Fundamental Limits-Achieving Polar Code Designs for Biometric Identification and Authentication

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

***Abstract*—In this work, we present polar code designs that offer a provably optimal solution for biometric identification and authentication systems under noisy enrollment for certain sources and observation channels. We consider a discrete memoryless biometric source and discrete symmetric memoryless observation channels. It is shown that the proposed polar code designs achieve the fundamental limits with privacy and secrecy constraints. Depending on how the secret keys are extracted and whether the privacy leakage rate should be close to zero, we consider four related setups, which are (i) the generated secret key system, (ii) the chosen secret key system, (iii) the generated secret key system with zero leakage, and (iv) the chosen secret key system with zero leakage. For the first two setups, (i) and (ii), the privacy level is characterized by the privacy leakage rate. For the last two setups (iii) and (iv), private keys are additionally employed to achieve close to zero privacy leakage rate. In setups (i) and (iii), it is assumed that the secret keys are generated, i.e., extracted from biometric information. While in setups (ii) and (iv), secret keys provided to the system are chosen uniformly at random from some trustful source. This work provides the first examples of fundamental limits-achieving code designs for identification and authentication. Moreover, since the code designs are based on polar codes and many existing works study low-complexity and short block-length polar coding, the proposed code designs in this work provide the code design structure and a framework for the application of biometric identification and authentication.**

***Index Terms*—Biometrics, identification systems, noisy enrollment, polar codes, privacy, strong secrecy.**

I. Introduction

With recent developments of biometric recognition technology, biometrics is increasingly used in various applications. Two common application areas are biometric authentication and identification, which offer several advantages. For example, biometric features are stable, e.g. the face or fingerprint features are always carried by an individual and they do not change over a period of time. Moreover, biometric authentication and identification require the person to be present at the time of authentication and identification, which is a stringent requirement and helps ensure security. Due to the advantages of employing biometrics in identification and authentication systems, they are more and more used in various smart technology and devices. However, biometric data need to be carefully protected since biometrics is permanently associated with an individual so that a breach would have severe consequences.

A biometric identification and authentication system can be modeled as follows. Consider a biometric identification and authentication system in Fig. 1 with $M_I = 2^{NR_I}$ users. In the enrollment phase, each user $i \in [1 : M_I]$ presents its biometric sequence $X^N(i)$ to the system and is enrolled through a noisy channel. The system maps the noisy enrollment $\tilde{X}^N(i)$ to the helper data $J(i) \in [1 : 2^{NR_J}]$, which is stored in a public helper database and later used for identification. In the meantime, the system produces a secret key $S(i)$ from the set $[1 : 2^{NR_S}]$, which is stored securely and later used for authentication of the identified user. In the identification and authentication phase, an unknown but previously enrolled user $W \in [1 : M_I]$ tries to access the system. The biometric sequence of the user is observed via a noisy channel, and a noisy observation Y^N is obtained. The system uses the observation Y^N and the public helper database $\{J(i)\}_{i=1}^{M_I}$ to identify the user as $\hat{W}$. The system also outputs an estimated secret key $\hat{S}$. If an authentication procedure is required, the system compares the estimated secret key $\hat{S}$ with the stored secret key $S(\hat{W})$ of the guessed user $\hat{W}$. If they are the same, i.e., $\hat{S} = S(\hat{W})$, the user is authenticated successfully, e.g., granted access to the system.

A usage scenario of the system described above is the biometric identification and authentication in a closed working area, i.e., the users in the working area are all enrolled in the system. Due to security reasons, the working area is divided into different sections and only certain users have access to each section. In such a system, the helper data are stored publicly, e.g. in the central cloud. The secret key of each user is stored in the corresponding section locally or handed to the user, e.g. stored on a card or token. When a user tries to access a section, the local system in that section uses the observation and the public helper database to identify the user identity. Furthermore, the local system estimates the corresponding secret key. If the guessed user belongs to the system and the estimated secret key matches the true one stored in the local system, the user is granted access to that section.

As the use of biometric data brings convenience, it also invokes privacy and secrecy issues if the biometric information and the secret keys are compromised. Since the identification and authentication are based on biometrics and the extracted secret keys, compromised biometric data and secret keys can lead to unauthorized access to the system. Hence, the following attack model regarding secrecy and privacy aspects is

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.

Linghui Zhou, Tobias J. Oechtering, and Mikael Skoglund are with the Division of Information Science and Engineering, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden (e-mail: linghui@kth.se; oech@kth.se; skoglund@kth.se).

Part of the results was presented in the 2021 Information Theory Workshop.

included in the design. We assume that the attacker has access to the public database but no access to the secure database which stores the secret keys used for authentication. The attacker is interested in inferring information about the secret keys and the biometrics. We do not consider the sophisticated attack where an attacker creates and stores its own enrollment data. We require that the amount of information leaked about the secret key from the public helper database, which is the secrecy leakage, is negligible. We also characterize the amount of information leaked about the biometrics from the public helper database, which is the privacy leakage. Additionally, the use of the public helper database avoids storing the raw biometric information directly, which helps to achieve a more efficient design in storage. Therefore, the use of a helper database helps to minimize security and privacy risks.

Biometric systems for authentication and identification have been studied from various perspectives in the literature, among which characterizing the fundamental limits is an important aspect. By characterizing the fundamental limits, the optimality performance of a biometric system can be assessed from a systematic information-theoretic perspective. The capacity of biometric identification systems, which relates to the maximal number of users that can be reliably identified, is firstly characterized in [1]. Taking search complexity into consideration, hierarchical identification is studied in [2]–[5]. The fundamental limits of biometric systems considering privacy and secrecy aspects in various scenarios are studied in [5]–[11]. The problem of the hypothesis testing of the identification problem and deciding whether the observation belongs to the system is investigated in [12], where it has been shown that the codes for an identification system aiming at identifying only the previous enrolled users can be also utilized for deciding if the user belongs to the system. In [13], fundamental limits and relations of identification problems have been studied in general, including uncertainty models. Additional to the aforementioned aspects, multi-factor authentication is studied to improve the security and usability of an authentication system. Combining two authentication factors, the authentication method studied in [14] is shown to perform better in security and usability. Based on extended chaotic-maps for mobile lightweight devices, a practical and provably secure three-factor authentication protocol that balances security and utility better is studied in [15]. In [16], three typical biometric encryption approaches, including fuzzy vault, fuzzy commitment, and fuzzy extractor are unified to achieve three-factor authentication without privacy leakage.

Considering the optimal performance of a biometric system, the fundamental limits studied in the literature do provide insights on designing and analyzing biometric systems, but most of the analysis is based on the concept of typical sequences and cannot be implemented directly. Practical schemes of biometric identification that achieve fundamental limits, to the best of our knowledge, have not been studied before. In this work, we propose polar code designs that are provably optimal. We consider the setup studied [10] and [11], in which both secrecy and privacy aspects are characterized in a biometric identification and authentication system. In [10], the identification and authentication are based on biometrics. Though applying external keys is less convenient compared to using biometrics only, it helps to preserve the privacy of the biometric system. In [11], both biometrics and external private keys are used to satisfy a more stringent privacy constraint.

Information-theoretic security and privacy, which aim at providing practical code designs to the aforementioned problems, is of high interest. Polar codes, as firstly studied by Arikan in [17], are the first codes that provably achieve the capacity of binary symmetric memoryless channels with efficient encoding and decoding operations. Due to the capacity-achieving performance of polar codes, the concepts have been quickly extended so that polar codes have been developed for many applications. Polar codes have been generalized to asymmetric channels in [18] and arbitrary alphabets in [19]. It is shown in [20] that channel polarization can be generalized to source polarization, which achieves the compression bound in Shannon's lossless source coding theorem. By additionally considering side information at the eavesdropper, polar codes for secure Wyner-Ziv coding are studied in [21]. Polar coding for the wiretap channel with a shared secret key is studied in [22]. Polar coding schemes for secure transmission and key agreement are studied in [23]. Secret-key capacity-achieving polar code schemes based on a sequential strategy are studied in [24]. Code constructions based on polar codes for biometric secrecy systems are studied in [25]. Secret key generation over biased physical unclonable functions with polar codes is studied in [26]. In this work, we consider biometric-based secret key generation using polar codes. We additionally include the identification problem and study the polar code design. Further, low-complexity polar coding schemes are currently intensively studied, see for instance in [27]–[29], which are important next development steps for the practical application of polar codes.

This work aims at developing fundamental limits-achieving polar code designs for the following setups: (i) the *generated secret key system*; (ii) the *chosen secret key system*; (iii) the *generated secret key system with zero leakage*; (iv) the *chosen secret key system with zero leakage*. Setups (i) and (iii) consider the case that secret keys are extracted from the biometrics. In setups (ii) and (iv), secret keys are assumed to be produced from some trustful source uniformly at random. For setups (i) and (ii), the identification and authentication are based on the biometrics only, which are studied in [10]. For setups (iii) and (iv), private keys are additionally included and a more stringent privacy leakage requirement can be satisfied, which are studied in [11]. For all the above four setups, we consider the binary case and symmetric memoryless channels.

The main contributions of this work are listed as follows:

- Fundamental limits-achieving designs that involve both identification and authentication are proposed.
- Strong secrecy is achieved for free with the proposed designs.

The rest of the paper is organized as follows. Section II gives the problem formulations of the four setups. Section III presents the proposed polar code designs for the four setups. In Section IV, we summarize this paper.

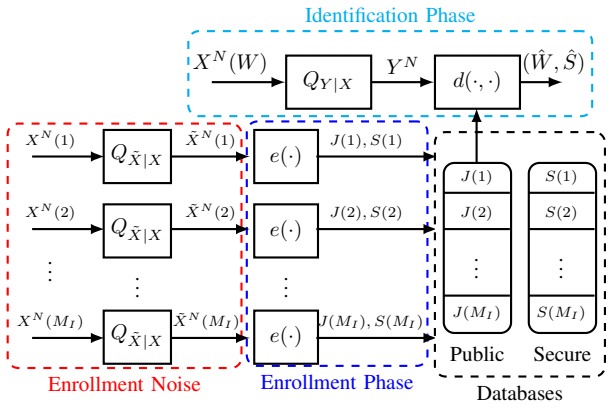

Fig. 1: Model of a generated secret key system. Note that the authentication step is not depicted in the figure. During an authentication procedure, the system compares the estimated secret key $\hat{S}$ with the stored secret key $S(\hat{W})$ in the secure database of the guessed user $\hat{W}$.

Notations: The set $\{1,2,...,M\}$ is denoted with $[1:M]$. For $n \in \mathbb{N}^+$ and $N \triangleq 2^n$, let $G_N \triangleq \left[\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right]^{\otimes n} B_N$ denote the source polarization transformation defined in [20], where "$\otimes n$" denotes the nth Kronecker power and B_N denotes the "bit-reversal" permutation matrix [17]. The matrix B_N can be interpreted as a bit-reversal operator: if $v_1^N = u_1^N B_N$, then $v_{b_1...b_n} = u_{b_n...b_1}$. We denote the variational distance and the Kullback-Leibler divergence by $\mathbb{V}(\cdot,\cdot)$ and $\mathbb{D}(\cdot||\cdot)$, i.e.,

$$\mathbb{D}(p||q) = \sum_{x\in\mathcal{X}} p(x)\log\frac{p(x)}{q(x)}, \tag{1}$$

$$\mathbb{V}(p,q) = \frac{1}{2}||p-q||_1. \tag{2}$$

Let $h_2(p)$ denote the binary entropy function $h_2(p) = -p\log_2 p - (1-p)\log_2(1-p)$. Let $\oplus$ and $\ominus$ denote modulo addition and subtraction, which in our binary case are the XOR operation. Let $V^N[\mathcal{A}] \longleftarrow \mathbf{a}$ denote assignment of vector $\mathbf{a}$ on coefficients of V^N with indices in $\mathcal{A}$. For a set $\mathcal{A} \triangleq \{a_j\}_{j=1}^{|\mathcal{A}|}$, define $U^N[\mathcal{A}](w) = (U_{a_1}(w), U_{a_2}(w), ..., U_{a_{|\mathcal{A}|}}(w))$. For a set $\mathcal{A}$ or a random variable (RV) A, we use $|\mathcal{A}|$ and $|A|$ to denote the size of them, respectively.

II. Problem Formulation

In this section, we describe the problem formulation.

A. Enrollment Noise

Assume that there are M_I users in the system with indices $\{1,2,...,M_I\}$. We further assume that the biometric sequence $x^N(w)$ for each user $w \in [1:M_I]$ is identically independently distributed (i.i.d.) according to the probability mass function (p.m.f.) $Q_X(\cdot)$ on a finite alphabet $\mathcal{X}$. In the enrollment phase, for each user $w \in [1:M_I]$, the biometric sequence $x^N(w)$ is observed via a noisy discrete memoryless enrollment channel $Q_{\tilde{X}|X}(\cdot)$, which generates a noisy enrollment sequence $\tilde{x}^N(w)$. The enrollment channel $Q_{\tilde{X}|X}(\cdot)$ is a model to incorporate the noisy measurement or other noise.

B. Generated Secret Key System

Consider a generated secret key system depicted in Fig. 1. In the enrollment phase, the enrollment mapping $e(\cdot)$ maps $\tilde{x}^N(w)$ to helper data $j(w) \in [1:M_J]$ and secret key $s(w) \in [1:M_S]$, i.e.,

$$(j(w), s(w)) = e(\tilde{x}^N(w)), \tag{3}$$

where $j(w)$ and $s(w)$ are stored in a public helper database and a secure database at the location w, respectively.

In the identification phase, an unknown user w with biometric source sequence $x^N(w)$ is observed via a discrete memoryless channel $Q_{Y|X}(\cdot)$ and an noisy observation y^N is generated. The observed user's index w is assumed to be a realization of RV W that is uniformly distributed on $[1:M_I]$. The identification mapping $d(\cdot,\cdot)$ uses y^N and $\{j(i)\}_{i=1}^{M_I}$ to guess the user index denoted as $\hat{w}$ and estimate the secret key denoted as $\hat{s}$, i.e.,

$$(\hat{w}, \hat{s}) = d(y^N, \{j(i)\}_{i=1}^{M_I}). \tag{4}$$

During an authentication procedure, the system compares the estimated secret key $\hat{s}$ with the stored secret key $s(\hat{w})$ of the guessed user $\hat{w}$. If they are the same, i.e., $\hat{s} = s(\hat{w})$, the user is authenticated successfully.

The achievability of a rate tuple in a generated secret key system is defined as follows.

Definition 1: A rate tuple $(R_I, R_S, R_L, R_J) \in \mathbb{R}_+^4$ of the identification rate, the secret key rate, the privacy leakage rate, and the helper data rate is *achievable* in a generated secret key system if for all $\delta > 0$ there exists some $N_0(\delta) \geq 1$ such that for all $N \geq N_0(\delta)$ there exists enrollment and identification mappings such that the following conditions are satisfied

$$\Pr\{(\hat{W},\hat{S}) \neq (W, S(W))\} \leq \delta, \tag{5}$$

$$H(S(W)) + \delta \geq \log M_S \geq N(R_S - \delta), \tag{6}$$

$$\log M_I \geq N(R_I - \delta), \tag{7}$$

$$I(S(W); \{J(i)\}_{i=1}^{M_I}) \leq \delta, \tag{8}$$

$$I(X^N(W); \{J(i)\}_{i=1}^{M_I}) \leq N(R_L + \delta), \tag{9}$$

$$\log M_J \leq N(R_J + \delta). \tag{10}$$

Let $\mathcal{R}_{\mathrm{g}}$ denote the capacity region that contains *all* achievable rate tuples (R_I, R_S, R_L, R_J) in a generated secret key system.

Note that (8) ensures that strong secrecy holds, while [10] requires weak secrecy only. Since strong secrecy implies weak secrecy, by proving the achievability of strong secrecy, we have also shown that weak secrecy can be achieved.

Equations (8) and (9) describe the secrecy preservation and privacy preservation requirements against an attacker that has access to the public database, respectively: (8) requires that the secrecy leakage is sufficiently small such that the attacker can only infer a negligible amount of information about the secret key from the public helper database; (9) requires that the privacy leakage rate is bounded by the privacy leakage rate R_L such that the attacker cannot infer more than the amount $N(R_L+\delta)$ of information about the biometric sequence from the public helper database.

Theorem 1 ([10]): The capacity region of a generated secret key system with weak secrecy is given by

$$\begin{aligned}\mathcal{R}_{\mathrm{g}} = \{&(R_I, R_S, R_L, R_J) : R_I + R_S \le I(U;Y),\\ &R_L \ge I(U;X) - I(U;Y) + R_I,\\ &R_J \ge I(U;\tilde{X}) - I(U;Y) + R_I,\\ &\text{for some } P_{UX\tilde{X}Y} = Q_X Q_{\tilde{X}|X} Q_{Y|X} P_{U|\tilde{X}}\}. \end{aligned} \quad (11)$$

From Theorem 1, we see that the capacity region of the generated secret keys system involves a test channel $P_{U|\tilde{X}}$ and the corresponding auxiliary RV U. We also see that, given a test channel, the minimal achievable privacy leakage rate relates to the mutual information $I(U;X)$ while the minimal achievable helper data rate relates to $I(U;\tilde{X})$. This difference can be interpreted as that the privacy leakage is characterized with respect to the original biometric source instead of the noisy version of it. Therefore, the original biometric source RV X is involved when characterizing R_L. While for the helper data, they are generated based on the noisy enrollment of the biometric sequence, and hence $\tilde{X}$ is involved in the characterization of R_J.

According to (6) in Definition 1, we obtain that given a secret key rate R_S, the maximal entropy of the secret key is $H(S(W)) = N(R_S - \delta)$, where $\delta > 0$ is a sufficiently small number. Moreover, from (11), we obtain that the maximal secret key rate that can be achieved is $R_S = I(U;Y)$ by setting $R_I = 0$ and U is a RV satisfying $U - \tilde{X} - X - Y$. By setting $U = \tilde{X}$, the maximal achievable secret key rate is $R_S = I(\tilde{X};Y)$. Consequently, the maximal entropy of the secret key is $H(S(W)) = N(I(\tilde{X};Y) - \delta)$, which directly relates to the entropy of the noisy biometrics. As investigated in the literature, biometrics-based secret keys give higher entropy than other authentication factors. For instance, it is shown that the entropy of biometric EEG feature is at best 83 bits [30]; human face template is at best 75 bits [31]; human voice is 18-30 bits [32]. However, the entropy of other authentication methods, e.g. PINs and passwords are much smaller than the entropy of biometric features. For example, it is shown that entropy of human chosen 4-digit PINs is 8.41 bits and 6-digit PINs 13.21 bits [33]; human chosen passwords is 20-22 bits [34]. Therefore, biometrics-based authentication is considered more secure than other authentication factors. Experimenting with actual biometric data and analyzing the actual entropies can be an interesting next research step.

C. Chosen Secret Key System

A chosen secret key system is illustrated in Fig. 2. For each user $w \in [1 : M_I]$, the secret key $s(w)$ is generated uniformly at random from $[1 : M_S]$ and independent of both the biometric sequences and the private keys. In the enrollment phase, the enrollment mapping $e(\cdot)$ decides on the helper data $j(w) \in [1 : M_J]$ using $\tilde{x}^N(w)$ and $s(w)$, i.e.,

$$j(w) = e(\tilde{x}^N(w), s(w)), \quad (12)$$

where $j(w)$ and $s(w)$ are stored in a public helper database and a secure database at the location w, respectively.

The identification and authentication procedures are similar to that of the generated secret key system. After an unknown

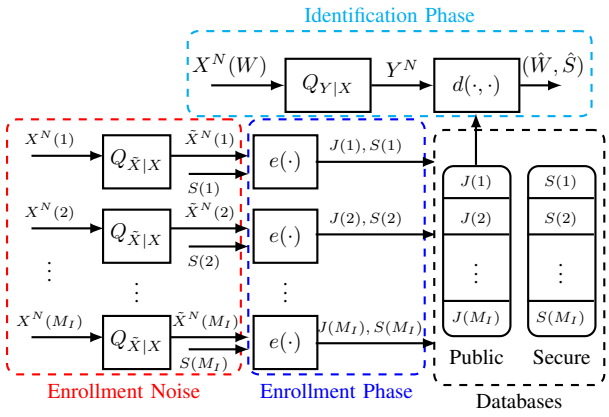

Fig. 2: Model of a chosen secret key system.

user w, which is a realization of a RV W uniformly distributed on $[1 : M_I]$, is observed, and noisy observation y^N is generated, the identification mapping $d(\cdot,\cdot)$ uses y^N and $\{j(i)\}_{i=1}^{M_I}$ to guess the user index denoted as $\hat{w}$ and estimate the secret key denoted as $\hat{s}$, i.e.,

$$(\hat{w}, \hat{s}) = d(y^N, \{j(i)\}_{i=1}^{M_I}). \quad (13)$$

During authentication, the system operates the same authentication procedure as in the generated secret key system.

The achievability of a rate tuple in a chosen secret key system is defined as follows.

Definition 2: A rate tuple $(R_I, R_S, R_L, R_J) \in \mathbb{R}_+^4$ of the identification rate, the secret key rate, the privacy leakage rate, and the helper data rate is *achievable* in a chosen secret key system if for all $\delta > 0$ there exists some $N_0(\delta) \ge 1$ such that for all $N \ge N_0(\delta)$ there exists enrollment and identification mappings such that the conditions in (5), (7), (8), (9), (10), and

$$\log M_S \ge N(R_S - \delta) \quad (14)$$

are satisfied. We use $\mathcal{R}_{\mathrm{c}}$ to denote the capacity region that contains *all* achievable rate tuples (R_I, R_S, R_L, R_J) in a chosen secret key system.

As before, (8) ensures that strong secrecy holds, while [10] requires weak secrecy only.

Theorem 2 ([10]): The capacity region of a chosen secret key system with weak secrecy is given by

$$\begin{aligned}\mathcal{R}_{\mathrm{c}} = \{&(R_I, R_S, R_L, R_J) : R_I + R_S \le I(U;Y),\\ &R_L \ge I(U;X) - I(U;Y) + R_I,\\ &R_J \ge I(U;\tilde{X}),\\ &\text{for some } P_{UX\tilde{X}Y} = Q_X Q_{\tilde{X}|X} Q_{Y|X} P_{U|\tilde{X}}\}. \end{aligned} \quad (15)$$

D. Generated Secret Key System with Zero Leakage

A generated secret key system with zero leakage is depicted in Fig. 3. Compared to the generated secret key system in Fig. 1, a private key $p(w)$ is additionally provided for each user $w \in [1 : M_I]$. We assume that the private key is uniformly distributed on $[1 : M_P]$ and independent of the user index and the biometric information. The private key is also later used

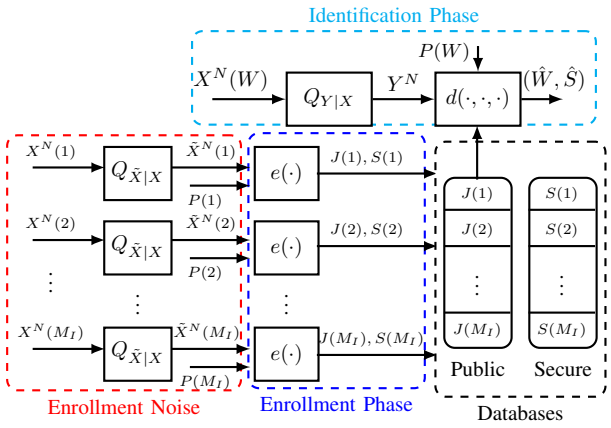

Fig. 3: Model of a generated secret key system with zero leakage.

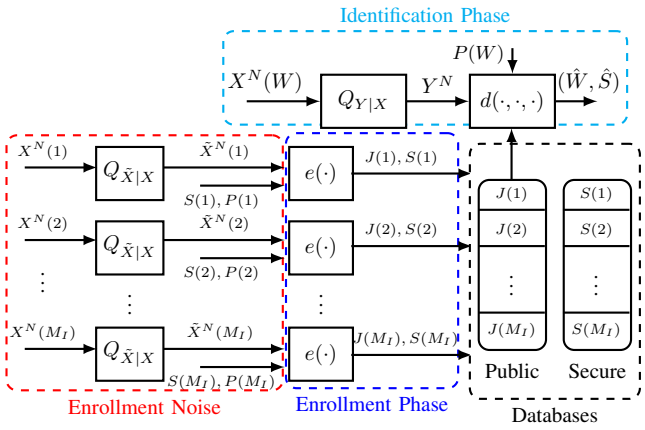

Fig. 4: Model of a chosen secret key system with zero leakage.

for identification and authentication. In the enrollment phase, the enrollment mapping $e(\cdot)$ maps $\tilde{x}^N(w)$ and $p(w)$ onto the helper data $j(w) \in [1 : M_J]$ and secret key $s(w) \in [1 : M_S]$, i.e.,

$$(j(w), s(w)) = e(\tilde{x}^N(w), p(w)), \tag{16}$$

where $j(w)$ and $s(w)$ are stored in a public helper database and a secure database at the location w, respectively.

In the identification phase, an unknown user w, which is a realization of a RV W uniformly distributed on $[1 : M_I]$, is observed via the observation channel and the noisy observation y^N is generated. The user provides its own private key $p(w)$ to the system. Then the identification mapping $d(\cdot,\cdot,\cdot)$ uses y^N, $p(w)$, and $\{j(i)\}_{i=1}^{M_I}$ to guess the user index denoted as $\hat{w}$ and estimate the secret key denoted as $\hat{s}$, i.e.,

$$(\hat{w}, \hat{s}) = d(y^N, p(w), \{j(i)\}_{i=1}^{M_I}). \tag{17}$$

During authentication, the system operates the same authentication procedure as in the generated secret key system.

The achievability of a rate tuple in a generated secret key system with zero leakage is defined as follows.

Definition 3: A rate tuple $(R_I, R_S, R_P, R_J) \in \mathbb{R}_+^4$ of the identification rate, the secret key rate, the private key rate, and the helper data rate is *achievable* in a generated secret key system with zero leakage if for all $\delta > 0$ there exists some $N_0(\delta) \geq 1$ such that for all $N \geq N_0(\delta)$ there exists enrollment and identification mappings such that the conditions in (5), (6), (7), (8), (10), and

$$\log M_P \leq N(R_P + \delta), \tag{18}$$
$$I(X^N(W); \{J(i)\}_{i=1}^{M_I}) \leq \delta \tag{19}$$

are satisfied. We use $\mathcal{R}_{\mathrm{g}}^0$ to denote the capacity region that contains *all* achievable rate tuples (R_I, R_S, R_P, R_J) in a generated secret key system with zero leakage.

Again, note that (8) ensures that strong secrecy holds, while [11] requires weak secrecy only.

Theorem 3 ([11]): The capacity region of a generated secret key system with weak measures and zero leakage is given by

$$\mathcal{R}_{\mathrm{g}}^0 = \{(R_I, R_S, R_P, R_J) : R_I + R_S \leq R_P + I(U;Y),$$
$$R_P \geq I(U;X) - I(U;Y) + R_I,$$
$$R_J \geq I(U;\tilde{X}) - I(U;Y) + R_I,$$
$$\text{for some } P_{UX\tilde{X}Y} = Q_X Q_{\tilde{X}|X} Q_{Y|X} P_{U|\tilde{X}}\}. \tag{20}$$

E. Chosen Secret Key System with Zero Leakage

A chosen secret key system with zero leakage is depicted in Fig. 4. For each user $w \in [1 : M_I]$, the secret key $s(w)$ is generated uniformly at random from $[1 : M_S]$ and independent of both the biometric sequences and the private keys. A private key $p(w)$ is additionally provided for each user, which is uniformly distributed on $[1 : M_P]$ and independent of the user index, the biometric sequences and the chosen secret keys. The private key is also later used for identification and authentication. In the enrollment phase, the enrollment mapping $e(\cdot)$ maps $\tilde{x}^N(w)$, $p(w)$, and $s(w)$ onto the helper data $j(w) \in [1 : M_J]$, i.e.,

$$j(w) = e(\tilde{x}^N(w), p(w), s(w)), \tag{21}$$

where $j(w)$ and $s(w)$ are stored in a public helper database and a secure database at the location w, respectively.

In the identification phase, an unknown user w, which is a realization of a RV W uniformly distributed on $[1 : M_I]$, is observed via the observation channel and noisy observation y^N is generated. In the meantime, the user provides its own private key $p(w)$ to the system. The identification mapping $d(\cdot,\cdot,\cdot)$ uses y^N, $p(w)$, and $\{j(i)\}_{i=1}^{M_I}$ to guess the user index denoted as $\hat{w}$ and estimate the secret key denoted as $\hat{s}$, i.e.,

$$(\hat{w}, \hat{s}) = d(y^N, p(w), \{j(i)\}_{i=1}^{M_I}). \tag{22}$$

During authentication, the system operates the same authentication procedure as in the generated secret key system.

The achievability of a rate tuple in a chosen secret key system with zero leakage is defined as follows.

Definition 4: A rate tuple $(R_I, R_S, R_P, R_J) \in \mathbb{R}_+^4$ of the identification rate, the secret key rate, the private key rate, and the helper data rate is *achievable* in a chosen secret key system with zero leakage if for all $\delta > 0$ there exists some $N_0(\delta) \geq 1$ such that for all $N \geq N_0(\delta)$ there exists enrollment and identification mappings such that the conditions in (5), (7), (8), (10), (14), (18), (19) are satisfied. Let $\mathcal{R}_{\mathrm{c}}^0$ denote

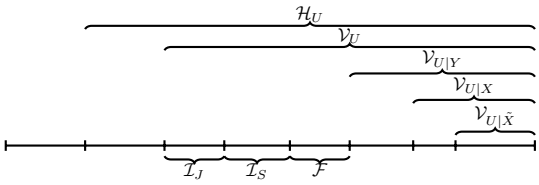

Fig. 5: An illustration of the subset structure of the index sets $\mathcal{V}_{U|\tilde{X}}$, $\mathcal{V}_{U|X}$, $\mathcal{V}_{U|Y}$, $\mathcal{V}_U$, $\mathcal{H}_U$, $\mathcal{F}$, $\mathcal{I}_S$, and $\mathcal{I}_J$. The line represents the indices, which are not ordered from 1 to N.

the capacity region that contains *all* achievable rate tuples (R_I, R_S, R_P, R_J) in a chosen secret key system with zero leakage.

As before, we require strong secrecy, which is a more stringent secrecy constraint than that in [11].

Theorem 4 ([11]): The capacity region of a chosen secret key system with weak measures and zero leakage is given by

$$\begin{aligned}\mathcal{R}_{\mathrm{c}}^0 = \{&(R_I, R_S, R_P, R_J) : R_I + R_S \le R_P + I(U;Y),\\ &R_P \ge I(U;X) - I(U;Y) + R_I,\\ &R_J \ge I(U;\tilde{X}) - I(U;Y) + R_I + R_S,\\ &\text{for some } P_{UX\tilde{X}Y} = Q_X Q_{\tilde{X}|X} Q_{Y|X} P_{U|\tilde{X}}\}. \end{aligned} \quad (23)$$

III. Polar Code Design

In this section, we present polar code designs that achieve the fundamental limits[1] given in Theorems 1, 2, 3, and 4.

A. Polar Code Design for Generated Secret Key System

Codebook Generation: Let block length $N \triangleq 2^n$, where $n \in \mathbb{N}^+$. Introduce auxiliary RVs U and V. Since we consider the binary case with memoryless symmetric channels in this work, both U and V are assumed to be binary. Fix an auxiliary p.m.f. $P_{U|\tilde{X}}$[2], which results in a joint p.m.f. $P_{UX\tilde{X}Y} = P_{U|\tilde{X}} Q_{\tilde{X}|X} Q_{Y|X} Q_X$. Fix a sufficiently small $\epsilon > 0$. Assume that there are $M_I = 2^{NR_I}$ users. For each user $w \in [1 : M_I]$, generate a vector $u^N(w)$:

$$u^N(w) \sim \prod_{i=1}^{N} P_{U|\tilde{X}}(u_i(w)|\tilde{x}_i(w)). \quad (24)$$

Denote $v^N(w)$ as the polar-code transformed vector, i.e.,

$$v^N(w) = u^N(w) G_N. \quad (25)$$

For $\delta_N \triangleq 2^{-N^\beta}$ with an arbitrary but fixed $\beta \in (0, 1/2)$, define the following sets on $[1 : N]$:

$$\mathcal{H}_U \triangleq \{i : H(V^i|V^{i-1}) \ge \delta_N\}, \quad (26)$$

$$\mathcal{V}_U \triangleq \{i : H(V^i|V^{i-1}) \ge 1 - \delta_N\} \subset \mathcal{H}_U, \quad (27)$$

$$\mathcal{V}_{U|Y} \triangleq \{i : H(V^i|V^{i-1}, Y^N) \ge 1 - \delta_N\} \subset \mathcal{V}_U, \quad (28)$$

$$\mathcal{V}_{U|X} \triangleq \{i : H(V^i|V^{i-1}, X^N) \ge 1 - \delta_N\} \subset \mathcal{V}_{U|Y}, \quad (29)$$

$$\mathcal{V}_{U|\tilde{X}} \triangleq \{i : H(V^i|V^{i-1}, \tilde{X}^N) \ge 1 - \delta_N\} \subset \mathcal{V}_{U|X}, \quad (30)$$

where $Y^N \sim \prod_{i=1}^N Q_{Y|X}(y_i|x_i(w))$ and the inclusions $\mathcal{V}_{U|\tilde{X}} \subset \mathcal{V}_{U|X} \subset \mathcal{V}_{U|Y}$ are due to the Markov chain $(U, V) - \tilde{X} - X - Y$. We pick a rate pair (R_I, R_S) such that

$$N(R_I + R_S + 2\epsilon) = |\mathcal{V}_U \backslash \mathcal{V}_{U|Y}|. \quad (31)$$

Next we define the following sets.

Definition 5: Let $\mathcal{F}$ be a subset of $\mathcal{V}_U \backslash \mathcal{V}_{U|Y}$ such that

$$|\mathcal{F}| = N\epsilon, \quad (32)$$

$$H(V^i|V^{i-1}, Y^N) \ge H(V^j|V^{j-1}, Y^N), \quad (33)$$

hold for all $i \in \mathcal{F}$ and any $j \in \mathcal{V}_U \backslash (\mathcal{V}_{U|Y} \cup \mathcal{F})$. That is, $\mathcal{F}$ includes the indices in $\mathcal{V}_U \backslash \mathcal{V}_{U|Y}$ with the largest conditional entropy. Let $\mathcal{I}_S$ be any subset of $\mathcal{V}_U \backslash (\mathcal{V}_{U|Y} \cup \mathcal{F})$ with size NR_S. Let $\mathcal{I}_J$ denote $\mathcal{V}_U \backslash (\mathcal{V}_{U|Y} \cup \mathcal{F} \cup \mathcal{I}_S)$ with size satisfying

$$|\mathcal{I}_J| = |\mathcal{V}_U \backslash \mathcal{V}_{U|Y}| - N(\epsilon + R_S) = N(R_I + \epsilon). \quad (34)$$

The sets defined above have been illustrated in Fig. 5.

The above index sets of $v^N(w)$ can be interpreted as follows: (a) $\mathcal{H}_U$ is the index set of not sufficiently close to zero entropy given the previous bits; (b) $\mathcal{V}_U$ is the high entropy index set given the previous bits; (c) $\mathcal{V}_{U|Y}$ is the high entropy index set given the previous bits and Y^N; (d) $\mathcal{V}_{U|X}$ is the high entropy index set given the previous bits and X^N; (e) $\mathcal{V}_{U|\tilde{X}}$ is the high entropy index set given the previous bits and $\tilde{X}^N$; (f) $\mathcal{F}$ is a small set to ensure the source polar coding works; (g) the bits in the sequence $v^N(w)$ related to indices in $\mathcal{I}_J$ and $\mathcal{I}_S$ are used for generating the helper data and the secret key, respectively.

Enrollment: The enrollment procedure is given in Algorithm 1. The system first constructs a source representation $\tilde{v}^N(w)$ of $\tilde{x}^N(w)$ for each user $w \in [1 : M_I]$. The system uses part of the bits in the sequence $\tilde{v}^N(w)$ to generate the helper data $j(w)$ and the secret key $s(w)$, where the helper data $j(w)$ consists two parts, i.e., $j_1(w)$ and $j_2(w)$.

Identification and Authentication: The identification and authentication procedure is given in Algorithm 2. After observing y^N generated by an unknown user, the system iteratively constructs an estimate $\hat{v}^N(\hat{w})$ of $\tilde{v}^N(\hat{w})$ for all $\hat{w} \in [1 : M_I]$ using the first part $j_1(\hat{w})$ of the public helper data and the vector $\mathbf{a}$ defined in Algorithm 1. The iterative operation is based on the successive cancellation decoder as described in [20] by successively calculating the likelihood ratio for bit-wise decoding. That is, the bit $\hat{v}_i$ is decoded based on $(y^N, \hat{v}^{i-1})$ and then the bit $\hat{v}_{i+1}$ is decoded based on $(y^N, \hat{v}^i)$. With the estimated sequence $\hat{v}^N(\hat{w})$, the system compares $\hat{v}^N(\hat{w})$ with the second part $j_2(\hat{w})$ of the public helper data. If they match, i.e., $j_2(\hat{w}) = \hat{v}^N[\mathcal{I}_J](\hat{w})$, the system outputs the guessed user index $\hat{w}$ and the estimated secret key $\hat{s}$. Otherwise, the system continues comparing with the next user when $\hat{w} < M_I$ or reports an error when $\hat{w} = M_I$.

[1]Since polar codes are designed for block length $N = 2^n$ with $n \in \mathbb{N}^+$, we formally do not provide codes for any block length, but still for infinitely many block lengths, which corresponds to achievability in the optimistic sense.

[2]The distribution $P_{U|\tilde{X}}$ is the so called test channel. It is the designing freedom, i.e., any $P_{U|\tilde{X}}$ such that the Markov chain $U - \tilde{X} - X - Y$ holds will result in a polar code with the corresponding performances as stated in Theorems 5, 6, 7, and 8. The test channel can be chosen according to the desired capacity of the system, e.g. how many users can be enrolled and how long the secret key can be.

$$p_{\tilde{V}^{j}(w)|\tilde{V}^{j-1}(w),\,\tilde{X}^{N}(w)}(\tilde{v}^{j}(w)|\tilde{v}^{j-1}(w),\tilde{x}^{N}(w))=\begin{cases}p_{V^{j}(w)|V^{j-1}(w),\,\tilde{X}^{N}(w)}(\tilde{v}^{j}(w)|\tilde{v}^{j-1}(w),\tilde{x}^{N}(w)),\ \text{if } j\in\mathcal{H}_U\backslash\mathcal{V}_{U|\tilde{X}}\\ p_{V^{j}(w)|V^{j-1}(w)}(\tilde{v}^{j}(w)|\tilde{v}^{j-1}(w)),\ \text{if } j\in\mathcal{H}_U^c\end{cases}\tag{35}$$

Algorithm 1 Enrollment of the Generated Secret Key System

Input: Biometric sequence $\tilde{x}^N(w)$ for each user $w\in[1:M_I]$; a rate pair (R_I,R_S), the sets $\mathcal{F}$, $\mathcal{I}_S$, and $\mathcal{I}_J$ defined in Definition 5; vector $\mathbf{a}$, which is a realization of a binary uniformly distributed RV of size $|\mathcal{V}_{U|\tilde{X}}|$.

Output: Secret key $s(w)$ and helper data $j(w)$ for all users $w\in[1:M_I]$.

for $w=1:M_I$ **do**
  $\tilde{v}^N[\mathcal{V}_{U|\tilde{X}}](w)\leftarrow\mathbf{a}$
  Given $\tilde{x}^N(w)$, successively draw the remaining bits of $\tilde{v}^N(w)$ according to (35).
  $j_1(w)\leftarrow\tilde{v}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}})\cup\mathcal{F}](w)$
  $j_2(w)\leftarrow\tilde{v}^N[\mathcal{I}_J](w)$
  $s(w)\leftarrow\tilde{v}^N[\mathcal{I}_S](w)$
  Store $j(w)=(j_1(w),j_2(w))$ and $s(w)$ in the public and secure databases at location w, respectively.
return $\{j(w)\}_{w=1}^{M_I},\{s(w)\}_{w=1}^{M_I}$

Algorithm 2 Identification and Authentication of the Generated Secret Key System

Input: Observation sequence y^N; the public helper database $\{j(w)\}_{w=1}^{M_I}$; the vector $\mathbf{a}$ and sets $\mathcal{F}$, $\mathcal{I}_S$, and $\mathcal{I}_J$ from Algorithm 1.

Output: Guessed user index $\hat{w}$ and estimated secret key $\hat{s}$.

$\hat{w}\leftarrow 0$
do
  $\hat{w}\leftarrow\hat{w}+1$
  $\hat{v}^N[\mathcal{V}_{U|\tilde{X}}](\hat{w})\leftarrow\mathbf{a}$
  $\hat{v}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}})\cup\mathcal{F}](\hat{w})\leftarrow j_1(\hat{w})$
  Given y^N and $\hat{v}^N[\mathcal{V}_{U|Y}\cup\mathcal{F}](\hat{w})$, obtain an estimate $\hat{v}^N(\hat{w})$ of $\tilde{v}^N(\hat{w})$ with successive cancellation decoder of [20].
while $\hat{v}^N[\mathcal{I}_J](\hat{w})\neq j_2(\hat{w})$ and $\hat{w}\leq M_I$
if $\hat{v}^N[\mathcal{I}_J](\hat{w})=j_2(\hat{w})$ **then**
  return $\hat{w},\hat{s}\leftarrow\hat{v}^N[\mathcal{I}_S](\hat{w})$
else
  return error

The performance of the algorithms is ensured as follows.

Theorem 5: In a generated secret key system, any rate tuple $(R_I,R_S,R_L,R_J)\in\mathcal{R}_{\mathrm{g}}$ can be achieved by the polar code design in Algorithm 1 and Algorithm 2, whose complexity is $\mathcal{O}(M_IN\log N)$.

Proof: See Appendix A-A.

B. Polar Code Design for Chosen Secret Key System

The enrollment algorithm is given in Algorithm 3, which extends Algorithm 1 with a masking procedure. The authentication and identification algorithm is given in Algorithm 4.

Algorithm 3 Enrollment of the Chosen Secret Key System

Input: Biometric sequence $\tilde{x}^N(w)$ for each user $w\in[1:M_I]$; a rate pair (R_I,R_S) and the sets $\mathcal{F}$, $\mathcal{I}_S$, and $\mathcal{I}_J$ defined in Definition 5; chosen secret keys $\{s(w)\}_{w=1}^{M_I}$; vector $\mathbf{a}$, which is a realization of uniformly distributed RV of size $|\mathcal{V}_{U|\tilde{X}}|$.

Output: Secret key $s(w)$ and helper data $j(w)$ for each user $w\in[1:M_I]$.

for $w=1:M_I$ **do**
  $\tilde{v}^N[\mathcal{V}_{U|\tilde{X}}](w)\leftarrow\mathbf{a}$
  Given $\tilde{x}^N(w)$, successively draw the remaining bits of $\tilde{v}^N(w)$ according to (35).
  $j_1(w)\leftarrow\tilde{v}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}})\cup\mathcal{F}](w)$
  $j_2(w)\leftarrow\tilde{v}^N[\mathcal{I}_J](w)$
  $j_3(w)\leftarrow s(w)\oplus\tilde{v}^N[\mathcal{I}_S](w)$
  Store $j(w)=(j_1(w),j_2(w),j_3(w))$ and $s(w)$ in the public and secure databases at location w, respectively.
return $\{j(w)\}_{w=1}^{M_I},\{s(w)\}_{w=1}^{M_I}$

Algorithm 4 Identification and Authentication of the Chosen Secret Key System

Input: Observation sequence y^N; the public helper database $\{j(w)\}_{w=1}^{M_I}$; the vector $\mathbf{a}$ and sets $\mathcal{F}$, $\mathcal{I}_S$, and $\mathcal{I}_J$ from Algorithm 3.

Output: Guessed user index $\hat{w}$ and estimated secret key $\hat{s}$.

$\hat{w}\leftarrow 0$
do
  $\hat{w}\leftarrow\hat{w}+1$
  $\hat{v}^N[\mathcal{V}_{U|\tilde{X}}](\hat{w})\leftarrow\mathbf{a}$
  $\hat{v}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}})\cup\mathcal{F}](\hat{w})\leftarrow j_1(\hat{w})$
  Given Y^N and $\hat{v}^N[\mathcal{V}_{U|Y}\cup\mathcal{F}](\hat{w})$, obtain an estimate $\hat{v}^N(\hat{w})$ of $\tilde{v}^N(\hat{w})$ with successive cancellation decoder of [20].
while $\hat{v}^N[\mathcal{I}_J](\hat{w})\neq j_2(\hat{w})$ and $\hat{w}\leq M_I$
if $\hat{v}^N[\mathcal{I}_J](\hat{w})=j_2(\hat{w})$ **then**
  return $\hat{w},\hat{s}\leftarrow j_3(\hat{w})\ominus\hat{v}^N[\mathcal{I}_S](\hat{w})$
else
  return error

The performance of the algorithms is ensured as follows.

Theorem 6: In a chosen secret key system, any rate tuple $(R_I,R_S,R_L,R_J)\in\mathcal{R}_{\mathrm{c}}$ can be achieved by the polar code design in Algorithm 3 and Algorithm 4, whose complexity is $\mathcal{O}(M_IN\log N)$.

Proof: See Appendix A-B.

Algorithm 5 Enrollment of the Generated Secret Key System with Zero Leakage

Input: Biometric sequence $\tilde{x}^N(w)$ for each user $w \in [1 : M_I]$; a rate pair (R_I, R_S) and the sets $\mathcal{F}$ and $\mathcal{I}$ defined in Definition 6; the private key $p(w) = (p_1(w), p_2(w), p_3(w))$, where $p_1(w) \in [1 : 2^{|(\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|X}) \cup \mathcal{F}|}]$, $p_2(w) \in [1 : 2^{N(R_I+\epsilon)}]$, and $p_3(w) \in [1 : 2^{N(R_P - R_I - \epsilon) - |(\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|X}) \cup \mathcal{F}|}]$; vector $\mathbf{a}$, which is a realization of uniformly distributed over $|\mathcal{V}_{U|\tilde{X}}|$.

Output: Secret key $s(w)$ and helper data $j(w)$ for each user $w \in [1 : M_I]$.

for $w = 1 : M_I$ **do**
  $\tilde{v}^N[\mathcal{V}_{U|\tilde{X}}](w) \leftarrow \mathbf{a}$
  Given $\tilde{x}^N(w)$, successively draw the remaining bits of $\tilde{v}^N(w)$ according to (35).
  $j_{11}(w) \leftarrow \tilde{v}^N[(\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|X}) \cup \mathcal{F}](w) \oplus p_1(w)$
  $j_{12}(w) \leftarrow \tilde{v}^N[\mathcal{V}_{U|X} \backslash \mathcal{V}_{U|\tilde{X}}](w)$
  $j_2(w) \leftarrow p_2(w)$
  $s(w) \leftarrow \tilde{v}^N[\mathcal{V}_U \backslash \mathcal{V}_{U|X}](w), p_3(w))$
  Store $j(w) = (j_{11}(w), j_{12}(w), j_2(w))$ and $s(w)$ in the public and secure databases at location w, respectively.
return $\{j(w)\}_{w=1}^{M_I}, \{s(w)\}_{w=1}^{M_I}$

C. Polar Code Design for Generated Secret Key System with Zero Leakage

Pick a rate triple (R_I, R_S, R_P) satisfying

$$NR_P \geq |\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|X}| + N(R_I + 2\epsilon), \tag{36}$$

$$N(R_I + R_S + 2\epsilon) = NR_P + |\mathcal{V}_U \backslash \mathcal{V}_{U|Y}|. \tag{37}$$

Similar to Definition 5, we define a set $\mathcal{F}$ as follows.

Definition 6: Let $\mathcal{F}$ be a subset of $\mathcal{V}_U \backslash \mathcal{V}_{U|Y}$ such that (32) and (33) hold for any $i \in \mathcal{F}$ and any $j \in \mathcal{V}_U \backslash (\mathcal{V}_{U|Y} \cup \mathcal{F})$. That is, $\mathcal{F}$ includes the indices with the largest conditional entropy.

The enrollment algorithm is given in Algorithm 5, which extends Algorithm 1 by including the private key with three purposes: (i) masking part of the helper data to ensure close to zero private leakage rate; (ii) constructing public helper data to identify more users; (iii) generating a longer secret key to increase the secret key rate. The authentication and identification algorithm is given in Algorithm 6.

The performance of the algorithms is ensured as follows.

Theorem 7: In a generated secret key system with zero leakage, any rate tuple $(R_I, R_S, R_P, R_J) \in \mathcal{R}_{\mathrm{g}}^0$ can be achieved by the polar code design in Algorithm 5 and Algorithm 6, whose complexity is $\mathcal{O}(N \log N)$.

Proof: See Appendix A-C.

D. Polar Code Design for Chosen Secret Key System with Zero Leakage

The polar code design of the chosen secret key system with zero leakage is a similar extension of the chosen secret key system as done for the generated secret key system in Sec III-B. The enrollment algorithm is given in Algorithm 7. The authentication and identification algorithm is given in Algorithm 8.

Algorithm 6 Identification and Authentication of the Generated Secret Key System with Zero Leakage

Input: Observation sequence y^N; the private key $p = (p_1, p_2, p_3)$ of the observed user; the public helper database $\{j(w)\}_{w=1}^{M_I}$; the vector $\mathbf{a}$ and sets $\mathcal{F}$ and $\mathcal{I}$ from enrollment.

Output: Guessed user index $\hat{w}$ and estimated secret key $\hat{s}$.

$\hat{w} \leftarrow 0$
do
  $\hat{w} \leftarrow \hat{w} + 1$
while $p_2 \neq j_2(\hat{w})$ and $\hat{w} \leq M_I$
if $p_2 = j_2(\hat{w})$ **then**
  $\hat{v}^N[\mathcal{V}_{U|\tilde{X}}](\hat{w}) \leftarrow \mathbf{a}$
  $\hat{v}^N[(\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|X}) \cup \mathcal{F}](\hat{w}) \leftarrow j_{11}(\hat{w}) \ominus p_1$
  $\hat{v}^N[\mathcal{V}_{U|X} \backslash \mathcal{V}_{U|\tilde{X}}](\hat{w}) \leftarrow j_{12}(\hat{w})$
  Given y^N and $\hat{v}^N[\mathcal{V}_{U|Y} \cup \mathcal{F}](\hat{w})$, obtain an estimate $\hat{v}^N(\hat{w})$ of $\tilde{v}^N(\hat{w})$ with successive cancellation decoder of [20].
  return $\hat{w}, \hat{s} \leftarrow \left(\hat{v}^N[\mathcal{V}_U \backslash \mathcal{V}_{U|X}](\hat{w}), p_3\right)$
else
  return error

Algorithm 7 Enrollment of the Chosen Secret Key System with Zero Leakage

Input: Biometric sequence $\tilde{x}^N(w)$ and the secret key $s(w)$ for each user $w \in [1 : M_I]$; a rate pair (R_I, R_S) and the sets $\mathcal{F}$ and $\mathcal{I}$ defined in Definition 6; the private key $p(w) = (p_1(w), p_2(w), p_3(w))$, where $p_1(w) \in [1 : 2^{|(\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|X}) \cup \mathcal{F}|}]$, $p_2(w) \in [1 : 2^{N(R_I+\epsilon)}]$, and $p_3(w) \in [1 : 2^{N(R_P - R_I - \epsilon) - |(\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|X}) \cup \mathcal{F}|})]$; vector $\mathbf{a}$, which is a realization of uniformly distributed over $|\mathcal{V}_{U|\tilde{X}}|$.

Output: Helper data $j(w)$ for each user $w \in [1 : M_I]$.

for $w = 1 : M_I$ **do**
  $\tilde{v}^N[\mathcal{V}_{U|\tilde{X}}](w) \leftarrow \mathbf{a}$
  Given $\tilde{x}^N(w)$, successively draw the remaining bits of $\tilde{v}^N(w)$ according to (35).
  $j_{11}(w) \leftarrow \tilde{v}^N[(\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|X}) \cup \mathcal{F}](w) \oplus p_1(w)$
  $j_{12}(w) \leftarrow \tilde{v}^N[\mathcal{V}_{U|X} \backslash \mathcal{V}_{U|\tilde{X}}](w)$
  $j_2(w) \leftarrow p_2(w)$
  $j_3(w) \leftarrow s(w) \oplus \left(\tilde{v}^N[\mathcal{V}_U \backslash \mathcal{V}_{U|X}](w), p_3(w)\right)$
  Store $j(w) = (j_{11}(w), j_{12}(w), j_2(w), j_3(w))$ and $s(w)$ in the public and secure databases at location w, respectively.
return $\{j(w)\}_{w=1}^{M_I}, \{s(w)\}_{w=1}^{M_I}$

The performance of the algorithms is ensured as follows.

Theorem 8: In a chosen secret key system with zero leakage, any rate tuple $(R_I, R_S, R_P, R_J) \in \mathcal{R}_{\mathrm{c}}^0$ can be achieved by the polar code design in Algorithm 7 and Algorithm 8, whose complexity is $\mathcal{O}(N \log N)$.

Proof: See Appendix A-D.

IV. Conclusion

In this work, we show that polar codes can be developed for implementation in the considered biometric identification and

Algorithm 8 Identification and Authentication of the Chosen Secret Key System with Zero Leakage

Input: Observation sequence y^N; the private key $p = (p_1, p_2, p_3)$ of the observed user; the public helper database $\{j(w)\}_{w=1}^{M_I}$; the vector $\mathbf{a}$ and sets $\mathcal{F}$ and, $\mathcal{I}$ from enrollment.

Output: Guessed user index $\hat{w}$ and estimated secret key $\hat{s}$.

$\hat{w} \leftarrow 0$
do
  $\hat{w} \leftarrow \hat{w} + 1$
while $p_2 \neq j_2(\hat{w})$ and $\hat{w} \leq M_I$
if $p_2 = j_2(\hat{w})$ **then**
  $\hat{v}^N[\mathcal{V}_{U|\tilde{X}}](\hat{w}) \leftarrow \mathbf{a}$
  $\hat{v}^N[(\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|X}) \cup \mathcal{F}](\hat{w}) \leftarrow j_{11}(\hat{w}) \ominus p_1$
  $\hat{v}^N[\mathcal{V}_{U|X} \backslash \mathcal{V}_{U|\tilde{X}}](\hat{w}) \leftarrow j_{12}(\hat{w})$
  Given y^N and $\hat{v}^N[\mathcal{V}_{U|Y} \cup \mathcal{F}](\hat{w})$, obtain an estimate $\hat{v}^N(\hat{w})$ of $\tilde{v}^N(\hat{w})$ with successive cancellation decoder of [20].
  return $\hat{w}, \hat{s} \leftarrow j_3(\hat{w}) \ominus \left(\hat{v}^N[\mathcal{V}_U \backslash \mathcal{V}_{U|X}](\hat{w}), p_3\right)$
else
  return error

authentication systems. Source polarization is implemented on biometrics to extract helper data and secret keys. Four different but closely related setups are studied depending on the secret key generation method and the privacy preservation requirement, which analogies in the setup are also reflected in the polar code design. The algorithms of the chosen secret key system extend that of the generated secret key system by including a masking procedure. The systems with zero leakage additionally include private keys to ensure close to zero leakage. As the polar coding principle has been shown to achieve the fundamental bounds for various setups, our work shows that they can be also developed for this setting. Moreover, the proposed code designs not only achieve optimal performance but also satisfy a more stringent secrecy preservation requirement, which is achieved for free. This ensures that proposed code designs result in better protection of secrecy. Recently, several methods and approaches have been developed for the design of low-complexity polar coding schemes that outperform state-of-the-art performance. Given those large ongoing research efforts to develop low-complexity polar coding schemes, one can expect that several of those ideas can be transferred to this problem as well, which is a direction to develop the technology further. For this, the polar code designs developed in this work will serve as the basic polar coding principles enabling the design and implementation of efficient low-complexity polar codes for identification and authentication.

Lastly, it is worth mentioning that our code designs achieve optimal performance asymptotically. This is due to the fact that the capacity-achieving performance of polar codes holds for sufficiently long sequence length. Moreover, for simplicity, we consider the binary case of the code design. Fortunately, there is a large amount of work going on for constructing practical polar codes that perform well for short block-lengths and non-binary codes, see e.g. [35]–[37]. Thus although the length of biometric sequences is finite and non-binary in real-life scenarios, with the development of more advanced techniques of constructing short block-length and non-binary polar codes, our work can be developed further to design efficient biometric systems.

APPENDIX A
PROOF OF THEOREMS

A. Proof of Theorem 5

a) Rate Analysis: From (31), we have that

$$\begin{aligned} N(R_I + R_S) &= |\mathcal{V}_U \backslash \mathcal{V}_{U|Y}| - 2N\epsilon \\ &\overset{(a)}{=} |\mathcal{V}_U| - |\mathcal{V}_{U|Y}| - 2N\epsilon, \end{aligned} \tag{38}$$

where (a) holds because $\mathcal{V}_{U|Y} \subset \mathcal{V}_U$.

By [24, Lemma 1], we obtain that

$$\lim_{N\to\infty} \frac{|\mathcal{V}_U|}{N} = H(U). \tag{39}$$

By [20, Theorem 1], we obtain that

$$\lim_{N\to\infty} \frac{|\mathcal{V}_{U|Y}|}{N} = H(U|Y). \tag{40}$$

Combining (38), (39), and (40), we have that

$$\lim_{N\to\infty} R_I + R_S = I(U;Y) - 2\epsilon. \tag{41}$$

The helper data rate R_J is

$$\begin{aligned} R_J = \frac{|J(W)|}{N} &= \frac{|\mathcal{V}_{U|Y} \backslash \mathcal{V}_{U|\tilde{X}}| + |\mathcal{F}| + |\mathcal{I}_J|}{N} \\ &\overset{(a)}{=} \frac{|\mathcal{V}_{U|Y}| - |\mathcal{V}_{U|\tilde{X}}| + N(R_I + 2\epsilon)}{N}, \end{aligned} \tag{42}$$

where (a) follows from $\mathcal{V}_{U|\tilde{X}} \subset \mathcal{V}_{U|Y}$ and (34).

By [20, Theorem 1], we obtain that

$$\lim_{N\to\infty} \frac{|\mathcal{V}_{U|\tilde{X}}|}{N} = H(U|\tilde{X}). \tag{43}$$

Combining (40), (42), and (43), we have that

$$\lim_{N\to\infty} R_J = I(U;\tilde{X}) - I(U;Y) + R_I + 2\epsilon. \tag{44}$$

b) Error Events Analysis: Assume that user W is observed. Let $J(W) = (J_1(W), J_2(W))$ and $S(W)$ denote the actual helper data and the secret key of user W. Define the following error events

$$\begin{aligned} \mathcal{E}_1 &= \{\hat{V}^N[\mathcal{I}_J](W) \neq J_2(W)\}, \\ \mathcal{E}_2 &= \{\exists \hat{w} \neq W : \hat{V}^N[\mathcal{I}_J](\hat{w}) = J_2(\hat{w}), \\ &\qquad \hat{V}^N[\mathcal{I}_S](\hat{w}) = S(W)\}, \\ \mathcal{E}_3 &= \{\hat{V}^N[\mathcal{I}_J](W) = J_2(W), \hat{V}^N[\mathcal{I}_S](W) \neq S(W)\}, \\ \mathcal{E}_4 &= \{\exists \hat{w} \neq W : \hat{V}^N[\mathcal{I}_J](\hat{w}) = J_2(\hat{w}), \\ &\qquad \hat{V}^N[\mathcal{I}_S](\hat{w}) \neq S(W)\}. \end{aligned} \tag{45}$$

The first two events are identification error events that the correct user is not identified: $\mathcal{E}_1$ denotes the case that the true user does not satisfy the equality condition in the comparison procedure; $\mathcal{E}_2$ denotes the case that there exists another user

that satisfies the equality condition and the estimated secret key matches the true one. The third and fourth error events are the authentication errors that the estimated secret key does not match the true one: $\mathcal{E}_3$ denotes the case that the guessed user $\hat{w}$ is the same as the true one but the estimated secret key does not match the true secret key $S(W)$; $\mathcal{E}_4$ denotes the case that neither the guessed user index nor the estimated secret key is the same as the true helper data index and the secret key of the observed user. The identification and authentication are reliable if and only if none of the above events happen. We define the error event

$$\mathcal{E} = \{(\hat{W}, \hat{S}) \neq (W, S(W))\} \tag{46}$$

and it holds that

$$\mathcal{E} \subset \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4. \tag{47}$$

Before proceeding with bounding the probability of $\mathcal{E}$, we include the following lemmas.

Lemma 1: Assume that user w is observed, then it holds that

$$\Pr\{\hat{V}^N(w) \neq \tilde{V}^N(w)\} \xrightarrow{N\to\infty} 0. \tag{48}$$

Proof: See Appendix B-A.

Lemma 2: For any $w \in [1 : M_I]$, $\tilde{V}^N(w)$ resulting from Algorithm 1 has a joint p.m.f. $p_{\tilde{X}^N(w)\tilde{V}^N(w)}$ with $\tilde{X}^N(w)$ such that

$$\mathbb{D}(p_{\tilde{X}^N(w)V^N(w)} || p_{\tilde{X}^N(W)\tilde{V}^N(w)}) \leq N\delta_N. \tag{49}$$

Hence, by Pinsker's inequality [38, p. 44], we have

$$\mathbb{V}(p_{\tilde{X}^N(w)V^N(w)}, p_{\tilde{X}^N(w)\tilde{V}^N(w)}) \leq \sqrt{2\ln 2}\sqrt{N\delta_N}. \tag{50}$$

Consequently, we obtain that

$$\mathbb{D}(p_{V^N(w)} || p_{\tilde{V}^N(w)}) \leq N\delta_N, \tag{51}$$

$$\mathbb{V}(p_{V^N(w)}, p_{\tilde{V}^N(w)}) \leq \sqrt{2\ln 2}\sqrt{N\delta_N}. \tag{52}$$

Proof: See Appendix B-B.

Lemma 3: For any user $w \in [1 : M_I]$ and any $i \in \mathcal{V}_U$, it holds that

$$H(\tilde{V}_i(w)|\tilde{V}^{i-1}(w)) \xrightarrow{N\to\infty} 1.$$

Proof: See Appendix B-C.

In the following, due to symmetry and without loss of generality, assume that $W = 1$. Now we consider the probability of $\mathcal{E}_1 \cup \mathcal{E}_3 | W = 1$ as follows

$$\begin{aligned}
&\Pr\{\mathcal{E}_1 \cup \mathcal{E}_3 | W = 1\} \\
&= \Pr\{(\hat{V}^N[\mathcal{I}_J](1), \hat{V}^N[\mathcal{I}_S](1)) \neq (J_2(1), S(1)) | W = 1\} \\
&= \Pr\{\hat{V}^N[\mathcal{I}_J \cup \mathcal{I}_S](1) \neq \tilde{V}^N[\mathcal{I}_J \cup \mathcal{I}_S](1) | W = 1\} \\
&\xrightarrow{N\to\infty} 0,
\end{aligned} \tag{53}$$

where the last step follows from Lemma 1.

In the following, let $a_1, a_2, ..., a_{|\mathcal{I}_J|}$ denotes the elements of the set $\mathcal{I}_J$ such that $a_1 < a_2 < ... < a_{|\mathcal{I}_J|}$. Let $\mathcal{I}_J^i = \{a_1, a_2, ..., a_i\}$. The probability of $\mathcal{E}_2 \cup \mathcal{E}_4 | W = 1$ can be bounded as follows

$$\begin{aligned}
&\Pr\{\mathcal{E}_2 \cup \mathcal{E}_4 | W = 1\} \\
&= \Pr\{\exists \hat{w} \neq 1 : \hat{V}^N[\mathcal{I}_J](\hat{w}) = J_2(\hat{w}) | W = 1\} \\
&\leq \sum_{\hat{w}\neq 1} \Pr\{\hat{V}^N[\mathcal{I}_J](\hat{w}) = J_2(\hat{w}) | W = 1\} \\
&= \sum_{\hat{w}\neq 1} \Pr\{\forall i \in \mathcal{I}_J : \hat{V}_i(\hat{w}) = \tilde{V}_i(\hat{w}) | W = 1\} \\
&= \sum_{\hat{w}\neq 1} \prod_{i=1}^{|\mathcal{I}_J|} \Pr\{\hat{V}_{a_i}(\hat{w}) = \tilde{V}_{a_i}(\hat{w}) | W = 1, \\
&\qquad \hat{V}^{a_{i-1}}[\mathcal{I}_J^{i-1}](\hat{w}) = \tilde{V}^{a_{i-1}}[\mathcal{I}_J^{i-1}](\hat{w})\}.
\end{aligned} \tag{54}$$

To bound the probability above, we firstly consider the following conditional entropy. For $\hat{w} \neq 1$, we obtain that

$$\begin{aligned}
&H(\hat{V}_{a_i}(\hat{w}) \oplus \tilde{V}_{a_i}(\hat{w}) | \\
&\qquad \hat{V}^{a_{i-1}}[\mathcal{I}_J^{i-1}](\hat{w}) = \tilde{V}^{a_{i-1}}[\mathcal{I}_J^{i-1}](\hat{w}), W = 1) \\
&\overset{(a)}{\geq} H(\hat{V}_{a_i}(\hat{w}) \oplus \tilde{V}_{a_i}(\hat{w}) | \hat{V}^{a_{i-1}}(\hat{w}) = \tilde{V}^{a_{i-1}}(\hat{w}), W = 1) \\
&\overset{(a)}{\geq} H(\hat{V}_{a_i}(\hat{w}) \oplus \tilde{V}_{a_i}(\hat{w}) | \hat{V}^{a_i}(\hat{w}), \tilde{V}^{a_{i-1}}(\hat{w}), Y^N, W = 1) \\
&= H(\tilde{V}_{a_i}(\hat{w}) | \hat{V}^{a_i}(\hat{w}), \tilde{V}^{a_{i-1}}(\hat{w}), Y^N, W = 1) \\
&\overset{(b)}{=} H(\tilde{V}_{a_i}(\hat{w}) | \tilde{V}^{a_{i-1}}(\hat{w}), Y^N, W = 1) \\
&\overset{(c)}{=} H(\tilde{V}_{a_i}(\hat{w}) | \tilde{V}^{a_{i-1}}(\hat{w})) \\
&\overset{(a)}{\geq} H(\tilde{V}_{a_i}(\hat{w}) | \tilde{V}^{a_i - 1}(\hat{w})) \\
&\overset{(d)}{\to} 1,
\end{aligned} \tag{55}$$

as $N \to \infty$; where (a) holds since conditioning reduces entropy; (b) holds because $\hat{V}^{a_i}(\hat{w})$ is a function of $(\tilde{V}^{a_{i-1}}(\hat{w}), Y^N)$ due to the successive cancellation operation; (c) holds because Y^N is the observation of user W and thus is independent of $(X^N(\hat{w}), \tilde{V}^N(\hat{w}))$ for $\hat{w} \neq 1$; (d) follows from Lemma 3. Since $\hat{V}_{a_i}(\hat{w}) \oplus \tilde{V}_{a_i}(\hat{w})$ is binary, (55) implies that for any $\epsilon' > 0$ and sufficiently large N, we have

$$\begin{aligned}
&\Pr\{\hat{V}_{a_i}(\hat{w}) \oplus \tilde{V}_{a_i}(\hat{w}) = 0 | W = 1, \\
&\qquad \hat{V}^{a_{i-1}}[\mathcal{I}_J^{i-1}](\hat{w}) = \tilde{V}^{a_{i-1}}[\mathcal{I}_J^{i-1}](\hat{w})\} < \frac{1}{2} + \epsilon'.
\end{aligned} \tag{56}$$

Now we consider $\epsilon' \in (0, \frac{2^{\frac{\epsilon}{R_I+\epsilon}} - 1}{2})$ for the fixed $\epsilon > 0$. Combining (54) and (56), we obtain that

$$\begin{aligned}
&\Pr\{\mathcal{E}_2 \cup \mathcal{E}_4 | W = 1\} \leq \sum_{\hat{w}\neq 1} \prod_{i=1}^{|\mathcal{I}_J|} (\frac{1}{2} + \epsilon') = 2^{NR_I} (\frac{1}{2} + \epsilon')^{|\mathcal{I}_J|} \\
&\overset{(a)}{=} 2^{NR_I} (\frac{1}{2} + \epsilon')^{N(R_I+\epsilon)} = \frac{(1 + 2\epsilon')^{N(R_I+\epsilon)}}{2^{N\epsilon}} \\
&= \Big(\frac{(1 + 2\epsilon')^{(R_I+\epsilon)}}{2^{\epsilon}}\Big)^N \overset{(b)}{\longrightarrow} 0,
\end{aligned} \tag{57}$$

as $N \to \infty$; where (a) follows from (34); (b) follows from $\frac{(1+2\epsilon')^{(R_I+\epsilon)}}{2^{\epsilon}} < 1$ since $\epsilon' < \frac{2^{\frac{\epsilon}{R_I+\epsilon}} - 1}{2}$.

Combining (53) and (57), we can conclude that

$$\Pr\{\mathcal{E}\} \to 0, \tag{58}$$

when $N \to \infty$. Therefore, we can conclude that there exists a suitable codebook $\mathcal{C} = C$ such that (58) holds.

c) Uniformity of Secret Keys: We include the following lemmas to show the uniformity of secret keys. Let $\mathcal{D}$ denote $\mathcal{V}_U \backslash \mathcal{V}_{U|\tilde{X}}$ for simplicity.

Lemma 4: Let 1-dim distribution $q_{\mathcal{U}_{|\mathcal{D}|}}$ denote a uniform distribution over $[1 : 2^{|\mathcal{D}|}]$. For any user $w \in [1 : M_I]$, the value of the vector $\tilde{V}^N[\mathcal{D}](w)$ is close to uniformly distributed, i.e.,

$$\mathbb{V}(p_{\tilde{V}^N[\mathcal{D}](w))}, q_{\mathcal{U}_{|\mathcal{D}|}}) \leq 2\sqrt{2\ln 2}\sqrt{N\delta_N}. \tag{59}$$

Proof: See Appendix B-D.

Lemma 5: For any user $w \in [1 : M_I]$, it holds that

$$|\mathcal{D}| - H(\tilde{V}^N[\mathcal{D}](w)) \leq \delta, \tag{60}$$

where $\delta \to 0$ as $N \to \infty$. Consequently, for any subset $\mathcal{D}_S$ of $\mathcal{D}$, we have that

$$|\mathcal{D}_S| - H(\tilde{V}^N[\mathcal{D}_S](w)) \leq \delta, \tag{61}$$

$$I(\tilde{V}^N[\mathcal{D}_S](w); \tilde{V}^N[\mathcal{D}\backslash\mathcal{D}_S](w)) \leq \delta. \tag{62}$$

Taking the subset as $\mathcal{D}_S = \mathcal{I}_S$ or $\mathcal{D}_S = \mathcal{I}_J$, we obtain that

$$|\mathcal{I}_S| - H(\tilde{V}^N[\mathcal{I}_S](w)) \leq \delta, \tag{63}$$

$$I(\tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}}) \cup \mathcal{F} \cup \mathcal{I}_J](w); \tilde{V}^N[\mathcal{I}_S](w)) \leq \delta. \tag{64}$$

Proof: See Appendix B-E.

Due to (63), the secret key for each user is close to uniform as required in (6). Now we consider the following

$$\begin{aligned} H(S(W)|\mathcal{C} = C) &= \sum_{w=1}^{M_I} \Pr(W = w) H(S(w)|\mathcal{C} = C) \\ &\overset{(a)}{\geq} \sum_{w=1}^{M_I} \Pr(W = w)|S(w)| - \delta = |S(W)| - \delta, \end{aligned} \tag{65}$$

where $\delta \to 0$ as $N \to \infty$; (a) holds due to (63) in Lemma 5. Therefore, $S(W)$ is close to uniform as required in (6).

d) Secrecy Analysis: It holds that

$$\begin{aligned} &I(S(W); \{J(i)\}_{i=1}^{M_I}|\mathcal{C} = C) \\ &\overset{(a)}{=} I(S(W); J_1(W), J_2(W)|\mathcal{C} = C) \\ &= \sum_{w=1}^{M_I} \Pr(W = w) I(S(w); J_1(w), J_2(w)|\mathcal{C} = C) \\ &= \sum_{w=1}^{M_I} \frac{I(\tilde{V}^N[\mathcal{I}_S](w); \tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}}) \cup \mathcal{I}_J](w)|\mathcal{C} = C)}{M_I} \\ &\overset{(b)}{\leq} \delta \end{aligned} \tag{66}$$

where $\delta \to 0$ as $N \to \infty$; (a) follows from the fact that the secret key of user W is independent of the helper data of the other users; (b) follows from (64) in Lemma 5.

e) Privacy Analysis: Before proceeding with the privacy leakage analysis, we include the following lemmas.

Lemma 6: For any $w \in [1 : M_I]$, $\tilde{V}^N(w)$ resulting from Algorithm 1 has a joint p.m.f. $p_{X^N(w)\tilde{V}^N(w)}$ with $X^N(w)$ such that

$$\mathbb{D}(p_{X^N(w)V^N(w)}||p_{X^N(W)\tilde{V}^N(w)}) \leq N\delta_N. \tag{67}$$

Hence, by Pinsker's inequality [38, p. 44], we have

$$\mathbb{V}(p_{X^N(w)V^N(w)}, p_{X^N(w)\tilde{V}^N(w)}) \leq \sqrt{2\ln 2}\sqrt{N\delta_N}. \tag{68}$$

Proof: See Appendix B-F.

Lemma 7: It holds that

$$I(\tilde{V}^N[\mathcal{V}_{U|X}](W); X^N(W)|\mathcal{C} = C) \leq \delta, \tag{69}$$

where $\delta \to 0$ as $N \to \infty$.

Proof: See Appendix B-G.

Now we consider the privacy leakage as follows

$$\begin{aligned} &I(\{J(i)\}_{i=1}^{M_I}; X^N(W)|\mathcal{C} = C) \\ &\overset{(a)}{=} I(J(W); X^N(W)|\mathcal{C} = C) \\ &= I(\tilde{V}^N[\big(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}}\big) \cup \mathcal{F} \cup \mathcal{I}_J](W); X^N(W)|\mathcal{C} = C) \\ &= I(\tilde{V}^N[\big(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}\big) \cup \big(\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}\big) \cup \mathcal{F} \cup \mathcal{I}_J](W); \\ &\qquad X^N(W)|\mathcal{C} = C) \\ &= I(\tilde{V}^N[\big(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}\big) \cup \mathcal{F} \cup \mathcal{I}_J](W); X^N(W)| \\ &\qquad \tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W), \mathcal{C} = C) \\ &\qquad + I(\tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W); X^N(W)|\mathcal{C} = C) \\ &\overset{(b)}{\leq} H(\tilde{V}^N[\big(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}\big) \cup \mathcal{F} \cup \mathcal{I}_J](W)|\mathcal{C} = C) + \delta \\ &\leq |\big(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}\big) \cup \mathcal{F} \cup \mathcal{I}_J| + \delta \\ &= |\mathcal{V}_{U|Y}| - |\mathcal{V}_{U|X}| + |\mathcal{F}| + |\mathcal{I}_J| + \delta, \end{aligned} \tag{70}$$

where $\delta \to 0$ as $N \to \infty$ and (a) follows from the fact that the biometric sequence $X^N(W)$ of user W is independent of the helper data of other users; (b) follows from Lemma 7. For size of $|\mathcal{V}_{U|X}|$, by [20, Theorem 1], it holds that

$$\lim_{N\to\infty} \frac{|\mathcal{V}_{U|X}|}{N} = H(U|X). \tag{71}$$

Combining (34), (40), (70), and (71), we have that

$$\begin{aligned} &\frac{I(\{J(i)\}_{i=1}^{M_I}; X^N(W)|\mathcal{C} = C)}{N} \\ &\qquad \xrightarrow{N\to\infty} I(U; X) - I(U; Y) + R_I + 2\epsilon, \end{aligned} \tag{72}$$

where we used $U - X - Y$.

f) Complexity Analysis: From [17], we know that the complexity of successive cancellation operation is $\mathcal{O}(N \log N)$. For the identification and authentication procedure in Algorithm 2, the system operates an exhaustive search until a matched user is found. Therefore, at most M_I times the successive cancellation is operated, which results in complexity no larger than $\mathcal{O}(M_I N \log N)$.

Combining (41), (44), (58), (65), (66), and (72), we complete the proof of Theorem 5.

B. Proof of Theorem 6

a) Rate Analysis: Following the analysis in (38), (39), and (40), we obtain that

$$\lim_{N\to\infty} R_I + R_S = I(U; Y) - 2\epsilon. \tag{73}$$

For the helper data rate R_J, it holds that

$$R_J = \frac{J(W)}{N} = \frac{|\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}}| + |\mathcal{I}_J| + |\mathcal{F}|}{N} + R_S$$

$$
\begin{aligned}
&= \frac{|\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}}|}{N} + R_I + R_S + 2\epsilon \\
&\overset{(a)}{=} \frac{|\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}}|}{N} + I(U;Y) \\
&\overset{(b)}{\to} I(U;\tilde{X}),
\end{aligned} \tag{74}
$$

as $N \to \infty$; where (a) holds due to (73); (b) follows from (34), (40), and (43).

b) Error Events Analysis: Comparing Algorithm 2 and Algorithm 4, the identification procedures, i.e., guessing the user's index, are the same for the generated secret key system and the chosen secret key system. From the error events analysis in the proof of Theorem 5, we can obtain that for the chosen secret key system, the identification is reliable, i.e.,

$$
\Pr\{\hat{W} \neq W\} \xrightarrow{N\to\infty} 0. \tag{75}
$$

Then for the authentication error probability, we have that

$$
\begin{aligned}
&\Pr\{\hat{S} \neq S(W)|\hat{W} = W\} \\
&\quad = \Pr\{J_3(\hat{W}) \ominus \hat{V}^N[\mathcal{I}_S](\hat{W}) \neq J_3(W) \ominus \tilde{V}^N[\mathcal{I}_S](W) \\
&\qquad\qquad |\hat{W} = W\} \\
&\quad = \Pr\{\hat{V}^N[\mathcal{I}_S](W) \neq \tilde{V}^N[\mathcal{I}_S](W)\} \xrightarrow{N\to\infty} 0,
\end{aligned} \tag{76}
$$

where the last step follows from Lemma 1. Combining (75) and (76), we obtain that there exists a suitable codebook $\mathcal{C} = C$ such that

$$
\Pr\{(\hat{W}, \hat{S}) \neq (W, S(W))\} \xrightarrow{N\to\infty} 0. \tag{77}
$$

c) Secrecy Analysis: It holds that

$$
\begin{aligned}
&I(S(W); \{J(i)\}_{i=1}^{M_I}|\mathcal{C} = C) = I(S(W); J(W)|\mathcal{C} = C) \\
&\overset{(a)}{=} I(S(W); J_1(W), J_2(W), J_3(W)|\mathcal{C} = C) \\
&= I(S(W); J_1(W), J_2(W)|\mathcal{C} = C) \\
&\quad + I(S(W); J_3(W)|J_1(W), J_2(W), \mathcal{C} = C) \\
&\overset{(b)}{=} H(J_3(W)|J_1(W), J_2(W), \mathcal{C} = C) \\
&\quad - H(J_3(W)|S(W), J_1(W), J_2(W), \mathcal{C} = C) \\
&\leq NR_S - H(\tilde{V}^N[\mathcal{I}_S](W)|S(W), J_1(W), J_2(W), \mathcal{C} = C) \\
&\overset{(b)}{=} NR_S - H(\tilde{V}^N[\mathcal{I}_S](W)|J_1(W), J_2(W), \mathcal{C} = C) \\
&= NR_S - H(\tilde{V}^N[\mathcal{I}_S](W)) \\
&\quad + I(\tilde{V}^N[\mathcal{I}_S](W); J_1(W), J_2(W)|\mathcal{C} = C) \\
&\overset{(c)}{\leq} 2\delta,
\end{aligned} \tag{78}
$$

where $\delta \to 0$ as $N \to \infty$; (a) holds since the secret key of user W is independent of the helper data of the other users; (b) follows from the fact that $S(W)$ is independent of $\tilde{X}^N(W)$ and thus independent of $(J_1(W), J_2(W), V^N[\mathcal{I}_S](W))$ given the codebook $\mathcal{C} = C$; (c) follows from (63) and (64).

d) Privacy Analysis: It holds that

$$
\begin{aligned}
&I(X^N(W); \{J(i)\}_{i=1}^{M_I}|\mathcal{C} = C) \\
&\overset{(a)}{=} I(X^N(W); J(W)|\mathcal{C} = C) \\
&= I(X^N(W); J_1(W), J_2(W), J_3(W)|\mathcal{C} = C) \\
&= I(X^N(W); J_1(W), J_2(W)|\mathcal{C} = C) \\
&\quad + I(X^N(W); J_3(W)|J_1(W), J_2(W), \mathcal{C} = C) \\
&\leq I(X^N(W); J_1(W), J_2(W)|\mathcal{C} = C) + H(J_3(W)|\mathcal{C} = C) \\
&\quad - H(J_3(W)|X^N(W), \tilde{X}(W), J_1(W), J_2(W), \mathcal{C} = C) \\
&\leq I(X^N(W); J_1(W), J_2(W)|\mathcal{C} = C) + NR_S \\
&\quad - H(S(W)|X^N(W), \tilde{X}(W), J_1(W), J_2(W), \mathcal{C} = C) \\
&\overset{(b)}{=} I(X^N(W); J_1(W), J_2(W)|\mathcal{C} = C),
\end{aligned} \tag{79}
$$

where (a) follows from the fact that the biometric sequence $X^N(W)$ of user W is independent of the helper data of other users; (b) follows from $S(W)$ is independent of $(X^N(W), \tilde{X}^N(W), J_1(W), J_2(W))$ and that $S(W)$ is uniformly distributed on $[1 : 2^{NR_S}]$.

Combining (72) and (79), we obtain that

$$
\begin{aligned}
&\lim_{N\to\infty} \frac{I(\{J(i)\}_{i=1}^{M_I}; X^N(W)|\mathcal{C} = C)}{N} \\
&\qquad \leq I(U;X) - I(U;Y) + R_I + 2\epsilon.
\end{aligned} \tag{80}
$$

Combining (73), (74), (77), (78), and (80), it completes the proof of Theorem 2.

C. Proof of Theorem 7

a) Rate Analysis: Combining (36), (37), (39), (40), and (71), we obtain that

$$
\lim_{N\to\infty} R_P \geq I(U;X) - I(U;Y) + R_I + 2\epsilon, \tag{81}
$$

$$
\lim_{N\to\infty} R_I + R_S \leq I(U;Y) + R_P - 2\epsilon. \tag{82}
$$

The helper data rate can be bounded as

$$
\begin{aligned}
R_J = \frac{J(W)}{N} &= \frac{|\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|\tilde{X}}| + |\mathcal{F}| + |P_2(W)|}{N} \\
&\overset{(a)}{\to} I(U;\tilde{X}) - I(U;Y) + R_I + 2\epsilon,
\end{aligned} \tag{83}
$$

as $N \to \infty$; where (a) follows from (40) and (43).

b) Error Events Analysis: Assume that user W is observed. Let $J(W) = (J_1(W), J_2(W))$, $P(W) = (P_1(W), P_2(W), P_3(W))$, and $S(W)$ denote the actual helper data, private key and the secret key of the observed user W. Define the following error events

$$
\begin{aligned}
\mathcal{E}_1^0 &= \{\exists \hat{w} \neq W : J_2(\hat{w}) = P_2(W), \\
&\qquad \big(\hat{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](\hat{w}), P_3(\hat{w})\big) = S(W)\}, \\
\mathcal{E}_2^0 &= \{\big(\hat{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](W), P_3(W)\big) \neq S(W)\}, \\
\mathcal{E}_3^0 &= \{\exists \hat{w} \neq W : J_2(\hat{w}) = P_2(W), \\
&\qquad \big(\hat{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](\hat{w}), P_3(\hat{w})\big) \neq S(W)\},
\end{aligned} \tag{84}
$$

The error event $\mathcal{E}_1^0$ denotes the identification error that there exists another user satisfying the required conditions and the estimated secret key matches the true one. The second and third error events are the authentication errors that the estimated secret key does not match the true one: $\mathcal{E}_2^0$ denotes the case that the estimated secret key of the observed user does not match the true one $S(W)$; $\mathcal{E}_3^0$ denotes the case that neither the guessed user index nor the estimated secret key is the same as the true one. The identification and authentication

are reliable if and only if none of the above events happen. We define the error event

$$\mathcal{E}^0 = \{(\hat{W}, \hat{S}) \neq (W, S(W))\} \tag{85}$$

and it holds that

$$\mathcal{E}^0 \subset \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4. \tag{86}$$

In the following, due to symmetry and without loss of generality, assume that $W = 1$. We firstly consider $\mathcal{E}_2^0|W = 1$ as follows

$$\begin{aligned}
&\Pr\{\mathcal{E}_2^0|W=1\} \\
&= \Pr\{\hat{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](1) \neq \tilde{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](1)|W=1\} \\
&\overset{(a)}{=} \Pr\{\hat{V}^N[\mathcal{V}_U(\backslash\mathcal{V}_{U|Y} \cup \mathcal{F})](1) \\
&\qquad \neq \hat{V}^N[\mathcal{V}_U(\backslash\mathcal{V}_{U|Y} \cup \mathcal{F})](1)|W=1\} \overset{(b)}{\to} 0,
\end{aligned} \tag{87}$$

as $N \to \infty$; where (a) holds because $\hat{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W) = J_{11}(W) \ominus P_1(W) = \tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W)$ as given in Algorithm 6; (b) follows from Lemma 1.

The probability of $\mathcal{E}_1^0 \cup \mathcal{E}_3^0|W = 1$ can be bounded as follows

$$\begin{aligned}
&\Pr\{\mathcal{E}_1^0 \cup \mathcal{E}_3^0|W=1\} \\
&=\Pr\{\exists \hat{w} \neq 1 : P_2(\hat{w}) = P_2(1)|W=1\} \\
&\overset{(a)}{\leq} \sum_{\hat{w}\neq 1} \Pr\{P_2(\hat{w}) = P_2(1)|W=1\} \\
&\overset{(b)}{\leq} 2^{NR_I} 2^{-N(R_I+\epsilon)} = 2^{-N\epsilon} \xrightarrow{N\to\infty} 0,
\end{aligned} \tag{88}$$

where (a) follows from the union bound; (b) holds since the private keys of different users are independent.

Combining (87) and (88), we obtain that there exists a suitable codebook $\mathcal{C} = C$ such that

$$\Pr\{\mathcal{E}^0\} \xrightarrow{N\to\infty} 0. \tag{89}$$

c) Uniformity of Secret Keys: Consider the following entropy

$$\begin{aligned}
H(S(W)|\mathcal{C}=C) &= H(\tilde{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](W), P_3(W)) \\
&\overset{(a)}{=} H(\tilde{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](W)) + H(P_3(W)) \\
&\overset{(b)}{\geq} |\mathcal{V}_U\backslash\mathcal{V}_{U|X}| - \delta + |P_3(W)| \\
&= |S(W)| - \delta,
\end{aligned} \tag{90}$$

where $\delta \to 0$ as $N \to \infty$; (a) holds because $P(W)$ is independent of $\tilde{X}^N(W)$ and $\tilde{V}^N(W)$; (b) follows from (60) in Lemma 5. This proves the uniformity of secret keys.

d) Secrecy Analysis: It holds that

$$\begin{aligned}
&I(S(W); \{J(i)\}_{i=1}^{M_I}|\mathcal{C}=C) \\
&\overset{(a)}{=} I(S(W); J(W)|\mathcal{C}=C) \\
&= I(\tilde{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](W), P_3(W); \\
&\qquad \tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W) \oplus P_1(W), \\
&\qquad\quad \tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W), P_2(W)|\mathcal{C}=C) \\
&\overset{(b)}{=} I(\tilde{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](W); \\
&\qquad \tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W) \oplus P_1(W), \\
&\qquad\quad \tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W)|\mathcal{C}=C) \\
&\overset{(c)}{\leq} I(\tilde{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](W); \\
&\qquad \tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W) \oplus P_1(W), \\
&\qquad\quad |\tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W), \mathcal{C}=C) + \delta \\
&\leq \delta + H(\tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W) \oplus P_1(W)) \\
&\quad - H(\tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W) \oplus P_1(W) \\
&\qquad |\tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W), \tilde{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](W), \mathcal{C}=C) \\
&\leq \delta + |(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}| - H(P_1(W)| \\
&\qquad \tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W), \tilde{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|X}](W), \mathcal{C}=C) \\
&\overset{(b)}{=} \delta,
\end{aligned} \tag{91}$$

where $\delta \to 0$ as $N \to \infty$; (a) follows from the fact that the biometric sequence $X^N(W)$ of user W is independent of the helper data of other users; (b) holds because $(P_1(W), P_2(W), P_3(W))$ are mutually independent and they are all independent of $\tilde{V}^N(W)$; (c) follows from (62) in Lemma 5 and letting $\mathcal{D}_S = \mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}$.

e) Privacy Analysis: It holds that

$$\begin{aligned}
&I(X^N(W); \{J(i)\}_{i=1}^{M_I}|\mathcal{C}=C) \\
&\overset{(a)}{=} I(X^N(W); J(W)|\mathcal{C}=C) \\
&\overset{(b)}{\leq} I(X^N(W); \tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W) \oplus P_1(W), \\
&\qquad P_2(W)|\tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W), \mathcal{C}=C) + \delta \\
&\overset{(c)}{\leq} I(X^N(W); \tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W) \oplus P_1(W) \\
&\qquad |\tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W), P_2(W), \mathcal{C}=C) + \delta \\
&\leq H(P_1(W)) - H(\tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W) \oplus P_1(W) \\
&\qquad |X^N(W), \tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W), P_2(W), \mathcal{C}=C) + \delta \\
&\overset{(d)}{\leq} H(P_1(W)) - H(P_1(W)|\tilde{V}^N[(\mathcal{V}_{U|Y}\backslash\mathcal{V}_{U|X}) \cup \mathcal{F}](W), \\
&\qquad X^N(W), \tilde{V}^N[\mathcal{V}_{U|X}\backslash\mathcal{V}_{U|\tilde{X}}](W), P_2(W), \mathcal{C}=C) + \delta \\
&\overset{(c)}{=} \delta,
\end{aligned} \tag{92}$$

where $\delta \to 0$ as $N \to \infty$; (a) follows from the fact that the biometric sequence $X^N(W)$ of user W is independent of the helper data of other users; (b) follows from Lemma 7; (c) holds because $P_1(W)$, $P_2(W)$, and $(X^N(W), \tilde{V}^N(W))$ are mutually independent; (d) holds since conditioning reduces entropy.

Combining (81), (82), (83), (89), (90), (91), and (92), we complete the proof of Theorem 7.

D. Proof of Theorem 8

The proof of Theorem 8 can be obtained by extending the proof of Theorem 7 with including a masking procedure, which is similar to the proof idea of Theorem 6.

Appendix B
Supplementary Proof

A. Proof of Lemma 1

In the identification and authentication algorithms, $\hat{V}[\mathcal{I}_J \cup \mathcal{I}_S](w)$ is obtained by successive cancellation operation with

Y^N and $\tilde{V}[\mathcal{V}_{U|Y}\cup\mathcal{F}](w)$. Next we consider the size of $\mathcal{V}_{U|Y}\cup\mathcal{F}$.

From (40), we have that for sufficiently small $\epsilon>0$ there exists $N_0(\epsilon)\in\mathbb{N}^+$ such that for any $N=2^n\geq N_0(\epsilon)$ for some $n\in\mathbb{N}^+$, it holds that $|\mathcal{V}_{U|Y}|/N > H(U|Y)-\epsilon$. Therefore, for sufficiently large N, it holds that

$$|\mathcal{V}_{U|Y}\cup\mathcal{F}| = |\mathcal{V}_{U|Y}| + |\mathcal{F}| > NH(U|Y). \tag{93}$$

Applying [20, Theorem 3], we can obtain (48) from (93).

B. Proof of Lemma 2

The proof idea is similar to [24, Appendix B-B]. For any user $w\in[1:M_I]$, we have that

$$\begin{aligned}
&\mathbb{D}(p_{\tilde{X}^N(w)V^N(w)}||p_{\tilde{X}^N(w)\tilde{V}^N(w)})\\
&\overset{(a)}{=}\mathbb{D}(p_{V^N(w)|\tilde{X}^N(w)}||p_{\tilde{V}^N(w)|\tilde{X}^N(w)})\\
&\overset{(a)}{=}\sum_{j}^{N}\mathbb{D}(p_{V^j(w)|V^{j-1}(w),\tilde{X}^N(w)}||p_{\tilde{V}^j(w)|\tilde{V}^{j-1}(w),\tilde{X}^N(w)})\\
&\overset{(b)}{=}\sum_{j\in\mathcal{V}_{U|\tilde{X}}\cup\mathcal{H}_U^c}\mathbb{D}(p_{V^j(w)|V^{j-1}(w),\tilde{X}^N(w)}\\
&\qquad\qquad||p_{\tilde{V}^j(w)|\tilde{V}^{j-1}(w),\tilde{X}^N(w)})\\
&\overset{(c)}{=}\sum_{j\in\mathcal{V}_{U|\tilde{X}}}(1-H(V^j(w)|V^{j-1}(w),\tilde{X}^N(w)))\\
&\quad+\sum_{j\in\mathcal{H}_U^c}(H(V^j(w)|V^{j-1}(w))\\
&\qquad-H(V^j(w)|V^{j-1}(w)),\tilde{X}^N(w))\\
&\overset{(d)}{\leq}|\mathcal{V}_{U|\tilde{X}}|\delta_N+\sum_{j\in\mathcal{H}_U^c}H(V^j(w)|V^{j-1}(w))\\
&\leq|\mathcal{V}_{U|\tilde{X}}|\delta_N+|\mathcal{H}_U^c|\delta_N\leq N\delta_N,
\end{aligned}\tag{94}$$

where (a) follows from the chain rule of Kullback-Leibler divergence; (b) follows from (35) and thus $p_{\tilde{V}^j(w)|\tilde{V}^{j-1}(w),\tilde{X}^N(w)}=p_{V^j(w)|V^{j-1}(w),\tilde{X}^N(w)}$ for any $j\in(\mathcal{H}_U\backslash\mathcal{V}_{U|\tilde{X}})$; (c) follows from (35) and the i.i.d. uniformly distributed assignment of $\tilde{V}^N[\mathcal{V}_{U|\tilde{X}}](W)$; (d) follows from (26). Therefore, we obtain that

$$\begin{aligned}
\mathbb{D}(p_{V^N(w)}||p_{\tilde{V}^N(w)}) &\overset{(a)}{=}\mathbb{D}(p_{\tilde{X}^N(w)V^N(w)}||p_{\tilde{X}^N(W)\tilde{V}^N(w)})\\
&\quad-\mathbb{D}(p_{V^N(w)|\tilde{X}^N(w)}||p_{\tilde{V}^N(w)|\tilde{X}^N(W)})\\
&\overset{(b)}{\leq}N\delta_N,
\end{aligned}\tag{95}$$

where (a) follows from the chain rule and the non-negativity of Kullback-Leibler divergence, respectively.

C. Proof of Lemma 3

For any user $w\in[1:M_I]$ and any $i\in\mathcal{V}_U$, we have

$$\begin{aligned}
&H(\tilde{V}_i(w)|\tilde{V}^{i-1}(w))\\
&\overset{(a)}{\geq}H(\tilde{V}_i(w)|\tilde{V}^{i-1}(w))-H(V_i(w)|V^{i-1}(w))\\
&\qquad+1-\delta_N\\
&\geq H(\tilde{V}_i(w),\tilde{V}^{i-1}(w))-H(\tilde{V}^{i-1}(w))+1-\delta_N\\
&\qquad-\big(H(V_{a_i}(w),V^{i-1}(w))-H(V^{i-1}(w))\big)\\
&=\big(H(\tilde{V}_i(w),\tilde{V}^{i-1}(w))-H(V_i(w),V^{i-1}(w))\big)\\
&\qquad-\big(H(\tilde{V}^{i-1}(w))-H(V^{i-1}(w))\big)+1-\delta_N\\
&\geq-|H(\tilde{V}_i(w),\tilde{V}^{i-1}(w))-H(V_i(w),V^{i-1}(w))|\\
&\qquad-|H(\tilde{V}^{i-1}(w))-H(V^{i-1}(w))|+1-\delta_N\\
&\overset{(b)}{\geq}-h_2\big(\mathbb{V}(p_{\tilde{V}_i(w),\tilde{V}^{i-1}(w)},p_{V_i(w),V^{i-1}(w)})\big)\\
&\qquad-\frac{\mathbb{V}(p_{\tilde{V}_i(w),\tilde{V}^{i-1}(w)},p_{V_i(w),V^{i-1}(w)})}{2}\log(i-1)\\
&\qquad-h_2\big(\mathbb{V}(p_{\tilde{V}^{i-1}(w)},p_{V^{i-1}(w)})\big)\\
&\qquad-\frac{\mathbb{V}(p_{\tilde{V}^{i-1}(w)},p_{V^{i-1}(w)})}{2}\log(i-2)+1-\delta_N\\
&\overset{(c)}{\geq}-2h_2\big(\mathbb{V}(p_{\tilde{V}^N(w)},p_{V^N(w)})\big)\\
&\qquad-\mathbb{V}(p_{\tilde{V}^N(w)},p_{V^N(w)})\log N+1-\delta_N\\
&\overset{(d)}{\geq}-2h_2\big(2\sqrt{2\ln 2}\sqrt{N\delta_N}\big)\\
&\qquad-2\sqrt{2\ln 2}\sqrt{N^3\delta_N}+1-\delta_N\\
&\overset{(e)}{\to}1,
\end{aligned}\tag{96}$$

as $N\to\infty$; where (a) follows from (27); (b) follows from [39, Theorem 6]; (c) holds because the binary entropy function $h_2(x)$ is increasing for small $x>0$; (d) follows from (52) in Lemma 2; (e) holds due to the fact that $N^3\delta_N\to 0$[3] as $N\to\infty$ and $h_2(x)\to 0$ as $x\to 0$.

D. Proof of Lemma 4

Similar to the proof idea in [24, Lemma 7], we have that

$$\begin{aligned}
&\mathbb{V}(p_{\tilde{V}^N[\mathcal{V}_U\backslash\mathcal{V}_{U|\tilde{X}}](w))},q_{\mathcal{U}_{|\mathcal{V}_U\backslash\mathcal{V}_{U|\tilde{X}}|}})\\
&\overset{(a)}{\leq}\mathbb{V}(p_{\tilde{V}^N[\mathcal{V}_U](w)},q_{\mathcal{U}_{|\mathcal{V}_U|}})\\
&\overset{(b)}{\leq}\mathbb{V}(p_{\tilde{V}^N[\mathcal{V}_U](w)},p_{V^N[\mathcal{V}_U](w)})+\mathbb{V}(p_{V^N[\mathcal{V}_U](w)},q_{\mathcal{U}_{|\mathcal{V}_U|}})\\
&\overset{(c)}{\leq}\sqrt{2\ln 2}\sqrt{N\delta_N}+\mathbb{V}(p_{V^N[\mathcal{V}_U](w)},q_{\mathcal{U}_{|\mathcal{V}_U|}})\\
&\overset{(d)}{\leq}\sqrt{2\ln 2}\sqrt{N\delta_N}+\sqrt{2\ln 2}\sqrt{\mathbb{D}(p_{V^N[\mathcal{V}_U](w)}||q_{\mathcal{U}_{|\mathcal{V}_U|}})}\\
&=\sqrt{2\ln 2}\sqrt{N\delta_N}+\sqrt{2\ln 2}\sqrt{|\mathcal{V}_U|-H(V^N[\mathcal{V}_U](w))}\\
&\overset{(e)}{\leq}2\sqrt{2\ln 2}\sqrt{N\delta_N},
\end{aligned}\tag{97}$$

where (a) follows by defining $q_{\mathcal{U}_{|\mathcal{V}_U|}}$ the uniform distribution on $[1:2^{|\mathcal{V}_U|}]$; (b) follows from the triangle inequality; (c) follows from Lemma 2; (d) follows from Pinsker's inequality [38, p. 44]; (e) holds because for any $w\in[1:M_I]$, we have that

$$\begin{aligned}
&|\mathcal{V}_U|-H(V^N[\mathcal{V}_U](w)))\\
&=|\mathcal{V}_U|-\sum_{i\in\mathcal{V}_U}H(V_i(w)|V^{i-1}(w))
\end{aligned}$$

[3]This can be proved by L'Hospital's rule. Since $\delta_N=2^{-\beta N}$, it holds that $\lim_{N\to\infty}N^3\delta_N=\lim_{N\to\infty}\frac{N^3}{2^{\beta N}}=\lim_{N\to\infty}\frac{6}{2^{\beta N}\ln^3(2^\beta)}$=0.

$$\overset{(f)}{\leq} |\mathcal{V}_U| - \sum_{i\in\mathcal{V}_U}(1-\delta_N) = |\mathcal{V}_U|\delta_N \leq N\delta_N, \tag{98}$$

where (f) follows from (27).

E. Proof of Lemma 5

Let $\mathcal{D}$ denote $\mathcal{V}_U\backslash\mathcal{V}_{U|\tilde{X}}$ for simplicity. Similar to [24, Appendix B-C], if N is sufficiently large, then we have

$$\begin{aligned}
&|\mathcal{D}| - H(\tilde{V}^N[\mathcal{D}](w)) \\
&\overset{(a)}{\leq} \mathbb{V}(p_{\tilde{V}^N[\mathcal{D}](w)}, q_{\mathcal{U}_{|\mathcal{D}|}}) \times \log_2 \frac{|\mathcal{D}|}{\mathbb{V}(p_{\tilde{V}^N[\mathcal{D}](w)}, q_{\mathcal{U}_{|\mathcal{D}|}})} \\
&\overset{(b)}{\leq} N\mathbb{V}(p_{\tilde{V}^N[\mathcal{D}](w)}, q_{\mathcal{U}_{|\mathcal{D}|}}) \\
&\quad - \mathbb{V}(p_{\tilde{V}^N[\mathcal{D}](w)}, q_{\mathcal{U}_{|\mathcal{D}|}}) \times \log_2 \mathbb{V}(p_{\tilde{V}^N[\mathcal{D}](w)}, q_{\mathcal{U}_{|\mathcal{D}|}}) \\
&\overset{(c)}{\leq} 2\sqrt{2\ln 2}\sqrt{N\delta_N}\big(N - \log_2(2\sqrt{2\ln 2}\sqrt{N\delta_N})\big),
\end{aligned} \tag{99}$$

where (a) follows from [38, Lemma 2.7]; (b) follows from the fact that $\log|\mathcal{D}| \leq N$; (c) holds due to Lemma 4 and the fact that $x\log_2 x$ is decreasing for sufficiently small $x > 0$.

Let δ'_N denote $2\sqrt{2\ln 2}\sqrt{N\delta_N}(N - \ln(2\sqrt{2\ln 2}\sqrt{N\delta_N}))$. Then for any subset $\mathcal{D}_S$ of $\mathcal{D}$, we have that

$$\begin{aligned}
&|\mathcal{D}| - H(\tilde{V}^N[\mathcal{D}](w)) \\
&= \Big(|\mathcal{D}_S| - H(\tilde{V}^N[\mathcal{D}_S](w))\Big) + \Big(|\mathcal{D}\backslash\mathcal{D}_S| \\
&\quad - H(\tilde{V}^N[\mathcal{D}\backslash\mathcal{D}_S](w))\Big) + I(\tilde{V}^N[\mathcal{D}_S](w); \tilde{V}^N[\mathcal{D}\backslash\mathcal{D}_S](w)) \\
&\leq \delta'_N.
\end{aligned} \tag{100}$$

Combining (100) with the fact that

$$|\mathcal{D}_S| - H(\tilde{V}^N[\mathcal{D}_S](w)) \geq 0, \tag{101}$$

$$|\mathcal{D}\backslash\mathcal{D}_S| - H(\tilde{V}^N[\mathcal{D}\backslash\mathcal{D}_S](w)) \geq 0, \tag{102}$$

$$I(\tilde{V}^N[\mathcal{D}_S](w); \tilde{V}^N[\mathcal{D}\backslash\mathcal{D}_S](w)) \geq 0, \tag{103}$$

we obtain that (61) and (62) hold.

Now we show that $\delta'_N \to 0$ as $N \to \infty$. As $\delta_N = 2^{-\beta N}$, from L'Hospital's rule, we have that

$$\begin{aligned}
\lim_{N\to\infty}\sqrt{N\delta_N}N &= \lim_{N\to\infty}\frac{N^{\frac{3}{2}}}{(2^{\beta/2})^N} = \lim_{N\to\infty}\frac{\frac{3}{2}N^{\frac{1}{2}}}{\ln(2^{\beta/2})(2^{\beta/2})^N} \\
&= \lim_{N\to\infty}\frac{3/4}{N^{1/2}\ln^2(2^{\beta/2})(2^{\beta/2})^N} = 0.
\end{aligned} \tag{104}$$

Since $\sqrt{N\delta_N} \to 0$ and $\lim_{x\to 0} x\log_2 x = 0$, it holds that

$$\lim_{N\to\infty}\sqrt{N\delta_N}\log_2\sqrt{N\delta_N} = 0. \tag{105}$$

Combining (99), (104), and (105), we obtain that

$$|\mathcal{D}| - H(\tilde{V}^N[\mathcal{D}](w)) \leq \delta'_N \xrightarrow{N\to\infty} 0. \tag{106}$$

F. Proof of Lemma 6

It holds that

$$\begin{aligned}
&\mathbb{D}(p_{X^N(w)V^N(w)}||p_{X^N(W)\tilde{V}^N(w)}) \\
&\leq \mathbb{D}(p_{X^N(w)\tilde{X}^N(w)V^N(w)}||p_{X^N(w)\tilde{X}^N(w)\tilde{V}^N(w)}) \\
&= \sum_{x^N}\sum_{\tilde{x}^N}\sum_{v^N} p_{X^N(w)\tilde{X}^N(w)V^N(w)}(x^N,\tilde{x}^N,v^N) \\
&\quad \times \log\frac{p_{X^N(w)\tilde{X}^N(w)V^N(w)}(x^N,\tilde{x}^N,v^N)}{p_{X^N(w)\tilde{X}^N(w)\tilde{V}^N(w)}(x^N,\tilde{x}^N,v^N)} \\
&= \sum_{x^N}\sum_{\tilde{x}^N}\sum_{v^N} p_{X^N(w)\tilde{X}^N(w)V^N(w)}(x^N,\tilde{x}^N,v^N) \\
&\quad \times \log\frac{p_{V^N(w)|X^N(w)\tilde{X}^N(w)}(v^N|x^N,\tilde{x}^N)}{p_{\tilde{V}^N(w)|X^N(w)\tilde{X}^N(w)}(v^N|x^N,\tilde{x}^N)} \\
&\overset{(a)}{=} \sum_{x^N}\sum_{\tilde{x}^N}\sum_{v^N} p_{X^N(w)\tilde{X}^N(w)V^N(w)}(x^N,\tilde{x}^N,v^N) \\
&\quad \times \log\frac{p_{V^N(w)|\tilde{X}^N(w)}(v^N|\tilde{x}^N)}{p_{\tilde{V}^N(w)|\tilde{X}^N(w)}(v^N|\tilde{x}^N)} \\
&= \sum_{\tilde{x}^N}\sum_{v^N} p_{\tilde{X}^N(w)V^N(w)}(\tilde{x}^N,v^N) \\
&\quad \times \log\frac{p_{\tilde{X}^N(w)V^N(w)}(\tilde{x}^N,v^N)}{p_{\tilde{X}^N(w)\tilde{V}^N(w)}(\tilde{x}^N,v^N)} \\
&= \mathbb{D}(p_{\tilde{X}^N(w)V^N(w)}||p_{\tilde{X}^N(w)\tilde{V}^N(w)}) \\
&\overset{(b)}{\leq} N\delta_N \xrightarrow{N\to\infty} 0,
\end{aligned} \tag{107}$$

where (a) follows from the Markov chain $(\tilde{V}^N(w), V^N(w)) - \tilde{X}^N(w) - X^N(w)$; (b) follows from Lemma 1.

G. Proof of Lemma 7

We firstly consider the following mutual information

$$\begin{aligned}
&I(V^N[\mathcal{V}_{U|X}](w); X^N(w)|\mathcal{C}=C) \\
&\leq |\mathcal{V}_{U|X}| - H(V^N[\mathcal{V}_{U|X}](w)|X^N(w),\mathcal{C}=C) \\
&\overset{(a)}{\leq} |\mathcal{V}_{U|X}| - \sum_{i\in\mathcal{V}_{U|X}} H(V^i(w)|V^{i-1}(w),X^N(w),\mathcal{C}=C) \\
&\overset{(b)}{\leq} |\mathcal{V}_{U|X}| - \sum_{i\in\mathcal{V}_{U|X}}(1-\delta_N) \\
&= |\mathcal{V}_{U|X}|\delta_N \leq N\delta_N \xrightarrow{N\to\infty} 0,
\end{aligned} \tag{108}$$

where (a) follows from the fact that conditioning reduces entropy; (b) follows from (29). Now we consider the following

$$\begin{aligned}
&I(\tilde{V}^N[\mathcal{V}_{U|X}](w); X^N(w)|\mathcal{C}=C) \\
&\quad - I(V^N[\mathcal{V}_{U|X}](w); X^N(w)|\mathcal{C}=C) \\
&= H(V^N[\mathcal{V}_{U|X}](w), X^N(w)|\mathcal{C}=C) \\
&\quad - H(\tilde{V}^N[\mathcal{V}_{U|X}](w), X^N(w)|\mathcal{C}=C) \\
&\quad + H(\tilde{V}^N[\mathcal{V}_{U|X}](w)|\mathcal{C}=C) - H(V^N[\mathcal{V}_{U|X}](w)|\mathcal{C}=C) \\
&\leq |H(V^N[\mathcal{V}_{U|X}](w), X^N(w)|\mathcal{C}=C) \\
&\quad - H(\tilde{V}^N[\mathcal{V}_{U|X}](w), X^N(w)|\mathcal{C}=C)| \\
&\quad + |H(\tilde{V}^N[\mathcal{V}_{U|X}](w)|\mathcal{C}=C) - H(V^N[\mathcal{V}_{U|X}](w)|\mathcal{C}=C)| \\
&\overset{(a)}{\leq} h_2\big(\mathbb{V}(p_{X^N(w)V^N[\mathcal{V}_{U|X}](w)}, p_{X^N(w)\tilde{V}^N[\mathcal{V}_{U|X}](w)})\big) \\
&\quad + \frac{\mathbb{V}(p_{X^N(w)V^N[\mathcal{V}_{U|X}](w)}, p_{X^N(w)\tilde{V}^N[\mathcal{V}_{U|X}](w)})}{2}\log(2N) \\
&\quad + h_2\big(\mathbb{V}(p_{V^N[\mathcal{V}_{U|X}](w)}, p_{\tilde{V}^N[\mathcal{V}_{U|X}](w)})\big) \\
&\quad + \frac{\mathbb{V}(p_{V^N[\mathcal{V}_{U|X}](w)}, p_{\tilde{V}^N[\mathcal{V}_{U|X}](w)})}{2}\log(2N)
\end{aligned}$$

$$\overset{(b)}{\leq} 2h_2\big(\mathbb{V}(p_{X^N(w)V^N(w)}, p_{X^N(w)\tilde{V}^N(w)})\big) + \mathbb{V}(p_{X^N(w)V^N(w)}, p_{X^N(w)\tilde{V}^N(w)}) \log(2N)$$
$$\overset{(c)}{\leq} 2h_2(N\delta_N) + 4N^2\delta_N \xrightarrow{N\to\infty} 0, \tag{109}$$

where (a) follows from [39, Theorem 6]; (b) holds since $h_2(x)$ is increasing for small $x > 0$; (c) follows from Lemma 6.

Combining (111) and (109), we obtain that

$$I(\tilde{V}^N[\mathcal{V}_{U|X}](w); X^N(w)|\mathcal{C} = C) \leq N\delta_N + 2h_2(N\delta_N) + 4N^2\delta_N \overset{(a)}{\to} 0, \tag{110}$$

as $N \to \infty$; where (a) holds due to the fact that $N^2\delta_N \to 0$ as $N \to \infty$ and $h_2(x) \to 0$ as $x \to 0$.

Then we have that

$$I(\tilde{V}^N[\mathcal{V}_{U|X}](W); X^N(W)|\mathcal{C} = C) = \sum_{w=1}^{M_I} \Pr(W = w) I(\tilde{V}^N[\mathcal{V}_{U|X}](w); X^N(w)|\mathcal{C} = C) \overset{(a)}{\leq} 2h_2(N\delta_N) + 4N^2\delta_N \xrightarrow{N\to\infty} 0, \tag{111}$$

where (a) follows from (110).

REFERENCES

[1] 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 (ISIT)*. IEEE, Jun. 2003, pp. 82–87.

[2] E. Tuncel, "Capacity/storage tradeoff in high-dimensional identification systems," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2097–2106, 2009.

[3] 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.

[4] 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, 2019.

[5] L. Zhou, M. T. Vu, T. J. Oechtering, and M. Skoglund, "Two-stage biometric identification systems without privacy leakage," *IEEE J. Sel. Areas Inf. Theory*, 2021.

[6] 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.

[7] ——, "Fundamental limits for privacy-preserving biometric identification systems that support authentication," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5583–5594, Oct 2015.

[8] 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.

[9] 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.

[10] V. Yachongka and H. Yagi, "A new characterization of the capacity region of identification systems under noisy enrollment," in *54th Annu. Conf. Inf. Sciences Systems*. IEEE, 2020, pp. 1–6.

[11] L. Zhou, M. T. Vu, T. J. Oechtering, and M. Skoglund, "Privacy-preserving identification systems with noisy enrollment," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 3510–3523, 2021.

[12] M. T. Vu, T. J. Oechtering, and M. Skoglund, "Hypothesis testing and identification systems," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 3765–3780, 2021.

[13] M. T. Vu, "Perspectives on identification systems," Ph.D. dissertation, Division of Information Science and Engineering, KTH Royal Institute of Technology, 2019.

[14] D. Wang and P. Wang, "Two birds with one stone: Two-factor authentication with security beyond conventional bound," *IEEE Trans. Dependable Secure Comput.*, vol. 15, no. 4, pp. 708–722, 2016.

[15] S. Qiu, D. Wang, G. Xu, and S. Kumari, "Practical and provably secure three-factor authentication protocol based on extended chaotic-maps for mobile lightweight devices," *IEEE Trans. Dependable Secure Comput.*, 2020.

[16] Q. Jiang, N. Zhang, J. Ni, J. Ma, X. Ma, and K.-K. R. Choo, "Unified biometric privacy preserving three-factor authentication and key agreement for cloud-assisted autonomous vehicles," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 9390–9401, 2020.

[17] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[18] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric models," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7829–7838, 2013.

[19] E. Sasoglu *et al.*, *Polarization and polar codes*. Citeseer, 2012.

[20] E. Arikan, "Source polarization," in *2010 IEEE Int. Symp. Inf. Theory*. IEEE, 2010, pp. 899–903.

[21] T. V. Minh, T. J. Oechtering, and M. Skoglund, "Polar code for secure wyner-ziv coding," in *2016 IEEE Int. Workshop Inf. Forensics Security (WIFS)*. IEEE, 2016, pp. 1–6.

[22] H. Wang, X. Tao, N. Li, and Z. Han, "Polar coding for the wiretap channel with shared key," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 6, pp. 1351–1360, 2017.

[23] O. O. Koyluoglu and H. El Gamal, "Polar coding for secure transmission and key agreement," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 5, pp. 1472–1483, 2012.

[24] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6213–6237, 2015.

[25] O. Günlü, O. İşcan, 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, 2019.

[26] B. Chen and F. M. Willems, "Secret key generation over biased physical unclonable functions with polar codes," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 435–445, 2018.

[27] S. A. Hashemi, C. Condo, and W. J. Gross, "Fast and flexible successive-cancellation list decoders for polar codes," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5756–5769, 2017.

[28] S. Cammerer, B. Leible, M. Stahl, J. Hoydis, and S. ten Brink, "Combining belief propagation and successive cancellation list decoding of polar codes on a gpu platform," in *2017 IEEE Int. Conf. Acoust. Speech Signal Process (ICASSP)*. IEEE, 2017, pp. 3664–3668.

[29] F. Ercan, T. Tonnellier, N. Doan, and W. J. Gross, "Simplified dynamic sc-flip polar decoding," in *ICASSP 2020-2020 IEEE Int. Conf. Acoust. Speech Signal Process (ICASSP)*. IEEE, 2020, pp. 1733–1737.

[30] K. Sadeghi, A. Banerjee, J. Sohankar, and S. K. Gupta, "Geometrical analysis of machine learning security in biometric authentication systems," in *2017 16th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*. IEEE, 2017, pp. 309–314.

[31] Y. C. Feng and P. C. Yuen, "Binary discriminant analysis for generating binary face template," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 613–624, 2011.

[32] K. Inthavisas and D. Lopresti, "Secure speech biometric templates for user authentication," *IET biometrics*, vol. 1, no. 1, pp. 46–54, 2012.

[33] D. Wang, Q. Gu, X. Huang, and P. Wang, "Understanding human-chosen pins: characteristics, distribution and security," in *Proceedings of the 2017 ACM on Asia Conf. Comput. Commun. Security*, 2017, pp. 372–385.

[34] D. Wang, H. Cheng, P. Wang, X. Huang, and G. Jian, "Zipf's law in passwords," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 11, pp. 2776–2791, 2017.

[35] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6562–6582, 2013.

[36] ——, "List decoding of polar codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2213–2226, 2015.

[37] R. Mori and T. Tanaka, "Non-binary polar codes using reed-solomon codes and algebraic geometry codes," in *2010 IEEE Inf. Theory Workshop*. IEEE, 2010, pp. 1–5.

[38] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.

[39] S.-W. Ho and R. W. Yeung, "The interplay between entropy and variational distance," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 5906–5929, 2010.