On the Error Probability of RPA Decoding of Reed-Muller Codes over BMS Channels

Dorsa Fathollahi
Department of Electrical Engineering
Stanford University, CA, USA
email: dorsafth@stanford.edu

V. Arvind Rameshwar
Department of Electrical Engineering
IIT Madras, India
email: arvind@ee.iitm.ac.in

V. Lalitha
SPCRC
IIIT Hyderabad, India
email: lalitha.v@iiit.ac.in

***Abstract*—We analyze the performance of the Recursive Projection-Aggregation (RPA) decoder of Ye and Abbe (2020), for Reed-Muller (RM) codes, over general binary memoryless symmetric (BMS) channels. Our work is a significant generalization of a recent result of Rameshwar and Lalitha (2025) that showed that the RPA decoder provably achieves vanishing error probabilities for "low-rate" RM codes, over the binary symmetric channel (BSC). While a straightforward generalization of the proof strategy in that paper will require additional, restrictive assumptions on the BMS channel, our technique, which employs an equivalence between the RPA projection operation and a part of the "channel combining" phase in polar codes, requires no such assumptions. Interestingly, such an equivalence allows for the use of a generic union bound on the error probability of the first-order RM code (the "base case" of the RPA decoder), under maximum-likelihood decoding, which holds for any BMS channel. We then exploit these observations in the proof strategy outlined in the work of Rameshwar and Lalitha (2025), and argue that, much like in the case of the BSC, one can obtain vanishing error probabilities, in the large n limit (where n is the blocklength), for RM orders that scale roughly as $\log \log n$, for all BMS channels.**

I. Introduction

Reed-Muller (RM) codes are a well-studied family of binary linear codes that are obtained by the evaluations of Boolean polynomials on the points of the Boolean hypercube [1], [2]. Recent breakthrough theoretical progress that has shown that RM codes are capacity-achieving for the binary erasure channel [3], and more generally, for BMS channels [4], [5] (we refer the reader to the survey [6] for a detailed treatment of the properties of RM codes).

Over the past few decades, much work has been dedicated to finding practical (low-complexity) decoding algorithms for RM codes. The earliest such algorithm by Reed [1] is capable of correcting bit-flip errors up to half the minimum distance of the code. For the case of first-order RM codes, a Fast Hadamard Transform-based (or FHT-based) decoder was designed in [7], [8], which is an efficient implementation of a maximum likelihood (ML) decoding procedure. We also refer the reader to decoding algorithms for RM codes of higher code that display good performance at moderate blocklengths in the works [9]–[14].

A much more recent decoding algorithm, variants of which were shown to achieve near-ML performance at moderate blocklengths over the binary symmetric channel (BSC), is the Recursive Projection-Aggregation (RPA) decoder of Ye and Abbe [15]. Later works [16], [17] presented procedures for reducing the complexity of the RPA decoder, by making use of a subset of the subspaces employed by the RPA decoder, for projection. More recent work [18], however, argues that limiting the number of subspaces used could suffer from significant performance loss, due to the presence of "coset error patterns". In this work, we hence work with the original RPA decoder of Ye and Abbe, which makes use of all subspaces of a fixed dimension, for projection.

Our main efforts in this paper are directed towards obtaining *analytical* performance guarantees for RPA decoding, via explicit bounds on the probability of error, over general binary memoryless symmetric (BMS) channels. Recent work [19] has obtained theoretical upper bounds on this error probability for the case when the channel is a binary symmetric channel (BSC). In this work, we generalize the results in [19] to the broad class of BMS channels, of which the BSC is a part. We mention however that such an extension does not follow straightforwardly from the proof strategy of [19] – indeed, while somewhat direct modifications can be carried out by placing additional restrictions on the channel (such as requiring a bounded output alphabet size and bounded log-likelihood ratios, for all outputs) – the problem of deriving general bounds for *arbitrary* BMS channels requires different techniques.

Our main result (Theorem 4) shows that RPA decoding for general BMS channels guarantees vanishing error for RM orders roughly logarithmically in the parameter $m = \log n$, where n is the code blocklength – a scaling that is asymptotically identical to that obtained in [19]. Our proof proceeds via the identification of an equivalence between the projection operation in RPA decoding and a certain channel combining operation in polar code construction, which allows for the use of a standard union bound on the ML error probability of first-order RM codes, which form the "base case" of the RPA decoder. Crucially, our analysis retains an important element of the proof strategy of [19], which restricts attention to the case when one iteration of the RPA decoder suffices for convergence. It hence appears that relaxing this restriction and performing an analysis of RPA decoding in the presence of the correlations introduced via multiple iterations is key to obtaining asymptotic improvements in error probability.

II. Notation and Preliminaries

A. Notation

Random variables are denoted by capital letters, e.g., X, Y, and small letters, e.g., x, y, denote their instantiations. Log-likelihood vectors are however denoted as L, following standard notation.The notation $\mathbf{0}$ denotes the all-zeros vector, whose length can be inferred from the context. Natural logarithms are denoted as $\ln$. The notations $O(\cdot), o(\cdot), \Omega(\cdot), \omega(\cdot)$ are used to refer to members of the standard Bachmann–Landau family of asymptotic notations. The indicator function $\mathbb{1}[\cdot]$ takes the value 1 when the argument is true, and 0, otherwise.

B. Reed-Muller Codes

Consider the polynomial ring $\mathbb{F}_2[x_1, x_2, \ldots, x_m]$ in m variables. For a polynomial $f \in \mathbb{F}_2[x_1, x_2, \ldots, x_m]$ and a binary vector $\mathbf{z} = (z_1, \ldots, z_m) \in \mathbb{F}_2^m$, we write $f(\mathbf{z}) = f(z_1, \ldots, z_m)$ as the evaluation of f at $\mathbf{z}$. Let $\mathbb{F}_2^{\leq r}[x_1, x_2, \ldots, x_m]$ denote the collection of polynomials of degree at most r. The evaluation points are ordered according to the standard lexicographic order on strings in $\mathbb{F}_2^m$, i.e., if $\mathbf{z} = (z_1, \ldots, z_m)$ and $\mathbf{z}' = (z'_1, \ldots, z'_m)$ are two evaluation points, then, $\mathbf{z}$ occurs before $\mathbf{z}'$ iff for some $i \geq 1$, we have $z_j = z'_j$ for all $j < i$, and $z_i < z'_i$. Now, let $\text{Eval}(f) := (f(\mathbf{z}) : \mathbf{z} \in \mathbb{F}_2^m)$ be the evaluation vector of f, where the coordinates $\mathbf{z}$ are ordered according to the standard lexicographic order.

Definition 1 (see Ch. 13 in [20], or [21])**.** *For $0 \leq r \leq m$, the r^{th}-order binary Reed-Muller code RM(m, r) is defined as*

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

All through, we set $n := 2^m$.

C. BMS Channels and RPA Decoding

We refer the reader to [22, Ch. 4] for the definition of binary memoryless symmetric (BMS) channels.

Definition 2. *The Bhattacharyya parameter $Z(W)$ of a BMS channel W with output alphabet $\mathcal{Y}$ is defined as*

$$Z(W) := \sum_{y \in \mathcal{Y}} \sqrt{P_{Y|0}(z|0) P_{Y|1}(z|1)},$$

when $\mathcal{Y}$ is finite, with the conditional p.m.f.s above replaced by corresponding probability densities and the summation above replaced by an integral, when the channel laws correspond to continuous probability distributions.

The RPA decoding algorithm for general BMS channels is briefly recapitulated as Algorithm 1 (see also [15, Alg. 3])[1]. In all that follows in this paper, we assume that the RPA decoder uses one-dimensional subspaces for projection.

In Algorithm 1, L denotes the vector of log-likelihood ratios (LLRs) with $L = (L(z) : \ z \in \{0,1\}^m)$, where $L(z) := \log \frac{W(y_z|0)}{W(y_z|1)}$, and $\mathbf{y} = (y_z : \ z \in \{0,1\}^m)$ is the received vector at the end of the BMS channel. In Step 2, we use

$$\begin{aligned} L^{(v)}(T) := \log \big(&\exp(L(z) + L(z \oplus v)) + 1 \big) \\ &- \log \big(\exp(L(z)) + \exp(L(z \oplus v)) \big). \end{aligned}$$

In Step 4, the subroutine FHTDecoder refers to the standard Fast Hadamard Transform decoder [7], [8] for ML decoding of first-order RM codes. Furthermore, in this paper, we do not set a "tolerance threshold" θ as in the original work [15], and instead only run the decoder for a fixed number N_{max} of iterations. However, for the purpose of analysis, as in [19], we restrict attention to the setting where *one* iteration of RPA decoding suffices for returning the correct decoded estimate.

Given the code RM(m, r), let C denote a (random) codeword that is drawn uniformly at random from the code. Now, let $\mathbf{Y}$ be the received sequence at the end of the BMS channel W and let $\widehat{C}$ denote the estimate of the input codeword obtained by using the RPA decoder in [15, Alg. 3]. The probability of error of RM(m, r) under RPA decoding is then defined as $P_{\text{err}}(\text{RM}(m, r)) := \Pr\left[\widehat{C} \neq C\right]$.

[1] Algorithm 3 of [15] in fact is stated for general binary-input memoryless channels that are not necessarily symmetric.

Algorithm 1 RPA Decoder for $\mathrm{RM}(m, r)$

```
Input: LLR vector $L \in \mathbb{R}^n$, max iter. $N_{\max}$
Output: $\widehat{c} \in \{0,1\}^n$.
for $j = 1$ to $N_{\max}$ do
    Projection: For each $v \in \mathbb{F}_2^m \backslash \{0\}$ and each coset $T = \{z, z \oplus v\}$, compute $L^{(v)}(T)$.
    if $r \geq 2$ then Compute $\widehat{y}^{(v)} \leftarrow \mathrm{RPA}(m-1, r-1, L^{(v)})$.
    else Set $\widehat{c} \leftarrow \mathrm{FHTDecoder}(L)$ break
    Aggregation: For each $z \in \mathbb{F}_2^m$, compute
        $\widehat{L}(z) \leftarrow \frac{1}{2^m - 1} \sum_{v \neq 0} \Big(1 - 2\,\widehat{y}^{(v)}([z + \langle v \rangle])\Big) L(z \oplus v).$
    Set $L \leftarrow \widehat{L}$.
Set $\widehat{c}(z) \leftarrow \mathbf{1}[L(z) < 0]$ for all $z$; return $\widehat{c}$.
```

Definition 3 (RPA Recursion Tree)**.** *The RPA decoding of* $\mathrm{RM}(m, r)$ *induces a recursion tree* $\mathcal{T}$ *with the following structure:*

- *The tree has depth r, with orders indexed by $i \in \{1, 2, \ldots, r\}$.*
- *Nodes are uniquely identified by a pair $v = (i, j)$ where:*
  - *$i \in \{1, 2, \ldots, r\}$ is the* height *of the node, where the root is at height r*
  - *$j \in \{1, 2, \ldots, N_i\}$ is the* index *at height i; there exists a one-one correspondence between j and the subspace v used for projection*
- *A node $v = (i, j)$ at order i corresponds to the RM code* $\mathrm{RM}(m - r + i, i)$.

Owing to the close relationship between the recursions in a single iteration of the RPA decoder and the tree structure defined above, we let $L_i^{(j)}$ denote the LLR vector $L^{(v)}$ at the $(r-i)^{\text{th}}$ step of the recursion, using the subspace v associated with index j of the node (i, j), for $i \in \{1, \ldots, r\}$. The estimates $\widehat{L}_i^{(j)}$ and $\widehat{y}_i^{(j)}$ are similarly defined.

The next lemma considers the channel

$$W^-(y_1, y_2 \mid s) = \frac{1}{2} \sum_{\substack{u_1, u_2 \in \{0,1\} \\ u_1 \oplus u_2 = s}} W(y_1 \mid u_1)\, W(y_2 \mid u_2), \quad s \in \{0,1\},$$

which, as we argue, is precisely the channel induced via the projection operation.

Lemma 1. *Let $W \colon \{0,1\} \to \mathcal{Y}$ be a binary-input memoryless symmetric (BMS) channel. Then:*

1) *W^- is a BMS channel.*
2) *For any nonzero $v \in \mathbb{F}_2^m$, the pair $\big(Y(z), Y(z \oplus b)\big)$ conditioned on $C(z) \oplus C(z \oplus b)$ has distribution W^-.*
3) *The Bhattacharyya parameter satisfies $Z(W^-) \leq 1 - \big(1 - Z(W)\big)^2$.*

Proof. (1) Follows directly from [23, Prop. 13].
(2) Let $U_1, U_2 \sim \mathrm{Bern}(1/2)$ be independent and set $S = U_1 \oplus U_2$. It can easily be checked that $\mathbb{P}\{Y_1 = y_1, Y_2 = y_2 \mid S = s\} = W^-(y_1, y_2 \mid s)$, for $s \in \{0, 1\}$. Any projection $v \neq \mathbf{0}$ selects a coset $\{z, z \oplus v\}$, and because the channel uses are memoryless, the distribution of $\big(Y(z), Y(z \oplus b)\big)$ given the parity $\mathcal{C}(z) \oplus \mathcal{C}(z \oplus b)$ is W^-. Thus the induced channel for projection along b is W^-.
(3) Follows directly from [23, Prop. 5]. □

Following Lemma 1, we see that the channel induced by projecting along any one-dimensional subspace $\{0, v\}$, which we denote $W^{(v)}$, is exactly W^-. Following previous notation, we let W_i stand for the (common) channel induced after any projection at the $(r - i + 1)^{\text{th}}$ step of the RPA recursion, $i \in \{1, \ldots, r\}$. Let Z_i be the Bhattacharyya parameter of channel W_i.

Lemma 2. *We have that $Z_i \leq 1 - (1 - Z_{i+1})^2$.*

Proof. Fix any node at height $i + 1$, and consider the projection along a one-dimensional subspace $B = \{0, v\}$. By Lemma 1 (Items 2 and 3), we obtain that $Z(W_{i+1}^-) \leq 1 - \big(1 - Z(W_{i+1})\big)^2$. Since $Z_i = Z(W_i) = Z(W^{(v)})$, we obtain the desired bound. □

Lemma 3. *We have that for any $i \in \{1, \ldots, r\}$, $Z_i \leq 1 - (1 - Z_r)^{2^{r-i}}$.*

Proof. From Lemma 2, we know that $Z_i \leq 1 - (1 - Z_{i+1})^2$. Define $Y_i = 1 - Z_i$. Then, $Y_i = (1 - Z_{i+1})^2 = Y_{i+1}^2$. Therefore, $Z_i = 1 - Y_i = 1 - Y_r^{2^{r-i}} = 1 - (1 - Z_r)^{2^{r-i}}$. □

III. Main Result

As in [19], we use [15, Prop. 2] to focus our analysis on the case when the input codeword is fixed to be $C = \mathbf{0} \in \{0,1\}^N$, since, via the symmetry of the channel, we have that $P_{\text{err}}(\mathrm{RM}(m, r))$ equals the error probability of the all-zeros codeword. In what follows, we let $Z =: Z(W) \in (0, 1)$ denote the Bhattacharyya parameter of the BMS channel W of interest[2].

[2] We restrict our attention in this paper to "non-degenerate" BMS channels whose Bhattacharyya parameters are strictly bounded away from 0 and 1.

Theorem 4. *Let $\lambda := -\ln(1-Z) \in (0,\infty)$. For $r < \log_2 m - \log_2 \lambda$, we have $P_{err}(RM(m,r)) \xrightarrow{m\to\infty} 0$.*

Remark 1. *Consider the setting where the BMS channel W is the BSC(p), with the cross-over probability $p \in (0,1/2)$. Theorem III.1 of [19] shows that for $r < \ln m + \ln\left(\frac{\ln 2}{\ln(1-2p)}\right)$, the RPA decoder achieves vanishing error probabilities over the BSC(p). Using the fact that for this channel, we have $Z = 2\sqrt{p(1-p)}$, it can be checked via numerical comparisons that the claim in [19] is stronger than Theorem 4, for all $p \in (0,1/2)$. However, both claims provide identical asymptotic guarantees on the growth rate of r with m (i.e., r growing roughly logarithmically in m) for vanishing error probabilities under RPA decoding.*

IV. Helper Lemmas

In order to prove Theorem 4, we shall first establish an upper bound on the error probability at each stage of recursion in terms of the error probabilities of the previous stages. We then unroll this recursion to obtain the final block error probability. But first, we require some more notation. For $1 \le i \le r$, let $N_i := 2^{m-r+i}$. Thus, the number of subspaces used for projection at any node at height i is $N_i - 1$.

Definition 4. *For any $1 \le i \le r$ and $j \le N_i$, let $\mathcal{Q}_i^{(j)}$ be the event that the decoded estimate $\widehat{Y}_i^{(j)}$ at node (i,j) of $\mathcal{T}$ is incorrect (does not equal $\mathbf{0}$). Further, let $\mathcal{Q}_i := \cup_{j=1}^{N_i} Q_i^{(j)}$.*

We are interested in obtaining an upper bound on $\mathbb{P}\{\mathcal{Q}_r\}$, which directly yields an upper bound on $P_{\text{err}}(\text{RM}(m,r))$.

Definition 5. *For a node (i,j) at height $i \ge 2$, define the event $\mathcal{G}_i^{(j)}$ as the event that all children of node (i,j) with order $i+1$ are decoded correctly.*

A. Recurrence Relations for $\mathbb{P}\{\mathcal{Q}_i\}$

We proceed with deriving a recurrence relation for $\mathbb{P}\{\mathcal{Q}_i\}$, $1 \le i \le r$, in terms of $\mathbb{P}\{\mathcal{Q}_j\}$, $j < i$; in all that follows, we implicitly condition on the fact that the all-zeros codeword was transmitted. We then "unroll" this recurrence to yield a closed-form upper bound on $\mathbb{P}\{\mathcal{Q}_i\}$. For any event $\mathcal{E}$, we let $\overline{\mathcal{E}}$ denote its complement, where the universe can be inferred from the context.

First, we obtain an upper bound on $\mathbb{P}\{\mathcal{Q}_1\}$, which forms the "base case" of our recursive analysis of error probabilities. While the upper bound can also be obtained via the application of a standard union bound argument for the ML error probability (see, e.g., [24, Sec. 2.1] or [22, Problem 1.21]), we provide a direct proof, using properties of first-order RM codes, which could be of independent interest.

Lemma 5. *We have that $\mathbb{P}\{\mathcal{Q}_1\} \le (2^{m-r+2}-1)Z_1^{2^{m-r}}$.*

Proof. Via standard arguments (see, e.g., [19, Sec. IV-A]), there exists a one-one correspondence between the codewords of RM$(m-r+1,1)$ with the functions $\sigma \cdot \chi_s$ where $s \in \{0,1\}^{m-r+1}$ and $\sigma \in \{\pm 1\}$, with $\chi_s(x) := (-1)^{x\cdot s}$, $x \in \{0,1\}^{m-r+1}$. The Fast Hadamard Transform (FHT) used by the RPA decoder then computes

$$(\sigma,s) = \text{argmax}_{\sigma,s}\langle L, \sigma\cdot\chi_s\rangle.$$

Here, for functions $f,g : \{0,1\}^n \to \{-1,1\}$, we define their inner product $\langle f,g\rangle := \frac{1}{2^n}\cdot\sum_{x\in\{0,1\}^n} f(x)g(x)$, and L denotes the vector of log-likelihood ratios $(L(z))$, for $z \in \{0,1\}^{m-r+1}$, with $L(z) := \log\frac{W(Y_z|0)}{W(Y_z|1)}$. It can be checked that the "correct" codeword for any leaf node of $\mathcal{T}$ is the all-zeros codeword $\mathbf{0}$, which corresponds to $(\sigma,s) = (1,\mathbf{0})$. For any pair $(\sigma,s) \ne (1,\mathbf{0})$, define the event, $\mathcal{E}_{\sigma,s} = \mathbb{1}[\langle L,\sigma\chi_s\rangle \ge \langle L,\chi_{\mathbf{0}}\rangle]$. The FHT decoder makes an error iff $\mathcal{E}_{\sigma,s}$ occurs, for some $(\sigma,s) \ne (1,\mathbf{0})$. Thus, via a union bound,

$$\mathbb{P}\left\{\overline{\mathcal{Q}_1}\right\} \le \sum_{(\sigma,s)\ne(1,\mathbf{0})} \mathbb{P}\left\{\mathcal{E}_{\sigma,s}\right\}.$$

Observe that $\langle L,\sigma\cdot\chi_s\rangle - \langle L,\chi_{\mathbf{0}}\rangle = -\sum_{z:\sigma\cdot\chi_s(z)=-1} L(z)$. Hence the event $\mathcal{E}_{\sigma,s}$ is equivalent to,

$$\mathcal{E}_{\sigma,s} = \mathbb{1}\Big[\sum_{z:\sigma\cdot\chi_s(z)=-1} L(z) < 0\Big] = \mathbb{1}\big[e^{-\frac{\sum_z L(z)}{2}} > 1\big]. \tag{1}$$

Therefore by applying the Markov inequality to (1), we have

$$\begin{aligned}\mathbb{P}\left\{\mathcal{E}_{\sigma,s}\right\} &\le \mathbb{E}\left[e^{-\frac{\sum_z L(z)}{2}}\right] \\ &= \prod_{z:\sigma\cdot\chi_s(z)=-1} \mathbb{E}\left[e^{-\frac{L(z)}{2}}\right].\end{aligned} \tag{2}$$

We reiterate that all probabilities and expected values above are conditioned on $\mathbf{0}$ being transmitted. The last step holds since $L(z)$ are i.i.d. across $z \in \{0,1\}^{m-r+1}$. Now, observe that

$$\begin{aligned}\mathbb{E}\left[e^{-L(z)/2}\right] &= \sum_y W(y|0)\sqrt{\frac{W(y|1)}{W(y|0)}} \\ &= \sum_y \sqrt{W(y|0)W(y|1)} = Z_1.\end{aligned}$$

Therefore, following on from (2),

$$\prod_{z:\sigma\chi_s(z)=-1} \mathbb{E}\left[e^{-\frac{L(z)}{2}}\right] = Z_1^{|\{z:\sigma\cdot\chi_s(z)=-1\}|}.$$

We now bound the exponent $|\{z : \sigma \cdot \chi_s(z) = -1\}|$. For any nonzero s, χ_s has equal number of $+1$ and -1. As a result, $|\{z : \sigma \cdot \chi_s(z) = -1\}| = 2^{m-r}$ for all $s \neq \mathbf{0}$ because multiplying by σ flips the signs globally. Moreover, when $s = \mathbf{0}$, we have that when $\sigma = -1$, $|\{z : \sigma\chi_s(z) = -1\}| = 2^{m-r+1}$. Therefore,

$$\mathbb{P}\{\mathcal{E}_{\sigma,s}\} \leq Z_1^{2^{m-r}}, \text{ for all } (\sigma, s) \neq (1, 0).$$

Finally, via (IV-A), we obtain that $\mathbb{P}\left\{\overline{\mathcal{Q}_1}\right\} \leq (2^{m-r+2} - 1)Z_1^{2^{m-r}}$. □

Lemma 6. *For $i \geq 2$, we have*

$$\mathbb{P}\{\mathcal{Q}_i\} \leq N_i Z_i^{N_i-1} + (N_i - 1)\mathbb{P}\{Q_{i-1}\}.$$

Proof. We know that

$$\mathbb{P}\{\mathcal{Q}_i\} \leq \mathbb{P}\{\mathcal{Q}_i|\mathcal{G}_i\} + \mathbb{P}\left\{\overline{\mathcal{G}_i}\right\}. \tag{3}$$

We will prove the lemma by showing the following two claims hold: Firstly, that

$$\mathbb{P}\left\{\overline{\mathcal{G}_i}\right\} \leq (N_i - 1)\mathbb{P}\{Q_{i-1}\}, \tag{4}$$

and secondly, that

$$\mathbb{P}\{\mathcal{Q}_i|\mathcal{G}_i\} \leq N_i Z_i^{N_i-1}. \tag{5}$$

To prove the first claim, notice that the event $\mathcal{G}_i$ is equivalent to the event $\cap_{j'\in\mathbf{ch}(i,j)} Q_{i+1}^{j'}$, where $\mathbf{ch}(i, j)$ denotes the collection of children of node (i, j). Hence by the union bound, we get that

$$\mathbb{P}\left\{\overline{\mathcal{G}_i}\right\} \leq \sum_{j'\mathbf{ch}(i,j)} \mathbb{P}\left\{Q_{i+1}^{(j)}\right\} \leq (N_i - 1)\mathbb{P}\{Q_{i+1}\},$$

thereby proving (4).

Now we move on to proving the second claim (5). Let $\mathcal{F}_z$ for $z \in \mathbb{F}_2^{m-r+i}$ be the event that $\widehat{Y}_i^{(j)}(z) \neq 0$. Now, via arguments similar to those in [19, Sec. IV-B], it can be checked that conditioned on the event $\mathcal{G}_i$, the "aggregated" LLR vector $\widehat{L}_i^{(j)}$ at node (i, j) obeys

$$\widehat{L}_i^{(j)}(z) = \frac{1}{2^{m-r+i} - 1} \sum_{z'\neq z} L_i^{(j)}(z').$$

Let $\overline{L}_i^{(j)}(z) := (2^{m-r+i} - 1) \cdot \widehat{L}_i^{(j)}(z)$. We then have via the Markov inequality that

$$\begin{aligned}
\mathbb{P}\{\mathcal{F}_z|\mathcal{G}_i\} &= \mathbb{P}\left\{\overline{L}_i^{(j)}(z) < 0\right\} \\
&\leq \mathbb{E}\left[e^{-\overline{L}_i^{(j)}(z)/2}\right] \\
&= \prod_{z\neq z'} \mathbb{E}\left[e^{L_i^{(j)}(z')/2}\right] = Z(W_i)^{2^{m-r+i}-1}.
\end{aligned}$$

Now, conditioned on $\mathcal{G}_i$, note that we have $Q_i = \cup_{z\in\mathbb{F}_2^{m-r+i}} \mathcal{F}_z$. Hence,

$$\begin{aligned}
\mathbb{P}\{\mathcal{Q}_i|\mathcal{G}_i\} &\leq \sum_{z\in\mathbb{F}_2^{m-r+i}} \mathbb{P}\{\mathcal{F}_z|\mathcal{G}_i\} \\
&= 2^{m-r+i} Z(W_i)^{2^{m-r+i}-1},
\end{aligned}$$

thereby proving our second claim. □

B. Explicit Upper Bound on $\mathbb{P}\{\mathcal{Q}_i\}$

The following lemma then follows by "unrolling" the recursion in Lemma 6.

Lemma 7. *Define for each $t \in \{2, \ldots, r\}$,*

$$A_t := 2^{m-r+t}\, Z_t^{2^{m-r+t}-1}, \qquad B_t := 2^{m-r+t} - 1.$$

Then for every $i \in \{2, \ldots, r\}$,

$$\mathbb{P}\{\mathcal{Q}_i\} \leq \sum_{t=2}^{i} \left(A_t \prod_{s=t+1}^{i} B_s\right) + \mathbb{P}\{\mathcal{Q}_1\} \prod_{s=2}^{i} B_s.$$

In particular,

$$\mathbb{P}\{\mathcal{Q}_r\} \leq \sum_{t=2}^{r} \left(A_t \prod_{s=t+1}^{r} B_s\right) + \mathbb{P}\{\mathcal{Q}_1\} \prod_{s=2}^{r} B_s.$$

V. Proof of Main Result

In this section we will prove Thm. 4.

Proof of Thm. 4. For each $t \in \{2, \ldots, r\}$ from Lemma 3 we have that $A_t := 2^{m-r+t} Z_t^{2^{m-r+t}-1}$ and $B_t := 2^{m-r+t} - 1 \leq 2^{m-r+t}$. Observe that for any $u \in \{2, \ldots, r\}$, we have

$$\begin{aligned}
\prod_{s=u}^{r} B_s &= \prod_{s=u}^{r} 2^{m-r+s} \\
&= 2^{(r-u+1)(m-r)+\sum_{s=u}^{r} s} \leq 2^{O(mr)}.
\end{aligned}$$

Thus, from Lemma 7,

$$\mathbb{P}\{\mathcal{Q}_r\} \leq 2^{O(mr)}\mathbb{P}\{\mathcal{Q}_1\} + \sum_{t=2}^{r} 2^{O(mr)} Z_t^{2^{m-r+t}-1}. \tag{6}$$

We will individually show that each term above approaches zero in the large m limit, for $r \leq \log_2 m - \log_2 \lambda - \delta$, for some small constant δ. For the first term we will show that $\mathbb{P}\{\mathcal{Q}_1\}$ is very small compared to its multiplicative pre-factor that is $2^{O(mr)}$. To this end, recall that $\lambda := -\ln(1 - Z) = -\ln(1 - Z_r)$. From Lemma 3, we know that $(1 - Z_t) \geq (1 - Z_r)^{2^{r-t}} = e^{-\lambda\cdot 2^{r-t}}$. Now, from Lemma 5, we obtain using the previous inequality that

$$\mathbb{P}\{\mathcal{Q}_1\} \leq 2^{m-r+2} e^{-(2^{m-r}\cdot e^{-\lambda\cdot 2^{r-1}})}.$$

Here, we use the fact that for any positive integer k, $Z_t^k \le e^{-k(1-Z_t)}$. Notice that in the regime $r \le \log_2 m - \log_2 \lambda - \delta$, we have $\lambda \cdot 2^{r-1} = O(m)$. Therefore, $2^{m-r} \cdot \exp(-\lambda 2^{r-1}) = \exp(\Omega(m))$. Plugging this into the first term in (6), we get that

$$\begin{aligned} 2^{O(mr)}\mathbb{P}\left\{Q_1\right\} &\le 2^{O(mr)} 2^{m-r+2} e^{-\left(2^{m-r} e^{-\lambda 2^{r-1}}\right)} \\ &\le 2^{O(mr)} e^{-e^{\Omega(m)}} \xrightarrow{m\to\infty} 0. \end{aligned}$$

Now we focus on the second summation in (6); we will show that each summand approaches zero. So fix any t, define $b_t = 2^{m-r+t} - 1$. Then, since $Z_t^{b_t} \le e^{-b_t(1-Z_t)}$, we have that

$$\begin{aligned} (1-Z_t)b_t &\ge \exp(-\lambda 2^{r-t})(2^{m-r+t}-1) \\ &\ge 2^{m-r+t-1} \cdot \exp(-\lambda \cdot 2^{r-t}) \\ &= \exp\left((m-r+t-1)\ln 2 - \lambda \cdot 2^{r-t}\right). \end{aligned}$$

Again using the fact that $\lambda \cdot 2^{r-t} \le \lambda \cdot 2^{r-1} = o(m)$, we get that

$$(m-r+t-1)\ln 2 - \lambda \cdot 2^{r-t} = \Theta(m) - o(m) = \Omega(m).$$

Plugging this into any term in the summation in (6), we get that

$$\begin{aligned} 2^{O(mr)} Z_t^{2^{m-r+t}-1} &\le 2^{O(mr)} e^{-\exp(m-r+t-1)\ln 2 - \lambda 2^{r-t})} \\ &\xrightarrow{m\to\infty} 0. \end{aligned}$$

This concludes the proof of Theorem 4. □

Acknowledgments

V. A. Rameshwar acknowledges support from the New Faculty Initiation Grant, IIT Madras. D. Fathollahi acknowledges support from the National Science Foundation (NSF) under Grant No. CCF-2231157. The authors thank Harshithanjani Athi for helpful discussions on the problem.

References

[1] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *Transactions of the IRE Professional Group on Information Theory*, vol. 4, no. 4, pp. 38–49, 1954.

[2] D. E. Muller, "Application of boolean algebra to switching circuit design and to error detection," *Transactions of the I.R.E. Professional Group on Electronic Computers*, vol. EC-3, no. 3, pp. 6–12, 1954.

[3] S. Kudekar, S. Kumar, M. Mondelli, H. D. Pfister, E. Sasoglu, and R. L. Urbanke, "Reed-Muller codes achieve capacity on erasure channels," *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 4298–4316, 2017.

[4] G. Reeves and H. D. Pfister, "Reed–Muller codes on BMS channels achieve vanishing bit-error probability for all rates below capacity," *IEEE Transactions on Information Theory*, pp. 1–1, 2023.

[5] E. Abbe and C. Sandon, "A proof that Reed-Muller codes achieve Shannon capacity on symmetric channels," in *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, 2023, pp. 177–193.

[6] E. Abbe, O. Sberlo, A. Shpilka, and M. Ye, "Reed-muller codes," *Found. Trends Commun. Inf. Theory*, vol. 20, no. 1–2, p. 1–156, Jan. 2023. [Online]. Available: https://doi.org/10.1561/0100000123

[7] R. R. Green, "A serial orthogonal decoder," *JPL Space Programs Summary*, vol. 37–39-IV, p. 247–253, 1966.

[8] Y. Be'ery and J. Snyders, "Optimal soft decision block decoders based on fast Hadamard transform," *IEEE Transactions on Information Theory*, vol. 32, no. 3, pp. 355–364, 1986.

[9] V. M. Sidel'nikov and A. S. Pershakov, "Decoding of Reed-Muller codes with a large number of errors," *Problemy Peredachi Informatsii*, vol. 28, no. 3, p. 80–94, 1992.

[10] B. Sakkour, "Decoding of second order Reed-Muller codes with a large number of errors," in *IEEE Information Theory Workshop*, 2005, p. 3.

[11] I. Dumer, "Recursive decoding and its performance for low-rate Reed-Muller codes," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 811–823, 2004.

[12] ——, "Soft-decision decoding of Reed-Muller codes: a simplified algorithm," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 954–963, 2006.

[13] I. Dumer and K. Shabunov, "Soft-decision decoding of Reed-Muller codes: recursive lists," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1260–1266, 2006.

[14] M. Burnashev and I. Dumer, "Error exponents for recursive decoding of Reed–Muller codes on a binary-symmetric channel," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 4880–4891, 2006.

[15] M. Ye and E. Abbe, "Recursive projection-aggregation decoding of Reed-Muller codes," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 2064–2068.

[16] J. Li, S. M. Abbas, T. Tonnellier, and W. J. Gross, "Reduced complexity RPA decoder for Reed-Muller codes," in *2021 11th International Symposium on Topics in Coding (ISTC)*, 2021, pp. 1–5.

[17] D. Fathollahi, N. Farsad, S. A. Hashemi, and M. Mondelli, "Sparse multi-decoder recursive projection aggregation for Reed-Muller codes," in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 1082–1087.

[18] B. Zhang, F. Chen, and Q. Huang, "Coset error pattern in projection-aggregation decoding," *IEEE Transactions on Information Theory*, vol. 71, no. 8, pp. 5920–5934, 2025.

[19] V. A. Rameshwar and V. Lalitha, "An upper bound on the error probability of rpa decoding of reed-muller codes over the bsc," in *2025 IEEE International Symposium on Information Theory (ISIT)*, 2025, pp. 1–6.

[20] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, 2nd ed. North-Holland, 1978.

[21] E. Abbe, A. Shpilka, and M. Ye, "Reed-Muller codes: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 67, no. 6, pp. 3251–3277, 2021.

[22] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.

[23] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[24] I. Sason and S. Shamai, "Performance analysis of linear codes under maximum-likelihood decoding: a tutorial," *Commun. Inf. Theory*, vol. 3, no. 1/2, p. 1–222, Jul. 2006. [Online]. Available: https://doi.org/10.1561/0100000009