Capacity and Random-Coding Error Exponent for Public Fingerprinting Game

Ying Wang and Pierre Moulin

Department of Electrical and Computer Engineering

University of Illinois at Urbana-Champaign

Urbana, IL 61801 USA

Email: {ywang11, moulin}@ifp.uiuc.edu

Abstract—Capacity and random-coding error exponent formulas are derived for a *public* fingerprinting (traitor tracing) game. The original media copy is available to the encoder, but not to the decoder. We derive the random-coding error exponent for a stacked binning scheme. The exponent is strictly positive at all rates below capacity. The converse part of the capacity proof is based on the Gel'fand-Pinsker technique.

I. INTRODUCTION

Content fingerprinting is one of many applications of information hiding. The same media covertext—image, video, audio, or text—is distributed to many users. A fingerprint, a mark unique to each user, is embedded into each copy of the distributed covertext. In a collusion attack, two or more users may combine their copies in an attempt to “remove” their fingerprints and forge a pirated copy. The distortion between the pirated copy and the colluding copies is bounded by a certain tolerance level. To trace the forgery back to members of the coalition, we need fingerprinting codes that can reliably identify fingerprints of those members. Essentially, from a communication viewpoint, the fingerprinting problem is a multiuser version of the watermarking problem, where the attack is by one user based on one single copy.

Depending on the availability of the original covertext to the decoder, there are two types of fingerprinting applications: private and public. In the *private fingerprinting* setup, the covertext is available to both the encoder and decoder. *Public fingerprinting*, where the original covertext is available to the encoder but not to the decoder, has some advantages over private fingerprinting. For example, public fingerprinting does not require the vast storage and computation resources that are needed for media registration in a large database. Also, it offers the flexibility of distributed detection, e.g., a DVD player can detect fingerprints from a movie disc and refuse to play it if fingerprints other than the owner's are present.

The capacity and error exponent limits of private fingerprinting have been derived by Somekh-Baruch and Merhav [1][2]. In this paper, we study the capacity and error exponent limits of *public* fingerprinting games. The specific setup we consider here is a discrete memoryless collusion channel and a detect-all decoder, which means there is an error if not all members of the coalition are caught. A binning scheme is used in the capacity-achievability proof because, unlike in the private setup, the availability of side information to the encoder

and decoder is asymmetric. To optimize the balance between the probabilities of encoding error and decoding error, we use a **stacked binning** scheme that we previously proposed for single-user channel coding with side information problems [3]. We obtain a single-letter lower bound on the random coding error exponent that is strictly positive at rate below capacity. Also, we derive a single-letter expression of capacity whose converse proof is based on the telescoping technique used by Gel'fand and Pinsker [4].

II. STATEMENT OF THE PROBLEM

Public fingerprinting with side information at the encoder is modelled in Fig. 1. Let $\mathcal{S}$, $\mathcal{X}$, and $\mathcal{Y}$ be three finite alphabet sets. The covertext sequence $\mathbf{S} = (S_1, \dots, S_N)$ consists of independent and identically distributed (i.i.d.) samples drawn from a probability mass function (p.m.f.) $p_S(s)$, $s \in \mathcal{S}$. The fingerprinting encoder generates $|\mathcal{M}| = \lceil 2^{NR} \rceil$ fingerprinted copies $\mathbf{X}_m = f_N(\mathbf{S}, m)$, where $f_N : \mathcal{S}^N \times \mathcal{M} \rightarrow \mathcal{X}^N$. Each fingerprint m corresponds to a user. Assume that the collusion size $L = 2$ ¹. Let M_1 and M_2 be two users selected independently and uniformly from the set $\mathcal{M}$. They collude to produce a pirated copy, or forgery, $\mathbf{Y}$, through an attack channel $p_{\mathbf{Y}|\mathbf{X}_1\mathbf{X}_2}$. The decoder does not know $p_{\mathbf{Y}|\mathbf{X}_1\mathbf{X}_2}$ selected by the two colluders and does not have access to the original covertext. The decoder produces estimates $(\hat{M}_1, \hat{M}_2) = g_N(\mathbf{Y}) \in \mathcal{M}^2$ in attempt to catch the two colluders. We allow the encoder/decoder pair (f_N, g_N) to be *randomized*, and the choice of (f_N, g_N) is a function of a random variable known to the encoder and decoder but not to the colluders. We can think of this random variable as a *secret key*. Formally, the randomized code will be denoted by (F_N, G_N) .

A. Notation

We use uppercase letters for random variables, lowercase letters for individual values, and boldface fonts for sequences. Entropy of a random variable X is denoted by $H(X)$, mutual information between two random variables X and Y is denoted by $I(X; Y)$, and Kullback-Leibler divergence between two p.m.f.'s p and q is denoted

¹In this paper, we focus on the case of two colluders for the simplicity of notation. Although cumbersome, it is conceptually straightforward to extend to the case where there are more than two colluders.

Fig. 1. Fingerprinting games.

by $D(p||q)$. We denote by $D(p_{Y|X}||q_{Y|X}|p_X) = \sum_x p_X(x) \sum_y p_{Y|X}(y|x) \log \frac{p_{Y|X}(y|x)}{q_{Y|X}(y|x)}$ the conditional Kullback-Leibler divergence of $p_{Y|X}$ and $q_{Y|X}$ with respect to p_X . Let p_x denote the type of a sequence $\mathbf{x} \in \mathcal{X}^N$ (p_x is a p.m.f. over $\mathcal{X}$) and T_x the type class associated with p_x , i.e., the set of all sequences of type p_x . Likewise, we define the joint type $p_{\mathbf{xy}}$ of a pair of sequences $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^N \times \mathcal{Y}^N$ (a p.m.f. over $\mathcal{X} \times \mathcal{Y}$) and $T_{\mathbf{xy}}$ the type class associated with $p_{\mathbf{xy}}$, i.e., the set of all sequences of type $p_{\mathbf{xy}}$. Finally, we define the conditional type $p_{\mathbf{y}|\mathbf{x}}$ of a pair of sequences $(\mathbf{x}, \mathbf{y})$ as $\frac{p_{\mathbf{xy}}(\mathbf{x}, \mathbf{y})}{p_{\mathbf{x}}(\mathbf{x})}$ for all $\mathbf{x} \in \mathcal{X}$ such that $p_{\mathbf{x}}(\mathbf{x}) > 0$. The conditional type class $T_{\mathbf{y}|\mathbf{x}}$, given $\mathbf{x}$, is the set of all sequences $\tilde{\mathbf{y}}$ such that $(\mathbf{x}, \tilde{\mathbf{y}}) \in T_{\mathbf{xy}}$. We denote by $H(\mathbf{x})$ the entropy of the p.m.f. p_x and by $I(\mathbf{x}; \mathbf{y})$ the mutual information for the joint p.m.f. $p_{\mathbf{xy}}$. We let $\mathcal{P}(\mathcal{X})$ and $\mathcal{P}^N(\mathcal{X})$ represent the set of all p.m.f.'s and all empirical p.m.f.'s, respectively, on the alphabet $\mathcal{X}$. Likewise, $\mathcal{P}(\mathcal{Y}|\mathcal{X})$ and $\mathcal{P}^N(\mathcal{Y}|\mathcal{X})$ denote the set of all conditional p.m.f.'s and all empirical conditional p.m.f.'s on the alphabet $\mathcal{Y}$. We define $|t|^+ = \max(0, t)$.

B. Constrained Fingerprint Codes with Side Information

A distortion function $d_1 : \mathcal{S} \times \mathcal{X} \rightarrow \mathbb{R}^+$ is defined to quantify the distortion between a fingerprint symbol x and a covertext symbol s . This definition is extended to N -vectors using $d_1^N(\mathbf{s}, \mathbf{x}) = \frac{1}{N} \sum_{i=1}^N d_1(s_i, x_i)$.

We now define a class of codes that satisfy maximum distortion constraints (Def. 2.1).

Definition 2.1: A length- N , rate- R , randomized code with side information and **maximum distortion** D_1 is a triple $(\mathcal{M}, F_N, G_N)$, where

- $\mathcal{M}$ is the fingerprint set of cardinality $|\mathcal{M}| = \lceil 2^{NR} \rceil$;
- (F_N, G_N) has joint distribution $p(f_N, g_N)$;
- $f_N : \mathcal{S}^N \times \mathcal{M} \rightarrow \mathcal{X}^N$ is the encoder mapping the covertext sequence $\mathbf{s}$ and fingerprint m to the fingerprinted sequence $\mathbf{x}(\mathbf{s}, m) = f_N(\mathbf{s}, m)$. The mapping is subject to the distortion constraint

$$d_1^N(\mathbf{s}, f_N(\mathbf{s}, m)) \leq D_1 \quad \text{almost surely } (p_{\mathcal{S}} p_{F_N} p_{\mathcal{M}}); \quad (2.1)$$

- $g_N : \mathcal{Y}^N \rightarrow \mathcal{M}^2$ is the decoder mapping the received sequence $\mathbf{y}$ to decoded fingerprints $(\hat{m}_1, \hat{m}_2) = g_N(\mathbf{y})$.

We use randomized codes in anticipation of collusion attacks with arbitrary memory [2]; in this paper, however, only memoryless strategies will be allowed for the coalition.

Definition 2.2: A transmit channel $p_{XU|S}$ is feasible if

$$\sum_{x, u, s} p_{XU|S}(x, u|s) p_S(s) d_1(s, x) \leq D_1$$

where $U \in \mathcal{U}$ is an arbitrary auxiliary random variable. We denote by $\mathcal{P}_{XU|S}(D_1)$ the set of feasible transmit channels.

C. Constrained Collusion Attack Channels

We define a constrained discrete memoryless collusion channel (DMCC) as follows.

Definition 2.3: A constrained DMCC is a p.m.f. $p_{Y|X_1 X_2}$ subject to a **fairness** constraint

$$p_{Y|X_1 X_2} = p_{Y|X_2 X_1} \quad (2.2)$$

and the **expected distortion** constraint

$$\sum_{x_1, x_2, y} p_{X_1, X_2}(x_1, x_2) p_{Y|X_1 X_2}(y|x_1, x_2) d(x_i, y) \leq D_2, \quad i = 1, 2. \quad (2.3)$$

Its N -dimensional extension is

$$p_{Y|\mathbf{X}_1 \mathbf{X}_2}(\mathbf{y}|\mathbf{x}_2 \mathbf{x}_1) = \prod_{i=1}^N p_{Y|X_1 X_2}(y_i|x_{1i}, x_{2i}). \quad (2.4)$$

We denote the set of feasible DMMCs by $\mathcal{P}_{Y|X_1 X_2}(D_2)$.

We pose the fairness constraint and the D_2 constraint on both colluders because neither one of the colluders wants to take more risks than the other.

D. Probability of Error

In this paper, we only consider the case of detecting all colluders. An error occurs when the decoded two fingerprints do not match the fingerprints of real colluders:

$$(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2).$$

The average probability of error for a deterministic code (f_N, g_N) when a DMCC $p_{Y|X_1 X_2}$ is in effect is denoted by $P_e(f_N, g_N, p_{Y|X_1 X_2})$. For a randomized code the expression above is averaged with respect to $p(f_N, g_N)$. The minmax probability of error for the class of randomized codes and the class of constrained DMCCs considered is given by

$$\begin{aligned} P_{e,N} &= \min_{p(f_N, g_N)} \max_{p_{Y|X_1 X_2}} \sum_{f_N, g_N} p(f_N, g_N) P_e(f_N, g_N, p_{Y|X_1 X_2}) \\ &= \min_{F_N, G_N} \max_{p_{Y|X_1 X_2}} P_e(F_N, G_N, p_{Y|X_1 X_2}) \end{aligned} \quad (2.5)$$

Definition 2.4: A rate R is said to be achievable if $P_{e,N} \rightarrow 0$ as $N \rightarrow \infty$.

Definition 2.5: The capacity $C(D_1, D_2)$ is the supremum of all achievable rates.

Definition 2.6: The reliability function for the class of randomized codes and attack channels considered is defined as

$$E(R) = \limsup_{N \rightarrow \infty} \left[-\frac{1}{N} \log P_{e,N} \right]. \quad (2.6)$$

where $P_{e,N}$ is given in (2.5).

III. MAIN RESULTS

The main tool used to prove the random coding error exponent bound in this paper is the method of types.

Theorem 3.1: For the DMCC model (Def. 2.3), the reliability function is lower-bounded by the random-coding error exponent

$$E_r^{DMCC}(R) = \min_{\tilde{p}_S \in \mathcal{P}_S} \sup_{\mathcal{U}} \min_{\tilde{p}_{X_2 U_2 | X_1 U_1 S}} \max_{\substack{p_{X_1 U_1 | S} = p_{X_2 U_2 | S} \\ \in \mathcal{P}_{XU|S}(D_1)}} \min_{\substack{\tilde{p}_{Y | X_1 X_2 U_1 U_2 S} \\ \tilde{p}_{Y | X_1 X_2} \in \mathcal{P}_{Y | X_1 X_2}(D_2)}} \left[D(\tilde{p}_S p_{X_1 U_1 | S} \tilde{p}_{X_2 U_2 | X_1 U_1 S} \tilde{p}_{Y | X_1 X_2 U_1 U_2 S} \| p_S p_{X_1 U_1 | S} p_{X_2 U_2 | S} p_{Y | X_1 X_2}) + J(\tilde{p}_S, p_{X_1 U_1 | S}, \tilde{p}_{X_2 U_2 | X_1 U_1 S}, \tilde{p}_{Y | X_1 X_2 U_1 U_2 S}) \right] \quad (3.7)$$

for all $R \leq C$, where $U_1 \in \mathcal{U}$, $U_2 \in \mathcal{U}$,

$$\sum_{X_1, U_1} \tilde{p}_{X_2 U_2 | X_1 U_1 S} \cdot p_{X_1 U_1 | S} = p_{X_2 U_2 | S} = p_{X_1 U_1 | S},$$

and

$$J(\tilde{p}_S, p_{X_1 U_1 | S}, \tilde{p}_{X_2 U_2 | X_1 U_1 S}, \tilde{p}_{Y | X_1 X_2 U_1 U_2 S}) = \left| \min \left\{ I(U_1; Y | U_2) - I(U_1; S | U_2) + I(U_1; U_2 | S) - R, \right. \right. \\ \left. I(U_2; Y | U_1) - I(U_2; S | U_1) + I(U_1; U_2 | S) - R, \right. \\ \left. I(U_1, U_2; Y) - I(U_1, U_2; S) + I(U_1; U_2 | S) - 2R \right\} \right|^+.$$

Due to symmetry— $p_{Y | X_1 X_2} = p_{Y | X_2 X_1}$ and $p_{X_1 U_1 | S} = p_{X_2 U_2 | S}$ —either one of the first two terms in the functional J can actually be dropped. $E_r^{DMCC}(R)$ is strictly positive at all rates below capacity and determined by the worst type class $T_{\mathbf{s} \mathbf{u}_1 \mathbf{x}_1 \mathbf{u}_2 \mathbf{x}_2 \mathbf{y}}$. The random-coding error exponent (3.7) is achieved by conditionally constant-composition codes using a **stacked binning** technique and a **maximum penalized mutual information** (MPMI) decoder [3]. A sketch of proof appears in Sec. IV.

Theorem 3.2: For the DMCC model (Def. 2.3), the capacity is given by

$$C(D_1, D_2) = \sup_{\mathcal{U}} \max_{p_{XU|S} \in \mathcal{P}_{XU|S}(D_1)} \min_{p_{Y | X_1 X_2} \in \mathcal{P}_{Y | X_1 X_2}(D_2)} \min \left\{ I(U_1; Y | U_2) - I(U_1; S | U_2), I(U_2; Y | U_1) - I(U_2; S | U_1), \right. \\ \left. \frac{1}{2} [I(U_1, U_2; Y) - I(U_1, U_2; S)] \right\},$$

where $U_1 \in \mathcal{U}$, $U_2 \in \mathcal{U}$, and $p_{X_1 U_1 | S} = p_{X_2 U_2 | S} = p_{XU|S}$.

The proof of the direct part of Theorem 3.2 is an immediate consequence of the achievability theorem 3.1 for DMMC's, and the proof of the converse part is given in Sec. V.

IV. SKETCH OF PROOF OF THEOREM 3.1

The encoder is the stacked random binning scheme of [3].

Let $\mathcal{U}$ be a finite set. We define

$$\rho(p_S) \triangleq I(\mathbf{u}; \mathbf{s}) + \eta_N,$$

Fig. 2. A binning scheme with a stack of variable-size arrays indexed by the covert text sequence type.

where $\mathbf{u} \in \mathcal{U}^N$, the joint type $p_{\mathbf{u}\mathbf{s}}$ is a function of type p_S , and η_N is an arbitrary positive sequence such that $\eta_N \rightarrow 0$ and $\eta_N \gg \frac{\log N}{N}$ as $N \rightarrow \infty$.

A random codebook $\mathcal{C}$ is the union of codebooks $\mathcal{C}(p_S)$ indexed by the covert text sequence type p_S . The codebook $\mathcal{C}(p_S)$ is obtained by drawing $2^{N(R+\rho(p_S))}$ random vectors independently and uniformly from the corresponding $T_{\mathbf{u}(p_S)}$, and arranging them in an array with $2^{N\rho(p_S)}$ rows and 2^{NR} columns indexed by fingerprint messages.

Given a covert text sequence $\mathbf{s}$ and a fingerprint message m , the encoder finds in $\mathcal{C}(p_S)$ the smallest l such that $\mathbf{u}(l, m) \in T_{\mathbf{u}|\mathbf{s}}$. Let $\mathbf{u} = \mathbf{u}(l, m)$. If no such l is available, generate $\mathbf{u}$ uniformly from the conditional type class $T_{\mathbf{u}|\mathbf{s}}$. Then generate $\mathbf{x}$ by uniformly drawing from the optimized conditional type class $T_{\mathbf{x}|\mathbf{u}\mathbf{s}}(\mathbf{u}, \mathbf{s})$. Notice that the conditional joint type $T_{\mathbf{x}\mathbf{u}|\mathbf{s}}$ is optimized and we later will give the cost function for the optimization; and the distortion between $\mathbf{x}$ and $\mathbf{s}$ satisfies the maximum distortion constraint (2.1). This procedure is repeated for all $|\mathcal{M}| = \lceil 2^{NR} \rceil$ users.

Given a forgery $\mathbf{y}$, the decoder seeks a covert text type p_S and a pair $\hat{\mathbf{u}}_1 \in \mathcal{C}(p_S)$, $\hat{\mathbf{u}}_2 \in \mathcal{C}(p_S)$ that maximizes the **penalized mutual information** criterion

$$\max_{p_S} \max_{\mathbf{u}_1 \in \mathcal{C}(p_S)} \max_{\mathbf{u}_2 \in \mathcal{C}(p_S)} [I(\mathbf{u}_1 \mathbf{u}_2; \mathbf{y}) - 2\rho(p_S) + I(\mathbf{u}_1; \mathbf{u}_2)]. \quad (4.8)$$

The decoder outputs the pair of column indices $(\hat{m}_1, \hat{m}_2)$ corresponding to $(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2)$. If there exist maximizers with different pairs of column indices, the decoder declares an error.

Without loss of generality, we assume that the two colluders have fingerprint indices $(m_1, m_2) = (1, 2)$.

Conditioned on a covert text $\mathbf{s}$, an encoding error happens under either of the following two error events:

$$\mathcal{E}_{e,1} = \{(\mathcal{C}, \mathbf{s}) : (\mathbf{u}(l, 1), \mathbf{s}) \notin T_{\mathbf{u}\mathbf{s}} \text{ for } 1 \leq l \leq 2^{N\rho(p_S)}\} \quad (4.9)$$

and

$$\mathcal{E}_{e,2} = \{(\mathcal{C}, \mathbf{s}) : (\mathbf{u}(l, 2), \mathbf{s}) \notin T_{\mathbf{u}\mathbf{s}} \text{ for } 1 \leq l \leq 2^{N\rho(p_S)}\}. \quad (4.10)$$

Therefore,

$$\begin{aligned}
Pr(\text{encoding error} | \mathbf{s}) &= Pr[\mathcal{E}_{e,1} \text{ or } \mathcal{E}_{e,2}] \leq 2 \cdot Pr[\mathcal{E}_{e,1}] \\
&= 2 \cdot (1 - Pr[\mathbf{U} \in T_{\mathbf{u}|\mathbf{s}} | \mathbf{U} \sim \mathbb{U}(T_{\mathbf{u}})])^{2^{N\rho(p_s)}} \\
&\doteq 2 \cdot (1 - 2^{-NI(\mathbf{u}|\mathbf{s})})^{2^{N\rho(p_s)}} \\
&\leq 2 \cdot \exp \left\{ -2^{N[\rho(p_s) - I(\mathbf{u}|\mathbf{s})]} \right\} = 2 \cdot \exp \left\{ -2^{N\eta_N} \right\},
\end{aligned}$$

which vanishes super-exponentially since $\eta_N \gg \frac{\log N}{N}$. The second and third equations follow since each $\mathbf{u}(l, 1)$ is uniformly distributed over $T_{\mathbf{u}}$ and the probability that such a vector belongs to $T_{\mathbf{u}|\mathbf{s}}$ is equal to $2^{-NI(\mathbf{u}|\mathbf{s})}$ on the exponential scale. The second inequality follows from $1 + a \leq e^a$.

Then we have

$$\begin{aligned}
P_{e,N} &= \min_{F_N, G_N} \max_{p_{Y|X_1 X_2}} P_e(F_N, G_N, p_{Y|X_1 X_2}, (m_1, m_2) = (1, 2)) \\
&= \min_{F_N, G_N} \max_{p_{Y|X_1 X_2}} \sum_{\mathbf{s}} p_{\mathbf{S}}^N(\mathbf{s}) \left[Pr(\text{encoding error} | \mathbf{s}) + \right. \\
&\quad \left. Pr(\text{decoding error} | \mathcal{E}_{e,1}^c \text{ and } \mathcal{E}_{e,2}^c) \right] \\
&\doteq \min_{F_N, G_N} \max_{p_{Y|X_1 X_2}} \sum_{\mathbf{s}} p_{\mathbf{S}}^N(\mathbf{s}) \cdot Pr(\text{decoding error} | \mathcal{E}_{e,1}^c \text{ and } \mathcal{E}_{e,2}^c) \\
&\doteq \min_{F_N, G_N} \max_{p_{Y|X_1 X_2}} \sum_{\mathbf{s}} p_{\mathbf{S}}^N(\mathbf{s}) \cdot \sum_{\mathbf{u}_1, \mathbf{x}_1, \mathbf{u}_2, \mathbf{x}_2} P_{\mathbf{X}_1 \mathbf{U}_1 | \mathbf{S}}(\mathbf{x}_1, \mathbf{u}_1 | \mathbf{s}) \cdot \\
&\quad P_{\mathbf{X}_2 \mathbf{U}_2 | \mathbf{S}}(\mathbf{x}_2, \mathbf{u}_2 | \mathbf{s}) \cdot \sum_{\mathbf{y}} p_{Y|X_1 X_2}^N(\mathbf{y} | \mathbf{x}_1 \mathbf{x}_2) \cdot \\
&\quad Pr(\text{decoding error} | \mathbf{s}, \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}, \mathcal{E}_{e,1}^c \text{ and } \mathcal{E}_{e,2}^c) \\
&\doteq \min_{F_N, G_N} \max_{p_{Y|X_1 X_2}} \sum_{T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}} Pr[T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}] \cdot \\
&\quad Pr(\text{decoding error} | T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}, \mathcal{E}_{e,1}^c \text{ and } \mathcal{E}_{e,2}^c) \\
&\leq \min_{F_N, G_N} \sum_{T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}} \max_{p_{Y|X_1 X_2}} Pr[T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}] \cdot \\
&\quad Pr(\text{decoding error} | T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}, \mathcal{E}_{e,1}^c \text{ and } \mathcal{E}_{e,2}^c), (4.11)
\end{aligned}$$

where $P_{\mathbf{X}_1 \mathbf{U}_1 | \mathbf{S}} = P_{\mathbf{X}_2 \mathbf{U}_2 | \mathbf{S}} = P_{\mathbf{X} \mathbf{U} | \mathbf{S}}$ is determined by the encoder and

$$\begin{aligned}
Pr[T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}] &= p_{\mathbf{S}}^N P_{\mathbf{X}_1 \mathbf{U}_1 | \mathbf{S}} P_{\mathbf{X}_2 \mathbf{U}_2 | \mathbf{S}} p_{Y|X_1 X_2}^N(T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}) \\
&\doteq 2^{-ND(p_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}} || p_{\mathbf{S}} p_{\mathbf{X}_1 \mathbf{U}_1 | \mathbf{S}} p_{\mathbf{X}_2 \mathbf{U}_2 | \mathbf{S}} p_{Y|X_1 X_2})}. (4.12)
\end{aligned}$$

Now, we turn to bound the above decoding error term. As we will show next, $Pr(\text{decoding error} | T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}, \mathcal{E}_{e,1}^c \text{ and } \mathcal{E}_{e,2}^c)$ is constant for each member $(\mathbf{s}, \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y})$ of the joint type $T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}$.

We distinguish five kinds of decoding error events. The sum of their probabilities is an upper bound of the decoding error probability. Let us use the shorthand

$$PMI(\mathbf{u}_1, \mathbf{u}_2) = I(\mathbf{u}_1 \mathbf{u}_2; \mathbf{y}) - 2\rho(p_s) + I(\mathbf{u}_1; \mathbf{u}_2),$$

with $\mathbf{u}_1 \in T_{\mathbf{u}(p_s)}$, and $\mathbf{u}_2 \in T_{\mathbf{u}(p_s)}$. Also, to simplify the notation, let the subscript of $\mathbf{u}$ denote the index of fingerprints.

The five kinds of error events are

$$\mathcal{E}_{d,1} = \left\{ (\tilde{p}_s, \tilde{\mathbf{u}}_{m'_1 > 2}, \tilde{\mathbf{u}}_{m'_2 > 2}) : PMI(\tilde{\mathbf{u}}_{m'_1}, \tilde{\mathbf{u}}_{m'_2}) \geq PMI(\mathbf{u}_1, \mathbf{u}_2), \right. \\
\left. \tilde{\mathbf{u}}_{m'_1} \in \mathcal{C}(\tilde{p}_s), \tilde{\mathbf{u}}_{m'_2} \in \mathcal{C}(\tilde{p}_s), m'_1 \neq m'_2 \right\},$$

$$\mathcal{E}_{d,2} = \left\{ (\tilde{p}_s, \tilde{\mathbf{u}}_{m'_1 > 2}, \tilde{\mathbf{u}}_2) : PMI(\tilde{\mathbf{u}}_{m'_1}, \tilde{\mathbf{u}}_2) \geq PMI(\mathbf{u}_1, \mathbf{u}_2), \right. \\
\left. \tilde{\mathbf{u}}_{m'_1} \in \mathcal{C}(\tilde{p}_s), \tilde{\mathbf{u}}_2 \in \mathcal{C}(\tilde{p}_s), \tilde{\mathbf{u}}_2 \neq \mathbf{u}_2 \right\},$$

$$\mathcal{E}_{d,3} = \left\{ (\tilde{p}_s, \tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_{m'_2 > 2}) : PMI(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_{m'_2}) \geq PMI(\mathbf{u}_1, \mathbf{u}_2), \right. \\
\left. \tilde{\mathbf{u}}_1 \in \mathcal{C}(\tilde{p}_s), \tilde{\mathbf{u}}_1 \neq \mathbf{u}_1, \tilde{\mathbf{u}}_{m'_2} \in \mathcal{C}(\tilde{p}_s) \right\},$$

$$\mathcal{E}_{d,4} = \left\{ (\tilde{\mathbf{u}}_{m'_1 > 2}, \mathbf{u}_2) : PMI(\tilde{\mathbf{u}}_{m'_1}, \mathbf{u}_2) \geq PMI(\mathbf{u}_1, \mathbf{u}_2), \right. \\
\left. \tilde{\mathbf{u}}_{m'_1} \in \mathcal{C}(p_s) \right\},$$

and

$$\mathcal{E}_{d,5} = \left\{ (\mathbf{u}_1, \tilde{\mathbf{u}}_{m'_2 > 2}) : PMI(\mathbf{u}_1, \tilde{\mathbf{u}}_{m'_2}) \geq PMI(\mathbf{u}_1, \mathbf{u}_2), \right. \\
\left. \tilde{\mathbf{u}}_{m'_2} \in \mathcal{C}(p_s) \right\}.$$

The bounds are given by

$$Pr[\mathcal{E}_{d,1}] \leq 2^{-N \left| I(\mathbf{u}_1 \mathbf{u}_2; \mathbf{y}) - I(\mathbf{u}_1 \mathbf{u}_2; \mathbf{s}) + I(\mathbf{u}_1; \mathbf{u}_2 | \mathbf{s}) - 2R - 2\eta_N \right|_+}, (4.13)$$

$$Pr[\mathcal{E}_{d,2}] \leq 2^{-N \left| I(\mathbf{u}_1 \mathbf{u}_2; \mathbf{y}) - I(\mathbf{u}_1 \mathbf{u}_2; \mathbf{s}) + I(\mathbf{u}_1; \mathbf{u}_2 | \mathbf{s}) - R - 2\eta_N \right|_+}, (4.14)$$

$$Pr[\mathcal{E}_{d,3}] \leq 2^{-N \left| I(\mathbf{u}_1 \mathbf{u}_2; \mathbf{y}) - I(\mathbf{u}_1 \mathbf{u}_2; \mathbf{s}) + I(\mathbf{u}_1; \mathbf{u}_2 | \mathbf{s}) - R - 2\eta_N \right|_+}, (4.15)$$

$$Pr[\mathcal{E}_{d,4}] \leq 2^{-N \left| I(\mathbf{u}_1; \mathbf{y} | \mathbf{u}_2) - I(\mathbf{u}_1; \mathbf{s} | \mathbf{u}_2) + I(\mathbf{u}_1; \mathbf{u}_2 | \mathbf{s}) - R - \eta_N \right|_+}, (4.16)$$

and

$$Pr[\mathcal{E}_{d,5}] \leq 2^{-N \left| I(\mathbf{u}_2; \mathbf{y} | \mathbf{u}_1) - I(\mathbf{u}_2; \mathbf{s} | \mathbf{u}_1) + I(\mathbf{u}_1; \mathbf{u}_2 | \mathbf{s}) - R - \eta_N \right|_+}. (4.17)$$

Derivations are omitted due to space limitations. Combining (4.11)–(4.17), we have

$$\begin{aligned}
P_{e,N} &\leq \min_{F_N, G_N} \sum_{T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}} \max_{p_{Y|X_1 X_2}} Pr[T_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}}] \cdot \max\{Pr[\mathcal{E}_{d,i}]_{i=1}^5\} \\
&\doteq \max_{p_s} \max_{p_{\mathbf{x}_2 \mathbf{u}_2 | \mathbf{x}_1 \mathbf{u}_1 \mathbf{s}} p_{\mathbf{x}_1 \mathbf{u}_1 | \mathbf{s}} = p_{\mathbf{x}_2 \mathbf{u}_2 | \mathbf{s}} p_{Y|X_1 X_2}} \min_{p_{Y|X_1 X_2}} \max_{p_{\mathbf{x}_2 \mathbf{u}_2 | \mathbf{x}_1 \mathbf{u}_1 \mathbf{s}} p_{Y|X_1 X_2}} \max_{p_{\mathbf{x}_1 \mathbf{u}_1 | \mathbf{s}} p_{Y|X_1 X_2}} \\
&\quad 2^{-ND(p_{\mathbf{su}_1 \mathbf{u}_2 \mathbf{x}_1 \mathbf{x}_2 \mathbf{y}} || p_{\mathbf{S}} p_{\mathbf{x}_1 \mathbf{u}_1 | \mathbf{s}} p_{\mathbf{x}_2 \mathbf{u}_2 | \mathbf{s}} p_{Y|X_1 X_2})} \cdot \max\{Pr[\mathcal{E}_{d,i}]_{i=1}^5\}, (4.18)
\end{aligned}$$

where the second equality holds because the encoding procedure optimizes the joint type $p_{\mathbf{x}_1 \mathbf{u}_1 | \mathbf{s}} = p_{\mathbf{x}_2 \mathbf{u}_2 | \mathbf{s}}$. The exponent from (4.18) converges to the limit $E_r^{DMMC}(R)$ in (3.7) as $N \rightarrow \infty$, due to continuity of the functionals D and I .

V. SKETCH OF PROOF OF THEOREM 3.2

The proof of the converse theorem is an extension of Gel'fand and Pinsker's proof in [4].

Let $P_e = \Pr\{(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)\}$ be the decoding error. The joint p.m.f of $(M_1, M_2, \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y})$ is given by

$$p_{M_1 M_2 \mathbf{S} \mathbf{X}_1 \mathbf{X}_2 \mathbf{Y}} = p_{M_1} p_{M_2} \mathbf{1}_{\{\mathbf{X}_1 = f_N(\mathbf{S}, M_1)\}} \mathbf{1}_{\{\mathbf{X}_2 = f_N(\mathbf{S}, M_2)\}} \cdot \prod_{i=1}^N p_{Y|X_1 X_2}(y_i | x_{1,i} x_{2,i}). \quad (5.19)$$

For any rate- R encoder f_N and DMCC $p_{Y|X_1 X_2}$, we consider three inequalities.

Inequality I: Due to Fano's inequality, we have

$$2NR = H(M_1, M_2) = H(M_1, M_2 | \mathbf{Y}) + I(M_1, M_2; \mathbf{Y}) \leq 1 + P_e \cdot 2NR + I(M_1, M_2; \mathbf{Y}).$$

Hence,

$$P_e \geq 1 - \frac{1}{2NR} - \frac{I(M_1, M_2; \mathbf{Y})}{2NR}. \quad (5.20)$$

The error probability P_e is not bounded away from 0 only if

$$2NR \leq I(M_1, M_2; \mathbf{Y}). \quad (5.21)$$

Define random variables

$$V_{1,i} = (M_1, S_{i+1}^N), \quad V_{2,i} = (M_2, S_{i+1}^N), \\ W_{1,i} = (V_{1,i}, Y^{i-1}), \quad W_{2,i} = (V_{2,i}, Y^{i-1}),$$

where $S_{i+1}^N = (S_{i+1}, \dots, S_N)$ and $Y^{i-1} = (Y_1, \dots, Y_{i-1})$. Then we have

$$V_{1,i-1} = (V_{1,i}, S_{1,i}), \quad V_{1,1} = W_{1,1}, \quad V_{1,N} = M_1, \\ V_{2,i-1} = (V_{2,i}, S_{2,i}), \quad V_{2,1} = W_{2,1}, \quad V_{2,N} = M_2.$$

Since $(M_1, M_2, S_{i+1}^N, Y^{i-1}) \rightarrow X_{1,i}, X_{2,i} \rightarrow Y_i$ forms a Markov chain for any $1 \leq i \leq N$ according to (5.19), so does $W_{1,i}, W_{2,i} \rightarrow X_{1,i}, X_{2,i} \rightarrow Y_i$.

Due to limited space, we omit the derivation of the following inequality

$$I(V_{1,i}, V_{2,i}; Y^i) - I(V_{1,i}, V_{2,i}; S^i) \leq I(W_{1,i}, W_{2,i}; Y_i) - I(W_{1,i}, W_{2,i}; S_i) \\ + I(V_{1,i-1}, V_{2,i-1}; Y^{i-1}) - I(V_{1,i-1}, V_{2,i-1}; S^{i-1}) \quad (5.22)$$

Summing up both sides of the inequality (5.22) from $i = 2$ to $i = N$, we have

$$2NR \leq I(M_1, M_2; \mathbf{Y}) = I(M_1, M_2; \mathbf{Y}) - \underbrace{I(M_1, M_2; \mathbf{S})}_{=0} \\ \leq \sum_{i=1}^N [I(W_{1,i}, W_{2,i}; Y_i) - I(W_{1,i}, W_{2,i}; S_i)] \\ = N [I(W_1, W_2; Y | T) - I(W_1, W_2; S | T)] \\ \leq N [I(W_1, W_2, T; Y) - I(W_1, W_2, T; S)] \\ = N [I(U_1, U_2; Y) - I(U_1, U_2; S)], \quad (5.23)$$

where the time-sharing random variable T is uniformly distributed over $\{1, \dots, N\}$ and independent of any other random variable; the random variable $(S, W_1, W_2, X_1, X_2, Y)$ is

defined as $(S_T, W_{1,T}, W_{2,T}, X_{1,T}, X_{2,T}, Y_T)$; $U_1 = (W_1, T)$ and $U_2 = (W_2, T)$.

Inequality II: For any rate- R encoder f_N and DMCC $p_{Y|X_1 X_2}$, we also have

$$NR = H(M_1) = H(M_1 | \mathbf{Y}) + I(M_1; \mathbf{Y}) \\ \leq H(M_1 M_2 | \mathbf{Y}) + I(M_1; \mathbf{Y}) \\ \leq 1 + P_e \cdot 2NR + I(M_1; \mathbf{Y}),$$

which leads to

$$P_e \geq \frac{1}{2} - \frac{1}{2NR} - \frac{I(M_1; \mathbf{Y})}{2NR}. \quad (5.24)$$

The error probability P_e is not bounded away from 0 only if

$$NR \leq I(M_1; \mathbf{Y}). \quad (5.25)$$

Still using the telescoping technique—although there are some tricks in the omitted derivation—we obtain

$$NR \leq N [I(U_1; Y | U_2) - I(U_1; S | U_2)]. \quad (5.26)$$

Inequality III: Similarly to Inequality II, we also have $NR \leq I(M_2; \mathbf{Y})$. Following the same steps as in Inequality II and switching the subscripts 1 and 2, we obtain

$$NR \leq N [I(U_2; Y | U_1) - I(U_2; S | U_1)]. \quad (5.27)$$

Combining (5.23), (5.26), (5.27) and the optimization taken by the encoder and colluders, we conclude that any achievable rate R should satisfy

$$R \leq \sup_{\mathcal{U}} \max_{p_{XU|S} \in \mathcal{P}_{XU|S}(D_1)} \min_{p_{Y|X_1 X_2} \in \mathcal{P}_{Y|X_1 X_2}(D_2)} \\ \min \left\{ I(U_1; Y | U_2) - I(U_1; S | U_2), I(U_2; Y | U_1) - I(U_2; S | U_1), \right. \\ \left. \frac{1}{2} [I(U_1, U_2; Y) - I(U_1, U_2; S)] \right\},$$

where $U_1 \in \mathcal{U}$, $U_2 \in \mathcal{U}$, and $p_{X_1 U_1 | S} = p_{X_2 U_2 | S} = p_{XU|S}$. This proves the converse part of Theorem 3.2.

ACKNOWLEDGMENT

This work was supported by NSF under grant CCR 03-25924.

REFERENCES

- [1] A. Somekh-Baruch and N. Merhav, "On the capacity game of private fingerprinting systems under collusion attacks," *IEEE Trans. Inform. Theory*, vol. 51, no. 3, pp. 884-899, Mar. 2005.
- [2] A. Somekh-Baruch and N. Merhav, "Achievable error exponents for the private fingerprinting game," submitted to *IEEE Trans. Inform. Theory*, Feb. 2005.
- [3] P. Moulin and Y. Wang, "Error exponents for channel coding with side information," submitted to *IEEE Trans. Inform. Theory*, Sept. 2004.
- [4] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Prob. of Control and Inform. Theory*, vol. 9, no. 1, pp. 19-31, 1980.
- [5] A. Somekh-Baruch and N. Merhav, "On the random coding error exponents of the single-user and the multiple-access Gel'fand-Pinsker channels," *Proc. IEEE Int. Symp. Info. Theory*, p. 448, Chicago, IL, Jul. 2004.