Today

- Locally decodable codes.
- Local decoding of Reed-Muller codes.

Sub-linear time decoding?

- What is the fastest time for decoding one can hope for?
- \( \text{Exp} \rightarrow \text{Poly} \rightarrow \text{Linear} \rightarrow \text{Sublinear} \)?
- “Clearly can’t get last step!” . Don’t have enough time to read input/write output!
- But can if we allow:
  - Implicit representation of input/output.
  - Randomization + low-error probability.

Local Decodability

Defn: \([n, k, d]_q\) Code \( C \) is \((\ell, \varepsilon)\)-locally decodable upto relative error \( \delta \) if there exists an algorithm \( A \) that behaves as follows:

- Takes input \( i \in [n] \).
- Has oracle access to received vector \( r \in \Sigma^n \).
- Tosses some random coins \( \$ \).
- Makes at most \( \ell \) queries to \( r \).
- Soundness: If there exists codeword \( c \in C \) with \( \Delta(r, c) \leq \delta \cdot n \), then \( \Pr_{\$}[A(i) \neq c_i] \leq \varepsilon \).

Will skip \( \varepsilon \) to imply such an \( \varepsilon < 1 - 1/q \) exists.

Complementary Property: Local Testability

- Local Decodability promises decoding if received vector is close to a codeword.
- What if vector not close to a codeword? Do we get to tell? No such guarantee!
- Detecting if close to codeword is a complementary property. We won’t discuss today.
Why local decodability?

- Possibly first interesting sub-linear time algorithm!
- Self-correcting programs and average-case complexity of the permanent.
- Permanent of a matrix.
  - Definition.
  - Complexity.
- Observation: Permanent is a multivariate polynomial. So written as a truth-table, it is a codeword of some enormous Reed-Muller code. If Reed-Muller code is locally decodable, then it implies permanent is hard to compute on random instances.

Local decoding of Hadamard Codes

- For today Hadamard codes will be homogenous polynomials of degree 1 in \( k \) variables. So they are \([2^k, k, 2^{k-1}]_2\) codes.
- Codeword is a function \( f : \mathbb{F}_2^k \to \mathbb{F}_2 \), given by coefficients \( a_1, \ldots, a_k \) and \( f(x) = \sum_i a_i x_i \).
- Local Decoding Question: Given oracle access to \( r : \mathbb{F}_2^k \to \mathbb{F}_2 \) that is \( \delta \)-close to \( f \), and input \( x \in \mathbb{F}_2^k \) can you compute \( f(x) \)?
- Points to be noted:
  - Oracle access is to \( r \), not \( f \).

Local decodability

- Reed-Solomon \([n, k, d]\) code is not \( k \)-locally decodable.
- Proposition: If a linear code is \((\ell, \epsilon)\) locally decodable, then its dual code must have distance less than or equal to \( \ell + 1 \).
- So what kind of codes are locally decodable?
- Hadamard codes? Dual is a Hamming code - so in principle 2-locally decodable.
- Reed-Muller codes? Duals are supposedly also Reed-Muller codes, but only under severe restrictions. In any case have nice

Output needs to be \( f(x) \), not \( r(x) \).
- \( r(x) \) usually equals \( f(x) \), but this probability is over \( x \) - not good enough for defn. of local decoding.
Local decoding algorithm

- Key idea: For codeword $f$, we have $f(x) = f(x + y) - f(y)$ for every $x, y$.
- $f(y)$ usually equals $r(y)$.
- $f(x+y)$ usually equals $r(x+y)$; Prob. only over $y$, not $x$!
- Union bound, bounds probability of either event not happening.

Algorithm & Analysis.

- Algorithm: Given $x$, Pick $y$ at random. Output $r(x + y) - r(y)$.
- Analysis:
  - $\Pr_y[f(y) \neq r(y)] \leq \delta$.
  - $\Pr_y[f(x + y) \neq r(x + y)] \leq \delta$.
  - $\Pr_y[\text{Either of above}] \leq 2\delta$.
  - If $\delta < 1/4$, then answer correct w.p. more than $1/2$.
- Conclude: These Hadamard codes are 2-locally decodable up to nearly half their minimum distance!

Reed-Muller Codes

- What was the basic idea above?
- Restrict attention of code to small dimensional (linear/affine) subspace containing point of interest, and infer value of codeword at the point of interest, based on its value at other points in subspace.
- Hadamard case: Subspace = $\{0, x, y, x + y\}$.
- Reed-Muller Case: Subspace = Lines = $\{x, x + y, x + 2y, \ldots, x + ty, \ldots\}$.

Lines/Small dimensional subspaces in $\mathbb{F}_m$

- Algebraic Property: Low-degree poly restricted to subspace is a low-degree polynomial.
- Randomness Property: Random $t$-dimensional subspace containing $t - 1$ fixed points, is mostly a collection of random points.
Decoding Algorithm

- Problem: Given oracle $r : \mathbb{F}^m \to \mathbb{F}$ s.t. \( \exists f : \mathbb{F}^m \to \mathbb{F} \) of degree $D$ that is $\delta$-close to $r$. Also, given $x$ and $D$. Find $f(x)$.

- Algorithm: Let $\alpha_1, \ldots, \alpha_{D+1} \in \mathbb{F}$ be non-zero and distinct. Pick $y \in \mathbb{F}^m$ at random. Let $y_i = r(x + \alpha_i y)$. Compute univ. degree $D$ poly $p(t) \text{ s.t. } p(\alpha_i) = y_i$. Output $p(0)$.

- Analysis:
  
  - $\Pr_y[ r(x + \alpha_i y) \neq f(x + \alpha_i y) ] = \delta$.
  
  - $\Pr_y[\exists i \text{ s.t. } r(x + \alpha_i y) \neq f(x + \alpha_i y)] \leq (D + 1)\delta$.
  
  - W.p. $1 - (D + 1)\delta$, $p(\cdot) = f|_L(\cdot)$. So $p(0) = f|_L(0) = f(x + 0 \cdot y) = f(x)$.

Some range of parameters

- If $D = \log^c k$ and $m = \Omega(\log k/((c - 1) \log \log k))$, then \# coefficients = $k$.

- Pick field size = $2D$ to get encoding size $n = (2D)^m = k^{c/(c-1)}$ (= poly rate).

- Get $D$-local decodability = poly log $n$.

- Pretty good. Almost best known.

- Error-tolerance not so good. Will do better next time.