

# Vanishing Spaces of Random Sets and Applications to Reed-Muller Codes

Siddharth Bhandari\*    Prahladh Harsha<sup>†</sup>    Ramprasad Saptharishi<sup>‡</sup>  
Srikanth Srinivasan<sup>§</sup>

## Abstract

We study the following natural question on random sets of points in  $\mathbb{F}_2^m$ :

Given a random set of  $k$  points  $Z = \{z_1, z_2, \dots, z_k\} \subseteq \mathbb{F}_2^m$ , what is the dimension of the space of degree at most  $r$  multilinear polynomials that vanish on all points in  $Z$ ?

We show that, for  $r \leq \gamma m$  (where  $\gamma > 0$  is a small, absolute constant) and  $k = (1 - \varepsilon) \cdot \binom{m}{\leq r}$  for any constant  $\varepsilon > 0$ , the space of degree at most  $r$  multilinear polynomials vanishing on a random set  $Z = \{z_1, \dots, z_k\}$  has dimension exactly  $\binom{m}{\leq r} - k$  with probability  $1 - o(1)$ . This bound shows that random sets have a much smaller space of degree at most  $r$  multilinear polynomials vanishing on them, compared to the worst-case bound (due to Wei (IEEE Trans. Inform. Theory, 1991)) of  $\binom{m}{\leq r} - \binom{\log_2 k}{\leq r} \gg \binom{m}{\leq r} - k$ .

Using this bound, we show that high-degree Reed-Muller codes (RM( $m, d$ )) with  $d > (1 - \gamma)m$  “achieve capacity” under the Binary Erasure Channel in the sense that, for any  $\varepsilon > 0$ , we can recover from  $(1 - \varepsilon) \cdot \binom{m}{\leq m-d-1}$  random erasures with probability  $1 - o(1)$ . This also implies that RM( $m, d$ ) is also efficiently decodable from  $\approx \binom{m}{\leq m-(d/2)}$  random errors for the same range of parameters.

## 1 Introduction

The Reed-Muller (RM) code is one of the most basic error-correcting codes studied in coding theory, first introduced by Muller [?] and Reed [?] in 1954. Stated in the language of polynomials, the RM code with parameters  $m$  and  $r$  for positive integers  $m > r$ , denoted by RM( $m, r$ ), is the code

\*Simons Institute for the Theory of Computing, Berkeley, USA, [siddharth.bhandari@berkeley.edu](mailto:siddharth.bhandari@berkeley.edu). Research supported by the Simons-Berkeley Postdoctoral Fellowship.

<sup>†</sup>Tata Institute of Fundamental Research, Mumbai, India, [prahladh@tifr.res.in](mailto:prahladh@tifr.res.in). Research done when the author was visiting the Simons Institute for the Theory of Computing. Research supported in part by the Department of Atomic Energy, Government of India, under project 12-R&D-TFR-5.01-0500 and the Swarnajayanti Fellowship.

<sup>‡</sup>Tata Institute of Fundamental Research, Mumbai, India, [ramprasad@tifr.res.in](mailto:ramprasad@tifr.res.in). Research supported by the Department of Atomic Energy, Government of India, under project 12-R&D-TFR-5.01-0500 and the Ramanujan Fellowship of the DST.

<sup>§</sup>Aarhus University, Aarhus, Denmark, [srikanth@cs.au.dk](mailto:srikanth@cs.au.dk). Supported by Startup grant from Aarhus University.

whose codewords are evaluations of  $m$ -variate multilinear polynomials of degree at most  $r$  over the vector space  $\mathbb{F}_2^m$ . Despite these being one of the earliest codes discovered in coding theory, several properties of these codes (weight-distribution, capacity-achieving on any binary memory-less symmetric (BMS) channel) are not yet fully-understood.

A natural problem that arises while investigating the Shannon-capacity of RM code is the following:

*Given a set  $Z$  of  $k$  points  $Z = \{z_1, z_2, \dots, z_k\} \subseteq \mathbb{F}_2^m$ , what is the dimension of the space  $\mathbb{I}_r(Z)$  of degree at most  $r$  multilinear polynomials that vanish on all points in  $Z$ ?*

To understand the connection to RM codes, we give an equivalent description of this problem in terms of the parity check matrix of the RM code. Let  $E(m, r)$  be the  $\binom{m}{\leq r} \times 2^m$ -matrix whose columns are indexed by elements in  $\mathbb{F}_2^m$  and rows by  $m$ -variate multilinear monomials of degree at most  $r$ . The  $z^{\text{th}}$  column of  $E(m, r)$  for  $z \in \mathbb{F}_2^m$  is the column vector  $z^{(r)}$  consisting of evaluations of all multilinear  $m$ -variate degree  $\leq r$  monomials at the point  $z$ . It is not hard to see that  $E(m, r)$  is the parity check matrix of the Reed-Muller code  $\text{RM}(m, m - r - 1)$ . An equivalent formulation of the above question in terms of the matrix  $E(m, r)$  is the following: given a set  $Z$  of  $k$  points  $Z = \{z_1, z_2, \dots, z_k\} \subseteq \mathbb{F}_2^m$ , what is the rank of the  $\binom{m}{\leq r} \times |Z|$  sub-matrix  $E(m, r)_Z$  of  $E(m, r)$  obtained by picking the columns corresponding to the points in  $Z$ .

While studying the generalized Hamming weights of RM codes over  $\mathbb{F}_2$ , Wei [?] proved worst-case bounds for the above problem. In its simplest form, these worst-case bounds show that if  $|Z| = 2^\ell$  for some  $0 \leq \ell \leq m$ , then  $\dim(\mathbb{I}_r(Z)) \leq \binom{m}{\leq r} - \binom{\ell}{\leq r}$ . These worst-case bounds were then generalized to arbitrary fields by Keevash and Sudakov [?] and have had applications in combinatorics as well as theoretical computer science. Ben-Eliezer, Hod and Lovett [?] reproved the above form of this bound and used it to study the correlation of random polynomials with lower-degree polynomials. Nie and Wang [?] used these bounds bound to prove extensions of the Kakeya Theorem. Chen, De, and Vijayaraghavan [?] used these bounds to analyze their algorithms for learning mixtures of subspaces over finite fields. In the context of RM codes, Abbe, Shpilka and Wigderson [?] used these worst-case bounds to show that RM codes of high rate ( $r = m - o(\sqrt{m/\log m})$ ) achieve Shannon-capacity over the BEC channel.

These bounds due to Wei only prove extremal (i.e., worst-case) bounds on  $\dim(\mathbb{I}_r(Z))$ . The tightness of this bound is witnessed when  $Z$  is a subcube of size  $2^\ell$ , where  $\dim(\mathbb{I}_r(Z)) = \binom{m}{\leq r} - \binom{\ell}{\leq r}$  which is considerably larger than  $\binom{m}{\leq r} - |Z| = \binom{m}{\leq r} - 2^\ell$ . We note that  $\binom{m}{\leq r} - |Z|$  is a lower bound on the dimension of  $\mathbb{I}_r(Z)$  since each point in  $Z$  can reduce the dimension by at most 1. It is thus natural to ask how large is  $\dim(\mathbb{I}_r(Z))$  for sets  $Z$  other than a subcube. In particular, how does  $\dim(\mathbb{I}_r(Z))$  compare to  $\binom{m}{\leq r} - |Z|$  when  $Z$  is a random set of points of size  $K < \binom{m}{\leq r}$ . Our main result shows that for small  $r$  (more precisely, for  $r \leq \gamma m$  for a small, but absolute, constant  $\gamma > 0$ ) and  $|Z| = (1 - \varepsilon) \cdot \binom{m}{\leq r}$  for any constant  $\varepsilon \in (0, 1)$ , a random set  $Z$  behaves very differently from a subcube and the dimension of  $\mathbb{I}_r(Z)$  is as small as it can be, namely  $\binom{m}{\leq r} - |Z|$ .

**Theorem 1.1** (dimension of degree- $r$  vanishing space of random set). *There exists a constant  $\gamma_0 > 0$  such that for all  $\varepsilon > 0$  the following is true. Let  $K = (1 - \varepsilon) \cdot \binom{m}{\leq r}$  and  $r < \gamma_0 m$ . Then,*

$$\Pr_{Z \subseteq \mathbb{F}_2^m} \left[ \dim(\mathbb{I}_r(Z)) = \binom{m}{\leq r} - K \right] = 1 - o(1)$$

where  $Z = \{z_1, \dots, z_K\}$  is a uniformly random set of  $K$  distinct points in  $\mathbb{F}_2^m$ . In other words, the probability that the set of columns  $\{z_1^{(r)}, z_2^{(r)}, \dots, z_K^{(r)}\}$  is not linearly independent is  $o(1)$ , where  $z^{(r)}$  denotes the vector consisting of evaluations of all multilinear  $m$ -variate degree  $\leq r$  monomials at the point  $z$ .

Abbe, Shpilka and Wigderson [?] also proved a similar result for a smaller range of  $r$ , namely  $r = o(\sqrt{m/\log m})$  using Wei's extremal bound. We note that if  $\gamma_0$  is chosen sufficiently small such that  $K \leq \binom{m}{\leq r} = o(2^{m/2})$ , then it does not matter if the points  $z_1, \dots, z_K$  are chosen with or without replacement since the two distributions are then  $o(1)$ -indistinguishable. When the points are chosen with replacement, then one cannot expect to improve the range of  $r$  for which the above theorem holds (up to the choice of the constant  $\gamma_0$ ) since if  $K \gg 2^{m/2}$ , then with high probability two of the  $z_i$ 's are equal in which case  $\mathbb{I}_r(Z)$  will definitely have dimension strictly larger than  $\binom{m}{\leq r} - K$ .

As one would expect, this average-case bound on  $\dim(\mathbb{I}_r(Z))$  leads to a better understanding of the Shannon-capacity of RM codes over the high-rate regime.

**Shannon-capacity of Reed-Muller codes** Reed-Muller codes were shown to achieve Shannon-capacity over the BEC channel in the extremal regimes (low rate  $r = o(m)$  and high rate  $r = m - o(m)$ ) by Abbe, Shpilka and Wigderson [?] and in the constant rate regime ( $r = m/2 \pm O(\sqrt{m})$ ) by Kudekar et al [?]. More recently, Reeves and Pfister [?] showed that RM codes over the constant rate regime ( $r = m/2 \pm O(\sqrt{m})$ ) achieve Shannon-capacity over any BMS channel. Despite all this progress, the general question of whether RM codes achieve capacity over the entire regime ( $1 \leq r < m - 1$ ) and over any BMS channel remains open.

The results of Kudekar et al [?] and Reeves and Pfister [?] are obtained using Boolean function analysis while the results of Abbe, Shpilka and Wigderson for the Shannon-capacity over the BEC channel of RM codes in the low-rate regime ( $r = o(m)$ ) were obtained using the weight-distribution bound of Kaufman, Lovett and Porat [?]. Subsequent improvements in the weight distribution due to Sberlo and Shpilka [?] led to the Shannon-capacity of RM codes over the BEC channel for a wider range of the degree parameter  $r$ , more precisely for  $r = \gamma m$  for  $\gamma < 1/70$ .

The Shannon-capacity of RM over the high-rate regime ( $r = m - o(\sqrt{m/\log m})$ ) were also proved by Abbe, Shpilka and Wigderson using Wei's worst-case bounds on  $\dim(\mathbb{I}_r(Z))$ . In particular, these latter results do not use the weight distribution of RM codes and hence the subsequent improvements in our understanding of the weight distribution of RM codes due to Samorodnitsky [?] and Sberlo-Shpilka [?] did not lead to an improved understanding of the Shannon-capacity

of RM codes in the high-rate regime. We rectify this gap by giving an alternate bound for the Shannon-capacity in terms of the our average-case bound on  $\dim(\mathbb{I}_r(Z))$  and widen the range of the degree parameter  $r$  for which high-rate RM codes achieve Shannon-capacity over the BEC channel.

**Theorem 1.2** (high-degree RM codes under erasures). *There exists a constant  $\gamma_0 > 0$  such that for all  $\varepsilon > 0$  the following is true. Let  $m, d$  be growing parameters with  $d > m(1 - \gamma_0)$ . Then, the code  $\text{RM}(m, d)$  can correct  $K = (1 - \varepsilon) \binom{m}{\leq m-d-1}$  random errors with probability  $1 - o(1)$ .*

Abbe, Shpilka and Wigderson [?] and Saptharishi, Shpilka and Volk [?] showed how to reduce the resilience of certain RM codes under the BSC to the resilience of appropriate RM codes under the BEC. Using this reduction, we obtain the following corollary of the above theorem for the BSC channel.

**Corollary 1.3** (high-degree RM codes under errors). *There exists a constant  $\gamma_0 > 0$  such that for all  $\varepsilon > 0$  the following is true. Let  $m, r$  be growing parameters with  $r < \gamma_0 m$ . Then, the code  $\text{RM}(m, m - 2r - 2)$  can be efficiently decode from  $K = (1 - \varepsilon) \binom{m}{\leq r}$  random errors.*

See [Section 4](#) for further discussion on Shannon-capacity and what it means for high-rate RM codes to achieve it over the BEC and BSC channels.

We remark that our proof methods work for any linear code, not necessarily the RM code, provided one has good bounds on the weight distribution of the dual code.

## 1.1 Proof Overview

Recall that [Theorem 1.1](#) is a statement about the dimension of the degree- $r$  vanishing space of a uniformly random subset  $Z$  of  $\mathbb{F}_2^m$  of size  $K$  where  $K = (1 - \varepsilon) \cdot \binom{m}{\leq r}$ . In the regime of parameters we are interested in (i.e.,  $r < \gamma_0 m$  for a small enough constant  $\gamma_0$ ), this is more or less equivalent to<sup>1</sup> choosing  $K$  points  $z_1, \dots, z_K$  independently and uniformly from  $\mathbb{F}_2^m$ . We assume that  $Z$  is defined this way for the rest of this section.

To argue about the dimension of the degree- $r$  vanishing space  $\mathbb{I}_r(Z)$ , we instead argue about the size  $S$  of  $\mathbb{I}_r(Z)$ . Note that since  $\dim(\mathbb{I}_r(Z)) \geq D := \binom{m}{\leq r} - K$ , this set has size at least  $2^D$ , and has size at least  $2^{D+1}$  if  $\dim(\mathbb{I}_r(Z)) > D$ . In light of this, it is sufficient to show that

$$\mathbb{E}[S] = \exp_2(D)(1 + o(1)).$$

Estimating  $\mathbb{E}[S]$  turns out to be very closely related to a recent result of Sberlo and Shpilka [?], who prove strong results on the parameters of capacity-achieving Reed-Muller codes in the low-degree setting.

---

<sup>1</sup>In the sense that the two distributions have small statistical distance.

More precisely, we can easily see that the probability that a uniformly random polynomial  $P \in \text{RM}(m, r)$  belongs to  $\mathbb{I}_r(Z)$  is exactly  $(1 - \text{wt}(P))^K$  where  $\text{wt}(P)$  is the *fractional Hamming weight* of  $P$  (i.e., the fraction of points where it does not vanish). Thus, we have

$$\mathbb{E}[S] = \sum_{P \in \text{RM}(m, r)} (1 - \text{wt}(P))^K.$$

Now, while there are polynomials  $P \in \text{RM}(m, r)$  of very small weight<sup>2</sup>, *most* polynomials in  $\text{RM}(m, r)$  have weight close to  $1/2$ , which should indicate that the sum is close to  $2^{-K}$  as required. However, a careful analysis is required, as the contribution of a polynomial  $P$  increases exponentially as its weight decreases.

It turns out that Sberlo and Shpilka [?] analyzed a very similar quantity in their recent work. More precisely they showed that for any constant  $\delta > 0$  (and also a large range of sub-constant  $\delta$ ), we have

$$\sum_{P \in \text{RM}(m, r) \setminus \{0\}} (1 - \text{wt}(P))^{(1+\delta) \cdot \binom{m}{\leq r}} = o(1).$$

While this is very closely related to our result, we are not able to recover the exact bound we need using this inequality.

However, the technical lemmas used to derive the above result are strong results on the *weight distribution* of  $\text{RM}(m, r)$ , which are upper bounds on the number of polynomials of small weight. Using these upper bounds and carrying out the relevant computations yields [Theorem 1.1](#).

The application to resilience under the BEC and BSC ([Theorem 1.2](#) and [Corollary 1.3](#)) follow from [Theorem 1.1](#) in a straight-forward manner from the works of Abbe, Shpilka and Wigderson [?] and Saptharishi, Shpilka and Volk [?].

## Organisation

We discuss some notation and preliminaries in [Section 2](#) and proceed to the proof of [Theorem 1.1](#) in [Section 3](#). We then present the applications to proving the resilience of RM codes under BEC and BSC in [Section 4](#).

---

<sup>2</sup>It is a standard fact (e.g. by the Schwartz-Zippel lemma) that the minimum weight of a non-zero polynomial from  $\text{RM}(m, r)$  is  $2^{-r}$ .

## 2 Notation and preliminaries

1. We use  $[m]$  to represent the set  $\{1, 2, 3, \dots, m\}$  and the expression  $\binom{m}{\leq r}$  to denote the sum

$$\binom{m}{0} + \dots + \binom{m}{r}.$$

2. We denote by  $\binom{[m]}{r}$  and  $\binom{[m]}{\leq r}$  the sets  $\{S \subseteq [m]: |S| = r\}$  and  $\{S \subseteq [m]: |S| \leq r\}$  respectively.
3. We shall abuse notation and use  $\text{RM}(m, r)$  to denote the vector space of all  $m$ -variate multivariate polynomials in  $\mathbb{F}_2[x_1, \dots, x_m]$  of degree at most  $r$ . Throughout the paper, the parameter  $m$  will be unchanged and we will avoid mentioning it for brevity. Let  $n = 2^m$ .
4. For parameters  $m, r$ , define the matrix  $E(m, r)$  as the  $\binom{m}{\leq r} \times 2^m$ -matrix whose columns are indexed by elements in  $\mathbb{F}_2^m$  and rows by multilinear  $m$ -variate monomials of degree at most  $r$  and whose  $z$ -th column is the vector consisting of the *evaluation* of all multilinear  $m$ -variate degree  $\leq r$  monomials on  $z$ .
5. For a polynomial  $P \in \text{RM}(m, r)$ , we use  $\text{wt}(P)$  to denote the *fractional Hamming weight*:

$$\text{wt}(P) := \frac{|\{z \in \mathbb{F}_2^m : P(z) \neq 0\}|}{2^m}$$

We shall say that a polynomial  $P \in \text{RM}(m, r)$  is  $\eta$ -biased if  $|\text{wt}(P) - 1/2| \geq \eta$ . We shall also use  $\text{RM}_\eta(m, r)$  to denote the set of all  $\eta$ -biased polynomials in  $\text{RM}(m, r)$ .

6. For a real number  $\alpha \in (0, 1)$ , we denote by  $\text{WtDist}_{m,r}(\alpha)$  the number of polynomials of (fractional) weight at most  $\alpha$  in  $\text{RM}(m, r)$ , i.e.,  $\text{WtDist}_{m,r}(\alpha) := |\{P \in \text{RM}(m, r) : \text{wt}(P) \leq \alpha\}|$ .
7. All complexity notations used in the paper are with respect to  $m$  as the growing parameter.

### 2.1 Weight distribution bounds

We use the following bounds on the weight-distribution of Reed-Muller codes due to Sberlo and Shpilka [?].

**Theorem 2.1** (Sberlo and Shpilka [?]: bounds for low-weight codewords). *For any  $\ell \geq 1$ , we have*

$$\text{WtDist}_{m,r}(2^{-\ell}) \leq \exp_2 \left( O(m^4) + 17 \cdot (c_\gamma \ell + d_\gamma) \cdot \gamma^{\ell-1} \cdot \binom{m}{\leq r} \right)$$

where  $c_\gamma = 1/(1 - \gamma)$  and  $d_\gamma = \frac{(2-\gamma)}{(1-\gamma)^2}$ .

In particular, if  $\gamma \leq (1/2)$ , we have

$$\text{WtDist}_{m,r}(2^{-\ell}) \leq \exp_2 \left( O(m^4) + O(\ell \cdot \gamma^{\ell-1}) \cdot \binom{m}{\leq r} \right). \quad (2.2)$$

**Theorem 2.3** (Sberlo and Shpilka [?]: bounds for medium-weight codewords). *Assume that  $\gamma \in (0, (1/2) - \Omega(1))$ . For a positive integer  $\ell$  such that  $\ell/m$  upper bounded by a small enough constant<sup>3</sup>, we have*

$$\text{WtDist}_{m,r} \left( \frac{1}{2} - 2^{-\ell} \right) \leq \exp_2 \left( O(m^4) + (1 - 2^{-c(\gamma,\ell)}) \cdot \binom{m}{\leq r} \right) \quad (2.4)$$

where  $c(\gamma, \ell) = O(\max\{\gamma^2 \ell, \gamma\})$ .

### 3 Degree- $r$ vanishing spaces and closures

We first define the notion of a *vanishing space*. This is similar to the notion of *vanishing ideals* in basic algebraic geometry but we refer to them as *vanishing space* instead to stress that we are studying them as a vector space and not an ideal.

**Definition 3.1** (degree- $r$  vanishing spaces). *For a set  $Z \subseteq \mathbb{F}_2^m$ , we use  $\mathbb{I}_r(Z)$  to denote the degree- $r$  vanishing space defined as*

$$\mathbb{I}_r(Z) := \{P \in \text{RM}(m, r) : P(z) = 0 \text{ for all } z \in Z\}. \quad \diamond$$

Related to the vanishing ideals is also the notion of the *degree- $r$  closure* (similar to the Zariski closure in standard algebraic geometry but restricted to the setting of  $\mathbb{F}_2^m$ ), which is the set of points on which every polynomial in the degree- $r$  vanishing space vanishes.

**Definition 3.2** (degree- $r$  closure). *For a set  $Z \subseteq \mathbb{F}_2^m$ , we use  $\text{Closure}_r(Z)$  to denote the degree- $r$  closure of  $Z$  defined as*

$$\text{Closure}_r(Z) := \{u \in \mathbb{F}_2^m : P(u) = 0 \text{ for all } P \in \mathbb{I}_r(Z)\}. \quad \diamond$$

The above notion can be equivalently defined as the set of all  $u \in \mathbb{F}_2^m$  such that column of  $E(m, r)$  indexed by  $u$  is in the span of columns of  $E(m, r)$  indexed by  $z \in Z$ . That is,

$$\text{Closure}_r(Z) = \left\{ u \in \mathbb{F}_2^m : u^{(r)} \in \text{span} \left\{ z^{(r)} : z \in Z \right\} \right\},$$

where  $z^{(r)}$  denotes the  $\binom{m}{\leq r}$ -dimension vector of evaluations of all  $m$ -variate degree at most  $r$  mono-

---

<sup>3</sup>[?, Theorem 1.3] only guarantees this for  $\ell = o(m)$  but [?, Remark 1.1] after the theorem statement says that in this setting it holds for  $\ell = \Omega(m)$ .

mials at the point  $z$ .

For any set  $Z \subseteq \mathbb{F}_2^m$  of size  $k$ , note that  $\mathbb{I}_r(Z)$  is a vector space of dimension at least  $\binom{m}{\leq r} - k$ , as each constraint  $P(z) = 0$  adds one homogeneous linear constraint on the ambient space  $\text{RM}(m, r)$ . In fact, it is easy to see that  $\mathbb{I}_r(Z)$  has rank  $\binom{m}{\leq r} - k$  if and only if each point of  $Z$  is not in the closure of the previous points.

**Observation 3.3** (vanishing ideals of minimal rank). *Let  $Z = \{z_1, \dots, z_k\} \subseteq \mathbb{F}_2^m$ . Then,*

$$\dim \mathbb{I}_r(Z) = \binom{m}{\leq r} - k \iff (\forall i = 1, \dots, k-1 : z_i \notin \text{Closure}_r(\{z_1, \dots, z_{i-1}\})). \quad \square$$

### 3.1 Dimension of the degree- $r$ vanishing spaces of random sets

In this section, we will be interested in studying the vanishing spaces and closures of a random set  $Z$  obtained by picking  $K = (1 - \varepsilon)\binom{m}{\leq r}$  points from  $\mathbb{F}_2^m$  (where  $\varepsilon > 0$  is some constant). We restate [Theorem 1.1](#) below.

**Theorem 1.1** (dimension of degree- $r$  vanishing space of random set). *There exists a constant  $\gamma_0 > 0$  such that for all  $\varepsilon > 0$  the following is true. Let  $K = (1 - \varepsilon) \cdot \binom{m}{\leq r}$  and  $r < \gamma_0 m$ . Then,*

$$\Pr_{Z \subseteq \mathbb{F}_2^m} \left[ \dim(\mathbb{I}_r(Z)) = \binom{m}{\leq r} - K \right] = 1 - o(1)$$

where  $Z = \{z_1, \dots, z_K\}$  is a uniformly random set of  $K$  distinct points in  $\mathbb{F}_2^m$ . In other words, the probability that the set of columns  $\{z_1^{(r)}, z_2^{(r)}, \dots, z_K^{(r)}\}$  is not linearly independent is  $o(1)$ , where  $z^{(r)}$  denotes the vector consisting of evaluations of all multilinear  $m$ -variate degree  $\leq r$  monomials at the point  $z$ .

We will use the following lemma to prove the above theorem.

**Lemma 3.4** (size of degree  $r$ -vanishing space of random set). *There exists a constant  $\gamma_0 > 0$  such that for all  $\varepsilon > 0$  the following is true. Let  $K = (1 - \varepsilon) \cdot \binom{m}{\leq r}$  and  $r < \gamma_0 m$ . Then,*

$$\mathbb{E} \left[ \frac{|\mathbb{I}_r(Z)|}{\exp_2 \left( \binom{m}{\leq r} \right)} \right] = 2^{-K} (1 + o(1))$$

where  $Z = \{z_1, \dots, z_K\}$  is a uniformly random set of  $K$  distinct points in  $\mathbb{F}_2^m$ .

We first use the above lemma to prove [Theorem 1.1](#). [Lemma 3.4](#) is proved in [Section 3.2](#).

*Proof of Theorem 1.1.* Let  $p$  be the probability that  $\dim(\mathbb{I}_r(Z)) = \binom{m}{\leq r} - K$ . Note that  $\dim(\mathbb{I}_r(Z))$  is always at least  $\binom{m}{\leq r} - K$ , as the space  $\mathbb{I}_r(Z)$  is a subspace of  $\text{RM}(m, r)$  defined by  $K$  linear equations.



Thus,

$$\begin{aligned} \mathbb{E} \left[ \frac{|\mathbb{I}_r(Z)|}{\exp_2 \binom{m}{\leq r}} \right] &= \mathbb{E} \left[ \exp_2 \left( \dim(\mathbb{I}_r(Z)) - \binom{m}{\leq r} \right) \right] \\ &\geq p \cdot \exp_2(-K) + (1-p) \cdot 2 \cdot \exp_2(-K). \end{aligned}$$

Combining this with [Lemma 3.4](#) we get that

$$p \cdot \exp_2(-K) + (1-p) \cdot 2 \cdot \exp_2(-K) \leq 2^{-K}(1 + o(1))$$

which gives  $1-p \leq o(1)$ . Hence, we get that  $p = 1 - o(1)$ . □

### 3.2 Proof of [Lemma 3.4](#)

We start with reformulating the problem of bounding size of  $\mathbb{I}_r(Z)$  to computing a certain weighted sum of the polynomials in  $\text{RM}(m, r)$ .

Let  $Q$  be a uniformly random polynomial chosen from  $\text{RM}(m, r)$ . Then,

$$\begin{aligned} \mathbb{E} \left[ \frac{|\mathbb{I}_r(Z)|}{\exp_2 \binom{m}{\leq r}} \right] &= \mathbb{E}_{Z, Q} [\mathbb{1}[Q \in \mathbb{I}_r(Z)]] \\ &= \mathbb{E}_Q \left[ \mathbb{E}_Z [\mathbb{1}[Q \in \mathbb{I}_r(Z)]] \right] \\ &= \sum_{P \in \text{RM}(m, r)} \exp_2 \left( -\binom{m}{\leq r} \right) \cdot \frac{\binom{(1-\text{wt}(P))2^m}{K}}{\binom{2^m}{K}} \\ &= \exp_2(-K) \cdot \left[ \sum_{P \in \text{RM}(m, r)} \exp_2 \left( -\varepsilon \cdot \binom{m}{\leq r} \right) \cdot \frac{\binom{(1-\text{wt}(P))2^m}{K}}{\binom{2^m}{K}} \right]. \end{aligned}$$

To complete the proof of [Lemma 3.4](#) we therefore need the following claim.

**Claim 3.5.**

$$\sum_{P \in \text{RM}(m, r)} \exp_2 \left( -\varepsilon \cdot \binom{m}{\leq r} \right) \cdot \frac{\binom{(1-\text{wt}(P))2^m}{K}}{\binom{2^m}{K}} \leq 1 + o(1).$$

Given [Claim 3.5](#), we see that

$$\begin{aligned} \mathbb{E} \left[ \frac{|\mathbb{I}_r(Z)|}{\exp_2 \binom{m}{\leq r}} \right] &= \exp_2(-K) \cdot \left[ \sum_{P \in \text{RM}(m, r)} \exp_2 \left( -\varepsilon \cdot \binom{m}{\leq r} \right) \cdot \frac{\binom{(1-\text{wt}(P))2^m}{K}}{\binom{2^m}{K}} \right] \\ &\leq \exp_2(-K) \cdot (1 + o(1)), \end{aligned}$$

which concludes the proof of [Lemma 3.4](#).

Next, we proceed to prove [Claim 3.5](#).

*Proof of Claim 3.5.* Let  $u = \binom{m}{\leq r}$  in the following. Recall that  $K = (1 - \varepsilon)u$ . Also,

$$\sum_{P \in \text{RM}(m,r)} \exp_2(-\varepsilon u) \cdot \frac{\binom{(1-\text{wt}(P))2^m}{K}}{\binom{2^m}{K}} \leq \sum_{P \in \text{RM}(m,r)} \exp_2(-\varepsilon u) \cdot (1 - \text{wt}(P))^{(1-\varepsilon)u}.$$

Set  $\delta = \left(\frac{1}{\binom{m}{\leq r}}\right)^2$  and let  $\text{RM}_\delta(m,r) = \{P \in \text{RM}(m,r) : |\text{wt}(P) - 1/2| \geq \delta/2\}$ , i.e., the set of polynomials with bias at least  $\delta$ .

Now,

$$\begin{aligned} & \sum_{P \in \text{RM}_\delta(m,r)} \exp_2(-\varepsilon u) \cdot (1 - \text{wt}(P))^{(1-\varepsilon)u} \\ & \leq 2 \cdot \sum_{P: \text{wt}(P) \leq 1/2 - \delta/2} \exp_2(-\varepsilon u) \cdot (1 - \text{wt}(P))^{(1-\varepsilon)u}, \end{aligned}$$

since the term corresponding to a polynomial  $P$  of weight greater than  $1/2 + \delta/2$  can be upper bounded by the term corresponding to the polynomial  $1 + P$  which has weight at most  $1/2 - \delta/2$ .

Since the RHS of the above equation involves polynomials whose weights are in the interval  $[1/2^r, 1/2 - \delta]$ , we will split this interval into sub-intervals and analyse the contribution from each.

$$\begin{aligned} \text{Low}_i &:= [1/2^{i+1}, 1/2^i] && \text{for } i = 2, 3, \dots, r-1, \\ \text{Med}_i &:= [(1/2 - 1/2^i), (1/2 - 1/2^{i+1})] && \text{for } i = 2, 3, \dots, t = \log \frac{1}{\delta}. \end{aligned}$$

Let us use the following quantities to denote the number of polynomials with weights in the above intervals:

$$\begin{aligned} L_i &:= |\{P \in \text{RM}(m,r) : \text{wt}(P) \in \text{Low}_i\}| \\ M_i &:= |\{P \in \text{RM}(m,r) : \text{wt}(P) \in \text{Med}_i\}|. \end{aligned}$$

Therefore,

$$\begin{aligned} & \sum_{P \in \text{RM}_\delta(m,r)} \exp_2(-\varepsilon u) \cdot (1 - \text{wt}(P))^{(1-\varepsilon)u} \leq 2 \cdot \sum_{P: \text{wt}(P) \leq 1/2 - \delta/2} \exp_2(-\varepsilon u) \cdot (1 - \text{wt}(P))^{(1-\varepsilon)u} \\ & \leq 2 \left( \sum_{i=2}^{r-1} L_i \cdot (1 - 1/2^{i+1})^k + \sum_{i=2}^t M_i \cdot (1/2 + 1/2^i)^k \right) \cdot \frac{1}{\exp_2\left(\binom{m}{\leq r} - K\right)}. \end{aligned}$$

By [Claim 3.6](#) proved below using the weight distribution bounds of Sberlo and Shpilka ([Theo-](#)

rem 2.1 and Theorem 2.3), we have that

$$\left( \sum_{i=2}^{r-1} L_i \cdot (1 - 1/2^{i+1})^k + \sum_{i=2}^t M_i \cdot (1/2 + 1/2^i)^k \right) \cdot \frac{1}{\exp_2 \left( \binom{m}{\leq r} - K \right)} \leq O(\delta^2).$$

Hence,

$$\begin{aligned} & \sum_{P \in \text{RM}(m,r)} \exp_2(-\varepsilon u) \cdot (1 - \text{wt}(P))^{(1-\varepsilon)u} \\ &= \sum_{P \in \text{RM}_\delta(m,r)} \exp_2(-\varepsilon u) \cdot (1 - \text{wt}(P))^{(1-\varepsilon)u} \\ & \quad + \sum_{P \notin \text{RM}_\delta(m,r)} \exp_2(-\varepsilon u) \cdot (1 - \text{wt}(P))^{(1-\varepsilon)u} \\ &\leq O(\delta^2) + \sum_{P \notin \text{RM}_\delta(m,r)} \exp_2(-\varepsilon u) \cdot (1/2 + \delta/2)^{(1-\varepsilon)u} \\ &\leq O(\delta^2) + (1 + \delta)^{(1-\varepsilon)u} \\ &\leq O(\delta^2) + \exp(\delta u) \\ &\leq 1 + O(1/u) = 1 + o(1). \end{aligned}$$

□

It remains to prove the following technical claim.

**Claim 3.6.** 1. For all  $i = 2, \dots, r-1$ , we have

$$\frac{L_i \cdot (1 - 1/2^{i+1})^k}{\exp_2 \left( \binom{m}{\leq r} - k \right)} \leq \delta^3.$$

2. For all  $i = 2, \dots, t = \log \frac{1}{\delta}$ , we have

$$\frac{M_i \cdot (1/2 + 1/2^i)^k}{\exp_2 \left( \binom{m}{\leq r} - k \right)} \leq \delta^3.$$

*Proof of Claim 3.6(1).* Note that  $(1 - 1/2^{i+1})^k \leq \exp(-k/2^{i+1}) \leq \exp_2(-k/2^{i+1})$ . Hence,

$$\frac{L_i \cdot (1 - 1/2^{i+1})^k}{\exp_2 \left( \binom{m}{\leq r} - k \right)} \leq \frac{\text{WtDist}_{m,r}(2^{-i}) \cdot \exp_2(-k/2^{i+1})}{\exp_2 \left( \binom{m}{\leq r} - k \right)}.$$

Using [Theorem 2.1](#) to bound  $\text{WtDist}_{m,r}(2^{-i})$ , we get

$$\begin{aligned} \frac{L_i \cdot (1 - 1/2^{i+1})^k}{\exp_2 \left( \binom{m}{\leq r} - k \right)} &\leq \exp_2 \left( O(m^4) + O(i\gamma^{i-1}) \cdot \left( \binom{m}{\leq r} - k/2^{i+1} - \left( \binom{m}{\leq r} \right) + k \right) \right) \\ &= \exp_2 \left( O(m^4) - \left( \binom{m}{\leq r} \right) \cdot (1 - O(i\gamma^{i-1})) + k \cdot (1 - 2^{-i-1}) \right), \end{aligned}$$

where  $\gamma = r/m$ . If this  $\gamma$  is a small enough absolute constant, we can see that the  $O(i\gamma^{i-1})$  term is always at most  $2^{-i-1}$ . Using this and continuing the computation above, we get

$$\begin{aligned} \frac{L_i \cdot (1 - 1/2^{i+1})^k}{\exp_2 \left( \binom{m}{\leq r} - k \right)} &\leq \exp_2 \left( O(m^4) - \left( \binom{m}{\leq r} - k \right) \cdot (1 - 2^{-i-1}) \right) \\ &\leq \exp_2 \left( O(m^4) - \varepsilon \left( \binom{m}{\leq r} \right) \cdot (1 - 2^{-i-1}) \right) \\ &\leq \exp_2 \left( O(m^4) - \frac{\varepsilon}{2} \cdot \left( \binom{m}{\leq r} \right) \right) \leq \exp_2 \left( -\Omega \left( \binom{m}{\leq r} \right) \right) \leq \delta^3 \end{aligned}$$

where for the second inequality above, we used the fact that  $k \leq K = (1 - \varepsilon) \cdot \left( \binom{m}{\leq r} \right)$ . □

*Proof of [Claim 3.6\(2\)](#).* Note that

$$\begin{aligned} (1/2 + 1/2^i)^k &= 2^{-k} \cdot (1 + 1/2^{i-1})^k \\ &\leq 2^{-k} \cdot \exp(k/2^{i-1}) = 2^{-k} \cdot \exp_2(k/2^{i-1} \cdot \log_2 e) \\ &= \exp_2(-k \cdot (1 - 2 \log_2 e \cdot 2^{-i})). \end{aligned}$$

Using the fact that  $M_i \leq \text{WtDist}_{m,r}(1/2 - 1/2^{i+1})$ , we get

$$\begin{aligned} &\frac{\text{WtDist}_{m,r}(1/2 - 1/2^{i+1}) \cdot \exp_2(-k \cdot (1 - 2 \log_2 e \cdot 2^{-i}))}{\exp_2 \left( \binom{m}{\leq r} - k \right)} \\ &\leq \frac{\exp_2 \left( O(m^4) + (1 - 2^{-O(\max\{\gamma^2 i, \gamma\})}) \cdot \left( \binom{m}{\leq r} - k \cdot (1 - 2 \log_2 e \cdot 2^{-i}) \right) \right)}{\exp_2 \left( \binom{m}{\leq r} - k \right)} \\ &= \exp_2 \left( O(m^4) - \frac{1}{2^{O(\max\{\gamma^2 i, \gamma\})}} \cdot \left( \binom{m}{\leq r} \right) + k \cdot \frac{1}{2^{-(i-1 - \log_2 \log_2 e)}} \right) \end{aligned}$$

where  $\gamma = r/m$ , and we used [Theorem 2.3](#) for the second inequality.

Note that as long as  $\gamma$  is a small enough absolute constant, the  $O(\max\{\gamma^2 i, \gamma\})$  term above is at most  $i - 1 - \log_2 \log_2 e$  for each  $i \in \{2, \dots, t\}$ . Using this bound, we continue the above

computation as follows.

$$\begin{aligned}
\frac{M_i \cdot (1/2 + 1/2^i)^k}{\exp_2 \left( \binom{m}{\leq r} - k \right)} &\leq \exp_2 \left( O(m^4) - 2^{-O(\max\{\gamma^2 i, \gamma\})} \cdot \left( \binom{m}{\leq r} - k \right) \right) \\
&\leq \exp_2 \left( O(m^4) - 2^{-O(\max\{\gamma^2 i, \gamma\})} \cdot \left( \binom{m}{\leq r} - K \right) \right) \\
&= \exp_2 \left( O(m^4) - 2^{-O(\max\{\gamma^2 i, \gamma\})} \cdot \varepsilon \cdot \binom{m}{\leq r} \right) \\
&\leq \exp_2 \left( O(m^4) - \Omega \left( \sqrt{\binom{m}{\leq r}} \right) \right) \leq \delta^3,
\end{aligned}$$

where for the last inequality we used the fact that  $\binom{m}{\leq r} \geq 2^{\Omega(\gamma \log(1/\gamma)m)}$  and hence for small enough absolute constant  $\gamma$ , we have

$$\begin{aligned}
2^{-O(\max\{\gamma^2 i, \gamma\})} \cdot \binom{m}{\leq r} &\geq 2^{-O(\gamma^2 t)} \cdot \binom{m}{\leq r} \geq \exp_2(\Omega(\gamma \log(1/\gamma)m) - O(\gamma^2 \log(1/\delta))) \\
&= \exp_2(\Omega(\gamma \log(1/\gamma)m) - O(\gamma^3 \log(1/\gamma)m)) = \exp_2(\Omega(m)). \quad \square
\end{aligned}$$

This completes the proof of [Lemma 3.4](#).

## 4 Resilience of Reed-Muller codes under erasures and errors

For parameters  $m, r$ , recall that  $E(m, r)$  is the  $\binom{m}{\leq r} \times 2^m$ -matrix where the columns are indexed by elements in  $\mathbb{F}_2^m$ , and the  $z$ -th column is the column vector consisting of *evaluations* of all multilinear  $m$ -variate degree  $\leq r$  monomials on  $z$ . It is well-known that the matrix  $E(m, r)$  is the generator matrix of the  $\text{RM}(m, r)$  code, and is also the parity-check matrix of the  $\text{RM}(m, m - r - 1)$  code.

[Theorem 1.1](#) can be re-interpreted as a statement about random sub-matrices of  $E(m, r)$  in the regime where  $r \leq \gamma m$  for a small absolute constant  $\gamma > 0$ .

**Corollary 4.1** (rank of random sub-matrices of  $E(m, r)$ ). *Let  $\gamma > 0$  be an absolute constant and  $m, r$  growing parameters with  $r \leq \gamma m$ . Then, for any constant  $\varepsilon > 0$ , a random set of  $K = (1 - \varepsilon)\binom{m}{\leq r}$  columns of  $E(m, r)$  are linearly independent with probability  $1 - o(1)$ .*

*Proof.* A set of columns of  $E(m, r)$  indexed by  $Z = \{z_1, \dots, z_K\}$  are linearly dependent if and only if  $\mathbb{I}_r(Z)$  has dimension strictly larger than  $\binom{m}{\leq r} - K$ . [Theorem 1.1](#) asserts that this happens with only  $o(1)$  probability.  $\square$

## 4.1 Channels under consideration

The above corollary can be used to talk about the resilience of Reed-Muller codes under the *erasure* and *error* channels. We first define the precise definitions to be able to state the results accurately.

**Definition 4.2** (binary erasure channel (BEC)). *The binary erasure channel with parameter  $\alpha$ , denoted by  $\text{BEC}_\alpha$ , is the channel with input alphabet  $\{0, 1\}$  where the each binary symbol is “erased” (replaced by the ‘?’ symbol) independently with probability  $\alpha$ .*  $\diamond$

A closely related model is where we have a fixed number of random erasures instead of each coordinate being erased with a fixed probability. We will refer to this as the  $\text{BEC}^*$  model although this isn’t a channel in the traditional sense of altering each coordinate independently.

**Definition 4.3** (capped binary erasure channel ( $\text{BEC}^*$ )). *The capped binary erasure channel with parameter  $K$ , denoted by  $\text{BEC}_K^*$ , refers to the channel with input alphabet  $\{0, 1\}^n$  that replaces a random subset of at most  $K$  of the coordinates by the ‘?’ symbol.*  $\diamond$

The Binary Symmetric Channel deals with *errors* or *corruptions* as opposed to erasures in the Binary Erasure Channel.

**Definition 4.4** (binary symmetric channel (BSC)). *The binary symmetric channel with parameter  $\alpha$ , is the channel where the each binary symbol is “flipped” (that is, 0 changed to 1 and vice-versa) independently with probability  $\alpha$ .*  $\diamond$

In similar spirit to [Definition 4.3](#), we define the capped binary symmetric channel.

**Definition 4.5** (capped binary symmetric channel ( $\text{BSC}^*$ )). *The capped binary symmetric channel with parameter  $K$ , denoted by  $\text{BSC}_K^*$ , refers to the channel with input alphabet  $\{0, 1\}^n$  that “flips” a random subset of at most  $K$  of the coordinates.*  $\diamond$

In most cases, resilience with respect to  $\text{BEC}_\alpha$  (or  $\text{BSC}_\alpha$ ) is the same as resilience with respect to  $\text{BEC}_K^*$  (or  $\text{BSC}_K^*$ ) for  $K = \alpha n$ , by standard concentration inequalities but this might require some subtlety when dealing with codes of rate very close to zero or close to one. For a concrete example, the code  $\text{RM}(m, m - 1)$  has rate  $R = 1 - 1/2^m$  and is *not* resilient under  $\text{BEC}_\alpha$  for  $\alpha = (1 - R)/2 = 1/2^{m+1}$  (with constant probability we may have two coordinates erased, and this is not recoverable), but is vacuously resilient under  $\text{BEC}_K^*$  for  $K = \alpha 2^m$ . To avoid these nuances, we will *only* be dealing with the setting of capped channels although this does not make much difference with most of our range of parameters.

### Notion of “capacity achieving” codes with respect to $\text{BEC}_\alpha$

Shannon’s seminar work [?] showed that, for any constant rate, the supremum of rates of random codes that is resilient to  $\text{BEC}_\alpha$  is exactly  $R = 1 - \alpha$ . That is, for any  $\varepsilon > 0$ , there are codes of rate  $1 - \alpha - \varepsilon$  that can decode from  $\text{BEC}_\alpha$  with decoding error  $1 - o(1)$ . However, when the rate is very close to zero or one, the situation becomes more nuanced due to the asymptotics involved.

Dealing specifically with Reed-Muller codes, let us consider the code  $\text{RM}(m, m - r - 1)$  with  $r \leq m/2$ . This has rate  $R = 1 - \frac{\binom{m}{\leq r}}{2^m}$  and can recover from a maximum of  $\binom{m}{\leq r}$  errors. Abbe, Shpilka and Wigderson [?] defined the notion of “capacity achieving under BEC” to mean that the code can decode (with  $1 - o(1)$  probability) from  $(1 - \varepsilon)\binom{m}{\leq r}$  erasures for any constant  $\varepsilon > 0$ . We refer the reader to [?, Section 1] and [?, Section 2.2] for a nuanced discussion on this.

## 4.2 Reed-Muller codes under BEC\*

An immediate consequence of the above corollary is that  $\text{RM}(m, d)$  with  $d \geq m(1 - \gamma)$  for a small but absolute constant  $\gamma > 0$  achieves capacity (as defined above) under the Binary Erasure Channel.

**Theorem 1.2** (high-degree RM codes under erasures). *There exists a constant  $\gamma_0 > 0$  such that for all  $\varepsilon > 0$  the following is true. Let  $m, d$  be growing parameters with  $d > m(1 - \gamma_0)$ . Then, the code  $\text{RM}(m, d)$  can correct  $K = (1 - \varepsilon)\binom{m}{\leq m-d-1}$  random errors with probability  $1 - o(1)$ .*

*Proof.* The code  $\text{RM}(m, d)$  can recover from erasures on coordinates indexed by  $Z \subseteq \mathbb{F}_2^m$  if and only if the corresponding columns of the parity check matrix are linearly independent. Since the parity-check matrix of  $\text{RM}(m, d)$  is  $E(m, r)$  for  $r = m - d - 1$ , we have from Corollary 4.1 that a random set of  $K$  columns are linearly independent with probability  $1 - o(1)$ .  $\square$

## 4.3 Reed-Muller codes under BSC\*

The following theorem of Abbe, Shpilka and Wigderson [?] provides a method to derive the resilience of certain Reed-Muller codes under the BSC by using the resilience of appropriate Reed-Muller codes under the BEC. Subsequent work of Saptharishi, Shpilka and Volk also showed that these codes also have efficient decoding procedures to recover from random errors.

**Theorem 4.6** ([?], [?]). *For any growing parameters  $m, r > 0$ , if the code  $\text{RM}(m, m - 2r - 2)$  can recover from a subset  $Z \subseteq [2^m]$  of erasures, then the code  $\text{RM}(m, m - r)$  can efficiently recover from errors on the subset  $Z$ .*

*In particular, if the code  $\text{RM}(m, m - 2r - 2)$  can recover  $K$  random erasures with probability  $1 - o(1)$ , then the code  $\text{RM}(m, m - r)$  can efficiently decode from  $K$  random errors.*

Applying the above theorem to Theorem 1.2 yields the following corollary.

**Corollary 1.3** (high-degree RM codes under errors). *There exists a constant  $\gamma_0 > 0$  such that for all  $\varepsilon > 0$  the following is true. Let  $m, r$  be growing parameters with  $r < \gamma_0 m$ . Then, the code  $\text{RM}(m, m - 2r - 2)$  can be efficiently decode from  $K = (1 - \varepsilon)\binom{m}{\leq r}$  random errors.*

**Corollary 4.7.** *There exists a constant  $\gamma_0 > 0$  such that for all  $\gamma < \gamma_0$  and  $\varepsilon > 0$  the following is true. Let  $m, r$  be growing parameters with  $r = \gamma m$ . Then, the code  $\text{RM}(m, m - 2r - 2)$  can be efficiently decode from  $K = (1 - \varepsilon)\binom{m}{\leq r}$  random errors.*  $\square$

It is worth noting that the minimum distance of the code  $\text{RM}(m, m - 2r - 2)$  is merely  $2^{2r+2} \ll \binom{m}{\leq r}$  when  $r = \gamma m$  for a small enough  $\gamma > 0$ . Thus, the above corollary shows that high-degree (or high-rate) Reed-Muller codes are resilient to random errors well beyond their minimum distance, and efficiently so.

## References

- [ASW15] EMMANUEL ABBE, AMIR SHPILKA, and AVI WIGDERSON. *Reed-Muller codes for random erasures and errors*. IEEE Trans. Inform. Theory, 61(10):5229–5252, 2015. (Preliminary version in 47th STOC, 2015). [arXiv:1411.4590](#).
- [BHL12] IDO BEN-ELIEZER, RANI HOD, and SHACHAR LOVETT. *Random low-degree polynomials are hard to approximate*. Comput. Complexity, 21(1):63–81, 2012. (Preliminary version in 13th RANDOM, 2009). [eccc:2008/TR08-080](#).
- [CDV21] AIDAO CHEN, ANINDYA DE, and ARAVINDAN VIJAYARAGHAVAN. *Learning a mixture of two subspaces over finite fields*. In VITALY FELDMAN, KATRINA LIGETT, and SIVAN SABATO, eds., *Algorithmic Learning Theory (ALT)*, volume 132 of *Proceedings of Machine Learning Research*, pages 481–504. 2021. [arXiv:2010.02841](#).
- [KKMPŞU17] SHRINIVAS KUDEKAR, SANTHOSH KUMAR, MARCO MONDELLI, HENRY D. PFISTER, EREN EREN ŞAŞOĞLU, and RÜDIGER L. URBANKE. *Reed-Muller codes achieve capacity on erasure channels*. IEEE Trans. Inform. Theory, 63(7):4298–4316, 2017. (Preliminary version in 48th STOC, 2016). [arXiv:1601.04689](#).
- [KLP12] TALİ KAUFMAN, SHACHAR LOVETT, and ELY PORAT. *Weight distribution and list-decoding size of Reed-Muller codes*. IEEE Trans. Inform. Theory, 58(5):2689–2696, 2012. (Preliminary version in 1st ICS, 2010). [arXiv:1411.4590](#).
- [KS05] PETER KEEVASH and BENNY SUDAKOV. *Set systems with restricted cross-intersections and the minimum rank of inclusion matrices*. SIAM J. Discrete Math., 18(4):713–727, 2005.
- [Mul54] DAVID E. MULLER. *Application of Boolean algebra to switching circuit design and to error detection*. Trans. IRE Prof. Group Electron. Comput., 3(3):6–12, 1954.
- [NW15] ZIPEI NIE and ANTHONY Y. WANG. *Hilbert functions and the finite degree Zariski closure in finite field combinatorial geometry*. J. Comb. Theory, Ser. A, 134:196–220, 2015. [arXiv:1402.3018](#).
- [Ree54] IRVING S. REED. *A class of multiple-error-correcting codes and the decoding scheme*. Trans. IRE Prof. Group Inf. Theory, 4:38–49, 1954.
- [RP21] GALEN REEVES and HENRY D. PFISTER. *Reed-Muller codes achieve capacity on BMS channels*, 2021. (manuscript). [arXiv:2110.14631](#).
- [Sam20] ALEX SAMORODNITSKY. *An upper bound on  $\ell_q$  norms of noisy functions*. IEEE Trans. Inform. Theory, 66(2):742–748, 2020. [arXiv:1809.09696](#), [eccc:2018/TR01-168](#).
- [Sha48] CLAUDE ELWOOD SHANNON. *A mathematical theory of communication*. The Bell System Technical Journal, 27:379–423, 623–656, 1948.



- [SS20] ORI SBERLO and AMIR SHPILKA. *On the performance of Reed-Muller codes with respect to random errors and erasures*. In SHUCHI CHAWLA, ed., *Proc. 31st Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 1357–1376. 2020. [arXiv:1811.12447](#).
- [SSV17] RAMPRASAD SAPTHARISHI, AMIR SHPILKA, and BEN LEE VOLK. *Efficiently decoding Reed-Muller codes from random errors*. *IEEE Trans. Inform. Theory*, 63(4):1954–1960, 2017. (Preliminary version in *48th STOC*, 2016). [arXiv:1503.09092](#).
- [Wei91] VICTOR K.-W. WEI. *Generalized Hamming weights for linear codes*. *IEEE Trans. Inform. Theory*, 37(5):1412–1418, 1991.