Area and Energy Efficient VLSI Architectures for Low-Density Parity-Check Decoders using an On-the-fly Computation

Kiran Gunnam
Texas A&M University
October 11 2006

Outline

- Introduction of LDPC
- Problem Statement
- On-the-fly computation for QC-LDPC
- Multi-rate TPMP Decoder
- Multi-rate Layered Decoder
- Irregular QC-LDPC (IEEE802.16e, IEEE 802.11n,…)
- Fixed code TPMP Decoder
- Parallel Layered Decoder
- Results and performance comparison

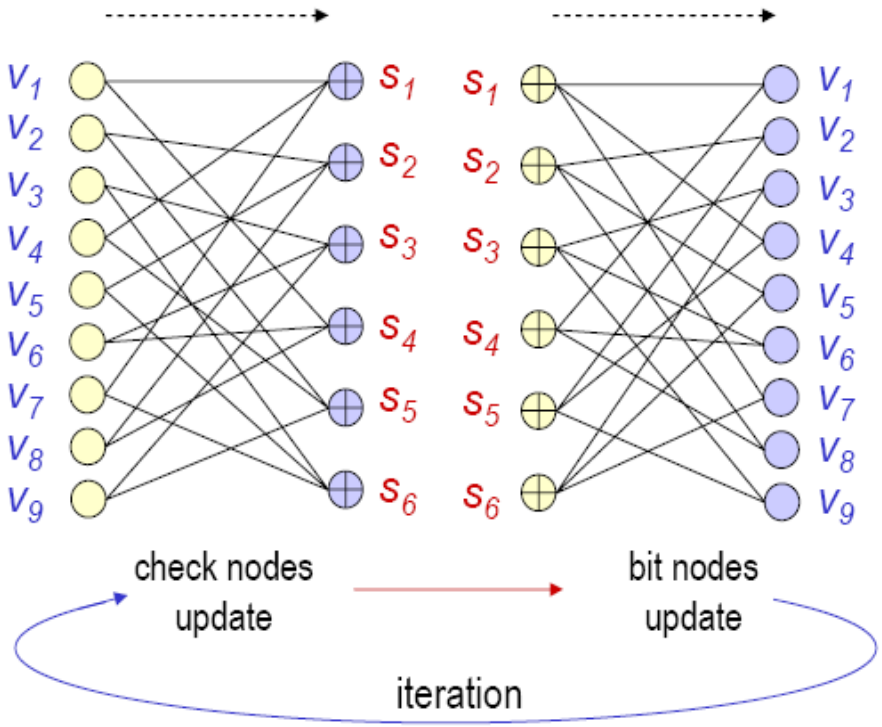
Example LDPC Code

check nodes

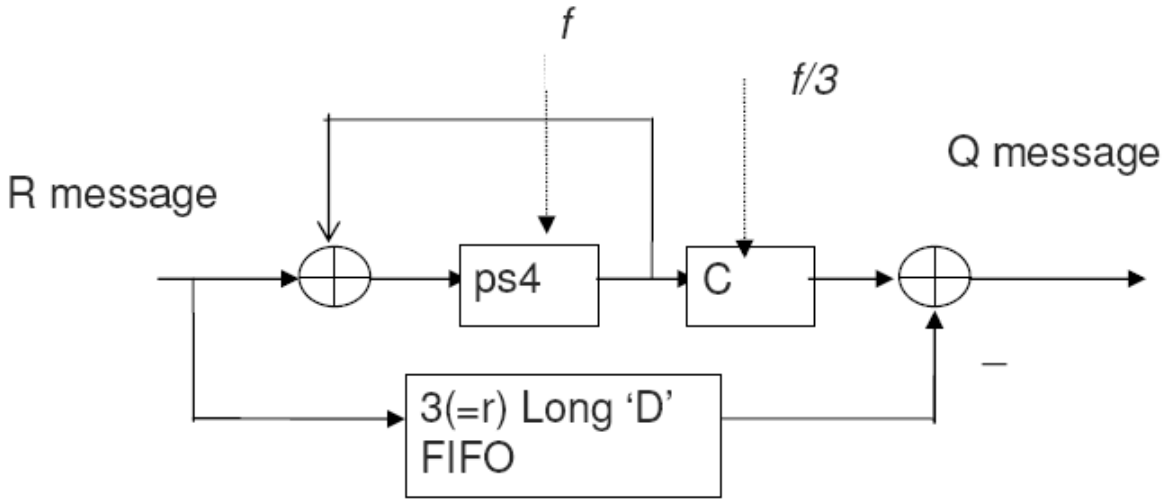
$$H=\begin{bmatrix}1&0&0&1&0&0&1&0&0\\0&1&0&0&1&0&0&1&0\\0&0&1&0&0&1&0&0&1\\1&0&0&0&0&1&0&1&0\\0&1&0&1&0&0&0&0&1\\0&0&1&0&1&0&1&0&0\end{bmatrix}\begin{matrix}s_1\\s_2\\s_3\\s_4\\s_5\\s_6\end{matrix}$$

$v_1 v_2 v_3 v_4 v_5 v_6 v_7 v_8 v_9$

variable nodes



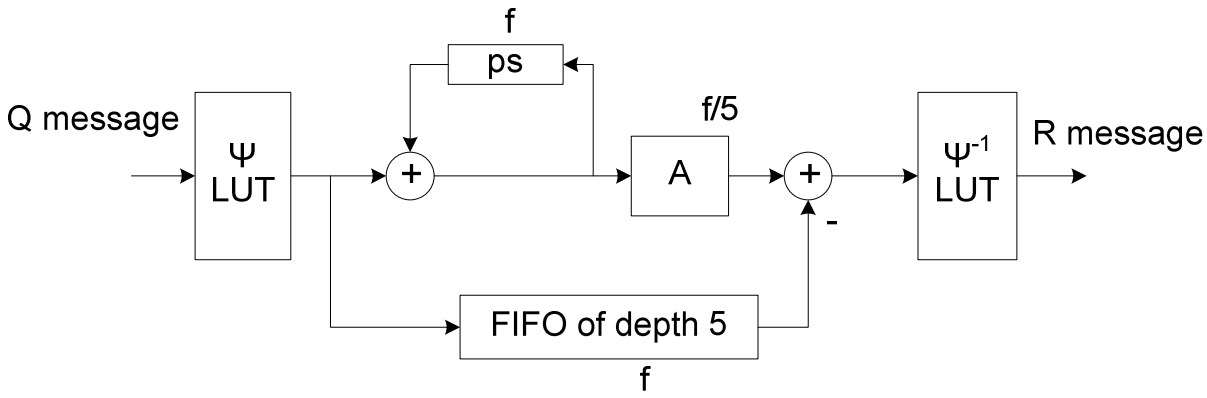
Micro-Architecture for VNU



Serial Variable Node Unit (VNU) to compute the variable node messages (Q) using check node messages (R)

$$Q_{bi,cj} = \left(\sum_{j' = Col[bi][1]}^{Col[bi][r]} R_{j',bi} \right) - R_{cj,bi} + \wedge(bi)$$

Micro-Architecture for CNU



A serial Check Node Unit (CNU) for Sum of Products algorithm.

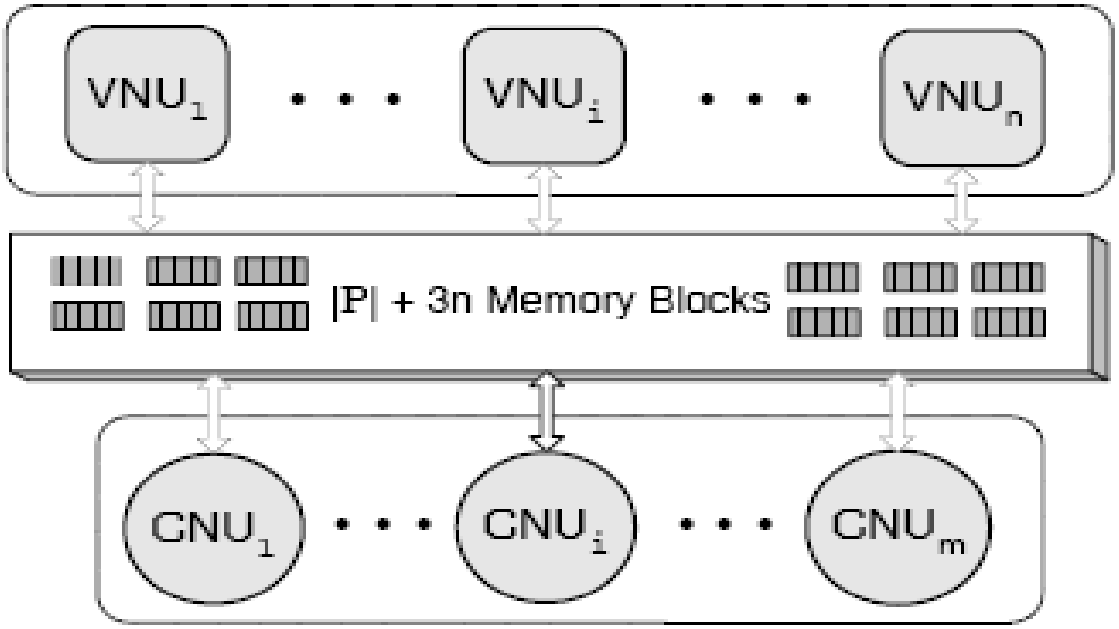
$$R_{cj,bi} = \psi^{-1}\left[\left(\sum_{i'=Row[cj][1]}^{Row[cj][c]} \psi\left(Q_{i',cj}\right)\right) - \psi\left(Q_{bi,cj}\right)\right].\delta(cj,bi)$$

Decoder Architectures

- Fully Parallel Architecture:
  - All the check updates in one clock cycle and all the bit updates in one more clock cycle.
  - Huge Hardware resources and routing congestion.
- Serial Architecture
  - All Check updates and bit updates in a serial fashion.
  - Huge Memory requirement. Memory in critical path.

Semi-Parallel Architecture

Check updates and bit updates using several units.
Partitioned memory by imposing structure on H matrix.
Practical solution for most of the applications.
Complexity differs based on architecture and scheduling



Semi Parallel Architecture

[1] T. Zhang and K. K. Parhi, Joint (3,k)-Regular LDPC Code and Decoder/Encoder Design, IEEE Transactions on Signal Processing April, 2004.

Outline

- Introduction of LDPC
- Problem Statement
- On-the-fly computation for QC-LDPC
- Multi-rate TPMP Decoder
- Multi-rate Layered Decoder
- Irregular QC-LDPC (IEEE802.16e, IEEE 802.11n,…)
- Fixed code TPMP Decoder
- Parallel Layered Decoder
- Results and performance comparison

Problem Statement

- The authors in [2] reported that 95% of power consumption of the decoder chip developed in [1] results from memory accesses.
- The authors in [2] reported that 50% of their decoder power is from memory accesses.
- Memory access is a bottleneck in preventing full utilization of units.

Component	Power (%)
Memory	95.4
CNU	2.5
VNU	1.5
Router	0.6

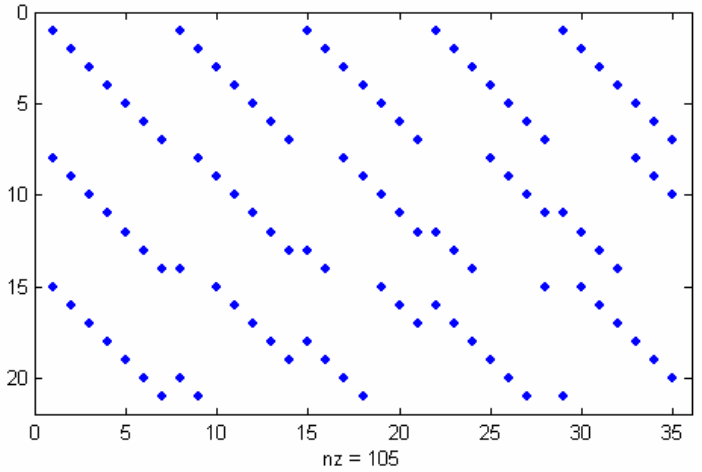
[2] Yijun Li et al, "Power efficient architecture for (3,6)-regular low-density parity-check code decoder," IEEE ISCAS 2004

[3] Mansour et al "A 640-Mb/s 2048-Bit Programmable LDPC Decoder Chip"-IEEE Journal of Solid-State Circuits, March 2006

Quasi-cyclic LDPC Codes

Array Codes

$$H = \begin{bmatrix} I & I & I & \dots & I \\ I & \sigma & \sigma^2 & \dots & \sigma^{r-1} \\ I & \sigma^2 & \sigma^4 & \dots & \sigma^{2(r-1)} \\ : & & & & \\ I & \sigma^{c-1} & \sigma^{(c-1)2} & \dots & \sigma^{(c-1)(r-1)} \end{bmatrix}$$



$$\sigma = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & .0 & 0 \\ : & & & & \\ 0 & 0 & \dots & .1 & 0 \end{bmatrix}$$

Example H Matrix

r row/ check node degree=5

c columns/variable node degree=3

P=7

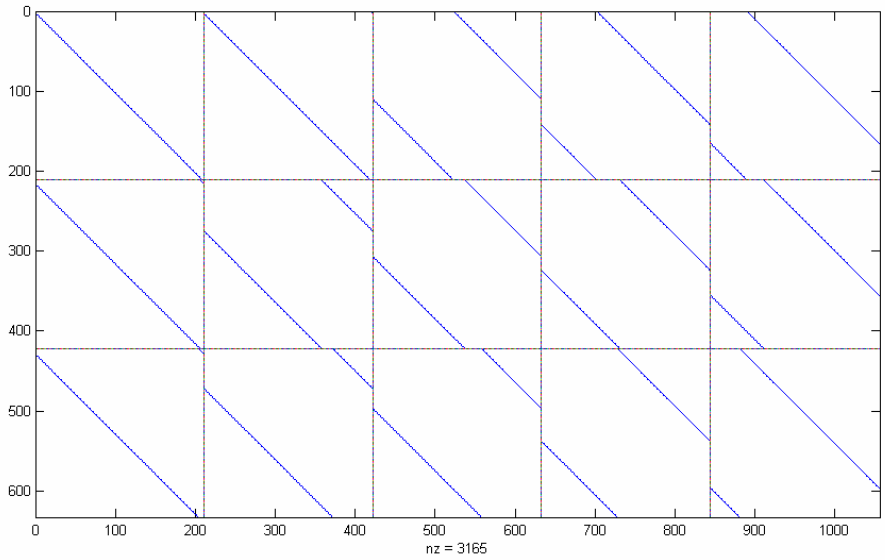
*N=P*r=35*

[4] J. L. Fan, "Array-codes as low-density parity-check codes", In Proc. TPP,2000

Quasi-cyclic LDPC Codes

Cyclotomic Cosets

$$S_{3\times 5} = \begin{bmatrix} 2 & 3 & 110 & 142 & 165 \\ 5 & 64 & 96 & 113 & 144 \\ 7 & 50 & 75 & 116 & 174 \end{bmatrix}$$



Example H Matrix

r row degree=5

c column degree =3

P=211

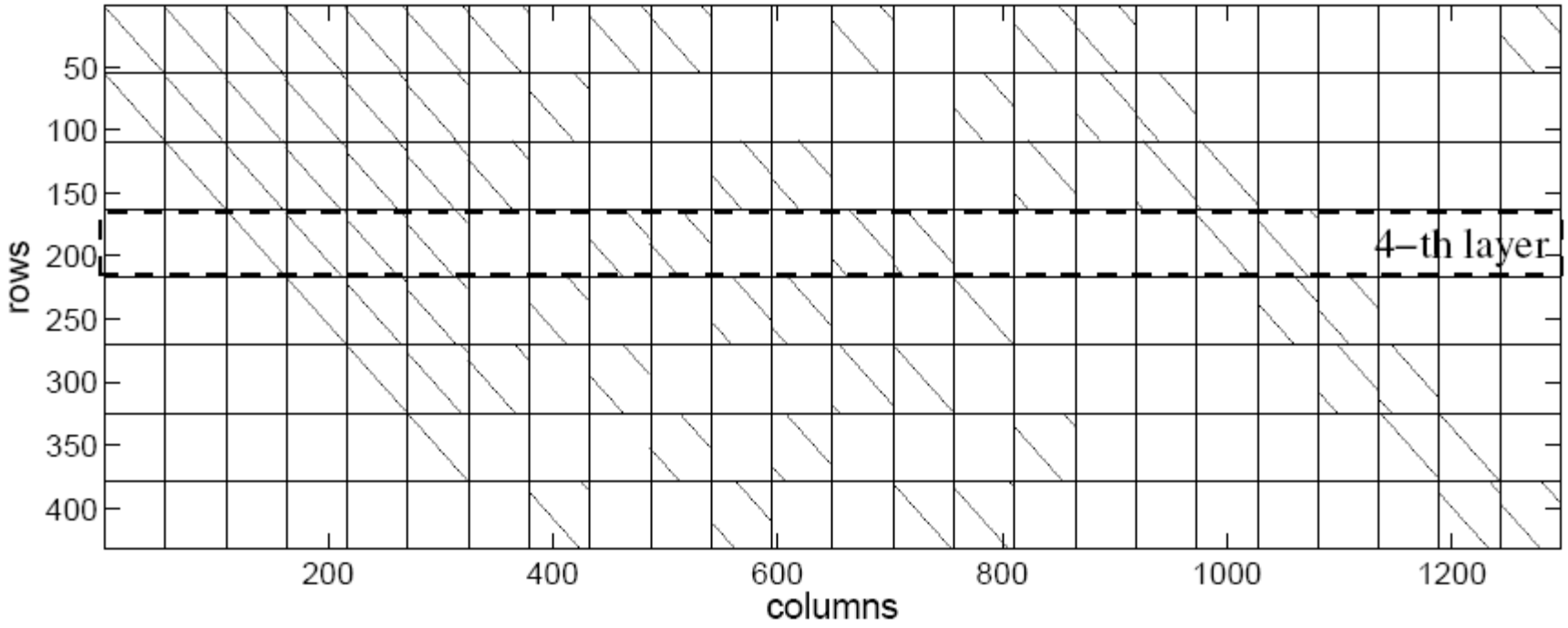
*N=P*r=1055*

M. M. Mansour and N. R. Shanbhag, "Low-power VLSI decoder architectures for LDPC codes," in Proc. Int. Symp. on Low Power Electronics and Design (ISLPED) Monterey, CA, Aug. 2002, pp. 284-289.

Irregular QC-LDPC

Regular QC-LDPC: High throughput and low error floor applications.

Irregular QC-LDPC Have zero-matrices also. Better BER performance. Proposed for all wireless standards.



Outline

- Introduction of LDPC
- Problem Statement
- On-the-fly computation for QC-LDPC
- Multi-rate TPMP Decoder
- Multi-rate Layered Decoder
- Irregular QC-LDPC (IEEE802.16e, IEEE 802.11n,…)
- Fixed code TPMP Decoder
- Parallel Layered Decoder
- Results and performance comparison

On-the-fly computation

This research introduces the following concepts to LDPC decoder implementation

1. Block serial scheduling
2. Value-reuse,
3. Scheduling of layered processing,
4. Out-of-order block processing,
5. Master-slave router,
6. Dynamic state,
7. Speculative Computation
8. Run-time Application Compiler [support for different LDPC codes with in a class of codes. Class:802.11n,802.16e,Array, etc. Off-line re-configurable for several regular and irregular LDPC codes]
9. Statistical Buffering of input and output memories to serve maximum iterations while having the decoder parallelization is based on the average number of iterations required.

All these concepts are termed as On-the-fly computation as the core of these concepts are based on minimizing memory and re-computations by employing just in-time scheduling.

Block Independence, Q Vector computation

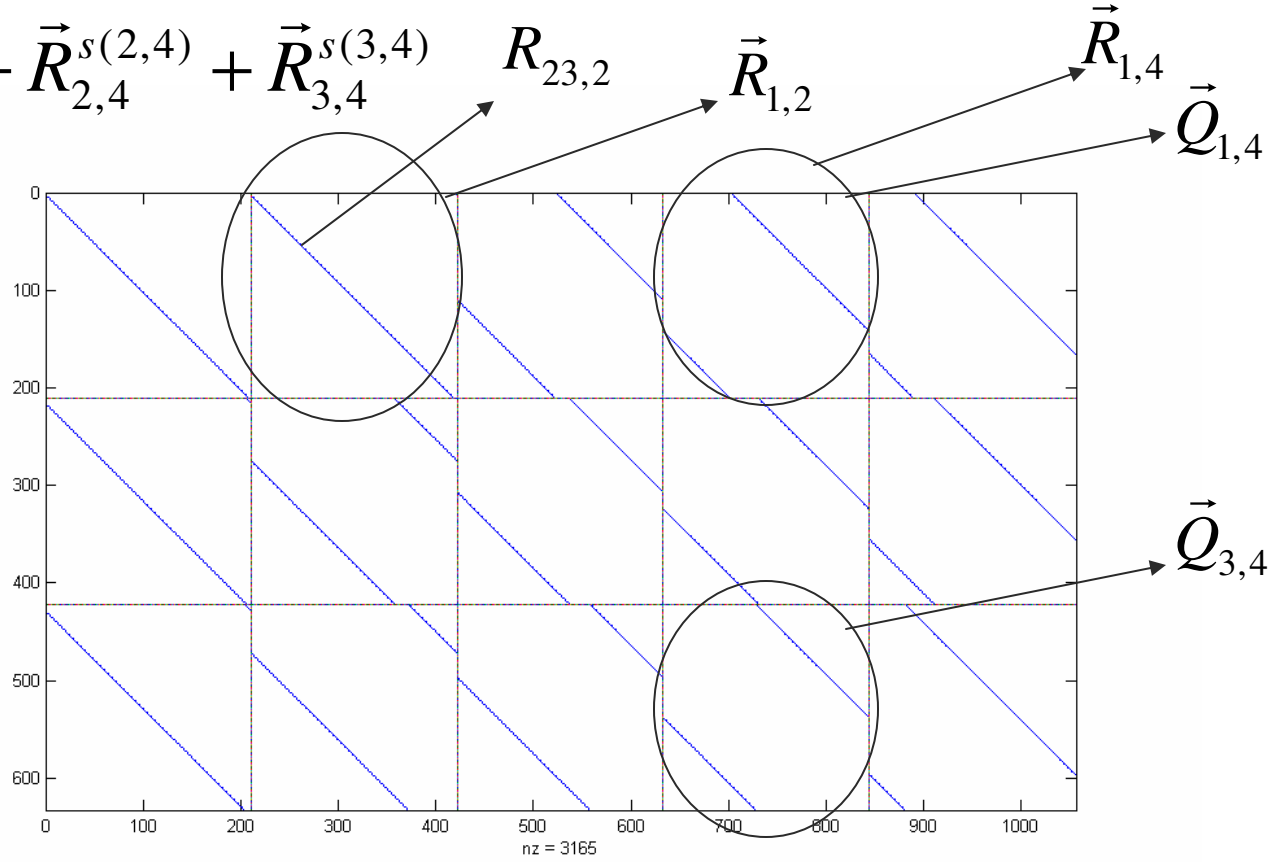
The bit nodes in each block column has support only on the one block column edges of the check nodes.

$$\vec{C}_4 = \vec{R}_{1,4}^{s(1,4)} + \vec{R}_{2,4}^{s(2,4)} + \vec{R}_{3,4}^{s(3,4)}$$

Block Row 1

Block Row 2

Block Row 3

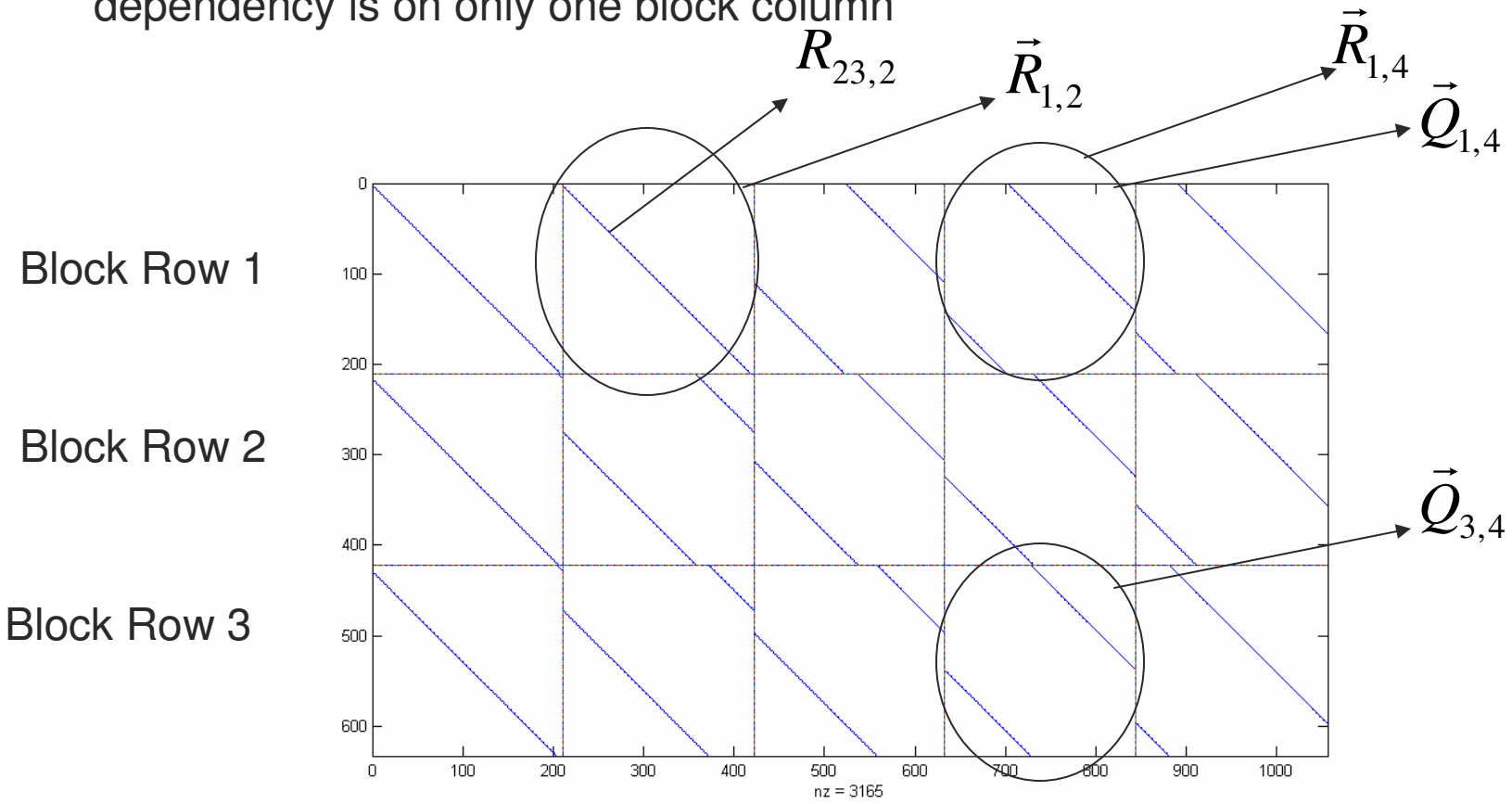


$$\vec{D}_{j,4} = \vec{R}_{j,4}^{s(j,4)}$$

$$\vec{Q}_{j,4} = \vec{C}_k - \vec{D}_{j,4} + \vec{\lambda}_k$$

Block Independence, R Vector computation

The check nodes has support on all block columns. Once the computation is split into accumulate and subtraction stages, the dependency is on only one block column



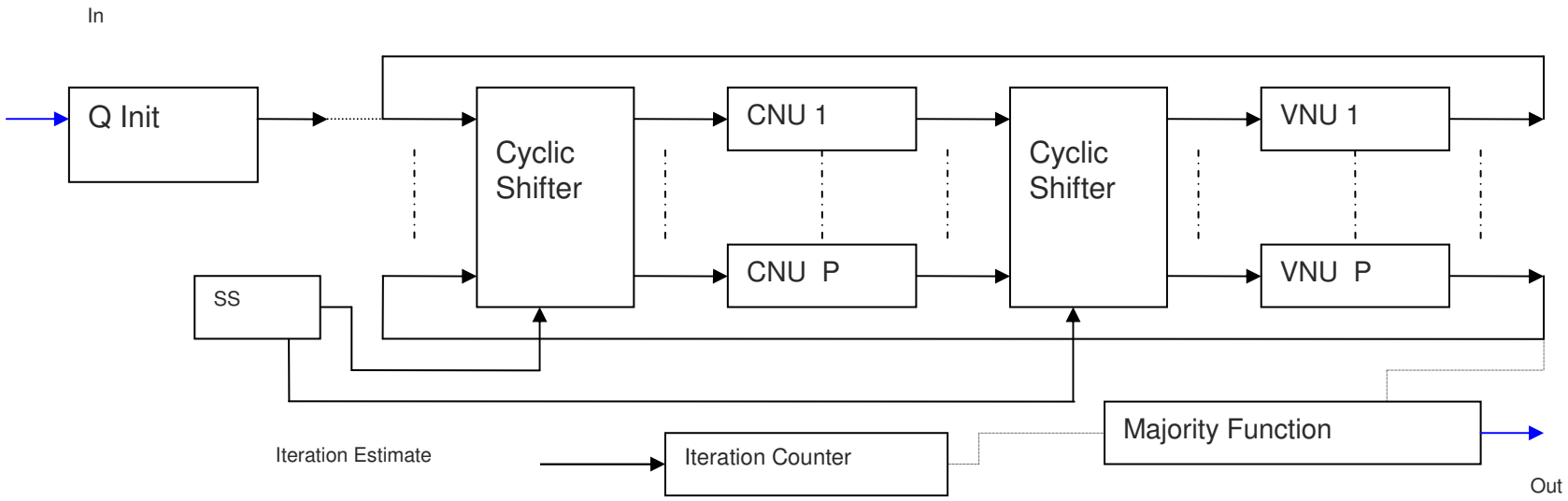
Proposed Scheduling schemes for no Message Passing Memory

- *Doing Computations in Block Column fashion and Block row serial fashion will result in on-the-fly computation*
- *P serial CNU and VNU along with 2 PxP Cyclic shifters [On-the-Fly Type 1].*
- *P serial CNU and VNU along with1 PxP Cyclic shifter [On-the-Fly Type 1]. Layered Decoding*
- *P*c serial CNU and P Parallel VNU: without any routers and any routing complexity [On-the-Fly Type 2].*
- *Transpose: Doing Computations in Block row fashion P parallel CNU and P serial VNU without any routers and any routing complexity. Layered Decoding [On-the-Fly Type 3].*

Outline

- Introduction of LDPC
- Problem Statement
- On-the-fly computation for QC-LDPC
- Multi-rate TPMP Decoder
- Multi-rate Layered Decoder
- Irregular QC-LDPC (IEEE802.16e, IEEE 802.11n,…)
- Fixed code TPMP Decoder
- Parallel Layered Decoder
- Results and performance comparison

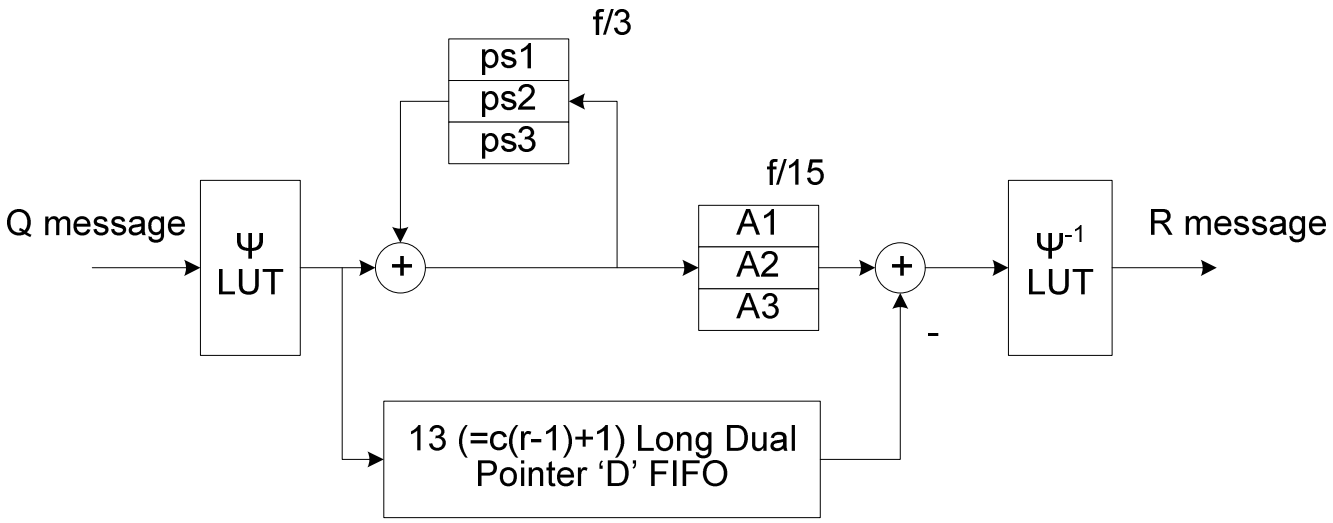
Proposed Architecture for Memory Less Decoding



Decoder for (3, 5) – structured LDPC code of length 1055. No message communication memory is needed. Possible due to structured property and scheduling.

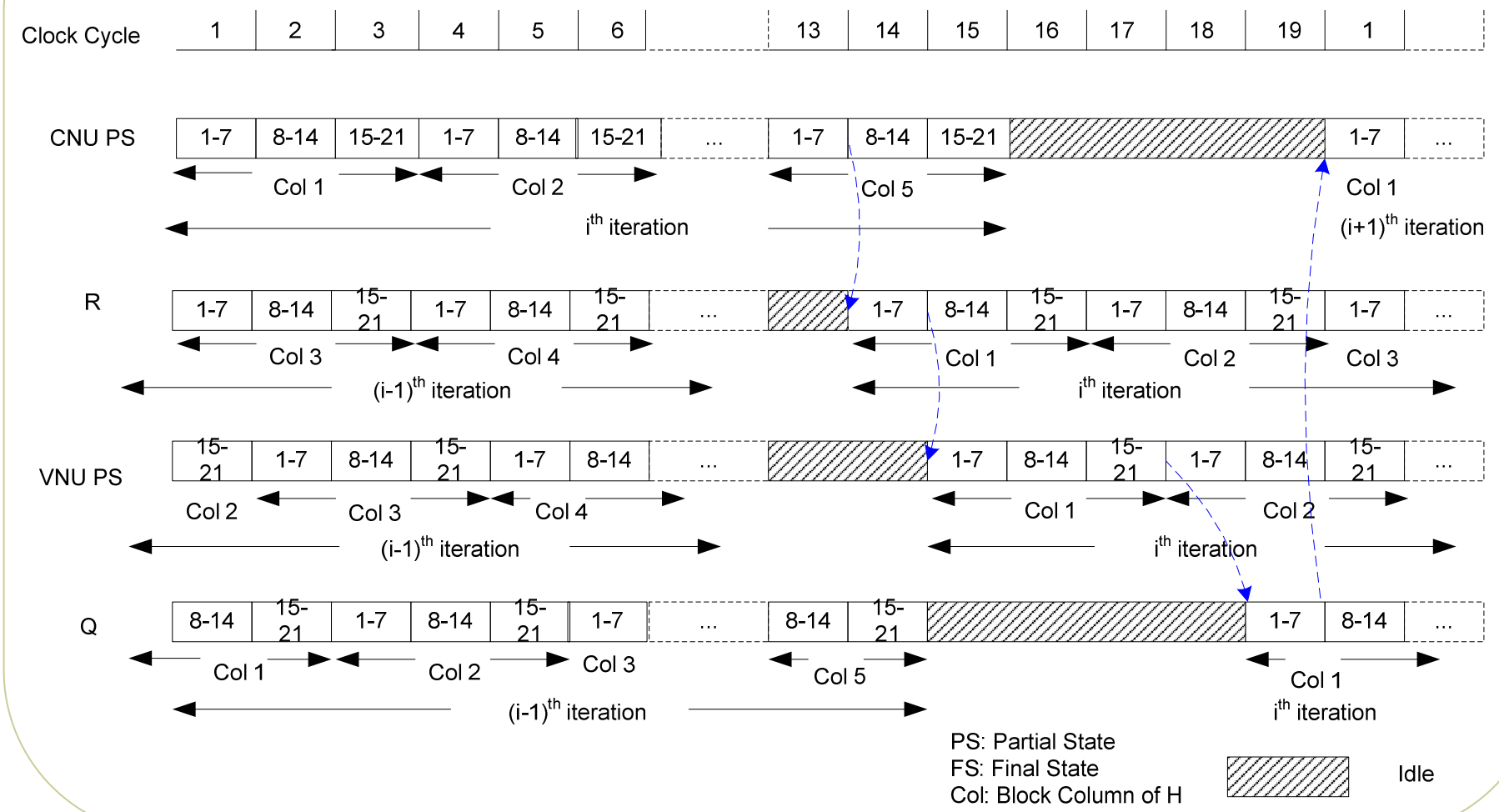
[5] K. Gunnam, G. Choi and M. B. Yeary, "An LDPC Decoding Schedule for Memory Access Reduction", IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)

Micro-Architecture for CNU



The Check Node processing unit computes the partial sum for each block row in a multiplexed fashion to produce the R messages in block column fashion. The registers ps1,ps2 and ps3 correspond to the partial sum for block row 1,2 and 3 respectively.

Pipeline Diagram for (3,5) code and p=7



Check Node update based on Offset Min-Sum

$$R_{cj,bi} = \psi^{-1}\left[\left(\sum_{i^{'}=Row[cj][1]}^{Row[cj][c]} \psi\left(Q_{i^{'},cj}\right)\right) - \psi\left(Q_{bi,cj}\right)\right].\delta(cj,bi)$$

$$R_{cj,bi} = \delta_{cj,bi} \max\left(\beta - \kappa_{cj,bi}\right)$$

$$\kappa_{cj,bi} = \min_{i^{'} \in Row[cj][bi] \setminus i} \left|Q_{i^{'},cj}\right|.$$

- No need of non-linear function, psi.
- Achieves almost the same performance as that of sum of products. The performance degradation is less than 0.1 dB with 5-bit uniform quantization.

[6] Jinghu Chen, Ajay Dholakia, Evangelos Eleftheriou, Marc Fossorier and Xiao-Yu Hu, "Reduced-complexity decoding of LDPC codes," IEEE Transactions on Communications, vol. 53, pp. 1288-1299, Aug. 2005

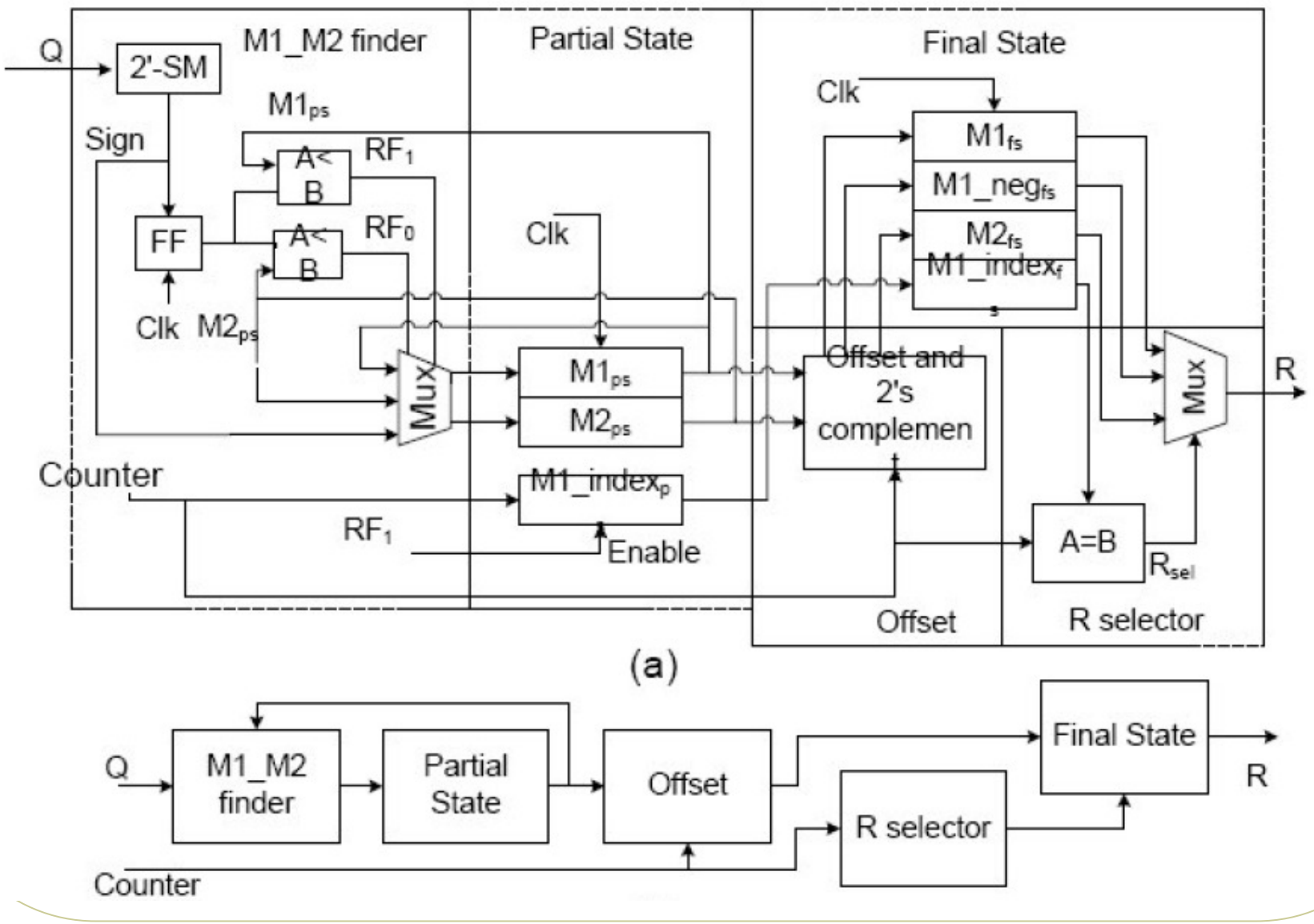
Proposed Micro Architecture for CNU min-sum based on Value-reuse

$$M1_m^{(i)} = \min_{n' \in \mathrm{N}(m)} \left| Q_{mn'}^{(i-1)} \right|. \qquad M2_m^{(i)} = 2nd \min_{n' \in \mathrm{N}(m)} \left| Q_{mn'}^{(i-1)} \right|.$$

$$\left| R_{mn}^{(i)} \right| = M1_m^{(i)}, \ \forall n \in \mathrm{N}(m) \setminus M1\_index$$
$$= M2_m^{(i)}, n = M1\_index$$

- Simplifies the number of comparisons required as well as the memory needed to store CNU outputs.
- The correction has to be applied to only two values instead of distinct values.
- Need to apply 2's complement to only 1 or 2 values instead of values at the output of CNU.

CNU Micro-architecture, OMS



CNU Operation

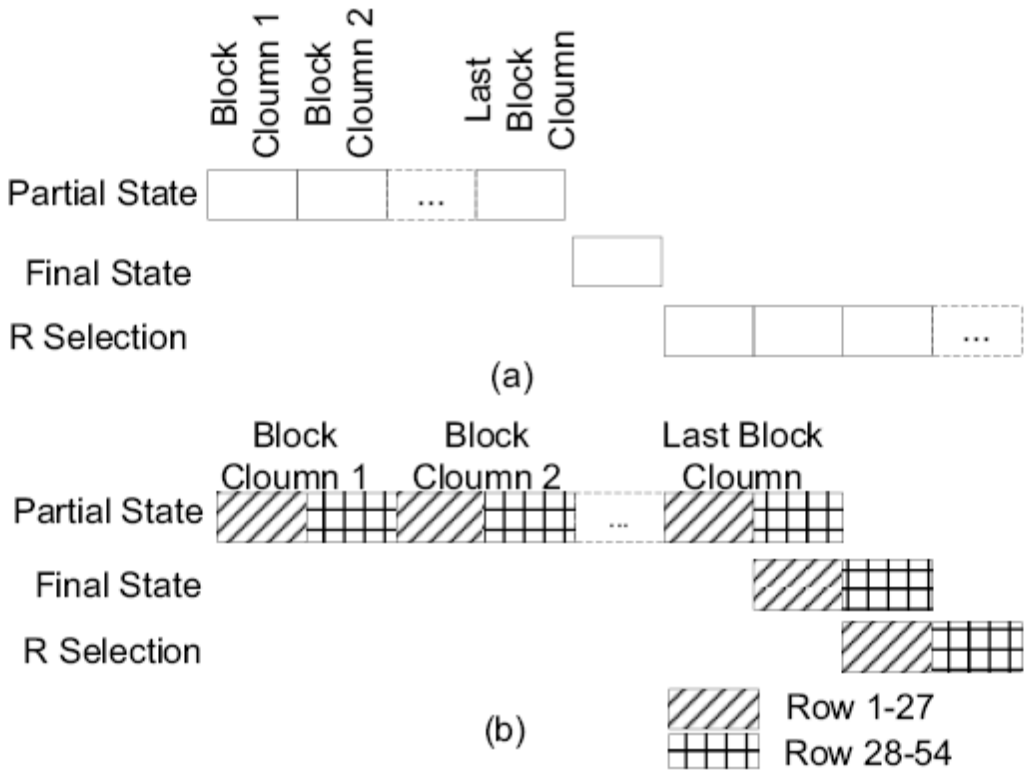


Fig. 3. Operation of CNU (a) no time-division multiplexing (b) time-division multiplexing

Outline

- Introduction of LDPC
- Problem Statement
- On-the-fly computation for QC-LDPC
- Multi-rate TPMP Decoder
- Multi-rate Layered Decoder
- Irregular QC-LDPC (IEEE802.16e, IEEE 802.11n,…)
- Fixed code TPMP Decoder
- Parallel Layered Decoder
- Results and performance comparison

Layered Decoding

- The message passing described before is called as standard message passing or Two Phase Message Passing (TPMP).
- QC-LDPC with j block rows can be viewed as concatenation of j layers or constituent sub-codes. The decoding based on this is called Turbo Decoding Message Passing (TDMP) or layered decoding [3]. This has 2x improvement in convergence.
- Transpose of above is also possible by applying the turbo principle on block columns [6].

[3] Mansour et al "A 640-Mb/s 2048-Bit Programmable LDPC Decoder Chip"-IEEE Journal of Solid-State Circuits, March 2006

[6] Juntan Zhang; Fossorier, M.P.C., "Shuffled iterative decoding," IEEE Transactions on Communications, Volume 53, Issue 2, Feb. 2005

Layered Decoding

$$\vec{R}_{l,n}^{(0)} = 0, l = 1,2,\cdots, j$$ State Initialization

$$\vec{P}_n = \vec{\Lambda}_n$$ Frame Initialization

$$\forall i = 1, 2, \cdots$$ Iteration Loop

$$\forall l = 1, 2, \cdots, j$$ Loop for sub-iterations/layers

$$\forall n = 1, 2, \cdots, k$$ Decoding of a layer

$$\vec{Q}_{l,n}^{(i)} = \vec{P}_n - \left[\vec{R}_{l,n}^{(i-1)}\right]^{S(l,n)}$$

$$\vec{R}_{l,n}^{(i)} = f\left(\left[\vec{Q}_{l,n'}^{(i)}\right]^{-S(l,n)}, \forall n' = 1,2,...,k\right)$$

$$\vec{P}_n = \vec{P}_n - \left[\vec{R}_{l,n}^{(i-1)}\right]^{S(l,n)} + \left[\vec{R}_{l,n}^{(i)}\right]^{S(l,n)} = \vec{Q}_{l,n}^{(i)} + \left[\vec{R}_{l,n}^{(i)}\right]^{S(l,n)}$$

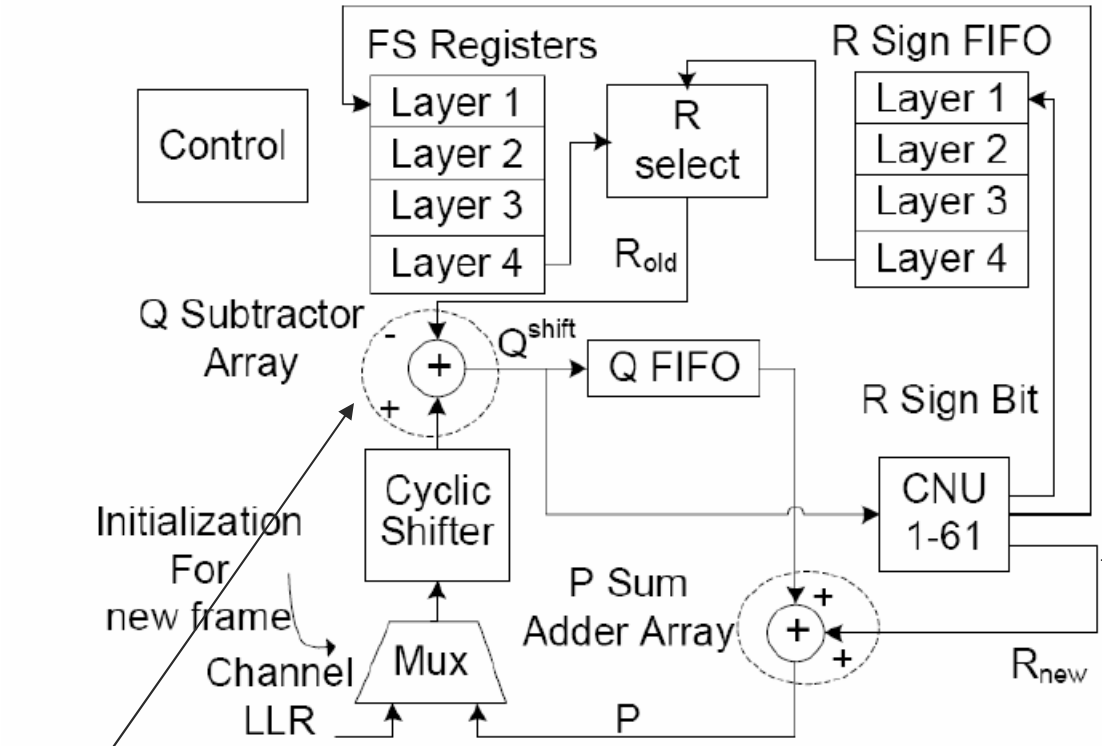
Transformation for Layered Decoding

$$\left[\vec{Q}_{l,n}^{(i)}\right]^{S(l,n)} = \left[\vec{P}_n\right]^{S(l,n)} - \vec{R}_{l,n}^{(i-1)}$$

$$\vec{R}_{l,n}^{(i)} = f\left(\left[\vec{Q}_{l,n'}^{(i)}\right]^{S(l,n')}, \forall n' = 1,2,\cdots,k\right)$$

$$\left[\vec{P}_n\right]^{S(l,n)} = \left[\vec{Q}_{l,n}^{(i)}\right]^{S(l,n)} + \vec{R}_{l,n}^{(i)}$$

New Dataflow Graph for Layered Decoding

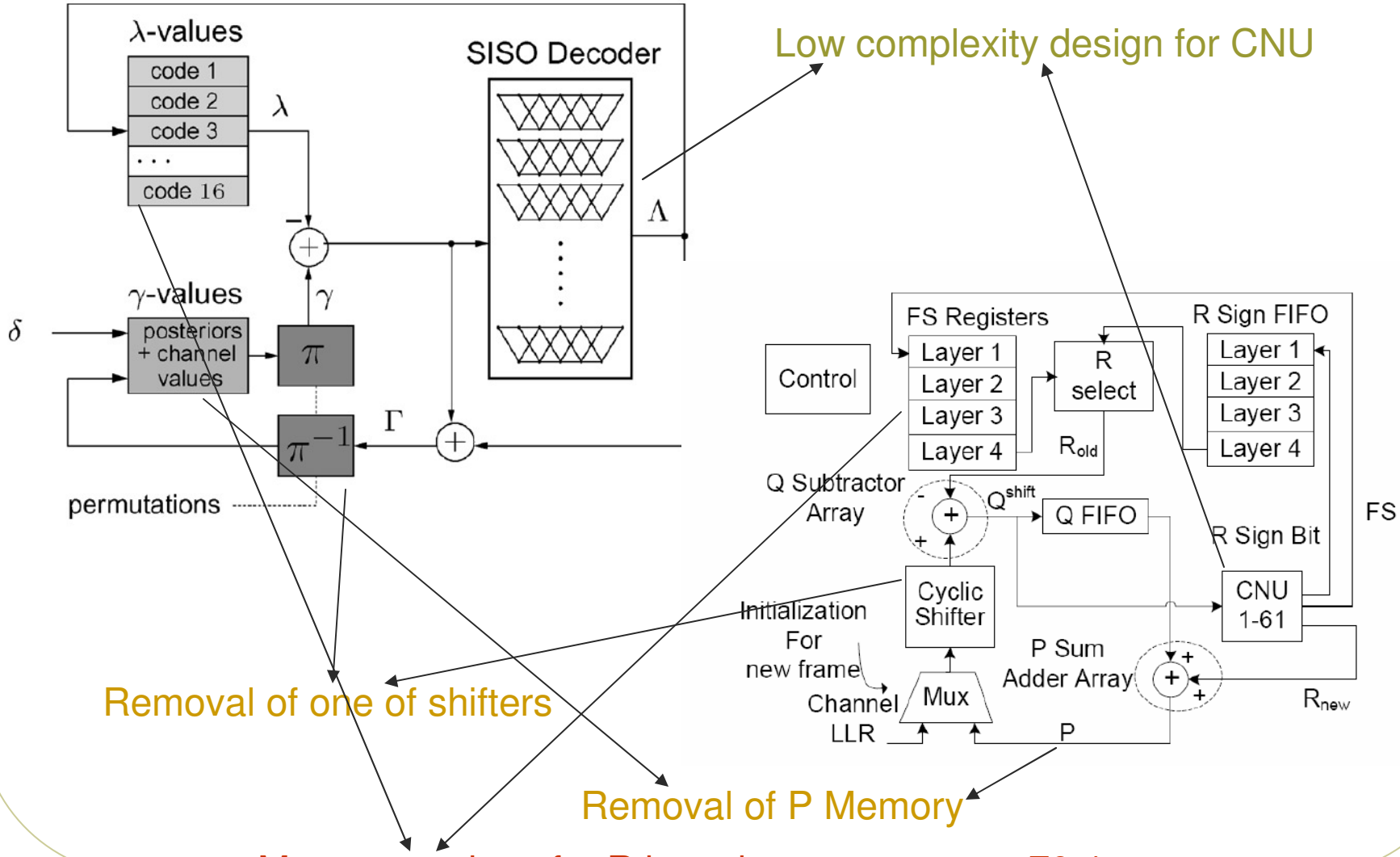


$$\left[\vec{Q}_{l,n}^{(i)}\right]^{S(l,n)} = \left[\vec{P}_n\right]^{S(l,n)} - \vec{R}_{l,n}^{(i-1)}$$

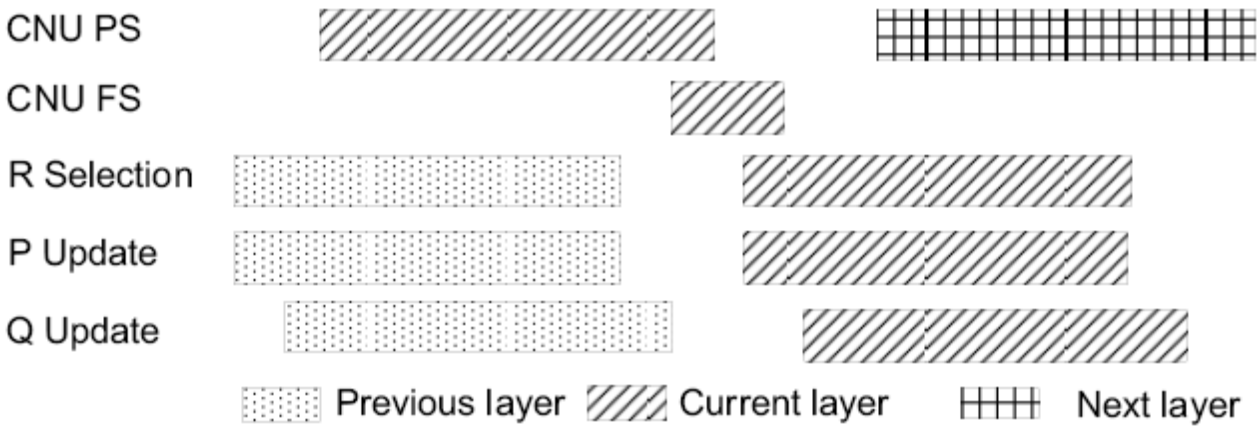
$$\left[\vec{P}_n\right]^{S(l,n)} = \left[\vec{Q}_{l,n}^{(i)}\right]^{S(l,n)} + \vec{R}_{l,n}^{(i)}$$

$$\vec{R}_{l,n}^{(i)} = f\left(\begin{array}{l}\left[\vec{Q}_{l,n'}^{(i)}\right]^{S(l,n')}, \\ \forall n' = 1,2,\cdots,k\end{array}\right)$$

[3] and Proposed



Decoder operation



Outline

- Introduction of LDPC
- Problem Statement
- On-the-fly computation for QC-LDPC
- Multi-rate TPMP Decoder
- Multi-rate Layered Decoder
- Irregular QC-LDPC (IEEE802.16e, IEEE 802.11n,…)
- Fixed code TPMP Decoder
- Parallel Layered Decoder
- Results and performance comparison

Irregular LDPC codes

$$\mathbf{H} = \begin{bmatrix} \mathbf{P}_{0,0} & \mathbf{P}_{0,1} & \mathbf{P}_{0,2} & \cdots & \mathbf{P}_{0,n_b-2} & \mathbf{P}_{0,n_b-1} \\ \mathbf{P}_{1,0} & \mathbf{P}_{1,1} & \mathbf{P}_{1,2} & \cdots & \mathbf{P}_{1,n_b-2} & \mathbf{P}_{1,n_b-1} \\ \mathbf{P}_{2,0} & \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \cdots & \mathbf{P}_{2,n_b-2} & \mathbf{P}_{0,n_b-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{P}_{m_b-1,0} & \mathbf{P}_{m_b-1,1} & \mathbf{P}_{m_b-1,2} & \cdots & \mathbf{P}_{m_b-1,n_b-2} & \mathbf{P}_{m_b-1,n_b-1} \end{bmatrix} = \mathbf{P}^{H_b}$$

Different base matrices to support different rates.

Different expansion factors (z) to support multiple lengths.

All the shift coefficients for different codes for a given rate are obtained from the same base matrix using modulo arithmetic.

Irregular LDPC codes

Rate 1/2:

-1	94	73	-1	-1	-1	-1	-1	55	83	-1	-1	7	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	27	-1	-1	-1	22	79	9	-1	-1	-1	12	-1	0	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	24	22	81	-1	33	-1	-1	-1	0	-1	-1	0	0	-1	-1	-1	-1	-1	-1	-1	-1
61	-1	47	-1	-1	-1	-1	-1	65	25	-1	-1	-1	-1	-1	0	0	-1	-1	-1	-1	-1	-1	-1
-1	-1	39	-1	-1	-1	84	-1	-1	41	72	-1	-1	-1	-1	-1	0	0	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	46	40	-1	82	-1	-1	-1	79	0	-1	-1	-1	-1	0	0	-1	-1	-1	-1	-1
-1	-1	95	53	-1	-1	-1	-1	-1	14	18	-1	-1	-1	-1	-1	-1	-1	0	0	-1	-1	-1	-1
-1	11	73	-1	-1	-1	2	-1	-1	47	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	-1	-1	-1
12	-1	-1	-1	83	24	-1	43	-1	-1	-1	51	-1	-1	-1	-1	-1	-1	-1	-1	0	0	-1	-1
-1	-1	-1	-1	-1	94	-1	59	-1	-1	70	72	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0	-1
-1	-1	7	65	-1	-1	-1	-1	39	49	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0	0
43	-1	-1	-1	-1	66	-1	41	-1	-1	-1	26	7	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	0

Rate 2/3 A code:

3	0	-1	-1	2	0	-1	3	7	-1	1	1	-1	-1	-1	-1	1	0	-1	-1	-1	-1	-1	-1
-1	-1	1	-1	36	-1	-1	34	10	-1	-1	18	2	-1	3	0	-1	0	0	-1	-1	-1	-1	-1
-1	-1	12	2	-1	15	-1	40	-1	3	-1	15	-1	2	13	-1	-1	-1	0	0	-1	-1	-1	-1
-1	-1	19	24	-1	3	0	-1	6	-1	17	-1	-1	-1	8	39	-1	-1	-1	0	0	-1	-1	-1
20	-1	6	-1	-1	10	29	-1	-1	28	-1	14	-1	38	-1	-1	0	-1	-1	-1	0	0	-1	-1
-1	-1	10	-1	28	20	-1	-1	8	-1	36	-1	9	-1	21	45	-1	-1	-1	-1	-1	0	0	-1
35	25	-1	37	-1	21	-1	-1	5	-1	-1	0	-1	4	20	-1	-1	-1	-1	-1	-1	-1	0	0
-1	6	6	-1	-1	-1	4	-1	14	30	-1	3	36	-1	14	-1	1	-1	-1	-1	-1	-1	-1	0

Irregular LDPC codes

Rate 3/4 A code:

6	38	3	93	-1	-1	-1	30	70	-1	86	-1	37	38	4	11	-1	46	48	0	-1	-1	-1	-1
62	94	19	84	-1	92	78	-1	15	-1	-1	92	-1	45	24	32	30	-1	-1	0	0	-1	-1	-1
71	-1	55	-1	12	66	45	79	-1	78	-1	-1	10	-1	22	55	70	82	-1	-1	0	0	-1	-1
38	61	-1	66	9	73	47	64	-1	39	61	43	-1	-1	-1	-1	95	32	0	-1	-1	0	0	-1
-1	-1	-1	-1	32	52	55	80	95	22	6	51	24	90	44	20	-1	-1	-1	-1	-1	-1	0	0
-1	63	31	88	20	-1	-1	-1	6	40	56	16	71	53	-1	-1	27	26	48	-1	-1	-1	-1	0

Rate 5/6 code:

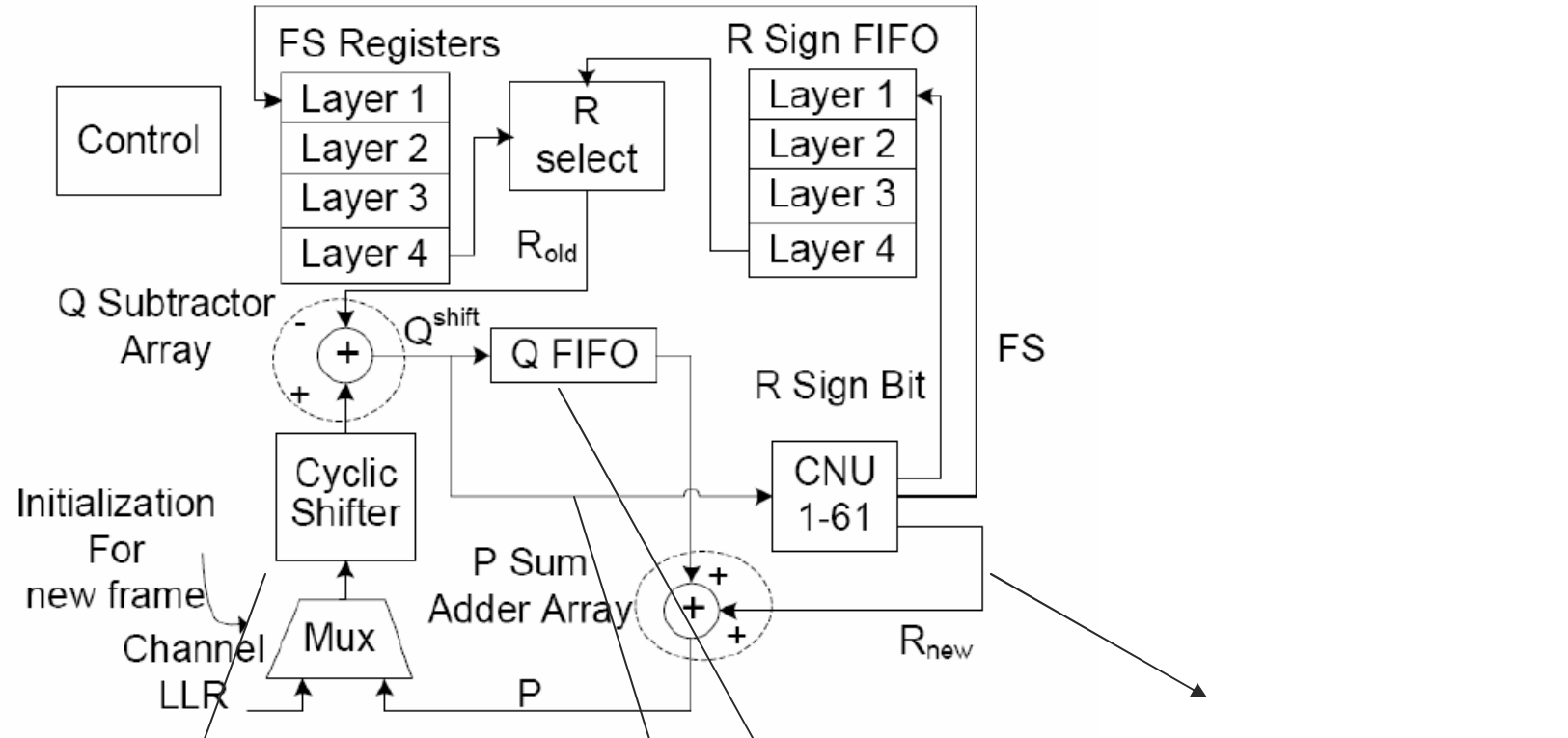
1	25	55	-1	47	4	-1	91	84	8	86	52	82	33	5	0	36	20	4	77	80	0	-1	-1
-1	6	-1	36	40	47	12	79	47	-1	41	21	12	71	14	72	0	44	49	0	0	0	0	-1
51	81	83	4	67	-1	21	-1	31	24	91	61	81	9	86	78	60	88	67	15	-1	-1	0	0
50	-1	50	15	-1	36	13	10	11	20	53	90	29	92	57	30	84	92	11	66	80	-1	-1	0

Irregular LDPC codes

- Existing implementations [4] show that these are more complex to implement.
- However these codes have the better BER performance and selected for IEEE 802.16e and IEEE 802.11n.
- It is anticipated that these codes will be the default choice for most of the standards.
- My research shows that with out-of-order processing and scheduling of layered processing, it is possible to design very efficient architectures based on a simple extension of the proposed layered decoder architecture for array LDPC codes!!

[7] Hocevar, D.E., "A reduced complexity decoder architecture via layered decoding of LDPC codes," IEEE Workshop on Signal Processing Systems, 2004. SIPS 2004. .pp. 107- 112, 13-15 Oct. 2004

Changes needed in Decoder architecture for regular QC-LDPC to accommodate irregular QC-LDPC codes



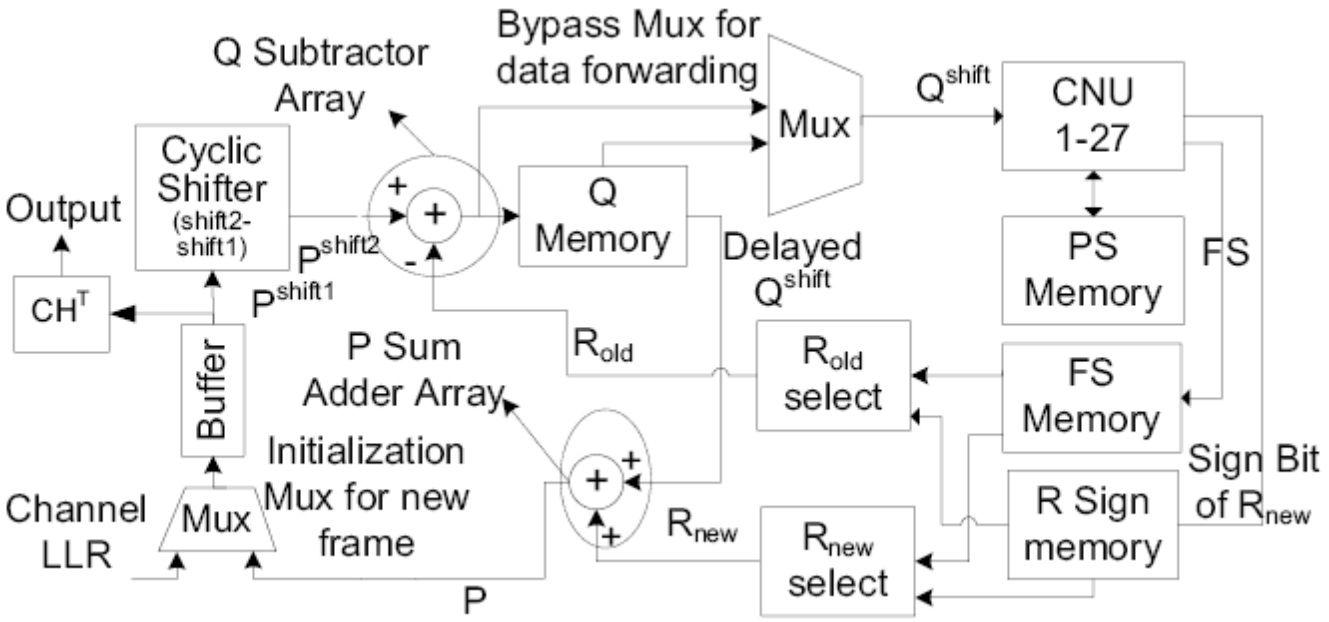
Rnew is generated out-of-order

Q FIFO becomes Q Memory

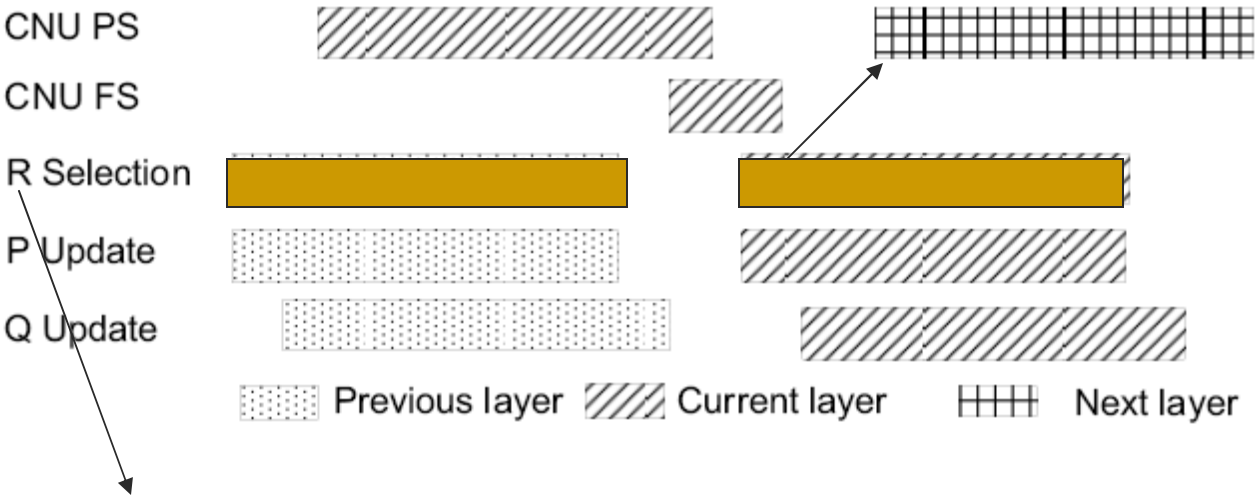
Need re-configurable shifter

Need data forwarding

Decoder for Irregular codes



Pipeline for Irregular codes



R selection for R_{new} operates out-of-order to feed the data for PS processing of next layer

Out-of-order processing for R Selection

Rate 2/3 A code:

3	0	-1	-1	2	0	-1	3	7	-1	1	1	-1	-1	-1	-1	1	0	-1	-1	-1	-1	-1	-1
-1	-1	1	-1	36	-1	-1	34	10	-1	-1	18	2	-1	3	0	-1	0	0	-1	-1	-1	-1	-1
-1	-1	12	2	-1	15	-1	40	-1	3	-1	15	-1	2	13	-1	-1	-1	0	0	-1	-1	-1	-1
-1	-1	19	24	-1	3	0	-1	6	-1	17	-1	-1	-1	8	39	-1	-1	-1	0	0	-1	-1	-1
20	-1	6	-1	-1	10	29	-1	-1	28	-1	14	-1	38	-1	-1	0	-1	-1	-1	0	0	-1	-1
-1	-1	10	-1	28	20	-1	-1	8	-1	36	-1	9	-1	21	45	-1	-1	-1	-1	-1	0	0	-1
35	25	-1	37	-1	21	-1	-1	5	-1	-1	0	-1	4	20	-1	-1	-1	-1	-1	-1	-1	0	0
-1	6	6	-1	-1	-1	4	-1	14	30	-1	3	36	-1	14	-1	1	-1	-1	-1	-1	-1	-1	0

PS processing R selection

R selection is out-of-order so that it can feed the data required for the PS processing of the second layer

In addition, for some of the base matrices, scheduling of layered processing is employed to avoid more than 3 accesses to one bank of Q memory

Out-of-order block processing for Partial State

Rate 2/3 A code:

3	0	-1	-1	2	0	-1	3	7	-1	1	1	-1	-1	-1	-1	1	0	-1	-1	-1	-1	-1	-1
-1	-1	1	-1	36	-1	-1	34	10	-1	-1	18	2	-1	3	0	-1	0	0	-1	-1	-1	-1	-1
-1	-1	12	2	-1	15	-1	40	-1	3	-1	15	-1	2	13	-1	-1	-1	0	0	-1	-1	-1	-1
-1	-1	19	24	-1	3	0	-1	6	-1	17	-1	-1	-1	8	39	-1	-1	-1	0	0	-1	-1	-1
20	-1	6	-1	-1	10	29	-1	-1	28	-1	14	-1	38	-1	-1	0	-1	-1	-1	0	0	-1	-1
-1	-1	10	-1	28	20	-1	-1	8	-1	36	-1	9	-1	21	45	-1	-1	-1	-1	-1	0	0	-1
35	25	-1	37	-1	21	-1	-1	5	-1	-1	0	-1	4	20	-1	-1	-1	-1	-1	-1	-1	0	0
-1	6	6	-1	-1	-1	4	-1	14	30	-1	3	36	-1	14	-1	1	-1	-1	-1	-1	-1	-1	0

PS processing

R selection

Re-ordering of block processing . While processing the layer 2,

the blocks which depend on layer 1 will be processed last to allow for the pipeline latency.

In the above example, the pipeline latency can be 5.

The vector pipeline depth is 5.so no stall cycles are needed while processing the layer 2 due to the pipelining. [In other implementations, the stall cycles are introduced – which will effectively reduce the throughput by a huge margin.]

Also we will sequence the operations in layer such that we process the block first that has dependent data available for the longest time.

This naturally leads us to true out-of-order processing across several layers. In practice we wont do out-of-order partial state processing involving more than 2 layers.

Scheduling of Layered Processing or Re-ordering of Layers

- Normal practice: Do the decoding of H matrix in the original layer order.
- It is possible to do the decoding using a different sequence of layers instead of processing the layers from 1 to j which is typically used to increase the parallelism such that it is possible to process two block rows simultaneously as indicated in IEEE Std.
- We propose to use the concept of re-ordering of layers for increased parallelism as well as for low complexity memory implementation and also for inserting additional pipeline stages without incurring overhead.
- There are j factorial layer sequences possible.

Scheduling of Layered Processing or Re-ordering of Layers

- One simple way to generate valid layer sequences is to apply a modulo step to get a new layer sequence and check whether all the layers are present in the modified sequence. The idea behind this is if we choose the layers with a modulo step, the adjacent layers have minimum dependencies.

Original layer sequence: 1 2 3 4 5 6 7 8 9 10 11 12
Modulo layer sequence with step size of 1: (all are unique)
1 2 3 4 5 6 7 8 9 10 11 12 (same as original sequence)
Modulo layer sequence with step size of 2: (all are unique)
1 3 5 7 9 11 2 4 6 8 10 12
Modulo layer sequence with step size of 3: (not a valid sequence)
1 4 7 10 1 4 7 10 1 4 7 10
Modulo layer sequence with step size of 4: (not a valid sequence)
1 5 9 1 5 9 1 5 9 1 5 9
Modulo layer sequence with step size of 5: (all are unique .so valid sequence!!)
1 6 11 4 9 2 7 12 5 10 3 8
And so on..we have 11 sequences to test and we will take a subset of them which are valid sequences.

Overall Schedule Optimization

- The decoder hardware architecture is proposed to support out-of-order processing to remove pipeline and memory accesses or to satisfy any other performance or hardware constraint. Remaining hardware architectures won't support out-of-order processing without further involving more logic and memory.
- For the above hardware decoder architecture, the optimization of decoder schedule belongs to the class of NP-complete problems. So there are several classic optimization algorithms such as dynamic programming that can be applied. We apply the following classic approach of optimal substructure.
- Step 1: We will try different layer schedules(j! i.e j factorial of j if there are j layers). For simplicity, we will try only a subset of possible sequences so as to have more spread between the original layers. (slide 43-44)
- Step 2:Given a layer schedule or a re-ordered H matrix, we will optimize the processing schedule of each layer. For this, we use the classic approach of optimal substructure i.e. the solution to a given optimization problem can be obtained by the combination of optimal solutions to its subproblems. So first we optimize the processing order to minimize the pipeline conflicts. Then we optimize the resulting processing order to minimize the memory conflicts. So for each layer schedule, we are measuring the number of stall cycles (our cost function).(slides 40-42)
- Step 3: We choose a layer schedule which minimizes the cost function i.e. meets the requirements with less stall cycles due to pipeline conflicts and memory conflicts and also minimizes the memory accesses (such as FS memory accesses to minimize the number of ports needed and to save the access power and to minimize the more muxing requirement and any interface memory access requirements.

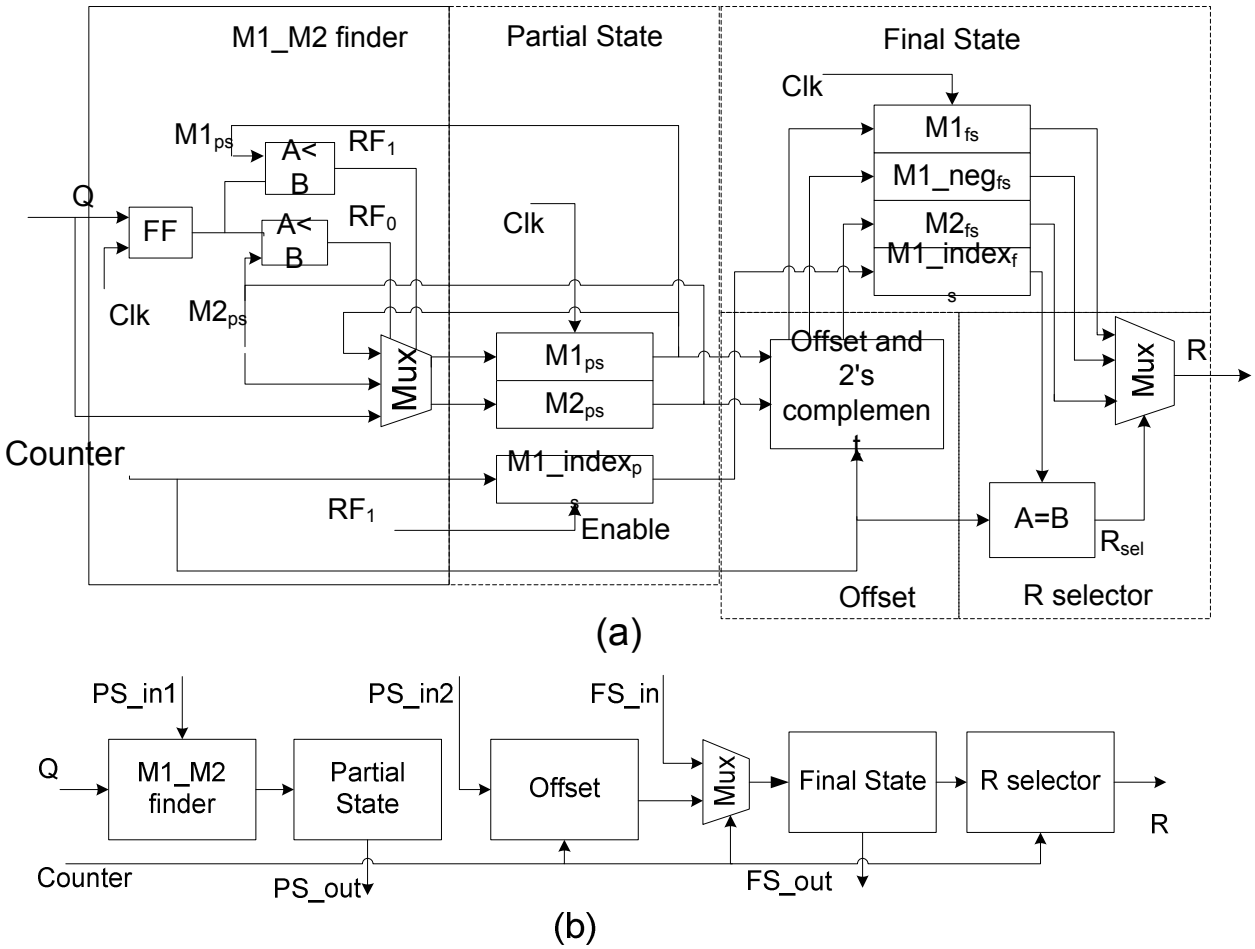
Outline

- Introduction of LDPC
- Problem Statement
- On-the-fly computation for QC-LDPC
- Multi-rate TPMP Decoder
- Multi-rate Layered Decoder
- Irregular QC-LDPC (IEEE802.16e, IEEE 802.11n,…)
- Fixed code TPMP Decoder
- Parallel Layered Decoder
- Results and performance comparison

High Throughput Fixed Code Architectures

- Some applications need very high throughput.
- Flexibility is not a factor.
- Magnetic recording channels is an example application
- Two new architectures based on regular array LDPC codes
- Block Serial TPMP Decoder
- Parallel Layered Decoder

CNU Micro Architecture for min-sum



New micro architecture a) Static unit b) Dynamic Unit

Static unit- assigned to one row in each block row

Dynamic unit- does the partial computations for *r (*check node degree*)* rows to achieve shifts

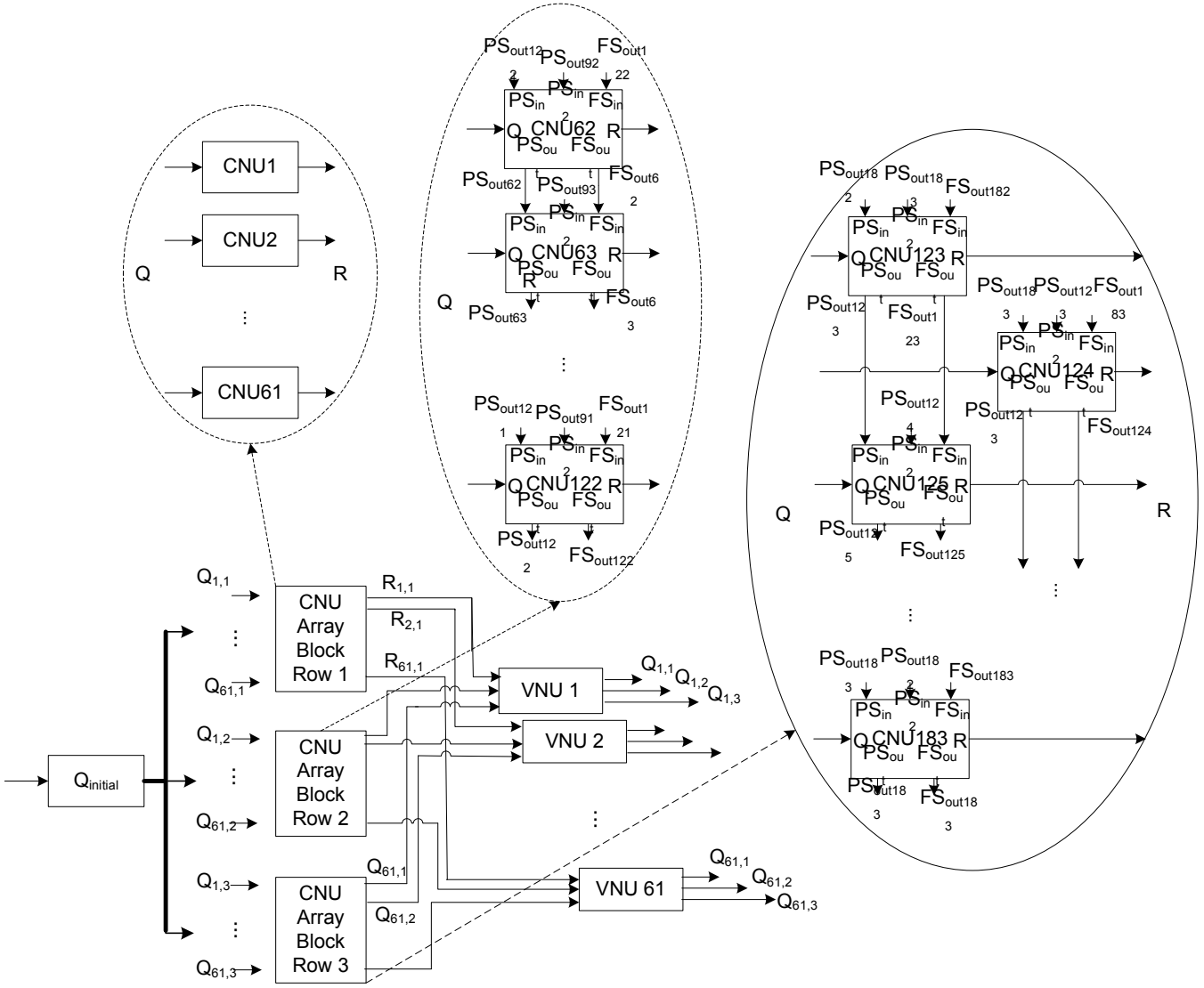
Fixed Code TPMP architecture

$$H = \begin{bmatrix} I & I & I & \dots & I \\ I & \sigma & \sigma^2 & \dots & \sigma^{r-1} \\ I & \sigma^2 & \sigma^4 & \dots & \sigma^{2(r-1)} \\ : & & & & \\ I & \sigma^{c-1} & \sigma^{(c-1)2} & \dots & \sigma^{(c-1)(r-1)} \end{bmatrix}$$

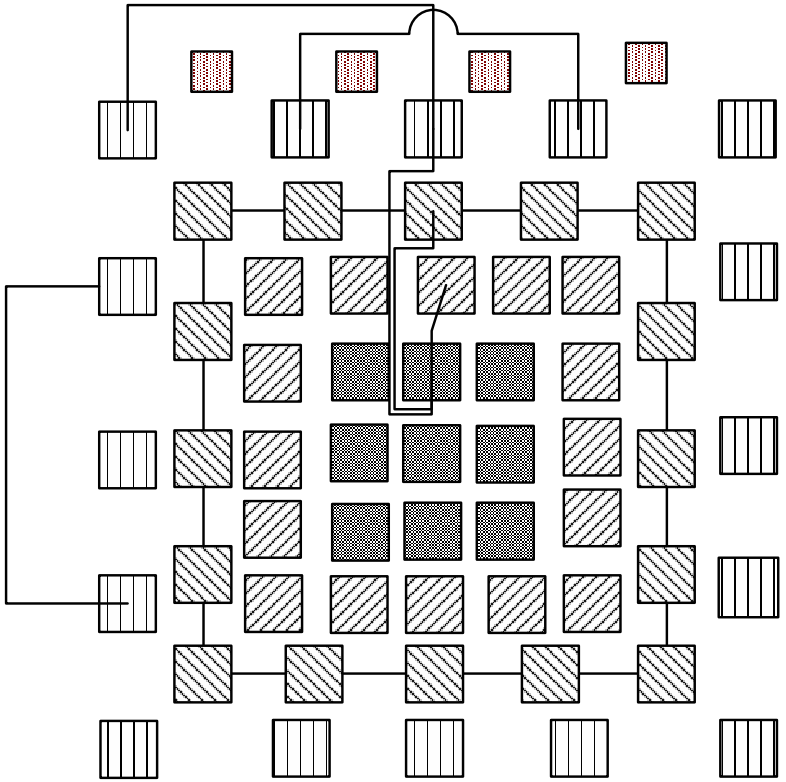
Constant shift across block rows.

Use of the dynamic unit concept.

Fixed Code TPMP architecture



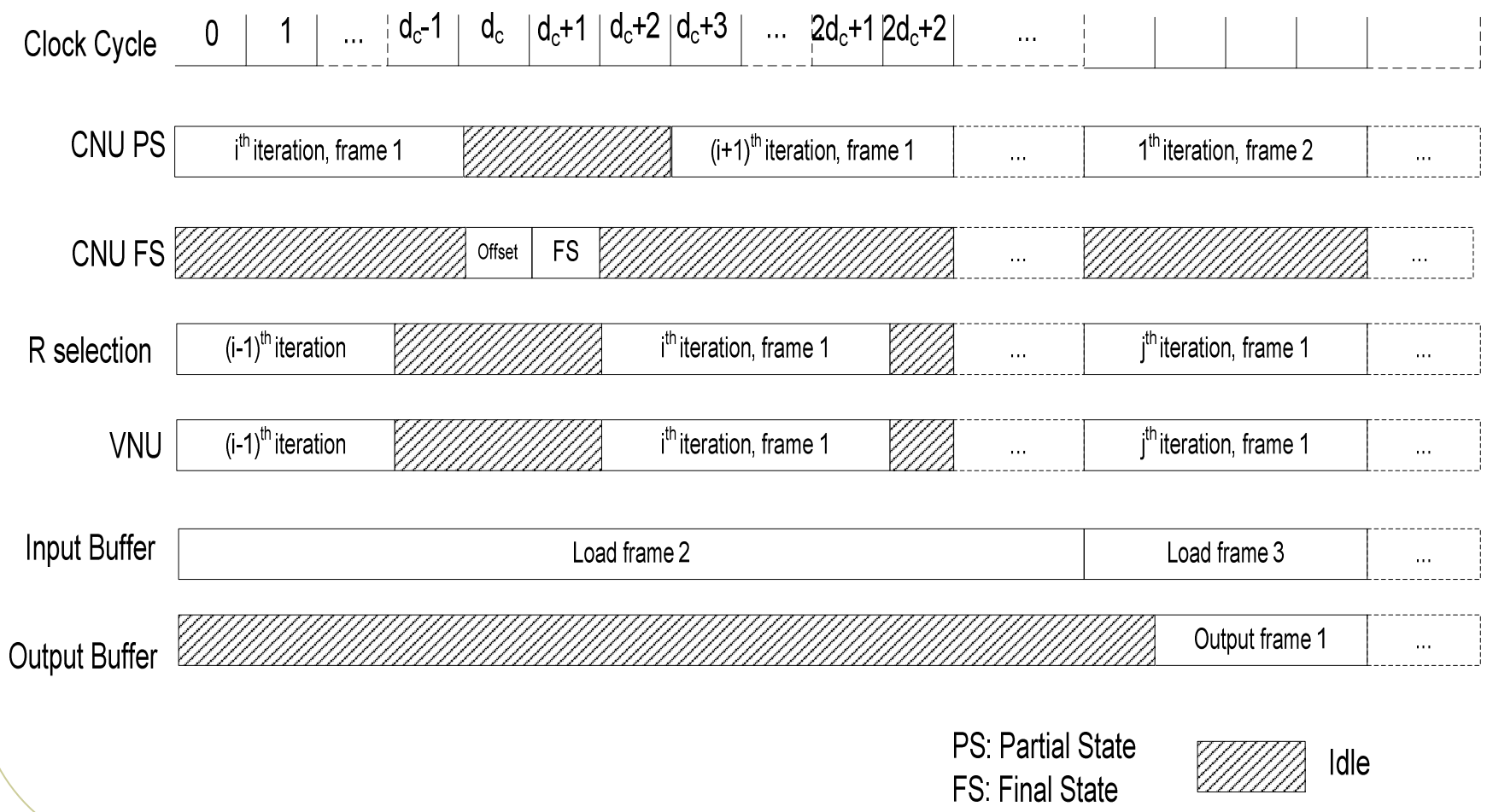
ASIC Floor planning



Input Memory
(split into banks)
VNU
CNU Block Row 1
CNU Block Row 2
CNU Block Row 3

Not all the connections shown

Pipeline


Clock Cycle
0
1
...
d_c-1
d_c
d_c+1
d_c+2
d_c+3
...
2d_c+1
2d_c+2
...
CNU PS
i^{th} iteration, frame 1
$(i+1)^{th}$ iteration, frame 1
...
1^{th} iteration, frame 2
...
CNU FS
Offset
FS
...
...
R selection
$(i-1)^{th}$ iteration
i^{th} iteration, frame 1
...
j^{th} iteration, frame 1
...
VNU
$(i-1)^{th}$ iteration
i^{th} iteration, frame 1
...
j^{th} iteration, frame 1
...
Input Buffer
Load frame 2
Load frame 3
...
Output Buffer
Output frame 1
...
PS: Partial State
FS: Final State
Idle

Outline

- Introduction of LDPC
- Problem Statement
- On-the-fly computation for QC-LDPC
- Multi-rate TPMP Decoder
- Multi-rate Layered Decoder
- Irregular QC-LDPC (IEEE802.16e, IEEE 802.11n,…)
- Fixed code TPMP Decoder
- Parallel Layered Decoder
- Results and performance comparison

Fixed Code TDMP architecture

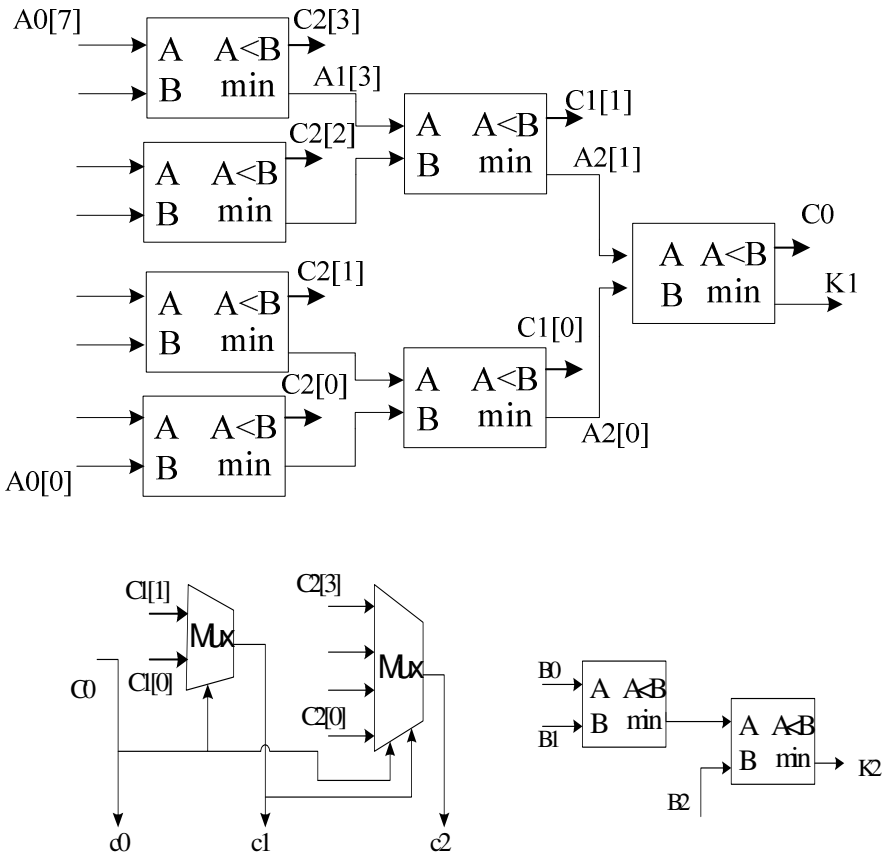
$$H = \begin{bmatrix} I & I & I & \dots & I \\ I & \sigma & \sigma^2 & \dots & \sigma^{r-1} \\ I & \sigma^2 & \sigma^4 & \dots & \sigma^{2(r-1)} \\ : & & & & \\ I & \sigma^{c-1} & \sigma^{(c-1)2} & \dots & \sigma^{(c-1)(r-1)} \end{bmatrix}$$

Constant shift across block columns.

Due to the layered scheduling, only two shifts are needed.

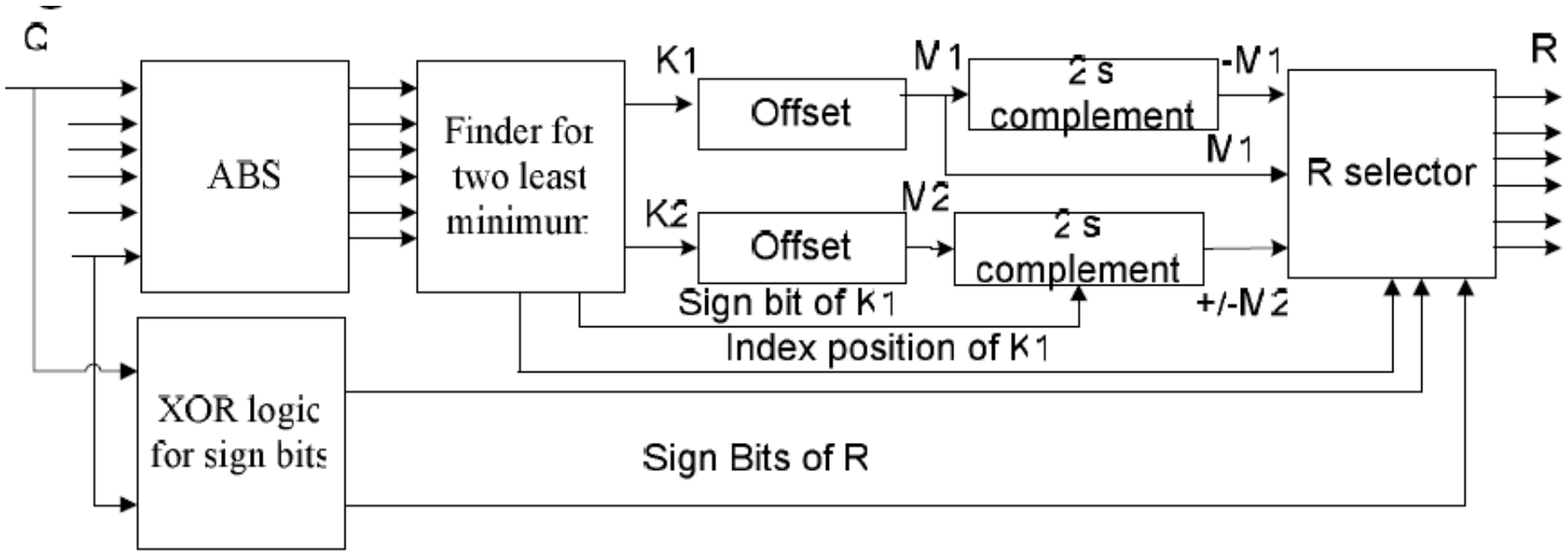
These are realized using lay-out aware design.

Low complexity Parallel CNU

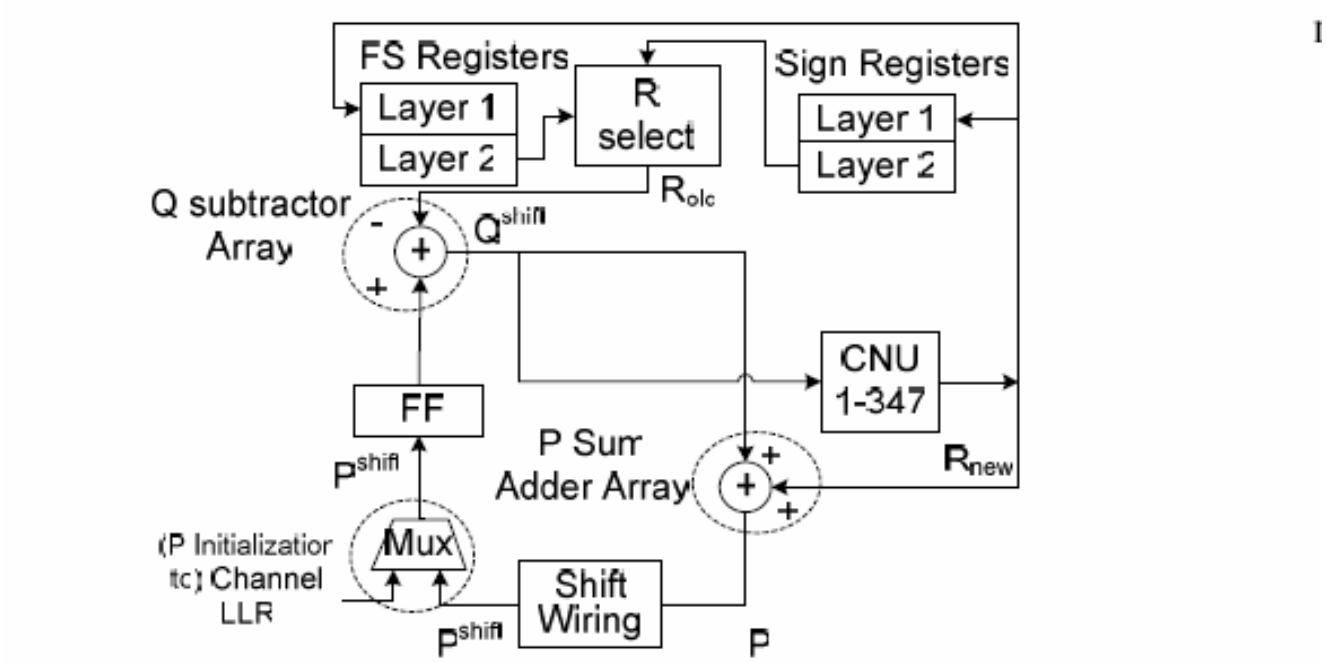


Finder for the two least minimum in CNU (a) Binary tree to find the least minimum. (b) Trace-back multiplexers and comparators on survivors to find the second minimum. Multiplexers for selecting survivors are not shown.

Low complexity Parallel CNU



Decoder architecture



Same as the Multi-rate layered decoder data flow graph.

Parallel CNU instead of serial CNU

Since only 2 shifts are needed, wiring and one layer of multiplexers to accomplish the required 2 shifts.

Decoder Floor planning

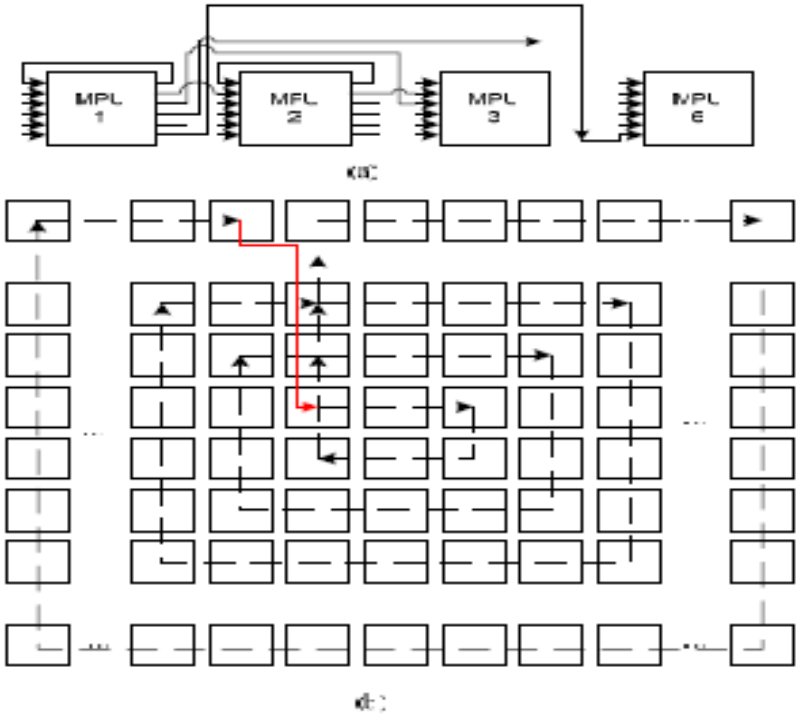
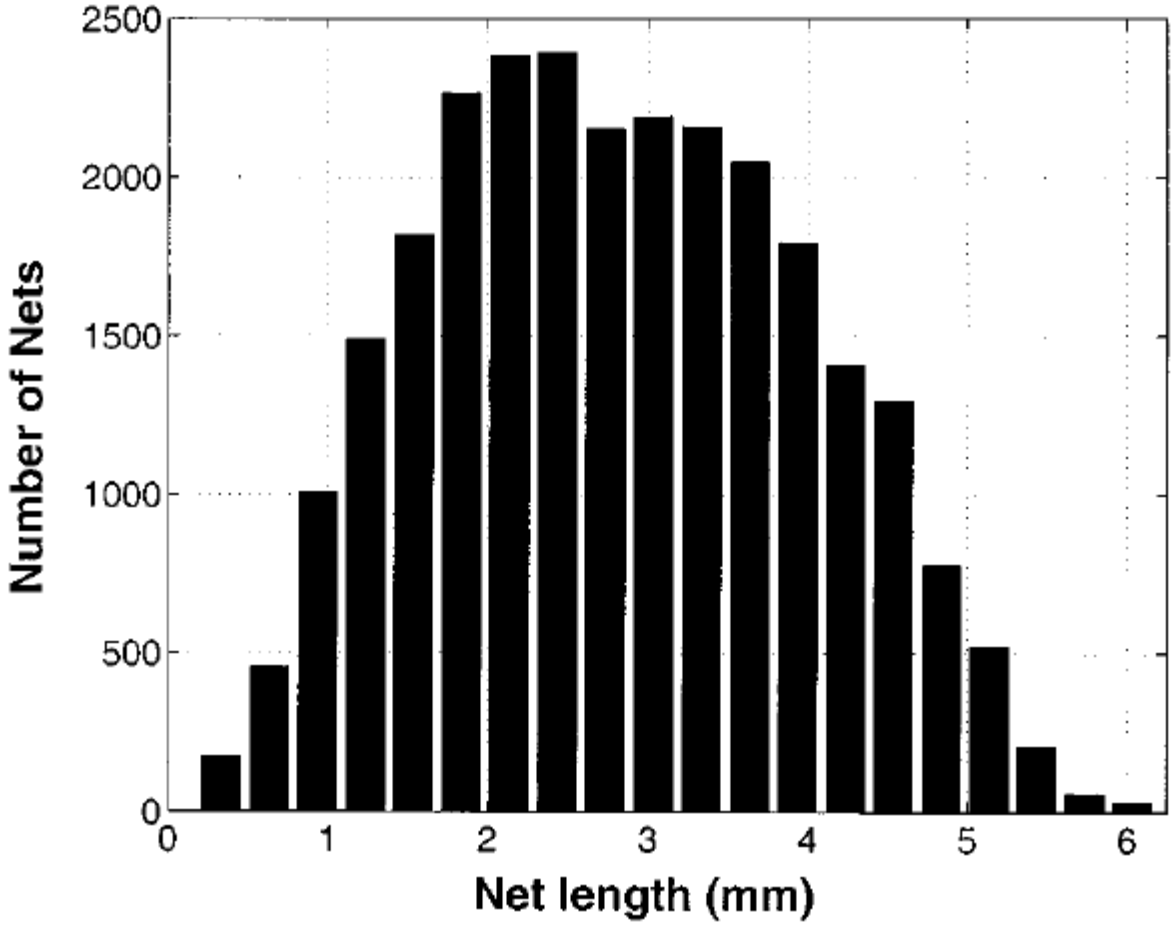


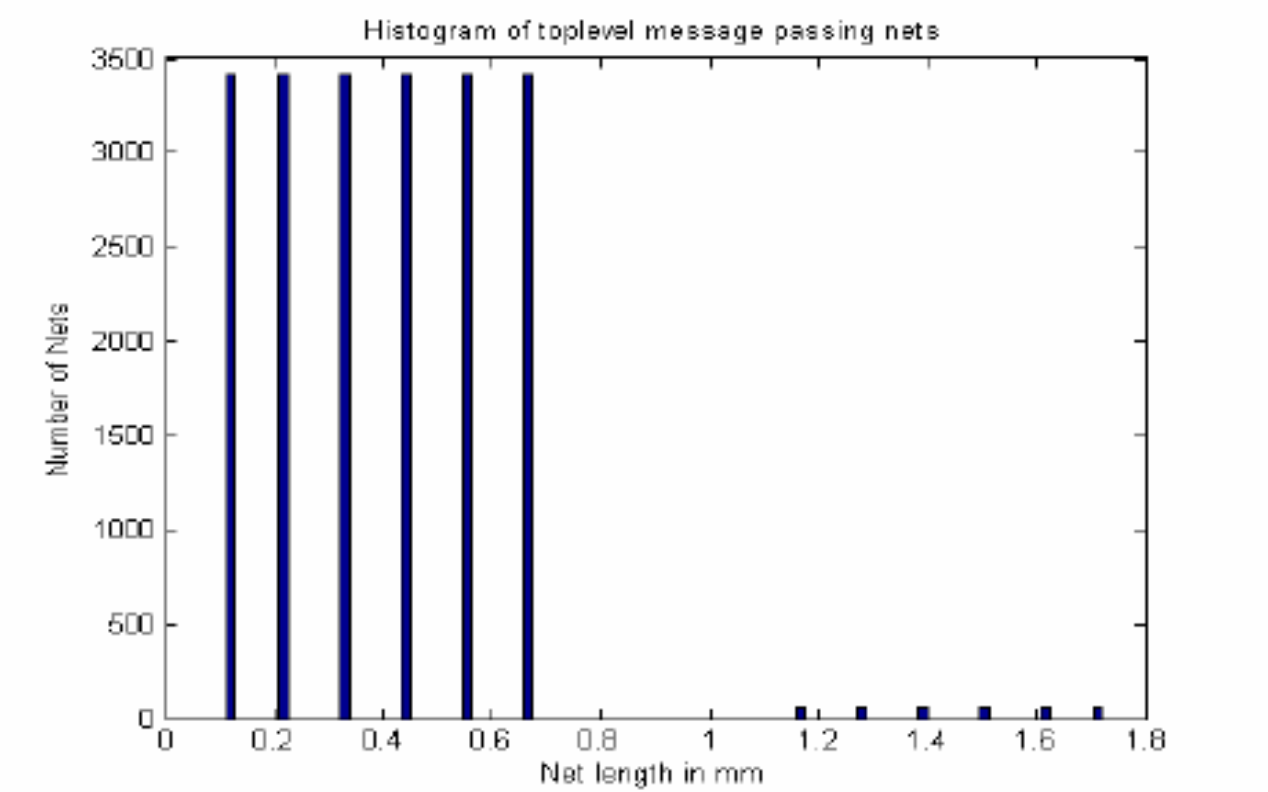
Figure 4: a)Illustration of connections between Message Processing units to achieve cyclic down shift of (n-1) on each block column n b) Concentric layout to accommodate 347 message processing units. Connections for cyclic up shift of 2n(=(j-1).n) are not shown.

1024-bit Parallel decoder [8] wiring



[8] Blanksby, A.J.; Howland ,C.J, A 690-mW 1-Gb/s 1024-b, rate-1/2 low-density parity-check code decoder", *Solid-State Circuits, IEEE Journal of*, Vol.37, Iss.3, Mar 2002 Pages:404-412

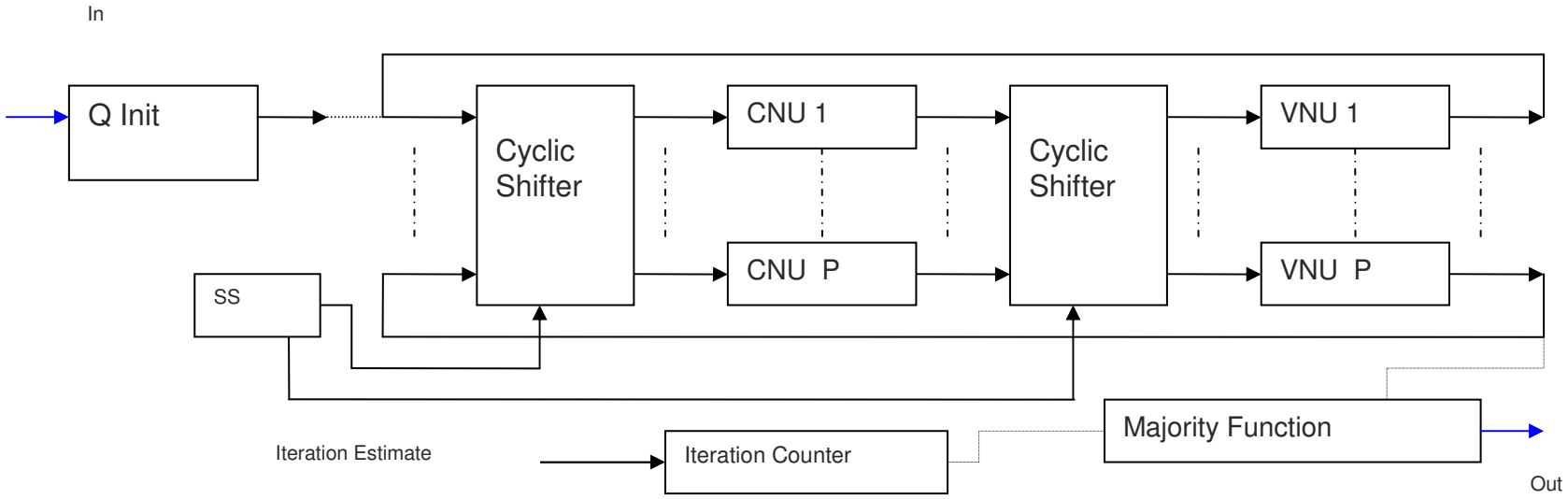
2082-bit Parallel Decoder Wiring



Outline

- Introduction of LDPC
- Problem Statement
- On-the-fly computation for QC-LDPC
- Multi-rate TPMP Decoder
- Multi-rate Layered Decoder
- Irregular QC-LDPC (IEEE802.16e, IEEE 802.11n,…)
- Fixed code TPMP Decoder
- Parallel Layered Decoder
- Results and performance comparison

TPMP, QC-LDPC, Multi-rate



Decoder for (3, 5) – structured LDPC code of length 1055. No message communication memory is needed. Possible due to structured property and scheduling.

Results

ASIC IMPLEMENTATION OF THE PROPOSED TPMP MULTI-RATE DECODER ARCHITECTURE

	Semi-Parallel multi-rate LDPC decoder [3]	Multi-rate TPMP Architecture regular QC-LDPC (Chapter 3)
LDPC Code	AA-LDPC, (3,6) code, rate 0.5, length 2048	(3,*k*) rate compatible array codes p=347. k=6,7,..12. length =$pk(2082,..,4164)$
Decoded Throughput, t_d,	640 Mbps	2.37 Gbps
Area	14.3 mm^2	7.62 mm^2
Frequency	125 MHz	500 MHz
Nominal Power Dissipation	787 mW	821 mW
Memory	51,680 bits	62,465 bits (including the channel LLR memory)
CMOS Technology	0.18 μ , 1.8V	0.13 μ , 1.2V
Decoding Schedule	TDMP, BCJR, it_{max}=10	TPMP, SP, it_{max}=20
Area Efficiency for t_d,	44.75 Mbps/mm^2	311 Mbps/ mm^2
Energy Efficiency for t_d,	123 pJ/Bit/Iteration	17 pJ/Bit/Iteration
Est. Area for 0.18u	14.3 mm^2	14.6 mm^2
Est. Frequency for 0.18u	125 MHz	360 MHz
Decoded Throughput(t_d) ,0.18u	640 Mbps	1.71 Gbps
Area Efficiency for t_d, 0.18u	44.75 Mbps/mm^2	117 Mbps/ mm^2
Energy Efficiency for t_d, 0.18 u	123 pJ/Bit/Iteration	38.25 pJ/Bit/Iteration

[3] Mansour et al "A 640-Mb/s 2048-Bit Programmable LDPC Decoder Chip"- IEEE Journal of Solid-State Circuits, March 2006

Results

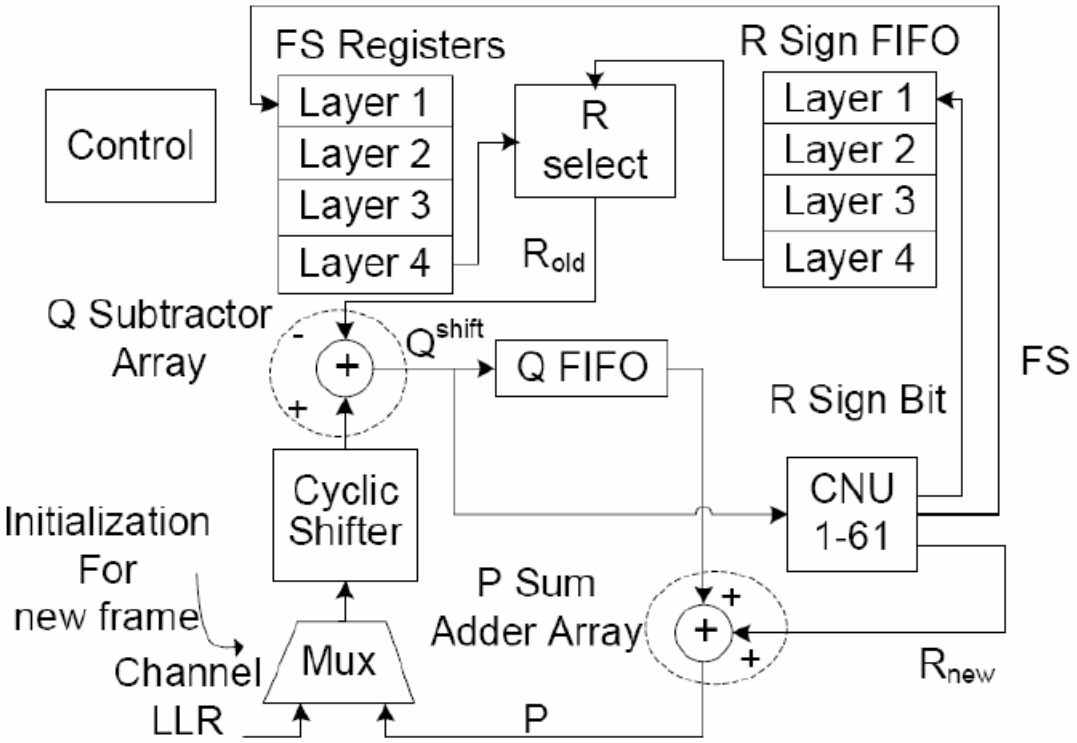
AREA DISTRIBUTION OF THE CHIP FOR (3, *k*) RATE COMPATIBLE ARRAY CODES, 0.13*u*

	Area (mm^2)
CNU Array (FIFO is not included)	2.1887
VNU Array	1.4274
Message Passing Memory+ Channel LLR memory	2.2768
2 Cyclic shifters	1.3703
Total chip area	7.6263

POWER DISTRIBUTION OF THE CHIP FOR (3,*k*) RATE COMPATIBLE ARRAY CODES, 0.13*u*

	Power (mW)
Logic(CNU,VNU and shifters)	482.5
Memory	199.7
Leakage	0.27
Clock	96.5
Wiring	48.2
Total	827.2

Layered Decoder architecture- QC-LDPC



$$\left[\vec{Q}_{l,n'}^{(i)}\right]^{-S(l,n)} = \left[\vec{P}_{n}\right]^{-S(l,n)} - \vec{R}_{l,n}^{(i-1)}$$

$$\vec{R}_{l,n}^{(i)} = f\left(\left[\vec{Q}_{l,n'}^{(i)}\right]^{-S(l,n)}\right),\ \forall n' = 1,2,\cdots k$$

$$\left[\vec{P}_{n}\right]^{-S(l,n)} = \left[\vec{Q}_{l,n'}^{(i)}\right]^{-S(l,n)} + \vec{R}_{l,n}^{(i)}$$

Performance comparison

	Semi-Parallel multi-rate LDPC decoder [3]	Multi-rate TDMP Architecture regular QC-LDPC
LDPC Code	AA-LDPC, (3,6) code, rate 0.5, length 2048	((5,k) rate compatible array codes $p=61$. k=10,11,..61length $=pk(610,.,3721)$
Decoded Throughput, t_d,	640 Mbps	590 Mbps
Area	14.3 mm^2	1.6 mm^2
Frequency	125 MHz	500 MHz
Nominal Power Dissipation	787 mW	257 mW
Memory	51,680 bits	37,210 bits
CMOS Technology	0.18 μ , 1.8V	0.13 μ , 1.2V
Decoding Schedule	TDMP, BCJR, it_{max}=10	TDMP, OMS, it_{max}=10
Area Efficiency for t_d,	44.75 Mbps/mm^2	369 Mbps/ mm^2
Energy Efficiency for t_d,	123 pJ/Bit/Iteration	44.2 pJ/Bit/Iteration
Est. Area for 0.18u	14.3 mm^2	~3.06 mm^2
Est. Frequency for 0.18u	125 MHz	~360 MHz
Decoded Throughput(t_d) ,0.18u	640 Mbps	426 Mbps
Area Efficiency for t_d, 0.18u	44.75 Mbps/mm^2	139.2 Mbps/mm^2
Energy Efficiency for t_d, 0.18 u	123 pJ/Bit/Iteration	99.45 pJ/Bit/Iteration
Application	Multi-rate application as well as fixed code application	Multi-rate application as well as fixed code application Rate-compatible array codes are considered for DSL applications.
Bit error rate Performance	Good	Good and similar to AA-LDPC

[3] Mansour et al "A 640-Mb/s 2048-Bit Programmable LDPC Decoder Chip"- IEEE Journal of Solid-State Circuits, March 2006

Design Statistics

TABLE V

AREA DISTRIBUTION OF THE CHIP FOR (5, *k*) RATE COMPATIBLE ARRAY CODES, 0.13*u*

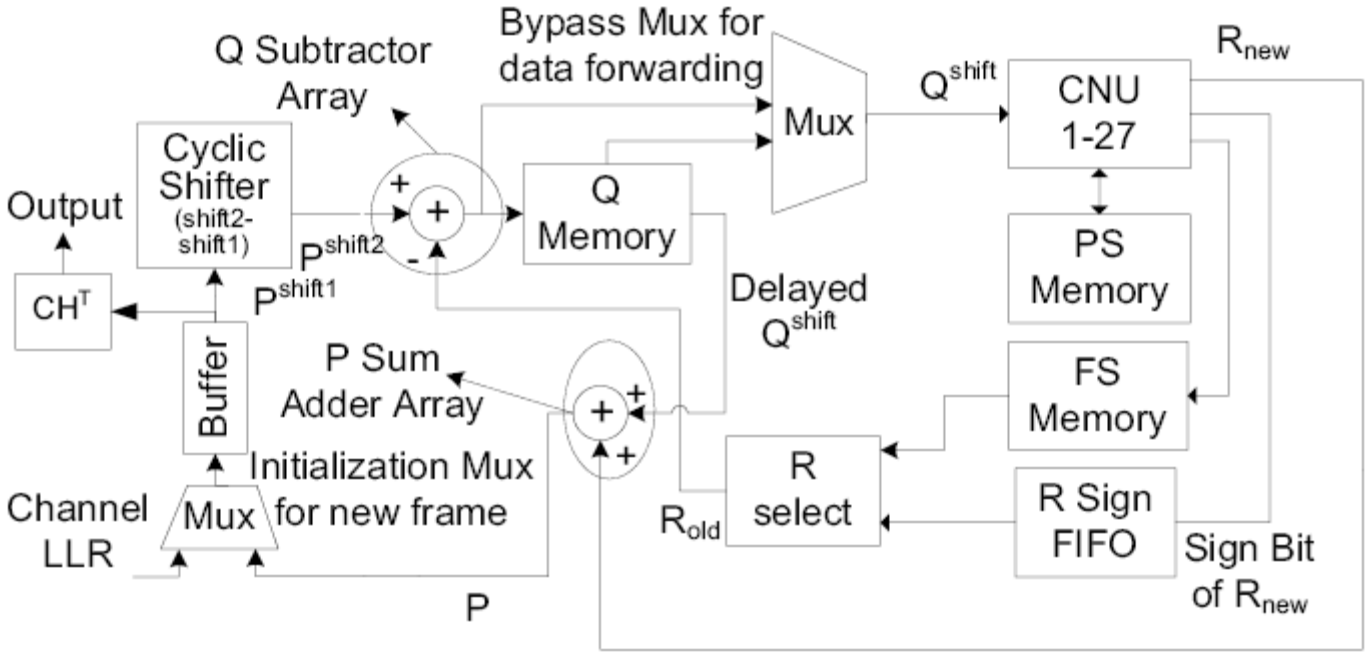
	Area (mm^2)
CNU Array	0.6691
VNU Array	0.0502
Memory	0.7053
Pipeline flip-flops	0.0167
Cyclic shifter	0.0803
Wiring	0.0761
Total chip area	1.5977

TABLE VI

POWER DISTRIBUTION OF THE CHIP FOR (3,*k*) RATE COMPATIBLE ARRAY CODES, 0.13*u*

	Power (mW)
Logic(CNU,VNU and shifters)	162.7115
Memory	45.4360
Leakage	0.0655
Clock	32.5423
Wiring	16.2712
Total	257.0

Layered Decoder for Irregular QC-LDPC codes



ASIC IMPLEMENTATION OF THE PROPOSED TPMP MULTI-RATE DECODER ARCHITECTURE

	Semi-Parallel multi-rate LDPC decoder [3]	Multi-rate TPMP Architecture regular QC-LDPC
LDPC Code	AA-LDPC, (3,6) code, rate 0.5, length 2048	Irregular codes up to length 2304 IEEE 802.16e WiMax LDPC codes
Decoded Throughput, t_d,	640 Mbps	1.37 Gbps
Area	14.3 mm^2	2.1 mm^2
Frequency	125 MHz	500 MHz
Nominal Power Dissipation	787 mW	282 mW
Memory	51,680 bits	60,288 bits
CMOS Technology	0.18 μ , 1.8V	0.13 μ , 1.2V
Decoding Schedule	TDMP, BCJR, it_{max}=10	TDMP, OMS, it_{max}=10
Area Efficiency for t_d,	44.75 Mbps/mm^2	649.5 Mbps/mm^2
Energy Efficiency for t_d,	123 pJ/Bit/Iteration	21 pJ/Bit/Iteration
Est. Area for 0.18u	14.3 mm^2	~4.02 mm^2
Est. Frequency for 0.18u	125 MHz	~360 MHz
Decoded Throughput(t_d) ,0.18u	640 Mbps	989 Mbps
Area Efficiency for t_d, 0.18u	44.75 Mbps/mm^2	246 Mbps/mm^2
Energy Efficiency for t_d, 0.18 u	123 pJ/Bit/Iteration	47.25 pJ/Bit/Iteration
Application	Multi-rate application as well as fixed code application	IEEE 802.16e Multi-rate application .
Bit error rate Performance	Good	Very good and close to capacity

Design Statistics

AREA DISTRIBUTION OF THE CHIP FOR WIMAX LDPC CODES

Architecture 1 :: supports z=24,48 and 96 and all the code rates)

Architecture 2: Fully Compliant to WiMax supports z=24,28,32,…,and 96 and all the code rates. The only difference is the replacement of logarithmic cyclic shifter with Master-slave Benes router. This will increase the complexity of router by almost 5x and increase the power dissipation of the decoder by 70% when compared to the Architecture 1.

	Architecture 1,Area (mm^2)	Architecture 2,Area (mm^2)
CNU Array	0.5265	0.5265
VNU Array	0.0790	0.0790
Memory	1.2317	1.2317
Pipeline flip-flops	0.0263	0.0263
Cyclic shifter	0.1474	0.7371
Wiring	0.1005	0.1300
Total chip area	2.1115	2.7307

Design Statistics

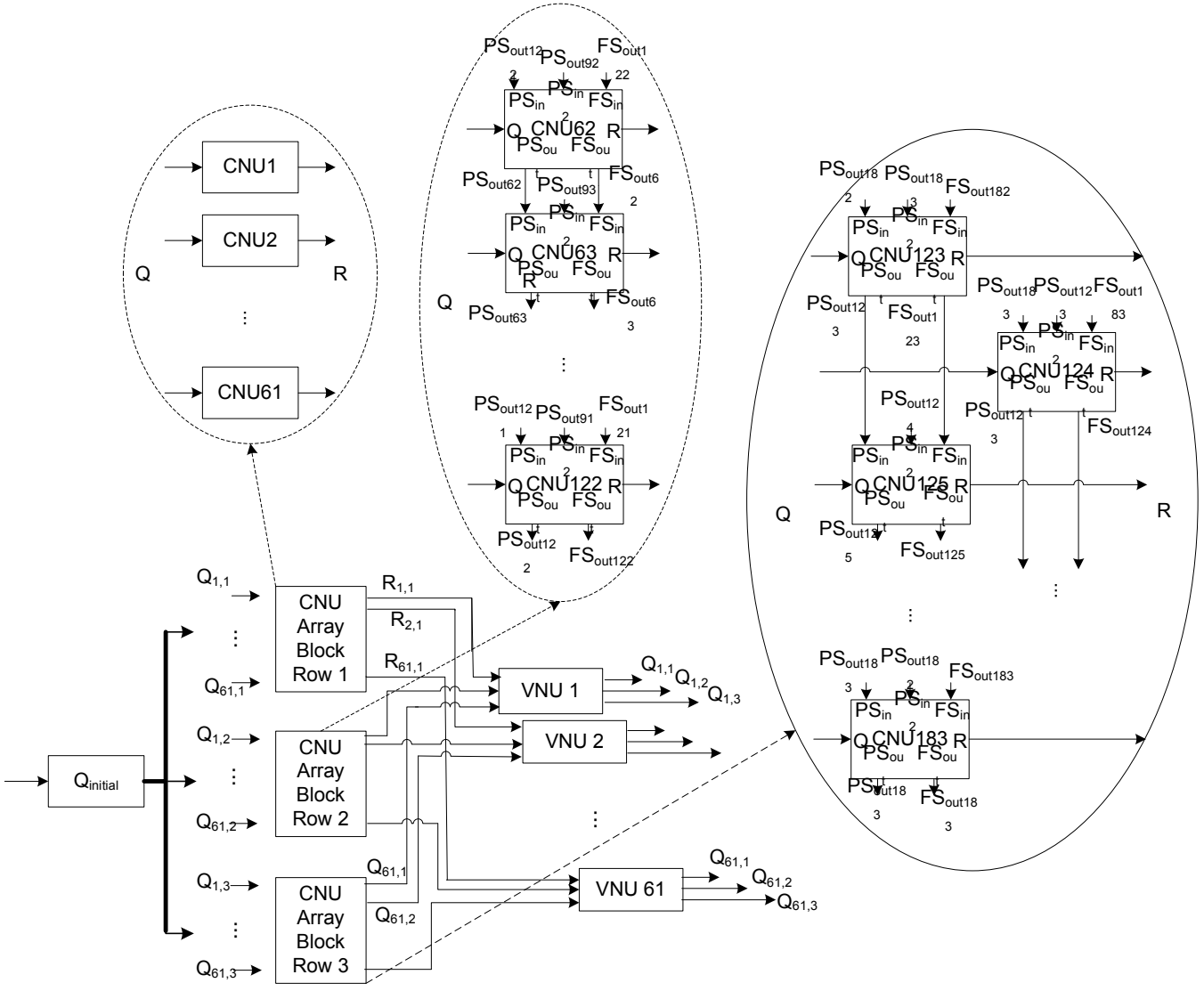
TABLE VI

POWER DISTRIBUTION OF THE CHIP FOR (3,*k*) RATE COMPATIBLE ARRAY CODES, 0.13*u*

(supports z=24,48 and 96 and all the code rates)

	Architecture 1,Power (mW)	Architecture 2, Power (mW)
Logic(CNU,VNU and shifters)	160.4127	273.3087
Memory	73.8816	73.8816
Leakage	0.0866	0.1120
Clock	32.0825	54.6617
Wiring	16.0413	27.3309
Total	282.5	429.3

TPMP-Array LDPC. Fixed Code



Comparison of Fixed Code TPMP

SUMMARY OF THE PROPOSED FIXED-CODE DECODER ARCHITECTURE, CODE 1

	Fully Parallel LDPC decoder [8]	TPMP Architecture regular Array QC-LDPC
Decoded Throughput, t_d,	1 Gbps	5.78 Gbps
Area	52.5 mm^2	9.9 mm^2
Frequency	64 MHz	500 MHz
Power Dissipation	690 mW	695 mW
Memory	34816 bits (scattered flip-flops)	20820 bits for 2 input buffers 6246 bits for sign memory
LDPC Code	Random LDPC code, rate 0.5, length 1024	(3,6) array code, rate 0.5, length 2082
CMOS Technology	0.16 μ, 1.5V	0.13 μ, 1.2V
Decoding Schedule	TPMP, SP, it_{max}=64	TPMP, SP, it_{max}=20
Area Efficiency for t_d,	19 Mbps/mm^2	585.2 Mbps/mm^2
Energy Efficiency for t_d,	10.1 pJ/Bit/Iteration	6.01 pJ/Bit/Iteration
Scalability of Design for other code parameters and longer lengths	No	Yes

[8] Blanksby, A.J.; Howland ,C.J, A 690-mW 1-Gb/s 1024-b, rate-1/2 low-density parity-check code decoder", Solid-State Circuits, IEEE Journal of, Vol.37, Iss.3, Mar 2002 Pages:404-412

Parallel Layered Decoder architecture

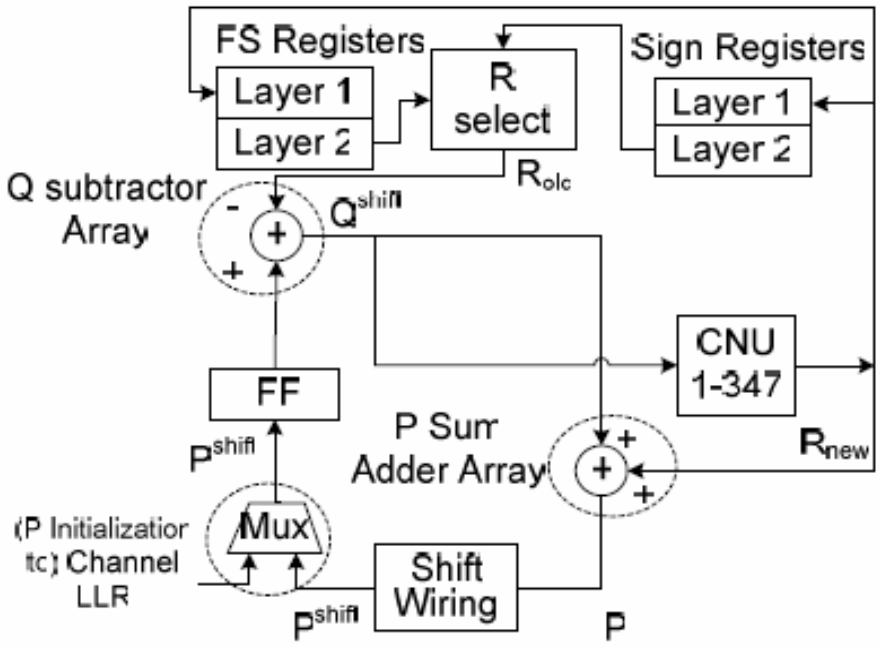


Figure 3: Dataflow graph of the proposed parallel architecture for layered decoder

Comparison of Parallel Decoder

	[3]	[8]	[9]	This work
Decoded Throughput, t_d	640 Mbps	1.0 Gbps	3.2 Gbps	6.9 Gbps
Area	14.3 mm^2	52.5 mm^2	17.64 mm^2	5.29 mm^2
Decoder's Internal memory	51680 bits (SRAM) 9216 bits (flip-flops)	34816 bits (scattered flip-flops)	98944 bits (scattered flip-flops)	27066 bits (scattered flip-flops)
Router/Wiring	3.28 mm^2-Network	26.25 mm^2-Wiring	Details unknown	0.89 mm^2-Wiring
Frequency, f	125 MHz	64 MHz	100 MHz	100 MHz
Nominal Power Dissipation	787 mW	690 mW	NA	75 mW
Area Efficiency for t_d,	44.7 Mbps/mm^2	19.04 Mbps/mm^2	181.4 Mbps/mm^2	493.0 Mbps/mm^2
Energy Efficiency for t_d,	123 pJ/Bit/Iteration	10.1 pJ/Bit/Iteration	NA	1.1 pJ/Bit/Iteration
LDPC Code	AA-LDPC, (3,6) code, rate 0.5	Random and Irregular code, rate 0.5	RS-LDPC, (6,32) code, rate 0.8413	Array code, (3,6) code, rate 0.5
Check Node Update	BCJR	SP	SP	Offset Min-Sum,OMS
Decoding Schedule	TDMP, it_{max}=10	TPMP, it_{max}=64	TPMP, it_{max}=32	TDMP, it_{max}=10
Block Length, N	2048	1024	2048	2082
SNR(E_b/N_o) for BER of 1e-6	2.4 dB	2.8 dB	4.0 dB	2.6 dB
Average CCI due to pipelining	40	1	1	3
CMOS Technology	0.18μ, 1.8V	0.16μ, 1.5V	0.18μ, 1.8V	0.13μ, 1.2V
Est. Area for 0.18u	14.3 mm^2	~66.4 mm^2	17.64 mm^2	~10.1 mm^2
Est. Frequency for 0.18u	125 MHz	~56.8 MHz	100 MHz	~72 MHz
Decoded Throughput(t_d) ,0.18u	640 Mbps	887.5 Mbps	3.2 Gbps	4.98 Gbps
Area Efficiency for t_d, 0.18u	44.7 Mbps/mm^2	13.36 Mbps/mm^2	181.4 Mbps/mm^2	493.0 Mbps/mm^2
Energy Efficiency for t_d, 0.18u	123 pJ/Bit/Iteration	10.1 pJ/Bit/Iteration	NA	1.1 pJ/Bit/Iteration
Scalability of Design	Yes	No	No	Yes

[9] A. Darabiha, A. C. Carusone and F. R. Kschischang, "Multi-Gbit/sec low density parity check decoders with reduced interconnect complexity," IEEE Int. Symp. on Circuits and Systems (ISCAS), Kobe, Japan, May 2005.

Comparison of Fixed code decoders

	Fully Parallel LDPC decoder [8]	TPMP Architecture regular Array QC-LDPC (Chapter 5)	TDMP Architecture for regular Array QC-LDPC (Chapter 8)
Decoded Throughput, t_d,	1 Gbps	1.5 Gbps	6.94 Gbps
Area	52.5 mm^2	3.39 mm^2	5.39 mm^2
Frequency	64 MHz	500 MHz	100 MHz
Nominal Power Dissipation	690 mW	156.5 mW	75 mW
LDPC Code	Random LDPCr code, rate 0.5, length 1024	(4,30) array code of length 1830	(3,6) array code of length 2082
CMOS Technology	0.16 μ, 1.5V	0.13 μ, 1.2V	0.13 μ, 1.2V
Decoding Schedule	TPMP, SP, it_{max}=64	TPMP, SP, it_{max}=20	TDMP, OMS, it_{max}=10
Area Efficiency for t_d,	19 Mbps/mm^2	510 Mbps/mm^2	1288 Mbps/mm^2
Energy Efficiency for t_d,	10.1 pJ/Bit/Iteration	5.6 pJ/Bit/Iteration	1.1 pJ/Bit/Iteration
Scalability of Design for other code parameters and longer lengths	No	Yes	Yes
Application	Fixed code application	Very High throughput and low error-floor applications such as magnetic recording channels, Ethernet and optical links	Very High throughput and low error-floor applications such as magnetic recording channels, Ethernet and optical links.
Bit error rate Performance	Good	Good	Good

Comparison of Multi-rate decoders

	Semi-Parallel multi-rate LDPC decoder [3]	Multi-rate TPMP Architecture regular QC-LDPC (Chapter 3)	Multi-rate TDMP Architecture for regular QC-LDPC (Chapter 5)	Multi-rate TDMP Architecture for irregular QC-LDPC (Chapter 6)
LDPC Code	AA-LDPC, (3,6) code, rate 0.5, length 2048	(3,*k*) rate compatible array codes *p*=347. *k*=6,7,..12	(5,*k*) rate compatible array codes *p=61*. *k*=10,11,..61	Irregular codes up to length 2304 IEEE 802.16e WiMax LDPC codes
Decoded Throughput, t_d,	640 Mbps	2.37 Gbps	590 Mbps	1.37 Gbps
Area	14.3 mm^2	7.62 mm^2	1.6 mm^2	2.1 mm^2
Frequency	125 MHz	500 MHz	500 MHz	500 MHz
Nominal Power Dissipation	787 mW	821 mW	257 mW	282 mW
CMOS Technology	0.18 μ, 1.8V	0.13 μ, 1.2V	0.13 μ, 1.2V	0.13 μ, 1.2V
Decoding Schedule	TDMP, BCJR, it_{max}=10	TPMP, SP, it_{max}=20	TDMP, OMS, it_{max}=10	TDMP, OMS, it_{max}=10
Area Efficiency for t_d,	44.75 Mbps/mm^2	311 Mbps/ mm^2	369 Mbps/ mm^2	649.5 Mbps/ mm^2
Energy Efficiency for t_d,	123 pJ/Bit/Iteration	17 pJ/Bit/Iteration	44.2 pJ/Bit/Iteration	21 pJ/Bit/Iteration
Application	Multi-rate application as well as fixed code application	DSL, Wireless	DSL,Wireless	Wireless, IEEE 802.11n, IEEE 802.16e, IEEE 802.22
Bit error rate Performance	Good	Good	Good	Very good and close to capacity

Contributions

- An area (logic and memory) and power efficient multi-rate architecture for Sum of Products Decoding of LDPC
- An area (logic and memory) and power efficient multi-rate architecture for Layered decoding of regular QC- LDPC
- An area (logic and memory) efficient multi-rate architecture for Layered decoding of irregular QC- LDPC for IEEE 802.11n, 802.16e and 802.22
- A memory efficient and router less architecture for min-sum decoding of Decoding of LDPC
- A very high throughput area (logic and memory) efficient parallel layered decoder for array LDPC for low error floor applications
- FPGA prototyping and ASIC design clearly illustrates the advantages of the proposed decoder architectures

Publications

Published

1. Gunnam, K.; Gwan Choi; Yeary, M.; “An LDPC decoding schedule for memory access reduction,” Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on Volume 5, 17-21 May 2004 Page(s):V - 173-6 vol.5

Accepted

2. Gunnam, Kiran K.; Choi, Gwan S.; Wang, Weihuang; Kim, Euncheol; Yeary, Mark B.; “Decoding of Quasi-cyclic LDPC Codes Using an On-the-Fly Computation,” Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on Oct.-Nov. 2006

3. Gunnam, K.K.; Choi, G.S.; Yeary, M.B.; “A Parallel VLSI Architecture for Layered Decoding for Array LDPC Codes,” VLSI Design, Jan 2007.

In Review

4. Gunnam, KK; Choi, G. S.; Yeary, M. B.; Atiquzzaman, M.; “VLSI Architectures for Layered Decoding for Irregular LDPC Codes of WiMax,” Submitted for ICC 2007.

5. Gunnam, K.; Gwan Choi; Weihuang Wang; Yeary, M.; “Multi-Rate Layered Decoder Architecture for Block LDPC Codes of the IEEE 802.11n Wireless Standard,” Submitted for ISCAS 2007

6. Gunnam, K.; Weihuang Wang; Gwan Choi; Yeary, M.; “VLSI Architectures for Turbo Decoding Message Passing Using Min-Sum for Rate-Compatible Array LDPC Codes,” Submitted for International Symposium on Wireless Pervasive Computing, 2007.

Other Publications

1. **K.Gunnam**, K.Chadha and M.B.Yeary, “New Optimizations for Carrier Synchronization in Single Carrier Systems”, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005).

2. J. Valasek, **K.Gunnam**, J. Kimmett, D. Hughes and J. Junkins., "Vision Based Sensor and Navigation System for Autonomous Aerial Refueling," ***Journal of Guidance and Control***, October 2005.

3. **K.Gunnam**, D.C.Hughes, J.L.Junkins and N.Kehtarnavaz,."A Vision Based DSP Embedded Optical Navigation Sensor” **IEEE Sensors Journal**, vol.2.pp 428-442,Oct 2002.

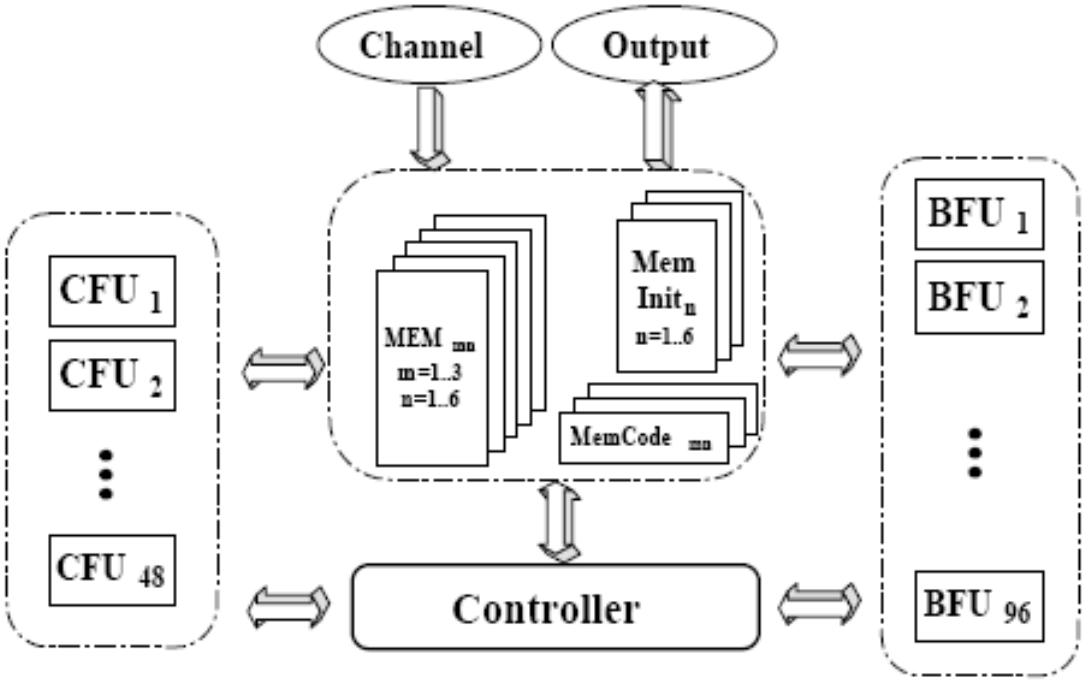
4. **K.Gunnam**, D.C.Hughes, J.L.Junkins and N.Kehtarnavaz,."A DSP Embedded Optical Navigation System" Proceedings of Sixth IEEE International Conference on Signal Processing, ICSP 2002.

- Thank you !

Backup Slides

Index of Backup slides

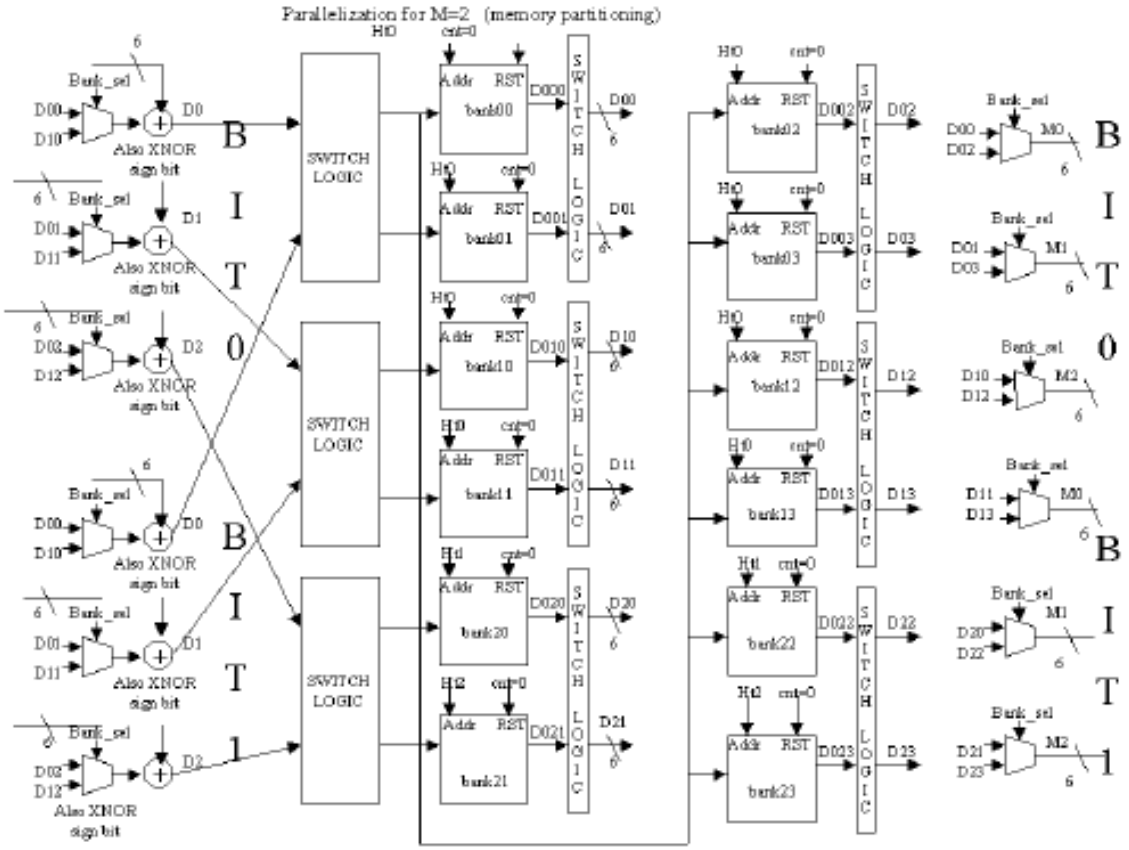
Semi-Parallel Architecture



Semi Parallel Architecture

[10] Karkooti etal. Semi-parallel reconfigurable architectures for real-time LDPC decoding ; Proceedings. ITCC 2004.

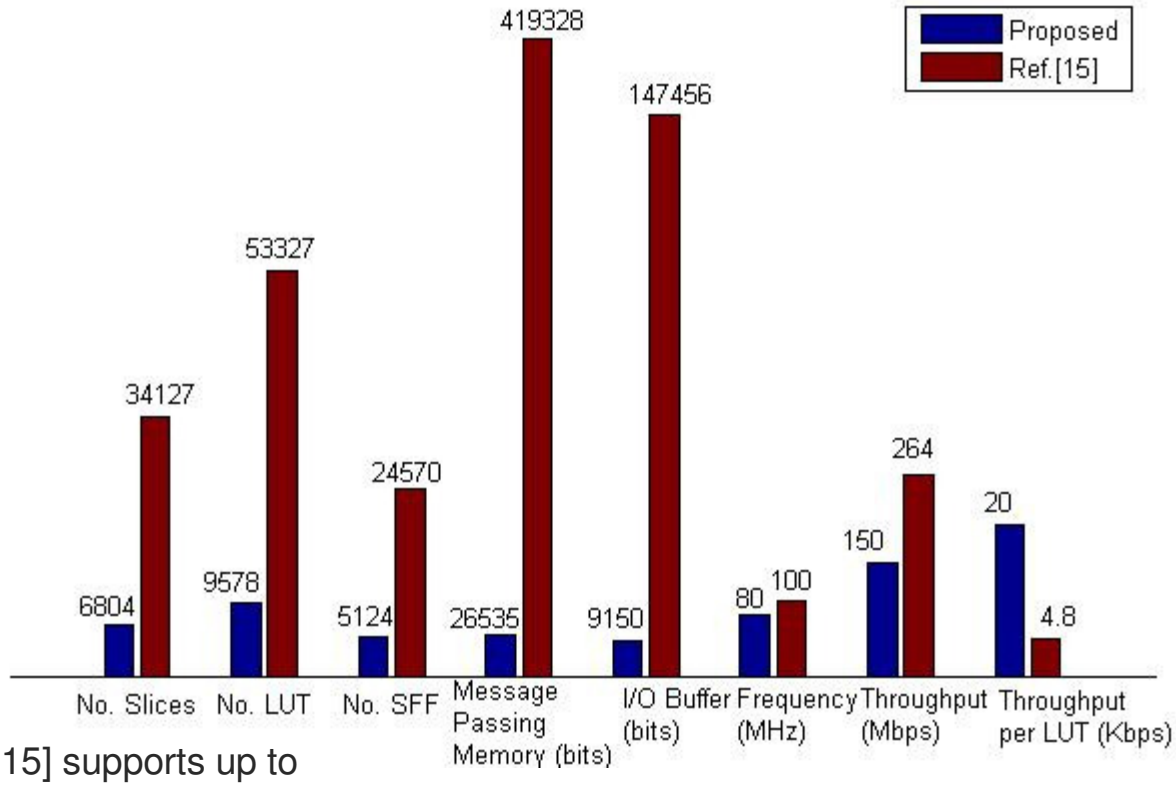
Semi-Parallel Architecture



Semi Parallel Architecture

A. Selvarathinam, G.Choi, K. Narayanan, A.Prabhakar, E. Kim, "A Massively Scalable Decoder Architecture for Low-Density Parity-Check Codes", in proceedings of ISCAS'2003

TPMP, Multi-rate decoder FPGA Results



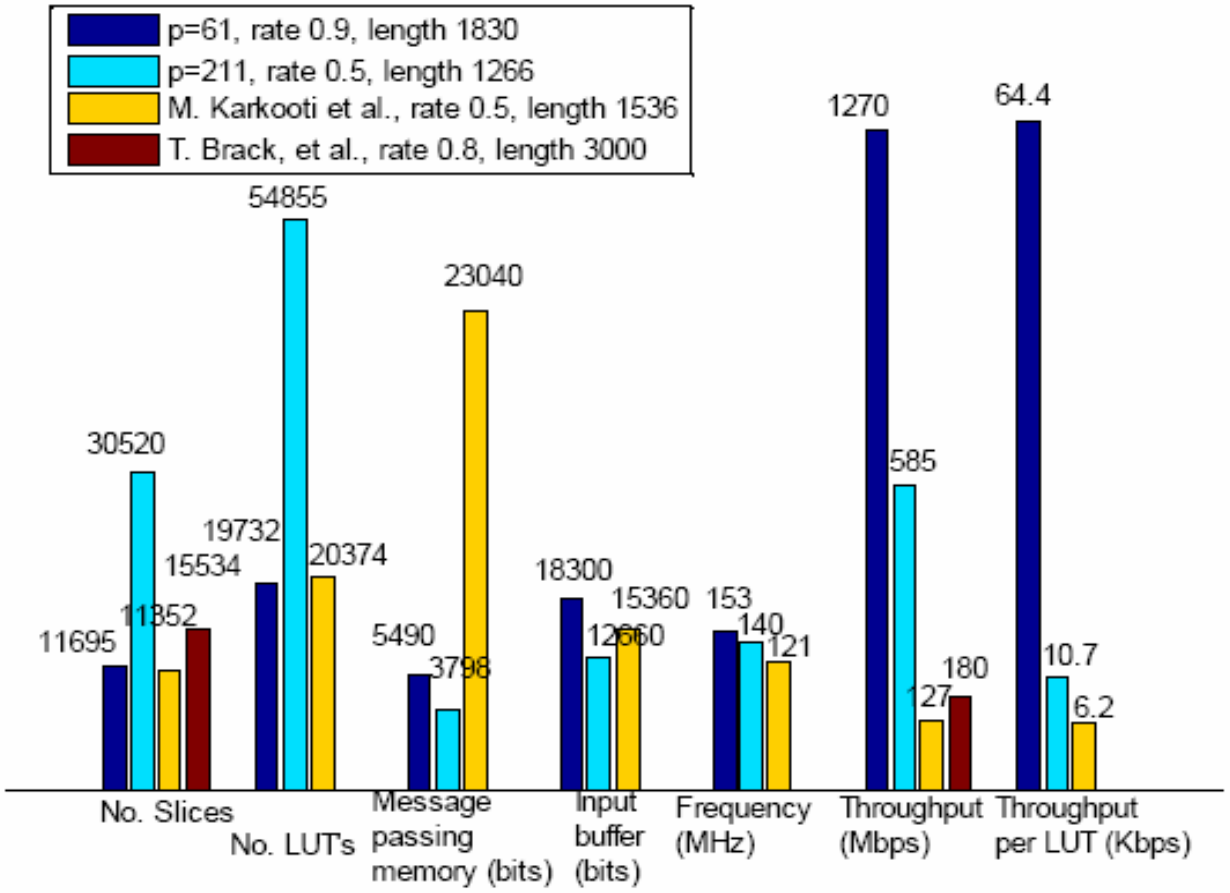
The authors in [15] supports up to code length 10,000. When Normalized to length of 1830, memory in [15] is 3x bigger.

Throughput compared for

max 6 iterations.

[15] Lei Yang etal., "An FPGA implementation of low-density parity-check code decoder with multi-rate capability," *Proceedings of the ASP-DAC 2005.*

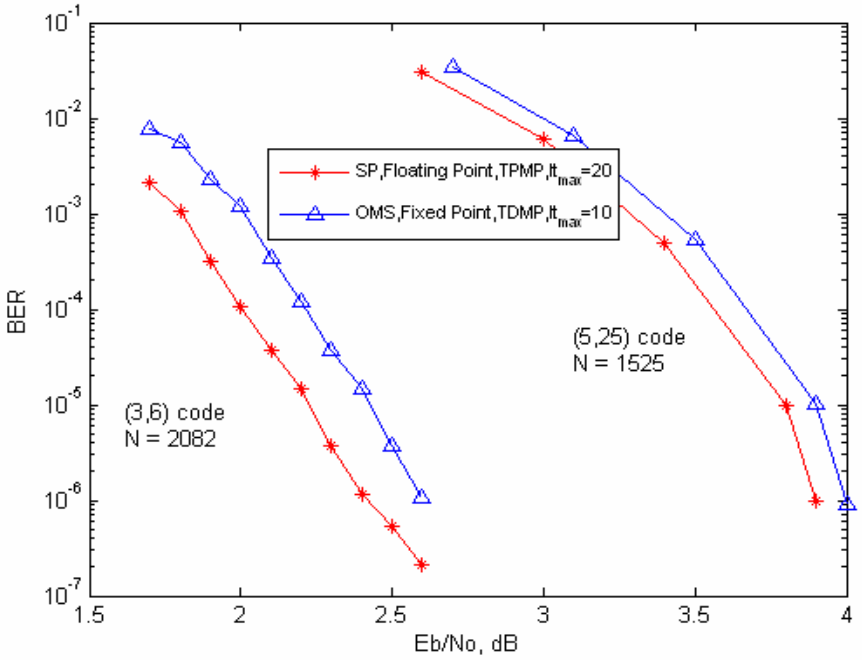
TPMP-Fixed Code architecture FPGA Results

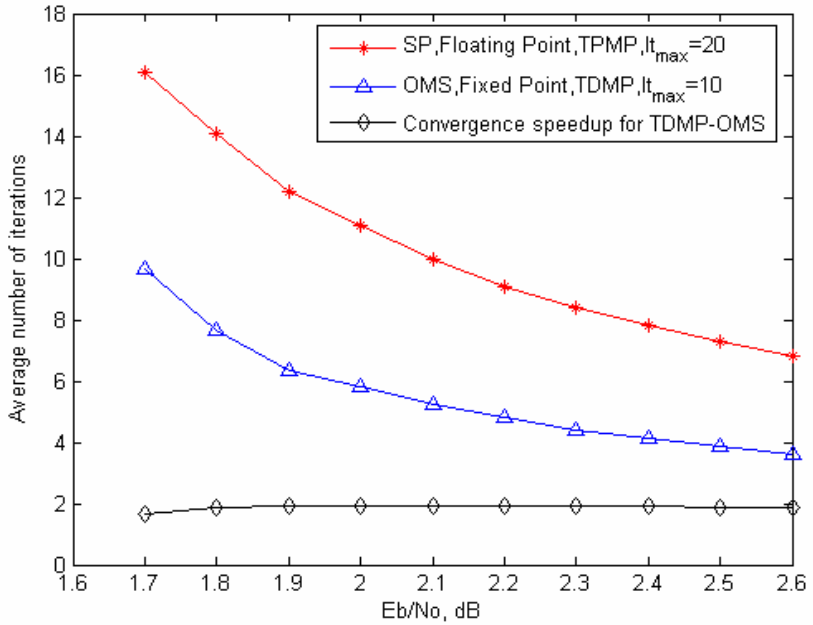


[10] Karkooti etal. Semi-parallel reconfigurable architectures for real-time LDPC decoding ; Proceedings. ITCC 2004.

[11] T. Brack etal. "Disclosing the LDPC Code Decoder Design Space" Design, Automation and Test in Europe (DATE) Conference 2006, March 2006, Munich, Germany

Layered Decoding for array codes, BER results





a)Bit Error Rate performance of the proposed TDMP Decoder using OMS(j=3,k=6,p=347,q=0) array LDPC code of length N=2082 and (j=5,k=25,p=61,q=0) array LDPC code of length N=1525.
(b) Convergence speed up of TDMP-OMS over TPMP-SP.
Results shown for (j=3,k=6,p=347,q=0) array LDPC code of length N=2082.
Here Itmax= Maximum number of iterations.

Layered Decoder for Irregular codes,802.16e

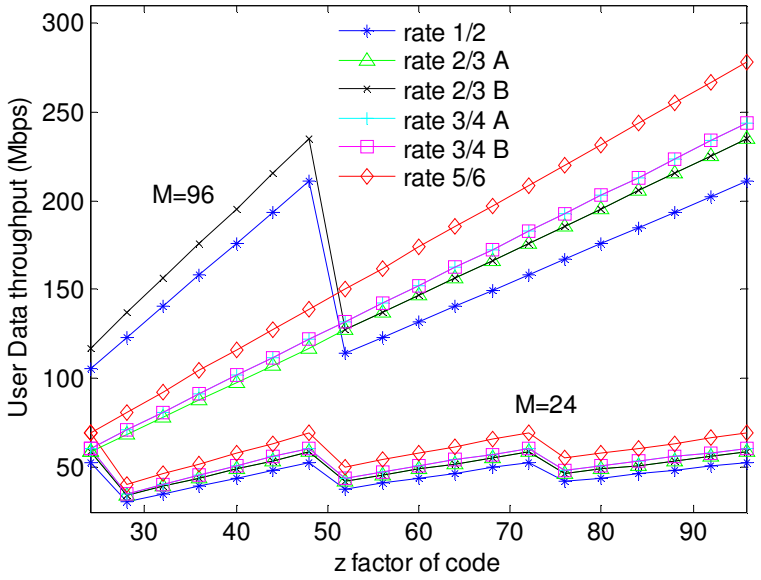
TABLE II: FPGA IMPLEMENTATION RESULTS
(Device, *Xilinx 2V8000ff152-5,* frequency *110MHz)*

	Used			Available
	M=24	M=48	M=96	
Slices	1640	3239	6568	46592
LUT	2982	5664	11028	93184
SFF	1582	3165	6330	93184
BRAM	38	73	100	168
Memory (bits)	65760	65760	60288	
Through-put(Mbps)	41~70	57~139	61~278	

TABLE III: IMPLEMENTATION COMPARISON

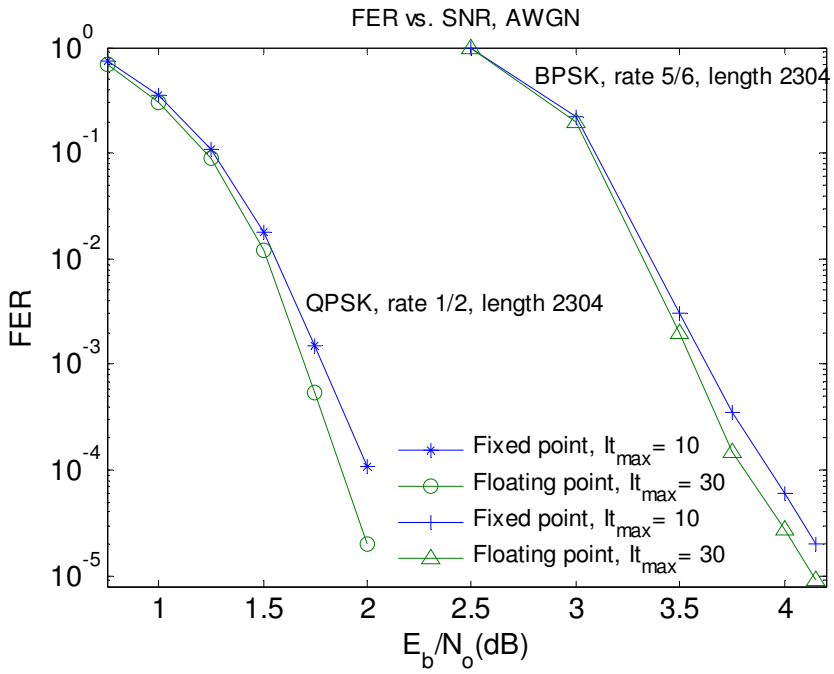
	M. Karkooti *et al*	T. Brack *et al.*	Proposed (M=24)
Slices	11352	14475	1640
LUT	20374	N/A	2982
Slice FF	N/A	N/A	1582
BRAM	66	165	38
Throughput (Mbps)	127	180	41~70

Layered Decoder Throughput Results- FPGA, 802.16e



User data throughput of the proposed decoder vs. the expansion factor of the code,z, for different numbers of decoder parallelization,M..

Layered Decoder, FER results, 802.16e



Frame-error rate results.

Layered Decoder Throughput Results- FPGA, 802.11n

FPGA IMPLEMENTATION RESULTS THE MULTI-RATE DECODER. FULLY COMPLIANT TO IEEE 802.11N(SUPPORTS $z_0 = 27, 54, 81$ AND ALL THE CODE RATES) (DEVICE, XILINX 2V8000FF152-5, FREQUENCY 110MHZ)

	$M = 27$	$M = 54$	$M = 81$	Available
Slices	1836	3647	5514	46592
LUT	3317	6335	9352	93184
SFF	1780	3560	5341	93184
BRAM	33	65	97	168
Memory(bits)	23376	39360	55344	
Throughput(Mbps)				
$z_0 = 81$	119	238	356	
$z_0 = 54$	119	238	178	
$z_0 = 27$	119	119	119	

Layered Decoder Throughput Results- ASIC, 802.11n

ASIC IMPLEMENTATION RESULTS THE MULTI-RATE DECODER FOR $M = 81$ (0.13 μ M TECHNOLOGY [15], FREQUENCY 500MHZ)

Resource	Area (mm^2)	Component	Power (mW)
CNU	0.444	Memory	62.3
VNU	0.067	leakage	0.07
Storage	1.039	Clock	27.1
Flip-flop	0.022	wiring	13.5
Shifter	0.124	active power	135.35
wiring	0.085	total power	238.4
total	1.7816		
Throughput(Mbps)	541, 1082 and 1618 for $z_0 = 27, 54$ and 81		

Proposed decoder takes around 100K logic gates and 55344 memory bits.

[16] takes 375 K logic gates and 88452 RAM bits for memory for a throughput of 940 Mbps

[16] Rovini, M.; L'Insalata, N.E.; Rossi, F.; Fanucci, L., "VLSI design of a high-throughput multi-rate decoder for structured LDPC codes," Digital System Design, 2005. Proceedings. 8th Euromicro Conference on , vol., no.pp. 202- 209, 30 Aug.-3 Sept. 2005

Cyclic Shifter

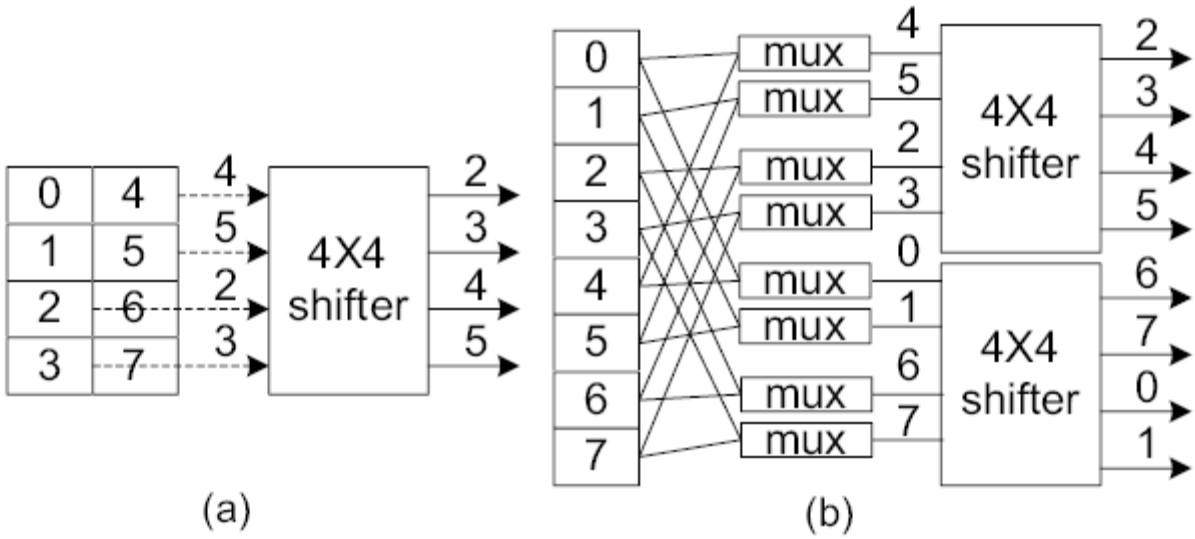


Fig. 5. (a) 8×8 shifting with 4×4 cyclic shifter. (b) 8×8 shifting with two 4×4 cyclic shifters.

This arrangement can support the base matrices having the expansion factors multiples of z by using z x z cyclic shifters.

Works for 802.11n in which the expansion factors are 27,54,81

Works for limited configurations of 802.16e in which the expansion factors are 24,28,..,96.[24 x 24 shifter works for 24,48 and 96]

Benes Network

Omega network is blocking network which can do cyclic permutations and other limited permutations.

Benes network is a non-blocking network which can do any permutation on the input vector. It is essentially combination of two omega n/w.

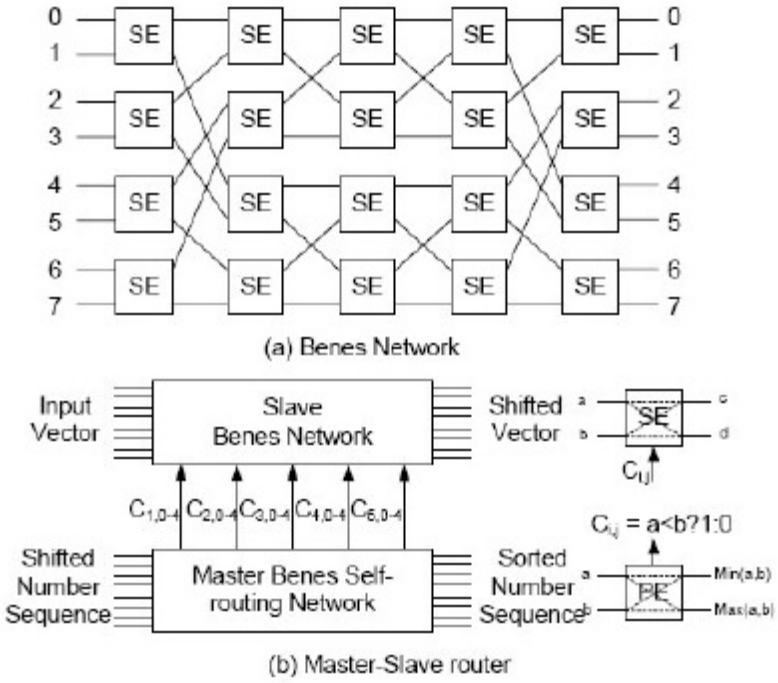
However the control complexity is more for both the networks.

In the existing work, the control signals are pre-computed and stored assuming only one *H* matrix need to be supported.

[3] Mansour et al "A 640-Mb/s 2048-Bit Programmable LDPC Decoder Chip"-IEEE Journal of Solid-State Circuits, March 2006

[12] Malema, G.; Liebelt, M., "Interconnection Network for Structured Low-Density Parity-Check Decoders," Communications, 2005 Asia-Pacific Conference on , vol., no.pp. 537- 540, 03-05 Oct. 2005

Master-Slave Router



In 802.16e in which the expansion factors are 24,28,..,96.

Assume that we have a monolithic 96 x 96 cyclic shifter. It can not do cyclic shift on vector size less than 96.

To accommodate all different vector sizes, we need to use a generic router such as Bene's n/w.

Control signals can be generated by using another self routing n/w

Master-Slave Router

Note that this memory for providing control signals to this network is equal to $\frac{M}{2}(2\log 2(M)-1)$ bits for every shift value that needs to be supported.

This will be a very huge requirement for supporting all the WiMax codes. Note that, the memory needed for storing control signals for Omega network is around 1.22 mm2 in out of the decoder chip area of 14.1 mm2.[3].

To support 19 different expansion factors and 6 types of base *H* matrices *in run time*, the control signal memory needs approximately 139.08 mm2.

Assume that we need to perform a cyclic shift of 2 on a message vector of length 4 using a 8 x8 Slave Benes network.

Supply the integers(2,3,0,1,4,5,6,7) to the Master Benes network which is always configured to sort the inputs and output (0,1,2,…7).

Though the sorting is based on [13], the new proposal is the conception of master-slave router.

[13] Gocal, B. 1996. Bitonic Sorting on Bene Networks. In Proceedings of the 10th international Parallel Processing Symposium (April 15 - 19, 1996). IPPS. IEEE Computer Society, Washington, DC, 749-753.

Min-Sum algorithm

- Original min-sum algorithm has performance degradation of up to 0.5 dB.
- Offset min-sum reduces this to 0.1 dB
- So offset min-sum is the reduced complexity algorithm among the available check node updates [6].

[14] M. P. C. 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, no. 5, pp. 673–680, May 1999.

[15] J. Chen and M. Fossorier, ``Near Optimum Universal Belief Propagation Based Decoding of Low-Density Parity Check Codes", IEEE Transactions on Communications, vol. COM-50, pp. 406-414, March 2002.

[6] J. Chen, A. Dholakia, E. Eleftheriou, M. Fossorier and X. Hu, "Reduced-complexity decoding of LDPC codes," IEEE Transactions on Communications, vol. 53, pp. 1288-1299, Aug. 2005