CSE 533: Error-Correcting Codes **(Autumn 2006)**

Lecture 9: Concatenated Codes

October 27, 2006

Lecturer: Venkatesan Guruswami *Scribe: Younsun Kim*

1 Last Lecture

- Reed Muller code distance properties
- Majority logic decoding
- Binary codes based on Reed Solomon code
- Field extension

2 Concatenated Codes [Forney '66]

Concatenated codes are generated from an operation of combining 2 codes to obtain, $C_1 \diamond C_2$, where C_1 is an $(N, K, D)_Q$ code and C_2 is an $(n, log_q Q, d)_q$ code. Figure 1 shows the conceptual operation of the outer code, C_1, and the inner code, C_2, in generating in concatenated code.

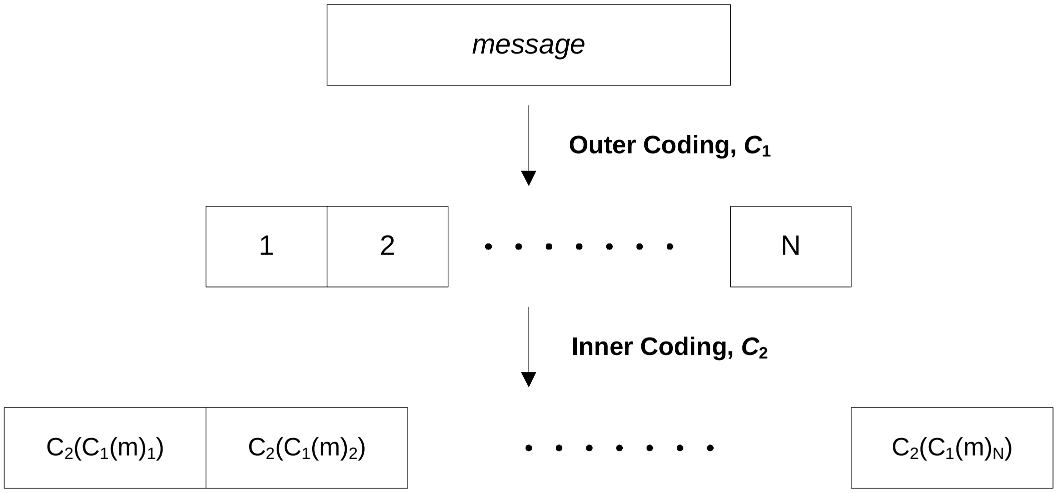

Figure 1: Conceptual operation of code concatenating.

The concatenated code generated from an outer code of $(N, K, D)_Q$ and an inner code of $(n, log_q Q, d)_q$ is an $(Nn, log_q Q, \geq Dd)$ code. Note that these numbers are similar to the code generated from the tensor product of a $(N, K, D)_Q$ code and a $(n, log_qQ, d)_q$ code. The only difference is that distance of the concatenated code is equal or greater than the product of the individual distances whereas the distance of the code generated from tensor product is equal to the product of the individual distances.

Proof. Distance of concatenated code $\geq$ *Dd*. Assume an outer code, C_1, of $(N, K, D)_Q$ and an inner code, C_2, of $(n, log_qQ, d)_q$. For two messages m_1 and m_2 ($m_1 \neq m_2$),

$$\triangle\left(C_1(m_1), C_1(m_2)\right) \geqq D \tag{1}$$

since the minimum distance of C_1 is *D*. The outer coding produces a distance equal to or larger than *D* in each of the *N* blocks and while the inner coding produces a distance of equal to or greater than *d* in all these blocks. Therefore, the distance of the concatenated code is equal to or greater than *D*.

If $C_1 \diamond C_2$ is linear and $Q = q^m$ where *m* is an integer, then $C_1 \diamond C_2$ is also linear over $\mathbb{F}_q$.

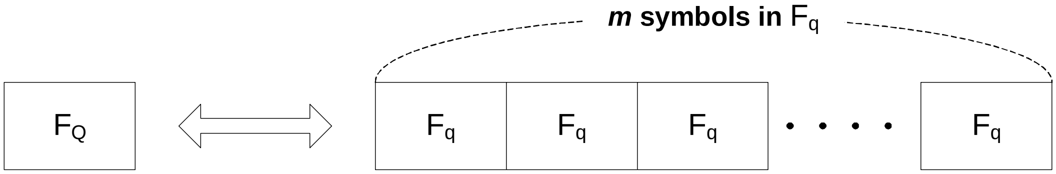

Figure 2: $\sigma : \mathbb{F}_Q \rightarrow \mathbb{F}_{q^m}$.

The transition from $\mathbb{F}_Q$ to $\mathbb{F}_q$ denoted by σ is possible if

- $\sigma\left(x + y\right) = \sigma\left(x\right) + \sigma\left(y\right)$
- $\sigma\left(\alpha x\right) = \alpha\sigma\left(x+\right)$

For example, $\mathbb{F}_{2^m}$ is an m dimensional vector space over $\mathbb{F}_2$ such that $\mathbb{F}_{2^m} = \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m \mid \alpha_i \in \mathbb{F}_2\}$.

In order to discuss the distance properties of the concatenated code based on an RS code and a binary linear code. The following assumptions will be made.

- Outer code, C_1: RS code of $[N, K, N - K + 1]_{2^m}$ and code rate $R = K/N$.
- Inner code, C_2: Binary linear code of $[n, m, d]_2$ and code rate$r = m/n$.

The GV bound can be expressed as

$$r = 1 - H(\delta) \tag{2}$$

$$\frac{d}{n} = H^{-1}(1 - r) - \varepsilon \tag{3}$$

The concatenated code $C_1 \diamond C_2$ is a binary linear code of $[Nn, km, \geq D(N-K+1)]_2$. The rate and the relative distance of the concatenated code can be expressed as

$$rate = Rr \tag{4}$$

$$\textit{relative distance} \geq (1-R)(H^{-1}(1-r)-\varepsilon) \tag{5}$$

Note that $H^{-1}(y)$ with $0 \leq y \leq 1$ has a unique $x \in [0, 1/2]$ such that $H(x) = y$. Since $rate = Rr$, this gives a binary linear codes of rate R_0 and relative distance at least

$$\delta_{Zyablov}(R_0) = \max_{R_0<r<1} \left(1 - \frac{R_0}{r}\right) H^{-1}(1-r) \tag{6}$$

This bound is known as the Zyablov bound and can be plotted as shown below.

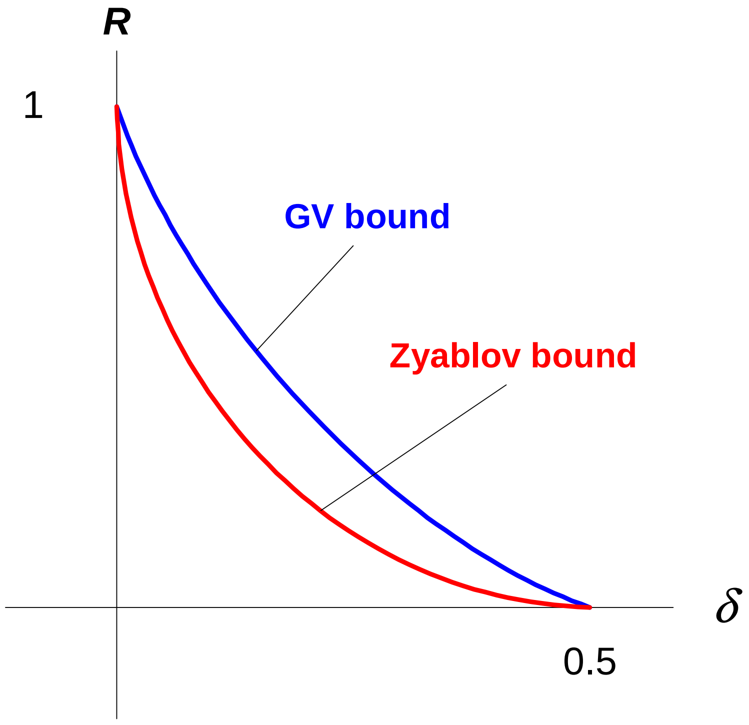

Figure 3: GV bound vs Zyablov bound.

The significance of the concatenated code is that it is possible to construct a linear binary code meeting the Zyablov bound in polynomial, $2^{0(m)}$, time.

Another way of writing the Zyablov bound is shown below.

$$R_{Zyablov}(\delta) = \max_{\delta_{in}\delta_{out}=\delta} (1-\delta_{out})(1-H(\delta_{in})) \tag{7}$$

For $\delta = 1/2 - \varepsilon$ ($R \approx 0$ but positive)

- GV bound: $R(\frac{1}{2} - \varepsilon) = 1 - H(\frac{1}{2} - \varepsilon) = 1 - (1 - (1 - 0(\varepsilon^2))) = \Omega(\varepsilon^2)$
- Zyablov bound: $\delta_{out} = 1 - \varepsilon, \delta_{in} = \frac{1}{2} - \varepsilon/2 \Rightarrow \delta_{in}\delta_{out} = \frac{1}{2} - \varepsilon + \varepsilon^2/2 \geq \frac{1}{2} - \varepsilon$

Note that the rate concatenated code in the above case is

$$rate = (1 - (1 - \varepsilon))(1 - H(\frac{1}{2} - \frac{\varepsilon}{2}) = \Omega(\varepsilon^3) \tag{8}$$

which is off from the GV bound by a factor of ε. Getting the rate asymptotically better than $\Omega(\varepsilon^3)$ for $\delta = \frac{1}{2} - \varepsilon$ with polynomial construction time is an open question.

Applying $\delta = 1/2 - \varepsilon$ to some of the bounds that were discussed in class

- Elias bound: $R(\delta) = 1 - H(\frac{1-\sqrt{1-2\delta}}{2}) \Rightarrow R(\delta = \frac{1}{2} - \varepsilon) = 1 - H(\frac{1-\sqrt{\varepsilon}}{2}) \leq 0(\varepsilon)$
- Plotkin bound: $R(\delta) \leq 1 - 2\delta \leq 2\varepsilon$
- Record bound (known as MRRW or LP bound, 1978):
  $R(\delta) \leq H(\frac{1}{2} - \sqrt{\delta(1-\delta)}) \Rightarrow R(\delta = \frac{1}{2} - \varepsilon) \leq H\left(\frac{1}{2} - \sqrt{\frac{1}{4} - \varepsilon^2}\right) \approx 2\varepsilon^2 log\frac{1}{q}$

3 Justesen's Code [Justesen '73]

Asymptotically good codes for binary refers to a family of (*R*,δ) codes with *R*, $\delta > 0$. Such a family of binary codes was found by Justesen. The main idea behind Justesen's code is that

- All inner codes do not have to be the same.
- Not all the inner goods have to be good (i.e., on the GV bound). Just need an ensemble of *N* codes most of which are good.

In Forney's version of the concatenated code, an asymptotically good code was found and applied to all the code symbols produced by an outer code. But the codes applied to different coordinates need not to be the same. What is needed is just an ensemble of *N* codes most of which are good. Justesen defined a generalized notion of concatenation as follows. Given an $[N, K, D]_Q$ code *C* and a vector of inner $[n, k, d_i]_q$ codes $(C_0, C_1, \cdots, C_{N-1})$ such that for all but εN inner codes $d_i \geq d$, he defined $C \cdot (C_0, C_1, \cdots, C_{N-1})$ to be the code obtained by applying C_i on i^{th} coordinate of outer code (and concatenating the results).

A rate $\frac{1}{2}$ code construction. Let $N = 2^m - 1$ with the following outer and inner code

- Outer code, RS $[N = 2^m - 1, K, N - K + 1]_{2^m}$
- Inner code, $\{C_{in}^{\alpha}\}_{\alpha \in \mathbb{F}_{2^m} - \{0\}}$, each C_{in}^{α} is an $[2m, m, ?]_2$ binary linear code

The overall rate of the concatenated code becomes $\frac{1}{2}$ and the Justesen code can be expressed as

$$f(x) \rightarrow (f(\alpha_0), f(\alpha_1) \cdots f(\alpha_{N-1})) \tag{9}$$

$$xf(x) \rightarrow (\alpha_0 f(\alpha_0), \alpha_1 f(\alpha_1) \cdots \alpha_{N-1} f(\alpha_{N-1})) \tag{10}$$

To be continued in next lecture.