X_i^n}(x_i^n) P_{Y_i^n|X_i^n}(y_i^n|x_i^n) P_{Z^n|X_i^n}(z^n|x_i^n) \\ &= \prod_{k=1}^{n} P_{Y|X}(y_{ik}|x_{ik}) P_X(x_{ik}) P_{Z|X}(z_k|x_{ik}), \end{aligned} \tag{2.19}$$

where the last equation in (2.19) is due to the i.i.d. structure of each symbol. Then, from the marginal logic, it can be easily derived that

$$\begin{aligned} P_{Y_i^n Z^n}(y_i^n, z^n) &= P_{Y_i^n}(y_i^n) P_{Z^n|Y_i^n}(z^n|y_i^n) \\ &= \prod_{k=1}^{n} P_Y(y_{ik}) P_{Z|Y}(z_k|y_{ik}), \end{aligned} \tag{2.20}$$

where P_Y and $P_{Z|Y}$ can be computed as

$$P_Y(y) = \sum_{x \in \mathcal{X}} \sum_{z \in \mathcal{Z}} P_X(x) P_{Y|X}(y|x) P_{Z|X}(z|x), \tag{2.21}$$

$$P_{Z|Y}(z|y) = \sum_{x \in \mathcal{X}} \frac{P_X(x) P_{Y|X}(y|x) P_{Z|X}(z|x)}{P_Y(y)}. \tag{2.22}$$

Equation (2.20) implies that $P_{Z|Y} : \mathcal{Y} \to \mathcal{Z}$ also forms a DMC and each enrollment output sequence y_i^n can be viewed as an i.i.d source sequence.

Remark 2.1. *Willems et al. [76] observed that the set of all the enrollment output sequences (a database) can be seen as a code whose codeword corresponding to message i is y_i^n. In the identification phase, these codewords y_i^n $(1 \leq i \leq M_I)$ are observed via a DMC $P_{Z|Y} : \mathcal{Y} \to \mathcal{Z}$.*

Remark 2.1 indicates that we can use the same techniques for analyzing the error probability in channel coding systems to bound the error probability in the BIS. The identification rate of the BIS with the block length n and M_I individuals is denoted by

$$R_I = \frac{1}{n} \log M_I. \tag{2.23}$$

Definition 2.4. *An identification rate R_I $(R_I \geq 0)$ is said to be achievable if for $\delta > 0$ and large enough n there exists a decoder d such that*

$$\max_{i \in \mathcal{I}} \Pr\{\widehat{W} \neq W | W = i\} \leq \delta, \tag{2.24}$$

$$\frac{1}{n} \log M_I \geq R_I - \delta. \tag{2.25}$$

Moreover, the identification capacity C of the BIS is the supremum of all achievable identification rates R_I. That is,

$$C = \sup\{R_I \mid R_I \text{ is achievable}\}.$$

□

Now, we are in a position to introduce the identification capacity. In [76], Willems et al. proved the following theorem.

Theorem 2.1. *(Willems et al. [76])*

The identification capacity of the BIS is given by

$$C = I(Y;Z), \tag{2.26}$$

where $I(Y;Z)$ denotes the mutual information between RVs Y and Z with the joint probability distribution $P_{ZY}(z,y) = \sum_{x \in \mathcal{X}} P(x) P_{Y|X}(y|x) P_{Z|X}(z|x)$ $(\forall y \in \mathcal{Y},\ \forall z \in \mathcal{Z})$. □

The above theorem is a convincing result because the set of $\{Y_1^n, \cdots, Y_{M_I}^n\}$ can be viewed as a random code. When a codeword of this random code is sent via the channel $P_{Z|Y}$, which is a compound channel consisting of the backward channel of the channel in the enrollment phase and the identification channel, the maximum achievable rate that the index of the sent message can be reliably estimated is the mutual information between RVs Y and Z. This is a simple result, but it has been a huge influence for all the relevant studies taken place later on.

Chapter 3

BISs Supporting Authentication

As mentioned in Chapter 1, the BIS supporting authentication was first investigated by Ignatenko and Willems for the generated-secret model [32] and for both the generated- and chosen-secret BIS models [33]. They aimed to maximize the identification and secrecy rates under a privacy-leakage constraint for the VSM. However, in real-life applications, when user enrolls her bio-data for future identification, the identity needs to be scanned and sent into the system database. During these processes, it is likely that noise is added to the original bio-data. Therefore, it is more natural to analyze the BIS under the circumstance that the enrollment channel is noisy. Another interesting thing is that when the model switches from the VSM to the HSM, the problem becomes more challenging. Especially, the evaluation of the privacy-leakage rate as the template is not a function of original bio-data, but a noisy version of it. Basically, many techniques used to investigate the fundamental trade-off in the VSM are not directly applicable to the HSM as claimed in [21]. Indeed, we are greatly motivated by these facts to improve the models considered in [33] to include the HSM.

In this chapter, we focus on clarifying the fundamental trade-off among identification, secrecy, template, and privacy-leakage rates of the generated- and chosen-secret BIS models. Both models are analyzed under the assumption of the HSM. Here, we wish to maximize the identification and secrecy rates under privacy and storage constraints. In order to get closer to practical system, we analyze the region by imposing the following requirements:

1) the enrollment channel is a noisy,
2) a scheme of both protecting privacy (e.g., [33], [21]) and compressing template (e.g., [69], [81]) is considered,
3) the capacity region is analyzed under the condition that the prior distribution of an identified individual is unknown.

In the achievability proof, we deploys layered binning technique to reduce the rate of database, and this enable us to make the error probability arbitrarily small. To handle the difficulties of bounding the privacy-leakage in the achievability proof, we introduce a *virtual* system with a *partial* decoder, which outputs only the secret data of individual, and use Lemma 2.2. In the converse proof, we relax the problem to be the one where the prior distribution of the identified user is uniform, and use Fano's inequality together with the assistance of auxiliary RVs, which is quite standard technique for analyzing the outer bound of the capacity region. We show that there are two different ways to express the capacity regions of the generated- and chosen-secret BIS models. An expression uses only a single auxiliary RV and another requires two auxiliary RVs. Later, we will demonstrate that the two regions (regions with one and two auxiliary RVs) are technically identical in Remark 3.5. Although there are two different aspects, we provide the proof of our main result based on the one employing two auxiliary RVs. Some benefits of deriving via two auxiliary RVs are that the achievability proof can be done in a simpler form since each rate constraint is addressed individually. The characterization of the capacity regions of the models is basically similar to the ones given in [21], [33], and [81]. As special cases, it can be checked that our characterization reduces to the one given by Ignatenko and Willems [33] where the enrollment channel is noiseless and there is no constraint on the template rate, and it also coincides with the result derived by Günlü and Kramer [21] where there is only one individual, and thus individual's estimation is not necessary.

The rest of this chapter is organized as follows. In Section 3.1, we describe the details of the system model considered in this chapter. In Section 3.2, we present our main results. Next, we provide the detailed proofs of the main results in Section 3.4 and Section 3.5. Finally, in Section 3.6, we give summary of results and discussion.

3.1 System Model

In this section, we explain the system models considered in this chapter. For the detailed explanations of the processes of generating bio-data sequence X_i^n and observing Y_i^n and Z^n, the readers should refer to Section 2.3. Here, we only provides the new parts. We start with describing the generated-secret BIS model and then the chosen-secret BIS model.

3.1.1 Generated-Secret BIS Model

The generated-secret BIS model investigated in this chapter are shown in Fig. 3.1. As we have previously mentioned, it consists of two phases: (I) Enrollment Phase and (II) Identification Phase. Next we explain the details of each phase.

(I) Enrollment Phase:

Let $\mathcal{I} = [1, M_I]$, $\mathcal{J} = [1, M_J]$, and $\mathcal{S} = [1, M_S]$ be the sets of indexes of users, indexes of templates, and secret data of users, respectively. The observed bio-data sequence Y_i^n via $P_{Y|X}$ with in put X_i^n is encoded into template $J(i) \in \mathcal{J}$ and secret data $S(i) \in \mathcal{S}$ as

$$(J(i), S(i)) = e(Y_i^n) \quad (i \in \mathcal{I}), \tag{3.1}$$

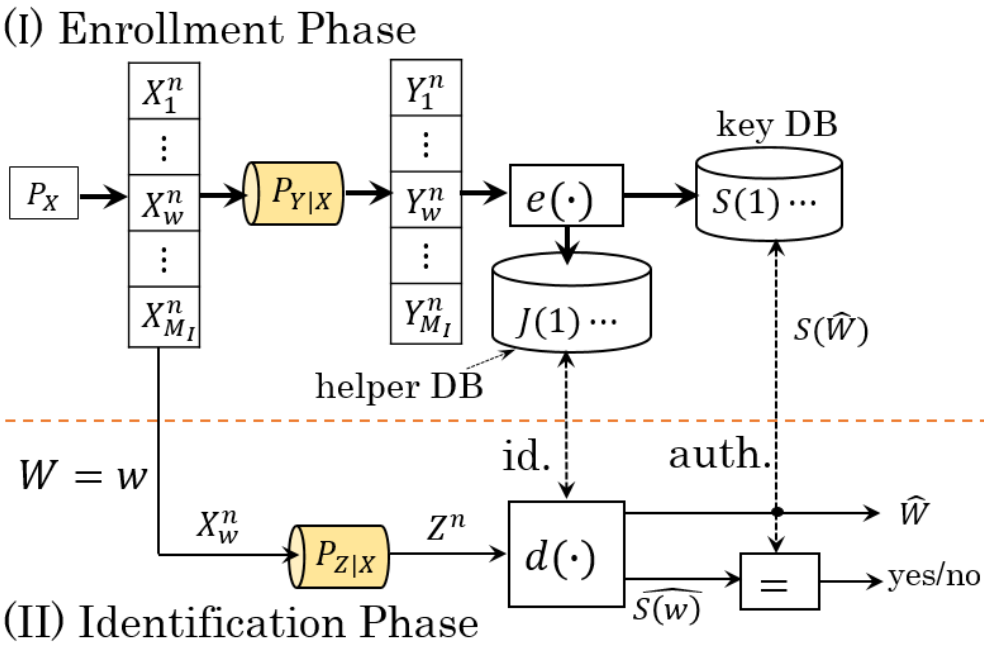

Fig. 3.1 Generated-secret BIS model; This figure illustrates the data flows in the generated-secret BIS model. In the enrollment phase, the encoder generates secret keys and templates from the bio-data sequence of individuals. The secret keys are used for authenticating purpose. The templates are stored in the system database so as to help the decoder to estimate the index and secret key based on the presented bio-data.

where $e:\ \mathcal{Y}^n \longrightarrow \mathcal{J}\times\mathcal{S}$ denotes encoding function. The corresponding template $J(i)$ is a compressed version of sequence Y_i^n and stored at position i in a public database $\boldsymbol{J} = \{J(1),\cdots,J(M_I)\}$, which can be accessed by the decoder. On the other hand, the secret data $S(i)$ is saved at position i in the key database, which is installed in a secure location. Note that both $J(i)$ and $S(i)$ are functions of index i.

(II) Identification Phase:

Suppose that the user w presents his identity to the BIS. The decoder observers the identified sequence Z^n, the noisy sequence of the identified user X_w^n, and estimates the pair of index and secret key by comparing Z^n with all templates $\boldsymbol{J}$ in the database.

$$(\widehat{W},\widehat{S(w)}) = d(Z^n,\boldsymbol{J}), \tag{3.2}$$

where $d:\ \mathcal{Z}^n\times\mathcal{J} \longrightarrow \mathcal{I}\times\mathcal{S}$ denotes decoding function.

3.1.2 Chosen-Secret BIS Model

The chosen-secret BIS model analyzed in this chapter is illustrated in Fig. 3.2.

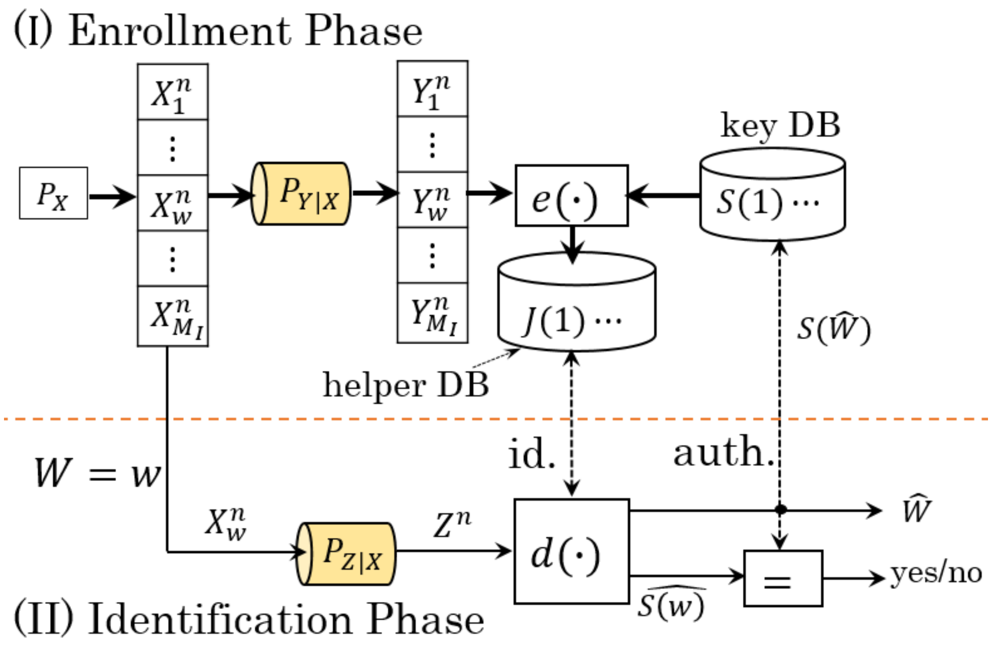

Fig. 3.2 Chosen-secret BIS model; This is the chosen-secret BIS model and the difference from the generated-secret BIS model is that the secret key is given to the encoder from a independent source. In the enrollment phase, the encoder uses the secret key and the bio-data sequence of individuals to form the template.

(I) Enrollment Phase:

In this model, it is assumed that the secret key $S(i)$ for the user i is uniformly distributed on $\mathcal{S}$. That is,

$$P_{S(i)}(s(i)) = \frac{1}{M_S}. \tag{3.3}$$

Also, the secret key is picked independently of other RVs, e.g., (X_i^n, Y_n^n, Z^n, W), from the key database. Upon observing the sequence Y_i^n and the secret key $S(i)$, the encoder forms a template as

$$J(i) = e(Y_i^n, S(i)) \quad (i \in \mathcal{I}), \tag{3.4}$$

where $e: \mathcal{Y}^n \times \mathcal{S} \longrightarrow \mathcal{J}$.

(I) Identification Phase:

Seeing the sequence Z^n, the decoder reconstructs the index and secret key of the identified user based on the information inside the database $\boldsymbol{J}$ as follows:

$$(\widehat{W}, \widehat{S(w)}) = d(Z^n, \boldsymbol{J}). \tag{3.5}$$

Remark 3.1. *Note that the distribution of P_X, $P_{Y|X}$, and $P_{Z|X}$ are assumed to be known or fixed and RV W is independent of $(X_i^n, Y_i^n, J(i), S(i), Z^n)$ for all $i \in \mathcal{I}$ like previous studies. However, in this* *thesis, we assume neither that the identified individual index W are uniformly distributed over $\mathcal{I}$ nor that there is a prior distribution of W.*

Again a motivation of analyzing performances of the BIS provided that the distribution of W is unknown is that the identified frequencies of each individual are likely different. For example, it is hard to think that each user comes to use a bank teller at the same rate, and thus for real applications, this assumption is important to take care of.

3.2 Problem Formulation and Main Results

In this section, formal definitions and main results are provided in details. We begin with stating the formal definition of the generated-secret BIS. $(W, S(W))$ and $(\widehat{W}, \widehat{S(W)})$, and denote the RV corresponding to the pair of index and secret key of the identified individual $(w, s(w))$, its estimated values $(\widehat{w}, \widehat{s(w)})$, respectively.

Definition 3.1. *(Generated-secret BIS model)*

The tuple of an identification, secrecy, template, privacy-leakage rates (R_I, R_S, R_J, R_L) is said to be achievable for the generated-secret BIS if for any $\delta > 0$ and large enough n there exist pairs of encoders and decoders that satisfy

$$\max_{i\in\mathcal{I}} \Pr\{(\widehat{W}, \widehat{S(W)}) \neq (W, S(W)) | W = i\} \leq \delta, \tag{3.6}$$

$$\frac{1}{n}\log M_I \geq R_I - \delta, \tag{3.7}$$

$$\min_{i\in\mathcal{I}} \frac{1}{n} H(S(i)) \geq R_S - \delta, \tag{3.8}$$

$$\frac{1}{n}\log M_J \leq R_J + \delta, \tag{3.9}$$

$$\max_{i\in\mathcal{I}} \frac{1}{n} I(S(i); J(i)) \leq \delta, \tag{3.10}$$

$$\max_{i\in\mathcal{I}} \frac{1}{n} I(X_i^n; J(i)) \leq R_L + \delta. \tag{3.11}$$

Moreover, $\mathcal{R}_{GS}$ is defined as the closure of the set of all achievable rate tuples for the generated-secret BIS, called the capacity region. □

In Definition 3.1, (3.6) is the condition of the maximum error probability of an individual i, which is arbitrarily small. Equations (3.7)–(3.9) are the constraints related to identification, secrecy, and template rates, respectively. In term of the privacy protection perspective, we measure the information leakage of individual i by (3.10) and (3.11). Condition (3.10) measures the secrecy-leakage between the template in the database and the secret data of individual i, and it requires that the maximum leaked amount is not greater than δ. Condition (3.11) measures the amount of privacy-leakage of original bio-data X_i^n from template $J(i)$ and its maximum value must be smaller than or equal to $R_L + \delta$. Later, we will see that the evaluation for R_L is the most *intricate* task.

Remark 3.2. *In [21], and [40], the constraint on the helper data is called storage rate. However, we stay away from calling it* storage rate *and instead we name it* template rate *because besides the database of helper data, there exists a database of secret keys, which is considered to be a secure storage. Here, we wish to minimize the storage of the helper data and maximize the database of secret key at the same time. Therefore, the entire storage is not being minimized and it is more proper to avoid such misleading key word.*

Remark 3.3. *In [33], a stronger requirement that the distribution of secret data of every individual must be almost uniform, i.e.* $\frac{1}{n}H(S(i))+\delta \geq \frac{1}{n}\log M_S$*, is included in (3.8). However, this requirement was not actually necessary in the general problem formulation, which will be seen in the proof of Theorem 3.1.*

The definition of the chosen-secret BIS model is given below.

Definition 3.2. *(Chosen-secret BIS model)*

The tuple of an identification, secrecy, template, privacy-leakage rates (R_I, R_S, R_J, R_L) *is said to be achievable for the chosen-secret BIS if for any* $\delta > 0$ *and large enough n there exist pairs of encoders and decoders satisfying*

$$\max_{i\in\mathcal{I}} \Pr\{(\widehat{W},\widehat{S(W)}) \neq (W,S(W))|W=i\} \leq \delta, \tag{3.12}$$

$$\frac{1}{n}\log M_I \geq R_I - \delta, \tag{3.13}$$

$$\frac{1}{n}\log M_S \geq R_S - \delta, \tag{3.14}$$

$$\frac{1}{n}\log M_J \leq R_J + \delta, \tag{3.15}$$

$$\max_{i\in\mathcal{I}} \frac{1}{n}I(S(i);J(i)) \leq \delta, \tag{3.16}$$

$$\max_{i\in\mathcal{I}} \frac{1}{n}I(X_i^n;J(i)) \leq R_L + \delta. \tag{3.17}$$

Moreover, the capacity region $\mathcal{R}_{CS}$ *is defined as the closure of the set of all achievable rate tuples for the chosen-secret BIS.* □

Due to the assumption of (3.3), the entropy in the left-handed side of (3.8) in Definition 3.1 becomes $\frac{1}{n}\log M_S$ in (3.14) (the entropy is maximized).

Remark 3.4. *Equations* (3.8) *and* (3.14) *imply that the size or length of the secret key should be maximized. We aim to extract as large as possible size of the secret key from the bio-data sequence at the encoder, and to estimate the key at the decoder reliably. The estimated key may be utilized as, e.g., authentication password or encryption key, in the later stage based on the identified user's purpose. Here, however, we do not discuss how it can be applied in real-life applications after estimated. On the other hand, in cryptography, the encryption key is used to transform the plaintext into ciphertext* *(encryption) and vice versa (decryption). Also, its size should be made as small as possible because of the computational complexity for encryption and decryption.*

Before stating the main results of this chapter, we define the following 4 new regions.

$$\begin{aligned}\mathcal{A}_1 = \{(R_I,R_S,R_J,R_L):\ & R_I+R_S \le I(Z;U),\\ & R_J \ge I(Y;U)-I(Z;U)+R_I,\\ & R_L \ge I(X;U)-I(Z;U)+R_I,\\ & R_I \ge 0, R_S \ge 0 \text{ for some } U \text{ s.t. } Z-X-Y-U\},\end{aligned} \tag{3.18}$$

$$\begin{aligned}\mathcal{A}_2 = \{(R_I,R_S,R_J,R_L):\ & R_I+R_S \le I(Z;U),\\ & R_J \ge I(Y;U),\\ & R_L \ge I(X;U)-I(Z;U)+R_I,\\ & R_I \ge 0, R_S \ge 0 \text{ for some } U \text{ s.t. } Z-X-Y-U\},\end{aligned} \tag{3.19}$$

$$\begin{aligned}\mathcal{A}_3 = \{(R_I,R_S,R_J,R_L):\ & 0 \le R_I \le I(Z;V),\\ & 0 \le R_S \le I(Z;U)-I(Z;V),\\ & R_J \ge I(Y;U)-I(Z;U)+I(Z;V),\\ & R_L \ge I(X;U)-I(Z;U)+I(Z;V),\\ & \text{for some } U \text{ and } V \text{ s.t. } Z-X-Y-U-V\},\end{aligned} \tag{3.20}$$

$$\begin{aligned}\mathcal{A}_4 = \{(R_I,R_S,R_J,R_L):\ & 0 \le R_I \le I(Z;V),\\ & 0 \le R_S \le I(Z;U)-I(Z;V),\\ & R_J \ge I(Y;U),\\ & R_L \ge I(X;U)-I(Z;U)+I(Z;V),\\ & \text{for some } U \text{ and } V \text{ s.t. } Z-X-Y-U-V\},\end{aligned} \tag{3.21}$$

where auxiliary RVs U and V take values in some finite alphabets $\mathcal{U}$ and $\mathcal{V}$ with $|\mathcal{U}| \le (|\mathcal{Y}|+2)(|\mathcal{Y}|+3)$ and $|\mathcal{V}| \le |\mathcal{Y}|+3$.

Remark 3.5. *It can be verified that*

$$\mathcal{A}_1 = \mathcal{A}_3, \tag{3.22}$$

$$\mathcal{A}_2 = \mathcal{A}_4, \tag{3.23}$$

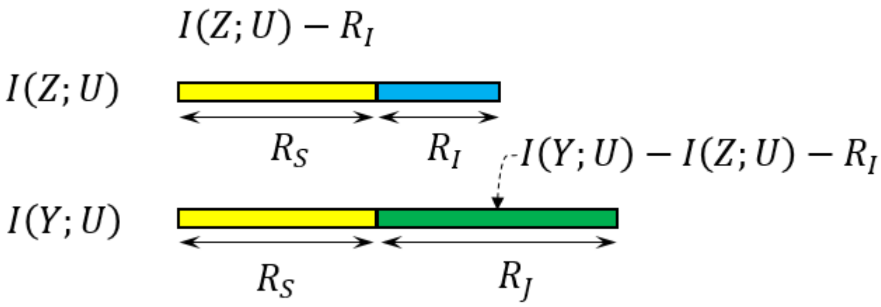

Fig. 3.3 Explanation of each rate constraint in Theorem 3.1

respectively, for which the proof is given in Appendix A.1. In this thesis, we prove Theorem 3.1 and 3.2 based on the rate constraints of the regions $\mathcal{A}_3$ and $\mathcal{A}_4$ instead of $\mathcal{A}_1$ and $\mathcal{A}_4$. In other words, we prove the main results of this chapter via two auxiliary RVs. □

Now we are at the position to introduce our main results of this chapter.

Theorem 3.1. *(Generated-secret BIS model)*

The capacity region for the generated-secret BIS is given by

$$\mathcal{R}_{GS}=\mathcal{A}_1. \tag{3.24}$$

□

The meaning of each rate constraint in Theorem 3.1 is shown in Fig. 3.3. In the top figure, $I(Z;U)$ is the maximum rate that user's identities can be estimated correctly at the decoder. Since the index and the secret key are reconstructed at the decoder, the sum of the identification and secrecy rates should be less than or equal to this value, and they are in a trade-off relation under $I(Z;U)$. In the bottom one, $I(Y;U)$ is the rate that we need to generate auxiliary random sequences for encoding. The first part (yellow part) represents the secrecy rate, and the second half is the rate of the sequences that are shared between the encoder and decoder to help estimation of the index and secret key, corresponding the template rate. Storing templates at this rate in the database results in leaking the user's privacy at least $I(X;U)-I(Z;U)+R_I$, and this quantity emerges in the last constraint in Theorem 3.1.

Theorem 3.2. *(Chosen-secret BIS model)*

The capacity region for the chosen-secret BIS is given by

$$\mathcal{R}_{CS}=\mathcal{A}_2. \tag{3.25}$$

□

Remark 3.6. *Likewise the observation in [21], $\mathcal{R}_{GS}$ is clearly wider than $\mathcal{R}_{CS}$, which is due to the bound on R_J. A remark given in [82] indicated that in case where the enrollment channel is noiseless*

$(X = Y)$, the minimum required amounts of the template and privacy-leakage rates are identical for the generated-secret BIS model. However, this claim does not apply to the chosen-secret BIS model.

In the chosen-secret BIS model (Theorem 3.2), the quantity of the identification and secrecy rates is the same as seen in Theorem 3.1 (cf. the top figure of Fig. 3.4). However, the minimum requirement of template rate (storage) becomes larger, which is $I(Y;U)$ (cf. the bottom figure of Fig. 3.4), as we need to store the information related to the secret key, chosen independently of other RVs, together with the secret key at the database. For the privacy-leakage rate, the minimum value does not vary. Indeed, in the chosen-secret BIS model, the chosen secret key might be used as an extra randomness seed to make the privacy-leakage decrease, but that is not seen in the final condition of the region $\mathcal{R}_C$. This is because the length of the chosen secret key is too small compared to the template rate, and it can be used to partially conceal the storage. However, the unconcealed part for the storage at rate $I(Y;U) - I(Z;U) + R_I$ is still exposed publicly. This is identical to the template rate of the generated-secret BIS model, and thus it is not surprised that the privacy-leakage rates of the two models are bounded by the same value.

In case there are no generation of the secret key ($R_S = 0$), and the template and privacy-leakage allow to be large enough ($R_J, R_L \to \infty$), the maximum achievable identification rate is $I(Y;Z)$. This is due to the Markov chain $Z - Y - U$ and the value is exactly the identification capacity shwon in Theorem 2.1. Moreover, if there are only one user ($R_I = 0$), $R_J, R_L \to \infty$, noise-free enrollment $(X = Y)$, the largest possible secrecy rate becomes $I(Z;X)$, corresponding to the secrecy capacity for one-way communication of two terminals in [2].

As we have previously mentioned, one can check that the characterizations of Theorem 3.1 and 3.2 coincide with the regions characterized by Ignatenko and Willems [33] in two steps: first replace Y by X and then remove the constraint R_J from (3.18). Also, this results correspond to the regions given by Günlü and Kramer [21] for the generated- and chosen-secret BIS with only one user. It is easy to check this claim by just setting $R_I = 0$. Moreover, in the case where there are no assumption of the adversary and the enrollment channel is noise-free, it is not hard to confirm that the characterization of the generated-secret BIS model in this chapter is equivalent to the result of [40, Theorem 1] by similar arguments in Appendix A.1.

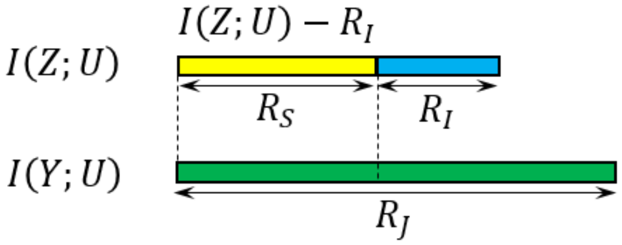

Fig. 3.4 Explanation of each rate constraint in Theorem 3.2

3.3 Numerical Example and Overviews of the Proof

3.3.1 Simplified Rate Region

In this section, a numerical example of the rate region of the generated-secret BIS for binary hidden source is given. We consider the case where $P_X(0) = P_X(1) = 0.5$ and the crossover probabilities at the encoder $0 \leq q_E \leq 0.5$ and at the decoder $0 \leq q_D \leq 0.5$. First, we simplify the capacity region for this case by applying Mrs. Gerber Lemma (MGL) [79]. From the right-hand side of (3.18), we obtain that

$$I(Z;U) = 1 - H(Z|U), \tag{3.26}$$

$$I(Y;U) - I(Z;U) + R_I = H(Z|U) - H(Y|U) + R_I, \tag{3.27}$$

$$I(X;U) - I(Z;U) + R_I = H(Z|U) - H(X|U) + R_I. \tag{3.28}$$

The above relations indicate that to simplify the capacity region, it is required to maximize $H(Y|U)$ and minimize $H(Z|U)$ for fixed $H(X|U)$.

First, observe that since $1 \geq H(X|U) \geq H(X|Y) = H_b(p_E)$, there must exist an γ satisfying that $H(X|U) = H_b(\gamma * p_E)$, where $\gamma \in [0, 0.5]$. By applying MGL to the Markov chain $U - X - Z$, we have

$$H(Z|U) \geq H_b(H_b^{-1}(H(X|U)) * p_D) = H_b(\gamma * p_E * p_D). \tag{3.29}$$

Again, in opposite direction, if the MGL is applied to the Markov chain $U - Y - X$, it follows that

$$H(X|U) \geq H_b(H_b^{-1}(H(Y|U)) * p_E). \tag{3.30}$$

As $H(X|U) = H_b(\gamma * p_E)$, (3.30) yields that

$$H_b(\gamma * p_E) \geq H_b(H_b^{-1}(H(Y|U)) * p_E) \tag{3.31}$$

and thus

$$\gamma * p_E \geq H_b^{-1}(H(Y|U)) * p_E \tag{3.32}$$

Therefore, we obtain

$$H(Y|U) \leq H_b(\gamma). \tag{3.33}$$

In (3.29) and (3.33) for binary symmetric (U, Y) with crossover probability γ, the minimum $H(Z|U) = H_b(\gamma * p_E * p_D)$ and the maximum $H(Y|U) = H_b(\gamma)$ are achieved. Therefore, the following corollary is obtained.

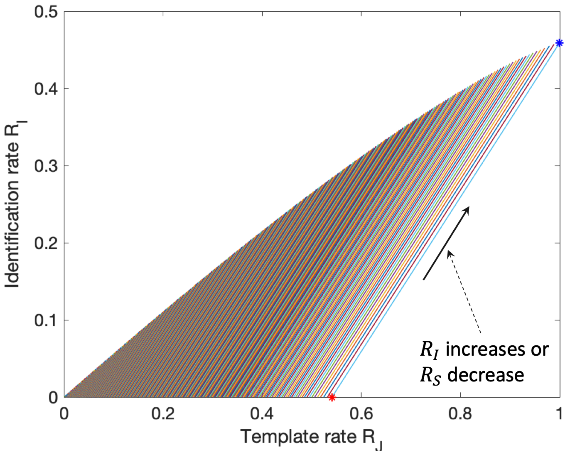

Fig. 3.5 The projection onto $R_J R_I$-plane

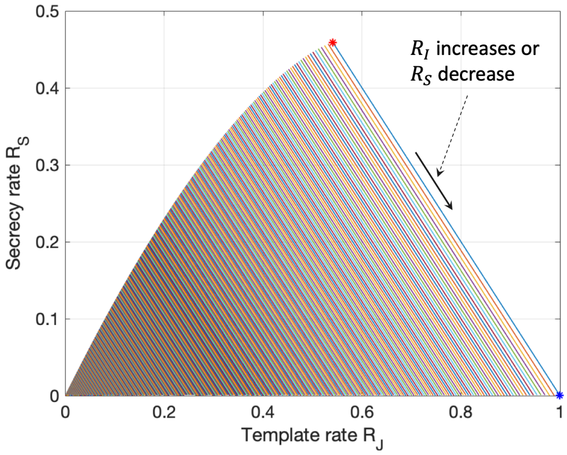

Fig. 3.6 The projection onto $R_J R_S$-plane

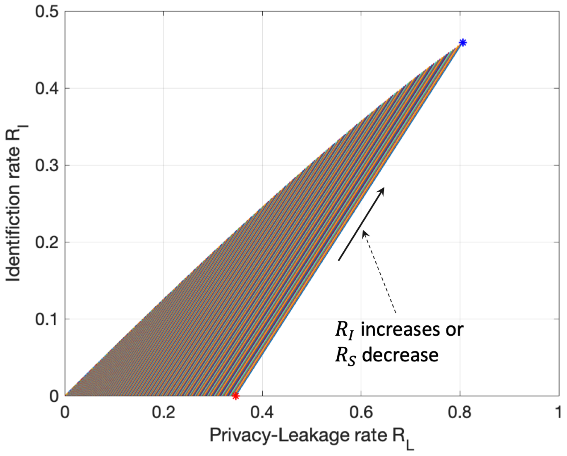

Fig. 3.7 The projection onto $R_L R_I$-plane

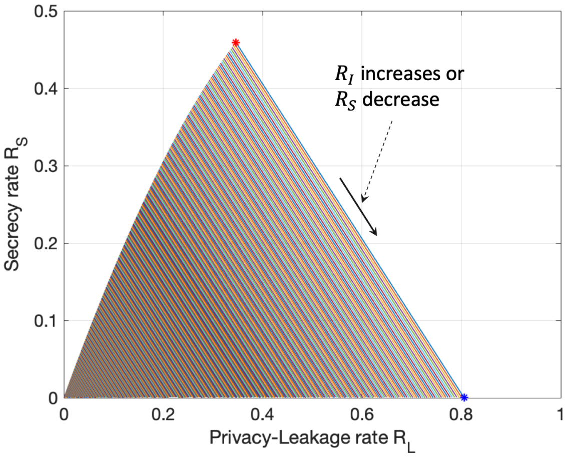

Fig. 3.8 The projection onto $R_L R_S$-plane

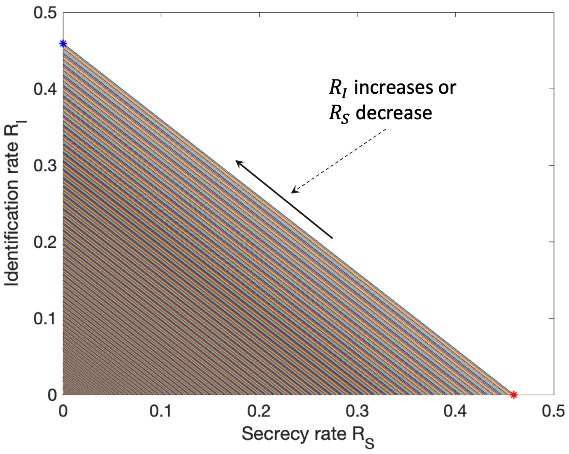

Fig. 3.9 The projection onto $R_S R_I$-plane

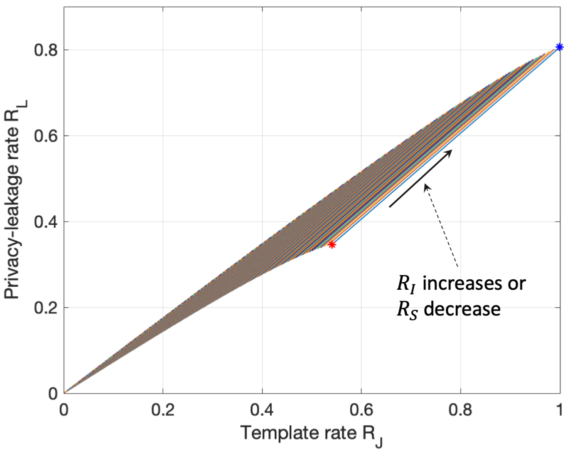

Fig. 3.10 The projection onto $R_J R_L$-plane

Corollary 3.1. *For binary source, Theorem 3.1 reduces to*

$$\begin{aligned}
&R_J = H_b(\gamma * p_E * p_D) - H_b(\gamma) + R_I,\\
&R_L = H_b(\gamma * p_E) - H_b(\gamma) + R_I,\\
&\text{For some } \gamma \in [0, 0.5] \text{ satisfying } R_I + R_S = 1 - H_b(\gamma * p_E * p_D), R_I \geq 0,\\
&\text{and } R_I \leq 1 - H_b(\gamma * p_E * p_D).
\end{aligned} \tag{3.34}$$

3.3.2 Numerical Example

We calculate the above rate region for $p_E = 0.03$ and $p_D = 0.1$, considered to be realistic values for bio-data [21],[33]. The numerical results are shown in Fig. 3.5–3.10 and the painted areas represent the achievable rate regions for each projection.

In Fig. 3.5, if we look at the blue point in R_J-axis, the optimal template rate at $R_I = 0$ (R_S is optimal), as R_I rises, the value of the template rate also increases gradually along the boundary in the direction of arrow, and eventually reaches the red point, the optimal point for identification and template rates (R_S is zero). Clearly, it implies that a greater value of the identification rate results in a larger value of the template rate. In contrast, when we take a look at the relation of the secrecy and template rates in Fig. 3.6, as the secrecy rate decreases (sliding from the blue point to the red point in the bottom), the template rate becomes larger increasingly. This is because the decrease of secrecy rate leads to a bigger gain for the identification rate due to the trade-off relation between them (cf. Fig. 3.9), and this quantity reflects to the value of template rate. Similarly, Fig. 3.7 and 3.8 show the trade-off of the identification rate versus the privacy-leakage rate and the secrecy rate versus the privacy-leakage rate, respectively. The general behaviors are similar to Fig. 3.5 and 3.6. Finally, one can see that the identification and secrecy rates are in trade-off relation from Fig. 3.9, and from Fig. 3.10, when the identification rate rises, both the privacy-leakage rate and the template rate increase.

3.3.3 Overviews on the Proofs of Theorem 3.1 and 3.2

The proofs of Theorem 3.1 (generated-secret BIS model) and 3.2 (chosen-secret BIS model) are provided in Section 3.4 and 3.5, respectively. Each proof contains two parts; achievability and converse parts. We prove these theorems based on a technique involving two auxiliary RVs U and V, i.e., the constraints in the regions (3.20) and (3.21). Basically, the proof of Theorem 3.1 covers the derivation of Theorem 3.2. The difference is that one time-pad operation is used as an extra layer to mask the chosen-secret key for secure transmission. Here, we mainly mention the proof of Theorem 3.1. The converse part follows from standard arguments where the assistance of auxiliary RVs and Fano's inequality plays an essential role. For deriving the cardinality of auxiliary RVs U and V, we apply the support lemma, introduced in [15] and simplified in [19], to find their upper bounds. In the achievability part, we make use of a combination of random coding and binning, where the binning is used to decrease the rate of sequences associating to the secret keys. In the proof, the auxiliary RVs U and V correspond to the secret keys and the templates of users, respectively.

3.4 Proof of Theorem 3.1

We take a standard information theoretic approach, in which the proof is divided into two parts: the achievability (direct) and converse parts.

3.4.1 Achievability (Direct) Part

First, we fix $\delta > 0$ arbitrarily small, and a block length n. We also fix test channels $P_{U|Y}$ and $P_{V|U}$. We set[1] $R_I = I(Z;V) - \delta$, $R_S = I(Z;U|V) - \delta$, $R_J = I(Y;U) - I(Z;U) + I(Z;V) + 3\delta$, and $R_L = I(X;U) - I(Z;U) + I(Z;V) + 3\delta$. We also set $M_I = 2^{nR_I}$, $M_S = 2^{nR_S}$, and $M_J = 2^{nR_J}$, respectively.

Random Code Generation:

Sequences v_m^n are generated i.i.d. from P_V for $m \in [1, N_V]$, where$N_V = 2^{n(I(Y;V)+\delta)}$. For each m, sequences $u_{k|m}^n$ are generated from the memoryless channel $P_{U^n|V^n=v_m^n}$ for $k \in [1, N_U]$, where $N_U = 2^{n(I(Y;U|V)+\delta)}$. Divide these sequences equally from the first index into $N_B = 2^{n(I(Y;U|V)-I(Z;U|V)+2\delta)}$ bins. That is, the first bin contains $\{u_{1|m}^n, \cdots, u_{M_S|m}^n\}$, the second bin contains $\{u_{M_S+1|m}^n, \cdots, u_{2M_S|m}^n\}$, and so on. Consequently, each bin contains exactly M_S codewords. Bins are indexed by $b \in [1, N_B]$ and codewords inside a certain bin are indexed by $s \in \mathcal{S}$. Without loss of generality, there exists a one-to-one mapping between k and the pair (b, s).

Encoding (Enrollment):

When encoder f observes the bio-data sequence y_i^n, the encoder looks for (m, k) such that $(y_i^n, v_m^n, u_{k|m}^n) \in \mathcal{T}_\varepsilon^n(YVU)$. In case there are more than one such pairs, the encoder picks one of them uniformly at random. Assume that the encoder found a corresponding pair $(m, k) = (m(i), k(i))$ satisfying the jointly typical condition above. We set the template $j(i) = (m(i), b(i))$ and the secret data to be the corresponding codeword's index $s(i)$ in bin $b(i)$ [2]. $j(i)$ is stored at position i in the database and $s(i)$ is handed back to individual i. If there do not exist such m and k, then we set $j(i) = (1, 1)$ and $s(i) = 1$.

Decoding (Identification):

The decoder has access to all records in the database $\{(m(1), b(1)), \cdots (m(M_I), b(M_I))\}$. When decoder g sees z^n, the noisy version of identified individual sequence x_w^n, it checks whether the codeword pair $(v_{m(i)}^n, u_{b(i),s|m(i)}^n)$ is jointly typical with z^n or not for all $i \in \mathcal{I}$ with some $s \in \mathcal{S}$, i.e. $(z^n, v_{m(i)}^n, u_{b(i),s|m(i)}^n) \in \mathcal{T}_\varepsilon^n(ZVU)$. If there exists a unique pair (i, s) for which this condition holds, then the decoder outputs $(\widehat{w}, \widehat{s(w)}) = (i, s)$ as the estimated index and secret data, respectively. Otherwise, the decoder outputs the index of the template $(1, 1)$ as $\hat{w}$ and $\widehat{s(w)} = 1$ if (i) there does

[1] Due to the Markov chain $V - U - Z$, we have $I(Z;U) - I(Z;V) = I(Z;UV) - I(Z;V) = I(Z;V) + I(Z;U|V) - I(Z;V) = I(Z;U|V)$. In the proof, we use this fact without explanation.

[2] Since there is a one-to-one mapping between k and (b, s), we identify $k(i)$ with $(b(i), s(i))$.

not exist such a pair (i,s), (ii) such a pair (i,s) exists but there are some $s' \neq s$ ($s' \in \mathcal{S}$) such that $(z^n, v^n_{m(i)}, u^n_{b(i),s'|m(i)}) \in \mathcal{T}^n_\varepsilon(ZVU)$ satisfies, or (iii) such a pair (i,s) exists but there are some $i' \neq i$ such that the pair $(v^n_{m(i')}, u^n_{b(i'),s'|m(i')})$ is jointly typical with z^n for some $s' \in \mathcal{S}$.

Analysis of Error Probability:

We evaluate the ensemble average of the error probability, where the average is taken over randomly chosen codebook $\mathcal{C}_n$, which is defined as the set $\{V^n_m, U^n_{k|m} : m \in [1,N_V], k \in [1,N_U]\}$. Let the pair $(M(i),K(i)) = (M(i),B(i),S(i))$ denote the RVs corresponding to the index pair $(m(i),k(i)) = (m(i),b(i),s(i))$ of sequences V^n_m and $U^n_{k|m}$ determined by the encoder for Y^n_i. For individual $W = i$, an possible event of errors occurs at the encoder is:

$$\mathcal{E}_1\colon\ \{(Y^n_i, V^n_m, U^n_{k|m}) \notin \mathcal{T}^n_\varepsilon(YVU) \text{ for all } m \in [1,N_V] \text{ and } k \in [1,N_U]\},$$

and those at the decoder are:

$$\begin{aligned}
&\mathcal{E}_2\colon\ \{(Z^n, V^n_{M(i)}, U^n_{B(i),S(i)|M(i)}) \notin \mathcal{T}^n_\varepsilon(ZVU)\},\\
&\mathcal{E}_3\colon\ \{\exists s' \neq S(i) \text{ s. t. } (Z^n, V^n_{M(i)}, U^n_{B(i),s'|M(i)}) \in \mathcal{T}^n_\varepsilon(ZVU)\},\\
&\mathcal{E}_4\colon\ \{\exists i' \neq i \text{ and } \exists s' \text{ s. t. } (Z^n, V^n_{M(i')}, U^n_{B(i'),s'|M(i')}) \in \mathcal{T}^n_\varepsilon(ZVU)\}.
\end{aligned}$$

Note that the authentication process is guaranteed to be successful if the genuine index and secret key of the identified user are correctly estimated at the decoder, indicating that it is sufficient to focus on assessing the probability of incorrect estimation for the pair at the decoder. Then, the error probability can be bounded as

$$\begin{aligned}
\max_{w\in\mathcal{I}} \Pr\{(\widehat{W}, \widehat{S(W)}) \neq (W, S(W)) | W = i\}\\
&= \Pr\{\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4\}\\
&\overset{\text{(a)}}{\leq} \Pr\{\mathcal{E}_1\} + \Pr\{\mathcal{E}_2|\mathcal{E}^c_1\} + \Pr\{\mathcal{E}_3\} + \Pr\{\mathcal{E}_4\}, \qquad (3.35)
\end{aligned}$$

where (a) follows because $\Pr\{\mathcal{E}_1, \mathcal{E}_2\} = \Pr\{\mathcal{E}_1\} + \Pr\{\mathcal{E}_2 \cap \mathcal{E}^c_1\} \leq \Pr\{\mathcal{E}_1\} + \Pr\{\mathcal{E}_2|\mathcal{E}^c_1\}$.

$\Pr\{\mathcal{E}_1\}$ can be made smaller than δ for large enough n by utilizing the covering lemma [19, Lemma 3.3] because $\frac{1}{n}\log N_V = I(Y;V) + \delta > I(Y;V)$ and $\frac{1}{n}\log N_U = I(Y;U|V) + \delta > I(Y;U|V)$. For $\Pr\{\mathcal{E}_2|\mathcal{E}^c_1\}$, it can also be made smaller than δ by the Markov lemma [14, Lemma 15.8.1]. By applying the packing lemma [19, Lemma 3.1], $\Pr\{\mathcal{E}_3\}$ and $\Pr\{\mathcal{E}_4\}$ are arbitrarily small for large enough n since $\frac{1}{n}\log M_S = I(Z;U|V) - \delta < I(Z;U|V)$ and $\frac{1}{n}\log M_I + \frac{1}{n}\log M_S = I(Z;U) - 2\delta < I(Z;UV)$, respectively.

Therefore, the ensemble average of the error probability can be made that

$$\max_{w\in\mathcal{I}} \Pr\{(\widehat{W}, \widehat{S(W)}) \neq (W, S(W)) | W = i\} \leq 4\delta \qquad (3.36)$$

for large enough n.

Intermediate Steps:

We consider a *virtual* system, where a *partial* decoder g_i is employed, for deriving the upper bound on the privacy-leakage rate. In this system, knowing index i and seeing Z_i^n (defined as the output sequence of X_i^n via $P_{Z|X}$), the partial decoder g_i estimates only the secret data of individual i as $\widehat{S(i)} = d_i(Z_i^n, J(i))$. Note that this system is just for analysis, and the partial decoder is not actually used during the decoding process.

For any given $i \in \mathcal{I}$, the partial decoder d_i operates as follows: observing z_i^n and the template $j(i) = (m(i), b(i))$ in the database, it looks for $s \in \mathcal{S}$ such that $(z_i^n, v_{m(i)}^n, u_{b(i),s|m(i)}^n) \in \mathcal{T}_\varepsilon^n(ZVU)$. It sets $\widehat{s(i)} = s$ if there exists a unique s. Otherwise, it outputs $\widehat{s(i)} = 1$. The potential events of error probability for this case are $\mathcal{E}_2$ and $\mathcal{E}_3$. Letting $P_e(i)$ be the error probability of d_i, we readily see that

$$P_e(i) \le \Pr\{(\widehat{W}, \widehat{S(W)}) \ne (W, S(W)) | W = i\} \le 4\delta, \tag{3.37}$$

where the middle term in (3.37) denotes the error probability of d (in the original BIS) for individual $W = i$.

The function of this partial decoder enables us to bound the following conditional entropy

$$H(S(i)|Z_i^n, J(i), \mathcal{C}_n) \overset{\text{(b)}}{\le} H(S(i)|\widehat{S(i)}) \overset{\text{(c)}}{\le} n\delta_n, \tag{3.38}$$

where

(b) follows because conditioning reduces entropy,

(c) follows because Fano's inequality and (3.37) are applied, and $\delta_n = \frac{1}{n}(1 + 4\delta \log M_I M_S)$.

The bound in (3.38) will be used in the analysis of the privacy-leakage rate.

Lemma 3.1. *For any $i \in \mathcal{I}$, it holds that*

$$\frac{1}{n} H(Y_i^n | J(i), S(i), \mathcal{C}_n) \le H(Y|U) + \delta_n', \tag{3.39}$$

where $\delta_n' > 0$ and $\delta_n' \downarrow 0$.

(Proof) The proof is provided in Appendix A.2. □

Due to the fact that we set $M_S = 2^{nR_S}$ and $M_J = 2^{nR_J}$, the following inequalities hold

$$\frac{1}{n} H(S(i)|\mathcal{C}_n) \le R_S = I(Z;U|V) - \delta, \tag{3.40}$$

$$\frac{1}{n} H(J(i)|\mathcal{C}_n) \le R_J = I(Y;U) - I(Z;U|V) + 3\delta \tag{3.41}$$

with equality when $S(i)$ and $J(i)$ are uniformly distributed on $\mathcal{S}$ and $\mathcal{J}$, respectively, for any codebook $\mathcal{C}_n$.

Hereafter, we shall check the bounds of identification, secrecy, secrecy-leakage, template, and privacy-leakage rates averaged over randomly chosen codebook $\mathcal{C}_n$. In the following analyses, the index i is arbitrarily fixed on $\mathcal{I}$ since we need to show that all conditions in Definition 1 are satisfied.

Analyses of Identification and Template Rates:

From the parameter settings of achievability scheme, it is straight-forward that the conditions (3.7) and (3.9) hold.

Analysis of Secrecy Rate:

The secrecy rate can be evaluated as follows:

$$\begin{aligned}\frac{1}{n}H(S(i)|\mathcal{C}_n) &= \frac{1}{n}\Big\{H(Y_i^n,J(i),S(i)|\mathcal{C}_n) - H(J(i)|S(i),\mathcal{C}_n) - H(Y_i^n|J(i),S(i),\mathcal{C}_n)\Big\}\\ &\overset{\text{(d)}}{\geq} \frac{1}{n}\Big\{H(Y_i^n) - H(J(i)|\mathcal{C}_n) - H(Y_i^n|J(i),S(i),\mathcal{C}_n)\Big\}\\ &\overset{\text{(e)}}{\geq} H(Y) - (I(Y;U) - I(Z;U) + I(Z;V) + 3\delta) - (H(Y|U) + \delta_n')\\ &= I(Z;U) - I(Z;V) - 3\delta - \delta_n'\\ &\overset{\text{(f)}}{=} R_S - 2\delta - \delta_n',\end{aligned} \tag{3.42}$$

where

(d) holds because $(J(i),S(i))$ is a function of Y_i^n,

(e) follows because (3.41) and Lemma 3.1 are applied,

(f) holds because we set $R_S = I(Z;U) - I(Z;V) - \delta$.

Analysis of Secrecy-Leakage:

The amount of leaked information about $S(i)$ from $J(i)$ can be expanded as

$$\begin{aligned}\frac{1}{n}I(J(i);S(i)|\mathcal{C}_n) =& \frac{1}{n}\{H(S(i)|\mathcal{C}_n) + H(J(i)|\mathcal{C}_n) - H(Y_i^n,J(i),S(i)|\mathcal{C}_n)\\ &+ H(Y_i^n|J(i),S(i),\mathcal{C}_n)\}\\ =& \frac{1}{n}H(S(i)|\mathcal{C}_n) + \frac{1}{n}H(J(i)|\mathcal{C}_n) - \frac{1}{n}H(Y_i^n)\\ &+ \frac{1}{n}H(Y_i^n|J(i),S(i),\mathcal{C}_n)\\ \overset{\text{(g)}}{\leq}& I(Z;U|V) - \delta + I(Y;U) - I(Z;U|V) + 3\delta - H(Y) + H(Y|U) + \delta_n'\\ =& 2\delta + \delta_n',\end{aligned} \tag{3.43}$$

where (g) follows because equation (3.41) and Lemma 3.1 are applied.

Analysis of Privacy-Leakage Rate:

In view of (3.11), we start by expanding the privacy-leakage rate $\frac{1}{n}I(X_i^n;J(i)|\mathcal{C}_n)$ as

$$\begin{aligned}\frac{1}{n}I(X_i^n;J(i)|\mathcal{C}_n) &= \frac{1}{n}H(J(i)|\mathcal{C}_n) - \frac{1}{n}H(J(i)|X_i^n,\mathcal{C}_n)\\ &\leq I(Y;U) - I(Z;U) + I(Z;V) + 3\delta - \frac{1}{n}H(J(i)|X_i^n,\mathcal{C}_n), \end{aligned} \tag{3.44}$$

where (3.44) is due to (3.41).

Next, let us focus solely on the conditional entropy in (3.44). It can be evaluated as

$$\begin{aligned}\frac{1}{n}H(J(i)|X_i^n,\mathcal{C}_n) &= \frac{1}{n}H(Y_i^n,J(i)|X_i^n,\mathcal{C}_n) - \frac{1}{n}H(Y_i^n|J(i),X_i^n,\mathcal{C}_n)\\
&\overset{\text{(h)}}{=} \frac{1}{n}H(Y_i^n|X_i^n,\mathcal{C}_n) - \frac{1}{n}H(Y_i^n|M(i),B(i),X_i^n,\mathcal{C}_n)\\
&\overset{\text{(i)}}{=} H(Y|X) - \frac{1}{n}H(Y_i^n|M(i),B(i),S(i),X_i^n,\mathcal{C}_n) - \frac{1}{n}I(S(i);Y_i^n|M(i),B(i),X_i^n,\mathcal{C}_n)\\
&\geq H(Y|X) - \frac{1}{n}H(Y_i^n|M(i),B(i),S(i),X_i^n,\mathcal{C}_n) - \frac{1}{n}H(S(i)|M(i),B(i),X_i^n,\mathcal{C}_n)\\
&\overset{\text{(j)}}{=} H(Y|X) - \frac{1}{n}H(Y_i^n|M(i),B(i),S(i),U_i^n,X_i^n,\mathcal{C}_n) - \frac{1}{n}H(S(i)|M(i),B(i),X_i^n,Z_i^n,\mathcal{C}_n)\\
&\overset{\text{(k)}}{\geq} H(Y|X) - \frac{1}{n}H(Y_i^n|U_i^n,X_i^n,\mathcal{C}_n) - \frac{1}{n}H(S(i)|M(i),B(i),Z_i^n,\mathcal{C}_n)\\
&\overset{\text{(l)}}{\geq} H(Y|X) - H(Y|X,U) - (\delta_n + \delta_n')\\
&= I(Y;U|X) - (\delta_n + \delta_n')\\
&\overset{\text{(m)}}{=} H(U|X) - H(U|Y) - (\delta_n + \delta_n'),\end{aligned} \tag{3.45}$$

where

(h) follows since $J(i)$ is a function of Y_i^n and we have $J(i) = (M(i),B(i))$,

(i) follows because Y_i^n and X_i^n are independent of $\mathcal{C}_n$,

(j) follows because $U^n(B(i),S(i)|M(i))$ is denoted by U_i^n and it is a function of $(M(i),B(i),S(i))$ for the second term, and the Markov chain $S(i) - (M(i),B(i),X_i^n) - Z_i^n$ holds for a given codebook in the last term,

(k) follows because conditioning reduces entropy,

(l) follows as (2.13) in Lemma 1 and Fano's inequality in (3.38) are applied,

(m) holds since we have $H(U|Y,X) = H(U|Y)$ by the Markov chain $U - Y - X$.

From (3.44) and (3.45), we obtain

$$\begin{aligned}\frac{1}{n}I(X_i^n;J(i)|\mathcal{C}_n) &\le H(U)-H(U|Y)-I(Z;U)+I(Z;V)\\&\quad+H(U|Y)-H(U|X)+3\delta+\delta_n+\delta_n'\\&\le I(X;U)-I(Z;U)+I(Z;V)+3\delta+\delta_n+\delta_n'\\&\le R_L+\delta \qquad (3.46)\end{aligned}$$

for all sufficiently large n.

Finally, applying Lemma 2.3 to all results shown above (i.e., Eqs. (3.36), (3.42), (3.43), and (3.46)), there exists at least a good codebook satisfying all the conditions in Definition 3.1 for all large enough n. □

3.4.2 Converse Part

For the converse proof, we consider a more relaxed case where identified individual index W is *uniformly* distributed over $\mathcal{I}$, and (3.6), (3.8), (3.10), and (3.11) in Definition 3.1 are replaced by

$$\Pr\{(\widehat{W},\widehat{S(W)})\neq(W,S(W))\}\le\delta, \qquad (3.47)$$

$$\frac{1}{n}H(S(W)|W)\ge R_S-\delta, \qquad (3.48)$$

$$\frac{1}{n}I(S(W);J(W)|W)\le\delta, \qquad (3.49)$$

$$\frac{1}{n}I(X_W^n;J(W)|W)\le R_L+\delta, \qquad (3.50)$$

respectively. We shall show that the capacity region, which is not smaller than the original one $\mathcal{R}_{GS}$, is contained in the right-hand side of (3.20).

We assume that a rate tuple (R_I,R_S,R_J,R_L) is achievable so that there exists a pair of encoder and decoder (e,d) such that all conditions in Definition 3.1 with replacing (3.6), (3.8), (3.10), and (3.11) by (3.36)–(3.50) are satisfied for any $\delta>0$ and large enough n.

Here, we provide other key lemmas used in this part. For $t\in[1,n]$, we define auxiliary RVs U_t and V_t as

$$U_t=(Z^{t-1},J(W),S(W),W) \qquad (3.51)$$

$$V_t=(Z^{t-1},J(W),W), \qquad (3.52)$$

respectively. We denote a sequence of RVs

$$X_W^n=(X_1(W),\cdots,X_n(W)), \qquad (3.53)$$

$$Y_W^n=(Y_1(W),\cdots,Y_n(W)). \qquad (3.54)$$

Lemma 3.2. *The following Markov chains hold*

$$Z^{t-1} - (Y^{t-1}(W), J(W), S(W), W) - Y_t(W), \tag{3.55}$$

$$Z^{t-1} - (X^{t-1}(W), J(W), S(W), W) - X_t(W). \tag{3.56}$$

(Proof) The proofs are available in Appendix A.3. □

Lemma 3.3. *There exist some RVs U and V satisfying the Markov chain $Z - X - Y - U - V$ and*

$$\sum_{t=1}^{n} I(Z_t; V_t) = nI(Z; V), \tag{3.57}$$

$$\sum_{t=1}^{n} I(Z_t; U_t) = nI(Z; U), \tag{3.58}$$

$$\sum_{t=1}^{n} I(Y_t(W); U_t) = nI(Y; U), \tag{3.59}$$

$$\sum_{t=1}^{n} I(X_t(W); U_t) = nI(X; U). \tag{3.60}$$

(Proof) The proofs are provided in Appendix A.4. □

In the subsequent analyses, we fix auxiliary RVs U and V specified in Lemma 3.3.

Analysis of Identification Rate:

Again note that we are considering the case where W is uniformly distributed in the converse part, and we have

$$\begin{aligned}
\log M_I &= H(W) \\
&= H(W|\boldsymbol{J}, Z^n) + I(W; \boldsymbol{J}, Z^n) \\
&\overset{\text{(a)}}{=} H(W|\boldsymbol{J}, Z^n, \widehat{W}, \widehat{S(W)}) + I(W; \boldsymbol{J}, Z^n) \\
&\overset{\text{(b)}}{\leq} H(W|\widehat{W}, \widehat{S(W)}) + I(W; \boldsymbol{J}, Z^n) \\
&\leq H(W, S(W)|\widehat{W}, \widehat{S(W)}) + I(W; \boldsymbol{J}, Z^n) \\
&\overset{\text{(c)}}{\leq} n\delta_n + I(W; \boldsymbol{J}, Z^n),
\end{aligned} \tag{3.61}$$

where

(a) holds because $(\widehat{W}, \widehat{S(W)})$ is function of $\boldsymbol{J}$ and Z^n,

(b) follows because conditioning reduces entropy,

(c) by applying Fano's inequality with $\delta_n = \frac{1}{n}(1 + \delta \log M_I M_S)$ as in (3.38).

Continue bounding the second term in (3.61),

$$\begin{aligned}
I(W;\boldsymbol{J},Z^n) &= I(W;\boldsymbol{J}) + I(W;Z^n|\boldsymbol{J}) \\
&\overset{\text{(d)}}{=} I(W;Z^n|\boldsymbol{J}) \\
&= H(Z^n|\boldsymbol{J}) - H(Z^n|\boldsymbol{J},W) \\
&\overset{\text{(e)}}{=} H(Z^n|J(W)) - H(Z^n|J(W),W) \\
&\overset{\text{(f)}}{\leq} H(Z^n) - H(Z^n|J(W),W) \\
&= H(Z^n) - H(Z^n|J(W),W) \\
&= \sum_{t=1}^{n} \Big\{ H(Z_t) - H(Z_t|Z^{t-1},J(W),W) \Big\} \\
&= \sum_{t=1}^{n} I(Z_t;Z^{t-1},J(W),W) \\
&= \sum_{t=1}^{n} I(Z_t;V_t) \\
&\overset{\text{(f)}}{=} nI(Z;V),
\end{aligned} \tag{3.62}$$

where

(d) follows because W is independent of other RVs,

(e) follows because only $J(W)$ is possibly dependent on Z^n,

(f) follows because conditioning reduces entropy,

(g) follows because of (3.57) in Lemma 3.3.

Thus, from (3.7), (3.61), and (3.62), we obtain

$$R_I \leq I(Z;V) + \delta + \delta_n, \tag{3.63}$$

where $\delta_n = \frac{1}{n}(1 + \delta \log M_I M_S)$ and[3] $\delta_n \downarrow 0$ as $n \to \infty$ and $\delta \downarrow 0$.

Analysis of Secrecy Rate:

[3] Willems et al. [76] characterized the identification capacity of the system, where the decoder estimates only the user index, and showed that $\frac{1}{n}\log M_I \leq I(Y;Z) + \delta$ for all sufficiently large n. Since the constraints imposed on the system addressed in this thesis are more rigorous than the ones in [76], it is trivial that $\frac{1}{n}\log M_I$ for this system cannot be larger than $I(Y;Z) + \delta$. Moreover, it holds that $\frac{1}{n}\log M_S \leq \log|\mathcal{Y}|$ because $S(i)$ is a function of Y_i^n. Therefore, for large enough n, we have that $\delta_n = \frac{1}{n} + \frac{\delta}{n}\log M_I M_S \leq \frac{1}{n} + \delta(\log|\mathcal{Y}||\mathcal{Z}| + \delta)$, and it converges to zero when $n \to \infty$ and $\delta \downarrow 0$.

This analysis is similar to the analysis of identification rate, which we have already seen above. We begin by considering the entropy of secret data as follows:

$$\begin{aligned}
H(S(W)|W) &= H(S(W)|\boldsymbol{J},Z^n,W)+I(S(W);\boldsymbol{J},Z^n|W)\\
&= H(S(W)|\boldsymbol{J},Z^n,\widehat{W},\widehat{S(W)})+I(S(W);\boldsymbol{J},Z^n|W)\\
&\le H(S(W)|\widehat{W},\widehat{S(W)})+I(S(W);\boldsymbol{J},Z^n|W)\\
&\le H(W,S(W)|\widehat{W},\widehat{S(W)})+I(S(W);\boldsymbol{J},Z^n|W)\\
&= H(W,S(W)|\widehat{W},\widehat{S(W)})+I(S(W);\boldsymbol{J}|W)+I(S(W);Z^n|\boldsymbol{J},W)\\
&\overset{\text{(g)}}{=} H(W,S(W)|\widehat{W},\widehat{S(W)})+I(S(W);J(W)|W)+I(S(W);Z^n|J(W),W),
\end{aligned} \tag{3.64}$$

where (g) follows because bio-data sequence of each individual is generated independently, so only $J(W),S(W)$, and Z^n are possibly dependent on each other.
For the third term in (3.64),

$$\begin{aligned}
&I(S(W);Z^n|J(W),W)\\
&\quad= H(Z^n|J(W),W)-H(Z^n|J(W),S(W),W)\\
&\quad= H(Z^n)-H(Z^n|J(W),S(W),W)-(H(Z^n)-H(Z^n|J(W),W))\\
&\quad\overset{\text{(h)}}{=} \sum_{t=1}^{n}\Big\{H(Z_t)-H(Z_t|Z^{t-1},J(W),S(W),W)\Big\}-\sum_{t=1}^{n}\Big\{H(Z_t)-H(Z_t|Z^{t-1},J(W),W)\Big\}\\
&\quad= \sum_{t=1}^{n}\Big\{I(Z_t;U_t)-I(Z_t;V_t)\Big\}\\
&\quad\overset{\text{(i)}}{=} n(I(Z;U)-I(Z;V)),
\end{aligned} \tag{3.65}$$

where

(h) holds because each symbol of Z^n is i.i.d,

(i) holds due to (3.57) and (3.58) in Lemma 3.3.

Therefore, from (3.48), (3.49), (3.64), (3.65), and Fano's inequality, we have

$$R_S \le I(Z;U)-I(Z;V)+2\delta+\delta_n. \tag{3.66}$$

Analysis of Template Rate:

It follows from (3.9) that

$$\begin{aligned}
n(R_J+\delta) &\ge \log M_J \ge \max_{w\in\mathcal{I}} H(J(w)) \ge H(J(W)|W)\\
&= I(Y_W^n;J(W)|W)\\
&= I(Y_W^n;J(W),S(W),Z^n|W)-I(Y_W^n;Z^n|J(W),W)-I(Y_W^n;S(W)|J(W),Z^n,W).
\end{aligned} \tag{3.67}$$

Now let us focus on each term in (3.67) separately. For the first term,

$$\begin{aligned}
I(Y_W^n;J(W),S(W),Z^n|W) &= I(Y_W^n;J(W),S(W)|W)+I(Y_W^n;Z^n|J(W),S(W)|W) \\
&= \sum_{t=1}^{n}\Big\{H(Y_t(W))-H(Y_t(W)|Y^{t-1}(W),J(W),S(W),W)\Big\} \\
&\quad +H(Z^n|J(W),S(W),W)-H(Z^n|J(W),S(W),Y_W^n,W) \\
&\overset{\text{(j)}}{=} \sum_{t=1}^{n}\Big\{H(Y_t(W))-H(Y_t(W)|Z^{t-1},Y^{t-1}(W),J(W),S(W),W)\Big\} \\
&\quad +\sum_{t=1}^{n}H(Z_t|Z^{t-1},J(W),S(W),W)-H(Z^n|Y_W^n,W) \\
&\overset{\text{(k)}}{\geq} \sum_{t=1}^{n}\Big\{H(Y_t(W))-H(Y_t(W)|Z^{t-1},J(W),S(W),W)\Big\} \\
&\quad +\sum_{t=1}^{n}H(Z_t|U_t)-nH(Z|Y) \\
&= \sum_{t=1}^{n}\Big\{I(Y_t(W);U_t)+H(Z_t|U_t)\Big\}-nH(Z|Y), \qquad (3.68)
\end{aligned}$$

where

(j) holds from (3.55) in Lemma 3.2 and $(S(W),J(W))$ is a function of Y_W^n,

(k) follows because conditioning reduces entropy.

For the second term,

$$\begin{aligned}
I(Y_W^n;Z^n|J(W)) &= H(Z^n|J(W),W)-H(Z^n|J(W),Y_W^n,W) \\
&= \sum_{t=1}^{n}H(Z_t|Z^{t-1},J(W),W)-H(Z^n|Y_W^n) \\
&= \sum_{t=1}^{n}H(Z_t|V_t)-nH(Z|Y). \qquad (3.69)
\end{aligned}$$

For the last one,

$$\begin{aligned}
I(Y_W^n;S(W)|J(W),Z^n,W) &\leq H(S(W)|J(W),Z^n,W) \\
&= H(S(W)|\boldsymbol{J},Z^n,W) \\
&= H(S(W)|\boldsymbol{J},Z^n,\widehat{W},\widehat{S(W)}) \\
&\overset{\text{(l)}}{\leq} H(S(W)|\widehat{W},\widehat{S(W)}) \\
&\overset{\text{(m)}}{\leq} n\delta_n, \qquad (3.70)
\end{aligned}$$

where

(l) follows because conditioning reduces entropy,

(m) follows due to Fano's inequality.

Finally, substituting (3.68)–(3.70) into (3.67), the last terms in (3.68) and (3.69) cancel out each other, and we obtain

$$\begin{aligned}
R_J+\delta &\geq \frac{1}{n}\sum_{t=1}^{n}\{I(Y_t(W);U_t)+H(Z_t|U_t)-H(Z_t|V_t)\}-\delta_n \\
&= \frac{1}{n}\sum_{t=1}^{n}\{I(Y_t(W);U_t)-I(Z_t;U_t)+I(Z_t;V_t)\}-\delta_n \\
&= I(Y;U)-I(Z;U)+I(Z;V)-\delta_n, \qquad (3.71)
\end{aligned}$$

where (3.71) follows due to (3.57)–(3.59) in Lemma 3.3.

Analysis of Privacy-Leakage Rate:

From (3.50), it follows that

$$\begin{aligned}
n(R_L+\delta) &\geq I(X_W^n;J(W)|W) \\
&= I(X_W^n;J(W),S(W),Z^n|W)-I(X_W^n;Z^n|J(W),W) \\
&\quad -I(X_W^n;S(W)|J(W),Z^n,W). \qquad (3.72)
\end{aligned}$$

Likewise in the analysis of template rate, let us focus on each term in (3.72) separately. For the first term,

$$\begin{aligned}
I(X_W^n;J(W),S(W),Z^n|W) &= I(X_W^n;J(W),S(W)|W)+I(X_W^n;Z^n|J(W),S(W),W) \\
&\overset{\text{(n)}}{\geq} I(X_W^n;J(W),S(W)|W)+H(Z^n|J(W),S(W),W) \\
&\quad -H(Z^n|J(W),X_W^n,W) \\
&\overset{\text{(o)}}{\geq} \sum_{t=1}^{n}\Big\{H(X_t(W))-H(X_t(W)|Z^{t-1},X^{t-1}(W),J(W),S(W),W)\Big\} \\
&\quad +\sum_{t=1}^{n}H(Z_t|Z^{t-1},J(W),S(W),W)-H(Z^n|J(W),X_W^n,W) \\
&\overset{\text{(p)}}{\geq} \sum_{t=1}^{n}\Big\{H(X_t(W))-H(X_t(W)|Z^{t-1},J(W),S(W),W)\Big\} \\
&\quad +\sum_{t=1}^{n}H(Z_t|U_t)-H(Z^n|J(W),X_W^n,W) \\
&= \sum_{t=1}^{n}\Big\{I(X_t(W);U_t)+H(Z_t|U_t)\Big\}-H(Z^n|J(W),X_W^n,W), \qquad (3.73)
\end{aligned}$$

where

(n) follows because conditioning reduces entropy,

(o) holds from (3.56) in Lemma 3.2,

(p) follows because conditioning reduces entropy.

For the second term,

$$\begin{aligned} I(X_W^n;Z^n|J(W),W) &= H(Z^n|J(W),W) - H(Z^n|J(W),X_W^n,W) \\ &= \sum_{t=1}^{n} H(Z_t|Z^{t-1},J(W),W) - H(Z^n|J(W),X_W^n,W) \\ &= \sum_{t=1}^{n} H(Z_t|V_t) - H(Z^n|J(W),X_W^n,W), \end{aligned} \tag{3.74}$$

and the last term can be bounded by the same quantity as seen in (3.70):

$$I(X_W^n;S(W)|J(W),Z^n,W) \le n\delta_n. \tag{3.75}$$

Finally, substituting (3.73)–(3.75) into (3.72) and taking similar steps as in (3.71), we obtain

$$\begin{aligned} R_L + \delta &\ge \frac{1}{n}\sum_{t=1}^{n}\{I(X_t(W);U_t) - I(Z_t;U_t) + I(Z_t;V_t)\} - \delta_n \\ &= I(X;U) - I(Z;U) + I(Z;V) - \delta_n, \end{aligned} \tag{3.76}$$

where (3.76) follows due to (3.57), (3.58), and (3.60) in Lemma 3.3.

To complete the proof of Theorem 3.1, we discuss the bounds on the cardinalities of auxiliary RVs. For proving the bound on the cardinality of alphabet $\mathcal{U}$ in the regions $\mathcal{A}_1$ and $\mathcal{A}_2$, we use the support lemma [15], [19, Appendix C] to show that RV U should have $|\mathcal{Y}|-1$ elements to preserve P_Y and add three more elements to preserve $I(Z;U)$, $I(Y;U)$, and $I(X;U)$. This implies that it suffices to take $|\mathcal{U}| \le |\mathcal{Y}|+2$ for preserving the regions. Similarly, to bound the cardinalities of alphabets $\mathcal{U}$ and $\mathcal{V}$ in the region $\mathcal{A}_3$ and $\mathcal{A}_4$, we also utilize the same lemma to show that $|\mathcal{V}| \le |\mathcal{Y}|+3$ and $|\mathcal{U}| \le (|\mathcal{Y}|+2)(|\mathcal{Y}|+3)$ suffice to preserve P_Y, $I(Z;V)$, $I(Z;U)$ $(= I(Z;U,V))$, $I(Y;U)$, and $I(X;U)$.

Eventually, letting $n \to \infty$ and $\delta \downarrow 0$ in (3.63), (3.66), (3.71), and (3.76), we can see that the capacity region is contained in the right-hand side of (3.20).

□

3.5 Proof of Theorem 3.2

In this section, we only give a guideline of how to prove Theorem 3.2. The theorem can be mostly derived by the same arguments of proving Theorem 3.1. The difference is that an one-time pad operation is used to mask the chosen secret key for secure transmission between the encoder and decoder.

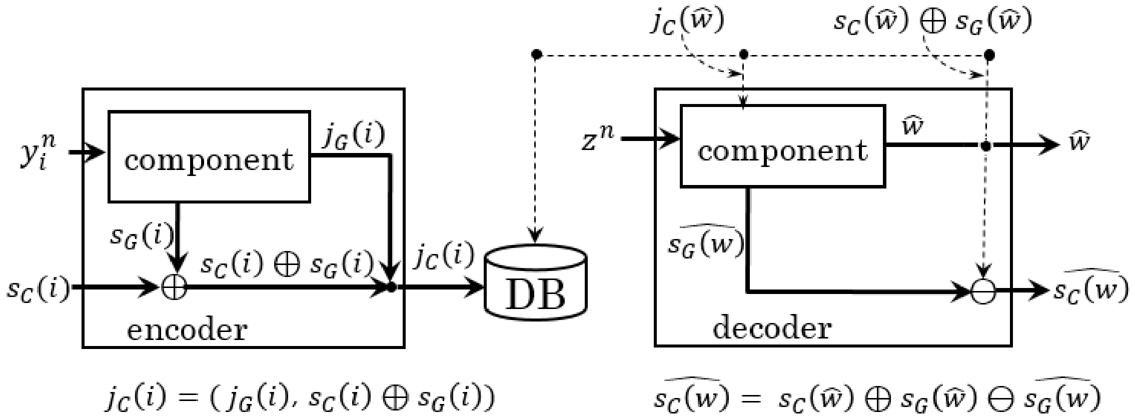

Fig. 3.11 Encoder and decoder of chosen-secret BIS model for the achievability scheme; The components are the encoder and decoder of the generated-secret BIS model and their descriptions are clearly written in Section 3.4.1.

3.5.1 Achievability (Direct) Part

In order to avoid the confusion in the following arguments, we introduce some new notations which are used only in this part. The pairs $(J_C(i), S_C(i))$ and $(J_G(i), S_G(i))$ denote the template and the secret key of individual i for chosen- and generated-secret BIS encoders, respectively. Moreover, M_{J_C} and M_{J_G} denote the number of templates of the chosen- and generated-secret BIS models[4].

Overviews of achievability proof:

The proof idea of this part is based on the achievability proof of the generated-secret BIS model provided in Section 3.4. The difference is that the encoder and decoder of the generated-secret BIS model are used as components inside the encoder and decoder of the chosen-secret BIS model as shown in Fig. 3.11. For encoding in the chosen-secret BIS model, a so-called masking layer (one-time pad operation) is used to mask $s_C(w) \in \mathcal{S}$ for secure transmission by using $s_G(w) \in \mathcal{S}$ as $s_C(w) \oplus s_G(w)$. The template $j_C(w)$ is the combined information of $j_G(i)$ and the masked data $s_C(w) \oplus s_G(w)$, i.e.,

$$j_C(w) = (j_G(w), s_C(w) \oplus s_G(w)). \tag{3.77}$$

For decoding, it first uses the decoder of the generated-secret BIS model to estimate the pair $(\widehat{w}, \widehat{s_G(w)})$ and afterward the secret key is retrieved by

$$\widehat{s_C(w)} = s_C(\widehat{w}) \oplus s_G(\widehat{w}) \ominus \widehat{s_G(i)}, \tag{3.78}$$

where $\oplus$ and $\ominus$ denote addition and subtraction modulo M_S. One-time pad system was proposed by Verman [73] and this technique is also used in many studies such as [2], [32], [29], [21].

[4] Normally, $J_C(i)$, $S_C(i)$, and M_{J_C} are denoted by $J(i)$, $S(i)$, and M_J in other sections of this chapter.

Parameter Settings:

First, we define R_{J_G} and R_{J_C} as the template rates in the generated- and chosen-secret BIS models encoders, respectively. Let δ be a small enough positive and fix a block length n. We choose test channels $P_{U|Y}$ and $P_{V|U}$. Next, we set

$$\begin{aligned} R_I &= I(Z;V)-\delta \\ R_S &= I(Z;U)-I(Z;V)-\delta \\ R_{J_C} &= I(Y;U)+\delta \\ R_L &= I(X;U)-I(Z;U)+I(Z;V)+2\delta. \end{aligned} \tag{3.79}$$

We also set the number of individuals $M_I = 2^{nR_I}$, the number of secret key $M_S = 2^{nR_S}$, and the number of templates $M_{J_C} = 2^{nR_{J_C}}$ for the chosen-secret BIS encoder and $M_{J_G} = \frac{M_{J_C}}{M_S} = 2^{n(I(Y;U)-I(Z;U)+I(Z;V)+2\delta)}$ for the generated-secret BIS encoder, respectively.

Random Code Generation:

The operation is the same as the one we have seen in the random code generation of the achievability proof of Theorem 3.1, so we omit the details.

Encoding (Enrollment):

When the generated-secret BIS encoder, deployed as a component inside the chosen-secret BIS encoder, observes the bio-data sequence $y_i^n \in \mathcal{Y}^n$, the component looks for (m,k) such that $(y_i^n, v_m^n, u_{k|m}^n) \in \mathcal{T}_\varepsilon^{(n)}(YVU)$. In case there are more than one such pairs, the component picks one of them uniformly at random. Assume that the component found a corresponding pair (m,k), denoted as $(m(i),k(i)) = (m(i),b(i),s(i))$, satisfying the jointly typical condition above. Then, the component sets $j_G(i) = (m(i),b(i))$ and $s_G(i) = s(i)$ and shares them to the chosen-secret BIS encoder. After that, the chosen-secret BIS encoder uses $s_G(i)$ to mask the secret $s_C(i)$ as $s_C(i) \oplus s_G(i)$. This masked information is combined with $j_G(i)$ to form the template $j_C(i)$ as

$$j_C(i) = (j_G(i), s_C(i) \oplus s_G(i)) = (m(i), b(i), s_C(i) \oplus s_G(i)) \tag{3.80}$$

The template is stored at position i in the database. If there do not exist such m and k, the component shares $j_G(i) = (1,1)$ and $s_G(i) = 1$ to the chosen-secret BIS encoder. In this case, the chosen-secret BIS encoder declares error.

Decoding (Identification):

The generated-secret BIS decoder, embedded as a component inside the chosen-secret BIS decoder, has access to all records in the database $\{(m(1),b(1),s_C(1) \oplus s_G(1)), \cdots, (m(M_I),b(M_I),s_C(M_I) \oplus s_G(M_I))\}$ (the chosen-secret BIS decoder also can). When the component receives z^n (the noisy version of identified individual sequence x_w^n), it checks if the codeword pair $(v_{m(i)}^n, u_{b(i),s|m(i)}^n)$ is jointly typical with z^n for all $i \in \mathcal{I}$ with some $s \in \mathcal{S}$, i.e. $(z^n, v_{m(i)}^n, u_{b(i),s|m(i)}^n) \in \mathcal{T}_\varepsilon^{(n)}(ZVU)$. If there exists a unique pair (i,s) for which this condition holds, then the component sets $(\widehat{w}, \widehat{s_G(w)}) = (i,s)$ and forwards the pair $(\widehat{w},\widehat{s_G(w)})$ to the chosen-secret BIS decoder. After getting it, the chosen-secret BIS decoder outputs $\widehat{w}=i$ and $\widehat{s_C(w)}$ as the result of $s_C(\widehat{w})\oplus s_G(\widehat{w})\ominus\widehat{s_G(w)}$. Otherwise, the component shares the index of the template $(1,1)$ and $\widehat{s_G(w)}=1$ to the chosen-secret BIS decoder. Upon detecting these information, the chosen-secret BIS decoder declares error.

Next we check that the conditions of (3.12)–(3.16) in Definition 3.1 averaged over randomly chosen codebook $\mathcal{C}_n$, which is defined as the set $\{V_m^n,U_{k|m}^n,\Pi_m : m\in[1,N_V],k\in[1,N_U]\}$.

Analysis of Error Probability:

For individual $W=i$, the operation at the decoder (3.78) means that $\widehat{S_C(W)}=S_C(W)$ iff $\widehat{S_G(W)}=S_G(W)$. In (3.36), it was revealed that the error probability of the identified individual i for the generated-secret BIS model can be made that $\Pr\{(\widehat{W},\widehat{S_G(W)})\neq(W,S_G(W))|W=i\}\leq 4\delta$. The detailed proof is provided in the analysis of Theorem 3.1.

Therefore, it follows that the error probability of individual i for the chosen-secret BIS model can also be bounded by

$$\Pr\{(\widehat{W},\widehat{S_C(W)})\neq(W,S_C(W))|W=i\}\leq 4\delta \tag{3.81}$$

for large enough n.

Analyses of Identification and Secrecy Rates:

It is easy to confirm that (3.13), (3.15), and (3.14) hold from the parameter settings.

Analysis of Storage Rate:

$$\begin{aligned}\frac{1}{n}\log M_{J_C} &\leq \frac{1}{n}\log M_{J_G}+\frac{1}{n}\log M_S\\ &= I(Y;U)-I(Z;U)+I(Z;V)+2\delta+I(Z;U)-I(Z;V)-\delta\\ &= I(Y;U)+\delta\\ &\leq R_{J_C}+\delta.\end{aligned} \tag{3.82}$$

Analysis of Privacy-Leakage Rate:

It can be proved that

$$I(X_i^n;J_C(i)|\mathcal{C}_n)=I(X_i^n;J_G(i)|\mathcal{C}_n). \tag{3.83}$$

To verify this, first one can easily see that

$$\begin{aligned}I(X_i^n;J_C(i)|\mathcal{C}_n) &= I(X_i^n;J_G(i),S_C(i)\oplus S_G(i)|\mathcal{C}_n)\\ &= I(X_i^n;J_G(i)|\mathcal{C}_n)+I(X_i^n;S_C(i)\oplus S_G(i)|J_G(i),\mathcal{C}_n)\\ &\geq I(X_i^n;J_G(i)|\mathcal{C}_n).\end{aligned} \tag{3.84}$$

Meanwhile, it can be shown that

$$\begin{aligned}
I(X_i^n;J_C(i)|\mathcal{C}_n) &= I(X_i^n;J_G(i),S_C(i)\oplus S_G(i)|\mathcal{C}_n)\\
&= I(X_i^n;J_G(i)|\mathcal{C}_n)+I(X_i^n;S_C(i)\oplus S_G(i)|J_G(i),\mathcal{C}_n)\\
&= I(X_i^n;J_G(i)|\mathcal{C}_n)+H(S_C(i)\oplus S_G(i)|J_G(i),\mathcal{C}_n)\\
&\quad -H(S_C(i)\oplus S_G(i)|X_i^n,J_G(i),\mathcal{C}_n)\\
&\overset{\text{(a)}}{\leq} I(X_i^n;J_G(i)|\mathcal{C}_n)+\log M_S-H(S_C(i)\oplus S_G(i)|X_i^n,J_G(i),S_G(i),\mathcal{C}_n)\\
&= I(X_i^n;J_G(i)|\mathcal{C}_n)+\log M_S-H(S_C(i)|X_i^n,J_G(i),S_G(i),\mathcal{C}_n)\\
&\overset{\text{(b)}}{=} I(X_i^n;J_G(i)|\mathcal{C}_n)+\log M_S-\log M_S\\
&= I(X_i^n;J_G(i)|\mathcal{C}_n),
\end{aligned} \tag{3.85}$$

where

(a) follows as conditioning reduces entropy,

(b) follows because $S_C(i)$ is chosen uniformly from $\mathcal{S}$ and independent of other RVs.

From (3.84) and 3.85, (3.83) clearly holds. By using a result shown in (3.46) of Theorem 3.1, the privacy-leakage of the generated-secret BIS model can be bounded by $\frac{1}{n}I(X_i^n;J_G(i)|\mathcal{C}_n) \leq I(X;U)-I(Z;U)+I(Z;V)+3\delta$ for large enough n. Then, the privacy-leakage of the chosen-secret BIS model can also be made that

$$\begin{aligned}
\frac{1}{n}I(X_i^n;J_C(i)|\mathcal{C}_n) &\leq I(X;U)-I(Z;U)+I(Z;V)+3\delta\\
&= R_L+\delta
\end{aligned} \tag{3.86}$$

for large enough n.

Analysis of Secrecy-Leakage:

It holds that

$$\begin{aligned}
I(J_C(i);S_C(i)|\mathcal{C}_n) &= I(J_G(i),S_C(i)\oplus S_G(i);S_C(i)|\mathcal{C}_n)\\
&= I(J_G(i);S_C(i)|\mathcal{C}_n)+I(S_C(i)\oplus S_G(i);S_C(i)|J_G(i),\mathcal{C}_n)\\
&= I(J_G(i);S_C(i)|\mathcal{C}_n)+H(S_C(i)\oplus S_G(i)|J_G(i),\mathcal{C}_n)\\
&\quad -H(S_C(i)\oplus S_G(i)|J_G(i),S_C(i),\mathcal{C}_n)\\
&\leq I(J_G(i);S_C(i)|\mathcal{C}_n)+\log M_S-H(S_G(i)|J_G(i),S_C(i),\mathcal{C}_n)\\
&\overset{\text{(c)}}{=} \log M_S-H(S_G(i)|J_G(i),\mathcal{C}_n)\\
&= I(J_G(i);S_G(i)|\mathcal{C}_n)+\log M_S-H(S_G(i)|\mathcal{C}_n),
\end{aligned} \tag{3.87}$$

where (c) holds because $S_C(i)$ is chosen independently of $(S_G(i), J_G(i))$ for given $\mathcal{C}_n$. In the analyses of the uniformity of secret key (cf, (3.42)) and secrecy-leakage (cf. (3.43)) of Theorem 3.1, it is shown that

$$\frac{1}{n}H(S_G(i)) \geq \log M_S - 2\delta, \tag{3.88}$$

$$\frac{1}{n}I(J_G(i); S_G(i)|\mathcal{C}_n) \leq 2\delta \tag{3.89}$$

for large enough n. Substituting (3.88) and (3.89) into (3.87), the secrecy-leakage of the chosen-secret BIS model is bounded by

$$\frac{1}{n}I(J_C(i); S_C(i)|\mathcal{C}_n) \leq 4\delta \tag{3.90}$$

for large enough n.

Finally, by applying Lemma 2.3 to above results, there exists at least a good codebook satisfying all conditions in Definition 3.2 for large enough n. □

3.5.2 Converse Part

Similar to the converse part of Theorem 3.1, we consider a more relaxed case where identified individual index W is *uniformly* distributed over $\mathcal{I}$ and (3.12), (3.16), and (3.17) in Definition 3.1 are replaced with the average error criterion

$$\Pr\{(\widehat{W}, \widehat{S(W)}) \neq (W, S(W))\} \leq \delta, \tag{3.91}$$

$$\frac{1}{n}I(S(W); J(W)|W) \leq \delta, \tag{3.92}$$

$$\frac{1}{n}I(X_W^n; J(W)|W) \leq R_L + \delta, \tag{3.93}$$

respectively. We shall show that the capacity region, which is not smaller than the original one $\mathcal{R}_{CS}$, is contained in the right-hand side of (3.21). We assume that a rate tuple (R_I, R_S, R_J, R_L) is achievable.

For $t \in [1, n]$, like the converse part of Theorem 3.1, we define auxiliary RVs U_t and V_t as

$$U_t = (Z^{t-1}, J(W), S(W), W), \tag{3.94}$$

$$V_t = (Z^{t-1}, J(W), W), \tag{3.95}$$

respectively. Though we do not provide the detailed proof, it can be verified that Lemma 3.2 and Lemma 3.3 still hold even for the case where the secret key is chosen independently of bio-data sequences by the same argument shown in Appendix A.3. In the following arguments, we fix auxiliary RVs U and V specified in Lemma 3.3.

Analysis of Identification and Secrecy Rates:

It can be shown that

$$R_I \le I(Z;V) + \delta + \delta_n, \tag{3.96}$$

$$R_S \le I(Z;U) - I(Z;V) + 2\delta + \delta_n, \tag{3.97}$$

where $\delta_n = \frac{1}{n}(1 + \delta \log M_I M_S)$ and $\delta_n \downarrow 0$ as $n \to \infty$. The proofs can be done by similar arguments of the analysis of identification and secrecy rates in the converse part of Theorem 3.1.

Analysis of Template Rate:

From (3.15), it holds that

$$\begin{aligned}
n(R_J + \delta) &\ge \log M_J \\
&\ge \max_{w \in \mathcal{I}} H(J(w)) \\
&\ge H(J(W)|W) \\
&= I(J(W); S(W), Y_W^n | W) \\
&\ge I(J(W); Y_W^n | S(W), W) \\
&= H(Y_W^n | S(W)) - H(Y_W^n | J(W), S(W)) \\
&\overset{\text{(a)}}{=} \sum_{t=1}^{n} \Big\{ H(Y_t(W)) - H(Y_t(W)|J(W), S(W), Y^{t-1}(W)) \Big\} \\
&\overset{\text{(b)}}{=} \sum_{t=1}^{n} \Big\{ H(Y_t(W)) - H(Y_t(W)|J(W), S(W), Y^{t-1}(W), Z^{t-1}) \Big\} \\
&\overset{\text{(c)}}{\ge} \sum_{t=1}^{n} I(Y_t(W); Z^{t-1}, J(W), S(W)) \\
&= \sum_{t=1}^{n} I(Y_t(W); U_t)) \\
&\overset{\text{(d)}}{=} nI(Y;U),
\end{aligned} \tag{3.98}$$

where

(a) holds because $S(W)$ is independent of Y_W^n and each symbol of Y_W^n is i.i.d.,

(b) is due to (3.55) in Lemma 3.2,

(c) follows because conditioning reduces entropy,

(d) holds due to (3.59) in Lemma 3.3.

Thus, we obtain

$$R_J \ge I(Y;U) - \delta. \tag{3.99}$$

Analysis of Privacy-Leakage Rate:

It can be proved that

$$R_L + \delta \geq I(X;U) - I(Z;U) + I(Z;V) - \delta_n. \tag{3.100}$$

For detailed proof, the readers should refer to the analysis of privacy-leakage rate in the converse part of Theorem 3.1 since similar approach is taken.

The cardinality bounds of $\mathcal{U}$ and $\mathcal{V}$ can be derived by the same arguments seen in the previous section.

Finally, by letting $n \to \infty$ and $\delta \downarrow 0$, we obtain that the capacity region is contained in the right-hand side of (3.21) from (3.96), (3.97), (3.99), and (3.100). □

3.6 Summary of Results and Discussion

In this chapter, we deployed a method using two auxiliary RVs to characterize the capacity regions of identification, secrecy, template, and privacy-leakage rates for both the generated- and chosen-secret BIS under the condition that that the prior distribution of the identified individual is unknown. We demonstrated that the characterizations using two auxiliary RVs reduce to the ones using only an auxiliary RV. Compared to the model proposed in [69] and [33], what we newly imposed on our models are:

- treating a noisy channel in the enrollment phase,
- considering a scheme of both compressing template (as in [69] and [81]) and protecting privacy (as in [33]),
- analyzing the capacity region provided that the prior distribution of the identified individual is unknown.

As special cases, it can be checked that our characterizations reduce to the one in [33] where the enrollment channel is noiseless and there is no constraint on the template rate, and also coincide with the ones derived by Günlü and Kramer [21] where there is only one individual.

After showing the capacity regions of the generated- and chosen-secret BIS models, we learned that the results are actually derivable with single auxiliary RV, too. For this case, the converse part can be proved similarly, but the achievability scheme needs to be adapted, especially, the encoding and decoding rules. The detailed proofs are provided in [85].

In [30, Section 3.2.2] and [31, Section 3.4], the constraint on the privacy-leakage is replaced by a conditional version, i.e., $\frac{1}{n}I(X_i^n;J(i)|S(i)) \leq R_L + \delta$. For the VSM, it is shown that the minimum required amount of the privacy-leakage rate for the generated-secret BIS model with unconditional or conditional privacy constraint is the same form. However, for the HSM, it seems that this claim dose not hold. As we require the secrecy-leakage $\frac{1}{n}I(S(i);J(i))$ should be negligible, compared to the unconditional privacy-leakage (3.11) in Definition 3.1, the conditional version is more rigorous. That is obvious from $\frac{1}{n}I(X_i^n;J(i)|S(i)) = \frac{1}{n}I(X_i^n,S(i);J(i)) - \frac{1}{n}I(S(i);J(i)) \sim \frac{1}{n}I(X_i^n,S(i);J(i)) =$

$\frac{1}{n}I(X_i^n;J(i)) + \frac{1}{n}I(S(i);J(i)|X_i^n)$. In VSM, $\frac{1}{n}I(S(i);J(i)|X_i^n)$ is zero since $(S(i);J(i))$ are a function of X_i^n, but this can not be applied to the HSM case where the pair $(S(i);J(i))$ is generated from the sequence Y_i^n, a noisy version of X_i^n. Therefore, the mutual information is likely positive and the minimum amount of the privacy-leakage rate is greater than the one seen in Theorem 3.1.

On the other hand, in the chosen-secret model, the minimum amount of the privacy-leakage rate under unconditional or conditional privacy is characterized differently (cf. [31, Theorem 3.2],[31, Theorem 3.4]) even for the VSM, and this conclusion is possibly applied to the HSM as well. Nevertheless, there are still rooms for investigating these models under the conditional privacy constraint.

Chapter 4

BISs With Both Chosen and Generated Secrecy: DMS

In this chapter, we investigate the fundamental limits of the BIS with a combined usage of chosen- and generated-secret keys. We also allows the two secret keys to be correlated, and the reason of this is because we wish to achieve a higher sum of the identification, chosen- and generated-secrecy rates. In the enrollment phase, for each user, the encoder generates a secret key (generated-secret key) and a template (helper data) by using another secret key (chosen-secret key), chosen independently of biometric identifiers and the bio-data sequence. In the identification phase, observing biometric data sequence, the decoder should estimate index, chosen- and generated-secret keys of the identified user reliably.

In the previous studies such as [21], [29], [33], and [85], the chosen- and generated-secret keys are assumed in the separate models, namely, chosen- and generated-secret BIS models, respectively. However, an interesting question is when the two keys are used in the same system, how the chosen- and generated-secrecy rates affect the fundamental performances of the BIS. The answer to this question has not yet been known, and it is not trivial from the results of the previous studies. A possible application of this model may be the system supporting two-factor authentication based on biometrics as the estimated index can be used to claim who the identified user is, and the chosen- and generated-secret keys may be used for the first and second rounds of authentications. In the present chapter, we are interested in characterizing the optimal trade-off of identification, chosen- and generated-secrecy rates under privacy and storage constraints for the BIS with exponentially many users. In the derivation, the hard part is the evaluation of the privacy-leakage rate in the converse part, and we establish a new lemma for dealing with the difficulty. As a result, the characterization shows that identification, chosen- and generated-secrecy rates are in a trade-off relation, and a larger sum of these rates is achievable compared to the result in [86]. The template rate (storage space) requires to be larger as identification and chosen-secrecy rates rise, similar to an observation for the chosen-secret BIS model in [21], [29], and [33], but it is not affected by the generated-secrecy rate. Unlike the template rate, the privacy-leakage rate increases or decrease in accordance with only the changes

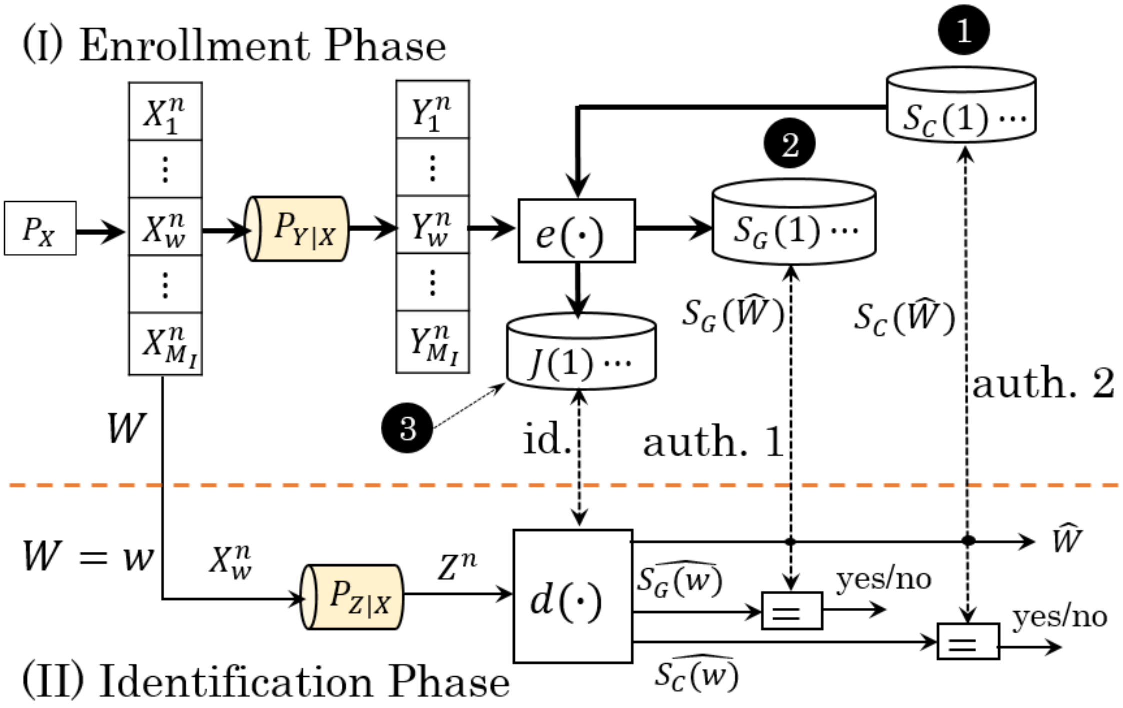

Fig. 4.1 BIS with both chosen and generated secrecy; One can see that two secret keys $S_C(i)$ and $S_G(i)$ appear in the model, and these secret keys and index of the identified user should be estimated reliably at the decoder.

of the identification rate. As special cases, this result reduces to several known characterizations provided in previous studies.

The organization of this chapter is as follow. We describe the basic settings of system model considered in this chapter in Section 4.1, state our main result in Section 4.2, and look into connections of the main result and the results in previous studies. The proof of main result is given in Section 4.4, and a short summary of results and discussion for this chapter follows in Section 4.5.

4.1 Basic Settings of the System Model

The system model considered in this paper is illustrated in Fig. 4.1. ❶, ❷, and ❸ represent the databases of chosen- and generated-secret keys, and templates, respectively. In order to avoid the notation confusion, we call $S_C(i)$ and $S_G(i)$ the chosen- and generated-secret keys, respectively. Let $\mathcal{S}_C = [1 : M_C]$ and $\mathcal{S}_G = [1 : M_G]$ be the sets of the chosen- and generated-secret keys. Lowercase letters $s_C(i) \in \mathcal{S}_C$, $s_G(i) \in \mathcal{S}_G$, and $j(i) \in \mathcal{J}$ stand for the realizations of the two keys and template, respectively. Here, as we have seen in the analysis of the chosen-secret BIS model in Section 3.1, it is also assumed that the chosen-secret key is uniformly distributed on $\mathcal{S}_C$, i.e.,

$$P_{S_C(i)}(s_C(i)) = \frac{1}{M_C} \tag{4.1}$$

for all $i \in \mathcal{I}$ and $s_C(i) \in \mathcal{S}_C$. Observing sequence Y_i^n and $S_C(i)$, provided independently of other RVs from ❶, the encoder e generates $S_G(i)$ and $J(i)$ as

$$(J(i), S_G(i)) = e(Y_i^n, S_C(i)) \tag{4.2}$$

for all $i \in \mathcal{I}$. $J(i)$ is stored at position i in ❸ and $S(i)$ is saved in ❷. This operation is repeated for all users. In the Identification Phase, observing Z^n, the decoder d reconstructs index and both secret keys by

$$(\widehat{W}, \widehat{S_C(w)}, \widehat{S_G(w)}) = d(Z^n, \boldsymbol{J}). \tag{4.3}$$

Moreover, to pass the process of identification, it is required that the first and second authentications must be accepted. In the first authentication, it checks that whether

$$\widehat{S_C(w)} = S_C(\widehat{W}) \tag{4.4}$$

, and in the second authentication phase, it again checks that if

$$\widehat{S_G(w)} = S_G(\widehat{W}) \tag{4.5}$$

or not, and if the condition of each step is matched, decoding is successful. Again note that it is not required that the identified user W must be uniformly distributed on $\mathcal{I}$ as in [40] and [76].

4.2 Problem Formulation and Main Result

In this section, we provide the formal definition of the system and the main result of this study. After that we take a look into the connection between our result and the ones characterized in previous studies. For simplicity, let $E(W)$ and $\widehat{E(W)}$ represent the tuples $(W, S_C(W), S_G(W))$ and $(\widehat{W}, \widehat{S_C(W)}, \widehat{S_G(W)})$, respectively.

Definition 4.1. *A tuple of identification, chosen- and generated-secrecy, template, and privacy-leakage rates* $(R_I, R_C, R_G, R_J, R_L)$ *is said to be* Γ*-achievable for a DMS if for any*[1] $0 \le \Gamma \le \min\{R_C, R_G\}$,

[1] In (4.8) and (4.9), we aim to achieve as large as possible R_C and R_G, indicating these values are very close to their entropies. Thus, it is natural that the mutual information in (4.12) is bounded by the smaller one of the two rates. For the case where $\Gamma > \min\{R_C, R_G\}$, it is impossible to achieve such a value since Γ becomes larger than $\frac{1}{n}H(S_C(i))$ or $\frac{1}{n}H(S_G(i))$, but the degree of correlation for the two secret keys is at most equal to the smaller one of these entropies.

$\delta > 0$, and large enough n there exist pairs of encoders and decoders that satisfy

$$\max_{i\in\mathcal{I}} \Pr\{\widehat{E(W)} \neq E(W)|W=i\} \leq \delta, \tag{4.6}$$

$$\frac{1}{n}\log M_I \geq R_I - \delta, \tag{4.7}$$

$$\frac{1}{n}\log M_C \geq R_C - \delta, \tag{4.8}$$

$$\min_{i\in\mathcal{I}} \frac{1}{n} H(S_G(i)) \geq R_G - \delta, \tag{4.9}$$

$$\frac{1}{n}\log M_J \leq R_J + \delta, \tag{4.10}$$

$$\max_{i\in\mathcal{I}} \frac{1}{n} I(X_i^n; J(i)) \leq R_L + \delta, \tag{4.11}$$

$$\max_{i\in\mathcal{I}} \frac{1}{n} I(S_C(i); S_G(i)) \leq \Gamma, \tag{4.12}$$

$$\max_{i\in\mathcal{I}} \frac{1}{n} I(S_C(i), S_G(i); J(i)) \leq \delta. \tag{4.13}$$

Moreover, $\mathcal{R}^D(\Gamma)$ is defined as the closure of the set of all Γ-achievable rate tuples, called the Γ-capacity region, of the BIS for discrete memoryless sources.

The first main result of this paper is presented below.

Theorem 4.1. *The Γ-capacity region for the system for discrete memoryless source is given by*

$$\begin{aligned}\mathcal{R}^D(\Gamma) = \{(R_I, R_C, R_G, R_J, R_L):\ & R_I + R_C \leq I(Z;U),\\ & R_I + R_C + R_G \leq I(Z;U) + \Gamma,\\ & R_J \geq I(Y;U) - I(Z;U) + R_I + R_C,\\ & R_L \geq I(X;U) - I(Z;U) + R_I,\\ & R_I \geq 0,\ R_C \geq \Gamma \geq 0,\ R_G \geq 0,\\ & \text{for some } U \text{ s. t. } Z - X - Y - U\},\end{aligned} \tag{4.14}$$

where auxiliary RV U takes values in a finite alphabet $\mathcal{U}$ with $|\mathcal{U}| \leq |\mathcal{Y}| + 2$. □

Note that a constraint of $R_I + R_G \leq I(Z;U)$ is redundant in (4.14) due to the fact that it is obvious from the second condition as $R_C \geq \Gamma$. Using a similar technique shown in [29, Sect. IV-A], one can easily check that $\mathcal{R}^D(\Gamma)$ is a convex region.

An explanation of rate constraints in (4.14) is illustrated in Fig. 4.2. In this setting, the decoder is required to reconstruct the index and both secret keys. In [86], Yachongka and Yagi showed that the sum of identification, generated- and chosen-secrecy rates cannot be larger than $I(Z;U)$ if the condition (4.12) is imposed by the perfect secrecy. However, since we permit the chosen- and generated-secret keys to be correlated (non-perfect secrecy), the recognizable value for the sum of these rates exceeds $I(Z;U)$, and the increased quantity is equal to the the degree of correlation between

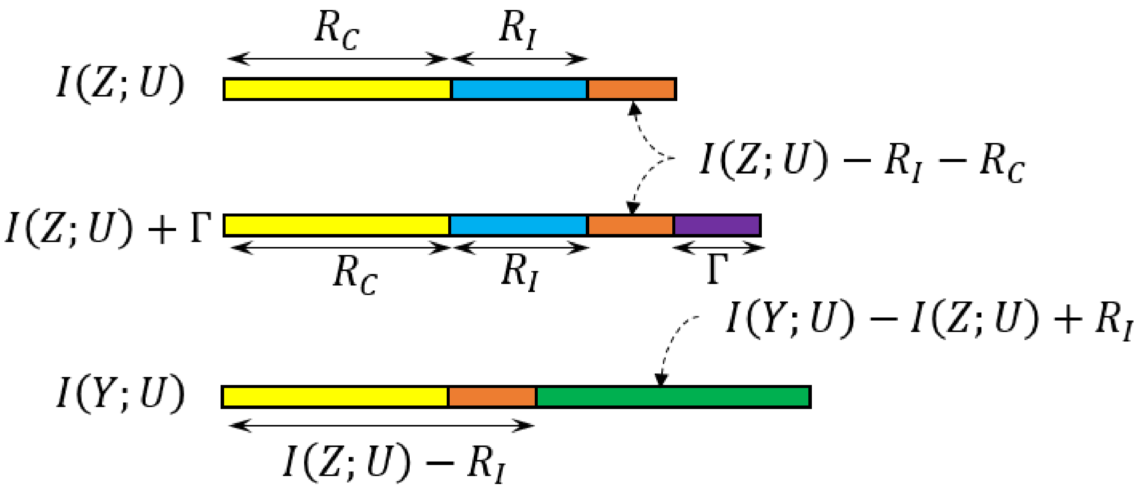

Fig. 4.2 Explanation of rate constraints for Theorem 4.1

the two keys' Γ (cf. the middle band graph in Fig. 4.2). More precisely, R_I and R_C can be any value in the range of $[0, I(Z;U)]$ under a constraint that their sum should be less than $I(Z;U)$ as shown in the top band graph. The generated-secrecy rate R_G can take values up to $I(Z;U)+\Gamma-(R_I+R_C)$, which was originally achieved up to $I(Z;U)-(R_I+R_C)$ for the case of perfect secrecy [86].

The minimum required amount of the template rate is larger than the one of the generated-secret BIS model seen in Theorem 3.1, which is identical to the sum of the rates of yellow part (R_C) and green part ($I(Y;U)-I(Z;U)+R_I$) in the bottom band graph of Fig. 4.2. The constraint of template rate R_J depends on both the identification rate R_I and the chosen-secrecy rate R_C. This is because the storage space increases with the number of users, and we need to attach the information related to the chosen-secret key with the templates in some form.

Fundamentally, the privacy-leakage rate is proportional to the template rate. Compared to the constraint of the template rate in Theorem 3.1, the one in Theorem 4.1 is lower bounded by a bigger value. Therefore, we expect that this increment might lead to leaking a larger amount of the privacy. Surprisingly, the minimum amount of the privacy-leakage rate is characterized in the same form in both theorems. As a matter of fact, the chosen-secrecy rate does not involve, and only the changes of identification rate affect the minimum required amount of the privacy-leakage rate. This is because the portion related to the chosen-secret key stored in the database should be made perfectly confidential, e.g., by using the one-time pad operation, and this information makes no contribution to the privacy-leakage. Similar to the conclusion of Theorem 3.2, the template rate that can be openly observed by the adversary is at least $I(Y;U)-I(Z;U)+R_I$, and for this reason, the minimum value of the privacy-leakage rate becomes $I(X;U)-I(Z;U)+R_I$.

4.3 Special Cases and Overviews of the Proof of Theorem 4.1

4.3.1 Connection to the Results of the Previous Studies and Example

Next, we will take a look into a few special cases. One can check that Theorem 4.1 covers the results provided in previous studies. For instance, in the case of no chosen-secrecy, that is $R_C = 0$, $\mathcal{R}^D(\Gamma)$ naturally reduces to the one given in Theorem 3.1. In the case of no secrecy generation, that is $R_G = 0$, the capacity region, denoted by $\mathcal{R}'$ in this case, is given in the following corollary.

Corollary 4.1.

$$\begin{aligned}\mathcal{R}' = \{(R_I,R_C,R_J,R_L):\ & R_I \geq 0, R_C \geq 0,\\ & R_I + R_C \leq I(Z;U),\\ & R_J \geq I(Y;U) - I(Z;U) + R_I + R_C,\\ & R_L \geq I(X;U) - I(Z;U) + R_I,\\ & \text{for some } U \text{ s. t. } Z - X - Y - U\},\end{aligned} \tag{4.15}$$

where $|\mathcal{U}| \leq |\mathcal{Y}| + 2$. □

Although the expression of $\mathcal{R}'$ and the one given in Theorem 3.2 are different, it can be checked that both are identical. The proof is available in Appendix B.1.

Moreover, in the case where we set R_I to be zero (single user case), the capacity region, denoted by $\mathcal{R}''$ in this case, is obtained.

Corollary 4.2.

$$\begin{aligned}\mathcal{R}'' = \{(R_C,R_G,R_J,R_L):\ & R_C \geq 0, R_G \geq 0,\\ & R_C + R_G \leq I(Z;U),\\ & R_J \geq I(Y;U) - I(Z;U) + R_C,\\ & R_L \geq I(X;U) - I(Z;U),\\ & \text{for some } U \text{ s. t. } Z - X - Y - U\},\end{aligned} \tag{4.16}$$

where $|\mathcal{U}| \leq |\mathcal{Y}| + 2$. □

When $R_C = 0$ (no provision of secret key), one can easily see that $\mathcal{R}''$ is equivalent to the one given in [21, Theorem 1]. Moreover, in case $R_G = 0$ (no generation of secret key), it can also be shown that $\mathcal{R}''$ matches with the region provided in [21, Theorem 2] by a similar argument of proving that $\mathcal{R}'$ and the region in Theorem 3.1 are the same.

Applying the similar arguments in Section 3.3, we can show the following corollary.

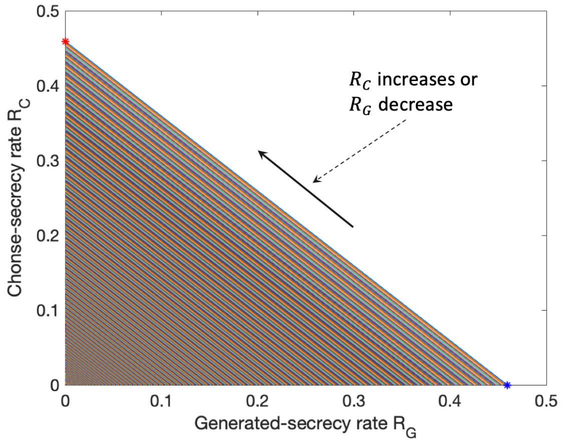

Fig. 4.3 The projection onto $R_G R_C$-plane.

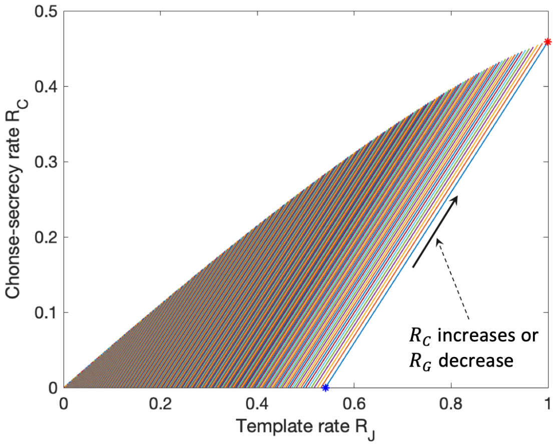

Fig. 4.4 The projection onto $R_J R_C$-plane.

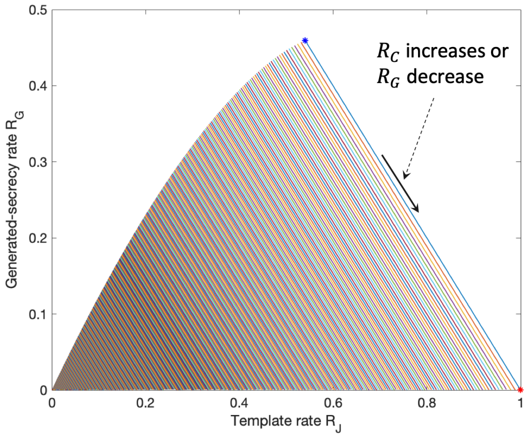

Fig. 4.5 The projection onto $R_J R_G$-plane.

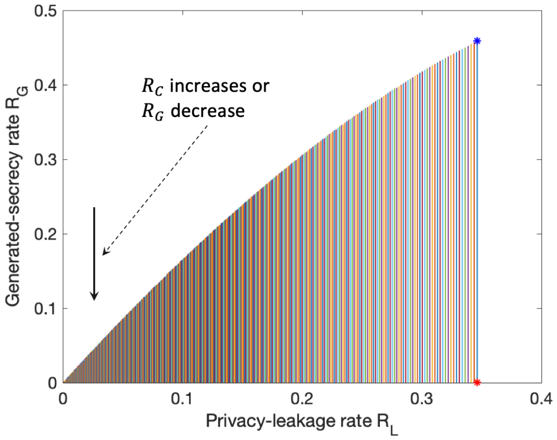

Fig. 4.6 The projection onto $R_L R_G$-plane.

Corollary 4.3. *For binary hidden source, Theorem 4.1 reduces to*

$$R_J = H_b(\gamma * p_E * p_D) - H_b(\gamma) + R_C + R_I,$$

$$R_L = H_b(\gamma * p_E) - H_b(\gamma) + R_I,$$

$$\text{For some } \gamma \in [0, 0.5] \text{ satisfying } R_I + R_C + R_G = 1 - H_b(\gamma * p_E * p_D) + \Gamma,\ R_I \geq 0,\ R_C \geq \Gamma \geq 0,$$

$$\text{and } R_I + R_C \leq 1 - H_b(\gamma * p_E * p_D). \tag{4.17}$$

We calculate the above rate region for $p_E = 0.03$ and $p_D = 0.1$, which is the same setting as Section 3.3, in the case where $R_I = 0$ and $\Gamma = 0$. The results are shown in Fig. 4.3–4.6 and the painted areas represent the achievable rate regions. Blue and red points marked with an asterisk correspond to the points where $R_G = 0$ and $R_C = 0$, respectively. One can see that the chosen- and generated-secrecy rates are in trade-off relation from Fig. 4.3. In Fig. 4.4, it is evident that a greater value of chosen-secrecy rate results in a larger value of template rate. In contrast, when we take a look at the relation of the generated-secrecy and template rates in Fig. 4.5, as the generated-secrecy rate decreases, the template rate is increasingly large. This is due to increase of the chosen-secrecy rate. Finally, Fig. 4.6 show the optimal trade-off of the generated-secrecy and privacy-leakage rates, and the optimal value of R_L is not affected by the change of chosen- or generated-secrecy rates.

4.3.2 Overviews on the Proof of Theorem 4.1

Unlike the technique seen in deriving Theorem 3.1 and 3.2, the main result of this chapter is proved based on a single auxiliary RV U. We also take a standard approach in which the proof of Theorem 4.1 is divided into two parts: achievability and converse parts. The converse part is based on Markov properties of the auxiliary RV and Fano's inequality, and the cardinality bound of auxiliary RV U follows by applying the support lemma [15], [19]. In the achievability part, the argument of random coding is used, and auxiliary sequences of RV U corresponds to the secret keys and the templates of users. Every generated sequence u^n is assigned with three indexes (s_1, s_2, m), where s_1 and m represent the generated-secret key and the dummy message shared between the encoder and decoder to help the estimation of both secret keys. On the other hand, s_2 acts as a random seed to mask the chosen-secret key by one time-pad operation, and this masked information is stored together with the dummy message in the database.

4.4 Proof of Theorem 4.1

In this section, we provide the detailed proof of Theorem 4.1. The proof begins with showing the converse part and then follows by the achievability part.

4.4.1 Converse Part

A same approach as seen in the proofs of the converse part of Theorem 3.1 and 3.2, we assume that W is uniformly distributed on $\mathcal{I}$, and (4.6), (4.9), (4.11), (4.12), and (4.13) are replaced by

$$\Pr\{\widehat{E(W)} \neq E(W)\} \leq \delta, \tag{4.18}$$

$$\frac{1}{n}H(S_G(W)|W) \geq R_G - \delta, \tag{4.19}$$

$$\frac{1}{n}I(X_W^n; J(W)|W) \leq R_L + \delta, \tag{4.20}$$

$$\frac{1}{n}I(S_C(W); S_G(W)|W) \leq \Gamma, \tag{4.21}$$

$$\frac{1}{n}I(S_C(W), S_G(W); J(W)|W) \leq \delta, \tag{4.22}$$

respectively. We demonstrate that even this more relaxed condition, the outer bound of the capacity region coincides with its inner bound derived under the circumstance that the prior distribution of W is unknown.

Suppose that a rate tuple $(R_I, R_C, R_G, R_J, R_L)$ is achievable. Before proceeding to detailed proofs, we provide some useful lemmas. For $t \in [1:n]$, we define an auxiliary RV

$$U_t = (Z^{t-1}, T(W)), \tag{4.23}$$

where $T(W) = (J(W), S_C(W), S_G(W), W)$.

Lemma 4.1. *The following Markov chains hold*

$$Z^{t-1} - (Y^{t-1}(W), T(W)) - Y_t(W), \tag{4.24}$$

$$Z^{t-1} - (X^{t-1}(W), T(W)) - X_t(W). \tag{4.25}$$

(Proof): The proofs can be done by similar arguments shown in Appendix A.3. □

The following lemma plays a key role in the analysis of privacy-leakage, which will be seen in the sequel.

Lemma 4.2. *It holds that*

$$\frac{1}{n} I(Z^n; \boldsymbol{J}, W) \geq R_I - (\delta + \delta_n), \tag{4.26}$$

where $\delta_n = \frac{1}{n}(1 + \delta \log M_I M_C M_G)$*, and* $\delta_n \downarrow 0$ *as* $n \to \infty$ *and* $\delta \downarrow 0$*.*

(Proof): We can prove the above lemma by a few steps as follows:

$$\begin{aligned}\frac{1}{n} I(Z^n; \boldsymbol{J}, W) &\geq \frac{1}{n} I(Z^n; W|\boldsymbol{J}) \\ &= \frac{1}{n} H(W|\boldsymbol{J}) - \frac{1}{n} H(W|Z^n, \boldsymbol{J}) \\ &\overset{\text{(a)}}{\geq} \frac{1}{n} H(W) - \delta_n \\ &= \frac{1}{n} \log M_I - \delta_n \\ &\overset{\text{(b)}}{\geq} R_I - (\delta + \delta_n),\end{aligned} \tag{4.27}$$

where

(a) follows because Fano's inequality is applied and W is statistically independent of all templates,

(b) follows since (4.7) is applied, and we have that $\delta_n = \frac{1}{n}(1 + \log M_I M_C M_G)$ goes to zero as $\delta \downarrow 0$ and $n \to \infty$.

□

Analysis of Identification, Chosen- and Generated-Secrecy Rates: We begin with considering the joint entropy of $E(W) = (W, S_G(W), S_C(W))$ as

$$\begin{aligned}
H(E(W)) &= H(E(W)|Z^n, \boldsymbol{J}) + I(E(W); Z^n, \boldsymbol{J}) \\
&\overset{\text{(c)}}{\leq} n\delta_n + I(E(W); \boldsymbol{J}) + I(E(W); Z^n|\boldsymbol{J}) \\
&\leq n\delta_n + I(W; \boldsymbol{J}) + I(S_C(W), S_G(W); \boldsymbol{J}|W) + H(Z^n|\boldsymbol{J}) \\
&\quad - H(Z^n|\boldsymbol{J}, S_C(W), S_G(W), W) \\
&\overset{\text{(d)}}{=} n\delta_n + I(S_C(W), S_G(W); J(W)|W) + H(Z^n|J(W)) - H(Z^n|T(W)) \\
&\overset{\text{(e)}}{\leq} n(\delta_n + \delta) + H(Z^n) - H(Z^n|T(W)) \\
&= \sum_{t=1}^{n} \Big\{ H(Z_t) - H(Z_t|Z^{t-1}, T(W)) \Big\} + n(\delta + \delta_n) \\
&= \sum_{t=1}^{n} I(Z_t; U_t) + n(\delta + \delta_n) \\
&\overset{\text{(f)}}{=} n(I(Z;U) + \delta + \delta_n),
\end{aligned} \tag{4.28}$$

where

(c) follows because Fano's inequality is applied,

(d) holds because W is independent of other RVs and only $J(W)$ is possibly dependent on Z^n, $S_C(W)$, and $S_G(W)$,

(e) follows because (4.22) is used and conditioning reduces entropy,

(f) holds due to (3.58) in Lemma 3.3.

In the opposite direction, we can also derive the following relation

$$\begin{aligned}
H(E(W)) = H(W, S_C(W), S_G(W)) &= H(W) + H(S_C(W), S_G(W)|W) \\
&\overset{\text{(g)}}{=} H(W) + H(S_C(W)) + H(S_G(W)|W) - I(S_C(W); S_G(W)|W) \\
&\overset{\text{(h)}}{\geq} \log M_I + \log M_C + H(S_G(W)|W) - I(S_C(W); S_G(W)|W) \\
&\overset{\text{(i)}}{\geq} n(R_I + R_C + R_G - \Gamma - 3\delta),
\end{aligned} \tag{4.29}$$

where

(g) holds because W and $S_C(W)$ are independent of each other,

(h) follows since W and $S_C(W)$ is uniformly distributed on $\mathcal{I}$ and $\mathcal{S}_C$, respectively,

(i) is due to (4.7), (4.8), (4.19), and (4.21).

From (4.28) and (4.29), we obtain

$$R_I + R_C + R_G \leq I(Z;U) + \Gamma + 4\delta + \delta_n. \tag{4.30}$$

Analysis of Identification and Chosen-secrecy Rates: It holds that

$$H(W, S_C(W)) \leq H(W, S_C(W), S_G(W)) \leq n(I(Z;U) + \delta + \delta_n), \tag{4.31}$$

and we have

$$H(W, S_C(W)) \overset{\text{(j)}}{=} H(W) + H(S_C(W)) = \log M_I + \log M_C \overset{\text{(k)}}{\geq} R_I + R_C - 2\delta, \tag{4.32}$$

where

(j) follows because W is independent of $S_C(W)$,

(k) is due to (4.7) and (4.8).

From (4.7) and (4.8), we obtain

$$R_I + R_C \leq I(Z;U) + 3\delta + \delta_n. \tag{4.33}$$

Analysis of Template Rate: We have that

$$\begin{aligned}
n(R_J + \delta) &\geq \log M_J \geq \max_{w \in \mathcal{I}} H(J(w)) \geq \frac{1}{M_I} \sum_{w=1}^{M_I} H(J(W)|W = w) = H(J(W)|W) \\
&= I(Y_W^n, S_C(W); J(W)|W) \\
&\overset{\text{(l)}}{=} H(Y_W^n) + H(S_C(W)) - H(Y_W^n, S_C(W), S_G(W)|J(W), W) \\
&\overset{\text{(m)}}{=} H(Y_W^n) - H(Y_W^n|T(W)) + \log M_C - H(S_C(W), S_G(W)|J(W)) \\
&= \sum_{t=1}^{n} \Big\{ H(Y_t(W)) - H(Y_t(W)|Y^{t-1}(W), T(W)) \Big\} + \log M_C \\
&\quad - H(S_C(W), S_G(W)) + I(S_C(W), S_G(W)); J(W)) \\
&\overset{\text{(n)}}{\geq} \sum_{t=1}^{n} \Big\{ H(Y_t(W)) - H(Y_t(W)|Z^{t-1}, Y^{t-1}(W), T(W)) \Big\} \\
&\quad + n(R_C - \delta) - H(S_C(W), S_G(W)) \\
&\overset{\text{(o)}}{\geq} \sum_{t=1}^{n} \Big\{ H(Y_t(W)) - H(Y_t(W)|Z^{t-1}, T(W)) \Big\} + n(R_C - \delta) - n(I(Z;U) - R_I + 2\delta + \delta_n) \\
&= \sum_{t=1}^{n} I(Y_t(W); U_t) - n(I(Z;U) - R_I - R_C + 3\delta + \delta_n) \\
&\overset{\text{(p)}}{=} n(I(Y;U) - I(Z;U) + R_I + R_C - 3\delta - \delta_n),
\end{aligned} \tag{4.34}$$

where

(l) holds because Y_W^n is independent of $S_C(W)$ and $S_G(W)$ is a function of $(Y_W^n, S_C(W))$,

(m) holds since $S_C(W)$ is uniformly distributed on $\mathcal{S}_C$ and W is independent of other RVs,

(n) follows due to (4.8) and (4.24),

(o) follows as conditioning reduces entropy and from (4.28), we have $H(S_C(w), S_G(w)) \leq n(I(Z;U) + \delta + \delta_n) - H(W) \leq n(I(Z;U) - R_I + 2\delta + \delta_n)$,

(p) holds due to (3.59) in Lemma 3.3.

Therefore,

$$R_J \geq I(Y;U) - I(Z;U) + R_C + R_I - 4\delta - \delta_n. \tag{4.35}$$

Analysis of Privacy-Leakage Rate: We expand the left-hand side of (4.20) as

$$\begin{aligned}
I(X_W^n; J(W)|W) &= I(X_W^n; J(W), S_C(W), S_G(W), Z^n|W) - I(X_W^n; S_C(W), S_G(W), Z^n|J(W), W) \\
&= I(X_W^n; J(W), S_C(W), S_G(W)|W) + I(X_W^n; Z^n|J(W), S_C(W), S_G(W), W) \\
&\quad - H(S_C(W), S_G(W), Z^n|J(W), W) + H(S_C(W), S_G(W), Z^n|J(W), W, X_W^n) \\
&= \sum_{t=1}^{n} I(X_t; X^{t-1}(W), T(W)) + H(Z^n|T(W)) - H(Z^n|T(W), X_W^n) \\
&\quad - H(Z^n|J(W), W) - H(S_C(W), S_G(W)|J(W), W, Z^n) \\
&\quad + H(S_C(W), S_G(W)|J(W), W, X_W^n) + H(Z^n|T(W), X_W^n) \\
&\overset{\text{(q)}}{\geq} \sum_{t=1}^{n} I(X_t; Z^{t-1}, X^{t-1}(W), T(W)) - (H(Z^n) - H(Z^n|T(W))) \\
&\quad + (H(Z^n) - H(Z^n|J(W), W)) - n\delta_n \\
&\overset{\text{(r)}}{\geq} \sum_{t=1}^{n} I(X_t(W); Z^{t-1}, T(W)) - \sum_{t=1}^{n} I(Z_t; Z^{t-1}, T(W)) + I(Z^n; \boldsymbol{J}, W) - n\delta_n \\
&\overset{\text{(s)}}{\geq} \sum_{t=1}^{n} \Big\{ I(X_t(W); U_t) - I(Z_t; U_t) \Big\} + n(R_I - (\delta + \delta_n)) - n\delta_n \\
&\overset{\text{(t)}}{=} n(I(X;U) - I(Z;U) + R_I - \delta - 2\delta_n),
\end{aligned} \tag{4.36}$$

where

(q) follows because (4.25) is used and in the right above equality, the third and the last terms cancel out each other, Fano's inequality is applied for the fifth term, and the sixth term is eliminated,

(r) follows since conditioning reduces entropy and $Z^n - (J(W), W) - \boldsymbol{J} \backslash J(W)$ is applied,

(s) follows due to Lemma 4.2,

(t) holds due to (3.58) and (3.60) in Lemma 3.3.

Dividing both sides of (4.36) by n, from (4.20), it yields that

$$R_L \geq I(X;U) + I(Z;U) + R_I - 2(\delta + \delta_n). \tag{4.37}$$

For the cardinality bound $|\mathcal{U}| \leq |\mathcal{Y}| + 2,$ it can be derived by using the support lemma [19, Lemma 3.4].

Finally, by letting $n \to \infty$ and $\delta \downarrow 0$, we complete the proof of the converse part. □

4.4.2 Achievability Part

Parameter Settings: First, fix the test channel $P_{U|Y}$. Let δ be a small enough positive value and fix a block length n. We set

$$R_I > 0, R_C > 0, \quad (R_I + R_C < I(Z;U)) \tag{4.38}$$

$$\Gamma < R_C, \tag{4.39}$$

$$R_G = I(Z;U) + \Gamma - (R_I + R_C) - \delta, \tag{4.40}$$

$$R_M = I(Y;U) - I(Z;U) + R_I + 2\delta, \tag{4.41}$$

$$R_J = I(Y;U) - I(Z;U) + R_I + R_C + 2\delta, \tag{4.42}$$

$$R_L = I(X;U) - I(Z;U) + R_I + 2\delta, \tag{4.43}$$

where R_M denotes the rate of dummy message shared between the encoder and decoder. We also set $\mathcal{S}_C = [1 : 2^{nR_C}]$, $\mathcal{S}_G = [1 : 2^{nR_G}]$, and $\mathcal{J} = [1 : 2^{nR_J}]$. We define four new sets $\mathcal{S}_\Gamma = [1 : 2^{n\Gamma}]$, $\mathcal{S}_{C\bar{\Gamma}} = [1 : 2^{n(R_C - \Gamma)}]$, $\mathcal{S}_{G\bar{\Gamma}} = [1 : 2^{n(R_G - \Gamma)}]$, and $\mathcal{M} = [1 : 2^{nR_M}]$, representing the sets of shared bits, unshared bits in chosen-secret key, unshared bits in generated-secret key, dummy message, respectively. Without loss of generality, we have that

- There exists one-to-one mapping between l and a pair (m,n), where $l \in \mathcal{S}_C$, $m \in \mathcal{S}_\Gamma$, and $n \in \mathcal{S}_{C\bar{\Gamma}}$.
- There exists one-to-one mapping between p and (q,r), where $p \in \mathcal{S}_G$, $q \in \mathcal{S}_\Gamma$, and $r \in \mathcal{S}_{G\bar{\Gamma}}$.

Codebook Generation: Generate $2^{n(I(Y;U)+\delta)}$ sequences of $u^n(s_1, s_2, m)$, which are i.i.d. from P_U, where $s_1 \in \mathcal{S}_C$, $s_2 \in \mathcal{S}_{G\bar{\Gamma}}$, and $m \in \mathcal{M}$.

Encoding (Enrollment): Note that the encoder knows the user index i beforehand and the chosen-secret key $s_C(i) \in \mathcal{S}_C$ is given and there is a one-to-one mapping between $s_C(i)$ and a pair $(s_{C1}(i), s_{C2}(i))$, where $s_{C1}(i) \in \mathcal{S}_\Gamma$, and $s_{C2}(i) \in \mathcal{S}_{C\bar{\Gamma}}$. The first $n\Gamma$ information bits of $s_C(i)$, which is $s_{C1}(i)$, are shared with the generated-secret key as displayed in Fig. 4.7.

Observing the measurement y_i^n and $s_C(i)$ chosen from ❸, the encoder finds index tuples (s_1, s_2, m) such that $(y_i^n, u^n(s, s_2, m)) \in \mathcal{T}_\varepsilon^{(n)}(YU)$. If there exists multiple tuples satisfying the joint typicality

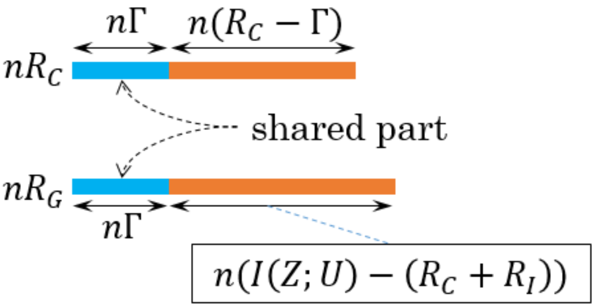

Fig. 4.7 Shared bits; The blue parts describe information bits shared between chosen- and generated-secret keys.

above, it picks one of them at random. Otherwise, error is declared. Let $(s_1(i), s_2(i), m(i))$ denote the tuple chosen for given y_i^n. Then, the encoder generates a secret key and a template as follows:

$$j(i) = (m(i), s_C(i) \oplus s_1(i))), \tag{4.44}$$

$$s_G(i) = (s_{C1}(i), s_2(i)), \tag{4.45}$$

where $\oplus$ denotes the addition modulo M_C. $j(i)$ is stored at location i in ❶, which can be accessed by the decoder, and $s_G(i)$ is saved at position i in ❷.

Decoding (Identification): Seeing z^n, the decoder looks for the index tuple $(s_1, s_2, m(i))$ such that $(z^n, u^n(s_1, s_2, m(i))) \in \mathcal{T}_\varepsilon^{(n)}(ZU)$ for all i with some $s_1 \in \mathcal{S}_C$ and $s_2 \in \mathcal{S}_{G\bar{\Gamma}}$. If such i, s_1, and s_2 are unique, the decoder sets $(\widehat{s_1(w)}, \widehat{s_2(w)}, \widehat{m(w)}) = (s_1, s_2, m(i))$. Otherwise, it declares error. Assume that i, s_2, and s_2 are uniquely found. Then, the decoder outputs the index $\widehat{w} = i$ and the chosen-secret keys as

$$\widehat{s_C(w)} = s_C(\widehat{w}) \oplus s_1(\widehat{w}) \ominus \widehat{s_1(w)}. \tag{4.46}$$

where $s_1(\hat{w}) \oplus s_C(\hat{w})$ is the latter half of the template $j(\hat{w})$ and $\ominus$ denotes the subtraction modulo M_C. After that, the decoder determines the corresponding pair $(\widehat{s_{C1}(w)}, \widehat{s_{C2}(w)})$ from the one-to-one mapping tables, and use $\widehat{s_{C1}(w)}$ to estimate the generated-secret key as $\widehat{s_G(w)} = (\widehat{s_{C1}(w)}, \widehat{s_2(w)})$. Finally, the decoder checks again that whether $\widehat{s_G(w)} = s_G(\widehat{w})$ and $\widehat{s_G(w)} = s_G(\widehat{w})$ or not. If they match, decoding is successful.

Next, we shall check that all the conditions (4.6)-(4.13) in Definition 4.1 satisfy under random codebook $\mathcal{C}_n = \{U^n(s_1, s_2, m) : s_1 \in \mathcal{S}_C, s_2 \in \mathcal{S}_{G\bar{\Gamma}}, m \in \mathcal{M}\}$. We denote the corresponding index tuple of individual i chosen by the encoder for given Y_i^n as $(S_1(i), S_2(i), M(i))$. For simplicity, we denote the sequence $U^n(S_1(i), S_2(i), M(i))$ as U_i^n.

Analysis of Error Probability: For $W = i$, an error event possibly happens at the encoder is

$$\mathcal{E}_1 \; : \{(Y^n, U^n(s_1, s_2, m)) \notin \mathcal{T}_\varepsilon^n(YU) \text{ for all } s_1 \in \mathcal{S}_C,\; s_2 \in \mathcal{S}_{G\bar{\Gamma}},\; m \in \mathcal{M}\},$$

and those at the decoder are:

$\mathcal{E}_2$: $\{(Z^n, U_i^n) \notin \mathcal{T}_\varepsilon^n(ZU)\}$,

$\mathcal{E}_3$: $\{(Z^n, U^n(S_1(i), s_2', M(i)) \in \mathcal{T}_\varepsilon^n(ZU)$ for $\exists s_2' \neq S_2(i)$ $(s_2' \in \mathcal{S}_{G\bar{\Gamma}})\}$,

$\mathcal{E}_4$: $\{(Z^n, U^n(s_1', S_2(i), M(i)) \in \mathcal{T}_\varepsilon^n(ZU)$ for $\exists s_1' \neq S_1(i)$ $(s_1' \in \mathcal{S}_C)\}$,

$\mathcal{E}_5$: $\{(Z^n, U^n(s_1', s_2', M(i)) \in \mathcal{T}_\varepsilon^n(ZU)$ for $\exists s_1' \neq S_1(i)$ $(s_1' \in \mathcal{S}_C)$ and $\exists s_2' \neq S_2(i)$ $(s_2' \in \mathcal{S}_{G\bar{\Gamma}})\}$,

$\mathcal{E}_6$: $\{(Z^n, U^n(s_1', s_2', M(i')) \in \mathcal{T}_\varepsilon^n(ZU)$ for $\exists i' \neq i$ $(i' \in \mathcal{I})$, $s_1' \in \mathcal{S}_C$, and $s_2' \in \mathcal{S}_{G\bar{\Gamma}}\}$.

Note that it suffices to concentrate on assessing the probability of incorrect estimation for the index, chosen- and generated-secret keys at the decoder. If they are correctly estimated, it is guaranteed that the first and second authentications are successful. The error probability of the model can be further analyzed as

$$\begin{aligned}\Pr\left\{\widehat{E(W)} \neq E(W)\right\} &= \Pr\{\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4 \cup \mathcal{E}_5 \cup \mathcal{E}_6\} \\ &\leq \Pr\{\mathcal{E}_1\} + \Pr\{\mathcal{E}_2 | \mathcal{E}_1^c\} + \Pr\{\mathcal{E}_3 \cup \mathcal{E}_4 \cup \mathcal{E}_5 \cup \mathcal{E}_6\}, \qquad (4.47)\end{aligned}$$

where (4.47) follows from the same reason of (a) in (3.35).

By using the covering lemma [19, Lemma 3.3], $\Pr\{\mathcal{E}_1\}$ can be made smaller than δ since $R_C + R_G - \Gamma + R_M > I(Y;U)$. $\Pr\{\mathcal{E}_2 | \mathcal{E}_1^c\}$ can also be made small enough by the Markov lemma [14, Lemma 15.8.1]. The last term vanishes as well by applying the packing lemma [19, Lemma 3.1] since we have $R_I \geq 0, R_C \geq 0, R_G \geq 0$, and $R_I + R_C + R_G - \Gamma < I(Z;U)$. Overall, the error probability can be bounded by

$$\Pr\{\widehat{E(W)} \neq E(W) | W = i\} \leq 3\delta \qquad (4.48)$$

for large enough n.

We first introduce some useful lemmas, which are used in the evaluation of the conditions, and then dive into the core part of the discussion.

Lemma 4.3. *It holds that*

$$H(Y_i^n | S_1(i), S_2(i), M(i), \mathcal{C}_n) \leq n(H(Y|U) + \delta_n), \qquad (4.49)$$

$$H(Y_i^n | X_i^n, S_1(i), S_2(i), M(i), \mathcal{C}_n) \leq n(H(Y|X,U) + \delta_n), \qquad (4.50)$$

where δ_n is a positive value satisfying $\delta_n \downarrow 0$.

Proof: Since the tuple $(S_1(i), S_2(i), M(i))$ determines U_i^n, the above lemma follows by applying Lemma 2.2. A similar proof can be found in Appendix A.2. □

Lemma 4.4. *We have that*

$$H(S_1(i)|\mathcal{C}_n) \geq n(R_C - \delta - \delta_n), \tag{4.51}$$

$$H(S_2(i)|\mathcal{C}_n) \geq n(R_G - \Gamma - \delta - \delta_n), \tag{4.52}$$

$$H(S_1(i), S_2(i)|\mathcal{C}_n) \geq n(R_C + R_G - \delta - \delta_n), \tag{4.53}$$

$$I(S_1(i), S_2(i); M(i)|\mathcal{C}_n) \leq n(\delta + \delta_n), \tag{4.54}$$

$$I(S_1(i); S_2(i), M(i)|\mathcal{C}_n) \leq n(\delta + \delta_n). \tag{4.55}$$

Proof: The proofs are provided in Appendix B.2. □

Analyses of Identification and Chosen-Secrecy Rates: From the parameter settings, (4.7) and (4.8) are trivial.

Analysis of Generated-Secrecy Rate: From the left-hand side of (4.9), we have

$$\frac{1}{n}H(S_G(i)|\mathcal{C}_n) = \frac{1}{n}H(S_{C1}(i), S_2(i)|\mathcal{C}_n) \overset{\text{(a)}}{=} \frac{1}{n}H(S_{C1}(i)|\mathcal{C}_n) + \frac{1}{n}H(S_2(i)|\mathcal{C}_n) \overset{\text{(b)}}{\geq} R_G - \delta - \delta_n, \tag{4.56}$$

where

(a) holds as $S_{C1}(i)$ is independent of $S_2(i)$,

(b) follows because $S_{C1}(i)$ is uniformly distributed on $\mathcal{S}_\Gamma$ and (4.52) is applied.

Analysis of Template Rate: The total required storage rate is

$$\frac{1}{n}\log M_J \leq R_M + R_C = I(Y;U) - I(Z;U) + R_I + 2\delta + R_C = R_J + \delta. \tag{4.57}$$

Analysis of Privacy-leakage Rate: We apply the techniques developed in [21] with some proper extensions. By invoking the same arguments around (3.83)–(3.85) in the analysis of the privacy-leakage rate of the chosen-secret BIS model, we obtain that

$$I(X_i^n; J(i)|\mathcal{C}_n) = I(X_i^n; M(i)|\mathcal{C}_n). \tag{4.58}$$

From (4.58), we have

$$\begin{aligned} I(X_i^n; M(i)|\mathcal{C}_n) &= I(X_i^n; S_1(i), S_2(i), M(i)|\mathcal{C}_n) - I(X_i^n; S_1(i), S_2(i)|M(i), \mathcal{C}_n) \\ &= H(S_1(i), S_2(i), M(i)|\mathcal{C}_n) - H(S_1(i), S_2(i), M(i)|X_i^n, \mathcal{C}_n) - H(S_1(i), S_2(i)|M(i), \mathcal{C}_n) \\ &\quad + H(S_1(i), S_2(i)|M(i), X_i^n, \mathcal{C}_n) \end{aligned}$$

$$\begin{aligned}
&\overset{\text{(c)}}{\leq} n(I(Y;U)+\delta) - H(Y_i^n, S_1(i), S_2(i), M(i)|X_i^n, \mathcal{C}_n) + H(Y_i^n|X_i^n, S_1(i), S_2(i), M(i), \mathcal{C}_n)\\
&\quad - H(S_1(i), S_2(i)|\mathcal{C}_n) + I(S_1(i), S_2(i); M(i)|\mathcal{C}_n) + \delta_n'\\
&\overset{\text{(d)}}{\leq} nI(Y;U) - H(Y_i^n|X_i^n, \mathcal{C}_n) + H(Y_i^n|X_i^n, S_1(i), S_2(i), M(i), \mathcal{C}_n)\\
&\quad - n(R_C + R_G - \delta - \delta_n) + n(\delta + \delta_n) + n(\delta + \delta_n')\\
&\overset{\text{(e)}}{\leq} nI(Y;U) - nH(Y|X) + n(H(Y|X,U) + \delta_n)\\
&\quad - n(I(Z;U) - R_I - 2\delta - 2\delta_n - \delta_n')\\
&= n(I(Y;U) - I(Y;U|X)) - n(I(Z;U) - R_I - 2\delta - 3\delta_n - \delta_n')\\
&= n(H(U) - H(U|Y) - H(U|X) + H(U|Y,X))\\
&\quad - n(I(Z;U) - R_I - 2\delta - 3\delta_n - \delta_n')\\
&\overset{\text{(f)}}{=} n(I(X;U) - I(Z;U) + R_I + 2\delta + 3\delta_n + \delta_n'),
\end{aligned} \tag{4.59}$$

where

(c) follows because a similar argument of the virtual system in Intermediate Steps of the achievability proof in Theorem 3.1 is applied,

(d) follows because conditioning reduces entropy, and (4.53) and (4.54) in Lemma 4.4 are applied,

(e) follows since (4.50) in Lemma 4.3 is applied, (Y_i^n, X_i^n) are independent of $\mathcal{C}_n$, and $R_C + R_G = I(Z;U) - R_I - \delta$,

(f) holds due to the Markov chain $X - Y - U$ (cf. (4.14)) and thus $H(U|Y,X) = H(U|Y)$.

Therefore, from (4.58) and (4.59), it follows that

$$\frac{1}{n} I(X_i^n; J(i)|\mathcal{C}_n) \leq I(X;U) - I(Z;U) + R_I + 3\delta = R_L + \delta \tag{4.60}$$

for large enough n.

Analysis of Information Leakage between Chosen- and Generated-Secret Keys: We have that

$$\begin{aligned}
I(S_C(i); S_G(i)|\mathcal{C}_n) &= I(S_{C1}(i), S_{C2}(i); S_{C1}(i), S_2(i)|\mathcal{C}_n)\\
&= H(S_{C1}(i), S_{C2}(i)|\mathcal{C}_n) - H(S_{C1}(i), S_{C2}(i)|S_{C1}(i), S_2(i), \mathcal{C}_n)\\
&\overset{\text{(g)}}{=} H(S_{C1}(i), S_{C2}(i)|\mathcal{C}_n) - H(S_{C2}(i)|\mathcal{C}_n)\\
&= H(S_{C1}(i)|\mathcal{C}_n) + H(S_{C2}(i)|\mathcal{C}_n) - H(S_{C2}(i)|\mathcal{C}_n)\\
&= n\Gamma,
\end{aligned} \tag{4.61}$$

where (g) follows because $S_{C2}(i)$ is chosen independently of $S_2(i)$. Thus, we obtain that

$$\frac{1}{n} I(S_C(i); S_G(i)|\mathcal{C}_n) \leq \Gamma. \tag{4.62}$$

Analysis of Secrecy-leakage: From the left-hand side of (4.12), it follows that

$$
\begin{aligned}
&I(S_C(i),S_G(i);J(i)|\mathcal{C}_n)\\
&\qquad\overset{\text{(h)}}{=} I(S_C(i),S_2(i);M(i),S_1(i)\oplus S_C(i)|\mathcal{C}_n)\\
&\qquad= H(M(i),S_1(i)\oplus S_C(i)|\mathcal{C}_n) - H(M(i),S_1(i)\oplus S_C(i)|S_C(i),S_2(i),\mathcal{C}_n)\\
&\qquad= H(M(i)|\mathcal{C}_n) + H(S_1(i)\oplus S_C(i)|M(i),\mathcal{C}_n) - H(M(i)|S_C(i),S_2(i),\mathcal{C}_n)\\
&\qquad\quad - H(S_1(i)\oplus S_C(i)|M(i),S_C(i),S_2(i),\mathcal{C}_n)\\
&\qquad\le H(M(i)|\mathcal{C}_n) + nR_C - H(M(i)|S_C(i),S_2(i),\mathcal{C}_n)\\
&\qquad\quad - H(S_1(i)|M(i),S_C(i),S_2(i),\mathcal{C}_n)\\
&\qquad\overset{\text{(i)}}{=} H(M(i)|\mathcal{C}_n) + nR_C - H(M(i)|S_2(i),\mathcal{C}_n) - H(S_1(i)|M(i),S_2(i),\mathcal{C}_n)\\
&\qquad= nR_C - H(S_1(i)|\mathcal{C}_n) + I(S_2(i);M(i)|\mathcal{C}_n) + I(S_1(i);S_2(i),M(i)|\mathcal{C}_n)\\
&\qquad\overset{\text{(j)}}{\le} 2n\delta + 3n\delta_n, \qquad (4.63)
\end{aligned}
$$

where

(h) due to the fact that $S_G(i) = (S_{C2}(i), S_2(i))$ and $S_{C2}(i)$ is the second half of the chosen-secret key $S_C(i)$,

(i) holds since $S_C(i)$ is independent of other RVs,

(j) follows because (4.51), (4.54), and (4.55) in Lemma 4.4 are applied.

Thus, the secrecy-leakage can be bounded as

$$
\frac{1}{n} I(S_C(i),S_G(i);J(i)|\mathcal{C}_n) \le 3\delta \qquad (4.64)
$$

for large enough n.

By applying Lemma 2.3 to all results shown above (i.e., Eqs. (4.48), (4.56), (4.57), (4.60), (4.61) and (4.64)), there exists at least a good codebook satisfying all the conditions in Definition 4.1 for all large enough n. □

4.5 Summary of Results and Discussion

In this chapter, we characterized the capacity region among identification, chosen- and generated-secrecy, template, and privacy-leakage rates for the BIS. The characterizations showed that identification, chosen- and generated-secrecy rates are in a trade-off relation, and by permitting the correlation of the two secrecy keys, a larger sum of these rates was achievable. In addition, larger memory space for the database is required when the sum of identification and chosen-secrecy rates increases. Unlike the template rate, only the identification rate contributes to the minimum required amount of the privacy-leakage rate and the chosen-secrecy rate does not. As special cases, this characterization reduces to the results seen in Chapter 3, and the ones provided in [21].

Actually, the models considered in Chapter 3 can also be applied to two-factor authentication if we consider partitioning the secret key into two parts. This leads to have two new secret keys with smaller sizes and these keys may be used for the first and second rounds in authentications. However, it seems impossible to achieve the secrecy rate that is larger than $I(Z;U)$ since there is no shared information bits from other sources. Another case could be the model considered in [44] where a user enrolls two times in different systems. However, in the settings of [44], the decoder of each system has no permission to access the other systems' database, meaning that it can only estimate one secret key. We need to adapt the settings in [44] by letting the decoder to access all databases so that it can reconstruct two secret keys at once. In this way, the system becomes capable of performing two-factor authentication by using these estimated keys.

Chapter 5

BIS With Both Chosen and Generated Secrecy: Gaussian Source

For DMS settings, the fundamental performances of the BIS are extensively analyzed in the literature [29]–[33],[40] for the VSM and in [21],[81],[85] for the HSM. However, the studies under Gaussian setting are not so many. For example, the optimal trade-off between secrecy and privacy-leakage was clarified in [77] and in order to speed up search complexity, hierarchical identification was taken into account in [74]. A common stand in [77], [74] is that the VSM was assumed.

In this study, we extend the BIS assuming the HSM in Chapter 4 to i.i.d. Gaussian sources and channels. This is motivated by the fact that the signal vectors of bio-data sequences are basically represented by continuous values in real-life applications and most communication links can be modeled as white addictive Gaussian channels. What is more, when the model is switched from the VSM to the HSM, the evaluation becomes more challenging [21], [83],[85] and many existing techniques for deriving the results of the VSM are not directly applicable. Thus, the extension is of both theoretical and practical interest. Our goal is to look for the optimal trade-off of identification, chosen- and generated-secrecy rates under privacy and storage constraints for Gaussian settings. We demonstrate that an idea of converting the system to another one where the data flow of each user is in the same direction, which enables us to characterize the capacity region. More specifically, in establishing the outer bound of the region, the converted system allows us to use the well-known EPI [65] twice in two opposite directions, and its property facilitates the derivation of the inner bound. In [21] and Chapter 3, MGL was applied twice, too, to simplify the rate region of the HSM for binary sources without converting the BIS. That was possible due to the uniformity of the sources, and the backward channel of the enrollment channel is also the binary symmetric channel with the same crossover probability. However, this claim is no longer true in the Gaussian case, so it is necessary to formulate the general behavior of the backward channel. We also provide numerical calculations of three different examples. As a consequence, we may conclude that it is difficult to achieve both high secrecy and small privacy-leakage rates at the same time. To achieve a small privacy-leakage rate, the secrecy rate is scarified somehow. Furthermore, as a by-product of our result, the capacity regions of the BIS analyzed in [21] (the BIS with a single user) is obtained, and as special cases, it can be checked that this characterization reduces to the results given in [76], [77].

This chapter is organized as follow. In Section 5.1, we briefly go through the system model and introduce an idea of converting the system for analysis. The main result and numerical examples are given in Section 5.2 and 5.3, respectively. The proof of the main result is available in Section 5.4 and finally, a short summary of results and discussion follows in Section 5.4.

5.1 System Model and Converted System

In this section, we explain system model analyzed in this chapter and introduce an idea of converting the system.

5.1.1 System Model

In this setting, we analyze the same model argued in Chapter 4 under the situation that the bio-data sequences are generated from i.i.d. Gaussian sources. For $i \in \mathcal{I}$ and $k \in [1:n]$, we assume $X_{ik} \sim \mathcal{N}(0,1)$. Note that Gaussian RV with mean zero and unit variance can be obtained by applying a scaling technique. The enrollment channel $P_{Y|X}$ and the identification channel $P_{Z|X}$ are modeled as follows:

$$Y_{ik} = \rho_1 X_{ik} + N_1, \tag{5.1}$$

$$Z_k = \rho_2 X_{ik} + N_2, \tag{5.2}$$

where $|\rho_1| < 1$, $|\rho_2| < 1$ are the Pearson's correlation coefficients, and $N_1 \sim \mathcal{N}(0, 1-\rho_1^2)$ and $N_2 \sim \mathcal{N}(0, 1-\rho_2^2)$ are i.i.d. Gaussian RVs, independent of each other and bio-data sequences. From (5.2), Y and Z are Gaussian with zero mean and unit variance, and the Markov chain $Y - X - Z$ holds. Then, the PDF corresponding to the tuple (X_i^n, Y_i^n, Z^n) is given by

$$f_{X_i^n Y_i^n Z^n}(x_i^n, y_i^n, z^n) = \prod_{k=1}^{n} f_{XYZ}(x_{ik}, y_{ik}, z_k), \tag{5.3}$$

where for $x, y, z \in \mathbb{R}$,

$$\begin{aligned} f_{XYZ}(x,y,z) &= f_X(x) \cdot f_{Y|X}(y|x) \cdot f_{Z|X}(z|x), && (5.4)\\ &= \frac{1}{\sqrt{(2\pi)^3(1-\rho_1^2)(1-\rho_2^2)}} \exp\left(-\left(\frac{x^2}{2} + \frac{(y-\rho_1 x)^2}{2(1-\rho_1^2)} + \frac{(z-\rho_2 x)^2}{2(1-\rho_2^2)}\right)\right). && (5.5) \end{aligned}$$

The bio-data sequences X_i^n $(i \in \mathcal{I})$ are generated i.i.d. from PDF $f_{X_i^n}$, a marginal PDF of $f_{X_i^n Y_i^n Z^n}$. Like what we have seen in the settings of Section 3.1.2 or Section 4.1 in the previous chapter, the chosen-secret key is chosen uniformly and independently from the set $\mathcal{S}_C$. The operations of encoder and decoder of this chapter are exactly the same as those given in Section 4.1, and therefore the detailed descriptions are omitted.

5.1.2 Converted System

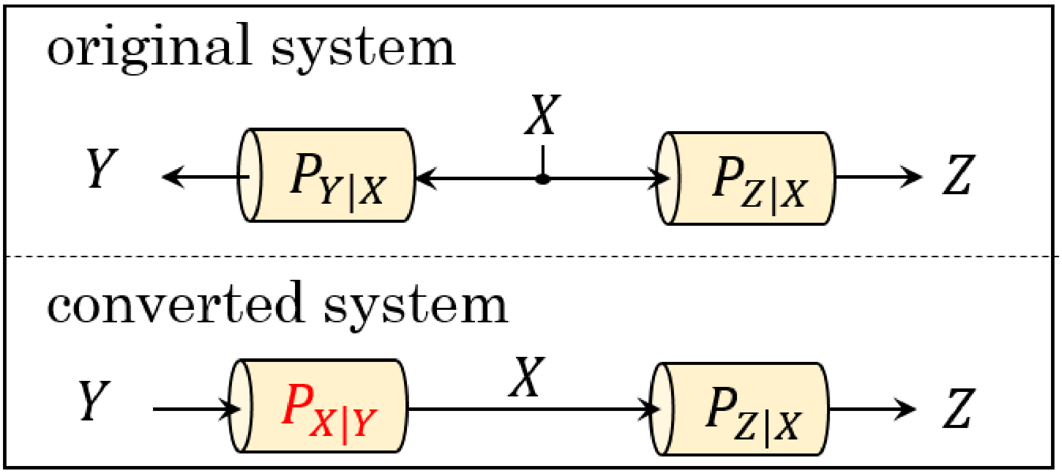

Fig. 5.1 Original and converted systems; The top figure shows the data flow of the bio-data in the original system and the below one is the converted system, where Y becomes virtual input and the data flow is a one-way direction from Y to Z.

The original system, having X as input source and Y,Z as outputs, is illustrated in the top figure in Fig. 5.1. There are two main obstacles toward characterizing the capacity regions directly from this system. (I) In establishing the converse proof, a tight upper bound regarding RV Y for a fixed condition of RV X is needed, but it is laborious to pursue the desired bound since applying EPI to the first relation in (5.2) only produces a lower bound. (II) It seems difficult to prove the achievability part based on generating auxiliary sequences from edge X, e.g., the rate settings. To overcome these bottlenecks, we introduce an idea of converting the original system to a new one in which the data flow of each user is one-way from Y to Z without losing its general properties. The image of this idea is shown in the bottom figure of Fig. 5.1, where Y becomes input virtually. To achieve this objective, knowing the property of the backward channel $P_{X|Y}$, namely, how X correlates to the virtual input Y, is crucial and we explore that in the rest of this section.

Due to the Markov chian $Y-X-Z$, the joint pdf of RVs X, Y, and Z of equation (5.4) can also be expanded in the following form.

$$f_{XYZ}(x,y,z) = f_Y(y)\cdot f_{X|Y}(x|y)\cdot f_{Z|X}(z|x) \tag{5.6}$$

for $x,y,z \in \mathbb{R}$.

Observe that

$$\begin{aligned}\frac{x^2}{2}+\frac{(y-\rho_1 x)^2}{2(1-\rho_1^2)} &= \frac{x^2}{2}+\frac{y^2}{2(1-\rho_1^2)}-\frac{\rho_1 xy}{1-\rho_1^2}+\frac{(\rho_1 x)^2}{2(1-\rho_1^2)} \\ &= \frac{y^2}{2(1-\rho_1^2)}+\frac{x^2}{2(1-\rho_1^2)}-\frac{\rho_1 xy}{1-\rho_1^2} \\ &= \frac{y^2}{2(1-\rho_1^2)}-\frac{(\rho_1 y)^2}{2(1-\rho_1^2)}+\frac{1}{2(1-\rho_1^2)}(x-\rho_1 y)^2 \\ &= \frac{y^2}{2}+\frac{(x-\rho_1 y)^2}{2(1-\rho_1^2)}. \end{aligned} \tag{5.7}$$

Without loss of generality, the equation (5.5) can be rearranged as

$$f_{XYZ}(x,y,z)=\frac{1}{\sqrt{(2\pi)^3(1-\rho_1^2)(1-\rho_2^2)}}\exp\left(-\left(\frac{y^2}{2}+\frac{(x-\rho_1 y)^2}{2(1-\rho_1^2)}+\frac{(z-\rho_2 x)^2}{2(1-\rho_2^2)}\right)\right). \tag{5.8}$$

From (5.6) and (5.8), we may conclude that the following equations hold.

$$X_{ik}=\rho_1 Y_{ik}+N_1', \tag{5.9}$$

$$Z_k=\rho_2 X_{ik}+N_2=\rho_1\rho_2 Y_{ik}+\rho_2 N_1'+N_2 \tag{5.10}$$

with some Gaussian RV $N_1' \sim \mathcal{N}(0,1)$. Equations (5.9) and (5.10) describe the outputs of the backward channel and the compound channel between the backward and identification channels, respectively, for virtual input Y. The above relations play key roles for solving the problem of the HSM, and indeed we use them in many steps during the analysis in this chapter. In [74] and [77], the concept of this transformation is not seen because the enrollment channel does not exist due to the assumption of VSM as mentioned before.

Remark 5.1. *In case there is no operation of scaling, equations* (5.9) *and* (5.10) *are settled as follows. Suppose that $X_{ik} \sim \mathcal{N}(0,\sigma_x^2)$ with $\sigma_x^2 < \infty$, $Y_{ik}=X_{ik}+D_1$, and $Z_k = X_{ik}+D_2$, where $D_1 \sim \mathcal{N}(0,\sigma_1^2)$ and $D_2 \sim \mathcal{N}(0,\sigma_2^2)$ are i.i.d. Gaussian RVs, and independent of each other and other RVs. By applying the arguments around* (5.6)–(5.8)*, we obtain that*

$$X_{ik}=\frac{\sigma_x^2}{\sigma_x^2+\sigma_1^2}Y_{ik}+D_1' \tag{5.11}$$

$$Z_k=X_{ik}+D_2=\frac{\sigma_x^2}{\sigma_x^2+\sigma_1^2}Y_{ik}+D_1'+D_2 \tag{5.12}$$

with some Gaussian RV $D_1' \sim \mathcal{N}(0,\frac{\sigma_x^2\sigma_1^2}{\sigma_x^2+\sigma_1^2})$ is Gaussian and independent of other RVs. The capacity region of the model consider in this study can also be characterized from (5.11) *and* (5.12)*. However, equation developments need more space and do not look so neat. Herein, we pursue our result based on the method that RVs X, Y, and Z are standardized (cf.* (5.9) *and* (5.10)*).*

Now from (5.9) and (5.10), it is not difficult to calculate that

$$I(X;Y) = \frac{1}{2}\ln\left(\frac{1}{1-\rho_1^2}\right), \tag{5.13}$$

$$I(Z;Y) = \frac{1}{2}\ln\left(\frac{1}{1-\rho_1^2\rho_2^2}\right), \tag{5.14}$$

where (5.14) is attained because the variance of the noise term $\rho_2 N_1' + N_2$ in (5.10) is equal to $1-\rho_1^2\rho_2^2$.

5.2 Problem Formulation and Main Results

In this section, we provide the formal definitions of the BIS with both chosen- and generated-secrecy under Gaussian sources, and state the main result.

Definition 5.1. *A tuple $(R_I, R_C, R_G, R_J, R_L)$ is said to be Γ-achievable for a Gaussian source if there exist pairs of encoders and decoders that satisfy all the requirements in Definition 4.1 for any $\delta > 0$, $0 \le \Gamma \le \min\{R_C, R_G\}$, and large enough n. In addition, let $\mathcal{R}^G(\Gamma)$ denote the Γ-capacity region for this case.*

□

The main result of this chapter is given below.

Theorem 5.1. *The Γ-capacity region of the BIS with both chosen and generated secrecy for a Gaussian source is given by*

$$\begin{aligned}\mathcal{R}^G(\Gamma) = \Big\{(R_I, R_C, R_G, R_J, R_L) \ :\ R_I + R_C &\le \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}\right),\\ R_I + R_C + R_G &\le \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}\right) + \Gamma,\\ R_J &\ge \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}{\alpha}\right) + R_C + R_I,\\ R_L &\ge \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}{\alpha\rho_1^2 + 1 - \rho_1^2}\right) + R_I,\\ R_I \ge 0,\ R_C \ge \Gamma,\ &R_G \ge 0 \text{ for some } 0 < \alpha \le 1\Big\}. \end{aligned} \tag{5.15}$$

□

It can be shown that $\mathcal{R}_G$ is convex region. The proof is given in Appendix C.1. The region $\mathcal{R}^G(\Gamma)$ can also be expressed in the form of using auxiliary RV like $\mathcal{R}^D(\Gamma)$. However, the issue is that we can not compute the behavior of the region for this expression due to the unbounded cardinality of the auxiliary RV. Here, instead of using auxiliary RV, e.g., U, we characterize the capacity region with a parameter α, where it lies in the range of $(0,1]$. Since the parameter varies within a limited range, the region $\mathcal{R}^G(\Gamma)$ becomes computable.

Remark 5.2. *If there is no scaling as in* (5.11) *and* (5.12) *in Remark 5.1, the* Γ*-capacity region of the BIS, denoted by* $\mathcal{R}'_G$*, becomes*

$$
\begin{aligned}
\overline{\mathcal{R}^G}(\Gamma) = \Big\{(R_I,R_C,R_G,R_J,R_L) : R_I + R_C &\le \frac{1}{2}\ln\left(\frac{(\sigma_x^2+\sigma_1^2)(\sigma_x^2+\sigma_2^2)}{\alpha\sigma_x^4+\sigma_x^2\sigma_1^2+\sigma_1^2\sigma_2^2+\sigma_2^2\sigma_x^2}\right), \\
R_I + R_C + R_G &\le \frac{1}{2}\ln\left(\frac{(\sigma_x^2+\sigma_1^2)(\sigma_x^2+\sigma_2^2)}{\alpha\sigma_x^4+\sigma_x^2\sigma_1^2+\sigma_1^2\sigma_2^2+\sigma_2^2\sigma_x^2}\right)+\Gamma, \\
R_J &\ge \frac{1}{2}\ln\left(\frac{\alpha\sigma_x^4+\sigma_x^2\sigma_1^2+\sigma_1^2\sigma_2^2+\sigma_2^2\sigma_x^2}{\alpha(\sigma_x^2+\sigma_1^2)(\sigma_x^2+\sigma_2^2)}\right)+R_I+R_C, \\
R_L &\ge \frac{1}{2}\ln\left(\frac{\alpha\sigma_x^4+\sigma_x^2\sigma_1^2+\sigma_1^2\sigma_2^2+\sigma_2^2\sigma_x^2}{(\alpha\sigma_x^2+\sigma_1^2)(\sigma_x^2+\sigma_2^2)}\right)+R_I, \\
R_I \ge 0,\ \ R_C &\ge \Gamma,\ \ R_S \ge 0 \text{ for some } 0 < \alpha \le 1\Big\}. \qquad (5.16)
\end{aligned}
$$

It can be verified that $\mathcal{R}^G(\Gamma)$ *is equivalent to* $\overline{\mathcal{R}^G}(\Gamma)$ *if we set* $\rho_1^2 = \frac{\sigma_x^2}{\sigma_x^2+\sigma_1^2}$ *and* $\rho_2^2 = \frac{\sigma_x^2}{\sigma_x^2+\sigma_2^2}$.

Similar to a conclusion in Section 4.2, the larger sum of identification, generated- and chosen-secrecy rates is obtained due to allowing the correlation of secret keys. The lower bound on the template rate R_J involves both R_I and R_C. This means the minimum required amount of R_J rises in accordance with the increase of R_I and R_C. This is because the number of users is proportional to the increase of storage, and the information related to the chosen-secret key of each user needs to be saved together with the templates in the database so that it can be reconstructed at the decoder. Unlike the template rate R_J, the bound on the privacy-leakage rate R_L only relies on R_I, and this implies the randomness (independence) of the chosen-secret keys make no contribution to the privacy-leakage. Here, we omit the meaning of each rate constraint since it can be explained similarly to that of Theorem 4.1 (cf. Figure 4.2) if one thinks of $I(Z;U) = \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right)$, $I(Y;U) = \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}\right)$, and $I(X;U) = \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha\rho_1^2+1-\rho_1^2}\right)$.

Next, let see how Theorem 5.1 associates with the results in previous studies. When the chosen- and generated-secrecy rates are zero ($R_C = R_G = 0$), and the template and privacy-leakage rates are large enough ($R_J, R_L \to \infty$), the maximum value of the identification rate R_I is $\frac{1}{2}\ln(\frac{1}{1-\rho_1^2\rho_2^2})$. This value is exactly the identification capacity $I(Y;Z)$ (cf. (5.14)) derived in [76], and it is achieved when $\alpha \downarrow 0$. Moreover, when $R_I = R_C = \Gamma = 0$, $R_J \to \infty$, and the enrollment channel is noiseless $(\rho_1 = 1)$, one can see that Theorem 5.1 naturally reduces to the characterizations of [77, Theorem 1]. In slightly different condition, when $R_I = R_G = \Gamma = 0$, $R_J \to \infty$, and the enrollment channel is noiseless $(\rho_1 = 1)$, it can be checked that Theorem 5.1 matches with [77, Theorem 2].

In [21], the capacity regions of the generated- and chosen-secret BIS models with a single user for DMS were characterized. As by-products, when the models studied in [21] is extended to Gaussian sources and channels, the capacity regions of these models are simply special cases of Theorem 5.1. Also, the capacity regions of the generated- and chosen-secret BIS models analyzed in Chapter 3 for Gaussian sources and channels are special cases to this theorem. To explain that, we define four following regions.

$$
\begin{aligned}
\mathcal{R}_1^G = \Big\{(R_G,R_J,R_L)\ :\ & R_G \leq \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right),\\
& R_J \geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}\right),\\
& R_L \geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha\rho_1^2+1-\rho_1^2}\right),\\
& R_G \geq 0 \text{ for some } 0<\alpha\leq 1\Big\}. && (5.17)\\
\mathcal{R}_2^G = \Big\{(R_C,R_J,R_L)\ :\ & R_C \leq \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right),\\
& R_J \geq \frac{1}{2}\ln\left(\frac{1}{\alpha}\right),\\
& R_L \geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha\rho_1^2+1-\rho_1^2}\right),\\
& R_C \geq 0 \text{ for some } 0<\alpha\leq 1\Big\}. && (5.18)\\
\mathcal{R}_3^G = \Big\{(R_I,R_G,R_J,R_L)\ :\ & R_I+R_G \leq \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right),\\
& R_J \geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}\right)+R_I,\\
& R_L \geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha\rho_1^2+1-\rho_1^2}\right)+R_I,\\
& R_I \geq 0, R_G \geq 0 \text{ for some } 0<\alpha\leq 1\Big\}. && (5.19)\\
\mathcal{R}_4^G = \Big\{(R_I,R_C,R_J,R_L)\ :\ & R_I+R_C \leq \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right),\\
& R_J \geq \frac{1}{2}\ln\left(\frac{1}{\alpha}\right),\\
& R_L \geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha\rho_1^2+1-\rho_1^2}\right)+R_I,\\
& R_I \geq 0, R_C \geq 0 \text{ for some } 0<\alpha\leq 1\Big\}. && (5.20)
\end{aligned}
$$

First, we give a remark related to the results of the models considered in [21] for Gaussian sources and channels. By the similar argument of proving Theorem 5.1, the following remark is obtained.

Remark 5.3. *The capacity regions of the generated- and chosen-secret BIS models with single user for a Gaussian source are given by $\mathcal{R}_1^G$ and $\mathcal{R}_2^G$, respectively.*

To show that Remark 5.3 is special cases for Theorem 5.1, we let $\mathcal{R}_{GSS}^{G}$ and $\mathcal{R}_{CSS}^{G}$ denote the special cases of Theorem 5.1 for $\Gamma = R_I = R_C = 0$ and $\Gamma = R_I = R_G = 0$, respectively. Indeed, it can be verified that

$$\mathcal{R}_{GSS}^{G} = \mathcal{R}_{1}^{G}, \tag{5.21}$$

$$\mathcal{R}_{CSS}^{G} = \mathcal{R}_{2}^{G}. \tag{5.22}$$

For the first case, one can easily see that $\mathcal{R}_{GSS}^{G}$ and $\mathcal{R}_{1}^{G}$ is the same, implying that (5.21) holds. For the second case, $\mathcal{R}_{CSS}^{G}$ becomes

$$\begin{aligned}\mathcal{R}_{CSS}^{G} = \Big\{(R_C, R_J, R_L) \ : \ R_C &\le \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}\right),\\ R_J &\ge \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}{\alpha}\right) + R_C,\\ R_L &\ge \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}{\alpha\rho_1^2 + 1 - \rho_1^2}\right),\\ R_C &\ge 0 \text{ for some } 0 < \alpha \le 1\Big\}.\end{aligned} \tag{5.23}$$

The concrete proof of (5.22) is provided in Appendix C.2.

Next, we discuss about the generated- and chosen-secret BIS models with the presence of exponentially many users considered in chapter 3. Note that the difference of these models to the one in this chapter is that the chosen- and generated-secret keys are treated in separate models. The capacity regions of the generated- and chosen-secret BIS models with the presence of exponentially many users were characterized in [85] for DMS and [87] for Gaussian sources. For Gaussian settings, it was demonstrated that the capacity regions of these were given by $\mathcal{R}_{3}^{G}$ (cf. [87, $\mathcal{R}_G$ in Theorem 1]) and $\mathcal{R}_{4}^{G}$ (cf. [87, $\mathcal{R}_C$ in Theorem 1]), respectively. The regions $\mathcal{R}_{3}^{G}$ and $\mathcal{R}_{4}^{G}$ are other special cases of Theorem 5.1. Now let $\mathcal{R}_{GSM}^{G}$ and $\mathcal{R}_{CSM}^{G}$ represent the special cases of the region $\mathcal{R}^{G}(\Gamma)$ where $\Gamma = R_C = 0$ and $\Gamma = R_G = 0$. Then, it is clear that $\mathcal{R}_{GSM}^{G}$ and $\mathcal{R}_{3}^{G}$ coincide, but the second condition of which $\Gamma = R_G = 0$, Theorem 5.1 reduces to the following region.

$$\begin{aligned}\mathcal{R}_{CSM}^{G} = \Big\{(R_I, R_C, R_J, R_L) \ : \ R_I + R_C &\le \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}\right),\\ R_J &\ge \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}{\alpha}\right) + R_I + R_C,\\ R_L &\ge \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}{\alpha\rho_1^2 + 1 - \rho_1^2}\right) + R_I,\\ R_C &\ge 0 \text{ for some } 0 < \alpha \le 1\Big\}.\end{aligned} \tag{5.24}$$

It can be shown that the regions $\mathcal{R}_{CSM}^{G}$ and $\mathcal{R}_{4}^{G}$ are identical by applying the similar approaches given in Appendix C.2.

5.3 Examples and Overviews of the Proof

5.3.1 Some Important Behaviors of the Capacity Region

For given $R_I, R_C \geq 0$, $R_I + R_C \leq \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right)$, and $0 \leq \Gamma \leq R_C$, we define two rate functions

$$R_G^*(R_J) = \max_{(R_I,R_C,R_S,R_J,R_L)\in\mathcal{R}_G} R_G, \tag{5.25}$$

$$R_L^*(R_J) = \min_{(R_I,R_C,R_S,R_J,R_L)\in\mathcal{R}_G} R_L, \tag{5.26}$$

where (5.25) and (5.26) are the maximum secrecy rate and the minimum privacy-leakage rate, respectively, for given $R_J \geq R_I + R_C$. Moreover, we define

$$R_J^\alpha = \frac{1}{2}\ln(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}) + R_I + R_C \tag{5.27}$$

so that we can write

$$R_G^*(R_J^\alpha) = \frac{1}{2}\ln\left(\frac{1-\rho_1^2\rho_2^2/e^{2(R_J^\alpha-R_I-R_C)}}{1-\rho_1^2\rho_2^2}\right) - R_C - R_I + \Gamma, \tag{5.28}$$

$$R_L^*(R_J^\alpha) = \frac{1}{2}\ln\left(\frac{1-\rho_1^2\rho_2^2}{1-\rho_1^2+\rho_1^2(1-\rho_2^2)/e^{2(R_J^\alpha-R_I-R_C)}}\right) + R_I, \tag{5.29}$$

For the sake of succinct discussion, we only concentrate on the condition at which $R_I = 0$ (one user), $R_C = 0$, and $\Gamma = 0$, corresponding to the region $\mathcal{R}_1^G$. For these conditions, (5.28) and (5.29) can be further simplified as

$$R_G^*(R_J^\alpha) = \frac{1}{2}\ln\left(\frac{1-\rho_1^2\rho_2^2/e^{2R_J^\alpha}}{1-\rho_1^2\rho_2^2}\right), \tag{5.30}$$

$$R_L^*(R_J^\alpha) = \frac{1}{2}\ln\left(\frac{1-\rho_1^2\rho_2^2}{1-\rho_1^2+\rho_1^2(1-\rho_2^2)/e^{2R_J^\alpha}}\right). \tag{5.31}$$

We first look over some special points of secrecy and privacy-leakage rates when storage rate becomes extremely low or large. As the template late is large enough, i.e., $R_J \to \infty$, we see that the asymptotic optimal secrecy and privacy-leakage rates

$$\begin{aligned}\lim_{R_J^\alpha\to\infty} R_G^*(R_J^\alpha) &= \lim_{R_J^\alpha\to\infty}\frac{1}{2}\ln\left(\frac{1-\rho_1^2\rho_2^2/e^{2R_J^\alpha}}{1-\rho_1^2\rho_2^2}\right)\\ &= \frac{1}{2}\ln\left(\frac{1}{1-\rho_1^2\rho_2^2}\right)\\ &= I(Y;Z), \end{aligned}\tag{5.32}$$

$$\begin{aligned}
\lim_{R_J^\alpha \to \infty} R_L^*(R_J^\alpha) &= \lim_{R_J^\alpha \to \infty} \frac{1}{2} \ln \left(\frac{1-\rho_1^2\rho_2^2}{1-\rho_1^2+\rho_1^2(1-\rho_2^2)/e^{2R_J^\alpha}} \right) \\
&= \frac{1}{2} \ln \left(\frac{1-\rho_1^2\rho_2^2}{1-\rho_1^2} \right) \\
&= \frac{1}{2} \ln \left(\frac{1}{1-\rho_1^2} \right) - \frac{1}{2} \ln \left(\frac{1}{1-\rho_1^2\rho_2^2} \right) \\
&= I(X;Y) - I(Z;Y). \qquad (5.33)
\end{aligned}$$

The result (5.32) corresponds to the optimal asymptotic secrecy rate [77, Sect. III-B] and in order to achieve this rate, it is required to take the storage rate to infinity and allow to leak the user's privacy up to rate $I(X;Y) - I(Z;Y)$.

In contrast, when $R_J \downarrow 0$, it is evident that R_S and R_L become zero as well, which does not carry much information. However, to investigate the BIS that achieves high secrecy and small privacy-leakage rates in the low storage rate regime, the zero-rate slopes of secrecy and privacy-leakage rates, namely, how fast they converge to zero, are important indicators. In views of (5.30), the first derivative of the generated-secrecy rate can be determined as follows:

$$\begin{aligned}
\frac{dR_S^*(R_J^\alpha)}{dR_J^\alpha} &= \frac{d}{dR_J^\alpha} \left[\frac{1}{2} \ln \left(\frac{1-\rho_1^2\rho_2^2/e^{2R_J^\alpha}}{1-\rho_1^2\rho_2^2} \right) \right] \\
&= \frac{1}{2} \frac{d}{dR_J^\alpha} \left[\ln \left(1-\rho_1^2\rho_2^2/e^{2R_J^\alpha} \right) - \ln \left(1-\rho_1^2\rho_2^2 \right) \right] \\
&= \frac{1}{2} \frac{d}{dR_J^\alpha} \left[\ln \left(1-\rho_1^2\rho_2^2/e^{2R_J^\alpha} \right) - \ln \left(1-\rho_1^2\rho_2^2 \right) \right] \\
&= \frac{1}{2} \left[\frac{\frac{d}{dR_J^\alpha} \left(1-\rho_1^2\rho_2^2/e^{2R_J^\alpha} \right)}{1-\rho_1^2\rho_2^2/e^{2R_J^\alpha}} \right] \\
&= \frac{1}{2} \left[\frac{2\rho_1^2\rho_2^2/e^{2R_J^\alpha}}{1-\rho_1^2\rho_2^2/e^{2R_J^\alpha}} \right] \\
&= \frac{\rho_1^2\rho_2^2/e^{2R_J^\alpha}}{1-\rho_1^2\rho_2^2/e^{2R_J^\alpha}}. \qquad (5.34)
\end{aligned}$$

Likewise, from (5.31), the first derivative of the privacy-leakage rate can be determined as follows:

$$\begin{aligned}
\frac{dR_L^*(R_J^\alpha)}{dR_J^\alpha} &= \frac{d}{dR_J^\alpha} \left[\frac{1}{2} \ln \left(\frac{1-\rho_1^2\rho_2^2}{1-\rho_1^2+\rho_1^2(1-\rho_2^2)/e^{2R_J^\alpha}} \right) \right] \\
&= \frac{1}{2} \frac{d}{dR_J^\alpha} \left[\ln \left(1-\rho_1^2\rho_2^2 \right) - \ln \left(1-\rho_1^2+\rho_1^2(1-\rho_2^2)/e^{2R_J^\alpha} \right) \right] \\
&= \frac{1}{2} \left[\frac{\frac{d}{dR_J^\alpha} \left(1-\rho_1^2+\rho_1^2(1-\rho_2^2)/e^{2R_J^\alpha} \right)}{1-\rho_1^2+\rho_1^2(1-\rho_2^2)/e^{2R_J^\alpha}} \right] \\
&= \frac{\rho_1^2(1-\rho_2^2)/e^{2R_J^\alpha}}{1-\rho_1^2+\rho_1^2(1-\rho_2^2)/e^{2R_J^\alpha}}. \qquad (5.35)
\end{aligned}$$

Table 5.1 The secrecy and privacy-leakage rates when $R_J \to \infty$.

Cases	The optimal secrecy rate			Privacy-leakage rate		
	a)	b)	c)	a)	b)	c)
Ex. 1	0.5	0.63	0.70	0.5	0.87	1.29
Ex. 2	0.5	1.12	1.41	0.5	0.54	0.59
Ex. 3	0.5	0.79	0.87	0.5	0.20	0.13

Table 5.2 The slopes of secrecy and privacy-leakage rates at $R_J \downarrow 0$.

Cases	The slope of secrecy rate			The slope of privacy-leakage rate		
	a)	b)	c)	a)	b)	c)
Ex. 1	1.0	1.40	1.67	0.5	0.7	0.83
Ex. 2	1.0	3.71	6.11	0.5	0.53	0.56
Ex. 3	1.0	2.0	2.33	0.5	0.25	0.17

Therefore, from (5.34) and (5.35), the slopes of secrecy and privacy-leakage rates at $R_J \downarrow 0$ can be determined as follows:

$$\left.\frac{dR_S^*(R_J^\alpha)}{dR_J^\alpha}\right|_{R_J^\alpha=0} = \frac{\rho_1^2\rho_2^2}{1-\rho_1^2\rho_2^2}, \tag{5.36}$$

$$\begin{aligned}\left.\frac{dR_L^*(R_J^\alpha)}{dR_J^\alpha}\right|_{R_J^\alpha=0} &= \frac{\rho_1^2(1-\rho_2^2)}{1-\rho_1^2\rho_2^2} \\ &= \frac{\rho_1^2\rho_2^2}{1-\rho_1^2\rho_2^2}\cdot\frac{1-\rho_2^2}{\rho_2^2},\end{aligned} \tag{5.37}$$

where (5.36) is equal to the signal-to-noise ratio of the compound channel from Y to Z. This value multiplied by the reverse of the signal-to-noise ratio of the channel $P_{Z|X}$ appears in the slope of privacy-leakage rate in (5.37).

5.3.2 Numerical Examples

Next, we give numerical computations of three different examples and take a look into behaviors of the special points.

Ex. 1: a) $\rho_1^2 = 3/4,\ \rho_2^2 = 2/3$, b) $\rho_1^2 = 7/8,\ \rho_2^2 = 2/3$, c) $\rho_1^2 = 15/16,\ \rho_2^2 = 2/3$,

Ex. 2: a) $\rho_1^2 = 3/4,\ \rho_2^2 = 2/3$, b) $\rho_1^2 = 9/10,\ \rho_2^2 = 7/8$, c) $\rho_1^2 = 15/16,\ \rho_2^2 = 11/12$,

Ex. 3: a) $\rho_1^2 = 3/4,\ \rho_2^2 = 2/3$, b) $\rho_1^2 = 3/4,\ \rho_2^2 = 8/9$, c) $\rho_1^2 = 3/4,\ \rho_2^2 = 14/15$.

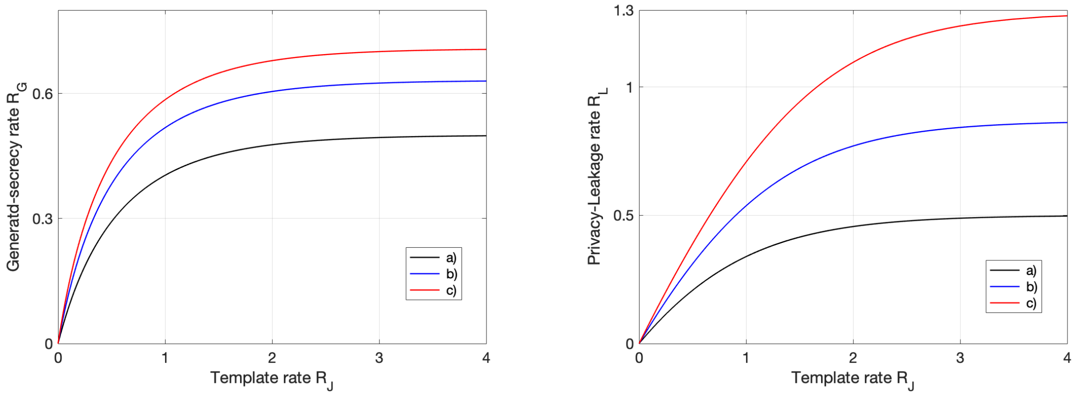

Fig. 5.2 Example 1; The left-hand side figure is the projection of the rate region onto the plane of template and generated-secrecy rates ($R_J R_G$-plane), and the right-hand side one is the projection of the rate region onto the plane of template and privacy-leakage rates ($R_J R_L$-plane).

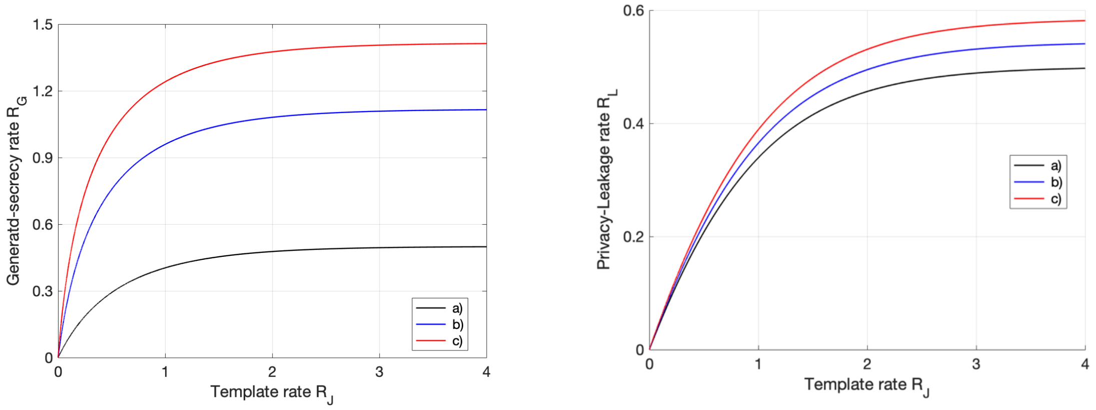

Fig. 5.3 Example 2: The left-hand side figure is the projection of the rate region onto $R_J R_G$-plane, and the right-hand side one is the projection of the rate region onto $R_J R_L$-plane.

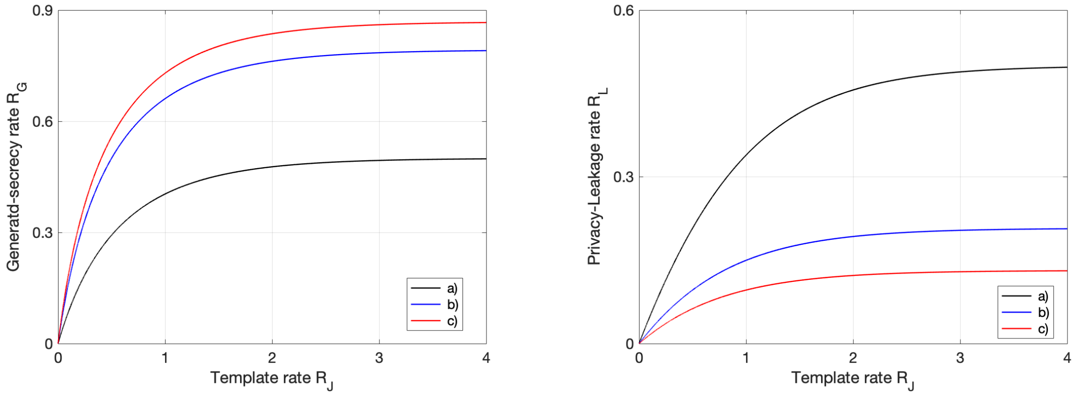

Fig. 5.4 Example 3; The left-hand side figure is the projection of the rate region onto $R_J R_G$-plane, and the right-hand side one is the projection of the rate region onto $R_J R_L$-plane.

Note that as ρ_1^2, ρ_2^2 are large, the noises added to the bio-data sequences at encoder and decoder become small. Example 1 is the case where the noise at encoder is gradually small from a) to c), but the noise added to the bio-data sequences at the decoder stays constant for each round. Example 2 is the case in which the noises added to the bio-data sequences at both encoder and decoder are improved increasingly from a) to c). Example 3 is opposite to Example 1. The calculated results of the generated-secrecy and privacy-leakage rates for these cases are summarized in Table 5.1 and 5.2, and Fig. 5.2–5.4.

It is ideal to keep the privacy-leakage rate small, while produce high secrecy rate, but Example 1 works out in the opposite way (cf. the rows of Ex. 1 in Table I and II, and Fig 5.2), so this is not a preferable choice. Example 2 realizes a high secrecy rate, but the amount of privacy-leakage remains high at some level, too (cf. the rows of Ex. 2 in Table I and II, and Fig. 5.3). On the other hand, in Example 3, the privacy-leakage rate declines, but the secrecy rate becomes small compared to Example 3 (cf. the rows of Ex. 3 in Table I and II, and Fig. 5.4). From these behaviors, we may conclude that it is unmanageable to achieve both a high secrecy rate and small privacy-leakage at the same time. If one aims to achieve a high secrecy rate, it is important to diminish the noises at both encoder and decoder, e.g., deploying quantizers with high quality, but this could result in leaking more user's privacy. In different circumstances, to achieve a small privacy-leakage rate, it is preferable to maintain a certain level of noise at encoder and pay sufficient attention for processing the noise at decoder. In this way, however, the gain of the secrecy rate may be dropped.

5.3.3 Overviews on the Proof of Theorem 4.1

In this section, a brief summary regarding the proof of Theorem 5.1 is described. It consists of two parts: achievability and converse parts. We first demonstrate the converse proof and then show the achievability part. The converse proof follows by applying Fano's inequality [14] and the conditional EPI [7, Lemma II] doubly in two opposite directions. In the achievability part, the modified set (cf. Lemma 2.3), giving the so-called Markov lemma for weak typicality, and Gaussian typicality [14, Section 8.2] help us determine the inner bound of the capacity region. Though a more general version of the Markov lemma for the Gaussian source, including lossy reconstruction, is shown in [54] and [55], we found out that the two properties of the modified set are handy tools for checking all conditions in Definition 4.1, and thus we provide our proof of the achievability based on this modified set. To evaluate the uniformity of secret keys (4.8), privacy-leakage (4.11), and information leakage (4.12), (4.13), we extend Lemma 2.2 to incorporate continuous RVs so that the extended one can be used to derive the upper bounds on conditional differential entropies of jointly typical sequences, appearing in these evaluations. This lemma is used in several analyses of the achievability part.

The detailed proofs are given in the following section. The proof begins with the converse part and follows by the achievability.

5.4 Proof of Theorem 5.1

5.4.1 Converse Part

Here, we also assume that W is uniformly distributed on $\mathcal{I}$, and the conditions in Definition 5.1 are replaced with the average error criterion, e.g.,

$$\Pr\{\widehat{E(W)} \neq E(W)\} \leq \delta, \tag{5.38}$$

$$\frac{1}{n}H(S_G(W)|W) \geq R_G - \delta, \tag{5.39}$$

$$\frac{1}{n}I(X_W^n;J(W)|W) \leq R_L + \delta, \tag{5.40}$$

$$\frac{1}{n}I(S_C(W);S_G(W)|W) \leq \Gamma, \tag{5.41}$$

$$\frac{1}{n}I(S_C(W),S_G(W);J(W)|W) \leq \delta. \tag{5.42}$$

In this part, we show that the capacity region for this case, which contains $\mathcal{R}^G(\Gamma)$, is contained in (5.15). Assume that a rate tuple (R_I,R_C,R_G,R_J,R_L) is achievable.

Analysis of Secrecy Rate: We begin with considering the join entropy of $E(W)$ as

$$\begin{aligned}
H(E(W)) &= H(E(W)|Z^n,\boldsymbol{J}) + I(E(W);Z^n,\boldsymbol{J}) \\
&\overset{\text{(a)}}{=} H(E(W)|\widehat{E(W)},Z^n,\boldsymbol{J}) + I(E(W);Z^n,\boldsymbol{J}) \\
&\overset{\text{(b)}}{=} H(E(W)|\widehat{E(W)}) + I(E(W);Z^n,\boldsymbol{J}) \\
&\overset{\text{(c)}}{\leq} n\delta_n + I(E(W);\boldsymbol{J}) + I(E(W);Z^n|\boldsymbol{J}) \\
&\leq n\delta_n + I(W;\boldsymbol{J}) + I(S_C(W),S_G(W);\boldsymbol{J}|W) + I(E(W);Z^n|\boldsymbol{J},W) \\
&\overset{\text{(d)}}{\leq} n\delta_n + I(S_C(W),S_G(W);J(W)|W) + I(E(W);Z^n|J(W)) \\
&\overset{\text{(e)}}{\leq} n(\delta_n+\delta) + h(Z^n|J(W)) - h(Z^n|J(W),S_G(W),S_C(W),W) \\
&\overset{\text{(f)}}{\leq} n(\delta_n+\delta) + h(Z^n) - h(Z^n|T(W)) \\
&= I(Z^n;T(W)) + n(\delta_n+\delta),
\end{aligned} \tag{5.43}$$

where

(a) holds since $\widehat{E(W)}$ are function of $(Z^n,\boldsymbol{J})$,

(b) follows because conditioning reduces entropy,

(c) follows due to Fano's inequality, where $\delta_n = \frac{1}{n}(1+\ln M_I M_C M_G)$, and $\delta_n \downarrow 0$ as $\delta \downarrow 0$ and $n \to \infty$,

(d) follows because only $J(W)$ is possibly dependent on Z^n and $E(W)$, and W is independent of other RVs

(e) is due to (5.42),

(f) follows because conditioning reduces entropy.

Also, from (5.43), it is trivial that

$$H(W,S_C(W)) \le H(E(W)) \le I(Z^n;T(W)) + n(\delta_n+\delta). \tag{5.44}$$

Analysis of Template Rate: From (4.10), we have that

$$\begin{aligned}
n(R_J+\delta) &\ge \ln M_J \\
&\ge \max_{w\in\mathcal{I}} H(J(w)) \\
&\ge H(J(W)|W) \\
&\overset{\text{(g)}}{=} I(Y_W^n,S_C(W);J(W)|W) \\
&= I(Y_W^n;J(W)|S_C(W),W) + I(S_C(W);J(W)|W) \\
&\overset{\text{(h)}}{=} h(Y_W^n) - h(Y^n|J(W),S_C(W),S_G(W),W) - I(S_G(W);Y_W^n|J(W),S_C(W),W) \\
&\overset{\text{(i)}}{=} I(Y_W^n;T(W)) - H(S_G(W)|J(W),S_C(W),W) \\
&\overset{\text{(j)}}{\ge} I(Y_W^n;T(W)) - H(S_G(W)|S_C(W),W) \\
&\overset{\text{(k)}}{\ge} I(Y_W^n;T(W)) - (I(Z^n;T(W)) - n(R_I+R_C-(3\delta+\delta_n)) \\
&= I(Y_W^n;T(W)) - I(Z^n;T(W)) + n(R_I+R_C-3\delta-\delta_n),
\end{aligned} \tag{5.45}$$

where

(g) holds as $J(W)$ is a function of $(Y_W^n,S_C(W))$,

(h) holds since W and $S_C(W)$ are independent of Y_W^n,

(i) holds because $S_G(W)$ is a function of $(Y_W^n,S_C(W))$,

(j) follows since conditioning reduces entropy,

(k) is due to the following equations (5.46) and (5.47), from (5.43), we have that

$$\begin{aligned}
I(Z^n;T(W)) + n(\delta_n+\delta) &\ge H(S_G(W),S_C(W),W) \\
&= H(W) + H(S_C(W)|W) + H(S_G(W)|S_C(W),W) \\
&\overset{(*)}{=} \ln M_I + \ln M_C + H(S_G(W)|S_C(W),W) \\
&\ge n(R_I+R_C-2\delta) + H(S_G(W)|S_C(W),W),
\end{aligned} \tag{5.46}$$

where $(*)$ holds as W and $S_C(W)$ are independent and W and $S_C(W)$ are uniformly distributed on $\mathcal{I}$ and $\mathcal{S}_C$, respectively, and thus

$$H(S_G(W)|S_C(W),W) \le I(Z^n;T(W)) - n(R_I+R_C-(3\delta+\delta_n)). \tag{5.47}$$

Analysis of Privacy-Leakage Rate: Equation (5.40) can be expanded as

$$\begin{aligned}
n(R_L+\delta) &\geq I(X_W^n;J(W)|W) \\
&= I(X_W^n;J(W),S_G(W),S_C(W),Z^n|W) - I(X_W^n;S_G(W),S_C(W),W,Z^n|J(W),W) \\
&= I(X_W^n;J(W),S_G(W),S_C(W),W) + I(X_W^n;Z^n|J(W),S_G(W),S_C(W),W) \\
&\quad - I(X_W^n;S_G(W),S_C(W),W|J(W),W) - I(X_W^n;Z^n|J(W),S_G(W),S_C(W),W) \\
&= I(X_W^n;J(W),S_G(W),S_C(W),W) - I(X_W^n;S_G(W),S_C(W)|J(W),W) \\
&\geq I(X_W^n;T(W)) - H(S_G(W),S_C(W)|J(W),W) \\
&\overset{\text{(l)}}{\geq} I(X_W^n;T(W)) - H(S_G(W),S_C(W)|W) \\
&\overset{\text{(m)}}{\geq} I(X_W^n;T(W)) - I(Z^n;T(W)) + n(R_I - (\delta_n + 2\delta)),
\end{aligned} \tag{5.48}$$

where

(l) follows since conditioning reduces entropy,

(m) is due to the following relation, which can be obtained by a similar reason seen in equations (5.46) and (5.47),

$$H(S_G(W),S_C(W)|W) \leq I(Z^n;T(W)) - n(R_I - (2\delta + \delta_n)). \tag{5.49}$$

For further evaluations of (5.43)–(5.48), we scrutinize a tight lower bound of $h(Z^n|T(W))$ and a tight upper bound of $h(Y_W^n|T(W))$ under fixed $h(X_W^n|T(W))$ by applying the conditional EPI [7, Lemma II]. It is a key to set

$$\frac{1}{n}h(X_W^n|T(W)) = \frac{1}{2}\ln\left(2\pi e(\alpha\rho_1^2 + 1 - \rho_1^2)\right), \tag{5.50}$$

where $0 < \alpha \leq 1$.

Actually, this is reasonable setting because $\frac{1}{2}\ln(2\pi e) \geq \frac{1}{n}h(X_W^n|T(W)) \geq \frac{1}{2}\ln(2\pi e(1-\rho_1^2))$. The lower bound is obtained from

$$\begin{aligned}
\frac{1}{n}h(X_W^n|T(W)) &= \frac{1}{n}h(X_W^n|J(W),S_G(W),S_C(W),W) \\
&\overset{\text{(l)}}{\geq} \frac{1}{n}h(X_W^n|Y_W^n,T(W)) \\
&= \frac{1}{n}h(X_W^n|Y_W^n,S_C(W),W) \\
&\overset{\text{(m)}}{=} \frac{1}{n}h(X_W^n|Y_W^n) \\
&= h(X|Y) = \frac{1}{2}\ln(2\pi e(1-\rho_1^2)),
\end{aligned} \tag{5.51}$$

where

(n) follows since conditioning reduces entropy, and $(J(W), S_G(W))$ is a function of $(Y_W^n, S_C(W))$,

(o) follows as both $(S_C(W), W)$ are chosen independently of X_W^n and Y_W^n.

First, we have

$$\begin{aligned}\frac{1}{n}I(X_W^n;T(W)) &= h(X) - \frac{1}{n}h(X^n|T(W)) \\ &= \frac{1}{2}\ln(2\pi e) - \frac{1}{2}\ln\left(2\pi e(\alpha\rho_1^2 + 1 - \rho_1^2)\right) \\ &= \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2 + 1 - \rho_1^2}\right). \end{aligned} \tag{5.52}$$

From the direction of X to Z, by applying the conditional EPI [7, Lemma II] to the first equality in (5.10), it follows that

$$\begin{aligned} e^{\frac{2}{n}h(Z^n|T(W))} &\geq e^{\frac{2}{n}h(\rho_2 X_W^n|T(W))} + e^{\frac{2}{n}h(N_2^n|T(W))}, \\ &\overset{\text{(n)}}{=} \rho_2^2 e^{\frac{2}{n}h(X_W^n|T(W))} + e^{\frac{2}{n}h(N_2^n)}, \\ &= \rho_2^2\left(2\pi e(\alpha\rho_1^2 + 1 - \rho_1^2)\right) + 2\pi e(1-\rho_2^2), \\ &= 2\pi e(\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2), \end{aligned} \tag{5.53}$$

where (n) holds as N_2^n is independent of $T(W)$, and as a deduction,

$$\frac{1}{n}h(Z^n|T(W)) \geq \frac{1}{2}\ln(2\pi e(\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2)). \tag{5.54}$$

Then, we can calculate

$$\begin{aligned}\frac{1}{n}I(Z^n;T(W)) &= h(Z) - \frac{1}{n}h(Z^n|T(W)) \\ &\leq \frac{1}{2}\ln(2\pi e) - \frac{1}{2}\ln(2\pi e(\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2)) \\ &= \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}\right). \end{aligned} \tag{5.55}$$

Next, we are interested in finding a tight upper bound on $\frac{1}{n}h(Y^n|T(W))$. In the opposite direction (from X to Y), by again applying the conditional EPI [7, Lemma II] to (5.9), we have that

$$e^{\frac{2}{n}h(X_W^n|T(W))} \geq e^{\frac{2}{n}h(\rho_1 Y_W^n|T(W))} + e^{\frac{2}{n}h(N_1'^n|T(W))}, \tag{5.56}$$

meaning that

$$2\pi e(\alpha\rho_1^2 + 1 - \rho_1^2) \geq \rho_1^2 e^{\frac{2}{n}h(Y_W^n|T(W))} + 2\pi e(1-\rho_1^2), \tag{5.57}$$

where (5.57) follows because $N_1'^n$ is independent of $T(W)$, and thus

$$e^{\frac{2}{n}h(Y_W^n|T(W))} \leq 2\pi e\alpha. \tag{5.58}$$

Hence, it follows that

$$\frac{1}{n}h(Y_W^n|T(W)) \leq \frac{1}{2}\ln(2\pi e\alpha), \tag{5.59}$$

which is not derivable from the first equation in (5.2) of the original system. In (5.59), $\alpha = 0$ is not achievable as the point implies that the entropy of $H(T(W))$ should be infinity (∞), which is impossible for finite sets of $\mathcal{I}$, $\mathcal{S}_C$, $\mathcal{S}_G$, and $\mathcal{J}$. From (5.59), we have that

$$\begin{aligned}\frac{1}{n}I(Y_W^n;T(W)) &= h(Y) - \frac{1}{n}h(Y_W^n|T(W))\\ &\geq \frac{1}{2}\ln(2\pi e) - \frac{1}{2}\ln(2\pi e\alpha) = \frac{1}{2}\ln\left(\frac{1}{\alpha}\right).\end{aligned} \tag{5.60}$$

By similar arguments around (4.29), and from (5.43), (5.44), (5.55), we have that

$$R_I + R_C \leq \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}\right) + 3\delta + \delta_n \tag{5.61}$$

$$R_I + R_C + R_G \leq \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}\right) + \Gamma + 4\delta + \delta_n. \tag{5.62}$$

From (5.52), (5.55) and (5.60), we

$$R_J \geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}{\alpha}\right) + R_I + R_C - (\delta + 4\delta_n), \tag{5.63}$$

$$R_L \geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2 + 1 - \rho_1^2\rho_2^2}{\alpha\rho_1^2 + 1 - \rho_1^2}\right) + R_I - (\delta + 3\delta_n). \tag{5.64}$$

Finally, by letting $\delta \downarrow 0$ and $n \to \infty$ in (5.61)–(5.64), one can see that the capacity region for the case where W is uniformly distributed on $\mathcal{I}$ is contained in the right-hand side of (5.15). This completes the proof of converse part. □

5.4.2 Achievability Part

First, let $0 < \alpha \leq 1$. Fix $\delta > 0$ (small enough positive), and the joint PDF of (U, Y, X, Z) such that the Markov chain $U - Y - X - Z$ holds. Let U be Gaussian with mean zero and variance $1 - \alpha$. Now consider that

$$Y_{ik} = U + \Phi, \tag{5.65}$$

where Φ, independent of U, is Gaussian with mean zero and variance α. From (5.9) and (5.10) of the converted system, we have that

$$X_{ik} = \rho_1 U + \rho_1 \Phi + N_1', \tag{5.66}$$

$$Z_k = \rho_1\rho_2 U + \rho_1\rho_2 \Phi + \rho_2 N_1' + N_2. \tag{5.67}$$

Hence, we readily see that

$$\begin{aligned} I(Y;U) &= \frac{1}{2}\ln\frac{1}{\alpha}, \\ I(X;U) &= \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2+1-\rho_1^2}\right), \\ I(Z;U) &= \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right). \end{aligned} \tag{5.68}$$

Similar to the parameter setting of DMS in the achievability proof of Chapter 4, we pick R_I and R_C as follows:

$$\begin{aligned} &R_I > 0,\ R_C > 0, \quad \left(R_I + R_C < I(Z;U) = \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right)\right), \\ &\Gamma < R_C, \end{aligned}$$

and[1] we set

$$\begin{aligned} R_G &= I(Z;U) + \Gamma - (R_I+R_C) - 2\delta = \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right) + \Gamma - (R_I+R_C) - 2\delta, \\ R_M &= I(Y;U) - I(Z;U) + R_I + 6\delta = \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}\right) + R_I + 6\delta, \\ R_J &= I(Y;U) - I(Z;U) + R_I + R_C + 6\delta = \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}\right) + R_I + R_C + 6\delta, \\ R_L &= I(X;U) - I(Z;U) + R_I + 6\delta = \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha\rho_1^2+1-\rho_1^2}\right) + R_I + 6\delta. \end{aligned} \tag{5.69}$$

We also set $\mathcal{S}_C = [1 : e^{nR_C}]$, $\mathcal{S}_G = [1 : e^{nR_G}]$, $\mathcal{J} = [1 : e^{nR_J}]$, $\mathcal{S}_\Gamma = [1 : e^{n\Gamma}]$, $\mathcal{S}_{C\bar{\Gamma}} = [1 : e^{n(R_C-\Gamma)}]$, $\mathcal{S}_{G\bar{\Gamma}} = [1 : e^{n(R_G-\Gamma)}]$, and $\mathcal{M} = [1 : e^{nR_M}]$.

For forming the codebook, we generate $e^{n(I(Y;U)+\delta)}$ sequences of $u^n(s_1,s_2,m)$ from f_U, in which each symbol of these sequences is i.i.d. Gaussian with mean zero and variance $1-\alpha$, and $s_1 \in \mathcal{S}_C$, $s_2 \in \mathcal{S}_{G\bar{\Gamma}}$, and $m \in \mathcal{M}$.

The encoding and decoding schemes are similar to the ones based on strong typicality, which we have already seen in Section 4.2. However, the differences is that the modified and weak typical sets are used as the encoding and decoding rules, respectively, since the strongly δ-typical set is not

[1] By setting R_C, R_G in this way, it is always guaranteed that $\Gamma \le \min\{R_C, R_G\}$.

applicable in this argument since X,Y,Z, and U are all continuous RVs. Here, we shortly provide the entire description of the scheme as follows:

For encoding, seeing y_i^n and $s_C(i)$, the encoder (enrollment) finds a tuple (s_1,s_2,m) satisfying $(y_i^n,u^n(s_1,s_2,m)) \in \mathcal{B}_\delta^{(n)}(YU)$. If there are multiple such pairs, it selects one at random. Otherwise, error is declared. We denote the chosen tuple as $(s_1(i),s_2(i),m(i))$. Finally, it generates the template $j(i) = (m(i), s_C(i) \oplus s_1(i))$ and the generated-secret key $s_G(i) = (s_{C1}(i), s_2(i))$.

To decode the identified user z^n, a noisy measurement of x_w^n, the decoder looks for the index tuple $(s_1,s_2,m(i))$ such that $(z^n,u^n(s_1,s_2,m(i))) \in \mathcal{A}_\delta^{(n)}(ZU)$ for all i with some $s_1 \in \mathcal{S}_C$ and $s_2 \in \mathcal{S}_{G\bar{\Gamma}}$. If such (i,s_1,s_2) are unique, the decoder sets $(\widehat{w}, \widehat{s_1(w)}, \widehat{s_2(w)}) = (i,s_1,s_2)$. Otherwise, error occurs at the decoder. Assume that (i,s_1,s_2) are uniquely determined. Then, the decoder outputs the index $\widehat{w} = i$, the chosen-secret key $\widehat{s_C(w)} = s_C(\widehat{w}) \oplus s_1(\widehat{w}) \ominus \widehat{s_1(w)}$, and the generated-secrecy key $\widehat{s_G(w)} = (\widehat{s_{C1}(w)}, \widehat{s_2(w)})$. Finally, the decoder checks if $\widehat{s_C(w)} = s_C(\widehat{w})$ and $\widehat{s_G(w)} = s_G(\widehat{w})$, and decoding is successful if these conditions are satisfied.

Next, we check all conditions in Definition 5.1 hold for a random codebook $\mathcal{C}_n = \{U^n(s_1,s_2,m),\ s_1 \in \mathcal{S}_C,\ s_2 \in \mathcal{S}_{G\bar{\Gamma}},\ m \in \mathcal{M}\}$.

For $W = i$, a possible error event at encoder is

$$\mathcal{E}_1 \ :\{(Y_i^n,U^n(s_1,s_2,m)) \notin \mathcal{B}_\delta^{(n)}(YU) \text{ for all } s_1 \in \mathcal{S}_C,\ s_2 \in \mathcal{S}_{G\bar{\Gamma}},\ m \in \mathcal{M}\},$$

and those at the decoder are:

$$\mathcal{E}_2 \ : \{(Z^n,U_i^n) \notin \mathcal{A}_\delta^{(n)}(ZU)\},$$

$$\mathcal{E}_3 \ : \{(Z^n,U^n(S_1(i),s_2',M(i))) \in \mathcal{A}_\delta^{(n)}(ZU) \text{ for } \exists s_2' \neq S_2(i)\ (s_2' \in \mathcal{S}_{G\bar{\Gamma}})\},$$

$$\mathcal{E}_4 \ : \{(Z^n,U^n(s_1',S_2(i),M(i))) \in \mathcal{A}_\delta^{(n)}(ZU) \text{ for } \exists s_1' \neq S_1(i)\ (s_1 \in \mathcal{S}_C)\},$$

$$\mathcal{E}_5 \ : \{(Z^n,U^n(s_1',s_2',M(i))) \in \mathcal{A}_\delta^{(n)}(ZU) \text{ for } \exists s_1' \neq S_1(i)\ (s_1 \in \mathcal{S}_C) \text{ and } \exists s_2' \neq S_2(i)\ (s_2' \in \mathcal{S}_{G\bar{\Gamma}})\},$$

$$\mathcal{E}_6 \ : \{(Z^n,U^n(s_1,s_2,M(i'))) \in \mathcal{A}_\delta^{(n)}(ZU) \text{ for } \exists i' \neq i\ (i' \in \mathcal{I}), \text{ and } s_1 \in \mathcal{S}_C \text{ and } s_2 \in \mathcal{S}_{G\bar{\Gamma}}\}.$$

Note that the authentication process is guaranteed to be successful if the genuine index and secret key of the identified user are correctly estimated at the decoder, indicating that it is sufficient to focus on assessing the probability of incorrect estimation for the pair at the decoder. Then, the error probability can be further evaluated as

$$\begin{aligned}\Pr\left\{\widehat{E(W)} \neq E(W)|W = i\right\} &= \Pr\{\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4 \cup \mathcal{E}_5 \cup \mathcal{E}_6\} \\ &\leq \Pr\{\mathcal{E}_1\} + \Pr\{\mathcal{E}_2 \cap \mathcal{E}_1^c\} + \Pr\{\mathcal{E}_3\} + \Pr\{\mathcal{E}_4\} + \Pr\{\mathcal{E}_5\} + \Pr\{\mathcal{E}_6\}.\end{aligned} \tag{5.70}$$

By applying the similar arguments of [33, Appendix A-B], it can be shown that the entire error probability vanishes. Nonetheless, we provide the details for completeness of the proof. The first term

$\Pr\{\mathcal{E}_1\}$ can be evaluated as

$$\begin{aligned}
\Pr\{\mathcal{E}_1\} &= \Pr\left[\bigcap_{s_1\in\mathcal{S}_C, s_2\in\mathcal{S}_{G\bar{\Gamma}}, m\in\mathcal{M}} (Y_i^n, U^n(s_1,s_2,m)) \notin \mathcal{B}_\delta^{(n)}(YU)\right] \\
&= \prod_{s=1}^{|\mathcal{S}|}\prod_{s_2=1}^{|\mathcal{S}_{G\bar{\Gamma}}|}\prod_{m=1}^{|\mathcal{M}|} \Pr\{(Y_i^n, U^n(s_1,s_2,m)) \notin \mathcal{B}_\delta^{(n)}(YU)\} \\
&\overset{\text{(a)}}{=} \prod_{s_1=1}^{|\mathcal{S}|}\prod_{s_2=1}^{|\mathcal{S}_{G\bar{\Gamma}}|}\prod_{m=1}^{|\mathcal{M}|} \int f_{Y_i^n}(y^n)\Pr\{U^n(s_1,s_2,m) \notin \mathcal{B}_\delta^{(n)}(U|y^n)\}dy^n \\
&= \int f_{Y_i^n}(y^n) \prod_{s_1=1}^{|\mathcal{S}|}\prod_{s_2=1}^{|\mathcal{S}_{G\bar{\Gamma}}|}\prod_{m=1}^{|\mathcal{M}|} \left\{\int_{\mathcal{B}_\delta^{(n)}(U|y^n)^c} f_{U^n}(u^n)du^n\right\}dy^n \\
&= \int f_{Y_i^n}(y^n)\left\{1-\int_{\mathcal{B}_\delta^{(n)}(U|y^n)} f_{U^n}(u^n)du^n\right\}^{|\mathcal{S}_C\times\mathcal{S}_{G\bar{\Gamma}}\times\mathcal{M}|} dy^n \\
&= \int f_{Y_i^n}(y^n)\left(1-\int_{\mathcal{B}_\delta^{(n)}(U|y^n)} f_{U^n}(u^n)du^n\right)^{|\mathcal{S}_C\times\mathcal{S}_{G\bar{\Gamma}}\times\mathcal{M}|} dy^n \\
&\overset{\text{(b)}}{\leq} \int f_{Y_i^n}(y^n)\left(1-e^{-n(I(U;Y)+3\delta)}\int_{u^n\in\mathcal{B}_\delta^{(n)}(U|y^n)} f_{U^n|Y_i^n}(u^n|y^n)du^n\right)^{|\mathcal{S}_C\times\mathcal{S}_{G\bar{\Gamma}}\times\mathcal{M}|} dy^n \\
&\overset{\text{(c)}}{\leq} \int f_{Y_i^n}(y^n)\left(1-\int_{u^n\in\mathcal{B}_\delta^{(n)}(U|y^n)} f_{U^n|Y_i^n}(u^n|y^n)du^n + e^{-|\mathcal{S}_C\times\mathcal{S}_{G\bar{\Gamma}}\times\mathcal{M}|e^{-n(I(U;Y)+3\delta)}}\right)dy^n \\
&\overset{\text{(d)}}{=} \iint_{\mathcal{B}_\delta^{(n)}(U|y^n)^c} f_{U^nY_i^n}(u^n,y^n)du^ndy^n + e^{-e^{n\delta}}\cdot\int_{y_i^n} f_{Y^n}(y^n)dy^n \\
&\overset{\text{(e)}}{\leq} 2\delta
\end{aligned} \tag{5.71}$$

for large enough n, where

(a) is due to the fact that Y_i^n and $U^n(s_1,s_2,m)$ are mutually independent,

(b) is obtained by applying Property 1 of Lemma 2.3, suggesting that if $(y^n,u^n)\in\mathcal{B}_\delta^{(n)}(YU)$, (y^n,u^n) is also a member of $\mathcal{A}_\delta^{(n)}(YU)$, and thus

$$\begin{aligned}
f_{U^n}(u^n) &= f_{U^n|Y_i^n}(u^n|y^n)\frac{f_{U^n}(u^n)\cdot f_{Y_i^n}(y^n)}{f_{U^nY_i^n}(u^n,y^n)} \\
&\geq f_{U^n|Y_i^n}(u^n|y^n)\frac{e^{-n(h(U)+\delta)}\cdot e^{-n(h(Y)+\delta)}}{e^{-n(h(Y,U)+\delta)}} \\
&= f_{U^n|Y_i^n}(u^n|y^n)e^{-n(I(Y;U)+3\delta)},
\end{aligned} \tag{5.72}$$

(c) follows because $(1-\alpha\beta)^m \leq 1-\alpha+e^{-m\beta}$ [20] is applied,

(d) since $\frac{1}{2}\ln|\mathcal{S}_C|+\frac{1}{2}\ln|\mathcal{S}_{G\bar{\Gamma}}|+\frac{1}{2}\ln|\mathcal{M}| = I(Y;U)+4\delta$,

(e) follows by applying Property 2 of Lemma 2.3.

For the second term, it follows that

$$
\begin{aligned}
\Pr\{\mathcal{E}_2 \cap \mathcal{E}_1^c\} &= \Pr\{(Z^n, U_i^n) \notin \mathcal{A}_\delta^{(n)}(ZU) \cap (Y_i^n, U_i^n) \in \mathcal{B}_\delta^{(n)}(YU)\} \\
&\leq \Pr\{(Z^n, Y_i^n, U_i^n) \notin \mathcal{A}_\delta^{(n)}(ZYU) \cap (Y_i^n, U_i^n) \in \mathcal{B}_\delta^{(n)}(YU)\} \\
&= \iint_{\mathcal{B}_\delta^{(n)}(YU)} f_{Y_i^n U_i^n}(y^n, u^n) \cdot \Pr\{Z^n \notin \mathcal{A}_\delta^{(n)}(Z|y^n, u^n) | (Y_i^n, U_i^n) = (y^n, u^n)\} d(y^n, u^n) \\
&\overset{\text{(f)}}{\leq} \delta \iint_{\mathcal{B}_\delta^{(n)}(YU)} f_{Y_i^n U_i^n}(y^n, u^n) d(y^n, u^n) \\
&\leq \delta,
\end{aligned} \tag{5.73}
$$

where (f) follows from the definition of the modified δ-typical set (cf. Lemma 2.3) due to the Markov chain $Z - Y - U$.

To bound the third term $\Pr\{\mathcal{E}_3\}$, we apply [19, Lemma 11.1], which guarantees that $\Pr\{\mathcal{E}_3\} \leq \Pr\{(Z^n, U^n(1, s_2', 1)) \in \mathcal{A}_\delta^{(n)}(ZU)$ for some $s_2' \neq S_2(i)$, $s_2' \in \mathcal{S}_{G\bar{\Gamma}}\}$. Since $U^n(1, s_2', 1)$ is independent of Z^n and both have i.i.d. structure, we have that

$$
\Pr\{\mathcal{E}_3\} \leq \sum_{s'=1}^{|\mathcal{S}_{G\bar{\Gamma}}|} \Pr\{(Z^n, U^n(1, s_2', 1)) \in \mathcal{A}_\delta^{(n)}(ZU)\} \leq \sum_{s_2'=1}^{|\mathcal{S}_{G\bar{\Gamma}}|} e^{-n(I(Z;U)-\delta)} = e^{-n(R_I + R_C + \delta)}. \tag{5.74}
$$

For $\Pr\{\mathcal{E}_4\}$–$\Pr\{\mathcal{E}_6\}$, they can be made negligible by similar techniques seen in (5.74). Consequently,

$$
\Pr\{\widehat{E(W)} \neq E(W) | W = i\} \leq 7\delta \tag{5.75}
$$

for large enough n.

The evaluations of (4.7)–(4.13) for Gaussian source can be checked by the same arguments as in the DMS setting. As we have seen, the discussions were build up based on the support of Lemma 4.3. To the end of this part, we provide an extended version of Lemma 4.3, associating continuous RVs. Since U_i^n can be determined uniquely by the tuple $S_1(i), S_2(i), M(i)$ for a given codebook $\mathcal{C}_n$, this lemma also become an extension of Lemma 2.2. Likewise the proof of achievability proof of Chapter 4, this lemma is a fundamental tool for deriving the bounds of generated-secrecy rate (4.9), privacy-leakage (4.11), and information leakage (4.12) and (4.13) for Gaussian sources.

Lemma 5.1. *For continuous RVs (X, Y, U), if (X_i^n, Y_i^n, U_i^n) are jointly typical with high probability, it holds that*

$$
\frac{1}{n} h(Y_i^n | S_1(i), S_2(i), M(i), \mathcal{C}_n) \leq h(Y|U) + \delta_n, \tag{5.76}
$$

$$
\frac{1}{n} h(Y_i^n | X_i^n, S_1(i), S_2(i), M(i), \mathcal{C}_n) \leq h(Y|X, U) + \delta_n. \tag{5.77}
$$

Proof: The connection between the modified δ-typical set $\mathcal{B}_\delta^{(n)}(\cdot)$ and the weakly δ-typical set $\mathcal{A}_\delta^{(n)}(\cdot)$ is useful for proving the above lemma. We first prove (5.76).

Define an RV T as follows:

$$T = \begin{cases} 1 \text{ if } (Y_i^n, U_i^n) \in \mathcal{B}_\delta^{(n)}(YU), \\ 0 \text{ otherwise.} \end{cases} \tag{5.78}$$

In the analysis of error probability, we have already demonstrated that $P_T(0) \leq 2\delta$, or $(Y_i^n, U_i^n) \in \mathcal{B}_\delta^{(n)}(YU)$ with high probability. From the left-hand side of (5.76),

$$\begin{aligned}
h(Y_i^n|S_1(i),S_2(i),M(i),\mathcal{C}_n) &\overset{\text{(h)}}{=} h(Y_i^n|U_i^n,S_1(i),S_2(i),M(i),\mathcal{C}_n) \\
&\overset{\text{(i)}}{\leq} h(Y_i^n|U_i^n) \leq h(Y_i^n,T|U_i^n) \\
&\leq H(T) + h(Y_i^n|U_i^n,T) \\
&\leq 1 + P_T(0)h(Y_i^n|U_i^n,T=0) + P_T(1)h(Y_i^n|U_i^n,T=1) \\
&\overset{\text{(j)}}{\leq} n\varepsilon_n + h(Y_i^n|U_i^n,T=1) \\
&= n\varepsilon_n + \int h(Y_i^n|U_i^n = u^n, T=1)dF(u^n) \\
&= n\varepsilon_n + \int \int_{\mathcal{B}_\delta^{(n)}(Y|u^n)} P_{Y_i^n|U_i^n,T}(y^n|u^n,1) \ln \frac{1}{P_{Y_i^n|U_i^n,T}(y^n|u^n,1)} dy^n dF(u^n) \\
&\overset{\text{(k)}}{\leq} n\varepsilon_n + \int \ln\left(\int_{\mathcal{B}_\delta^{(n)}(Y|u^n)} P_{Y_i^n|U_i^n,T}(y^n|u^n,1) \frac{1}{P_{Y_i^n|U_i^n,T}(y^n|u^n,1)} dy^n \right) dF(u^n) \\
&= n\varepsilon_n + \int \ln\left(\int_{\mathcal{B}_\delta^{(n)}(Y|u^n)} dy^n \right) dF(u^n) \\
&= n\varepsilon_n + \int \ln\left(\mathrm{Vol}\left(\mathcal{B}_\delta^{(n)}(Y|u^n)\right)\right) dF(u^n) \\
&\overset{\text{(l)}}{\leq} n\varepsilon_n + n(h(Y|U)+\delta)) \int dF(u^n) \\
&\leq n(h(Y|U)+\delta+\varepsilon_n),
\end{aligned} \tag{5.79}$$

where

(h) follows as $(J(i),S(i))$ determines U_i^n,

(i) follows because conditioning reduces entropy,

(j) follows as $h(Y_i^n|U_i^n,T=0) \leq h(Y_i^n) = \frac{n}{2}\log(2\pi e)$, and we define $\varepsilon_n = \frac{1}{n} + \delta \log(2\pi e)$,

(k) follows by applying Jensen's inequality,

(l) follows since $\mathrm{Vol}\left(\mathcal{B}_\delta^{(n)}(Y|u^n)\right) \leq \mathrm{Vol}\left(\mathcal{A}_\delta^{(n)}(Y|u^n)\right) \leq e^{n(h(Y|U)+\delta))}$ (cf. [14]).

Therefore, from (5.79), we obtain that

$$\frac{1}{n}h(Y_i^n|S_1(i),S_2(i),M(i),\mathcal{C}_n) \leq h(Y|U)+\delta_n, \tag{5.80}$$

where $\delta_n = 2\delta + \varepsilon_n$ and $\delta_n \downarrow 0$ as $n \to \infty$ and $\delta \downarrow 0$.

Next, we briefly summarize how to show (5.77). The left-hand side of (5.77) can be developed as $h(Y_i^n|X_i^n,S_1(i),S_2(i),M(i),\mathcal{C}_n) = h(Y_i^n|X_i^n,U_i^n,S_1(i),S_2(i),M(i),\mathcal{C}_n) \leq h(Y_i^n|X_i^n,U_i^n,\mathcal{C}_n)$, where the first equality and second inequality follow due to the same reasons of (h) and (i) in (5.79), respectively. By applying the definition of the modified typical set [33, Appendix A-A], it can be concluded that $\Pr\{(X_i^n,Y_i^n,U_i^n) \in \mathcal{A}_\delta^{(n)}(XYU)\} \to 1$ as $n \to \infty$ (cf. (5.73)) due to the Markov chain $X-Y-U$ and $(Y_i^n,U_i^n) \in \mathcal{B}_\delta^{(n)}(YU)$ with high probability. This implies $\Pr\{(X_i^n,U_i^n) \in \mathcal{A}_\delta^{(n)}(XU)\} \to 1$ and $\Pr\{Y_i^n \in \mathcal{A}_\delta^{(n)}(Y|x^n,u^n)|(X_i^n,U_i^n)=(x^n,u^n)\} \to 1$ as $n \to \infty$ as well. Based on this observation, the rest of proof for (5.77) can be done similarly by the arguments seen in [41, Appendix C], and therefore the details are omitted.

5.5 Summary of Results and Discussion

We extended to the system considered in Chapter 4 to Gaussian sources and channels and characterized the capacity region among identification, generated- and chosen-secrecy, storage, and privacy-leakage rates for the BIS under this setting. This was motivated by the truth that the signal vectors of bio-data sequences are represented by continuous values and transmission channels can be modeled as additive white Gaussian channels. Therefore, considering BIS with Gaussian sources and channels brings a step closer to real applications. We showed that an idea for deriving the capacity region is to convert the system to one where the data flow of each user are in one-way direction. Moreover, numerical computations of three different examples for the capacity region were provided, and from these results, it appeared that achieving both high secrecy and small privacy-leakage rates simultaneously is unlikely manageable.

In this chapter, we did not discuss how the models with Gaussian sources and channels can be applied to real-life BIS. The point was mentioned as a note in [77]. Here, we also leave such discussion for future studies. Also, an investigation to characterize the capacity regions of the BIS for Gaussian vector sources and channels is of sufficient interest.

Chapter 6

Conclusions and Future Directions

6.1 Conclusions

Chapter 3 dealt with two different models: BIS supporting authentication and BIS with both chosen and generated secrecy. The fundamental limits of the BIS supporting authentication (generated- and chosen-secret BIS models) were discussed in Chapter 3. We demonstrated that there are two different ways to express the capacity regions for these models. An expression uses a single auxiliary RV and the other requires two auxiliary RVs, but the two regions are technically identical. We provided the proofs of our main results based on the one employing two auxiliary RVs. We showed that a combination use of random coding and binning is optimal scheme for proving the achievability. As a result, as mentioned in [33], identification and secrecy rates are in a trade-off relation. The minimum values of the template and privacy-leakage rates increase when the identification rate rises, and like an observation in [21], the minimum amount of the template rate is always greater than the bound of the privacy-leakage rate.

Chapter 4 studied the BIS model with both chosen and generated secrecy for DMS. In the analysis, we allowed the two secret keys to be correlated at some level, and this led to obtain a greater sum of identification, chosen- and generated-secrecy rates compared to the result derived in [86]. In addition, the template rate involves both identification and chosen-secrecy rates, but it is not affected by the generated-secrecy rate. Unlike the template rate, the privacy-leakage rate varies in accordance with the change of identification rate, but it has nothing to do with the chosen-secrecy rate.

In Chapter 5, we extended the model considered in Chapter 4 to Gaussian sources and channels. We gave a complete characterization of the capacity region of the model for Gaussian settings. We showed that an idea of deriving the capacity region is to convert the BIS to another one where the data flow of each user is in the same direction. Moreover, we provided numerical computations of three different examples for the derived region, and as a consequence, it seems hard to achieve both high secrecy and small privacy-leakage rates at the same time. When we manage to obtain small privacy-leakage rate, the gain of the secrecy rate is also declined.

6.2 Future Directions

In this section, we mention some possible future directions of this work. An intuitive extension is to build practical codes that provide as good performance as the derived theoretical bounds. Recently, in the BIS with a single user, code constructions for secrecy, template, and privacy-leakage rates were considered in [22] with the use of linear codes and nested polar codes. It was improved in [24] by deploying randomized nested polar subcodes, a combined usage of a polar code as a vector quantizer and a polar subcode as error correcting code. Another approach can be found in [38] by using nested tailbiting convolutional codes. However, the gap between achievable point by this method and the theoretical bound is still quite big. Moreover, there has not yet been any studies considering code constructions for the BIS with multiple users. In this case, one has to take care of identification rate, and hence the problem may become more complicated.

Secrecy amplification is another natural extension. As seen in the definitions of each chapter, we solely focused on deriving the results under the weak secrecy criterion in terms of secrecy-leakage. In the BIS with a single user, it has been shown that it is possible to make the secrecy-leakage negligible under the strong secrecy criterion [11], [22], and [23]. More specifically, in [22] and [23], the capacity regions are derived via the technique of output statistic of random binning [88], [51], or resolvability [28] combined with likelihood encoder [67]. A distinct technique can be seen found in [11] based on source polarization of polar codes [4], [12]. These methods are promising tools for analyzing the capacity region of the BIS with multiple users under the strong secrecy criterion, too. In fact, when the criterion is switched from weak to strong, we have to check only the achievability proof since fundamentally the outer bound established under the weak secrecy criterion results in an outer bound for a more rigorous one, e.g., the strong secrecy criterion. In the achievability proof, it suffices to concentrate only on the user with the worst performance due to the fact that the prior distribution of the identified user is assumed to be unknown. In this fashion, the analysis can be proved similarly as the proof of the BIS with single user, and thus the techniques mentioned above is possibly applicable. Moreover, establishing a technique, which has few results compared to DMS settings, to investigate the BIS for Gaussian sources under the strong secrecy criterion is an interesting and challenging open problem.

Finally, for the sake of simplicity in the analysis, we assume that the bio-data sequences are generated from i.i.d. sources, but this assumption is still far from a realistic model. A good example for this is fingerprints. There are similarities in the patterns of fingerprint for a user, and thus adjacent elements in a quantized vector of the bio-data sequences are possibly correlated. To get closer to practical circumstance, it is important to analyze the fundamental trade-off of the models for non-i.i.d. sources, e.g., the Markov source. Furthermore, all results in this thesis were derived under the condition that the block-length trends to infinite (asymptotic arguments). The discussion in the finite block-length regime still remains as an attractive and important topic.

References

[1] Wikipedia contributors, “Biometrics”, https://en.wikipedia.org/wiki/Biometrics

[2] R. Ahlswede and I. Csiszár, “Common randomness in information theory and cryptography - part I: Secret sharing,” *IEEE Trans. Inf. Theory*, vol. 39, pp. 1121–1132, Jul. 1993

[3] A. Antonelli, R. Cappelli, D. Maio, and D. Maltoni, “Fake finger detection by skin distortion analysis,” *IEEE Trans. Inf. Forensics and Security*, vol. 1, no. 3, pp. 360-373, Sep. 2006.

[4] E. Arikan, “Source polarization,” *in Proc. IEEE Int. Symp. Inf. Theory*, Austin, Texas, U.S.A., pp. 899–903, Jun. 2010.

[5] S. Arimoto, “On the converse to the coding theorem for discrete memoryless channels (Corresp.),” *IEEE Trans. Inf. Theory*, vol. 19, no. 3, pp. 357—359, May 1973.

[6] T. Berger, “Multiterminal source coding,” *the Information Theory Approach to Communications*, vol. 229 of CISM Courses and Lectures, pp. 171-231, Springer-Verlag, 1978.

[7] P. Bergmans, “A simple converse for broadcast channels with additive white Gaussian noise (Corresp.),” *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 279–280, Mar. 1974.

[8] S. Bharadwaj, M. Vatsa, and R. Singh, “Biometric quality: A review of fingerprint, iris, and face,” *EURASIP Journal on Image and Video Processing*, vol. 34, Jul. 2014.

[9] N. M. Blachman, “The convolution inequality for entropy powers,” *IEEE Trans. Inf. Theory*, vol. 11, no. 2, pp. 267–271, Apr. 1965.

[10] M. Bloch and J. Barros, *Physical-Layer Security*, Cambridge, U.K.: Cambridge Univ. Press, 2011.

[11] R. Chou, “Biometric systems with multiuser access structures,” *in Proc. IEEE Int. Symp. Inf. Theory*, Paris, France, Jul. 2019.

[12] R. Chou, M. Bloch, and E. Abbe, “Polar coding for secret-key generation,” *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6213–6237, Nov. 2015

[13] R. Clarke, “Human identification in information systems: Management challenges and public policy issues,” *Information Technology & People*, vol. 7, no. 4, pp. 6–37, 1994.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd ed.*, John Wiley & Sons, New Jersy, 2006.

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

[16] 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.

[17] A. Dembo, T. M. Cover, and J. A. Thomas, “Information theoretic inequalities,” *IEEE Trans. Inf. Theory*, vol. 37, no. 6, pp. 1501–1518, Nov. 1991.

[18] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, “Fuzzy extractors: How to generate strong keys from biometrics and other noisy data”, SIAM J. Comput., vol. 38, no. 1, pp. 97–139, Jan. 2008.

[19] A. El Gamal and Y.-H. Kim, *Network Information Theory*, Cambridge, U.K.: Cambridge Univ. Press, 2011.

[20] R. G. Gallager, *Information Theory and Reliable Communication*, New York: Wiley, 1968.

[21] O. Günlü and G. Kramer, “Privacy, secrecy, and storage with multiple noisy measurements of identifiers,” *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 11, pp. 2872–2883, Nov. 2018.

[22] O. Günlü, O. Iscan, V. Sidorenko, and G. Kramer, “Code constructions for physical unclonable functions and biometric secrecy systems,” *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 11, pp. 2848–2858, Nov. 2019.

[23] O. Günlü, R. F. Schaefer, and G. Kramer, “Private authentication with physical identifiers through broadcast channel measurements,” *in Proc. IEEE Inf. Theory Workshop*, Visby, Sweden, Aug. 2019.

[24] O. Günlü, P. Trifonov, M. Kim, R. F. Schaefer, and V. Sidorenko, “Randomized nested polar subcode constructions for privacy, secrecy, and storage,” *ISITA2020*, Hawaii, USA, pp. 475–479, Oct. 2020.

[25] D. Guo, S. Shamai (Shitz), and S. Verdú, “Proof of entropy power inequalities via MMSE,” *in Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, USA, Jul. 2006, pp. 1011–1015.

[26] R. V. L. Hartley, “Transmission of information,” *The Bell System Technical Journal*, vol. 7, no. 3, pp. 535–563, Jul. 1928.

[27] T.-S. Han, *Information-Spectrum Methods in Information Theory*, Berlin, Germany: Springer-Verlag, 2003.

[28] J. Hou and G. Kramer, “Informational divergence approximations to product distributions,” *in Canadian Workshop Inf. Theory*, Toronto, Canada, pp. 76–81, Jun. 2013.

[29] T. Ignatenko and F.M.J. Willems, “Biometric systems: Privacy and secrecy aspects,” *IEEE Trans. Inf. Forensics Security,* vol. 4, no. 4, pp.956–973, Dec. 2009.

[30] T. Ignatenko, “Secret-key rates and privacy leakage in biometric systems,” Ph.D. dissertation, Technical University of Eindhoven, Eindhoven, The Netherlands, 2009.

[31] T. Ignatenko and F.M.J. Willems, “Biometric security from an information-theoretical perspective,” *Foundations Trends in Communications and Information Theory*, vol. 7, no. 2-3, pp. 135–216, 2010.

[32] T. Ignatenko and F.M.J. Willems, “Fundamental limits for biometric identification with a database containing protected templates,” *in Proc. 2018 Int. Symp. Inf. Theory and Its Appl.*, Taichung, Taiwan, pp.54–59, Oct. 2010.

[33] T. Ignatenko and F.M.J. Willems, “Fundamental limits for privacy-preserving biometric identification systems that support authentication,” *IEEE Trans. Inf. Theory,* vol. 61, no. 10, pp.5583–5594, Oct. 2015.

[34] A. K. Jain, L. Hong, and S. Pankanti, “Biometric identification,” *Communications of the ACM*, vol. 43, no. 2, pp. 90–98, Feb. 2000.

[35] A. K. Jain, K. Nandakumar, and A. Nagar, “Biometric template security,” *EURASIP Journal on Advances in Signal Processing*, 2008.

[36] A. K. Jain, P. Flynn, and A. Ross, *Handbook of Biometrics*, New York, Springer-Verlag, 2009.

[37] A. K. Jain, R. M. Bolle, and S. Pankanti, *Biometrics: Personal Identification in a Networked Society*, New York, Springer-Verlag, 2006.

[38] T. Jerkovits, O.Günlü, V. Sidorenko, and G. Kramer “Nested tailbiting convolutional codes for secrecy, privacy, and storage,” *arXiv:2004.13095*, Apr. 2020.

[39] A. Juels and M. Wattenberg, “A fuzzy commitment scheme,” in Proc. 6th ACM Conf. Comput. Commun. Security, pp. 28–36, Nov. 1999.

[40] K. Kittichokechai and G. Caire, “Secret key-based identification and authentication with a privacy constraint,” *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6189–6203, Nov. 2016.

[41] K. Kittichokechai, T. J. Oechtering, M. Skoglund, and Y.-K. Chia, “Secure source coding with action-dependent side information,” *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6444–6464, Dec. 2015.

[42] M. Koide and H. Yamamoto, “Coding theorems for biometric systems,” *in Proc. IEEE Int. Symp. Inf. Theory*, Texas, USA, pp. 2647–2651, Jun. 2010.

[43] 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.

[44] L. Lai, S.-W. Ho, and H. V. Poor, “Privacy-security trade-offs in biometric security systems–part II: Multiple use case,” *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 140–151, Mar. 2011.

[45] Y. Liang, H. V. Poor, and S. Shamai, “Information theoretic security,” *Found. and Trends in Commun. and Inf. Theory*, vol. 5, no. 4–5, pp. 355–580, 2008.

[46] E. C. Lee, K. R. Park, and J. Kim, “Fake iris detection by using purkinje image,” *in Proc. of Int. Conf. on Advances in Biometrics*, vol. 3832 of Lecture Notes in Computer Science, pp. 397–403, Hong Kong, Jan. 2006.

[47] R. Luis-Garcia, C. Alberola-Lopez, O. Aghzout, and J. Ruiz-Alzola, “Biometric identification systems,” *Signal Processing*, vol. 83, no. 12, pp. 2539–2557, Dec. 2003.

[48] D. Maltoni, D. Maio, A. K. Jain, S. Prabhakar, *Handbook of Fingerprint Recognition*, New York, Springer-Verlag, 2003.

[49] U. M. Maurer, “Secret key agreement by public discussion from common information,” *IEEE Trans. Inf. Theory*, vol. 39, no. 03, pp. 733–742, May 1993.

[50] U. Maurer, “Information-theoretic cryptography,” *in Advances in Cryptology–CRYPTO’99 (Lecture Notes in Computer Science)*, Berlin, Germany: Springer-Verlag, vol. 1666, pp. 47–64, 1999.

[51] M. Nafea and A. Yener, “A new wiretap channel model and its strong secrecy capacity,” *IEEE Trans. Inf. Theory*, vol. 64, no. 03, pp. 2077–2092, Mar. 2014.

[52] S. Prabhakar, S. Pankanti, and A. K. Jain, “Biometric recognition: Security and privacy concerns,” *IEEE SECURITY & PRIVACY*, pp. 33–42 2008.

[53] R. Pappu, “Physical one-way functions,” Ph.D. dissertation, M.I.T., Cambridge, MA, Oct. 2001.

[54] Y. Oohama, “Gaussian multiterminal source coding,” *IEEE Trans. Inf. Theory*,vol. 43, no. 6, pp. 1912–1923, Nov. 1997.

[55] Y. Oohama, “The rate-distortion function for the quadratic gaussian CEO problem,” *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1057–1070, May 1998.

[56] J. A. O'Sullivan and N. A. Schmid, "Large deviations performance analysis for biometrics recognition," *in Proc. 40th Annual Allerton Conf. on Communication, Control, and Computing*, Allerton House, IL, USA, Oct. 2002.

[57] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM Systems Journal*, vol. 40, no. 3, pp. 614–634, 2001.

[58] N. k. Ratha, S. Chikkerur, J. Connell, R. Bolle, "Generating cancelable fingerprint templates," *IEEE Trans. Pattern Anal. Machine Intell*, vol. 29, pp. 561–572, 2007.

[59] A. Rényi, "On measures of information and entropy," *in Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability*, vol. 1, pp. 547—561, 1961.

[60] O. Rioul, "A simple proof of the entropy-power inequality via properties of mutual information," *in Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, pp. 46–50, Jun. 2007.

[61] O. Rioul, "Information theoretic proofs of entropy power inequalities," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 33–55, Jan. 2011.

[62] O. Rioul, "Yet another proof of the entropy power inequality," *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3595–3599, Jun. 2017.

[63] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol. 21, no. 2, pp. 120–126, Feb. 1978.

[64] B. Schneier, "Inside risks: The uses and abuses of biometrics," *Commun. ACM* , vol. 42, no. 8, p. 136, Aug. 1999.

[65] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 623-656, Oct. 1948.

[66] A. Smith, "Maintaining secrecy when information leakage is unavoidable", *Ph.D. dissertation, Massachusetts Inst. of Technology*, Cambridge, 2004.

[67] E. C. Song, P. Cuff, and H. V. Poor, "The likelihood encoder for lossy compression" *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1836–1849, Apr. 2016.

[68] A. J. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Inf. Control*, vol. 2, no. 2, pp. 101–112, Jun. 1959.

[69] E. Tuncel, "Capacity/Storage tradeoff in high-dimensional identification systems," *in Proc. IEEE Int. Symp. Inf. Theory*, Seattle, USA, pp. 1929–1933, Jul. 2006.

[70] E. Tuncel, "Capacity/Storage tradeoff in high-dimensional identification systems," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2097–2016, May 2009.

[71] E. Tuncel and D. Gündüz, "Identification and lossy reconstruction in noisy databases," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 822–831, Feb. 2014.

[72] S. Verdú and D. Guo, "A simple proof of the entropy-power inequality," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2165–2166, May 2006.

[73] G.S. Vernam, "Cipher printing telegraph systems for secret wire and radio telegraphic communications," *Trans. of the American Institute of Electrical Engineers*, vol. 55, pp. 295-301, 1926.

[74] M. T. Vu, T. J. Oechtering, and M. Skoglund. "Gaussian hierarchical identification with preprocessing". *in Proc. IEEE Data Compression Conference,* Snowbird, UT, USA, pp. 277—286, Mar. 2018.

[75] M. T. Vu, T. J. Oechtering, and M. Skoglund "Hierarchical identification with pre-processing," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 82–113, Jan. 2020.

[76] F. M. J. Willems, T. Kalker, S. Baggen, and J. P. Linnartz, "On the capacity of a biometric identification system," *in Proc. IEEE Int. Symp. Inf. Theory,* Yokohama, Japan, p.82, Jun./Jul. 2003.

[77] F. M. J. Willems and T. Ignatenko, "Quantization effects in biometric systems," *In Proc. Inf. Theory and App. Workshop*, San Diego, CA, pp. 372–379, Feb. 2009.

[78] J. Wolfowitz, *Coding Theorems of Information Theory*, Berlin: Springer-Verlag, 1961.

[79] A. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequences and applications–I," *IEEE Trans. Inf. Theory*, vol. 19, no. 6, pp. 769–772, Nov. 1973.

[80] V. Yachongka and Y. Yagi, "Reliability function and strong converse of biometrical identification systems," *in Proc. 2016 Int. Symp. Inf. Theory and Its Appl.*, Montorey, CA, USA, Oct.–Nov. 2016.

[81] V. Yachongka and H. Yagi, "Fundamental trade-off among identification, secrecy and template rates in identification system," *in Proc. 2018 Int. Symp. on Inf. Theory and Its Appl.*, Singapore, Oct. 2018.

[82] V. Yachongka and H. Yagi, "Fundamental tradeoff among identification, secrecy and compression rates in biometric identification system," *Journal of Signal Processing*, vol. 22, no. 6, pp.337–342, Nov. 2018.

[83] V. Yachongka and H. Yagi, "Fundamental limits of biometric identification system under noisy enrollment," *IEICE Trans. Fundamentals*, vol. E104-A, no. 1, pp. 283-294, Jan. 2019.

[84] V. Yachongka and H. Yagi, “Fundamental limits of identification system with secret binding under noisy enrollment,” *arXiv:1905.03598*, May 2019.

[85] V. Yachongka and H. Yagi, “A new characterization of the capacity region of identification systems under noisy enrollment,” *54th Annual Conference on Information Sciences and Systems (CISS2020)*, Princeton, NJ, Mar. 2020.

[86] V. Yachongka and H. Yagi, “Biometric identification systems with both chosen and generated secrecy,” *ISITA2020*, Kapolei, Hawaii, USA, pp. 417–421, Oct. 2020.

[87] V. Yachongka, H. Yagi, and Y. Oohama, “Biometric identification systems with noisy enrollment for Gaussian source,” (to appear) *in Proc. IEEE Inf. Theory Workshop*, Apr. 2021.

[88] M. H. Yassaee, M. R. Aref, and A. Gohari, “Achievability proof via output statistics of random binning,” *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6760–6786, Nov. 2014.

[89] S. Yu, “Big privacy: Challenges and opportunities of privacy study in the age of big data,” *IEEE Access*, vol. 4, pp. 2751–2763, Jun. 2016.

[90] L. Zhou, M. T. Vu, T. Oechtering, and M. Skoglund, “Fundamental limits for biometric identification systems without privacy leakage,” *in Proc. 57th Annual Allerton Conf. on Commun., Control, and Comput.*, Monticello, USA, Sep. 2019.

[91] L. Zhou, V.Y.F. Tan, and M. Motani, “Strong converse for content identification with lossy recovery,” *in Proc. IEEE Int. Symp. Inf. Theory,* Aachen, Germany, pp. 928–934, Jun. 2017.

[92] L. Zhou, V.Y.F. Tan, L. Yu, and M. Motani, “Strong converse for content identification with lossy recovery,” *IEEE Trans. Inf. Theory*, vol. 64, no. 8, pp. 5879–5897, Aug. 2018.

Appendix A

Supplementary Proofs for Chapter 3

A.1 Proof of Remark 3.5

We show only the relation of $\mathcal{A}_1 = \mathcal{A}_3$ since the proof that $\mathcal{A}_2$ and $\mathcal{A}_4$ can be done similarly. In the proof, we prove the equivalence of $\mathcal{A}_1$ and $\mathcal{A}_3$ by removing the cardinality bounds of auxiliary RVs U and V from the two regions. Once the equivalence without the cardinality bounds is established, the cardinality bounds follow from the standard arguments (cf. [19, Appendix C]).

It is obvious that $\mathcal{A}_3 \subseteq \mathcal{A}_1$, so we shall show that $\mathcal{A}_3 \supseteq \mathcal{A}_1$. We assume that $(R_I, R_S, R_J, R_L) \in \mathcal{A}_1$, meaning that (R_I, R_S, R_J, R_L) satisfies all conditions in (3.18) for some $P_{U|Y}$. Especially, we have $R_I + R_S \leq I(Z;U)$. We choose the test channel $P_{V|U}$ satisfying that

$$R_I = I(Z;V). \tag{A.1}$$

Such $P_{V|U}$ always exists since $I(Z;U) \geq I(Z;V) \geq 0$ and $I(Z;V)$ is a continuous function of $P_{V|U}$. Under that condition, it is easy to check that (R_I, R_S, R_J, R_L) is also an element lying in the region $\mathcal{A}_3$. □

A.2 Proof of Lemma 3.1

In [29], a similar result of this lemma is used without the proof. Here, we will provide a proof for readers' sake. Note that $J(i) = (M(i), B(i))$. We start by considering the conditional entropy in the left-hand side of (3.39) as

$$\begin{aligned}\tfrac{1}{n}H(Y_i^n|J(i),S(i),\mathcal{C}_n) &= \tfrac{1}{n}H(Y_i^n|M(i),B(i),S(i),\mathcal{C}_n)\\ &\overset{\text{(a)}}{=} \tfrac{1}{n}H(Y_i^n|M(i),B(i),S(i),U_i^n,\mathcal{C}_n)\\ &\overset{\text{(b)}}{\leq} \tfrac{1}{n}H(Y_i^n|U_i^n,\mathcal{C}_n)\\ &\overset{\text{(c)}}{\leq} H(Y|U)+\delta_n' \end{aligned} \tag{A.2}$$

where

(a) holds because we denote $U^n(B(i),S(i)|M(i))$ as U_i^n for simplicity and the tuple $(M(i),B(i),S(i))$ determines U_i^n for a given codebook,

(b) follows because conditioning reduces entropy,

(c) follows because Y_i^n and U_i^n are jointly typical with high probability and (2.12) in Lemma 2.2 is applied.

□

A.3 Proof of Lemma 3.2

First, we prove that (3.55) holds. The joint distribution among $Z^{t-1}, Y^t(W), J(W)$, and $S(W)$ can be developed as

$$\begin{aligned}&P_{Z^{t-1},Y^t(W),J(W),S(W)}(z^{t-1},y_w^t,j(w),s(w))\\ &= \sum_{y_{w,t+1}^n\in\mathcal{Y}^{n-t}} \Big\{P_{Y_W^n}(y_w^n)\cdot P_{J(W),S(W)|Y_W^n}(j(w),s(w)|y_w^n)\\ &\quad \cdot P_{Z^{t-1}|Y_W^n,J(W),S(W)}(z^{t-1}|y_w^n,j(w),s(w))\Big\}\\ &\overset{\text{(d)}}{=} \sum_{y_{w,t+1}^n\in\mathcal{Y}^{n-t}} \Big\{P_{Y_W^n}(y_w^n)\cdot P_{J(W),S(W)|Y_W^n}(j(w),s(w)|y_w^n)\cdot P_{Z^{t-1}|Y_W^n}(z^{t-1}|y_w^n)\Big\}\\ &= \sum_{y_{w,t+1}^n\in\mathcal{Y}^{n-t}} \Big\{P_{Y_W^n}(y_w^n)\cdot P_{J(W),S(W)|Y_W^n}(j(w),s(w)|y_w^n)\Big\}\cdot P_{Z^{t-1}|Y^{t-1}(W)}(z^{t-1}|y_w^{t-1})\\ &= P_{Y^t(W),J(W),S(W)}(y_w^t,j(w),s(w))\cdot P_{Z^{t-1}|Y^{t-1}(W)}(z^{t-1}|y_w^{t-1})\\ &\overset{\text{(e)}}{=} P_{Y^{t-1}(W),J(W),S(W)}(y_w^{t-1},j(w),s(w))\cdot P_{Y_t(W)|Y^{t-1}(W),J(W),S(W)}(y_{wt}|y_w^{t-1},j(w),s(w))\\ &\quad \cdot P_{Z^{t-1}|Y^{t-1}(W),J(W),S(W)}(z^{t-1}|y_w^{t-1},j(w),s(w)),\end{aligned} \tag{A.3}$$

where

(d) holds because $(J(W),S(W))$ is a function of Y_W^n,

(e) follows because of the Markov chain $Z^{t-1} - Y^{t-1}(W) - (J(W), S(W))$.

Similarly, equation (3.56) can be shown as follows:

$$\begin{aligned}
&P_{Z^{t-1},X^t(W),J(W),S(W)}(z^{t-1}, x_w^t, j(w), s(w)) \\
&\quad = \sum_{y_w^n \in \mathcal{Y}^n} \Big\{ P_{Y_W^n}(y_w^n) \cdot P_{J(W),S(W)|Y_W^n}(j(w), s(w)|y_w^n) \\
&\qquad \cdot P_{X^t(W)|Y_W^n,J(W),S(W)}(x_w^t|y_w^n, j(w), s(w)) \\
&\qquad \cdot P_{Z^{t-1}|X^t(W),Y_W^n,J(W),S(W)}(z^{t-1}|x_w^t, y_w^n, j(w), s(w)) \Big\} \\
&\overset{\text{(f)}}{=} \sum_{y_w^n \in \mathcal{Y}^n} \Big\{ P_{Y_W^n}(y_w^n) \cdot P_{J(W),S(W)|Y_W^n}(j(w), s(w)|y_w^n) \\
&\quad \cdot P_{X^t(W)|Y_W^n,J(W),S(W)}(x_w^t|y_w^n, j(w), s(w)) \cdot P_{Z^{t-1}|X^t(W),Y_W^n}(z^{t-1}|x_w^t, y_w^n) \Big\} \\
&\overset{\text{(g)}}{=} \sum_{y_w^n \in \mathcal{Y}^n} \Big\{ P_{Y_W^n}(y_w^n) \cdot P_{J(W),S(W)|Y_W^n}(j(w), s(w)|y_w^n) \\
&\quad \cdot P_{X^t(W)|Y_W^n,J(W),S(W)}(x_w^t|y_w^n, j(w), s(w)) \Big\} \cdot P_{Z^{t-1}|X^{t-1}(W)}(z^{t-1}|x_w^{t-1}) \\
&= P_{X^t(W),J(W),S(W)}(x_w^t, j(w), s(w)) \cdot P_{Z^{t-1}|X^{t-1}(W)}(z^{t-1}|x_w^{t-1}) \\
&\overset{\text{(h)}}{=} P_{X^{t-1}(W),J(W),S(W)}(x_w^{t-1}, j(w), s(w)) \cdot P_{X_t(W)|X^{t-1}(W),J(W),S(W)}(x_{wt}|x_w^{t-1}, j(w), s(w)) \\
&\quad \cdot P_{Z^{t-1}|X^{t-1}(W),J(W),S(W)}(z^{t-1}|x_w^{t-1}, j(w), s(w)),
\end{aligned} \tag{A.4}$$

where

(f) holds because $(J(W), S(W))$ is a function of Y_W^n,

(g) follows due to the i.i.d. property of each symbol and the Markov chain $Z^{t-1} - X^{t-1}(W) - Y^{t-1}(W)$,

(h) follows because of the Markov chain $Z^{t-1} - X^{t-1}(W) - (J(W), S(W))$.

□

A.4 Proof of Lemma 3.3

We will prove only (3.57) by the well-known argument (cf. [14]). We introduce a timesharing variable Q which is uniformly distributed over $\{1, 2, \cdots, n\}$ and is independent of all other RVs. The left-hand side of (3.57) can be rewritten as

$$\begin{aligned}\sum_{t=1}^{n} I(Z_t;V_t) &= n\left\{\frac{1}{n}\sum_{t=1}^{n} I(Z_t;V_t|Q=t)\right\}\\ &= nI(Z_Q;V_Q|Q)\\ &= n[I(Z_Q;V_Q,Q)-I(Z_Q;Q)]\\ &= nI(Z_Q;V_Q,Q). \qquad \text{(A.5)}\end{aligned}$$

By denoting $V=(V_Q,Q)$ and $Z=Z_Q$, (3.57) obviously holds. The proof of (3.58)–(3.60) can be done similarly by setting $X=X_Q$ and $Y=Y_Q$.

To complete the proof, we need to verify that $Z_t - X_t(W) - Y_t(W) - U_t - V_t$ holds. We shall first check that $Z_t - X_t(W) - Y_t(W) - U_t$ holds for any $t \in [1,n]$. To prove this claim, we have to verify that

$$\begin{aligned}&Z_t - X_t(W) - Y_t(W), &\text{(A.6)}\\ &X_t(W) - Y_t(W) - U_t, &\text{(A.7)}\\ &Z_t - (X_t(W),Y_t(W)) - U_t. &\text{(A.8)}\end{aligned}$$

Indeed, Eqs. (A.6) and (A.7) clearly hold so the remaining task is to check if the last one also holds. Before checking that, we show that the Markov chain $Z_t - (Z^{t-1},X_t(W),Y_t(W)) - (J(W),S(W),W)$, which will be used to confirm (A.8), holds.

$$\begin{aligned}&I(Z_t;J(W),S(W),W|Z^{t-1},X_t(W),Y_t(W))\\ &\quad= H(Z_t|Z^{t-1},X_t(W),Y_t(W)) - H(Z_t|Z^{t-1},X_t(W),Y_t(W),J(W),S(W),W)\\ &\quad\overset{\text{(i)}}{\leq} H(Z_t|Z^{t-1},X_t(W),Y_t(W)) - H(Z_t|Z^{t-1},X_t(W),Y_W^n,J(W),S(W),W)\\ &\quad\overset{\text{(j)}}{=} H(Z_t|Z^{t-1},X_t(W),Y_t(W)) - H(Z_t|Z^{t-1},X_t(W),Y_W^n,W)\\ &\quad\overset{\text{(k)}}{=} H(Z_t|X_t(W)) - H(Z_t|X_t(W))\\ &\quad= 0, \qquad \text{(A.9)}\end{aligned}$$

where

(i) follows because conditioning reduces entropy,

(j) holds because $(J(W),S(W))$ is a function of Y_W^n,

(k) holds because each symbol of bio-data sequences is i.i.d., W is independent of other RVs, and we have $Z_t - X_t(W) - Y_t(W)$.

From (A.9), the conditional mutual information is zero and the claim is valid.

Equation (A.8) can be checked as follows:

$$
\begin{aligned}
&I(Z_t;U_t|X_t(W),Y_t(W))\\
&\quad= H(U_t|X_t(W),Y_t(W)) - H(U_t|X_t(W),Y_t(W),Z_t)\\
&\quad= H(Z^{t-1},J(W),S(W),W|X_t(W),Y_t(W))\\
&\qquad - H(Z^{t-1},J(W),S(W),W|X_t(W),Y_t(W),Z_t)\\
&\quad= H(Z^{t-1}|X_t(W),Y_t(W)) + H(J(W),S(W),W|X_t(W),Y_t(W),Z^{t-1})\\
&\qquad - H(Z^{t-1}|X_t(W),Y_t(W),Z_t) - H(J(W),S(W),W|X_t(W),Y_t(W),Z_t,Z^{t-1}) \qquad \text{(A.10)}\\
&\quad\overset{\text{(l)}}{=} H(J(W),S(W),W|X_t(W),Y_t(W),Z^{t-1}) - H(J(W),S(W),W|X_t(W),Y_t(W),Z^{t-1},Z_t)\\
&\quad\overset{\text{(m)}}{=} H(J(W),S(W),W|X_t(W),Y_t(W),Z^{t-1}) - H(J(W),S(W),W|X_t(W),Y_t(W),Z^{t-1})\\
&\quad= 0, \qquad \text{(A.11)}
\end{aligned}
$$

where

(l) holds because every symbol of bio-data sequences is i.i.d. generated so the first and third terms in (A.10) cancel each other,

(m) follows because $Z_t - (Z^{t-1},X_t(W),Y_t(W)) - (J(W),S(W))$ holds (cf. (A.9)).

Thus, $Z_t - X_t(W) - Y_t(W) - U_t$ holds, and since V_t is a function of U_t, it follows that $Z_t - X_t(W) - Y_t(W) - U_t - V_t$ also forms a Markov chain. □

Appendix B

Supplementary Proofs for Chapter 4

B.1 Equivalence of the Region in (4.15) and in (3.2)

One can easily see that $\mathcal{R}'$ is contained in the region in (3.2) due to the range of R_J. For proving the opposite relation, we choose a new test channel $P_{U'|U}$ satisfying that $R_I + R_C = I(U';Z)$. We can pick such channel since $I(Z;U) \geq I(Z;U') \geq 0$ and $I(Z;U')$ is a continuous function. The bounds of the template and privacy-leakage rates become

$$\begin{aligned} R_J &\geq I(Y;U) - I(Z;U) + I(Z;U') \\ &\overset{\text{(a)}}{\geq} I(Y;U') - I(Z;U') + I(Z;U') \\ &= I(Y;U') \qquad \text{(B.1)} \\ R_L &\geq I(X;U) - I(Z;U) + R_I \\ &\overset{\text{(b)}}{\geq} I(X;U') - I(Z;U') + R_I, \qquad \text{(B.2)} \end{aligned}$$

where (a) and (b) follow from the face that $I(Y;U|Z) \geq I(Y;U'|Z)$ and $I(X;U|Z) \geq I(X;U'|Z)$, respectively. Hence, there always exists an auxiliary U' where an achievable rate tuple (R_I, R_C, R_J, R_L) in the region in (3.2) is also included in $\mathcal{R}'$.

□

B.2 Proof of Lemma 4.4

We show only the proofs of (4.51) and (4.54). We omit the proofs of the others because (4.52) and (4.53) can be proved by similar arguments of (4.51), and (4.55) follows similarly from the arguments of (4.54). We begin with checking equation (4.51).

$$\begin{aligned}\frac{1}{n}H(S_1(i)|\mathcal{C}_n) &= \frac{1}{n}H(Y_i^n, S_1(i), S_2(i), M(i)|\mathcal{C}_n) - \frac{1}{n}H(S_2(i), M(i)|S_1(i), \mathcal{C}_n)\\ &\quad - \frac{1}{n}H(Y_i^n|S_1(i), S_2(i), M(i), \mathcal{C}_n)\\ &\overset{(a)}{\geq} \frac{1}{n}H(Y_i^n) - \frac{1}{n}H(S_2(i)|\mathcal{C}_n) - \frac{1}{n}H(M(i)|\mathcal{C}_n) - \frac{1}{n}H(Y_i^n|V(i), \mathcal{C}_n)\\ &\overset{(b)}{\geq} H(Y) - (I(Z;U) - R_I - R_C - \delta) - (I(Y;U) - I(Z;U) + R_I + 2\delta)\\ &\quad - (H(Y|U) + \delta_n)\\ &\geq R_C - \delta - \delta_n, \end{aligned} \tag{B.3}$$

where

(a) follows because conditioning reduces entropy and Y_i^n is independent of $\mathcal{C}_n$,

(b) follows because (4.49) in Lemma 4.3 is applied.

Next, we prove (4.54). From the left-hand side of the equation, we have that

$$\begin{aligned}&\frac{1}{n}I(S_1(i), S_2(i); M(i)|\mathcal{C}_n)\\ &\quad= \frac{1}{n}H(S_1(i), S_2(i)|\mathcal{C}_n) + \frac{1}{n}H(M(i)|\mathcal{C}_n) + \frac{1}{n}H(Y_i^n, S_1(i), S_2(i), M(i)|\mathcal{C}_n)\\ &\quad\quad + \frac{1}{n}H(Y_i^n|S_1(i), S_2(i), M(i), \mathcal{C}_n)\\ &\quad\overset{(c)}{\leq} \frac{1}{n}H(S_1(i)|\mathcal{C}_n) + H(S_2(i)|\mathcal{C}_n) + \frac{1}{n}H(M(i)|\mathcal{C}_n) - \frac{1}{n}H(Y_i^n) + \frac{1}{n}H(Y_i^n|V(i), \mathcal{C}_n)\\ &\quad\overset{(d)}{\leq} R_C + (I(Z;U) - R_I - R_C - \delta) + (I(Y;U) - I(Z;U) + R_I + 2\delta)\\ &\quad\quad - H(Y) + (H(Y|U) + \delta_n)\\ &\quad= \delta + \delta_n, \end{aligned} \tag{B.4}$$

where

(c) follows because conditioning reduces entropy,

(d) follows (4.49) in Lemma 4.3 is applied.

□

Appendix C

Supplementary Proofs for Chapter 5

C.1 Convexity of $\mathcal{R}_G$

In this appendix, we prove that the region $\mathcal{R}_G$ is convex. First we define $\eta = \frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}$. and then it follows that $\alpha = \frac{1}{\rho_1^2\rho_2^2}\left(\frac{1}{\eta} - (1-\rho_1^2\rho_2^2)\right)$. Therefore, the right-hand sides of R_J and R_L in (5.15) can be transformed as follows:

$$\begin{aligned}
R_J &\geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}\right)+R_I+R_C \\
&= \frac{1}{2}\ln\left(\frac{\frac{1}{\rho_1^2\rho_2^2}\left(\frac{1}{\eta}-(1-\rho_1^2\rho_2^2)\right)\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\frac{1}{\rho_1^2\rho_2^2}\left(\frac{1}{\eta}-(1-\rho_1^2\rho_2^2)\right)}\right)+R_I+R_C \\
&= \frac{1}{2}\ln\left(\frac{\rho_1^2\rho_2^2}{1-(1-\rho_1^2\rho_2^2)\eta}\right)+R_I+R_C \\
&= -\frac{1}{2}\ln\left(1-(1-\rho_1^2\rho_2^2)\eta\right)+\ln|\rho_1\rho_2|+R_I+R_C
\end{aligned} \tag{C.1}$$

and

$$\begin{aligned}
R_L &\geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha\rho_1^2+1-\rho_1^2}\right)+R_I \\
&= \frac{1}{2}\ln\left(\frac{\frac{1}{\rho_1^2\rho_2^2}\left(\frac{1}{\eta}-(1-\rho_1^2\rho_2^2)\right)\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\frac{1}{\rho_1^2\rho_2^2}\left(\frac{1}{\eta}-(1-\rho_1^2\rho_2^2)\right)\rho_1^2+1-\rho_1^2}\right)+R_I \\
&= \frac{1}{2}\ln\left(\frac{\rho_2^2}{1-(1-\rho_2^2)\eta}\right)+R_I \\
&= -\frac{1}{2}\ln\left(1-(1-\rho_2^2)\eta\right)+\ln|\rho_2|+R_I.
\end{aligned} \tag{C.2}$$

Since $|\rho_1|, |\rho_2| < 1$, and $0 < \alpha \leq 1$, we have that $\frac{1-\rho_2^2}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2} \leq \frac{1-\rho_1^2\rho_2^2}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2} < 1$, indicating the values of $1-(1-\rho_1^2\rho_2^2)\eta$ and $1-(1-\rho_2^2)\eta$ are positive. Now the region in (5.15) can also be expressed as follows:

$$\begin{aligned}\mathcal{R}_G = \Big\{(R_I,R_C,R_G,R_J,R_L)\ :\ & R_I+R_C \le \frac{1}{2}\ln\eta,\\ & R_I+R_C+R_G \le \frac{1}{2}\ln\eta+\Gamma,\\ & R_J \ge -\frac{1}{2}\ln\big(1-(1-\rho_1^2\rho_2^2)\eta\big)+\ln|\rho_1\rho_2|+R_I+R_C,\\ & R_L \ge -\frac{1}{2}\ln\big(1-(1-\rho_2^2)\eta\big)+\ln|\rho_2|+R_I,\\ & R_I\ge 0,\ \ R_C\ge\Gamma,\ \ R_G\ge 0 \text{ for some } 1\le\eta<\frac{1}{1-\rho_1^2\rho_2^2}\Big\}.\end{aligned} \tag{C.3}$$

Suppose that $\boldsymbol{R}_1=(R_I^1,R_C^1,R_G^1,R_J^1,R_L^1)$ and $\boldsymbol{R}_2=(R_I^2,R_C^2,R_G^2,R_J^2,R_L^2)$ are achievable tuples for η_1 and η_2, respectively. Without loss of generality, we assume that $1\le\eta_1\le\eta_2<\frac{1}{1-\rho_1^2\rho_2^2}$. Next, let consider linear combination of these tuples. For $0\le\lambda\le 1$, we have that

$$\begin{aligned}\lambda(R_I^1+R_C^1)+(1-\lambda)(R_I^2+R_C^2) &\le \frac{1}{2}(\lambda\ln\eta_1+(1-\lambda)\ln\eta_2)\\ &\overset{(a)}{\le} \frac{1}{2}\ln(\lambda\eta_1+(1-\lambda)\eta_2)\\ &\overset{(b)}{=} \frac{1}{2}\ln\eta',\end{aligned} \tag{C.4}$$

where

(a) is due to $\log x$ $(x>0)$ is a convex upward function,

(b) holds as we define $\eta'=\lambda\eta_1+(1-\lambda)\eta_2$.

Similarly, we can show that

$$\lambda(R_I^1+R_C^1+R_G^1)+(1-\lambda)(R_I^2+R_C^2+R_G^2)\le\frac{1}{2}\ln\eta'+\Gamma. \tag{C.5}$$

Let take a look into the template rate.

$$\begin{aligned}\lambda R_J^1+(1-\lambda)R_J^2 &\ge -\lambda\frac{1}{2}\ln\big(1-(1-\rho_1^2\rho_2^2)\eta_1\big)-(1-\lambda)\frac{1}{2}\ln\big(1-(1-\rho_1^2\rho_2^2)\eta_2\big)\\ &\quad+\ln|\rho_1\rho_2|+R_I+R_C\\ &\overset{(c)}{\ge} -\frac{1}{2}\ln\big(1-(1-\rho_1^2\rho_2^2)(\lambda\eta_1+(1-\lambda)\eta_2)\big)+\ln|\rho_1\rho_2|+R_I+R_C\\ &= -\frac{1}{2}\ln\big(1-(1-\rho_1^2\rho_2^2)\eta'\big)+\ln|\rho_1\rho_2|+R_I+R_C,\end{aligned} \tag{C.6}$$

where (c) follows because $f(x) = -\ln(1-x)$ $(x < 1)$, is a convex downward function. Likewise, we can also show that

$$\lambda R_L^1 + (1-\lambda)R_L^2 \geq -\frac{1}{2}\ln\left(1-(1-\rho_2^2)\eta'\right) + \ln|\rho_2| + R_I. \tag{C.7}$$

From (C.4)–(C.7), we see that there exists an η', where $\eta_1 \leq \eta' \leq \eta_2$, that satisfies $\lambda \boldsymbol{R}_1 + (1-\lambda)\boldsymbol{R}_2 \in \mathcal{R}_G$. This indicates that the region $\mathcal{R}_G$ is convex. □

C.2 Verification of $\mathcal{R}_G^2 = \mathcal{R}_G''$

It is easy to see that $\mathcal{R}_G^2 \in \mathcal{R}_G''$ due to the range of the template rate. It is clear that

$$\begin{aligned}\frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}\right) + R_C &\leq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}\right) + \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right)\\ &= \frac{1}{2}\ln\left(\frac{1}{\alpha}\right).\end{aligned} \tag{C.8}$$

To prove that $\mathcal{R}_G'' \in \mathcal{R}_G^2$, we assume that $(R_C, R_J, R_L) \in \mathcal{R}_G''$. For a given α, we pick another α' $(\alpha \leq \alpha' \leq 1)$ that satisfies

$$R_C = \frac{1}{2}\ln\left(\frac{1}{\alpha'\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right), \tag{C.9}$$

which is possible since $\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right)$ is continuous function for $0 \leq \alpha \leq 1$ and it is guaranteed that $\frac{1}{2}\ln\left(\frac{1}{\alpha'\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right) \leq \frac{1}{2}\ln\left(\frac{1}{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right)$. Thus, we have

$$\begin{aligned}R_J &\geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha}\right) + R_C,\\ &\geq \frac{1}{2}\ln\left(\frac{\alpha'\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha'}\right) + \frac{1}{2}\ln\left(\frac{1}{\alpha'\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}\right),\\ &= \frac{1}{2}\ln\left(\frac{1}{\alpha'}\right)\end{aligned} \tag{C.10}$$

$$R_L \geq \frac{1}{2}\ln\left(\frac{\alpha\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha\rho_1^2+1-\rho_1^2}\right) \geq \frac{1}{2}\ln\left(\frac{\alpha'\rho_1^2\rho_2^2+1-\rho_1^2\rho_2^2}{\alpha'\rho_1^2+1-\rho_1^2}\right). \tag{C.11}$$

From (C.9)–(C.11), we see that (R_C, R_J, R_L) also lies in $\mathcal{R}_G^2$. □