



<http://www.diva-portal.org>

This is the published version of a paper published in *IEEE Transactions on Information Forensics and Security*.

Citation for the original published paper (version of record):

Zhou, L., Vu, M T., Oechtering, T J., Skoglund, M. (2021)
Privacy-Preserving Identification Systems With Noisy Enrollment
IEEE Transactions on Information Forensics and Security, 16: 3510-3523
<https://doi.org/10.1109/TIFS.2021.3078297>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-298636>

Privacy-Preserving Identification Systems with Noisy Enrollment

Linghui Zhou, *Student Member, IEEE*, Minh Thanh Vu,
Tobias J. Oechtering, *Senior Member, IEEE*, and Mikael Skoglund, *Fellow, IEEE*.

Abstract—In this paper, we study fundamental trade-offs in privacy-preserving biometric identification systems with noisy enrollment. The proposed identification systems include helper data, secret keys, and private keys. Helper data are stored in a public database and used for identification. Secret keys are either stored in a secure database or provided to the user, and can be used in a next step, e.g. for authentication. Private keys are provided by users, and are also used for identification. In this paper, we impose a noisy enrollment channel and an arbitrarily small privacy and secrecy leakage rate. We characterize the optimal trade-off among the identification, secret key, private key, and helper data rates. Depending on how secret keys are produced, we study two cases of the proposed privacy-preserving identification systems, where the secret keys are *generated* and *chosen* respectively. By introducing private keys, it is shown that the identification system achieves close to zero privacy leakage rate in both *generated* and *chosen* secret key settings. The results also show that the identification rate and the secret key rate can be enlarged by increasing the private key rate. This work provides a framework for analyzing privacy-preserving identification systems and an insight on the design of optimal systems.

Index Terms—Biometrics, identification systems, noisy enrollment, privacy, secrecy.

I. INTRODUCTION

With recent advances in technology and smart devices, biometrics are more and more deployed to reinforce traditional authentication or identification that uses keys, passwords, etc. Compared to traditional methods, biometric features have the advantage that they are more stable and individualized. Nowadays, the most recognized biometric technologies are fingerprint mapping, face recognition, and retina scans. The use of biometric features makes authentication or identification more convenient, but the abuse of biometric information could invoke serious privacy issues. A recent breach of BioStar, which is a biometric identification system using facial recognition and fingerprinting technology, leads to a compromise of millions records containing personal information of sensitive nature. As stated in [1], once biometric information is compromised, the privacy can not be restored. Therefore, it is important to take privacy into account when designing identification or authentication systems using biometrics.

The work was partially supported by the Swedish Research Council under grant 2016-03853, the Digital Futures research center, and the Strategic Research Agenda Program, Information and Communication Technology - The Next Generation (SRA ICT - TNG), through the Swedish Government.

The authors are with the Division of Information Science and Engineering, KTH Royal Institute of Technology and KTH Digital Future Center, 100 44 Stockholm, Sweden (e-mail: linghui@kth.se; mtvu@kth.se; oech@kth.se; skoglund@kth.se).

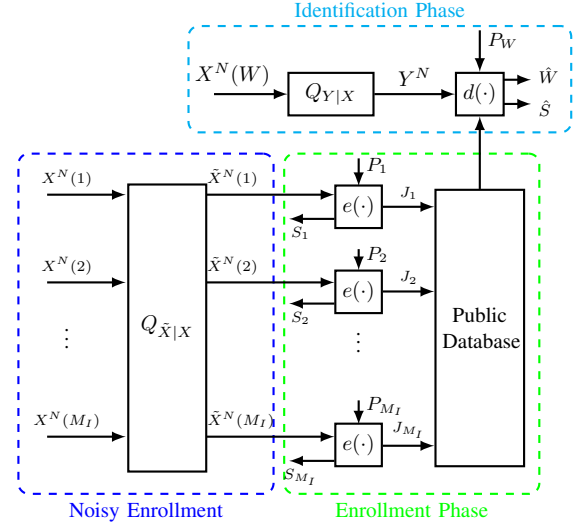


Fig. 1. *Generated Secret Key System*: In the enrollment phase, each user $w \in [1 : M_I]$ is observed via a noisy enrollment channel $Q_{\tilde{X}|X}$ and generates an observation $\tilde{X}^N(w)$. Then the enrollment mapping maps the observation $\tilde{X}^N(w)$ and private key P_w to the helper data J_w and secret key S_w . In the identification phase, assume a previously enrolled user is observed. The system compares the observation Y^N with the database and uses the private key to guess the user index and the secret key.

Several aspects of identification and authentication problems have been studied. The fundamental limits of an identification system was firstly studied by Willems *et al.* in [2], where the identification capacity of a biometric identification system was characterized. Tuncel *et al.* analyzed in [3] the identification capacity and the storage trade-off in a biometric identification system. The identification rate, search and memory complexity trade-offs are studied in [4]. In [5], several assumptions in the identification problems are relaxed and the corresponding fundamental bounds regions are derived. The information theoretic perspectives of robust authentication systems is investigated in [6]. Hypothesis testing in identification and authentication system is studied in [7] and [8], respectively. The privacy and secrecy aspects of biometric systems are studied in [9], [10]. Algorithmic computability of the secret key and authentication with constraints is discussed in [11]. The problem of controllable authentication with privacy and storage constraints on the source sequence is considered in [12]. In [13], Kittichokechai *et al.* investigated the secret-based identification and authentication with a privacy constraint. Fundamental trade-offs in biometric identification systems that supports authentication are studied in [14].

For the privacy aspect, several privacy metrics have been studied in the literature. Often, one aims for privacy against statistical inference, such as membership privacy or reconstruction of certain data. In this work, we take an information-theoretic approach and use mutual information rate as privacy leakage measure. By bounding the mutual information, we limit the amount of information leaked about the biometric source from the public data. Since the generation of the public helper data can be seen as source coding process, bounding the mutual information leakage directly relates to the rate-distortion problem, which therefore protects against adversarial reconstruction of the biometric data. The achieved privacy level in [9], [13], [14] is also measured using mutual information rate where, however, the privacy leakage is not necessarily small. In this work, we consider the case that the privacy leakage is required to be close to zero, i.e., the mutual information rate should be negligible.

Additional to a more restrictive privacy constraints, the noise in the observation in both the enrollment and the identification phases has to be included in the design. For example, when scanning the fingerprints or faces, it is inevitable that there would be random noise due to the devices, as well as different angles or positions when one scans the physical features. The noisy enrollment can be interpreted as an additional fixed privacy filter that protects the true biometric source so that the privacy leakage measures the total protection of the system. Therefore, considering the noisy enrollment is closer to real life scenarios and has a positive side effect on protecting the true biometric data. Biometric systems with noisy enrollment and without zero privacy leakage constraints are studied in [15]–[17], in which it is assumed that the true biometric source is hidden and only noisy versions of biometric source are available to the system. Therein, the public helper data are generated based on the noisy enrollment whereas in our work the public label is generated from the combination of noisy information and a private key. Additionally, in our work, the privacy leakage is evaluated with respect to the noise-free biometric source instead of the noisy enrollment, which makes the privacy leakage analysis with noisy enrollment more challenging.

This work extends the single-user problem with close to zero privacy leakage rate in [9] to noisy enrollment and the problem of identification that allows authentication. In other words, by introducing an extra private key as in [9], close to zero privacy leakage rate is achieved. This work is also an extension of [18], while in this work we additionally include noisy enrollment channel. Due to the existence of noise in the enrollment phase, the proofs are significantly changed and results are more general. Moreover, depending on how the secret keys are produced, we consider two variations of identification systems, where the secret keys are generated and chosen respectively.

The rest of the paper is organized as follows. In Section II, we introduce the problem formulation for two settings, the *generated secret key system* and the *chosen secret key system*. In Section III, the main results are presented, where we will determine the optimal regions for both settings. Finally, in Section IV, we summarize the work and provide a conclusion.

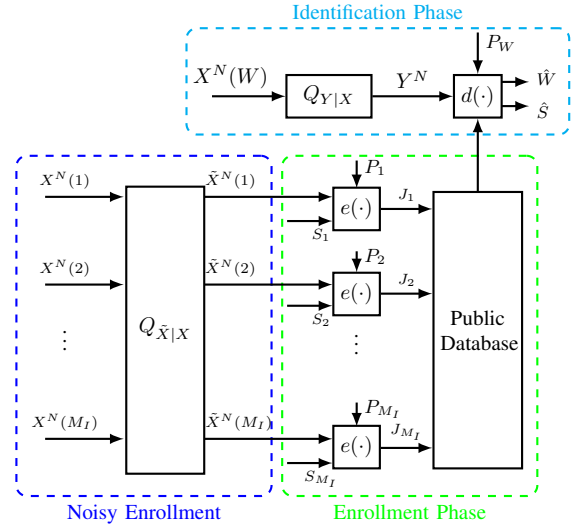


Fig. 2. *Chosen Secret Key System*: In the enrollment phase, each user $w \in [1 : M_I]$ is observed via a noisy enrollment channel $Q_{\tilde{X}|X}$ and generates an observation $\tilde{X}^N(w)$. Then the enrollment mapping maps the observation $\tilde{X}^N(w)$, the secret key S_w and private key P_w to the helper data J_w . In the identification phase, assume a previously enrolled user W is observed. The system compares the observation Y^N with the database and uses the private key to guess the user index and the secret key.

Notations: We denote random variables, their realizations and alphabets by upper cases, lower cases and calligraphic letters. We use X^N to denote a vector (X_1, X_2, \dots, X_N) . $H(X)$, $I(X; Y)$, $Q_X(x)$ and $Q_{Y|X}(y|x)$ denote the entropy, mutual information, marginal and conditional probability distributions respectively. The strong typical set is denoted by \mathcal{T}_ϵ^N . We use $|\mathcal{A}|$ to denote the cardinality of a finite set \mathcal{A} .

II. PROBLEM FORMULATION

A. Noisy Enrollment

Fig. 1 depicts the *generated secret key system*. Assume that there are M_I users indexed by $w \in [1 : M_I]$. We use $x^N(w)$ to denote the biometric sequence of user $w \in [1 : M_I]$, which is assumed to be identically independently distributed (i.i.d.) according to the mass probability function (p.m.f.) $Q_X(\cdot)$ defined on the finite alphabet \mathcal{X} .

In the enrollment phase, for each user w , the biometric sequence $x^N(w)$ is observed via a noisy memoryless enrollment channel $Q_{\tilde{X}|X}(\cdot)$ and an observation $\tilde{x}^N(w)$ is generated. Hence, given $x^N(w)$, the observation $\tilde{x}^N(w)$ in the enrollment phase occurs with probability

$$\begin{aligned} \Pr\{\tilde{X}^N(w) = \tilde{x}^N(w) | X^N(w) = x^N(w)\} \\ = \prod_{i=1}^N Q_{\tilde{X}|X}(\tilde{x}_i(w) | x_i(w)). \end{aligned} \quad (1)$$

B. Generated Secret Key Systems

The enrollment mapping $e(\cdot)$ maps the noisy biometric sequence $\tilde{x}^N(w)$ and the private key $p_w \in [1 : M_P]$ to generate the helper data $j_w \in [1 : M_J]$ and a secret key $s_w \in [1 : M_S]$, i.e.,

$$(j_w, s_w) = e(\tilde{x}^N(w), p_w). \quad (2)$$

We assume that the private key for every user $w \in [1 : M_I]$ is independent and uniformly distributed on $[1 : M_P]$. The private key p_w is also presented to the system when the user w wants to be identified in the identification phase.

In the identification phase, an unknown user w is observed via the discrete memoryless channel (DMC) $Q_{Y|X}(\cdot)$. It is assumed that the user index w is uniformly distributed over $[1 : M_I]$.

Lastly, after an unknown user with index w has been observed, the observation y^N and its corresponding private key p_w are presented to the system. The identification mapping $d(\cdot)$ estimates the user index $\hat{w} \in [1 : M_I]$ and the secret key $\hat{s} \in [1 : M_S]$ as

$$(\hat{w}, \hat{s}) = d(y^N, \mathbf{j}, p_w), \quad (3)$$

where $\mathbf{j} = (j_i)_{i=1}^{M_I}$ denotes data of the database.

We are interested in the optimal trade-off among the achievable identification, secret key, private key, and helper data rates such that the identification system is able to: (a) return the true user index and the secret key with high probability; (b) distributions of secret keys are approximately uniform; (c) preserve privacy and secrecy such that the privacy leakage rate and secrecy leakage rate are arbitrarily small. Moreover, we want the identification rate and secret key rate as large as possible, and the private key rate and helper data rate as small as possible. Accordingly, we define the achievability for the generated secret key system as follows.

Definition 1: An identification, secret key, private key, and helper data rate tuple $(R_I, R_S, R_P, R_J) \in \mathbb{R}_+^4$ is *achievable* in a generated secret key system if, given any $\delta > 0$ there exists some $N_0(\delta) \geq 1$, enrollment mapping $e(\cdot)$, and identification mapping $d(\cdot)$ such that for any $N \geq N_0(\delta)$, the following conditions are satisfied

$$\Pr\{(\hat{W}, \hat{S}) \neq (W, S_W)\} \leq \delta, \quad (4a)$$

$$\log M_I \geq N(R_I - \delta), \quad (4b)$$

$$H(S_W) + N\delta \geq \log M_S \geq N(R_S - \delta), \quad (4c)$$

$$\log M_P \leq N(R_P + \delta), \quad (4d)$$

$$\log M_J \leq N(R_J + \delta), \quad (4e)$$

$$I(S_W; J_W) \leq N\delta, \quad (4f)$$

$$I(X^N(W); J_W) \leq N\delta. \quad (4g)$$

The capacity region \mathcal{R}_g is the closure of the set of all achievable identification, secret key, private key, and helper data rate tuples for a generated secret key system.

The capacity region can be interpreted as follows: (4a) indicates that the identification system is able to return the true user index and the secret key with only negligible error probability; (4b), (4c), (4d) and (4e) put constraints on the identification, secret, private key and helper data rate respectively; especially, (4c) also states that distributions of secret keys are approximately uniform; (4f) and (4g) require the system is secrecy-preserving and privacy-preserving in a weak sense, respectively.

C. Chosen Secret Key System

Similarly, in a *chosen secret key system*, as illustrated in Fig. 2, in the enrollment phase, the system observes the noisy

enrolled biometric sequence $\tilde{x}^N(w)$ of user w and private key p_w . Additionally, the system also observes a chosen secret key s_w . It is assumed that the secret key s_w and private key p_w are chosen uniformly at random from $[1 : M_S]$ and $[1 : M_P]$ respectively.

In the identification phase, an unknown user w is observed via a discrete memoryless channel (DMC) $Q_{Y|X}(\cdot)$, where w is uniformly distributed over $[1 : M_I]$. After an unknown user with index w is observed, the observation y^N and its private key p_w are provided to the system. The identification mapping $d(\cdot)$ estimates the user index $\hat{w} \in [1 : M_I]$ and the secret key $\hat{s} \in [1 : M_S]$ as

$$(\hat{w}, \hat{s}) = d(y^N, \mathbf{j}, p_w), \quad (5)$$

where $\mathbf{j} = (j_i)_{i=1}^{M_I}$. We define the achievability of a chosen secret key system as follows.

Definition 2: An identification, secret key, private key, and helper data rates tuple $(R_I, R_S, R_P, R_J) \in \mathbb{R}_+^4$ is *achievable* in a chosen secret key system if, given any $\delta > 0$ there exists some $N_0(\delta) \geq 1$, enrollment and identification mappings such that for any $N \geq N_0(\delta)$ the conditions (4a), (4b), (4d), (4e), (4f), and (4g) and

$$\log M_S \geq N(R_S - \delta), \quad (6)$$

are satisfied.

The capacity region \mathcal{R}_c is the closure of the set of all achievable identification, secret key, private key, and helper data rate tuples for a chosen secret key system.

III. OPTIMAL TRADE-OFFS

Now we give the capacity regions of generated secret key systems and chosen secret key systems. A binary example is also provided to illustrate the optimal trade-offs.

A. Capacity Regions and Discussion

Theorem 1: For a privacy-preserving identification system using *generated secret keys* with noisy enrollment, the capacity region \mathcal{R}_g is given by

$$\mathcal{R}_g = \{(R_I, R_S, R_P, R_J) \in \mathbb{R}_+^4 :$$

$$R_I + R_S \leq R_P + I(U; Y), \quad (7a)$$

$$R_I + I(U; X) \leq R_P + I(U; Y), \quad (7b)$$

$$R_I + I(U; \tilde{X}) \leq R_J + I(U; Y), \quad (7c)$$

$$\text{for some } P_{U\tilde{X}XY} = Q_X Q_{\tilde{X}|X} Q_{Y|X} P_{U|\tilde{X}}, \quad (7d)$$

$$\text{and } |\mathcal{U}| \leq |\tilde{\mathcal{X}}| + 2\}. \quad (7e)$$

Theorem 2: For a privacy-preserving identification system using *chosen secret keys* with noisy enrollment, the capacity region \mathcal{R}_c is given by

$$\mathcal{R}_c = \{(R_I, R_S, R_P, R_J) \in \mathbb{R}_+^4 :$$

$$R_I + R_S \leq R_P + I(U; Y), \quad (8a)$$

$$R_I + I(U; X) \leq R_P + I(U; Y), \quad (8b)$$

$$R_I + I(U; \tilde{X}) + R_S \leq R_J + I(U; Y), \quad (8c)$$

$$\text{for some } P_{U\tilde{X}XY} = Q_X Q_{\tilde{X}|X} Q_{Y|X} P_{U|\tilde{X}}, \quad (8d)$$

$$\text{and } |\mathcal{U}| \leq |\tilde{\mathcal{X}}| + 2\}. \quad (8e)$$

In the following, we provide some discussion of the results.

Corollary 1: If a rate tuple $(R_I, R_S, R_P, R_J) \in \mathcal{R}_g$, then $(R_I, R_S, R_P, R_J + R_S) \in \mathcal{R}_c$.

Comparing (7c) and (8c), we can conclude that, to achieve the same identification rate, secret key rate and private key rate tuple (R_I, R_S, R_P) , the chosen secret key system needs a larger minimum helper data rate, which is the sum of the minimum helper data rate in the generated secret key system and the secret key rate.

Corollary 2: If $(R_I, R_S, R_P, R_J) \in \mathcal{R}_g$, then for any $r_1 \geq 0$ and $r_2 \geq 0$, $(R_I + r_1, R_S + r_2, R_P + r_1 + r_2, R_J + r_1) \in \mathcal{R}_g$ and $(R_I + r_1, R_S + r_2, R_P + r_1 + r_2, R_J + r_1 + r_2) \in \mathcal{R}_c$ hold.

On one hand, the above result can be interpreted as a rate transfer argument. An extra private key can be used to increase identification rate or secret key rate. On the other hand, to increase the identification rate, the helper data rate has to be enlarged correspondingly. Further, for the chosen secret key system, the helper data rate also increases as the secret rate increase.

In the following analysis, for simplicity, we use the following notations. For $R_J \geq 0$, denote $\mathcal{R}_g(R_J)$ and $\mathcal{R}_c(R_J)$ to be the set of rate triple (R_I, R_S, R_P) such that $(R_I, R_S, R_P, R_J) \in \mathcal{R}_g$ and $(R_I, R_S, R_P, R_J) \in \mathcal{R}_c$, respectively. Moreover, we define the region \mathcal{R}^* as follows

$$\begin{aligned} \mathcal{R}^* = \{ & (R_I, R_S, R_P) \in \mathbb{R}_+^3 : \\ & R_I + R_S \leq R_P + I(U; Y), \\ & R_I + I(U; X) \leq R_P + I(U; Y) \}. \end{aligned} \quad (9)$$

Corollary 3: If $R_J^1 \leq R_J^2$, then $\mathcal{R}_g(R_J^1) \subseteq \mathcal{R}_g(R_J^2)$ and $\mathcal{R}_c(R_J^1) \subseteq \mathcal{R}_c(R_J^2)$. Additionally, there exists a R_J^0 such that for every $R_J \geq R_J^0$, $\mathcal{R}_g(R_J) = \mathcal{R}_c(R_J) = \mathcal{R}^*$.

If we increase the helper data rate, the capacity region is larger. Further, if R_J is increased to be sufficiently large, then the constraints (7c) and (8c) are not active. Since the remaining constraints are the same, the generated and chosen secret key systems achieve the same capacity region.

In the following analysis, for simplicity and without losing generality, we consider the above case that the storage can be arbitrarily large, and the generated and chosen secret key systems both achieve the region \mathcal{R}^* .

Corollary 4: When $\tilde{X} = X$, then

$$\begin{aligned} \mathcal{R}^*|_{\tilde{X}=X} = \{ & (R_I, R_S, R_P) \in \mathbb{R}_+^3 : \\ & R_I + R_S \leq R_P + I(U; Y), \\ & R_I + I(U; X) \leq R_P + I(U; Y) \}. \end{aligned} \quad (10)$$

The region above corresponds to the region for biometric identification system without privacy leakage and with clean enrollment channel, which is derived in [18]. Therefore, the results in Theorem 1 and 2 give more general results for the biometric identification systems without privacy leakage.

Corollary 5: When $X = \tilde{X}$ and further $R_I = 0$, then

$$\begin{aligned} \mathcal{R}^*|_{X=\tilde{X}, R_I=0} = \{ & (R_S, R_P) \in \mathbb{R}_+^2 : \\ & R_S \leq R_P + I(U; Y), \end{aligned}$$

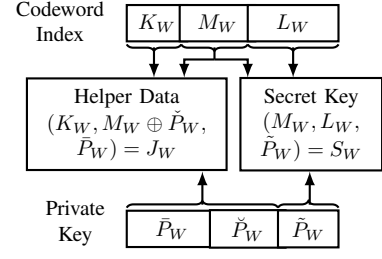


Fig. 3. Illustration of generating helper data and secret keys in the generated secret key system. Note that the block size does not necessarily reflect the true sequence length.

$$I(U; X) \leq R_P + I(U; Y)\}. \quad (11)$$

This region corresponds to the region for biometric authentication without privacy leakage derived in [9]. Therefore, authentication is a special case of identification that allows authentication.

B. Overview of Proofs

The proofs of Theorem 1 and Theorem 2 consist of two parts, i.e., the achievability part and the converse part. The detailed achievability proof of Theorem 1 and Theorem 2 are provided in Appendix A and Appendix B, respectively. The achievability for Theorem 2 is an extension of Theorem 1 adding an extra masking layer. In this layer, the one-time pad is used to mask a generated secret key. Here we provide a sketch of the achievability proof of Theorem 1.

We first fix a conditional p.m.f. $P_{U|\tilde{X}}$, which determines the joint p.m.f.

$$P_{U\tilde{X}XY} = P_{U|\tilde{X}}Q_{\tilde{X}|X}Q_XQ_{Y|X}. \quad (12)$$

Here, U is an auxiliary variable that describes the codebook. Then we generate roughly $2^{NI(U;\tilde{X})}$ codebook sequences u^N . Each sequence is assigned to a bin. We have roughly $2^{N(I(U;\tilde{X})-I(U;X))}$ bins with $2^{NI(U;X)}$ sequences in each bin. Each sequence in each bin is additionally assigned to one out of $2^{N(I(U;X)-I(U;Y))}$ sub-bins with $2^{NI(U;Y)}$ sequences in each sub-bin.

In the enrollment phase, data from M_I users are enrolled. For each user $w \in [1 : M_I]$, the system observes a noisy biometric sequence $\tilde{x}^N(w)$ and receives a private key p_w . The private key is divided into three parts, \tilde{p}_w , \check{p}_w and \tilde{p}_w . We assume roughly 2^{NR_I} values of \tilde{p}_w , $2^{N(I(U;X)-I(U;Y))}$ values of \check{p}_w , and $2^{N(R_P-I(U;X)+I(U;Y))}$ values of \tilde{p}_w . A sequence $u^N(k_w, m_w, l_w)$ is looked for such that it is jointly typical with $\tilde{x}^N(w)$. A helper data is generated as $j_w = (k_w, m_w \oplus \tilde{p}_w, \tilde{p}_w)$, which is stored in a public database. Then the enrollment mapping generates a secret key $s_w = (m_w, l_w, \tilde{p}_w)$. Fig. 3 illustrates the process of generating helper data and secret key of user W .

Therefore, in the enrollment mapping, the private key can be used for three purposes: (i) masking the codeword index to ensure zero leakage; (ii) generating helper data additional to the masked codeword index and thus enable identifying more users; (iii) generating secret key additional to the codeword

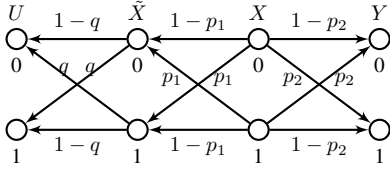


Fig. 4. U , \tilde{X} , X and Y form a Markov chain $U - \tilde{X} - X - Y$.

index and hence increasing the secrecy level. This is consistent with the discussion in Corollary 2.

It can be shown that the labels K_W , M_W , and L_W are close to uniformly distributed. Using this property and the fact that the private key is uniformly distributed, we can show that the generated secret key is close to uniformly distributed and the secrecy leakage rate $\frac{1}{N}I(S_W; J_W)$ becomes sufficiently small for large N . Using again the fact that the private key is uniformly distributed, we can also have that the privacy leakage rate $\frac{1}{N}I(J_W; X^N(W))$ is close to zero.

During the identification phase, the system observes a noisy biometric sequence y^N and receives the user's corresponding private key p . The system checks the database and looks for a unique triple $(\hat{w}, \hat{k}, \hat{m}, \hat{l})$ such that $u^N(\hat{k}, \hat{m}, \hat{l})$ is jointly typical with y^N and $j_{\hat{w}} = (\hat{k}, \hat{m} \oplus \hat{p}, \hat{p})$. The identification mapping also provides an estimate of the secret key as $\hat{s} = (\hat{m}, \hat{l}, \hat{p})$. It can be shown that the identification mapping can reliably identify the user index and guess the secret key if and only if the conditions in Theorems 1 and 2 are satisfied.

Based on the codebook generation method described above, the privacy leakage analysis should be treated carefully due to the noisy enrollment. To bound the privacy leakage rate, we bring in Fano's inequality (which usually is used in the converse) for the achievability proof. This proof technique is reflected in (37).

C. Binary Case Example

In the following we present a binary example that illustrates the trade-off relationships. We assume a binary symmetric source, i.e., the source X is Bernoulli distributed with probability $\frac{1}{2}$. Let the noisy enrollment and observation channels be binary symmetric channels (BSCs) with crossover probability p_1 and p_2 respectively. Moreover, choose U to be an output of a BSC with crossover probability q and input \tilde{X} . We depict the relations among U , \tilde{X} , X and Y in Fig. 4. According to Mrs. Gerber's Lemma [19], we know that if $H(\tilde{X}|U) = h_2(q)$ for some $q \in (0, \frac{1}{2})$, then $H(X|U) \geq h_2(q * p_1)$, where $q * p_1 = q(1 - p_1) + p_1(1 - q)$. It similarly holds $H(Y|U) \geq h_2(q * p_1 * p_2)$, where $q * p_1 * p_2 = (q(1 - p_1) + p_1(1 - q))(1 - p_2) + p_2(1 - q(1 - p_1) - p_1(1 - q))$.

In the following example, the private key rate is fixed as $R_P = 1$, and we choose $q = 0.2$, $p_1 = 0.1$ and $p_2 = 0.15$. The achievable regions of the generated and chosen secret key system are then given as follows:

$$\begin{aligned} \mathcal{R}_g|_{R_P=1} = \{ & (R_I, R_S, R_J) \in \mathbb{R}_+^3 : \\ & R_I + R_S \leq 2 - h_2(q * p_1 * p_2), \\ & R_I - h_2(q * p_1) \leq 1 - h_2(q * p_1 * p_2), \end{aligned}$$

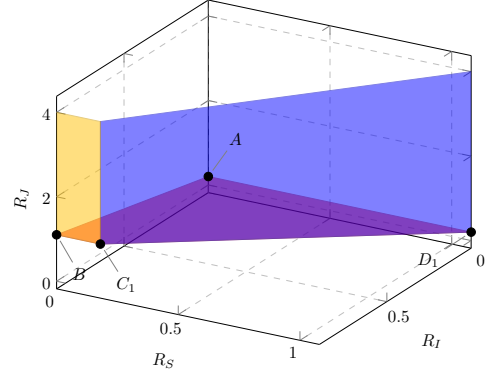


Fig. 5. Generated secret key system subset capacity region boundary with $q = 0.2$, $p_1 = 0.1$ and $p_2 = 0.15$.

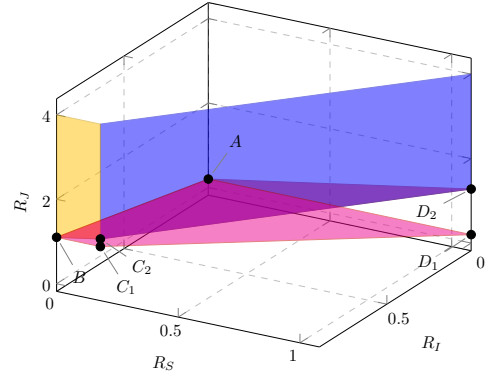


Fig. 6. Chosen secret key system subset capacity region boundary with $q = 0.2$, $p_1 = 0.1$ and $p_2 = 0.15$. The facet ABC_1D_1 is the bottom facet from the region in Fig. 5.

$$\begin{aligned} R_I - h_2(q) &\leq R_J - h_2(q * p_1 * p_2), \\ \text{for some } q &\in (0, \frac{1}{2}), \end{aligned} \quad (13)$$

and

$$\begin{aligned} \mathcal{R}_c|_{R_P=1} = \{ & (R_I, R_S, R_J) \in \mathbb{R}_+^3 : \\ & R_I + R_S \leq 2 - h_2(q * p_1 * p_2), \\ & R_I - h_2(q * p_1) \leq 1 - h_2(q * p_1 * p_2), \\ & R_I - h_2(q) + R_S \leq R_J - h_2(q * p_1 * p_2), \\ & \text{for some } q &\in (0, \frac{1}{2}), \end{aligned} \quad (14)$$

Fig. 5 and Fig. 6 illustrate subsets of capacity regions' boundaries of the identification rate, secret key rate and helper data rate for the generated and chosen secret key systems, respectively. Any rate triples with smaller identification rate, or smaller secret key rate, or larger helper data rate than the rate triples on the boundaries are achievable. From comparison of Fig. 5 and Fig. 6, we can see that when the private key rate is fixed, the generated and chosen secret key system show similar trade-off relationships among the rate triple of the identification rate, the secret key rate and helper data rate. However, we can also observe the difference between these two settings. As the secret key rate increases, the chosen secret key system needs a larger helper data rate. To further compare the generated and the chosen secret key systems, in

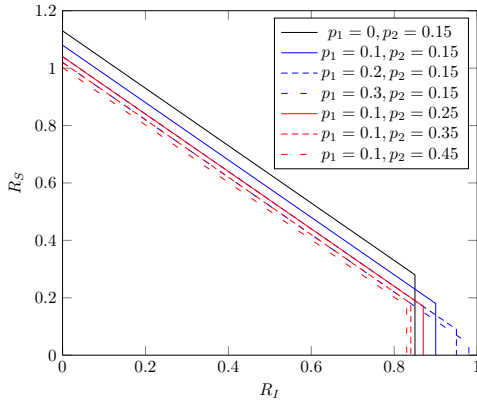


Fig. 7. Identification rate and secret key rate projection of the capacity region boundary with $q = 0.2$.

Fig. 6, we also plot the bottom facet ABC_1D_1 of the region boundary in Fig. 5. Observe the bottom facet ABC_1D_1 of the generated secret key system and the bottom facet ABC_2D_2 of the chosen secret key system in Fig. 6, we see that for the same identification rate and the secret key rate, the chosen secret key system requires larger minimum achievable helper data rate. Therefore, in order to enroll and identify the same number of users as well as employing secret keys of the same rate, the chosen secret key system needs more storage to store the helper data than the generated secret key system.

To further illustrate the trade-off between the identification and the secret key rate, in Fig. 7 we depict the identification rate and secret key rate plane section of the capacity region boundary. Moreover, as in Corollary 3, if helper data rate is not considered, the generated and chosen secret key systems achieve the same region, therefore we do not distinguish whether the secret key is generated or chosen in the following discussion. First, we fix $q = 0.2$ and let p_1 and p_2 vary to investigate how the projection changes. We also include the case that the enrollment channel is noise-free, i.e., $p_1 = 0$, which is depicted with the black curve. For noisy enrollment channels, we see that the secret key rate is smaller while the identification rate is enlarged compared with noise-free enrollment channel. Moreover, when we fix $p_2 = 0.15$ and vary p_1 , i.e., the enrollment channel changes, we can see that the maximal achievable secret key rate decreases while the identification rate increases as the enrollment channel becomes more noisy. If the enrollment channel is fixed, in this example we have $p_1 = 0.1$, and the observation channel quality is varied, then we can observe that the maximal achievable identification rate and secret key rate both decrease as the channel is more noisy.

IV. CONCLUSION

The fundamental trade-off for privacy-preserving identification systems with noisy enrollment has been characterized. It shows that for reliable identification and authentication, as well as close to zero privacy leakage rate and secrecy leakage rate, under noisy enrollment, certain helper data rate and private key rate are necessary. Noisy enrollment is a significant assumption especially for biometric systems. We considered

two variations of the proposed biometric identification system, where the secret keys are assumed to be generated and chosen, respectively. We illustrated capacity region with a binary example. The results show that, a higher minimum helper data rate is necessary if the secret key is chosen rather than generated. If there is no restriction on the helper data rate, i.e., the database can be arbitrarily large, the generated secret key and chosen secret key systems have the same capacity region.

APPENDIX A PROOF OF THEOREM 1

1) Achievability Part: Fix a conditional probability distribution $P_{U|\tilde{X}}$. Thus we have the joint p.m.f. $P_{U\tilde{X}XY} = Q_X Q_{\tilde{X}|X} Q_{Y|X} P_{U|\tilde{X}}$. Let $M_I = 2^{NR_I}$ denote the number of enrolled users. Further, fix small enough $\epsilon > 0$, δ_1 and δ such that $\delta_1 > 2\delta > 0$.

Codebook generation: For fixed δ and δ_1 , pick a rate pair tuple $(R_I, R_S, R_P, R_J) \in \mathcal{R}_g$ such that $R_P \geq R_I + I(U; X) - I(U; Y) + 2\delta_1$, $R_J = R_I + I(U; \tilde{X}) - I(U; Y) + 3\delta_1$ and $R_S = R_P - R_I + I(U; Y) - 2\delta_1$. Randomly and independently generate $2^{N(I(U; \tilde{X}) + \delta_1)}$ i.i.d. codewords $u^N(k, m, l)$ according to $\prod_{i=1}^N P_{U|u_i}$. We distribute the codewords uniformly at random into $2^{N(I(U; \tilde{X}) - I(U; X) + \delta_1)}$ bins indexed by k , and each bin consists of $2^{NI(U; X)}$ codewords. We further distribute the codewords in a bin into $2^{N(I(U; X) - I(U; Y) + \delta)}$ subbins indexed by m , and each subbin consists of $2^{N(I(U; Y) - \delta)}$ codewords indexed by l .

Enrollment: For each user $w \in [1 : 2^{NR_I}]$, a codeword $u^N(k_w, m_w, l_w)$ is looked for such that $(u^N(k_w, m_w, l_w), \tilde{x}^N(w)) \in \mathcal{T}_\epsilon^N$. If no such (k_w, m_w, l_w) exists, an index triple (k_w, m_w, l_w) is randomly drawn from $[1 : 2^{NR_1}] \times [1 : 2^{NR_2}] \times [1 : 2^{NR_3}]$, where $R_1 = I(U; \tilde{X}) - I(U; X) + \delta_1$, $R_2 = I(U; X) - I(U; Y) + \delta_1$ and $R_3 = I(U; Y) - \delta_1$. If there are more than one such index triple (k_w, m_w, l_w) , one of them is selected uniformly at random. The private key of the user p_w is divided into three parts, \bar{p}_w , \check{p}_w and \tilde{p}_w , such that $p_w = (\bar{p}_w, \check{p}_w, \tilde{p}_w)$, $\bar{p}_w \in [1 : 2^{N(R_I + \delta_1)}]$, $\check{p}_w \in [1 : 2^{N(I(U; X) - I(U; Y) + \delta_1)}]$ and $\tilde{p}_w \in [1 : 2^{N(R_P - (R_I + \delta_1) - (I(U; X) - I(U; Y) + \delta_1))}]$. A helper data index is generated as $j_w = (k_w, m_w \oplus \check{p}_w, \bar{p}_w)$, which is stored in the database at the location w . Lastly, a secret key s_w is generated as $s_w = (m_w, l_w, \tilde{p}_w)$.

Identification and Authentication: After user w is observed, the observation y^N and the user's private key $p = (\bar{p}, \check{p}, \tilde{p})$ are provided to the system. The identification mapping searches for a unique index tuple $(\hat{w}, \hat{k}, \hat{m}, \hat{l})$ such that $j_{\hat{w}} = (\hat{k}, \hat{m} \oplus \check{p}, \bar{p})$ and $(u^N(\hat{k}, \hat{m}, \hat{l}), y^N) \in \mathcal{T}_\epsilon^N$. An estimate of the secret key is given by $\hat{s} = (\hat{m}, \hat{l}, \tilde{p})$. If there is no such index triple or more than one index tuple, an error is declared.

Error Events Analysis: Assume that the user with index W is observed. Let (K_W, M_W, L_W) be the corresponding codeword index triple determined by the enrollment mapping. Let P_W, S_W and J_W be the actual private key, secret key, and helper data stored in the database. Let \hat{W} and \hat{S} denote the estimated user index and secret key, respectively. We use \mathcal{C} to denote the codebook and define the following events:

$$A(\hat{w}, \hat{k}, \hat{m}, \hat{l}) = \{J_{\hat{w}} = (\hat{k}, \hat{m} \oplus \check{P}_W, \bar{P}_W),$$

$$(U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N \quad (15)$$

We have the following error events:

$$\begin{aligned} \mathcal{E}_1 &= \{(U^N(k, m, l), \tilde{X}^N(W)) \notin \mathcal{T}_\epsilon^N, \\ &\quad \forall (k, m, l) \in [1 : 2^{NR_1}] \times [1 : 2^{NR_2}] \times [1 : 2^{NR_3}]\}, \\ \mathcal{E}_2 &= \{(U^N(K_W, M_W, L_W), Y^N) \notin \mathcal{T}_\epsilon^N\}, \\ \mathcal{E}_3 &= \bigcup_{\hat{w} \neq W} \bigcup_{\hat{k}} \bigcup_{\hat{m}} \bigcup_{\hat{l}} A(\hat{w}, \hat{k}, \hat{m}, \hat{l}), \\ \mathcal{E}_4 &= \bigcup_{\hat{l} \neq L_W} A(W, K_W, M_W, \hat{l}). \end{aligned} \quad (16)$$

The first error event corresponds to the enrollment error, i.e., there is no codeword jointly typical with the noisy biometric sequence. The error event \mathcal{E}_2 is an error in the identification phase that the true codeword is not jointly typical with the observation Y^N . The error event \mathcal{E}_3 denote the identification error that there exists another user index \hat{w} that fulfills all conditions. The error event \mathcal{E}_4 is the authentication error that the estimated user index is correct while the estimated secret key is not the same with the true one. Taking the codeword indices into account, the identification error event \mathcal{E}_3 can be covered by the following five error events:

$$\begin{aligned} \mathcal{E}_{31} &= \bigcup_{\hat{w} \neq W} \bigcup_{\hat{k} \neq K_W} \bigcup_{\hat{m} \neq M_W} \bigcup_{\hat{l}} A(\hat{w}, \hat{k}, \hat{m}, \hat{l}), \\ \mathcal{E}_{32} &= \bigcup_{\hat{w} \neq W} \bigcup_{\hat{k} \neq K_W} \bigcup_{\hat{l}} A(\hat{w}, \hat{k}, M_W, \hat{l}), \\ \mathcal{E}_{33} &= \bigcup_{\hat{w} \neq W} \bigcup_{\hat{m} \neq M_W} \bigcup_{\hat{l}} A(\hat{w}, K_W, \hat{m}, \hat{l}), \\ \mathcal{E}_{34} &= \bigcup_{\hat{w} \neq W} \bigcup_{\hat{l} \neq L_W} A(\hat{w}, K_W, M_W, \hat{l}), \\ \mathcal{E}_{35} &= \bigcup_{\hat{w} \neq W} A(\hat{w}, K_W, M_W, L_W). \end{aligned} \quad (17)$$

If none of above error events happens, the identification is successful, i.e., the estimated user index is the same with the true user index. We obtain that $\mathcal{E}_3 = \mathcal{E}_{31} \cup \mathcal{E}_{32} \cup \mathcal{E}_{33} \cup \mathcal{E}_{34} \cup \mathcal{E}_{35}$.

As for the authentication error, note that due to the identification and authentication mapping, the estimated user index is correct $\hat{w} = W$ implies that the estimated helper data $J_{\hat{w}} = (\hat{k}, \hat{m} \oplus \check{P}_W, \bar{P}_W)$ and the true user's helper data $J_W = (K_W, M_W \oplus \check{P}_W, \bar{P}_W)$ are the same. Further, from $J_{\hat{w}} = J_W$, we can obtain that $\hat{k} = K_W$ and $\hat{m} = M_W$ always hold. Therefore, under the condition that the estimated user index is correct, i.e., $\hat{w} = W$, the authentication error happens when $\hat{l} \neq L_W$.

If none of the events occurs, the identification and authentication will be successful. Thus, we have the following error event

$$\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_{31} \cup \mathcal{E}_{32} \cup \mathcal{E}_{33} \cup \mathcal{E}_{34} \cup \mathcal{E}_{35} \cup \mathcal{E}_4. \quad (18)$$

Since $R_1 + R_2 + R_3 > I(U; \tilde{X})$, we obtain that $\Pr(\mathcal{E}_1) \rightarrow 0$ as $N \rightarrow \infty$ due to the covering lemma [20, Lemma 3.3].

The event $\{\tilde{X}^N(W) = \tilde{x}^N, U^N = u^N\}$ implies $Y^N \sim \prod_{i=1}^N P_{Y|\tilde{X}}(\cdot|\tilde{x}_i)$. By the Markov lemma [20, p.27], we obtain that $\Pr(\mathcal{E}_1^c \cap \mathcal{E}_2) \rightarrow 0$ as $N \rightarrow \infty$.

In the following analysis of bounding the error events probabilities, we use the arguments as follows: (a): the union bound; (b): expanding helper data $J_{\hat{w}} = (K_{\hat{w}}, M_{\hat{w}} \oplus \check{P}_{\hat{w}}, \bar{P}_{\hat{w}})$; (c): $K_{\hat{w}}, M_{\hat{w}} \oplus \check{P}_W$ and \bar{P}_W are mutually independent as \bar{P}_W and \bar{P}_W are uniformly distributed; (d): $\tilde{X}^N(w)$ and Y^N are either jointly typical with the same codeword or two different codewords; (e): Lemma 1 in [21]; (f): $\check{P}_{\hat{w}}$ is independent of $(\hat{m}, \check{P}_1, M_{\hat{w}})$.

Define the event $\mathcal{E}_{1,\hat{w}}$ similarly as \mathcal{E}_1 by replacing W with w , i.e.,

$$\begin{aligned} \mathcal{E}_{1,\hat{w}} &= \{(U^N(k, m, l), \tilde{X}^N(\hat{w})) \notin \mathcal{T}_\epsilon^N, \\ &\quad \forall (k, m, l) \in [1 : 2^{NR_1}] \times [1 : 2^{NR_2}] \times [1 : 2^{NR_3}]\}. \end{aligned} \quad (19)$$

The probability $\Pr(\mathcal{E}_{31}|W = 1)$ can be bounded as

$$\begin{aligned} &\Pr(\mathcal{E}_{31}|W = 1) \\ &= \Pr\{J_{\hat{w}} = (\hat{k}, \hat{m} \oplus \check{P}_1, \bar{P}_1), (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\ &\quad \text{for some } \hat{w} \neq 1, \hat{k} \neq K_1, \hat{m} \neq M_1, \text{ and } \hat{l}|W = 1\} \\ &\stackrel{(a)}{\leq} \sum_{\hat{w} \neq 1} \Pr\{J_{\hat{w}} = (\hat{k}, \hat{m} \oplus \check{P}_1, \bar{P}_1), (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\ &\quad \text{for some } \hat{k} \neq K_1, \hat{m} \neq M_1, \text{ and } \hat{l}|W = 1\} \\ &\leq \sum_{\hat{w} \neq 1} \left(\Pr\{J_{\hat{w}} = (\hat{k}, \hat{m} \oplus \check{P}_1, \bar{P}_1), \right. \\ &\quad \left. (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \right. \\ &\quad \left. \text{for some } \hat{k} \neq K_1, \hat{m} \neq M_1, \text{ and } \hat{l}, \mathcal{E}_{1,\hat{w}}^c|W = 1\} \right. \\ &\quad \left. + \Pr\{\mathcal{E}_{1,\hat{w}}|W = 1\} \right), \end{aligned} \quad (20)$$

Due to the covering lemma [20, Lemma 3.3], we have that

$$\Pr\{\mathcal{E}_{1,\hat{w}}|W = 1\} \rightarrow 0, \quad (21)$$

double exponentially. Therefore, the second term in (20) goes to 0. As for the first term, without loss of generality, we condition on $B = \{W = 1, K_1 = 1, M_1 = 1, L_1 = 1, \bar{P}_1 = 1, \check{P}_1 = 1, \bar{P}_1 = 1\}$. Then we have that

$$\begin{aligned} &\Pr\{J_{\hat{w}} = (\hat{k}, \hat{m} \oplus \check{P}_1, \bar{P}_1), (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\ &\quad \text{for some } \hat{k} \neq K_1, \hat{m} \neq M_1, \text{ and } \hat{l}, \mathcal{E}_{1,\hat{w}}^c|B\} \\ &= \Pr\{J_{\hat{w}} = (\hat{k}, \hat{m} \oplus 1, 1), (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\ &\quad \text{for some } \hat{k} \neq 1, \hat{m} \neq 1, \text{ and } \hat{l}, \mathcal{E}_{1,\hat{w}}^c|B\} \\ &\stackrel{(a)}{\leq} \sum_{\hat{k} \neq 1} \sum_{\hat{m} \neq 1} \sum_{\hat{l}} \Pr\{J_{\hat{w}} = (\hat{k}, \hat{m} \oplus 1, 1), \\ &\quad (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \mathcal{E}_{1,\hat{w}}^c|B\} \\ &\stackrel{(b)}{=} \sum_{\hat{k} \neq 1} \sum_{\hat{m} \neq 1} \sum_{\hat{l}} \Pr\{K_{\hat{w}} = \hat{k}, M_{\hat{w}} \oplus \check{P}_{\hat{w}} = \hat{m} \oplus 1, \\ &\quad \bar{P}_{\hat{w}} = 1, (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \mathcal{E}_{1,\hat{w}}^c|B\} \\ &\leq \sum_{\hat{k} \neq 1} \sum_{\hat{m} \neq 1} \sum_{\hat{l}} \Pr\{(U^N(\hat{k}, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\ &\quad \text{for some } m' \text{ and } l', m' \oplus \check{P}_{\hat{w}} = \hat{m} \oplus 1, \\ &\quad \bar{P}_{\hat{w}} = 1, (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N|B\} \end{aligned}$$

$$\begin{aligned}
&\stackrel{(c)}{\leq} \sum_{\hat{k} \neq 1} \sum_{\hat{m} \neq 1} \sum_{\hat{l}} \sum_{m'} \sum_{l'} \\
&\quad \Pr\{(U^N(\hat{k}, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N | B\} \\
&\quad \times \Pr\{\bar{P}_{\hat{w}} = 1\} \Pr\{\check{P}_{\hat{w}} = \hat{m} \oplus 1 \ominus m' | B\}. \tag{22}
\end{aligned}$$

Consider and bound the following term:

$$\begin{aligned}
&\sum_{m'} \sum_{l'} \Pr\{(U^N(\hat{k}, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N | B\} \\
&\stackrel{(d)}{=} \Pr\{(U^N(\hat{k}, \hat{m}, \hat{l}), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N | B\} \\
&\quad + \sum_{(m', l') \neq (\hat{m}, \hat{l})} \Pr\{(U^N(\hat{k}, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad (U^N(\hat{k}, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N | B\} \\
&= \sum_{u^N} \sum_{y^N} \Pr\{U^N(\hat{k}, \hat{m}, \hat{l}) = u^N, Y^N = y^N | B\} \\
&\quad \underbrace{\Pr\{(u^N, y^N) \in \mathcal{T}_\epsilon^N | U^N(\hat{k}, \hat{m}, \hat{l}) = u^N, Y^N = y^N, B\}}_{\stackrel{(e)}{\leq} (1+\hat{\epsilon})2^{-N(I(U;Y)-\delta)}} \\
&\quad \underbrace{\Pr\{(u^N, \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N | U^N(\hat{k}, \hat{m}, \hat{l}) = u^N, B\}}_{\leq 2^{-N(I(U;\tilde{X})-\delta)}} \\
&\quad + \sum_{(m', l') \neq (\hat{m}, \hat{l})} \sum_{u_1^N} \sum_{u_2^N} \\
&\quad \Pr\{U^N(\hat{k}, \hat{m}, \hat{l}) = u_2^N, Y^N = y^N | B\} \\
&\quad \underbrace{\Pr\{(u_2^N, y^N) \in \mathcal{T}_\epsilon^N | U^N(\hat{k}, \hat{m}, \hat{l}) = u_2^N, Y^N = y^N, B\}}_{\stackrel{(e)}{\leq} (1+\hat{\epsilon})2^{-N(I(U;Y)-\delta)}} \\
&\quad \times \Pr\{U^N(\hat{k}, m', l') = u_1^N | U^N(\hat{k}, \hat{m}, \hat{l}) = u_2^N, \\
&\quad Y^N = y^N, B\} \\
&\quad \times \Pr\{(u_1^N, \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N | U^N(\hat{k}, m', l') = u_1^N, \\
&\quad \underbrace{U^N(\hat{k}, \hat{m}, \hat{l}) = u_2^N, Y^N = y^N, B\}}_{\leq 2^{-N(I(U;\tilde{X})-\delta)}} \\
&\leq (1+\hat{\epsilon})2^{-N(I(U;Y)-\delta)}2^{-N(I(U;\tilde{X})-\delta)} \\
&\quad + 2^{N(I(U;X)-I(U;Y)+\delta_1)}2^{N(I(U;Y)-\delta_1)} \\
&\quad \times 2^{-N(I(U;\tilde{X})-\delta)}(1+\hat{\epsilon})2^{-N(I(U;Y)-\delta)} \\
&= (1+\hat{\epsilon})(2^{-NI(U;X)} + 1)2^{-N(I(U;\tilde{X})-I(U;X)+I(U;Y)-2\delta)}. \tag{23}
\end{aligned}$$

where $\hat{\epsilon} > 0$ is a fixed number.

Consider the following term in (23) have that

$$\begin{aligned}
&\Pr\{\bar{P}_{\hat{w}} = 1\} \Pr\{\check{P}_{\hat{w}} = \hat{m} \oplus 1 \ominus m' | B\} \\
&\stackrel{(f)}{\leq} 2^{-N(R_I+\delta_1)}2^{-N(I(U;X)-I(U;Y)+\delta_1)}. \tag{24}
\end{aligned}$$

Combining the above results and define $\delta'' \rightarrow 0$, the probability of $\Pr(\mathcal{E}_{31}|W=1)$ can be bounded as follows

$$\Pr(\mathcal{E}_{31}|W=1)$$

$$\begin{aligned}
&\leq \sum_{\hat{w} \neq 1} \sum_{\hat{k} \neq K_1} \sum_{\hat{m} \neq M_1} \sum_{\hat{l}} (1+\hat{\epsilon})(2^{-NI(U;X)} + 1) \\
&\quad \times 2^{-N(I(U;\tilde{X})-I(U;X)+I(U;Y)-2\delta)} \\
&\quad \times 2^{-N(R_I+\delta_1)}2^{-N(I(U;X)-I(U;Y)+\delta_1)} + \delta'' \\
&\leq 2^{NR_I}2^{N(I(U;\tilde{X})-I(U;X)+\delta_1)}2^{N(I(U;X)-I(U;Y)+\delta_1)} \\
&\quad \times 2^{N(I(U;Y)-\delta_1)}(1+\hat{\epsilon})(2^{-NI(U;X)} + 1) \\
&\quad \times 2^{-N(I(U;\tilde{X})-I(U;X)+I(U;Y)-2\delta)} \\
&\quad \times 2^{-N(R_I+\delta_1)}2^{-N(I(U;X)-I(U;Y)+\delta_1)} + \delta'' \\
&= (1+\hat{\epsilon})(2^{-NI(U;X)} + 1)2^{-N(\delta_1-2\delta)} + \delta'', \tag{25}
\end{aligned}$$

where $\hat{\epsilon} > 0$ is a fixed number and $\delta'' \rightarrow 0$. Therefore $\Pr(\mathcal{E}_{31}|W=1) \rightarrow 0$ as $N \rightarrow \infty$.

Follow similar analysis in bounding $\Pr(\mathcal{E}_{32}|W=1)$, we can obtain that the probability of $\mathcal{E}_{32}|W=1$ can be bounded as follows:

$$\begin{aligned}
&\Pr(\mathcal{E}_{32}|W=1) \\
&= \Pr\{J_{\hat{w}} = (\hat{k}, M_1 \oplus \check{P}_1, \bar{P}_1), (U^N(\hat{k}, M_1, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } \hat{w} \neq 1, \hat{k} \neq K_1, \text{ and } \hat{l}|W=1\} \\
&\stackrel{(*)}{\leq} \sum_{\hat{w} \neq 1} \Pr\{J_{\hat{w}} = (\hat{k}, M_1 \oplus \check{P}_1, \bar{P}_1), \\
&\quad (U^N(\hat{k}, M_1, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } \hat{k} \neq K_1 \text{ and } \hat{l}, \mathcal{E}_{1,\hat{w}}^c|W=1\} + \delta'' \\
&\stackrel{(b)}{\leq} \sum_{\hat{w} \neq 1} \sum_{\hat{k} \neq K_1} \sum_{\hat{l}} \Pr\{(U^N(\hat{k}, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N \\
&\quad \text{for some } m' \text{ and } l', M_{\hat{w}} \oplus \check{P}_{\hat{w}} = M_1 \oplus \check{P}_1, \bar{P}_{\hat{w}} = \bar{P}_1, \\
&\quad (U^N(\hat{k}, M_1, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N|W=1\} + \delta'' \\
&\stackrel{(c)}{\leq} \sum_{\hat{w} \neq 1} \sum_{\hat{k} \neq K_1} \sum_{\hat{l}} \sum_{m'} \sum_{l'} \Pr\{(U^N(\hat{k}, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad (U^N(\hat{k}, M_1, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N|W=1\} \\
&\quad \times \Pr\{\bar{P}_{\hat{w}} = \bar{P}_1\} \Pr\{\check{P}_{\hat{w}} = M_1 \oplus \check{P}_1 \ominus m' | W=1\} + \delta'' \\
&\stackrel{(**)}{\leq} \sum_{\hat{w} \neq 1} \sum_{\hat{k} \neq K_1} \sum_{\hat{l}} (1+\hat{\epsilon})(2^{-NI(U;X)} + 1) \\
&\quad \times 2^{-N(I(U;\tilde{X})-I(U;X)+I(U;Y)-2\delta)} \\
&\quad \times 2^{-N(R_I+\delta_1)}2^{-N(I(U;X)-I(U;Y)+\delta_1)} + \delta'' \\
&\leq 2^{NR_I}2^{N(I(U;\tilde{X})-I(U;X)+\delta_1)}2^{N(I(U;Y)-\delta_1)} \\
&\quad (1+\hat{\epsilon})(2^{-NI(U;X)} + 1)2^{-N(I(U;\tilde{X})-I(U;X)+I(U;Y)-2\delta)} \\
&\quad \times 2^{-N(R_I+\delta_1)}2^{-N(I(U;X)-I(U;Y)+\delta_1)} + \delta'' \\
&= (1+\hat{\epsilon})(2^{-NI(U;X)} + 1)2^{-N(I(U;X)-I(U;Y)+2\delta_1-2\delta)} + \delta''. \tag{26}
\end{aligned}$$

where $\hat{\epsilon} > 0$ is a fixed number and $\delta'' \rightarrow 0$; $(*)$ follows the analysis in (20) and (21); $(**)$ follows the analysis in (23) and (24). Therefore $\Pr(\mathcal{E}_{32}|W=1) \rightarrow 0$ as $N \rightarrow \infty$.

Consider $\Pr(\mathcal{E}_{33}|W=1)$, when N is sufficiently large, then we obtain that

$$\Pr(\mathcal{E}_{33}|W=1)$$

$$\begin{aligned}
&= \Pr\{J_{\hat{w}} = (K_1, \hat{m} \oplus \check{P}_1, \bar{P}_1), (U^N(K_1, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } \hat{w} \neq 1, \hat{m} \neq K_1, \text{ and } \hat{l}|W=1\} \\
&\stackrel{(\star)}{\leq} \sum_{\hat{w} \neq 1} \Pr\{J_{\hat{w}} = (K_1, \hat{m} \oplus \check{P}_1, \bar{P}_1), \\
&\quad (U^N(K_1, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } \hat{m} \neq M_1 \text{ and } \hat{l}, \mathcal{E}_{1,\hat{w}}^c|W=1\} + \delta'' \\
&\stackrel{(b)}{\leq} \sum_{\hat{w} \neq 1} \sum_{\hat{m} \neq M_1} \sum_{\hat{l}} \Pr\{(U^N(K_1, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } m' \text{ and } l', m' \oplus \check{P}_{\hat{w}} = \hat{m} \oplus \check{P}_1, \\
&\quad \bar{P}_{\hat{w}} = \bar{P}_1, (U^N(1, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \mathcal{E}_{1,\hat{w}}^c|W=1\} + \delta'' \\
&\stackrel{(c)}{\leq} \sum_{\hat{w} \neq K_1} \sum_{\hat{m} \neq M_1} \sum_{\hat{l}} \sum_{m'} \sum_{l'} \\
&\quad \Pr\{(U^N(K_1, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad (U^N(K_1, \hat{m}, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N|W=1\} \\
&\quad \times \Pr\{\bar{P}_{\hat{w}} = \bar{P}_1\} \Pr\{\check{P}_{\hat{w}} = \hat{m} \oplus \check{P}_1 \ominus m'|W=1\} + \delta'' \\
&\stackrel{(\star\star)}{\leq} \sum_{\hat{w} \neq 1} \sum_{\hat{m} \neq M_1} \sum_{\hat{l}} (1 + \hat{\epsilon})(2^{-NI(U;X)} + 1) \\
&\quad \times 2^{-N(I(U;\tilde{X}) - I(U;X) + I(U;Y) - 2\delta)} \\
&\quad \times 2^{-N(R_I + \delta_1)} 2^{-N(I(U;X) - I(U;Y) + \delta_1)} + \delta'' \\
&\leq 2^{NR_I} 2^{N(I(U;X) - I(U;Y) + \delta_1)} 2^{N(I(U;Y) - \delta_1)} \\
&\quad (1 + \hat{\epsilon})(2^{-NI(U;X)} + 1) 2^{-N(I(U;\tilde{X}) - I(U;X) + I(U;Y) - 2\delta)} \\
&\quad \times 2^{-N(R_I + \delta_1)} 2^{-N(I(U;X) - I(U;Y) + \delta_1)} + \delta'' \\
&= (1 + \hat{\epsilon})(2^{-NI(U;X)} + 1) 2^{-N(I(U;\tilde{X}) - I(U;Y) + 2\delta_1 - 2\delta)} + \delta''. \tag{27}
\end{aligned}$$

where $\hat{\epsilon} > 0$ is a fixed number and $\delta'' \rightarrow 0$; (\star) follows the analysis in (20) and (21); $(\star\star)$ follows the analysis in (23) and (24). Therefore $\Pr(\mathcal{E}_{33}|W=1) \rightarrow 0$ as $N \rightarrow \infty$.

As for the probability of the error event $\Pr\{\mathcal{E}_{34}|W=1\}$, for sufficiently large N , we have

$$\begin{aligned}
&\Pr(\mathcal{E}_{34}|W=1) \\
&= \Pr\{J_{\hat{w}} = (K_1, M_1 \oplus \check{P}_1, \bar{P}_1), \\
&\quad (U^N(K_1, M_1, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } \hat{w} \neq 1 \text{ and } \hat{l} \neq L_1|W=1\} \\
&\stackrel{(\star)}{\leq} \sum_{\hat{w} \neq 1} \Pr\{J_{\hat{w}} = (K_1, M_1 \oplus \check{P}_1, \bar{P}_1), \\
&\quad (U^N(K_1, M_1, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } \hat{l} \neq L_1, \mathcal{E}_{1,\hat{w}}^c|W=1\} + \delta'' \\
&\stackrel{(b)}{\leq} \sum_{\hat{w} \neq 1} \sum_{\hat{l} \neq L_1} \Pr\{(U^N(K_1, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } m' \text{ and } l', m' \oplus \check{P}_{\hat{w}} = M_1 \oplus \check{P}_1, \\
&\quad \bar{P}_{\hat{w}} = \bar{P}_1, (U^N(K_1, M_1, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N|W=1\} + \delta'' \\
&\stackrel{(c)}{\leq} \sum_{\hat{w} \neq 1} \sum_{\hat{l} \neq L_1} \sum_{m'} \sum_{l'} \Pr\{(U^N(K_1, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N,
\end{aligned}$$

$$\begin{aligned}
&\quad (U^N(K_1, M_1, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N|W=1\} \\
&\quad \Pr\{\bar{P}_{\hat{w}} = \bar{P}_1\} \Pr\{\check{P}_{\hat{w}} = \hat{m} \oplus \check{P}_1 \ominus m'|W=1\} + \delta'' \\
&\stackrel{(\star\star)}{\leq} \sum_{\hat{w} \neq 1} \sum_{\hat{l} \neq L_1} (1 + \hat{\epsilon})(2^{-NI(U;X)} + 1) \\
&\quad \times 2^{-N(I(U;\tilde{X}) - I(U;X) + I(U;Y) - 2\delta)} \\
&\quad \times 2^{-N(R_I + \delta_1)} 2^{-N(I(U;X) - I(U;Y) + \delta_1)} + \delta'' \\
&\leq 2^{NR_I} 2^{N(I(U;Y) - \delta_1)} (1 + \hat{\epsilon})(2^{-NI(U;X)} + 1) \\
&\quad \times 2^{-N(I(U;\tilde{X}) - I(U;X) + I(U;Y) - 2\delta)} \\
&\quad \times 2^{-N(R_I + \delta_1)} 2^{-N(I(U;X) - I(U;Y) + \delta_1)} + \delta'' \\
&= (1 + \hat{\epsilon})(2^{-NI(U;X)} + 1) 2^{-N(I(U;\tilde{X}) - I(U;Y) + 3\delta_1 - 2\delta)} + \delta''. \tag{28}
\end{aligned}$$

where $\hat{\epsilon} > 0$ is a fixed number and $\delta'' \rightarrow 0$; (\star) follows the analysis in (20) and (21); $(\star\star)$ follows the analysis in (23) and (24). Therefore $\Pr(\mathcal{E}_{34}|W=1) \rightarrow 0$ as $N \rightarrow \infty$.

The probability $\Pr\{\mathcal{E}_{35}|B\}$ can be bounded as follows for sufficiently large enough N

$$\begin{aligned}
&\Pr(\mathcal{E}_{35}|W=1) \\
&= \Pr\{J_{\hat{w}} = (K_1, M_1 \oplus \check{P}_1, \bar{P}_1), \\
&\quad (U^N(K_1, M_1, L_1), Y^N) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } \hat{w} \neq 1|\mathcal{E}_1^c, W=1\} \\
&\stackrel{(\star)}{\leq} \sum_{\hat{w} \neq 1} \Pr\{J_{\hat{w}} = (K_1, M_1 \oplus \check{P}_1, \bar{P}_1), \\
&\quad (U^N(K_1, M_1, L_1), Y^N) \in \mathcal{T}_\epsilon^N, \mathcal{E}_{1,\hat{w}}^c|W=1\} + \delta'' \\
&\stackrel{(b)}{\leq} \sum_{\hat{w} \neq 1} \Pr\{(U^N(K_1, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad \text{for some } m' \text{ and } l', m' \oplus \check{P}_{\hat{w}} = M_1 \oplus \check{P}_1, \bar{P}_{\hat{w}} = \bar{P}_1, \\
&\quad (U^N(K_1, M_1, L_1), Y^N) \in \mathcal{T}_\epsilon^N, \mathcal{E}_{1,\hat{w}}^c|W=1\} + \delta'' \\
&\stackrel{(c)}{\leq} \sum_{\hat{w} \neq 1} \sum_{m'} \sum_{l'} \Pr\{(U^N(K_1, m', l'), \tilde{X}^N(\hat{w})) \in \mathcal{T}_\epsilon^N, \\
&\quad (U^N(K_1, M_1, L_1), Y^N) \in \mathcal{T}_\epsilon^N, \mathcal{E}_{1,\hat{w}}^c|W=1\} \\
&\quad \Pr\{\bar{P}_{\hat{w}} = \bar{P}_1\} \Pr\{\check{P}_{\hat{w}} = \hat{m} \oplus \check{P}_1 \ominus m'|W=1\} + \delta'' \\
&\stackrel{(\star\star)}{\leq} \sum_{\hat{w} \neq 1} (1 + \hat{\epsilon})(2^{-NI(U;X)} + 1) \\
&\quad \times 2^{-N(I(U;\tilde{X}) - I(U;X) + I(U;Y) - 2\delta)} \\
&\quad \times 2^{-N(R_I + \delta_1)} 2^{-N(I(U;X) - I(U;Y) + \delta_1)} + \delta'' \\
&\leq 2^{NR_I} (1 + \hat{\epsilon})(2^{-NI(U;X)} + 1) \\
&\quad \times 2^{-N(I(U;\tilde{X}) - I(U;X) + I(U;Y) - 2\delta)} \\
&\quad \times 2^{-N(R_I + \delta_1)} 2^{-N(I(U;X) - I(U;Y) + \delta_1)} + \delta'' \\
&= (1 + \hat{\epsilon})(2^{-NI(U;X)} + 1) 2^{-N(I(U;\tilde{X}) + 2\delta_1 - \delta)} + \delta'', \tag{29}
\end{aligned}$$

where $\hat{\epsilon} > 0$ is a fixed number and $\delta'' \rightarrow 0$; (\star) follows the analysis in (20) and (21); $(\star\star)$ follows the analysis in (23) and (24). Therefore we obtain that $\Pr(\mathcal{E}_{35}|W=1) \rightarrow 0$ as $N \rightarrow \infty$.

As for the probability of the last error event $\Pr\{\mathcal{E}_4|B\}$, for sufficiently large enough N , we have

$$\begin{aligned}\Pr(\mathcal{E}_4|W=1) &\leq \sum_{\hat{l} \neq L_1} \Pr\{(U^N(K_1, M_1, \hat{l}), Y^N) \in \mathcal{T}_\epsilon^N|B\} \\ &\stackrel{(e)}{\leq} 2^{N(I(U;Y)-\delta_1)}(1+\hat{\epsilon})2^{-N(I(U;Y)-\delta)} \\ &= (1+\hat{\epsilon})2^{N(\delta-\delta_1)},\end{aligned}\quad (30)$$

where $\hat{\epsilon} > 0$ is a fixed number. Therefore $\Pr(\mathcal{E}_4|W=1) \rightarrow 0$ as $N \rightarrow \infty$.

As a consequence, due to the union bound, we obtain that

$$\Pr(\mathcal{E}) \rightarrow 0, \text{ as } N \rightarrow \infty. \quad (31)$$

Therefore, from the random coding argument, we can conclude that there exists a suitable codebook \mathcal{C} such that (31) holds.

Uniformity of Secret Key Analysis: The proof of uniformity of secret key is similar to the analysis in [18]. Due to symmetry, it is sufficient to consider the case when the user index is $W = 1$. For any sequence $u^N(k_1, m_1, l_1)$ where $k \in [1 : 2^{NR_1}]$, $m \in [1 : 2^{NR_2}]$, and $l \in [1 : 2^{NR_3}]$, we have

$$\Pr\{\tilde{X}^N(1) \in \mathcal{T}_\epsilon^N(\tilde{X}|u^N(k_1, m_1, l_1))\} \leq 2^{-N(I(U;\tilde{X})-\delta)}. \quad (32)$$

Let E be a random variable such that $E = 0$ when there exists a codeword index triple (k_1, m_1, l_1) for user $W = 1$ satisfying $(X^N(1), u^N(k_1, m_1, l_1)) \in \mathcal{T}_\epsilon^N$. If no such codeword index exists, then $E = 1$. Let γ denote $\Pr(E = 1)$. According to the error events analysis, we obtain that $\gamma \rightarrow 0$ as $N \rightarrow \infty$.

Similar to [9, (105)-(106)], consider the following joint entropy

$$\begin{aligned}H(K_1, M_1, L_1) &= H(K_1, M_1, L_1, E) - H(E|K_1, M_1, L_1) \\ &\geq H(K_1, M_1, L_1, E) - H(E) \\ &\stackrel{(a)}{\geq} - \sum_k \sum_m \sum_l \Pr(K_1 = k, M_1 = m, L_1 = l, E = 0) \\ &\quad \times \log(\Pr\{\tilde{X}^N(1) \in \mathcal{T}_\epsilon^N(\tilde{X}|u^N(k, m, l))\}) \\ &\quad - h_2(\Pr(E = 0)) \\ &\geq \sum_k \sum_m \sum_l \Pr(K_1 = k, M_1 = m, L_1 = l, E = 0) \\ &\quad \times N(I(U; \tilde{X}) - \delta) - h_2(\Pr(E = 0)) \\ &= N(I(U; \tilde{X}) - \delta)(1 - \gamma) - h_2(1 - \gamma),\end{aligned}\quad (33)$$

where we use $h_2(\cdot)$ to denote the binary entropy function. (a) follows from that if $(K_1 = k, M_1 = m, L_1 = l)$ and $E = 0$ hold, then $\tilde{X}^N(1) \in \mathcal{T}_\epsilon^N(\tilde{X}|u^N(k, m, l))$ is valid. Thus (K_W, M_W, L_W) is close to uniformly distributed. Moreover, the private key is uniformly distributed and independent of the user index and biometrics. Since the secret key is a combination of (M_W, L_W) and part of the private key, this proves that the secret key is close to uniformly distributed.

Total Leakage Analysis: We first consider the secrecy leakage, which can be bounded as follows

$$\begin{aligned}I(S_W; J_W|\mathcal{C} = C) &= I(M_W, L_W, \tilde{P}_W; K_W, M_W \oplus \tilde{P}_W, \bar{P}_W|\mathcal{C} = C) \\ &\stackrel{(a)}{=} I(M_W, L_W; K_W|\mathcal{C} = C) \stackrel{(b)}{\leq} N\delta,\end{aligned}\quad (34)$$

where (a) follows from that \bar{P}_W , \tilde{P}_W , \check{P}_W , and (K_W, M_W, L_W) are mutually independent given the codebook; (b) follows from that (K_W, M_W, L_W) is close to uniformly distributed, which can be obtained from (33).

Now we consider the privacy leakage. When N is sufficiently large, the privacy leakage can be bounded as

$$\begin{aligned}I(X^N(W); J_W|\mathcal{C} = C) &= I(X^N(W); K_W, M_W \oplus \tilde{P}_W, \bar{P}_W|\mathcal{C} = C) \\ &= I(X^N(W); K_W, \bar{P}_W|\mathcal{C} = C) \\ &\quad + H(M_W \oplus \tilde{P}_W|K_W, \bar{P}_W, \mathcal{C} = C) \\ &\quad - H(M_W \oplus \tilde{P}_W|X^N(W), K_W, \bar{P}_W, \mathcal{C} = C) \\ &\leq I(X^N(W); K_W|\mathcal{C} = C) + NR_2 \\ &\quad - H(\check{P}_W|\tilde{X}^N(W), K_W, \bar{P}_W, \mathcal{C} = C) \\ &\stackrel{(a)}{=} I(X^N(W); K_W|\mathcal{C} = C) \\ &= H(K_W|\mathcal{C} = C) - H(K_W|X^N(W), \mathcal{C} = C) \\ &\leq NR_1 - H(K_W, |X^N(W), \mathcal{C} = C) \\ &= NR_1 - H(K_W, \tilde{X}^N(W)|X^N(W), \mathcal{C} = C) \\ &\quad + H(\tilde{X}^N(W)|K_W, X^N(W), \mathcal{C} = C) \\ &\stackrel{(b)}{=} NR_1 - H(\tilde{X}^N(W)|X^N(W), \mathcal{C} = C) \\ &\quad + H(\tilde{X}^N(W)|K_W, X^N(W), \mathcal{C} = C) \\ &\stackrel{(c)}{=} NR_1 - H(\tilde{X}^N(W)|X^N(W), \mathcal{C} = C) \\ &\quad + H(\tilde{X}^N(W), M_W, L_W|K_W, X^N(W), \mathcal{C} = C) \\ &\stackrel{(d)}{\leq} NR_1 - H(\tilde{X}^N(W)|X^N(W), \mathcal{C} = C) \\ &\quad + H(M_W, L_W|K_W, X^N(W), \mathcal{C} = C) \\ &\quad + H(\tilde{X}^N(W)|u^N(K_W, M_W, L_W), X^N(W), \mathcal{C} = C) \\ &\stackrel{(e)}{\leq} N(R_1 - H(\tilde{X}|X) + H(\tilde{X}|U, X) + \delta_\epsilon) \\ &\quad + H(M_W, L_W|K_W, X^N(W), \mathcal{C} = C) \\ &\stackrel{(f)}{\leq} N(\delta_1 + \delta_\epsilon) + H(M_W, L_W|K_W, X^N(W), \mathcal{C} = C),\end{aligned}\quad (35)$$

where (a) follows from that \bar{P}_W is independent of $(X^N(W), K_W)$ given the codebook, and the fact that \tilde{P}_W is independent of $(\tilde{X}^N(W), K_W, \bar{P}_W)$; (b) follows because, given the codebook $\mathcal{C} = C$, \tilde{X}^N determines K_W ; (c) follows because, given the codebook $\mathcal{C} = C$, \tilde{X}^N determines (M_W, L_W) ; (d) follows from that given the codebook $\mathcal{C} = C$, (K_W, M_W, L_W) determines $u^N(K_W, M_W, L_W)^1$; (e) follows for sufficiently large N from [22, Lemma 4]; (f) follows from the Markov chain $U - \tilde{X} - X - Y$ and the choice of $R_1 = I(U; \tilde{X}) - I(U; X) + \delta_1$.

¹Note that when the codebook is also fixed, the codeword $u^N(K_W, M_W, L_W)$ is fixed and therefore written using lower case.

Next, we provide the details of analysis for (f). Let E be a random variable such that $E = 0$ when $(X^N(W), U^N(K_W, M_W, L_W), \tilde{X}^N(W)) \in \mathcal{T}_\epsilon^N$ and $E = 1$ when $(X^N(W), U^N(K_W, M_W, L_W), \tilde{X}^N(W)) \notin \mathcal{T}_\epsilon^N$. Due to (31), we obtain that $\Pr(E = 0 | \mathcal{C} = C) \rightarrow 1$ as $N \rightarrow \infty$. Then for sufficiently large N , we have that

$$\begin{aligned}
& H(\tilde{X}^N(W) | u^N(K_W, M_W, L_W), X^N(W), \mathcal{C} = C) \\
& \leq H(\tilde{X}^N(W), E | u^N(K_W, M_W, L_W), X^N(W), \mathcal{C} = C) \\
& \leq H(E | \mathcal{C} = C) + \Pr(E = 0 | \mathcal{C} = C) \times H(\tilde{X}^N(W) | \\
& \quad u^N(K_W, M_W, L_W), X^N(W), E = 0, \mathcal{C} = C) \\
& \quad + \Pr(E = 1 | \mathcal{C} = C) H(\tilde{X}^N(W) | u^N(K_W, M_W, L_W), \\
& \quad X^N(W), E = 1, \mathcal{C} = C) \\
& \stackrel{(a)}{\leq} N\delta'_\epsilon + \Pr(E = 0 | \mathcal{C} = C) \times H(\tilde{X}^N(W) | \\
& \quad u^N(K_W, M_W, L_W), X^N(W), E = 0, \mathcal{C} = C) \\
& \stackrel{(b)}{\leq} N(\delta'_\epsilon + \delta''_\epsilon) + NH(\tilde{X} | U, X), \tag{36}
\end{aligned}$$

where (a) follows from that $\Pr(E = 1 | \mathcal{C} = C) \rightarrow 0$ and therefore $H(E) \rightarrow 0$ as $N \rightarrow \infty$; (b) follows from that, given the codebook $\mathcal{C} = C$ and $E = 0$, $(X^N(W), u^N(K_W, M_W, L_W), \tilde{X}^N(W)) \in \mathcal{T}_\epsilon^N$ is valid so that we can use Lemma 4 in [22].

The term $H(M_W, L_W | K_W, X^N(W), \mathcal{C} = C)$ in (35) can be bounded as follows

$$\begin{aligned}
& H(M_W, L_W | K_W, X^N(W), \mathcal{C} = C) \\
& \stackrel{(a)}{=} H(M_W, L_W | K_W, X^N(W), W, \mathcal{C} = C) \\
& \stackrel{(b)}{=} H(M_W, L_W | K_W, X^N(W), W, P_W, \mathcal{C} = C) \\
& \stackrel{(c)}{=} H(M_W, L_W | K_W, J_W, X^N(W), W, P_W, \mathcal{C} = C) \\
& \stackrel{(d)}{=} H(M_W, L_W | K_W, (J_i)_{i=1}^{M_I}, X^N(W), W, P_W, \mathcal{C} = C) \\
& \stackrel{(e)}{\leq} H(S_W | \hat{S}, \mathcal{C} = C) \\
& \stackrel{(f)}{\leq} 1 + \Pr\{\hat{S} \neq S_W | \mathcal{C} = C\} \log M_S \\
& \stackrel{(g)}{\leq} NR_S\delta'', \tag{37}
\end{aligned}$$

where (a) follows from that W is independent of $(M_W, K_W, L_W, X^N(W))$ given the codebook $\mathcal{C} = C$; (b) follows from that P_W is independent of $(M_W, K_W, L_W, X^N(W), W)$; (c) follows from that J_W is a function of (K_W, P_W) ; (d) follows from that given W and the codebook $\mathcal{C} = C$, $(M_W, K_W, L_W, X^N(W), P_W)$ is independent of the helper data of the other users; (e) follows from that \hat{S} is a function of $((J_i)_{i=1}^{M_I}, Y^N(W), P_W)$ and conditioning reduces entropy; (f) follows from Fano's inequality; (g) follows from defining a parameter δ'' that is small with large N and small ϵ due to (31) and the choice of the codebook C .

Therefore, we obtain that

$$\begin{aligned}
& I(S_W; J_W | \mathcal{C} = C) + I(X^N(W); J_W | \mathcal{C} = C) \\
& \leq N(\delta + \delta_1 + \delta_\epsilon + R_S\delta''). \tag{38}
\end{aligned}$$

Combining the above results, the direct part of the proof is completed.

2) *Converse*: Let \mathbf{J} denote $(J_i)_{i=1}^{M_I}$. Define auxiliary random variable $U_n = (W, S_W, P_W, J_W, X^{n-1}(W))$ for $n \in [1 : N]$. We assume that there exists a sequence of codes \mathcal{C} with identification rate R_I , secret key rate R_S , private key rate R_P , and helper data rate R_J such that the identification and authentication error probability vanishes as $N \rightarrow \infty$. For such code, Fano's inequality implies that $H(W, S_W | \hat{W}, \hat{S}) \leq F$, where $F \triangleq 1 + \Pr\{(\hat{W}, \hat{S}) \neq (W, S_W)\} \log(M_I M_S)$. Therefore $\frac{F}{N} \rightarrow 0$ as $\Pr\{(\hat{W}, \hat{S}) \neq (W, S_W)\} \rightarrow 0$ and $N \rightarrow \infty$.

Similar to [14, Equation (38)], consider the joint entropy

$$\begin{aligned}
& H(W, S_W) \\
& = I(W, S_W; P_W, \mathbf{J}, Y^N) + H(W, S_W | P_W, \mathbf{J}, Y^N) \\
& \stackrel{(a)}{\leq} I(W, S_W; P_W, \mathbf{J}, Y^N) + F \\
& = I(W, S_W; \mathbf{J}) + I(W, S_W; Y^N, P_W | \mathbf{J}) + F \\
& \stackrel{(b)}{=} I(S_W; \mathbf{J} | W) + I(W, S_W; Y^N, P_W | \mathbf{J}) + F \\
& \stackrel{(c)}{\leq} I(S_W; J_W, W) + I(W, S_W; Y^N, P_W | \mathbf{J}) + F \\
& \stackrel{(d)}{=} I(S_W; J_W) + I(W, S_W; Y^N, P_W | \mathbf{J}) + F \\
& \leq N\delta + I(W, S_W; Y^N, P_W | \mathbf{J}) + F \\
& = N\delta + I(W, S_W; P_W | \mathbf{J}) + I(W, S_W; Y^N | P_W, \mathbf{J}) + F \\
& \leq N\delta + H(P_W) + I(W, S_W, \mathbf{J}, P_W; Y^N) + F \\
& \stackrel{(e)}{=} N\delta + H(P_W) + I(Y^N; W, S_W, P_W, J_W) + F \\
& \leq N(2\delta + R_P) + I(Y^N; W, S_W, P_W, J_W) + F \\
& \leq N(2\delta + R_P) + F \\
& \quad + \sum_{n=1}^N I(Y_n; W, S_W, P_W, J_W, Y^{n-1}, X^{n-1}(W)) \\
& \stackrel{(f)}{=} N(2\delta + R_P) + \sum_{n=1}^N I(Y_n; U_n) + F, \tag{39}
\end{aligned}$$

where (a) follows from the fact that (\hat{W}, \hat{S}) are functions of (P_W, \mathbf{J}, Y^N) , and Fano's inequality; (b) holds since W is independent of \mathbf{J} ; (c) is valid since given the user index W , S_W is independent of the helper data of the other users; (d) follows from that W is independent of both J_W and S_W ; (e) holds as Y^N is only dependent with the helper data of the true user in the database; (f) holds due to $Y^{n-1} - (W, S_W, P_W, J_W, X^{n-1}(W)) - Y_n$.

Combining the above results, we obtain that

$$\begin{aligned}
\log M_I M_S & \stackrel{(a)}{\leq} \log M_I + \min_{w=1,2,\dots,M_I} H(S_w) + N\delta \\
& \leq H(W) + H(S_W | W) + N\delta = H(W, S_W) + N\delta \\
& \leq NR_P + \sum_{n=1}^N I(Y_n; U_n) + F + 3N\delta, \tag{40}
\end{aligned}$$

where (a) follows from the uniformity of the secret key (4c).

Combining (40) with (4b) and (4c), we obtain that

$$\begin{aligned}
R_I + R_S & \leq \frac{\log M_I M_S}{N} + 2\delta \\
& \leq R_P + \frac{1}{N} \sum_{n=1}^N I(Y_n; U_n) + \frac{F}{N} + 5\delta. \tag{41}
\end{aligned}$$

Next we consider the privacy leakage,

$$\begin{aligned}
I(X^N(W); J_W) &= I(X^N(W); \mathbf{J}|W) \\
&\stackrel{(a)}{=} I(X^N(W), W; \mathbf{J}) \\
&\geq H(X^N(W), W) - H(X^N(W), S_W, W|\mathbf{J}) \\
&= H(W) + H(X^N(W)) - H(S_W, W|\mathbf{J}, Y^N, P_W) \\
&\quad - I(Y^N, P_W; S_W, W|\mathbf{J}) - H(X^N(W)|S_W, W, \mathbf{J}) \\
&\stackrel{(b)}{\geq} H(W) + I(X^N(W); S_W, W, \mathbf{J}) - H(S_W, W|\hat{S}, \hat{W}) \\
&\quad - I(Y^N, P_W; S_W, W, \mathbf{J}) \\
&\stackrel{(c)}{\geq} H(W) + I(X^N(W); S_W, W, J_W) - F \\
&\quad - I(Y^N, P_W; S_W, W, J_W) \\
&\stackrel{(d)}{=} N(R_I - \delta) + I(X^N(W); S_W, W, J_W|Y^N) - F \\
&\quad - I(P_W; S_W, W, J_W|Y^N) \\
&= N(R_I - \delta) + I(X^N(W); S_W, W, J_W|Y^N) - F \\
&\quad - H(P_W|Y^N) + H(P_W|S_W, W, J_W, Y^N) \\
&\geq N(R_I - R_P - 2\delta) \\
&\quad + I(X^N(W); S_W, W, J_W, P_W|Y^N) - F \\
&\stackrel{(d)}{=} N(R_I - R_P - 2\delta) + I(X^N(W); S_W, W, J_W, P_W) \\
&\quad - I(Y^N; S_W, W, J_W, P_W) - F \\
&\stackrel{(e)}{\geq} N(R_I - R_P - 2\delta) \\
&\quad + \sum_{n=1}^N I(X_n(W); U_n) - \sum_{n=1}^N I(Y_n; U_n) - F, \quad (42)
\end{aligned}$$

where (a) follows from that W is independent of \mathbf{J} ; (b) follows from the fact that (\hat{W}, \hat{S}) are functions of (P_W, \mathbf{J}, Y^N) ; (c) follows from that the $X^N(W)$ and Y^N only correlate to the helper data J_W of user W in the database, and Fano's inequality; (d) holds due to the Markov chain $Y^N - X^N(W) - \tilde{X}^N(W) - (S_W, W, J_W, P_W)$ holds; (e) is valid due to $Y^{n-1} - (S_W, W, J_W, P_W, X^{n-1}(W)) - Y_n$.

Combining the above result with (4g), we obtain that

$$\begin{aligned}
R_I + \frac{1}{N} \sum_{n=1}^N I(X_n(W); U_n) \\
\leq R_P + \frac{1}{N} \sum_{n=1}^N I(Y_n; U_n) + \frac{F}{N} + 3\delta. \quad (43)
\end{aligned}$$

We bound the helper data rate as follows

$$\begin{aligned}
N(R_J + \delta) &\geq H(J_W|W) \\
&= I(\tilde{X}^N(W); J_W|W) + H(J_W|\tilde{X}^N(W), W) \\
&\stackrel{(a)}{=} I(\tilde{X}^N(W); \mathbf{J}|W) + H(J_W|\tilde{X}^N(W), W) \\
&= I(\tilde{X}^N(W), W; \mathbf{J}) + H(J_W|\tilde{X}^N(W), W) \\
&= H(W) + H(\tilde{X}^N(W)) - H(\tilde{X}^N(W), W, S_W|\mathbf{J}) \\
&\quad + H(S_W|\tilde{X}^N(W), W, \mathbf{J}) + H(J_W|\tilde{X}^N(W), W) \\
&= H(W) + H(\tilde{X}^N(W)) - H(\tilde{X}^N(W)|\mathbf{J}, W, S_W) \\
&\quad - H(W, S_W|\mathbf{J}) + H(S_W|\tilde{X}^N(W), W, \mathbf{J}) \\
&\quad + H(J_W|\tilde{X}^N(W), W)
\end{aligned}$$

$$\begin{aligned}
&= H(W) + I(\tilde{X}^N(W); S_W, W, \mathbf{J}) \\
&\quad - H(S_W, W|\mathbf{J}, Y^N, P_W) - I(Y^N, P_W; S_W, W|\mathbf{J}) \\
&\quad + H(S_W, J_W|\tilde{X}^N(W), W) \\
&\geq H(W) + I(\tilde{X}^N(W); S_W, W, \mathbf{J}) \\
&\quad - H(S_W, W|\hat{S}, \hat{W}) - I(Y^N, P_W; S_W, W|\mathbf{J}) \\
&\quad + H(S_W, J_W|\tilde{X}^N(W), W) \\
&\stackrel{(b)}{\geq} H(W) + I(\tilde{X}^N(W); S_W, W, J_W|Y^N) - F \\
&\quad - I(P_W; S_W, W, J_W|Y^N) + I(Y^N, P_W; \mathbf{J}) \\
&\quad + H(S_W, J_W|\tilde{X}^N(W), W) \\
&\geq H(W) + I(\tilde{X}^N(W); S_W, W, J_W, P_W|Y^N) - F \\
&\quad + H(P_W|S_W, W, J_W, \tilde{X}^N(W), Y^N) \\
&\quad - H(P_W|Y^N) + H(S_W, J_W|\tilde{X}^N(W), W) \\
&\stackrel{(c)}{=} H(W) + I(\tilde{X}^N(W); S_W, W, J_W|Y^N) - F \\
&\quad + H(P_W|S_W, W, J_W, \tilde{X}^N(W)) - H(P_W|Y^N) \\
&\quad + H(S_W, J_W|\tilde{X}^N(W), W) \\
&= H(W) + I(\tilde{X}^N(W); S_W, W, J_W|Y^N) - F \\
&\quad + H(P_W|\tilde{X}^N(W), W) \\
&\quad - I(P_W; S_W, J_W|\tilde{X}^N(W), W) \\
&\quad - H(P_W|Y^N) + H(S_W, J_W|\tilde{X}^N(W), W) \\
&= H(W) + I(\tilde{X}^N(W); S_W, W, J_W|Y^N) - F \\
&\quad + H(P_W|\tilde{X}^N(W), W) - H(P_W|Y^N) \\
&\quad + H(S_W, J_W|\tilde{X}^N(W), W, P_W) \\
&\stackrel{(d)}{=} H(W) + I(\tilde{X}^N(W); S_W, W, J_W|Y^N) - F \\
&\stackrel{(c)}{=} H(W) + I(\tilde{X}^N(W); S_W, W, J_W, P_W) \\
&\quad - I(Y^N(W); S_W, W, J_W, P_W) - F \\
&\stackrel{(e)}{\geq} N(R_I - \delta) - F \\
&\quad + \sum_{n=1}^N I(\tilde{X}_n(W); U_n) - \sum_{n=1}^N I(Y_n; U_n), \quad (44)
\end{aligned}$$

where (a) follows from the fact that, given the user index W , the enrolled biometric sequence $\tilde{X}^N(W)$ is independent of the helper data of the other users; (b) follows from Fano's inequality; (c) follows from the Markov chain $Y^N - X^N(W) - \tilde{X}^N(W) - (S_W, W, J_W, P_W)$; (d) follows from that P_W is independent of $(\tilde{X}^N(W), W, Y^N)$, and (J_W, S_W) is a function of $(\tilde{X}^N(W), P_W)$; (e) follows from $X^{n-1}(W) - (S_W, W, J_W, P_W, \tilde{X}^{n-1}(W)) - X_n(W)$ and $Y^{n-1} - (S_W, W, J_W, P_W, X^{n-1}(W)) - Y_n$.

Therefore, we obtain the rate condition

$$\begin{aligned}
R_I + \frac{1}{N} \sum_{n=1}^N I(\tilde{X}_n(W); U_n) \\
\leq R_J + \frac{1}{N} \sum_{n=1}^N I(Y_n; U_n) + \frac{F}{N} + 2\delta. \quad (45)
\end{aligned}$$

Let Q be a uniform random variable on $[1 : N]$ and independent of everything else. Define $U = (U_Q, Q)$ and

$P_{UX_QY_Q} = Q_{Y_Q|X_Q}P_{U|X_Q}P_{X_Q}$. Note that $U - X_Q(W) - Y_Q$ still holds. As (X_Q, Y_Q) has the same joint distribution as (X, Y) , with $\delta \rightarrow 0$, we obtain that $(R_I, R_S, R_P, R_J) \in \mathcal{R}_g$ from (41), (43), and (45). This completes the proof of the backward direction.

APPENDIX B PROOF OF THEOREM 2

The proof of Theorem 2 is based on that of Theorem 1.

1) *Achievability*: We use the similar codebook generation as in Theorem 1, while the difference is that we fix $R_J = R_S + R_I + I(U; \tilde{X}) - I(U; Y) + \delta + \delta_1$. As in the proof of Theorem 1, for each user w , we look for a triple (k_w, m_w, l_w) and generate the corresponding helper data $j_w^g = (k_w, m_w \oplus \tilde{p}_w, \tilde{p}_w)$ and a secret key $s_w^g = (m_w, l_w, \tilde{p}_w)$, where $J_W^g \in [1 : M_J^g]$. Theorem 1 states that for any $\delta > 0$, there exist some $N \geq 1$, enrollment mapping and identification mapping such that

$$\begin{aligned} \Pr\{(\hat{W}, \hat{S}^g) \neq (W, S_W^g)\} &\leq \delta, \\ \log M_I &\geq N(R_I - \delta), \\ H(S_W^g) + N\delta &\geq \log M_S \geq N(R_S - \delta), \\ \log M_P &\leq N(R_P + \delta), \\ \log M_J^g &\leq N(R_J^g + \delta), \\ I(S_W^g; J_W^g) + I(X^N(W); J_W^g) &\leq N\delta. \end{aligned} \quad (46)$$

Here, we include another masking layer. Let the chosen secret key s_w^c mask the previously generated secret key s_w^g , where s_w^c is generated from $[1 : M_S]$ uniformly at random. We then obtain the masked secret key $j_w^a = s_w^g \oplus s_w^c$, which is additional helper data. The masked secret keys $(j_w^a)_{w=1}^{M_I}$ can be stored in a public database instead of a secure database. Therefore, we consider the joint helper data (j_w^g, j_w^a) . In the identification phase, the secret key can be estimated as

$$\hat{s}^c = j_w^a \oplus \hat{s}^g = s_w^g \oplus s_w^c \oplus \hat{s}^g. \quad (47)$$

Following a similar analysis as in [14, (48)], we obtain that

$$\begin{aligned} I(S_W^c; J_W^g, J_W^a) &= I(S_W^c; J_W^g, S_W^c \oplus S_W^g) \\ &= I(S_W^c; J_W^g) + I(S_W^c; S_W^c \oplus S_W^g | J_W^g) \\ &\stackrel{(a)}{=} H(S_W^c \oplus S_W^g | J_W^g) - H(S_W^c \oplus S_W^g | J_W^g, S_W^c) \\ &\leq H(S_W^c \oplus S_W^g) - H(S_W^c \oplus S_W^g | J_W^g, S_W^c) \\ &= H(S_W^c \oplus S_W^g) - H(S_W^g | J_W^g, S_W^c) \\ &\stackrel{(b)}{\leq} \log M_S - H(S_W^g) + I(S_W^g; J_W^g), \end{aligned} \quad (48)$$

where (a) holds since S_W^c is independent of J_W^g ; (b) follows from that S_W^c is independent of both S_W^g and J_W^g .

Following [14, (48)], the privacy leakage can be bounded as

$$\begin{aligned} I(X^N(W); J_W^g, J_W^a) &= I(X^N(W); J_W^g, S_W^c \oplus S_W^g) \\ &= I(X^N(W); J_W^g) + I(X^N(W); S_W^c \oplus S_W^g | J_W^g) \\ &\leq I(X^N(W); J_W^g) + H(S_W^c \oplus S_W^g) \\ &\quad - H(S_W^c \oplus S_W^g | X^N(W), J_W^g, S_W^g) \\ &\leq I(X^N(W); J_W^g) + \log M_S - H(S_W^c | X^N(W), J_W^g, S_W^g) \end{aligned}$$

$$= I(X^N(W); J_W^g). \quad (49)$$

Therefore, the total leakage can be bounded as

$$\begin{aligned} I(S_W^c; J_W^g, J_W^a | \mathcal{C} = C) &+ I(X^N(W); J_W^g, J_W^a | \mathcal{C} = C) \\ &\leq \log M_S - H(S_W^g) + I(S_W^g; J_W^g | \mathcal{C} = C) \\ &\quad + I(X^N(W); J_W^g | \mathcal{C} = C) \\ &\stackrel{(a)}{\leq} N\delta + I(S_W^g; J_W^g | \mathcal{C} = C) + I(X^N(W); J_W^g | \mathcal{C} = C) \\ &\stackrel{(b)}{\leq} N(2\delta + \delta_1 + \delta_\epsilon + R_S\delta''), \end{aligned} \quad (50)$$

where (a) holds since S_W^g is close to uniformly distributed on $[1 : M_S]$; (b) follows from (38).

In the scenario of the chosen secret key system, as we are using the masking layer, we can obtain that $\hat{S}^c = S^c$ holds only if $\hat{S}^g = S^g$. Thus, $\Pr\{(\hat{W}, \hat{S}^g) \neq (W, S_W^g)\} \leq \delta$ implies $\Pr\{(\hat{W}, \hat{S}^c) \neq (W, S_W^c)\} \leq \delta$. And we also have that

$$\begin{aligned} \log M_I &\geq N(R_I - \delta), \\ H(S_W^c) &= \log M_S \geq N(R_S - \delta), \\ \log M_P &\leq N(R_P + \delta), \\ \log M_J^g + \log M_S &\leq N(R_J^g + \delta), \\ I(S_W^g; J_W^a, J_W^g) + I(X^N(W); J_W^a, J_W^g) &\leq 2N\delta, \end{aligned} \quad (51)$$

for a chosen secret key system.

Consequently, if the rate tuple (R_I, R_S, R_P, R_J) is achievable for a generated secret key system, then the rate tuple $(R_I, R_S, R_P, R_J + R_S)$ is achievable for a chosen secret key system.

2) *Converse*: Define the auxiliary random variable $U_n = (W, S_W, P_W, J_W, X^{n-1}(W))$ for $n \in [1 : N]$. Following (39) and (40), we can obtain that

$$R_I + R_S \leq R_P + \frac{1}{N} \sum_{n=1}^N I(Y_n; U_n) + \frac{F}{N} + 5\delta. \quad (52)$$

Similar to the analysis in (42) and (43), we can obtain that

$$\begin{aligned} R_I + \frac{1}{N} \sum_{n=1}^N I(X_n(W); U_n) \\ \leq R_P + \frac{1}{N} \sum_{n=1}^N I(Y_n; U_n) + \frac{F}{N} + 3\delta. \end{aligned} \quad (53)$$

For the helper data rate, we obtain that

$$\begin{aligned} N(R_J + \delta) &\geq H(J_W | W) \\ &= I(\tilde{X}^N(W), S_W; J_W | W) + H(J_W | \tilde{X}^N(W), S_W, W) \\ &= I(\tilde{X}^N(W), S_W; \mathbf{J} | W) + H(J_W | \tilde{X}^N(W), S_W, W) \\ &= I(\tilde{X}^N(W), S_W, W; \mathbf{J}) + H(J_W | \tilde{X}^N(W), S_W, W) \\ &\stackrel{(a)}{=} H(\tilde{X}^N(W)) + H(S_W) + H(W) \\ &\quad - H(\tilde{X}^N(W), S_W, W | \mathbf{J}) + H(J_W | \tilde{X}^N(W), S_W, W) \\ &= H(\tilde{X}^N(W)) + H(S_W) + H(W) - H(S_W, W | \mathbf{J}) \\ &\quad - H(\tilde{X}^N(W) | \mathbf{J}, S_W, W) + H(J_W | \tilde{X}^N(W), S_W, W) \\ &\stackrel{(b)}{=} I(\tilde{X}^N(W); S_W, W, J_W) + H(S_W) + H(W) \\ &\quad - H(S_W, W | \mathbf{J}) + H(J_W | \tilde{X}^N(W), S_W, W) \end{aligned}$$

$$\begin{aligned}
&= I(\tilde{X}^N(W); S_W, W, J_W) + H(S_W) + H(W) \\
&\quad - H(S_W, W | \mathbf{J}, Y^N, P_W) - I(W, S_W, \mathbf{J}; P_W, Y^N) \\
&\quad + I(\mathbf{J}; Y^N, P_W) + H(J_W | \tilde{X}^N(W), S_W, W) \\
&\stackrel{(c)}{\geq} I(\tilde{X}^N(W); S_W, W, J_W | Y^N) + H(S_W) + H(W) \\
&\quad - H(S_W, W | \hat{S}, \hat{W}) - I(W, S_W, \mathbf{J}; P_W | Y^N) \\
&\quad + H(J_W | \tilde{X}^N(W), S_W, W) \\
&\stackrel{(d)}{\geq} I(\tilde{X}^N(W); S_W, W, J_W, P_W | Y^N) + H(S_W) \\
&\quad + H(W) - F + H(P_W | Y^N, \tilde{X}^N, W, S_W, J_W) \\
&\quad - H(P_W | Y^N) + H(J_W | \tilde{X}^N(W), S_W, W) \\
&\stackrel{(e)}{=} I(\tilde{X}^N(W); S_W, W, J_W, P_W | Y^N) + H(S_W) \\
&\quad + H(W) - F + H(P_W | \tilde{X}^N, W, S_W, J_W) \\
&\quad - H(P_W | Y^N) + H(J_W | \tilde{X}^N(W), W, S_W) \\
&= I(\tilde{X}^N(W); S_W, W, J_W, P_W | Y^N) + H(S_W) \\
&\quad + H(P_W | \tilde{X}^N(W), W, S_W) - H(P_W | Y^N) \\
&\quad + H(W) - F + H(J_W | \tilde{X}^N(W), W, S_W, P_W) \\
&\stackrel{(f)}{=} I(\tilde{X}^N(W); S_W, W, J_W, P_W | Y^N) + H(S_W) \\
&\quad + H(W) - F \\
&\stackrel{(e)}{=} H(S_W) + H(W) + I(\tilde{X}^N(W); S_W, W, J_W, P_W) \\
&\quad - I(Y^N(W); S_W, W, J_W, P_W) - F \\
&\stackrel{(g)}{\geq} N(R_I + R_S - 2\delta) - F \\
&\quad + \sum_{n=1}^N I(\tilde{X}_n(W); U_n) - \sum_{n=1}^N I(Y_n; U_n), \tag{54}
\end{aligned}$$

where (a) follows from the fact that $\tilde{X}^N(W)$, S_W and W are mutually independent; (b) follows from $\tilde{X}^N(W)$ being independent of the helper data of other users; (c) follows from the fact that (\hat{S}, \hat{W}) are functions of (\mathbf{J}, Y^N, P_W) ; (d) follows from Fano's inequality; (e) follows from the Markov chain $Y^N - \tilde{X}^N(W) - (S_W, W, J_W, P_W)$; (f) follows from P_W being independent of $(\tilde{X}^N, W, S_W, Y^N)$, and J_W is a function of (\tilde{X}^N, S_W, P_W) ; (g) follows from $X^{n-1}(W) - (S_W, W, J_W, P_W, \tilde{X}^{n-1}(W)) - X_n(W)$ and $Y^{n-1} - (S_W, W, J_W, P_W, X^{n-1}(W)) - Y_n$. Therefore, we obtain that

$$\begin{aligned}
R_I + R_S + \frac{1}{N} \sum_{n=1}^N I(\tilde{X}_n(W); U_n) \\
\leq R_J + \frac{1}{N} \sum_{n=1}^N I(Y_n; U_n) + \frac{F}{N} + 3\delta. \tag{55}
\end{aligned}$$

Using the cardinality bounding argument, we can complete the proof for the backward direction.

REFERENCES

- [1] B. Schneier, "Inside risk: The uses and abuses of biometrics," *Comm. ACM*, vol. 42, no. 8, pp. 136–199, 1999.
- [2] F. Willems, T. Kalker, J. Goseling, and J. P. Linnartz, "On the capacity of a biometrical identification system," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2003, pp. 82–87.
- [3] E. Tuncel, "Capacity/storage tradeoff in high-dimensional identification systems," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 1929–1933.
- [4] F. Farhadzadeh and F. M. J. Willems, "Identification rate, search and memory complexity tradeoff: Fundamental limits," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6173–6188, Nov. 2016.
- [5] M. T. Vu, "Perspectives on identification systems," Ph.D. dissertation, KTH Royal Institute of Technology, 2019.
- [6] A. Grigorescu, H. Boche, and R. F. Schaefer, "Robust biometric authentication from an information theoretic perspective," *Entropy*, vol. 19, no. 9, p. 480, 2017.
- [7] M. T. Vu, T. J. Oechtering, and M. Skoglund, "Testing in identification systems," in *2018 IEEE Inf. Theory Workshop (ITW)*. IEEE, 2018, pp. 1–5.
- [8] N. Merhav, "False-accept/false-reject trade-offs for ensembles of biometric authentication systems," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 4997–5006, 2019.
- [9] T. Ignatenko and F. M. J. Willems, "Biometric systems: Privacy and secrecy aspects," *IEEE Tran. Inf. Forensics Security*, vol. 4, no. 4, pp. 956–973, Dec. 2009.
- [10] L. Lai, S. 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, 2011.
- [11] H. Boche, R. F. Schaefer, S. Baur, and H. V. Poor, "On the algorithmic computability of the secret key and authentication capacity under channel, storage, and privacy leakage constraints," *IEEE Trans. Signal Process.*, vol. 67, no. 17, pp. 4636–4648, 2019.
- [12] O. Günlü, R. F. Schaefer, and H. V. Poor, "Biometric and physical identifiers with correlated noise for controllable private authentication," *CoRR*, vol. abs/2001.00847, 2020. [Online]. Available: <http://arxiv.org/abs/2001.00847>
- [13] 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, 2016.
- [14] 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.
- [15] 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, 2018.
- [16] V. Yachongka and H. Yagi, "Identification, secrecy, template, and privacy-leakage of biometric identification system under noisy enrollment," *arXiv preprint arXiv:1902.01663*, 2019.
- [17] —, "Fundamental limits of identification system with secret binding under noisy enrollment," *arXiv preprint arXiv:1905.03598*, 2019.
- [18] L. Zhou, M. T. Vu, T. J. Oechtering, and M. Skoglund, "Fundamental limits for biometric identification systems without privacy leakage," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2019, pp. 1105–1112.
- [19] A. Wyner, "A theorem on the entropy of certain binary sequences and applications-i," *IEEE Trans. Inf. Theory*, vol. 19, no. 6, pp. 769–772, 1973.
- [20] A. E. Gamal and Y. H. Kim, *Network information theory*. Cambridge university press, 2011.
- [21] P. Minero, S. H. Lim, and Y.-H. Kim, "A unified approach to hybrid coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1509–1523, 2015.
- [22] K. Kittichokechai, T. J. Oechtering, M. Skoglund, and Y. Chia, "Secure source coding with action-dependent side information," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6444–6464, Dec 2015.