Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 69 (2004)

Improving the alphabet-size in expander based code constructions.

Eran Rom, Amnon Ta-Shma

April 7, 2005

Abstract

In this note we revisit the construction of high noise, almost optimal rate list decodable code of Guruswami [Gur04]. Guruswami showed that based on optimal extractors one can build a $(1-\epsilon, O(\frac{1}{\epsilon}))$ list decodable codes of rate $\Omega(\frac{\epsilon}{log\frac{1}{\epsilon}})$ and alphabet size $2^{O(\frac{1}{\epsilon}\cdot log\frac{1}{\epsilon})}$. We show that if one replaces the expander component in the construction with an unbalanced disperser, than one can improve the alphabet size to $2^{O(log^2\frac{1}{\epsilon})}$ while keeping all other parameters the same.

In this revision we point out that the same type of improvement can be done in various other expander based constructions given by Guruswami and Indyk in [GI01]. We give here one example. [GI01] give an explicit encodable/decodable unique decoding with rate ϵ, relative distance $(1-\epsilon)$ and alphabet size $2^{O(\frac{1}{\epsilon})}$. we improve the alphabet size to $2^{2^{polyloglog(\frac{1}{\epsilon})}}$[1]. Unlike the construction in [Gur04], where the expanding graph is used only for its expansion properties, the construction studied here requires both expansion and mixing. Thus, the improvement is achieved by replacing the balanced expander used with an unbalanced extractor.

1 Introduction

List decoding was defined independently by Elias [Eli57] and Wozencraft [Woz58] as a generalization of the unique decoding problem. In unique decoding a codeword is transmitted over a noisy channel such that if not too many errors occur, one can recover the transmitted word from the received word. Unique decoding is possible only when the number of errors is guaranteed to be less than half the minimum distance of the code. In particular, unique decoding is not possible when the error rate is greater than half. In list decoding we give up unique decoding and instead only require that in any Hamming ball of relatively large radius ("large" error rate), there are not too many codewords ("small" list containing all possible transmitted codewords).

More formally, we say that $C : \Sigma^n \rightarrow \Sigma^N$ is (p, L)-list decodable if for every

[1] assuming one can build an explicit optimal extractor the alphabet size can get to $poly(\frac{1}{\epsilon})$

ISSN 1433-8092

$r \in \Sigma^N$, $|\{x \in \Sigma^n | \Delta(C(x), r) \leq pN\}| \leq L$, where $\Delta(x, y)$ denotes the Hamming distance between strings x, y. That is, the number of codewords which agree with r, on at least $(1-p)N$ coordinates is smaller than L.
We focus on the high noise regime, where $p = 1 - \epsilon$, for $\epsilon > 0$ being a very small constant. Simple probabilistic argument shows that $(1-\epsilon, O(\frac{1}{\epsilon}))$-list decodable codes with $rate = \Omega(\epsilon)$, and $|\Sigma| = O(\frac{1}{\epsilon^2})$ exists. Until recently, the best known explicit construction achieved rate of ϵ^2, which has been a "barrier" for the rate of list decoding from $1-\epsilon$ fraction of errors. Recently, Guruswami ([Gur04]) used an expander based construction to give the first explicit $(1-\epsilon, O(\frac{1}{\epsilon}))$-list decodable code having $rate = \Omega(\frac{\epsilon}{\log \frac{1}{\epsilon}})$.

1.1 Our results

[Gur04] uses a strong extractor in his constuction, we skip the definition (the interested reader can look in section 2 for the definition). [Gur04] proves:

Theorem 1. *[Gur04] For every $0 \leq K = K(N) \leq N$ and every $\epsilon > 0$, if a family of $(K(N), \frac{1}{4})$-strong extractors $E : [N] \times [D] \to [M]$ where $M = \Theta(\frac{1}{\epsilon})$ and $D = \log n \cdot plogM$ can be explicitly constructed, then one can construct a family of $(1-\epsilon, K(N))$-list decodable codes of rate $\Omega(\frac{\epsilon}{plog(\frac{1}{\epsilon})})$ and alphabet of $2^{O(\epsilon^{-1}\log(\frac{1}{\epsilon}))}$.*

We show that there is an explicit construction with a better alphabet size, while all other parameters match:

Theorem 2. *For every $0 \leq K = K(N) \leq N$ and every $\epsilon > 0$ if a family of $(K(N), \frac{1}{4})$-strong extractors $E : [N] \times [D] \to [M]$ where $M = \Theta(\frac{1}{\epsilon})$ and $D = \log n \cdot plogM$ can be explicitly constructed, then one can construct a family of $(1-\epsilon, K(N))$-list decodable codes of block length N, rate $= \Theta(\frac{\epsilon}{\log^{O(1)} \frac{1}{\epsilon}})$, and alphabet size of $2^{2^{ploglog(\frac{1}{\epsilon})}}$*

For strong extractor constructions with optimal entropy loss $K(N) = K = \Theta(M)$, and near optimal degree $D = O(\log N)$, both the construction of [Gur04], and ours achieve $(1-\epsilon, O(\frac{1}{\epsilon}))$-list decodable code. The alphabet size, however, in [Gur04] is $2^{O(\epsilon^{-1}\log(\frac{1}{\epsilon}))}$ while we acheive alphabet size of $2^{O(\log^2(\frac{1}{\epsilon}))}$.

1.2 The technique

1.2.1 Extractor Codes

One underlying component is an extractor code [TSZ01]. An extractor takes as input a sample drawn from a weak random source, and using a short seed of truly random bits, outputs an almost uniform string. In a strong extractor, the output is almost uniform even if the truly random bits are made public. An extractor code is obtained by viewing the input as an information message, and taking the encoded message to be the extractor's output as the seed varies over all possible values.

Extractor codes can list decode from a large error rate. Specifically, if the extractor error is ϵ, and its output length is M, then the extractor code can decode from $1-(\epsilon+\frac{1}{M})$ fraction of errors. The drawback of extractor codes lies in the $\Omega(\log\frac{1}{\epsilon^2})$ lower bound on its degree [RTS00]. This lower bound translates to an $O(\epsilon^2)$ upper bound on the code's rate. On the other hand, as observed in [TSZ01], extractor codes have a property stronger than list decoding, also known as list recovering [GI02]. Roughly speaking, list recovering deals with the situation where the receiver only knows that the i^{th} symbol received, belongs to some set S_i, whose size is a non negligible fraction of the alphabet size. Specifically, if ϵ is the error of the extractor and the sets S_i are of size $\alpha|\Sigma|$, then the error rate from which recovery is possible is $\epsilon+\alpha$.

1.2.2 Amplification using expanders

The technique of code amplification through expanders was introduced in [ABN+92], where it is used to amplify Justesen code. Justesen code rate vanishes as the error rate grows. [ABN+92] take a Justesen code of constant error rate and amplify it using an expander to get a code with large distance and constant rate over a large alphabet. The optimality of the expander is important as the rate achieved using amplification is proportional to $\frac{1}{D}$, where D is the degree of the expander.

In [Gur04] it is shown how to use the list recovering property of extractor codes to bypass the degree lower bound "rate barrier" of $O(\epsilon^2)$ using the expander amplification technique. The idea is to use a strong extractor with a constant error, thus bypassing the "rate barrier" on the expense of worse error rate. Now, the expander amplification technique can be used to improve the error rate on the expense of the alphabet size.

Looking back, the amplification in [ABN+92] can be done using any disperser (and good expander is just a special case), as the expansion property needed for the amplification is to expand "large" sets (representing the agreement larger than pN), rather then expanding small "sets".

What we do is replace the expander component in [Gur04] with a good disperser. Studying the problem we discover that what is needed is a disperser for the high min–entropy rate, that has optimal entropy loss, and a surprisingly small degree. Fortunately, an explicit construction of such a graph was given recently [RVW00]. Using such a graph shows that our improvement over the construction in [Gur04] can be made explicit. For every code built upon Guruswami's scheme the expander component can be replaced with the explicit disperser and improve the alphabet size. As the disperser is explicit, the decoding scheme mentioned in [Gur04] and the time it takes does not change.

2 preliminaries

We first describe the components used in the construction.

2.1 Extractors and Dispersers

An extractor is a function which extracts the randomness of a defective random source using truly random bits as a catalyst. In a strong extractor, the same holds even when the catalyst is made public. Formally,

Definition 1. *$f : [N] \times [D] \to [M]$ is a (K, ζ_{ext})-strong extractor if for every $X \subseteq [N]$, $|X| \geq K$, the distribution $Y \circ f(X, Y)$ is ζ_{ext}-close to the uniform distribution over $[D] \times [M]$, where Y is taken uniformly at random from $[D]$. The entropy loss of the strong extractor is $\frac{K}{M}$.*

Ta-Shma and Radhakrishnan [RTS00] show that a (K, ζ_{ext})-strong extractor $f : [N] \times [D] \to [M]$ must have degree $D = \Omega(\frac{1}{\zeta_{ext}^2} \log \frac{N}{K})$, and entropy loss $\frac{K}{M} = O(\frac{1}{\zeta_{ext}^2})$. Also shown in [RTS00] are matching implicit upper bounds.

An important property of extractors is mixing (see, [ASE92], Chap 9). For that we now introduce some notation. For $x \in [N]$ we define $\Gamma_f(x)$ to be the ordered neighbors of x. Formally,

$$\Gamma_f(x) \quad = \quad \{(i, f(x, i) \mid i \in [D]\}$$

The mixing property says that:

Fact 1. *If $f : [N] \times [D] \to [M]$ is a (K, ζ_{ext}) strong extractor, then for every $S \subseteq [D] \times [M]$, there are at most K elements $x \in [N]$ satisfying*

$$\frac{|\Gamma_f(x) \cap S|}{D} - \frac{|S|}{D \cdot M} \quad \geq \quad \zeta_{ext}$$

A disperser is the one-sided variant of an extractor. Instead of requiring that the output is ϵ-close to the uniform distribution, we require that the disperser's output covers at least a $1 - \epsilon$ fraction of the target set.

Definition 2. *$g : [L] \times [T] \to [D]$ is a (H, ζ_{disp})-disperser if for every $X \subseteq [L]$ with $|X| \geq H$ we have $|\Gamma_g(X)| \geq (1 - \zeta_{disp})D$. The entropy loss of the disperser is $\frac{HT}{D}$.*

[RTS00] show matching lower bound and non-explicit upper bound for dispersers. Specifically, a (H, ζ_{disp})-disperser $g : [L] \times [T] \to [D]$ must have $T = \Omega(\frac{1}{\zeta_{disp}} \log \frac{L}{H})$, and entropy loss $\frac{HT}{D} = \Omega(log \frac{1}{\zeta_{disp}})$.

We again define the neighbor set $\Gamma_g(\ell)$ of ℓ, except that the dispersers we work with are not strong, and so the set is not ordered . For $\ell \in [L]$ we define

$$\Gamma_g(\ell) \quad = \quad \{g(\ell, j) \mid \ j \in [T]\}$$

For a subset $H \subseteq [L]$ we define $\Gamma_g(H) = \bigcup_{\ell \in H} \Gamma_g(\ell)$.

2.2 List Decodable Codes

Definition 3. *A Code $C : \Sigma_1^n \to \Sigma_2^N$ is (ϵ, K)-List Decodable, if for every $r \in \Sigma_2^N$, $|\{x \in \Sigma_1^n | \Delta(r, C(x) \le \epsilon N\}| \le K$*

Intuitively, this means that even if an $(1-\epsilon)$ fraction of the symbols in a codeword are noisy, the size of the decoding set is upper bounded by K.

The rate of the code, which captures the amount of redundancy added is $\frac{n \log |\Sigma_1|}{N \log |\Sigma_2|}$. The goal in constructing a list decodable code in the "high noise" situation, where $\epsilon > 0$ is treated as a very small constant, is to minimize K and the alphabet size, while maintaining a high rate. It is known that $(1-\epsilon, O(\frac{1}{\epsilon}))$ codes of rate $\Omega(\epsilon)$, and $|\Sigma| = O(\frac{1}{\epsilon^2})$ exists. Also the rate must be $\Omega(\epsilon)$, and $|\Sigma| = \Omega(\frac{1}{\epsilon})$.

3 The construction

The construction is basically Guruswami's construction, except that Guruswami uses a balanced, good expander and we use a slightly unbalanced good disperser.

- $f : [N] \times [D] \to [M]$ be a (K, ζ_{ext})-strong extractor, and,
- $g : [L] \times [T] \to [D]$ be a (H, ζ_{disp})-disperser.

We define the code $C_{f,g} : [N] \to [M^T]^L$ as follows:

1. Given $x \in [N]$, denote by $\overline{y} = (y_1, \ldots, y_D) \in [M]^D$ where $y_i = f(x, i)$.
2. Put the symbols $(y_1, \ldots, y_D) \in [M]^D$ along g's range $[D]$. Each element $\ell \in [L]$ has T neighbors in $[D]$. Collect from each neighbor the symbol that was put along it. I.e., for each $\ell \in [L]$ define $\overline{z}_\ell = (z_{\ell,1}, \ldots, z_{\ell,T}) \in [M]^T$, where $z_{\ell,t} = y_{g(\ell,t)}$.
3. The encoding of x is defined to be

$$C_{f,g}(x) \quad = \quad (\overline{z}_1, \ldots, \overline{z}_L)$$

See Figure 1 for a figure illustrating the construction. We claim:

Lemma 1. *If the extractor f and the disperser g are as above, and if $M \cdot D \ge \frac{L \cdot T}{1 - \zeta_{ext} - \zeta_{disp}}$, then $C_{f,g}$ is $(1 - \frac{H}{L}, K)$-list decodable code.*

Proof. Let $z = (\overline{z}_1, \ldots, \overline{z}_L) \in [M^T]^L$ be an arbitrary word in $[M^T]^L$. From z we build a set S as follows. For each $1 \le \ell \le L$, we look at $\overline{z}_\ell = z_{\ell,1}, \ldots, z_{\ell,T}$ and we build the set $S_\ell \subseteq [D] \times [M]$ by

$$S_\ell \quad = \quad \{(g(\ell, t), z_{\ell,t}) \mid 1 \le t \le T\}$$

I.e., we build a subset of what $\overline{z}_\ell$ thinks $y_1, \ldots, y_D$ are in locations $g(\ell, 1), \ldots, g(\ell, T)$. We define the set S of $z = (\overline{z}_1, \ldots, \overline{z}_L)$ to be $\bigcup_{\ell=1}^{L} S_\ell$.

Suppose a codeword $C_{f,g}(x) \in [M^T]^L$ agrees with z on at least H coordinates. Now, if $C_{f,g}(x)$ agrees with z on the ℓ'th coordinate, then $(i, f(x,i)) \in S_\ell$ for every $i \in \Gamma_g(\ell)$. As g is a (H, ζ_{disp})-disperser, the set of neighbors of H has at least $(1-\zeta_{disp})D$ elements. Hence, $|\Gamma_f(x) \cap S| \geq (1-\zeta_{disp})D$.

Noting that $|S| \leq L \cdot T$, and using the assumption $M \cdot D \geq \frac{L \cdot T}{1-\zeta_{ext}-\zeta_{disp}}$, we see that $\frac{|S|}{MD} \leq 1 - \zeta_{ext} - \zeta_{disp}$ and together

$$\frac{|\Gamma_f(x) \cap S|}{D} - \frac{|S|}{MD} \quad \geq \quad \zeta_{ext}$$

By Fact 1 we conclude that there are at most K such x's, hence the number of close codewords is at most K. □

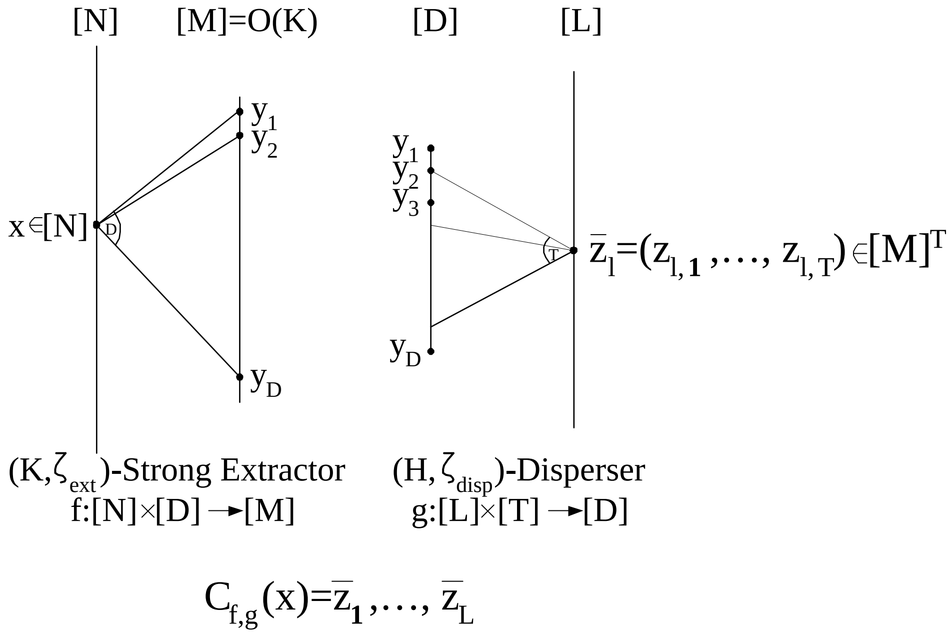

Figure 1: *The neighbors of $x \in [N]$: $(y_1, \ldots, y_D)$, are "put" along the disperser's output $[D]$, defining for each $z_l \in [L]$, a T ordered vector of its neighbors: $\overline{z}_l = (z_{l,1}, \ldots, z_{l,T})$. The vector $\overline{z}_l$ is the l^{th} symbol in the codeword $C_{f,g}(x)$.*

4 Analyzing the parameters

4.1 The constraints

First, we write down all the constraints. The bounds we give are both lower bounds, and achievable by non-explicit constructions. We have:

$$
\begin{aligned}
D &= \Omega(\frac{1}{\zeta_{ext}^2} \cdot \log \frac{N}{K}) && (1) \\
M &= O(K\zeta_{ext}^2) && (2) \\
T &= \Omega(\frac{1}{\zeta_{disp}} \cdot \log \frac{L}{H}) && (3) \\
D &= O(\frac{HT}{\log \frac{1}{\zeta_{disp}}}) && (4) \\
M \cdot D &\geq \frac{L \cdot T}{1 - \zeta_{ext} - \zeta_{disp}} && (5)
\end{aligned}
$$

where the first two equations are the degree and entropy loss of the extractor, the third and fourth are the degree and entropy loss of the disperser, and the fifth is the construction bound that guarantees that the set S is small in $[D] \times [M]$.

4.2 A specific choice of parameters

We now choose parameters. We first set $\zeta_{ext}, \zeta_{disp}$ to be small constants, say we set both to be $\frac{1}{4}$. In order to get a $(1-\epsilon, O(\frac{1}{\epsilon}))$ we set K to be $K = O(\frac{1}{\epsilon})$. With these choices we have $D = \Theta(\log(N))$, and $M = \Theta(K) = \Theta(\frac{1}{\epsilon})$. We also set $\frac{H}{L} = \epsilon$. This implies that $T = \Theta(\log(\frac{L}{H})) = \Theta(\log \frac{1}{\epsilon})$. To satisfy Equation (4) we need to take $H = \Theta(\frac{D}{T}) = \Theta(\frac{\log(N)}{\log \frac{1}{\epsilon}})$ which implies that $L = \frac{H}{\epsilon} = \Theta(\frac{\log(N)}{\epsilon \cdot \log(\frac{1}{\epsilon})})$. Finally, we check Equation (5). We see that $M \cdot D = \Theta(\frac{\log(N)}{\epsilon})$ and $L \cdot T = \Theta(\frac{\log(N)}{\epsilon})$, so with the proper choice of constants the equation holds.

We let $N = 2^n$ and $\epsilon > 0$ be our basic parameters. We summarize all other parameters as functions in n and ϵ. We have,

$$
\begin{aligned}
K &= \Theta(\frac{1}{\epsilon}) && (6) \\
D &= \Theta(n) && (7) \\
M &= \Theta(\frac{1}{\epsilon}) && (8) \\
L &= \Theta(\frac{n}{\epsilon \cdot \log(\frac{1}{\epsilon})}) && (9) \\
H &= \Theta(\frac{n}{\log(\frac{1}{\epsilon})}) && (10) \\
T &= \Theta(\log \frac{1}{\epsilon}) && (11)
\end{aligned}
$$

The rate of the code is given by

$$rate \quad = \quad \frac{\log N}{L \cdot T \log M} \quad = \Theta(\frac{n}{\frac{n}{\epsilon \cdot \log(\frac{1}{\epsilon})} \cdot \log(\frac{1}{\epsilon}) \cdot \log(\frac{1}{\epsilon})}) \; = \; \Theta(\frac{\epsilon}{\log(\frac{1}{\epsilon})})$$

The size of the alphabet is $|\Sigma| = M^T = (\frac{1}{\epsilon})^{O(\log(\frac{1}{\epsilon}))}$. This proves:

Corollary 1. *For every fixed positive integer K, and arbitrary $\epsilon > 0$*

1. *If a family of $(K, \frac{1}{4})$-strong extractors $f : [N] \times [D] \to [M]$, with degree $D = O(\log N)$, and with optimal entropy loss can be explicitly constructed, and,*

2. *a $(\epsilon L, \frac{1}{4})$ disperser $g : [L] \times [T] \to [D]$ with degree $T = \Omega(\log \frac{L}{H})$ and optimal entropy loss can be explicitly constructed*

then we can explicitly construct $(1 - \epsilon, O(\frac{1}{\epsilon}))$-list decodable code of rate $\Omega(\frac{\epsilon}{\log \frac{1}{\epsilon}})$ over an alphabet size of $2^{O(\log^2(\frac{1}{\epsilon}))}$.

4.3 On the optimality of the parameters choice

The parameters chosen in section 4.2 give good rate and alphabet size, but are sub optimal with respect to the non explicit construction. We now show that in the suggested construction this choice of parameters is optimal. In all cases below we consider the high noise regime of $1 - \epsilon$ fraction of errors, namely, $H = \epsilon L$.

4.3.1 $(1-\epsilon, O(\frac{1}{\epsilon}))$ list-decoding implies sub optimal rate and alphabet size

Claim 1. *In the construction given in section 3, for any choice of parameters satisfying $(1 - \epsilon, \frac{1}{\epsilon})$-list decoding with rate bounded away from zero, the resulting rate and alphabet size cannot be better (up to constant factor) than those in section 4.2.*

Proof. We first show that M must be $\Theta(\frac{1}{\epsilon})$:

- Decoding list of size $\frac{1}{\epsilon}$ implies that $K = \frac{1}{\epsilon}$, and by constraint (2) $M = O(K) = O(\frac{1}{\epsilon})$
- By constraint (5) $M \geq \frac{L \cdot T}{D}$. $L = \frac{H}{\epsilon}$, and constraint (4) implies $T = \Omega(\frac{D}{H})$. Altogether, $M \geq \frac{L \cdot T}{D} = \Omega(\frac{1}{\epsilon})$

By the construction, $rate = \frac{\log N}{L \cdot T \log M}$, constraint (5) implies $L \cdot T \leq M \cdot D$, and so $rate \geq \frac{\log N}{M \cdot D \log M} = \Theta(\frac{\epsilon \log N}{D \log \frac{1}{\epsilon}})$. Thus, in order to bound the rate away from zero, we must take $D = O(\log N)$, which gives $rate = \Omega(\frac{\epsilon}{\log \frac{1}{\epsilon}})$. As for the alphabet size $|\Sigma| = M^T$. $M = \Omega(\frac{1}{\epsilon})$, and by constraint (3) $T = \Omega(\log \frac{1}{\epsilon})$. Altogether, $|\Sigma| = (\frac{1}{\epsilon})^{\Omega(\log \frac{1}{\epsilon})}$ □

4.3.2 Rate close to ϵ implies an almost optimal extractor degree

Regardless of the decoding list size and the alphabet size, having rate close to ϵ, requires an almost optimal extractor degree.

Claim 2. *In the construction given in section 3, for any choice of parameters satisfying list decoding from $1-\epsilon$ error fraction with rate $= \frac{\epsilon}{\log \frac{1}{\epsilon}}$, it must be that $D = O(\log N)$.*

Proof. We show that $\frac{\log N}{D} = \Omega(1)$

- The construction constraint suggests that $L \cdot T \leq M \cdot D$, and so $rate = \frac{\log N}{L \cdot T \log M} \geq \frac{\log N}{M \cdot D \log M}$. Thus, $rate = \frac{\epsilon}{\log \frac{1}{\epsilon}}$ implies $\frac{\log N}{D} \leq M \log M \frac{\epsilon}{\log \frac{1}{\epsilon}}$
- $rate = \frac{\log N}{L \cdot T \log M} = \frac{\epsilon \log N}{H \cdot T \log M}$, as $L = \frac{1}{\epsilon} H$. Assuming $rate = \frac{\epsilon}{\log \frac{1}{\epsilon}}$, and noting that by constraint (4) $H \cdot T \geq \Theta(D)$, we have $\frac{\log N}{D} \geq \Theta(\frac{\log M}{\log \frac{1}{\epsilon}})$

Altogether we have $\Theta(\frac{\log M}{\log \frac{1}{\epsilon}}) \leq \frac{\log N}{D} \leq M \log M \frac{\epsilon}{\log \frac{1}{\epsilon}}$, and so $M \geq \Theta(\frac{1}{\epsilon})$ and $\frac{\log N}{D} = \Omega(1)$ □

We mention that [TSZS01] show an explicit strong extractor with a degree very close to $O(\log N)$. However, the entropy loss of the extractor is high. We quote its exact parameters in the next section.

4.3.3 Rate close to ϵ, and alphabet size of $(\frac{1}{\epsilon})^{\log \frac{1}{\epsilon}}$ implies optimal entropy loss disperser

Regardless of the list size, if we wish to decode from $1-\epsilon$ fraction of errors and achieve the rate and alphabet size as in section 4.2, then we must take an optimal entropy loss disperser.

Claim 3. *In the construction given in section 3, for any choice of parameters satisfying list decoding from $1-\epsilon$ error fraction with rate $= \frac{\epsilon}{\log \frac{1}{\epsilon}}$, and $|\Sigma| = (\frac{1}{\epsilon})^{\log \frac{1}{\epsilon}}$ it must be that $\frac{H \cdot T}{D} = O(1)$.*

Proof. $rate = \frac{\log N}{L \cdot T \log M} = \frac{\epsilon \log N}{H \cdot T \log M}$, as $L = \frac{1}{\epsilon} H$. By the construction constraint, $L \cdot T \leq M \cdot D$, and so $\frac{\epsilon \log N}{H \cdot T \log M} \geq \frac{\log N}{M \cdot D \log M}$, giving $\frac{H \cdot T}{D} \leq \epsilon M$. Now, by constraint (3) $T \geq \Theta(\frac{1}{\epsilon})$, and $|\Sigma| = M^T = (\frac{1}{\epsilon})^{\log \frac{1}{\epsilon}}$ implies $M \leq \frac{1}{\epsilon}$. Altogether, $\frac{H \cdot T}{D} = O(1)$ □

4.3.4 $(1-\epsilon, O(\frac{1}{\epsilon}))$ list-decoding implies disperser and extractor optimal entropy loss

Claim 4. *In the construction given in section 3, for any choice of parameters satisfying $(1-\epsilon, \frac{1}{\epsilon})$-list decoding the disperser and extractor must have optimal entropy loss (namely, $\frac{HT}{D} = O(1)$ and $\frac{K}{M} = O(1)$).*

Proof. The list size implies $K = \frac{1}{\epsilon}$, the use of strong extractor implies $M = O(K) = O(\frac{1}{\epsilon})$, and the error fraction implies $H = \epsilon L$. By constraint (5) $D \geq \frac{L \cdot T}{M}$. Altogether, $D \geq \Omega(H \cdot T)$ or $\frac{HT}{D} = O(1)$. Interpreting constraint (5) otherwise, $M \geq \frac{L \cdot T}{D}$, and by the disperser's lower bound $D \leq \Theta(H \cdot T)$. Altogether, we have $\Theta(\frac{1}{\epsilon}) \leq M \leq K$ (as $\frac{L}{H} = \frac{1}{\epsilon}$). Now in order to get a list size of $\frac{1}{\epsilon}$, it must be that $M = \Theta(K)$. ☐

We mention that [RVW00] give an explicit disperser with high min–entropy and optimal entropy loss, as imposed by the above claim. We quote the exact parameters of the explicit disperser in the next section.

5 Explicit Constructions

As before, we set the extractor and disperser errors to be constants, say $\zeta_{ext} = \zeta_{disp} = \frac{1}{4}$. We note that the construction constraint now becomes, $M \cdot D \geq 2 \cdot L \cdot T$

5.1 Using explicit high min–entropy optimal loss disperser

As mentioned in 4.3.4, the high noise regime and the good rate we wish, imply the need of a high min–entropy disperser with optimal entropy loss. [RVW00] give an explicit construction of a disperser for high min–entropies with $O(1)$ entropy loss. Specifically:

Fact 2 ([RVW00]). *For every L and $\epsilon > \frac{1}{\sqrt{L}}$, and for every $\zeta_{disp} > 0$, there exists an explicit construction of $(\epsilon L, \zeta_{disp})$-disperser $g : [L] \times [T] \to [D]$, with $T = 2^{ploglog(\frac{1}{\epsilon})}$, and entropy loss $\frac{\epsilon L \cdot T}{D} = O(1)$.*

We now plug in the above disperser in the construction:

Corollary 2. *If for $\epsilon > 0$, and for $K = K(N) < N$ a family of $(K, \frac{1}{4})$-strong extractors $f : [N] \times [D] \to \Theta(\frac{1}{\epsilon})$ can be constructed, then for every $\epsilon > 0$ one can construct a family of $(1 - \epsilon, K)$-list decodable codes of block length N, with rate $= \Omega(\frac{\epsilon \cdot \log N}{D \cdot \log \frac{1}{\epsilon}})$, and $|\Sigma| = 2^{2^{ploglog(\frac{1}{\epsilon})}}$*

Proof. Let $\epsilon > 0$, $M = \Theta(\frac{1}{\epsilon})$, and $T = 2^{ploglog(\frac{1}{\epsilon})}$. We now choose N so that $\epsilon^2 \cdot T \cdot D \geq \sqrt{D}$. By the assumption, if $K(N) < N$, then there is a $(K, \frac{1}{4})$-strong extractor $f : [N] \times [D] \to \Theta(\frac{1}{\epsilon})$, with degree $D(N)$. By the choice of N, there is a $(\epsilon L, \frac{1}{4})$-disperser $g : [L] \times [T] \to [D]$, with $\frac{\epsilon L \cdot T}{D} = O(1)$, and so choosing a $\Theta(\cdot)$ constant small enough for M, we have $M \cdot D \geq 2 \cdot L \cdot T$. Thus we can apply lemma 1 to get a $(1 - \epsilon, K)$-list decodable code, of block length N, with the above rate and alphabet size. ☐

Assuming one can construct a family of optimal entropy loss strong extractors with near optimal degree (namely, $K = O(M) = O(\frac{1}{\epsilon})$, and $D = O(\log N)$ in the above), then one can get $(1 - \epsilon, O(\frac{1}{\epsilon}))$-list decodable code with *rate* $= \Omega(\frac{\epsilon}{\log \frac{1}{\epsilon}})$, and the above alphabet size.

5.2 Using the above disperser with an explicit extractor

As mentioned by Guruswami, and in section 4.3.2 we need an extractor with degree $D = O(\log N)$. The best explicit construction to date of a strong extractor, which achieves the required degree is due to [TSZS01].

Fact 3 ([TSZS01]). *For Every $m = m(n)$, $k = k(n)$ and $\zeta = \zeta(n)$ such that $3m\sqrt{n \log(n/\zeta)} \leq k \leq n$, there is an explicit family of (k, ζ) strong extractors $E_n : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $d = \log n + O(\log \frac{m}{\zeta})$.*

Denoting $N = 2^n$, $K = 2^k$, $D = 2^d$, and $M = 2^m$ Plugging the above extractor in the construction, we get:

Corollary 3. *For every $\epsilon > 0$, there is an explicit constructible family of $(1 - \epsilon, (\frac{1}{\epsilon})^{\Theta(\sqrt{n \log n})})$-list decodable codes of block length N, rate $= \Theta(\frac{\epsilon}{\log^{O(1)} \frac{1}{\epsilon}})$, and $|\Sigma| = 2^{2^{ploglog(\frac{1}{\epsilon})}}$*

Proof. Apply corollary 2 with $K = M^{\Theta(\sqrt{n \log n})} = (\frac{1}{\epsilon})^{\Theta(\sqrt{n \log n})}$, and $D = \log N \cdot \log^{O(1)} M = \Theta(\log N \cdot \log^{O(1)}(\frac{1}{\epsilon}))$. □

This proves Theorem 2.

6 Unique Decoding with Rate $\Omega(\epsilon)$, and Relative Distance $(1 - \epsilon)$

The example we study here as well as all constructions from [GI01] share some basic form described below.

6.1 The Abstract Expander Based Construction

The expander based constructions given in [GI01] have the following general form. Take some (N, rN) code C over some alphabet Σ and a bipartite balanced expanding graph $G = (A, B, E)$, with $|A| = |B| = N$. Put the symbols of a codeword $C(x)$ along one side of G, and redistribute the symbols along the edges of G. Collect the symbols on the other side of G to get the resulting codeword. See figure 2. [ABN$^+$92] gave a similar construction to amplify the error of Justesen code. The property needed from the expanding graph in [ABN$^+$92] was that of a disperser, namely, every large enough subset of vertices in one side misses only a small fraction of vertices on the other side. The properties needed from the expanding graph in [GI01] are both expansion and mixing properties. We point out the fact that the graph used in [GI01] is balanced. We show that by replacing it with an unbalanced graph, the alphabet size can be dramatically improved.

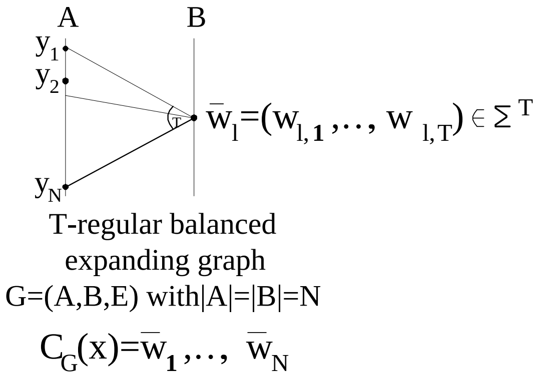

Figure 2: *$y_1, \ldots, y_N \in \Sigma^N$ are the coordinates of some codeword $C(x)$. The coordinates of $C(x)$ are "put" along the graph's left side A, defining for each $l \in B$ an ordered vector of its neighbors $\overline{w}_l = (w_{l,1}, \ldots, w_{l,T}) \in \Sigma^T$. The vector $\overline{w}_l$ is the l^{th} symbol in the overall codeword $C_G(x)$.*

6.2 The Actual Components Used

For the specific example we bring here of unique decoding with rate $\Omega(\epsilon)$, relative distance $(1-\epsilon)$, and alphabet size $2^{O(\frac{1}{\epsilon})}$ with encoding time $N \log^{O(1)} N$ and almost linear decoding time [GI01] use the above scheme with the following code C and expanding graph G:

1. ([GI01] Claim 1) For every $\alpha > 0$ there exists a prime power $q = q_\alpha$ of order $O(\frac{1}{\alpha^2})$, which may be assumed to be a power of 2, such that for all $\beta > 0$, the following holds. there is an explicitly specified code family with constant rate $r_{\alpha,\beta} > 0$ over an alphabet of size q with the property that a code of block length N in the family can be list decoded from up to $(1-\alpha)$ fraction of errors in $O(N^{1+\beta})$ time, and can be encoded in $O(N \log^{O(1)} N)$

2. A T-regular balanced expander graph $G = (A, B, E)$, with $|A| = |B| = N$, satisfying for every $\epsilon < \delta' < \delta$:

   - $\forall Y \subset A$, $|Y| = \frac{1}{2}N$ we have $|\Gamma_G(Y)| \geq (1-\epsilon)N$
   - $\forall Y \subset A$, $|Y| = \frac{1}{2}N$, $\forall X \subset B$, $|X| = \delta N$, we have $\frac{|\Gamma_G(Y) \cap X|}{|Y|T} > \delta'$

   Using Ramanujan graph, a degree $T = O(\frac{1}{\epsilon} + \frac{1}{\delta})$ is enough to achieve the above two properties.

Plugging the above building blocks in the construction above we get a code with rate:

$$r_{\alpha,\beta}\frac{1}{T} \tag{12}$$

and with alphabet size:

$$O((\frac{1}{\alpha^2})^T) \tag{13}$$

Using these two building blocks [GI01] prove the following:

Theorem 3 ([GI01] Theorem 8). *For any $\beta, \delta > 0$, there is a constant $B > 1$ such that for all small enough $\epsilon > 0$ there is an explicitly specified code family with the properties:*

1. *It has rate $(\frac{\epsilon}{B})$, relative distance at least $(1-\epsilon)$ and alphabet size $2^{O(\frac{1}{\epsilon})}$.*
2. *A code of block length N in the family can be list decoded in $O(N^{1+\beta})$ time from up to a $(1-\delta)$ errors, and can be encoded in $O(N \log^O(1)N)$ time.*

Given $\delta > 0$, the parameters in the theorem are achieved by taking $\delta' = (\frac{\delta}{4})$, $\alpha = (\frac{\delta}{8})$ and any $\epsilon < \delta'$. Thus, the required degree T is $O(\frac{1}{\epsilon})$. Noting that δ, β are independent of ϵ (12), (13) imply rate $\Omega(\epsilon)$, and alphabet size $2^{O(\frac{1}{\epsilon})}$

6.3 Replacing the Balanced Expander with Unbalanced Extractor

We now claim that the balanced expanding graph from above can be replaced with an unbalanced extractor $g : [L] \times [T] \to [N]$ with regular right degree $Q = \frac{LT}{N}$ with $[N]$ replacing A, and $[L]$ replacing B, see figure 3.

Lemma 2. *For every $\delta > 0$, and $\epsilon < \delta$ if $g : [L] \times [T] \to [N]$ is an $(\epsilon L, \frac{1}{4})$-extractor with regular right degree $Q = \frac{LT}{N}$ we have:*

- $\forall Y \subset [N]$, $|Y| = \frac{1}{2}N$ *we have* $|\Gamma_G(Y)| \geq (1-\epsilon)L$
- $\forall Y \subset [N]$, $|Y| = \frac{1}{2}N$, $\forall X \subset [L]$, $|X| = \delta L$, *we have* $\frac{|\Gamma_G(Y) \cap X|}{|Y|Q} > \delta'$

Proof. For the expansion property we note that any extractor is also a disperser. Thus, for every $X \subseteq [L]$, $|X| = \epsilon L$, we have $|\Gamma_g(X)| > \frac{3}{4}N$. This means that for every $Y \subset [N]$, $|Y| = \frac{1}{2}N$, there are at most ϵL vertices in $[L]$ which do not have any neighbor in Y (Otherwise, we have a set of ϵL vertices in L with neighbor set of size $\frac{3}{4}N$ which have no neighbor in Y of size $\frac{1}{2}N$). Thus, for every $Y \subset [N]$, $|Y| = \frac{1}{2}N$ there are at least $(1-\epsilon)L$ vertices with at least one neighbor in Y. In other words $|\Gamma_g(Y)| > (1-\epsilon)L$.

For the second property we have by the mixing property of extractors that $\forall Y \subseteq [N]$, $\forall X \subseteq [L]$, $|X| \geq \epsilon L$ it holds that:

$$\left| \frac{|\Gamma_g(X) \cap Y|}{T|X|} - \frac{|Y|}{N} \right| < \frac{1}{4} \tag{14}$$

$$\frac{Q|Y|}{T|X|} \cdot \left| \frac{|\Gamma_g(X) \cap Y|}{Q|Y|} - \frac{|X||Y|T}{Q|Y|N} \right| < \frac{1}{4} \tag{15}$$

$$\left| \frac{|\Gamma_g(X) \cap Y|}{Q|Y|} - \frac{|X|T}{QN} \right| < \frac{T|X|}{Q|Y|} \cdot \frac{1}{4} \tag{16}$$

$$\left| \frac{|\Gamma_g(X) \cap Y|}{Q|Y|} - \frac{|X|}{L} \right| < \frac{N|X|}{|Y|L} \cdot \frac{1}{4} \tag{17}$$

$$\frac{|\Gamma_g(Y) \cap X|}{Q|Y|} > \left(1 - \frac{N}{|Y|}\frac{1}{4}\right)\frac{|X|}{L} \tag{18}$$

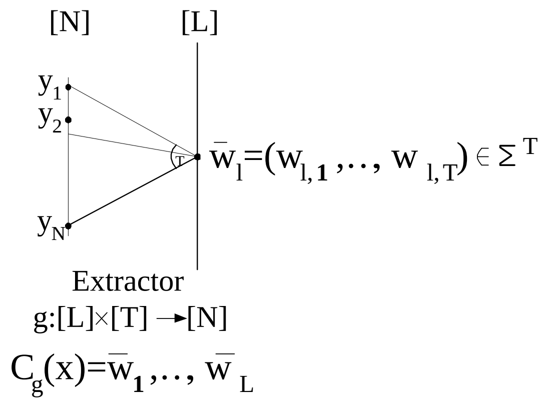

Figure 3: *Exactly as with the previous construction* $y_1, \ldots, y_N \in \Sigma^N$*, the coordinates of some codeword* $C(x)$ *are "put" along the extractor's output* $[N]$*, defining for each* $l \in [L]$ *an ordered vector of its neighbors* $\overline{w}_l = (w_{l,1}, \ldots, w_{l,T}) \in \Sigma^T$*. The vector* $\overline{w}_l$ *is the* l^{th} *symbol in the overall codeword* $C_g(x)$*.*

Now, taking $|X| = \delta L$ for $\delta > \epsilon$ we get $\frac{|\Gamma_g(Y) \cap X|}{Q|Y|} > \frac{1}{2}\delta$ □

Plugging in the same code as before with the extractor instead of the balanced expander in the above construction we get a code with rate

$$r_{\alpha,\beta}\frac{N}{LT} \tag{19}$$

and with alphabet size:

$$O((\frac{1}{\alpha^2})^T) \tag{20}$$

Once again taking $\delta' = (\frac{\delta}{4})$, $\alpha = (\frac{\delta}{8})$ and any $\epsilon < \delta'$, and noting that an optimal extractor with constant error satisfies:

$$T = \Theta(\log(\frac{1}{\epsilon})) \tag{21}$$

and

$$N = \Theta(\epsilon LT) \tag{22}$$

We get by (19) rate $\Omega(\epsilon)$ and by (20) alphabet size $poly(\frac{1}{\epsilon})$.
For an explicit result we can take the Zig-Zag extractor of [RVW00] which when used with constant error it satisfies:

$$T = \Theta(2^{polylolog(\frac{1}{\epsilon})}) \tag{23}$$

and

$$N = \Theta(\epsilon LT) \tag{24}$$

Thus, for $L > (\frac{1}{\epsilon^2})$ we get again rate $\Omega(\epsilon)$, and alphabet size $2^{2^{polyloglog(\frac{1}{\epsilon})}}$.

References

[ABN$^+$92] N. Alon, J. Bruck, J. Naor, M. Naor, and R. Roth. Construction of asymptotically good, low-rate error-correcting codes through pseudo-random graphs. *IEEE Transactions on Information Theory*, 38:509–516, 1992.

[ASE92] N. Alon, J. H. Spencer, and P. Erdős. *The Probabilistic Method.* Wiley–Interscience Series, John Wiley & Sons, Inc., New York, 1992.

[Eli57] P. Elias. List decoding for noisy channels. In *1957-IRE WESCON Convention Record, Pt. 2*, pages 94–104, 1957.

[GI01] Venkatesan Guruswami and Piotr Indyk. Expander-based constructions of efficiently decodable codes. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 658–667, 2001.

[GI02] Venkatesan Guruswami and Piotr Indyk. Near-optimal linear-time codes for unique decoding and new list-decodable codes over smaller alphabets. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 812–821. ACM Press, 2002.

[Gur04] Venkatesan Guruswami. Better extractors for better codes? In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 436–444. ACM Press, 2004.

[RTS00] J. Radhakrishnan and A. Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13(1):2–24, 2000.

[RVW00] O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag product, and new constant-degree expanders and extractors. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, 2000.

[TSZ01] A. Ta-Shma and D. Zuckerman. Extractor codes. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 193–199, 2001.

[TSZS01] Amnon Ta-Shma, David Zuckerman, and Shmuel Safra. Extractors from Reed-Muller codes. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 638–647, October 2001.

[Woz58] J.M. Wozencraft. List decoding. In *Quarterly Progress Report*, volume 48, pages 90–95. Research Laboratory of Electronics, MIT, 1958.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'