

High rate fingerprinting codes and the fingerprinting capacity

Ehsan Amiri*

Gábor Tardos†

Abstract

Including a unique code in each copy of a distributed document is an effective way of fighting intellectual piracy. Codes designed for this purpose that are secure against *collusion attacks* are called fingerprinting codes.

In this paper we consider fingerprinting with the marking assumption and design codes that achieve much higher rates than previous constructions. We conjecture that these codes attain the maximum possible rate (the fingerprinting capacity) for any fixed number of pirates. We prove new upper bounds for the fingerprinting capacity that are not far from the rate of our codes. On the downside the accusation algorithm of our codes are much slower than those of earlier codes.

We introduce the novel model of *weak fingerprinting codes* where one pirate should be caught only if the identity of all other pirates are revealed. We construct fingerprinting codes in this model with improved rates but our upper bound on the rate still applies. In fact, these improved codes achieve the fingerprinting capacity of the weak model by a recent upper bound.

Using analytic techniques we compare the rates of our codes in the standard model and the rates of the optimal codes in the weak model. To our surprise these rates asymptotically agree, that is, their ratio tends to 1 as t goes to infinity. Although we cannot prove that each one of our codes in the standard model achieves the fingerprinting capacity, this proves that *asymptotically* they do.

1 Introduction

To ensure protection of their copyright, content producers often make each copy of their productions unique by embedding a distinct code in each. For this to work they have to be able to hide the positions where the code is embedded. A *collision attack* is performed by a group of malicious users (the *pirates*), who compare their copies and identify the positions where they differ as a position of the embedded code. They can then arbitrarily change the code in these positions. We assume however that they do not notice the positions of the hidden code where all their codes agree and therefore they cannot alter these positions. This is the *marking assumption*.

A *fingerprinting code* is a carefully selected collec-

tion of codewords (in fact a randomized procedure to obtain them) that makes the distributor able to catch at least one of the pirates who perform such an collusion attack. The mathematical definition (see below) was first given by Boneh and Shaw [3].

1.1 The model: We start with introducing simple notation. We fix a finite alphabet Σ . In this paper we consider the *binary alphabet* $\Sigma = \{0, 1\}$ but fingerprinting is studied over larger alphabets too (although there seems to be little advantage in using them unless one replaces the marking assumption used here with a stronger assumption). For a positive integer n we denote by $[n]$ the set $\{1, \dots, n\}$ of positive integers not exceeding n . For a sequence X of length n and $i \in [n]$ we denote the i th entry in X by X_i , i.e., $X = (X_1, \dots, X_n)$.

An (*arbitrary digit*) *pirate-strategy* for a set T of *pirates* and for codes of length n over an alphabet Σ is any (deterministic or randomized) algorithm that takes as input the codewords $X^v \in \Sigma^n$ of the pirates $v \in T$ and outputs a *forged codeword* $F \in \Sigma^n$ that satisfies the so called *marking assumption*, i.e., if for some $j \in [n]$ the digits X_j^v agree for all $v \in T$, then we also have $F_j = X_j^v$.

A *fingerprinting code* of length n over the alphabet Σ for the users in the set U consist of an *accusation algorithm* and a randomized procedure that generate codewords $X^v \in \Sigma^n$ for users $v \in U$ (and possibly some additional input for the accusation algorithm). The accusation algorithm takes the output of this procedure and a forged codeword $F \in \Sigma^n$ and outputs a set $A(F) \subseteq U$ of *accused users*. If F was obtained by a set $T \subseteq U$ of users performing a pirate-strategy and $v \in A(F)$ for some user $v \notin T$, then we say that v is *falsely accused*, while if T and $A(F)$ are disjoint we say the pirates *are not caught*. (One may assume A always accuses a single user and then the two types of error coincide, but it is sometimes better to allow $A(F) = \emptyset$ because false accusations are considered worse than declaring failure.)

We call a fingerprinting code ϵ -*secure against t pirates* if for any set $T \subseteq U$ of users of size $|T| \leq t$ using any pirate-strategy the probability that either the pirates are not caught or some user is falsely accused is at most ϵ . The bound is on the overall probability of

*School of Computing Science, Simon Fraser University. E-mail: eamiri@fas.sfu.ca

†School of Computing Science, Simon Fraser University and Rényi Institute, Budapest. Supported by NSERC grant 329527 and by OTKA grants T-046234, AT048826 and NK-62321. E-mail: tardos@cs.sfu.ca

either type of error.

The *rate* of the fingerprinting code is $R = \frac{\log_{|\Sigma|} |U|}{n}$.

A *fingerprinting scheme* of rate R over the alphabet Σ is an infinite sequence of fingerprinting codes C_i over Σ for the set U_i of users such that $|U_i|$ goes to infinity and the rate R_i of C_i tends to R . We say that a fingerprinting scheme is *t-secure* if C_i is ϵ_i -secure against t pirates with ϵ_i tending to 0. The *t-fingerprinting capacity* is the maximum achievable rate of t -secure fingerprinting schemes.

The goal of fingerprinting research is to find efficient and secure fingerprinting codes. The paramount problem in the application of fingerprinting codes is the high cost of embedding every single digit of the code. This makes it important to design secure fingerprinting codes that are short, or equivalently, have high rate. In particular, recent research focused on finding or estimating the t -fingerprinting capacity for various values of t .

1.2 Previous work: Boneh and Shaw [3] were first to define fingerprinting secure against collusion attacks. They proposed such codes of length $O(t^4 \log(N/\epsilon) \log(1/\epsilon))$ for N users that are ϵ -secure against t pirates. This translates to a t -secure fingerprinting scheme with rate of $\Theta(1/t^4)$. The same paper gave a lower bound of $\Omega(t \log(1/(t\epsilon)))$ for the length of the fingerprinting code with the same parameters. As this bound does not depend on the number of users it does not give an upper bound on the rates of t -secure fingerprinting schemes.

Later the second author [9] constructed fingerprinting codes of length at most $100t^2 \log(N/\epsilon)$ for N users that are ϵ -secure against t pirates. This construction yields a t -secure fingerprinting scheme of rate $1/(100t^2)$. The same paper gave an $\Omega(t^2 \log(1/\epsilon))$ bound on the length of any fingerprinting code with the above parameters. Although this is within a constant factor of the length achieved by the construction for small error rates $\epsilon < 1/N$, as this lower bound (just like that of Boneh and Shaw) does not depend on the number of users it does not directly translate to any upper bound on the rates of t -secure fingerprinting schemes.

The constant 100 in the length $100t^2 \log(N/\epsilon)$ was subsequently improved by several papers. In particular all the papers [8, 7, 2, 6, 5] go through the code and the proof in [9] and optimize various parameters to improve various aspects of the code without fundamentally changing the basic structure of the code construction and the accusation algorithm. We are concerned here with improvements of the code length that translates directly to improved rates. While some of these papers rely on reasonable, but not fully justified assumptions and others are estimating the improvement fac-

tor based on experimental evidence, the most significant and mathematical rigorous improvement is achieved by Nuida et al. [5]. For large t their rates are slightly below 19 times larger than those in [9].

Anthapadmanabhan, Barg and Dumer [1] choose a different, *information theoretic* approach. They construct t -secure fingerprinting schemes whose rates for $t = 2$ and 3 are much higher than previously obtained rates but the rate of their schemes deteriorates exponentially with t . They are the first to prove upper bounds on the rates of t -secure fingerprinting schemes. The upper bounds in their paper is given in terms of a hard to evaluate information theoretic minimax formula. They estimate this formula and prove strong upper bounds on the t -fingerprinting capacity for small values of t (namely 2 and 3) and an $O(1/t)$ asymptotic bound.

1.3 Our results: In this paper we combine the two approaches represented by [9] and [1] to obtain fingerprinting codes of rates higher than those achieved with either method separately.

Our method for *constructing the codewords* is very similar to the construction in [9]: There one uses a continuous distribution to select certain bias probabilities P_j used in the construction of the codeword, here we replace this continuous distribution by a discrete distribution fine-tuned for the given number t of pirates. (Note that Nuida et al. [6] also propose a discrete distribution but a different one and there the primary goal is also different: to reduce memory usage.) The optimal distribution is found through a game-theoretic equilibrium.

As opposed to the codeword construction, our *accusation algorithm* is fundamentally different from that of [9] and it is inspired by the paper [1]. Unlike the simple and efficient accusation algorithm in [9] that uses only a user's own codeword to decide whether to accuse him, this algorithm (like the codes in [1]) consider all t -tuple of users in order to find who should be accused. It seems that this considerable slowdown is the price one has to pay for the better rates.

For $t = 2$ our codes coincide with the codes given in [1] but for $t \geq 3$ our rates are better than the rates of previous codes. We numerically estimate the achieved rates for $t \leq 7$ and prove the $R_t \geq t^{-2}/(2 \ln 2) + o(t^{-2})$ asymptotic bound on the rate of our t -secure codes. This improves the rates of the codes in [9] by a factor over 72 for large values of t and thus overshadows previous improvements that were mostly numerical and kept the basic structure of that code. More importantly than the actual numeric factor is that we believe these rates are optimal: they achieve the fingerprinting capacity.

We give a very simple deduction of an $O(t^{-2})$ upper bound on the t -fingerprinting capacity from results in [9]. This bound attains the correct order of magnitude and applies for arbitrary alphabet sizes and even to the *unreadable digit* model of fingerprinting (see, e.g., [9]). We also derive a similar upper bound with a better constant factor from analytic evaluation of a bound proved in [1] improving the $O(t^{-1})$ bound of that paper.

In section 5 we introduce a new model we call *weak fingerprinting*. In this model the accusation algorithm receives substantial help: when it tries to find one of the pirates the identity of all the other pirates are revealed. We construct fingerprinting codes of improved rates in this model. As established recently these improved codes attain optimal rates. This result will be published separately, but here we explore its consequences.

The weak fingerprinting model may be considered unreasonably lenient toward the accuser/distributor, so it may come as a surprise that the rate of our fingerprinting codes in the standard and the weak models asymptotically agree. As the codes for the weak model are optimal, this implies that the fingerprinting capacities of the two models also agree asymptotically and our codes in the standard model are close to optimal: for any $\epsilon > 0$ and large enough t their rates are within a factor of $1 - \epsilon$ of the fingerprinting capacity.

1.4 Proofs: For lack of space we omit most proofs in this version of the paper. We give outlines of those proofs we feel most important. Please refer to the upcoming full journal version for the detailed proofs.

2 The construction

Before presenting our fingerprinting codes we motivate the construction with introducing a few relevant concepts. We start with simple probabilistic pirate-strategies we call *channel-based* and a simple method for code generation we call *bias-based*.

Let us stress that although these restricted classes are useful they are not restricting our results in any way. Both types of fingerprinting codes we construct use bias-based code generation but our upper bounds apply to arbitrary fingerprinting codes, even to those using different code generation algorithms.

Similarly, channel-based pirate-strategies are useful sources of intuition but we prove that our codes are secure against any pirate-strategy, even against non channel-based strategies.

A *t-channel* or *channel* for short is a typically randomized procedure that produces an output bit f from an input $x \in \{0, 1\}^t$. It is determined by the function $S : \{0, 1\}^t \rightarrow [0, 1]$ defined as $S(z) = \Pr[f = 1 | x = z]$. We use the function S and the channel it

determines interchangeably. We say that the pirates u_1, \dots, u_t use the channel S if they apply a randomized pirate-strategy that produces each bit of the forged codeword F independently, obtaining F_j through the channel S from the input $x = (x_1, \dots, x_t)$, where x_i is the i th bit of the codeword of the pirate u_i .

We call the channel S *eligible* if $S(0 \dots 0) = 0$ and $S(1 \dots 1) = 1$. Clearly, if a t -tuple of pirates use an eligible channel, then they obey the marking assumption. We call a pirate-strategy *channel-based* if the pirates use an eligible channel.

Note that in a general pirate-strategy each individual bit of the forged codeword may depend on the entire codewords of the pirates. The channel-based pirate-strategies are those where bit j of the forged codeword only depends on the j th bit of the codewords, and furthermore the pirates use the same channel to obtain each bit of the forged codeword independently. Such strategies are often referred to as “multiple user memoryless channels”. channel-based pirate-strategies with a symmetric channel (see definition in Section 4.1) are called *bias-strategies* in [9].

A *bias-based code generation* is a two phase process determined by the probability distribution D on the interval $[0, 1]$ (the *bias distribution*) and the codeword length n . Given D and n first we select the *bias vector* $P = (P_1, \dots, P_n)$ by selecting the individual biases $P_j \in [0, 1]$ independently for $j \in [n]$, each according to the distribution D . In the second phase we select the bits of the codewords X^v for all users v . We select each bit of each such codeword independently. For the bit X_j^v we use the bias P_j , i.e., we have $\Pr[X_j^v = 1] = P_j$.

The fingerprinting codes constructed in [9] use bias-based code generation with a continuous bias distribution. Here we use a discrete bias distribution fine tuned for channel-based pirate-strategies. We will then establish that they are also secure against an arbitrary pirate-strategy.

Throughout this section we consider the positive integer parameter t fixed. It will eventually represent the bound on the number of pirates. For $p \in [0, 1]$ and a channel S we define the distribution $B_{p,S}$ on the binary vector $x \in \{0, 1\}^t$ and the binary variable $f \in \{0, 1\}$ by choosing the individual digits x_i of x independently for $i \in [t]$ with an identical distribution of expectation p and finally obtaining f from x through the channel S . We define $I_{p,S}$ to be the mutual information $I(x; f)$ in this distribution. For the definition of mutual information and other information theoretic terms (entropy, conditional entropy, conditional mutual information and relative entropy) we refer the reader to Chapter 1.1 of the textbook [4]. Note that we use binary logarithm in defining all these quantities.

To motivate the above definition and the game theoretic approach below assume that in a fingerprinting code the code generation is bias-based and we have performed the first phase of this code generation (so the bias P_j is fixed for all j) but the second phase is yet to be performed. Also assume that t pirates u_1, \dots, u_t will use channel S to obtain the forged codeword F . Fix $j \in [n]$ and consider the vector $x \in \{0, 1\}^t$ given by $x_i = X_j^{u_i}$ and the binary value $f = F_j$. The pair (x, f) is distributed here according to $B_{P_j, S}$ so it is natural to say that $I_{P_j, S} = I(x; f)$ is the amount of information the bit F_j of the forged codeword F reveals of the identity of the pirates. Although this statement is not precise it gives a good intuition why the pirates want to minimize this quantity and why the distributor tries to maximize it and motivates the following two person game.

Paula (representing the pirates) chooses an eligible channel S and David (the distributor) chooses a probability $p \in [0, 1]$. The players choose simultaneously without seeing the other player's choice. After making their choice they learn about each other's choices and Paula pays David $I_{p, S}$.

The proof (omitted here) is based on a continuous version of the minimax theorem of game theory and the observation that Paula is always better off using a *pure strategy* (i.e., a deterministic choice of an eligible channel) than a *mixed strategy* (i.e., a randomized choice of the channel). The equilibrium of the game, whose existence is stated in the lemma gives us the optimal bias distribution for the code generation part of our fingerprinting codes and it also determines the rate achievable by these codes.

LEMMA 2.1. *We have*

$$\min_S \max_p I_{p, S} = \max_D \min_S E_{p \in D} [I_{p, S}],$$

where $p \in [0, 1]$, S is an eligible channel, D is a distribution over $[0, 1]$ and $E_{p \in D}[\cdot]$ represents the expectation as p is distributed according to D .

We use S^t to denote a channel S minimizing $\max_p I_{p, S}$ and D^t to denote a distribution D on $[0, 1]$ maximizing $\min_S E_{p \in D} [I_{p, S}]$. We let V_t stand for the common value of this minimum and maximum and set $R_t = V_t/t$. For technical reasons we need that D^t is concentrated to a finite number of values. See Section 7 estimates on the size of the support.

The novelty in the fingerprinting code is in the involved accusation algorithm. In effect we look for a set of users such that the forged codeword "looks like" it was created by that set of users using a channel-based pirate-strategy. We then accuse a carefully selected subset of these users. Pirates may use complicated,

non channel-based strategies, but, as we will show in Lemma 3.1 below, no matter what they do, with high probability the distributor will find such a channel-based strategy producing a similar forged codeword.

Given a t -tuple $u = (u_1, \dots, u_t)$ of distinct users, their codewords X^{u_i} , the bias vector $P \in [0, 1]^n$ and a forged codeword $F \in \{0, 1\}^n$ we define a distribution B_u on (x, f, p) where $x \in \{0, 1\}^t$, $f \in \{0, 1\}$ and $p \in [0, 1]$. We obtain B_u by choosing $j \in [n]$ uniformly at random and setting $p = P_j$, $f = F_j$ and $x_i = X_j^{u_i}$ for $i \in [t]$.

B_u is, in fact, the *joint type* of the codewords X^{u_i} for $i \in [t]$, F and the bias vector P . Here the *type* of a sequence is simply the distribution of a uniform random element of the sequence. If A^1, \dots, A^i are sequences of the same length n , then their *joint type* is the distribution of $(A_j^1, A_j^2, \dots, A_j^i)$ for a uniform random $j \in [n]$. See more on types in [4, Chapter 1.2]. Here we only mention that types employ randomness "on a different level" than the construction of the fingerprinting codes. When speaking about the types of codewords or about B_u we treat codewords as fixed sequences. But in reality codewords themselves are constructed in a randomized fashion. This makes their types and also the joint type B_u (themselves distributions) random variables.

For a distribution D over $[0, 1]$ and an eligible channel S we define $B_{D, S}$ to be the distribution over (x, f, p) , where $p \in [0, 1]$ is chosen according to D and then (x, f) is chosen according to $B_{p, S}$.

Note that if the codeword generation is bias-based with bias distribution D and the pirates u_1, \dots, u_t use a channel-based pirate-strategy with the channel S to obtain the forged codeword F , then the distribution of the vector $(X_j^{u_1}, \dots, X_j^{u_t}, F_j, P_j)$ is exactly $B_{D, S}$ for every fixed $j \in [n]$ and these vectors are independent.

We are now ready to define our fingerprint codes denoted by $E_{t, R, \delta, n}$. Here t is the number pirates it is designed to be secure against, $R > 0$ is the rate, n is the codelength and $\delta > 0$ is a security parameter.

Codeword generation: We use bias-based codeword generation with the bias distribution D^t . We let the set U of users be of cardinality $|U| = N = \lfloor 2^{Rn} \rfloor$.

Accusation algorithm: To accuse users based on the forged codeword F we select a t -tuple u of distinct users from U and an eligible channel S satisfying that the total variation distance between the distributions B_u and $B_{D^t, S}$ is at most δ . If there is no suitable choice for S and u we declare failure. After fixing S and u we select a minimal nonempty set $Q \subseteq [t]$ with the property that in the distribution $B_{D^t, S}$ we have $I(x_Q; f|p) \geq |Q|R_t$, where x_Q is the collection of the variables x_i with $i \in Q$. Finally we accuse the users u_i for the indices $i \in Q$. Note that if suitable u and S are found then we will be

able to find a suitable set Q too, as for the set $Q = [t]$ we have $x_Q = x$, $|Q| = t$ and $I(x; f|p) = E_{p \in D^t} [I_{p,S}] \geq tR_t$ by Lemma 2.1.

THEOREM 2.1. *For any t and arbitrary $R < R_t$ there exists a positive δ such that the fingerprinting codes $E_{t,R,\delta,n}$ for $n \geq 1$ form a t -secure fingerprinting scheme of rate R .*

For the proof we have to show that if δ is small enough for t and R , then the error rate of the codes $E_{t,R,\delta,n}$ tend to zero. We state the error bound separately for the two types of error (failure and falsely accusing a user) in the next section.

It is useful (although not required in the definition) that the error bound on falsely accusing somebody applies even if the size of the pirate coalition exceeds t . This ensures that if the distributor's assumption on the maximal number of malicious users is violated, then only the probability of failure (not accusing anybody) goes up but innocent users are still not likely to get framed. The codes described in [9] have this favourable property regardless of the coalition size. Our codes here have almost the same property. For technical reasons we require some bound on the pirate coalition, but anything *subexponential* (i.e., $2^{o(n)}$) suffices here instead of the constant bound t .

COROLLARY 2.1. *The t -fingerprinting capacity is at least R_t .*

CONJECTURE 2.1. *The t -fingerprinting capacity is exactly R_t .*

For numerical and asymptotic estimates on R_t see Section 4. For upper bounds on t -fingerprinting capacity see Section 6.

3 Error bound for the codes

We will show that the probability of either form of error is exponentially small. We call a probability *exponentially small* if it can be bounded by 2^{-cn} for some $c > 0$. We consider the parameters t , δ and R fixed, c may depend on them, but not on the running parameter n .

Given a t -tuple u of users and the joint type B_u the *perceived strategy* of u is the channel S defined by $S(b) = E[f|x = b]$ for all $b \in \{0,1\}^t$ with $\Pr[x = b] > 0$, where the expectation and the probability are in the distribution B_u . For values of b with $\Pr[x = b] = 0$ we arbitrarily define $S(b) = 0$ with the exception of $S(1 \dots 1) = 1$ (to make S eligible if possible). With notation from [4] this perceived strategy is the *conditional type* of the forged codeword given the codewords of the users in u .

Our first lemma bounds the probability of the pirates are not caught.

LEMMA 3.1. *Consider the fingerprinting code $E_{t,R,\delta,n}$ and a set of at most t pirates performing any pirate-strategy. Let u be a t -tuple of distinct users including all the pirates. Then the perceived strategy S of u is an eligible channel and the probability that the total variation distance between B_u and $B_{D^t,S}$ exceeds δ is exponentially small.*

Proof. In the analysis of our fingerprinting codes (i.e., here and in the proof of Lemma 3.2) we always assume that the pirates use a deterministic strategy to compute the forged codeword from their codewords. Security against such pirates implies the same degree of security against pirates using a randomized strategy. See for example [9] for the simple argument. Note that our accusation algorithm is based on finding a channel-based pirate-strategy outputting a similar forged codeword and this hypothetical channel-based strategy is typically probabilistic.

Note that S is indeed eligible by the marking assumption.

Consider the joint distribution of x and p in the two distributions (i.e., ignore f for now). They take a finite number of possible values. The bias P_j is chosen according to D^t independently for every $j \in [n]$ and then $X_j^{u_i}$ is chosen independently and with expected value P_j . As (x, p) is distributed the same way in $B_{D^t,S}$ (for any S) the distribution of (x, p) in B_u (the joint type of the codewords X^{u_i} for $i \in [t]$ and P) is the distribution obtained by n independent samples of the distribution of (x, p) in $B_{D^t,S}$. By the Chernoff bound the probability that the relative frequency of any of the possible values differs from its probability in $B_{D^t,S}$ by any constant amount is exponentially small. Thus the probability that the distributions of (x, p) in B_u and $B_{D^t,S}$ differ by at least $\delta/2$ in total variation is also exponentially small.

We need to show that this distance is not significantly increased by considering f too. The main point to note here is that the pirates produce the forged codeword without having access to the bias probabilities beyond observing their codewords that depend on them. Thus any correlation between f and p given x is B_u happens by coincidence and not by the design of the pirates. Now if we let f truly independent from p by defining the distribution B^* on (x, f, p) where (x, p) is distributed as in B_u and f is obtained from x through the channel S , then, clearly, the total variation distance between B^* and $B_{D^t,S}$ is the same as the total variation distance between the two distributions on (x, p) bounded above. We finish the proof by bounding the

total variation distance between B^* and B_u and the statement of the lemma follows by the triangle inequality.

For this second part of the proof we reverse the order of picking the bias vector and picking the the codewords for the pirates. Formally we consider the following four phase process:

1. First we pick a preliminary bias vector P' by choosing P'_j independently for each $j \in [n]$, each value according to the distribution D^t .
2. Next we pick the binary codewords X^{u_i} for $i \in [t]$ by selecting all digits $X_j^{u_i} \in \{0, 1\}$ independently for $i \in [t]$ and $j \in [n]$ such that $E[X_j^{u_i}] = P'_j$.
3. Next we pick the real bias vector P selecting uniformly at random among all such vectors satisfying that the joint type of the codewords X^{u_i} and P is the same as the joint type of the codewords X^{u_i} and P' . Using the notation of [4] the requirement is that the conditional types of P and P' given the codewords X^{u_i} should agree.
4. Finally we pick the codewords not selected in phase 2 by independently picking all the digits $X_j^v \in \{0, 1\}$ for $v \in U \setminus \{u_1, \dots, u_t\}$, $j \in [n]$ such that $E[X_j^v] = P_j$.

Notice that if in the crucial third phase we picked $P = P'$ the construction would be identical to that defining our fingerprinting code. We allow however some randomness for the choice of the bias vector. Notice that all bias vectors we allow to be chosen at this point are obtained in the original construction with exactly the same chance. Each yields the codewords X^{u_i} in the original construction with the same probability. Thus choosing uniformly randomly among them does not break this symmetry and our twisted construction above yields the same distribution on the codewords and the bias vector as the original construction. Therefore this twisted construction is a legitimate alternative method for generating the same fingerprinting code.

Notice that after phase 2 the codewords of the pirates are fixed and thus the forged codeword F they produce is also determined and so is the channel S . (Recall that we consider deterministic pirate-strategies.) The distribution B^* seem to depend on P but, in fact, it is also determined after phase 2 as our restriction on the choice of P in phase 3 makes sure that B^* is the same as it were with the $P = P'$ choice. In contrast the joint type B_u really depends on the bias vector P to be determined only in phase 3.

We claim that however the first two phases of the construction go, the (conditional) probability that B_u

(as determined in the third phase) will not be within $\delta/2$ of the fixed distribution B^* is exponentially small. This is clearly enough to finish the proof of the lemma.

As B^* and B_u are concentrated on a finite number of possible values it is enough to bound the probability of a choice in phase 3 that makes the probability of a fixed value differ by more than some constant. Because of the random choice in phase 3 this event can be combinatorially expressed as the size of the intersection $H \cap H'$ significantly deviating from its expectation, where H is a fixed set and H' is a uniform random subset of given size of a certain base set. The proof is finished by a standard application of the Chernoff bound. See details in the full version. ■

LEMMA 3.2. *For any $t, R < R_t$ and small enough $\delta > 0$ the following holds for the fingerprinting codes $E_{t,R,\delta,n}$. For an arbitrary subexponential size set of users (the pirates) performing an arbitrary pirate-strategy the probability that anybody gets falsely accused is exponentially small.*

The game theoretic definition of D^t and R^t yields relatively easily that that B_u is not likely to be close to any $B_{D^t,S}$ for a t -tuple u of innocent users. Avoiding this similarity is necessary to avoid falsely accusing one of the members of u .

Unfortunately, similar statement can not be said about mixed t -tuples consisting of some innocent users and some pirates. The involved selection of the subset of u to be accused is designed to separate some pirates from the non-pirates in such mixed company. Proving that this selection is done right and the code does not falsely accuse innocent users is technically more challenging and for lack of space we only sketch the proof here. Notice that the proof does not use the marking assumption, so applies even if it is violated.

Proof. We fix the bias vector P and the codewords of the pirates arbitrarily. This fixes the forged codeword F too, but the codewords of the innocent users are still obtained via the process defining our fingerprinting code. We claim that the statement of Lemma 3.2 holds regardless of these initial choices. This implies the same bound overall.

If somebody is accused, then they are the users u_i for $i \in Q$, where a t -tuple u of users, an eligible channel S and $Q \subseteq [t]$ is chosen by the accusation algorithm. Let Q_0 be the subset of Q consisting of the indices i with u_i a pirate and let $Q_1 = Q \setminus Q_0$. The falsely accused users are u_i with $i \in Q_1$. We have to bound the probability of $Q_1 \neq \emptyset$.

Recall that for an index set $H \subseteq [t]$ we write x_H for the sequence of values x_i with $i \in H$. If $Q_1 \neq \emptyset$ then we

have $I(x_{Q_0}; f|p) \leq |Q_0|R_t$ in $B_{D^t, S}$ by the minimality of Q . We also have $|Q|R_t \leq I(x_Q; f|p) = I(x_{Q_0}; f|p) + I(x_{Q_1}; f|x_{Q_0}, p)$. We conclude that $I(x_{Q_1}; f|x_{Q_0}, p) \geq |Q_1|R_t$.

All this computation regards the probability space $B_{D^t, S}$ but we are more interested in the joint type B_u . All functions considered take on a bounded number of values (recall that we consider t fixed). As the mutual information depends continuously on the probabilities with which the combinations of these values are obtained we conclude that if the total variation distance between B_u and $B_{D^t, S}$ is small enough almost the same bound holds in B_u . More precisely the inequality

$$(3.1) \quad I(x_{Q_1}; f|x_{Q_0}, p) \geq |Q_1|R_t + \delta'$$

holds in B_u for any $\delta' > 0$ as long as δ is sufficiently small as a function of t and δ' . In what follows we bound the probability of (3.1) holding for some u , $Q_1 \neq \emptyset$ and Q_0 .

Notice that the mutual information above is determined by Q_0 , Q_1 and the joint distribution of x_Q , f and p in B_u , i.e., by the joint type of the codewords X^{u_i} for $i \in Q$, F and P . In other words we can safely ignore the codewords X^{u_i} with $i \notin Q$ that also appear in the joint type B_u .

Let us fix now Q and u (and thus Q_0 and Q_1). We bound the probability that a given type B is obtained as the joint type of the codewords X^{u_i} for $i \in Q$, F and P . Recall that we consider now the bias vector P and the codewords of the pirates (i.e., X^{u_i} for $i \in Q_0$) and also the forged codeword F fixed. If obtaining B as the joint type with these fixed values is possible at all, then B will be the joint type for $2^{nH(x_{Q_1}|x_{Q_0}, f, p) + O(\log n)}$ different choices for the codewords X^{u_i} with $i \in Q_1$. This computation is considered standard in information theory and we are counting here the so called *V-shell*. See [4, Chapter 1.2]. Here (and later in this paragraph) H refers to entropy in B and the hidden constant in the O notation depends only on t . A similar argument gives that the number of choices for these codewords that lead to the joint type of X^{u_i} for $i \in Q_1$ and P being the same as the distribution of (x_{Q_1}, p) in B (if at all possible) is $2^{nH(x_{Q_1}|p) + O(\log n)}$. As the probability of a particular choice for these codewords is determined by the joint type of the codewords and the bias vector P all these choices are equiprobable. This yields

$$(3.2) \quad P[Z_B] \leq P[Z_B|Z'_B] = 2^{-nI(x_{Q_1}; x_{Q_0}, f|p) + O(\log n)}$$

where Z_B denotes the event that B is the joint type of X^{u_i} for $i \in Q$, F and P , while Z'_B denotes larger event that the joint type of X^{u_i} for $i \in Q_1$ and P is the distribution of (x_{Q_1}, p) in B , and I denotes the mutual information in B .

The probability that (3.1) holds for some fixed Q and z is the sum of the probabilities of the events Z_B for the joint types B satisfying (3.1). For such a distribution B we have

$$\begin{aligned} P[Z_B] &\leq 2^{-nI(x_{Q_1}; x_{Q_0}, f|p) + O(\log n)} \\ &\leq 2^{-nI(x_{Q_1}; f|x_{Q_0}, p) + O(\log n)} \\ &\leq 2^{-n|Q_1|R_t + \delta'n + O(\log n)}. \end{aligned}$$

Here the first inequality is (3.2), the second follows from a general property of the entropy and the last is an application of (3.1). The number of possible joint types B is polynomial in n , therefore $2^{-n|Q_1|R_t + \delta'n + O(\log n)}$ also bounds the total probability of (3.1) holding in B_u for some fixed u and Q .

Finally we add up the above estimate for all choices of u and Q with $Q_1 \neq \emptyset$ and find that the probability of accusing an innocent user is at most $2^{(R - R_t + \delta')n + o(n)}$. This is exponentially small as claimed if $\delta' < R_t - R$. ■

4 Evaluation of the rates

4.1 Few pirates: We make the computational search for the rates R_t simpler by restricting attention to (strongly) symmetric channels and distributions as defined below. This reduces the number of parameters in the minimization for S^t from $2^t - 2$ to $\lceil t/2 \rceil - 1$.

Let us fix t as in the previous section. For a binary vector $x \in \{0, 1\}^t$ we write $|x| = \sum_{i=1}^t x_i$. We call the channel S *symmetric* if $S(b) = S(c)$ whenever $b, c \in \{0, 1\}^t$ satisfy $|b| = |c|$. S is *strongly symmetric* if we further have $S(b) + S(c) = 1$ whenever $|b| + |c| = t$. We call a discrete distribution D on $[0, 1]$ *symmetric* if for every $x \in [0, 1]$ the probabilities of x and $1 - x$ are equal.

Recall that S^t is an eligible channel minimizing the left hand side of the equation in Lemma 2.1, while D^t is a (discrete) distribution on $[0, 1]$ maximizing the right hand side.

LEMMA 4.1. *We can choose S^t to be a strongly symmetric channel and D^t to be a symmetric distribution.*

For the simple proof see the full version of this paper.

Using the lemma we can computationally find the rate R_t and the corresponding channel S^t and distribution D^t . See our results in Table 1.

For $t = 2$ we don't even have to search for S^t as there is only a single strongly symmetric eligible channel. The corresponding distribution D^2 is concentrated on $1/2$ yielding uniform random codewords and a rate of $R_2 = 1/4$. The same simple code has also been proposed in [1].

For $t \geq 3$ the codes we construct are new and achieve better rates than previously suggested codes.

t	Support of D^t	S^t	R_t
2	0.5	0, 0.5, 1	0.25
3	0.26, 0.74	0, 0.38, 0.62, 1	0.0975
4	0.227, 0.773	0, 0.27, 0.5, 0.73, 1	0.0545
5	0.18, 0.82	0, 0.23, 0.36, 0.64, 0.77, 1	0.0338
6	0.15, 0.85	0, 0.19, 0.31, 0.5, 0.69, 0.81, 1	0.0232
7	0.13, 0.50, 0.87	0, 0.16, 0.26, 0.44, 0.56, 0.74, 0.84, 1	0.0168

Table 1: Optimal channels and distributions and the rate for t up to 7. The third column lists values of $S^t(x)$ in increasing order of $|x|$. All numbers for $t \geq 3$ are numeric approximations.

4.2 Many pirates (asymptotics): The next theorem establishes exact asymptotics for the rates achieved by our codes. For large t the improvement over the codes in [9] is over 72-fold. More importantly, they are asymptotically optimal as we will see in Section 6.

THEOREM 4.1.

$$R_t \geq \frac{1}{(2 \ln 2)t^2} - o(t^{-2}) = \frac{0.721 \dots + o(1)}{t^2}$$

See the proof in the full version of the paper. For lack of space we only give a very rough sketch here.

Our goal is to give a lower bound for $V_t = tR_t$ using $V_t = \max_p I_{p,S^t}$. Here I_{p,S^t} is a mutual information and as such can be expressed as the expectation of a relative entropy:

$$I_{p,S^t} = E_p[h(S^t(x)||g(p))],$$

where $h(a||b)$ denotes the relative entropy of the Bernoulli random variables with expectations a and b , $E_p[\cdot]$ denotes expectation in the distribution B_{p,S^t} and $g(p) = E_p[f]$.

We use a strong form of Pinsker inequality (that we could not find in the literature) to relate this relative entropy to $(S^t(x) - g(p))^2$.

We use Cauchy-Schwarz to bound $E_p[(S^t(x) - g(p))(|x| - pt)]$ and use the surprising identity

$$E_p[(S^t(x) - g(p))(|x| - pt)] = (p - p^2)g'(p)$$

to bound the derivative $g'(p)$ of $g(p)$. This yields a differential equation-like inequality bounding $g'(p)$. The boundary conditions $g(0) = 0$ and $g(1) = 1$ follow from marking assumption, which, together with the differential equation give the lower bound part of the theorem.

5 Weak fingerprinting

In this section we introduce a model for fingerprinting in which the accusation algorithm gets a huge help: when searching for one of the pirates it learns from an ‘‘oracle’’ the identity of all the other pirates. See

the formal definition below. It is not surprising that we can adapt our fingerprint codes to this setup and increase their rate somewhat. We introduce this model mainly because the information theoretic upper bounds on the capacity work in this model too and show that the capacities of the weak and the standard fingerprinting models agree asymptotically as t goes to infinity. As the upper bound on the rate (to be published separately) matches the capacity of the new codes we have exact one letter characterisation of the weak t -fingerprinting capacity for every t , see Theorem 6.3.

Definition: A weak fingerprinting code consists of a *distribution algorithm*, an *oracle* and an *accusation algorithm*. The distribution algorithm works the same as in the standard model and so is the type of pirate-strategies we allow. After a set of pirates output a forged codeword the oracle chooses one of the pirates and reveals the identity of all *other* pirates. The oracle has full access to the outputs of the distribution algorithm, the forged codeword and also to the set of pirates. Finally the accusation algorithm has access to the output of the distribution algorithm and the oracle and the forged codeword and tries to guess the pirate kept hidden by the oracle.

Note that the oracle is part of the fingerprinting code, so it collaborates with the accusation algorithm: it will choose the pirate that is most vulnerable.

Secure weak fingerprinting schemes are defined as in the standard model and the *weak t -fingerprinting capacity* is the maximal rate achievable by a t -secure weak fingerprinting scheme.

In the rest of this section we adapt our fingerprint codes constructed in Section 2 to the weak model while increasing their rate. Both the minimax result (Lemma 5.1) the construction is based on and the correctness (Theorem 5.1) are proved the same way as in the standard model, so we omit these proofs.

For $i \in [t]$, $p \in [0, 1]$ and S an eligible channel we define $I_{p,S}^{(i)}$ to be the conditional mutual information $I(x_i; f|x_{(i)})$ in the probability space $B_{p,S}$. Here $x_{(i)} := (x_1 \dots x_{i-1} x_{i+1} \dots x_t)$.

Now we consider the game in which Paula chooses a channel S and David chooses a pair (p, i) with $p \in [0, 1]$ and $i \in [t]$. After their simultaneous choices Paula pays $I_{p,S}^{(i)}$ to David. As in Section 2 we argue that Paula is always better off using a pure strategy than a mixed one. As for Lemma 4.1 we use symmetry to prove that David can choose an optimal mixed strategy (distribution over (p, i)) in which p and i are independent and i is distributed uniformly over $[t]$.

LEMMA 5.1. *We have*

$$\max_D \min_S E_{p \in D, i \in U_t} [I_{p,S}^{(i)}] = \min_S \max_{p,i} I_{p,S}^{(i)},$$

where $p \in [0, 1]$, $i \in [t]$, S is an eligible channel, D is a distribution over $[0, 1]$ and U_t is the uniform distribution over $[t]$.

We use S_w^t to denote a channel S minimizing the right hand side and D_w^t to denote a distribution D maximizing the left hand side. We let W_t stand for the common value of this minimum and maximum.

Let $F_{t,R,\delta,n}$ be the following weak fingerprint code. The parameters have the same meaning as in $E_{t,R,\delta,n}$. For simplicity we assume here that the number of pirates is exactly t . As the accusation algorithm learns the number of pirates anyway, it is easy to remove this simplifying assumption.

Codeword generation: We use bias-based codeword generation with the bias distribution D_w^t . The code-length is n and the number of users is $\lfloor 2^{Rn} \rfloor$.

Oracle: Let $u = (u_1, \dots, u_t)$ be the t -tuple of pirates and let S be their perceived strategy. The oracle chooses a pirate u_i with $E_{p \in D_t} [I_{p,S}^{(i)}] \geq W_t$ and outputs the other pirates. By the definition of W_t and D_t such index i always exists.

Accusation algorithm: Let $T' = \{u_1, \dots, u_{t-1}\}$ be the output of the oracle. For any user $u_t \in U \setminus T'$ we consider the tuple $u = (u_1, \dots, u_{t-1}, u_t)$ and the perceived strategy S of u . We accuse u_t if S is eligible, the total variation distance between the distributions B_u and $B_{D_t,S}$ is at most δ and $E_{p \in D_t} [I_{p,S}^{(t)}] \geq W_t$. In case no user or more than one user satisfies these conditions the accusation algorithm fails.

THEOREM 5.1. *For any t and $R < W_t$ there exists a positive δ such that the fingerprinting codes $E_{t,R,\delta,n}$ for $n \geq 1$ form a t -secure fingerprinting scheme of rate R .*

6 Upper bounds on the fingerprinting capacity

Recall that [9] contains a lower bound on the length of fingerprinting codes but as it does not depend on the number of users it does not translate to an upper bound on the rate. To fix this we make this bound depend on

the number of users by first reducing the error bound while restricting the code to a small set of users, then applying the original bound to this restricted code. As the upper bound in [9] applies very generally, so does this bound: it applies to the so called *unreadable digit model* and it also shows that using non-binary alphabets do not make fingerprinting schemes shorter. We state the result in its simplest form here. For the proof see the full version of this paper.

THEOREM 6.1. *The t -fingerprinting capacity is $\Theta(1/t^2)$.*

The only previous upper bound on the t -fingerprinting capacity was $O(1/t)$ by Anthapadmanabhan, Barg and Dumer [1]. This was derived from what they call the “weak converse upper bound”: $R \leq \min_S \max_p I_{p,S}$. This bound gives $R \leq V_t = tR_t = \Theta(1/t)$.

In the same paper they also prove the “strong converse upper bound” on the t -fingerprinting capacity:

$$(6.3) \quad R \leq C_t := \min_S \max_D \max_i I(x_i; f|x_{(i)}),$$

where the minimization is over eligible channels S , the maximization is for distributions D over the independent Bernoulli random variables x_i and I denotes conditional mutual information in the probability space obtained by choosing the variables x_j according to D and getting the 0-1 variable f from $x = (x_1, \dots, x_t)$ through the channel S . Recall that $x_{(i)} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_t)$. The proof of this bound also applies to the weak fingerprinting capacity.

The strong converse upper bound is indeed stronger than the previous one but it is harder to evaluate it asymptotically and in [1] C_t is only estimated for $t = 2$ and 3 (by setting S to be the *uniform channel*: $S(x) = |x|/t$). We find exact asymptotics using approximation theory. See the proof in the full version.

THEOREM 6.2. *The (weak) t -fingerprinting capacity is at most C_t with*

$$C_t = \frac{\pi^2}{8 \ln 2} t^{-2} + o(t^{-2}) = \frac{1.77 \dots + o(1)}{t^2}.$$

The strongest upper bound is obtained by strengthening the bound (6.3) of [1] by limiting the maximization for iid. random Bernoulli variables x_i . This simplifies the formula and makes it equal to the one defining W_t . Therefore the upper bound and the rate of our weak codes coincide.

THEOREM 6.3. *The weak t -fingerprinting capacity is $W_t = \min_S \max_{p,i} I_{p,S}^{(i)}$. We have*

$$W_t = \frac{1}{(2 \ln 2)t^2} + o(t^{-2}).$$

The proof of this theorem is to be published separately, we state it here as it implies the optimality of the weak fingerprinting codes constructed in this paper and the asymptotic optimality of our standard fingerprinting codes.

t	2	3	4	5	6	7
R_t	0.250	0.097	0.054	0.034	0.023	0.017
W_t	0.311	0.137	0.078	0.050	0.034	0.025

Table 2: Lower and upper bounds on the t -fingerprinting capacity for t up to 7

7 Concluding remarks

In this section we list further directions of research related to the present paper.

Clearly, we would like to find the exact value of the t -fingerprinting capacity. In particular, we would like to see our Conjecture 2.1 proved establishing that our codes achieve optimal rate.

It would also be desirable to improve the analysis of the codes in this paper to have useful estimates on the achievable rates with prescribed length and error probabilities. Our result here give tight estimates of the limit of the rate as the codelength goes to infinity but we do not say anything about the rates for specific finite parameters. Note that [9] gave codes and bounds for all settings of the parameters.

To obtain such explicit rate bounds one has to make all computations in this paper explicit. This job is straightforward for most of the paper, here we mention a single point that requires more elaboration: The size of the probability space plays an important role in most of our estimates. Our variables are binary except for the bias: we choose the biases according to the discrete distribution D^t . Therefore we need to estimate the size of the support of D^t . For this we have:

LEMMA 7.1. *The size k_t of the support of the distribution D^t satisfies*

$$\sqrt{\frac{t}{(4 \ln 2) \log t}} \leq k_t \leq \frac{5}{2}t - 1.$$

See the proof in the journal version of this paper.

Finally we consider the computational complexity of our fingerprinting codes. We find that the code generation is very efficient but the accusation algorithm is rather slow. This algorithm has to guess a t -tuple u of users and an eligible channel S for which the joint type B_u is very close to the distribution $B_{D^t, S}$. Finding the pair with the smallest distance would mean optimizing

for both u and S , but by Lemma 3.1 it is enough to optimize for u and take S to be the perceived strategy of u . Computing the perceived strategy S and the total variation distance between $B_{D^t, S}$ and B_u can be done efficiently in time $O(tn + 2^t k_t)$. The bottleneck is that all t -tuples of users must be considered. This is much slower than the efficient algorithm in [9] that decides in time linear in the codelength whether a given user is to be accused.

It would be interesting to see if it is possible to combine the high rates of the codes in this paper with the efficiency and simplicity of the accusation algorithm in [9] or there is an interesting type of tradeoff between the rates of fingerprinting codes and the efficiency of their accusation algorithm.

References

- [1] N.P. Anthapadmanabhan, A. Barg, I. Dumer, Fingerprinting capacity under the marking assumption. Submitted to IEEE Transaction on Information Theory - Special Issue on Information-theoretic Security. Available from arXiv:cs/0612073v3. Preliminary version appeared in the Proceedings of the 2007 IEEE International Symposium on Information Theory, (ISIT 2007), 2007.
- [2] O. Blayer, T. Tassa, Improved versions of Tardos' fingerprinting scheme. Submitted.
- [3] D. Boneh, J. Shaw, Collusion-secure fingerprinting for digital data, *IEEE Transactions of Information Theory* **44** (1988), 480–491.
- [4] I. Csiszár, J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, Academic Press, London, 1981.
- [5] K. Nuida, S. Fujitsu, M. Hagiwara, T. Kitagawa, H. Watanabe, K. Ogawa, H. Imai, An improvrmrny of the discrete Tardos fingerprinting codes, *Cryptology ePrint Archive*, Report 2008/338.
- [6] K. Nuida, M. Hagiwara, H Watanabe, H. Imai, Optimization of memory usage in Tardos's fingerprinting codes. Available from arXiv:cs/0610036.
- [7] B. Skoric, S. Katzenbeisser, M.U. Celik, Symmetric Tardos fingerprinting codes for arbitrary alphabet sizes. To appear in Designs, Codes and Cryptography. Available at <http://members.home.nl/skoric/security/symmetricTardos.pdf>.
- [8] B. Skoric, T.U. Vladimirova, M. Celik, J.C. Talstra, Tardos fingerprinting is better than we thought. Available from arXiv:cs/0607131.
- [9] G. Tardos, Optimal probabilistic fingerprint codes, *Journal of the ACM*, to appear. Preliminary version appeared in Proceedings of the 35th Annual ACM Symposium on Theory of Computing, (STOC 2003), 116–125.