SHIRAZ UNIVERSITY
OF TECHNOLOGY

Iranian Association
of Electrical Engineering
Sciences
License 2011

2022 30th International Conference on Electrical Engineering

Area-Efficient Partially-Pipelined Architecture for Fast-SSC Decoding of Polar Codes

Presented by:

Matin Hashemi

Authors:

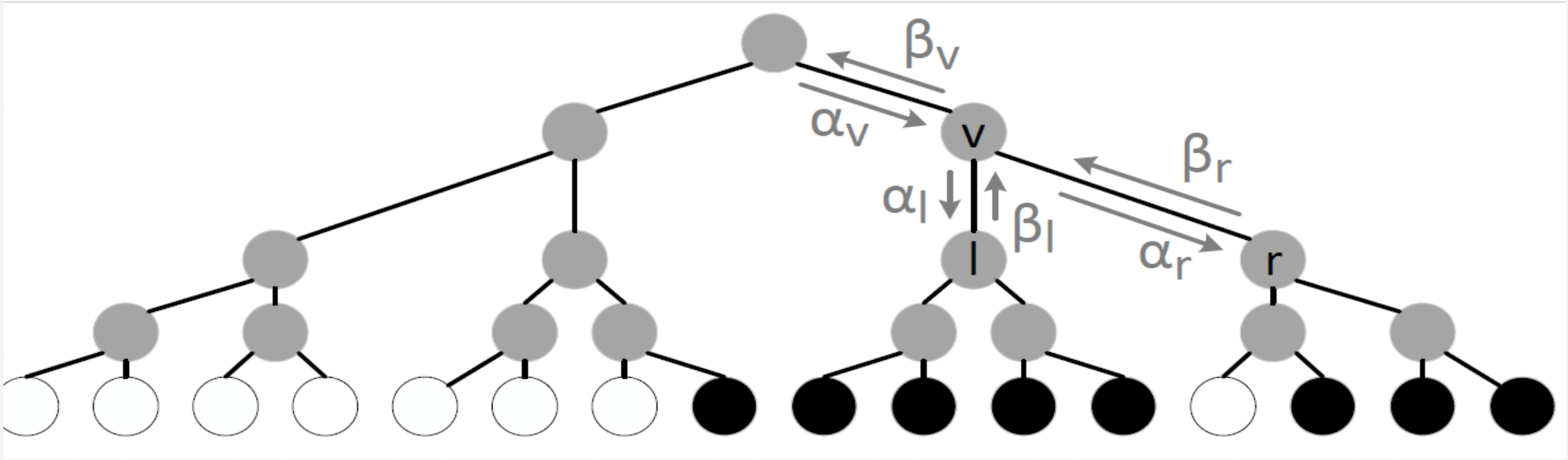
Mehdi Saeidi, Matin Hashemi
Sharif University of Technology

Motivation:

- Conventional fully-unrolled partially-pipelined polar decoders suffer from low utilization of processing elements
- In our proposed method, the number of processing elements is reduced, but they process data more often
- No degradation in throughput or latency

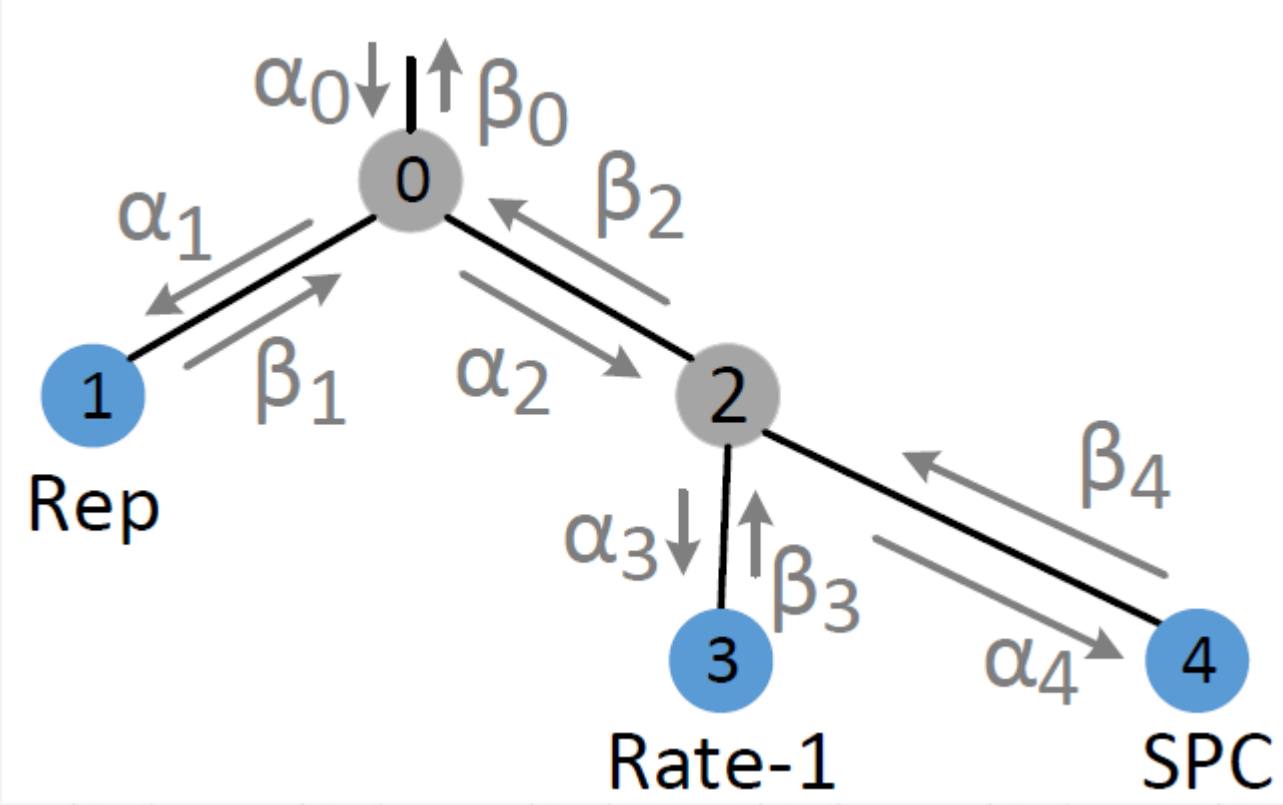
Successive Cancellation (SC) Algorithm:

- The operations to traverse the tree are called F, G, and Combine.
- This algorithm suffers from high latency in practical codeword sizes
- High latency makes the pipelined architecture infeasible



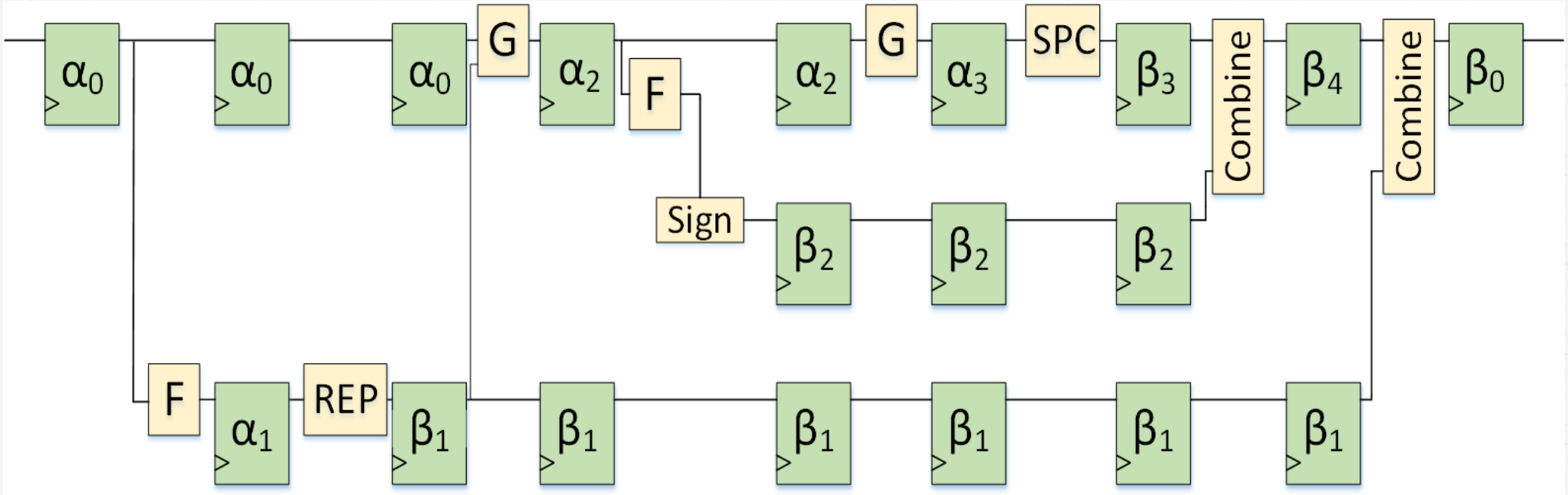
Fast-SSC Algorithm:

- Fast-SSC algorithm replaces specific sub-trees with one node
- This results in reduced latency, which makes the pipelined architecture feasible



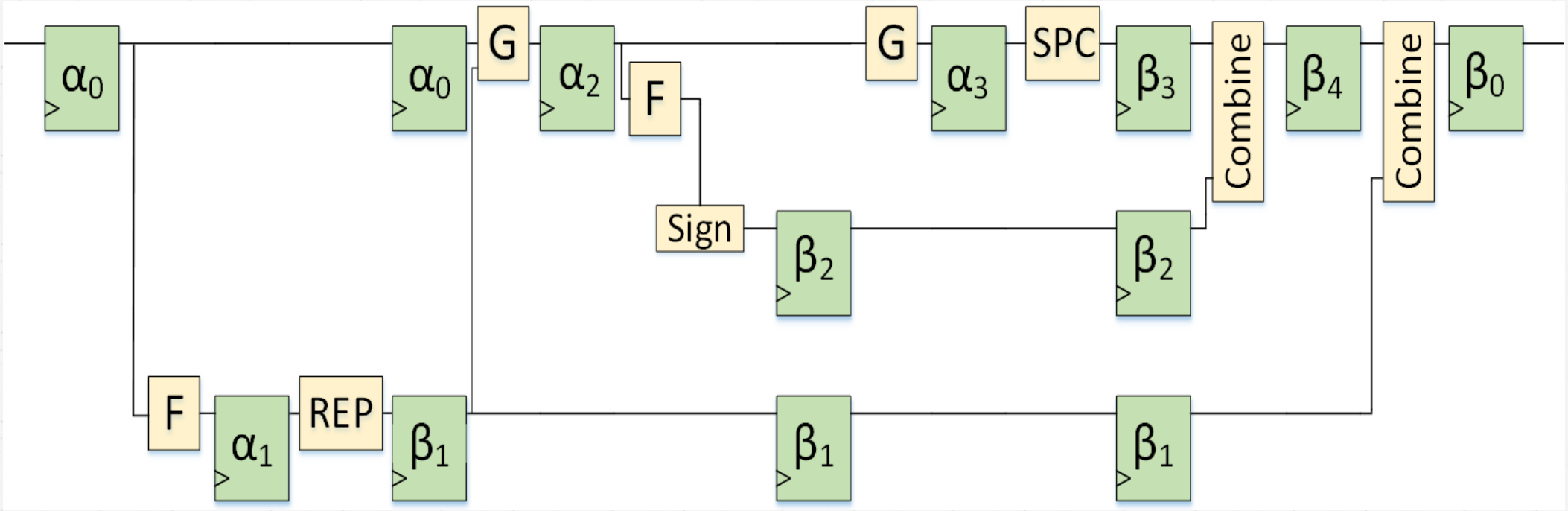
Fully-Unrolled Deeply-Pipelined Architecture:

- **Some of the LLR vectors need to cross many pipeline stages**
- **This results in high register requirement**



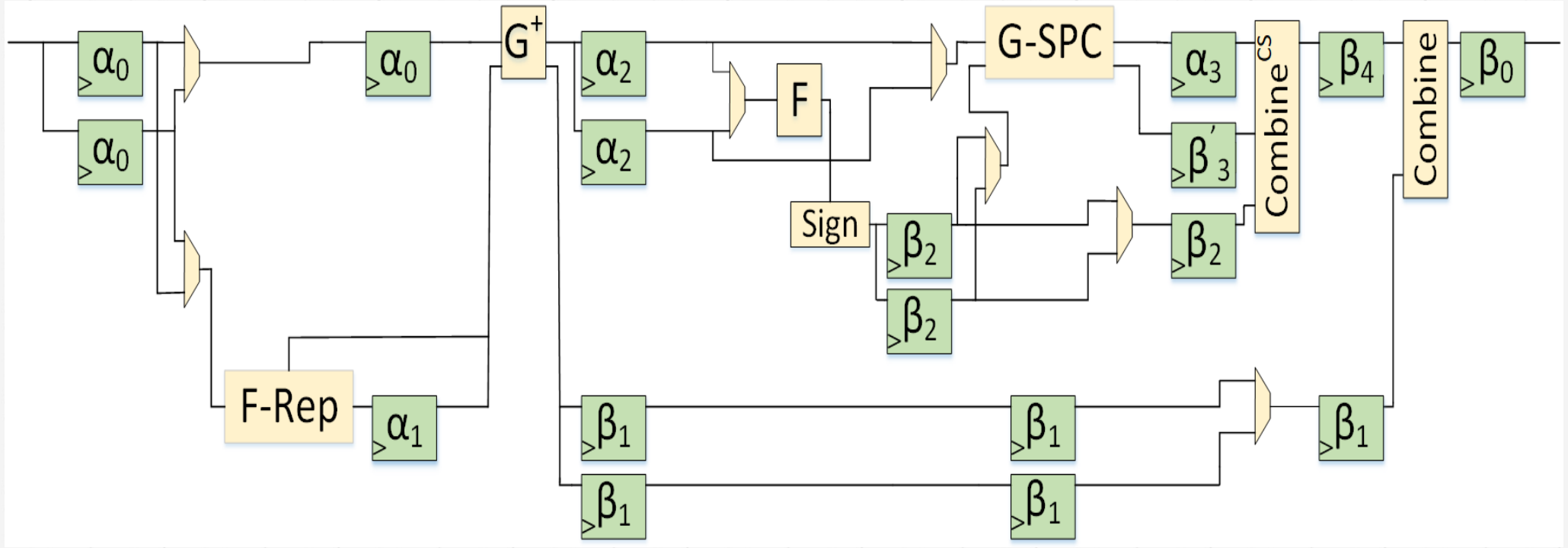
Fully-Unrolled Partially-Pipelined Architecture:

- To reduce the number of required registers, the Initiation Interval (II) is increased
- The throughput, as well as the number of registers become 1 / II
- However, the processing elements remain the same, which results in lower area efficiency



Reduced-Area Partially-Pipelined Architecture:

- **Because of the increased Initiation Interval (II), there is more time to prepare inputs to the next stages**
- **As a result, the number of processing elements in each block can be reduced**



Reduced-Area Partially-Pipelined Architecture:

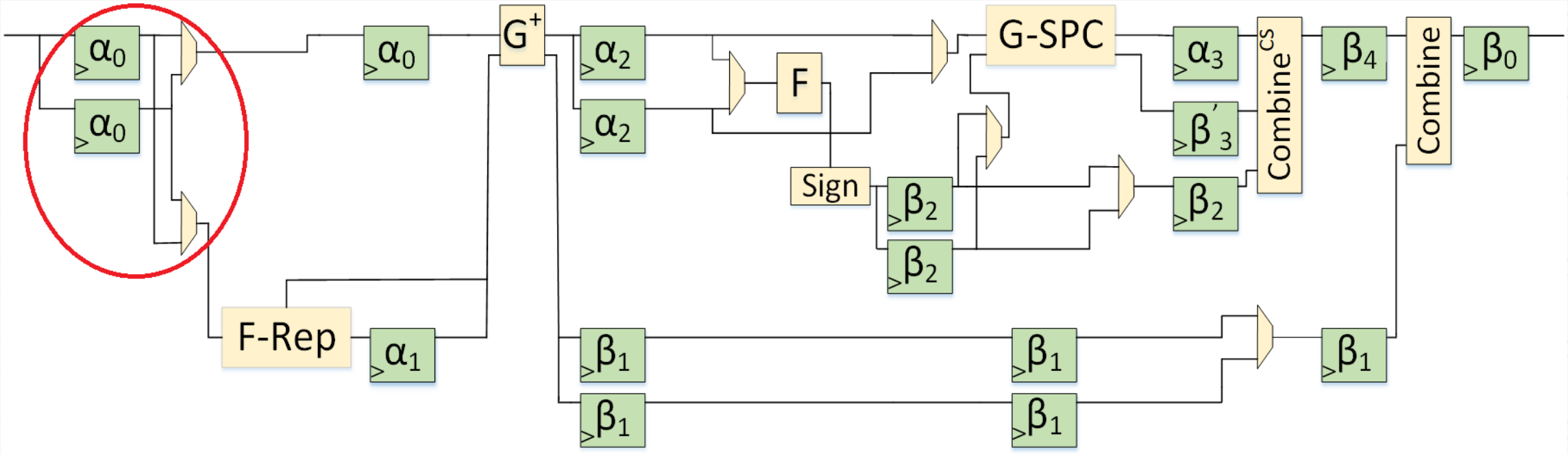
- **In F and G blocks, the number of processing elements can be halved**
- **These blocks can process odd indices at one clock cycle, and even indices at the next clock cycle**
- **This is because in the output vector of F and G operations, the elements with odd (even) index only depend on the inputs with odd (even) index**

$$\begin{aligned}\forall i \in [0, N_v/2) : \alpha_l[i] &= F\big(\alpha_v[i], \alpha_v[i+N_v/2]\big) \\ &= \text{sign}\big(\alpha_v[i]\big) \times \text{sign}\big(\alpha_v[i+N_v/2]\big) \\ &\times \min\big(\alpha_v[i], \alpha_v[i+N_v/2]\big)\end{aligned}$$

$$\begin{aligned}\forall i \in [0, N_v/2) : \alpha_r[i] &= G\big(\alpha_v[i], \alpha_v[i+N_v/2], \beta_l[i]\big) \\ &= \alpha_v[i+N_v/2] + (-1)^{\beta_l[i]}\alpha_v[i]\end{aligned}$$

Reduced-Area Partially-Pipelined Architecture:

- Each register in the conventional decoder is replaced with a pair of registers
- One for storing the elements with odd index, and one for storing elements with even index
- Multiplexers are required to choose between odd and even elements



Reduced-Area Partially-Pipelined Architecture:

- **One challenge:**
  - **Repetition (Rep) and Single Parity Check (SPC) nodes require the whole input to generate the output**

$$\forall i \in [0, N_x) : \beta_x[i] = \begin{cases} 0 & \text{if } \left(\sum_{j \in [0, N_x)} \alpha_x[j] \right) \geq 0 \\ 1 & \text{otherwise} \end{cases}$$

Repetition

$$j = \arg\min |\alpha_x[i]|$$

$$\text{parity} = \text{sign}(\alpha_x[0]) \oplus \cdots \oplus \text{sign}(\alpha_x[N_x - 1])$$

$$\beta_x[i] = \begin{cases} \text{sign}(\alpha_x[i]) \oplus \text{parity} & \text{if } i = j \\ \text{sign}(\alpha_x[i]) & \text{otherwise} \end{cases}$$

Single Parity Check

Reduced-Area Partially-Pipelined Architecture:

- It is possible to wait for an additional clock cycle for the output of SPC and Rep nodes
  - This will result in higher latency, and thus, more pipeline stages
- In order to avoid higher latency, SPC and Rep operations are merged with their previous operation
- The merge is performed without making the critical path longer

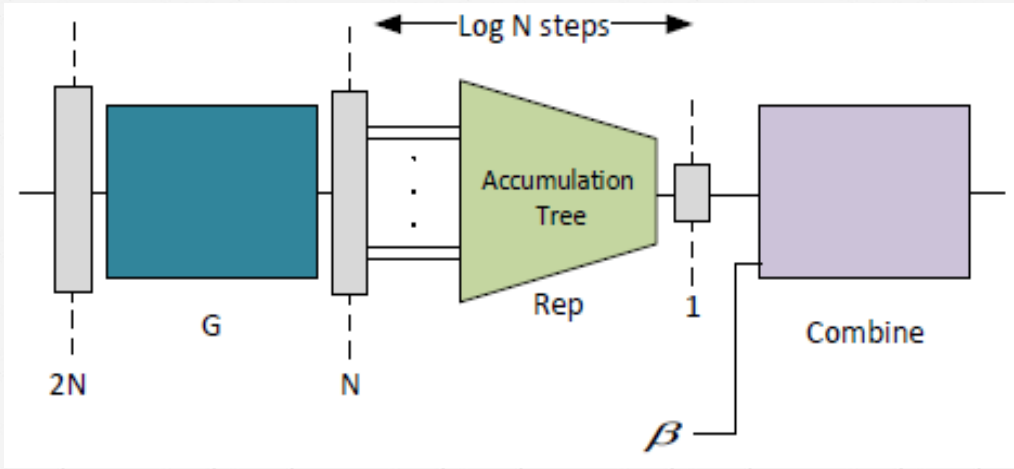
Reduced-Area Partially-Pipelined Architecture:

- **The only possible sequence of operations that contain Rep or SPC:**
  - **If the Rep node is the left child of its parent: F-Rep-G**
  - **If the Rep node is the right child of its parent: G-Rep-Combine**
  - **If the SPC node is the left child of its parent: F-SPC-G**
  - **If the SPC node is the right child of its parent: G-SPC-Combine**

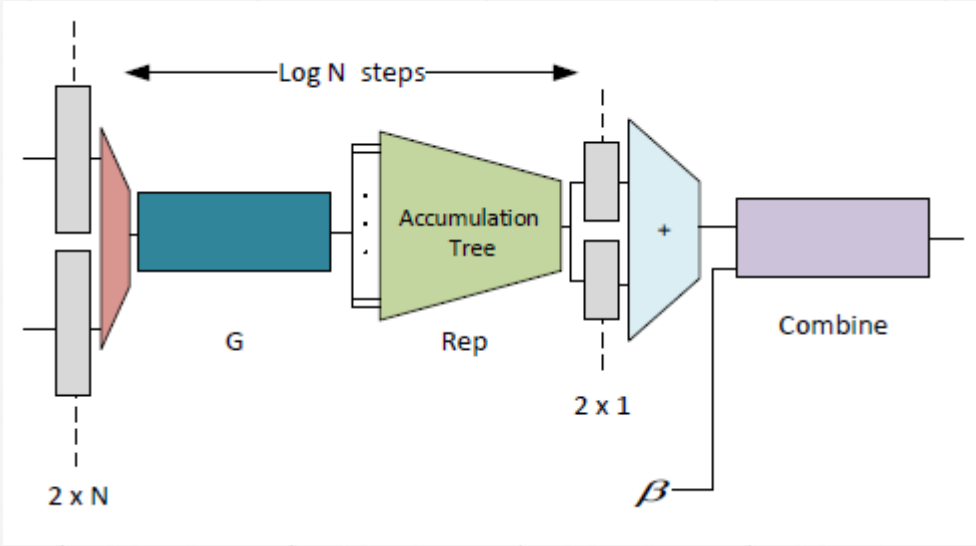
- **We show how to merge SPC or Rep with their previous operation without making the critical path longer**

Reduced-Area Partially-Pipelined Architecture:

- Optimizing G-Rep-Combine sequence:



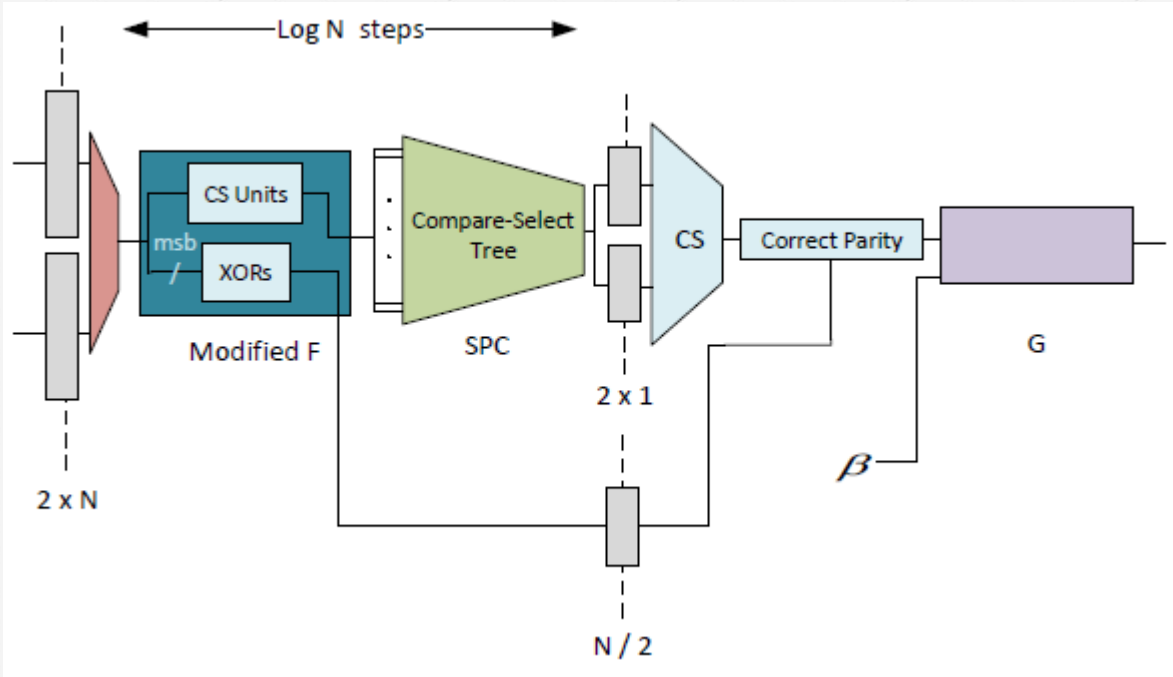
Conventional decoder



Area efficient decoder

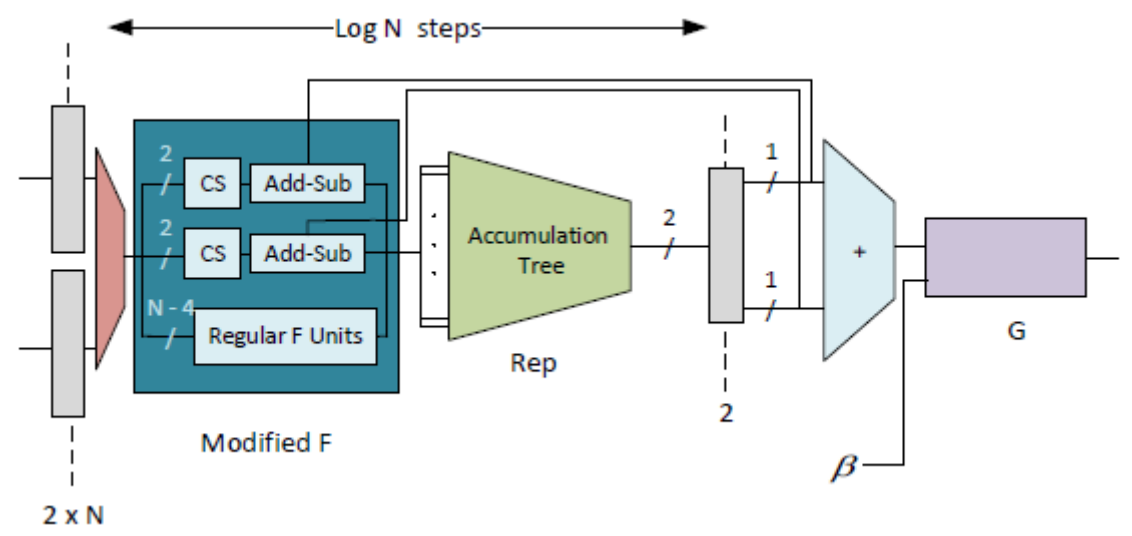
Reduced-Area Partially-Pipelined Architecture:

- Optimizing F-SPC-G sequence:



Reduced-Area Partially-Pipelined Architecture:

- Optimizing F-Rep-G sequence:



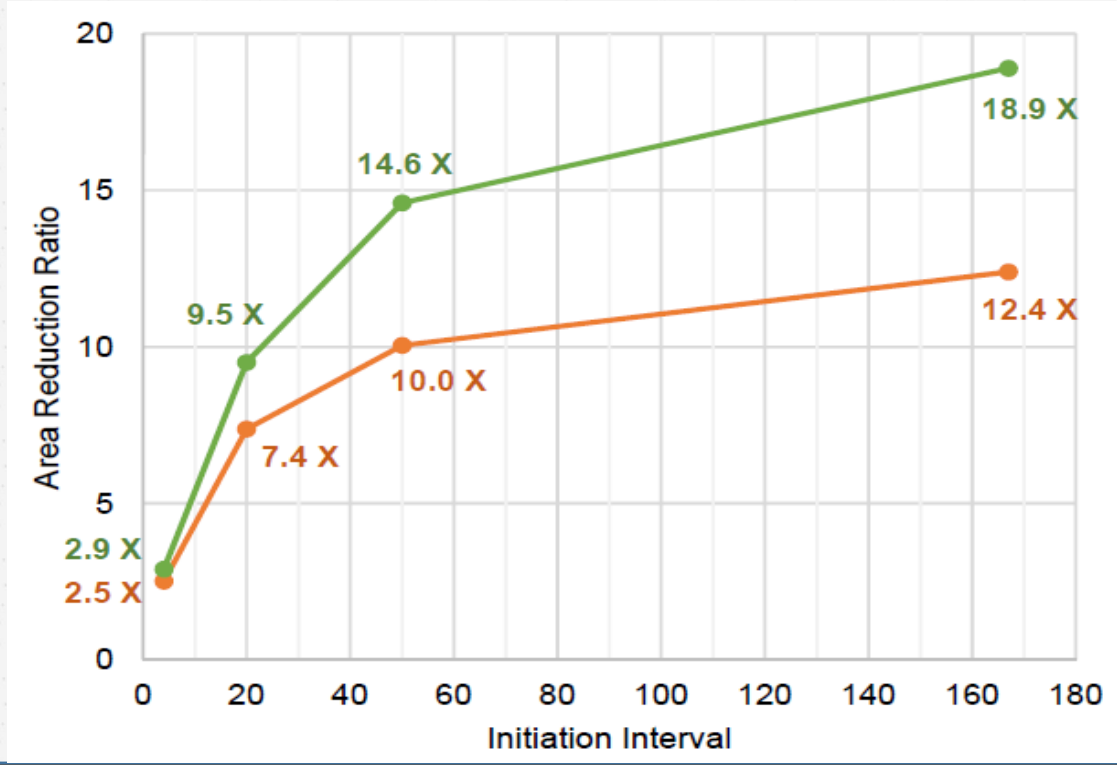
Reduced-Area Partially-Pipelined Architecture:

- Synthesis results:
  - 28% higher area-efficiency in comparison to the fully-unrolled partially-pipelined architecture

	AE-PPA	PPA [1]	[2]	[3]
Polar Code	(1024, 512)	(1024, 512)	(1024, 512)	(1024, k)
Decoding Algorithm	Fast-SSC	Fast-SSC	Fast-SSC	BP
Total Area (mm^2)	1.30	1.68	0.38	2.19
Coded Throughput (Gbps)	25.6	25.6	1.56	11.9
Initiation Interval	20	20	-	-
Frequency (MHz)	500	500	-	1164
Latency (microseconds)	0.73	0.73	0.66	0.08
Area Efficiency ($Gbps/mm^2$)	19.69	15.27	4.10	5.44

Reduced-Area Partially-Pipelined Architecture:

- Synthesis results:
  - Comparison of area-reduction in area-efficient partially-pipelined architecture (green curve) and conventional partially-pipelined architecture (red curve)
  - The area reduction is measured against the fully-pipelined architecture



References

[1]: P. Giard et al., “Multi-mode unrolled architectures for polar decoders,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 63, no. 9, pp. 1443–1453, 2016.

[2]: F. Ercan, T. Tonnellier, and W. J. Gross, “Energy-efficient hardware architectures for fast polar decoders,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 67, no. 1, pp. 322–335, 2020.

[3]:R. Shrestha, P. Bansal, and S. Srinivasan, “High-throughput and highspeed polar-decoder vlsi-architecture for 5g new radio,” in 32nd International Conference on VLSI Design and 18th International Conference on Embedded Systems (VLSID), 2019, pp. 329–334.

Questions?

Thank you for your attention!