Fingerprinting With Equiangular Tight Frames

Dustin G. Mixon, Christopher J. Quinn, *Student Member, IEEE*, Negar Kiyavash, *Member, IEEE*, and Matthew Fickus, *Member, IEEE*

***Abstract*—Digital fingerprinting is a framework for marking media files, such as images, music, or movies, with user-specific signatures to deter illegal distribution. Multiple users can collude to produce a forgery that can potentially overcome a fingerprinting system. This paper proposes an equiangular tight frame fingerprint design which is robust to such collusion attacks. We motivate this design by considering digital fingerprinting in terms of compressed sensing. The attack is modeled as linear averaging of multiple marked copies before adding a Gaussian noise vector. The content owner can then determine guilt by exploiting correlation between each user's fingerprint and the forged copy. The worst case error probability of this detection scheme is analyzed and bounded. Simulation results demonstrate that the average-case performance is similar to the performance of orthogonal and simplex fingerprint designs, while accommodating several times as many users.**

***Index Terms*—Collusion attacks, digital fingerprinting, frames.**

I. Introduction

DIGITAL media protection has become an important issue in recent years, as illegal distribution of licensed material has become increasingly prevalent. A number of methods have been proposed to restrict illegal distribution of media and ensure only licensed users are able to access it. One method involves cryptographic techniques, which encrypt the media before distribution. By doing this, only the users with appropriate licensed hardware or software have access; satellite TV and DVDs are two such examples. Unfortunately, cryptographic approaches are limited in that once the content is decrypted (legally or illegally), it can potentially be copied and distributed freely.

An alternate approach involves marking each copy of the media with a unique signature. The signature could be a change in the bit sequence of the digital file or some noise-like distortion of the media. The unique signatures are called *fingerprints*, by analogy to the uniqueness of human fingerprints. With this approach, a licensed user could illegally distribute the file, only to be implicated by his fingerprint. The potential for prosecution acts as a deterrent to unauthorized distribution.

Fingerprinting can be an effective technique for inhibiting individual licensed users from distributing their copies of the media. However, fingerprinting systems are vulnerable when multiple users form a *collusion* by combining their copies to create a forged copy. This attack can reduce and distort the colluders' individual fingerprints, making identification of any particular user difficult. Some examples of potential attacks involve comparing the bit sequences of different copies, averaging copies in the signal space, as well as introducing distortion (such as noise, rotation, or cropping).

There are two principal approaches to designing fingerprints with robustness to collusions. The first approach uses the *marking* assumption [1]: that the forged copy only differs from the colluders' copies where the colluders' copies are different (typically in the bit sequence). In many cases, this is a reasonable assumption because modifying other bits would likely render the file unusable (such as with software).

Boneh and Shaw proposed the first-known fingerprint design that uses the marking assumption to identify a member of the collusion with high probability [1]. Boneh and Shaw's method incorporates the results of Chor *et al.*, who investigated how to detect users who illegally share keys for encrypted material [2]. Schaathun later showed that the Boneh–Shaw scheme is more efficient than initially thought and proposed further improvements [3]. Tardos also proposed a method with significantly shorter code lengths than those of the Boneh–Shaw procedure [4]. Several recent works investigate the relationship between the fingerprinting problem and multiple access channels, and they calculate the capacity of a "fingerprint channel" [5]–[8]. Barg *et al.* also developed "parent-identifying" codes under a relaxation of the marking assumption [9], and there have been a number of other works developing special classes of binary fingerprinting codes, including [10]–[15].

The second major approach uses the *distortion* assumption—this is the assumption we will use. In this regime, fingerprints are noise-like distortions to the media in signal space. In order to preserve the overall quality of the media, limits are placed on the magnitude of this distortion. The content owner limits the power of the fingerprint he adds, and the collusion limits the power of the noise they add in their attack. When applying the distortion assumption, the literature typically assumes that the collusion linearly averages their individual copies to forge the host signal. Also, while results in this domain tend to accommodate fewer users than those with the marking assumption, the distortion assumption enables a

Manuscript received November 14, 2011; revised August 28, 2012; accepted November 11, 2012. Date of publication January 25, 2013; date of current version February 12, 2013. This work was supported in part by the National Science Foundation under Grant DMS 1042701, in part by the Air Force Office of Scientific Research under Grants F1ATA00083G004, F1ATA00183G003, and FA9550-10-1-0345, and in part by the Naval Research Laboratory under Grant N00173-09-1-G033. D. G. Mixon was supported by the A. B. Krongard Fellowship. C. J. Quinn was supported by the Department of Energy Computational Science Graduate Fellowship under Grant DE-FG02-97ER25308. This paper was presented in part at the 2011 IEEE International Conference on Acoustics, Speech, and Signal Processing and in part at the 2011 International Society for Optical Engineers Conference.

D. G. Mixon and M. Fickus are with the Department of Mathematics and Statistics, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433 USA (e-mail: dustin.mixon@afit.edu; matthew.fickus@afit.edu).

C. J. Quinn is with the Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL 61801 USA (e-mail: quinn7@illinois.edu).

N. Kiyavash is with the Department of Industrial and Enterprise Systems Engineering, University of Illinois, Urbana, IL 61801 USA (e-mail: kiyavash@illinois.edu).

Communicated by D. Guo, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2012.2229781

0018-9448/$31.00 © 2013 IEEE

more natural embedding of the fingerprints, i.e., in the signal space.

Cox *et al.* introduced one of the first robust fingerprint designs under the distortion assumption [16]; the robustness was later analytically proven in [17]. Ergun *et al.* then showed that for any fingerprinting system, there is a tradeoff between the probabilities of successful detection and false positives imposed by a linear-average-plus-noise attack from sufficiently large collusions [18]. Specific fingerprint designs were later studied, including orthogonal fingerprints [19] and simplex fingerprints [20]. There are also some proposed methods motivated by CDMA techniques [21], [22]. Jourdas and Moulin demonstrated that a high rate can be achieved by embedding randomly concatenated, independent short codes [23]. Kiyavash and Moulin derived a lower bound on the worst case error probabilities for designs with equal-energy fingerprints [24]. An error probability analysis of general fingerprint designs with unequal priors on user collusion is given in [25]. Some works have also investigated fingerprint performance under attack strategies other than linear averaging [26], [27] and under alternative noise models [28].

When working with distortion-type fingerprints, some embedding process is typically needed. This process is intended to make it difficult for the colluders to identify or distort fingerprints without significantly damaging the corresponding media in the process. For instance, if the identity basis is used as an orthogonal fingerprint design, then three or more colluders can easily identify and remove their fingerprints through comparisons. Indeed, different signal dimensions have different perceptual significance to the human user, and one should vary signal power strengths accordingly. There is a large body of work in this area known as *watermarking*, and this literature is relevant since fingerprinting assigns a distinct watermark to each user. As an example, one method of fingerprint embedding is known as spread spectrum watermarking. Inspired by spread spectrum communications [29], this technique rotates the fingerprint basis to be distributed across the perceptually significant dimensions of the signal [16]. This makes fingerprint removal difficult while at the same time maintaining acceptable fidelity. For an overview of watermarking media, see [30] and [31].

This paper proposes equiangular tight frames (ETFs) for fingerprint design under the distortion assumption. To be clear, a *frame* is a collection of vectors $\{f_m\}_{m=1}^M \subseteq \mathbb{R}^N$ with *frame bounds* $0 < A \leq B < \infty$ that satisfy

$$A\|x\|^2 \leq \sum_{m=1}^{M} |\langle x, f_m\rangle|^2 \leq B\|x\|^2$$

for every $x \in \mathbb{R}^N$. If each f_m has unit norm, we say $\{f_m\}_{m=1}^M$ forms a unit norm frame. In the special case where $A = B$, the collection is said to be a *tight frame*. This leads to the following definition.

Definition 1: An ETF is a unit norm tight frame with the additional property that there exists a constant c such that

$$|\langle f_m, f_{m'}\rangle| = c$$

for every pair $m \neq m'$.

In considering ETFs for fingerprint design, we analyze their performance against the worst case collusion. Moreover, through simulations, we show that these fingerprints perform comparably to orthogonal and simplex fingerprints on average, while accommodating several times as many users. Li and Trappe [22] also used fingerprints satisfying the Welch bound, but did not use tight frames, designed the decoder to return the whole collusion, and did not perform worst case analysis. Use of compressed sensing for fingerprint design was first suggested in [32] and [33], followed by [34]. However, the work in [33] and [34] was focused on detection schemes with Gaussian fingerprints. Additionally, Colbourn *et al.* [35] examined combinatorial similarities of fingerprint codes and compressed sensing measurement matrices.

We present the description of the fingerprinting problem in Section II. In Section III, we discuss this problem from a compressed sensing viewpoint and introduce the ETF fingerprint design. Using this design, we consider a detector which determines guilt for each user through binary hypothesis testing, using the correlation between the forged copy and the user's fingerprint as a test statistic. In Section IV, we derive bounds on the worst case error probability for this detection scheme assuming a linear-average-plus-noise attack. Finally, in Section V, we provide simulations that demonstrate the average-case performance in comparison to orthogonal and simplex fingerprint designs.

II. Problem Setup

In this section, we describe the fingerprinting problem and discuss the performance criteria we will use in this paper. We start with the model we use for the fingerprinting and attack processes.

A. Mathematical Model

A content owner has a host signal that he wishes to share, but he wants to mark it with fingerprints before distributing. We view this host signal as a vector $s \in \mathbb{R}^N$, and the marked versions of this vector will be given to $M > N$ users. Specifically, the mth user is given by

$$x_m = s + f_m \tag{1}$$

where $f_m \in \mathbb{R}^N$ denotes the mth fingerprint. We assume the fingerprints have equal energy

$$\gamma^2 := \|f_m\|^2 = ND_f \tag{2}$$

i.e., D_f denotes the average energy per dimension of each fingerprint.

We wish to design the fingerprints $\{f_m\}_{m=1}^M$ to be robust to a linear averaging attack. In particular, let $\mathcal{K} \subseteq \{1, \ldots, M\}$ denote a group of users who together forge a copy of the host signal. Then, their linear averaging attack is of the form

$$y = \sum_{k\in\mathcal{K}} \alpha_k(s + f_k) + \epsilon, \qquad \sum_{k\in\mathcal{K}} \alpha_k = 1 \tag{3}$$

where ϵ is a noise vector introduced by the colluders. We assume ϵ is Gaussian noise with mean zero and variance $N\sigma^2$, i.e., σ^2

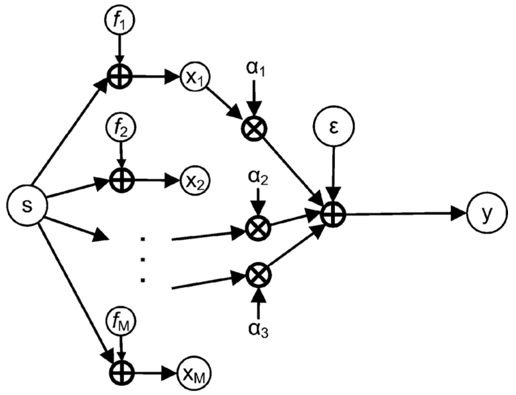


Fig. 1. Fingerprint assignment (1) and forgery attack (3) processes.

is the noise power per dimension. The relative strength of the attack noise is measured as the *watermark-to-noise* ratio (WNR)

$$\mathrm{WNR} := 10\log_{10}\Big(\frac{ND_f}{N\sigma^2}\Big). \tag{4}$$

This is analogous to the signal-to-noise ratio. See Fig. 1 for a schematic of this attack model.

B. Detection

Certainly, the ultimate goal of the content owner is to detect every member in a forgery coalition. This can prove difficult in practice, though, particularly when some individuals contribute little to the forgery, with $\alpha_k \ll 1$. However, in the real world, if at least one colluder is caught, then other members could be identified through the legal process. As such, we consider *focused* detection, where a test statistic is computed for each user, and we perform a binary hypothesis test for each user to decide whether that particular user is guilty.

With the cooperation of the content owner, the host signal can be subtracted from a forgery to isolate the fingerprint combination

$$z := \sum_{k\in\mathcal{K}} \alpha_k f_k + \epsilon.$$

We refer to $\sum_{k\in\mathcal{K}} \alpha_k f_k$ as the *noiseless* fingerprint combination. The test statistic for each user m is the normalized correlation function

$$T_m(z) := \frac{1}{\gamma^2}\langle z, f_m\rangle \tag{5}$$

where γ^2 is the fingerprint energy (2). For each user m, let $H_1(m)$ denote the guilty hypothesis $(m \in \mathcal{K})$ and $H_0(m)$ denote the innocent hypothesis $(m \notin \mathcal{K})$. Letting τ denote a correlation threshold, we use the following detector:

$$\delta_m(\tau) := \begin{cases} H_1(m), & T_m(z) \geq \tau \\ H_0(m), & T_m(z) < \tau. \end{cases} \tag{6}$$

To determine the effectiveness of our fingerprint design and focused detector, we will investigate the corresponding error probabilities.

C. Error Analysis

Due in part to the noise that the coalition introduced to the forgery, there could be errors associated with our detection method. One type of error we can expect is the false-positive error, whose probability is denoted $\mathrm{P_I}$, in which an innocent user m $(m \notin \mathcal{K})$ is found guilty $(T_m(z) \geq \tau)$. This could have significant ramifications in legal proceedings, so this error probability should be kept extremely low. The other error type is the false-negative error, whose probability is denoted $\mathrm{P_{II}}$, in which a guilty user $(m \in \mathcal{K})$ is found innocent $(T_m(z) < \tau)$. The probabilities of these two errors depend on the fingerprints $F = \{f_m\}_{m=1}^M$, the coalition $\mathcal{K}$, the weights $\alpha = \{\alpha_k\}_{k\in\mathcal{K}}$, the user m, and the threshold τ

$$\begin{aligned} \mathrm{P_I}(F, m, \tau, \mathcal{K}, \alpha) &:= \mathrm{Prob}[T_m(z) \geq \tau | H_0(m)] \\ \mathrm{P_{II}}(F, m, \tau, \mathcal{K}, \alpha) &:= \mathrm{Prob}[T_m(z) < \tau | H_1(m)]. \end{aligned}$$

We will characterize the *worst case* error probabilities over all possible coalitions and users.

We first define the probability of a "false alarm"

$$\mathrm{P_{fa}}(F, \tau, \mathcal{K}, \alpha) := \max_{m\notin\mathcal{K}} \mathrm{P_I}(F, m, \tau, \mathcal{K}, \alpha). \tag{7}$$

This is the probability of wrongly accusing the innocent user who looks most guilty. Equivalently, this is the probability of accusing at least one innocent user. The worst case type I error probability is given by

$$\mathrm{P_I}(F, \tau, \alpha) := \max_{\mathcal{K}} \mathrm{P_{fa}}(F, \tau, \mathcal{K}, \alpha). \tag{8}$$

Next, consider the probability of a "miss"

$$\mathrm{P}_m(F, \tau, \mathcal{K}, \alpha) := \min_{m\in\mathcal{K}} \mathrm{P_{II}}(F, m, \tau, \mathcal{K}, \alpha).$$

This is the probability of not accusing the most vulnerable guilty user. Equivalently, this is the probability of not detecting any colluders. Note that this event is the opposite of detecting at least one colluder

$$\mathrm{P_d}(F, \tau, \mathcal{K}, \alpha) := 1 - \mathrm{P}_m(F, \tau, \mathcal{K}, \alpha). \tag{9}$$

The worst case type II error probability is given by

$$\mathrm{P_{II}}(F, \tau, \alpha) := \max_{\mathcal{K}} \mathrm{P}_m(F, \tau, \mathcal{K}, \alpha). \tag{10}$$

The *worst case* error probability is the maximum of the two error probabilities (8) and (10)

$$\mathrm{P}_e(F, \tau, \alpha) := \max\big\{\mathrm{P_I}(F, \tau, \alpha), \mathrm{P_{II}}(F, \tau, \alpha)\big\}.$$

The threshold parameter τ can be varied to minimize this quantity, yielding the minmax error probability

$$\mathrm{P_{minmax}}(F, \alpha) := \min_{\tau} \mathrm{P}_e(F, \tau, \alpha). \tag{11}$$

In Section IV, we will analyze these error probabilities. We will also investigate average-case performance using simulations in Section V.

III. ETF Fingerprint Design

In this section, we introduce a fingerprint design based on ETFs. Before we discuss the design, we first motivate it in terms of a certain geometric figure of merit that originated in [20] to evaluate simplex fingerprints.

A. Geometric Figure of Merit

For each user m, consider the distance between two types of potential collusions: those of which m is a member, and those of which m is not. Intuitively, if every noiseless fingerprint combination involving m is distant from every fingerprint combination not involving m, then even with moderate noise, there should be little ambiguity as to whether the mth user was involved or not.

To make this precise, for each user m, we define the "guilty" and "not guilty" sets of noiseless fingerprint combinations

$$\mathcal{G}_m^{(K)} := \left\{ \frac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}} f_k : m \in \mathcal{K} \subseteq \{1,\ldots,M\}, |\mathcal{K}| \le K \right\}$$

$$\neg\mathcal{G}_m^{(K)} := \left\{ \frac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}} f_k : m \notin \mathcal{K} \subseteq \{1,\ldots,M\}, |\mathcal{K}| \le K \right\}.$$

In words, $\mathcal{G}_m^{(K)}$ is the set of size-K fingerprint combinations of equal weights ($\alpha_k = \frac{1}{K}$) which include m, while $\neg\mathcal{G}_m^{(K)}$ is the set of combinations which do not include m. Note that in our setup (3), the α_k's were arbitrary nonnegative values bounded by one. We will show in Section IV that the best attack from the collusion's perspective uses equal weights ($\alpha_k = \frac{1}{K}$) so that no single colluder is particularly vulnerable. For this reason, we use equal weights to obtain bounds on the distance between these two sets, which we define to be

$$\begin{aligned}&\mathrm{dist}(\mathcal{G}_m^{(K)}, \neg\mathcal{G}_m^{(K)}) \\ &\quad := \min\{\|x-y\| : x \in \mathcal{G}_m^{(K)}, y \in \neg\mathcal{G}_m^{(K)}\}. \qquad (12)\end{aligned}$$

In this section, we will find a lower bound on this distance by exploiting specially designed properties of the fingerprints.

B. Applying the Restricted Isometry Property (RIP)

In this section, we introduce some ideas from the compressed sensing literature. To motivate our use of these ideas, we take a brief detour from our goal of focused detection. For the moment, suppose we wish to identify the entire collusion, as opposed to just one member of the collusion. To do this, we consider a matrix-vector formulation of the attack (3) without noise, i.e., with $\epsilon = 0$. Specifically, let F denote the $N \times M$ matrix whose columns are the fingerprints $\{f_m\}_{m=1}^M$, and let the $M \times 1$ vector α denote the weights used in the collusion's linear average. Note that α_m is zero if user m is innocent; otherwise, it is given by the corresponding coefficient in (3). This gives

$$z = F\alpha.$$

Using this representation, the detection problem can be interpreted from a compressed sensing perspective. Namely, α is a K-sparse vector that we wish to recover. Under certain conditions on the matrix F, we may "sense" this vector with $N < M$ measurements in such a way that the K-sparse vector is recoverable; this is the main idea behind compressed sensing. In particular, the vector is recoverable when the matrix satisfies the following matrix property.

Definition 2: We say an $N \times M$ matrix F satisfies the RIP if there exists a constant $\delta_{2K} < \sqrt{2} - 1$ such that

$$(1-\delta_{2K})\|\alpha\|^2 \le \|F\alpha\|^2 \le (1+\delta_{2K})\|\alpha\|^2 \qquad (13)$$

for every $2K$-sparse vector $\alpha \in \mathbb{R}^M$.

The RIP was introduced in [36] to recover sparse vectors using linear programming. Since then, RIP has proven to be particularly ubiquitous across compressed sensing, enabling reconstruction guarantees for a variety of alternatives to linear programming, see, e.g., [37] and [38]. Thus, if F satisfies RIP (13), we can recover the K-sparse vector $\alpha \in \mathbb{R}^M$ using any of these techniques. In fact, for the attack model with adversarial noise, $z = F\alpha + \epsilon$, if F satisfies RIP (13), linear programming will still produce an estimate $\hat{\alpha}$ of the sparse vector [39]. However, the distance between $\hat{\alpha}$ and α will be on the order of ten times the size of the error ϵ. Due to potential legal ramifications of false accusations, this order of error is not tolerable. This is precisely the reason we prefer focused detection (6) over identifying the entire collusion.

Regardless, we can evaluate fingerprints with the RIP in terms of our geometric figure of merit. Without loss of generality, we assume the fingerprints are unit norm; since they have equal energy γ^2, the fingerprint combination z can be normalized by γ before the detection phase. With this in mind, we have the following lower bound on the distance (12) between the "guilty" and "not guilty" sets corresponding to any user m.

Theorem 3: Suppose fingerprints $F = [f_1, \ldots, f_M]$ satisfy the RIP (13). Then

$$\mathrm{dist}(\mathcal{G}_m^{(K)}, \neg\mathcal{G}_m^{(K)}) \ge \sqrt{\frac{1-\delta_{2K}}{K(K-1)}}. \qquad (14)$$

Proof: Take $\mathcal{K}, \mathcal{K}' \subseteq \{1,\ldots,M\}$ such that $|\mathcal{K}|, |\mathcal{K}'| \le K$ and $m \in \mathcal{K} \setminus \mathcal{K}'$. Then, the left-hand inequality of the RIP (13) gives

$$\begin{aligned}&\left\| \frac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}} f_k - \frac{1}{|\mathcal{K}'|}\sum_{k\in\mathcal{K}'} f_k \right\|^2 \\ &= \left\| \left(\tfrac{1}{|\mathcal{K}|} - \tfrac{1}{|\mathcal{K}'|}\right)\sum_{k\in\mathcal{K}\cap\mathcal{K}'} f_k + \tfrac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}\setminus\mathcal{K}'} f_k - \tfrac{1}{|\mathcal{K}'|}\sum_{k\in\mathcal{K}'\setminus\mathcal{K}} f_k \right\|^2 \\ &\ge (1-\delta_{|\mathcal{K}\cup\mathcal{K}'|})\left(|\mathcal{K}\cap\mathcal{K}'|\left(\tfrac{1}{|\mathcal{K}|} - \tfrac{1}{|\mathcal{K}'|}\right)^2 + \tfrac{|\mathcal{K}\setminus\mathcal{K}'|}{|\mathcal{K}|^2} + \tfrac{|\mathcal{K}'\setminus\mathcal{K}|}{|\mathcal{K}'|^2} \right) \\ &= \frac{1-\delta_{|\mathcal{K}\cup\mathcal{K}'|}}{|\mathcal{K}||\mathcal{K}'|}\Big(|\mathcal{K}| + |\mathcal{K}'| - 2|\mathcal{K}\cap\mathcal{K}'| \Big). \qquad (15)\end{aligned}$$

For a fixed $|\mathcal{K}|$, we will find a lower bound for

$$\frac{1}{|\mathcal{K}'|}\Big(|\mathcal{K}| + |\mathcal{K}'| - 2|\mathcal{K}\cap\mathcal{K}'| \Big) = 1 + \frac{|\mathcal{K}| - 2|\mathcal{K}\cap\mathcal{K}'|}{|\mathcal{K}'|}. \qquad (16)$$

Since we can have $|\mathcal{K}\cap\mathcal{K}'| > \frac{|\mathcal{K}|}{2}$, we know $\frac{|\mathcal{K}| - 2|\mathcal{K}\cap\mathcal{K}'|}{|\mathcal{K}'|} < 0$ when (16) is minimized. That said, $|\mathcal{K}'|$ must be as small as possible, i.e., $|\mathcal{K}'| = |\mathcal{K} \cap \mathcal{K}'|$. Thus, when (16) is minimized, we must have

$$\frac{1}{|\mathcal{K}'|}\Big(|\mathcal{K}| + |\mathcal{K}'| - 2|\mathcal{K} \cap \mathcal{K}'|\Big) = \frac{|\mathcal{K}|}{|\mathcal{K} \cap \mathcal{K}'|} - 1$$

i.e., $|\mathcal{K} \cap \mathcal{K}'|$ must be as large as possible. Since $m \in \mathcal{K} \setminus \mathcal{K}'$, we have $|\mathcal{K} \cap \mathcal{K}'| \leq |\mathcal{K}| - 1$. Therefore

$$\frac{1}{|\mathcal{K}'|}\Big(|\mathcal{K}| + |\mathcal{K}'| - 2|\mathcal{K} \cap \mathcal{K}'|\Big) \geq \frac{1}{|\mathcal{K}| - 1}. \tag{17}$$

Substituting (17) into (15) gives

$$\Bigg\|\tfrac{1}{|\mathcal{K}|}\sum_{k\in\mathcal{K}} f_k - \tfrac{1}{|\mathcal{K}'|}\sum_{k\in\mathcal{K}'} f_k\Bigg\|^2 \geq \frac{1-\delta_{|\mathcal{K}\cup\mathcal{K}'|}}{|\mathcal{K}|(|\mathcal{K}|-1)} \geq \frac{1-\delta_{2K}}{K(K-1)}.$$

Since this bound holds for every $m, \mathcal{K}$, and $\mathcal{K}'$ with $m \in \mathcal{K} \setminus \mathcal{K}'$, we have (14). ■

Note that (14) depends on δ_{2K}, and so it is natural to ask how small δ_{2K} can be for a given fingerprint design. Also, which matrices even satisfy RIP (13)? Unfortunately, certifying RIP is NP-hard in general [40], [41], but when K is sufficiently small, RIP can be demonstrated in terms of a certain coherence parameter. Specifically, the *worst case coherence* is the largest inner product between any two distinct fingerprints

$$\mu := \max_{i \neq j} |\langle f_i, f_j \rangle|. \tag{18}$$

Intuitively, we want this value to be small, as this would correspond to having the fingerprints spaced apart. In fact, the Gershgorin circle theorem can be used to produce the following well-known bound on δ_{2K} in terms of worst case coherence (see [42] for details):

Lemma 4: Given a matrix F with unit-norm columns, then

$$\delta_{2K} \leq (2K-1)\mu \tag{19}$$

where μ is the worst case coherence (18).

Combining Theorem 3 and Lemma 4 yields a coherence-based lower bound on the distance between the "guilty" and "not guilty" sets corresponding to any user m.

Theorem 5: Suppose fingerprints $F = [f_1, \ldots, f_M]$ are unit-norm with worst case coherence μ. Then

$$\operatorname{dist}(\mathcal{G}_m^{(K)}, \neg\mathcal{G}_m^{(K)}) \geq \sqrt{\frac{1-(2K-1)\mu}{K(K-1)}}. \tag{20}$$

We would like μ to be small, so that the lower bound (20) is as large as possible. But for a fixed N and M, the worst case coherence of unit-norm fingerprints cannot be arbitrarily small; it necessarily satisfies the Welch bound [43]

$$\mu \geq \sqrt{\frac{M-N}{N(M-1)}}.$$

Equality in the Welch bound occurs precisely in the case where the fingerprints form an ETF [44].

One type of ETF has already been proposed for fingerprint design: the simplex [20]. The simplex is an ETF with $M = N+1$ and $\mu = \frac{1}{N}$. In fact, Kiyavash *et al.* [20] give a derivation for the value of the distance (12) in this case

$$\operatorname{dist}(\mathcal{G}_m^{(K)}, \neg\mathcal{G}_m^{(K)}) = \sqrt{\frac{1}{K(K-1)}\frac{M}{M-1}}. \tag{21}$$

The bound (20) is lower than (21) by a factor of $\sqrt{1-\frac{2K}{N+1}}$, and for practical cases in which $K \ll N$, they are particularly close. Overall, ETF fingerprint design is a natural generalization of the provably optimal simplex design of [20].

C. Sufficiency of Deterministic ETFs

Theorem 3 gives that RIP fingerprints suffice to separate the "guilty" and "not guilty" sets for every user m. Intuitively, this allows law enforcement to deliver sound accusations, provided the noise used by the collusion is sufficiently small. As such, we naturally seek RIP matrices. To this end, random matrices are particularly successful: For any δ_{2K}, there exists a constant C such that if $N \geq CK\log(M/K)$, then matrices of Gaussian or Bernoulli (± 1) entries satisfy RIP with high probability. However, as mentioned in the previous section, checking if a given matrix satisfies RIP is computationally difficult [40], [41]; there is no known efficient procedure to verify that a randomly generated matrix satisfies RIP for the large values of K that probabilistic constructions provide.

Fortunately, there are some deterministic constructions of RIP matrices, such as [45] and [46]. However, the performance, measured by how large K can be, is not as good as the random constructions; the random constructions only require $N = \Omega(K\log^a M)$, while the deterministic constructions require $N = \Omega(K^2)$. The specific class of ETFs additionally requires $M \leq N^2$ [44].

Whether N scales as K versus K^2 is an important distinction between random and deterministic RIP matrices in the compressed sensing community [47], since users seek to minimize the number N of measurements needed to sense K-sparse signals. However, this difference in scaling offers no advantage for fingerprinting. Indeed, Ergun *et al.* showed that for any fingerprinting system, a collusion of size $K = O(\sqrt{N/\log N})$ is sufficient to overcome the fingerprints [18]. This means with such a K, the detector cannot, with high probability, identify any attacker without incurring a significant false-alarm probability P_{fa} (7). This constraint is more restrictive than deterministic ETF constructions, as $\sqrt{N/\log N} < \sqrt{N} \ll N/\log^a M$. In this way, it appears that random fingerprints are not necessary, as they offer no apparent advantage over deterministic ETF constructions. In fact, Section V illustrates that deterministic ETFs tend to perform slightly better than random fingerprints. This marks a significant difference between compressed sensing and the fingerprinting problem.

Ergun's bound indicates that with a large enough K, colluders can overcome any fingerprinting system and render our detector unreliable. As an aside, the following lemma shows that if K is sufficiently large, the colluders can *exactly* recover the original signal s.

$$F = \frac{1}{\sqrt{3}} \begin{bmatrix} + & - & + & - & + & - & + & - & & & & & & & & \\ + & + & - & - & & & & & + & - & + & - & & & & \\ + & - & - & + & & & & & & & & & + & - & + & - \\ & & & & + & + & - & - & + & + & - & - & & & & \\ & & & & + & - & - & + & & & & & + & + & - & - \\ & & & & & & & & + & - & - & + & + & - & - & + \end{bmatrix}.$$

Fig. 2. ETF constructed in Example 1.

Lemma 6: Suppose the real equal-norm fingerprints $\{f_k\}_{k\in\mathcal{K}}$ do not lie in a common hyperplane. Then, s is the unique minimizer of

$$g(x) := \sum_{k\in\mathcal{K}} \left(\|x-(s+f_k)\|^2 - \frac{1}{|\mathcal{K}|} \sum_{k'\in\mathcal{K}} \|x-(s+f_{k'})\|^2 \right)^2 .$$

Proof: Note that $g(x) \geq 0$, with equality precisely when $\|x-(s+f_k)\|^2$ is constant over $k \in \mathcal{K}$. Since the f_k's have equal norm, $g(s) = 0$. To show that this minimum is unique, suppose $g(x) = 0$ for some $x \neq s$. This implies that the values $\|x-(s+f_k)\|^2$ are constant, with each being equal to their average. Moreover, since

$$\|x-(s+f_k)\|^2 = \|x-s\|^2 - 2\langle x-s, f_k\rangle + \|f_k\|^2$$

we have that $\langle x-s, f_k\rangle$ is constant over $k \in \mathcal{K}$, contradicting the assumption that the fingerprints $\{f_k\}_{k\in\mathcal{K}}$ do not lie in a common hyperplane. ■

D. Construction of ETFs

Having established that deterministic RIP constructions are sufficient for fingerprint design, we now consider a particular method for constructing ETFs. Note that Definition 1 can be restated as follows: An ETF is an $N \times M$ matrix which has orthogonal rows of equal norm and unit-norm columns whose inner products have equal magnitude. This combination of row and column conditions makes ETFs notoriously difficult to construct in general, but a relatively simple approach was recently introduced in [46]. The approach uses a tensor-like combination of a Steiner system's adjacency matrix and a regular simplex, and it is general enough to construct infinite families of ETFs. We illustrate the construction with an example.

Example 1 (See [46]): To construct an ETF, we will use a simple class of Steiner systems, namely $(2, 2, v)$-Steiner systems, and Hadamard matrices with v rows. In particular, $(2, 2, v)$-Steiner systems can be thought of as all possible pairs of a v-element set [48], while a Hadamard matrix is a square matrix of ±1's with orthogonal rows [48].

In this example, we take $v = 4$. The adjacency matrix of the $(2, 2, 4)$-Steiner system has $\binom{4}{2} = 6$ rows, each indicating a distinct pair from a size-4 set

$$A = \begin{bmatrix} + & + & & \\ + & & + & \\ + & & & + \\ & + & + & \\ & + & & + \\ & & + & + \end{bmatrix}. \tag{22}$$

To be clear, we use "$+/-$" to represent ±1 and an empty space to represent a 0 value. Also, one example of a 4×4 Hadamard matrix is

$$H = \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{bmatrix}. \tag{23}$$

We now build an ETF by replacing each of the "$+$" entries in the adjacency matrix (22) with a row from the Hadamard matrix (23). This is done in such a way that for each column of A, distinct rows of H are used; in this example, we use the second, third, and fourth rows of H. After performing this procedure, we normalize the columns, and the result is a real ETF F of dimension $N = 6$ with $M = 16$ fingerprints (see Fig. 2).

We will use this method of ETF construction for simulations in Section V. Specifically, our simulations will let the total number of fingerprints M range from $2N$ to $7N$, a significant increase from the conventional $N+1$ simplex and N orthogonal fingerprints. Note, however, that unlike simplex and orthogonal fingerprints, there are limitations to the dimensions that admit real ETFs. For instance, the size of a Hadamard matrix is necessarily a multiple of 4, and the existence of Steiner systems is rather sporadic. Fortunately, identifying Steiner systems is an active area of research, with multiple infinite families already characterized and alternative construction methods developed [46], [49]. Also, it is important to note that the sporadic existence of ETFs will not help colluders determine the fingerprint marking, since the same random factors used in conventional orthogonal and simplex fingerprints can also be employed with ETFs in general. For example, we can randomly orient our ETF fingerprints F by randomly drawing an $N \times N$ orthogonal matrix Q and using the columns of QF as fingerprints; note that QF is still an ETF by Definition 1, but these new fingerprints enjoy some degree of randomness.

IV. Error Analysis

A. Analysis of Type I and Type II Errors

We now investigate the worst case errors involved with using ETF fingerprint design and focused correlation detection under linear averaging attacks. Recall that the worst case type I error probability (8) is the probability of falsely accusing the most guilty-looking innocent user. Also recall that the worst case type II error probability (10) is the probability of not accusing the most vulnerable guilty user.

Theorem 7: Suppose the fingerprints $F = \{f_m\}_{m=1}^{M}$ form an ETF. Then, the worst case type I and type II error probabilities, (8) and (10), satisfy

$$\mathrm{P_I}(F,\tau,\alpha) \le Q\left[\frac{\gamma}{\sigma}(\tau-\mu)\right]$$
$$\mathrm{P_{II}}(F,\tau,\alpha) \le Q\left[\frac{\gamma}{\sigma}\Big(\big((1+\mu)\max_{k\in\mathcal{K}}\alpha_k-\mu\big)-\tau\Big)\right]$$

where $Q(x) := \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-u^2/2}du$ and $\mu = \sqrt{\frac{M-N}{N(M-1)}}$.

Proof: Under hypothesis $H_0(m)$, the test statistic for the detector (6) is

$$\begin{aligned} T_z(m) &= \frac{1}{\gamma^2}\left\langle \sum_{n\in\mathcal{K}}\alpha_n f_n + \epsilon, f_m\right\rangle \\ &= \sum_{n\in\mathcal{K}}\alpha_n(\pm\mu) + \epsilon' \end{aligned}$$

where ϵ' is the projection of noise ϵ/γ onto the normed vector f_m/γ, and, due to symmetry of the variance of ϵ in all dimensions, $\epsilon' \sim \mathcal{N}(0,\sigma^2/\gamma^2)$. Thus, under hypothesis $H_0(m)$, $T_z(m) \sim \mathcal{N}(\sum_{n\in\mathcal{K}}\alpha_n(\pm\mu),\sigma^2/\gamma^2)$. We can subtract the mean and divide by the standard deviation to obtain

$$\begin{aligned} \mathrm{Prob}\left[T_m(z)\ge\tau|H_0(m)\right] &= Q\left(\frac{\gamma}{\sigma}\left[\tau-\left(\sum_{n\in\mathcal{K}}\alpha_n(\pm\mu)\right)\right]\right) \\ &\le Q(\frac{\gamma}{\sigma}(\tau-\mu)). \end{aligned} \tag{24}$$

Note that $Q(x)$ is a decreasing function. The bound (24) is obtained by setting all of the coefficients of the coherence to be positive.

Likewise, under hypothesis $H_1(m)$, the test statistic is

$$\begin{aligned} T_z(m) &= \frac{1}{\gamma^2}\left\langle \sum_{n\in\mathcal{K}}\alpha_n f_n + \epsilon, f_m\right\rangle \\ &= \alpha_m + \sum_{n\in\mathcal{K}\setminus\{m\}}\alpha_n(\pm\mu)+\epsilon'. \end{aligned}$$

Thus, under hypothesis $H_1(m)$

$$T_z(m) \sim \mathcal{N}\left(\alpha_m + \sum_{n\in\mathcal{K}\setminus\{m\}}\alpha_n(\pm\mu), \frac{\sigma^2}{\gamma^2}\right).$$

Since $1-Q(x) = Q(-x)$, the type II error probability can be bounded as

$$\begin{aligned} \mathrm{Prob}\left[T_m(z)<\tau\right] &= Q\left(-\frac{\gamma}{\sigma}\left[\tau-\left(\alpha_m+\sum_{n\in\mathcal{K}\setminus\{m\}}\alpha_n(\pm\mu)\right)\right]\right) \\ &\le Q\Big(\frac{\gamma}{\sigma}([\alpha_m(1+\mu)-\mu]-\tau)\Big). \end{aligned}$$

We can now evaluate the worst case errors (8) and (10). For type I errors, with $m\notin\mathcal{K}$

$$\begin{aligned} \mathrm{P_I}(F,\tau,\alpha) &= \max_{\mathcal{K}}\max_{m\notin\mathcal{K}}\mathrm{P_I}(F,m,\tau,\mathcal{K},\alpha) \\ &= \max_{\mathcal{K}}\max_{m\notin\mathcal{K}}\mathrm{Prob}\left[T_m(z)\ge\tau|H_0(m)\right] \\ &\le \max_{\mathcal{K}}\max_{m\notin\mathcal{K}}Q(\frac{\gamma}{\sigma}(\tau-\mu)) \\ &= Q(\frac{\gamma}{\sigma}(\tau-\mu)). \end{aligned}$$

For type II errors

$$\begin{aligned} \mathrm{P_{II}}(F,\tau,\alpha) &= \max_{\mathcal{K}}\min_{m\in\mathcal{K}}\mathrm{P_{II}}(F,m,\tau,\mathcal{K},\alpha) \\ &= \max_{\mathcal{K}}\min_{m\in\mathcal{K}}\mathrm{Prob}\left[T_m(z)<\tau\right] \\ &\le \max_{\mathcal{K}}\min_{m\in\mathcal{K}}Q(\frac{\gamma}{\sigma}([\alpha_m(1+\mu)-\mu]-\tau)) \\ &= Q(\frac{\gamma}{\sigma}([\max_{m\in\mathcal{K}}\alpha_m(1+\mu)-\mu]-\tau)). \end{aligned}$$

The last equation follows since once the α's are fixed, the actual coalition $\mathcal{K}$ does not matter. ■

We can further maximize over all possible weightings α

$$\begin{aligned} \mathrm{P_I}(F,\tau) &= \max_{\alpha}\mathrm{P_I}(F,\tau,\alpha) \\ &\le Q(\frac{\gamma}{\sigma}(\tau-\mu)) \\ \mathrm{P_{II}}(F,\tau) &= \max_{\alpha}\mathrm{P_{II}}(F,\tau,\alpha) \\ &\le \max_{\alpha}Q(\frac{\gamma}{\sigma}([\max_{m\in\mathcal{K}}\alpha_m(1+\mu)-\mu]-\tau)) \\ &= Q(\frac{\gamma}{\sigma}([\min_{\alpha}\max_{m\in\mathcal{K}}\alpha_m(1+\mu)-\mu]-\tau)) \\ &= Q(\frac{\gamma}{\sigma}([\frac{1}{K}(1+\mu)-\mu]-\tau)). \end{aligned}$$

Thus, the vector α which minimizes the value of its maximum element is the uniform weight vector. This motivates the attackers' use of a uniformly weighted coefficient vector α to maximize the probability that none of the members will be caught.

B. Minmax Error Analysis

From a detection standpoint, an important criterion is the minmax error probability (11). The threshold τ trades off type I and type II errors. Thus, there is a value of τ, denoted by τ^*, which minimizes the maximum of the probabilities of the two error types. Since the bound for the type I error probability is independent of α, and the bound for the type II error probability is maximized with a uniform weighting, assume this to be the case.

Theorem 8: The minmax probability of error (11) can be bounded as

$$Q\left(\frac{d^*_{\mathrm{low}}}{2}\right) \le \mathrm{P_{minmax}}(F,\alpha) \le Q\left(\frac{d^*_{\mathrm{up}}}{2}\right) \tag{25}$$

where

$$\begin{aligned} d^*_{\text{low}} &:= \frac{\sqrt{\frac{M}{M-1}}\sqrt{ND_f}}{\sigma\sqrt{K(K-1)}} \\ d^*_{\text{up}} &:= \frac{\sqrt{ND_f}}{\sigma K}\Big(1-(2K-1)\mu\Big). \end{aligned}$$

Note that for orthogonal and simplex fingerprints, the minmax errors are both of the form

$$\mathrm{P}_{\text{minmax}}(F,\alpha) = Q\left(\frac{d^*(K)}{2}\right).$$

For orthogonal fingerprints[20]

$$d^*(K) = \frac{\sqrt{ND_f}}{\sigma K}$$

which is better than $d^*_{\text{low}}(K)$ (the Q function is decreasing). For simplex fingerprints, $d^*(K)$ is slightly better than both [20]

$$d^*(K) = \frac{\sqrt{ND_f}}{\sigma K}\frac{M}{M-1}.$$

Proof: The lower bound is the sphere packing lower bound [24]. For the upper bound

$$\begin{aligned} P_e(F,\tau,\alpha) &= \max\{P_{\text{I}}(F,\tau,\alpha), P_{\text{II}}(F,\tau,\alpha)\} \\ &\le \max\{Q(\frac{\gamma}{\sigma}(\tau-\mu)), Q(\frac{\gamma}{\sigma}([\frac{1}{K}(1+\mu)-\mu]-\tau))\}. \end{aligned}$$

Since the test statistic $T_z(m)$ is normally distributed with the same variance under either of the hypotheses $H_0(m)$ and $H_1(m)$, the value of τ that minimizes this upper bound is the average of the means μ and $\frac{1+\mu}{K}-\mu$, namely $\tau^* := \frac{1+\mu}{2K}$. Using this τ^* and recalling (2), we have

$$\begin{aligned} \mathrm{P}_{\text{minmax}}(F,\alpha) = \min_{\tau} \mathrm{P}_e(F,\tau,\alpha) &= \mathrm{P}_e(F,\tau^*,\alpha) \\ &\le Q\left(\frac{\gamma}{\sigma}(\tau^*-\mu)\right) \\ &= Q\left(\frac{\sqrt{ND_f}}{2\sigma K}\Big(1-(2K-1)\mu\Big)\right) \\ &= Q(d^*_{\text{up}}(K)/2). \end{aligned}$$

■

Consider the regime where N is large, M grows linearly or faster than N, and WNR is constant (in particular 0, so $D_f = \sigma^2$). Then, $\mu \approx 1/\sqrt{N}$

$$d^*_{\text{low}} \approx \frac{\sqrt{N}}{K}, \quad \text{and} \quad d^*_{\text{up}} \approx \frac{\sqrt{N}}{K}\left(1-\frac{2K}{\sqrt{N}}\right) = \frac{\sqrt{N}}{K}-2.$$

If $K \ll \sqrt{N}$, then both bounds go to infinity so $\mathrm{P}_{\text{minmax}}(F,\alpha) \to 0$. Also, the geometric figure of merit (12) then behaves as

$$\underline{\text{dist}}(\mathcal{G}_m^{(K)}, \neg\mathcal{G}_m^{(K)}) \approx \frac{1}{K} \approx \sqrt{N} d^*_{\text{up}}.$$

If K is proportional to $\sqrt{N}$, then $\mathrm{P}_{\text{minmax}}(F,\alpha)$ is bounded away from 0.

We can also compute the error exponent for this test

$$\mathrm{e}(F,\tau^*,\alpha) := -\lim_{N\to\infty}\frac{1}{N}\ln \mathrm{P}_e(F,\tau^*,\alpha).$$

Corollary 9: If $M \gg N \gg K^2$, then the error exponent is

$$\mathrm{e}(F,\tau^*,\alpha) = \frac{1}{8K^2}. \tag{26}$$

Proof: The proof follows by applying the asymptotic equality $\ln Q(t) \sim -\frac{t^2}{2}$ as $t \to \infty$ to the bounds in (25). The bounds are asymptotically equivalent. ■

Note that this error exponent is the same as in the simplex case [20]. As $K \to \infty$, the error exponent (26) goes to zero.

V. Simulations

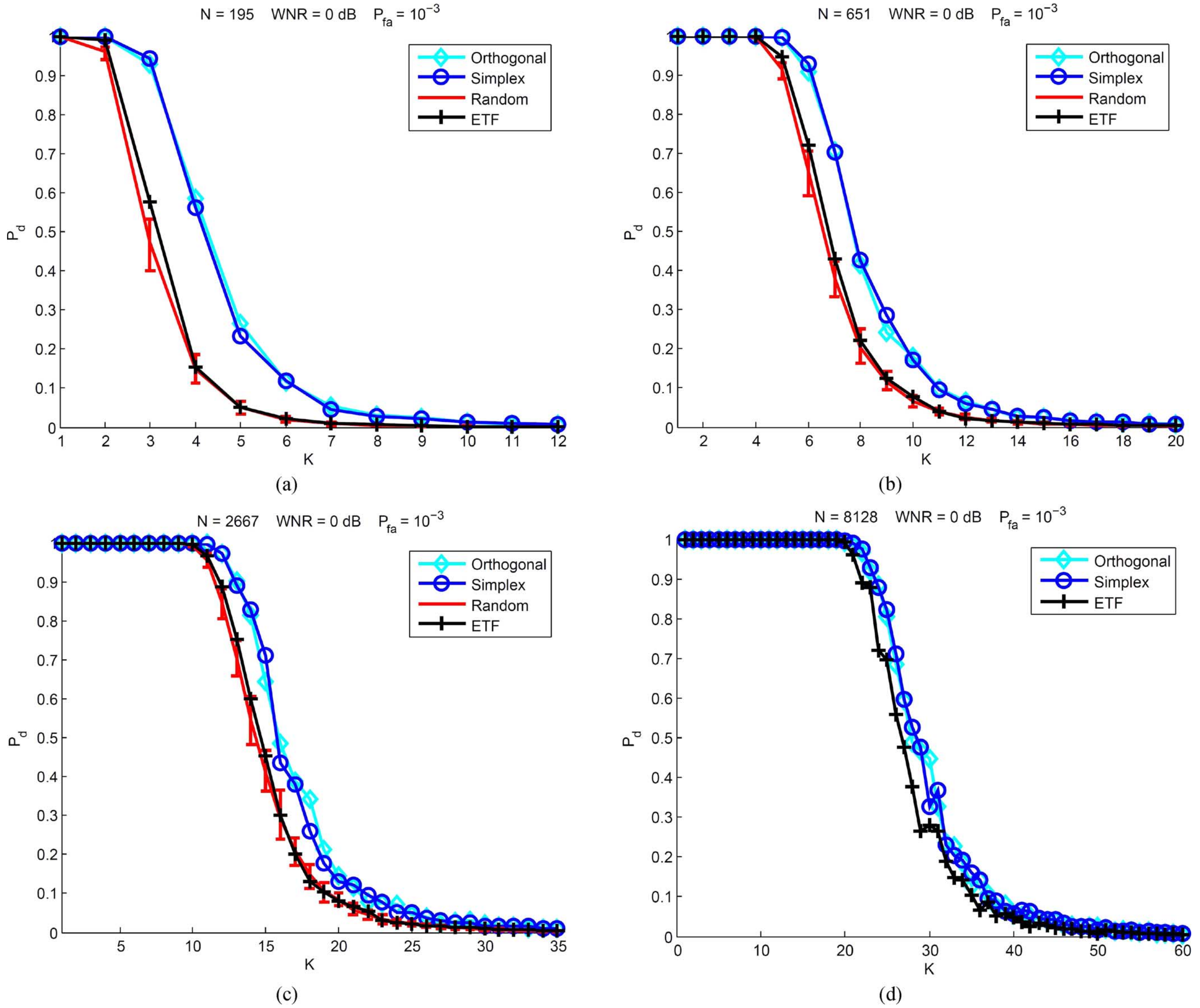
In Section IV, the worst case error probabilities were analyzed, where the worst case was over all collusions. Here, we investigate average-case behavior. We examine the probability of detecting at least one guilty user, P_{d} (9), as a function of K with the false alarm P_{fa} (7) fixed below a threshold. The threshold can be interpreted as the (legally) allowable limit for probability of false accusation. We compare the average-case performance of ETF fingerprints, simplex fingerprints [20], and orthogonal fingerprints [16] for four signal dimension sizes $N \in \{195, 651, 2667, 8128\}$. We also examine the performance of randomly generated ETF fingerprints for $N \in \{195, 651, 2667\}$. The results demonstrate that ETF fingerprints perform almost as well as both orthogonal and simplex fingerprints, while accommodating several times as many users. The results also show that randomly generated ETF fingerprints do not outperform the deterministic ETF fingerprints. We now describe the design of the simulations.

A. Design

For each signal dimension size N, ETF, orthogonal, and simplex fingerprints were created. The deterministic ETF fingerprint design was constructed using the method shown in Example 1. For $N = 195$, a $(2,7,91)$-Steiner system was used [49], yielding $M = 1456$ fingerprints. For $N = 651$ and $N = 2667$, the Steiner systems were constructed using projective geometry, with $M = 2016$ and $M = 8128$, respectively [46]. For $N = 8128$, a $(2,2,2^7)$-Steiner system was used, giving $M = 16\,384$ fingerprints [46].

The orthogonal fingerprint design was constructed using an identity matrix. As discussed at the end of Section III, random rotations of the basis vectors can be applied in practice to hide them from attackers [16], but this does not affect detection. For orthogonal fingerprint designs, $M = N$.

The simplex fingerprint design was constructed using the following method. The vectors of a regular simplex are equidistant from the center, having the same power, and have inner products equal to $-\frac{1}{N}$ [50]. Sequentially set the elements in each fingerprint to satisfy constraints. For the first fingerprint, set the first coordinate to satisfy the power constraint (e.g., γ)



(a)

(b)

(c)

(d)

Fig. 3. Plot of the probability of detecting at least one colluder ($\mathrm{P_d}$) as a function of the number of colluders (K). The threshold τ is picked to be the minimum threshold to fix $\mathrm{P_{fa}} \leq 10^{-3}$. The WNR is 0 dB. For the curves plotting performance of random fingerprints, the curve denotes probability of detection averaged over the random fingerprint ensembles, and the vertical bars denote maximum and minimum detection probabilities. (a) Plot for $N = 195$. The number of fingerprints M used were 195 for the orthogonal, 196 for the simplex, and 1456 for the ETF constructions. There were 2000 random ETF ensembles generated. (b) Plot for $N = 651$. The number of fingerprints M used were 651 for the orthogonal, 652 for the simplex, and 2016 for the ETF constructions. There were 250 random ETF ensembles generated. (c) Plot for $N = 2667$. The number of fingerprints M used were 2667 for the orthogonal, 2668 for the simplex, and 8128 for the ETF constructions. There were 30 random ETF ensembles generated. (d) Plot for $N = 8128$. The number of fingerprints M used were 8128 for the orthogonal, 8129 for the simplex, and 16 384 for the ETF constructions.

and the rest of the entries as zero. To meet the inner product constraint with the first fingerprint, every remaining fingerprint has the same value as its first coordinate (e.g., $-\frac{1}{\gamma N}$). For the second fingerprint, set its second coordinate to meet the power constraint and the rest to be zero. Every remaining fingerprint must now have the same second coordinate to satisfy the inner product constraint with the second fingerprint. Repeating this process fills the fingerprint matrix. For simplex fingerprint designs, $M = N + 1$.

For $N \in \{195, 651, 2667\}$, we also generated random ETF fingerprint ensembles with the same size $N \times M$ as the corresponding deterministic ETF. The random fingerprints were constructed by drawing i.i.d. samples from a $\mathcal{N}(0, 1/N)$ distribution for each element and then normalizing the fingerprints. For $N \in \{195, 651, 2667\}$, 2000, 250, and 30 such fingerprint ensembles were created, respectively.

For each fingerprint matrix and collusion size K, over 15 000 linear collusion attacks were simulated. For each attack, K of the M fingerprints were randomly chosen and uniformly averaged. Next, an i.i.d. Gaussian noise vector was added with per-sample noise power $\sigma^2 = D_f$, corresponding to a WNR (4) of 0 dB [19]. The test statistic $T_z(m)$ (5) was then computed for each user m. For each threshold τ, it was determined whether there was a detection event (at least one colluder with $T_z(m) > \tau$) and/or a false alarm (at least one innocent user with $T_z(m) > \tau$). The detection and false alarm counts were then averaged over all attacks. The minimal τ value was selected for which $\mathrm{P_{fa}} \leq 10^{-3}$. This induced the corresponding $\mathrm{P_d}$.

B. Results

We ran experiments with four different signal dimension sizes $N \in \{195, 651, 2667, 8128\}$. The noise level was kept at $\mathrm{WNR} = 0$ dB. The value of K varied between 1 and sufficiently large values, so $\mathrm{P_d}$ approached zero. Plots for the probability of detection $\mathrm{P_d}$ as a function of the size of the coalition K are shown in Fig. 3(a)–(d) for $N \in \{195, 651, 2667, 8128\}$, respectively. The largest values of K for which at least one attacker can be caught with probability (nearly) one under the $\mathrm{P_{fa}}$ constraint are about 2, 5, 11, and 21, respectively. Overall, ETF fingerprints perform comparably to orthogonal and simplex fingerprints while accommodating many more users. Also, the randomly generated fingerprints generally performed slightly worse than the deterministic ETF fingerprints, demonstrating that random constructions do not have an advantage in terms of focused detection.

Acknowledgment

The authors would like to thank the anonymous reviewers and the associate editor for their suggestions which improved the quality of this work. The authors thank the Air Force Summer Faculty Fellowship Program for making this collaboration possible. The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

References

[1] D. Boneh and J. Shaw, “Collusion-secure fingerprinting for digital data,” *IEEE Trans. Inf. Theory*, vol. 44, no. 5, pp. 1897–1905, Sep. 1998.

[2] B. Chor, A. Fiat, M. Naor, and B. Pinkas, “Tracing traitors,” *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 893–910, May 2000.

[3] H. Schaathun, “The Boneh–Shaw fingerprinting scheme is better than we thought,” *IEEE Trans. Inf. Forens. Security*, vol. 1, no. 2, pp. 248–255, Jun. 2006.

[4] G. Tardos, “Optimal probabilistic fingerprint codes,” *J. ACM*, vol. 55, no. 2, p. 10, 2008.

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

[6] A. Somekh-Baruch and N. Merhav, “Achievable error exponents for the private fingerprinting game,” *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1827–1838, May 2007.

[7] N. Anthapadmanabhan, A. Barg, and I. Dumer, “On the fingerprinting capacity under the marking assumption,” *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2678–2689, Jun. 2008.

[8] P. Moulin, Universal fingerprinting: capacity and random-coding exponents 2008 [Online]. Available: arXiv:0801.3837

[9] A. Barg, G. Blakley, G. Kabatiansky, and C. Tavernier, “Robust parent-identifying codes,” in *Proc. IEEE Inf. Theory Workshop*, 2010, pp. 1–4.

[10] S. Lin, M. Shahmohammadi, and H. El Gamal, “Fingerprinting with minimum distance decoding,” *IEEE Trans. Inf. Forens. Security*, vol. 4, no. 1, pp. 59–69, Mar. 2009.

[11] M. Cheng and Y. Miao, “On anti-collusion codes and detection algorithms for multimedia fingerprinting,” *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4843–4851, Jul. 2011.

[12] J. Cotrina-Navau and M. Fernández, “A family of asymptotically good binary fingerprinting codes,” *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5335–5343, Oct. 2010.

[13] D. Boneh, A. Kiayias, and H. Montgomery, “Robust fingerprinting codes: A near optimal construction,” in *Proc. 10th Annu. ACM Workshop Dig. Rights Manag.*, 2010, pp. 3–12.

[14] H. Koga and Y. Minami, “A digital fingerprinting code based on a projective plane and its identifiability of all malicious users,” *IEICE Trans. Fund. Electron., Commun. Comput. Sci.*, vol. 94, no. 1, pp. 223–232, 2011.

[15] W. Trappe, M. Wu, Z. Wang, and K. Liu, “Anti-collusion fingerprinting for multimedia,” *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 1069–1087, Apr. 2003.

[16] I. Cox, J. Kilian, F. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia,” *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.

[17] J. Kilian, F. Leighton, L. Matheson, T. Shamoon, R. Tarjan, and F. Zane, “Resistance of digital watermarks to collusive attacks,” in *Proc. IEEE Int. Symp. Inf. Theory*, 1998, p. 271.

[18] F. Ergun, J. Kilian, and R. Kumar, “A note on the limits of collusion-resistant watermarks,” in *Proc. Adv. Cryptol.—EUROCRYPT*, 1999, pp. 140–149.

[19] Z. Wang, M. Wu, H. Zhao, W. Trappe, and K. Liu, “Anti-collusion forensics of multimedia fingerprinting using orthogonal modulation,” *IEEE Trans. Image Process.*, vol. 14, no. 6, pp. 804–821, Jun. 2005.

[20] N. Kiyavash, P. Moulin, and T. Kalker, “Regular simplex fingerprints and their optimality properties,” *IEEE Trans. Inf. Forens. Security*, vol. 4, no. 3, pp. 318–329, Sep. 2009.

[21] N. Hayashi, M. Kuribayashi, and M. Morii, “Collusion-resistant fingerprinting scheme based on the CDMA-technique,” in *Proc. 2nd Int. Conf. Adv. Inf. Comput. Security*, 2007, pp. 28–43.

[22] Z. Li and W. Trappe, “Collusion-resistant fingerprints from WBE sequence sets,” in *Proc. IEEE Int. Conf. Commun.*, 2005, vol. 2, pp. 1336–1340.

[23] J. Jourdas and P. Moulin, “High-rate random-like spherical fingerprinting codes with linear decoding complexity,” *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 4, pp. 768–780, Dec. 2009.

[24] N. Kiyavash and P. Moulin, “Sphere packing lower bound on fingerprinting error probability,” *Proc. SPIE*, vol. 6505, 2007.

[25] O. Dalkilic, E. Ekrem, S. Varlik, and M. Mihcak, “A detection theoretic approach to digital fingerprinting with focused receivers under uniform linear averaging Gaussian attacks,” *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 4, pp. 658–669, Dec. 2010.

[26] Y. Wang and P. Moulin, “Capacity and optimal collusion attack channels for Gaussian fingerprinting games,” *Proc. SPIE*, vol. 6505, no. 1, p. 65050J, 2007.

[27] H. Ling, H. Feng, F. Zou, W. Yan, and Z. Lu, “A novel collusion attack strategy for digital fingerprinting,” in *Proc. Int. Conf. Dig. Watermarking*, 2011, pp. 224–238.

[28] N. Kiyavash and P. Moulin, “Performance of orthogonal fingerprinting codes under worst-case noise,” *IEEE Trans. Inf. Forens. Security*, vol. 4, no. 3, pp. 293–301, Sep. 2009.

[29] R. Pickholtz, D. Schilling, and L. Milstein, “Theory of spread-spectrum communications—A tutorial,” *IEEE Trans. Commun.*, vol. 30, no. 5, pp. 855–884, May 1982.

[30] I. Cox, M. Miller, and J. Bloom, *Digital Watermarking*. San Mateo, CA: Morgan Kaufmann, 2002.

[31] F. Hartung and M. Kutter, “Multimedia watermarking techniques,” *Proc. IEEE*, vol. 87, no. 7, pp. 1079–1107, Jul. 1999.

[32] W. Dai, N. Kiyavash, and O. Milenkovic, “Spherical codes for sparse digital fingerprinting,” presented at the Spring Central Meet. Amer. Math. Soc., 2008.

[33] D. Varodayan and C. Pépin, “Collusion-aware traitor tracing in multimedia fingerprinting using sparse signal approximation,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2008, pp. 1645–1648.

[34] H. Pham, W. Dai, and O. Milenkovic, “Compressive list-support recovery for colluder identification,” in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2010, pp. 4166–4169.

[35] C. Colbourn, D. Horsley, and V. Syrotiuk, “Frameproof codes and compressive sensing,” in *Proc. 48th Allerton Conf. Commun., Control, Comput.*, 2010, pp. 985–990.

[36] E. Candes and T. Tao, “Decoding by linear programming,” *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[37] D. Needell and J. A. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Appl. Comp. Harmon. Anal.*, vol. 26, pp. 301–321, 2008.

[38] M. A. Davenport and M. B. Wakin, “Analysis of orthogonal matching pursuit using the restricted isometry property,” *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4395–4401, Sep. 2010.

[39] E. Candes, J. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[40] A. S. Bandeira, E. Dobriban, D. G. Mixon, and W. F. Sawin, Certifying the restricted isometry property is hard 2012 [Online]. Available: arXiv:1204.1580

[41] A. M. Tillmann and M. E. Pfetsch, The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing 2012 [Online]. Available: arXiv:1205.2081

[42] D. G. Mixon, "Sparse signal processing with frame theory," Ph.D. dissertation, Princeton Univ., Princeton, NJ, 2012.

[43] L. Welch, "Lower bounds on the maximum cross correlation of signals (corresp.)," *IEEE Trans. Inf. Theory*, vol. 20, no. 3, pp. 397–399, May 1974.

[44] T. Strohmer and R. W. Heath, "Grassmannian frames with applications to coding and communication," *Appl. Comput. Harmon. Anal.*, vol. 14, no. 3, pp. 257–275, 2003.

[45] R. DeVore, "Deterministic constructions of compressed sensing matrices," *J. Complex.*, vol. 23, no. 4–6, pp. 918–925, 2007.

[46] M. Fickus, D. G. Mixon, and J. Tremain, "Steiner equiangular tight frames," *Linear Algebra Appl.*, vol. 436, no. 5, pp. 1014–1027, 2012.

[47] A. S. Bandeira, M. Fickus, D. G. Mixon, and P. Wong, The road to deterministic matrices with the restricted isometry property 2012 [Online]. Available: arXiv:1202.1234

[48] J. van Lint and R. Wilson, *A Course in Combinatorics*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[49] "Steiner systems," La Jolla Covering Repository 2011 [Online]. Available: http://www.ccrwest.org/cover/steiner.html

[50] J. Munkres, *Elements of Algebraic Topology*. Reading, MA: Addison-Wesley, 1984, vol. 2.

Dustin G. Mixon received the B.S. degree in mathematics from Central Washington University, Ellensburg, WA, in 2004, the M.S. degree in applied mathematics from the Air Force Institute of Technology, Wright-Patterson AFB, OH, in 2006, and the Ph.D. degree in applied and computational mathematics from Princeton University, Princeton, NJ, in 2012.

From 2006 to 2009, he was an applied mathematical analyst at the Air Force Research Laboratory, Brooks City-Base, TX. He is presently an Assistant Professor of Mathematics in the Department of Mathematics and Statistics at the Air Force Institute of Technology. His research interests include frame theory, signal processing, and compressed sensing.

Christopher J. Quinn (S'11) recieved a B.S. in Engineering Physics from Cornell University and M.S. in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign in 2008 and 2010 respectively.

He is currently a Ph.D. student in Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign. His research interests include information theory and statistical signal processing.

Negar Kiyavash (S'06–M'06) received the B.S. degree in electrical and computer engineering from the Sharif University of Technology, Tehran, in 1999, and the M.S. and Ph.D. degrees, also in electrical and computer engineering, both from the University of Illinois at Urbana-Champaign in 2003 and 2006, respectively. She is presently an Assistant Professor in the Electrical and Computer Engineering Department at the University of Illinois at Urbana-Champaign. Her research interests are in information theory and statistical signal processing with applications to computer, communication, and multimedia security.

Matthew Fickus (M'08) received a Ph.D. in Mathematics from the University of Maryland in 2001, and spent the following three years at Cornell University as an NSF VIGRE postdoc. In 2004, he started working in the Department of Mathematics and Statistics of the Air Force Institute of Technology, where he is currently an Associate Professor of Mathematics. His research focuses on frames and harmonic analysis, emphasizing their applications to signal and image processing.