Optimal Probabilistic Fingerprint Codes*

Gábor Tardos
Rényi Institute
Pf. 127, H-1354 Budapest, Hungary
tardos@renyi.hu

Abstract

We construct binary codes for fingerprinting. Our codes for n users that are ϵ-secure against c pirates have length $O(c^2 \log(n/\epsilon))$. This improves the codes proposed by Boneh and Shaw [3] whose length is approximately the square of this length. Our codes use the full power of randomization. This improvement carries over to works using the Boneh-Shaw code as a primitive, e.g. to the dynamic traitor tracing scheme of Tassa [16].

By proving matching lower bounds we establish that the length of our codes is best within a constant factor for reasonable error probabilities. This lower bound generalizes the bound found independently by Peikert, Shelat, and Smith [11] that applies to a limited class of codes. Our results also imply that randomized fingerprint codes over a binary alphabet are as powerful as over an arbitrary alphabet and the equal strength of two distinct models for fingerprinting.

1 Introduction

1.1 Motivation

The problem of making many copies of a digital document unique by embedding something like a serial number is a very natural one. For example, a software distributor may want to be able to trace any running copy of his software to the specific costumer (user) who bought that piece of software. Other applications include copyrighted digital documents of any form, e.g. digital images, audio or video. Leaking sensitive documents to the press can also be fought this way.

If the users do not cheat, this represents no problem, but a malicious user may try to erase the serial number (also called *fingerprint*) from his copy before distributing illegal copies. To prevent such fraud it is natural to distribute the

*Preliminary version of this paper appeared in *STOC'03* [15]. Work on this paper has been supported by the Hungarian National Research & Development Fund # 2/019/2001, the Hungarian Science Foundation grants OTKA T029255, OTKA T030059, and the grant AKP 2000-78 2.1.

digits of the fingerprint into locations of the digital document that are unknown to the users. The digits in these positions must be irrelevant with respect to the intended use of the document (e.g. the software must run correctly whatever the digits are on these positions), but the exact locations should be impossible to find for the user. This way the user cannot erase the fingerprint without risking to alter relevant bits of the document too. Here we consider the documents to be strings over a finite alphabet Σ. In this paper we do not consider the task of hiding the digits of the fingerprint. This is a highly nontrivial implementation challenge.

A further problem arises when a coalition of malicious users (we call them pirates) collaborate. Each of them has access to one fingerprinted copy of the document. Comparing these copies, they can isolate the positions where the copies differ, these positions hold digits of the fingerprint. They can erase these digits of the fingerprint or they can even introduce arbitrary digits in these positions. Such a strategy results in a document (pirated copy) that is not identical to any of the legitimate (fingerprinted) copies but is identical with all of them on relevant positions. In this scenario, we want the distributor to be able to identify at least one pirate of the guilty coalition. We assume that the pirates do not alter the digital document on positions where all of the copies they see agree. This is called the *marking condition*. The pirates may have an arbitrary strategy to fill in the positions where they detected disagreement. (See Section 6 for a slight relaxation of the marking condition.)

It is important that in the scenario we consider the pirates can put any digit in the document on positions where they detected difference. If we restrict the pirates to use a digit at any position that appears in the same position in one of their documents, we get another model that is more restrictive for the pirates if the alphabet is not binary. This alternate model allows for a deterministic solution, such codes are called IPP codes. This more restrictive assumption (while adequate in some applications) seems to be too strong in many fingerprinting applications. Our problem formulated above has no error-free solution if there are at least three users and any two of them can form the pirate coalition. We present here an efficient randomized scheme. See discussion of related results in Section 1.3.

1.2 The model

Since a deterministic solution does not exist, we turn to a *randomized procedure* to generate codewords and accuse users that works with high probability. In the formal definition below, we simplify the notation by ignoring the relevant positions of the document and concentrating on the fingerprint itself. Thus, the length of the fingerprint code is the number of irrelevant positions needed to embed such a code.

Definition 1.1. *A* fingerprint code *of length m for n users over the alphabet Σ is a distribution over the pairs (X, σ), where X is an n by m matrix over Σ and σ is an algorithm that takes a string $y \in \Sigma^m$ (the pirated copy) as input,*

and produces a subset $\sigma(y) \subseteq [n] := \{1, 2, \ldots, n\}$ *(the set of accused users). For* $\emptyset \neq C \subseteq [n]$ *a* C-strategy *is an algorithm* ρ *that takes the submatrix of* X *formed by the rows with indices in* C *as input, and produces a string* $y = \rho(X) \in \Sigma^m$ *as output*[1] *and satisfies the* marking condition *that, for all positions* $1 \leq i \leq m$*, if all the values* X_{ji} *for* $j \in C$ *agree with some letter* $s \in \Sigma$ *then* $y_i = s$*. We say that a fingerprint code is* ϵ-secure against coalitions of size c*, if for any* $C \subseteq [n]$ *of size* $|C| \leq c$ *and for any* C*-strategy* ρ*, the error probability*

$$P[\sigma(\rho(X)) = \emptyset \text{ or } \sigma(\rho(X)) \not\subseteq C]$$

is at most ϵ*.*

Our main results are a construction of short fingerprint codes (see Corollary 3) and a matching lower bound for the length of any fingerprint code (see Theorem 4). We state these results and give the construction itself in the next section.

Remarks

1. In the above definition, we do not have any complexity assumptions on the algorithms σ and ρ. Furthermore, we can restrict our attention to deterministic algorithms. Randomization in σ can be "moved" to the distribution over (X, σ), while for ρ one can suppose it chooses deterministically one of the strings that maximizes the error probability. Thus, considering randomized or deterministic algorithms here (or simply considering σ and ρ to be functions) leads to equivalent definitions. We assume all algorithms to be deterministic unless otherwise stated. Despite allowing algorithms of arbitrary complexity, our construction in the next section uses a very efficient algorithm σ: each accusation is determined by a linear constraint. The proof of the lower bound claimed in the next section is also based on very simple (randomized) algorithms ρ.

2. In the setting of the above definition, one can assume that always a single user is accused, i.e., that $|\sigma(y)| = 1$. Indeed, one can modify σ to accuse any one user from the set $\sigma(y)$ and an arbitrary user if $\sigma(y) = \emptyset$. This does not increase the error probability. However, later we will treat separately the error of accusing an innocent user and the error of not accusing any guilty one. For obvious reasons the former type of error (that we call "soundness error") is considered far worse. Our construction has the advantage that the bound on the soundness error is maintained even against arbitrarily large coalitions. To achieve this the algorithm σ need to be able not to accuse anybody if it is not sure.

3. The definition above assumes that the number n of users is known in advance. Our construction however does not need this assumption, codewords can be generated one by one as users appear.

4. In the real scenario of fingerprinting digital documents explained before the definition the pirates have just a little less information than in the setting of this definition. Indeed, they learn only about irrelevant positions where

[1] For simplicity, we denote the output of the C-strategy ρ by $\rho(X)$, despite the fact that the input is only a submatrix of X, ρ "does not see" the rows with indices outside C.

not all of their codes agree. Thus, they can reconstruct the submatrix of X consisting of their respective rows, but *missing* all columns that are constant in this submatrix. This subtle difference is not relevant though. Naturally, our construction is secure against these more restricted pirate coalitions, and also, our lower bound works in this more restricted case too, as the proof is based on very simple strategies ρ for cheating, where the ith digit of the output depends (in some randomized manner) only on the ith digits of their respective codewords (the ith column of the submatrix).

1.3 Earlier results

Fingerprinting was first studied by Wagner [18]. Fingerprinting resilient against pirate *coalitions* were studied by Blakley et al. [1]. Many different models for fingerprinting are studied in the literature, see for example Kilian et al. [8] for a model where the fingerprint can alter the document but the *distance* should be bounded.

IPP or *identifiable parent property codes* were introduced by Chor et al. [4]. These codes must work only against pirates who must output a pirated copy such that for any i the ith position of the pirated copy is identical to the ith position of a legitimate copy the pirates have access to. These codes and the related traitor tracing are widely studied, see e.g. [2, 9, 13, 14]. As we have already mentioned, this more restrictive assumption seems to be too strong in many fingerprinting applications. The *unreadable digit model* seems to be a more appropriate intermediate model. See Section 5 for the definition and for a comparison between these models.

The following is the standard argument to show that in our model, where the pirates could introduce arbitrary digits in positions their codewords differ, no deterministic fingerprint code exists for 3 players if any two of them can form a pirate coalition. Consider any three fingerprinted document X_1, X_2 and X_3 distributed to the players and let X be a document such that for any position i if the ith digits of at least two of X_1, X_2 and X_3 are some letter $s \in \Sigma$ then the ith digit of X is also s. (Over the binary alphabet X is determined by X_1, X_2 and X_3, it is their bitwise majority. Over larger alphabets X may not be determined uniquely by X_j but such X always exist.) No matter which two of the three users form the pirate coalition it is *possible* for them to come up with the pirated copy $y = X$. Thus no deterministic algorithm can accuse any of them for producing this copy without risking to accuse an innocent user. (In a related model Chung, Graham, and Leighton [5] get around this problem by accepting accusations of the form "two out of these three players are guilty" and even more complicated accusations for larger coalitions. But even in the model they study an exponential lower bound in the coalition size is immediate.)

Randomized fingerprint codes were introduced by Boneh and Shaw [3]. The fingerprint code they propose uses randomization in a restricted way. They first deterministically construct a code matrix and use randomization only for randomly permuting the columns of this matrix. We, on the other hand, use the full power of randomization allowed by Definition 1.1. Boneh and Shaw

constructed fingerprint codes of length $m = O(n^3 \log(n/\epsilon))$ for n users that are ϵ-secure against coalitions of any size. Against coalitions of size $c < n$ they constructed ϵ-secure fingerprint codes of length $m = O(c^4 \log(1/\epsilon) \log(n/\epsilon))$ for n users. In follow-up works Lindkvist [10] made minor improvement on the length not effecting the asymptotics, while Yacobi [17] designed a very efficient implementation of the Boneh-Shaw codes. The length of our codes presented in the next section is approximately the square root of the length of the Boneh-Shaw codes.

Dynamic traitor tracing was introduced by Fiat and Tassa [6]. This was originally a deterministic model requiring high alphabet size, but Tamir Tassa [16] introduced a more efficient probabilistic version. Tassa uses the Boneh-Shaw code as a primitive in his scheme. Substituting our codes presented in the next section substantially improves the convergence time of the Tassa scheme.

Boneh and Shaw also proves an $\Omega(c \log(1/(c\epsilon)))$ lower bound for the length of fingerprint codes. Our lower bound improves their bound significantly and matches the construction if ϵ is reasonably small. Peikert, Shelat, and Smith [11] prove a lower bound for a restricted type of fingerprint codes. Their bound is basically the same as our bound in Theorem 4, but it only applies for codes with a limited number of "column types". If all columns of the code matrix X differ their side condition on the number of column types is not met. For codes that use randomization in the limited way the Boneh-Show code does they prove that the original construction of Boneh and Show is almost optimal. The codes constructed in this paper do not satisfy the requirements needed for either of their bounds to apply. Nevertheless, the results in [11] also point toward the $c^2 \log(1/\epsilon)$ bound. The pirate strategy they employ in the proof is similar to our strategy, both are based on a carefully selected bias function.

The rest of the paper is organized as follows. In the next section we present our construction for fingerprint codes and summarize our results. In Sections 3 and 4 we prove Theorems 1 and 2, respectively, the two results stating the favorable properties of our fingerprint code. In Section 5 we introduce the *unreadable digit model* for fingerprinting, and we prove that any fingerprint code in the standard (arbitrary digit) model also works in this model. We state and prove our lower bound result (Theorem 5) for the unreadable digit model, and our original lower bound (Theorem 4) follows as a consequence. The matching length of the construction in the stronger model and the lower bound in the weaker establishes the equal strength of the two models. Similarly, the equal strength of fingerprint codes over binary and larger alphabets is a consequence of this lower bound. Section 6 contains a few concluding remarks.

2 Construction and results

Our main result is the construction of fingerprint codes of length $m = O(c^2 \log(n/\epsilon))$ that are ϵ-secure against coalitions of size c (see Corollary 3). After presenting the construction we motivate some of the seemingly arbitrary choices in it, then state its main properties in Theorems 1 and 2. These results are much

stronger than the requirements of Definition 1.1. A few comments on these added advantages of our codes follow. Theorem 4 states a lower bound for the length of fingerprint codes that matches our construction if the error bound ϵ is reasonably small.

2.1 The construction

In this paper, log always denotes the *natural logarithm.*

Let n and c be positive integers, $0 < \epsilon < 1$ and let $k = \lceil \log(1/\epsilon) \rceil$. We define the binary fingerprint code $F_{nc\epsilon}$ of length $m = 100c^2k$ for n users to be the following distribution over the pairs (X, σ).

We select the pair (X, σ) in two phases. First, let p_i be independent, identically distributed random variables from $[t, 1-t]$ for all $1 \leq i \leq m$. Here $t = 1/(300c)$ and $p_i = \sin^2 r_i$ is selected by picking uniformly at random the value $r_i \in [t', \pi/2 - t']$ with $0 < t' < \pi/4$, $\sin^2 t' = t$.

In the second phase, we select the code matrix X, by selecting each entry X_{ji} independently from the binary alphabet $\{0, 1\}$ with $P[X_{ji} = 1] = p_i$. Notice that independence of the entries X_{ji} holds only in the second phase, the overall random variables X_{ji} and $X_{j'i}$ are positively correlated as both of them tend to be 1 if p_i is large.

The accusation algorithm σ is determined by the values p_i and the matrix X, as follows. We define the n by m matrix U with entries

$$U_{ji} = \begin{cases} \sqrt{\frac{1-p_i}{p_i}} \text{ if } X_{ji} = 1, \\ -\sqrt{\frac{p_i}{1-p_i}} \text{ if } X_{ji} = 0. \end{cases}$$

Let σ accuse user j on the pirated copy $y \in \{0, 1\}^m$ as input if

$$\sum_{i=1}^{m} y_i U_{ji} > Z,$$

where $Z = 20ck$ is a threshold parameter. In other words, $\sigma(y)$ consists of the indices j for which the jth entry of Uy^T exceeds Z.

Remarks Having described the construction here we motivate some of our choices in it.

The formula defining U_{ji} is chosen so that after the first phase, it only depends on X_{ji}, it is positive if $X_{ji} = 1$ and it has expectation 0 and variance 1. For a motivation observe that having 1 as the ith digit in the pirated copy makes player j more suspicious if $X_{ji} = 1$ and less suspicious otherwise. Clearly, having a 1 in the pirated position where only a few players have that digit (p_i is small) makes these players even more suspicious.

Our choice of the distribution for p_i is biased toward the values close to 0 or 1 (as opposed to values close to 1/2). This is motivated by the marking condition. This is the only restriction on the pirates' strategy and it is more likely to apply

to these columns with a high bias. On the other hand, no fingerprint code can do totally without highly mixed columns, this is the basic idea of the Boneh-Shaw lower bound on the length of fingerprint codes.

Technically, the choice of the distribution of p_i is used only in the proof of "completeness" (Theorem 2) to show that the pirates' choice of strategy has only a minor effect on (an exponential average related to) their chance to be caught.

The cutoff points t and $1-t$ for the distribution on p_i are introduced for technical reasons. If p_i gets too close to either 0 or 1 then U_{ji} can have too high a positive or negative value and therefore this single position can have too much of an influence over the accusations.

The following two theorems bound the error probabilities of our codes $F_{nc\epsilon}$. Theorem 1 bounds the "soundness error" of accusing an innocent user, while Theorem 2 bounds the "completeness error" of not accusing any guilty one. For both theorems $n \geq c \geq 1$ and $0 < \epsilon < 1$ are arbitrary.

Theorem 1. *Let (X, σ) be distributed according to $F_{nc\epsilon}$. Let $j \in [n]$ be an arbitrary user, let $C \subseteq [n] \setminus \{j\}$ be a coalition of arbitrary size not containing j, and let ρ be any C-strategy. We have*

$$P[j \in \sigma(\rho(X))] < \epsilon.$$

Theorem 2. *Let (X, σ) be distributed according to $F_{nc\epsilon}$. Let $C \subseteq [n]$ be a coalition of size $|C| \leq c$, and let ρ be any C-strategy. We have*

$$P[C \cap \sigma(\rho(X)) = \emptyset] < \epsilon^{c/4}.$$

2.2 Advantages of the construction

Notice that Theorems 1 and 2 establish properties that are in most parts, much stronger then those required by the definition of ϵ-security. Most notably, Theorem 1 states that *innocent users are not likely to be accused even if all other users collaborate against them.* As the proof of Theorem 1 does not use the marking condition, innocent users are not likely to be accused even if the pirates can find the positions of the fingerprint code and thus they can break the marking condition. In other words, we can be reasonably sure that any accused user is a member of the group of pirates even if we do not know any bound on the size or power of this group. Our lower bound theorem (Theorem 4) tells us that $F_{nc\epsilon}$ (or any fingerprint code of the same length) cannot be secure against coalitions of size much larger than c. If much more than c users collaborate, they are able to come up with a strategy, so that in all likelihood none of them is accused. As Theorem 1 still applies, in this case *nobody* is accused. The distributor can use this property the following way. He chooses a reasonable value for c, assumes that at most c users collaborate and uses the fingerprint code $F_{nc\epsilon}$. Any user the code accuses will likely to be guilty regardless of the validity of the distributor's assumption. If the accusation algorithm accuses nobody,

this indicates that the pirate coalition is larger than c (or they can break the marking condition).

Another advantage of this code is the simple algorithm σ for accusation. To compute $\sigma(y)$ one only has to multiply y with a fixed matrix U and check which entries of the resulting vector exceed a threshold parameter Z.

Notice that the distributor does not need to know n in advance. Based only on c and ϵ, one can find the length m of the code and select the values p_i according to the required distribution. This is the first (preprocessing) phase. The next phase can be carried out separately (independently) for each user. Whenever a new user comes up, the distributor can generate his codeword (the corresponding row of the matrix X) and the rules for his accusation will be clear as the corresponding row of the matrix U is also defined.

Theorem 2 is stronger than required in its bound on the error probability. This is only of theoretical interest as the "soundness error" of Theorem 1 is higher and that type of error is considered worse.

An easy to fix weakness of Theorem 1 is that it bounds the probability of accusing a single innocent user and not the probability of accusing *some* innocent users. This is a natural consequence of the fact that the length of the code m does not depend on the number of users. If n is larger than 2^{m+1}, most users must share their codeword with another user, thus even if a single user distributes his copy, any algorithm that accuses him will accuse another user too. From Theorems 1 and 2 it clearly follows that

Corollary 3. *The fingerprint code $F_{nc\frac{\epsilon}{n}}$ is ϵ-secure against coalitions of size c if $c \geq 4$. The length of this code is $O(c^2 \log(n/\epsilon))$.*

2.3 The lower bound

Theorem 4. *Let F be a fingerprint code of length m over an arbitrary alphabet Σ for n users. Let $3 \leq c \leq n$ be an integer and $0 < \epsilon < 1/(100c^a)$ a real, where $a > 1$ is a constant. If F satisfies conditions (i) and (ii) below then*

$$m \geq d_a c^2 \log(1/\epsilon),$$

where $d_a > 0$ depends solely on a.

(i) For any coalition $C \subset [n]$ of size $|C| = c - 1$, for any C-strategy ρ, and for any user $j \in [n] \setminus C$, we have

$$P[j \in \sigma(\rho(X))] \leq \epsilon.$$

(ii) For any coalition $C \subseteq [n]$ of size $|C| = c$, and for any C-strategy ρ, we have

$$P[C \cap \sigma(\rho(X)) = \emptyset] < 0.99 \ .$$

While Theorems 1 and 2 claim properties much *stronger* than required for an ϵ-secure fingerprint code, our lower bound result, Theorem 4, assumes properties of the code that are somewhat *weaker* then those required for ϵ-secure

codes. This makes the matching lower and upper bounds even more interesting. Comparing the results of Theorems 1, 2, and 4 one can notice the following.

1. The length of our codes $F_{nc\epsilon}$ are optimal within a constant factor amongst codes satisfying the conditions of Theorem 4 if $\epsilon < 1/(100c^a)$ for a fixed $a > 1$. By Corollary 3, the length of the code $F_{nc\frac{\epsilon}{n}}$ is optimal within a constant factor amongst all codes for n users that are ϵ-secure against coalitions of size c if $\epsilon < 1/(100c^a)$ for a fixed $a > 1$ and $\epsilon < 1/n^b$ for a fixed $b > 0$. The assumption on ϵ seems to be reasonable, as in case $\epsilon \geq 1/c$, one can simply accuse everybody independently with probability ϵ and both conditions of Theorem 4 are satisfied with code length $m = 0$. See further remarks on ϵ-secure codes with high ϵ in Section 6.

2. Between the two types of error probabilities, the important one that has more effect on the code length is the "soundness error", the probability of accusing innocent users. The "completeness error", the probability of not accusing any of the pirates, can be arbitrarily chosen in a very wide interval without having a significant effect on the optimal code length. Making the completeness error vanish entirely seems to be difficult though. A code achieving that would mix some of the deterministic features of the IPP codes with the probabilistic properties of the Boneh-Shaw codes and the codes of this paper.

3. Our codes $F_{nc\epsilon}$ are binary, and have optimal length amongst codes over *arbitrary* alphabets. This answers the problem raised by Lindkvist [10] if binary codes are as good for fingerprinting as codes over larger alphabets. Lindkvist gives the same answer but only for a very limited class of fingerprint codes. This result is in sharp contrast with IPP codes that exist over larger alphabets, but do not exist over a binary alphabet.

4. In Section 5, we introduce another model for fingerprinting, the *unreadable digit model*, in which the pirates are more restricted in producing their illegitimate copy $\rho(X)$. In this model the pirates can put "unreadable digits" in the positions of the illegitimate copy where they detected disagreement but they cannot put a specific digit none of them has in that position. Naturally, Theorems 1 and 2 remain true in this model (the code $F_{nc\epsilon}$ remains secure against these more restricted pirate coalitions; see Lemma 5.3 for the precise statement). We prove the lower bound stated in Theorem 4 in the unreadable digit model (see Theorem 5) and get Theorem 4 as a corollary. Thus, we prove that the unreadable digit model and the model considered in this section are almost equivalent with respect to optimal code length. The importance of this fact comes from certain applications where the unreadable digit model seems to be more natural. If "digits" are implemented as complicated objects, we can safely assume that the pirates cannot create well-formed digits none of their documents contains, but they can simply put random noise in positions where they detected disagreement. For the distributor this random noise will be an unreadable digit.

5. The constant 0.99 in Theorem 4 is arbitrary. Our techniques work with $1 - \nu$ in place of 0.99 if $\epsilon < (\nu/c)^a$ for some constant $a > 1$. For $n = c$, $\epsilon = \nu/c$ we can have a code of length $m = 0$: simply accuse a random user, each with probability at most ϵ.

3 Why the innocent is not accused

Proving that our fingerprint code works consists of proving Theorems 1 and 2. In this section we prove Theorem 1, establishing that innocent users are not likely to be accused.

Proof of Theorem 1: Let n, c, ϵ, j, C, and ρ be as in the theorem. As $j \notin C$ we can consider performing the first phase of the construction of the $F_{nc\epsilon}$ codes (i.e., selecting the values p_i), performing the second phase for rows $j \in C$ (i.e., selecting the rows of X seen by ρ), and running the algorithm ρ all before selecting row j of X. This way $y = \rho(X)$ is fixed before the codeword of player j is selected. We claim that not only is the overall probability of the event $j \in \sigma(\rho(X))$ bounded by ϵ, but conditioned on any set of values p_i and y the probability of $j \in \sigma(y)$ is bounded by ϵ. Clearly, proving this stronger statement proves the theorem.

We have fixed values p_i from $[t, 1-t]$ and a fixed string $y \in \{0,1\}^m$. We choose X_{ji} from $\{0,1\}$ independently with $P[X_{ji} = 1] = p_i$ and define $u_i = U_{ji}$. Recall that $u_i = \sqrt{(1-p_i)/p_i}$ if $X_{ji} = 1$ and $u_i = -\sqrt{p_i/(1-p_i)}$ if $X_{ji} = 0$. Finally we set $S = \sum_{i=1}^{m} y_i u_i = \sum_{i:y_i=1} u_i$. User j is accused (i.e., $j \in \sigma(y)$) if $S > Z$, so we need to prove that $P[S > Z] < \epsilon$.

Consider the expected value $E[e^{\alpha S}]$ where e is the base of the natural logarithm, and $\alpha = 1/(10c)$. Using the independence of the random variables u_i we have

$$E\left[e^{\alpha S}\right] = E\left[\prod_{i:y_i=1} e^{\alpha u_i}\right] = \prod_{i:y_i=1} E\left[e^{\alpha u_i}\right].$$

Next we use $1+u \le e^u \le 1+u+u^2$, where the first inequality always holds, and the second inequality holds for $u < 1.7$. Notice that $u_i \le \sqrt{(1-t)/t} \le t^{-1/2}$ and thus $\alpha u_i < 1$. Using that u_i has expectation zero and variance 1 we get

$$E\left[e^{\alpha u_i}\right] \le E\left[1 + \alpha u_i + \alpha^2 u_i^2\right]$$

$$= 1 + \alpha E\left[u_i\right] + \alpha^2 E\left[u_i^2\right] = 1 + \alpha^2 \le e^{\alpha^2}.$$

$$E\left[e^{\alpha S}\right] = \prod_{i:y_i=1} E\left[e^{\alpha u_i}\right] \le \left(e^{\alpha^2}\right)^{|\{i:y_i=1\}|} \le e^{\alpha^2 m}.$$

Finally by the Markov inequality we have

$$P\left[S > Z\right] = P\left[e^{\alpha S} > e^{\alpha Z}\right] < \frac{E\left[e^{\alpha S}\right]}{e^{\alpha Z}} \le e^{\alpha^2 m - \alpha Z}.$$

Here the exponent is $\alpha^2 m - \alpha Z = -k = -\lceil \log(1/\epsilon) \rceil$ thus

$$P[S > Z] < e^{-k} \le \epsilon$$

as claimed. □

4 Why some guilty is accused

In this section we turn to Theorem 2 stating that our fingerprint code accuses one of the pirates with very high probability.

Proof of Theorem 2: Let n, c, ϵ, C, and ρ be as in the theorem. We assume without loss of generality that $C = [n]$, $n \leq c$ as the codewords of the users outside C are irrelevant. Let (X, σ) be distributed according to the code $F_{nc\epsilon}$. Here σ is determined by X and $\overline{p} = (p_1, \ldots, p_m)$. For simplicity we introduce $q_i = \sqrt{(1-p_i)/p_i}$ and recall the definition $U_{ji} = q_i$ if $X_{ji} = 1$ and $U_{ji} = -1/q_i$ if $X_{ji} = 0$. Let us set $y = \rho(X)$ and $S_j = \sum_{i=0}^{m} y_i U_{ji}$ for $j \in C$. Let

$$S = \sum_{j \in C} S_j = \sum_{i=1}^{m} y_i \left(x_i q_i - \frac{n - x_i}{q_i} \right), \tag{1}$$

where $x_i = \sum_{j=1}^{n} X_{ji}$ denotes the number of ones in column i of X. Recall that $j \in C$ is accused (i.e., $j \in \sigma(y)$) if $S_j > Z$. Thus if $S > nZ$ at least one pirate in C must be accused. It is enough to bound the probability

$$P[C \cap \sigma(\rho(X)) = \emptyset] \leq P[S \leq nZ].$$

The high level description of the proof is as follows. If the pirates would be able to produce $y = \rho(X)$ consisting of all zeros then we would have $S = 0$ and in particular our algorithm σ would accuse nobody. Unfortunately for the pirates, for indices i such that column i of X consists of all ones they must output $y_i = 1$ by the marking condition, and this definitely increases S. They may try to offset this increase by outputting some ones at indices i where column i of X is mixed. By outputting 1 they decrease S if the column contains fewer than the expected number of $p_i n$ ones, and increase S if the column contains more than that many ones. They know the number x_i of ones of the column but they do not know p_i. We chose the distribution of p_i such that the wins and the losses almost cancel out and their choice for y_i has almost no effect on the expectation of S (more precisely on the exponential average $E[e^{-\alpha S}]$ for a suitable α). The increase coming from the all one columns is thus impossible to offset, and it is enough to make $S > nZ$ with very high probability.

We set $\alpha = 1/(20c)$. In the first part of the proof we study the exponential average $E[e^{-\alpha S}]$ and in Equation (2) we find the largest value it can take for any C-strategy ρ. In the second part of the proof we study that formula closely. We establish that the two formulae of which M_x in Equation (2) is the maximum are very close to each other for $1 \leq x \leq n-1$. This represents establishing that the choice of the C-strategy ρ has only a minor effect on the expectation. Bounding the $x = n$ case corresponds to calculating the effect on S of the all one columns. In Equation (4) we establish a simple bound on the exponential average. We finish the proof by bounding the probability of $S \leq nZ$ (and thus the chance that nobody is accused by σ) using the Markov inequality.

Using the rules of the second phase of the code generation we have

$$E_{\overline{p},X}\left[e^{-\alpha S}\right] = E_{\overline{p}}\left[\sum_X \left(e^{-\alpha S}\prod_{i=1}^{m}\left(p_i^{x_i}(1-p_i)^{n-x_i}\right)\right)\right]$$

$$= \sum_X E_{\overline{p}}\left[e^{-\alpha S}\prod_{i=1}^{m}\left(p_i^{x_i}(1-p_i)^{n-x_i}\right)\right].$$

The expectation in this formula is for the choice of $\overline{p}$ in the first phase or for $\overline{p}$ and X as generated in the first and second phases of generating $F_{nc\epsilon}$ as indicated. The summation is for all n by m 0-1 matrices X. The number of ones in column i of X is denoted by x_i. Using Equation (1) we have

$$E_{\overline{p},X}[e^{-\alpha S}] = \sum_X E_{\overline{p}}\left[\prod_{i=1}^{m}\left(p_i^{x_i}(1-p_i)^{n-x_i}e^{-\alpha y_i(x_i q_i - \frac{n-x_i}{q_i})}\right)\right].$$

Here x_i and $y = \rho(X)$ is determined by X, while $q_i = \sqrt{(1-p_i)/p_i}$ is determined by $\overline{p}$. Notice that for fixed X term i of the product depends solely on p_i, thus these terms are independent. We have

$$E_{\overline{p},X}[e^{-\alpha S}] = \sum_X \prod_{i=1}^{m} E_{p_i}\left[p_i^{x_i}(1-p_i)^{n-x_i}e^{-\alpha y_i(x_i q_i - \frac{n-x_i}{q_i})}\right].$$

The expectation is taken for the random variable p_i. As each p_i is identically distributed we write p instead, where p is identically distributed with each p_i. We let $q = \sqrt{(1-p)/p}$. Each y_i is either 0 or 1, furthermore if $x_i = 0$ then $y_i = 0$ and if $x_i = n$ then $y_i = 1$ by the marking condition. Thus we have

$$E_{\overline{p},X}[e^{-\alpha S}] \leq \sum_X \prod_{i=1}^{m} \max^*(N_{0,i}, N_{1,i}),$$

where

$$N_{0,i} = E_p\left[p^{x_i}(1-p)^{n-x_i}\right],$$

$$N_{1,i} = E_p\left[p^{x_i}(1-p)^{n-x_i}e^{-\alpha(x_i q - \frac{n-x_i}{q})}\right],$$

and $\max^*$ denotes the first term $N_{0,i}$ if $x_i = 0$, the last term $N_{1,i}$ if $x_i = n$ and the maximum of the two terms otherwise. Notice that this last bound does not depend on the C-strategy ρ, and as the only assumption on $y = \rho(X)$ is the marking condition we have equality for some C-strategy ρ. As term i of the product only depends on x_i and the summation is for all 0-1 matrices X we can switch the summation and the product to get

$$E_{\overline{p},X}[e^{-\alpha S}] \leq \prod_{i=1}^{m}\sum_{x_i=0}^{n}\binom{n}{x_i}\max^*(N_{0,i}, N_{1,i}).$$

We conclude this part of the proof by the bound (still tight for some C-strategy ρ):

$$E_{\overline{p},X}[e^{-\alpha S}] \leq \left(\sum_{x=0}^{n} \binom{n}{x} M_x\right)^m, \tag{2}$$

where

$$M_0 = E_{0,0}, \quad M_n = E_{1,n},$$
$$M_x = max(E_{0,x}, E_{1,x}) \text{ for } 1 \leq x \leq n-1,$$

and for $0 \leq x \leq n$

$$E_{0,x} = E_p\left[p^x(1-p)^{n-x}\right],$$
$$E_{1,x} = E_p\left[p^x(1-p)^{n-x} e^{-\alpha\left(xq - \frac{n-x}{q}\right)}\right].$$

Here p is distributed as p_i in the construction and $q = \sqrt{(1-p)/p}$.

We use $e^u \leq 1 + u + u^2$ that holds for $u < 1.7$ to bound the exponential term in $E_{1,x}$. If $-\alpha(xq - (n-x)/q) < 1.7$ we have

$$e^{-\alpha\left(xq - \frac{n-x}{q}\right)} \leq 1 - \alpha\left(xq - \frac{n-x}{q}\right) + \alpha^2\left(xq - \frac{n-x}{q}\right)^2.$$

We make the bound work for all q by adding the extra term $\chi_x(p) e^{\alpha(n-x)/\sqrt{1-p}}$ to the right hand side. Here $\chi_x(p)$ is the characteristic function of the event $p \geq 1 - \alpha^2(n-x)^2$, which is implied by $-\alpha(xq - (n-x)/q) > 1$.

We remark that in the preliminary version [15] of this paper we chose $\alpha = \sqrt{t}/c$. This makes $-\alpha(xq - (n-x)/q) < 1$ always hold, thus we could make the proof simpler by getting rid of the term $\chi_x(p)$. Unfortunately, this small value for α makes the computation yield a somewhat weaker error bound. This weaker error bound is still more than enough to imply Corollary 3 for high c, but we strive here for the strongest bounds achievable by these methods.

We have

$$\begin{aligned} p^x(1-p)^{n-x} e^{-\alpha\left(xq - \frac{n-x}{q}\right)} \quad \leq \quad & p^x(1-p)^{n-x} - \\ & - \alpha p^x(1-p)^{n-x}\left(xq - \frac{n-x}{q}\right) + \\ & + \alpha^2 p^x(1-p)^{n-x}\left(xq - \frac{n-x}{q}\right)^2 + \\ & + \chi_x(p)(1-p)^{n-x} e^{\frac{\alpha(n-x)}{\sqrt{1-p}}}. \end{aligned}$$

Taking expectations we get

$$E_{1,x} \leq E_{0,x} - \alpha F_{1,x} + \alpha^2 F_{2,x} + R_x,$$

where

$$F_{1,x} = E_p\left[p^x(1-p)^{n-x}\left(xq - \frac{n-x}{q}\right)\right],$$

$$F_{2,x} = E_p\left[p^x(1-p)^{n-x}\left(xq - \frac{n-x}{q}\right)^2\right] \geq 0,$$

$$R_x = E_p\left[\chi_x(p)(1-p)^{n-x}e^{\frac{\alpha(n-x)}{\sqrt{1-p}}}\right] \geq 0.$$

The term $F_{1,x}$ is the most important. Our choice of the distribution for p makes sure that it is small for $1 \leq x \leq n-1$. The specific choice of the distribution is used for this bound only. Recall that $p = \sin^2 r$ with a uniform random $r \in [t', \pi/2 - t']$, where $\sin^2 t' = t$. We have $1 - p = \cos^2 r$, $q = \cot r$ and

$$F_{1,x} = \frac{1}{\pi/2 - 2t'}\int_{t'}^{\pi/2-t'} \sin^{2x} r \cos^{2n-2x} r(x\cot r - (n-x)\tan r)dr.$$

Notice that the primitive function of the integrand is $f(r) = 1/2\sin^{2x} r\cos^{2n-2x} r$, thus we have

$$F_{1,x} = \frac{f(\pi/2 - t') - f(t')}{\pi/2 - 2t'} = \frac{t^{n-x}(1-t)^x - t^x(1-t)^{n-x}}{\pi - 4t'}.$$

For this calculation the choice $t = 0$ (no cutoff) would be optimal yielding $F_{1,x} = 0$ for $1 \leq x \leq n-1$. We need $t > 0$ in other calculations of this proof and also in the proof of Theorem 1. The choice $t = 1/(300c)$ is a compromise yielding a small but nonzero value for $F_{1,x}$. For $1 \leq x \leq n-1$ we use

$$F_{1,x} \geq -\frac{t^x(1-t)^{n-x}}{\pi - 4t'} < 0$$

and get

$$M_x = \max(E_{0,x}, E_{1,x}) \leq E_{0,x} + \alpha\frac{t^x(1-t)^{n-x}}{\pi - 4t'} + \alpha^2 F_{2,x} + R_x.$$

We also have $M_0 = E_{0,0}$ and since $F_{1,n} = \frac{(1-t)^n - t^n}{\pi - 4t'}$

$$M_n = E_{1,n} \leq E_{0,n} - \alpha\frac{(1-t)^n - t^n}{\pi - 4t'} + \alpha^2 F_{2,n} + R_n.$$

Next we estimate the summation in Equation (2)

$$\begin{aligned}\sum_{x=0}^{n}\binom{n}{x}M_x \quad \leq \quad & \sum_{x=0}^{n}\binom{n}{x}E_{0,x} - \\ & - \alpha\frac{(1-t)^n - \sum_{x=1}^{n}\binom{n}{x}t^x(1-t)^{n-x}}{\pi - 4t'} + \\ & + \alpha^2\sum_{x=0}^{n}\binom{n}{x}F_{2,x} + \sum_{x=0}^{n}\binom{n}{x}R_x. \qquad (3)\end{aligned}$$

We bound each term separately:

$$\begin{aligned}\sum_{x=0}^{n}\binom{n}{x}E_{0,x} &= \sum_{x=0}^{n}\binom{n}{x}E_p\left[p^x(1-p)^{n-x}\right]\\ &= E_p\left[\sum_{x=0}^{n}\binom{n}{x}p^x(1-p)^{n-x}\right]\\ &= E_p[1]=1;\end{aligned}$$

$$(1-t)^n-\sum_{x=1}^{n}\binom{n}{x}t^x(1-t)^{n-x}=2(1-t)^n-1\geq 1-2nt;$$

$$\begin{aligned}\sum_{x=0}^{n}\binom{n}{x}F_{2,x} &= \sum_{x=0}^{n}\binom{n}{x}E_p\left[p^x(1-p)^{n-x}\left(xq-\frac{n-x}{q}\right)^2\right]\\ &= E_p\left[\sum_{x=0}^{n}\binom{n}{x}p^x(1-p)^{n-x}\left(xq-\frac{n-x}{q}\right)^2\right]\end{aligned}$$

To further simplify this expression let $0<p<1$ be fixed and consider the independent identically distributed random variables U_j for $j\in[n]$ with $P[U_j=q]=p$ and $P[U_j=-1/q]=1-p$, where $q=\sqrt{(1-p)/p}$. These random variables have expectation 0 and variance 1, so we have $E[(\sum_{j=1}^{n}U_j)^2]=n$. This expectation is for a fixed p with respect to the random variables U_j. Formally spelling out this expectation yields exactly the formula *inside the expectation for* p in the last displayed equation. This implies

$$\sum_{x=0}^{n}\binom{n}{x}F_{2,x}=E_p[n]=n.$$

For the last error term

$$R_x=E_p\left[\chi_x(p)(1-p)^{n-x}e^{\alpha\frac{n-x}{\sqrt{1-p}}}\right]$$

we have $R_x=0$ for $x>n-\sqrt{t}/\alpha$ as in this case $\chi_x(p)=0$ for $p\in[t,1-t]$. (Recall that choosing α somewhat smaller we can get rid of this error term entirely.) For any $0\leq x\leq n$ the function $(1-p)^{n-x}e^{\alpha(n-x)/\sqrt{1-p}}$ is monotone decreasing in $[t,1-t]$, so for $x\leq n-\sqrt{t}/\alpha$ the maximum of $\chi_x(p)(1-p)^{n-x}e^{\alpha(n-x)/\sqrt{1-p}}$ for $p\in[t,1-t]$ is at $p=1-\alpha^2(n-x)^2$. Thus we have

$$R_x\leq e(\alpha(n-x))^{2(n-x)}.$$

Using $\binom{n}{x}\leq(\frac{ne}{n-x})^{n-x}$ and $n\alpha\leq c\alpha=1/20$ we get

$$\sum_{x=0}^{n}\binom{n}{x}R_x \leq \sum_{x=0}^{\lfloor n-\sqrt{t}/\alpha\rfloor}\left(\frac{ne}{n-x}\right)^{n-x}e(\alpha(n-x))^{2(n-x)}$$

$$
\begin{aligned}
&= e \sum_{x=0}^{\lfloor n-\sqrt{t}/\alpha \rfloor} (e^3 n(n-x)\alpha^2)^{n-x} \\
&< e \sum_{x=0}^{\lfloor n-\sqrt{t}/\alpha \rfloor} 19^{-(n-x)} \\
&< 3 \cdot 19^{-\lceil \sqrt{t}/\alpha \rceil}.
\end{aligned}
$$

Now we can estimate each term in Equation (3):

$$
\sum_{x=0}^{n} \binom{n}{x} M_x < 1 - \alpha \frac{1-2nt}{\pi - 4t'} + \alpha^2 n + 3 \cdot 19^{-\lceil \sqrt{t}/\alpha \rceil} < 1 - \alpha/4.
$$

We used $n \leq c$, $\alpha = 1/(20c)$ and $t = 1/(300c)$ here. To be honest the last inequality above works for $c \geq 7$ only. For smaller values of c one needs to compute R_x exactly to prove the estimate.

By Equation (2) we have

$$
E_{\overline{p},X}[e^{-\alpha S}] < (1-\alpha/4)^m < e^{-\alpha m/4}. \tag{4}
$$

By the Markov inequality and $m = 100c^2k$, $Z = 20ck$ we get

$$
P[S \leq nZ] \leq P[S \leq cZ] \leq \frac{e^{-\alpha m/4}}{e^{-\alpha cZ}} = e^{-\alpha(m/4 - cZ)} \leq \epsilon^{c/4}.
$$

As mentioned in the beginning of this proof, if $S > nZ$ then $C \cap \sigma(\rho(X))$ is not empty, so the above bound proves the theorem. □

5 The unreadable digit model—the lower bound on code length

In this section we give the definition of the unreadable digit model of fingerprinting, which we have already mentioned in Sections 1 and 2. We compare it to the standard (arbitrary digit) model in Lemma 5.3. We prove our lower bound on the code length in this model, see Theorem 5. The lower bound in the standard model (Theorem 4) follows as a corollary. Note that for binary codes the two models are trivially equivalent.

Definition 5.1. *A* unreadable digit fingerprint code *of length m for n users over the alphabet Σ is a distribution over the pairs (X, σ), where X is an n by m matrix over Σ and σ is an algorithm that takes a string $y \in \Sigma'^m$ (the illegitimate copy) as input, and produces a subset $\sigma(y) \subseteq [n] := \{1, 2, \ldots, n\}$ (the set of accused users). Here $\Sigma' = \Sigma \cup \{?\}$, where $? \notin \Sigma$ represents the unreadable digit. For $\emptyset \neq C \subseteq [n]$ an* unreadable digit C-strategy *is an algorithm ρ that takes the submatrix of X formed by the rows with indices in C as input, and produces a string $y = \rho(X) \in \Sigma'^m$ as output and satisfies the following (strong marking)*

conditions. For all positions $1 \leq i \leq m$ *the digit* y_i *is either* ? *or one of the digits* X_{ji} *with* $j \in C$*. Furthermore, if for some* i *all the values* X_{ji} *for* $j \in C$ *agree then* $y_i \neq ?$*. We say that an unreadable digit fingerprint code is* ϵ-secure against coalitions of size c *if for any* $C \subseteq [n]$ *of size* $|C| \leq c$ *and any unreadable digit* C*-strategy* ρ *the error probability*

$$P[\sigma(\rho(X)) = \emptyset \text{ or } \sigma(\rho(X)) \not\subseteq C]$$

is at most ϵ*.*

At first the unreadable digit model may appear to be incomparable with the arbitrary digit model. It introduces a new possibility (creating unreadable digits) for the pirates and simultaneously restricts their choices with respects to digits in Σ. A closer look will tell however, that the pirates can replace any unreadable digit with any fixed digit $a \in \Sigma$ without increasing their chance to be caught. This simple observation is formalized in Lemma 5.3. We start with some definitions.

Definition 5.2. *Let* Σ *be a finite alphabet and let* $\Sigma' = \Sigma \cup \{?\}$ *for some* $? \notin \Sigma$ *and let* $a \in \Sigma$ *be arbitrary. Let us denote by* f_a *the transformation* $f_a : \Sigma'^* \to \Sigma^*$ *that replaces each occurrence of* ? *by* a *and leaves all other digits unchanged. Let* F *be an (arbitrary digit) fingerprint code over the alphabet* Σ*. By* F_a *we denote the unreadable digit fingerprint code* $(X, \sigma \circ f_a)$ *where* (X, σ) *is distributed according to* F*.*

Lemma 5.3. *If an (arbitrary digit) fingerprint code* F *over the alphabet* Σ *is* ϵ*-secure against any coalition of size* c *then the unreadable bit fingerprint code* F_a *(for* $a \in \Sigma$*) is also* ϵ*-secure against any coalition of size* c*. Moreover, if* C *is an arbitrary coalition and* j *is an arbitrary user then we have*

$$\max_{\rho} P[j \in \sigma(\rho(X))] \geq \max_{\rho'} P[j \in \sigma'(\rho'(X))],$$

$$\max_{\rho} P[C \cap \sigma(\rho(X)) = \emptyset] \geq \max_{\rho'} P[C \cap \sigma'(\rho'(X)) = \emptyset],$$

where the maxima are taken over C*-strategies* ρ*, and unreadable digit* C*-strategies* ρ'*, while the probabilities are according to the distributions* F *on* (X, σ) *and* F_a *on* (X, σ')*.*

Proof: All the complicated looking statements of the lemma follow from the simple observation, that for an unreadable digit C-strategy ρ' the function $\rho = f_a \circ \rho'$ is an (arbitrary digit) C-strategy and $\sigma(\rho(X)) = \sigma'(\rho'(X))$ for any X if $\sigma' = \sigma \circ f_a$. □

Lemma 5.3 tells us that the arbitrary digit model (studied in most of this paper) demands more of a fingerprint code than the unreadable digit model. In particular, the fingerprint code $F_{nc\epsilon}$ can be trivially extended to a unreadable digit fingerprint code $(F_{nc\epsilon})_0$ (we simply treat unreadable digits as zeros), and this code satisfies all the nice properties stated in Theorems 1, 2 and Corollary 3. Also by Lemma 5.3 the arbitrary digit and the unreadable digit models

are equivalent over a binary alphabet. Over larger alphabets such a direct equivalence does not hold but Lemma 5.3 tells us which model is stronger. Lindkvist [10] studied the relative power of fingerprinting over binary and larger alphabets and concluded that for a severely limited class of fingerprint codes a binary alphabet is just as powerful as arbitrary alphabets are. The main results of this paper (Theorems 1, 2, 5) answer both these questions in full generality: for reasonable error parameters the optimal code length is the same within a constant factor for both models of fingerprinting and over an alphabet of arbitrary size at least two.

To make our lower bound in Theorem 4 work in both models we state it in the unreadable digit model. By Lemma 5.3 Theorem 4 follows.

Theorem 5. *Let F be an unreadable digit fingerprint code of length m over an arbitrary alphabet Σ for n users. Let $3 \leq c \leq n$ be an integer and $0 < \epsilon < 1/(100c^a)$ a real, where $a > 1$ is a constant. If F satisfies the conditions (i) and (ii) below, then*

$$m \geq d_a c^2 \log(1/\epsilon),$$

where $d_a > 0$ depends solely on a.

(i) For any coalition $C \subset [n]$ of size $|C| = c-1$, any unreadable digit C-strategy ρ, and any user $\ell \in [n] \setminus C$

$$P[\ell \in \sigma(\rho(X))] \leq \epsilon.$$

(ii) For any coalition $C \subseteq [n]$ of size $|C| = c$ and any unreadable digit C-strategy ρ

$$P[C \cap \sigma(\rho(X)) = \emptyset] < 0.99$$

Independent of our paper Peikert, Shelat, and Smith in [11] prove a numerically almost identical lower bound for the length of binary fingerprint codes. Their result only applies to a limited class of codes with a strong bound on the number of column types: the number of non-equal columns of the matrix X produced by the code. In the code constructed by Boneh and Shaw [3] the number of columns types were severely limited. Our construction typically yields matrices with all the columns different. For such codes the bound in [11] is not applicable. Nevertheless, some of the techniques of the lower bound proofs in [11] and in this paper are similar and we shall comment on these similarities.

As the proof uses an esoteric measure of distance for distributions (Rényi divergence) we motivate the choice here.

Assume we have a fingerprint code (X, σ) satisfying conditions (i) and (ii) of Theorem 5. We concentrate on the set $[c]$ of the first c users only. Consider the pirate coalition $C_\ell = [c] \setminus \{\ell\}$ containing all these users but user $\ell \in [c]$. Our goal is to give a *randomized* C_ℓ-strategy ρ_ℓ to this coalition such that the output $\rho_\ell(X)$ of this strategy is almost the same for all $\ell \in [c]$. Here we think of X and σ as being fixed, and the randomization coming from the randomized strategy ρ_ℓ. (This simplification is not fully justified and will not be used in the

formal proof. Instead, we ensure the distributions of the triples $(X, \sigma, \rho_\ell(X))$ are close to each other.)

The randomized strategies ρ_ℓ we use are very simple. We call this type of randomized strategies *bias strategies*. In a bias C-strategy ρ the pirates decide independently for each digit y_i of $y = \rho(X)$ if it is ? or the most popular digit s_i they see on position i. The probability of $y_i = s_i$ is determined by a *bias function* based on how many of the pirates in C see s_i at position i in their codewords. The bias function must give $P[y_i = s_i] = 1$ if all of their codewords agree at position i to satisfy the marking condition, while if the most popular digit is not seen in the majority of the rows we have $P[y_i =?] = 1$ to accommodate for the case when the most popular digit is not unique.

Let us first see what happens if *identically* distributed outputs $\rho_\ell(X)$ were possible to achieve. What is the accused set $\sigma(\rho_\ell(X))$ in this case? According to condition (i) this set does not contain any of the players with probability more than ϵ, but according to (ii) it contains one of them with probability at least $1/100$. The contradiction is clear as $\epsilon < 1/(100c)$. Unfortunately, identical distributions are impossible to achieve as for $\ell \neq \ell'$ the pirate coalition C_ℓ and $C_{\ell'}$ may see a different number of the most popular digit at any given position. Fortunately, this difference is bounded by 1. Thus, we have to study some kind of distance between these distributions.

The proof technique of [11] is almost identical to ours up to this point. They use the same pirate coalitions, similar very simple strategies and even their bias function is similar to ours. Their solution to the non-identical distributions obtained is to designate a target distribution (the uniform distribution on the so called ideal words) and they prove that with some small probability the output will hit this target distribution. However, this small probability exponentially decreases with the number of column types, and it becomes useless if there are too many column types.

The simplest measure to consider is the usual distance in distribution. This is the maximal difference in the probabilities of any event according to the two distributions. It is easy to verify, that no matter how we choose the bias function the difference of 1 in the number of appearances of the most popular digit may cause a difference $1/c$ in the probability $P[y_i =?]$. Thus, the distribution of a single digit may differ by as much as $1/c$. The distance of the total distributions $\rho_\ell(X)$ and $\rho_{\ell'}(X)$ is at most the sum of these distances as in each pirated copy each digit is independent (recall, that we consider X fixed). Thus, if the number of positions $m = o(c)$ the resulting total distributions are close to each other. But for $m > c$ this approach gives nothing.

A better choice for the distance measure is the information theoretic *divergence*. For technical reasons we must consider divergence from a common target distribution (obtained by the coalition of all c players by a similar bias strategy). With the correct choice of the bias function one can guarantee that any individual digit contributes only $O(1/c^2)$ to the divergence. (This phenomenon can be best understood through the following example: Suppose you have a biased coin: it gives heads with probability $1/2 + 1/c$. The distribution of your coin is in distance $1/c$ from the fair distribution but you need $\Theta(c^2)$ coin flips

to realize the bias.) These divergences add up for the independent positions. Unfortunately, the properties (i) and (ii) of Theorem 5 do not guarantee a high divergence for the total distributions $\rho_\ell(X)$, these divergences can be as low as $O(\log(1/\epsilon)/c)$. We are thus back at a linear bound.

The correct choice of the distance measure is the higher order *Rényi divergence*. This esoteric version of informational divergence was introduced by Alfréd Rényi in [12]. It has seldom been used since. Again, we have to measure Rényi divergence from a common target distribution $\rho(X)$ that is obtained by the coalition of all c players by a similar bias $[c]$-strategy. Rényi divergence still has the property that each digit contributes $O(1/c^2)$, and these contributions simply add up for the independent digits. But with the correct choice of the parameters conditions (i) and (ii) of Theorem 5 now imply a total divergence of $\Omega(\log(1/\epsilon))$ between $\rho(X)$ and $\rho_\ell(X)$ for at least one value of ℓ. See the detailed calculation in the proof of Theorem 5.

Definition 5.4. *Rényi divergence of order $\alpha+1$ ($\alpha > 0$) between the discrete random variables Q and R is defined as*

$$H_{\alpha+1}(Q||R) = \frac{1}{\alpha}\log\left(\sum_x \frac{(P[Q=x])^{\alpha+1}}{(P[R=x])^\alpha}\right),$$

where the summation extends over the values x taken by the random variable Q with positive probability. The divergence is only defined if all these values are also taken with positive probability by the random variable R.

Notice that these divergences depend only on the separate distributions of the variables Q and R and not on their joint distribution. The following basic properties of Rényi divergences are well known and have straightforward one line proofs.

(a) If Q_1 and Q_2 are independent and R_1 and R_2 are independent, then

$$H_{\alpha+1}((Q_1,Q_2)||(R_1,R_2))$$
$$= H_{\alpha+1}(Q_1||R_1) + H_{\alpha+1}(Q_2||R_2).$$

(b)
$$e^{\alpha H_{\alpha+1}((Q,S)||(R,S))} = E[e^{\alpha H_{\alpha+1}((Q|S=s_0)||(R|S=s_0))}],$$

where the expectation is taken for the value s_0 of the random variable S.

(c) For any function f

$$H_{\alpha+1}(f(Q)||f(R)) \le H_{\alpha+1}(Q||R).$$

(d) If the random variables Q and R take values from $\{0,1\}$ and $P[Q=1]=q$, $P[R=1]=s$, $0<s<1$, then

$$H_{\alpha+1}(Q||R) \ge \frac{1}{\alpha}\log\left(\frac{q^{\alpha+1}}{s^\alpha}\right).$$

Furthermore, if $q/s < 10$, $(1-q)/(1-s) < 10$, then

$$H_{\alpha+1}(Q||R) = O\left(\frac{(q-s)^2}{s(1-s)}\right),$$

where the constant hidden in the O notation depends only on α.

Using the properties above we make the proof of the Theorem 5 outlined above formal.

Proof of Theorem 5: We start by describing the pirate coalitions C and the unreadable digit bias C-strategies the proof is based on. We concentrate on the first c users only. We apply condition (i) for $\ell \in [c]$, $C_\ell = [c] \setminus \{\ell\}$ and the probabilistic unreadable digit C_ℓ-strategy ρ_ℓ defined below. We use condition (ii) with the coalition $C_0 = [c]$ and the probabilistic unreadable digit C_0-strategy ρ_0 defined below.

Let us start with the $c = 3$ case, here the unreadable digit C_ℓ-strategies are simpler and somewhat different from the $c > 3$ case. For ρ_0 we take the deterministic algorithm producing $y = \rho_0(X)$, with ith digit $(1 \le i \le m)$ $y_i = s \in \Sigma$ if $X_{ji} = s$ for at least two of the three indices $j \in C_0$ and $y_i =?$ if all three values X_{ji} for $j \in C_0$ are distinct. For ρ_ℓ with $\ell \in [c]$ we take the randomized algorithm producing each output allowed for an unreadable digit C_ℓ-strategy with equal probability. In other words, for $\ell \in [c]$ the digits of the output $y = \rho_\ell(X)$ are independent and for $1 \le i \le m$ $y_i = s$ if $X_{ji} = s$ for both $j \in C_\ell$, and if the values X_{ji} are different for the two $j \in C_\ell$, then y_i takes one of the these two values or ? with probability 1/3 each.

For the definition of the strategies ρ_ℓ for $c > 3$ we need some preparations.

Let the real function f be defined by $f(x) = 0$ if $x \le 0$, $f(x) = 3x^2 - 2x^3$ if $0 \le x \le 1$, and $f(x) = 1$ for $x \ge 1$. This function was chosen for the following property that can be easily verified.

(*) If the reals u and v satisfy $0 < v < 1$, $u \le 3v$ and $1 - u \le 3(1-v)$ then $f(u) \le 9f(v)$, $1 - f(u) \le 9(1 - f(v))$ and

$$\frac{(f(u) - f(v))^2}{f(v)(1-f(v))} = O\left((u-v)^2\right).$$

For $0 \le \ell \le c$ and $1 \le i \le m$ let k_i^ℓ be the maximum multiplicity of a digit among the digits X_{ji} with $j \in C_\ell$, and let s_i^ℓ be one of the digits with this multiplicity.

We define the bias unreadable digit C_ℓ-strategy ρ_ℓ for $0 \le \ell \le c \ge 4$ with the following rules. For fixed X and ℓ, the digits of $y = \rho_\ell(X)$ are chosen independently from $y_i \in \{s_i^\ell, ?\}$ with

$$P[y_i = s_i^\ell] = \begin{cases} f\left(\frac{2k_i^\ell - c + 1}{c-1}\right) & \text{if } \ell > 0 \\ f\left(\frac{2k_i^0 - c - 1}{c-3}\right) & \text{if } \ell = 0. \end{cases}$$

To check that ρ_ℓ is indeed an unreadable digit C_ℓ-strategy we need to check the marking condition: if $k_i^\ell = |C_\ell|$ then $y_i = s_i^\ell$. Notice that $y_i = s_i^\ell$ happens with positive probability only if s_i^ℓ is the absolute majority of the digits X_{ji} with $j \in C_\ell$, and in this case there is no ambiguity in the definition of s_i^ℓ.

By condition (ii) $P[C \cap \sigma(\rho_0(X)) = \emptyset] < 0.99$. Here the probability is according to the distribution F on (X, σ) and according to the random choices taken in ρ_0. Thus there is a user $j \in [c]$ accused by $\sigma(\rho_0(X))$ with probability more than $1/(100c)$. Assume without loss of generality that this is true for user 1 and we have

$$P[1 \in \sigma(\rho_0(X))] > \frac{1}{100c}. \tag{5}$$

We contrast this with the bound given by condition (i):

$$P[1 \in \sigma(\rho_1(X))] \leq \epsilon. \tag{6}$$

Our goal is to finish the proof by showing that if the code F is not long enough then the distributions $(\rho_0(X), X, \sigma)$ and $(\rho_1(X), X, \sigma)$ are too close to each other to let the separation stated in Inequalities (5) and (6) happen.

Let α be a positive parameter to be set later depending solely on the constant a in the theorem.

Let us first consider the random variables $y = \rho_0(X)$ and $y' = \rho_1(X)$ for an arbitrary fixed X.

For $c = 3$ it is straightforward to see, that

$$H_{\alpha+1}(y_i||y_i') = O(1) = O(1/c^2).$$

Our first goal is to prove a similar bound for $c > 3$.

Suppose $c > 3$. For $1 \leq i \leq m$ we have $y_i \in \{s_i^0, ?\}$ and $P[y_i = s_i^0] > 0$ only if $k_i^0 \geq c/2 + 1$ in which case s_i^0 appears at least $k_i^0 - 1 \geq c/2$ times among the digits X_{ji} for $j \in C_1$, thus s_i^0 is an absolute majority here too, and $s_i^1 = s_i^0$. Thus, both y_i and y_i' take value from $\{s_i^1, ?\}$ and by the definition

$$q = P[y_i = s_i^1] = f(q_0) \text{ with } q_0 = \frac{2k_i^0 - c - 1}{c - 3},$$

$$r = P[y_i' = s_i^1] = f(r_0) \text{ with } r_0 = \frac{2k_i^1 - c + 1}{c - 1}.$$

Here $k_i^0 = k_i^1$ or $k_i^0 = k_i^1 + 1$. Now $r = 0$ implies $q = 0$ and similarly $r = 1$ implies $q = 1$, in both cases $H_{\alpha+1}(y_i||y_i') = 0$. If we have $0 < r < 1$ then we also have $q_0 \leq 3r_0$ and $(1 - q_0) \leq 3(1 - r_0)$, thus property (*) of the function f yields $q \leq 10r$, $1 - q \leq 10(1 - r)$ and

$$\frac{(q - r)^2}{r(1 - r)} = O((q_0 - r_0)^2).$$

Straitforward calculations yield the bound $|q_0 - r_0| = O(1/c)$. Using the last two observations and property (d) of the Rényi divergence we get

$$H_{\alpha+1}(y_i||y_i') = O\left(\frac{(q - r)^2}{r(1 - r)}\right) = O\left(\frac{1}{c^2}\right).$$

The hidden constant in the O notation here and elsewhere in this section depends only on α and thus on the exponent a in the theorem.

Next we apply property (a) of the Rényi divergence. Recall that we still consider X to be fixed, and thus all the digits of both y and y' are independent. So we have

$$H_{\alpha+1}(y||y') = O\left(\frac{m}{c^2}\right). \tag{7}$$

Our next goal is to consider (X, σ) to be distributed according to F and prove

$$H_{\alpha+1}\big((\rho_0(X), X, \sigma)||(\rho_1(X), X, \sigma)\big) = O\left(\frac{m}{c^2}\right). \tag{8}$$

Indeed, by property (b) of the Rényi divergence the above divergence is an exponential average of the corresponding divergences with fixed (X, σ). As Equation (7) bounds all those divergences, the bound also holds for their mean and Equation (8) is verified.

Now we apply property (c) of the Rényi divergence for the function

$$g(y, X, \sigma) = \chi_{1\in\sigma(y)} = \begin{cases} 1 \text{ if } 1 \in \sigma(y) \\ 0 \text{ if } 1 \notin \sigma(y) \end{cases},$$

that tells if user 1 is accused. From Equation (8) we get

$$H_{\alpha+1}(\chi_{1\in\sigma(\rho_0(X))}||\chi_{1\in\sigma(\rho_1(X))}) \leq H_{\alpha+1}((\rho_0(X), X, \sigma)||(\rho_1(X), X, \sigma)) = O\left(\frac{m}{c^2}\right).$$

Inequalities (5) and (6) and property (d) of the Rényi divergence show that the left hand side is at least

$$\frac{1}{\alpha}\log\left(\frac{1}{(100c)^{\alpha+1}\epsilon^\alpha}\right) \geq \frac{a-1}{2a+10}\log(1/\epsilon).$$

The last bound can be made true by setting $\alpha = 12/(a-1)$, where $a > 1$ is the exponent in the $\epsilon < 1/(100c^a)$ condition of the theorem.

Putting the last two displayed equations together we get

$$m = \Omega(c^2 \log(1/\epsilon))$$

with the constant in the Ω notation depending only on a, as required. □

6 Concluding remarks

1. Guth and Pfitzmann in [7] introduce a relaxation of the marking condition. They assume the following relaxed version of the marking condition: At any position where the codeword of all pirates agree the pirates still have a δ probability of being able to output a different digit. This happens independently for all the positions of agreement and if they can output a different digit they are not restricted at all in the digit they output. This models the situations where the users cannot detect the positions in a digital document where the

fingerprint is embedded but they are allowed to modify a δ fraction *of the entire document*, thus also modifying some digits of the fingerprint code where such modification is against the marking condition. If the fingerprint is embedded in digital images, audio or video files this relaxation seems to be natural.

Although the pirates are less restricted in their output in this case, they cannot fool our fingerprint codes $F_{nc\epsilon}$ much better. Indeed, the proof of Theorem 1 does not use the marking condition, thus remains valid in this model too. The proof of Theorem 2 however heavily depends on the marking condition. The proof is based on bounding the expectation of a random variable, and the main term in the bound comes from the contribution of the positions where all codewords coincide and the marking condition applies. A closer look however reveals that it is enough that the marking condition applies for a *large fraction* of these positions of agreement and thus exactly the same argument gives a similar bound in the relaxed model of Guth and Pfitzmann. More precisely, the following holds:

Theorem 2'. *Consider the $F_{nc\epsilon}$ code and let $\delta < 1/2$, let $C \subseteq [n]$ be a coalition of size $|C| \leq (1-2\delta)c$, and let ρ be any C-strategy in the relaxed model of Guth and Pfitzmann. We have*

$$P[C \cap \sigma(\rho(X)) = \emptyset] < \epsilon^{c/4},$$

where the probability is according to the distribution on (X, σ) defined by the code $F_{nc\epsilon}$.

Notice that for Theorem 2' to work for coalitions of size c we only have to consider the code $F_{nc'\epsilon}$ for $c' = \lceil c/(1-2\delta) \rceil$, a code that is only a constant factor longer than $F_{nc\epsilon}$ for any fixed $\delta < 1/2$. Also notice that for any binary fingerprint code if $\delta \geq 1/2$ even a single pirate can output a uniform random sequence, thus all fingerprinting is impossible in this case.

2. The situation when only a fraction of the fingerprint code can be retrieved from the illegitimate copy can be handled very similarly. If a random positive fraction of the fingerprint code is retrieved, then the code $F_{nc'\epsilon}$ with some $c' = O(c)$ is ϵ-secure against coalitions of size c. To apply the accusation algorithm, simply treat all unknown digits as zeros.

3. Consider the high-error case of $\epsilon \geq 1/c$. Assume that for an (unreadable digit) fingerprint code somebody is accused from all coalitions of size *at most* c with at least one percent probability and no fixed innocent person is accused with more than ϵ probability. Using coalitions of size substantially smaller than c Theorem 4 implies that the code length is $\Omega(1/\epsilon^b)$ for arbitrary $b < 2$ (the hidden constant depends on b). It is easy to see that if we augment the fingerprint code $F_{nc'\epsilon'}$ with $c' = \lfloor 2/\epsilon \rfloor$ and $\epsilon' = \epsilon/2$ with independently accusing everybody with probability $\epsilon/2$ one gets a fingerprint code satisfying the above requirements with length $m = O(\log(1/\epsilon)/\epsilon^2)$.

4. Let us end the paper with a philosophical remark on fingerprinting and cryptography. It seems that fingerprinting is a cryptographic primitive whose

mathematical analysis does not depend on complexity assumptions, and computational complexity does not seem to play any role here. Notice however, that the complexity assumption exists, it is hidden in the marking condition. The marking condition (or even its relaxation) is based on the assumption that the users cannot detect the positions in a digital document where the fingerprint is hidden unless they see a difference in their copies of the document. In most cases this assumption translates to some kind of a complexity assumption.

Acknowledgments

The author thanks Dezső Miklós for introducing him to the area of fingerprinting and János Pach for a lot of help in writing this paper.

References

[1] G. R. Blakley, C. Meadows and G. B. Purdy, Fingerprinting long forgiving messages, *Proc. of Crypto '85* Springer-Verlag Berlin, Heidelberg, 1985, pp. 180–189.

[2] D. Boneh and M. Franklin, An efficient public key traitor tracing scheme, *Proc. of Crypto '99* Springer-Verlag, Berlin, Heidelberg, 1999, pp. 338–353.

[3] D. Boneh and J. Shaw, Collusion-secure fingerprinting for digital data, *IEEE Transactions of Information Theory* **44** (1998), 480–491.

[4] B. Chor, A. Fiat and M. Naor, Tracing traitors, *Proc. of Crypto'94* LNCS 839, Springer-Verlag Berlin, Heidelberg, 1994, pp. 257–270.

[5] F. Chung, R. Graham, T. Leighton, Guessing secrets, *Electronic Journal of Combinatorics* **8** (1), 2001.

[6] A. Fiat, T. Tassa, Dynamic traitor tracing, *Journal of Cryptology* **14** (3) (2001), 211–223.

[7] J. Guth and B. Pfitzmann, Error- and collusion-secure fingerprinting for digital data, *Information Hiding (IH '99)* LNCS 1768, Springer-Verlag, Berlin 2000, pp. 134-145.

[8] J. Kilian, T. Leighton, L. Matheson, T. Shamoon, R. Tarjan, F. Zane, Resistance of digital watermarks to collusive attacks, in: *Proceedings of 1998 IEEE International Symposium on Information Theory*, p. 71.

[9] K. Kurosawa and Y. Desmedt, Optimum traitor tracing and asymmetric schemes, *Advances in cryptology—EUROCRYPT'98* LNCS 1403, Springer, Berlin, 1998, pp. 145–157.

[10] T. Lindkvist, Fingerprinting digital documents, Linköping Studies in Science and Technology, Thesis No. 798, 1999.

[11] C. Peikert, A. Shelat, A. Smith, Lower bounds for collusion-secure fingerprinting, in: *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) 2003*, pp. 472–479.

[12] A. Rényi, *Probability theory*, North-Holland Series in Applied Mathematics and Mechanics, Vol. 10. North-Holland Publishing Co., Amsterdam, London; American Elsevier Publishing Co., Inc., New York, 1970.

[13] R. Safavi-Naini and Y. Wang, Sequential traitor tracing, in: *Proc. of Crypto'2000*, LNCS 1880, Springer-Verlag Berlin, Heidelberg, 2000, pp. 316–332.

[14] J. N. Staddon, D. R. Stinson, R. Wei, Combinatorial properties of frameproof and traceability codes, *IEEE Transactions of Information Theory* **47** (2001), 1042–1049.

[15] G. Tardos, Optimal probabilistic fingerprint codes, in: *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, 2003, pp. 116–125.

[16] T. Tassa, Low bandwidth dynamic traitor tracing schemes, *Journal of Cryptology*, to appear.

[17] Y. Yacobi, Improved Boneh-Shaw content fingerprinting, in: *Topics in cryptology—CT-RSA 2001*, LNCS 2020, Springer-Verlag Berlin, 2001, 378–391.

[18] N. Wagner, Fingerprinting, *Proc. of the 1983 IEEE Symposium on Security and Privacy* (1983), pp. 18–22.