Information-Theoretic Analysis of Spherical Fingerprinting

Pierre Moulin
ECE Dept, Coord. Sci. Lab., Beckman Inst.
University of Illinois
Urbana, IL 61801, USA
Email: pmoulin@illinois.edu

Ying Wang
Qualcomm Flarion Technologies
Bridgewater, NJ 08807, USA
Email: yingw@qualcomm.com

***Abstract*—Information-theoretic performance limits of digital fingerprinting systems subject to almost-sure squared-error distortion constraints on the fingerprint embedder and the colluders are derived in this paper. The rate of the fingerprinting code is $R = \frac{1}{N}\log M$ where N is codelength and M is the number of users. No assumption is made on the host signal statistics, but the collusion channel is also subject to a location-invariant condition. The receiver knows neither the collusion channel nor even the number of colluders. Capacity is the supremum of achievable rates and is shown to be equal to $\frac{1}{2K}\log(1+\frac{D_{\rm f}}{KD_{\rm c}})$ where K is the number of colluders, and $D_{\rm f}$ and $D_{\rm c}$ are the L^2-distortion tolerance levels for the fingerprint embedder and the colluders, respectively. The worst collusion is shown to consist of uniform linear averaging of the coalition's marked copies followed by addition of independent spherical noise. Positive error exponents are achieved at all rates below capacity using random spherical fingerprinting codes and a new universal decoding criterion based on empirical Gaussian mutual information. It is also shown that minimum-distance decoding fails for this problem, and that a simple single-user decoder is almost as good as the universal decoder for large K. Geometric interpretations for all the results are given.**

Keywords. Digital fingerprinting, capacity, error exponents, universal coding, normalized correlation, randomized codes, Gaussian random variables, multiple-access channels, typical sets, model order selection.

I. Introduction

Digital fingerprinting systems can be used for traitor tracing or digital rights management applications. A length-N real-valued signal is to be protected and distributed to M users. Some of the users (K of them) may collude and process their copies to create a *pirated copy* that contains only weak traces of their fingerprints. This problem was first posed by Cox *et al.* [1] who proposed the use of *Gaussian fingerprints* for this purpose. Specifically, their fingerprints were i.i.d. (independent and identically distributed) Gaussian sequences; the fingerprint code is shared with the decoder but not revealed to the users.

A fundamental question is what are the optimal performance limits for detection of colluders. To make the problem non-trivial, one may assume embedding distortion constraints on the fingerprinter and the colluders. Examples of this analysis include [2]–[4] for the case of signals defined over finite alphabets, and [5], [6] for the case of real-valued signals. In the latter case, an obvious (but not necessarily optimal) strategy for the colluders is to perform a uniform linear average of their copies and add i.i.d. Gaussian noise; this strategy was examined in the above papers. Possible improvements for the attackers consist of developing (nonlinear) order-statistics attacks [7], as illustrated by numerical simulations [8]. The papers [5]–[8] did not establish whether stronger attacks or better decoders might exist.

An information-theoretic analysis of universal fingerprinting over finite alphabets was recently reported in [4]. The decoder has access to a pirated copy as well as to the host signal (nonblind detection) and returns a list of accused users. The decoder knows neither the collusion channel nor even the number of colluders. The cost functions are the false-positive and false-negative error probabilities, which should vanish for any admissible collusion strategy. The same universal fingerprinting setup is assumed in this paper. However the random coding techniques of [4] rely on the method of types and are not applicable here. A different approach is therefore needed. An extended version of this paper containing the proofs will be presented in [9].

Notation. Throughout this paper, we use boldface uppercase letters to denote random vectors, and uppercase letters for the components of the vectors. Mathematical expectation is denoted by the symbol $\mathbb{E}$. The shorthands $a_N \doteq b_N$ and $a_N \dot{\leq} b_N$ denote asymptotic relations in the exponential scale, respectively $\lim_{N\to\infty}\frac{1}{N}\log\frac{a_N}{b_N} = 0$ and $\limsup_{N\to\infty}\frac{1}{N}\log\frac{a_N}{b_N} \leq 0$. We define $\mathsf{K} \triangleq \{1,2,\cdots,K\}$. The Euclidean norm of a vector $\mathbf{x}$ is denoted by $\|\mathbf{x}\| = (\sum_i x_i^2)^{1/2}$. The N-sphere of radius r is defined as $\{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\|^2 = r^2\}$. If $r=1$, the N-sphere is conventionally denoted by $\mathcal{S}^{N-1}$. The $K\times K$ identity matrix is denoted by $\mathbf{I}_K$. For any collection of vectors $\{\mathbf{x}_1,\ldots,\mathbf{x}_M\}$, we denote by $\mathbf{x}_{\mathcal{K}} = \{\mathbf{x}_k,\, k\in\mathcal{K}\}$ the restriction of this collection to its components $k\in\mathcal{K}$. The Gaussian distribution with mean zero and covariance matrix $\mathbf{R}$ is denoted by $\mathcal{N}(0,\mathbf{R})$. The determinant of a matrix is denoted by $|\cdot|$. The differential entropy of a random variable X with probability density function (pdf) p is denoted by $h(X) = -\int p\log p$. The Kullback-Leibler divergence between two pdf's p and q is denoted by $D(p\|q) = \int p\log\frac{p}{q}$.

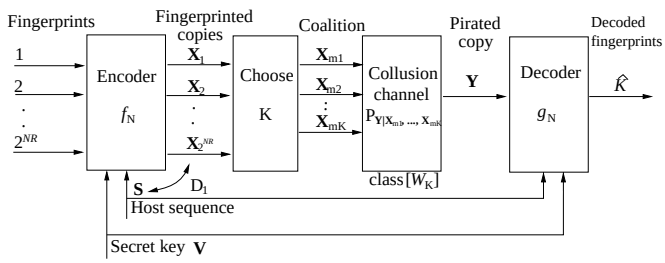

Fig. 1. The fingerprinting process and the collusion channel $p_{\mathbf{Y}|\mathbf{X}_{\mathcal{K}}}$.

II. Problem Statement

The mathematical setup of the problem is diagrammed in Fig. 1.

A. Fingerprint Generation and Embedding

The host signal is a sequence $\mathbf{S} = (S(1), \ldots, S(N))$ in $\mathbb{R}^N$, viewed as *deterministic* but *unknown* to the colluders. Fingerprints are added to $\mathbf{S}$, and the marked copies of the signal are distributed to 2^{NR} users. Specifically, user m is assigned a marked copy $\mathbf{X}_m = \mathbf{S} + \mathbf{U}_m$ where $m \in \{1, \ldots, 2^{NR}\} \triangleq \mathcal{M}_N$, and $\mathbf{U}_m \in \mathbb{R}^N$ is the fingerprint assigned to user m. The fingerprints are randomized by means of a random variable $\mathbf{V}$ shared between encoder and decoder but unknown to the colluders. The random variable $\mathbf{V}$ is independent of $\mathbf{S}$ and m, and the dependency of $\mathbf{U}_m$ on $\mathbf{V}$ is not indicated explicitly. The encoding function takes the form $\mathbf{X}_m = f_N(\mathbf{S}, m, \mathbf{V})$.

The fingerprints $\mathbf{U}_m$, $m \in \mathcal{M}_N$ form a $(N, 2^{NR})$ fingerprinting code $\mathcal{C}$ of rate R. In a typical signal fingerprinting application, $N \sim 10^3 - 10^9$ and $2^{NR} < 10^9$ (not to exceed the number of humans).

Fingerprints must satisfy the distortion constraint

$$\|\mathbf{U}_m\|^2 \leq ND_{\mathrm{f}}, \quad \forall m \in \mathcal{M}_N,$$

where D_{f} is the mean-squared distortion due to embedding.

As we shall see, only two kinds of randomization are needed here. The first is randomized rotation of a prototype fingerprint constellation $\widetilde{\mathcal{C}} = \{\widetilde{\mathbf{U}}_m, m \in \mathcal{M}_N\}$ via a parameter Θ that is uniformly distributed over the special orthonormal group $SO(N)$. The randomly rotated code is of the form $\mathcal{C}^{\Theta} = \{\Theta\widetilde{\mathbf{U}}_m, m \in \mathcal{M}_N\}$. The second kind of randomization is permutation of the users fingerprint assignments, to equalize error probabilities for all coalitions. The randomly permuted code is of the form $\mathcal{C}^{\pi} = \{\widetilde{\mathbf{U}}_{\pi^{-1}(m)}, m \in \mathcal{M}_N\}$ where π is drawn uniformly from the set of all $2^{NR}!$ permutations of $\mathcal{M}_N$. Combining both kinds of randomization yields a randomized code $\mathcal{C}^{\Theta,\pi} = \{\Theta\widetilde{\mathbf{U}}_{\pi^{-1}(m)}, m \in \mathcal{M}_N\}$. Thus $\mathbf{V} = (\Theta, \pi)$.

B. Attack Model

Denote by $\mathcal{K} \subseteq \mathcal{M}_N$ the *coalition*, i.e., the index set of the colluding users. Their coalition has cardinality K. They select a conditional pdf $p_{\mathbf{Y}|\mathbf{X}_{\mathcal{K}}}$, termed the **collusion channel**. Then they draw a pirated copy $\mathbf{Y} \in \mathbb{R}^N$ from that distribution. Consider the following constraints on the collusion channel.

(A1) *Location-Invariant constraint*:

$$p_{\mathbf{Y}|\mathbf{X}_{\mathcal{K}}}(\mathbf{y}|\mathbf{x}_{\mathcal{K}}) = p_{\mathbf{Y}|\mathbf{X}_{\mathcal{K}}}(\mathbf{y}-\mathbf{s}|(\mathbf{x}-\mathbf{s})_{\mathcal{K}}), \quad \forall \mathbf{s} \in \mathbb{R}^N.$$

(A2) *Almost-Sure Mean-Squared Distortion constraint*:

$$\|\mathbf{Y} - \overline{\mathbf{X}}\|^2 \leq ND_c$$

where

$$\overline{\mathbf{X}} = \frac{1}{K}\sum_{k \in \mathcal{K}} \mathbf{X}_k \tag{1}$$

is the mean of the colluders' fingerprints.

(A3) *Fairness constraint*:

$$p_{\mathbf{Y}|\mathbf{X}_{\mathcal{K}}}(\mathbf{y}|\mathbf{x}_{\pi\mathcal{K}}) = p_{\mathbf{Y}|\mathbf{X}_{\mathcal{K}}}(\mathbf{y}|\mathbf{x}_{\mathcal{K}})$$

for all permutations π of the index set $\mathcal{K}$.

The constraint **(A1)** precludes attacks involving filtering of host signal components. The motivation for this restriction is that it considerably simplifies the mathematical derivation and does not require a statistical model for the host $\mathbf{S}$. The restriction is relatively mild if embedding is done in a transform domain in which the components of the host $\mathbf{S}$ are approximately independent and are large relative to the embedding distortion. The motivation for **(A2)** is that distortion is best measured relative to the host $\mathbf{S}$, but $\mathbf{S}$ is not known to the coalition, so we replace $\mathbf{S}$ by its best linear unbiased estimate, $\overline{\mathbf{X}}$. Choosing an *expected distortion* constraint of the form $\mathbb{E}\|\mathbf{Y} - \mathbf{S}\|^2 \leq ND_c$ would allow impulsive noise strategies which blast independent and identically distributed (i.i.d.) additive noise $\mathcal{N}(0, ND_c)$ with probability $1/N$. Such attacks are extremely effective [12] as they result in zero error exponents. Finally, the fairness condition **(A3) will not be imposed but holds for optimal attacks** [4, Prop. 2.4]. All members of the coalition incur the same risk. The constraints **(A1)** and **(A2)** define a class $\mathscr{W}_K$ of feasible collusion channels. That class is convex and permutation-invariant. The subset of $\mathscr{W}_K$ consisting of the collusion channels that satisfy **(A3)** will be denoted by $\mathscr{W}_K^{\mathrm{fair}}$.

A collusion channel that will be of interest is

$$\mathbf{Y} = \overline{\mathbf{X}} + \mathbf{E} \tag{2}$$

where the noise $\mathbf{E}$ is independent of $\mathbf{X}_{\mathcal{K}}$ and uniformly distributed on the centered sphere with radius $\sqrt{ND_{\mathrm{c}}}$. (Note that if the attackers can retrieve the original signal $\mathbf{S}$, they will succeed in defeating the decoder.) The attack of (2) satisfies the requirements **(A1)—(A3)**.

C. Decoder

We study the nonblind scenario where the host signal $\mathbf{S}$ is available at the decoder. The decoder input is $(\mathbf{Y}, \mathbf{S}, \Theta, \pi)$ and its output is an estimated coalition

$$\hat{\mathcal{K}} = g_N(\mathbf{y}, \mathbf{s}, \theta, \pi) = \pi(\tilde{g}_N(\theta^{-1}(\mathbf{y} - \mathbf{s})))$$

where $\tilde{g}_N$ is the decoding function for the prototype code $\widetilde{\mathcal{C}}$. Let $\widetilde{\mathcal{D}}_m$ be the prototype code's decoding region for user m, i.e.,

$$\theta^{-1}(\mathbf{y}-\mathbf{s}) \in \widetilde{\mathcal{D}}_{\pi^{-1}(m)} \quad \Leftrightarrow \quad m \in \hat{\mathcal{K}}.$$

The decoder does not know the channel $p_{\mathbf{Y}|\mathbf{X}_\mathcal{K}}$ used by the colluders or even the exact number K of colluders.

D. Performance Metrics

As detailed in [4], there are three basic error probabilities of interest: the probability $P_{\rm FP}$ of false positives (accusing one or more innocent users), the probability $P_e^{\rm all}$ of failing to catch all colluders, and the probability $P_e^{\rm one}$ of failing to catch even one colluder. By our location-invariant assumption on the attack channel and the decoding regions, these error probabilities are independent of $\mathbf{s}$. Since user assignments are randomly permuted, these probabilities are also independent of $\mathcal{K}$. Their dependency on the code (f_N, g_N) and the collusion channel $p_{\mathbf{Y}|\mathbf{X}_\mathcal{K}}$ is indicated explicitly as follows:

$$\begin{aligned} P_{\rm FP}(f_N, g_N, p_{\mathbf{Y}|\mathbf{X}_\mathcal{K}}) &= Pr[\hat{\mathcal{K}} \setminus \mathcal{K} \neq \emptyset] \\ P_e^{\rm one}(f_N, g_N, p_{\mathbf{Y}|\mathbf{X}_\mathcal{K}}) &= Pr[\hat{\mathcal{K}} {\rm cap} \mathcal{K} = \emptyset] \\ P_e^{\rm all}(f_N, g_N, p_{\mathbf{Y}|\mathbf{X}_\mathcal{K}}) &= Pr[\mathcal{K} \not\subseteq \hat{\mathcal{K}}]. \end{aligned}$$

The worst-case error probabilities over $p_{\mathbf{Y}|\mathbf{X}_\mathcal{K}} \in \mathscr{W}_K$ are denoted by $P(f_N, g_N, \mathscr{W}_K)$ where P represents $P_{\rm FP}$, $P_e^{\rm one}$, or $P_e^{\rm all}$.

Definition 1: A rate R is achievable for embedding distortion $D_{\rm f}$, collusion class $\mathscr{W}_K$, and **detect-one** criterion if there exists a sequence of $(N, \lceil 2^{NR} \rceil)$ randomized codes (f_N, g_N) with maximum embedding distortion $D_{\rm f}$, such that both $P_e^{\rm one}(f_N, g_N, \mathscr{W}_K)$ and $P_{\rm FP}(f_N, g_N, \mathscr{W}_K)$ vanish as $N \to \infty$.

Definition 2: A rate R is achievable for embedding distortion $D_{\rm f}$, collusion class $\mathscr{W}_K$, and **detect-all** criterion if there exists a sequence of $(N, \lceil 2^{NR} \rceil)$ randomized codes (f_N, g_N) with maximum embedding distortion $D_{\rm f}$, such that both $P_e^{\rm all}(f_N, g_N, \mathscr{W}_K)$ and $P_{\rm FP}(f_N, g_N, \mathscr{W}_K)$ vanish as $N \to \infty$.

Definition 3: Fingerprinting capacities $C^{\rm one}(D_{\rm f}, \mathscr{W}_K)$ and $C^{\rm all}(D_{\rm f}, \mathscr{W}_K)$ are the suprema of all achievable rates with respect to the detect-one and detect-all criteria, respectively.

It was shown in [4, Theorems 3.3, 3.4] that $C^{\rm all}(D_{\rm f}, \mathscr{W}_K) = C^{\rm one}(D_{\rm f}, \mathscr{W}_K)$.

For a sequence of randomized codes (f_N, g_N), the error exponents are defined as

$$E(R, D_{\rm f}, \mathscr{W}_K) = \liminf_{N\to\infty} \left[-\frac{1}{N} \log P_e(f_N, g_N, \mathscr{W}_K) \right]$$

where E represents the random coding exponent $E_{\rm FP}$, $E^{\rm one}$, or $E^{\rm all}$. Moreover, $E^{\rm all}(R, D_{\rm f}, \mathscr{W}_K) \leq E^{\rm one}(R, D_{\rm f}, \mathscr{W}_K)$ because an error event for the detect-one problem is also an error event for the detect-all problem. We have $E^{\rm all} = 0$ if the class $\mathscr{W}_K$ includes channels in which one colluder can "stay out," i.e., not contribute to the pirated copy.

III. Spherically Symmetric Collusion Channel

Lemma 1: For any fingerprinting code with randomized rotation Θ, there is no loss of optimality for the colluders in choosing a spherically symmetric collusion channel: $p_{\mathbf{Y}|\mathbf{X}_\mathcal{K}}(\mathbf{y}|\mathbf{x}_\mathcal{K}) = p_{\mathbf{Y}|\mathbf{X}_\mathcal{K}}(\theta_c\mathbf{y}|\theta_c\mathbf{x}_\mathcal{K})$ for any $\theta_c \in SO(N)$.

Hence for the error probability analyses in this paper, we assume without loss of generality that a spherically symmetric collusion channel is used. Moreover it follows from **(A1)** that there is no loss of generality in assuming $\mathbf{S} = 0$. Thus $\mathbf{X}_m = \mathbf{U}_m$ for each m, and we refer to $\mathbf{X}_m$, $m \in \mathcal{M}_N$ as the fingerprints in the sequel.

IV. Mutual-Information Game

Fingerprinting capacity was obtained in [4] as the value of a mutual-information game. When $S = \emptyset$, capacity takes the form

$$C(K) = \sup_{p_W, p_{X|W}} \min_{p_{Y|X_\mathsf{K}}} \frac{1}{K} I(X_\mathsf{K}; Y|W) \tag{3}$$

where W plays the role of a time-sharing random variable, and the joint distribution of (W, X_K, Y) is given by

$$p(w, x_\mathsf{K}, y) = p_W(w) \left(\prod_{k\in\mathsf{K}} p_{X|W}(x_k|w) \right) p_{Y|X_\mathsf{K}}(y|x_\mathsf{K}).$$

The optimizations are subject to the distortion constraints

$$\sum_{w,x} p_W(w) p_{X|W}(x|w) d(x) \leq D_{\rm f}, \tag{4}$$

$$\begin{aligned} \sum_{w,x_\mathsf{K},y} p_W(w) \left(\prod_{k\in\mathsf{K}} p_{X|W}(x_k|w) \right) \\ \times p_{Y|X_\mathsf{K}}(y|x_\mathsf{K})\, d(f(x_\mathsf{K}), y) \leq D_{\rm c}, \end{aligned} \tag{5}$$

and possibly additional convex, permutation-invariant constraints on $p_{Y|X_\mathsf{K}}$. In (4) and (5), $d(x)$ and $d(f(x_\mathsf{K}), y)$ are the distortion functions for the fingerprint embedder and the colluders, respectively, and f is a permutation-invariant operator, typically representing an averaging. For our problem with quadratic distortion constraints, we have $d(x) = x^2$ in (4), and $f(x_\mathsf{K}) = \overline{x} = \frac{1}{K}\sum_k x_k$ and $d(\overline{x}, y) = (y - \overline{x})^2$ in (5). We also have the location-invariant constraint

$$p_{Y|X_\mathsf{K}}(y|x_\mathsf{K}) = p_{Y|X_\mathsf{K}}(y - s|(x - s)_\mathsf{K}), \quad \forall s \in \mathbb{R}.$$

While (3) was derived under the assumption of discrete alphabets, modifications of the proof can be made to extend the applicability of the formula to the case of abstract alphabets. In this section we solve the maxmin problem of (6) and derive a simple formula in terms of the number of colluders, K, and the distortion levels $D_{\rm f}$ and $D_{\rm c}$ for the fingerprint embedder and colluders.

Let p_X^* be the Gaussian pdf with mean 0 and variance $D_{\rm f}$, and $p_{Y|X_\mathsf{K}}^*$ the conditional Gaussian pdf with mean $\overline{X} = \frac{1}{K}\sum_k X_k$ and variance $D_{\rm c}$.

Theorem 2: The solution to the maxmin game of (3) is given by

$$C(K) = \frac{1}{2K} \log \left(1 + \frac{D_{\rm f}}{K D_{\rm c}} \right) \tag{6}$$

and is achieved by $W = \emptyset$, $X \sim \mathcal{N}(0, D_{\rm f})$, and $Y \sim \mathcal{N}(\overline{X}, D_{\rm c})$. Moreover $C(K) = \frac{1}{K} I(\overline{X}; Y)$.

Note that the time-sharing random variable is degenerate here. This is in contrast with discrete-alphabet problems such as the Boney-Shaw problem with binary alphabet where the presence of W makes an important difference [4, Sec. 3].

V. Limitations of Minimum-Distance Decoding

One of the difficulties with fingerprinting is that the number of colluders is unknown to the encoder and decoder. For simplicity we assume in this section that the colluders use the averaging plus additive white Gaussian noise (AWGN) attack:

$$\mathbf{Y} = \frac{1}{K} \sum_{k \in \mathcal{K}} \mathbf{X}_k + \mathbf{W} = \overline{\mathbf{X}} + \mathbf{W}$$

where $\mathbf{W}$ is i.i.d. $\mathcal{N}(0, D_{\rm c})$. The minimum-distance decoding rule takes the form

$$\min_{\mathcal{K}} \left\| \mathbf{y} - \frac{1}{K} \sum_{k \in \mathcal{K}} \mathbf{x}_k \right\|^2. \tag{7}$$

This would coincide with the maximum likelihood (ML) decoding rule if K were known and the minimization over coalitions $\mathcal{K}$ of size K. In conventional decoding problems, the ML decoder minimizes error probability and is therefore optimal. Since K is unknown here, (7) is a generalized ML decoding rule, which may be severely suboptimal.

To investigate this issue, consider a closely related joint typicality decoder, which is easier to analyze. Fix an arbitrarily small $\epsilon > 0$. The joint typicality decoder outputs $\mathcal{K}'$ if and only if $\mathcal{K}'$ is the only coalition that satisfies the following conditions:

- Typical noise energy:

$$\left| \frac{1}{N D_{\rm c}} \left\| \mathbf{y} - \frac{1}{K} \sum_{k \in \mathcal{K}'} \mathbf{x}_k \right\|^2 - 1 \right| \le \epsilon. \tag{8}$$

- Nearly orthogonal fingerprints:

$$\left| \frac{1}{N D_{\rm f}} \mathbf{x}_k \cdot \mathbf{x}_l - \mathbb{1}\{k = l\} \right| \le \epsilon, \quad \forall k, l \in \mathcal{K}'. \tag{9}$$

  The set of coalitions $\mathcal{K}'$ that satisfy (9) will be denoted by $\mathscr{K}(\epsilon)$. By convention, the empty coalition $\emptyset \in \mathscr{K}(\epsilon)$.

It may be verified that this decoder is plagued by false positives because K is unknown and unbounded. No positive rates are achievable.

However, if a maximum coalition size $K_{\max}$ is given to the decoder, all rates below $C(K_{\max})$ are achievable. Still, the decoder would fail if the coalition size is $K < K_{\max}$ and the code rate satisfies $C(K_{\max}) < R < C(K)$. This decoder

- does not adapt to the actual coalition size;
- does not provide a tradeoff between false positives and false negatives;
- rejects any slightly atypical coalition $\mathcal{K} \notin \mathscr{K}(\epsilon)$ from consideration, hence its error exponents are zero.

VI. Empirical Mutual Informations

Recall that the mutual information $I(X;Y) = D(p_{XY} \| p_X p_Y)$ between two Gaussian random variables X and Y with normalized correlation coefficient $\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} \in [-1, 1]$ is given by

$$I_G(X;Y) = -\frac{1}{2} \log(1 - \rho^2). \tag{10}$$

More generally, if $(X_{\mathcal{K}}, Y)$ are jointly Gaussian, then

$$I_G(X_{\mathcal{K}}; Y) = \frac{1}{2} \log \frac{|\mathrm{Cov}(X_{\mathcal{K}})| \, \mathrm{Var}(Y)}{|\mathrm{Cov}(X_{\mathcal{K}}, Y)|}$$

where $|\cdot|$ denotes matrix determinant. For X_k i.i.d. $\mathcal{N}(0, D_{\rm f})$, $\mathrm{Cov}(X_{\mathcal{K}})$ is $D_{\rm f}$ times the $K \times K$ identity matrix, and

$$I_G(X_{\mathcal{K}}; Y) = -\frac{1}{2} \log \left(1 - \sum_{k \in \mathcal{K}} \rho^2(X_k, Y) \right).$$

The mutual information of L random variables $U_1, \cdots, U_L$ with joint differential entropy $h(U_1, \cdots, U_L)$ is defined as

$$\begin{aligned} \mathring{I}(U_1; \cdots; U_L) &= \sum_{k=1}^{L} h(U_k) - h(U_1, \cdots, U_L) \\ &= D(p_{U_1 \cdots U_L} \| p_{U_1} \cdots p_{U_L}). \end{aligned} \tag{11}$$

Lemma 3: The mutual information of L jointly Gaussian random variables $U_1, \cdots, U_L$, is given by

$$\mathring{I}_G(U_1; \cdots; U_L) = -\frac{1}{2} \log |\mathbf{R}| \tag{12}$$

where $\mathbf{R}$ is the $L \times L$ matrix of normalized correlation coefficients whose entries are

$$\mathbf{R}_{jk} = \rho(U_j, U_k), \quad 1 \le j, k \le L.$$

The matrix $\mathbf{R}$ will be henceforth termed the *normalized-correlation matrix*. Note that $\mathring{I}_G(U_1; \cdots; U_L)$ is invariant to scalings of the individual random variables U_k, and that (10) is a special case of (12) with $L = 2$.

Empirical Mutual informations. Given two sequences $\mathbf{x}$ and $\mathbf{y}$ defined over *discrete alphabets*, the empirical mutual information between $\mathbf{x}$ and $\mathbf{y}$ is defined as the mutual information with respect to the empirical joint distribution (joint type) of $\mathbf{x}$ and $\mathbf{y}$. The empirical mutual information of L random variables is also defined with respect to their joint type. This definition does not admit a natural extension to the case of continuous alphabets, so we follow a different path.

Given two sequences $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^N$, the normalized empirical correlation coefficient between $\mathbf{x}$ and $\mathbf{y}$ is defined as $\rho(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|}$. Likewise, the empirical correlation between two sequences $\mathbf{x}$ and $\mathbf{y}$ is defined as $\widehat{\mathrm{Cov}}(\mathbf{x}, \mathbf{y}) = \frac{1}{N} \mathbf{x} \cdot \mathbf{y}$. The empirical correlation matrix $\widehat{\mathrm{Cov}}(\mathbf{u}_1, \cdots, \mathbf{u}_L)$ for L sequences $\mathbf{u}_1, \cdots, \mathbf{u}_L$ is the $L \times L$ matrix whose (j,k) element is given by $\widehat{\mathrm{Cov}}(\mathbf{u}_j, \mathbf{u}_k)$. The empirical normalized correlation matrix is defined as $\mathbf{R} = \{\rho(\mathbf{u}_j, \mathbf{u}_k)\}_{j,k=1}^{L}$.

We now define the *empirical Gaussian mutual information between* $\mathbf{x}$ *and* $\mathbf{y}$ as

$$I_G(\mathbf{x},\mathbf{y}) = -\frac{1}{2}\log(1-\rho^2(\mathbf{x},\mathbf{y})) \tag{13}$$

and the *empirical Gaussian mutual information of* $\mathbf{u}_1,\cdots,\mathbf{u}_L$ as

$$\overset{\circ}{I}_G(\mathbf{u}_1;\cdots;\mathbf{u}_L) = -\frac{1}{2}\log|\mathbf{R}| \tag{14}$$

where

$$|\mathbf{R}| = \left|[\rho(\mathbf{u}_j,\mathbf{u}_k)]_{j,k=1}^L\right| = \frac{|\widehat{\mathrm{Cov}}(\mathbf{u}_1,\cdots,\mathbf{u}_L)|}{\prod_k \frac{1}{N}\|\mathbf{u}_k\|^2}. \tag{15}$$

This expression is invariant to scalings of the individual sequences, and (13) is a special case of (14) with $L=2$.

Chain Rule. Similarly to [4, pp. 20,21], we shall use a chain rule for $\overset{\circ}{I}_G$. To this end, we define

$$\begin{aligned}
&\overset{\circ}{I}_G(\mathbf{u}_1;\cdots;\mathbf{u}_i;\mathbf{u}_{i+1}\cdots\mathbf{u}_L)\\
&\quad= \overset{\circ}{I}_G(\mathbf{u}_1;\cdots;\mathbf{u}_L) - \overset{\circ}{I}_G(\mathbf{u}_{i+1};\cdots;\mathbf{u}_L)\\
&\quad= -\frac{1}{2}\log\frac{|\mathbf{R}_{1:i}|}{|\mathbf{R}_{i+1:L}|}
\end{aligned}$$

where $\mathbf{R}_{i:j}$ is the normalized correlation matrix for the $j-i+1$ sequences $\mathbf{u}_i,\cdots,\mathbf{u}_j$.

VII. Decoders

We consider three decoders: a simple (single-user) thresholding decoder, a joint decoder that achieves positive error exponents at all rates below capacity, and a simplified capacity-achieving joint decoder based on typicality (this decoder has zero error exponents).

A. Thresholding Decoder

The simple decoder outputs the estimated coalition $\hat{\mathcal{K}}$ that consists of all user indices m such that

$$I_G(\mathbf{x}_m;\mathbf{y}) > R+\Delta \tag{16}$$

where $0<\Delta<C(K_{\max})-R$.

Let

$$\eta = \sqrt{1-2^{-2(R+\Delta)}} \tag{17}$$

hence $R+\Delta = -\frac{1}{2}\log(1-\eta^2)$. By (10), the simple decoder of (16) takes the equivalent form

$$m\in\hat{\mathcal{K}} \quad\Leftrightarrow\quad |\rho(\mathbf{x}_m,\mathbf{y})| > \eta \tag{18}$$

Note the similarity with the classical maximum normalized correlation decoder:

$$\hat{m} = \underset{1\le m\le 2^{NR}}{\operatorname{argmax}}\ \rho(\mathbf{x}_m,\mathbf{y}).$$

B. MPGMI Decoder

Denote by

$$\mathbf{R} = \begin{bmatrix} 1 & \cdots & \rho(\mathbf{x}_{m_1},\mathbf{x}_{m_K}) & \rho(\mathbf{x}_{m_1},\mathbf{y})\\ \rho(\mathbf{x}_{m_2},\mathbf{x}_{m_1}) & \vdots & \rho(\mathbf{x}_{m_2},\mathbf{x}_{m_K}) & \rho(\mathbf{x}_{m_2},\mathbf{y})\\ \vdots & & \vdots & \vdots\\ \rho(\mathbf{x}_K,\mathbf{x}_{m_1}) & \cdots & 1 & \rho(\mathbf{x}_{m_K},\mathbf{y})\\ \rho(\mathbf{y},\mathbf{x}_{m_1}) & \cdots & \rho(\mathbf{y},\mathbf{x}_{m_K}) & 1 \end{bmatrix} \tag{19}$$

the normalized correlation matrix for the sequences $\mathbf{x}_{m_1},\cdots,\mathbf{x}_{m_K},\mathbf{y}$. The joint decoder outputs $\hat{\mathcal{K}}$ that maximizes the maximum penalized mutual information criterion

$$\begin{aligned}
MPGMI(\mathcal{K}) &\triangleq \overset{\circ}{I}_G(\mathbf{x}_{\mathcal{K}};\mathbf{y}) - K(R+\Delta)\\
&= -\frac{1}{2}\log|\mathbf{R}| - K(R+\Delta)
\end{aligned} \tag{20}$$

over all possible coalitions $\mathcal{K}$. By convention, $MPGMI(\emptyset)=0$, and $\hat{\mathcal{K}}=\emptyset$ is an admissible decision.

C. $\widetilde{MPGMI}$ Joint Decoder

A simple approximation to $\mathbf{R}$ in (19) can be obtained based on the fact that $\rho(\mathbf{x}_j,\mathbf{x}_k)$ is equal to 1 for $j=k$ and converges in probability to zero for $j\neq k$, as $N\to\infty$. Write

$$\begin{aligned}
\widetilde{\mathbf{R}} &= \begin{vmatrix} 1 & \cdots & 0 & \rho(\mathbf{x}_{m_1},\mathbf{y})\\ 0 & \vdots & 0 & \rho(\mathbf{x}_{m_2},\mathbf{y})\\ \vdots & & \vdots & \vdots\\ 0 & \cdots & 1 & \rho(\mathbf{x}_{m_K},\mathbf{y})\\ \rho(\mathbf{y},\mathbf{x}_{m_1}) & \cdots & \rho(\mathbf{y},\mathbf{x}_{m_K}) & 1 \end{vmatrix}\\
&= 1-\sum_{m\in\mathcal{K}}\rho^2(\mathbf{x}_m,\mathbf{y})
\end{aligned} \tag{21}$$

which converges in probability to $1-\sum_{m\in\mathcal{K}}\rho^2(X_m,Y)$.

We will thus consider a *simplified decoder* based on the approximation (21). Plugging this approximation into (20), we obtain the criterion

$$\widetilde{MPGMI}(\mathcal{K}) = -\frac{1}{2}\log\left(1-\sum_{m\in\mathcal{K}}\rho^2(\mathbf{x}_m,\mathbf{y})\right) - K(R+\Delta) \tag{22}$$

where $\widetilde{MPGMI}(\emptyset)=0$ again. The $\widetilde{MPGMI}$ criterion is maximized over the set $\mathscr{K}(\epsilon)$ of typical coalitions; see (9).

VIII. Analysis of Thresholding Decoder

Define the function

$$E_{sp}(A,\tau) = \begin{cases} \infty & \text{if } 0<\tau\le\frac{1}{2}|\log A|^+\\ E_{\mathrm{cap}}(\rho^*) & \text{if } \frac{1}{2}|\log A|^+<\tau\le\frac{1}{2}\log(1+A)\\ 0 & \text{if } \tau\ge\frac{1}{2}\log(1+A), \end{cases} \tag{23}$$

where $|t|^+\triangleq\max\{t,0\}$, $E_{\mathrm{cap}}(\rho)\triangleq-\frac{1}{2}\log(1-\rho^2)$, and

$$\rho^* = A^{1/2}\,2^{-2\tau}\left(-1+\sqrt{1-2^{2\tau}(1-A^{-1}(2^{2\tau}-1))}\right) \in[-1,0]. \tag{24}$$

Lemma 4: Fix $\mathbf{x} \in \mathcal{S}^{N-1}$ and define the random sequence $\mathbf{Y} = \mathbf{x} + \mathbf{W}$ where $\mathbf{W}$ is uniformly distributed over the centered n-sphere of radius $A^{-1/2}$. Then

$$Pr[I_G(\mathbf{x};\mathbf{y}) < \tau] \doteq 2^{-nE_{sp}(A,\tau)}, \quad \forall \tau > 0.$$

Similarly to [13], the proof admits a simple geometric interpretation. Unlike the AWGN channel of [13], here the exponent is infinite if $\tau \leq \frac{1}{2}|\log A|^+$ because the event $I_G(\mathbf{x};\mathbf{y}) < \tau$ has zero probability.

Proposition 5: The threshold decoder of (16) achieves the following error exponents:

$$E_{\rm FP}(R,\Delta,K) = \Delta \tag{25}$$

$$E^{\rm one}(R,\Delta,K) \leq E_{sp}\left(\frac{D_{\rm f}/K^2}{D_{\rm c}+D_{\rm f}(K-1)/K^2}, R+\Delta\right). \tag{26}$$

Moreover, under **(A3)**, $E^{\rm all}(R,\Delta,K) = E^{\rm one}(R,\Delta,K)$. The supremum of the rates for which these error exponents are positive is

$$C_1(K) = \frac{1}{2}\log\left(1+\frac{D_{\rm f}}{D_{\rm f}(K-1)+D_{\rm c}K^2}\right) \tag{27}$$

and is achieved by letting $\Delta \to 0$.

The threshold $R+\Delta$ controls the fundamental tradeoff between false-positive and false-negative error probabilities. However $C_1(K)$ is strictly lower than $C(K)$, so this decoder is not capacity-achieving.

IX. Analysis of MPGMI Decoder

For notational convenience, the fingerprints $\mathbf{x}_m$, $m \in \mathcal{M}_N$ and the pirated copy $\mathbf{y}$ are scaled by $(ND_{\rm f})^{-1/2}$ in this section. Thus each $\mathbf{x}_m \in \mathcal{S}^{N-1}$.

False Negatives. We focus again on the detect-one criterion. Following the proof of Theorem 5.2 in [4], the error probability is upper-bounded by $Pr[\overset{\circ}{I}_G(\mathbf{x}_{\mathcal{K}};\mathbf{y}) < K(R+\Delta)]$. Since the random variable $\overset{\circ}{I}_G(\mathbf{x}_{\mathcal{K}};\mathbf{y})$ converges in probability to $\overset{\circ}{I}_G(X_{\mathcal{K}};Y) = I_G(X_{\mathcal{K}};Y)$, the false-negative probability vanishes provided that $R+\Delta < \frac{1}{K}\overset{\circ}{I}_G(X_{\mathcal{K}};Y) = C(K)$. Error exponents will be given below.

False Positives. Given $\mathbf{x}_{\mathcal{K}}$ and $\mathbf{y}$, the conditional probability of false positives is upper bounded by [4, p. 49]

$$P_{\rm FP}(\mathbf{x}_{\mathcal{K}},\mathbf{y}) \leq \sum_{\mathcal{B}\subseteq\mathcal{K}}\sum_{|\mathcal{A}|\geq 1} 2^{N|\mathcal{A}|R} \times Pr[\overset{\circ}{I}_G(\mathbf{x}_{\mathcal{A}};\mathbf{y}\mathbf{x}_{\mathcal{B}}) > |\mathcal{A}|(R+\Delta)] \tag{28}$$

where $\mathcal{A}$ is a set of innocent users: $\mathbf{x}_m$, $m \in \mathcal{A}$ are drawn i.i.d. uniformly on $\mathcal{S}^{N-1}$. Here $\overset{\circ}{I}_G(\mathbf{x}_{\mathcal{A}};\mathbf{y}\mathbf{x}_{\mathcal{B}})$ is obtained from (16).

Lemma 6: Let $L \geq 2$. Consider an $L \times L$ normalized correlation matrix $\mathbf{R}$ and the subset $\Omega(\mathbf{R})$ of the L-fold Cartesian product of $\mathcal{S}^{N-1}$ whose elements have normalized correlation matrix $\mathbf{R}$:

$$\Omega(\mathbf{R}) = \{(\mathbf{x}_1,\cdots,\mathbf{x}_L) \ : \ \mathbf{x}_k\cdot\mathbf{x}_l = R_{kl},\ 1\leq k,l\leq L\}. \tag{29}$$

Then

$$\frac{1}{N}\log\frac{|\Omega(\mathbf{R})|}{|\mathcal{S}^{N-1}|^L} \sim \frac{1}{2}\log|\mathbf{R}|, \quad \text{as } N\to\infty.$$

In the special case $L=2$, the determinant $|\mathbf{R}| = 1-R_{12}^2$. Given any $\mathbf{x}_1 \in \mathcal{S}^{N-1}$, the constraint $\mathbf{x}_1\cdot\mathbf{x}_2 = R_{12}$ defines a hyperplane whose intersection with $\mathcal{S}^{N-1}$ is a $N-1$ sphere of radius $\sqrt{1-R_{12}^2}$, centered at $R_{12}\mathbf{x}_1$. The proof of the lemma generalizes this geometric approach.

Corollary 7: Define the following subset of the L-fold Cartesian product of $\mathcal{S}^{N-1}$:

$$\Omega_{cap}(\nu) = \{(\mathbf{x}_1,\cdots,\mathbf{x}_L)\in\Omega(\mathbf{R}) \quad \text{for some } \mathbf{R} \text{ s.t.} \quad -\frac{1}{2}\log|\mathbf{R}| \geq \nu\}, \quad \nu > 0. \tag{30}$$

Then

$$\frac{1}{N}\log\frac{|\Omega_{cap}(\nu)|}{|\mathcal{S}^{N-1}|^L} \sim -\nu, \quad \text{as } N\to\infty.$$

This follows from Lemma 6 and the compactness of the set of normalized correlation matrices. The result is analogous to Sanov's theorem in large-deviations theory [14].

Lemma 8: Fix $\mathbf{y} \in \mathbb{R}^N$ and draw $\mathbf{x}_1,\cdots,\mathbf{x}_K$ i.i.d. uniformly from $\mathcal{S}^{N-1}$. Then

$$Pr[\overset{\circ}{I}_G(\mathbf{x}_{\mathcal{K}};\mathbf{y}) > \nu] \dot{\leq} 2^{-N\nu} \tag{31}$$

Proof. Let $L = K+1$ and $\mathbf{R}$ be the $L\times L$ normalized correlation matrix for $\mathbf{x}_1,\cdots,\mathbf{x}_K,\mathbf{z}$. Hence $\overset{\circ}{I}_G(\mathbf{x}_{\mathcal{K}};\mathbf{z}) = -\frac{1}{2}\log|\mathbf{R}|$, and the claim follows directly by application of Corollary 7. The result is analogous to [4, Eq. (10.12)].

Similarly, fix $\mathbf{y}$ and $\mathbf{x}_{\mathcal{B}}$ and draw $\mathbf{x}_1,\cdots,\mathbf{x}_{|\mathcal{A}|}$ i.i.d. uniformly from the unit sphere. Then

$$Pr[\overset{\circ}{I}_G(\mathbf{x}_{\mathcal{A}};\mathbf{y}\mathbf{x}_{\mathcal{B}}) > \nu] \dot{\leq} 2^{-N\nu}. \tag{32}$$

Substituting (31) and (32) into (28) yields $P_{\rm FP}(\mathbf{x}_{\mathcal{K}},\mathbf{y}) \dot{\leq} 2^{-N\Delta}$ as in [4, p. 49]. Integrating over $\mathbf{x}_{\mathcal{K}}$ and $\mathbf{y}$ yields $P_{\rm FP} \dot{\leq} 2^{-N\Delta}$.

Lemma 9: Given $0\leq\mu\leq 1$, let $\mathbf{x}_1,\cdots,\mathbf{x}_K \in \mathcal{S}^{N-1}$ form a regular simplex constellation centered at

$$\overline{\mathbf{x}} = \left(\underbrace{\frac{\mu}{\sqrt{K}},\cdots,\frac{\mu}{\sqrt{K}}}_{K \text{ times}},0,\cdots,0\right),$$

hence $\|\overline{\mathbf{x}}\| = |\mu| \leq 1$. The determinant of the empirical normalized correlation matrix for $\mathbf{x}_1,\cdots,\mathbf{x}_K$ is given by

$$|\mathbf{R}| = K\mu^2\left(\frac{K(1-\mu^2)}{K-1}\right)^{K-1} \tag{33}$$

and the corresponding empirical Gaussian mutual information by

$$\overset{\circ}{I}{}^*_G(\mu) = -\frac{1}{2}\log|\mathbf{R}| = -\log(\mu\sqrt{K}) - \frac{K-1}{2}\log\frac{K(1-\mu^2)}{K-1}. \tag{34}$$

Any constellation with the same mean vector length $\|\overline{\mathbf{x}}\| = \mu \geq 0$ satisfies

$$\overset{\circ}{I}_G(\mathbf{x}_{\mathcal{K}}) \geq \overset{\circ}{I}{}^*_G(\mu). \tag{35}$$

The function $\overset{\circ}{I}{}^*_G(\mu)$ is unimodal with a minimum equal to 0 achieved at $\mu = 1/\sqrt{K}$, and tends to ∞ for $\mu \to 0$ and $\mu \to 1$.

Corollary 10: Any constellation $\mathbf{x}_1, \cdots, \mathbf{x}_K \in \mathcal{S}^{N-1}$ such that $\overset{\circ}{I}_G(\mathbf{x}_{\mathcal{K}}) = \nu \geq 0$ satisfies $\mu_1 \leq \|\overline{\mathbf{x}}\| \leq \mu_2$ where $\mu_1 \in (0, \frac{1}{\sqrt{K}}]$ and $\mu_2 \in [\frac{1}{\sqrt{K}}, 1)$ are the solutions of $\overset{\circ}{I}{}^*_G(\cdot) = \nu$. Equality is achieved by the shifted regular simplex constellation.

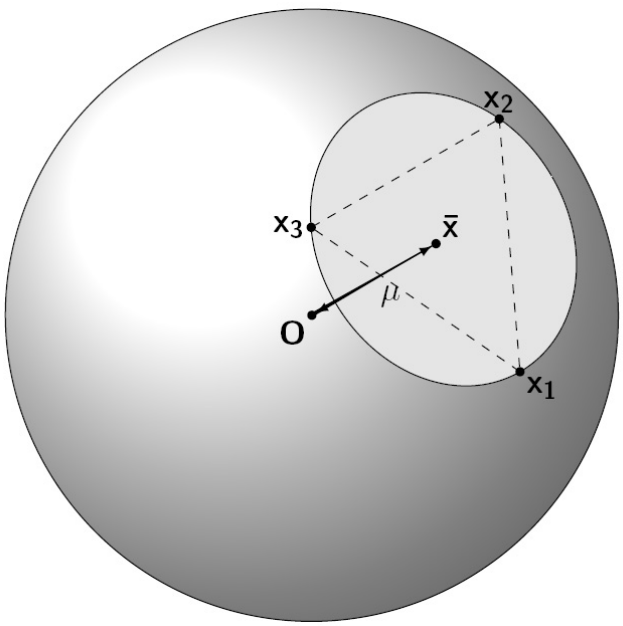

Fig. 2. The shifted simplex constellation of Lemma 9 for $K = 3$. The points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ form an equilateral triangle centered at $\overline{\mathbf{x}}$. Here the fingerprints are positively correlated: $\|\overline{\mathbf{x}}\| = \mu > 1/\sqrt{3}$.

Special cases of Lemma 9 include the orthogonal constellation, in which case $\mu = 1/\sqrt{K}$ and $|\mathbf{R}| = 1$; the simplex constellation, in which case $\mu = 0$ and $|\mathbf{R}| = 0$ (the matrix $\mathbf{R}$ has rank $K-1$), and the degenerate constellation consisting of identical vectors, in which case $\mu = 1$ and $|\mathbf{R}| = 0$ (the matrix $\mathbf{R}$ has rank 1). Also note that for $K = 2$ we have

$$|\mathbf{R}| = 1 - \rho^2(\mathbf{x}_1, \mathbf{x}_2) = 4\mu^2(1-\mu^2).$$

By analogy with the case $K = 2$, it will be useful to think of $0 \leq \mu < 1/\sqrt{K}$ and $1/\sqrt{K} < \mu \leq 1$ as corresponding to "negatively correlated" and "positively correlated" fingerprints, respectively. As we shall see, positively correlated fingerprints are easier to decode than negatively correlated fingerprints. The case $K = 3$, $\mu > 1/\sqrt{3}$ is depicted in Fig. 2.

Proposition 11: The MPGMI decoder of (20) achieves the following error exponents:

$$E_{\mathrm{FP}}(R, \Delta, K) = \Delta \tag{36}$$

$$E^{\mathrm{one}}(R, \Delta, K) = \min_{\underline{\mu} \leq \mu \leq 1/\sqrt{K}} \left[\overset{\circ}{I}{}^*_G(\mu) + E_{sp}\left(\mu^2 \frac{D_{\mathrm{f}}}{D_{\mathrm{c}}}, K(R+\Delta) - \overset{\circ}{I}{}^*_G(\mu)\right)\right]. \tag{37}$$

where $\underline{\mu}$ is the unique solution of $\overset{\circ}{I}{}^*_G(\mu) = K(R+\Delta)$ for $0 \leq \mu \leq 1/\sqrt{K}$. Moreover

- $E^{\mathrm{one}}(R, \Delta, K) > 0$ if and only if $R + \Delta < C(K)$;
- the supremum of all R for which the error exponents are positive is $C(K)$ in (6) and is achieved by letting $\Delta \to 0$;
- for any fair coalition, $E^{\mathrm{all}}(R, \Delta, K) = E^{\mathrm{one}}(R, \Delta, K)$.

The false-negative exponent formula (37) may be interpreted by viewing the error event as the intersection of two events: (1) the colluders' fingerprints $\mathbf{x}_{\mathcal{K}}$ may be correlated, in which case the length μ of the mean vector $\overline{\mathbf{x}}$ differs from the typical value $1/\sqrt{K}$; and (2) the component of the noise vector $\mathbf{w}$ along the direction of $\overline{\mathbf{x}}$ is significant. Each of these two events has exponentially vanishing probability, and each exponent is a function of μ. The minimand in (37) is the sum of these two exponents, and the typical value of $\|\overline{\mathbf{x}}\|$ for the error event is the minimizing μ. This minimizing value is below $1/\sqrt{K}$, i.e., the fingerprints are negatively correlated.

To illustrate this analysis, consider a fingerprinting system with 2^{20} ($\sim$ one million) users and a target false-positive error probability of $2^{-12} \approx 10^{-4}$. Hence $NR = 20$ and $N\Delta = 12$. These requirements are met using a random spherical fingerprinting code of *approximate length* $N = 20,000$ and rate $R = 10^{-3}$, and fixing our universal decoder's parameter at $\Delta = 6 \times 10^{-4}$. The error exponents (36) and (37) are plotted as a function of K in Fig. 3, for SNR = $D_{\mathrm{f}}/D_{\mathrm{c}} = 1$. While $E_{\mathrm{FP}} = \Delta$ is independent of K, E^{one} decreases with K. For $K = 2$ we have $E^{\mathrm{one}} = 0.3986$; for $K = 18$ we have $E^{\mathrm{one}} = 0.0008$; and for $K \geq 35$ we have $E^{\mathrm{one}} = 0$. Hence this random spherical code can resist 18 colluders with both P_{FP} and $P_e^{\mathrm{one}} \leq 10^{-4}$.

The minimizing values of μ in (37) when $K = 2$ and $K = 34$ are equal to $\underline{\mu} = 0.6975$ and $\mu = 0.1715$, respectively. Equivalently, $\mu\sqrt{K} = 0.9865$ and 0.99998, respectively. In both cases, the minimizing μ is close to the typical value.

Fig. 4 shows the maximum achievable rate

$$R_{\max}(K, \Delta) = C(K) - \Delta$$

for $\Delta = 0.0006$. Note that $R_{\max}(K, 0.0006) = 0$ for $K \geq 35$ while $C(K)$ is asymptotic to $0.72/K^2$ for large K (see Sec. XI).

X. Analysis of $\widetilde{MPGMI}$ Decoder

The $\widetilde{MPGMI}$ decoder is capacity-achieving. However its false-negative error exponent is zero because it limits its search to typical coalitions.

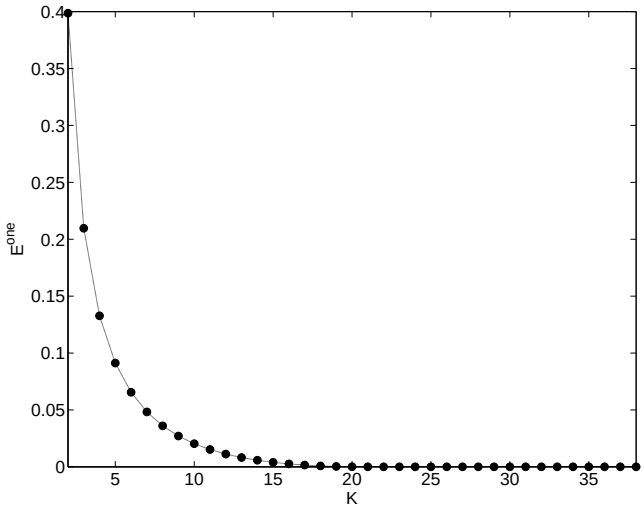

Fig. 3. False-negative error exponent $E^{\rm one}(R,\Delta,K)$ as a function of coalition size K, for fingerprinting code rate $R = 10^{-3}$, false-positive exponent $E_{\rm FP} = \Delta = 0.0006$, and SNR = $D_{\rm f}/D_{\rm c} = 1$.

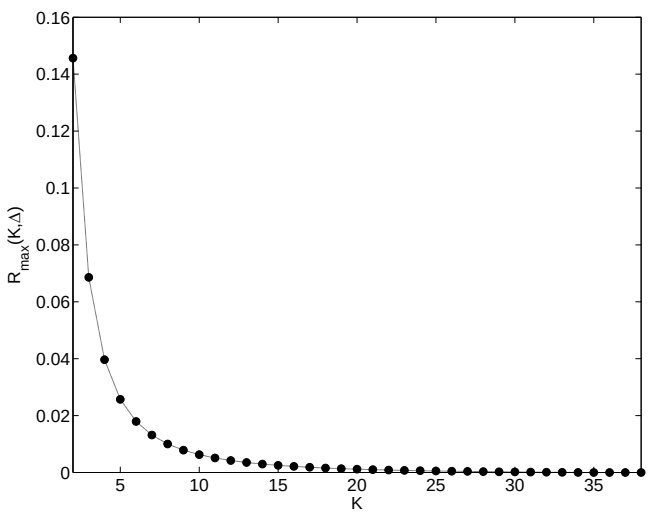

Fig. 4. Maximum achievable rate $R_{\max}(K,\Delta) = C(K) - \Delta$ (for false-positive error exponent $\Delta = 0.0006$) as a function of coalition size K. Here SNR = $D_{\rm f}/D_{\rm c} = 1$.

XI. Large Coalitions

Comparing $C_1(K) = \frac{1}{2}\log\left(1 + \frac{D_{\rm f}}{D_{\rm c}K^2 + D_{\rm f}(K-1)}\right)$ of (27) with $C(K) = \frac{1}{2K}\log\left(1 + \frac{D_{\rm f}}{KD_{\rm c}}\right)$ of (6), we observe that both asymptotically approach $\frac{D_{\rm f}}{2(\ln 2)K^2 D_{\rm c}}$ from below as $K \to \infty$. Hence, for large K, the simple decoder is nearly as good as the more complex joint decoder.

XII. Discussion

Closed-form solutions have been obtained for the capacity and random-coding exponents of a fingerprinting system subject to almost-sure squared distortion constraints. This study extends recent work presented in [4] for finite alphabets. Here the decoding metric is based on the determinant of the empirical normalized correlation matrix. The encoder and decoder do not need to know anything about the collusion channel and the coalition size K. Note that in the finite-alphabet case [4], the encoder needs to assume a nominal value for the coalition size, and a time-sharing random variable is used.

Acknowledgment. The authors are thankful to Yen-Wei Huang for preparing Figs. 2—4. This work was supported by NSF under grants CCR 03-25924, CCF 06-35137, and CCF 07-29061.

References

[1] I. J. Cox, J. Killian, F. T. Leighton and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE T-IP*, Vol. 6, pp. 1673—1687, Dec. 1997. (Also NEC Tech. Rep. 95-10, 1995).

[2] P. Moulin and J. A. O'Sullivan, "Information-Theoretic Analysis of Information Hiding," *IEEE T-IT*, Vol. 49, No. 3, pp. 563—593, 2003.

[3] A. Somekh-Baruch and N. Merhav, "On the Capacity Game of Private Fingerprinting Systems Under Collusion Attacks," *Proc. IEEE Int. Symp. on Information Theory*, Yokohama, Japan, p. 191, July 2003.

[4] P. Moulin, "Universal Fingerprinting: Capacity and Random Coding Exponents," *preprint*, Jan. 2008, revised Dec. 2008, available from `arXiv:0801.3837v2 [cs:IT]`. Short version in *Proc. ISIT*, Toronto, Canada, July 2008.

[5] J. Kilian, F. T. Leighton, L. R. Matheson, T. G. Shamoon, R. E. Tarjan, and F. Zane, "Resistance of digital watermarks to collusive attacks," *Proc. ISIT*, p. 271, Cambridge, MA, 1998.

[6] P. Moulin and A. Briassouli, "The Gaussian Fingerprinting Game," *Proc. CISS'02*, Princeton, NJ, March 2002.

[7] H. S. Stone, "Analysis of Attacks on Image Watermarks With Randomized Coefficients," *NEC TR 96-045*, Princeton, NJ, 1996.

[8] H. Zhao, M. Wu, Z. Wang and K. J. R. Liu, "Forensic Analysis of Nonlinear Collusion Attacks for Multimedia Fingerprinting," *IEEE T-IP*, Vol. 14, No. 5, pp. 646—661, May 2005.

[9] P. Moulin and Y. Wang, "The Spherical Fingerprinting Game," in preparation, to be posted in `Information Theory arxiv`, Feb. 2009.

[10] N. Kiyavash and P. Moulin, "A Framework for Optimizing Nonlinear Collusion Attacks on Fingerprinting Systems," *Proc. Conf. on Information Systems and Science*, Princeton, NJ, March 2006.

[11] P. Moulin and N. Kiyavash, "Performance of Random Fingerprinting Codes Under Arbitrary Nonlinear Collusion Attacks," *Proc. ICASSP*, Hawaii, Apr. 2007.

[12] N. Kiyavash and P. Moulin, "On Optimal Collusion Strategies for Fingerprinting," *Proc. ICASSP*, Toulouse, France, May 2006.

[13] C. Swannack and G. W. Wornell, "Reflections on the AWGN Error Exponent," *Proc. Allerton Conference*, Monticello, IL, Sep. 2005.

[14] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed., Springer-Verlag, New York, 1998.