FEC Proposal for NGEPON

Mark Laubach, Shaohua Yang, Ryan Hirth, Glen Kramer

15 May 2017

Introduction

- Various FEC codes and performance have been presented
  - effenberger_3ca_1_1116, RS and LDPC
  - houtsma_3ca_1_0916, RS and LDPC
  - laubach_3ca_2_0317, Folded BCH
  - vanveen_3ca_1_0317, RS
- PR30 has a power budget gap that needs to be closed, help from FEC is needed
  - johnson_3ca_1_0317 page 11
    - 5.8 dB for OLT TX DS Power Gap
    - 3.5 dB OLT RX US Power Gap
- FEC gains have suggested +1 dB to > +2 dB over 10G-EPON RS
  - Selecting an FEC code argues for higher NECG (Net Effective Coding Gain) that meets error performance for expected raw BER, appropriate noise model, and implementation complexity
- Having same FEC code for downstream and upstream is important
  - 10G-EPON vendor experience: speeds development, helps in troubleshooting

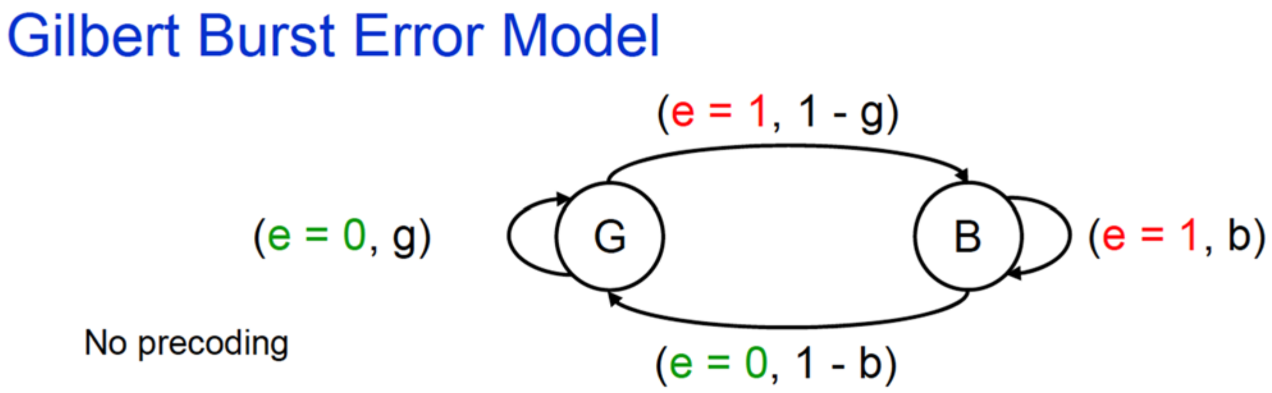
Raw BER and Noise Model

- 25G channel performance is more challenging
  - Argument for 10^{-2} rBER in laubach_3ca_3_0317
- Receiver burst noise becomes more of an issue
  - In reviewing with others in 802.3 WG, AWGN alone is not sufficient, depending on receiver design; e.g. DFE will propagate errors
  - vanveen_3ca_1_0317 discussed Gilbert-Elliot model
  - Other 802.3 efforts refer to Gilbert burst model (802.3 vice chair and secretary at March 2017 meeting)
    - 802.3bj cideciyan_02a_1111 (page 5)
- Burst error mitigation techniques are needed for LDPC
  - Reducing occurrence using pre-coding
  - Interleaving

BROADCOM

Gilbert Burst Error Model

- Attributed to account for DFE error propagation
- Error events are consecutive
- Recommended to use b=0.5
- Stationary state probability
  - $P_B = \frac{1-g}{1.5-g}$, $P_G = \frac{0.5}{1.5-g}$
- Bit error rate
  - $P_e = (1-g)P_G + bP_B = \frac{1-g}{1.5-g}$
    - $g = \frac{1-1.5P_e}{1-P_e}$

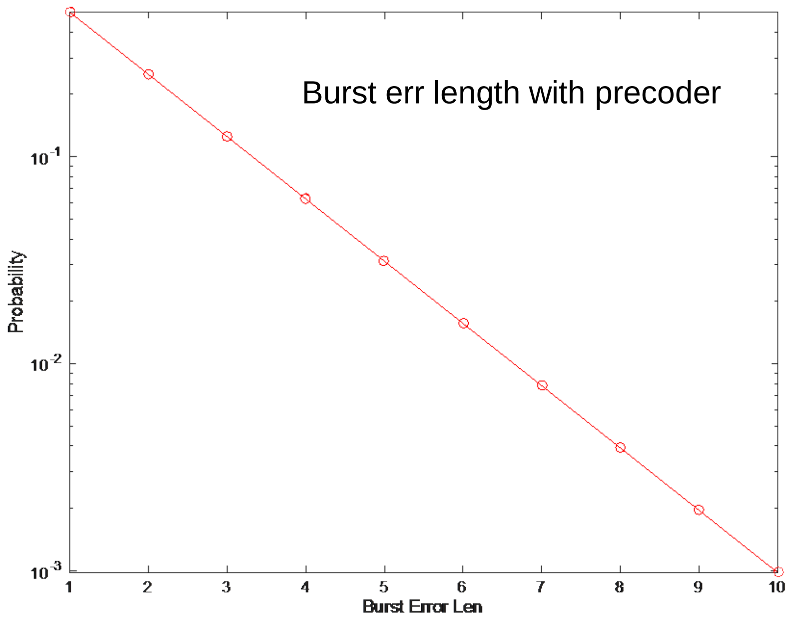


See: http://www.ieee802.org/3/bj/public/nov11/cideciyan_02a_1111.pdf

BROADCOM

1+D Pre-coder (see houtsma_3ca_1_0916, pages 3 and 4)

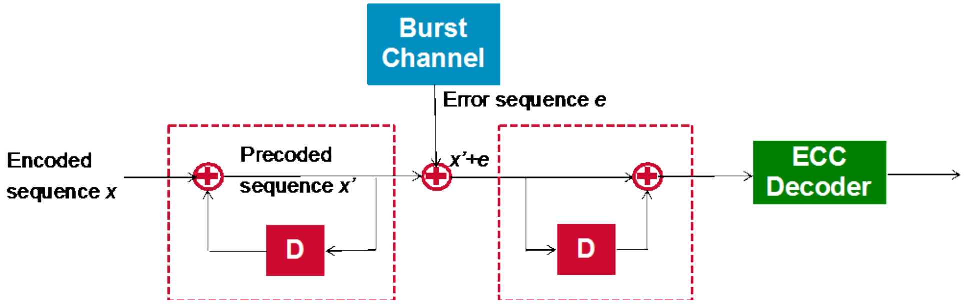
- Encoder: $y_k = x_k + y_{k-1}$ decoder: $x_k = y_k + y_{k-1}$
- Dual bit errors
  - example 1: 0eeeee0➔0e0000e
  - example 2: 0e0 ➔ 0ee
- Without pre-coder: average error event length = $\sum_{i=0}^{\infty}(i+1)(1-b)b^i$
  - b=0.5➔average error length = 2
- With 1+D pre-coder:
  - All burst error events contain two error bits
  - b=0.5➔ bit error rate $P_{ep} = P_e$
- Pre-coder reduces BER if b>0.5, and increases error rate if b<0.5

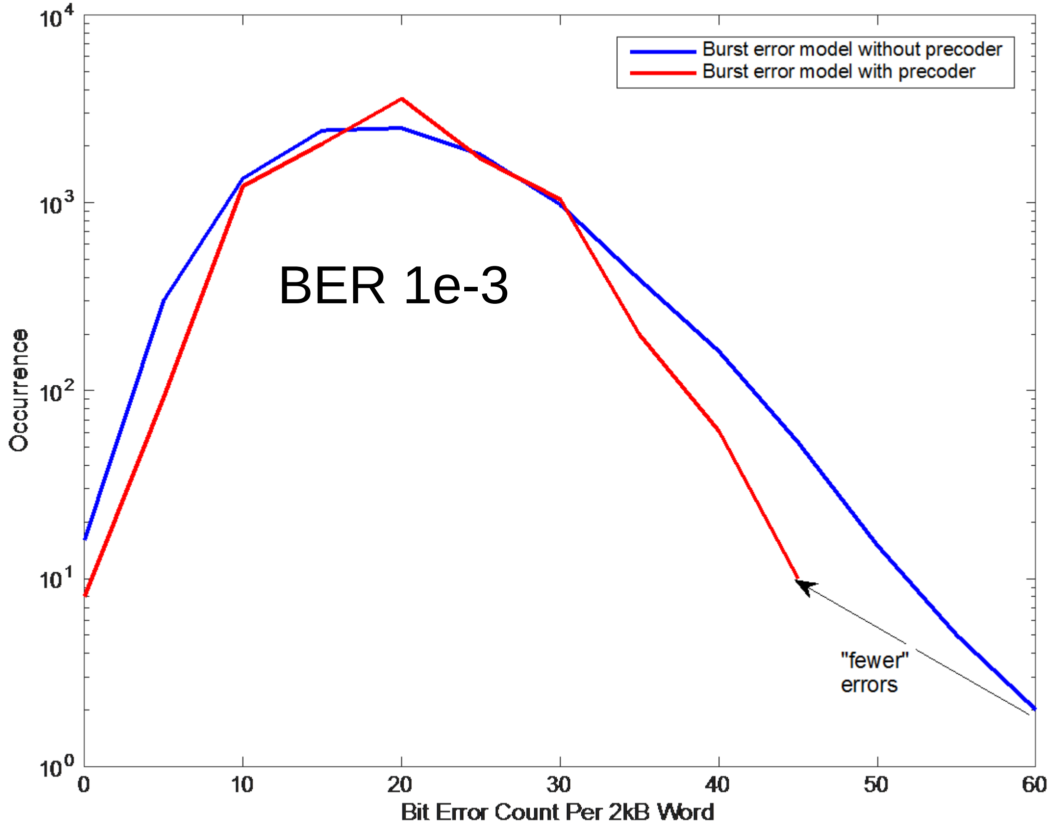


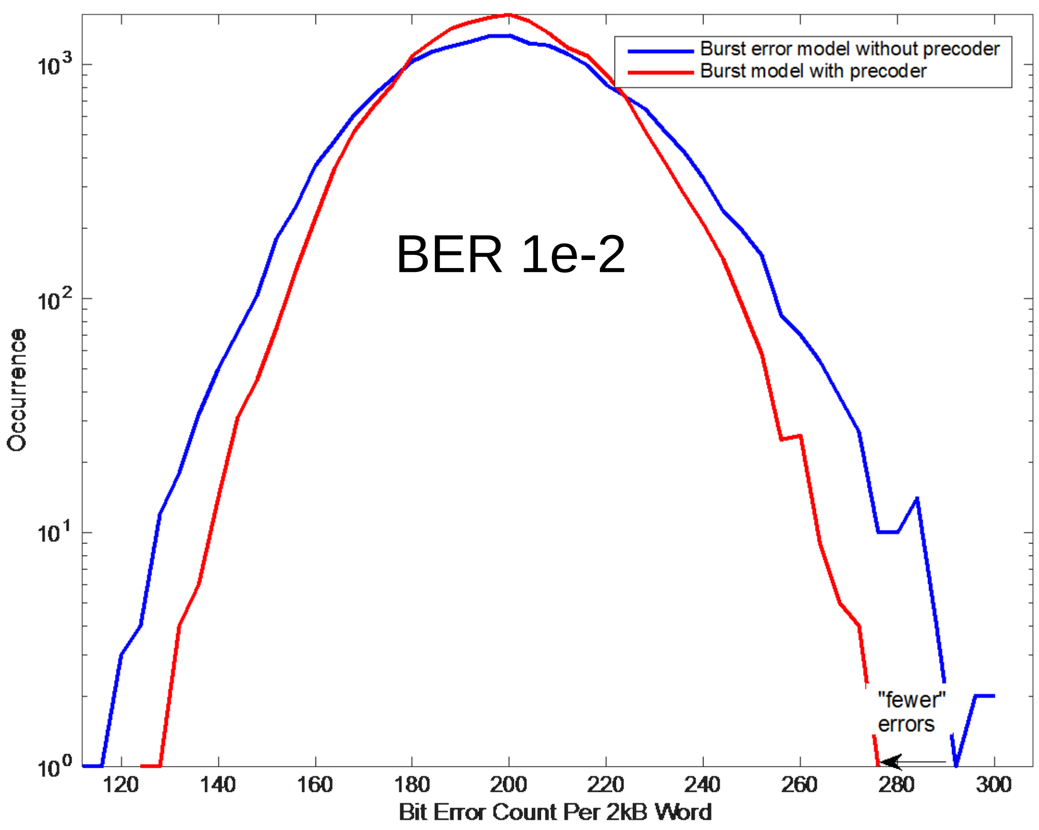
BROADCOM®

1+D Precoder and Bit Error Distribution

- Precoder maps any burst error event to a two bit error
  - example 1: 0eeeee0➔0e0000e
  - example 2: 0e0 ➔ 0ee



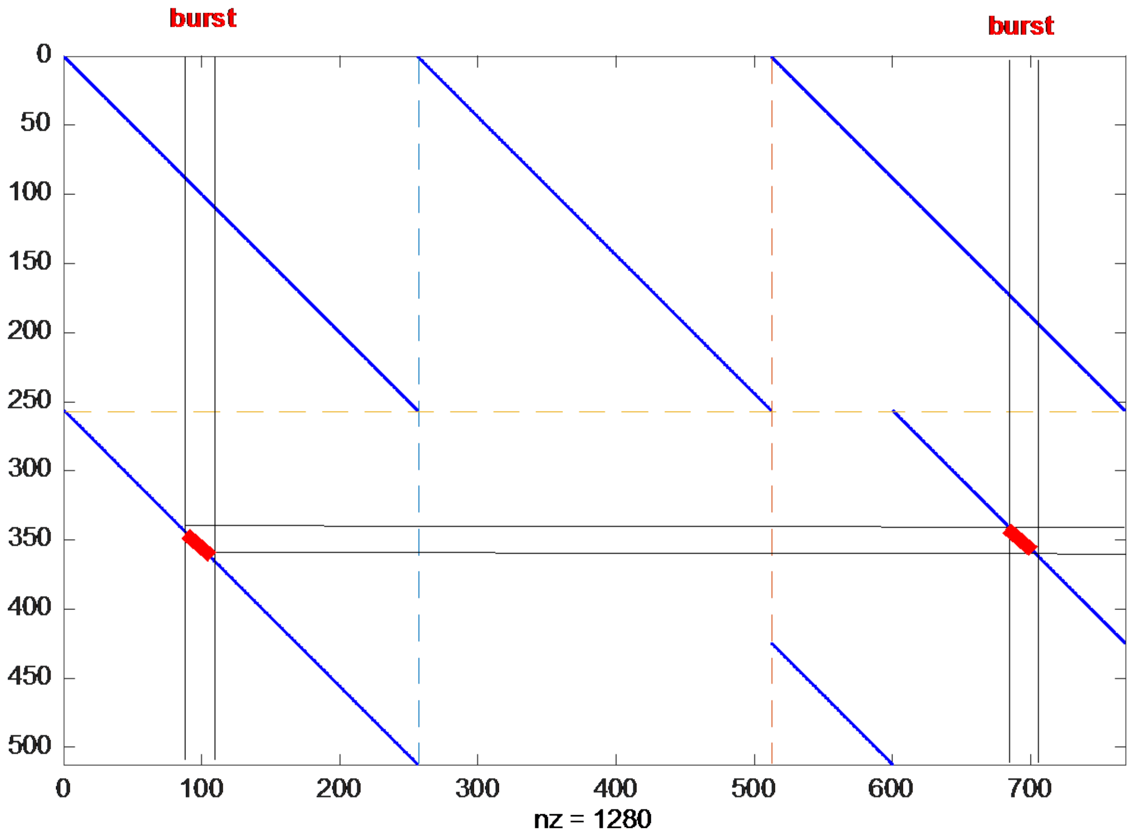




BROADCOM

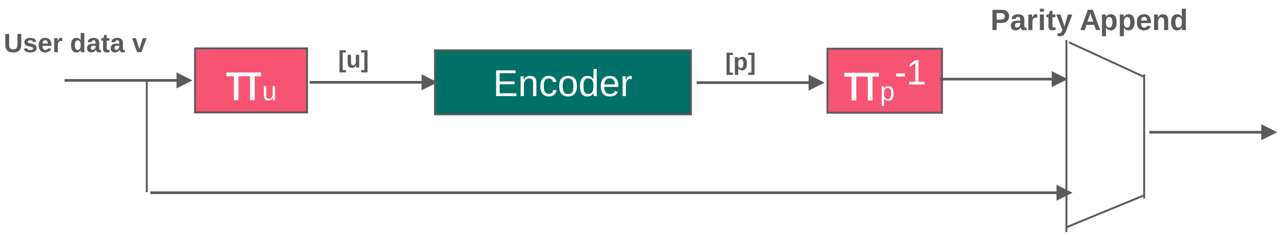
Benefit of Interleaver and LDPC

- Without precoder or interleaver, bursty errors would disable parity checks
  - This is specific to the quasi-cyclic parity code structure
  - Either need to remove the burst errors or apply an interleaver



BROADCOM

Basic Interleaver Concept



Channel order:
$[\pi_p^{-1}(p), \pi_u^{-1}(u)]$

$$v=\pi_u^{-1}(u)$$

Decoder order:
$[p, u]*H^T=0$

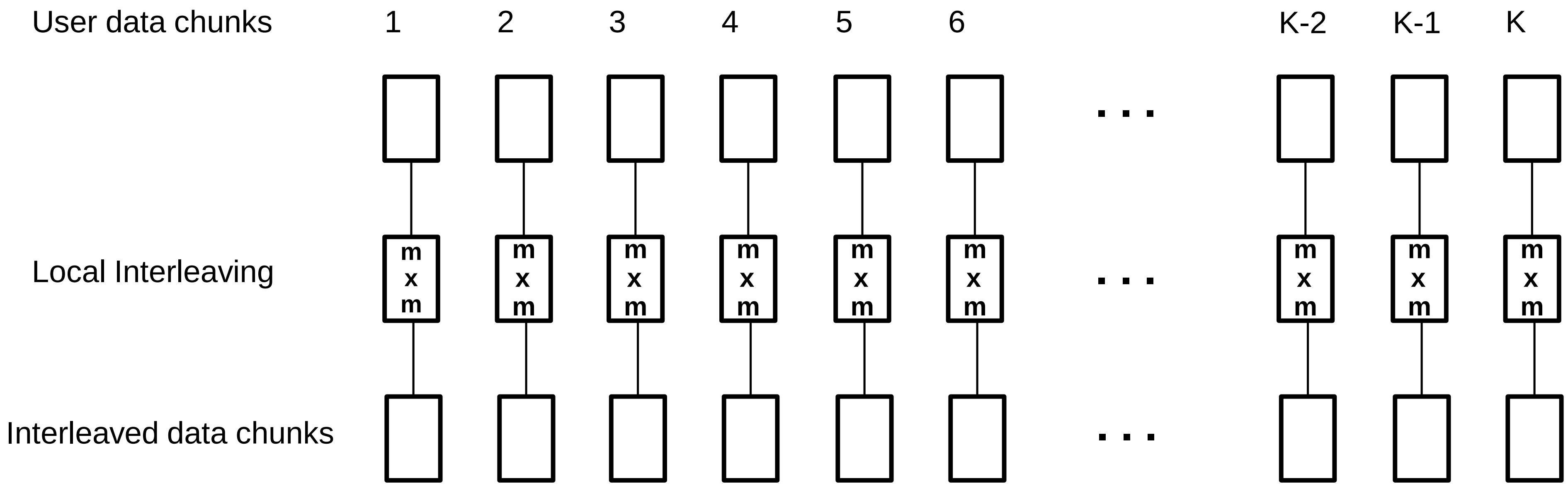
- Interleaving is used to protect data against burst errors. Burst errors can overwrite many bits in a row, so LDPC codes that expect errors to be more uniformly distributed can be overwhelmed. Interleaving is used to help stop this from happening.
- User data is first permuted using user symbol-level interleaver π_u, then fed to LDPC encoder. The parity symbols at encoder output are permuted using parity symbol de-interleaver π_p^{-1}.
- Systematic encoding, where the user data is explicitly written onto channel (or precoder if exists). The bit sequence in channel order is:

  $[\pi_p^{-1}$**(p)**, π_u^{-1}**(u)**$]$ or $[\pi_p^{-1}$**(p) v**$]$

BROADCOM

Local Interleaver

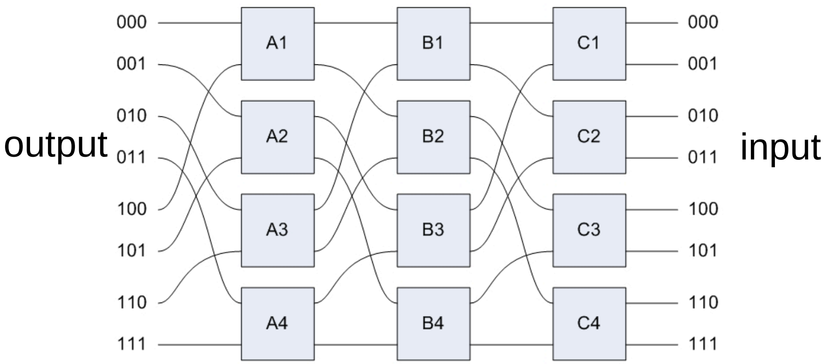
- For low hardware latency, a local interleaver is desirable
- User bits are divided into data chunks which are interleaved by K independent mxm local interleavers
- Locality is controlled by interleaver size (m), chosen to be multiple of circulant size

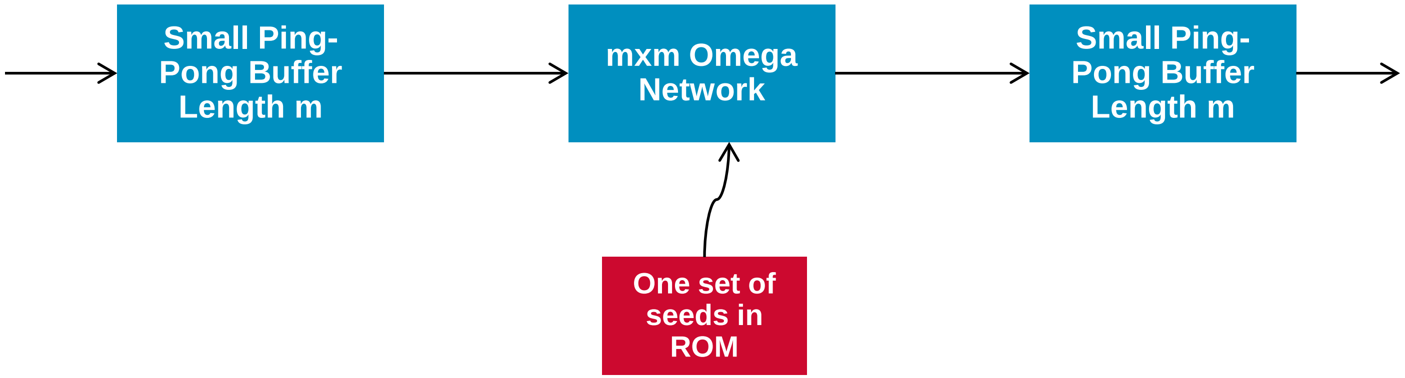


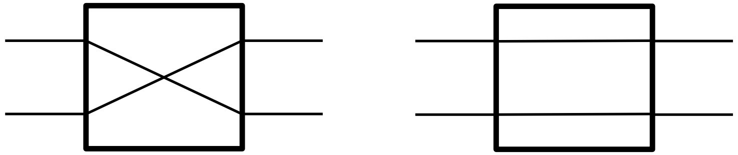
BROADCOM

Omega Switching Networks for Interleaving

- A multistage interconnection network using multiple stages of switches
- A mxm network needs:
  - $\log_2(m)$ stages
  - $m/2$ switches per stage
  - Can be fully controlled by a seed sequence of length $m/2$
  - One set of seeds for each mxm interleaver, stored in a ROM







Switch=1 Switch=0

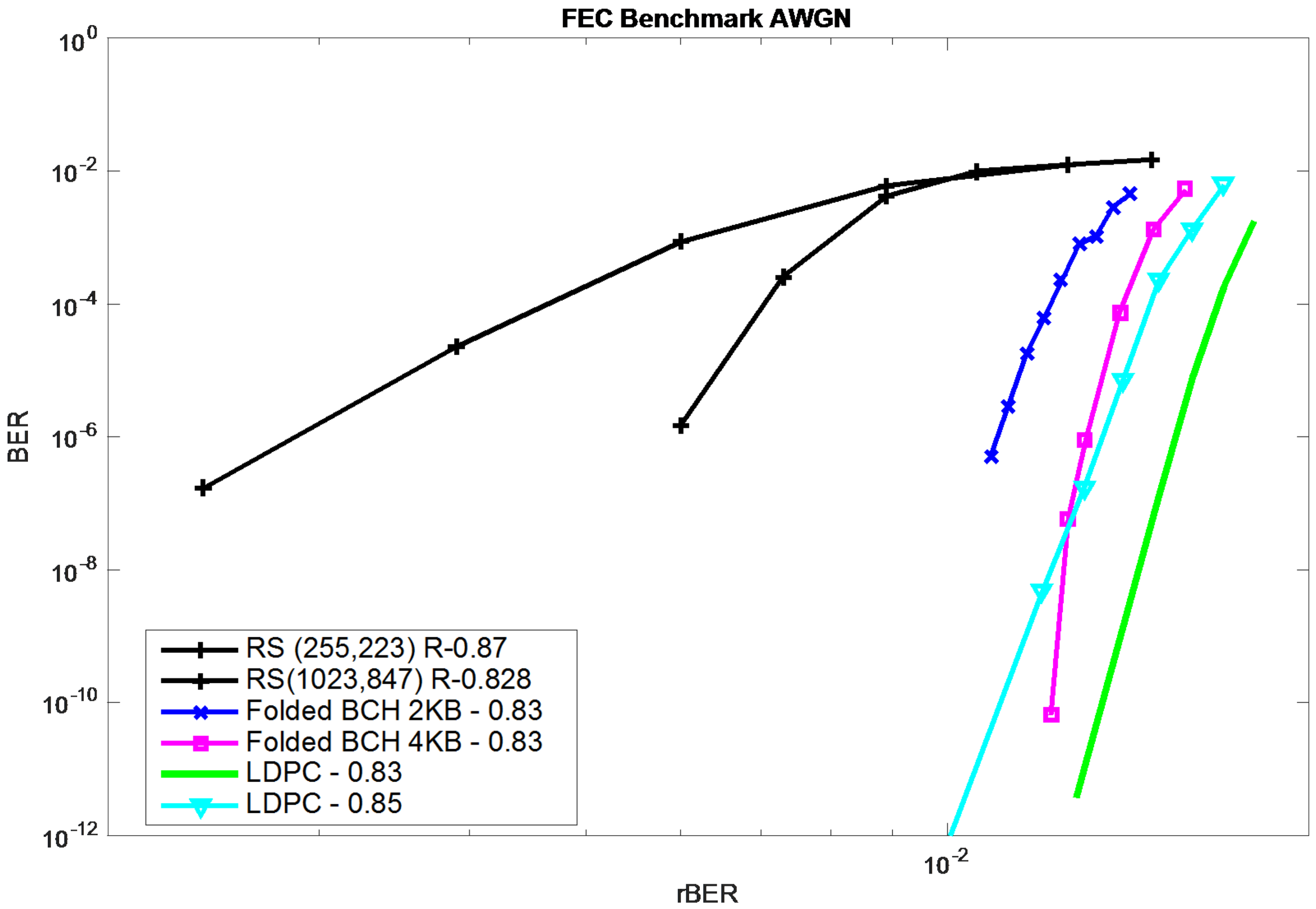
D. H. Lawrie, “Access and alignment of data in an array processor,” IEEE Trans. on Computers, vol. C- 24, pp. 1145-1155, Dec. 1975.

Omega Switching Network Interleave Seed Control

- Rate 0.85 LDPC:
  - 11x74ex256 LDPC matrix
  - 256x256 Omega Network
- Parameters:
  - There are 74 different 256x256 Omega Networks
  - Each network has 8 stages, 128 switches per stage (total 8x128 switches)
    - For each Omega network, ROM stores 128 seed values
    - The actual 8 stage switch values are wrap-around shifted from the stored seed value
      - The wrap around seed shifts are [17, 34, 51, 68, 85, 102, 119, 136]
  - In total, 74x8x128 switches, 74x128 binary seeds stored in ROM
  - Omega 256 Seed in ROM: laubach_3ca_2_0517.txt

BROADCOM

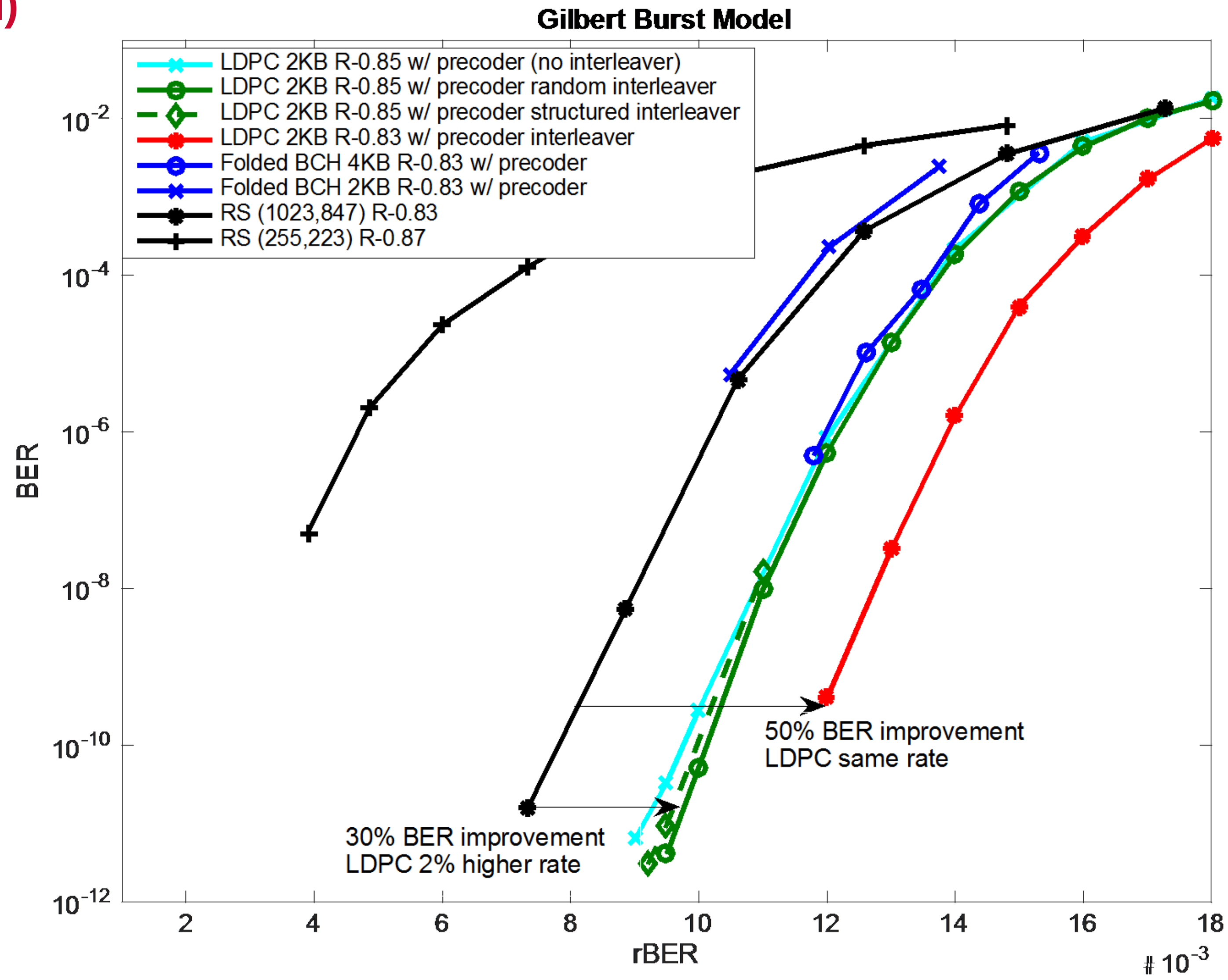
FEC Benchmark Over AWGN



Interleaver has no impact on AWGN performance.

BROADCOM

Simulation Benchmarks
(Gilbert Burst Model)



Note: error floor for both LDPC rates expected to be below 1×10^{-15}.

BROADCOM®

FEC Codes Studied

	Length	Rate	Parity	User	Encoded	NECG[1] (dB)		Notes
						AWGN	Gilbert Burst	
Folded BCH	2kB	0.83	3272	16576	19848	2.25	1.48	bits
	4kB	0.83	6064	30784	36848	2.6	1.78	bits
LDPC	2kB	0.848	2816	15677	18493	2.46	1.8	bits (18493,15677)
	2kB	0.833	3200	16000	19200	2.82	2.12	bits (19200,16000)
RS	(255,223)	0.8745	256	1784	2040	0	0	S=8, T=16 (10G-EPON)
	(1023,847)	0.828	1760	8470	10230	1.34	1.35	S=10, T=88

[1] electrical gain over RS(255,223)

BROADCOM

Simulation Results and Summary

- 1+D pre-coder and structured interleaver are beneficial for correcting bursty errors
- LDPC optimization involves degree distribution and deep trapping sets removal (refer to appendix)
- Hard input min-sum soft decoding, iterative decoding for the best NECG
- Folded BCH, while good for flash memory and AWGN, did not perform as well with Gilbert noise model (and also pre-coding)

- LDPC(18493,15677) 0.848 rate, using min-sum decoding and Omega 256 structured interleaver sufficiently provides a NECG that meets error performance objective using 10^{-2} raw input BER, Gilbert burst error model, and pre-coding.
  - On implementation complexity: ~15% less mm^2 than LPDC 0.83 rate code

BROADCOM

LDPC(18493,15677) Parity Code Matrix

- Rate=84.8% (18493,15677), m=11 rows x n=74 columns, L=256
  - Parity code matrix: laubach_3ca_3_0517.txt

```
0 -1 -1 -1 0 -1 -1 -1 0 -1 0 -1 0 -1 -1 -1 -1 -1 -1 0 -1 -1 0 -1 -1 -1 0 -1 -1 -1 0 -1 -1 0 -1 -1 -1 0 -1 -1 0 -1 -1 -1 0 -1 -1 -1 0 -1 -1 -1 0 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0;
-1 0 -1 -1 -1 0 -1 -1 0 -1 -1 0 -1 -1 -1 -1 0 -1 0 -1 -1 -1 -1 0 -1 -1 88 -1 -1 0 -1 -1 -1 62 -1 -1 -1 176 -1 -1 -1 0 -1 -1 -1 0 -1 -1 -1 0 -1 0 -1 -1 236 85 177 113 11 52 246 234 90 87 137 75 209 223 127 96 255 118 224 47;
-1 33 -1 -1 -1 -1 0 -1 -1 0 -1 -1 1 -1 -1 -1 160 0 -1 -1 -1 -1 -1 -1 0 -1 -1 0 -1 149 -1 -1 -1 -1 0 -1 -1 -1 0 -1 -1 -1 0 -1 -1 -1 0 -1 -1 -1 0 -1 243 -1 0 81 7 144 11 187 109 77 214 14 120 231 239 135 115 202 195 124 234 99;
7 -1 -1 -1 -1 -1 42 0 -1 -1 -1 115 -1 -1 0 -1 -1 -1 -1 250 -1 -1 -1 -1 -1 0 -1 207 -1 -1 180 -1 -1 -1 -1 0 -1 -1 39 -1 -1 -1 212 -1 170 -1 -1 -1 -1 -1 -1 36 -1 0 128 223 10 156 119 108 97 220 81 158 211 8 151 66 166 25 115 138 196 70;
-1 -1 0 -1 -1 -1 -1 210 2 -1 -1 -1 -1 0 -1 -1 -1 16 -1 36 -1 -1 -1 46 -1 -1 242 -1 -1 -1 -1 0 -1 -1 106 -1 -1 244 -1 -1 -1 -1 72 -1 -1 -1 129 -1 -1 -1 236 -1 -1 -1 42 223 83 49 108 194 228 98 248 87 121 216 16 160 181 162 60 234 63 187;
-1 -1 -1 0 -1 -1 244 -1 -1 -1 124 -1 -1 11 -1 -1 -1 -1 221 -1 0 -1 -1 -1 21 -1 -1 -1 0 -1 -1 -1 0 -1 -1 4 0 -1 -1 -1 -1 -1 -1 0 -1 82 -1 -1 -1 230 -1 -1 -1 255 18 211 255 196 25 77 17 142 158 203 167 244 88 217 21 225 236 52 186 121;
-1 -1 0 179 -1 -1 -1 -1 -1 -1 -1 165 1 -1 -1 0 -1 -1 -1 -1 -1 0 117 -1 -1 -1 -1 -1 -1 244 -1 17 -1 -1 -1 -1 0 -1 70 -1 -1 -1 -1 212 -1 -1 -1 0 184 -1 -1 -1 41 -1 150 146 162 142 158 239 206 94 194 13 213 164 222 66 207 40 132 110 199 251;
-1 -1 -1 112 96 -1 -1 -1 -1 50 -1 -1 -1 -1 220 -1 22 -1 -1 -1 30 -1 -1 251 -1 -1 -1 229 -1 -1 -1 -1 148 -1 -1 -1 0 -1 -1 0 -1 -1 -1 117 17 -1 -1 -1 -1 109 -1 233 -1 -1 218 204 76 48 21 133 172 28 215 214 43 36 70 193 5 123 191 84 66 170;
-1 19 -1 -1 23 -1 -1 12 -1 -1 -1 -1 -1 -1 -1 7 -1 76 -1 -1 159 -1 -1 -1 -1 197 -1 -1 152 -1 -1 176 -1 251 -1 -1 -1 -1 -1 255 72 -1 -1 -1 -1 -1 -1 251 -1 -1 147 -1 -1 -1 178 214 144 187 211 52 60 190 186 167 46 14 172 117 96 198 179 113 142 217;
166 -1 -1 -1 -1 116 -1 -1 -1 158 -1 -1 -1 113 -1 136 -1 -1 -1 -1 -1 64 -1 -1 -1 177 -1 -1 22 -1 136 -1 -1 -1 -1 38 -1 -1 -1 94 -1 165 -1 -1 -1 -1 177 -1 201 -1 -1 -1 -1 191 186 59 39 239 70 158 100 69 91 87 127 81 126 105 60 192 10 5 193 182;
-1 -1 1 -1 -1 98 -1 -1 -1 -1 31 -1 -1 -1 226 -1 -1 -1 188 -1 -1 94 202 -1 161 -1 -1 -1 -1 -1 -1 -1 75 -1 233 -1 -1 -1 -1 -1 236 29 -1 -1 -1 190 -1 19 -1 -1 -1 -1 -1 -1 136 88 33 241 129 177 70 82 47 154 60 145 234 98 173 219 152 186 62 124;
```

Proposed Motion

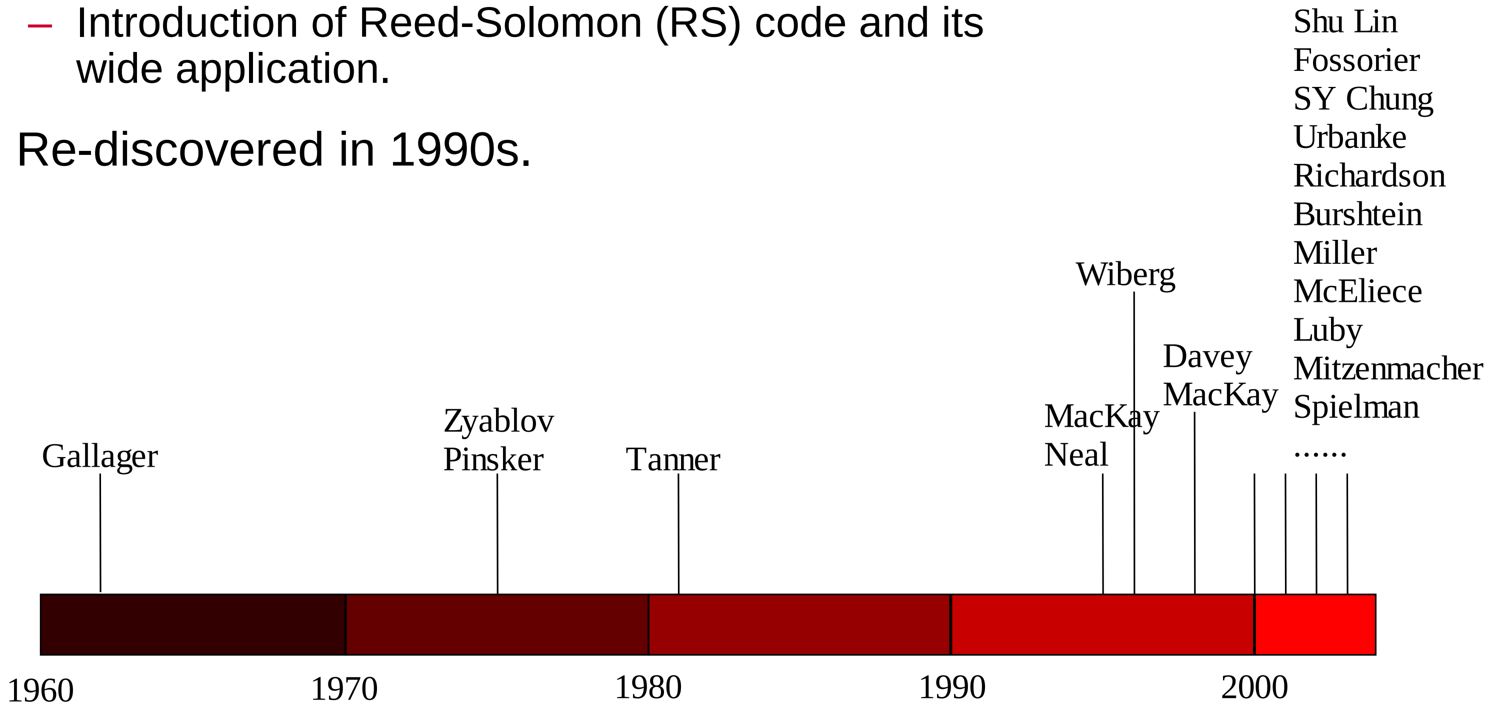
- Adopt the LDPC(18493,15677) 0.848 rate FEC code presented in laubach_3ca_1_0517.pdf and use of pre-coding for downstream and upstream channels. The Omega 256 seed code as in laubach_3ca_2_0517.txt. The LDPC parity code matrix as in laubach_3ca_3_0517.txt.

BROADCOM

Appendix

History of LDPC Codes

- LDPC was invented by Gallager in his PhD thesis, largely ignored due to
  - High complexity (in 1960s)
  - Introduction of Reed-Solomon (RS) code and its wide application.
- Re-discovered in 1990s.

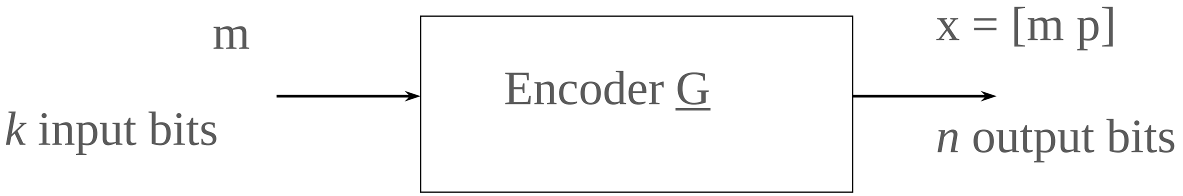


Error correction	introduced
check digit	1850s-1900s
Checksum	1940s-1960s
Hamming codes	1950s
Reed-Solomon	**1960s**
Turbo codes	1990s
LDPC codes	**1960s**

BROADCOM

Fundamentals of Linear Block Code

- A linear code C (over a finite field) can be defined in terms of either a generator matrix or parity-check matrix.
- Generator matrix G ($k \times n$)



- Parity-check matrix H ($n\text{-}k \times n$)

$$x \bullet H^T = 0$$

- The capacity of correcting symbol errors in a codeword is determined by the minimum distance (dmin)

Example: (7, 4) Hamming Code

- The generating set for the (7,4) Hamming code:

  1000 ===> 1101000;   0100 ===> 0110100

  0010 ===> 1110010;   0001 ===> 1010001

- Every codeword is a linear combination of these 4 codewords.

  That is: **x** = **m** •**G,** where

$$\mathbf{G} = \begin{bmatrix} \underbrace{\begin{matrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{matrix}}_{k\times(n-k)} & \underbrace{\begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}}_{k\times k} \end{bmatrix} = [\mathbf{P} \,|\, \mathbf{I}_{\mathbf{k}}]$$

- The (7,4) Hamming code has dmin=3.

Parity-Check Matrix

- For $\mathbf{G} = [\ \mathbf{P} \mid \mathbf{I}_k\ ]$, define the matrix $\mathbf{H} = [\mathbf{I}_{n\text{-}k} \mid \mathbf{P}^{\mathrm{T}}]$
- (The size of $\mathbf{H}$ is $(n\text{-}k)\mathrm{x}n$).
- It follows that $\mathbf{G}\mathbf{H}^{\mathrm{T}} = \mathbf{0}$.
- Since $\mathbf{x} = \mathbf{m}\bullet\mathbf{G}$, then $\mathbf{x}\bullet\mathbf{H}^{\mathrm{T}} = \mathbf{x}\bullet\mathbf{G}\mathbf{H}^{\mathrm{T}} = \mathbf{0}$.
- The parity check matrix of code C is the generator matrix of another code C_d, called the dual of C.

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

BROADCOM®

Low Density Parity Check Matrix

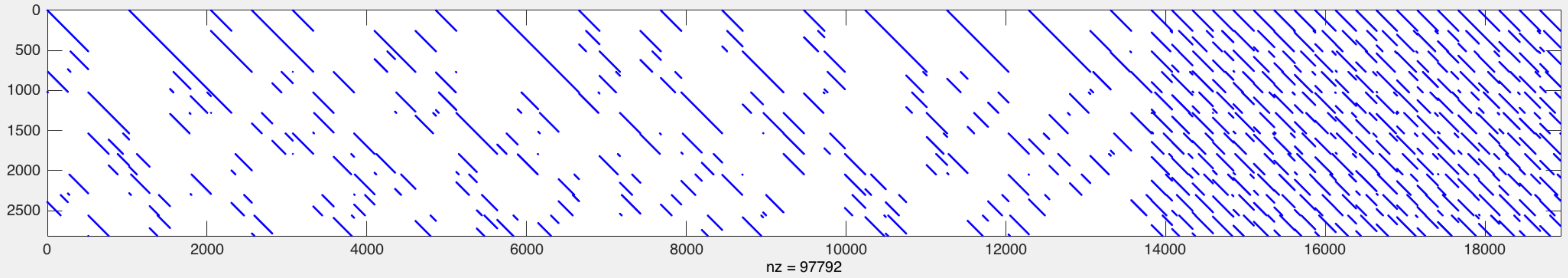
- # of 1s in each column is small, i.e., 3~11.
- Rate 85% H is 2816x18944 for 2k word size, density =0.18%

$$H = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & . & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & . & 0 & 1 & 0 \\ . & . & . & . & . & . & . & . & . \\ 1 & 0 & 0 & 0 & 1 & . & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & . & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & . & 0 & 1 & 0 \end{bmatrix}$$

- The linear code defined by the low density parity check matrix is a LDPC code.
- What is the magic of low density?
  - Iterative decoding
  - Soft decoding
  - Long block code

LDPC Code Optimized for EPON

- Rate 85% code
  - H is 2816x18944 for 2k word size
  - Quasi-cyclic structure (see next page)



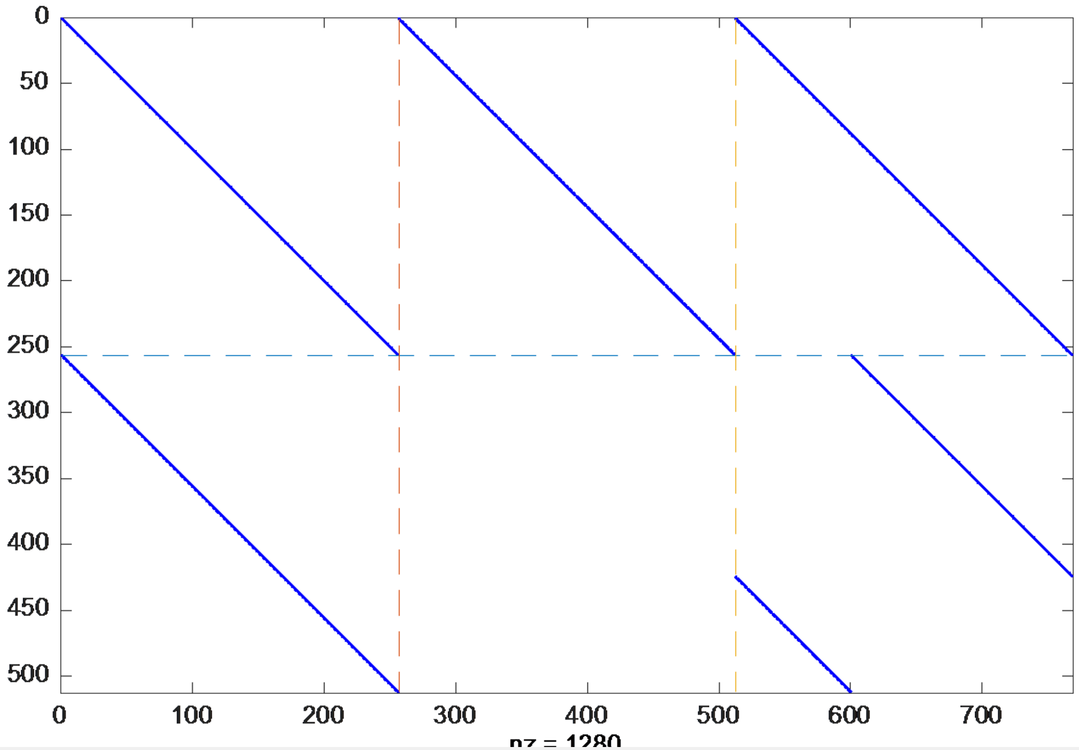
BROADCOM

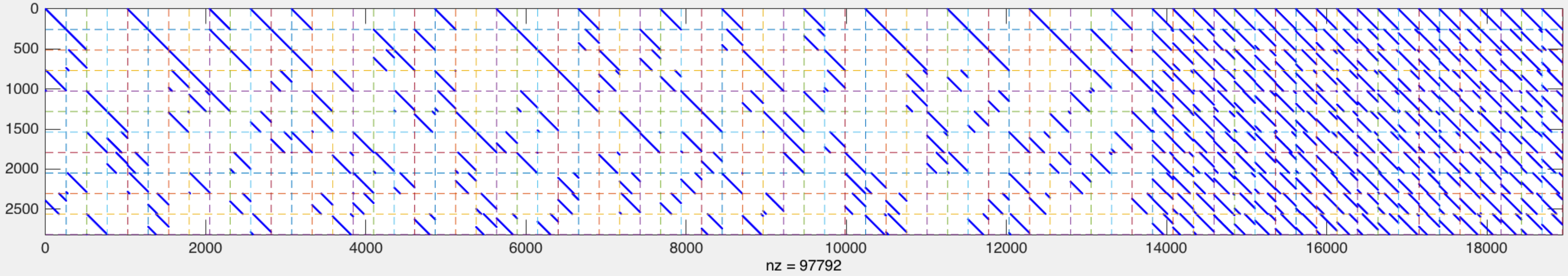
Quasi-Cyclic Structure

- Quasi-cyclic H matrix structure
  – 256x256 cyclic permutation submatrices

    $Q_k*[x_0, x_1, \ldots, x_m]^T = [x_k,\ldots,x_m, x_0, \ldots, x_{k-1}]^T$

    – A barrel shifter can be used for QC permutation
    – Circuit parallelism corresponds to the circulant size
  – H matrix has a low density
    – # of 1s in each column is small, 3 or 11, density =0.18%
    – Every bit is connected with 3 or 11 parity checks
    – Every parity check is connected with 35 bits



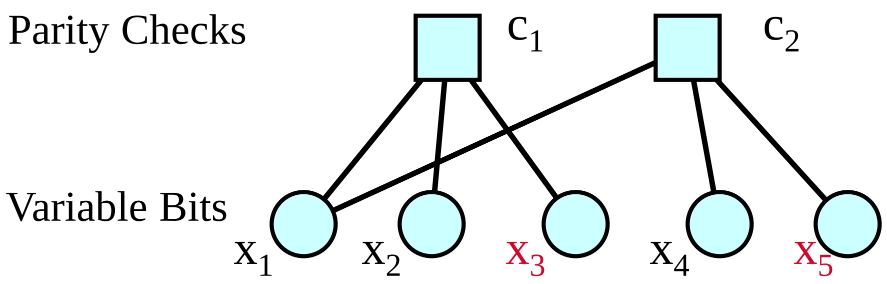


BROADCOM

Variable Node and Check Node

Example. (5, 3) toy code

$$x_1 \oplus x_2 \oplus x_3 = 0 \qquad x_1 \oplus x_4 \oplus x_5 = 0$$



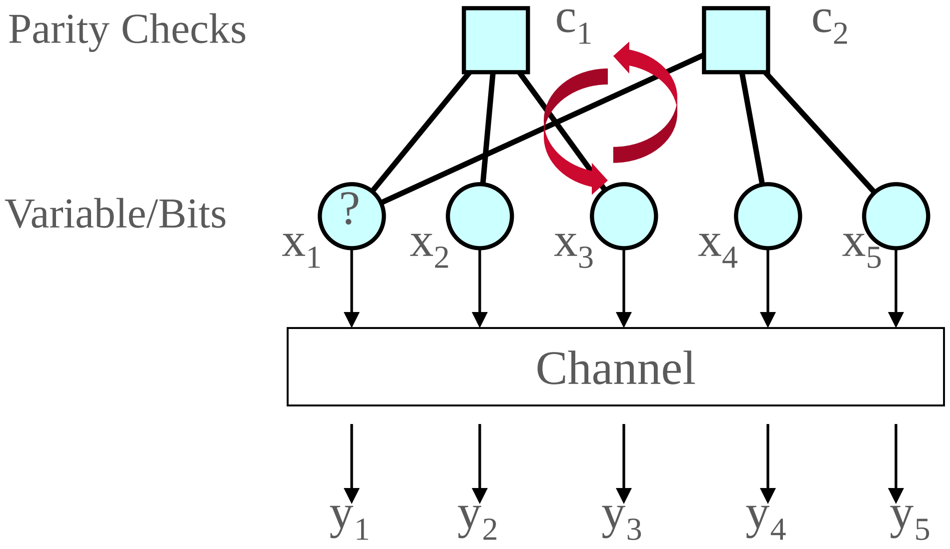
8 valid codewords

1	2	3	4	5
0	0	0	0	0
0	0	0	1	1
0	1	1	0	0
0	1	1	1	1
1	0	1	0	1
1	0	1	1	0
1	1	0	0	1
1	1	0	1	0

Min Sum Decoding

- Define log-likelihood ratio (LLR) as $L(x_k) = \log\left(\frac{P(y|x_k=0)}{P(y|x_k=1)}\right)$

$$\begin{cases} x_1 \oplus x_2 \oplus x_3 = 0 \\ x_1 \oplus x_4 \oplus x_5 = 0 \end{cases}$$



Notation:

Input LLR: $L_{x_i} \leftarrow$ 0:0.75, and 1:-0.75

Variable to check: $L_{x_i \to c_j}$

Check to variable: $L_{c_j \to x_i}$

Initialization:

$L_{c_j \to x_i} = 0$

Bit-to-Check: (sum)

$$L_{x_i \to c_j} = L_{x_i} + \sum_{m \neq j} L_{c_m \to x_i}$$

Check-to-Bit: (min)

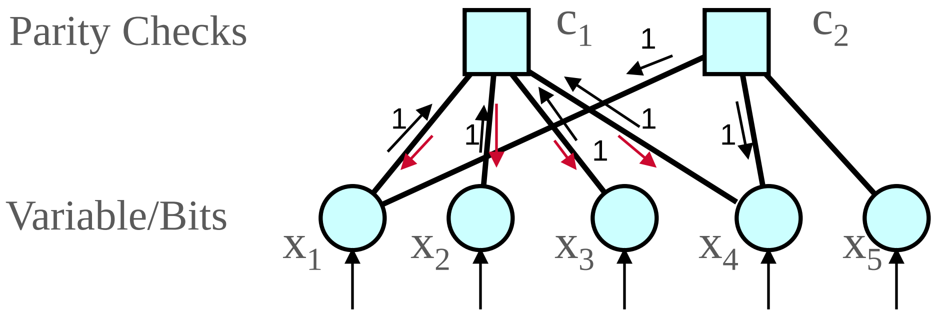
$$L_{c_j \to x_i} = \alpha \prod_{m \neq i} sign\left(L_{x_m \to c_j}\right) \min_{m \neq i}(|L_{x_m \to c_j}|)$$

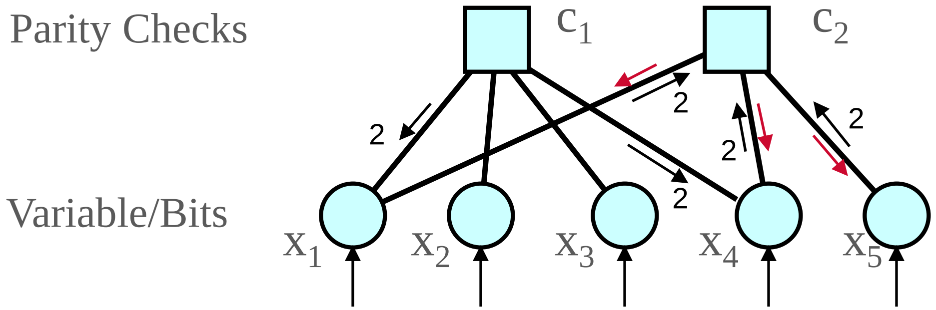
[1] M. Fossorier, M. Mihaljevic, and H. Imai, “Reduced complexity iterative decoding of low-density parity check codes based on belief propagation,” IEEE Trans. Commun., vol. 47, pp. 673-680, May 1999.

[2] J. Chen and M. Fossorier, “Density evolution for two improved BPsased decoding algorithms of LDPC codes,” IEEE Commun. Lett., vol. 6, pp. 208-210, May 2002.

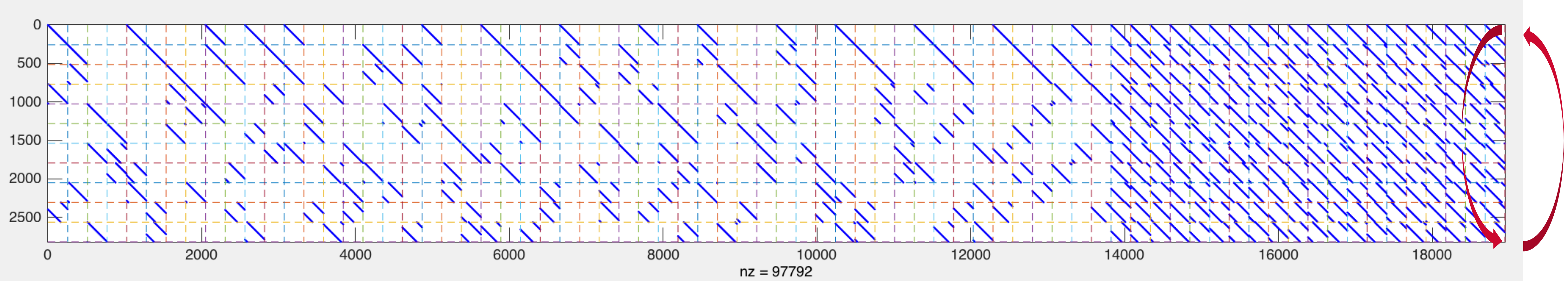
Layered Decoding Scheduling

– Non-layered: bit by bit, process all connected checks at the same time
– Layered: check by check, process all connected bits in groups





– Suitable for irregular code
– Each clock processes 512 bits (two circulants) on the same H matrix layer
– Decoder synthesis area: area is comparable to RS codec



[3] M. M. Mansour and N. R. Shanbhag, “Turbo decoder architectures for low-density parity-check codes,” in Proc. IEEE Global Telecommunications Conf. (IEEE GLOBECOM), Taipei, Taiwan, Nov. 2002, pp. 1383–1388

BROADCOM®

connecting everything®