Reed-Muller Codes Achieve Capacity on BMS Channels

Galen Reeves[1,2] and Henry Pfister[1,3]

[1]Department of ECE
[2]Department of Statistical Science
[3]Department of Mathematics
Duke University

Recent Advances in Coding Theory Workshop
54th Annual ACM Symposium on Theory of Computing
Rome, Italy
June 23, 2022

Outline

Reed-Muller (RM) Codes

Codes by Muller, efficient suboptimal decoder by Reed, both in 1954

$[N, K, D]$ binary code RM(r, m) is indexed by integers $0 \leq r \leq m$ with

$$N = 2^m, \qquad K = \sum_{i=0}^{r} \binom{m}{i}, \qquad D = 2^{m-r}$$

RM(r, m) is a multivariate polynomial evaluation code. defined by

$$\left\{ \boldsymbol{c} \in \mathbb{F}_2^N \,\middle|\, c_{\tau(\boldsymbol{v})} = f(\boldsymbol{v}), \boldsymbol{v} \in \mathbb{F}_2^m, f \in \mathcal{F}_{r,m} \right\},$$

where $\mathcal{F}_{r,m}$ is the set of multilinear polynomials in m vars $(v_0, \ldots, v_{m-1})$ with binary coefs and degree $\leq r$, and $\tau \colon \mathbb{F}_2^m \to [N]$ defines the bit order

Conjectured to Achieve Capacity...

- First discussed in 1990s by Lin, 1993 and Dumer, Farrell, 1994
- Stated explicitly for Gaussian noise by Costello, Forney, 2007
- Discussed further following polar codes Arikan, 2010
- Proved for BEC/BSC if rate $\to 0/1$ by Abbe, Shpilka, Widgerson, 2014
- Open problem at 2015 Simons Institute Program on Information Theory
- Proved for BEC by Kudekar, Kumar, Mondelli, Pfister, Şaşoğlu, Urbanke, 2016
- RM codes polarize and "Twin RM" codes achieve capacity Abbe, Ye, 2020
- Reliable on BSC but gap to capacity by Hązła, Samorodnitsky, Sberlo, 2021

This list is not exhaustive and we apologize for any neglected references

Binary Memoryless Symmetric (BMS) Channels

- Binary input, real output:

$$\boldsymbol{X} = (X_0, \ldots, X_{N-1}) \in \{\pm 1\}^N, \qquad \boldsymbol{Y} = (Y_0, \ldots, Y_{N-1}) \in \mathbb{R}^N$$

- Memoryless:

$$p(y_0, \ldots, y_{N-1} \mid x_0, \ldots, x_{N-1}) = \prod_{i=0}^{N-1} w(y_i \mid x_i)$$

- Symmetric:

$$w(y \mid +1) = w(-y \mid -1)$$

- Generated by IID multiplicative noise $\boldsymbol{Z} = (Z_0, \ldots, Z_{N-1}) \in \mathbb{R}^N$:

$$Y_i = X_i Z_i \quad \Leftrightarrow \quad \boldsymbol{Y} = \boldsymbol{X} \odot \boldsymbol{Z},$$

BSC(p): $p = \Pr(Z_i = -1) = 1 - \Pr(Z_i = 1)$

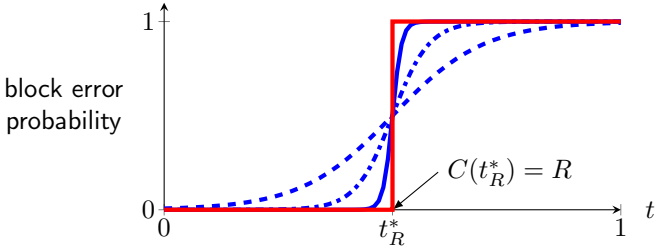
Capacity of Binary Memoryless Symmetric (BMS) Channels

- Defined by single channel use with uniform input $X \in \{\pm 1\}$ and output Y

$$C = 1 - H(X \mid Y)$$

- Consider a family of BMS channels $\{W_t : 0 \leq t \leq 1\}$ with capacity $C(t)$
  - Ordered by degradation from perfect ($t = 0$) to uninformative ($t = 1$)
  - Examples include BEC(t), BSC($t/2$), and BIAWGN with $\sigma^2 = \frac{t}{1-t}$

- Shannon's theorem: Random codes have sharp threshold for block error at t_R^*:



Main Result

Theorem

Consider a BMS channel with capacity $C \in (0,1)$. For every RM(r,m) code with rate $R < C$, the bit-error rate under bit-MAP decoding satisfies

$$\mathrm{BER}(X_i \mid \boldsymbol{Y}) \leq \frac{3\ln(m) + 17}{5\sqrt{m}\big(C - R\big)}$$

for all $i \in [N] := \{0, 1, \ldots, N-1\}$. Thus, there exists a sequence of RM codes with increasing blocklength $N = 2^m$ and rate converging to C such that the BER under bit-MAP decoding converges to zero.

This is proved in arXiv:2110.14631 and these slides outline some of the key steps.

Outline

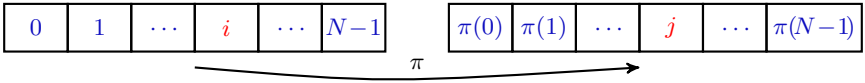
Group Symmetry Refresher

The Permutation Automorphism Group $\mathcal{G}$ of code $\mathcal{C}$ is defined to be

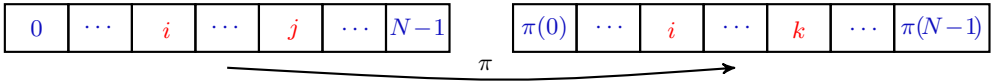
$$\mathcal{G} = \{\pi \in S_N \mid \forall \boldsymbol{c} \in \mathcal{C},\ \ (c_{\pi(0)}, c_{\pi(1)}, \ldots, c_{\pi(N-1)}) \in \mathcal{C}\}$$

Transitive Permutation Groups

- $\mathcal{G}$ is transitive if, for all $i, j \in [N]$, there exists $\pi \in \mathcal{G}$ such that $\pi(i) = j$



- $\mathcal{G}$ is doubly transitive if, for all distinct $i, j, k \in [N]$, there exists $\pi \in \mathcal{G}$ such that $\pi(i) = i$ and $\pi(j) = k$



Extrinsic Information Idea and Analysis

- Used to describe turbo decoding by Berrou, Glavieux, and Thitimajshima, 1993
- Focuses on recovering a single input X_i from all other outputs $Y_{\sim i}$
  - Independent of i for codes with transitive symmetry
  - Related to bit-error probability rather than block-error probability
- Analysis by ten Brink, 1999 and Ashikhmin, Kramer, ten Brink, 2004
  - The key idea is the Area Theorem for the BEC(t):

$$\frac{d}{dt}\frac{1}{N}H(\boldsymbol{X}|\boldsymbol{Y}(t)) = \underbrace{\frac{1}{N}\sum_{i=1}^{N} H(X_i \,|\, Y_{\sim i}(t))}_{\text{EXIT function}}$$

- Prove capacity via sharp threshold with jump localized by Area Theorem

$$\int_0^1 H\big(X_0 \,|\, Y_{\sim 0}(t)\big)\, dt = \frac{1}{N}H(\boldsymbol{X}) = \text{code rate.}$$

Proof for Binary Erasure Channel [KKMPSU16]

- Let $\boldsymbol{E} \in \{0,1\}^N$ be the erasure indicator vector (i.e., $E_i = 1$ iff $Y_i = ?$)
- Linear code + BEC(t) $\Rightarrow$ Conditioned on $Y_{\sim i}$ there are only two cases:

  X_i can be decoded or X_i is uniform (no information)

- "X_i is uniform" iff $g(E_{\sim i}) = 1$ for a monotone Boolean function g
  - $\mathbb{E}[g(E_{\sim i})] = h_i(t) := H(X_i | Y_{\sim i})$
  - Area Theorem: $\int_0^1 h_i(t)\, dt = R$ by transitive symmetry

- All symmetric monotone Boolean functions have a sharp threshold!
  - Stated by Friedgut, Kalai 1996 for monotone graph properties
- Thus, sequences of doubly transitive codes achieve capacity on BEC!

A Few Open Questions related to the BEC Result

- For RM codes on BEC, this proof fails for rate $1-N^{-\alpha}$ with $\alpha \in (0,1)$. Can one extend boolean function argument to a wider range?
- For RM codes, does the EXIT function have transition width $\tilde{O}(N^{-1/2})$?
- No known counterexamples for transitive codes with $D_{\min}, D_{\min}^{\perp} \to \infty$. Do all reasonable transitive code sequences achieve capacity?

Outline

Area Theorem for Generalized EXIT Function

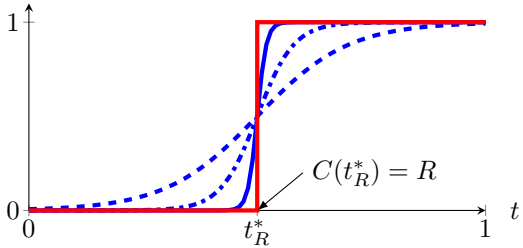
- Integral of derivative def by Méasson, Montanari, Richardson, Urbanke, 2009

$$NR = \underbrace{H(\boldsymbol{X} \mid \boldsymbol{Y}(t=1))}_{\log_2 \#\text{messages}} - \underbrace{H(\boldsymbol{X} \mid \boldsymbol{Y}(t=0))}_{=0} = \int_0^1 \frac{d}{dt} H(\boldsymbol{X} \mid \boldsymbol{Y}(t))\, dt$$

- Total derivative + chain rule for entropy + transitive symmetry gives

$$\frac{d}{dt} H(\boldsymbol{X} \mid \boldsymbol{Y}(t)) = N \times \underbrace{\frac{\partial}{\partial t_i} H(X_i \mid \boldsymbol{Y}(t))\Big|_{t_i=t}}_{\text{Generalized EXIT function}}$$

GEXIT function for X_i given $\boldsymbol{Y}$



Going Beyond the Erasure Channel

What measure of uncertainty to analyze?

- The "GEXIT function" satisfies the area theorem but lacks interpretability
- Standard generalizations of hypercontractivity do not imply a sharp threshold

Key observation:

- Nesting property: RM codes contain shorter RM codes of nearly the same rate

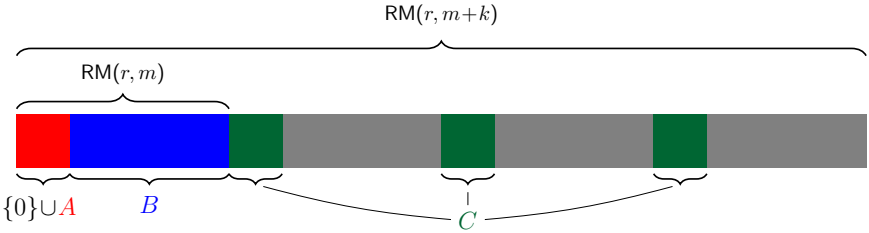
Our approach:

- Minimum mean-square error (MMSE) has almost all the right properties

$$\mathsf{mmse}(X_i|O) := \mathbb{E}\Big[\big(X_i - \mathbb{E}[X_i|O]\big)^2\Big]$$

- Use variance decomposition, generalized influence and estimation inequalities to analyze difference between short and long RM codes with similar rates
- New information inequalities connect the MMSE to entropy and code rate

Nested Structure of Reed–Muller Codes



- This diagram illustrates a copy of RM(r, m) inside RM$(r, m+k)$ for $k = 2$. It is generated by degree-$\leq r$ monomials in variables $v_0, \ldots, v_{m-1}$.
- $A = \{1, \ldots, 2^{m-k} - 1\}$ and $B = \{2^{m-k}, \ldots, 2^m - 1\}$ index bits in RM$(r, m+k)$
- A second copy is generated by degree-$\leq r$ monomials in variables $v_0, \ldots, v_{m-k-1}$, $v_m, \ldots, v_{m+k-1}$ and supported on $\{0\} \cup A \cup C$ with

$$C = \bigcup_{j=1}^{2^k-1} \left(\{0, 1, \ldots, 2^{m-k} - 1\} + j2^m \right)$$

Rate Difference for Nearby Reed–Muller Codes

- The rate of RM(r, m) is given by

$$R(r,m) \coloneqq \frac{1}{2^m} \sum_{i=0}^{r} \binom{m}{i}$$

  - If $Q_m \sim$ Binomial$(m, \frac{1}{2})$, then $R(r,m) = \Pr(Q_m \leq r)$
  - By the CLT, $(Q_m - \frac{m}{2})/\sqrt{\frac{m}{4}} \xrightarrow{w} \mathcal{N}(0,1)$

- For $k \geq 1$, improved CLT for symmetric binomial shows that

$$R(r,m) - R(r,m+k) \leq \frac{3k+4}{5\sqrt{m}}$$

  - Thus, the rate difference vanishes for sequences where $k = o(\sqrt{m})$.
  - This is quite surprising because it means that there is a puncturing pattern for RM$(r, m+k)$, keeping only a fraction 2^{-k} of the code bits, that does not appreciably change the code rate.

The Extrinsic MMSE – Definition and Properties

- Define the extrinsic MMSE as a function of channel parameter t:

$$M(t) := \mathsf{mmse}(X_i \mid Y_{\sim i}) = 1 - \left\| \mathbb{E}[X_i | Y_{\sim i}] \right\|_2^2,$$

  where $\| \cdot \|_p := \mathbb{E}[(\,\cdot\,)^p]^{1/p}$ and i can be dropped by transitive symmetry

- For a seq. of codes, the extrinsic MMSE has a sharp threshold if and only if

$$\lim_{N \to \infty} \int_0^1 M(t)(1 - M(t))\, dt = 0$$

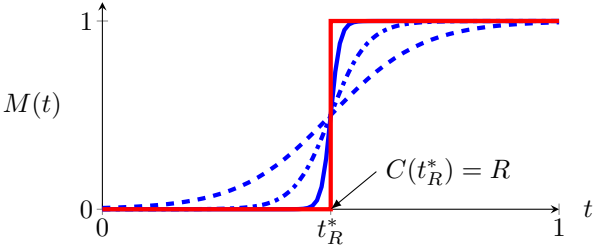
  i.e., $M(t)$ converges to a 0/1 step function

Extrinsic MMSE Analysis

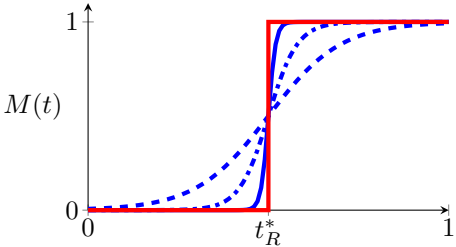
- Input $\boldsymbol{X} = (X_0 \dots X_{2^m-1})$ is an RM(r, m) codeword
- Output $\boldsymbol{Y}(t) = (Y_0(1) \dots Y_{2^m-1}(t))$ is from a family of BMS channels, ordered from perfect $(t = 0)$ to uninformative $(t = 1)$
- Extrinsic MMSE

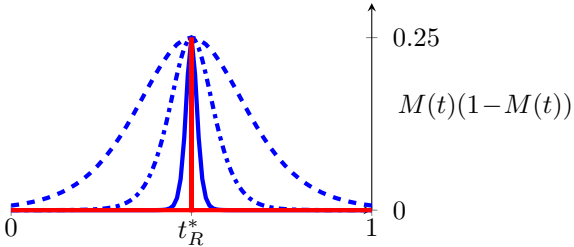
$$M(t) \coloneqq \text{mmse}(X_i \mid Y_{\sim i}) = 1 - \left\| \mathbb{E}[X_i | Y_{\sim i}] \right\|_2^2$$

- Goal is to show sharp threshold and then localize jump via "area theorem"



Visualizing a Sharp Threshold





$M(t) \to$ 0/1 step function $\quad \Leftrightarrow \quad \int_0^1 M(t)\big(1 - M(t)\big)\, dt \to 0$

Preliminary Intuition for Sharp Threshold Argument

- For $k \geq 1$, RM$(r, m+k)$ can be punctured in multiple ways to get RM(r, m)
- Consider estimation of X_0 using either the long code or two short codes
- Hand Waving Argument
  - Choose t so that $R(r, m+k) < C(t) < R(r, m)$
  - $C(t) < R(r, m)$ implies RM(r, m) estimate is imperfect (i.e., $M(t) > 0$)
  - For RM$(r, m+k)$, combine two RM(r, m) estimates to reduce MMSE
  - For the code sequence RM(r_m, m), we either have:
    (A) the liminf of the mmse$(X_0|Y_A, Y_B, Y_C)$ for RM$(r_m, m+k)$ will be strictly lower than the liminf of mmse$(X_0|Y_A, Y_B)$ for RM(r_m, m)
    (B) OR the two RM(r_m, m) estimates will be equal in the limit. More specifically, we will have $\|\mathbb{E}[X_0|Y_A, Y_B] - \mathbb{E}[X_0|Y_A, Y_C]\| \to 0$.
- Our current proof is quite different but comes from this rough intuition

Decomposition of Variance

- BMS + linear code gives the variance identity

$$M(t)(1-M(t)) = \frac{1}{2}\left\|\mathbb{E}[X_i|Y_{\sim i}] - \mathbb{E}[X_i|Y'_{\sim i}]\right\|_2^2$$

  where $\boldsymbol{Y}' = (Y'_1, \ldots, Y'_N)$ is resampled output with the same input.

- Resampling $\boldsymbol{Y}$ should only change estimate only if $M(t) \in [\delta, 1-\delta]$ for $\delta > 0$

- Efron-Stein Inequality: For any partition $\{A, B\}$ of $\{1, \ldots, N\} \backslash i$, the above term is bounded from above by

$$\underbrace{\frac{1}{2}\left\|\mathbb{E}[X_i|Y_A, Y_B] - \mathbb{E}[X_i|Y'_A, Y_B]\right\|_2^2}_{\mathrm{Inf}_A := \text{ influence of } A} + \underbrace{\frac{1}{2}\left\|\mathbb{E}[X_i|Y_A, Y_B] - \mathbb{E}[X_i|Y_A, Y'_B]\right\|_2^2}_{\mathrm{Inf}_B := \text{ influence of } B}.$$

- Can be seen as generalized influence functions for sets (See O'Donnell, 2014)

Proof of Sharp Threshold

For $A = \{1, \ldots, 2^{m-k} - 1\}$ and $B = \{2^{m-k}, \ldots, 2^m - 1\}$ we show that

$$\underbrace{\int_0^1 \mathrm{Inf}_A \, dt \lesssim 2^{-k}}_{\text{double transitivity}} \qquad \text{and} \qquad \underbrace{\int_0^1 \mathrm{Inf}_B \, dt \lesssim \frac{k}{\sqrt{m}}}_{\text{RM nesting property}}$$

Choosing $k = \lceil \frac{1}{2} \log_2 m \rceil$ yields

$$\int_0^1 M(t)(1 - M(t)) \lesssim \frac{\log m}{\sqrt{m}}$$

From Influence to Difference in MMSE

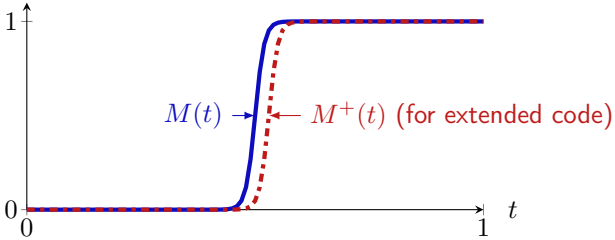
Lemma

The influence of set $B = \{2^{m-k}, \ldots, 2^m - 1\}$ satisfies

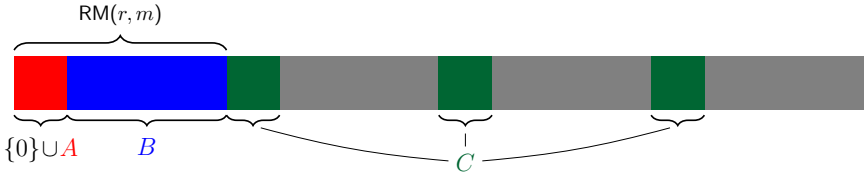
$$\mathrm{Inf}_B \leq 2\Big(\underbrace{\mathsf{mmse}(X_0 \mid Y_1 \ldots Y_{2^m-1})}_{M(t)} - \underbrace{\mathsf{mmse}(X_0 \mid Y_1 \ldots Y_{2^{m+k}-1})}_{M^+(t)}\Big) \quad (\heartsuit)$$

where $Y_0 \ldots Y_{2^{m+k}-1}$ are the outputs for the $\mathsf{RM}(r, m+k)$ codeword.

How Much Better is the Longer Code?



Proof of (♥) via Nesting Property



$$
\begin{aligned}
\sqrt{2\,\mathrm{Inf}_B} &= \left\| \mathbb{E}[X_0 \mid Y_A, Y_B] - \mathbb{E}[X_0 \mid Y_A, Y'_B] \right\|_2 \\
&= \left\| \mathbb{E}[X_0 \mid Y_A, Y_B] - \mathbb{E}[X_0 \mid Y_A, Y_C] \right\|_2 && \text{nesting property} \\
&= \| \mathbb{E}[X_0 \mid Y_A, Y_B] - \hat{X}_0 - \mathbb{E}[X_0 \mid Y_A, Y_C] + \hat{X}_0 \|_2 && \hat{X}_0 := \mathbb{E}[X_0 \mid Y_1 \dots Y_{2^{m+k}-1}] \\
&\le \| \mathbb{E}[X_0 \mid Y_A, Y_B] - \hat{X}_0 \|_2 + \| \mathbb{E}[X_0 \mid Y_A, Y_C] - \hat{X}_0 \|_2 && \text{triangle inequality} \\
&= 2 \left\| \mathbb{E}[X_0 \mid Y_1 \dots Y_{2^m-1}] - \mathbb{E}[X_0 \mid Y_1 \dots Y_{2^{m+k}-1}] \right\|_2 && \text{equal in distribution} \\
&= 2 \sqrt{\| \mathbb{E}[X_0 \mid Y_1 \dots Y_{2^{m+k}-1}] \|_2^2 - \| \mathbb{E}[X_0 \mid Y_1 \dots Y_{2^m-1}] \|_2^2} && \text{nested expectation} \\
&= 2 \sqrt{\mathsf{mmse}(X_0 \mid Y_1 \dots Y_{2^m-1}) - \mathsf{mmse}(X_0 \mid Y_1 \dots Y_{\text{Ł}-1})} && \text{definition of mmse}
\end{aligned}
$$

From Difference in MMSE to Difference in Entropy

Lemma

Consider a family of BMS channels $\{W_t : 0 \le t \le 1\}$:

- Ordered by degradation from perfect ($t = 0$) to uninformative ($t = 1$)
- $\frac{d}{dt}\,\mathsf{mmse}(X_0 \mid Y_0(t)) \ge \text{constant}$

Then

$$M(t) - M^+(t) \lesssim \frac{d}{dt}\left(\frac{H(\boldsymbol{X} \mid \boldsymbol{Y}(t))}{2^m} - \frac{H(\boldsymbol{X}^+ \mid \boldsymbol{Y}^+(t))}{2^{m+k}}\right) \qquad (\spadesuit)$$

where $(\boldsymbol{X}^+, \boldsymbol{Y}^+(t))$ is input-output pair for $\mathsf{RM}(r, m+k)$ code.

Extrinsic Series Expansion of GEXIT Function

- Law of total derivative + chain rule + transitive symmetry

$$\frac{d}{dt}\frac{H(\boldsymbol{X} \mid \boldsymbol{Y}(t))}{2^m} = \underbrace{\frac{\partial}{\partial s} H\big(X_0 \mid Y_0(s), Y_{\sim 0}(t)\big)\Big|_{s=t}}_{\text{GEXIT function}}$$

- Series expansion of entropy with positive coefficients $c_n > 0$.

$$H\big(X_0 \,|\, Y_0(s), Y_{\sim 0}(t)\big) = \sum_{n=1}^{\infty} c_n \Big(1 - \underbrace{\big\|\mathbb{E}[X_0 \,|\, Y_0(s)]\big\|_{2n}^{2n}}_{\text{direct}}\Big)\Big(1 - \underbrace{\big\|\mathbb{E}[X_0 \,|\, Y_{\sim 0}(t)]\big\|_{2n}^{2n}}_{\text{extrinsic}}\Big)$$

Proof of (♠) via Series Expansion + Channel Ordering

$$\frac{d}{dt}\left(\frac{H(\boldsymbol{X} \mid \boldsymbol{Y}(t))}{2^m} - \frac{H(\boldsymbol{X}^+ \mid \boldsymbol{Y}^+(t))}{2^{m+k}}\right)$$

$$= \sum_{n=1}^{\infty} c_n \Bigg(\underbrace{-\frac{d}{ds}\left\|\mathbb{E}[X_0 \mid Y_0(s)]\right\|_{2n}^{2n}\Big|_{s=t}}_{\geq 0 \text{ by channel ordering}} \Bigg) \Bigg(\underbrace{\overbrace{\left\|\mathbb{E}\big[X_0 \mid Y_{\sim 0}^+(t)\big]\right\|_{2n}^{2n}}^{\mathsf{RM}(r,\, m+k)} - \overbrace{\left\|\mathbb{E}[X_0 \mid Y_{\sim 0}(t)]\right\|_{2n}^{2n}}^{\mathsf{RM}(m,\, k)}}_{\geq 0 \text{ by data processing inq.}} \Bigg)$$

$$\geq c_1 \Bigg(\underbrace{-\frac{d}{ds}\left\|\mathbb{E}[X_0 \mid Y_0(s)]\right\|_{2}^{2}\Big|_{s=t}}_{\frac{d}{dt}\,\mathsf{mmse}(X_0 \mid Y_0(t))} \Bigg) \Bigg(\underbrace{\left\|\mathbb{E}\big[X_0 \mid Y_{\sim 0}^+(t)\big]\right\|_{2}^{2} - \left\|\mathbb{E}[X_0 \mid Y_{\sim 0}(t)]\right\|_{2}^{2}}_{M(t) - M^+(t)} \Bigg)$$

Putting the Pieces Together

$$\int_0^1 \mathrm{Inf}_B \, dt \lesssim \int_0^1 \big(M(t) - M^+(t) \big) \, dt \qquad (\heartsuit)$$

$$\lesssim \int_0^1 \frac{d}{dt}\Big(\frac{H(\boldsymbol{X} \mid \boldsymbol{Y}(t))}{2^m} - \frac{H(\boldsymbol{X}^+ \mid \boldsymbol{Y}^+(t))}{2^{m+k}} \Big) \, dt \qquad (\spadesuit)$$

$$= \frac{H(\boldsymbol{X})}{2^m} - \frac{H(\boldsymbol{X}^+)}{2^{m+k}} \qquad t=0 \text{ perfect, } t=1 \text{ useless}$$

$$= R(r,m) - R(r,m+k) \qquad \text{uniform codewords}$$

$$\leq \frac{3k+4}{5\sqrt{m}} \qquad \text{rate difference lemma}$$

Average influence of large set vanishes if $k = o(\sqrt{m})$

Summary

- This talk outlined some key steps to prove a sharp threshold for $M(t)$
- What parts did we skip?
  - Going from extrinsic MMSE to entropy requires the channel with $t = \text{mmse}(X_i|Y_i)$ and we skipped argument that extends to all channels.
  - Localizing the $M(t)$ jump to the capacity limit. For this, we derive a few simple bounds connecting $M(t)$ and the GEXIT function. The GEXIT area theorem allows one to localize the jump similar to the BEC case.
  - Bounding the contribution of each term in A set to the quantity $\int M(t)(1 - M(t))\, dt$. This uses an extra look argument and bounds the GEXIT (with an extra look) using an entropy series expansion. The final bound uses the GEXIT area theorem via symmetrization.

Outline

Conclusions and Open Questions

Summary of main results:

- Proof that RM codes achieve capacity on BMS channels
  i.e., vanishing fraction of incorrectly decoded inputs
- Approach based on analyzing extrinsic information
- Unlike previous result for erasure channel, does not rely on hypercontractivity
- Proof combines a nesting property of RM codes with new information inequalities for GEXIT functions and the extrinsic MMSE

Open Questions:

- Does the block error probability goes to zero? Prove it.
- Can we extend this to code families beyond Reed–Muller?

References I

[AKtB04] A. Ashikhmin, G. Kramer, and S. ten Brink.
Extrinsic information transfer functions: model and erasure channel properties.
IEEE Trans. Inform. Theory, 50(11):2657–2674, Nov. 2004.

[BGT93] Claude Berrou, Alain Glavieux, and Punya Thitimajshima.
Near Shannon limit error-correcting coding and decoding: Turbo-codes.
In *Proc. IEEE Int. Conf. Commun.*, volume 2, pages 1064–1070, Geneva, Switzerland, May 1993. IEEE.

[FK96] Ehud Friedgut and Gil Kalai.
Every monotone graph property has a sharp threshold.
Proc. Amer. Math. Soc., 124(10):2993–3002, 1996.

[MMRU09] C. Méasson, A. Montanari, T. J. Richardson, and R. Urbanke.
The generalized area theorem and some of its consequences.
IEEE Trans. Inform. Theory, 55(11):4793–4821, Nov. 2009.

[MMU08] Cyril Méasson, Andrea Montanari, and Rüdiger L. Urbanke.
Maxwell construction: The hidden bridge between iterative and maximum a posteriori decoding.
IEEE Trans. Inform. Theory, 54(12):5277–5307, Dec. 2008.

References II

[MU03] C Méasson and R Urbanke.
An upper-bound for the ML threshold of iterative coding systems over the BEC.
In *Proc. Annual Allerton Conf. on Commun., Control, and Comp.*, page 3, 2003.

[tB01] S. ten Brink.
Exploiting the chain rule of mutual information for the design of iterative decoding schemes.
In *Proc. 39th Annual Allerton Conf. on Commun., Control, and Comp.*, Monticello, IL, USA, Oct. 2001.