

## Lecture 25: Justesen Codes

October 26, 2007

Lecturer: Atri Rudra

Scribe: Kanke Gao

In the last lecture, we introduced concatenation code with an outer code  $C_{out}$  and an inner code  $C_{in} = [n, k, d]_q$ . We derived Zyablov bound by picking  $C_{out}$  on Singleton bound and  $C_{in}$  on the GV bound. We also presented a poly time construction of a code that achieves the zyablov bound (and hence, an asymptotically good code). A somewhat unsatisfactory aspect of this construction was the brute force search for a suitable inner code (which lead to the polynomial construction time). In today's lecture, we will study a "super" explicit construction of an asymptotically good code.

## 1 "super" explicit construction

As poly time construction of asymptotically good code was presented in the last lecture. A natural question left is if we can have a "super" explicit construction. Technically speaking, by super explicit construction, we mean a log space construction. However, we will not formally define this notation. We will now consider the so called *Justesen code* [1]. Justesen code is concatenation code with *different* linear inner codes, which is composed of an  $(N, K, D)_{q^k}$  outer code  $C_{out}$  and different  $(n, k, d)_q$  inner codes  $C_{in}^i : 1 \leq i \leq N$ . We will need the following result.

**Theorem 1.1.** *There exists an ensemble of inner code  $C_{in}^1, C_{in}^2, \dots, C_{in}^N$  of rate  $\frac{1}{2}$ , where  $N = q^k - 1$ , such that for at least  $(1 - \varepsilon)N$  values of  $i$ ,  $C_{in}^i$  has relative distance  $\geq H_q^{-1}(\frac{1}{2} - \varepsilon)$ ,  $\varepsilon > 0$ .*

In fact, this ensemble is the following. For  $\alpha \in \mathbb{F}_{q^k}^*$ , the inner code  $C_{in}^\alpha : \mathbb{F}_q^k \rightarrow \mathbb{F}_q^{2k}$ , is defined as  $C_{in}^\alpha(x) = (x, \alpha x)$ . This ensemble is due to Wozencraft and is called the Wozencraft ensemble.

## 2 Justesen code

For the Justesen code, the outer code  $C_{out}$  is Reed-Solomon code over  $\mathbb{F}_{q^k}$  evaluated over  $\mathbb{F}_{q^k}^*$  of rate  $R$ ,  $0 < R < 1$ . The outer code  $C_{out}$  have relative distance  $\delta_{out} = 1 - R$  and block length of  $N = q^k - 1$ . There are a set of inner codes  $\{C_{in}^\alpha\}_{\alpha \in \mathbb{F}_{q^k}^*}$ . So Justesen code is the concatenated code  $C^* = C_{out} \circ (C_{in}^1, C_{in}^2, \dots, C_{in}^N)$  with the rate  $\frac{R}{2}$ . The following proposition estimates the distance of  $C^*$ .

**Proposition 2.1.**  *$C^*$  has relative distance at least  $(1 - R - \varepsilon)H_q^{-1}(\frac{1}{2} - \varepsilon)$*

*Proof.* Consider  $\mathbf{m}^1 \neq \mathbf{m}^2 \in (\mathbb{F}_{q^k})^K$ . By the distance of outer codes  $|\mathbf{S}| \geq D$ , where  $\mathbf{S} = \{i | C_{out}(\mathbf{m}^1)_i \neq C_{out}(\mathbf{m}^2)_i\}$ . Call the  $i$ th inner code "good", if  $C_{in}^i$  has distance at least  $d \triangleq$

$H_q^{-1}(\frac{1}{2} - \varepsilon) \cdot 2k$ , otherwise bad. Note that by Theorem 1.1, there are at most  $\varepsilon N$  bad inner codes. Let  $\mathbf{S}_g$  be the set of all good inner codes in  $\mathbf{S}$ , while  $\mathbf{S}_b$  is the set of all bad inner codes in  $\mathbf{S}$ . Since  $\mathbf{S}_b \leq \varepsilon N$ ,

$$\begin{aligned} |\mathbf{S}_g| &= |\mathbf{S}| - |\mathbf{S}_b| \\ &\geq (1 - R - \varepsilon)N. \end{aligned} \tag{1}$$

For each good  $i \in S$

$$\Delta(C_{in}^i(C_{out}(\mathbf{m}^1)_i), C_{in}^i(C_{out}(\mathbf{m}^2)_i)) \geq d \tag{2}$$

Finally, from (1) and (2), we obtain that the distance of  $C^*$  is at least

$$(1 - R - \varepsilon)Nd \geq (1 - R - \varepsilon)H_q^{-1}(\frac{1}{2} - \varepsilon)N \cdot 2k,$$

as desired. □

**Corollary 2.2.** *The concatenated code  $C^*$  is an asymptotically good code and has a “super” explicit construction.*

Thus, we have now satisfactorily answered the question of whether explicit asymptotically good (binary) codes exist modulo Theorem 1.1, which we prove next.

## 2.1 Proof of Theorem 1.1

Fix  $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2) \in \mathbb{F}_q^{2k} \setminus \{\mathbf{0}\}$ . Note that this implies that  $\mathbf{y}_1 = \mathbf{0}$  and  $\mathbf{y}_2 = \mathbf{0}$  are not possible. We claim that  $\mathbf{y} \in C_{in}^\alpha$  for at most one  $\alpha \in \mathbb{F}_{2^k}^*$ . The proof is by a simple case analysis.

- Case 1:  $\mathbf{y}_1 \neq \mathbf{0}$  and  $\mathbf{y}_2 \neq \mathbf{0}$ , then  $\mathbf{y} \in C_{in}^\alpha$ , where  $\alpha = \frac{\mathbf{y}_2}{\mathbf{y}_1}$ .
- Case 2:  $\mathbf{y}_1 \neq \mathbf{0}$  and  $\mathbf{y}_2 = \mathbf{0}$ , then  $\mathbf{y} \notin C_{in}^\alpha$  for every  $\alpha \in \mathbb{F}_{2^k}^*$  (as  $\alpha\mathbf{y} \neq \mathbf{0}$ ).
- Case 3:  $\mathbf{y}_1 = \mathbf{0}$  and  $\mathbf{y}_2 \neq \mathbf{0}$ , then  $\mathbf{y} \notin C_{in}^\alpha$  for every  $\alpha \in \mathbb{F}_{2^k}^*$  (as  $\alpha\mathbf{y} = \mathbf{0}$ ).

Now assume that  $wt(\mathbf{y}) < H_q^{-1}(1 - \varepsilon)n$ . Note that if  $\mathbf{y} \in C_{in}^\alpha$ , then  $C_{in}^\alpha$  is “bad”(i.e. has relative distance  $< H_q^{-1}(\frac{1}{2} - \varepsilon)$ ). Since  $\mathbf{y} \in C_{in}^\alpha$  for at most one value of  $\alpha$ , the total number of bad codes is at most

$$\begin{aligned} |\{\mathbf{y} | wt(\mathbf{y}) < H_q^{-1}(\frac{1}{2} - \varepsilon) \cdot 2k\}| &\leq Vol_q(\mathbf{0}, H_q^{-1}(\frac{1}{2} - \varepsilon) \cdot 2k) \\ &\leq q^{H_q(H_q^{-1}(\frac{1}{2} - \varepsilon)) \cdot 2k} \\ &= q^{(\frac{1}{2} - \varepsilon) \cdot 2k} \\ &= \frac{q^k}{2\varepsilon k} \\ &< \varepsilon(q^k - 1) \text{ (for large enough R)} \\ &= \varepsilon N \end{aligned}$$

Thus for at least  $(1 - \varepsilon)N$  values of  $\alpha$ ,  $C_{in}^\alpha$  has relative distance at least  $H_q^{-1}(\frac{1}{2} - \varepsilon)$ , as desired. Concatenating an outer code of distance  $D$  and an inner code of distance  $d$ , we can obtain a code of distance  $\geq Dd$ , which is design distance. For asymptotically good codes, we have obtained poly time construction of such codes, as well as super explicit construction of similar codes and poly time encoding (since the codes are linear). However, the following natural question about decoding still remains unanswered. Can we correct  $\frac{dD}{2}$  errors in poly time? We will study this question in the next lecture.

## References

- [1] J. Justesen. Class of constructive asymptotically good algebraic codes. *IEEE Trans. Inform. Theory*, pages 652–656, Sep 1972.