

Lecture 29: Achieving capacity of  $BSC_p$

November 6, 2007

Lecturer: Atri Rudra

Scribe: Yang Wang

# 1 Derandomized GMD algorithm

We introduced the GMD algorithm in the last lecture. Notice that we give a randomized version last time and today we will introduce the derandomized version. Obviously we can obtain this by doing an exhaustive search for  $\theta$ . However, there are uncountable choices of  $\theta$  because  $\theta \in [0, 1]$ . But this can be solved by a clever observation.

Define  $Q = \{0, 1\} \cup \{\frac{2w_1}{d}, \dots, \frac{2w_N}{d}\}$ . Then because for each  $i$ ,  $w_i = \min(\Delta(\mathbf{y}'_i, \mathbf{y}_i), d/2)$ , we have

$$Q = \{0, 1\} \cup \{q_1, \dots, q_{\lceil \frac{d}{2} \rceil}\}$$

where  $q_1 < q_2 < \dots < q_{\lceil \frac{d}{2} \rceil}$ . Notice that for every  $\theta \in [q_i, q_{i+1})$ , the step 1 of GMD algorithm output the same  $\mathbf{y}'$ . We change the step 1(c) and step 2 of GMD algorithm as follows:

Step 1: (c) for every  $\theta \in Q$

- (1) if  $\theta < \frac{2w_i}{2}, y_i \leftarrow ?$
- (2) run error and erasures decoding algorithm for  $C_{out}$  and let  $C_\theta$  be the returned codeword.

Step 2: return  $C_\theta$  that is closest to  $\mathbf{y}$ .

As before we can decode half Zyablov bound.

# 2 Achieving capacity of $BSC_p$

The following table summarized the results we have learned:

	Shannon	Hamming Unique decoding	Hamming list decoding
Existence	$1 - H(p)$	$\geq$ GV bound $\leq$ MRRN	$1 - H(p)$
Explicit	?	Zyablov bound	?
Algorithm	?	half Zyablov bound	?

**Question 2.1.** Can we achieve reliable transmission over  $BSC_p$  with explicit codes that have rate of  $1 - H(p) - \epsilon, \epsilon > 0$ ?

We claim that there exist linear codes of rate  $1 - H(p) - \epsilon$  such that decoding error probability is not more than  $2^{-\delta n}, \delta = \Theta\epsilon^2$  realized by concatenated codes with the following property:

- (i)  $C_{out}$ : block length of  $n$  over  $F_{2^b}$ ,  $b = O(\log n)$ .
- (ii)  $C_{in}$ : dimension  $b$ , rate of  $1 - H(p) - \varepsilon/2$  and  $D_{in}$  is a decoding algorithm that runs in  $T_{in}(b)$  time and has decoding error probability no more than  $r/2$  over  $BSC_p$ .

Suppose  $C^* = C_{out} \circ C_{in}$ . Then

$$rate(C^*) = (1 - \frac{\varepsilon}{2})(1 - H(p) - \frac{\varepsilon}{2}) \geq 1 - H(p) - \varepsilon.$$

The decoding algorithm for  $C^*$  works as following. Given  $\mathbf{y} = (y_1, \dots, y_n) \in F_b^n$  received,

Step 1:  $\mathbf{y}'_i = D_{in}(y_i), 1 \leq i \leq n$ .

Step 2: Run  $D_{out}$  on  $\mathbf{y}'_i$ .

Then encoding takes time of  $O(n^2) + O(nb^2) = O(n^2)$ . Decoding takes time  $nT_{in}(b) + T_{out}(n) = n^{O(1)}$  as long as  $T_{out}(n) = n^{O(1)}, T_{in}(b) = 2^{O(b)}$ .