Latency-Optimized Stochastic LDPC Decoder for High-Throughput Applications

Di Wu, Yun Chen, Qichen Zhang, Lirong Zheng, Xiaoyang Zeng
State Key Laboratory of ASIC & system
Fudan University
Shanghai, China
{wud12, chenyun@fudan.edu.cn}

Yeong-luh Ueng
Department of Electrical Engineering
National Tsing Hua University
Taiwan
ylueng@ee.nthu.edu.tw

***Abstract*—Stochastic decoding can be applied to Low-Density Parity-Check codes in order to achieve high throughput with less area. However, most architectures suffer from large decoding latencies, due to the mechanism of stochastic computation. In this paper, three novel strategies, including the LUT-based initialization, the posterior-information-based hard decision and the Bit-Flipping-based post processing, are proposed in order to reduce decoding latency and hence improve throughput. For the standard IEEE 802.3an (2048, 1723) code, simulation indicates 75.7% reduction in average decoding cycles at 4.5 dB with satisfied bit error rate. Moreover, hardware implementation shows that the area of variable node units is reduced significantly in SMIC 65 nm technology.**

Keywords—Bit-Flipping Algorithm; Low-Density Parity-Check Codes; High-throughput Decoder; Stochastic decoding

I. INTRODUCTION

After the rediscovery of Low-Density Parity-Check (LDPC) codes based on Tanner graph, which were first introduced by Gallager in 1960s [1], LDPC codes are adopted in various communication standards, such as the IEEE 802.3an-2006 (10GBASE-T) standard [2]. With the increasing demand for higher throughput in recent years, fully-parallel decoder architectures for LDPC codes are desired. However, with the parallelism of a bit-parallel message-passing LDPC decoder augmenting, the wiring overhead increases enormously [3], limiting hardware clock frequency and chip size, constituting the bottleneck for the next generation communication systems with both high throughput and high reliability.

In recent years, stochastic decoding, with the single-bit-transmission characteristic, similar to the bit-serial decoders, efficiently reducing the wiring congestion, provides a promising alternative for optical communication systems which require throughput over 100Gb/s. Gaudet and Raley first adopted stochastic computation in LDPC decoding in [4]. In the following researches, LDPC decoders based on stochastic computation with Edge Memory (EM), Tracking-Forecast Memory (TFM), Majority-based Tracking-Forecast Memory (MTFM) and Sliding Window Method (SWM) were proposed in [5], [6], [7] and [8], respectively, utilizing different methods to improve the performance.

In terms of the data transmission in stochastic decoding, the latency of the decoder presented in [7] is 800ns, much larger than Min-Sum (MS) decoders on average [9]. For the objective of higher throughput, research challenges focus on reducing the average decoding cycles of stochastic decoders. In [10], the proposed decoder initializes the first bit transmitted to the hard decision unit with the sign bit of the channel Log Likelihood Ratio (LLR) in a (48, 24) LDPC code, while no demonstration for long codes is given. In [11], new hard decision method is proposed, reducing the entire decoding period by a few cycles.

In this paper, we propose a novel architecture with three strategies, aiming at the reduction of decoding cycles (DCs) of stochastic LDPC decoders. First, a new initialization method based on the LUT and counter is proposed, achieving the same Bit Error Rate (BER) as the EM based stochastic decoder. Second, a modern hard decision strategy is proposed, making use of the posterior information rather than the soft messages from VN (Variable Node) to CN (Check Node). Third, a simple post processing based on a modified Bit-Flipping (BF) algorithm is combined with the stochastic processing in order to help the stochastic LDPC decoders converge rapidly when the number of bit errors is less than a specific value. For the standard IEEE 802.3an (2048, 1723) code, simulation indicates 75.7% reduction in average decoding cycles at 4.5 dB with satisfied bit error rate. Moreover, hardware implementation shows that the area of VN units is reduced significantly in SMIC 65 nm technology.

The rest of this paper is presented as follows. Computation mechanism of the stochastic LDPC decoders is reviewed in Section II. The proposed architecture is depicted in detail in Section III. Simulation results and hardware implementation results are shown in Section IV. Finally, some conclusions are summarized in Section V.

II. STOCHASTIC DECODING

A. Algorithm

The stochastic LDPC decoding is deduced from the Sum-Product Algorithm (SPA). The SPA is an iterative decoding algorithm, exchanging data in the probability domain between VNs and CNs. Thus correct code words can be found after a certain number of iterations.

The stochastic decoding procedure between a degree-2 VN and a degree-3 CN acts similarly with the SPA. The difference is that the computation in stochastic decoders is done in each

Sponsored by the National Natural Science Foundation of China (Grant No. 61306043), the Shanghai Natural Science funding (Grant No. 13ZR140-1800). National Science and Technology Major Project of China (Grant No. 2011ZX03003-003-03), Natural Science Foundation of China (Grant No. 61234002).

clock cycle, but not in each iteration. Stochastic decoding is described as follows:

1) Initialization: Initialize each VN v_i with probability P_{init}^{i} calculated with channel LLR L_i according to (1).

$$P_{i\to j} = P_{init}^{i} = e^{L_i} / (e^{L_i} + 1) \tag{1}$$

2) CN update: Each VN v_i sends soft message $P_{i->j}$ to corresponding CN c_j, and CN message $Q_{j->i}$ updates as (2).

$$Q_{j\to i} = P_{l\to j}(1-P_{m\to j}) + P_{m\to j}(1-P_{l\to j}) \tag{2}$$

3) VN update: Each CN c_j sends soft message $Q_{j->i}$ to corresponding VN v_i, and VN message $P_{i->j}$ updates as (3).

$$P_{i\to j} = \frac{Q_{r\to i}Q_{s\to i}}{Q_{r\to i}Q_{s\to i} + (1-Q_{r\to i})(1-Q_{s\to i})} \tag{3}$$

4) Termination: Stop decoding if the decoded code word satisfies all the parity check constraints or the decoder reaches the maximum iteration. Otherwise, go back to *2)*.

B. Stochastic Decoder Elements

Messages in a stochastic decoder are represented in probability domain as Bernoulli sequences, in which the ratio of 1s equals to the probability. Schematics of the CN and VN function ((2) and (3)) are shown in Fig. 1 and Fig. 2, respectively.

The CN function is executed by a XOR gate, where two input Bernoulli sequences are fed to the XOR gate. Assumptions such as uniform distribution, adequate computation cycles and no relevance between two input Bernoulli sequences guarantee negligible errors between the results of Fig. 1 and (2).

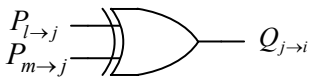

Fig. 1. Structure of a degree-3 CN unit

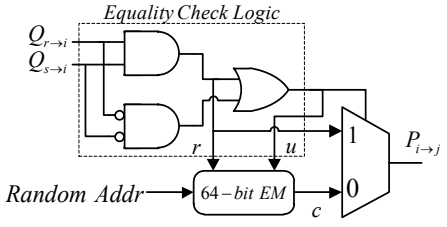

Fig. 2. Structure of a degree-2 VN unit

The structure of a VN unit with 64-bit EM is shown in Fig. 2, where the dashed rectangle is the Equality Check Logic (ECL). When the two inputs are the same ($u = 1$), the EM is updated with the regenerative bit r and the output equals to the input, corresponding to the non-hold state. When one input differs from the other ($u = 0$), the EM remains the same and the output equals to the conservative bit chosen from the EM according to a 6-bit random address, corresponding to the hold state. A degree-6 VN can be composed with two degree-3 subunits and one degree-2 subunit as mentioned in [5]. Here the degree-2 subunit utilizes the 2-bit internal memory instead of EM to reduce the number of hold states, and the degree-3 subunit utilizes the EM.

III. Proposed Decoder Architecture

Fig. 3 shows the proposed decoder architecture, where three optimizing strategies are introduced to solve the large latency problem in terms of three different aspects during decoding.

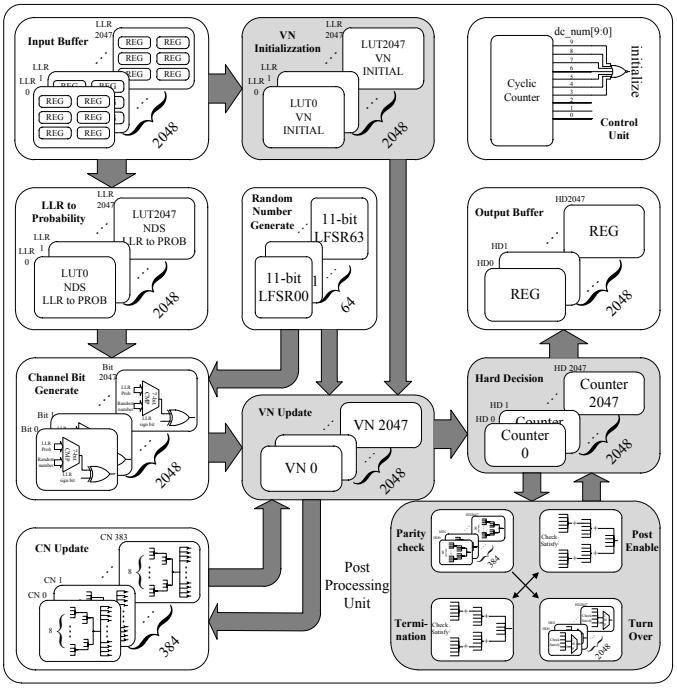

Fig. 3. Proposed Decoder Architecture

The four gray blocks in Fig. 3 are the modified modules in the proposed architecture. The rest white modules are the same as those in [5], [6], [7] and [8]. A LUT-based initialization procedure is employed in the VN initialization unit and the VN update unit, a posterior-information-based hard decision method is employed in the VN update unit, and a BF-based post processing procedure is employed in the hard decision unit and the post processing unit.

A. LUT-based Initialization

For a practical EM-based LDPC stochastic decoder, thousands of VN units are used and the EM occupies 61% area of a VN unit, leading to a large chip size. To reduce the area of an EM-based VN unit, TFM, MTFM and SWM structure were proposed. In this paper, we proposed a counter-based VN structure, with which the LUT-based one-step initialization and the posterior information based hard decision strategy can be adopted.

The proposed VN subunit is given in Fig. 4. Compared to the EM based VN unit, the 64-bit EM is replaced by a 7-bit

up/down saturation counter, which can be initialized according to the initialization signal generated by the control unit in Fig. 3 and the initialization information generated by the VN initialization unit described in the next section. Both the 7-bit up/down saturation counter and the TFM used in [5] record the regenerative bits in the past DCs in a linear mapping. However, the difference between them is that the internal data are stored in integer and probability domain in the counter and the TFM, respectively.

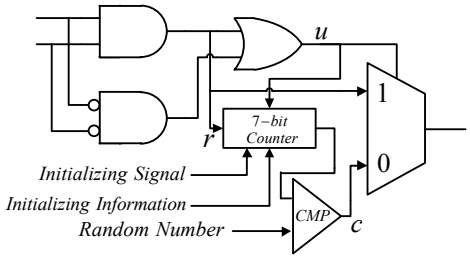

Fig. 4. Modified degree-2 subnode

In [5], an initialization procedure is recommended in EM structure to reduce the total DCs. At high SNRs, this stage prevents decoder from terminating in less total DCs. In order to reduce the length of initialization period, a one-step initializing structure based on LUTs is proposed. When the initializing signal is 1, the counter is initialized with the initializing information, which is related to the theoretical number of 1s in EM. The following comparator compares the random number and the counter value to generate the conservative bit as the output is in hold state. If the random number is smaller, the output is 1, otherwise 0.

Considering the hardware implementation, LUT is applied to each VN. For the proposed decoder, channel LLRs are quantized to 6-bit signed binary with 3-bit fraction, and the corresponding probabilities are 7-bit binary. These LUTs use channel LLRs as inputs, similar to the LLR-to-probability unit, and output certain numbers for initializing the VN counters.

A 7-bit counter-based VN unit has a similar performance as a 64-bit EM. Provided the number of 1s in the 64-bit EM is $x \in [0,64]$, the number of 0s is $64-x$, and the number inside the counter is $64+x-(64-x)=2x$.

B. Posterior-Information-Based Hard Decision

The hard decision unit makes use of the up/down saturation counters. To enable the post processing and the one-step initialization, the counters are modified. As to the one-step initialization, these counters are initialized as well, and the initial value of the hard decision counter is 1 ($LLR \geq 0$) or -1 ($LLR < 0$). In the post processing period, the posterior information bits stop updating the hard decision unit. Alternatively, a BF algorithm is employed. During this period, bits in a counter remain the same, except the sign bit. The sign bit during the post processing period and the post processing unit constitutes a modified BF decoder, helping to accelerate converge in much less DCs.

In the hard decision unit, the posterior information rather than the VN to CN soft messages can be used in stochastic decoders to improve performance. As shown in Fig. 5, one additional ECL is adopted in the whole VN unit. The output from a certain VN unit to its corresponding hard decision unit is the posterior bit. If the information bits from and to a certain CN (e.g. in_0 and out_0 in Fig. 5) are the same, the update signal of the proposed VN unit is set 1, and the hard decision unit is updated according to the posterior bit. Otherwise, the corresponding counter value in the hard decision unit remains the same in the next DC.

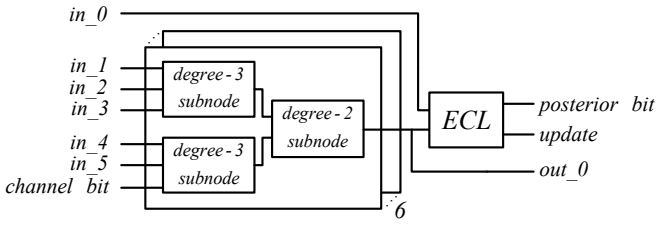

Fig. 5. Posterior bit for Hard Decision

C. BF-based Post Processing

However, simulation result shows that when the number of bit errors in a stochastic LDPC decoder is less than a certain amount, decoding process slows down due to the decrease in switching activity. To avoid this period, a BF-based post processing is adopted.

1) Parity Check Unit

The parity check unit is similar to the CN_UPDATE in Fig. 3, executing 384 parity checks and outputting the 384 results to the other three units.

2) Post Enable Unit

Consider statistics that when the number of bit errors is less than 7, the number of unsatisfied parity check equations is usually less than 25 for the 10GBASE-T LDPC code, the post enable unit adds all 384 parity check results in a 3-stage pipeline. If the result is less than 30, set the post enable signal to 1. Here the threshold 30 is chosen due to the pipeline.

3) Termination Unit

The termination signal is 1 if all the parity check results are 0. As the sum signal is given in the post enable unit, it can also be used to verify the termination. Termination signal is 1 if the sum equals to 0 in the post processing period.

4) Turn Over Unit

The turn over unit sums up all the 6 parity check results corresponding to each VN. If the sum is larger than 3, the hard decision register is turned over, different from BF algorithm, which turns over the bit with the most parity check errors. The proposed decoding procedure will terminate in less than 3 DCs after turning into the post processing state in most situation, and return back to the normal state if it does not terminate.

IV. Performance Evaluation

A. BER Performance

Fig. 6 shows the BER for the (2048, 1723) 10GbASE-T code, where AWGN channel and BPSK modulation are considered. The BER comparison among EM, EM with proposed hard decision, EM with both the proposed the hard

decision and post processing, the proposed architecture and the Normalized MS (NMS) decoders indicates that the proposed strategies can improve the error-rate performance. At low SNRs, the proposed design and the NMS float algorithm achieve a similar performance. At high SNRs, a 0.15 dB loss in coding gain is observed at the BER of 1.0e-6.

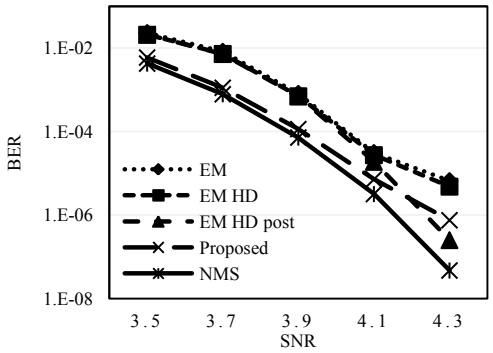

Fig. 6. BER performance

B. Average Decoding Cycles (ADCs) Reduction

Fig. 7 shows the reduction in ADCs when using the proposed three strategies progressively at the SNR of 4.5 dB. The proposed architecture saves 75.7% ADCs, leading to more than four times increase in throughput. It can be seen from Fig. 8 that the three proposed strategies can significantly reduce the ADCs at all different SNRs.

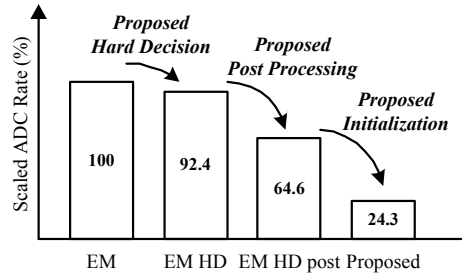

Fig. 7. ADC Reduction Rate

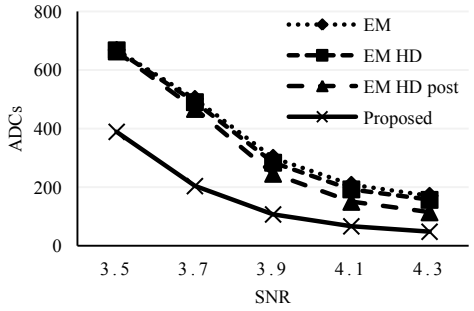

Fig. 8. Average Decoding Cycles Reduction

C. Hardware Implementation

The proposed stochastic decoder is synthesized in SMIC 65 nm technology. Result and comparison are given in TABLE I. At 5.5 dB, only 12.4 ADCs is needed for the proposed design. Compared with [5], significant area reduction (70%) in VNs is achieved. However, VNs still account for 48.4% of total area.

TABLE I. HARDWARE IMPLEMENTATION RESULT

Module	Parameter			
	Scaled Area	*Frequency*	*Code or Degree*	*Throughput*
Proposed	2.75mm^3	500MHz	RS based (2048, 1723)	82.58Gb/s @5.5dB
EM [5]	46097 Slices	222MHz	(1056,528)	1.66Gb/s @4.25dB
MTFM [7]	3.19mm^3	500MHz	RS based (2048, 1723)	61.3Gb/s @5.5dB
Proposed VN	1863.7um^2	500MHz	Degree-6	-
EM VN	6216.5um^2	500MHz	Degree-6	-
TFM VN [6]	5505.5um^2	-	Degree-6	-

V. CONCLUSION

We have presented three strategies, including the LUT-based initialization, the posterior-information-based hard decision and the BF-based post processing, to reduce the latency for stochastic LDPC decoders. Simulation shows that the average decoding cycles are reduced by 75.7% at the SNR of 4.5 dB in comparison with the original EM decoder in [7]. The proposed decoder achieves 82.58 Gb/s at the SNR of 5.5 dB, increasing by 34.7% compared with the MTFM decoder in [7]. Hardware implementation in SMIC 65 nm technology shows that the chip area is reduced due to the VN area reduced by 70% and the area of additional modules been ignored.

REFERENCES

[1] R. G. Gallager, *Low Density Parity Check Codes*. Cambridge, MA: MIT Press, 1963.

[2] The IEEE P802.3an 10GBASE-T Task Force [Online]. Available: www.ieee802.org/3/an

[3] Di Wu, Yun Chen, Yuebin Huang, Yeongluh Ueng, Lirong Zheng and Xiaoyang Zeng, "A high-throughput LDPC decoder for optical communication," *IEEE 10th International Conference on ASIC (ASICON)*, 2013, pp. 1-4.

[4] V. Gaudet and A. Rapley, "Iterative decoding using stochastic computation," *Electron. Lett.*, vol. 39, no. 3, pp. 299–301, Feb. 2003.

[5] S. Sharifi Tehrani, S. Mannor, and W. J. Gross, "Fully parallel stochastic LDPC decoders," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5692–5703, Nov. 2008.

[6] S. Sharifi Tehrani, A. Naderi, G.-A. Kamendje, S. Mannor, and W. J. Gross, "Tracking forecast memories in stochastic decoders," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 561–564.

[7] S. Sharifi Tehrani, A. Naderi, G.-A. Kamendje, S. Hemati, S. Mannor, and W. J. Gross, "Majority-based tracking forecast memories for stochastic LDPC decoding," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4883–4896, Sep. 2010.

[8] Jienan Chen and Jianhao Hu, "Sliding Window Method for stochastic LDPC decoder," *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2014, pp.1307-1310.

[9] Cheng Chung-Chao, Yang Jeng-Da, Lee Huang-Chang, Yang Chia-Hsiang, and Ueng Yeong-Luh, "A Fully Parallel LDPC Decoder Architecture Using Probabilistic Min-Sum Algorithm for High-Throughput Applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol.61, no.9, pp.2738-2746, Sep. 2014

[10] M. Maamoun, R. Bradai, A. Naderi, R. Beguenane, and M. Sawan, "Controlled start-up stochastic decoding of LDPC codes," *IEEE 11th International New Circuits and Systems Conference (NEWCAS)*, 2013, pp. 1-4.

[11] Huang Kuo-Lun, V. Gaudet, and M. Salehi, "Output decisions for stochastic LDPC decoders," *48th Annual Conference on Information Sciences and Systems (CISS)*, 2014, pp. 1-5.