An Analysis of RPA Decoding of Reed-Muller Codes Over the BSC

Lalitha Vadlamani, IIIT Hyderabad

HDX Workshop, ICTS-TIFR

5th May, 2025

Joint Work with

Dr. V. Arvind Rameshwar,
Research Fellow,
India Urban Data Exchange,
Indian Institute of Science

Reed-Muller Codes

- For $0 \leq r \leq m$, the r^{th}-order binary Reed-Muller code $\mathrm{RM}(m, r)$ is defined as

$$\mathrm{RM}(m, r) := \{\mathrm{Eval}(f) : f \in \mathbb{F}_2[x_1, x_2, \ldots, x_m],\ \deg(f) \leq r\},$$

  where $\deg(f)$ is the largest degree of a monomial in f, and the degree of a monomial $\prod_{j \in S : S \subseteq [m]} x_j$ is simply $|S|$.

- Length $N = 2^m$ and dimension $\binom{m}{\leq r} := \sum_{i=0}^{r} \binom{m}{i}$.

- Minimum Hamming distance $d_{\min}(\mathrm{RM}(m, r)) = 2^{m-r}$.

Recent Capacity Results

- [Kudekar etal 17] RM codes are capacity-achieving for the binary erasure channel
- [Reeves-Pfister23, Abbe-Sandon23] RM codes are capacity-achieving for BMS channels
- [Arikan08] Close cousins of polar codes

Decoding Algorithms

- [Reed54] Reed's Majority logic decoder
- [Sidel-Persha92, Sakkour05] Sidel'nikov and Pesharekov, Sakkour for second order RM codes
- [Dumer04, 06] Dumer's Recursive Decoding based on Plotkin decomposition
- [Ye-Abbe20] Recursive Projection Aggregation (RPA) Decoding

RPA Decoding - Projection Step

- Binary symmetric channel with cross-over probability p
- $\{\mathbb{B}_i \subseteq \{0,1\}^m\}$ - collection of all k dimensional subspaces
- $Y_{/\mathbb{B}_i}(T) := \bigoplus_{\mathbf{b} \in \mathbb{B}_i} Y_{\mathbf{x} \oplus \mathbf{b}}$, T is the coset containing $\mathbf{x}$
- Projection of $\mathbf{Y}$ onto the cosets of $\mathbb{B}_i$ as

$$\mathbf{Y}_{/\mathbb{B}_i} := \left(Y_{/\mathbb{B}_i}(T) : \; T \in \{0,1\}^m / \mathbb{B}_i \right),$$

  for some fixed ordering among cosets T.
- Projecting the codewords gives

$$\mathbf{c}_{/\mathbb{B}_i} := \left(c_{/\mathbb{B}_i}(T) : \; T \in \{0,1\}^m / \mathbb{B}_i \right),$$

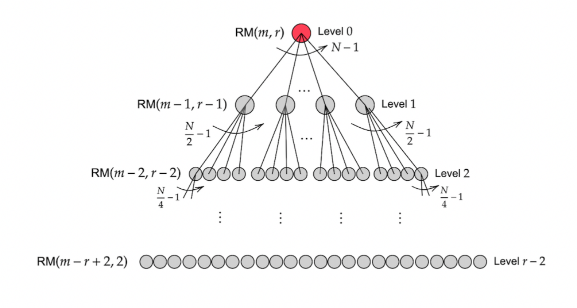
- $\mathbf{c}_{/\mathbb{B}_i} \in \mathrm{RM}(m-k, r-k)$, if $\mathbf{c} \in \mathrm{RM}(m, r)$.

RPA Decoding - Aggregation Step

For $\mathbf{x} \in \{0,1\}^m$

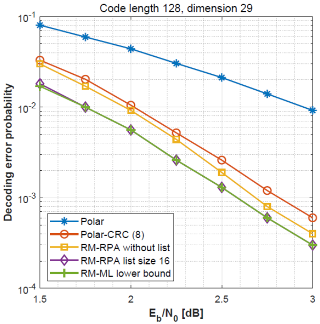
- Compute $\phi(\mathbf{x}) = \sum_{i=1}^{n_{k,m}-1} 1\{Y_{/\mathbb{B}_i}([\mathbf{x}+\mathbb{B}_i]) \neq \widehat{Y}_{/\mathbb{B}_i}([\mathbf{x}+\mathbb{B}_i])\}$
- Set $\text{Flip}(\mathbf{x}) = 1$, if $\phi(\mathbf{x}) > \frac{n_{k,m}}{2}$
- $\bar{Y} = Y \oplus \text{Flip}$

Projection Aggregation Tree

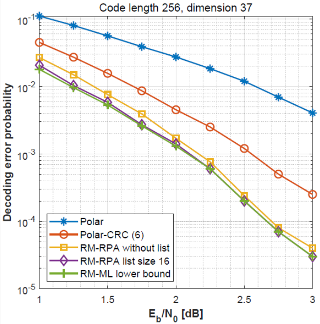


- Rooted tree, with the root (at level 0) being the code RM(m, r).
- Each node at level i has $\frac{N}{2^i} - 1$ children at level $i+1$, each of which is an RM$(m-i-1, r-i-1)$ code; here, $N := 2^m$.
- "Moving down" from a parent node to a child node corresponds to projection step and "moving up" corresponds to aggregation step.

Close to ML Performance



(a) $\mathcal{RM}(7,2)$ v.s. polar codes



(d) $\mathcal{RM}(8,2)$ v.s. polar codes

- Empirical performance of RPA decoder is close to ML bound for low values of r.

Image Courtesy: [Ye-Abbe20]

Main Result

Let $\overline{p} := \frac{1}{2} \cdot (1 - (1 - 2p)^{2^{r-2}})$ and $\eta(\overline{p}) := \frac{1}{2} \cdot (1 - 4\overline{p}(1 - \overline{p}))$.

Theorem

For any $0 < \epsilon < \eta(\overline{p})$, we have that for $r \geq 2$, using one-dimensional subspaces for projection,

$$P_{\text{err}}(\text{RM}(m, r)) \leq 32N^{r+1} \cdot \exp\left(-2^{-r-1}N\epsilon^2\right).$$

Let $c = c(p) := \frac{\log 2}{\log\left(\frac{1}{1-2p}\right)}$.

Corollary

For any $0 < \overline{c} < c$, we have that for all $r \leq \log(\overline{c}m)$,

$$\lim_{m \to \infty} P_{\text{err}}(\text{RM}(m, r)) = 0.$$

Analysis of Second Order RM Codes

- Assume that all-zero codeword is transmitted

Key Ingredients:

- Analyzing the FHT Decoder
- Analyzing the aggregation step

Base Case - FHT Decoder

- $\mathbf{Y} \sim \text{Ber}^{\otimes N}(p)$ implies $\mathbf{Y}_{/\mathbb{B}_i} \sim \text{Ber}^{\otimes(N/2)}(2p(1-p))$.
- $\mathbf{Y}_{/\mathbb{B}_i} \in \{0,1\}^{N/2}$ mapped to $\mathbf{Y}^{\pm}_{/\mathbb{B}_i} \in \{-1,1\}^{N/2}$
- Consider the family of functions $\left(\chi_{\mathbf{s}} : \mathbf{s} \in \{0,1\}^{m-1}\right)$, where $\chi_{\mathbf{s}}(\mathbf{x}) := (-1)^{\mathbf{x}\cdot\mathbf{s}}$, $\mathbf{x} \in \{0,1\}^{m-1}$.
- One-one correspondence between codewords of $\text{RM}(m-1,1)$ and the collection of vectors $\chi := \left(\chi_{\mathbf{s}} : \mathbf{s} \in \{0,1\}^{m-1}\right) \cup \left(-\chi_{\mathbf{s}} : \mathbf{s} \in \{0,1\}^{m-1}\right)$.
- ML decoder for RM(m-1,1) given by

$$\text{ML}(\mathbf{Y}^{\pm}_{/\mathbb{B}_i}) = \underset{\mathbf{s}\in\{0,1\}^{m-1},\ \sigma\in\{-1,1\}}{\arg\max} \langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, \sigma \cdot \chi_{\mathbf{s}} \rangle.$$

Analysis of FHT Decoder

- The all-zeros codeword $\mathbf{0} \in \mathrm{RM}(m-1,1)$ corresponds to $+\chi_{\mathbf{0}}$.

Concentration of the term with $+\chi_{\mathbf{0}}$:

Lemma

For all $\epsilon > 0$ and any $i \in [N-1]$, we have that

$$\Pr\left[\left|\langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, \chi_{\mathbf{0}} \rangle - (1 - 4p(1-p))\right| \leq \epsilon\right] \geq 1 - 2e^{-\frac{N\epsilon^2}{4}}.$$

Follows by concentration around mean using Hoeffding's Inequality

Concentration of the term with $\chi_{\mathbf{s}}$:

Lemma

For all $\epsilon > 0$ and any $i \in [N-1]$, we have that for any $\mathbf{s} \neq \mathbf{0}$,

$$\Pr\left[|\langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, \chi_{\mathbf{s}} \rangle| \leq \epsilon\right] \geq 1 - 4e^{-\frac{N\epsilon^2}{8}}.$$

$$\langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, \chi_{\mathbf{s}} \rangle = \frac{2}{N} \cdot \left[\sum_{j=1}^{N/4} Z'_j - \sum_{k=1}^{N/4} Z''_k\right],$$

$$\Pr\left[|\langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, \chi_{\mathbf{s}} \rangle| \leq \epsilon\right] = \Pr[|\alpha_1 - \alpha_2| \leq \epsilon]$$

$$\geq \Pr\left[\alpha_1 \in [\beta - \epsilon/2, \beta + \epsilon/2] \text{ and } \alpha_2 \in [\beta - \epsilon/2, \beta + \epsilon/2]\right]$$

$$\geq \left(1 - 2e^{-\frac{N\epsilon^2}{8}}\right)^2 \geq 1 - 4e^{-\frac{N\epsilon^2}{8}},$$

Combining the Cases

By union bound that with probability at least $1 - 8N \cdot e^{-\frac{N\epsilon^2}{8}}$, the following events occur:

- $\langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, \chi_{\mathbf{0}} \rangle \in [(1 - 4p(1-p)) - \epsilon, (1 - 4p(1-p)) + \epsilon]$
- $\langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, -\chi_{\mathbf{0}} \rangle \in [-(1 - 4p(1-p)) - \epsilon, -(1 - 4p(1-p)) + \epsilon]$
- For all $\mathbf{s} \neq \mathbf{0}$, we have $\langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, \chi_{\mathbf{s}} \rangle \in [-\epsilon, \epsilon]$ and $\langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, -\chi_{\mathbf{s}} \rangle \in [-\epsilon, \epsilon]$.

If $\epsilon < \eta(p) = \frac{1}{2}(1 - 4p(1-p))$

- With probability at least $1 - 8N \cdot e^{-\frac{N\epsilon^2}{8}}$,

$$\langle \max_{f \in \chi} \langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, f \rangle = \langle \mathbf{Y}^{\pm}_{/\mathbb{B}_i}, \chi_{\mathbf{0}} \rangle.$$

[Burnshev-Dumer06] have analysis of ML decoding of first order RM codes for BSC. Our approach is very different. The exponents match.

Analysis of the Aggregation Step

- Let $\mathcal{G}$ denote the event $\left\{\widehat{\mathbf{Y}}_{/\mathbb{B}_i} = \mathbf{0}, \text{for all } i \in [N-1]\right\}$.
- $\Pr[\mathcal{G}] \geq 1 - 8N(N-1) \cdot e^{-\frac{N\epsilon^2}{8}}$.

Upon conditioning on $\mathcal{G}$:

$\sum_{i=1}^{N-1} 1\{Y_{/\mathbb{B}_i}([\mathbf{x} + \mathbb{B}_i]) \neq \widehat{Y}_{/\mathbb{B}_i}([\mathbf{x} + \mathbb{B}_i])\}$ reduces to

$$\phi(\mathbf{x}) = \sum_{i=1}^{N-1} \left(Y_{\mathbf{x}} \oplus Y_{\mathbf{x}+\mathbf{b}_i}\right),$$

- Rewriting $\phi(\mathbf{x})$ as

$$\phi(\mathbf{x}) = \begin{cases} \sum_{\mathbf{z}\neq\mathbf{x}} Y_{\mathbf{z}}, & \text{if } Y_{\mathbf{x}} = 0, \\ N - 1 - \sum_{\mathbf{z}\neq\mathbf{x}} Y_{\mathbf{z}}, & \text{if } Y_{\mathbf{x}} = 1. \end{cases}$$

- $\overline{\phi}(\mathbf{x}) = \frac{\phi(\mathbf{x})}{N-1}$ concentrates around its mean $\overline{\phi}_\infty(\mathbf{x}) := p(1 - Y_{\mathbf{x}}) + (1-p)Y_{\mathbf{x}}$

Analysis of Aggregation Step

- If $\overline{\phi}(\mathbf{x})$ is close to $\overline{\phi}_\infty(\mathbf{x})$, then their indicators $1\{\overline{\phi}(\mathbf{x}) > \frac{1}{2}\}$ and $1\{\overline{\phi}_\infty(\mathbf{x}) > \frac{1}{2}\}$ are close.
- Only upon conditioning on $\mathcal{G}$:, $\text{Flip}^{(N)}(\mathbf{x}) = 1\{\overline{\phi}(\mathbf{x}) > \frac{1}{2}\}$
- $1\{\overline{\phi}(\mathbf{x}) > \frac{1}{2}\}$ and $1\{\overline{\phi}_\infty(\mathbf{x}) > \frac{1}{2}\}$ are close even after conditioning on $\mathcal{G}$ because $\mathcal{G}$ is a very high probability event.

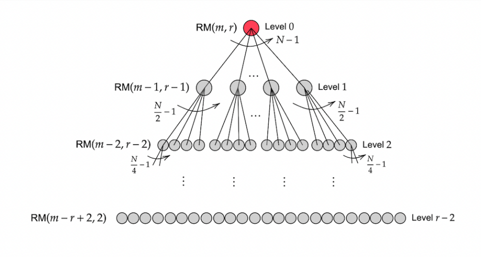
$$\Pr\left[\text{Flip}^{(N)}(\mathbf{x}) = 1\left\{\overline{\phi}_\infty(\mathbf{x}) > \frac{1}{2}\right\} \;\middle|\; \mathcal{G}\right]$$
$$\geq 1 - 16N(N-1) \cdot e^{-\frac{N\epsilon^2}{8}}.$$

Combining FHT Analysis and Aggregation Analysis

$$\Pr\left[\text{Flip}^{(N)}(\mathbf{x}) = 1\left\{\overline{\phi}_\infty(\mathbf{x}) > \frac{1}{2}\right\},\ \text{for all } \mathbf{x} \,\middle|\, \mathcal{G}\right]$$
$$\geq 1 - 16N^2(N-1)\cdot e^{-\frac{N\epsilon^2}{8}}.$$

$$\Pr\left[\text{Flip}^{(N)} = \mathbf{Y}\right]$$
$$\geq \Pr\left[\text{Flip}^{(N)} = \mathbf{Y} \,\middle|\, \mathcal{G}\right]\cdot \Pr[\mathcal{G}]$$
$$\geq \left(1 - 16N^2(N-1)\cdot e^{-\frac{N\epsilon^2}{8}}\right)\left(1 - 8N(N-1)\cdot e^{-\frac{N\epsilon^2}{8}}\right)$$
$$\geq 1 - 32N^3\cdot e^{-\frac{N\epsilon^2}{8}}.$$

Analysis of Higher Order RM Codes



- Analysis of FHT decoder required only at the leaf nodes
- Analysis of the aggregation step is done at all levels other than the last one

Theorem

For any $0 < \epsilon < \eta(\overline{p})$, we have that for $r \geq 2$, using one-dimensional subspaces for projection,

$$P_{\text{err}}(\text{RM}(m,r)) \leq 32N^{r+1} \cdot \exp\left(-2^{-r-1}N\epsilon^2\right).$$

Analysis of Higher Dimensional Projections

- Aggregation step is different
- $\phi(\mathbf{x}) = \sum_{i=1}^{\tilde{n}} \bigoplus_{\mathbf{b}\in\mathbb{B}_i} \tilde{Y}_{\mathbf{x}\oplus\mathbf{b}}$

$$\phi(\mathbf{x}) = \begin{cases} \sum_{i=1}^{\tilde{n}} \bigoplus\limits_{\mathbf{b}\in[\mathbf{x}+\mathbb{B}_i],\mathbf{b}\neq\mathbf{x}} Y_{\mathbf{b}}, \text{ if } Y_{\mathbf{x}} = 0, \\ \tilde{n} - \sum_{i=1}^{\tilde{n}} \bigoplus\limits_{\mathbf{b}\in[\mathbf{x}+\mathbb{B}_i],\mathbf{b}\neq\mathbf{x}} Y_{\mathbf{b}}, \text{ if } Y_{\mathbf{x}} = 1. \end{cases}$$

Unlike one dimensional projections, $\phi(x)$ is not sum of independent random variables. Hoeffding's Inequality cannot be applied.

Analysis of Higher Dimensional Projections

A more general concentration inequality:

[Raginsky-Sason18, Thm. 3.4.4]

Let $X_1, \dots, X_n$ be i.i.d. $\mathrm{Ber}(p)$ random variables. Then, for every Lipschitz function $f : \{0,1\}^n \to \mathbb{R}$ with Lipschitz constant c_f, we have for all $\alpha > 0$,

$$\Pr\left[f(X^n) - E[f(X^n)] > \alpha\right] \leq \exp\left(-\ln\left(\frac{1-p}{p}\right) \cdot \frac{\alpha^2}{nc_f^2 \cdot (1-2p)}\right).$$

Analysis of Higher Dimensional Projections

Lipschitz constant for $\phi(x)$ is

$$n_{k-1,m-1} = \begin{bmatrix} m-1 \\ k-1 \end{bmatrix} := \prod_{i=0}^{k-2} \frac{2^{m-1} - 2^i}{2^{k-1} - 2^i}.$$

Theorem

After a single iteration of the RPA decoder at the last level, for all $\epsilon < \eta(\bar{p})$, $\Pr[\overline{\mathbf{Y}} = \mathbf{0}] \geq 1 - 32 n_{k,m} \tilde{N}^2 \cdot \exp\left(-\ln\left(\frac{1-\bar{p}}{\bar{p}}\right) \cdot 2^{-2k-2} \tilde{N} \epsilon^2\right)$

- Analysis can be repeated on a projection aggregation tree with k-dimensional projection in each step
- The exponent doesn't improve compared to the one-dimensional projection

Open Problems

- Analysis by picking smaller number of subspaces is straightforward
- Analysis of RPA for general BMS channels
- Bounds for probability of error in the regime when r is growing increasing faster with m
- Performance of RPA with list decoding

Thanks!

`https://arxiv.org/abs/2412.08129`