

A 3.3-Gbps Bit-Serial Block-Interlaced Min-Sum LDPC Decoder in 0.13- μm CMOS

Ahmad Darabiha, Anthony Chan Carusone and Frank R. Kschischang
Department of Electrical and Computer Engineering, University of Toronto
Email: {ahmadd, tcc}@eecg.utoronto.ca, frank@comm.utoronto.ca

Abstract—A bit-serial architecture for multi-Gbps LDPC decoding is demonstrated to alleviate the routing congestion which is the main limitation for LDPC decoders. We report on a 3.3-Gbps 0.13- μm CMOS prototype. It occupies 7.3- mm^2 core area with 1416-mW maximum power consumption from a 1.2-V supply. We demonstrate how early termination and supply voltage scaling can improve the decoder energy efficiency. Finally, the same architecture is applied to a (2048, 1723) LDPC code compliant with the 10GBase-T standard.

I. INTRODUCTION

Low-density parity-check (LDPC) codes have been chosen for several emerging digital communication standards because of their excellent error correction performance and because they have the potential for high-throughput decoding using a highly-parallel iterative message-passing algorithm.

VLSI LDPC decoders can be categorized into two groups: memory-based and fully-parallel. Memory-based decoders [1], [2] communicate messages between shared processing units via a shared memory. Fully-parallel decoders directly instantiate all nodes of the code's Tanner graph in hardware. The focus of this work is on fully-parallel decoders because they provide higher decoding throughput.

The main challenge in implementing a fully-parallel decoder is the congestion that arises when routing interconnections between the processing nodes [3]. This routing congestion results in a large decoder with low area utilization, and poor timing and power performance due to the presence of long interconnects across the chip. These problems are exacerbated because word lengths of 4 to 8 bits are generally required to represent each message, so each graph edge requires 4 to 8 parallel wires.

To address this problem, in this work we report on a fully-parallel architecture in which the multi-bit extrinsic messages are transferred bit-serially between the processing units. Bit-serial message passing reduces node-to-node interconnections since each edge requires just a single wire. Furthermore, the MIN and SUM functions in iterative min-sum decoding have natural hardware-efficient bit-serial implementations. The result is smaller circuits (compared to the conventional fully-parallel scheme) and less global wiring. These advantages become more pronounced for longer word lengths and in deep submicron scaled CMOS. Finally, a bit-serial scheme permits efficient variable word length decoding [4].

In this paper we report on a 3.3-Gbps min-sum bit-serial LDPC decoder fabricated in a 0.13- μm CMOS process. This

decoder has the highest throughput among LDPC decoders reported in the literature to date. We will also show how the same architecture can be used to realize a decoder for the (2048, 1723) LDPC code in IEEE 10GBase-T standard. We will also discuss techniques to reduce power consumption by terminating the decoder's iterations early, and by supply voltage scaling.

After a brief introduction to LDPC codes and decoding schemes in Section II, Section III describes the bit-serial message-passing scheme and the interlaced scheduling used in this work. Section IV reports the measurement results for the implemented (660, 480) LDPC decoder. It also discusses how the same architecture is applicable to the (2048, 1723) LDPC code compliant with the 10GBase-T standard. Section V concludes the paper.

II. BACKGROUND

LDPC codes are a sub-class of linear error control codes with a sparse parity-check matrix, H . These codes can also be described by a bipartite graph, or Tanner graph, in which check nodes $\{c_1, c_2, \dots, c_C\}$ represent the rows of H and variable nodes $\{v_1, v_2, \dots, v_V\}$ represent the columns. An edge connects the check node c_m to the variable node v_n if and only if matrix entry H_{mn} is nonzero.

LDPC codes may be decoded with iterative message-passing algorithms such as the sum-product or min-sum algorithms. In these algorithms, each decoding iteration consists of a variable node update phase followed by a check node update phase. At the end of each phase, the updated messages are communicated between neighboring nodes in the code graph. Each message represents the decoder's *belief* about the value of a received bit, referred to as a log-likelihood ratio (LLR). Fig. 1 shows a fully-parallel LDPC decoder comprising V variable node update units (VNUs) and C check node update units (CNU). The VNUs and CNU are interconnected with hard wires based on the LDPC code's Tanner graph. The number of edges in the graph is $E = Vd_v = Cd_c$ where d_v is the average number of check nodes connected to each variable node and d_c is the average number of variable nodes connected to each check node. If LLR values are represented using q bits, a total of $2qE$ node-to-node wires are required to convey all check-to-variable and variable-to-check messages. For example, a decoder for the standard 10GBase-T LDPC code using 4-bit LLR messages requires 98304 hard wires.

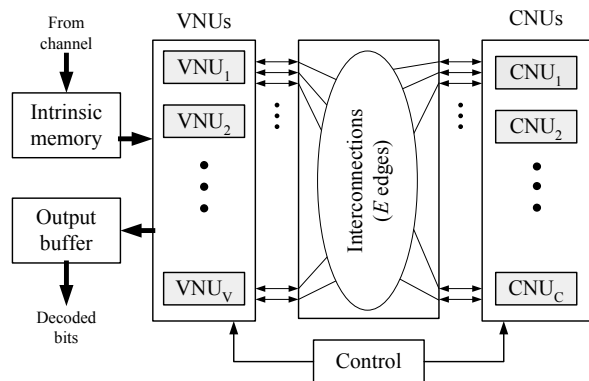


Fig. 1. A fully-parallel LDPC decoder.

III. BIT-SERIAL DECODING

To alleviate the routing problem caused by the large number of global wires in a fully-parallel LDPC decoder, in this section we describe a bit-serial message-passing scheme. Fig. 2(a) shows a conventional decoder where q -bit messages are calculated and communicated between VNUs and CNUs in parallel. Alternatively, Fig. 2(b) shows the proposed decoder in which messages are generated and transferred between nodes bit-serially over one wire in q clock cycles. An immediate advantage of bit-serial LDPC decoding is reduced routing complexity. In addition, since both the MIN and SUM operations in min-sum decoding can be naturally performed bit-serially, efficient VNU and CNU hardware implementations are possible. Since only 1-bit operations are performed in each clock cycle, the bit-serial circuits have shorter critical paths and can, therefore, operate at higher clock frequency thus providing a throughput similar to the conventional parallel implementation.

Figure 3 shows the schematic for a bit-serial CNU based on the approximate min-sum algorithm proposed in [5]. The CNU inputs arrive bit-serially in sign-magnitude MSB-first format. The d_c -input AND gate and the surrounding logic calculate the output magnitude and the XOR gates at the bottom calculate the output signs.

The SUM function in the VNUs is most efficiently performed on inputs in 2's-complement LSB-first format. Additional hardware is therefore required to convert between the sign-magnitude MSB-first and 2's-complement LSB-first formats. In our design the format conversion imposes very little logic overhead because the added storage registers are also used to facilitate block-interlacing.

Fig. 4 shows the timing diagram of a block-interlaced decoder, where two successive frames are simultaneously decoded. At any given time, the messages for one frame are in the VNUs and messages for the other are in the CNUs. After every q clock cycles, the VNUs and CNUs swap messages. This effectively pipelines the iterative decoding, providing a two-fold increase in throughput compared with non-interlaced designs such as [3].

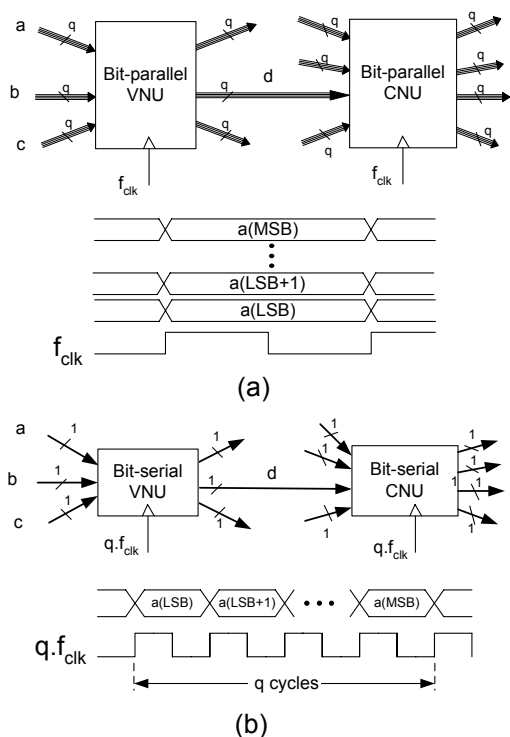


Fig. 2. Bit-serial vs. bit-parallel message passing.

IV. IMPLEMENTATIONS

The high-level architecture of the decoders reported in this section is shown in Fig. 1. All message updates and message transfers are performed bit-serially.

A. A (660, 480) LDPC decoder

A (4,15)-regular LDPC code was constructed using a progressive edge-growth algorithm [6] to minimize the number of short cycles in the code's Tanner graph. The block length was limited to 660 by the silicon area available for prototyping. The decoder employs a modified min-sum algorithm that decreases the total logic by about 20% with only a 0.2 dB degradation in error performance compared to conventional min-sum decoding.

The decoder was fabricated in a 0.13- μ m CMOS8M process and was tested using an Agilent 93000 high-speed digital tester. The 4-bit channel LLR values are input to the decoder via 44 input pins and the decoded outputs are read out via 44 output pins. With a 1.2-V supply, the measured maximum clock frequency is 300 MHz. The decoder performs 15 decoding iterations per frame (beyond which decoding performance increases negligibly) resulting in a total encoded data throughput of 3.3 Gbps. This throughput could be further increased by reducing the number of iterations, at the cost of slightly reduced decoding performance.

Fig. 6 shows the measured BER performance of the decoder. The results match bit-true simulation results. The figure also plots the decoder's power consumption with a 1.2-V supply

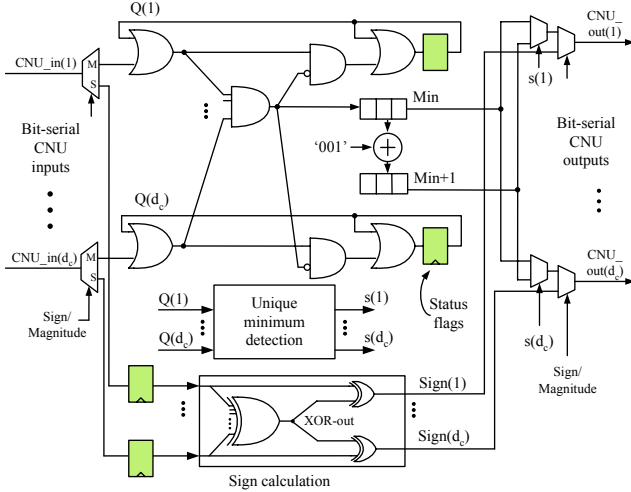


Fig. 3. Bit-serial CNU schematic.

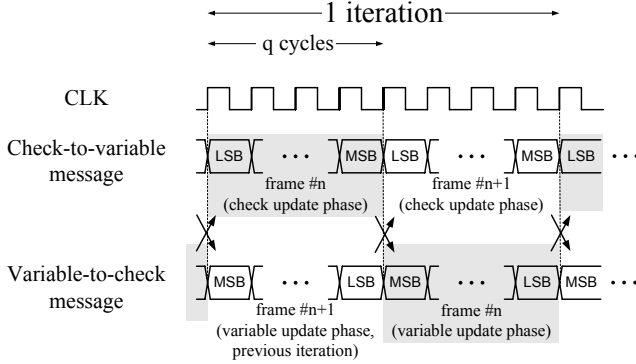


Fig. 4. Timing diagram for block-interlaced bit-serial decoding.

and 300-MHz clock frequency as a function of input SNR. It is observed that the power consumption peaks at SNRs where the iterative decoder struggles to converge, thus resulting in a slightly higher switching activity. However, the correlation between power consumption and input SNR is lower than in conventional fully-parallel decoders because block-interlacing and bit-serial message-passing techniques tend to maintain high switching activity on all nodes, even at high SNRs. Measurements show that approximately 20% of the total power dissipation is due to the clock tree. The static power consumption due to leakage current accounts for less than 1% of the total decoder power.

Fig. 5 shows the effect of scaling the decoder's supply voltage on its maximum clock frequency and the corresponding power consumption. The plot suggests that for lower-throughput applications, energy efficiency can be improved using the same bit-serial decoder but operating at a lower supply voltage [7]. For example, with a 0.6-V supply, the design has a measured energy efficiency of 0.148 nJ/bit –

better than previously reported designs specifically targeting low power [8].

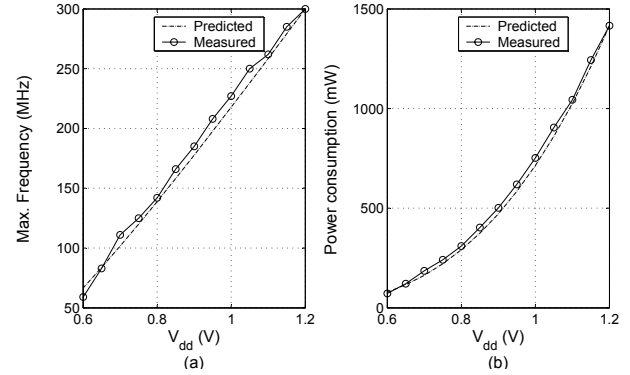


Fig. 5. Frequency and power dissipation vs. supply voltage.

Although a maximum of 15 iterations per frame is required for excellent low BER performance, the vast majority of frames need only few decoding iterations to be correctly decoded. As a result, a significant power saving can be achieved by detecting early convergence in the decoder and turning it off for the remaining iterations. This early termination feature is not included in the currently fabricated decoder but, synthesis results show that it adds less than 1% overhead in logic area while providing a large power saving as shown in Fig. 6. The values in this graph are obtained from laboratory power measurements of the prototype decoder (without early termination) combined with simulation results of the average required number of iterations per frame at different input SNRs. The plot shows that for the 660-bit LDPC decoder, early termination results in more than 60% power savings at low input SNRs and more than 70% savings at high input SNRs. There is no change in the decoder's BER performance.

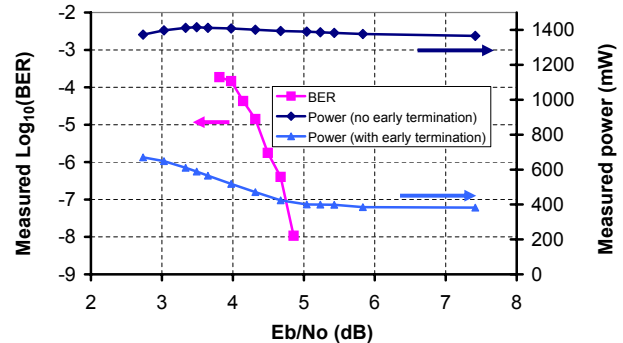


Fig. 6. Measured power and BER for the decoder in Section IV.A.

The decoder die photo is shown in Fig. 7. Table I summarizes the measured results. Fig. 8 compares the throughput and power efficiency of the decoder in this work with CMOS LDPC decoders reported in [1]–[3]. To take into account the

TABLE I

SUMMARY OF MEASURED RESULTS FOR THE (660, 480) LDPC DECODER.

Process	0.13- μ m CMOS	
LDPC Code	(4, 15)-regular 660-bit	
Code rate	0.73	
Decoding algorithm	Modified min-sum	
Core area /Total area	7.3mm ² /9mm ²	
Gate count	690 k	
Core area utilization	72%	
Iterations per frame	15	
Message word length	4 bits	
Supply	1.2 V	0.6 V
Maximum frequency (MHz)	300	59
Total throughput (Mbps)	3300	648
Information throughput (Mbps)	2440	480
Power @ (Eb/No=4dB) (mW)	1408	72
Power @ (Eb/No=5.5dB) (mW)	1383	71
Energy efficiency @ (Eb/No=4dB) (nJ/bit)	0.577	0.150
Energy efficiency @ (Eb/No=5.5dB) (nJ/bit)	0.566	0.148

varying number of iterations per frame and code rates in different implementations, the vertical axis is the information throughput normalized with respect to the number of iterations per frame. It can be seen that our implementation has the highest throughput and the second lowest energy consumption per information bit, even without early termination or supply voltage scaling. The figure shows how supply voltage scaling can be used to trade the throughput for even lower energy consumption per decoded information bit. It also shows the power saving achievable from the early termination.

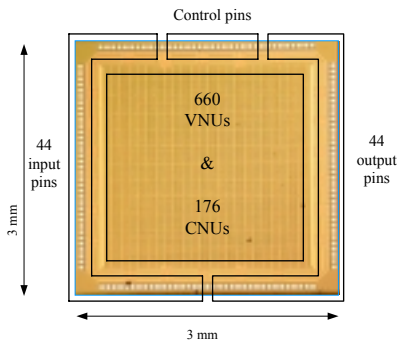


Fig. 7. Die photo of the (660, 480) LDPC decoder in Section IV.A.

B. (2048, 1723) LDPC decoder

Using the same bit-serial architecture, a decoder was synthesized for the (6,32)-regular (2048, 1723) LDPC code in the 10GBase-T 10-Gbps Ethernet standard. Again, $q=4$ bit quantization was used for LLR messages. In this decoder, the maximum number of iterations per frame is 8 since our simulations indicate that additional iterations provide less than 0.1-dB performance improvement.

Synthesis in a 90-nm CMOS library using Synopsys Design Compiler results in 9.8-mm² logic area (2.23M equivalent

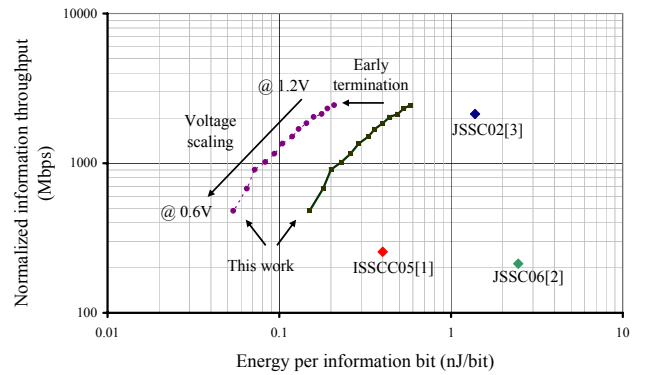


Fig. 8. Comparison to previously published work.

NAND Gates) and 250 MHz maximum clock frequency corresponding to 16-Gbps maximum decoding throughput. This throughput is significantly higher than that required by the 10GBase-T standard.

V. CONCLUSION

In this paper we have reported a bit-serial message-passing decoding architecture to alleviate the routing congestion in fully-parallel LDPC decoders. First, we presented the measurement results for a 3.3-Gbps bit-serial fully-parallel (660, 480) LDPC decoder fabricated in a 0.13- μ m CMOS process. The decoder has a higher throughput than any previously reported LDPC decoder and has the second best energy efficiency per decoded information bit. We showed how early termination and supply voltage scaling can further improve the decoder energy efficiency. Finally, we demonstrated how the same bit-serial architecture can be applied to design a decoder for the (2048, 1723) 10GBase-T LDPC code.

Acknowledgment: The authors thank Gennum Corporation, Canada, for supporting this work.

REFERENCES

- [1] H.-Y. Liu and et al, "A 480mb/s LDPC-COFDM-Based UWB baseband transceiver," in *IEEE Int. Solid-State Circuits Conference*, 2005.
- [2] M. M. Mansour and N. R. Shanbhag, "A 640-Mb/sec 2048-bit programmable LDPC decoder chip," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 3, Mar. 2006.
- [3] A. J. Blanksby and C. J. Howland, "A 690-mW 1-Gb/s 1024-b, rate-1/2 low-density parity-check decoder," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 3, pp. 404–412, Mar. 2002.
- [4] M. Ardakani and F. R. Kschischang, "Gear-shift decoding," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 1235–1242, July 2006.
- [5] A. Darabiha, A. Chan Carusone, and F. R. Kschischang, "A bit-serial approximate Min-Sum LDPC decoder and FPGA implementation," in *International Symposium on Circuits and Systems*, Kos, Greece, May 2006.
- [6] X.-Y. Hu, E. Eleftheriou, and D. M. Arnold, "Regular and irregular progressive edge-growth Tanner graphs," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 386–398, Jan. 2005.
- [7] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-power CMOS digital design," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 4, pp. 473–484, Apr. 1992.
- [8] S. Hemati, A. H. Banihashemi, and C. Plett, "A 0.18- μ m CMOS analog min-sum iterative decoder for a (32,8) low-density parity-check (LDPC) code," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 11, pp. 2531–2540, Nov. 2006.