Foreword

This chapter is based on lecture notes from coding theory courses taught by Venkatesan Guruswami at University at Washington and CMU; by Atri Rudra at University at Buffalo, SUNY and by Madhu Sudan at MIT.

This version is dated **April 28, 2013**. For the latest version, please go to

http://www.cse.buffalo.edu/ atri/courses/coding-theory/book/

The material in this chapter is supported in part by the National Science Foundation under CAREER grant CCF-0844796. Any opinions, findings and conclusions or recomendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

Chapter 8

What Cannot be Done-II

In this brief interlude of a chapter, we revisit the trade-offs between rate and relative distance for codes. Recall that the best (and only) lower bound on R that we have seen is the GV bound and the best upper bound on R that we have have seen so far is a combination of the Plotkin and Hamming bounds (see Figure 4.5). In this chapter, we will prove the final upper bound on R in this book due to Elias and Bassalygo. Then we will mention the best known upper bound on rate (but without stating or proving it). Finally, we will conclude by summarizing what we have seen so far and laying down the course for the rest of the book.

8.1 Elias-Bassalygo bound

We begin with the statement of a new upper bound on the rate called the Elias-Bassalygo bound.

Theorem 8.1.1 (Elias-Bassalygo bound)**.** *Every q-ary code of rate R, distance δ, and large enough block length, satisfies the following:*

$$R \le 1 - H_q\left(J_q(\delta)\right) + o(1)$$

See Figure 8.1 for an illustration of the Elias-Bassalygo bound for binary codes. Note that this bound is tighter than all the previous upper bounds on rate that we have seen so far.

The proof of Theorem 8.1.1 uses the following lemma:

Lemma 8.1.2. *Given a q-ary code, $C \subseteq [q]^n$ and $0 \le e \le n$, there exists a Hamming ball of radius e with at least $\frac{|C| Vol_q(e,n)}{q^n}$ codewords in it.*

Proof. We will prove the existence of the required Hamming ball by the probabilistic method. Pick a received word $\mathbf{y} \in [q]^n$ at random. It is easy to check that the expected value of $|B(\mathbf{y}, e) \cap C|$ is $\frac{|C| Vol_q(e,n)}{q^n}$. (We have seen this argument earlier in the proof of part (ii) of Theorem 7.4.2.)

This by the probabilistic method implies the existence of a $\mathbf{y} \in [q]^n$ such that

$$|B(\mathbf{y}, e) \cap C| \ge \frac{|C|\, Vol_q(e,n)}{q^n},$$

as desired. □

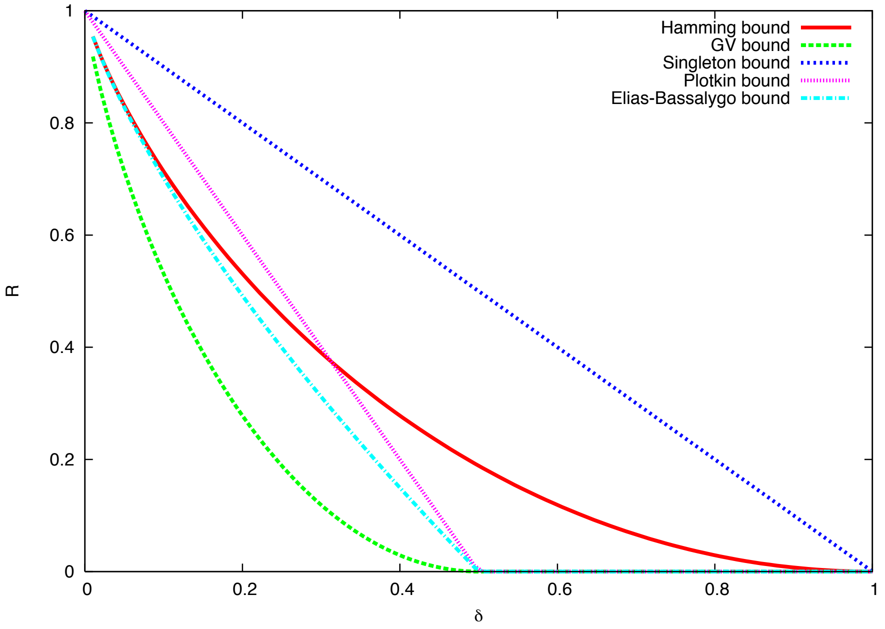

Figure 8.1: Singleton, Hamming, Plotkin, GV and Elias-Bassalygo bounds on rate versus distance for binary codes.

Proof of Theorem 8.1.1. Let $C \subseteq [q]^n$ be any code with relative distance δ. Define $e = nJ_q(\delta) - 1$. By Lemma 8.1.2, there exists a Hamming ball with $\mathscr{B}$ codewords such that the following inequality is true:

$$\mathscr{B} \geq \frac{|C|\, Vol_q(e,n)}{q^n}.$$

By our choice of e and the Johnson bound (Theorem 7.3.1), we have

$$\mathscr{B} \leq qdn.$$

Combining the upper and lower bounds on $\mathscr{B}$ implies the following

$$|C| \leq qnd \cdot \frac{q^n}{Vol_q(e,n)} \leq q^{n\left(1-H_q(J_q(\delta))+o(1)\right)},$$

where the second inequality follows from our good old lower bound on the volume of a Hamming ball (Proposition 3.3.1) and the fact that $qdn \leq qn^2 \leq q^{o(n)}$ for large enough n. This implies that the rate R of C satisfies:

$$R \leq 1 - H_q\left(J_q(\delta)\right) + o(1),$$

as desired. □

8.2 The MRRW bound: A better upper bound

The MRRW bound (due to McEliece, Rodemich, Rumsey and Welch) is based on a linear programming approach introduced by Delsarte to bound the rate of a code. The MRRW bound is a better upper bound than the Elias-Bassalygo bound (though we will not state or prove the bound in this book). However, there is a gap between the Gilbert-Varshamov (GV) bound and the MRRW bound. The gap still exists to this day. To give one data point on the gap, consider $\delta = \frac{1}{2} - \varepsilon$ (think of $\varepsilon \to 0$), the GV bound gives a lower bound on R of $\Omega\left(\varepsilon^2\right)$ (see Proposition 3.3.5), while the MRRW bound gives an upper bound on R of $O\left(\varepsilon^2 \log\left(\frac{1}{\varepsilon}\right)\right)$.

8.3 A Breather

Let us now recap the combinatorial results that we have seen so far. Table 8.1 summarizes what we have seen so far for binary codes in Shannon's world and Hamming's world (under both unique and list decoding settings).

Shannon	Hamming	
BSC$_p$	Unique	List
1-H$\left(p\right)$ is capacity	$R \geq 1 - H(\delta)$	$1 - H\left(p\right)$ is list decoding capacity
	$R \leq MRRW$	
Explicit codes at capacity?	Explicit Asymptotically good codes?	Explicit codes at capacity?
Efficient decoding algorithm?	Efficient decoding algorithms?	Efficient decoding algorithms?

Table 8.1: High level summary of results seen so far.

For the rest of the section, we remind the reader about the definition of explicit codes (Definition 6.3.1) and strongly explicit codes (Definition 6.3.2).

We begin with BSC$_p$. We have seen that the capacity of BSC$_p$ is $1 - H(p)$. The most natural open question is to obtain the capacity result but with explicit codes along with efficient decoding (and encoding) algorithms (Question 6.3.1).

Next we consider Hamming's world under unique decoding. For large enough alphabets, we have seen that Reed-Solomon codes (Chapter 5) meet the Singleton bound (Theorem 4.3.1). Further, the Reed-Solomon codes are strongly explicit[1]. The natural question then is

Question 8.3.1. *Can we decode Reed-Solomon codes up to half its distance?*

For smaller alphabets, especially binary codes, as we have seen in the last section, there is a gap between the best known lower and upper bounds on the rate of a code with a given relative

[1] The proof is left as an exercise.

distance. Further, we do not know of an explicit construction of a binary code that lies on the GV bound. These lead to the following questions that are still wide open:

Open Question 8.3.1. *What is the optimal trade-off between R and δ?*

Open Question 8.3.2.

Does there exist an explicit construction of (binary) codes on the GV bound?

If we scale down our ambitions, the following is a natural weaker version of the second question above:

Question 8.3.2. *Do there exist explicit asymptotically good binary codes?*

We also have the following algorithmic counterpart to the above question:

Question 8.3.3. *If one can answer Question 8.3.2, then can we decode such codes efficiently from a non-zero fraction of errors?*

For list decoding, we have seen that the list decoding capacity is $1 - H_q(p)$. The natural open questions are whether we can achieve the capacity with explicit codes (Question 7.4.1) along with efficient list decoding algorithms (Question 7.4.2).

In the remainder of the book, we will focus on the questions mentioned above (and summarized in the last two rows of Table 8.1).

8.4 Bibliographic Notes

The McEliece-Rodemich-Rumsey-Welch (MRRW) bound was introduced in 1977 in the paper [42].