CONJURA · OPEN PROBLEM

Reed-Muller Codes Efficiently Achieve Capacity for the Binary Erasure Channel

Capacity is known; the rate of convergence to it is not

Status. Statement: AI-written from Jarosław Błasiok, Paul Lou, Alon Rosen and Madhu Sudan, *Expanders Meet Reed-Muller: Easy Instances of Noisy k-XOR*, IACR ePrint 2026/664. Not yet checked by a human. This is their Conjecture 14. That Reed-Muller codes achieve capacity for the BEC is a theorem of Kudekar et al.; what is open is the *gap to capacity* — a polynomial rate of convergence. The source needs it to strengthen its refutation of the expansion-implies-hardness conjecture for noisy k-XOR from quasi-polynomially many constraints to polynomially many, and notes the question was raised explicitly in three prior works.

Category. *Coding Theory.*

Conjecture (informal). *Reed-Muller codes are the oldest family in coding theory and, it turns out, among the best: they achieve capacity on the binary erasure channel, which was proved only in 2016. Achieving capacity in the limit is not the same as reaching it fast, though, and for an explicit code with an efficient decoder the rate of convergence is much harder to control than the limit. For random codes it is easy — the gap closes like the inverse square root of the blocklength. The conjecture is that Reed-Muller codes do something of the same order: a gap to capacity that shrinks polynomially in the blocklength.*

1 The setting

Write $\mathrm{RM}(m, r)$ for the Reed-Muller code of order r in m variables: the evaluations on $\mathbb{F}_2^m$ of polynomials in $\mathbb{F}_2[X]$ of total degree at most r, a code of blocklength $M = 2^m$. Over the binary erasure channel each coordinate is erased independently with probability η, and the question is for which η the erased codeword is still recoverable with probability $1 - o(1)$.

Kudekar, Kumar, Mondelli, Pfister, Şaşoğlu and Urbanke proved that Reed-Muller codes achieve capacity over the BEC, recovering

from erasure rate $\varepsilon - o_M(1)$, where $\varepsilon := 1 - \text{Rate}$. That settles the limit. It does not say how fast the $o_M(1)$ term vanishes, and it is exactly that rate — the *scaling exponent* — which the source needs and which is open.

For calibration, the source records what is known elsewhere. For random codes this behaviour is true and relatively simple to show, with scaling exponent $1/\gamma = 2$. For explicit codes with efficient decoding it is much harder: the breakthrough results of Guruswami and Xing, and of Hassani, Alishahi and Urbanke, showed that polar codes efficiently achieve capacity for binary symmetric channels, i.e. have finite scaling exponent, and variants of polar codes can achieve $1/\gamma \to 2$. For Reed-Muller codes nothing of the kind is known.

2 Notation and parameters

- m is the number of variables, r the order, $M = 2^m$ the blocklength of $\mathrm{RM}(m, r)$.
- $\varepsilon := 1 - \text{Rate}$; the code has rate at most $1 - \varepsilon$.
- η is the erasure probability; the erasure pattern is sampled from $\mathsf{Ber}(\eta)^{2^m}$.
- γ is the constant the conjecture asserts exists; $1/\gamma$ is the scaling exponent.

3 The conjecture

Conjecture 1 (Reed-Muller codes efficiently achieve capacity for BEC, [BLRS26, Conjecture 14]). *There exists a constant γ such that for every $m, r \in \mathbb{N}$ for which $\mathrm{RM}(m, r)$ has rate at most $1 - \varepsilon$, a random erasure pattern sampled from $\mathsf{Ber}(\eta)^{2^m}$ is recoverable with probability $1 - o(1)$ provided that*

$$\eta \;<\; \varepsilon - \frac{1}{M^{\gamma}},$$

where $M = 2^m$ is the blocklength of the underlying code.

Remark 1 (An equivalent purely algebraic form). The source supplies one, and it is the cleanest way to attack or refute the conjec-

ture. Consider the linear space $\mathbb{F}_2[X]_{\leq r}$ of low-degree polynomials with dimension $(1-\varepsilon)M$, and a random subset S of $(1-\varepsilon)M + M^{1-\gamma}$ locations on the hypercube $\mathbb{F}_2^m$. Conjecture 1 is equivalent to the statement that with probability $1 - o(1)$ there is no non-zero polynomial in $\mathbb{F}_2[X]_{\leq r}$ that vanishes on all points of S. So the question is about how many random hypercube points are needed before they force a low-degree polynomial to be zero — a statement with no channel and no decoder in it.

Remark 2 (The weaker form that already suffices). The source offers a second conjecture, which it describes as "plausibly simpler-to-prove" and which is enough for its application.

Conjecture 16 (Weak Random Erasure Recovery). There exist constants $\gamma > 0$ and $\zeta \geq 1$ such that for every $m, r \in \mathbb{N}$, if $\mathrm{RM}(m,r)$ has rate at most $1 - \varepsilon$ for $M^{-\gamma} < \varepsilon < 1/2$, a random erasure pattern sampled from $\mathrm{Ber}(\eta)^{2^m}$ is recoverable with probability $1 - o(1)$ provided that $\eta < \varepsilon^{\zeta}$.

The source states that Conjecture 16 is implied by Conjecture 1, and proves its Main Theorem 2 assuming only Conjecture 16. Anyone attacking this should attack the weaker one first; a resolution should say which it settles, since they are not equivalent as far as the source claims.

Remark 3 (What it buys, and what is already unconditional). The source refutes the widely-held conjecture that expansion of the constraint graph implies hardness of noisy k-XOR. Its Main Theorem 1 is unconditional: for every constant $\alpha > 0$ there is an infinite family of k-left-regular constraint graphs with $M = 2^{\Theta(\log^2 N)}$, $k = (\log N)^{\Theta(1/\alpha)}$, which is $(N^{1-\alpha}, 1 - o(1))$-expanding and on which the η-noisy k-XOR distinguishing problem is solvable in time $\mathrm{poly}(M)$ at constant noise rate $\eta = 1/3$.

Under Conjecture 16 — hence under Conjecture 1 — the source's Main Theorem 2 upgrades this to $M = N^{\Theta(1)}$, i.e. the number of constraints polynomially related to the number of variables rather than quasi-polynomially, at noise rate $\eta = N^{-c}$ for any $c > 0$. The source's own summary: "To get a result for a polynomial number of equations in the number of variables we need much faster rate of convergence to capacity than is currently known."

So the refutation does not depend on this conjecture; only its strongest form does.

Remark 4 (Not the source's own question). The source records that "[t]he question of the gap-to-capacity results for RM codes for symmetric channels has been explicitly raised in [AY19, Section V], [Has+18; MHU14], and in [ASY21]." It is a coding-theory question that the source imports and gives a new consequence to. A resolution belongs to that older line.

Remark 5 (A separate question the source leaves open). Its construction needs the constraint graph to be an expander, which it gets from the explicit lossless expanders of Guruswami, Umans and Vadhan reinterpreted as coset graphs. The natural alternative — random subspaces — it cannot analyse: "we leave open the question of whether a random subspace construction of coset graphs is expanding." That is a different question and it is not this statement.

4 Bibliography

[BLRS26] Jarosław Błasiok, Paul Lou, Alon Rosen and Madhu Sudan. *Expanders meet Reed-Muller: easy instances of noisy k-XOR.* IACR ePrint 2026/664. The source. Conjecture 14, Remark 15 and Conjecture 16 are all on page 9; Main Theorems 1 and 2 are stated informally in Section 1.3 and proved as Theorems 37 and 38 in Section 6.

[Kud+16] Shrinivas Kudekar, Santhosh Kumar, Marco Mondelli, Henry D. Pfister, Eren Şaşoğlu and Rüdiger Urbanke. *Reed-Muller codes achieve capacity on erasure channels.* STOC 2016. The theorem that Reed-Muller codes achieve capacity over the BEC, whose rate of convergence this conjecture is about. [UNVERIFIED] — cited in the source as [Kud+16]; the venue here is inferred.

[GUV09] Venkatesan Guruswami, Christopher Umans and Salil Vadhan. *Unbalanced expanders and randomness extractors from Parvaresh-Vardy codes.* Journal of the ACM, 2009. The lossless expanders the source reinterprets as coset graphs, which is what lets it obtain expansion while retaining Reed-Muller decoding.

[GX15] Venkatesan Guruswami and Chaoping Xing. *Optimal rate list decoding over bounded alphabets using algebraic-geometric codes.* Cited by the source, with Hassani et al., for polar codes efficiently achieving capacity for binary symmetric channels — the comparison point for what is wanted here. [UNVERIFIED] — cited in the source as [GX15]; the bibliographic detail here is inferred and may name the wrong paper by these authors.

[Ale03] Michael Alekhnovich. *More on average case vs approximation complexity.* FOCS 2003. The origin of the belief that expansion implies hardness for noisy k-XOR, which the source refutes, and the source of the inverse-polynomial noise-rate regime.