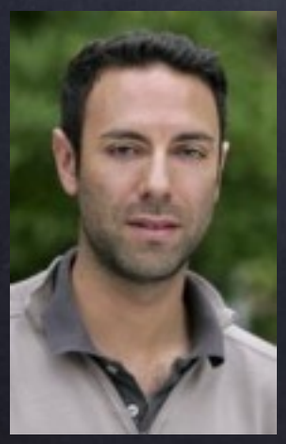
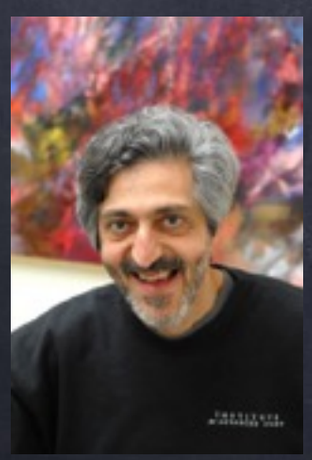
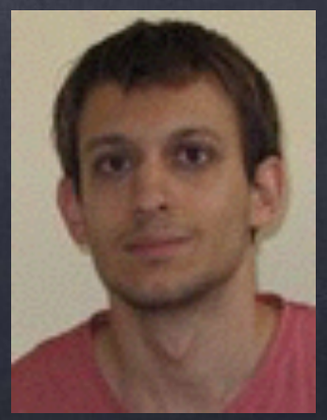
Reed-Muller codes for random errors and erasures

Based on:

Abbe-S-Wigderson
Saptharishi-S-Volk



Reed-Muller code - RM(m,r)

- Message:
  coefficient vector of a polynomial
  $f(x_1,\ldots,x_m)$ over $\mathbb{F}_2$ of degree $\leq r$
- Encoding: evaluations of f:
  $f \rightarrow (f(000),f(001),\ldots,f(111))$
- Distance: Hamming distance between any two code words is $\geq 2^{m-r}$
- Note: codewords form a linear space over $\mathbb{F}_2$

The Evaluation matrix (Sierpinski matrix)

$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ $\{0,1\}^m$ $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$

m-variate monomials in lexicographical order:

1

x_1

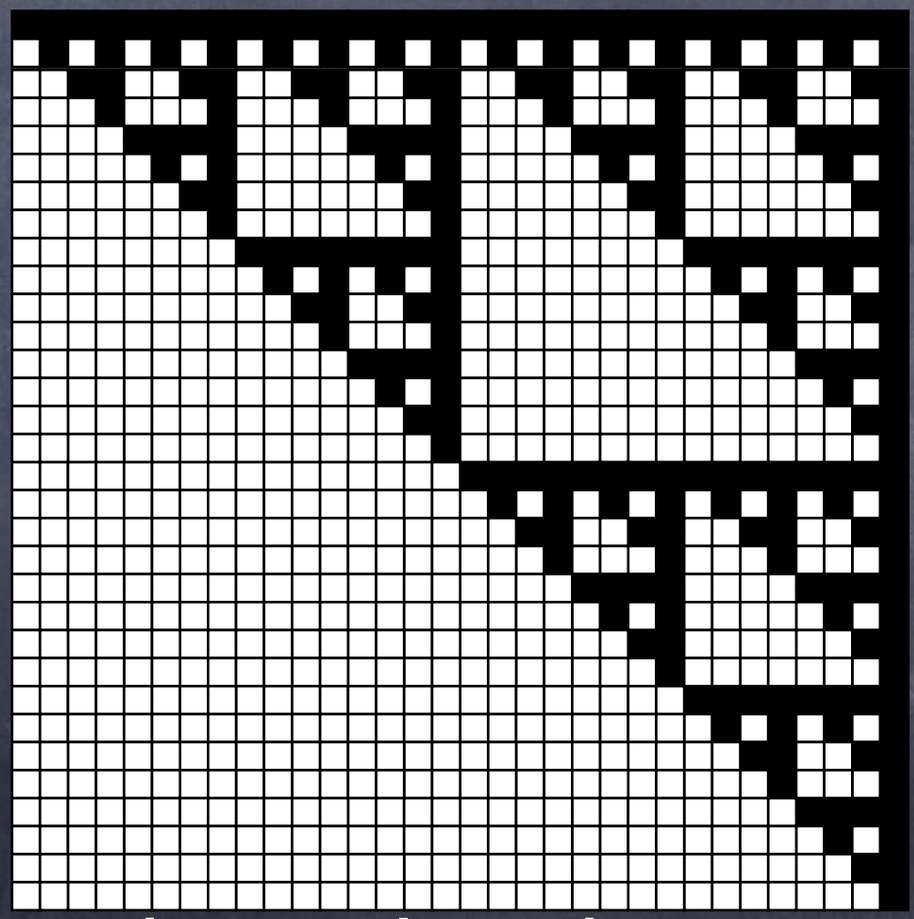
x_2

x_1x_2

x_3

x_1x_3

$1, x_1, x_2, x_1x_2, x_3, x_1x_3,$
…

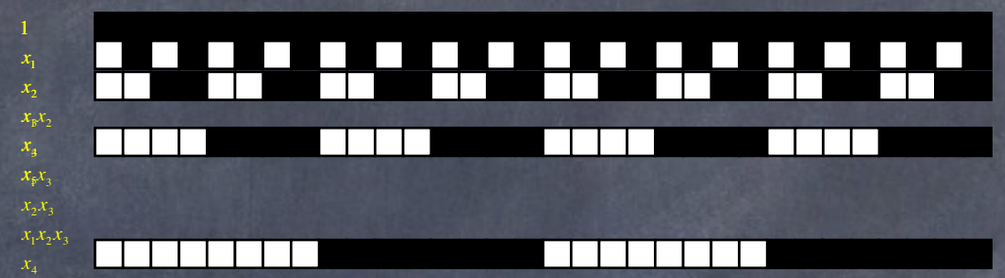


Keep rows corresponding to low degree monomials (heaviest rows)

The Evaluation matrix (Sierpinski matrix)

$\{0,1\}^m$

$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$



Generating matrix of Reed-Muller code:
codewords are linear combination of rows

Keep rows corresponding to low degree monomials (heaviest rows)

Reed-Muller code - RM(m,r)

- Message: coefficient vectors of a polynomial $f(x_1,\ldots,x_m)$ over $\mathbb{F}_2$ of degree $\leq r$
- Encoding: evaluations of f:
  $f \rightarrow (f(000),f(001),\ldots,f(111))$
- Minimum distance: $d=2^{m-r}$
- Most studied linear algebraic code
- Around over 50 years yet fundamental questions are still open!

Why care about RM codes?

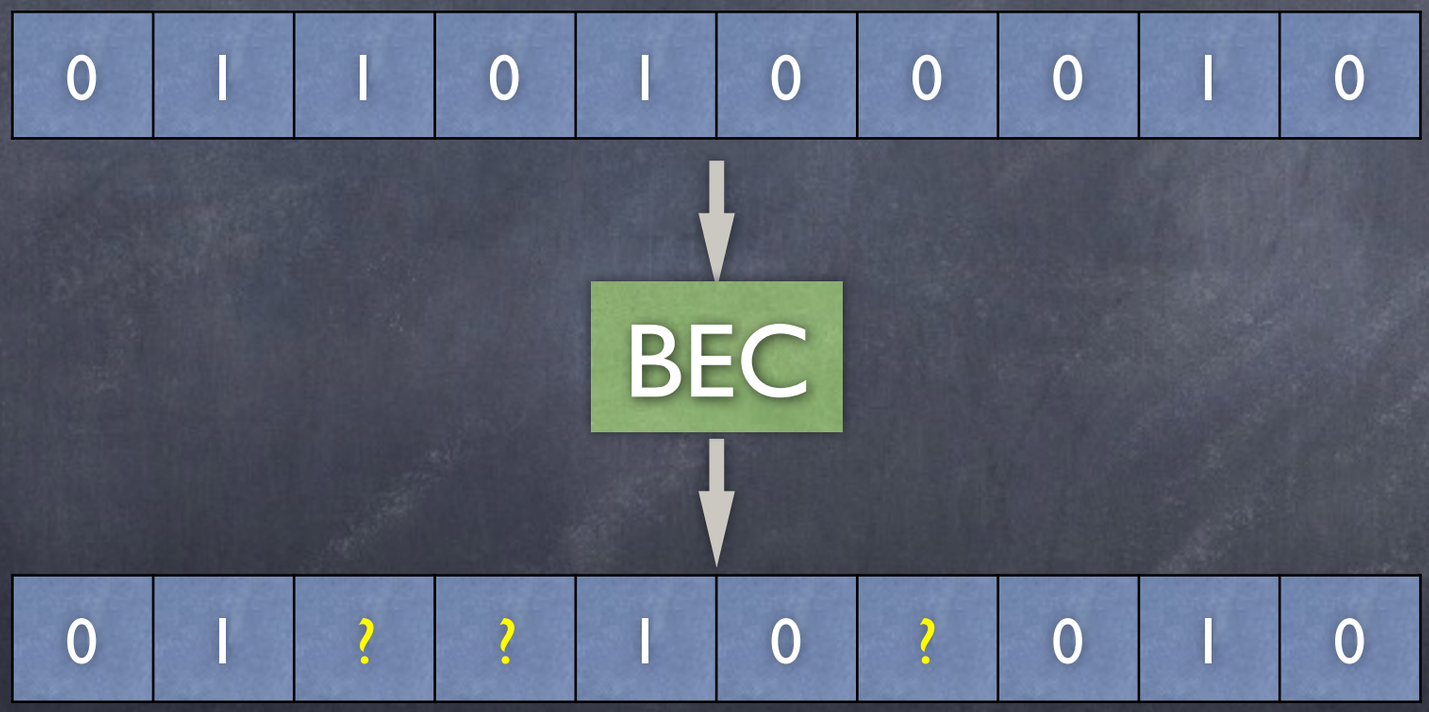
- Low degree polynomials ubiquitous in TCS:
  - lower bounds
  - derandomization
  - PCP
  - Hardness amplification
  - List decoding
  - Algorithms
  - Property testing
  - Extractors
  - …

Decoding RM codes

- Decoding problem: decode corrupted codewords
- Worst case behavior well understood:
  - Reed: Efficient decoding up to half min. distance
  - Gopalan-Klivans-Zuckerman, Bhowmick-Lovett: List decoding radius $\leq 2 \cdot$ dist
- Can we do better for random errors?
- Average case study of errors
- 50 years old open problem in coding theory:
  - How good are RM codes?
  - Do RM codes meet Shannon's bounds for random errors and erasures?

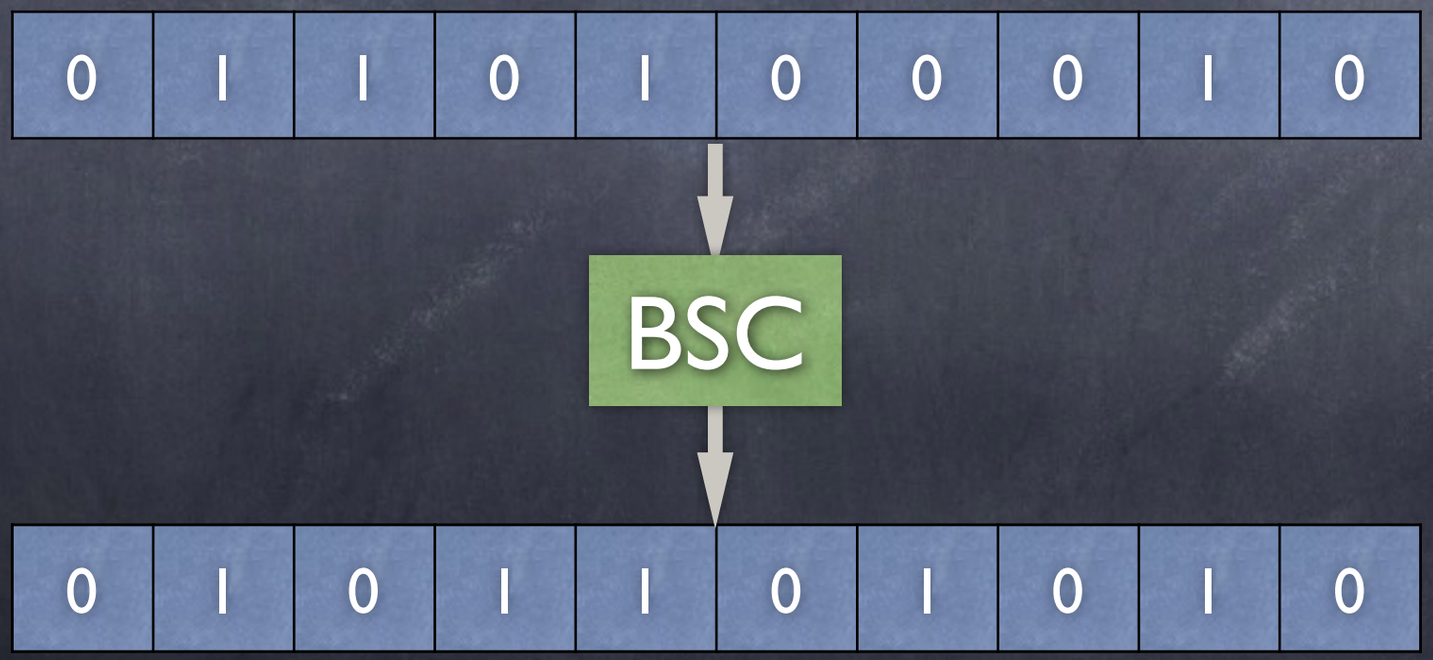
Error Model ↔ Channels

- Binary Erasure Channel:
  symbol is replaced with ? (erased) with probability p



Error Model ⟷ Channels

- Binary Erasure Channel:
  symbol is replaced with ? (erased) with probability p
- Binary Symmetric Channel:
  symbol is flipped with probability p



Error Model ⟷ Channels

- Binary Erasure Channel:
  symbol is replaced with ? (erased) with probability p
- Binary Symmetric Channel:
  symbol is flipped with probability p

- Shannon: maximal rate that enables decoding w.h.p. (capacity of channel) - best tradeoff between redundancy and robustness
  - BEC: $R = 1 - p$
  - BSC $R = 1 - h(p)$ $(h(x) = -x\log(x)-(1-x)\log(1-x))$
- Major goal: design explicit codes that meet Shannon's bound (with efficient encoding and decoding)

Average case behavior of RM

- Do Reed-Muller codes meet Shannon's bound?
  I.e., can RM codes of rate R handle the same fraction of errors/erasures that random codes of rate R handle.
- Problem related to
  - Rank of random evaluation matrices
  - Spaces of tensors over $\mathbb{F}_2$
  - Polynomial interpolation from noisy data
  - Sparse recovery
  - Learning
  - Polar codes
  - …

Polar Codes

- Introduced by Arikan 2009 and now get all the rage in the coding theory community
- Very similar to RM - messages are polynomials with respect to a different monomial basis (different choice of rows from the Serpiensky matrix)
- Achieve capacity for all channels!
- Due to strange basis (no simple description), Polar codes seem less natural than RM codes, yet are still morally similar
- It would be much nicer to work with RM codes rather than Polar!

Our Results

- Theorem 1 [ASW]: RM achieve capacity for BEC at low rate
- Theorem 2 [ASW]: RM achieve capacity for BSC at low rate
- Theorem 3 [ASW]: RM achieve capacity for BEC at high rate
- Theorem 4 [ASW]: RM decodable from "many" errors at high rate (super polynomial in min. distance and polynomially smaller than what capacity achieving code can do)
- Theorem 4' [ASW,SSV]: If RM(m,m-r-1) achieves capacity for BEC, then can efficiently correct $\binom{m}{\leq r}$ many errors in RM(m,m-2r-2).
- Important tool: understanding weight distribution of RM codes (how many codewords of any given weight are there)
- Theorem 5 [ASW]: Tighter bounds on weight distribution of RM codes (based upon and improves Kaufman-Lovett-Porat)

Results [ASW]

	BEC	BSC
$r = o(m)$	√	√
$r = m - O\left(\sqrt{m/\log(m)}\right)$	√	$\sqrt{(\#\text{ of errors})} \approx \sqrt{\binom{m}{r}}$

- Note: minimum distance of RM(m,m-s) is 2^s
- Prior work: r=1 (folklore). r=2 (Helleseth, Klove, Levenshtein '05)
- Kumar-Pfister, Kudekar-Mondelli-Sasoglu-Urbanke 15: RM achieve capacity for the BEC at constant Rate !

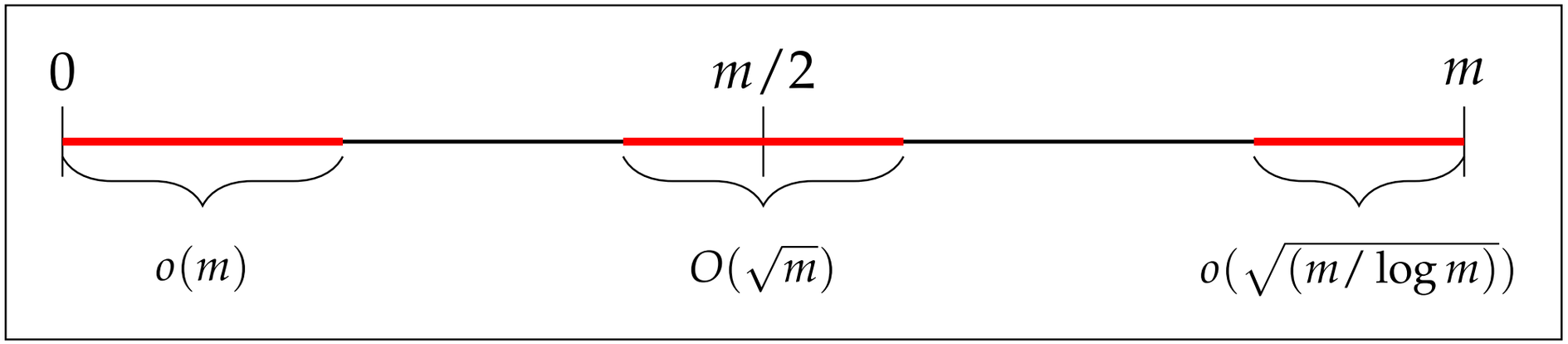


Figure 1: Regime of r for which $RM(m,r)$ is known to achieve capacity for the BEC

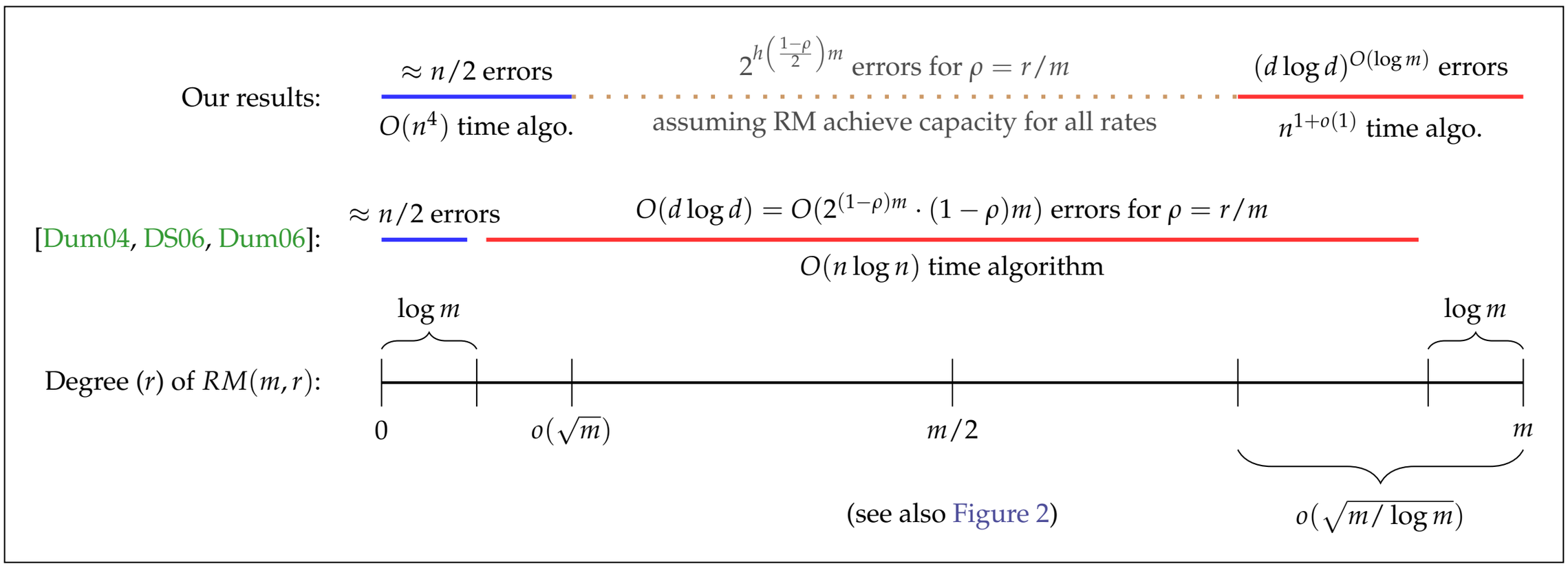


Figure 3: Comparison with [Dum04, DS06, Dum06]

The algorithm

Enter linear algebra

- Let $U = \{u_1,\ldots,u_t\} \subseteq \{0,1\}^m$ be the unknown locations of errors.
- $v \in U$ iff the following linear system, in undetermined coefficients of degree r polynomial f, is solvable:
- For every monomial M of degree at most r:

$$\sum_i f(u_i) = f(v) = 1 \qquad \text{(f not trivial)}$$

$$\sum_i M(u_i) \cdot f(u_i) = M(v) \qquad (v^r \text{ in span of } \{u_i^r\})$$

$$\sum_i M(u_i) \cdot x_j(u_i) \cdot f(u_i) = M(v) \qquad (\text{if } v_j = 1)$$

$$\sum_i M(u_i) \cdot (1-x_j)(u_i) \cdot f(u_i) = M(v) \qquad (\text{if } v_j = 0)$$

- Last two equations: for each j, v spanned by those that agree with it on j'th coordinate
- Crucial: Coefficients of equations can be computed given corrupted word!

Parity Check Matrix

Parity Check Matrix

- H such that v in Code iff Hv = 0
- H is generating matrix for the dual code
  $C^{\perp}$ = {u : u$\perp$v for all v in C}
- Fact $RM(m,r)^{\perp} = RM(m,m-r-1)$
- I.e., P.C.M is an evaluation matrix (monomials vs. points)

Example PCM of RM(4,1):

Dual of RM(4,1) is RM(4,4-1-1)=RM(4,2)

	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
x_1									1	1	1	1	1	1	1	1
x_2					1	1	1	1					1	1	1	1
x_3			1	1			1	1			1	1			1	1
x_4		1		1		1		1		1		1		1		1
x_1x_2													1	1	1	1
x_1x_3											1	1			1	1
x_1x_4										1		1		1		1
x_2x_3							1	1							1	1
x_2x_4						1		1						1		1
x_3x_4				1				1				1				1

Correcting erasures and PCM

- Want to solve linear system

$$\boxed{\text{PCM}} \times \begin{pmatrix} 0 \\ ? \\ ? \\ 1 \\ 1 \\ ? \\ 0 \\ ? \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

- Number of erasures slightly less than number of rows
- Unique solution (unique decoding) iff corresponding sub-matrix has full column-rank.

Back to the algorithm

- Let $U = \{u_1,\ldots,u_t\} \subseteq \{0,1\}^m$ be the unknown locations of errors.
- $v \in U$ iff the following linear system, in undetermined coefficients of degree r polynomial f, is solvable:
- For every monomial M of degree at most r:

$$\sum_i f(u_i) = f(v) = 1 \qquad \text{(f not trivial)}$$

$$\sum_i M(u_i) \cdot f(u_i) = M(v) \qquad (v^r \text{ in span of } \{u_i^r\})$$

$$\sum_i M(u_i) \cdot x_j(u_i) \cdot f(u_i) = M(v) \qquad (\text{if } v_j = 1)$$

$$\sum_i M(u_i) \cdot (1-x_j)(u_i) \cdot f(u_i) = M(v) \qquad (\text{if } v_j = 0)$$

- Last two equations: for each j, v spanned by those that agree with it on j'th coordinate
- Crucial: Coefficients of equations can be read from Hy !!

Summary

- Abbe-S-Wigderson: RM codes achieve capacity for BEC for rates close to 0 or 1 and for BSC at rate close to 0
- First improvement on a 50 years old problem
- Kumar-Pfister, Kudekar-Mondelli-Sasoglu-Urbanke: RM achieve BEC capacity for constant rate
- Saptharishi-S-Volk: efficient algorithm for correcting most error patterns of weight t in RM(m,m-2t)
- Open problem: Do RM codes achieve capacity for the BSC?
- Most interesting case: what can be said for the BSC for constant rate (degree $m/2 \pm O(\sqrt{m})$)?
- Open problem: Do RM codes achieve capacity for the BEC for all rates?

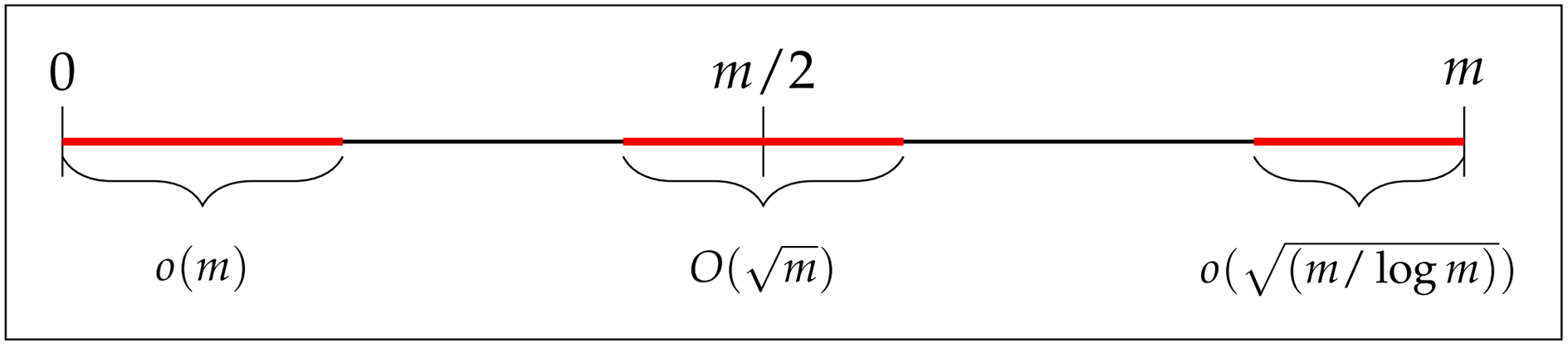


Figure 1: Regime of r for which $RM(m,r)$ is known to achieve capacity for the BEC

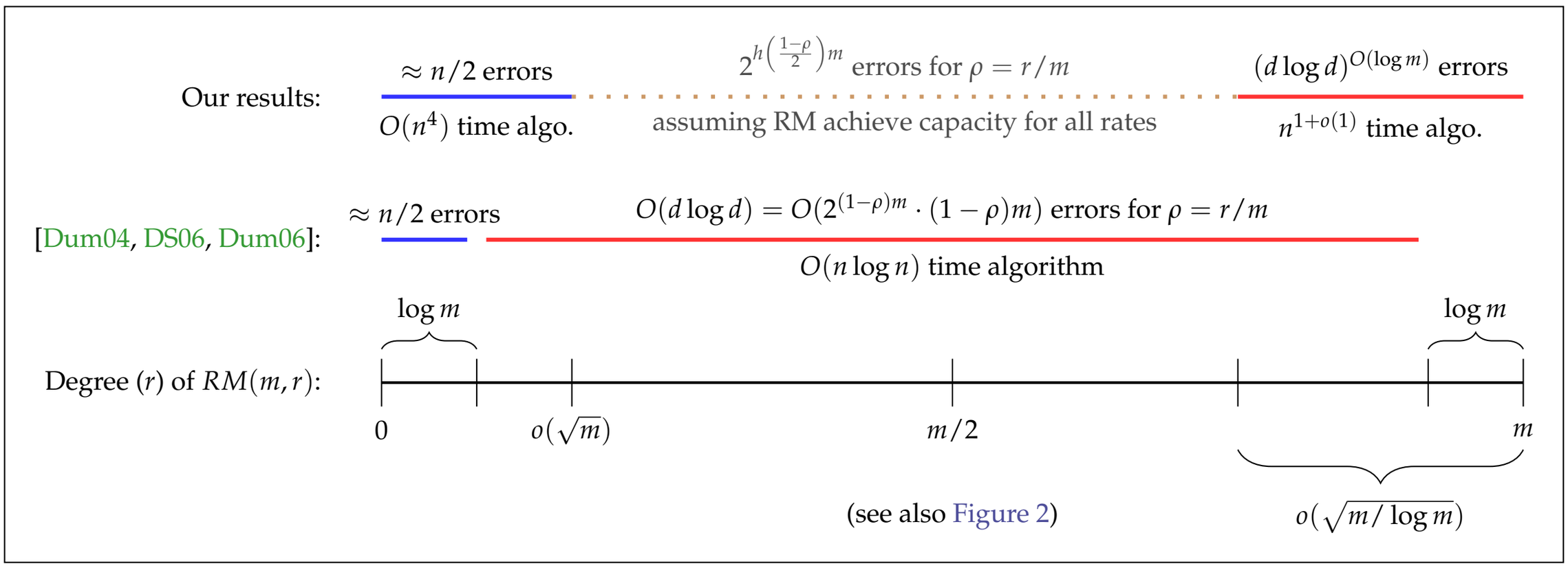


Figure 3: Comparison with [Dum04, DS06, Dum06]

Thank You!