Private codes
or
Succinct random codes that are (almost) perfect

Michael Langberg *

Abstract

Coding theory addresses the design and analysis of codes that enable communication over noisy channels. Two types of channels that have been extensively considered are the binary symmetric *channel and the* adversarial *channel. In a binary symmetric channel each bit of the sent message is flipped independently with some probability p, implying that the noise imposed by the channel is* random *in nature where the* amount *of noise is determined by p. In an adversarial channel the message is treated as a whole, and the noise may be an arbitrary (and malicious) function of the message being sent, as long as it does not effect more that a certain fraction (say p) of the bits transmitted.*

Roughly speaking, any code designed for an adversarial channel can be used on a corresponding binary symmetric channel successfully, whereas the contrary is not necessarily true. In this work we will present a construction that transforms the best *codes for binary symmetric channels into "codes" for corresponding adversarial channels.*

The "codes" we present assume that the sender and the receiver of the message have a joint secret random string (which is not known to the channel). These codes are referred to as private *codes. Intuitively, this private randomness allows a reduction between the random and adversarial channels. Such a reduction is simple once the size of the joint random string is $\Theta(n \log n)$ (here the codes are a subset of $\{0,1\}^n$).*

In this work we present private codes in which the size of the joint random string is $O(\log n)$. Moreover, we show that our result is tight. Namely, to design private codes that allow communication over adversarial channels that meet *the bounds achievable when communicating over binary symmetric channels, an amount of $\Omega(\log n)$ shared random bits are required. To the best of our knowledge, no prior results of this nature have been presented in the past. As part of our proof we establish a connection between list decodable codes and private codes which complements a recent result of Guruswami (CCC '03) on list decoding with side information.*

1 Introduction

Consider the case in which A (the sender) wants to send k bits of information to B (the receiver) over a channel which is noisy. The standard framework applied in such cases is to use an $[n,k]$ code $\mathcal{C} : \{0,1\}^k \longrightarrow \{0,1\}^n$, and a decoder $\mathcal{D} : \{0,1\}^n \longrightarrow \{0,1\}^k$ (here n is taken to be larger than k). The message $w \in \{0,1\}^k$ is first encoded by A into $\mathcal{C}(w)$ and only then transmitted over the noisy channel to B. Using the redundancy of the message $\mathcal{C}(w)$, the receiver B is now able to decode (via $\mathcal{D}$) the obtained corrupted message.

The design of $\mathcal{C}$ for which such a scheme succeeds depends strongly on the properties of the channel connecting A and B. Two channels that have been extensively studied in the past are the binary symmetric channel and the adversarial channel. In a binary symmetric channel (BSC), each bit of the message sent is flipped (from 0 to 1 or visa versa) independently with some probability p (we will denote these channels as BSC$_p$). Roughly speaking, the noise added when communicating over a BSC$_p$ can be viewed as a *random* vector in $\{0,1\}^n$ of Hamming[1] weight approximately pn. In an adversarial channel (which we denote as ADVC), the message is treated as a whole, and the error imposed by the channel may be an arbitrary (and malicious) function of the message being sent. As in the case of the BSC, we consider ADVC in which the hamming weight of the error is bounded by pn, we denote such channels as ADVC$_p$.

An $[n,k]$ code $\mathcal{C}$ is said to be of rate R if $R = \frac{k}{n}$. The design of rate R codes which allow communication over noisy channels is a central theme in coding theory. In his funda-

*Department of Computer Science, California Institute of Technology, Pasadena, California 91125. E-mail: mikel@caltech.edu.

[1]The Hamming weight of a vector is defined to be the number of entries in the vector which differ from 0.

mental work, Shannon [12] proved the existence of codes $\mathcal{C}$ and decoders $\mathcal{D}$ which enable[2] communication over a BSC_p where p satisfies the inequality $H(p) < 1 - R$. Here, H is the standard binary entropy function, R is assumed to be constant, and n is assumed to be sufficiently large. The value of p which satisfies this equation with equality will be referred to as the Shannon bound[3]. Moreover, Shannon proved that once p is larger than the Shannon bound, no pair of $\mathcal{C}$ and $\mathcal{D}$ exist that allow communication with non-negligible probability over BSC_p.

One would be interested to obtain similar bounds when considering adversarial channels. To obtain codes $\mathcal{C}$ that can be used when an adversarial channel is present, one must take into consideration the *minimum distance* of a code (*i.e.* the minimum Hamming distance between any two given codewords in the image of $\mathcal{C}$). Namely, it is well known that a code can be used under the presence of an ADVC_p iff it has minimum distance at least $2pn$. However, as apposed to the case in which the binary symmetric channel is present, the *best* adversarial codes (those with the largest minimum distance) are not necessarily known. This question has attracted a vast amount of study in the last several decades.

To match the results of Shannon regarding binary symmetric channels, when considering communication over adversarial ones, one must design codes with minimum distance which is arbitrarily close to twice the Shannon bound (times n). Roughly speaking, codes in which this distance is "exactly" twice the Shannon bound (times n) are known as *perfect* codes. The existence of rate R perfect codes implies the ability to communicate over an ADVC_p in which p matches the Shannon bound. Similarly, the existence of $[n, k]$ codes which are *almost* perfect (we defer the exact definition of *almost* until later in this discussion) would imply the ability to communicate over an ADVC_p in which p comes arbitrarily close to the Shannon bound (matching the results on BSC_p).

Unfortunately, such codes do not exist (at least not asymptotically). Namely, by well known bounds in coding theory (*e.g.* the Elias-Bassalygo bound [4, 14]), it follows that for any constant rate R, there do not exist infinite families of (almost) perfect codes of rate R. This rules out the possibility to match the results achievable on binary symmetric channels when considering communication over adversarial ones.

Nevertheless, in this work we address the existence of "codes" which are successful against an ADVC_p when p is arbitrarily close to the Shannon bound (or in other words, "codes" which are almost perfect). To do so, we consider *relaxing* the standard notion of codes. The relaxation we consider involves a shared random *key* which is known to both the sender A and the receiver B, but is assumed to be kept secret from the adversarial channel.

[2]We defer defining rigorous success criteria of a communication protocol to later in this extended abstract.

[3]Also referred to as the Gilbert-Varshamov bound.

Private codes More precisely, we assume that both A and B have a joint random string, say r, uniformly distributed in $\{0,1\}^m$. When A wants to send the message $w \in \{0,1\}^k$ to B, the random string r is used as input to the encoder. That is the encoder used by A is now a function $\mathcal{C} : \{0,1\}^k \times \{0,1\}^m \longrightarrow \{0,1\}^n$. On receiving a (corrupted) message c, B now decodes c using a decoder which takes the joint randomness r into account. Namely, the decoder is now defined as $\mathcal{D} : \{0,1\}^n \times \{0,1\}^m \longrightarrow \{0,1\}^k$. The adversarial channel remains a malicious function of the codeword being sent, but does not have any knowledge of the joint random string held by A and B.

We refer to these codes as *private* codes (and to these decoders as private decoders). A private code and decoder are said to be successful against an ADVC_p if with high probability over the joint random string r, B is able to decode the information sent by A (over the ADVC_p). Private codes have been defined and studied in the past. They first appear in [13] where they are referred to as *random* codes. Namely, the function $\mathcal{C} : \{0,1\}^k \times \{0,1\}^m \longrightarrow \{0,1\}^n$ is viewed as a distribution over the collection $\{\mathcal{C}_r \mid r \in \{0,1\}^m\}$ of standard $[n, k]$ codes, where $\mathcal{C}_r$ is defined as the restriction of $\mathcal{C}$ to r (*i.e.* $\mathcal{C}_r(w) = \mathcal{C}(w, r)$). We refer to such codes as "private" instead of "random" to avoid any mixup with the notion of codes $\mathcal{C}$ which are a *random* function $\{0,1\}^k \longrightarrow \{0,1\}^n$ (or equivalently, codes $\mathcal{C}$ for which their image is a random subset of $\{0,1\}^n$ of size 2^k).

Since their introduction, private codes have been considered in a number of works in the Information Theory community. For example in [5], the performance of standard and private codes used in the presence of several types of channels was studied. One channel studied in this work was the so called *arbitrary varying channel*, in which the noise imposed on each bit of the sent message is an independent Bernoulli trail governed by a (potentially) different value of p (such channels have a natural extension to larger alphabets as well). In [1, 6] the study of private codes in the presence of an arbitrarily varying channel is continued, and various relations between standard and private codes are presented in this setting. More recent works, which address how the two parties A and B get hold of a shared random string, include [2, 3]. In these works, an information theoretic approach is taken to the study of *secret sharing* under several different settings.

To the best of our knowledge, the amount of shared randomness needed in private codes that can be used in the presence of an adversarial channel has not been explicitly studied in the past.

Our results In this work, we study the minimum random key size required to enable communication via private codes on ADVC_p for values of p that come arbitrarily close to the Shannon bound. Intuitively, the private randomness shared by A and B allows a reduction between the random binary symmetric channel and the adversarial channel. Such a reduction is simple once the size of the joint random string is *large*. Loosely speaking, once one allows $\Theta(n \log n)$ random bits shared between A and B it can be shown that for any value of p, any code which allows communication in the presence of a BSC_p can be used to construct private codes and decoders which allow communication over an ADVC_p.

We address the question that now arises naturally: How small can the shared randomness be? We present matching upper and lower bounds. As for the upper bound, we show that $O(\log n)$ random bits is enough to obtain private codes which are privately decodable in the presence of an ADVC_p where p is arbitrarily close to the Shannon bound. Moreover, we present two types of "lower" bounds. First of all we show that no matter how large the random key is, there do not exist private codes which enable communication over an ADVC_p for values of p beyond the Shannon bound. Secondly, at least $\Omega(\log n)$ shared random bits are required in the design of private codes which come arbitrarily close to the Shannon bound.

1.1 Formal statement of our results

We now give a few formal definitions followed by the main results of this paper. For concreteness, throughout this work we consider codes over the binary alphabet $\{0,1\}$. Our results have a natural extension to general alphabets as well.

Definition 1.1 *An $[n,k,m]$ private code is a function $\mathcal{C} : \{0,1\}^k \times \{0,1\}^m \longrightarrow \{0,1\}^n$.*

Definition 1.2 *A private decoder of an $[n,k,m]$ private code is a function $\mathcal{D} : \{0,1\}^n \times \{0,1\}^m \longrightarrow \{0,1\}^k$.*

Definition 1.3 *An adversarial channel $\mathcal{S}$ is a function $\mathcal{S} : \{0,1\}^n \longrightarrow \{0,1\}^n$. $\mathcal{S}(c)$ corresponds to the error an adversary may induce on a given message $c \in \{0,1\}^n$. An adversarial channel $\mathcal{S}$ is said to be in $ADVC_p$ if for every c the Hamming weight of $\mathcal{S}(c)$ is less than pn.*

Notice that we have abused our previous notation. Namely, ADVC_p now denotes a set of channels $\mathcal{S} : \{0,1\}^n \longrightarrow \{0,1\}^n$ instead of the single adversarial channel. We now define the success criteria of communication via private codes. Loosely speaking, a private code $\mathcal{C}$ and a private decoder $\mathcal{D}$ enable the communication of w over a channel $\mathcal{S}$ if, for many values of r, the word w can be obtained (via $\mathcal{D}$) from the transmitted codeword $\mathcal{C}(w,r)$ corrupted by the noise $\mathcal{S}(\mathcal{C}(w,r))$ imposed by $\mathcal{S}$.

Definition 1.4 *Let $\mathcal{C}$ be an $[n,k,m]$ private code. Let $\mathcal{D}$ be a private decoding function. $\mathcal{C}$ is said to be privately decodable with probability $1-\delta$ within error parameter p by $\mathcal{D}$ iff $\forall w \in \{0,1\}^k$ and $\forall \mathcal{S} \in ADVC_p$ it holds that*

$$\Pr_{r \in \{0,1\}^m}[\mathcal{D}(\mathcal{C}(w,r) \oplus \mathcal{S}(\mathcal{C}(w,r)), r) = w] \geq 1-\delta.$$

Here the probability is taken over r uniformly distributed in $\{0,1\}^m$, and the $\oplus$ operator is bitwise XOR.

We now specify the main theorems proved in this work.

Theorem 1.5 (Upper bound on randomness) *Let $R \in (0,1)$ be constant. Let $p < 1/2$ satisfy $H(p) = 1-R$. Let $\varepsilon > 0$ be a sufficiently small constant. For every n sufficiently large and any $\delta > 2^{-R\frac{\varepsilon^2}{4}n}$ there exists an $\left[n, Rn, O\left(\log k + \log\left(\frac{1}{\delta}\right)\right)\right]$ private code which is privately decodable with probability $1-\delta$ within error parameter $p-\varepsilon$.*

Theorem 1.6 (Bound on rate) *Let $R \in (0,1)$ be constant. Let p satisfy $H(p) = 1-R$. Let $\varepsilon > 0$ and $\delta \in (0,1)$ be constant. Let n be sufficiently large. Let $\mathcal{C}$ be any private $[n,Rn,m]$ code. The private code $\mathcal{C}$ combined with any private decoder $\mathcal{D}$ is* **not** *privately decodable with probability $1-\delta$ within error parameter $p+\varepsilon$.*

Theorem 1.7 (Lower bound on randomness) *Let $R \in (0,1)$ be constant. Let p satisfy $H(p) = 1-R$. Let $\varepsilon > 0$ be a sufficiently small constant. Let $\delta < \frac{1}{8}$. For sufficiently large n, an $[n,Rn,m]$ private code that is privately decodable with probability $1-\delta$ within error parameter greater than $p-\varepsilon$ must have $m = \Omega\left(\log k + \log\left(\frac{1}{\delta}\right)\right)$.*

1.2 Proof techniques and list decodable codes

An additional relaxation to standard coding which allows communication over an adversarial channel for large values of p is *list decoding*. In list decoding, the decoding of a corrupted message is considered successful if the decoder $\mathcal{D}$ outputs a *small* list of words in $\{0,1\}^k$ which include the word transmitted. Viewing an $[n,k]$ code $\mathcal{C}$ as a subset of $\{0,1\}^n$ (the image of $\mathcal{C}$), $\mathcal{C}$ is said to be (ℓ, pn) list decodable if for every ball $\mathcal{B}$ of radius pn in $\{0,1\}^n$ the intersection $\mathcal{C} \cap \mathcal{B}$ is of size at most ℓ.

Upper bound In this work, we first establish a connection between list decodable codes and private codes. Roughly speaking, we show that any $[n,k]$ code $\mathcal{C}'$ which is (ℓ, pn) list decodable, can be used to construct an $[n, k', O(\log k + \log \ell)]$ private code $\mathcal{C}$ which is privately decodable within error parameter p (*i.e.* if ℓ is small the random key is only of logarithmic length). Here k' is slightly smaller than k ($k' \simeq k - \log k$).

We turn to describe our construction of $\mathcal{C}$. For every $r \in \{0,1\}^m$, let the standard $[n,k]$ code $\mathcal{C}_r$ be the restriction of $\mathcal{C}$ to r (*i.e.* $\mathcal{C}_r(w) = \mathcal{C}(w,r)$). To define $\mathcal{C}$ we must define $\mathcal{C}_r$ for every $r \in \{0,1\}^m$. In our reduction, given a list decodable code $\mathcal{C}'$, we take $\mathcal{C}_r$ to be a *large* subset of $\mathcal{C}'$. We need $\mathcal{C}_r$ to be large as to preserve the rate of the private code we are constructing. A natural construction would be to define the code $\mathcal{C}_r$ as a random subset of $\mathcal{C}'$ (independently for each r). However, it can be seen that this will not yield a private code $\mathcal{C}$ with the required properties. This follows from the fact that in such a construction, the codeword $\mathcal{C}(w,r)$ tends to reveal the value of r, which takes us back to the study of standard codes. Thus, one must govern the construction of $\mathcal{C}$ in a different manner. In our construction, for each r we define $\mathcal{C}_r$ according to a certain distribution on subsets of $\mathcal{C}'$. This distribution will imply dependencies between the codes $\mathcal{C}_r$ which in turn yield a private code $\mathcal{C}$ with the required properties.

Our random construction can be made *efficient* using a slight variation to the well known combinatorial design presented in [9]. Namely, using this combinatorial design, one can construct efficient private codes from efficient list decodable codes. Here we refer to a code as being efficient, if both $\mathcal{C}$ and $\mathcal{D}$ have polynomial (in n) time complexity[4]. Hence we obtain the following result of independent interest.

Theorem 1.8 *Let $R \in (0,1)$ be constant. Let n be sufficiently large, and let $k = Rn$. Let $\mathcal{C}'$ be an $[n,k]$ code which is (ℓ, pn) list decodable. Let $M = \frac{k^2\ell^2}{\delta^2}$. There exists an $[n, k - \frac{1}{2}\log M, \log M]$ private code which is privately decodable with probability $1-\delta$ within error parameter p. Moreover, if $\mathcal{C}'$ is efficiently encodable and list decodable, then the resulting private code is also efficiently encodable and decodable.*

Once this connection between list decoding and private decoding has been established. Our upper bound follows by the existence of codes of rate R which are (ℓ, pn) list decodable for small values of ℓ and values of p arbitrarily close to the Shannon bound.

"Lower" bounds As for our negative results, very roughly speaking they are obtained by a variety of averaging arguments, and rely strongly on the non-existence of almost perfect codes. Almost perfect codes are defined below.

Definition 1.9 *Let $R \in (0,1)$ be constant. Let p satisfy $H(p) = 1-R$. Let ε_R and ε_p be positive constants. A code $\mathcal{C}$, with range $\{0,1\}^n$, is said to be $(\varepsilon_R, \varepsilon_p)$-*perfect *with respect to R if it has rate at least $(1-\varepsilon_R)R$ and minimum distance greater than $2(p-\varepsilon_p)n$.*

[4] In this context, we do not allow $\mathcal{C}$ or $\mathcal{D}$ to have access to an exponentially large lookup table.

As mentioned previously, for small values of ε, it can be seen that the existence of $(\varepsilon, \varepsilon)$-perfect codes implies the existence of private codes **without any shared randomness** which come close to the Shannon bound. Thus in order for us to prove any lower bounds we must first rule out the existence of an infinite family of $(\varepsilon, \varepsilon)$-perfect codes. Luckily, this is known (*e.g.* the Elias-Bassalygo bound), and will be used to establish Theorem 1.7.

1.3 Connection to list decoding with side information

In a recent paper, Guruswami [10] considered a reduction (of nature similar to that presented in Theorem 1.8) between list decodable codes and error correction in the presence of *side information*. Namely, it was shown that if the sender A and the receiver B have access to both a noisy and noiseless channel, then list decoding can be reduced to unique decoding when only a logarithmic amount of bits are sent on the noiseless channel. Let $\mathcal{C}$ be a list decodable code. In the framework suggested by [10], given $w \in \{0,1\}^k$, the sender A sends $\mathcal{C}(w)$ on the noisy channel, and sends $f(w)$ on the noiseless channel. Here f is some (randomized) function of w. Upon receival of the corrupted message c and the side information $f(w)$, B now uses a decoder $\mathcal{D}$ that takes as input both c and $f(w)$ to recover w. In [10], matching upper and lower bounds on the size of $f(w)$ are presented.

How do our results compare to the results stated in [10]? First consider the difference between the model of private codes and the model of side information. In both models, the shared information is assumed to be unknown to the adversarial channel. However, the joint information between A and B in the case of private codes is random and **independent** of the message w being sent, whereas in the case of side information the joint information (may) depend strongly on the message w. Thus our upper bounds in Theorem 1.8 (almost) imply those presented in [10]. Why almost? Notice that in Theorem 1.8 we have a slight loss in the rate of our private code (compared to the original rate of the list decodable code), whereas no such loss appears in the side information scenario.

The discussion above also implies that any lower bounds proven on the size of *dependent* side information (as studied in [10]) will yield lower bounds on the size of *independent* side information (as studied in this paper). So are the lower bounds presented in this work a direct consequence of those presented in [10]? This is not the case. In the framework studied by Guruswami, the sender A is assumed to encode the message w using a **standard** code $\mathcal{C}$. That is, the code $\mathcal{C}$ is a function that depends only on w and not on the joint shared information $f(w)$ (recall that f may be randomized). Potentially, it may be the case that the lower bounds implied by those presented in [10] do not hold when one allows the stronger model in which $\mathcal{C}$ is dependent on both w and the side information $f(w)$. However, it turns out that this is not the case. Specifically, in Theorem 1.7, we prove our lower bound on the size of the side information $f(w)$ under the assumption that $f(w)$ is independent of w (*i.e.* $f(w)$ is just a random string). Only few modifications in the proof of Theorem 1.7 are needed to obtain similar bounds for general $f(w)$ as well (details are omitted from this abstract).

1.4 Organization

The remainder of the paper is organized as follows. In Section 2 we present some definitions and notations that will be used throughout our work. This section also includes the reduction between codes that allow communication over BSC_p and private codes, with $\Theta(n \log n)$ shared random bits, that allow communication over adversarial channels. The combinatorial design used in our reduction between list decodable codes and private codes is presented in this section as well. In Section 3 we present our reduction between list decodable codes and private codes. In Section 4 our positive results on private codes which (almost) meet the Shannon bound are presented. Finally, in Section 5 we present our negative results ("lower" bounds).

2 Preliminaries

We start by the following definitions and notation which will be used throughout our work. For a positive integer a, we denote the subset $\{1, 2, \dots, a\}$ by $[a]$.

Definition 2.1 (Hamming distance, Hamming ball) *Let $w = w_1 w_2 \dots w_n$ and $w' = w'_1 w'_2 \dots w'_n$ be two elements in $\{0,1\}^n$. The Hamming distance $\Delta(w, w')$ between w, w' is defined to be the number of positions $i \in [n]$ in which $w_i \neq w'_i$. The Hamming ball of radius d centered at $c \in \{0,1\}^n$ is defined to be $\{w \in \{0,1\}^n \mid \Delta(w,c) \leq d\}$.*

Notice that the size of a Hamming ball of radius $d \leq \frac{n}{2}$ in $\{0,1\}^n$ is $Vol(d) = \sum_0^d \binom{n}{i} \leq 2^{H(d/n)n}$ (where the last inequality appears for example in [11]). Here $H(x) = -x \log x - (1-x) \log(1-x)$ is the standard (binary) entropy function.

Our first reduction We now show the reduction between codes $\mathcal{C}'$ that allow communication over BSC_p and private codes $\mathcal{C}$ with *many* shared random bits that allow communication over ADVC_p. The main idea in the design of $\mathcal{C}$ is as follows. Given a word c from the image of the private code $\mathcal{C}$, we would like to hide as much information from the adversarial channel regarding the source of c (that is which w and r were used by A to yield $c = \mathcal{C}(w,r)$). This way, roughly speaking, any error imposed by the channel is *no better* than a random error. Hence, if we design such a private code $\mathcal{C}$ which also enables communication over BSC_p, we are done. More specifically, given $\mathcal{C}'$, we define $\mathcal{C}(w,r)$ to be a random permutation of $\mathcal{C}'(w)$, XOR'ed by a random string in $\{0,1\}^n$. The random permutation and string are governed by r chosen from a large enough domain. We would like to remark that both the random permutation, and the random shift (XOR) are needed to make the reduction general. We will use the following definitions.

Definition 2.2 *Let $\mathcal{C}$ be a (standard) $[n,k]$ code. Let $\mathcal{D}$ be a (standard) decoding function. $\mathcal{C}$ is said to be decodable in the presence of a BSC_p with probability $1-\delta$ by $\mathcal{D}$ iff $\forall w \in \{0,1\}^k$ it holds that*

$$\Pr_e[\mathcal{D}(\mathcal{C}(w) \oplus e) = w] \geq 1 - \delta.$$

Here the probability is taken over a random error $e \in \{0,1\}^n$ chosen by the distribution implied by BSC_p, and the $\oplus$ operator is bitwise XOR. $\mathcal{C}$ is said to be decodable on average *in the presence of a BSC_p with probability $1-\delta$ by $\mathcal{D}$ iff it holds that*

$$\Pr_{w,e}[\mathcal{D}(\mathcal{C}(w) \oplus e) = w] \geq 1 - \delta.$$

In this case, the probability is taken over a random $w \in \{0,1\}^k$ (chosen uniformly) and a random error e (chosen from the distribution implied by BSC_p).

Definition 2.3 *Let $\mathcal{C}$ be a (standard) $[n,k]$ code. A decoder $\mathcal{D}$ is said to be a* Maximum Likelihood *decoder corresponding to $\mathcal{C}$ if, given $c \in \{0,1\}^n$, $\mathcal{D}(c) = w$ where w is one of the vectors in $\{0,1\}^k$ satisfying $\Delta(\mathcal{C}(w), c) = \min_{w \in \{0,1\}^k} \Delta(\mathcal{C}(w), c)$.*

Proposition 2.4 *Let $p \in (0, 1/2)$ be constant. Let n be a sufficiently large constant. Let $\mathcal{C}'$ be an $[n,k]$ code which is decodable via a Maximum Likelihood decoder in the presence of a BSC_p with probability $1-\delta$. For $m = \Theta(n \log(n))$ there exists a private $[n,k,m]$ code $\mathcal{C}$ which is privately decodable with probability $1-3\delta$ within error parameter p.*

Proof: Let $\mathcal{C}'$ be as defined above, and let $\mathcal{D}'$ be a Maximum Likelihood decoder. Our private code $\mathcal{C}$ is defined as follows. We pick m to be of size $\log(n!2^n) = \Theta(n \log n)$, thus we may identify any $r \in \{0,1\}^m$ with a pair $\langle \sigma, a \rangle$ where σ is a permutation in S_n and $a \in \{0,1\}^n$. Denote the permutation and word corresponding to $r \in \{0,1\}^m$ as σ_r and a_r. Now define $\mathcal{C}(w,r)$ to be $\sigma_r(\mathcal{C}'(w)) \oplus a_r$. The decoder $\mathcal{D}$ of $\mathcal{C}$ is defined as $\mathcal{D}(c,r) = \mathcal{D}'\left(\sigma_r^{-1}(c \oplus a_r)\right)$, it is not hard to verify that $\mathcal{D}(\mathcal{C}(w,r), r) = w$.

Let $w \in \{0,1\}^k$. Consider any adversarial channel $\mathcal{S} \in \mathrm{ADVC}_p$. Let Γ be the random error distribution of the BSC_p. By our assumption on $\mathcal{C}'$, we have that $\Pr_{e\in\Gamma}[\mathcal{D}'(\mathcal{C}'(w) \oplus e) = w] \geq 1 - \delta$. Let $\Gamma(d)$ be this distribution conditioned on the event that the error $e \in \Gamma$ is of Hamming weight exactly d. The properties of $\mathcal{D}'$ and Γ imply that for every $d \leq pn$ we have $\Pr_{e\in\Gamma(d)}[\mathcal{D}'(\mathcal{C}'(w) \oplus e) = w] \geq 1 - 3\delta$.

For any $c \in \{0,1\}^n$ notice that there are exactly $n!$ values of r which satisfy $\mathcal{C}(w,r) = c$. Furthermore, each such value of r has a distinct corresponding permutation σ_r. Thus, $\Pr_{r\in\{0,1\}^m}[\mathcal{D}\left(\mathcal{C}(w,r) \oplus \mathcal{S}(\mathcal{C}(w,r)), r\right) = w]$

$= \frac{1}{2^n}\sum_{c\in\{0,1\}^n} \Pr_r[\mathcal{D}\left(c \oplus \mathcal{S}(c), r\right) = w \mid \mathcal{C}(w,r) = c]$

$= \frac{1}{2^n}\sum_{c} \Pr_r[\mathcal{D}\left(c \oplus \mathcal{S}(c), r\right) = w \mid \sigma_r(\mathcal{C}'(w)) \oplus a_r = c]$

$= \frac{1}{2^n}\sum_{c} \Pr_{\sigma\in S_n}[\mathcal{D}'\left(\mathcal{C}'(w) \oplus \sigma^{-1}\left(\mathcal{S}(c)\right)\right) = w]$

$= \frac{1}{2^n}\sum_{c} \Pr_{e\in\Gamma\left(|\mathcal{S}(c)|\right)}[\mathcal{D}'(\mathcal{C}'(w) \oplus e) = w] \geq 1 - 3\delta$ $\square$

The exitances of private codes which almost meet the Shannon bound with $\Theta(n\log n)$ shared random bits now follows from the existence of standard $[n,k]$ codes which are decodable in the presence of a BSC_p (for p arbitrarily close to the Shannon bound) by a Maximum Likelihood decoder (*e.g.* [7]).

A combinatorial design We now turn to define the combinatorial design that will be used later in Section 3. The design and its efficient construction are a slight variation to the well known construction of [9].

Definition 2.5 *Let A be a set of elements of size M. Let α, β be positive integers. A set system $\mathcal{A} = \{A_{ij}\}_{i\in[\alpha], j\in[\beta]}$ of subsets of A is said to be an $[\alpha,\beta,X,Y]$ set system iff*

1. *For each $i \in [\alpha]$ and $j \in [\beta]$ the size of A_{ij} is X.*
2. *For each $i \in [\alpha]$ the set system $\mathcal{A}_i = \{A_{i1},\ldots,A_{i\beta}\}$ is a partition of A.*
3. *For any i, i' and any j, j' it is the case that $|A_{ij} \cap A_{i'j'}| \leq Y$ (unless $i = i'$ and $j = j'$).*

Notice that using the definition above, the size of the base set M is equal to βX and thus does not appear explicitly in our the notation. We are interested in constructing *efficient* set systems. A set system is said to be efficient if there exist an efficient algorithm (time complexity polynomial in $\log\alpha$ and $\log M$), that given an element $a \in A$ and an index $i \in [\alpha]$ finds (the unique) $j \in [\beta]$ s.t. $a \in A_{ij}$. Using ideas of [9] we have

Lemma 2.6 *For infinitely many integers M and any small $\varepsilon \geq \frac{1}{M}$ there exists an efficient $[\alpha,\beta,X,Y]$ set system of $A = [M]$ for $\alpha = \sqrt{M}^{\frac{1}{\varepsilon}-1}$, $\beta = \sqrt{M}$, $X = \sqrt{M}$, $Y = \frac{1}{\varepsilon}$.*

Proof: Let $\mathcal{F}$ be a finite field of size $\sqrt{M}$ (we assume M is chosen to satisfy this constraint). Let $A = \mathcal{F} \times \mathcal{F}$. Let $d = \left\lfloor \frac{1}{\varepsilon} \right\rfloor$. Let $[\mathcal{F}]^{d+1}$ denote the set of polynomials of degree d over $\mathcal{F}$. For a polynomial $f \in [\mathcal{F}]^{d+1}$ let A_f be the graph of f. Namely let $A_f = \{(x, f(x)) \mid x \in \mathcal{F}\} \subseteq A$. In general, the set system $\mathcal{A}$ will consist of the subsets A_f. Each such subset is of size exactly $\sqrt{M}$. Moreover, for any $f \neq g \in [\mathcal{F}]^{d+1}$, the intersection between A_f and A_g is bounded by d. Thus, properties (1) and (3) of Definition 2.5 will be satisfied by $\mathcal{A}$. It is left to consider property (2).

Let $\alpha = |\mathcal{F}|^d$, and identify $i \in [\alpha]$ with the i'th element in $\mathcal{F}^d$ (ordered in lexicographic order). Denote this correspondence by $\alpha(i)$. Let $\beta = |\mathcal{F}|$, again we identify $j \in [\beta]$ with a corresponding element $\beta(j)$ in $\mathcal{F}$. Now define A_{ij} to be A_f where f in the polynomial with free coefficient $\beta(j)$ an the remainder of coefficients defined by $\alpha(i)$. It is now the case that for a given $i \in \alpha$, the set $\{A_{i1},\ldots,A_{i\beta}\}$ is a partition of A. $\square$

3 List Decoding vs. Private Decoding

In this section we present our reduction between list decodable codes and privately decodable codes. Namely, we show that any $[n,k]$ code which is (ℓ, pn) list decodable, for a small value of ℓ, can be used to construct an $[n, k', O(\log k)]$ private code which is privately decodable within error parameter p. Here $k' = k - O(\log k)$. These results will be used in Section 4 to obtain $[n,k]$ privately decodable codes for strong adversarial channels. We start by proving the following lemma, it may be helpful for the reader to recall the discussion given in the Introduction related to list decoding.

Lemma 3.1 *Let $\mathcal{C}' \subset \{0,1\}^n$ be an $[n,k]$ code which is (ℓ, pn) list decodable. Let A be a set of size $M = 2^m$, and let $\mathcal{A}$ be an $[\alpha,\beta,X,Y]$ set system of A such that $\alpha\beta = |\mathcal{C}'| = 2^k$. Using $\mathcal{C}'$ and $\mathcal{A}$ one can construct an $[n, k - \lceil\log\beta\rceil, m]$ private code $\mathcal{C}$ which is privately decodable with probability at least $1 - \frac{\ell Y}{X}$ within error parameter p.*

Proof: We start by constructing our private code $\mathcal{C}(w,r)$. We will then proceed to prove that it is privately decodable with the asserted probability and error parameter. To construct $\mathcal{C}$ we need to define $\mathcal{C}(w,r)$ for every value of w and r. Our goal can be roughly summarized as follows. On one hand, we would like to define $\mathcal{C}$ such that for every w and most values of r the codeword $\mathcal{C}(w,r)$ does not reveal (too much information on) r. Loosely speaking, if this is not the case, our study of private codes is reduced to that of standard error correcting codes. On the other, we would like $\mathcal{C}(w,r)$ to depend on r, otherwise our construction will not express the full power obtainable using the joint randomness r. With this in mind, it is not hard to verify that defining $\mathcal{C}$ to be a random function with range $\{0,1\}^n$, or even a random function with range $\mathcal{C}'$, will not suffice (here and throughout this proof we view codes as both functions and ordered subsets of $\{0,1\}^n$).

We proceed to describe our deterministic construction of $\mathcal{C}$. For convenience, assume that $\log\alpha$ is an integer and let $k' = \log\alpha = k - \log\beta$. Recall that $\alpha\beta = 2^k = |\mathcal{C}'|$. In our construction we use an arbitrary partition of the list decodable code $\mathcal{C}'$ into α disjoint subsets each of size β. We refer to each such subset as $\mathcal{C}'_i = \{c_{i1}, \ldots, c_{i\beta}\}$. Namely, $\mathcal{C}' = \{c_{ij}\}$ where $i \in [\alpha]$ and $j \in [\beta]$.

Let $w \in \{0,1\}^{k'}$ and $r \in \{0,1\}^m$. To define $\mathcal{C}(w,r)$ we identify elements in $w \in \{0,1\}^{k'}$ with integers in $i \in [\alpha]$. Namely, w has a corresponding i which we denote as $i(w)$. We now define $\mathcal{C}(w,r)$ to be an element in $\mathcal{C}'_{i(w)}$. Namely we define $\mathcal{C}(w,r)$ to be c_{ij} iff $i = i(w)$ and $r \in A_{ij}$. That is, w determines (via $i(w)$) the index i in c_{ij} and r determines the index j (via the partition $\{A_{i1} \ldots A_{i\beta}\}$). This definition and the properties of $\mathcal{A}$ now imply that $\mathcal{C}$ satisfies: (a) Every element c_{ij} in $\mathcal{C}'$ is the image under $\mathcal{C}$ of exactly X pairs (w,r) (where w satisfies $i = i(w)$). (b) Let c_{ij} and $c_{i'j'}$ be two distinct elements in $\mathcal{C}'$, and let $i = i(w)$, $i' = i(w')$. Then there are exactly Y values of r such that both $\mathcal{C}(w,r) = c_{ij}$ and $\mathcal{C}(w',r) = c_{i'j'}$ (*i.e.* $r \in A_{ij} \cap A_{i'j'}$). As we will see shortly, these properties will imply our assertions regarding the private code $\mathcal{C}$.

We turn to show that $\mathcal{C}(w,r)$ is privately decodable within error parameter p with a Maximum Likelihood decoder $\mathcal{D}$. Namely, given an element $r \in \{0,1\}^m$ let $\mathcal{C}_r$ be the code $\mathcal{C}$ restricted to r (*i.e.* $\mathcal{C}_r(w) = \mathcal{C}(w,r)$). Now, for a word $c \in \{0,1\}^n$ and an element $r \in \{0,1\}^m$, $\mathcal{D}(c,r)$ is equal to a message $w \in \{0,1\}^k$ for which $\mathcal{C}_r(w)$ is one of the codewords in $\mathcal{C}_r$ nearest to c (ties are broken arbitrarily).

Let $w \in \{0,1\}^{k'}$. Let $\mathcal{S}$ be any channel in ADVC_p. We would like to show that

$$\Pr_{r\in\{0,1\}^m}[\mathcal{D}(\mathcal{C}(w,r) \oplus \mathcal{S}(\mathcal{C}(w,r)), r) = w] \geq 1 - \frac{\ell Y}{X}.$$

The above is equal to $\frac{1}{\beta}$ times

$\sum_{j\in[\beta]} \Pr[\mathcal{D}(\mathcal{C}(w,r) \oplus \mathcal{S}(\mathcal{C}(w,r)), r) = w \mid \mathcal{C}(w,r) = c_{i(w)j}]$

Here we use the fact that $\Pr[r \in A_{i(w)j}] = \frac{1}{\beta}$.

We now bound the conditional probability. Let $i = i(w)$, $j \in [\beta]$, and let $c_{ij} = \mathcal{C}(w,r)$ be the (unique) codeword obtained for all $r \in A_{ij}$. Consider the ball $\mathcal{B}$ of radius pn centered at $c_{ij} \oplus \mathcal{S}(c_{ij})$. Given $r \in A_{ij}$, the corrupted codeword $c_{ij} \oplus \mathcal{S}(c_{ij})$ may be mistakenly decoded only if there exists a codeword $c_{i'j'}$ in $\mathcal{B}$ for which (a) $i' \neq i$ and (b) $r \in A_{i'j'}$. The number of codewords in $\mathcal{B}$ is bounded by ℓ (recall the properties of $\mathcal{C}'$). For each such codeword $c_{i'j'}$ it is the case that $|A_{i'j'} \cap A_{ij}| \leq Y$ (by the properties of $\mathcal{A}$). Thus the number of elements $r \in A_{ij}$ which satisfy condition (b) above is bounded by ℓY. The total size of A_{ij} is X, implying that for at most a fraction of $\frac{\ell Y}{X}$ elements in $r \in A_{ij}$, the corrupted codeword $c_{ij} \oplus \mathcal{S}(c_{ij})$ may be mistakenly decoded, which in turn implies our assertion. □

Theorem 1.8 *Let $R \in (0,1)$ be constant. Let n be sufficiently large, and let $k = Rn$. Let $\mathcal{C}'$ be an $[n,k]$ code which is (ℓ, pn) list decodable. Let $M = \frac{k^2\ell^2}{\delta^2}$. There exists an $[n, k - \frac{1}{2}\log M, \log M]$ private code which is privately decodable with probability $1-\delta$ within error parameter p. Moreover, if $\mathcal{C}'$ is efficiently encodable and list decodable, then the resulting private code is also efficiently encodable and decodable.*

Proof: Assume $m = \log M$ is an even integer. Let $\varepsilon = \frac{\log M}{2k}$. Notice that $\sqrt{M}^{\frac{1}{\varepsilon}} = 2^k$. Thus, by Lemma 2.6 there exists a $\left[\frac{2^k}{\sqrt{M}}, \sqrt{M}, \sqrt{M}, \frac{1}{\varepsilon}\right]$ set system of $\{0,1\}^m$. Let k' satisfy $2^{k'} = \frac{2^k}{\sqrt{M}}$. We have that $k' = k - \frac{1}{2}\log M$. By Lemma 3.1, one can construct an $[n, k', \log M]$ private code $\mathcal{C}$, which is privately decodable with probability $1 - \frac{\ell Y}{X}$ within error parameter p. For our choice of parameters $1 - \frac{\ell Y}{X} \geq 1 - \delta$.

It is left to show that $\mathcal{C}$ is efficient once $\mathcal{C}'$ is. This follows from the construction of $\mathcal{A}$ presented in Lemma 2.6 and by the definition of $\mathcal{C}$ and $\mathcal{D}$ (as presented in Lemma 3.1), details omitted. □

4 Privately Decodable codes that (almost) meet the Shannon bound

We now present the existence of private codes that use only a logarithmic number of private bits, which are privately decodable within (an arbitrarily small additive error of) the Shannon bound. We start by stating (and proving for completeness) a result appearing in [8] on the existence of standard $[n,k]$ codes with very good list decoding properties (results of similar nature which suffice for our purposes also appear in [15]). We then use Theorem 1.8 to obtain our result.

Lemma 4.1 ([8]) *Let $R \in (0,1)$ be constant. Let $\ell \geq 4$ be a fixed integer. Let n be any sufficiently large constant. There exists $[n, Rn]$ codes which are (ℓ, pn) list decodable with $R = 1 - H(p)(1 + 1/\ell)$.*

Proof: Consider a code $\mathcal{C}$ defined by a random sample of $\{0,1\}^n$ of size $2 \cdot 2^{Rn}$ (here we assume for convenience that 2^{Rn} is an integer). We will show that with positive probability, $\mathcal{C}$ contains a subcode $\mathcal{C}'$ of size 2^{Rn} which is (ℓ, pn) list decodable for p which satisfies $R = 1-H(p)(1+1/\ell)$. Consider a ball $\mathcal{B}$ of radius pn in $\{0,1\}^n$. The volume $Vol(pn)$ of such a ball is at most $2^{H(p)n}$. Thus, the probability that a specific subset of $\mathcal{C}$ of size $\ell+1$ happens to fall into $\mathcal{B}$ is at most $2^{-(\ell+1)(1-H(p))n}$. As there are $\binom{2\cdot 2^{Rn}}{\ell+1}$ such subsets and 2^n such balls, the expected amount of subsets of size $\ell+1$ that are included in some ball $\mathcal{B}$ of radius pn is

$$\binom{2\cdot 2^{Rn}}{\ell+1}2^n 2^{-(\ell+1)(1-H(p))n} \leq 2^{Rn}$$

for our setting of parameters. We conclude the existence of a code $\mathcal{C}$ with $2\cdot 2^{Rn}$ codewords with at most 2^{Rn} subsets of size $\ell+1$ contained in a ball of radius pn. Now, removing a single element from $\mathcal{C}$ for each such subset of size $\ell+1$, will result in the desired subcode $\mathcal{C}'$ of size 2^{Rn} which is (ℓ, pn) list decodable. $\square$

The following is now a Corollary of Theorem 1.8 (and was stated as Theorem 1.5 in the Introduction).

Corollary 4.2 *Let $R \in (0,1)$ be constant. Let $p < 1/2$ satisfy $H(p) = 1-R$. Let $\varepsilon > 0$ be a sufficiently small constant. For every n sufficiently large and any $\delta > 2^{-\frac{R\varepsilon^2}{4}n}$ there exists an $\left[n, Rn, O\left(\log k + \log\left(\frac{1}{\delta}\right)\right)\right]$ private code which is privately decodable with probability $1-\delta$ within error parameter $p-\varepsilon$.*

Proof: Let $\varepsilon > 0$ be a sufficiently small constant. Let $\ell = 2/\varepsilon^2$. Let $\delta > 2^{-\frac{R\varepsilon^2}{4}n}$. Let $k = Rn$. Let $k' = k + \log(1/\delta) + \log\ell + 2\log k$. For our definitions, it is not hard to verify that $k' < n$ (*i.e.* we take $\varepsilon^2 << 1-R$). By Lemma 4.1, we conclude the existence of an $[n, k']$ code which is $(\ell, p'n)$ list decodable for p' satisfying $k'/n = 1 - H(p')(1+1/\ell)$. Finally, by Theorem 1.8 there exist an $\left[n, k' - (\log k' + \log\ell + \log(1/\delta)), O\left(\log(\frac{k'\ell}{\delta})\right)\right]$ private code which is privately decodable with probability at least $1-\delta$ within error parameter p'. As n is sufficiently large, we have that $k' - (\log k' + \log\ell + \log(1/\delta)) \geq k$. Implying the existence of an $\left[n, k, O\left(\log k + \log\left(\frac{1}{\delta}\right)\right)\right]$ private code which is privately decodable within distance p' with probability $1-\delta$.

It is left to show that $p' \geq p - \varepsilon$. Recall that p satisfies $H(p) = 1-R$ and notice that p' satisfies $H(p') \geq 1 - k'/n - 1/\ell$ (the latter following from basic computations). Notice (for our choice of ℓ and k') that $H(p') \geq 1 - R - \varepsilon^2 = H(p) - \varepsilon^2$, implying $H(p') \geq H(p-\varepsilon)$ (where the final inequality also follows from basic calculations). We conclude, that $p' \geq p - \varepsilon$. $\square$

5 Lower bounds

We now turn to show that the result stated in Corollary 4.2 is tight. More specifically, we show two types of bounds. The first bound addresses, for any given rate $R = \frac{k}{n}$, the maximum size of p for which there exists a privately decodable $[n,k]$ private code (in the presence of ADVC$_p$). We show that no matter how large the shared random key is, for any rate $R = \frac{k}{n}$ one cannot construct private $[n,k]$ codes which are privately decodable beyond the Shannon bound. Our second bound addresses the number of shared random bits needed in order to construct an $[n,k]$ private code which is privately decodable arbitrarily close to the Shannon bound. In this case, we present a lower bound that matches our $O(\log k + \log(1/\delta))$ upper bound.

Theorem 1.6 *Let $R \in (0,1)$ be constant. Let p satisfy $H(p) = 1-R$. Let $\varepsilon > 0$ and $\delta \in (0,1)$ be constant. Let n be sufficiently large. Let $\mathcal{C}$ be any private $[n, Rn, m]$ code. The private code $\mathcal{C}$ combined with any private decoder $\mathcal{D}$ is* **not** *privately decodable with probability $1-\delta$ within error parameter $p+\varepsilon$.*

Proof: We will show that the existence of such a private code $\mathcal{C}$ and decoder $\mathcal{D}$ will contradict the fundamental theory of Shannon [12] which states

Theorem 5.1 ([12]) *Let $R \in (0,1)$ be constant. Let p satisfy $H(p) = 1-R$. Let $\varepsilon > 0$ and $\delta \in (0,1)$ be constant. Let n be sufficiently large. Let $\mathcal{C}$ be any $[n, Rn]$ code. The code $\mathcal{C}$ combined with any decoder $\mathcal{D}$ is* **not** *decodable on average with probability $1-\delta$ in the presence of a BSC$_{p+\varepsilon}$.*

Let $k = Rn$. Let $\mathcal{C}$ be an $[n,k,m]$ private code which is privately decodable via $\mathcal{D}$ within an error parameter of $p+\varepsilon$ with some constant probability, say $\frac{1}{2}$ (our proof generalizes to other δ). Namely $\forall w \in \{0,1\}^k$ $\forall \mathcal{S} \in$ ADVC$_{p+\varepsilon}$ $\Pr_r[\mathcal{D}(\mathcal{C}(w,r) \oplus \mathcal{S}(\mathcal{C}(w,r)), r) = w] \geq \frac{1}{2}$. Let $\varepsilon' = \frac{1}{2}\varepsilon$ be a positive constant. Let Γ be the random error distribution of the BSC$_{p+\varepsilon'}$. It is not hard to verify that

$$\Pr_{w,r,e\in\Gamma}[\mathcal{D}(\mathcal{C}(w,r) \oplus e, r) = w] \geq \frac{1}{4}.$$

Thus there exists an $r_0 \in \{0,1\}^m$ such that

$$\Pr_{w,e\in\Gamma}[\mathcal{D}(\mathcal{C}(w,r_0) \oplus e), r_0) = w] \geq \frac{1}{4}.$$

Let $\mathcal{C}_0$ be the $[n,k]$ code defined by the restriction of $\mathcal{C}$ to r_0. Namely $\mathcal{C}_0(w) = \mathcal{C}(w, r_0)$. By setting n sufficiently large, this contradicts the bounds presented in Theorem 5.1. $\square$

Theorem 1.7 *Let $R \in (0,1)$ be constant. Let p satisfy $H(p) = 1-R$. Let $\varepsilon > 0$ be a sufficiently small constant. Let $\delta < \frac{1}{8}$. For sufficiently large n, an $[n, Rn, m]$ private code that is privately decodable with probability $1-\delta$ within error parameter greater than $p-\varepsilon$ must have $m = \Omega\left(\log k + \log\left(\frac{1}{\delta}\right)\right)$.*

Proof: Let $k = Rn$. We start by stating the following proposition, which follows for example from the Elias-Bassalygo bound [4, 14]:

Theorem 5.2 ([4, 14]) *Let $R \in (0,1)$ be constant. Let $\varepsilon > 0$ be any constant. There exists n_0 such that for any $n > n_0$, and any $[n, Rn]$ code $\mathcal{C}$, the minimal distance d of $\mathcal{C}$ satisfies $H\left(\frac{1}{2}\left(1 - \sqrt{(1-2d/n)}\right)\right) \le 1 - R + \varepsilon$.*

Proposition 5.3 *Let $R \in (0,1)$ be constant. Let p satisfy $H(p) = 1 - R$. Let $\varepsilon_R, \varepsilon_p$ be sufficiently small positive constants. For sufficiently large n, there do not exist codes (with range $\{0,1\}^n$) which are $(\varepsilon_R, \varepsilon_p)$-perfect with respect to R.*

For the $\Omega\left(\log\left(\frac{1}{\delta}\right)\right)$ bound, let $\mathcal{C}$ be an $[n,k,m]$ private code that is privately decodable with probability $1-\delta$ within error parameter p' greater than $p - \varepsilon$. Assume that $m < \log\left(\frac{1}{\delta}\right)$. For any $r \in \{0,1\}^m$, consider the code $\mathcal{C}_r = \mathcal{C}(\cdot, r)$. By our definition of p' and by Proposition 5.3, in $\mathcal{C}_r$ there exists a ball $\mathcal{B}$ of radius $p'n$ centered at c such that $|\mathcal{B} \cap \mathcal{C}| = \ell > 1$ (as done before, we view codes as both functions and ordered subsets of $\{0,1\}^n$). Let $\{w_1, \dots, w_\ell\}$ be the words in $\{0,1\}^k$ that satisfy $\mathcal{C}(w_i, r) \in \mathcal{B}$ for $i \in [\ell]$. Let $\mathcal{D}$ be any private decoder. On receival of c, there exists a single word $w \in \{w_1, \dots, w_\ell\}$ such that $\mathcal{D}(c, r) = w$. Thus for $w' \in \{w_1, \dots, w_\ell\} \setminus \{w\}$, setting $\mathcal{S}(\mathcal{C}(w', r)) = \mathcal{C}(w', r) \oplus c$, we have that $\Pr[\mathcal{D}(\mathcal{C}(w', r) \oplus \mathcal{S}(\mathcal{C}(w', r)), r) = w'] \le 1 - \frac{1}{2^m}$. This implies (by our assumption) that $2^m \ge \frac{1}{\delta}$.

For the $\Omega(\log k)$ bound, assume (in sake of contradiction) the existence of an $[n,k,m]$ private code $\mathcal{C}$ that is privately decodable with probability $1 - \delta > \frac{7}{8}$ within error parameter p' greater than $p - \varepsilon$ for which $m \le \frac{1}{2}\log k - 1$. We will show that this implies for a specific $\mathcal{S} \in \mathrm{ADVC}_{p'}$ that the average decoding error (over $w \in \{0,1\}^k$ and $r \in \{0,1\}^m$) of any private decoder $\mathcal{D}$ is at least $\frac{1}{8}$. This in turn, implies that $\mathcal{C}$ is not privately decodable with probability greater than $1 - \delta$: a contradiction.

In what follows, we construct $\mathcal{S} \in \mathrm{ADVC}_{p'}$ which yields a high decoding error probability no matter which decoder $\mathcal{D}$ is used. Roughly speaking, this is done by finding many pairs of codewords in $\mathcal{C}$ that satisfy two properties. First of all we require that each pair of codewords is contained in some ball of radius $p'n$ in $\{0,1\}^n$. Secondly, we require that each pair of codewords is contained in $\mathcal{C}_r$ for many values of $r \in \{0,1\}^m$. Now for each such pair (c_1, c_2) let o be the center of the ball containing c_1 and c_2. Defining $\mathcal{S}(c_1) = c_1 \oplus o$ and $\mathcal{S}(c_2) = c_2 \oplus o$ will yield a decoding error on a large fraction of inputs w, r for which $\mathcal{C}(w,r) \in \{c_1, c_2\}$. As many such pairs (c_1, c_2) are found, we conclude a decoding error on a large fraction of possible inputs w, r.

We turn to proving our claim formally. Each distinct codeword $c \in \mathcal{C}$ has a set of pre-images $\{(w,r) \mid w \in \{0,1\}^k,\ r \in \{0,1\}^m,\ \mathcal{C}(w,r) = c\}$. In the following, we will assume for every c that its corresponding set of pre-images does not include a pair (w_1, r_1) and (w_2, r_2) such that $r_1 = r_2$ (we will briefly discuss the general case later in the proof). For each c, let $R_c \subseteq \{0,1\}^m$ be the subset $\{r \mid \exists w\ \mathcal{C}(w,r) = c\}$. Namely, all the pre-images of c consist of pairs (w,r) in which $r \in R_c$. In our proof, we refer to the subset $R_c \subseteq \{0,1\}^m$ that (partially) defines the pre-image of c as the subset corresponding to the codeword c. This correspondence also defines for each $R \subseteq \{0,1\}^m$ a subset of $\mathcal{C}$ consisting of the codewords $c \in \mathcal{C}$ for which R (partially) defines their pre-image (*i.e.* $R = R_c$). Denote these sets as $\mathcal{C}_R$ (notice that the sets $\mathcal{C}_R$ from a partition of $\mathcal{C}$).

We would like to bound the number of elements in $\mathcal{C}$ that are in subsets $\mathcal{C}_R$ which are *small*. Let N be the number of distinct elements in the image of $\mathcal{C}$. As the total number of subsets of $\{0,1\}^m$ is at most $2^{\sqrt{k}/2}$, at least $N - 2^{\sqrt{k}/2} 2^{k - \sqrt{k}}$ of the distinct codewords in $\mathcal{C}$ are in sets $\mathcal{C}_R$ which are of size $\ge 2^{k-\sqrt{k}}$ (these sets will be referred to as large sets). This implies that the total pre-image of the union of the small sets $\mathcal{C}_R$ (those smaller than $2^{k-\sqrt{k}}$) is at most $2^m 2^{k - \sqrt{k}/2} \le \frac{2^m 2^k}{2}$ (notice that this implies that the majority of codewords in $\mathcal{C}$ are in large subsets).

Consider any one of the large sets $\mathcal{C}_R$. By the definition of p', it is not hard to verify that in $\mathcal{C}_R$ there exist at least $\frac{|\mathcal{C}_R|}{4}$ disjoint pairs of codewords (c_1, c_2) which are each contained in a ball of radius $p'n$ (*i.e.* $\Delta(c_1, c_2) \le 2p'n$). This follows from Proposition 5.3 by the fact that $|\mathcal{C}_R| >> 2^{k(1-\varepsilon)}$ for any constant ε (given k is sufficiently large).

Consider the following adversarial strategy $\mathcal{S} \in \mathrm{ADVC}_{p'}$: For any large set $\mathcal{C}_R$, and each pair (c_1, c_2) described above, let o be the center of the ball containing the pair; for $i = 1, 2$ set $\mathcal{S}(c_i) = o \oplus c_i$. For the remaining codewords $c \in \mathcal{C}$, let $\mathcal{S}(c) = \mathbf{0}$.

Now consider any private decoder $\mathcal{D}$. We would like to bound (from below) the probability that $\mathcal{D}$ fails (over both $w \in \{0,1\}^k$ and $r \in \{0,1\}^m$) with $\mathcal{S}$. In every large set $\mathcal{C}_R$, at least $\frac{1}{4}$ of all pairs (w,r) such that $\mathcal{C}(w,r) \in \mathcal{C}_R$ will result in a decoding error. This follows from the pairing (c_1, c_2) discussed above. The total number of pairs (w,r) for which $\mathcal{C}(w,r)$ is in a large $\mathcal{C}_R$ is more than $\frac{1}{2} 2^m 2^k$. This implies that the total decoding failure probability of $\mathcal{D}$ is at least $\frac{1}{8}$, which completes our assertion.

We still need to justify our assumption regarding the pre-image of any codeword $c \in \mathcal{C}$. Consider a codeword c with a pre-image which includes two pairs (w_1, r_1) and (w_2, r_2) in which $r_1 = r_2$. It is not hard to verify that no matter what the adversarial channel $\mathcal{S}$ is and no matter what the private decoder does, at least one of these pre-images will be decoded incorrectly. Thus, this is actually in our favor, and implies (by the assumptions stated in the theorem) that considering only $\frac{7}{8}$ of the pairs $(w, r) \in \{0,1\}^k \times \{0,1\}^m$ will reduce us to the case analyzed above (some minor modifications will still be needed in the proof - details omitted). $\square$

6 Concluding remarks

In this work we have presented the existence of private codes with logarithmic shared randomness that allow communication over adversarial channels which come arbitrarily close to the Shannon bound. Namely, we have shown that it is possible to communicate, with high probability, over an adversarial channel with error parameter p arbitrarily close to the Shannon bound, once both the sender A and the receiver B have access to a logarithmic amount of random bits. We have also shown that our results are tight.

A number of remarks are in place. First of all, the results we present are existential. We present an efficient reduction between list decoding and private codes, however to obtain optimal (in rate and randomness) private codes we rely on the existence of optimal (in rate) list decodable codes (efficient encoding and decoding of the latter is a long standing open problem).

Secondly, assume the sender A wants to send more than a single message w to B. In the framework studied, the joint random bits used for the communication of the first message are assumed to be independent of the joint random bits used in the communication of the second message. This need not be the case. It can be shown, using a different set of parameters in Lemma 2.6, that there exist (optimal) private codes which use only an additional $O(\log(1/\delta))$ fresh random bits for every additional message sent (at a price of increasing the error probability with each message sent).

Finally, one may consider private codes in the cryptographic setting (in which the parties involved are computationally bounded). Roughly speaking, in this case, one may replace the shared random bits needed for the communication of several messages, by a series of pseudo-random bits (generated from a single shared seed).

7 Acknowledgements

I would like to thank Eli Ben-Sasson, Yuval Cassuto, Oded Goldreich, Venkat Guruswami, Kobbi Nissim and Chris Umans for their helpful discussions and comments regarding private codes. Special thanks goes to Shuki Bruck and Alex Sprintson for their interest and involvement in many stages of this work. Finally, I would like to thank the FOCS 2004 anonymous referees for their useful comments.

References

[1] R. Ahlswede. Elimination of correlation in random codes for arbitrary varying channels. *Z. Wahrsch. Verw. Gebiete*, 33:159–175, 1978.

[2] R. Ahlswede and I. Csiszár. Common randomness in information theory and cryptography - part 1: Secret sharing. *IEEE Transactions on Information Theory*, 39(4):1121–1132, 1993.

[3] R. Ahlswede and I. Csiszár. Common randomness in information theory and cryptography - part 2: CR capacity. *IEEE Transactions on Information Theory*, 44(1):225–240, 1998.

[4] L. Bassalygo. New upper bounds for error-correcting codes. *Problems of Information Transmission*, 1(1):32–35, 1965.

[5] D. Blackwell, L. Breiman, and A. Thomasian. The capacities of certain channel classes under random coding. *The Annals of Mathematical Statistics*, 31(3):558–567, 1960.

[6] I. Csiszár and P. Narayan. The capacity of the arbitrarily varying channel revisited: positivity, constraints. *IEEE Transactions on Information Theory*, 34(4):181–193, 1988.

[7] P. Elias. Coding for noisy channels. *IRE Convention Record, pt. 4*, pages 37–47, 1955.

[8] P. Elias. Error-correcting codes for list decoding. *IEEE Transactions on Information Theory*, 37(1):5–12, 1991.

[9] P. Erdös, P. Frankl, and Z. Füredi. Families of finite sets in which no set is covered by the union of r others. *Israel Journal of Mathematics*, 51:79–89, 1985.

[10] V. Guruswami. List decoding with side information. *In Proceedings of 18th IEEE Annual Conference on Computational Complexity*, pages 300–309, 2003.

[11] F. MacWilliams and N. Sloane. The theory of error-correcting codes. *North-Holland, Amsterdam*, 1977.

[12] C. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, 1948.

[13] C. Shannon. A note on partial ordering for communication channels. *Information and Control*, 1:390–397, 1958.

[14] C. Shannon, R. Gallager, and E. Berlekamp. Lower bounds to error probability for coding on discrete memoryless channels. *Information and Control*, 10:65–103 and 522–552, 1967.

[15] V. Zyablov and M. Pinsker. List cascade decoding. *Prob. Information Transmission*, 17(4):236–240, 1992.