Remark: The codewords of $RM(m,r)$ are thus the evaluations of $f: \deg(f) \le r$ at all points of the hypercube $\{0,1\}^m$, taken in some fixed order. A natural ordering of points of the hypercube is the lexicographic ordering.

$000\ldots000 \to 000\ldots001 \to 000\ldots010 \to 000\ldots011 \to \ldots \to 111\ldots111$

The following properties thus hold:

(0) $RM(m,r)$ is a linear code, $\forall r, m$. [Why?]

(i) The blocklength of $RM(m,r)$, for any $r \in [0:m]$, is $n = 2^m$.

(ii) $\dim(RM(m,r)) = \sum_{i=0}^{r} \binom{m}{i}$, since the # monomials of degree i is exactly $\binom{m}{i}$ [Why? How does this give rise to the dimension?]

(iii) $d_{min}(RM(m,r)) = 2^{m-r}$.

We now prove Property (iii) above. Observe that any polynomial $f(X_1,\ldots,X_m) \in \mathbb{F}_2[X_1,\ldots,X_m]$ can be decomposed as:

$$f(X_1,\ldots,X_m) = g(X_1,\ldots,X_{m-1}) + X_m \cdot h(X_1,\ldots,X_{m-1}).$$

Thus, any codeword $\underline{c} = Eval(f)$ (ordered in the lexicographic order) is such that $\underline{c} = (\underline{c}_1 | \underline{c}_2)$, where

$$\underline{c}_1 = Eval(g) \in RM(m-1,r)$$

and $$\underline{c}_2 = Eval(g) + Eval(h), \text{ with } Eval(h) \in RM(m-1,r-1).$$

This observation is called the "Plotkin decomposition of RM codes".

Lemma 2: We have that $\underline{c} \in RM(m,r)$ is such that $\underline{c} = (\underline{u} \mid \underline{u}+\underline{v})$, where $\underline{u} \in RM(m-1, r)$ and $\underline{v} \in RM(m-1, r-1)$.

Proof of Property (iii): We prove this by induction on m and r. The base case ($m=0$ and $r=0$) is easy to establish. Now, assume that the property holds for $\bar{m} = m-1$ and all $r \leq m-1$. We need to show that it holds at $\bar{m} = m$ and at order r.

In particular, we need to show that when $\underline{u} \in RM(m-1, r)$ and $\underline{v} \in RM(m-1, r-1)$, we have $w_H(\underline{u}) + w_H(\underline{u}+\underline{v}) \geq 2^{m-r}$, when $\underline{u}$ & $\underline{v}$ are not both all-zeros.

(a) Suppose that $\underline{v} = \underline{0}$: Clearly, then $w_H(\underline{u}) + w_H(\underline{u}+\underline{v})$

$$= 2 w_H(\underline{u}) \geq 2 \cdot 2^{m-1-r} = 2^{m-r}.$$

(b) Suppose that $\underline{v} \neq \underline{0}$: Then, by the Δ-ineq., we have

$$w_H(\underline{u}) + w_H(\underline{u}+\underline{v}) \geq w_H(\underline{v}) \geq 2^{m-1-(r-1)} = 2^{m-r}.$$

Further, there exists a codeword $\underline{c} = \text{Eval}\left(\prod_{i=1}^{r} x_i\right)$ that has Hamming weight exactly 2^{m-r}. ☒

Lecture: Decoding RM Codes and List Decoding

Last time, we discussed the RM family of codes and its basic properties. We shall now take a look at a decoder for RM codes that can correct up to $\left\lfloor \frac{d_{min}-1}{2} \right\rfloor$ errors – Reed's decoder.

We first make some observations about the structure of RM codewords. For a given subset $S \subseteq [m]$, let $\bar{S}$ denote $[m] \setminus S$ and let

$$V_S \triangleq \{ \underline{z} \in \{0,1\}^m : z_i = 0, \forall i \in \bar{S} \}$$

be the vector space [Why?] that consists of points (or coordinates) $\underline{z}$ that are 0 outside S. Clearly, $\dim(V_S) = |S|$ and there are hence $2^{m-|S|}$ cosets $(V_S + \underline{b})$ of V_S.

Now, note that for any $S \subseteq [m]$ and any $\underline{b} \in \{0,1\}^m$, we always have

$$\sum_{\underline{z} \in V_S + \underline{b}} \text{Eval}_{\underline{z}}(x_S) = 1 \quad . \quad \text{——} \quad (1).$$

To see why (1) is true, note that $\text{Eval}_{\underline{z}}(x_S) = 1$ iff $z_i = 1, \forall i \in S$, which happens at exactly one point $\underline{z} \in V_S + \underline{b}$.

Furthermore, we have for all $T \not\subseteq S$,

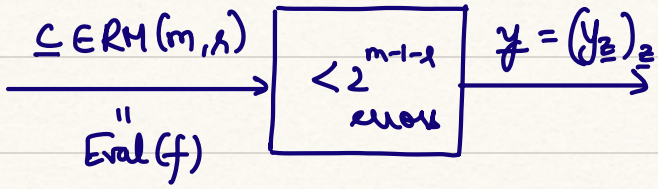
$$\sum_{\underline{z} \in V_S + \underline{b}} \text{Eval}_{\underline{z}}(x_T) = 0 \quad . \quad \text{——} \quad (2)$$

To see (2), note that since $T \not\supseteq S$, there exists $i \in S \setminus T$, with $z_i = 0$ or $z_i = 1$ not affecting $\text{Eval}_{\underline{z}}(x_T)$.

Let $f(x_1, \dots, x_m) = \sum_{S \subseteq [m] : |S| \le r} a_S x_S$ be s.t. $\text{Eval}(f) \in RM(m, r)$.

Reed's algorithm uses properties (1) & (2) to first decode a_S, for S s.t. $|S| = r$, and then $a_{S'}$, for $|S'| = r-1$, and so on.

Let $y = (y_{\underline{z}} : \underline{z} \in \{0,1\}^m)$ be the received sequence when $\underline{c} = \text{Eval}(f)$ is corrupted by $< \frac{d_{min}}{2} = 2^{m-1-r}$ errors.



Claim: We have that $\underset{\underline{b}}{\text{MAJ}}\left(\sum_{\underline{z} \in V_S + \underline{b}} y_{\underline{z}}\right) = a_S$, for $|S| = r$.

In other words, taking a majority vote among $\left(\sum_{\underline{z} \in V_S + \underline{b}} y_{\underline{z}} : \underline{b} \in \{0,1\}^m\right)$ yields a_S.

Proof: Suppose first that $y = \text{Eval}(f)$. In this case, we have that for any $\underline{b} \in \{0,1\}^m$,

$$\sum_{\underline{z} \in V_S + \underline{b}} y_{\underline{z}} = \sum_{\underline{z} \in V_S + \underline{b}} \text{Eval}_{\underline{z}}(f) = \sum_{T : |T| \le r} a_T \sum_{\underline{z} \in V_S + \underline{b}} \text{Eval}_{\underline{z}}(x_T)$$

Note that from (1) and (2), we have $\sum_{\underline{z} \in V_S + \underline{b}} \text{Eval}_{\underline{z}}(x_T) = 1$ iff $S = T$.

Hence, $\sum_{\underline{z} \in V_S + \underline{b}} y_{\underline{z}} = a_S$, $\forall\, \underline{b} \in \{0,1\}^m$.

Now, if there are at most 2^{m-r-1} errors in $\underline{y}$, then at most 2^{m-1-r} cosets $V_S + \underline{b}$ will be such that

$$\sum_{\underline{z} \in V_S + \underline{b}} y_{\underline{z}} \neq a_S. \quad \text{[Why ?]}$$

Hence, taking a majority vote $\underset{\underline{b}}{\text{MAJ}}\left(\sum_{\underline{z} \in V_S + \underline{b}} y_{\underline{z}}\right)$ yields a_S. ☑

Now, after decoding a_S, for all S s.t. $|S| = r$, Reed's algorithm obtains

$$\underline{y} \leftarrow \underline{y} - \text{Eval}\left(\sum_{S : |S| = r} a_S x_S\right),$$

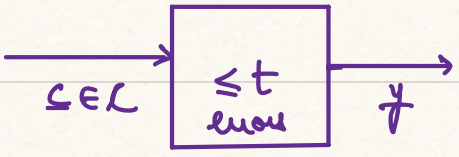
yielding a new "noisy" sequence $\underline{y}$ that can be seen as a corrupted version of a codeword $\underline{c}' \in RM(m, r-1)$. The algorithm then applies the same procedure as outlined in Claim 1 above to decode a_S, for S s.t. $|S| = r-1$, and so on.

Remark: Each step of majority voting is carried out on at most $2^{m-r} \leq n$ cosets, and there are at most $\binom{m}{\leq m/2} \leq 2^m = n$ monomials

of any fixed degree. Thus, Reed's algorithm works in poly(n) time.

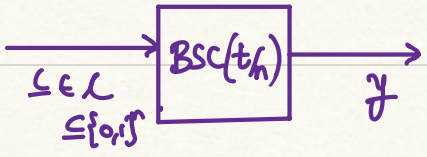
Further addenda:

We studied the family of RS codes, which we learnt were good since they achieve the Singleton bound [i.e., are Maximum Distance Separable (MDS) codes]. In other words, RS codes have the highest error-correcting capability, for errors introduced in an adversarial manner, i.e., when there is a bound $t \leq \left\lfloor \frac{n-k}{2} \right\rfloor$ on the # of errors a channel can introduce.



"Adversarial"/"Hamming" model of errors

In contrast, we have the stochastic/information-theoretic/Shannon model for errors:



Note that, "on average", one expects to see $\approx t$ errors in y [Why?]. RM codes, on the other hand, are provably good for the above stochastic noise channel, i.e., RM codes achieve the capacity [Reeves-Pfister (2023), Abbe-Sandon (2023)] of the channel above, when $\lim_n \frac{t}{n} = \tau \in (0,1)$.

Note however that while the relative distance δ of RS codes of positive rate $R<1$ positive (if $R = \lim_n \frac{k}{n} \in (0,1)$, then $\delta = \lim_n \frac{d}{n} = \lim_n \frac{n-k+1}{n} > 0$), the relative distance of RM codes of positive rate is asymptotically zero!

To see why, one must know that the rate

$$R(RM(m,r)) = \binom{m}{\leq r} / 2^m ,$$

with $\lim_{m\to\infty} R(RM(m,r_m)) > 0$ iff $r_m \approx \frac{m}{2} \pm o(m)$.

For such orders $(r_m)_{m\geq 1}$, we have $\delta = \lim_{m\to\infty} \frac{2^{m-r_m}}{2^m} = 0$.

Hence, it is NOT necessary for a code to have good performance over adversarial noise channels, for it to perform well over stochastic noise channels.

List decoding: An introduction

As seen above, there appears to be a "gap" between the fundamental limits of the adversarial and stochastic noise channels. More precisely, we make the following observations:

(i) [Adversarial model] $\lim_n \frac{\#\text{ errors tolerated}}{n} \overbrace{\leq}^{\triangleq \tau} \frac{1-R}{2}$. [Why?]

$$\equiv \quad R \leq 1 - 2\tau$$

(ii) [Shannon model] $R \leq 1 - h_b(\tau)$ [Recall Shannon's Theorem!]

$\approx 1 - \tau$, for small τ.

The goal of list decoding is to serve as a bridge between the above fundamental limits, via the introduction of a new decoding paradigm, where the decoder is allowed to output a (small) list of candidate codewords. The decoder makes an error if the true input codeword is not present in this list.