A Novel Systematic Representation of Reed-Muller Codes with an Application to Linear Block Feedback Encoding

Vinayak Suresh and David J. Love
School of Electrical and Computer Engineering
Purdue University
Email: suresh20@purdue.edu, djlove@purdue.edu

***Abstract*—Reed-Muller (RM) codes are an important and powerful class of codes with several applications in electrical engineering and computer science. In this work, we prove a useful property for RM codes namely that they admit a special systematic generator matrix where the parity component has embedded in it a surprisingly large triangle of zeros. Asymptotically, the size in sub-diagonals of this all-zero triangle is shown to approach the largest possible value $\Delta^\star = \min(K, N-K)$ where K and N are the code dimension and code length respectively. To demonstrate an application of this result, we introduce the concept of linear block feedback codes where an open loop codeword is combined linearly with the feedback signal during encoding at the transmitter. This is shown to allow strengthening of a weak code to be as good as any desired code. We then show that, by virtue of the above property of RM codes, they can be emulated from an uncoded system using linear feedback encoding against remarkably large feedback delays.**

I. Introduction

Reed-Muller (RM) codes are some of the oldest and well-studied code families that remain relevant even today. They have found application in many research areas such as cryptography, distributed computing, theory of randomness (e.g., see [1] and the references therein). The invention of polar codes [2] has rekindled intensive research into RM codes due to their close relationship. Polar codes are known to be provably capacity achieving for binary-input symmetric discrete memoryless channels (DMCs). Recent developments have shown that RM codes achieve capacity on the Binary Erasure Channel [3] and are also long believed to achieve capacity over the Binary Symmetric Channel, which is strongly supported by simulations [4], [5]. Table 1 in [6] provides the best known capacity results for RM codes to date. Systematic RM codes have been considered from an encoding perspective in [7] and decoding perspective in [8]. A comprehensive survey on RM codes, their applications and connections to other research problems can be found in [9].

In this work, we prove a useful property of RM codes. We show that RM codes admit a special kind of systematic generator matrix of the form

$$\mathbf{G} = [\mathbf{I}_K \ \mathbf{P}^{(\Delta)}] \tag{1}$$

This work was supported in part by the National Science Foundation (NSF) under grants CNS1642982 and CCF1816013.

where $\mathbf{P}^{(\Delta)}$ is a $K \times (N-K)$ binary matrix with the property that it has Δ consecutive all-zero sub-diagonals beginning at the lower left end. For instance,

$$\mathbf{P}^{(3)} = \begin{bmatrix} \times & \times & \times & \times & \cdots & \times \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \times & \times & \times & \times & \cdots & \times \\ 0 & \times & \times & \times & \cdots & \times \\ 0 & 0 & \times & \times & \cdots & \times \\ 0 & 0 & 0 & \times & \cdots & \times \end{bmatrix}.$$

In other words, the parity component $\mathbf{P}^{(\Delta)}$ consists of an 'all-zero triangle' of size Δ embedded from the lower left. Note that the largest such triangle that can be fit into the $K \times (N-K)$ matrix $\mathbf{P}^{(\Delta)}$ has size $\Delta^\star = \min(K, N-K)$ where K and N are the code dimension and code length respectively. We show that for RM codes, as the code length increases, the size of the all-zero triangle Δ approaches the largest possible size $\Delta^\star$.

Availability of feedback in a communication system is greatly useful and can be employed to either simplify the communication scheme or improve reliability (e.g., [10]–[14]). Advanced forms of feedback techniques such as Hybrid Automatic Repeat Request (ARQ) and soft combining are proven to be important in mobile standards to achieve good performance. For the point to point AWGN channel, Schalkwijk and Kailath in a seminal work [15], [16] described a simple linear scheme (the S-K scheme) that is low complexity, achieves capacity and has a doubly exponentially decaying probability of error. Their general idea of sending an *uncoded* signal in the first channel use and then performing *linear processing of the feedback signal* (mainly through linear encoding of the error realizations or current error estimate at the receiver) in subsequent channel uses has been extended and applied successfully to many other scenarios as well [17], [18].

Inspired by the above, we introduce the idea of *linear block feedback codes* where an open loop codeword is combined linearly with the feedback signal during encoding. We show that this enables a weak code to effectively be strengthened to be as good as any desired code. Analogous to the S-K scheme, our work implies that for binary codes, an uncoded transmission combined with linear feedback processing can

achieve capacity. We then consider feedback encoding for the case when there is a delay of Δ channel uses in the feedback link. A central problem that arises is to find a code with good performance that has a generator of the form (1). Our results demonstrate that RM codes turn out to be a good solution - it is possible to enhance an uncoded transmssion to a desired RM code with linear feedback processing against remarkably large feedback delays.

II. Main Results

Denote the vector space of all binary m-tuples as V_m. A boolean function in m variables $f(v_1, v_2, \cdots, v_m)$ is a mapping from V_m to $\{0,1\}$. By fixing an ordering on $\{(v_1, v_2, \cdots, v_m) \in V_m\}$, we can uniquely associate to function f a binary vector $\mathbf{f}$ of length 2^m whose components are the result of evaluating f at all possible ordered input combinations.

Definition. The r-th order Reed-Muller (RM) code of length $N = 2^m$ denoted $\mathcal{R}(r,m)$ is a linear code that consists of vectors associated to all boolean polynomials f of degree less than equal to r in m variables. The dimension of $\mathcal{R}(r,m)$ is $K(r,m) = \sum_{j=0}^{r} \binom{m}{j}$ and minimum distance is $d_{min}(r,m) = 2^{m-r}$ [19].

A. A formula for $\Delta(r,m)$

In this section, we prove that the code $\mathcal{R}(r,m)$ admits a generator of the form (1) with $\Delta = \Delta(r,m)$ given by

$$\Delta(r,m) = \begin{cases} \sum_{j=0}^{r} \binom{m}{j} - \sum_{j=0}^{r} \binom{2j}{j} & 0 \le r \le \lfloor \frac{m}{2} \rfloor \\ \Delta(m-r-1,m) & \lfloor \frac{m}{2} \rfloor < r < m \end{cases}. \quad (2)$$

We being by showing that if a RM code admits a certain value of Δ, then so does its dual.

Lemma 1. *Suppose that for (r,m) with $r \le \lfloor \frac{m}{2} \rfloor$, the Reed-Muller code $\mathcal{R}(r,m)$ admits a systematic form $[\mathbf{I}\ \mathbf{P}_1^{(\Delta)}]$. Then, its dual $\mathcal{R}(m-r-1,m)$ admits a form $[\mathbf{I}\ \mathbf{P}_2^{(\Delta)}]$, with the same Δ.*

Proof: Since $\mathcal{R}(m-r-1,m)$ is the dual code to $\mathcal{R}(r,m)$, a valid generator matrix is $\left[\left(\mathbf{P}_1^{(\Delta)}\right)^T\ \mathbf{I}\right]$. The rest of the steps are pictorially represented in Fig. 1 where the triangle of zeros is shown in grey. The steps are

(a) Permute columns as shown.
(b) Apply a row permutation matrix $\mathbf{M}$ on the left to rearrange rows to obtain the form shown.
(c) Apply column permutation matrix $\mathbf{M}^T$ on the right to only the first block of columns to obtain the identity matrix for this block.

■

We are now ready to state the main theorem.

Theorem 1. *For $0 \le r < m$, $\mathcal{R}(r,m)$, the r-th order Reed-Muller code of length 2^m admits a systematic generator matrix of the form*

$$\mathbf{G}_{(r,m)} = [\mathbf{I}\ \mathbf{P}^{(\Delta(r,m))}] \quad (3)$$

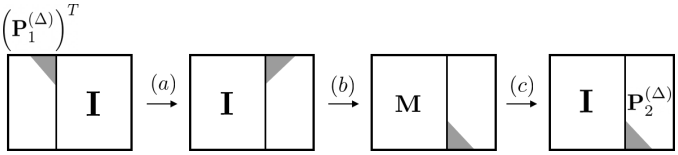

Fig. 1. Sequence of operations in the proof of Lemma 1.

where $\Delta(r,m)$ is given by (2).

Remark. Note that $\mathcal{R}(m,m)$ consists of all binary 2^m-tuples and its only systematic representation is simply $\mathbf{I}_{2^m}$, i.e., $\Delta(m,m) = 0$.

Remark. When $\lfloor \frac{m}{2} \rfloor < r < m$, we have $\Delta(r,m) = \Delta(m-r-1,m)$ which agrees with Lemma 1 since codes $\mathcal{R}(r,m)$ and $\mathcal{R}(m-r-1,m)$ are dual of one another.

Proof: We use an induction argument. Let $\mathcal{P}(r,m)$ be the proposition that $\mathcal{R}(r,m)$ admits a generator of the form given in (3). For the base case we prove $\mathcal{P}(0,m)$ and $\mathcal{P}(1,m)$. In the induction step, assuming $\mathcal{P}(r+1,m)$ and $\mathcal{P}(r,m)$ to be true, we prove $P(r+1,m+1)$ to conclude the proof.

Base Case: For $r = 0$, the generator matrix for $\mathcal{R}(0,m)$ is simply

$$\mathbf{G}_{(0,m)} = [1\ 1\ 1\ \cdots 1]$$

meaning that $\Delta(0,m) = 0\ \forall m$. This agrees with (2), where $\binom{0}{0}$ is understood to be 1. For $r = 1$, we need to show that $\Delta(1,m) = (m+1) - (1+2) = m-2$. This is most easily seen by induction. For the base case, $\mathcal{R}(1,3)$ has the generator matrix

$$\mathbf{G}_{(1,3)} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

The row operation to transform the above to a systematic form is to add rows $2, 3, \cdots, m+1$ to row 1 which gives

$$\mathbf{G}_{(1,3)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

Hence, $\Delta(1,3) = 1 = 3-2$. Now, suppose that the hypothesis is true for $\mathcal{R}(1,m)$. It is known that [19]

$$\mathcal{R}(1,m+1) = \{(\mathbf{u}, \mathbf{1}+\mathbf{u}), \mathbf{u} \in \mathcal{R}(1,m)\}$$

where $\mathbf{1}$ is the all-one codeword. Hence, the generator for $\mathcal{R}(1,m+1)$ after row operations and column permutations can be put in a form shown in Fig. 2. The 'triangle' of zeros for $\mathcal{R}(1,m)$ is of size $m-2$ and shown shaded. Finally, the column vector $[0,0,\cdots,1]^T$ is adjoined to $\mathbf{I}_m$ to obtain a systematic form for $\mathcal{R}(1,m+1)$. When $2^m - (m+1) > m-2$ which indeed holds for $\forall\ m \ge 3$, we see that the size of the triangle is guaranteed to increase by 1, i.e., $\Delta(1,m+1) = \Delta(1,m) + 1 = m-1 = (m+1)-2$ proving that $\mathcal{P}(1,m)$ is true.

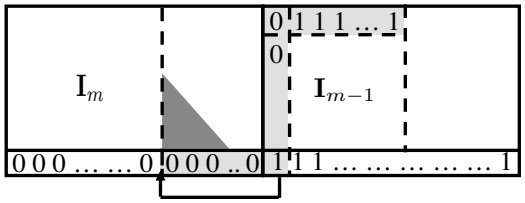

Fig. 2. Proof of base case proposition $\mathcal{P}(1, m)$ of Theorem 1.

Induction Step: Assume that $\mathcal{P}(r+1, m)$ and $\mathcal{P}(r, m)$ are true. By the well-known Plotkin $(\mathbf{u}, \mathbf{u}+\mathbf{v})$ construction, a valid generator for $\mathcal{R}(r+1, m+1)$ is

$$\mathbf{G}_{(r+1,m+1)} = \begin{bmatrix} \mathbf{G}_{(r+1,m)} & \mathbf{G}_{(r+1,m)} \\ \mathbf{0} & \mathbf{G}_{(r,m)} \end{bmatrix}.$$

It is clear that $\mathbf{G}_{(r+1,m+1)}$ can be put into a form shown in Fig. 3.(a) by row operations on the top and bottom block, followed by suitable column permutations. The next steps to obtain a systematic form are illustrated in Fig. 3:

I: Suitable row operations are done to zero out the matrix marked with a crosshatch pattern shown in 3.(a).

II: The newly obtained block $\begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}$ is then moved as shown in 3.(b) resulting in a systematic form shown in 3.(c).

The systematic form obtained admits a 'triangle' of zeros in its parity matrix component as illustrated in Fig. 3.(d). The guaranteed number of consecutive all-zero sub-diagonals, beginning at the lower left end can be seen to be

$$\Delta(r+1, m+1) = \min\{\Delta_1 + q, \Delta_2 + p\}, \tag{4}$$

where $\Delta_1 = \Delta(r+1, m)$, $\Delta_2 = \Delta(r, m)$ and

$$p = \sum_{j=r+2}^{m} \binom{m}{j}, \; q = \sum_{j=0}^{r} \binom{m}{j}.$$

Our goal is to show that the expression in (4) matches with the hypothesis (2) for all $0 \le r < m$. The proof will extensively use the well-known Pascal's identity

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}, \tag{5}$$

and the symmetry property of the binomial coefficients $\binom{n}{k} = \binom{n}{n-k}$ which are recalled here for ease of exposition.

Case 1: $r+1 \le \lfloor \frac{m}{2} \rfloor$

The above assumption implies that $r \le \lfloor \frac{m}{2} \rfloor$ and $r+1 \le \lfloor \frac{m+1}{2} \rfloor$. From (2) then, we have

$$\Delta_1 = \sum_{j=0}^{r+1} \binom{m}{j} - \sum_{j=0}^{r+1} \binom{2j}{j}, \; \Delta_2 = \sum_{j=0}^{r} \binom{m}{j} - \sum_{j=0}^{r} \binom{2j}{j}.$$

Note that

$$(p+\Delta_2) - (q+\Delta_1) = \sum_{j=r+2}^{m} \binom{m}{j} - \sum_{j=0}^{r+1} \binom{m}{j} + \binom{2r+2}{r+1} \overset{(i)}{\ge} 0,$$

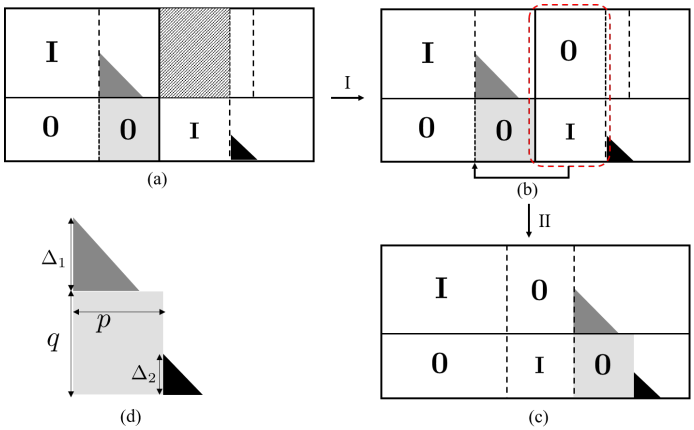

Fig. 3. Steps in the proof of Theorem 1.

where (i) is proved in the following Lemma.

Lemma 2. *For $\ell \le \lfloor \frac{m}{2} \rfloor$, we have that*

$$\sum_{j=\ell+1}^{m} \binom{m}{j} \ge \sum_{j=0}^{\ell} \binom{m}{j} - \binom{2\ell}{\ell}. \tag{6}$$

Proof: We have,

$$\sum_{j=\ell+1}^{m} \binom{m}{j} - \sum_{j=0}^{\ell} \binom{m}{j} = \sum_{j=0}^{m-\ell-1} \binom{m}{j} - \sum_{j=0}^{\ell} \binom{m}{j}.$$

For $\ell \le \lfloor \frac{m-1}{2} \rfloor$, $m-\ell-1 \ge \ell$ meaning that (6) holds. For odd m, $\lfloor \frac{m-1}{2} \rfloor = \lfloor \frac{m}{2} \rfloor$ and there is nothing left to prove. When m is even, say $m = 2q$, the only case left to prove is for $\ell = \lfloor \frac{m}{2} \rfloor = q$. This holds trivially since both the LHS and RHS in (6) simplify to $\sum_{j=q+1}^{2q} \binom{2q}{j}$. ∎

From (4) and (5) then,

$$\Delta(r+1, m+1) = q + \Delta_1 = \sum_{j=0}^{r+1} \binom{m+1}{j} - \sum_{j=0}^{r+1} \binom{2j}{j},$$

settling case 1.

Case 2: $r > \lfloor \frac{m}{2} \rfloor$

In this case, we have $r+1 > \lfloor \frac{m}{2} \rfloor$ and $r+1 > \lfloor \frac{m+1}{2} \rfloor$. From (2),

$$\Delta_1 = \sum_{j=0}^{m-r-2} \binom{m}{j} - \sum_{j=0}^{m-r-2} \binom{2j}{j}$$
$$\Delta_2 = \sum_{j=0}^{m-r-1} \binom{m}{j} - \sum_{j=0}^{m-r-1} \binom{2j}{j}.$$

Here, we have $(p+\Delta_2) - (q+\Delta_1) = -\sum_{j=m-r}^{r} \binom{m}{j} - \binom{2(m-r-1)}{m-r-1} < 0$. From (4) then,

$$\Delta(r+1, m+1) = p + \Delta_2 = \sum_{j=0}^{m-r-1} \binom{m+1}{j} - \sum_{j=0}^{m-r-1} \binom{2j}{j},$$

which is indeed the form in (2) and case 2 is settled.

Case 3: $r \leq \lfloor \frac{m}{2} \rfloor$, $r+1 > \lfloor \frac{m}{2} \rfloor$ and $m = 2s$ even.

The above assumptions simplify to $r = s$. We also have $r+1 = s+1 > s = \lfloor \frac{m+1}{2} \rfloor$. From (2),

$$\Delta_1 = \sum_{j=0}^{s-2} \binom{2s}{j} - \sum_{j=0}^{s-2} \binom{2j}{j}\ ,\ \Delta_2 = \sum_{j=0}^{s} \binom{2s}{j} - \sum_{j=0}^{s} \binom{2j}{j}.$$

We have $(p+\Delta_2) - (q+\Delta_1) = -\binom{2s}{s} - \binom{2(s-1)}{s-1} < 0$. Thus, from (4),

$$\Delta(r+1, m+1) = p + \Delta_2 = \sum_{j=0}^{s-1} \binom{2s+1}{j} - \sum_{j=0}^{s-1} \binom{2j}{j}$$

which agrees with the hypothesis.

Case 4: $r \leq \lfloor \frac{m}{2} \rfloor$, $r+1 > \lfloor \frac{m}{2} \rfloor$ and $m = 2t+1$ odd.

The assumptions imply $r = t$, $r+1 = \lfloor \frac{m+1}{2} \rfloor$ and

$$\begin{aligned} \Delta_1 &= \sum_{j=0}^{t-1} \binom{2t+1}{j} - \sum_{j=0}^{t-1} \binom{2j}{j} \\ \Delta_2 &= \sum_{j=0}^{t} \binom{2t+1}{j} - \sum_{j=0}^{t} \binom{2j}{j} \end{aligned}.$$

We have $(p+\Delta_2) - (q+\Delta_1) = -\binom{2t}{t} < 0$. Hence,

$$\Delta(r+1, m+1) = p + \Delta_2 = \sum_{j=0}^{t+1} \binom{2t+2}{j} - \sum_{j=0}^{t+1} \binom{2j}{j},$$

which is the desired form, hence settling all cases and completing the induction. ∎

B. Asymptotic scaling of $\Delta(r, m)$

In this section, we study how $\Delta(r, m)$ behaves asymptotically. Denote the coding rate by

$$\gamma(r, m) = \frac{K(r, m)}{2^m} = \frac{\sum_{i=0}^{r} \binom{m}{i}}{2^m}. \tag{7}$$

The implication of Theorem 1 is illustrated in Fig. 4. RM codes admit a systematic generator matrix where one can almost fit an all-zero triangle of size

$$\Delta^\star = \min\left(K(r, m), 2^m - K(r, m)\right)$$

in the parity component except that there is a gap $g(r, m)$ from $\Delta^\star$ given by

$$g(r, m) = \begin{cases} \sum_{j=0}^{r} \binom{2j}{j} & \gamma(r, m) \leq 0.5 \\ \sum_{j=0}^{m-r-1} \binom{2j}{j} & \gamma(r, m) > 0.5 \end{cases}. \tag{8}$$

We show that for long RM codes of constant rate, $\frac{g(r,m)}{\Delta^\star} \approx 0$. Note that $\gamma(r, m)$ in (7) can be interpreted to be the probability that a random binary m-tuple has Hamming weight at most r, i.e.,

$$\gamma(r, m) = Pr(X_1 + X_2 + \cdots X_m \leq r) \tag{9}$$

where $\{X_j\}$ are i.i.d. $Ber\left(\frac{1}{2}\right)$. Then, by the central limit theorem, long Reed-muller codes (i.e., $m \to \infty$) of constant rate $0 < \alpha < 1$ can be obtained by letting r to scale with m as

$$r = \frac{m}{2} + \frac{\sqrt{m}}{2}\Phi^{-1}(\alpha) \tag{10}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt \tag{11}$$

is the standard Gaussian CDF.

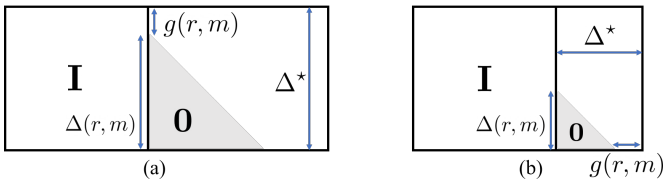

Fig. 4. A systematic form for RM codes where the parity component has a large triangle of zeros. Also shown is the gap $g(r, m)$ from $\Delta^\star$ (Eq. (8)) for code rates (a) $\gamma(r, m) < 0.5$ and (b) $\gamma(r, m) > 0.5$. For long RM codes, $\frac{g(r,m)}{\Delta^\star} \approx 0$.

Theorem 2. *1) For long RM codes $\mathcal{R}(r, m)$ of constant-rate $\alpha < 0.5$ with r scaling as* (10)*, we have*

$$\lim_{m\to\infty} \frac{\Delta(r, m)}{K(r, m)} = \lim_{m\to\infty} \frac{\Delta(r, m)}{\Delta^\star} = 1. \tag{12}$$

2) For long RM codes $\mathcal{R}(r, m)$ of constant-rate $\alpha > 0.5$ with r scaling as (10)*, we have*

$$\lim_{m\to\infty} \frac{\Delta(r, m)}{2^m - K(r, m)} = \lim_{m\to\infty} \frac{\Delta(r, m)}{\Delta^\star} = 1. \tag{13}$$

Proof: For $0 < \alpha < 0.5$, (10) becomes $r = \frac{m}{2} - \beta\frac{\sqrt{m}}{2}$ where $\beta = -\Phi^{-1}(\alpha) > 0$ and (12) is

$$\lim_{m\to\infty} \frac{\Delta(r, m)}{K(r, m)} = 1 - \lim_{m\to\infty} \frac{\sum_{j=0}^{r} \binom{2j}{j}}{K(r, m)}.$$

Since $\frac{K(r,m)}{2^m} \to \alpha$, all that remains to be shown is that $\frac{\sum_{j=0}^{r} \binom{2j}{j}}{2^m} \to 0$. To see this, simply note

$$\frac{\sum_{j=0}^{r} \binom{2j}{j}}{2^m} < \frac{\sum_{j=0}^{r} 4^j}{2^m} < \frac{4}{3} 2^{2r-m} \to 0$$

where the first inequality follows from the identity $4^n = (1+1)^{2n} = \sum_k \binom{2n}{k}$.

When $\alpha > 0.5$, we have $r = \frac{m}{2} + \beta\frac{\sqrt{m}}{2}$ where $\beta = \Phi^{-1}(\alpha) > 0$ and we get

$$\lim_{m\to\infty} \frac{\Delta(r, m)}{2^m - K(r, m)} = \lim_{m\to\infty} 1 - \frac{\sum_{j=0}^{m-r-1} \binom{2j}{j}}{K(m-r-1, m)} = 1,$$

for the same reasons as above, hence proving (13). ∎

Theorem 2 thus implies that asymptotically for constant-rate RM codes, the gap $g(r, m)$ relative to $\Delta^\star$ vanishes and $\Delta(r, m) \approx \Delta^\star$.

III. Feedback Encoding Application

In this section, we introduce the concept of linear block feedback codes and demonstrate an application of our results.

A. Linear Block Feedback Encoding

Consider a memoryless binary symmetric channel (BSC) with an input-output for each channel use i of

$$y_i = c_i \oplus e_i \tag{14}$$

where $y_i \in \{0,1\}$ is the received bit, $c_i \in \{0,1\}$ is the transmitted bit, $e_i \in \{0,1\}$ is memoryless additive noise with $Pr(e_i = 1) = 1 - Pr(e_i = 0) = p$, $\oplus$ is mod-2 addition. Traditional error control coding analysis is *open-loop*, meaning that the transmitted signal is independent of all past, current, and future noise realizations. Instead, consider the problem of *closed-loop* error control coding and assume that the transmitter has causal access to the received signal. This means that prior to channel use i, the transmitter has perfect knowledge of $\{y_j\}_{j=0}^{i-1}$. Because the transmitter perfectly knows $\{c_j\}_{j=0}^{i-1}$, this side information is equivalent to the transmitter having knowledge of the past noise realizations $\{e_j\}_{j=0}^{i-1}$ prior to channel use i.

Inspired by the successes of linear schemes for AWGN and related channels, consider the use of *linear block feedback codes* with encoding function of the form

$$c_i = \phi_{closed,i}\left(\mathbf{m}, \{e_j\}_{j=0}^{i-1}\right) = \sum_{j=0}^{K-1} m_j g_{j,i} \oplus \sum_{\ell=0}^{i-1} e_\ell f_{\ell,i}, \tag{15}$$

where $\mathbf{m} = [m_0, m_1, \cdots, m_{K-1}]$ is the message string and $\{g_{j,i}\}_{j,i},\ \{f_{\ell,i}\}_{\ell,i} \in \{0,1\}$. In particular, expressing (15) in vector form

$$\mathbf{c} = \underbrace{\mathbf{mG}}_{\text{open-loop component}} \oplus \underbrace{\mathbf{eF}}_{\text{noise-shaping}}, \tag{16}$$

$\mathbf{G}$ is the $K \times N$ open-loop generator matrix and $\mathbf{F}$ is a $N \times N$ binary matrix that represents feedback encoding in the form of *linear noise-shaping*. Since c_i cannot possibly depend on future errors $\{e_j\}_{j=i}^{N-1}$, we must have $f_{i,j} = 0\ \forall i \le j$. In other words, causality enforces $\mathbf{F}$ to be strictly upper-triangular. Using (16), the receiver observes

$$\mathbf{y} = \mathbf{c} \oplus \mathbf{e} = \mathbf{mG} \oplus \mathbf{e}\left(\mathbf{I} \oplus \mathbf{F}\right). \tag{17}$$

Eq. (17) indicates transmission of an open-loop codeword over a channel with a special kind of correlated Bernoulli noise. Since $\mathbf{F}$ is strictly upper-triangular, $(\mathbf{I} \oplus \mathbf{F})$ has linearly independent columns and is full rank over $GF(2)$. What is the effect of noise-shaping on the error detection and correction capabilities of the system? Let $\mathbf{H}$ denote a parity-check corresponding to $\mathbf{G}$, and suppose that the receiver calculates the syndrome [19] to be

$$\mathbf{s} = \mathbf{y}\mathbf{H}^T \overset{(a)}{=} \mathbf{e}\widetilde{\mathbf{H}}^T \tag{18}$$

where (a) follows from setting $\widetilde{\mathbf{H}} = \mathbf{H}\left(\mathbf{I} \oplus \mathbf{F}\right)^T$ and noting that $\mathbf{G}\mathbf{H}^T = \mathbf{0}$. From (18), it is clear that the use of feedback allows us to detect and correct error patterns as robustly as an open-loop code with the parity check matrix $\widetilde{\mathbf{H}}$. Since $\mathbf{G}(\mathbf{I} \oplus \mathbf{F})^{-1}\widetilde{\mathbf{H}}^T = \mathbf{G}\mathbf{H}^T = \mathbf{0}$, $\widetilde{\mathbf{G}} = \mathbf{G}(\mathbf{I} \oplus \mathbf{F})^{-1}$ and $\widetilde{\mathbf{H}} = \mathbf{H}\left(\mathbf{I} \oplus \mathbf{F}\right)^T$ can be thought of as the effective open-loop generator and open-loop parity-check matrices respectively that we hope is an improvement when $\mathbf{G}$ or $\mathbf{H}$ is relatively weak.

Suppose now that a system is equipped with a relatively weak (N, K) code with the generator $\mathbf{G}$. Without loss of generality, we assume that the $K \times K$ submatrix of $\mathbf{G}$ comprising of its first K columns is full rank and we can let $\mathbf{G}$ be of the form $\mathbf{G} = [\mathbf{I}_K\ \mathbf{P}_{K\times(N-K)}]$. To improve performance, we employ linear feedback encoding as in (16) by means of a noise-shaping matrix $\mathbf{F}$. We have the following result:

Theorem 3. *By means of noise-shaping, $\mathbf{G} = [\mathbf{I}\ \mathbf{P}]$ can be strengthened to have performance equivalent to that of any desired linear code.*

Proof: The resulting linear block feedback code has performance equivalent to the open-loop linear code with generator $\widetilde{\mathbf{G}} = \mathbf{G}(\mathbf{I} \oplus \mathbf{F})^{-1}$. Suppose that the desired linear code has generator $\mathbf{G}_0 = [\mathbf{I}\ \mathbf{Q}]$. For example, one could choose $\mathbf{G}_0$ to be the systematic generator of a capacity-achieving code. Then, we set

$$\mathbf{F} = \left[\begin{array}{c:c} \mathbf{0}_{K\times K} & \mathbf{Q} \oplus \mathbf{P} \\ \hdashline \mathbf{0}_{(N-K)\times K} & \mathbf{0}_{(N-K)\times(N-K)} \end{array}\right]. \tag{19}$$

For this specific choice, $\mathbf{F}^2 = \mathbf{0}_{N\times N}$ and $(\mathbf{I} \oplus \mathbf{F})^{-1} = (\mathbf{I} \oplus \mathbf{F})$. Hence we have $\widetilde{\mathbf{G}} = \mathbf{G}(\mathbf{I} \oplus \mathbf{F})^{-1} = \mathbf{G}(\mathbf{I} \oplus \mathbf{F}) = [\mathbf{I}\ \mathbf{Q}] = \mathbf{G}_0$ and we are done. ∎

Consider an uncoded system, i.e., $\mathbf{G} = [\mathbf{I}\ \mathbf{0}]$. Suppose we wish to enhance it to a code with generator $\widetilde{\mathbf{G}} = [\mathbf{I}\ \mathbf{Q}]$. As discussed in theorem 3, we set

$$\mathbf{F} = \left[\begin{array}{c:c} \mathbf{0} & \mathbf{Q} \\ \hdashline \mathbf{0} & \mathbf{0} \end{array}\right]. \tag{20}$$

From (16) then, the codewords are given by

$$\begin{aligned} \mathbf{c} &= \mathbf{m}[\mathbf{I}_K\ \mathbf{0}] \oplus \mathbf{e}\left[\begin{array}{c:c} \mathbf{0} & \mathbf{Q} \\ \hdashline \mathbf{0} & \mathbf{0} \end{array}\right] \\ &= [\ \underbrace{m_0\ m_1\ \cdots\ m_{K-1}}_{\text{Systematic Tx}}\ |\ \underbrace{[e_0\ e_1\ \cdots\ e_{K-1}]\mathbf{Q}}_{\text{Noise-Shaping}}\ ] \end{aligned} \tag{21}$$

where e_i is the noise bit that corrupts message bit m_i, for $i = 0, 1, \cdots, K-1$. *Systematic message bit transmission followed by noise-shaping of only the initial error bits is thus sufficient to mimic any arbitrary open-loop code.*

B. Delayed Feedback

In Section III-A, we assumed that any feedback sent after one channel use was immediately available to the transmitter to be used for the next subsequent channel use. Now suppose that there is a delay of Δ channel uses between the time that feedback is sent and when it can be used for encoding at the transmitter. The encoding function changes to

$$c_i = \phi_{closed,i}\left(\mathbf{m}, \{e_j\}_{j=0}^{i-1-\Delta}\right) = \sum_{j=0}^{K-1} m_j g_{j,i} \oplus \sum_{\ell=0}^{i-1-\Delta} e_\ell f_{\ell,i}.$$

In vector form, this is expressed as

$$\mathbf{c} = \underbrace{\mathbf{mG}}_{\text{open-loop component}} \oplus \underbrace{\mathbf{eF}}_{\text{noise-shaping}}$$

where the noise-shaping matrix $\mathbf{F}$ is strictly upper triangular additionally with Δ all-zero diagonals above the main diagonal, i.e., $f_{i,j} = 0 \ \forall i \leq j + k$ for $k = 0, 1, \cdots, \Delta$.

Let $\mathcal{S}^{(\Delta)}$ be the set of all (N, K) open-loop systematic linear codes induced by a generator matrix of the form

$$\mathbf{G} = [\mathbf{I}_K \ \mathbf{P}^{(\Delta)}] \tag{22}$$

as in (1). The class of linear codes that can be emulated from an uncoded system, when there is feedback delay is characterized by the following theorem.

Theorem 4. *Suppose that we start with an uncoded system with $\mathbf{G} = \left[\mathbf{I}_K \ \mathbf{0}_{K\times(N-K)}\right]$. For any choice of linear block feedback encoding with complete causal feedback and Δ units of feedback delay, the new code obtained is effectively equivalent to some code in $\mathcal{S}^{(\Delta)}$. Conversely, every code in the set $\mathcal{S}^{(\Delta)}$ can be emulated with suitable encoding.*

Proof: Let $\mathbf{F}$ be the noise-shaping matrix chosen for encoding. Denoting

$$(\mathbf{I} \oplus \mathbf{F}) = \left[\begin{array}{c|c} \mathbf{J}^{(1)}_{K\times K} & \mathbf{J}^{(2)}_{K\times(N-K)} \\ \hline \mathbf{0}_{(N-K)\times K} & \mathbf{J}^{(3)}_{(N-K)\times(N-K)} \end{array}\right], \tag{23}$$

the effective generator matrix is $\widetilde{\mathbf{G}} = [\mathbf{I} \ \mathbf{J}^{(2)}\mathbf{J}^{(3)^{-1}}]$. Since $\mathbf{F}$ has Δ all-zero diagonals above its main diagonal and $\mathbf{J}^{(3)^{-1}}$ is upper-triangular, the matrix $\mathbf{Q} = \mathbf{J}^{(2)}\mathbf{J}^{(3)^{-1}}$ has Δ all-zero sub-diagonals beginning at the lower left corner. In other words, $\widetilde{\mathbf{G}}$ essentially is of the form $\widetilde{\mathbf{G}} = [\mathbf{I} \ \mathbf{P}^{(\Delta)}]$ and the first part of the theorem is proved. To see the converse, consider emulating an arbitrary code from $\mathcal{S}^{(\Delta)}$ with generator $\widetilde{\mathbf{G}} = [\mathbf{I} \ \mathbf{P}_1^{(\Delta)}]$, and note that a valid choice of feedback encoding is to simply set

$$\mathbf{F} = \left[\begin{array}{c|c} \mathbf{0} & \mathbf{P}_1^{(\Delta)} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right].$$

■

From Theorem 4, an uncoded system can be transformed to a code of the form (22) (and no better) when there is a delay of Δ units in the feedback link. A central question is thus whether there exists a good code or a good class of codes that has a generator of the form (22) with a possibly large Δ. *The larger the value of Δ, the larger is the feedback delay that can be tolerated*. Our results in Theorems 1 and 2 imply that RM codes prove to be a good candidate in that they have excellent performance and admit very large values of Δ. *Thus, a RM code can be emulated from an uncoded system by means of linear noise-shaping against a feedback delay that is nearly as large as $\Delta^\star = \min(K, N - K)$.*

IV. Conclusion

In this work, we proved a novel result for RM codes, showing that they admit a systematic generator matrix whose parity component has a rather large number of contiguous all-zero sub-diagonals. We then introduced the concept of linear block feedback encoding with and without feedback delay and showed that it can be employed to emulate a stronger code from a weaker one. Our result on RM codes implies that they can be emulated from an uncoded system against large feedback delays. An interesting open question is whether for finite r and m, $\mathcal{R}(r, m)$ admits a systematic form (1) with a Δ larger than what is proved in Theorem 1.

References

[1] E. Abbe, A. Shpilka, and A. Wigderson, "Reed–Muller codes for random erasures and errors," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5229–5252, 2015.

[2] 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.

[3] S. Kudekar, S. Kumar, M. Mondelli, H. D. Pfister, E. Şaşoğlu, 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] E. Arikan, "A performance comparison of polar codes and Reed-Muller codes," *IEEE Communications Letters*, vol. 12, no. 6, pp. 447–449, 2008.

[5] M. Mondelli, S. H. Hassani, and R. L. Urbanke, "From polar to Reed-Muller codes: A technique to improve the finite-length performance," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3084–3091, 2014.

[6] O. Sberlo and A. Shpilka, "On the performance of Reed-Muller codes with respect to random errors and erasures," in *Proceedings of the Thirty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, 2020, pp. 1357–1376.

[7] E. Arikan, "Systematic polar coding," *IEEE communications letters*, vol. 15, no. 8, pp. 860–862, 2011.

[8] P. Hauck, M. Huber, J. Bertram, D. Brauchle, and S. Ziesche, "Efficient majority-logic decoding of short-length Reed-Muller codes at information positions," *IEEE transactions on communications*, vol. 61, no. 3, pp. 930–938, 2013.

[9] E. Abbe, A. Shpilka, and M. Ye, "Reed-Muller codes: Theory and algorithms," *arXiv preprint arXiv:2002.03317*, 2020.

[10] M. Horstein, "Sequential transmission using noiseless feedback," *IEEE Transactions on Information Theory*, vol. 9, no. 3, pp. 136–143, 1963.

[11] C. T. Li and A. El Gamal, "An efficient feedback coding scheme with low error probability for discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 61, no. 6, pp. 2953–2963, 2015.

[12] M. Agrawal, Z. Chance, D. J. Love, and V. Balakrishnan, "Using channel output feedback to increase throughput in Hybrid-ARQ," *IEEE transactions on signal processing*, vol. 60, no. 12, pp. 6465–6480, 2012.

[13] O. Shayevitz and M. Feder, "Optimal feedback communication via posterior matching," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1186–1222, 2011.

[14] M. Agrawal, D. J. Love, and V. Balakrishnan, "Communicating over filter-and-forward relay networks with channel output feedback," *IEEE Transactions on Signal Processing*, vol. 64, no. 5, pp. 1117–1131, 2015.

[15] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–i: No bandwidth constraint," *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 172–182, April 1966.

[16] J. Schalkwijk, "A coding scheme for additive noise channels with feedback–ii: Band-limited signals," *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 183–189, April 1966.

[17] S. Butman, "A general formulation of linear feedback communication systems with solutions," *IEEE Transactions on Information Theory*, vol. 15, no. 3, pp. 392–400, May 1969.

[18] Z. Chance and D. J. Love, "Concatenated coding for the AWGN channel with noisy feedback," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6633–6649, Oct 2011.

[19] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*. Elsevier, 1977, vol. 16.