Multi-Factor Pruning for Recursive Projection-Aggregation Decoding of RM Codes

Marzieh Hashemipour-Nazari  
Department of Electrical Engineering  
Eindhoven University of Technology  
Eindhoven, The Netherlands  
s.m.hashemipour.nazari@tue.nl

Kees Goossens  
Department of Electrical Engineering  
Eindhoven University of Technology  
Eindhoven, The Netherlands  
k.g.w.goossens@tue.nl

Alexios Balatsoukas-Stimming  
Department of Electrical Engineering  
Eindhoven University of Technology  
Eindhoven, The Netherlands  
a.k.balatsoukas.stimming@tue.nl

Abstract—The recently introduced recursive projection aggregation (RPA) decoding method for Reed-Muller (RM) codes can achieve near-maximum likelihood (ML) decoding performance. However, its high computational complexity makes its implementation challenging for time- and resource-critical applications. In this work, we present a complexity reduction technique called multi-factor pruning that reduces the computational complexity of RPA significantly. Our simulation results show that the proposed pruning approach with appropriately selected factors can reduce the complexity of RPA by up to 92% for RM(8, 3) while keeping a comparable error-correcting performance.

Index Terms—Reed-Muller codes, Recursive projection-aggregation (RPA) decoding, Projection pruning, Computational complexity

I. INTRODUCTION

Reed-Muller (RM) codes are linear block codes with a recursive structure introduced in 1954 by Muller [1]. Immediately after, Reed [2] introduced the first decoder for RM codes based on majority voting, which could correct bit errors in a codeword up to half of the minimum distance of the corresponding code. Since then, several efficient decoding methods for RM codes have been published [3]–[10]. Recently, there has been growing interest in RM codes as they have been shown to be capacity-achieving under the maximum-likelihood (ML) decoding for general symmetric channels [11]–[14]. However, since the complexity of ML decoding scales exponentially with the blocklength, ML decoding is not a practically feasible decoding method. Therefore, considerable effort has been put into proposing efficient near-ML decoding algorithms for RM codes with tractable complexity, which is especially attractive for short blocklength applications [15].

A wide variety of such proposed algorithms are taking advantage of the recursive structure of the RM codes to achieve near-ML decoding. For example, the recursive list decoder proposed in [10] provides efficient near-ML decoding performance for RM codes with reasonable list size. Nevertheless, similarly to all list-based decoders, it is challenging to achieve low-latency decoding even with a relatively small list size due to the required path selection operation that cannot be readily parallelized. Moreover, due to the similarity of RM codes to polar codes, the successive-cancellation (SC) decoding [11] and SC list (SCL) [16] decoding algorithms for polar codes are able to decode RM codes recursively with reasonable complexity.

Recently, a recursive algorithm called recursive projection-aggregation (RPA) [17] decoding was proposed for decoding RM codes. RM codes decoded with the RPA algorithm achieve near-ML decoding performance, outperforming polar codes with the same blocklength and rate under SCL decoding even with large list sizes. On a high level, the RPA decoder generates many smaller codewords by combining elements of the received vector and then decodes the generated codewords recursively to aggregate them into a reliable decision for the received vector. The authors of [17] mention that the computational complexity of RPA decoding scales like $O(n^r \log n)$ where $n$ is the blocklength and $r$ is the order of the RM code, making RPA decoding impractical for RM codes with large order $r$.

Several algorithms have been published to reduce the complexity of the RPA algorithm. For example, simplified RPA [17] and collapsed projection-aggregation (CPA) [18] reduce the number of recursive levels and project the received vectors directly onto smaller codes, thus reducing the number of projections. However, both simplified RPA and CPA lead to high-complexity projection and aggregation steps. The authors of [19] introduced a variant of the RPA decoder called sparse RPA that uses a small random subset of the projections, thus reducing the complexity of the RPA algorithm. However, the error-correcting performance of SRPA degrades significantly as the sparsity increases. To address this, the authors also proposed k-SRPA which employs $k$ sparse decoders to improve the error-correcting performance of single SRPA. A cyclic redundancy check (CRC) and Reed’s algorithm are used to select the most reliable candidate among all candidates decoded from the $k$ individual SRPA decoders. Unfortunately, the k-SRPA algorithm causes a reduction in the effective code rate due to the added CRC as well as hardware overhead due to the required randomization of projections, multiple control units, memories, and data paths for multiple decoders. The work of [20] proposed syndrome-based early stopping techniques along with a scheduling scheme to reduce the complexity of the RPA algorithm called reduced complexity RPA. Reduced complexity RPA is still challenging due to the computational overhead of syndrome checks and the variable decoding latency.

In this work, we further optimize RPA decoding with the ultimate goal of making it feasible for hardware implementation. More specifically, we propose a multi-factor pruning method...
that is inspired by observations we made about the relative
importance of projections at various recursion levels based
on simulations. Our method prunes the RPA decoder with
a varying level of aggressiveness for different iterations
and recursion levels. Our results show a computational complexity
reduction of 92% compared to the baseline RPA algorithm [17]
and 38% to 77% compared to the works of [20] and [19] with
no degradation in the error-correcting performance. Moreover,
in contrast to the works of [19] and [20], our proposed pruning
method can translate to a simpler hardware implementation.

The remainder of this paper is organized as follows. In
Section II, we give the preliminaries and background, including
a description of RM codes as well as the RPA algorithm
and we explain the issues with previously proposed pruning
methods in more detail. In Section III, we present our proposed
multi-factor pruning method. In Section IV, we present error-
correcting performance simulation results and computational
complexity comparisons with existing pruning methods. Finally,
we conclude the paper in Section V.

II. PRELIMINARIES AND BACKGROUND

A. Reed-Muller codes

Reed-Muller codes are linear block codes denoted by
RM(m, r), where n = 2^m is the blocklength, and r is the
order of the code. Similar to the other linear block codes, RM
codes are defined by a k × n generator matrix G_{(m,r)}, where
k = \sum_{i=0}^{r} (\binom{m}{i}) and the rate of the code is R = \frac{k}{n}.
A binary vector \mathbf{u} with k information bits is encoded into a binary vector
\mathbf{c} belonging to the RM(m, r) code as follows:

\begin{equation}
\mathbf{c} = \mathbf{u}G_{(m,r)}. (1)
\end{equation}

The generator matrix \( G_{(m,r)} \) of RM(m, r) code is obtained
from the \( n^{th} \) Kronecker power of \( F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \) by selecting the
rows with a Hamming weight of at least \( 2^{m-r} \). The resulting
generator matrix \( G_{(m,r)} \) has a universal and recursive structure
defined as:

\begin{equation}
G_{(m,r)} = \begin{bmatrix} G_{(m-1,r)} & G_{(m-1,r)} \\ 0 & G_{(m-1,r-1)} \end{bmatrix}, \quad G_{(1,1)} = F . (2)
\end{equation}

In addition, a codeword \( \mathbf{c} \in RM(m, r) \) can be be mapped to
an \( m \)-variate polynomial with degree less than or equal to \( r \)
in a vector space \( E := \mathbb{F}_2^m \). Consequently, any coordinate of the
codeword \( \mathbf{c} \) is indexed by an \( m \)-bit binary vector \( z \in E \), i.e.,
\( \mathbf{c} = (c(z), z \in E) \) [21].

B. Coset-based projections

Let \( \mathbb{B} \) be an \( s \)-dimensional subspace of \( E \). There are \( 2^{m-s} \)
different cosets \( T \) for \( \mathbb{B} \) making the quotient space \( E/\mathbb{B} \):

\begin{equation}
E/\mathbb{B} = \{ T := z + \mathbb{B}, z \in E \} . (3)
\end{equation}

Moreover, the projection of \( \mathbf{c} = (c(z), z \in E) \) on the cosets
of an \( s \)-dimensional subspace \( \mathbb{B} \) is defined as:

\begin{equation}
\mathbf{c}_/\mathbb{B} = \text{Proj}(\mathbf{c}, \mathbb{B}) := (c_/\mathbb{B}(T), T \in E/\mathbb{B}) , (4)
\end{equation}

where \( c_/\mathbb{B} = (c_/\mathbb{B}(T) := \oplus_{z \in T} c(z)) \) is a binary vector
generated by summing up all coordinates of \( c \) indexed by the
elements of each coset \( T \in E/\mathbb{B} \). The work of [17]
proved that \( c_\mathbb{B} \) is a codeword of \( RM(m-s, r) \) if \( c \) is a
codeword of \( RM(m, r) \). The RPA decoding algorithm, which
is the focus of this paper, exploits this property of RM codes
using one-dimensional subspaces, i.e., \( s = 1 \).

C. Recursive projection aggregation (RPA) Decoding

Let \( W : \{0,1\} \to \mathcal{W} \) denote a binary-input memoryless
channel, where \( \mathcal{W} \) is the output alphabet. The log-likelihood
ratio (LLR) of a channel output \( y \in \mathcal{W} \), denoted by \( L(y) \), is
defined as:

\begin{equation}
L(y) := \ln \left( \frac{W(y | 0)}{W(y | 1)} \right) . (5)
\end{equation}

Soft-decision RPA decoding uses the LLR values of the channel
output vector \( \mathbf{y} \), denoted by the vector \( \mathbf{L} \), as its input. In
particular, the RPA decoding algorithm decodes \( \mathbf{L} \) in three
steps: projection, recursive decoding, and aggregation.

As Fig. 1 shows, in the projection step the RPA algorithm
generates \( n - 1 \) vectors \( \mathbf{L}_/\mathbb{B}_i \), with length \( 2^{m-1} \) by projecting
\( \mathbf{L} \) onto \( n - 1 \) one-dimensional subspaces of the vector space
\( E \). Let \( \mathbb{B}_i = [0,i] \) be the \( i \)-th one-dimensional subspace of
\( E \). There are \( 2^{m-1} \) different cosets \( T \) for \( \mathbb{B}_i \). Therefore, the
quotient space \( E/\mathbb{B}_i \) is built based on (3). There are \( n - 1 \)
one-dimensional subspaces for the binary vector space \( E \) that
provide \( n - 1 \) projections for channel output vector \( \mathbf{L} \) given
their corresponding quotient space such that:

\begin{equation}
\mathbf{L}_/\mathbb{B}_i = \text{Proj}(\mathbf{L}, \mathbb{B}_i) := (L_/\mathbb{B}_i(T), T \in E/\mathbb{B}_i) , (6)
\end{equation}

where

\begin{equation}
L_/\mathbb{B}_i(T) = 2 \tanh^{-1} \left( \prod_{z \in T} \tanh \left( \frac{L(z)}{2} \right) \right) . (7)
\end{equation}

In the recursive decoding step, each projected vector \( \mathbf{L}_/\mathbb{B}_i \)
is recursively decoded by RPA for \( RM(m-1, r-1) \) until first-order
RM codes are reached. At this point, the RPA algorithm
uses an efficient decoder based on the fast Hadamard transform (FHT) to decode the first-order RM codes [3].

In the final aggregation step a per-coordinate average LLR is calculated from all the decoded codewords obtained from the previous step as follows:

\[
\hat{L}(z) = \frac{1}{n-1} \sum_{i=1}^{n-1} (1 - 2\hat{c}_i / \hat{c}_i + \hat{B}_i) L(z + z_i), \quad (8)
\]

where \(\hat{c}_i / \hat{c}_i + \hat{B}_i\) is the decoded binary codeword from recursive decoding, and \(\hat{L}(z)\) is the LLL value of \(z_i\) index of the estimated vector \(\hat{c}\) for the transmitted codeword \(c\).

The RPA decoding algorithm repeats the above steps up to \(N_{\text{max}}\) times or until the output converges to the input as:

\[
|\hat{L}(z) - L(z)| < \theta |L(z)|, \forall z \in E,
\]

where \(\theta\) is a small constant called the exiting threshold. The total number of FHT-based first-order decodings (FODs), which we use as a proxy for the computational complexity, is:

\[
\lambda_{\text{RPA}} = N_{\text{max}}^{r-1} \prod_{i=1}^{r-1} (2^m - i - 1),
\]

as mentioned in [20].

D. Existing complexity reduction methods

1) Sparse RPA decoder: The sparse RPA decoder (SRPA) [19] reduced the computational complexity of the RPA decoder by introducing a projection pruning factor \(0 < q < 1\). Specifically, it keeps \(q \times (n-1)\) random projections instead of all \(n-1\) projections at each recursive level. The numerical results of this decoder show that removing half of the projections for decoding an \(RM(m, 2)\) code does not degrade the error-correcting performance significantly. Moreover, for more aggressive pruning, k-SRPA was introduced to improve the decoding performance. It runs \(k\) SRPA decoders generating \(k\) estimated codewords for a single received vector and then uses a CRC check to choose the best candidate. It utilizes Reed’s decoder for \(k > 2\) to guarantee that the output is always a valid codeword. Numerical results reported in [19] showed that the computational complexity can be reduced between 50% to 70% compared to RPA. However, having multiple decoders and a Reed decoder for the final decision introduces overhead in terms of hardware implementation. In addition, added CRC bits compromise the effective rate of the RM codes and the randomization of projections introduces additional hardware implementation overhead.

2) Reduced complexity RPA decoder: The work of [20] proposed scheduled RPA, denoted by \(\text{RPA}_{\text{SCH}}\), that reduces the number of projections in successive iterations. \(\text{RPA}_{\text{SCH}}\) introduced a decaying parameter \(d > 1\) defining the number of projections at each iteration \(j (j \in [1, N_{\text{max}}])\) as \(\lceil \frac{n-1}{d^j} \rceil\). It keeps all projections for the first iteration since most of the errors get corrected in the first iteration [22]. After that, it reduces the number of projections by a factor of \(d\) at each iteration level. This scheduling technique lowers the computational complexity compared to the baseline RPA by up to 67%. However, there is still a lot of room to further reduce the number of projections, especially for RM codes with \(r > 2\) as we will show in the sequel.

III. Multi-factor pruned RPA algorithm

Both SRPA and \(\text{RPA}_{\text{SCH}}\) significantly reduce the computational complexity as measured by the number of FODs, especially for second-order RM codes. However, the number of FODs is still high for \(RM(m, r > 2)\) codes as both methods prune all the recursive levels uniformly. Moreover, \(\text{RPA}_{\text{SCH}}\) does not apply the decaying factor for the recursion levels in the iteration layers. In other words, the decoder being called recursively in \(\text{RPA}_{\text{SCH}}\) for decoding \(RM(m, r > 2)\) codes starts with all projections even in after the first iteration, resulting in a very large number of FODs for the recursion levels called in each iteration.

In this section, we first describe a general multi-factor projection pruning function and we explain the motivation behind the choice of each factor. We then use this function to introduce a multi-factor pruned RPA (MFP-RPA) algorithm where the pruning factor is a function of both on the iteration and the recursion level, resulting in significantly more aggressive pruning than existing methods while maintaining the error-correcting performance of the original RPA algorithm.

A. Projection pruning function

We define the following pruning function that uses three user-defined factors \(\gamma, \delta_{\text{itr}}, \text{ and } \delta_{\text{rec}}, \text{ where } 0 < \gamma, \delta_{\text{itr}}, \delta_{\text{rec}} \leq 1\), and has the iteration number \(j\) and the recursion level \(l\) as its inputs:

\[
\Delta(j, l, \gamma, \delta_{\text{itr}}, \delta_{\text{rec}}) = \gamma \times (\delta_{\text{itr}})^{j-1} \times (\delta_{\text{rec}})^{l-2},
\]

The output of the pruning function \(\Delta(j, l, \gamma, \delta_{\text{itr}}, \delta_{\text{rec}})\) is a pruning factor that dictates the fraction of projections that are kept at iteration \(j\) for recursion level \(l\) compared to RPA decoding. These parameters are set off-line. In the following, we explain the motivation behind each pruning factor \(\gamma, \delta_{\text{itr}}, \text{ and } \delta_{\text{rec}}\).

1) Starting pruning factor \(\gamma\): As mentioned in Section II-C, the RPA decoder starts with \(n - 1\) projections for decoding an \(RM(m, r)\) code. As [19] stated, having this number of projections is typically redundant and results in a very large computational complexity. As a result, inspired by [19], we define the starting pruning factor \(\gamma\), which reduces the number of projections even for the first iteration, in contrast to [20]. Moreover, contrary to the approach taken by [19], we allow \(\gamma\) to change in each level of recursion as a function of the other pruning factors, as shown in (11).

2) Iteration pruning factor \(\delta_{\text{itr}}\): Similar to [20], from our numerical simulation results of RPA, we also concluded that the first iteration is the most crucial one and the importance of further iterations quickly decreases. Therefore, the iteration pruning factor \(\delta_{\text{itr}}\) exponentially decreases with the iteration count, as shown in (11).
Algorithm 1: MFP-RPA

Input: \(L, n, r, N_{\text{max}}, \gamma, \delta_{\text{itr}}, \delta_{\text{rec}}\)

Output: Estimated decoded codeword \(\hat{c}\)

1. if \(r = 1\) then
   2. \(\hat{c} \leftarrow \text{FOD}(L, n)\)
   3. else for \(j = 1 : N_{\text{max}}\) do
      4. \(n_p = \lfloor \Delta(j, r, \gamma, \delta_{\text{itr}}, \delta_{\text{rec}}) \times (n - 1) \rfloor\)
      5. \(\gamma' \leftarrow \gamma \times (\delta_{\text{itr}})^{j-1}\)
      6. for \(t = 0 : n_p - 1\) do
         7. \(i = t \times \lfloor \frac{n - 1}{n_p} \rfloor + 1\)
         8. \(L_{/B_i} \leftarrow \text{Projection}(L, B_i)\)
         9. \(\hat{c}_{/B_i} \leftarrow \text{MFP-RPA}(L_{/B_i}, \frac{n}{2}, r - 1, N_{\text{max}}, \gamma', \delta_{\text{itr}}, \delta_{\text{rec}})\)
         10. \(L_{\text{accu}} \leftarrow \text{Aggregation}(L, c_{/B_i}, B_i)\)
      11. end
      12. \(L \leftarrow L_{\text{max}} / n_p\)
   13. end
   14. \(\hat{c} \leftarrow \frac{1 - \text{sign}(L)}{2}\) // hard-decision
15. return \(\hat{c}\);

3) Recursion pruning factor \(\delta_{\text{rec}}\): We designed an experiment to explore the impact of each recursive layer on the error-correcting performance of the pruned version of the RPA decoder for various RM\((m, 3)\) codes. We defined two parameters \(P_1\) and \(P_2\) representing the number of projections in level \(r = 3\) and \(r = 2\) for the RM\((m, 3)\) codes, respectively, as shown in Fig. 2. The product \(P_1 \times P_2\) determines the number of required FODs for the generated first-order RM codes. In this experiment, we set \(N_{\text{max}} = 1\) to remove the effect of iterations on the performance. As shown in Fig. 3, with the same overall number of FODs, aggressive pruning in the level \(r = 2\) leads to a worse error-correcting performance compared to the aggressive pruning in the level \(r = 3\). As a result, we can conclude that the effect of each recursion level on the overall error-correcting performance of the pruned version of the RPA decoder is clearly not the same. Motivated by this observation, the recursion pruning factor \(\delta_{\text{rec}}\) affects the recursion levels such that the highest number of projections is kept at the level \(r = 2\), meaning that the pruning becomes more aggressive for the higher recursion layer, as (11) shows.

B. Multi-factor pruned RPA algorithm

Algorithm 1 describes the overall proposed multi-factor pruned (MFP-RPA) decoding method for RM\((m, r)\) codes. Overall, it follows the logic of the baseline RPA decoding algorithm but, as noted in line 10, it selects \(n_p\) projections distributed among all \(n - 1\) projections uniformly, where \(n_p \ll n - 1\). Moreover, SRPA can be obtained using MFP-RPA by setting the tuple \((\gamma, \delta_{\text{itr}}, \delta_{\text{rec}})\) to \((q, 1, 1)\). Similarly, RPA\(_{\text{SCH}}\) can be obtained using MFP-RPA by setting the tuple \((\gamma, \delta_{\text{itr}}, \delta_{\text{rec}})\) to \((1, \frac{1}{3}, 1)\) for all recursion levels. Finally, the baseline RPA algorithm can also be obtained from MFP-RPA with \((\gamma, \delta_{\text{itr}}, \delta_{\text{rec}}) = (1, 1, 1)\).

C. Complexity of MFP-RPA

The number of FODs for MFP-RPA decoding the RM\((m, r)\) code can be calculated as follows:

\[
\lambda_{\text{MFP-RPA}} = \sum_{j=1}^{N_{\text{max}}} \prod_{t=2}^{r} \left[ \Delta(j, l, \gamma \delta_{\text{itr}}^{j-1}, \delta_{\text{itr}}, \delta_{\text{rec}}) \times \left( \frac{n}{2^{r-t}} - 1 \right) \right].
\]

Compared to (10), a good selection of \(\gamma, \delta_{\text{itr}}\), and \(\delta_{\text{rec}}\) can result in a significant reduction of computational complexity. This complexity reduction can lead to lower hardware resource requirements, lower latency, or both, depending on the exact hardware architecture that is used (e.g., fully parallel, sequential, or semi-parallel, respectively).

IV. SIMULATION RESULTS

We simulate the performance of the proposed MFP-RPA algorithm on RM\((7, 2)\) and RM\((8, 3)\) codes over the additive white Gaussian noise (AWGN) channel. Fig. 4 shows the
with the number of FODs significantly. Specifically, Table I illustrates the algorithm and its previously proposed pruning methods SRPA factors. Taken into consideration for selecting the appropriate pruning and we currently do not have a systematic way of choosing the pruning factors heuristically in a way that the performance loss compared to the non-pruned decoder is negligible and we currently do not have a systematic way of choosing these factors. Furthermore, the code’s parameters as well as required reliability, latency, and available hardware resources in the event of hardware implementation should generally be taken into consideration for selecting the appropriate pruning factors.

We observe that the proposed MFP-RPA algorithm has effectively identical error-correcting performance to the RPA algorithm and its previously proposed pruning methods SRPA and RPA\textsubscript{SCH}. However, as shown in Table I, it reduces the number of FODs significantly. Specifically, Table I illustrates the number of FODs required for the baseline RPA, RPA\textsubscript{SCH}, 2-SRPA, and our proposed MFP-RPA. We note that the numbers in Table I do not use early stopping controlled by \( \theta \) in (9) and we use the worst-case complexity numbers for SRPA and RPA\textsubscript{SCH} as a hardware implementation generally has to account for the worst case. However, early stopping can still be used in conjunction with MFP-RPA if desired to, e.g., reduce the energy consumption. We observe that, for the RM\((7,2)\) code, our proposed MFP-RPA reduces the number of FODs by 70% and 49% compared to the RPA and RPA\textsubscript{SCH} algorithms. Compared to 2-SRPA, MFP-RPA requires 18% more FODs, but it also has a better error-correcting performance. For the RM\((8,3)\) code, MFP-RPA reduces the number of FODs by 92%, 77%, and 38% compared to RPA, RPA\textsubscript{SCH}, and 2-SRPA, respectively. We proposed the update rule (11) based on the intuitions explained in Section III. However, other update rules are possible and they may lead to further complexity reduction. This is an interesting open problem.

V. CONCLUSION

In this work, we proposed a multi-factor pruning method for RPA decoding of RM codes that prunes projections as a function of both the iteration number and the recursion level. Our results show that significantly more aggressive projection pruning is possible compared to existing methods without degrading the error-correcting performance. Specifically, our proposed multi-factor pruning method leads to up to 92% and 77% lower computational complexity compared to the baseline RPA decoding algorithm and previously proposed complexity reduction techniques, respectively.

TABLE I

Comparison of the number of FODs required for RPA [17], RPA\textsubscript{SCH} \([20]\), 2-SRPA \([19]\), and our proposed MFP-RPA.

| Code \((k, n)\) | RPA | RPA\textsubscript{SCH} | 2-SRPA | MFP-RPA\((\gamma, \delta_{ur}, \delta_{rec})\) |
|---------------|-----|----------------------|--------|---------------------------------|
| RM\((7,2)\)   | 481 | 221                  | 96     | 118                             |
| RM\((8,3)\)   | 29445 | 98385 | 36433 | 22544 |

Fig. 4. FER comparison between RPA, 2-SRPA, RPA\textsubscript{SCH}, and MFP-RPA algorithms. The results for 2-SRPA and RPA\textsubscript{SCH} are taken from \([19]\) and \([20]\), respectively.

simulation results for the proposed MFP-RPA compared to RPA \([17]\), SRPA \([19]\), and RPA\textsubscript{SCH} \([20]\). From our simulations with \( N_{\text{max}} = \lceil \frac{m}{2} \rceil \), we observed that the last iteration does not impact the error-correcting performance. Therefore, we set the maximum number of iterations \( N_{\text{max}} = 3 \) for RPA and MFP-RPA. Moreover, we set the tuple \((\gamma, \delta_{ur}, \delta_{rec})\) of user-defined pruning factors of (11) to \((\frac{2}{7}, \frac{1}{7}, \frac{1}{2})\) for the RM\((7,2)\) code and \((\frac{2}{7}, \frac{3}{7}, \frac{2}{7})\) for the RM\((8,3)\) code. We note that we selected the pruning factors heuristically in a way that the performance loss compared to the non-pruned decoder is negligible and we currently do not have a systematic way of choosing these factors. Furthermore, the code’s parameters as well as required reliability, latency, and available hardware resources in the event of hardware implementation should generally be taken into consideration for selecting the appropriate pruning factors.

We observe that the proposed MFP-RPA algorithm has effectively identical error-correcting performance to the RPA algorithm and its previously proposed pruning methods SRPA and RPA\textsubscript{SCH}. However, as shown in Table I, it reduces the number of FODs significantly. Specifically, Table I illustrates the number of FODs required for the baseline RPA, RPA\textsubscript{SCH}, 2-SRPA, and our proposed MFP-RPA. We note that the numbers in Table I do not use early stopping controlled by \( \theta \) in (9) and we use the worst-case complexity numbers for SRPA and RPA\textsubscript{SCH} as a hardware implementation generally has to account for the worst case. However, early stopping can still be used in conjunction with MFP-RPA if desired to, e.g., reduce the energy consumption. We observe that, for the RM\((7,2)\) code, our proposed MFP-RPA reduces the number of FODs by 70% and 49% compared to the RPA and RPA\textsubscript{SCH} algorithms. Compared to 2-SRPA, MFP-RPA requires 18% more FODs, but it also has a better error-correcting performance. For the RM\((8,3)\) code, MFP-RPA reduces the number of FODs by 92%, 77%, and 38% compared to RPA, RPA\textsubscript{SCH}, and 2-SRPA, respectively. We proposed the update rule (11) based on the intuitions explained in Section III. However, other update rules are possible and they may lead to further complexity reduction. This is an interesting open problem.

V. CONCLUSION

In this work, we proposed a multi-factor pruning method for RPA decoding of RM codes that prunes projections as a function of both the iteration number and the recursion level. Our results show that significantly more aggressive projection pruning is possible compared to existing methods without degrading the error-correcting performance. Specifically, our proposed multi-factor pruning method leads to up to 92% and 77% lower computational complexity compared to the baseline RPA decoding algorithm and previously proposed complexity reduction techniques, respectively.

REFERENCES

[1] D. E. Muller, “Application of boolean algebra to switching circuit design and to error detection,” Transactions of the IRE Professional Group on Electronic Computers, vol. EC-3, pp. 6–12, Sep 1954.
[2] I. Reed, “A class of multiple-error-correcting codes and the decoding scheme,” Transactions of the IRE Professional Group on Information Theory, vol. 4, pp. 38–49, Sep 1954.
[3] Y. Be’ery and J. Snyders, “Optimal decision block decoders based on fast Hadamard transform,” IEEE Transactions on Information Theory, vol. 32, pp. 355–364, May 1986.
[4] A. E. Ashikhmin and S. N. Litsyn, “Fast decoding algorithms for first order Reed–Muller and related codes,” Designs, Codes and Cryptography, vol. 7, pp. 187–214, Mar 1996.
[5] R. Green, “A serial orthogonal decoder,” Jet Propulsion Laboratory (JPL) Space Programs Summary, vol. 37, pp. 247–253, 1966.
[6] V. M. Sidelnikov and A. S. Pershakov, “Decoding of Reed–Muller codes with a large number of errors,” Problemy Peredachi Informatsii, vol. 28, pp. 80–94, 1992.
[7] I. Dumer, “Recursive decoding and its performance for low-rate Reed–Muller codes,” IEEE Transactions on Information Theory, vol. 50, pp. 811–823, May 2004.
[8] B. Sakkour, “Decoding of second order Reed–Muller codes with a large number of errors,” in IEEE Information Theory Workshop, Aug. 2005.
[9] I. Dumer, “Soft-decision decoding of Reed–Muller codes: a simplified algorithm,” IEEE Transactions on Information Theory, vol. 52, pp. 954–963, Mar. 2006.
[10] I. Dumer and K. Shabunov, “Soft-decision decoding of Reed–Muller codes: Recursive lists,” IEEE Transactions on Information Theory, vol. 52, pp. 1260–1266, Mar. 2006.
[11] E. Arikan, “Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,” IEEE Transactions on Information Theory, vol. 55, pp. 3051–3073, Jul. 2009.
[12] N. Hussami, S. B. Korada, and R. Urbanke, “Performance of polar codes for channel and source coding,” in IEEE International Symposium on Information Theory, Jun. 2009.
[13] 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, pp. 4298–4316, Jul. 2017.
[14] E. Abbe, A. Shpilka, and A. Wigderson, “Reed–Muller codes for random erasures and errors,” in Proceedings of the forty-seventh annual ACM symposium on Theory of Computing, ACM, Jun. 2015.
[15] T. Tonnellier, M. Hashemipour-Nazari, N. Doan, W. J. Gross, and A. Balatsoukas-Stimming, “Towards practical near-maximum-likelihood decoding of error-correcting codes: An overview,” in IEEE International Conference on Acoustics, Speech and Signal Processing, Jun. 2021.

[16] I. Tal and A. Vardy, “List decoding of polar codes,” in IEEE International Symposium on Information Theory Proceedings, Jul. 2011.
[17] M. Ye and E. Abbe, “Recursive projection-aggregation decoding of Reed–Muller codes,” IEEE Transactions on Information Theory, vol. 66, pp. 4948–4965, Aug. 2020.
[18] M. Lian, C. Hager, and H. D. Pfister, “Decoding Reed–Muller codes using redundant code constraints,” in IEEE International Symposium on Information Theory, Jun. 2020.
[19] D. Fathollahi, N. Farsad, S. A. Hashemi, and M. Mondelli, “Sparse multi-decoder recursive projection aggregation for Reed–Muller codes,” in IEEE International Symposium on Information Theory, Jul. 2021.
[20] J. Li, S. M. Abbas, T. Tonnellier, and W. J. Gross, “Reduced complexity RPA decoder for Reed–Muller codes,” in IEEE International Symposium on Topics in Coding, Sep. 2021.
[21] E. Abbe and M. Ye, “Reed–Muller codes polarize,” IEEE Transactions on Information Theory, vol. 66, pp. 7311–7332, Dec. 2020.
[22] M. Hashemipour-Nazari, K. Goossens, and A. Balatsoukas-Stimming, “Hardware implementation of iterative projection-aggregation decoding of Reed–Muller codes,” in IEEE International Conference on Acoustics, Speech and Signal Processing, Jun. 2021.