Deception with Side Information in Biometric Authentication Systems

Wei Kang, Member, IEEE, Daming Cao, Nan Liu, Member, IEEE,

Abstract—In this paper, we study the probability of successful deception of an uncompressed biometric authentication system with side information at the adversary. It represents the scenario where the adversary may have correlated side information, e.g., a partial fingerprint or a DNA sequence of a relative of the legitimate user. We find the optimal exponent of the deception probability by proving both the achievability and the converse. Our proofs are based on a connection between the problem of deception with side information and the rate distortion problem with side information at both the encoder and decoder.

Index terms: Biometric authentication, blowing up lemma, information theoretical security, permutation, side information

I. INTRODUCTION

The biometric authentication problem has been studied extensively in recent years. In a biometric authentication system, a biometric feature, e.g., a finger print, a DNA sequence, etc., of a legitimate user is measured and the measurement, called enrollment, is stored in a database. Later this biometric feature of the same user is measured and compared with the enrollment for authentication. Due to the randomness in the process, different measurements of the same biometric features of the same person can not be exactly the same. Thus, the authentication system needs to tolerate a certain level of distortion between the measurements taken in the enrollment stage and the authentication stage.

In a biometric authentication system, successful deception happens when an adversary impersonates a legitimate user by faking a biometric feature close enough to the enrollment and deceives the authentication system. The first study on the deception probability in the biometric authentication system is [1], where the authors studied the deception in an authentication system where the enrollment is compressed. The authors obtained the optimal trade-off between the compression rate and the exponent of the probability of successful deception when the adversary has no side information. In the case with correlated side information at the adversary, achievability and converse results on the optimal trade-off were proposed in [1], however, they do not meet. A similar result was obtained in a recent paper [2], where the optimal exponent of the deception probability has been given in both cases of uncompressed and compressed enrollment with no side information at the adversary.

A different direction in studying the performance of a biometric authentication system is to study the maximum number of legitimate users, called capacity, allowed in a biometric authentication system under a given tolerated distortion level. In [3], the capacity is obtained if the enrollment is not compressed. Later, the capacity result is generalized to the case where the enrollment is compressed [4], and the trade-off between the compression rate and the capacity of the authentication system was studied. The threat of a deception from an adversary was not considered in this line of work.

In this paper, we study the optimal exponent of the probability of successful deception of a biometric authentication system with uncompressed enrollment and side information at the adversary. It represents the scenario where the adversary may have correlated side information, e.g., a partial fingerprint of the legitimate user or a DNA sequence of a relative of the legitimate user. We provide the optimal exponent of the deception probability by providing the proofs of both the achievability and the converse. Our proofs are based on a connection between the problem of deception with side information and the rate distortion problem with side information at both the encoder and decoder.

The remainder of the paper is as follows. In the next section, we state the problem formulation and the main result. The proofs of the achievability and the converse are given in Sections III and IV, respectively. Finally, a conclusion is given in Section V.

II. PROBLEM FORMULATION AND MAIN RESULT

A. Notations

Capital letters are used to denote random variables, e.g., $X$, $Y$, and calligraphic letters are used to denote sets, e.g., $\mathcal{P}$, $\mathcal{X}$. We use $X^n$ to represent a random sequence $(X_1, X_2, \ldots, X_n)$. $\mathcal{T}_P^n$ denotes the type with distribution $P$ and $\mathcal{T}_{[P]}_n$ denotes the typical set with distribution $P$, where both concepts are defined in [5, Chapter 2]. For a function $g$, we use $||g||$ to denotes the cardinality of the image of the function $g$. The logarithm function $\log(\cdot)$ and exponential function $\exp(\cdot)$ are of the same base.

B. Problem Formulation

Consider a pair of independent and identically distributed (i.i.d.) random sequences $X^n$ and $Y^n$, generated according to
a joint distribution \( P \), which is defined on a finite space \( \mathcal{X} \times \mathcal{Y} \). The random sequence \( X^n \) represents the biometric enrollment in the system and \( Y^n \) represents the side information at the adversary. We define a reconstruction space \( \mathcal{X} \) and a distortion function \( d : \mathcal{X} \times \mathcal{X} \rightarrow [0, \infty) \). The distortion between the sequences \( x^n \in \mathcal{X}^n \) and \( \hat{x}^n \in \hat{\mathcal{X}}^n \) is defined as

\[
d(x^n, \hat{x}^n) = \frac{1}{n} \sum_{i=1}^{n} d(x_i, \hat{x}_i). \tag{1}\]

The legitimate user is successfully identified if the distortion between the measurements in the enrollment stage and the authentication stage does not exceed a certain level, say \( \Delta \). We note that the probability of false rejection in our model is the same as in the model without side information at the adversary. Therefore, we refer the readers to [1], [2] for the derivation of the maximal probability of false rejection.

The adversary observes the side information \( Y^n \) and tries to impersonate the legitimate user using a deception function \( f : \hat{\mathcal{Y}}^n \rightarrow \mathcal{X}^n \). We define the achievable deception exponent as follows.

**Definition 1** A deception exponent \( E \) is achievable under the distortion constraint \( \Delta \) if for every \( \delta > 0 \) and every sufficiently large \( n \), there exists a deception function \( f \) such that

\[
-\frac{1}{n} \log \Pr(d(X^n, f(Y^n)) \leq \Delta) \leq E + \delta. \tag{2}\]

In this paper, we are interested in the minimal achievable deception exponent, which is the best the adversary can do. Based on the minimal achievable deception exponent, the designer of the biometric authentication system can choose an appropriate \( \Delta \) value that on one hand, limits the probability of successful deception below the tolerance level, and on the other hand, does not cause too large a probability of false rejection when the legitimate user is authenticated [1], [2].

**C. Rate-distortion with Side Information at Both the Encoder and Decoder**

It turns out that finding the minimal achievable deception exponent with side information at the adversary is intimately related to the rate distortion problem with side information at both the encoder and decoder [6]. Thus, in this subsection, we review the result for the rate distortion problem.

Assume a pair of i.i.d sequences \( X^n \) and \( Y^n \) is generated according to a joint distribution \( Q \) defined on \( \mathcal{X} \times \mathcal{Y} \), where \( Q \) is not necessarily equal to \( P \), defined in the Section II-B. The random sequence \( X^n \) is the source sequence to be reconstructed at the decoder under a certain distortion constraint and \( Y^n \) represents the side information available at both the encoder and decoder. The encoding function at the encoder is defined as \( g : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{1, 2, \ldots, M\} \), and the decoding function at the decoder is defined as \( \varphi : \{1, 2, \ldots, M\} \times \hat{\mathcal{Y}}^n \rightarrow \mathcal{X}^n \). We denote the minimal achievable rate under the distortion constraint \( \Delta \) in the rate distortion problem with side information at both the encoder and decoder as \( R_{SI}(Q, \Delta) \). From [6], we have

\[
R_{SI}(Q, \Delta) = \min_{V(\hat{x} | x, y) : Ed(x, \hat{x}) \leq \Delta} I(X; \hat{X}|Y). \tag{3}\]

**Remark:** The above rate distortion problem with side information at both the encoder and decoder can also be viewed as a special case of the Wyner-Ziv problem, i.e., the rate distortion problem with side information only at the decoder, as follows: in the Wyner-Ziv problem, view \((X^n, Y^n)\) jointly as the source sequence available at the encoder, view \( Y^n \) as the side information at the decoder, and take the the distortion function in the Wyner-ziv problem as \( d(x, \hat{x}) \), defined in Section II-B, i.e., the distortion of \( Y^n \) does not matter. By viewing the rate distortion problem with side information at both the encoder and decoder as a special case of the Wyner-Ziv problem, we can invoke the results of the Wyner-Ziv problem, more specifically [5, Theorem 16.5], in the later development. The reason why we prefer the Wyner-Ziv result over the rate distortion problem with side information at both the encoder and decoder is that [5, Theorem 16.5] provides a code satisfying distortion constraint with high probability, while the results for rate distortion with side information at both the encoder and decoder are usually given under the average distortion criterion, e.g., [6], [7].

**D. Main Result**

The main result of this paper is the following theorem.

**Theorem 1** The deception exponent \( E \) is achievable under the distortion constraint \( \Delta \) if and only if

\[
E \geq \min_Q \{D(Q||P) + R_{SI}(Q, \Delta)\}, \tag{4}\]

where \( R_{SI}(Q, \Delta) \) is given in (3), the distribution \( Q \) is defined on \( \mathcal{X} \times \mathcal{Y} \) and \( D(Q||P) \) represents the Kullback-Leibler divergence between the distributions \( Q \) and \( P \) [5].

In the next two sections, we will show the proofs of the achievability and the converse of Theorem 1 via the connection between the problem of deception with side information and the rate distortion problem with side information at both the encoder and decoder.

**III. THE ACHIEVABILITY**

In this section, we will show that there exists a deception function \( f \) that can achieve the deception exponent \( D(Q||P) + R_{SI}(Q, \Delta) \) for the distortion constraint \( \Delta \) and the distribution \( Q \) defined on \( \mathcal{X} \times \mathcal{Y} \). We will construct the deception function \( f \) from the rate distortion code with side information at both the encoder and decoder.

First, consider the rate distortion problem with side information at both the encoder and decoder as defined in subsection II-C. where \((X^n, Y^n)\) are generated i.i.d. according to the distribution \( Q \). It is shown in [5, Theorem 16.5] that for any sufficiently large \( n \), and \( 0 < \tau < 1 \), there exists a length \( n \) code that achieves the rate \( R_{SI}(Q, \Delta) \) and satisfies the distortion constraint with probability larger than \( 1 - \tau \). More specifically, for every \( \delta > 0 \) and \( 0 < \tau < 1 \), and sufficiently large \( n \), there exists a function pair \((g, \varphi)\) such that

\[
\frac{1}{n} \log |g| \leq R_{SI}(Q, \Delta) + \delta, \tag{5}\]

\[
\Pr \left( d(X^n, \varphi(g(X^n, Y^n), Y^n) \leq \Delta \right) \geq 1 - \tau. \tag{6}\]
Define $A \subset X^n \times Y^n$ as the set of sequences $(x^n, y^n)$ that satisfies the distortion constraint $\Delta$ under $(g, \varphi)$, i.e.,

$$A \triangleq \{(x^n, y^n) \in X^n \times Y^n : d(x^n, \varphi(g(x^n, y^n)), y^n) \leq \Delta\}. \tag{7}$$

The inequality in (6) is equivalent to

$$Q^n(A) \geq 1 - \tau, \tag{8}$$

where $Q^n(A)$ is the probability that an i.i.d. randomly generated $(X^n, Y^n)$ according to the distribution $Q$ falls in the set $A$, i.e.,

$$Q^n(A) \triangleq \sum_{(x^n, y^n) \in A} Q(x_i, y_i). \tag{9}$$

We further define $A(Q)$ as the intersection of the set $A$ with the typical set $T^n_{(Q)}$, where the definition of a typical set can be found in [5, Definition 2.8], i.e.,

$$A(Q) \triangleq A \cap T^n_{(Q)} \tag{10}$$

From [5, Lemma 2.12], we have

$$Q^n(T^n_{(Q)}) \geq 1 - \epsilon_n, \tag{11}$$

with $\epsilon_n \to 0$, and as a result, from (8) and (11), we have

$$Q^n(A(Q)) \geq Q^n(A) + Q^n(T^n_{(Q)}) - 1 \geq 1 - \epsilon_n - \tau. \tag{12}$$

Thus, from [5, Lemma 2.14], we have

$$\frac{1}{n} \log |A(Q)| \geq H(Q) - \epsilon_n, \tag{13}$$

with $\epsilon_n \to 0$. We further define $A_i(Q) \subset A(Q)$ for $i = 1, 2, \ldots, |g|$ as the set of $(x^n, y^n)$ sequences that satisfy the distortion constraint when mapped to index $i$ at the encoder, i.e.,

$$A_i(Q) \triangleq \{(x^n, y^n) \in T^n_{(Q)} : d(x^n, \varphi(i, y^n)) \leq \Delta\}. \tag{14}$$

Since all $(x^n, y^n)$ sequence pairs that satisfy the distortion constraint $\Delta$ under $(g, \varphi)$ have to be mapped to an index $g(x^n, y^n)$, we have

$$A(Q) = \bigcup_{i=1}^{||g||} A_i(Q). \tag{15}$$

We define $i^o$ as the index of the set $A_i(Q)$ with the largest cardinality, i.e.,

$$|A_{i^o}(Q)| = \max_{i \in \{1, 2, \ldots, ||g||\}} |A_i(Q)|. \tag{16}$$

Then, we have

$$|A_{i^o}(Q)| \geq \frac{|A(Q)|}{||g||}. \tag{17}$$

Therefore, we have

$$\frac{1}{n} \log P^n(A_{i^o}(Q)) \leq \frac{1}{n} \log \left( |A_{i^o}(Q)| \min_{(x^n, y^n) \in A_{i^o}(Q)} P^n((x^n, y^n)) \right)$$

$$\leq \frac{1}{n} \log \left( |A(Q)| \min_{(x^n, y^n) \in A_{i^o}(Q)} P^n((x^n, y^n)) \right)$$

$$\leq -H(Q) + \epsilon_n + R_{SI}(Q, \Delta) + \delta$$

$$- \frac{1}{n} \log \left( \min_{(x^n, y^n) \in A_{i^o}(Q)} P^n((x^n, y^n)) \right)$$

$$\leq -H(Q) + \epsilon_n + R_{SI}(Q, \Delta) + \delta$$

$$+ D(Q||P) + H(Q) + \epsilon_n$$

$$= D(Q||P) + R_{SI}(Q, \Delta) + \delta + 2\epsilon_n, \tag{18}$$

where (18) follows from (17), (19) follows from (5) and (13), and (20) follows because using [5, Lemma 2.6], we have that for any $(x^n, y^n)$,

$$\frac{1}{n} \log P^n((x^n, y^n)) = D(Q'||P) + H(Q'), \tag{22}$$

where $Q'$ is the empirical distribution of $(x^n, y^n)$. If $(x^n, y^n) \in T^n_{(Q)}$, we have [5, Definition 2.8]

$$|Q'(a, b) - Q(a, b)| \leq \delta, \quad \forall (a, b) \in X \times Y. \tag{23}$$

From delta-convention [5, Convention 2.11] and the continuity of the function $D(Q||P) + H(Q)$ in $Q$, we have

$$\frac{1}{n} \log P^n((x^n, y^n)) \leq D(Q||P) + H(Q) + \epsilon_n, \tag{24}$$

with $\epsilon_n \to 0$. Now, we construct the deception function $f$ according to $(g, \varphi)$ described above, i.e.,

$$f(y^n) = \varphi(i^o, y^n). \tag{25}$$

Then we have

$$\frac{1}{n} \log \Pr(d(X^n, f(Y^n)) \leq \Delta)$$

$$= \frac{1}{n} \log P^n \left( \{(x^n, y^n) : d(x^n, \varphi(i^n, y^n)) \leq \Delta\} \right) \tag{26}$$

$$\leq \frac{1}{n} \log P^n(A_{i^o}(Q)) \tag{27}$$

$$\leq D(Q||P) + R_{SI}(Q, \Delta) + \delta + 2\epsilon_n, \tag{28}$$

where (26) follows from the construction of $f$ in (25), (27) follows from the definition of $A_{i^o}(Q)$ in (14), and (28) follows from (21).

Thus, we have shown that there exists a deception function $f$, constructed according to the encoding and decoding function of the corresponding rate distortion problem with side information, that can achieve the deception exponent $D(Q||P) + R_{SI}(Q, \Delta)$ for any distribution $Q$ defined on $X \times Y$ and distortion constraint $\Delta$. This concludes the proof of the achievability.

### IV. The Converse

In this section, we will prove that for any deception function $f$, the deception exponent cannot be smaller than
min_\{D(Q||P) + R_{SI}(Q, \Delta)\}. This will be proven by contradiction. More specifically, we will show that if there is a deception function with the deception exponent achieving min_\{D(Q||P) + R_{SI}(Q, \Delta)\} − \alpha for some \alpha > 0, then we can construct a coding scheme in the rate distortion with side information problem with achievable rate smaller than \( R_{SI}(Q, \Delta) \), which is obviously false. In contrast to the achievability, in which we select one deception function \( f(y^n) \) from many rate distortion decoding functions \( \phi(i, y^n) \), \( i = 1, 2, \cdots, ||q|| \), i.e., (25), the converse requires us to construct many rate distortion decoding functions \( \phi(i, y^n) \), \( i = 1, 2, \cdots, ||q|| \) from one deception function \( f(y^n) \), which is more difficult.

We first assume a deception function \( f \), which achieves the deception exponent \( E \) under the distortion constraint \( \Delta \), as defined in (2). The proof of the converse includes the following three steps.

**Step 1 Type selection:** In this step, we will select one type among all the types, which contributes most to the deception probability for the function \( f \).

We define a set \( A \subset \mathcal{X}^n \times \mathcal{Y}^n \) as follows

\[
A \triangleq \{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : d(x^n, f(y^n)) \leq \Delta\}. \tag{29}
\]

Based on the definition of set \( A \) and the definition of deception exponent in (2), the deception exponent \( E \) is achievable means that

\[
-\frac{1}{n} \log P^n(A) \leq E + \delta. \tag{30}
\]

We further define \( \mathcal{P} \) as the set of all possible empirical distributions on \( \mathcal{X}^n \times \mathcal{Y}^n \), i.e.,

\[
\mathcal{P} \triangleq \left\{ Q : Q(x,y) = \frac{1}{n}, \forall(x,y) \in \mathcal{X} \times \mathcal{Y}, i \in \{0,1,2,\ldots,n\} \right\}. \tag{31}
\]

For any distribution \( Q \in \mathcal{P} \), we define the set \( A(Q) \) as the intersection between the set \( A \) and \( T_Q^n \), where \( T_Q^n \) is the type of \( Q \) as defined in [5, Definition 2.1], i.e.,

\[
A(Q) = A \cap T_Q^n. \tag{32}
\]

Thus, we have

\[
A = \bigcup_{Q \in \mathcal{P}} A(Q). \tag{33}
\]

Since the types \( T_Q^n \) for different empirical distributions \( Q \in \mathcal{P} \) are disjoint, we have

\[ P^n(A) = \sum_{Q \in \mathcal{P}} P^n(A(Q)). \tag{34}\]

We define the type \( Q^o \) as the type which contributes most to the deception probability, i.e.,

\[ P^n(A(Q^o)) = \max_{Q \in \mathcal{P}} P^n(A(Q)). \tag{35}\]

Let \( \beta \) be the number of different types in \( \mathcal{X}^n \times \mathcal{Y}^n \). From (34) and (35), we have

\[ P^n(A(Q^o)) \geq \frac{P^n(A)}{\beta}. \tag{36}\]

Thus, we have

\[
-\frac{1}{n} \log P^n(A(Q^o)) \leq -\frac{1}{n} \log \frac{P^n(A)}{\beta} \leq E + \delta + \frac{1}{n} \log \beta \tag{37}
\]

\[
\leq E + \epsilon_n + \delta, \tag{38}
\]

where (37) follows from (36), (38) follows from (30), and (39) follows from the fact that [5, Lemma 2.2]

\[
\beta \leq (n + 1)^{||X||||Y||} \leq \exp(n\epsilon_n), \tag{40}
\]

with \( \epsilon_n \to 0 \).

Since every sequence in the same type is equally probable, we have

\[ P^n(A(Q^o)) = |A(Q^o)| P^n(x^n, y^n), \quad \forall(x^n, y^n) \in T_Q^n. \tag{41}\]

Thus, for any sequence pair \((x^n, y^n) \in T_Q^n\), we have

\[
-\frac{1}{n} \log |A(Q^o)| = -\frac{1}{n} \log \frac{P^n(A(Q^o))}{P^n(x^n, y^n)} \leq -E - \epsilon_n - \delta - \frac{1}{n} \log P^n((x^n, y^n)) \tag{42}
\]

\[
= D(Q^n||P) + H(Q^o) - E - \epsilon_n - \delta. \tag{43}
\]

where (42) follows from (39), and (43) follows from [5, Lemma 2.6], i.e., for distribution \( Q^o \),

\[ -\frac{1}{n} \log P^n((x^n, y^n)) = D(Q^n||P) + H(Q^o), \quad \forall(x^n, y^n) \in T_Q^n. \]

**Step 2 Permutation:** In this step, we will construct a rate distortion code, restricted to the type \( Q^o \), with side information at both the encoder and decoder from the deception function \( f \) via permutations.

We consider the symmetric group \( S_n \), which consists of all the permutations on \( \{1,2,\ldots,n\} \). For a set \( S \subset \mathcal{X}^n \), and a permutation \( \pi \in S_n \), define the set \( \pi(S) \) as the set of sequences that is permuted from the sequences in set \( S \) by the permutation \( \pi \), i.e.,

\[ \pi(S) \triangleq \{ \bar{x^n} \in \mathcal{X}^n : \exists x^n \in S, \pi(x^n) = \bar{x^n} \}. \tag{44}\]

Thus, \( \pi(S) \) is a permuted version of the set \( S \).

We will use the following lemma, proved by Ahlswede [8] in 1980, to obtain a covering of \( T_Q^n \) by the permuted versions of \( A(Q^o) \).

**Lemma 1 (Covering Lemma) [8, Section 6.1]** For any set \( S \subset T_Q^n \), there exist permutations \( \pi_1, \pi_2, \ldots, \pi_k \in S_n \) with

\[ \bigcup_{i=1}^{k} \pi_i(S) = T_Q^n, \tag{45}\]
if 

\[ k > \frac{|T_Q^n|}{|S|} \log |T_Q^n|. \quad (46) \]

From the above lemma, by letting \( Q \) be \( Q^n \) and \( S \) be \( A(Q^n) \), we have that there exist permutations \( \pi_1, \pi_2, \ldots, \pi_k \in \mathcal{S}_n \) such that

\[ \bigcup_{i=1}^k \pi_i(A(Q^n)) = T_Q^n. \quad (47) \]

where \( k \) satisfies

\[ \frac{1}{n} \log k = \frac{1}{n} \log |T_Q^n| - \frac{1}{n} \log |A(Q^n)| + \frac{1}{n} \log (\log |T_Q^n|) + \epsilon_n \]

\[ \leq E - D(Q^n||P) + 2c_n + \delta + \frac{\log(nH(Q^n))}{n}, \quad (48) \]

where (48) follows from (43) and the fact that the size of the type \( |T_Q^n| \) satisfies [5, Lemma 2.3]

\[ |T_Q^n| \leq \exp(nH(Q^n)). \]

Based on (29) and (32), we have

\[ A(Q^n) = \{ (x^n, y^n) \in T_Q^n : d(x^n, f(y^n)) \leq \Delta \}. \quad (49) \]

Based on the definition in (1), we see that the same permutation of the two sequences does not change the distortion between the two sequences. Therefore we have for \( i = 1, 2, \ldots, k \)

\[ A(Q^n) = \{ (x^n, y^n) \in T_Q^n : d(\pi_i(x^n), \pi_i(f(y^n))) \leq \Delta \}. \quad (50) \]

From the definition of the permutation of a set in (44), we have

\[ \pi_i(A(Q^n)) = \{ (\bar{x}^n, \bar{y}^n) \in T_Q^n : \exists (x^n, y^n) \in A(Q^n) \] such that \((\pi_i(x^n), \pi_i(y^n)) = (\bar{x}^n, \bar{y}^n)\}. \quad (51) \]

Thus, by combining the above two equations, we can view the set \( \pi_i(A(Q^n)) \) as

\[ \pi_i(A(Q^n)) = \{ (\bar{x}^n, \bar{y}^n) \in T_Q^n : d(\bar{x}^n, \pi_i(f(\pi_i^{-1}(\bar{y}^n)))) \leq \Delta \}. \quad (52) \]

In other words, the permuted set \( \pi_i(A(Q^n)) \) can be characterized by a composite function \( \pi_i(f(\pi_i^{-1}(-))) \).

With the sets \( \pi_i(A(Q^n)) \) and the functions \( \pi_i(f(\pi_i^{-1}(-))) \) for \( i = 1, 2, \ldots, k \), we are ready to construct a rate distortion code with side information. We assume that \((X^n, Y^n)\) are generated i.i.d. according to distribution \( Q^n \). We will construct an encoding-decoding function pair \((g', \varphi')\) for all \((x^n, y^n) \in T_Q^n\) as follows.

We define \( g'(x^n, y^n) = i \) if \((x^n, y^n) \in \pi_i(A(Q^n))\). If there exist multiple sets \( \pi_i(A(Q^n)) \) to which \((x^n, y^n)\) belongs, we can arbitrarily pick one set and assign the index of the set to the output of the function \( g' \). Define \( Q^n_\pi \) as the marginal distribution of \( Q^n \) in \( \mathcal{Y} \). Then, for \( i = 1, 2, \ldots, k \) and \( y^n \in \mathcal{T}_{Q^n_\pi} \), we define the decoding function \( \varphi' \) as follows

\[ \varphi'(i, y^n) = \pi_i(f(\pi_i^{-1}(y^n))), \quad i = 1, 2, \ldots, k. \quad (53) \]

Due to the covering in (47), we obtain an encoding-decoding function pair \((g', \varphi')\) for the rate-distortion problem with side information available at both the encoder and decoder for every \((x^n, y^n) \in T_Q^n\), which satisfies

\[ ||g'|| = k, \quad (54) \]

\[ d(x^n, \varphi'(g'(x^n, y^n), y^n)) \leq \Delta. \quad (55) \]

**Step 3 Blowing-up:** In the previous step, we have constructed a code \((g', \varphi')\) for every sequence pair \((x^n, y^n)\) in the type \( T_Q^n \). In this step, we will expand the code, first to the neighborhood of the type \( T_Q^n \), and then to the whole space \( X^n \times Y^n \). The expansion uses the blowing-up lemma [5, Chapter 5].

First, we expand the code we described in the previous step to the neighborhood of the type \( T_Q^n \). To do so, let us first introduce the definition of the neighborhood of a set as follows.

**Definition 2** [5, Chapter 5] Given a set \( S \subset X^n \), we define the Hamming \( l \) neighborhood of \( S \) as

\[ \Gamma^l(S) \triangleq \{ x^n \in X^n : \exists x^n \in S, \text{ s.t. } d_H(x^n, x^n) \leq l \}, \quad (56) \]

where \( d_H(x^n, x^n) \) denotes the Hamming distance between sequences \( x^n \) and \( x^n \), i.e., the number of positions in which the two sequences differ.

Assume a sequence of positive integer \( l_n \) with \( l_n \) converging to 0. We consider the \( l_n \) neighborhood of the type \( T_Q^n \), i.e., \( \Gamma^{l_n}(T_Q^n) \). Based on Definition 2 and (47), we have

\[ \Gamma^{l_n}(T_Q^n) = \bigcup_{i=1}^k \Gamma^{l_n}(\pi_i(A(Q^n))). \quad (57) \]

For any \((x^n, y^n) \in \Gamma^{l_n}(T_Q^n)\), we will construct a rate distortion code with side information at both the encoder and decoder \((g, \varphi)\) from the rate distortion code \((g', \varphi')\) defined on the type \( T_Q^n \) described in the Step 2. The basic idea is that for \((x^n, y^n)\) in the neighborhood of \((x^n, y^n) \in T_Q^n\), we utilize the encoder-decoder pair \((g', \varphi')\) for \((x^n, y^n)\) to construct the encoder-decoder pair for \((x^n, y^n)\). The details are as follows.

The encoding function \( g \) on the set \( \Gamma^{l_n}(T_Q^n) \) includes two parts, i.e., \( g = (g_1, g_2) \). The function \( g_1 \) can be constructed in a similar way as we constructed the function \( g \) on the type \( T_Q^n \) in Step 2. More specifically, for any \((x^n, y^n) \in \Gamma^{l_n}(T_Q^n)\), we define \( g_1(x^n, y^n) = i \) if \((x^n, y^n) \in \Gamma^{l_n}(\pi_i(\pi(A(Q^n)))) \).

Based on (57), we know such \( i \) always exists. If there exist multiple sets \( \Gamma^{l_n}(\pi_i(\pi(A(Q^n)))) \) to which \((x^n, y^n)\) belongs, we can arbitrarily pick one set and assign the index of the set to the output of the function \( g_1 \). Then we have

\[ ||g_1|| = k. \quad (58) \]

Once we determine that \( g_1(x^n, y^n) = i \), we can assert that there exists a sequence pair \((\bar{x}^n, \bar{y}^n) \in \pi_i(A(Q^n))\) such that

\[ d_H((x^n, y^n), (\bar{x}^n, \bar{y}n)) \leq l_n. \quad (59) \]
We would like to adopt \( \varphi'(i, \tilde{y}^n) \) as the decoding function \( \varphi(i, y^n) \). However, \( \tilde{y}^n \) is determined based on sequence pair \((x^n, y^n)\). Therefore, the decoder, with only the knowledge of \( y^n \), can not determine \( \tilde{y}^n \) by itself. To solve this problem, we need to design \( g_2 \), i.e., the second part of encoding function, to inform the decoder of \( \tilde{y}^n \) as follows.

Let us consider \( y^n \), the second sequence in the pair \((x^n, y^n)\). We construct a \( l_n \) Hamming neighborhood \( \Gamma^n(y^n) \) around the sequence \( y^n \), which is called a Hamming ball, and give every sequence in \( \Gamma^n(y^n) \) an index. Obviously, the sequence \( \tilde{y}^n \) can be uniquely determined by the sequence \( y^n \) together with the index of \( \tilde{y}^n \) with respect to \( y^n \), say \( j \in \{1, 2, \ldots, |\Gamma^n(y^n)|\} \). Then \( g_2 \), the second part of the encoding function \( g \), can be defined as the above index, i.e.,

\[
g_2(x^n, y^n) = j, \quad (60)
\]

and we have

\[
||g_2|| = |\Gamma^n(y^n)|, \quad (61)
\]

where the size of the Hamming ball satisfies [5, Lemma 5.1] \( \frac{1}{n} \log |\Gamma^n(y^n)| \leq h \left( \frac{l_n}{n} \right) + \frac{l_n}{n} \log |Y|, \quad (62) \)

and \( h(\cdot) \) represents the binary entropy function. Therefore, the size of the function \( g \) is

\[
||g|| = ||g_1|| \cdot ||g_2|| = k |\Gamma^n(y^n)|. \quad (63)
\]

Note that for \((x^n, y^n) \in \Gamma^n(T_{Q^n})\), we have \( y^n \in \Gamma^n(T_{Q^n}) \). Thus, we expand the domain of the second argument of the function \( \varphi \) from the type \( T_{Q^n} \) to its \( l_n \) neighborhood \( \Gamma^n(T_{Q^n}) \).

Assume the decoder observes \( y^n \in \Gamma^n(T_{Q^n}) \), and obtain \( g(x^n, y^n) \) from the encoder as follows

\[
g(x^n, y^n) = (i, j). \quad (64)
\]

From \( y^n \) and \( j \), we can determine the sequence \( \tilde{y}^n \) in the type \( T_{Q^n} \). We then define \( \varphi'(i, j, y^n) = \varphi'(i, \tilde{y}^n) \) for \( i = 1, 2, \ldots, k \), where the decoding function \( \varphi'(i, \tilde{y}^n) \) was defined in the previous step.

Based on the definition of \( \Gamma^n(\pi_i(A(Q^n))) \), for every \((x^n, y^n) \in \Gamma^n(\pi_i(A(Q^n)))\), there exists a \((\tilde{x}^n, \tilde{y}^n) \in \pi_i(A(Q^n))\) such that

\[
d_H((x^n, y^n), (\tilde{x}^n, \tilde{y}^n)) \leq l_n, \quad (65)
\]

and from the definition of \( \pi_i(A(Q^n)) \), we know that

\[
d(\tilde{x}^n, \varphi'(i, \tilde{y}^n)) \leq \Delta. \quad (66)
\]

Therefore, we have

\[
d(x^n, \varphi(g(x^n, y^n), y^n)) = d(x^n, \varphi'(g'(\tilde{x}^n, \tilde{y}^n), \tilde{y}^n)) \leq d(\tilde{x}^n, \varphi'(g'(\tilde{x}^n, \tilde{y}^n), \tilde{y}^n)) + d_M \frac{l_n}{n}, \quad (67)
\]

where (67) is because \( x^n \) and \( \tilde{x}^n \) at most differ in \( l_n \) positions and

\[
d_M \Delta \leq \max_{(x, \tilde{x})} d(x, \tilde{x}). \quad (68)
\]

Hence, with the above definition of \((g, \varphi)\), we have that

\[
d(x^n, \varphi(g(x^n, y^n), y^n)) \leq \Delta + d_M \frac{l_n}{n}. \quad (70)
\]

Finally, we expand the rate-distortion code \((g, \varphi)\) to all \((x^n, y^n) \in X^n \times Y^n\). For \((x^n, y^n) \notin \Gamma^n(T_{Q^n})\), we define \( g(x^n, y^n) = 0 \), and we define \( \varphi(0, y^n) \) to be an arbitrary sequence in \( X^n \).

Therefore, for \((g, \varphi)\) defined on \( X^n \times Y^n \), we have

\[
\frac{1}{n} \log ||g|| \leq E - D(Q^n||P) + 3\epsilon_n + \delta + \frac{\log(nH(Q^n))}{n} + h \left( \frac{l_n}{n} \right) + \frac{l_n}{n} \log |Y|, \quad (71)
\]

where (71) follows from (48), (63) and (62).

We use the blowing up lemma to calculate the average distortion for the rate distortion code we constructed above. We restate the blowing up lemma as follows.

**Lemma 2 (Blowing up) [5, Lemma 5.4]** To any finite set \( X \) and sequence \( \epsilon_n \rightarrow 0 \), there exist a sequence of positive integers \( l_n \) with \( l_n \rightarrow 0 \) and a sequence \( \tau_n \rightarrow 0 \) such that for any distribution \( Q \) defined on \( X \) and every \( n, A \subset X^n \)

\[
Q^n(A) \geq \exp(-n\epsilon_n) \quad (72)
\]

implies

\[
Q^n(\Gamma^n(A)) \geq 1 - \tau_n. \quad (73)
\]

In our case, from [5, Lemma 2.3], we have that for any sequence \( \epsilon_n \rightarrow 0 \) and sufficiently large \( n \)

\[
(Q^n)^n(T_{Q^n}) \geq (n + 1)^{-|X|^3}|X| \geq \exp(-n\epsilon_n). \quad (74)
\]

Thus, from the blowing up lemma, we have that there exists a sequence \( \eta_n \rightarrow 0 \) such that

\[
(Q^n)^n(T_{Q^n}) \geq 1 - \eta_n. \quad (75)
\]

Therefore, for \((g, \varphi)\) defined on \( X^n \times Y^n \),

\[
\Pr \left( d(x^n, \varphi(g(x^n, y^n), y^n)) \leq \Delta + d_M \frac{l_n}{n} \right) \geq (Q^n)^n(T_{Q^n}) \geq 1 - \eta_n. \quad (76)
\]

where (76) follows from (67) and (75). The inequality in (76) leads to

\[
E(d(X^n, \varphi(g(X^n, Y^n), Y^n))) \leq \Delta + d_M \frac{l_n}{n} + d_M \eta_n. \quad (77)
\]

Thus, if we have a deception function, which under the distortion constraint \( \Delta \) can achieve the deception exponent

\[
E = \min_Q D(Q||P) + R_{ST}(Q, \Delta) - \alpha, \quad (78)
\]

for some \( \alpha > 0 \), then we can construct a rate distortion code with side information at both the encoder and decoder, where \((X^n, Y^n)\) is generated i.i.d. according to the distribution \( Q^n \),
that satisfies
\[
\frac{1}{n} \log |g| \leq \min_Q \{D(Q||P) + R_{SI}(Q, \Delta)\} - \alpha - D(Q^o||P) + 3\epsilon_n + \delta + \frac{\log(nH(Q^o))}{n} + h\left(\frac{l_n}{n}\right) + \frac{l_n}{n} \log |Y|
\]
\[
\leq R_{SI}(Q^o, \Delta) - \alpha + 3\epsilon_n + \delta + \frac{\log(nH(Q^o))}{n} + h\left(\frac{l_n}{n}\right) + \frac{l_n}{n} \log |Y|,
\]
and
\[
E(d(X^n, \varphi(g(X^n, Y^n)), Y^n)) \leq \Delta + d_M \frac{l_n}{n} + d_M \eta_n.
\]

Since the rate distortion function \(R_{SI}(Q^o, \Delta)\) is a continuous function of \(\Delta\), the above result contradicts with the result in the rate distortion problem with side information at both the encoder and decoder. This concludes the proof of the converse.

V. CONCLUSION

In this paper, we studied the probability of successful deception of an uncompressed biometric authentication system with side information at the adversary. We found the optimal exponent of the deception probability by providing the proofs of both the achievability and the converse. The results are proved by exploiting a connection between the problem of deception with side information and the rate distortion problem with side information at both the encoder and decoder.

REFERENCES


Wei Kang received the B. Eng. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2001, the M.Eng. degree in electrical engineering from McGill University, Montreal, Canada, in 2003, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 2008. He joined the School of Information Science and Engineering, Southeast University, Nanjing, China, in 2009, where he is currently an Associate Professor. His research mainly focuses on network information theory.

Daming Cao received the B. Eng. degree in information engineering from Southeast University, Nanjing, China, in 2013. He is currently a master student in Southeast University. His research mainly focuses on information theory and security.

Nan Liu received the B. Eng. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, P. R. China in 2001, and the Ph.D. degree in electrical and computer engineering from University of Maryland, College Park, MD in 2007.

From 2007-2008, she was a postdoctoral scholar in the Wireless Systems Lab, Department of Electrical Engineering, Stanford University. In 2009, she became a professor in the National Mobile Communications Research Laboratory, School of Information Science and Engineering in Southeast University, Nanjing, China. Her research interests are in network information theory for wireless networks, SON algorithms for next generation cellular networks and energy-efficient communications.