**CS 6815 Pseudorandomness and Combinatorial Constructions** **Fall 2019**

# Lecture 16: October 24th

*Lecturer: Eshan Chattopadhyay* *Scribe: Alexander Frolov*

## 16.1 Efficient List Decoding for Reed-Solomon Codes

Recall, that a code $C \subseteq \Sigma^n$ is $(\rho, L)$-list decodable if $\forall v \in \Sigma^n$, $|Ball(v, \rho_n) \cap C| \leq L$. Also recall the *Johnson Bound*, which states that if code $C$ has relative min distance $1 - \epsilon$, then it has list decoding radius $1 - \sqrt{\epsilon} - o(1)$. Thus the Reed-Solomon (RS) $[n, k, d = n - (k-1)]$ code has list-decoding radius at least $1 - \sqrt{\frac{k}{n}}$.

We will see a simpler version of Sudan's list decoding algorithm for RS codes that works till a relative radius of $1 - \frac{2k}{\sqrt{n}}$.

Recall that $Poly_{\leq m}$ denotes the set of polynomials of degree less than or equal to $m$. The list decoding problem can be framed as follows: given $f : \mathbb{F}_q \to \mathbb{F}_q$, find all $p \in Poly_{\leq k-1}$ such that $Pr_{x \sim \mathbb{F}_q}[f(x) = p(x)] \geq \gamma$, where $\gamma \geq 2k/\sqrt{n}$.

On a high level, the algorithm generalizes the Welch-Berlekamp algorithm.

We use the following algorithm:

Step $a$: Find a nonzero bivariate polynomial $Q(x, y)$ over $\mathbb{F}_q$ satisfying the following properties:

- $deg_x(Q) \leq \sqrt{n}$ ($deg_x(Q)$ is the degree of $Q$ with respect to $x$ (i.e. the degree of $Q$ as a univariate polynomial with $y$ held constant).
- $deg_y(Q) \leq \sqrt{n}$
- $\forall x \in \mathbb{F}_q$, $Q(x, f(x)) = 0$.

Step $b$: Find all factors of $Q(x, y)$ of the form $y - p(x)$. For each such factor, check if $p(x) \in poly_{\leq k-1}$ and $Pr[f(x) = p(x)] \geq \gamma$.
If this check results in $YES$, add $p$ to the output list $L$.

Now, we claim that this algorithm works correctly if $\gamma \geq \frac{2k}{\sqrt{n}}$.

The first part of this proof is showing that there exists a nontrivial $Q$. Write $Q(x, y) = \sum_{i=0,j=0}^{\sqrt{n},\sqrt{n}} \lambda_{ij} x^i y^j$. This equation has $(\sqrt{n}+1)(\sqrt{n}+1) > n$ unknowns, while the constraints on $Q$ determine $q = n$ equations. Thus, there exists nontrivial $Q$, since there are more unknowns than equations defining $Q$. We use the fact that the factor finding step can be done efficiently.

We now need the following result.

**Theorem 16.1** *Bezout's Theorem (special case): Given any two polynomials $A(x, y)$, $B(x, y)$ of degree $d_1$ and $d_2$ respectively, if $|\{(x, y) \in \mathbb{F}_q^2 : A(x, y) = 0, B(x, y) = 0\}| > d_1 d_2$, then both polynomials share a common factor.*

*Consider a codeword $p(x) \in Ball(z, (1-\gamma)n)$. Let the received word be $z = (f(0), f(1), ..., f(q-1))$. By the definition of the ball, $Pr[p(x) = f(x)] \geq \gamma$. Thus, we claim that $y - p(x)$ and $Q(x, y)$ have at least $\gamma n$ common zeros. This is because at each point where $f(x) = p(x)$, $y - p(x) = 0$ at the point $(x, f(x))$. Further, by definition, $Q$ is zero at all points $(x, f(x))$.*

*$deg(y - p(x)) \leq k - 1$, and $deg(Q(x, y)) \leq 2\sqrt{n}$, both by definition of these polynomials. Thus, choosing $\gamma$ such that $\gamma n \geq 2k\sqrt{n}$, $\gamma > \frac{2k}{\sqrt{n}}$, this would lead to $y - p(x)$ and $Q(x, y)$ having a common factor; hence $(y - p(x))|Q(x, y)$. This completes the proof.*

## 16.2 Seeded Extractors from error-correcting codes

***Lemma 16.2*** *Let $C$ be an $[n, k, d]_q$ code, with $d = n(1 - \gamma - \frac{1}{q})$. Define $Ext : \mathbb{F}_q^k \times [n] \to \mathbb{F}_q$, which is $Ext(x, y) = C(x)_y$ (slight notation overloading here, $C$ is both the code and the encoder function). This extractor is a $(log(\frac{1}{\delta}), \sqrt{2q\delta})$ strong seeded extractor.*

*The proof is very similar to that of the Leftover Hash Lemma (proved in Lecture 12). Let $X$ be a weak-source with min-entropy $\log(1/\delta)$ and $Y$ be uniformly distributed on $\mathbb{F}_q$. Let cp denote the collision probability. We have*

$$cp(Y, Ext(X, Y)) \leq cp(Y)(cp(X) + max_{x \neq x', x, x' \in \mathbb{F}_q, y \sim U_m} Pr[Ext(x, y) = Ext(x', y)].$$

*Considering the last quantity, $Pr[C(x)_y = C(x')_y] \leq 1 - \frac{d}{n}$ by the distance of the code and the definition of the extractor. Thus,*

$$cp(Y, Ext(X, Y)) \leq \frac{1}{nq}(1 + 2\gamma q).$$

*Using a lemma from a previous class (Lecture 11), we conclude $|(Y, Ext(X, Y) - (U_{[N]}, U_{\mathbb{F}_q})| \leq \sqrt{2q\delta}$.*