ETH zürich Research Collection

Report

A lower bound for a class of graph based loss resilient codes

Author(s):
Blömer, Johannes; Trachsler, Beat

Publication Date:
1998

Permanent Link:
https://doi.org/10.3929/ethz-a-006652860 →

Rights / License:
In Copyright - Non-Commercial Use Permitted →

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.

ETH Library

A Lower Bound for a Class of Graph Based Loss-Resilient Codes

Johannes Blömer Beat Trachsler[1]

May 5, 1998

Abstract. Recently, Luby et al. constructed information-theoretically almost optimal loss-resilient codes. The construction is based on a sequence of bipartite graphs. Using a probabilistic construction for the individual bipartite graphs, they obtain loss-resilient codes that have very efficient encoding and decoding algorithms. They left open the question whether there are other bipartite graphs leading to codes with even more efficient encoding and decoding algorithms. In this paper we show that if one follows the basic construction used by Luby et al., then their choice of bipartite graphs leads to codes with asymptotically optimal encoding and decoding algorithms.

Authors' addresses:
Johannes Blömer and Beat Trachsler
Institute for Theoretical Computer Science
ETH Zentrum
CH-8092 Zurich
Switzerland

Technical Report #298, Departement Informatik, ETH Zürich.
Electronically available from:
ftp://ftp.inf.ethz.ch/pub/publications/tech-reports/

[1]Work of both authors has been supported by Grant 21-49626.96 from Schweizer Nationalfonds (SNF).

1 Introduction

A $(1+\varepsilon)$ loss-resilient code encodes messages consisting of m symbols into an encoding consisting of cm symbols, $c > 1$, such that the m message symbols can be reconstructed from any subset of $(1+\varepsilon)m$ symbols of the encoding. If $\varepsilon = 0$ then these codes are usually called MDS codes (**M**aximum **D**istance **S**eparable). If the reconstruction is guaranteed to be successful only with high probability, then the code is called a probabilistic $(1+\varepsilon)$ loss-resilient code. Here the probability is with respect to random subsets of $(1+\varepsilon)m$ symbols of the encoding.

Loss-resilient codes have applications for example in networking. In real-time applications, where retransmission of lost packets is not feasible, they can be used to protect against the effects of packet losses. For example, many MPEG video players cannot deal with disrupted MPEG streams. Even if they can, video quality usually suffers significantly from packet losses.

In [2] Luby et al. introduced a family of probabilistic $(1+\varepsilon)$ loss-resilient codes with very efficient encoding and decoding algorithms. The codes are based on a sequence of bipartite graphs. The individual bipartite graphs are obtained by a probabilistic construction using a very special distribution on bipartite graphs. This left open the question, whether choosing different bipartite graphs can lead to codes with more efficient encoding and decoding algorithms.

In this paper we show that no matter how the individual bipartite graphs are chosen, codes based on a sequence of bipartite graphs as proposed in [2] lead to loss-resilient codes for which the encoding and decoding requires time $\Omega(m \ln(1/\varepsilon))$. To be more precise, if we want that with non-constant probability a random subset of $(1+\varepsilon)m$ symbols of the encoding allows to reconstruct the entire message, then the average degree of the bipartite graphs has to be $\Omega(\ln(1/\varepsilon))$. This lower bound on the average degree immediately leads to the lower bounds for the encoding and decoding times.

The paper is organized as follows. In Section 2 we review the codes of Luby et al. in more detail. In Sections 3 and 4 we state and prove our main results. In Section 5 we indicate how to refine the proofs of Section 4 in order to obtain lower bounds for the encoding and decoding times that almost match the upper bounds given in [2].

2 An Efficient Loss-Resilient Code

In this section we introduce the coding scheme by Luby et al. The goal is to give a short overview of the construction. Proofs can be found in [2, 1]. For an introduction to coding theory we refer to [3]. In the following we denote message symbols by bits although in practice one actually uses bit strings. It is straightforward to generalize the construction of Luby et al. to bit strings.

The main part of the scheme is a random directed acyclic graph whose nodes are partitioned into sets S_i, $0 \leq i \leq \ell$, such that $|S_{i+1}| = \beta\,|S_i|$, where $0 < \beta < 1$. The nodes in S_0 correspond to the message bits of the code. The number of message bits is denoted by m. The nodes in S_i, $1 \leq i \leq \ell$, correspond to the check bits. The edges of the graph go only from S_i to S_{i+1}, $0 \leq i \leq \ell - 1$. The construction allows for a very simple and efficient encoding algorithm which

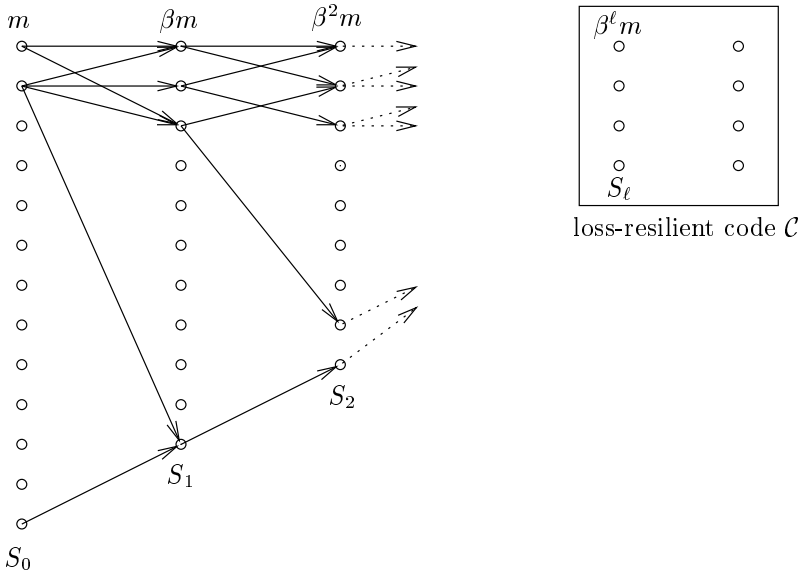

Figure 2.1: Coding scheme by Luby et al.

computes the check bits in the sets $S_2, \ldots, S_\ell$ from left to right by iterating the following operation.

Encoding. *The value of a node in S_{i+1} is computed as the XOR of the values of its neighbors in S_i.*

The running time of this part of the encoding algorithm is proportional to the number of edges in the graph.

The decoding is done from right to left. It turns out to be more problematic although the underlying operation is elementary as well.

Decoding. *Assume that the value of a node c in S_{i+1} and the values of all but one of its neighbors in S_i are known. Then the value of the missing neighbor in S_i is computed by the XOR of the value of c and the values of its known neighbors.*

For this procedure to recover all message bits, one has to guarantee that as long as there are nodes in S_i, whose values are unknown, there is a node c in S_{i+1} to which the basic decoding operation can be applied. In [2] this guarantee is obtained using random bipartite graphs with the following properties to connect S_i and S_{i+1}

1. With high probability the bits in S_i can be reconstructed by the decoding operation from above if all the bits in S_{i+1} are known and only a $(1-\varepsilon)\beta$ part of the bits in S_i is lost.

2. Approximately, the average out-degree of the nodes in S_i is $\ln(1/\varepsilon)$.

To start this procedure with levels S_ℓ and $S_{\ell-1}$ one has to guarantee that all of the bits at the last level S_ℓ are known. The idea is to choose ℓ in such a way that an efficient standard loss-resilient code $\mathcal{C}$ with rate $1-\beta$ can encode and decode S_ℓ in linear time with respect to m. Since in practice the most efficient loss-resilient codes have quadratic time encoding and decoding algorithms, we choose ℓ such that

$$|S_\ell|^2 \leq m.$$

Hence for this part of the encoding and decoding can be done in time linear in m. Since the encoding and the decoding in the graph take time proportional to the number of edges, the running time of the overall encoding and decoding algorithms is $\mathcal{O}(m \ln(1/\varepsilon))$.

From the discussion above it also follows that the code is a $(1+\varepsilon)$ loss-resilient code. Finally, we claim that the rate of the code is $1-\beta$. This follows from the fact that the encoding length is given by

$$\underbrace{\sum_{j=0}^{\ell} \beta^j m}_{\#\text{ nodes in } S_0 \cup \ldots \cup S_\ell} + \underbrace{\frac{\beta^{\ell+1} m}{1-\beta}}_{\#\text{ check bits of } \mathcal{C}} = \frac{m}{1-\beta}.$$

3 The Lower Bound

For the following arguments we only need the structure of the bipartite graph that connects the nodes in S_0 to the nodes in S_1. First, we make the following information-theoretical observation. It is the key argument for the results below.

Proposition 3.1. *Denote the set of nodes corresponding to the lost message bits by T. Then the set $\Gamma(T)$ of nodes connected to T has to be at least as large as T if the decoding algorithm terminates successfully.*

Proof. If $|\Gamma(T)| < |T|$, the check bits do not contain enough information to recover all the message bits. □

From this we can derive a lower bound for the average degree Δ_ℓ of the nodes in S_0.

Lemma 3.2. *Assume that a node in S_0 is contained in T uniformly and independently with probability $(1-\varepsilon)\beta$, where $0 < \varepsilon < 1$ is constant. If the average*

degree Δ_ℓ of the nodes in S_0 satisfies

$$\Delta_\ell < \frac{1}{4}(1-\beta)\ln\frac{1}{\varepsilon}, \tag{3.1}$$

then the probability that the set $\Gamma(T)$ is smaller than T is $\Omega(1)$.

The proof of this lemma will be given in the next section.

Remark 3.1. Notice that β is the information-theoretically optimal erasure probability for the message bits. A coding scheme that guarantees a successful decoding at this erasure probability leads to an optimal MDS code.

Proposition 3.1 and Lemma 3.2 imply the following lower bound for the running time of the encoding and decoding algorithm by Luby et al.

Theorem 3.3. *Assume that the message bits are lost uniformly and independently with probability $(1-\varepsilon)\beta$, where $0 < \varepsilon < 1$ is constant. If the probability that the decoding fails is required to be $o(1)$, then the encoding and the decoding takes time $\Omega(m \ln(1/\varepsilon))$.*

4 The Proof of Lemma 3.2

For the proof we use the following strategy. To show that for $\Delta_\ell < \frac{1}{4}(1-\beta)\ln\frac{1}{\varepsilon}$ the probability $\mathbf{Pr}[|\Gamma(T)| < |T|]$ is $\Omega(1)$, for all values of $|T|$ in some interval I, we estimate the expected size of the set $\Gamma(T)$ of neighbors of T. Then we apply Markov's inequality to obtain a constant lower bound on $\mathbf{Pr}[|\Gamma(T)| < |T| \,||T| = t]$. Choosing I such that $\mathbf{Pr}[|T| \in I]$ is constant, finishes the proof.

Since every node in S_0 lies in T with probability $\delta = (1-\varepsilon)\beta$, the expectation and the standard deviation are

$$\begin{aligned}\mathbf{E}[|T|] &= \delta m,\\ \boldsymbol{\sigma}[|T|] &= \sqrt{\delta(1-\delta)m}.\end{aligned}$$

Define I by

$$I = [t_{min}, t_{max}] \qquad \text{where} \qquad \begin{aligned} t_{min} &\stackrel{\text{def}}{=} (1-2\varepsilon)\beta m\\ t_{max} &\stackrel{\text{def}}{=} \beta m.\end{aligned}$$

Notice that the expectation of $|T|$ lies in the middle of I. By Chebyshev's inequality

$$\begin{aligned}\mathbf{Pr}\left[|T| \notin I\right] &= \mathbf{Pr}\left[\big||T| - \mathbf{E}[|T|]\big| > \varepsilon\beta m\right]\\ &= \mathbf{Pr}\left[\big||T| - \mathbf{E}[|T|]\big| > \tfrac{\varepsilon\beta\sqrt{m}}{\sqrt{\delta(1-\delta)}}\,\boldsymbol{\sigma}[|T|]\right]\\ &\le k(\varepsilon,\beta)\frac{1}{m} \qquad \text{where } k(\varepsilon,\beta) = \tfrac{\delta(1-\delta)}{\varepsilon^2\beta^2}.\end{aligned}$$

Hence

$$\mathbf{Pr}\left[|T| \in I\right] \geq 1 - k(\varepsilon,\beta)\frac{1}{m}.$$

We want to show that for any $t \in I$ the probability $\mathbf{Pr}\left[|\Gamma(T)| < |T| \,\middle|\, |T| = t\right]$ is $\Omega(1)$. For this we compute $\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right]$.

Fix an arbitrary $t \in I$. Conditioned on $|T| = t$ the probability that a fixed node $w \in S_1$ is not connected to a random set T of size t is $\binom{m-d_w}{t}/\binom{m}{t}$, where d_w denotes the in-degree of w. Hence

$$\mathbf{Pr}\left[w \in \Gamma(T) \,\middle|\, |T| = t\right] = 1 - \frac{\binom{m-d_w}{t}}{\binom{m}{t}}.$$

Since $|S_1| = \beta m$,

$$\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right] = \beta m - \sum_{w \in S_1} \frac{\binom{m-d_w}{t}}{\binom{m}{t}}.$$

Lemma 4.1. *If the average in-degree Δ_r in S_1 satisfies $\Delta_r \leq m - t$, then*

$$\sum_{w \in S_1} \frac{\binom{m-d_w}{t}}{\binom{m}{t}} \geq \beta m \left(1 - \frac{\Delta_r}{m-t}\right)^t.$$

Proof. Since $\binom{x}{t}$ is a convex function in x,

$$\sum_{w \in S_1} \frac{\binom{m-d_w}{t}}{\binom{m}{t}} \geq \beta m \frac{\binom{m-\Delta_r}{t}}{\binom{m}{t}}.$$

Now the proof is completed by the following calculation

$$\begin{aligned}
\frac{\binom{m-\Delta_r}{t}}{\binom{m}{t}} &= \frac{(m-\Delta_r)(m-\Delta_r-1)\cdot\ldots\cdot(m-\Delta_r-t+1)}{m(m-1)\cdot\ldots\cdot(m-t+1)} \\
&= \left(1 - \frac{\Delta_r}{m}\right)\cdot\left(1 - \frac{\Delta_r}{m-1}\right)\cdot\ldots\cdot\left(1 - \frac{\Delta_r}{m-t+1}\right) \\
&\geq \left(1 - \frac{\Delta_r}{m-t}\right)^t
\end{aligned}$$

□

Since $\Delta_r = \frac{\Delta_\ell}{\beta}$ and since Δ_ℓ is bounded by a constant, we can assume that $\Delta_r \leq \frac{1}{2}(m-t)$ for m large enough. By Lemma 4.1 and using $1 - x \geq e^{-2x}, 0 \leq x \leq 1/2$,

$$\begin{aligned}
\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right] &\leq \beta m \left(1 - \left(1 - \frac{\Delta_r}{m-t}\right)^t\right) \\
&\leq \beta m \left(1 - e^{-2\frac{t}{m-t}\Delta_r}\right).
\end{aligned}$$

Since $\frac{t}{m-t} \leq \frac{t_{max}}{m-t_{max}} = \frac{\beta}{1-\beta}$,

$$\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right] \leq \beta m \left(1 - e^{-2\frac{\beta}{1-\beta}\Delta_r}\right). \tag{4.1}$$

for an arbitrary $t \in I$.
Now we derive a suitable inequality for the conditional expectation of $|\Gamma(T)|$.

Lemma 4.2. *If* $\Delta_\ell < \frac{1}{4}(1-\beta)\ln\frac{1}{\varepsilon}$, *then*

$$\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right] < \left(1 - \sqrt{\varepsilon}\right)\beta m \tag{4.2}$$

for any $t \in I$.

Proof. Assume that

$$\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right] \geq \left(1 - \sqrt{\varepsilon}\right)\beta m.$$

Then with (4.1)

$$\sqrt{\varepsilon} \geq e^{-2\frac{\beta}{1-\beta}\Delta_r}.$$

Hence

$$\Delta_r \geq \frac{1}{4}\left(\frac{1-\beta}{\beta}\right)\ln\frac{1}{\varepsilon}$$

and finally

$$\Delta_\ell = \beta\Delta_r \geq \frac{1}{4}(1-\beta)\ln\frac{1}{\varepsilon}.$$

This contradicts the condition of the lemma. □

Now we can finish the proof of Lemma 3.2.

$$\mathbf{Pr}\left[|\Gamma(T)| \geq |T| \,\middle|\, |T| = t\right] \leq \mathbf{Pr}\left[|\Gamma(T)| \geq t_{min} \middle| \, |T| = t\right]$$

for an arbitrary $t \in I$. Since in Lemma 3.2 we assume $\Delta_\ell < \frac{1}{4}(1-\beta)\ln\frac{1}{\varepsilon}$, inequality (4.2) implies

$$\begin{aligned} t_{min} = (1-2\varepsilon)\beta m &= \frac{1-2\varepsilon}{1-\sqrt{\varepsilon}}(1-\sqrt{\varepsilon})\beta m \\ &> \frac{1-2\varepsilon}{1-\sqrt{\varepsilon}}\,\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right] \end{aligned}$$

for any $t \in I$. By Markov's inequality

$$\mathbf{Pr}\left[|\Gamma(T)| \geq |T| \,\middle|\, |T| = t\right] \leq \frac{1-\sqrt{\varepsilon}}{1-2\varepsilon}.$$

Thus

$$\mathbf{Pr}\left[|\Gamma(T)| < |T| \,\middle|\, |T| = t\right] \geq \frac{\sqrt{\varepsilon} - 2\varepsilon}{1 - 2\varepsilon}$$

for an arbitrary $t \in I$. Hence

$$\begin{aligned}\mathbf{Pr}\left[|\Gamma(T)| < |T| \,\middle|\, |T| \in I\right] &= \sum_{t \in I} \mathbf{Pr}\left[|\Gamma(T)| < |T| \,\middle|\, |T| = t\right] \mathbf{Pr}\left[|T| = t\right] \\ &\geq \frac{\sqrt{\varepsilon} - 2\varepsilon}{1 - 2\varepsilon} \sum_{t \in I} \mathbf{Pr}\left[|T| = t\right] \\ &= \frac{\sqrt{\varepsilon} - 2\varepsilon}{1 - 2\varepsilon} \mathbf{Pr}\left[|T| \in I\right].\end{aligned}$$

Therefore

$$\begin{aligned}\mathbf{Pr}\left[|\Gamma(T)| < |T|\right] &\geq \mathbf{Pr}\left[|\Gamma(T)| < |T| \,\middle|\, |T| \in I\right] \mathbf{Pr}\left[|T| \in I\right] \\ &\geq \frac{\sqrt{\varepsilon} - 2\varepsilon}{1 - 2\varepsilon} \left(\mathbf{Pr}\left[|T| \in I\right]\right)^2 \\ &\geq \frac{\sqrt{\varepsilon} - 2\varepsilon}{1 - 2\varepsilon} \left(1 - 2k(\varepsilon, \beta)\frac{1}{m}\right) = \Omega(1).\end{aligned}$$

This completes the proof.

5 A Degree Bound with Better Constants

In this section we want to improve the constants in Lemma 3.2.

Corollary 5.1. *Assume that a node in S_0 is contained in T uniformly and independently with probability $(1 - \varepsilon)\beta$, where $0 < \varepsilon < 1$ is constant. If the average degree Δ_ℓ of the nodes in S_0 satisfies*

$$\Delta_\ell < (1 - \beta)(1 - \varepsilon) \ln \frac{1}{\varepsilon} - C \tag{5.1}$$

where $C = (1 - \beta)(1 - \varepsilon) \ln 2$, then the probability that the set $\Gamma(T)$ is smaller than T is $\Omega(1)$.

Remark 5.1. In the construction by Luby et al. the average degree of the nodes in S_0 is at least $\ln \frac{1}{\varepsilon}$ (see [2]). This is for small ε up to a constant factor $(1 - \beta)$ the same as the right hand side of (5.1).

The proof of this corollary is almost the same as the proof of Lemma 3.2. Consider inequality (4.1). It can be replaced by

$$\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right] \leq \beta m \left(1 - e^{\frac{\beta}{c(1-\beta)}\Delta_r}\right) \tag{5.2}$$

for an arbitrary $t \in I$ where $0 < c < 1$. Instead of (4.2) we obtain

Lemma 5.2. *If* $\Delta_\ell < (1-\beta)(1-\varepsilon)\ln\frac{1}{2\varepsilon}$, *then*

$$\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right] < \left(1 - (2\varepsilon)^{\sqrt{1-\varepsilon}}\right)\beta m \tag{5.3}$$

for any $t \in I$.

Proof. Assume that

$$\mathbf{E}\left[|\Gamma(T)| \,\middle|\, |T| = t\right] \geq \left(1 - (2\varepsilon)^{\sqrt{1-\varepsilon}}\right)\beta m.$$

Then by (5.2) with $c = \sqrt{1-\varepsilon}$

$$(2\varepsilon)^{\sqrt{1-\varepsilon}} \geq e^{\frac{\beta}{\sqrt{1-\varepsilon}(1-\beta)}\Delta_r}.$$

Hence

$$\Delta_r \geq \left(\frac{1-\beta}{\beta}\right)(1-\varepsilon)\ln\frac{1}{2\varepsilon}$$

and finally

$$\Delta_\ell = \beta\Delta_r \geq (1-\beta)(1-\varepsilon)\ln\frac{1}{2\varepsilon}.$$

This contradicts the condition of the lemma. □

The rest of the proof follows along the same lines as the proof in Section 4.

References

[1] Michael G. Luby, Michael Mitzenmacher, and M. Amin Shokrollahi. Analysis of random processes via and-or tree evaluation. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 364–373, San Francisco, California, 25–27 January 1998.

[2] Michael G. Luby, Michael Mitzenmacher, M. Amin Shokrollahi, Daniel A. Spielman, and Volker Stemann. Practical loss-resilient codes. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 150–159, El Paso, Texas, 4–6 May 1997.

[3] J. H. van Lint. *Introduction to Coding Theory.* Springer-Verlag, 1982.