UNIVERSAL FINGERPRINTING: CAPACITY AND RANDOM-CODING EXPONENTS

Pierre Moulin

Beckman Inst., Coord. Sci. Lab and ECE Department
University of Illinois at Urbana-Champaign, USA

ABSTRACT

Bounds on fingerprinting capacity have been derived in recent literature. In this paper we present an exact capacity formula and a universal fingerprinting scheme. Our problem setup unifies the signal-distortion and Boneh-Shaw formulations of fingerprinting. The proposed scheme has four useful properties: (1) the receiver does not need to know the coalition size and collusion channel; (2) a tunable parameter $\Delta$ trades off false-positive and false-negative error exponents; (3) the receiver provides a reliability metric for its decision; and (4) the decoder is capacity-achieving when the false-positive exponent $\Delta$ tends to zero.

The new random coding scheme uses a “time-sharing” randomized sequence and produces conditionally constant-composition fingerprints. The decoder is a minimum penalized equivocation decoder, where the penalty term is proportional to coalition size.

1. INTRODUCTION

Content fingerprinting (a.k.a. digital fingerprinting, or traitor tracing) is essentially a multiuser version of watermarking. A covertext — such as image, video, audio, text, or software — is to be distributed to many users. Prior to distribution, each user is assigned a fingerprint that is embedded into the covertext. In a collusion attack, a coalition of users combine their marked copies, creating a pirated copy that contains only weak traces of their fingerprints. The pirated copy is subject to a fidelity requirement relative to the coalition’s copies. The fidelity requirement may take the form of a distortion constraint, which is a natural model for media fingerprinting applications [1–4]; or it may take the form of Boneh and Shaw’s marking assumption, which is a popular model for software fingerprinting [5–7]. To trace the forgery back to the coalition members, the receiver needs to decode the colluders’ fingerprints from the pirated copy.

The fingerprinting problem presents two key challenges.

1. The number of colluders may be large (dozens of users), which makes it easier for the colluders to mount a strong attack. The difficulty of the decoding problem is compounded by the fact that the number of colluders and the attack channel are unknown to the decoder.

2. There are two fundamental types of error probabilities, namely false positives, by which innocent users are wrongly accused, and false negatives, by which one or more colluders escape detection. The first type of error is often considered to be more costly.

This paper derives information-theoretic performance limits for a model that captures the above features of fingerprinting systems. Complete derivations are given in [8]. Prior art is reviewed below.

The basic performance metric is capacity, which is defined with respect to a class of attack channels. A closely related problem (multiuser data hiding) was analyzed by Moulin and O’Sullivan [1], and capacity expressions were obtained assuming expected-distortion constraints for the fingerprint distributor and the coalition, and noncooperating, single-user decoders. Although this problem presents clear mathematical similarities with the standard fingerprinting setup, it is quite different from the setup adopted in other fingerprinting papers and in this manuscript. This more standard setup was studied by Somekh-Baruch and Merhav [2, 3], and connections with the problem of coding for the multiple-access channel (MAC) were explored. No constraints were imposed on the probability of false positives, which could approach 1. Lower bounds on capacity were obtained using a restrictive encoding strategy (random constant-composition codes without time sharing) and almost-sure distortion constraints between the pirated copy and one [2] or all [3] of the coalition’s copies. The decoders in [1–3] assume a fixed number of colluders, and admit no obvious extension to cope with the two challenges listed above.

Connections between the MAC problem and fingerprinting under the Boneh-Shaw assumption have also been studied recently by Anthapadmanabhan et al. [7]. The covertext is degenerate, and side information does not appear in the information-theoretic formulation of this problem. Bounds on capacity were also presented in [7].

In order to cope with unknown collusion channels and unknown number of colluders, a special kind of universal decoder should be designed, with universality holds not only with respect to some set of channels, but also with respect to an unknown number of inputs. One would expect such a universal decoder to feature a tunable parameter that trades off the two fundamental types of error probability. When the number of colluders is unknown, two extreme instances of this tradeoff are to accuse all users or accuse none of them.
Recently Tardos [6] proposed a random fingerprinting scheme using what amounts to an auxiliary random sequence for encoding fingerprints. While this scheme is presented at an algorithmic level (and no optimization is involved in its construction), in our game-theoretic setting the auxiliary random variable appears fundamentally as part of a randomized strategy for a game whose payoff function is nonconcave with respect to the maximizing variable (the fingerprint distribution). It should also be mentioned that the decoding procedure in [6], which correlates the pirated copy with individual fingerprints and outputs the highest-scoring one, is simple but suboptimal (not capacity-achieving).

Another issue that is resolved in our game-theoretic setting is the optimality of coalition strategies that are invariant to permutations of the colluders. While one may heuristically expect that such strategies are optimal, a proof of this property is established in this paper. The approach used in previous papers was to assume that coalitions employ such strategies.

1.1. Notation

We use uppercase letters for random variables, lowercase letters for their individual values, calligraphic letters for finite alphabets, and boldface letters for sequences. The probability mass function (PMF) of a random variable \( X \in \mathcal{X} \) is denoted by \( p_X(x) \). The entropy of a random variable \( X \) is denoted by \( H(X) \), and the mutual information between two random variables \( X \) and \( Y \) is denoted by \( I(X;Y) \). The Kullback-Leibler divergence between two PMFs \( p \) and \( q \) is denoted by \( D(p||q) \); the conditional Kullback-Leibler divergence of \( p_{Y|X} \) and \( q_{Y|X} \) given \( p_X \) is denoted by

\[
D(p_{Y|X}||q_{Y|X}|p_X) = D(p_{Y|X}|p_X) \cdot \frac{q_{Y|X}}{p_{Y|X}}.
\]

Mathematical expectation is denoted by the symbol \( \mathbb{E} \). All logarithms are in base 2 unless specified otherwise.

Let \( p_X \) denote the empirical PMF (type) induced by a sequence \( x \in \mathcal{X}^N \). The type class \( T_x \) associated with \( p_x \) is the set of all sequences of type \( p_x \). Likewise, we define the joint type \( p_{xy} \) of a pair of sequences \((x,y) \in \mathcal{X}^N \times \mathcal{Y}^N \) and \( T_{xy} \) the type class associated with \( p_{xy} \). Also, we define the conditional type \( p_{y|x} \) of a pair of sequences \((x,y) \) as \( p_{y|x}(x,y) = p_{xy}(x,y) / p_x(x) \) for all \( x \) such that \( p_x(x) > 0 \). The conditional type class \( T_{y|x} \) given \( x \) is the set of all sequences \( y \) such that \( (x,y) \in T_{xy} \). We denote by \( H(x) \) the empirical entropy of the PMF \( p_X \) and by \( I(x;y) \) the empirical mutual information for the joint PMF \( p_{xy} \).

We use the calligraphic fonts \( \mathcal{P}_X \) and \( \mathcal{P}^N_X \) to represent the set of all PMFs and all empirical PMFs, respectively, on the alphabet \( \mathcal{X} \). Likewise, \( \mathcal{P}_{Y|X} \) and \( \mathcal{P}_{Y|X}^N \) denote the set of all conditional PMFs and all empirical conditional PMFs on the alphabet \( \mathcal{Y} \).

1.2. Mutual Information of \( k \) Random Variables

The mutual information of \( k \) random variables \( X_1, \ldots, X_k \) is defined as the sum of their individual entropies minus their joint entropy [11, p. 57] or equivalently, the divergence between their joint distribution and the product of their marginals:

\[
\hat{I}(X_1;\cdots;X_k) = \sum_{i=1}^k H(X_i) - H(X_1,\cdots,X_k) = D(p_{X_1,\cdots,X_k}||\prod_i p_{X_i}).
\]

Similarly, we define the empirical mutual information \( \hat{I}(x_1;\cdots;x_k) \) between \( k \) sequences \( x_1,\ldots,x_k \) as the mutual information with respect to the joint type of \( x_1,\ldots,x_k \). We have \( \hat{I}(x_1;\cdots;x_k) = \sum_{i=1}^k H(x_i) - H(x_1,\cdots,x_k) \).

This leads to the following alternative interpretation of the minimum-equivocation decoder of Liu and Hughes [9]. If \( x_1,\ldots,x_k \) are codewords from a constant-composition code \( \mathcal{C} \), then \( H(x_i) \) is the same for all \( i \), and the minimum-equivocation decoder is equivalent to a maximum-mutual-information decoder (using \( I \) instead of \( I^o \)):

\[
\min_{x_1,\ldots,x_k} H(x_1,\ldots,x_k | y) \iff \max_{x_1,\ldots,x_k} \hat{I}(x_1;\cdots;x_k;y).
\]

2. STATEMENT OF THE PROBLEM

2.1. Overview

Our model for digital fingerprinting is diagrammed in Fig. 1. Let \( \mathcal{S} \), \( \mathcal{X} \), and \( \mathcal{Y} \) be three finite alphabets. The coverttext sequence \( \mathbf{S} = (S_1,\ldots,S_N) \in \mathcal{S}^N \) consists of \( N \) independent and identically distributed (i.i.d.) samples drawn from a PMF \( p_S(s), s \in \mathcal{S} \). There are \( 2^{NR} \) users, each of which receives a fingerprinted copy:

\[
X_m = f_N(S,m), \quad 1 \leq m \leq 2^{NR}.
\]

Our model for digital fingerprinting is diagrammed in Fig. 1. Let \( \mathcal{S} \), \( \mathcal{X} \), and \( \mathcal{Y} \) be three finite alphabets. The coverttext sequence \( \mathbf{S} = (S_1,\ldots,S_N) \in \mathcal{S}^N \) consists of \( N \) independent and identically distributed (i.i.d.) samples drawn from a PMF \( p_S(s), s \in \mathcal{S} \). There are \( 2^{NR} \) users, each of which receives a fingerprinted copy:

\[
X_m = f_N(S,m), \quad 1 \leq m \leq 2^{NR}.
\]

Let \( d : \mathcal{S} \times \mathcal{X} \to \mathbb{R}^+ \) be the distortion measure and \( d^N(s,x) = \frac{1}{N} \sum_{i=1}^N d(s_i,x_i) \) be the extension of this measure to length-\( N \) sequences. The code \( f_N \) is subject to the distortion constraint

\[
d^N(s,x_m) \leq D_1 \quad 1 \leq m \leq 2^{NR}.
\]

Fig. 1. Model for fingerprinting game. In the Boneh-Shaw setup, the host sequence \( \mathbf{S} \) is degenerate.
Let $\mathcal{K} \triangleq \{m_1, m_2, \ldots, m_K\}$ be a coalition of $K$ users; no constraints are imposed on the formation of coalitions. The coalition uses their copies $X_K \triangleq \{X_m, m \in \mathcal{K}\}$ to produce a pirated copy $Y \in \mathcal{Y}^N$. Without loss of generality, we assume that $Y$ is generated stochastically according to a conditional pmf $p_Y|X_K$. Fidelity constraints are imposed on $p_Y|X_K$ to ensure that $Y$ is “close” to the fingerprinted copies $X_m$, $m \in \mathcal{K}$. These constraints can take the form of distortion constraints (analogously to (2.2)), or alternatively, a constraint that will be referred to as the Boneh-Shaw constraint. The formulation of these constraints is detailed below and results in the definition of a feasible class $\mathcal{W}_K$ of attacks.

The decoder knows neither $K$ nor $p_Y|X_K$ selected by the $K$ colluders and has access to the pirated copy $Y$ and the host $S$. It produces an estimate $\hat{K} = g_N(Y, S)$ of the coalition. A possible decision is the empty set, $\hat{K} = \emptyset$, reflecting the possibility that the signal submitted to the decoder is unrelated to the fingerprints. If this possibility was not allowed, an innocent user would be accused. Another good reason to allow $\hat{K} = \emptyset$ is that reliable detection is impossible when there are too many colluders, and the probability of false positives should remain small for any coalition size.

The encoder/decoder pair $(f_N, g_N)$ is randomized, i.e., the choice of $(f_N, g_N)$ is a function of a random variable known to the encoder and decoder but not to the adversary. This random variable is independent of all other random variables and plays the role of a secret key. The randomized code is denoted by $(F_N, G_N)$.

Our randomized codes are obtained using permutations of the letters $\{1, 2, \ldots, N\}$ (to eliminate apparent correlations across positions in the marked sequences), permutations of the fingerprint assignments (to ensure that error probabilities are the same for all coalitions), as well a special randomized sequence. An example of the latter was given by Tardos [6]. For binary alphabets $\mathcal{S}, \mathcal{X}$, and $\mathcal{Y}$, iid random variables $w_i \in (0, 1)$, $1 \leq i \leq N$, were generated according to some fixed pmf, and next the fingerprint letters $X_i(m)$ were generated as independent Bernoulli$(w_i)$ random variables. Here $\{w_i, 1 \leq i \leq N\}$ is the secret key shared by encoder and decoder.

For an embedding distortion $D_1$ and a coalition of size $K$ using collusion channel $p_{Y|X_K}$ in class $\mathcal{W}_K$, there corresponds a capacity $C(D_1, \mathcal{W}_K)$ which is the supremum of all achievable $R$, under a prescribed error criterion. By fixing $R$, $D_1$, and a minimum value $\Delta$ for the false-positive error exponent, one aims at guaranteeing resistance to any coalition whose size does not exceed some number $K_{\max}$ that depends on $R$, $D_1$, $\Delta$, and $(\mathcal{W}_K, k \geq 1)$. In the remainder of this section, we detail the attack models and define the relevant error probabilities, capacities, and error exponents.

2.2. Collusion Channels

The fidelity constraint on the coalition is of the form $p_Y|X_K \in \mathcal{W}_K$, i.e., the empirical pmf of the pirated copy given the marked copies is restricted in a sense that depends on the application. This model can be used to impose hard distortion constraints on the coalition or to enforce the Boneh-Shaw marking assumption when $X = Y$. Under this assumption, the colluders are not allowed to modify their samples at any location where these samples agree. The appropriate choice for $\mathcal{W}_K$ in this case is the set of all $p_Y|X_K$ that satisfies

$$x_1 = \cdots = x_K \Rightarrow y = x_1. \quad (2.3)$$

Hence the only constraint on $y$ is that $y_i = x_{m_1,i}$ for any position $1 \leq i \leq N$ such that $x_{m_1,i} = \cdots = x_{m_K,i}$.

We denote by $\mathcal{W}_K^{\text{fair}}$ the subset of $\mathcal{W}_K$ that consists of fair collusion channels, i.e., channels that are invariant to permutations of the colluders.

2.3. Error Probabilities

There are several error probabilities of interest: the probability of false positives (one or more innocent users are accused), $P_{FP}(K) = Pr[\hat{K} \subseteq \mathcal{K}]$, the probability of missed detection for a specific coalition member $m \in \mathcal{K}$, $P_{e,m}(K) = Pr[m \notin \mathcal{K}]$, the probability of failing to catch a single co-founder, $P_{e^{\text{false}}}(K) = Pr[\forall m \in \mathcal{K} : m \notin \hat{K}]$, and the probability of failing to catch the full coalition: $P_{e^{\text{full}}}(K) = Pr[K \notin \hat{K}]$. These probabilities are obtained by averaging over all sources of randomness, i.e., the randomized code $(F_N, G_N)$ and the collusion channel $p_{Y|X_K}$. Clearly $P_{e^{\text{false}}}(K) \leq P_{e^{\text{full}}}(K)$.

Let $P_e(F_N, G_N, \mathcal{W}_K) = \max_{p_{Y|X_K} \in \mathcal{W}_K} P_e(F_N, G_N, p_{Y|X_K})$. The tradeoff between false positives and false negatives can be put in the context of statistical detection theory (the Neyman-Pearson problem) as well as list decoding [10]. For the latter problem, the false-negative error exponent is larger if the list size is large, and approaches the sphere packing exponent if the list size grows subexponentially with $N$. (There is no notion of false positives for classical list decoding because errors are defined by the event that the transmitted message is missing from the decoder’s output list). Likewise, the error exponents obtained in [2,3] are obtained for false negatives only, using a fixed list size.

2.4. Capacity

**Definition 2.1** A rate $R$ is achievable for embedding distortion $D_1$, collusion class $\mathcal{W}_K$, and detect-one criterion if there exists a sequence of $(N, \lfloor 2^{NR} \rfloor)$ randomized codes $(F_N, G_N)$ with maximum embedding distortion $D_1$, such that both $P_{e^{\text{false}}}(F_N, G_N, \mathcal{W}_K)$ and $P_{d_{FP}}(F_N, G_N, \mathcal{W}_K) \to 0$ as $N \to \infty$.

An analogous definition is given under the detect-all criterion, using $P_{e^{\text{full}}}$ in place of $P_{e^{\text{false}}}$. The encoder does not know $K$.

**Definition 2.2** Fingerprinting capacities $C_{e^{\text{false}}}(D_1, \mathcal{W}_K)$ and $C_{e^{\text{full}}}(D_1, \mathcal{W}_K)$ are the suprema of all achievable rates with respect to the detect-one and detect-all criteria, respectively.

We have $C_{e^{\text{false}}}(D_1, \mathcal{W}_K) \leq C_{e^{\text{false}}}(D_1, \mathcal{W}_K)$ because an error event for the detect-one problem is also an error event for the detect-all problem.
2.5. Random-Coding Exponents

For a sequence \((F_N, G_N)\) of random codes the error exponents are given by \(E(R, D_1, W_K) = \lim_{N \to \infty} \frac{1}{N} \log P_e(F_N, G_N, W_K)\) where \(E\) represents the random coding exponent \(E_{FP}, E_{one}\), or \(E_{all}\), and \(P_e\) the corresponding error probability. Again \(E_{all}(R, D_1, W_K) \leq E_{one}(R, D_1, W_K)\) because an error event for the detect-one problem is also an error event for the detect-all problem.

3. FINGERPRINTING CAPACITY

We now present a computable expression for fingerprinting capacity. Consider an auxiliary random variable \(W\) defined over an alphabet \(W = \{1, 2, \ldots, L\}\). Define the set

\[
\mathcal{P}_{X,W|S}(p_S, L, D_1) = \left\{ p_{X_i|S} = pw \prod_{k \in K} p_{X_k|SW} : p_{X_1|S} = \cdots = p_{X_K|S}, Ed(S, X_1) \leq D_1 \right\} \tag{3.1}
\]

and the functions

\[
C_{one}^{\text{un}}(D_1, W_K) = \max_{p_{X,K,W|S} \in \mathcal{P}_{X,W|S}(p_S, L, D_1)} \min_{p_{Y|X,K} \in \mathcal{W}_K} \frac{1}{K} I(X_K; Y|SW) \tag{3.2}
\]

\[
C_{all}^{\text{un}}(D_1, W_K) = \max_{p_{X,K,W|S} \in \mathcal{P}_{X,W|S}(p_S, L, D_1)} \min_{A \subseteq K} \frac{1}{|A|} I(X_A; Y|S, X_{K\setminus A}, W). \tag{3.3}
\]

The quantities \(C_{one}^{\text{un}}(D_1, W_K)\) and \(C_{all}^{\text{un}}(D_1, W_K)\) are nondecreasing functions of \(L\) and converge to finite limits:

\[
C_{one}^{\text{un}}(D_1, W_K) = \lim_{L \to \infty} C_{one}^{\text{un}}(D_1, W_K), \tag{3.4}
\]

\[
C_{all}^{\text{un}}(D_1, W_K) = \lim_{L \to \infty} C_{all}^{\text{un}}(D_1, W_K). \tag{3.5}
\]

Moreover, the gap to the limit may be bounded by a polynomial function of \(L\), hence this limit may be computed within any desired accuracy.

**Theorem 3.1** Fingerprinting capacity is given by \(C_{all}^{\text{un}}(D_1, W_K)\) under the “detect-all” criterion. If the colluders select a fair collusion channel, as is their collective interest, the minimization is restricted to \(\mathcal{W}_K^{fair}\) in (3.3), and then \(C_{all}^{\text{un}}(D_1, W_K) = C_{one}^{\text{un}}(D_1, W_K)\).

The converse is proved in [8]. A scheme that achieves private fingerprinting capacity is presented in Sec. 4.

Using a strong converse, we have also shown that capacity under the “detect-one” criterion for a fairly broad class of codes is given by \(C_{one}^{\text{un}}(D_1, W_K)\). The lower bounds on fingerprinting capacity derived in [2, 3] are of the form (3.2) with \(L = 1\), i.e., the random variable \(W\) is degenerate. Since the payoff function \(I(X_K; Y|S)\) is generally nonconcave with respect to the variable \(p_{X|S}\) of (3.1), a randomized strategy in which \(p_{X|S}\) is randomized will generally outperform a deterministic strategy in which \(p_{X|S}\) is fixed. The auxiliary random variable \(W\) plays the role of selector of \(p_{X|S}\) in this game.

Apparently the benefits of this randomization can be dramatic for large \(K\). For the Boneh-Shaw problem with \(L = 1\), we have derived \(C_{one}^{\text{un}}(D_1, W_K) = K^{-1}q^{-2(K-1)}\); however Tardos’ scheme achieves a rate \(O(K^{-2})\) which is therefore a lower bound on \(C_{one}^{\text{un}}(D_1, W_K)\).

4. RANDOM CODING SCHEME

Our random coding scheme is parameterized by three parameters \(K_{nom}, L\) and \(\Delta\). We present the scheme here and give its error exponents in the next section. It is also shown that fingerprinting capacity is achieved in the limit as \(\Delta \to 0\).

The encoder assumes a nominal value \(K_{nom}\) for coalition size. An arbitrarily large \(L\) is selected, defining an alphabet \(W = \{1, 2, \ldots, L\}\). A sequence \(w \in W^N\) is drawn independently of \(s\) and uniformly from a type class \(T_w\) to be optimized. This sequence is shared with the receiver. A rate-\(R\) conditionally constant-composition code \(\mathcal{C}(s, w) = \{x(m)\}, 1 \leq m \leq 2^{NR}\) is generated for each \(s, w\) by drawing \(2^{NR}\) sequences independently from the uniform distribution over a conditional type class \(\mathcal{T}_w\) optimized. User \(m\) is assigned codeword \(x(m)\). The parameter \(K_{nom}\) is used to optimize the joint types \(T_w\) and \(T_{w|x}\); see (5.5).

The decision rule is parameterized by \(\Delta \geq 0\) that controls the fundamental tradeoff between probabilities of false positives and false negatives.

The restriction of \(X_A\) to a subset \(A\) of \(\mathcal{M}\) will be denoted by \(X_A = \{x_m, m \in A\}\). For disjoint sets \(A = \{m_1, \ldots, m_{|A|}\}\) and \(B\), we denote by \(I(X_A; Y X_B|S, W) \triangleq \sum_{m_{|A|}} I(x_{m_1}; \ldots; x_{m_{|A|-1}}; Y X_B|S, W)\) the mutual information of \(|A|+1\) random variables \(X_{m_1}, \ldots, X_{m_{|A|-1}}(Y, X_B)\), conditioned on \((S, W)\). With this notation, we define the function

\[
\max_{x_m \in \mathcal{C}(s, w), m \in K} \left( I(x_K; Y|SW) - k(R + \Delta) \right)
\]

where \(k = |K|\), and \(I(x_K; Y|SW) = kH(x|SW) - H(x_K|YSW)\) is the empirical mutual information of the \(k+1\) sequences \(x_1, \ldots, x_k, y\), conditioned on \((s, w)\), as defined in Sec. 1.2. Our joint fingerprint decoder is a Maximum Penalized Mutual Information (MPMI) decoder:

\[
\max_{k \geq 0} \text{MPMI}(k).
\]

The decoder seeks the coalition size \(k\) and the codewords \(x_m, m \in K, \) in \(\mathcal{C}(s, w)\) that achieve the MPMI criterion...
above. The indices of these codewords form the decoded coalition \( \hat{K} \). If the minimizing \( k \) in (4.2) is zero, the receiver outputs \( K = \emptyset \). The MPMI criterion (4.2) can equivalently be formulated as a Minimum Penalized Equiprobation decoder. The value of the criterion can serve as a confidence measure for the level of guilt of the coalition.

5. ERROR EXPONENTS

Recall \( \mathcal{P}_{X_K|W|S} \) defined in (3.1). Define now the following set of conditional pmf’s for \( X_K \) given \( W \) whose conditional pmf \( p_{X_K|SW} \) is the same for each \( X_m, m \in K \):

\[
\mathcal{M}(p_{X|SW}) = \{ p_{X_K|SW} : p_{X_K|SW} = p_{X|SW}, m \in K \}.
\]

Define for each \( A \subseteq K \) the set of conditional pmf’s

\[
\mathcal{P}_{Y|X_K|SW}(p_{W}, \tilde{p}_{S|W}, p_{X|SW}, \mathcal{W}_K, R, L, A) \triangleq \left\{ \tilde{p}_{X_K|SW} : \tilde{p}_{X_K|SW} \in \mathcal{M}(p_{X|SW}), \tilde{p}_{Y|X_K} \in \mathcal{W}_K, \right. \left. \frac{1}{|A|} \int_{p_{W}} \tilde{p}_{S|W} \tilde{p}_{Y|X_K|SW}(X_A; Y|X_K, A | | SW) \leq R \right\}
\]

and the pseudo sphere packing exponent

\[
\tilde{E}_{psp,A}(R, L, p_{W}, \tilde{p}_{S|W}, p_{X|SW}, \mathcal{W}_K) = \min_{\tilde{p}_{Y|X_K|SW} \in \mathcal{P}_{Y|X_K|SW}(p_{W}, \tilde{p}_{S|W}, p_{X|SW}, \mathcal{W}_K, R, L, A)} D(\tilde{p}_{Y|X_K|SW} \| \tilde{p}_{Y|X_K} p_{X_K|SW} P_{S|W} | p_{W}).
\]

Taking the maximum and the minimum of \( \tilde{E}_{psp,A} \) above over all subsets of \( K \), we define

\[
\tilde{E}_{psp}(R, L, p_{W}, \tilde{p}_{S|W}, p_{X|SW}, \mathcal{W}_K) = \max_{R, L, p_{W}, \tilde{p}_{S|W}, p_{X|SW}, \mathcal{W}_K} \tilde{E}_{psp,K}(R, L, p_{W}, \tilde{p}_{S|W}, p_{X|SW}, \mathcal{W}_K),
\]

\[
\tilde{E}_{psp}(R, L, p_{W}, \tilde{p}_{S|W}, p_{X|SW}, \mathcal{W}_K) = \min_{A \subseteq K} \tilde{E}_{psp,A}(R, L, p_{W}, \tilde{p}_{S|W}, p_{X|SW}, \mathcal{W}_K).
\]

which are equal if \( \mathcal{W}_K = \mathcal{W}_K^{fair} \). Now define

\[
E_{psp}(R, L, D_1, \mathcal{W}_K) = \min_{p_{X|SW} \in \mathcal{P}_{X|SW}(p_{W} \tilde{p}_{S|W}, L, D_1)} \max_{p_{W} \in \mathcal{P}_{W} \tilde{p}_{S|W} \in \mathcal{P}_{S|W}} \tilde{E}_{psp,K}(R, L, p_{W}, \tilde{p}_{S|W}, p_{X|SW}, \mathcal{W}_K, \mathcal{W}_K^{fair}).
\]

Denote by \( \tilde{p}_W \) and \( \tilde{p}^*_X|SW \) the maximizers in (5.5), where the latter is to be viewed as a function of \( \tilde{p}_S|W \). Finally, define

\[
\tilde{E}_{psp}(R, L, D_1, \mathcal{W}_K) = \min_{\tilde{p}_S|W} \tilde{E}_{psp}(R, L, \tilde{p}_W, \tilde{p}_S|W, \tilde{p}^*_X|SW, \mathcal{W}_K),
\]

\[
E_{psp}(R, L, D_1, \mathcal{W}_K) = \min_{\tilde{p}_S|W} E_{psp}(R, L, D_1, \mathcal{W}_K),
\]

\[
\tilde{E}_{psp}(R, L, p_{W}, \tilde{p}_S|W, p_{X|SW}, \mathcal{W}_K).
\]

**Theorem 5.1** The decision rule (4.2) yields the following error exponents, uniformly over \( \{ \mathcal{W}_K, K \geq 1 \} \).

(i) The false-positive error exponent is

\[
E_{FP}(R, D_1, \mathcal{W}_K, \Delta) = \Delta.
\]

(ii) The error exponent for the (false negative) probability that the decoder fails to catch even one member of the coalition (misses every single colluder) is

\[
E^{one}(R, L, D_1, \mathcal{W}_K, \Delta) = E_{psp}(R + \Delta, L, D_1, \mathcal{W}_K).
\]

(iii) The error exponent for the (false negative) probability that the decoder fails to catch all members of coalition (misses some of them) is

\[
E^{all}(R, L, D_1, \mathcal{W}_K, \Delta) = E_{psp}(R + \Delta, L, D_1, \mathcal{W}_K).
\]

(iv) The error exponent for the (false negative) probability that the decoder fails to catch all members of coalition (misses some of them) is

\[
E^{all}(R, L, D_1, \mathcal{W}_K, \Delta) = E^{one}(R, L, D_1, \mathcal{W}_K^{fair}, \Delta).
\]

(v) If \( K = K_{non} \), the supremum of \( R \) such that \( E_{FP} \) and \( E^{one} \) (resp. \( E^{all} \)) are positive is \( C^{one}(D_1, \mathcal{W}_K^{fair}) \) (resp. \( C^{all}(D_1, \mathcal{W}_K) \)) and is obtained by letting \( \Delta \to 0 \) and \( L \to \infty \).

6. REFERENCES

[1] P. Moulin and J. A. O’Sullivan, “Information-Theoretic analysis of information hiding,” *IEEE Trans. on Information Theory*, Vol. 49, No. 3, pp. 563—593, March 2003.

[2] A. Somchek-Baruch and N. Merhav, “On the capacity game of private fingerprinting systems under collusion attacks,” *IEEE Trans. Information Theory*, vol. 51, no. 3, pp. 884—899, Mar. 2005.

[3] A. Somchek-Baruch and N. Merhav, “Achievable error exponents for the private fingerprinting game,” *IEEE Trans. Information Theory*, Vol. 53, No. 5, pp. 1827—1838, May 2007.

[4] P. Moulin and N. Kiyavash, “Expurgated Gaussian Fingerprinting Codes,” *Proc. IEEE Int. Symp. on Information Theory*, Nice, France, June 2007.

[5] D. Boneh and J. Shaw, “Collusion—Secure Fingerprinting for Digital Data,” in *Advances in Cryptology: Proc. CRYPTO’95*, Springer-Verlag, New York, 1995.

[6] G. Tardos, “Optimal Probabilistic Fingerprinting Codes,” *STOC*, 2003.

[7] N. P. Anthapadmanabhan, A. Barg and I. Dumer, “On the Fingerprinting Capacity Under the Marking Assumption,” submitted to *IEEE Trans. Information Theory*, arXiv:0611207v3, July 2007.

[8] P. Moulin, “Universal Fingerprinting: Capacity and Random-Coding Exponents,” preprint, arXiv:0801.3837, Jan. 2008.

[9] Y.-S. Liu and B. L. Hughes, “A new universal random coding bound for the multiple-access channel,” *IEEE Trans. Information Theory*, vol. 42, no. 2, pp. 376—386, Mar. 1996.

[10] G. D. Forney, Jr., “Exponential Error Bounds for Erasure, List, and Decision Feedback Schemes,” *IEEE Trans. Information Theory*, Vol. 14, No. 2, pp. 206—220, 1968.

[11] I. Csiszar and J. Korner, *Information Theory: Coding Theory for Discrete Memoryless Systems*, Academic Press, NY, 1981.

[12] A. Das and P. Narayan, “Capacities of Time-Varying Multiple-Access Channels With Side Information,” *IEEE Trans. Information Theory*, Vol. 48, No. 1, pp. 4—25, Jan. 2002.