A Modified Min-Sum Algorithm for Quantized LDPC Decoders

Homayoon Hatami*, David G. M. Mitchell†, Daniel J. Costello, Jr.*, and Thomas Fuja*
*Dept. of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana, USA
Email: {hhatami,dcostel1,tfuja}@nd.edu
†Klipsch School of Electrical and Computer Engineering, New Mexico State University, Las Cruces, USA
Email: dgmm@nmsu.edu

Abstract—It is well known that for decoding low-density parity-check (LDPC) codes, the attenuated min-sum algorithm (AMSA) and the offset min-sum algorithm (OMSA) can outperform the conventional min-sum algorithm (MSA) at low signal-to-noise-ratios (SNRs). In this paper, we demonstrate that, for quantized LDPC decoders, although the MSA achieves better high SNR performance than the AMSA and OMSA, each of the MSA, AMSA, and OMSA all suffer from a relatively high error floor. Therefore, we propose a novel modification of the MSA for decoding quantized LDPC codes with the aim of lowering the error floor. Compared to the quantized MSA, the proposed modification is also helpful at low SNRs, where it matches the waterfall performance of the quantized AMSA and OMSA. The new algorithm is designed based on the assumption that trapping/absorbing sets (or other problematic graphical objects) are the major cause of the error floor for quantized LDPC decoders, and it aims to reduce the probability that these problematic objects lead to decoding errors.

I. INTRODUCTION

Low-density parity-check (LDPC) codes are a class of linear block codes for which the performance of iterative message passing (MP) decoding can approach that of much more complex maximum likelihood (ML) decoding. The min-sum algorithm (MSA) [1] is a simplified version of the sum-product algorithm (SPA) [2] that is commonly used for iterative MP decoding of LDPC codes, where the check node computation is approximated (and hence significantly easier to perform) at the cost of some performance loss when compared to the SPA. The simplification is particularly desirable for hardware decoder implementations. Moreover, unlike the SPA, no estimation of the channel signal-to-noise ratio (SNR) is needed at the receiver for an additive white Gaussian noise (AWGN) channel.

In [3], two modifications of the MSA, denoted attenuated MSA (AMSA) and offset MSA (OMSA) were introduced to account for the error in the approximation and improve performance. Also, [4] independently introduced the OMSA. These algorithms reduce the magnitudes of the log-likelihood ratios (LLRs) computed at the check nodes of the Tanner graph representation of the parity-check matrix \( \mathbf{H} \) of an LDPC code. Since the LLR magnitudes computed at the check nodes for the SPA are smaller than or equal to those of the MSA, making the LLR magnitudes of the check node to variable node messages in the MSA smaller can achieve better waterfall performance (closer to the SPA) compared to the conventional MSA.

Practical implementation of LDPC decoders requires a finite precision (quantized) representation of the LLRs. In [3], quantized density evolution (DE) was used to find the optimum attenuation and offset parameters for the AMSA and OMSA, in the sense that DE calculates the iterative decoding threshold, which determines the waterfall performance. At high SNRs, quantization typically causes an early error floor. In [5]–[7], it was shown that certain objects, called trapping sets, leafless elementary trapping sets, or absorbing sets, in the Tanner graph cause the decoding process to get stuck, resulting in decoding errors at high SNRs. (Hereafter, we refer to the sub-graphs induced by these sets, as well as similar sets, as problematic objects.) Following [3], [4], several papers, e.g., [8]–[10], focused on further improving the waterfall performance of the MSA. Strategies have also been proposed to lower the error floor of quantized LDPC decoders, including quantizer design [9], [11], [12], modifications to iterative decoding [13]–[16], and post-processing [17]–[20].

In this paper, we propose a new check node update modification of quantized MSA that is straightforward to implement and reduces the error floor when compared to the methods proposed in [3], [4]. First, we show that the AMSA and OMSA, with parameters optimized for waterfall performance, can exhibit worse (higher) error floors than the MSA. We then introduce a novel modification to the MSA that applies the strategies from the AMSA and the OMSA selectively, i.e., it applies attenuation/offset when it would be helpful and does not apply it otherwise. Assuming that there exist problematic objects that cause most of the decoding failures in the high SNR regime, we show that our new MSA modification causes these objects to become less prone to decoding failures. As a result, the new algorithm matches the waterfall performance of the AMSA and OMSA, while improving the error floor. In addition, no information about the location or structure of the problematic objects is required.

II. BACKGROUND

Let \( V = \{v_1, v_2, \ldots, v_n\} \) and \( C = \{c_1, c_2, \ldots, c_m\} \) represent the sets of variable nodes and check nodes, re-
spectively, of a bipartite Tanner graph representation of an LDPC parity-check matrix $H$. Assume that a codeword $u = (u_1, u_2, \ldots, u_n)$ is binary phase shift keyed (BPSK) modulated such that each zero is mapped to $+1$ and each one is mapped to $-1$. The modulated signal is transmitted over an AWGN channel with mean 0 and standard deviation $\sigma$. The received signal is $r = 1 - 2u + n$, where $n$ is the channel noise. We denote the quantized version of $r$ as $R = (r_1, r_2, \ldots, r_n).

A. The Min-Sum Algorithm and its Modifications

The MSA is an iterative MP algorithm that is simpler to implement than the SPA. Unlike the SPA, no channel noise information is needed to calculate the channel LLRs. The SPA is optimum for codes without cycles, but for finite length codes and finite precision LLRs, the SPA is not necessarily optimum, particularly with respect to error floor performance [16]. Let $V_{ij}$ represent the LLR passed from variable node $v_i$ to check node $c_j$ and let $C_{ji}$ represent the LLR passed from $c_j$ to $v_i$. The check nodes that are neighbors to $v_i$ are denoted $N(v_i)$, and the variable nodes that are neighbors to $c_j$ are denoted $N(c_j)$. To initialize decoding, each variable node $v_i$ passes $r_i$ to the check nodes in $N(v_i)$, i.e.,

$$V_{ij} = r_i,$$  \hspace{1cm} (1)

where the $V_{ij}$'s computed throughout the decoding process are referred to as the variable node LLRs.\(^1\) The check node operation to calculate the LLRs sent from check node $c_j$ to variable node $v_i$ is given by

$$C_{ji} = \left[ \prod_{i' \in N(c_j) \setminus i} \text{sign} (V_{i'j}) \right] \cdot \min_{i' \in N(c_j) \setminus i} |V_{i'j}|,$$  \hspace{1cm} (2)

where the $C_{ji}$'s computed throughout the decoding process are referred to as the check node LLRs. After each iteration, the hard decision estimate $\hat{u}$ is checked to see if it is a valid codeword, where $\hat{u}_i = 0$ iff

$$r_i + \sum_{j' \in N(v_i)} C_{j'i} > 0.$$  \hspace{1cm} (3)

If $\hat{u}$ is a valid codeword, or if the iteration number has reached $I_{\text{max}}$, decoding stops. Otherwise, the variable node LLRs are calculated as

$$V_{ij} = r_i + \sum_{j' \in N(v_i) \setminus j} C_{j'i},$$  \hspace{1cm} (4)

and decoding continues using (2).

In [3], two modified versions of the MSA, called attenuated MSA (AMSA) and offset MSA (OMSA), were introduced to reduce the waterfall performance loss of the MSA compared to the SPA. The modified check node computations are given by

$$C_{ji} = \alpha \left[ \prod_{i' \in N(c_j) \setminus i} \text{sign} (V_{i'j}) \right] \cdot \min_{i' \in N(c_j) \setminus i} |V_{i'j}|,$$  \hspace{1cm} (5)

1Computing an LLR value from the received value requires multiplying by $2/\sigma^2$, where $\sigma^2$ is the channel noise variance. However, this normalization, which is required for the SPA, is not required for min-sum decoding and its variants, so we omit it.

and

$$C_{ji} = \left[ \prod_{i' \in N(c_j) \setminus i} \text{sign} (V_{i'j}) \right] \cdot \max_{i' \in N(c_j) \setminus i} \left| V_{i'j} - \beta, 0 \right|,$$  \hspace{1cm} (6)

respectively, where $\alpha, \beta > 0$ are constants. In both algorithms, the check node LLR magnitudes are modified to be smaller than those of MSA, as noted earlier. This reduces the negative effect of overestimating the LLR magnitudes in the MSA, whose larger check node LLR magnitudes compared to the SPA cause additional errors in decoding. In a quantized decoder, the operations in (1) – (6) have finite precision, i.e., the values are quantized to a set of numbers ranging from $-\ell_{\text{max}}$ to $\ell_{\text{max}}$, with step size $\Delta$, where the resulting quantizer thresholds are set from $-\ell_{\text{max}} + \frac{\Delta}{2}$ to $\ell_{\text{max}} - \frac{\Delta}{2}$. The attenuation and offset parameters $\alpha$ and $\beta$ in (5) and (6) that have the best iterative decoding thresholds were found using quantized density evolution (DE) in [3]. In [4], the effects of these modifications on unquantized, clipped, and quantized versions of the MSA for three different codes were studied using extensive simulations.

B. Trapping Sets and Error Floors

Let $A$ denote a subset of $V$ of cardinality $a$. Let $E(A)$ and $O(A)$ represent the subsets of check nodes connected to variable nodes in $A$ with even and odd degrees, respectively, where $|O(A)| = b$. Here $A$ is called an $(a,b)$ trapping set [5]. A is defined to be an $(a,b)$ absorbing set if each variable node in $A$ is connected to fewer check nodes in $O(A)$ than $E(A)$ [7]. These sets, along with similar objects such as leafless elementary trapping sets, are known to cause most of the decoding errors at high SNRs for LDPC decoders [7]. In Fig. 1, the sub-graph $G(A)$ induced by a $(4,2)$ absorbing set $A$ is shown. In the next section, we will explain how the check node LLRs in (2) can be modified to improve decoding performance at high SNR, i.e., to lower the error floor, by considering decoder behavior in the presence of such problematic objects.

III. THRESHOLD ATTENUATED/OFFSET MSA

A. Motivation and Rationale

As discussed in the previous section, it is known that applying attenuation or offset when computing the check node LLRs can typically improve performance in the low SNR (waterfall) region for quantized decoders. On the other hand, since high SNR performance is tied to problematic graphical objects, the AMSA and OMSA do not necessarily achieve a good error floor. For example, Fig. 2 presents the simulated
bit-error-rate (BER) and frame-error-rate (FER) performance of the (8000,4000) code of [21] (which was used in [3] and [4]) with a 5-bit uniform quantizer, $\Delta = 0.15$, and $\ell_{\text{max}} = 2.25$, decoded using the MSA, AMSA, and OMSA. As shown, the AMSA and OMSA gain about 0.7dB in the waterfall compared to the MSA. However, all the algorithms eventually exhibit an error floor at higher SNRs. In this section, we focus on improving the high SNR performance of an LDPC code decoded with the MSA, AMSA, or OMSA, assuming that a problematic object causes the error floor.

At high SNRs, for a received vector $r$, decoding is successful most of the time. In the case of unsuccessful decoding, it is typically a small number of problematic objects that cause errors - objects containing variable nodes with unreliable (small magnitude) LLRs [7]. The channel LLRs for the variable nodes outside the problematic objects are, however, mostly reliable and have large magnitudes. In other words, the outside LLRs are typically initially large (and reliable) and continue to grow quickly to even larger values (often $\ell_{\text{max}}$), while the inside LLRs initially have smaller (often unreliable) magnitudes, but also grow quickly to larger values. The cause of errors in iterative decoding is that the unreliable (small magnitude) LLRs of the problematic objects increase to larger values during the iterations, but they typically contain at least one cycle, which does not allow for correction.

To improve the probability of correcting errors occurring in a problematic object $G(A)$ at high SNR, we note that the LLR magnitudes sent from a check node $c_j$ to variable nodes inside $A$ should grow more slowly (be attenuated) if $c_j$ receives at least one unreliable (small) LLR, so that the incorrect messages received from the channel in $A$ are not reinforced. If a check node $c_j$ (inside or outside $G(A)$) only receives large magnitude LLRs, on the other hand, these can be helpful for decoding and hence should not be attenuated. These two factors combined can lead to correct decoding of a received vector $r$ that would not occur otherwise.

### B. A Threshold Attenuated/Offset MSA

In our modified approach, we leverage a relationship observed at high SNRs between the variable node LLR magnitudes $|V_{ij}|$ received by check node $c_j$ and the likelihood of the check node $c_j$ being inside a problematic object $G(A)$. Therefore, the problem of locating errors affected by $G(A)$ is mapped into merely considering the variable node LLR magnitudes $|V_{ij}|$ received at check node $c_j$, i.e., we rely on $|V_{ij}|$ to indicate if $c_j$ is inside $G(A)$ and potentially causing decoding failures. At high SNR, the check node LLRs outside $G(A)$ typically grow faster than the LLRs inside. Therefore, if a check node $c_j$ receives at least one small LLR, i.e., $\min_{i' \in N(c_j)} |V_{ij}| \leq \tau$, where $\tau$ is some threshold, it is likely that $c_j$ is inside $G(A)$. Consequently, to improve the error floor performance, we propose the following check node computation to replace (2):

$$C_{ji} = \begin{cases} \prod_{i' \in N(c_j) \setminus i} \text{sign}(V_{ij}) & \frac{\min_{i' \in N(c_j)} |V_{ij}|}{|V_{ij}|} > \tau, \\ \alpha' \prod_{i' \in N(c_j) \setminus i} \text{sign}(V_{ij}) & \frac{\min_{i' \in N(c_j)} |V_{ij}|}{|V_{ij}|} \leq \tau, \end{cases}$$

(7)

where $\alpha' < 1$ is an attenuation parameter designed to reduce the check node LLR magnitudes sent from a check node $c_j$ inside $G(A)$ to the variable nodes of $A$. We denote this modified check node update algorithm as the threshold attenuated MSA (TAMSA). We will see in Sec. IV that, with a proper choice of parameters, the TAMSA is capable of correctly decoding some of the problematic objects that cause errors in the AMSA or OMSA.

In (7), we propose using $\alpha'$ to make the check node LLR magnitudes smaller when $\min_{i' \in N(c_j)} |V_{ij}| \leq \tau$. As an alternative (or together), an offset parameter $\beta'$ can be used to serve the same purpose, where

$$C_{ji} = \begin{cases} \prod_{i' \in N(c_j) \setminus i} \text{sign}(V_{ij}) & \frac{\min_{i' \in N(c_j)} |V_{ij}|}{|V_{ij}|} > \tau, \\ \prod_{i' \in N(c_j) \setminus i} \text{sign}(V_{ij}) & \frac{\min_{i' \in N(c_j)} |V_{ij}|}{|V_{ij}|} \leq \tau. \end{cases}$$

\begin{align*}
\frac{\min_{i' \in N(c_j)} |V_{ij}|}{|V_{ij}|} & \leq \tau, \\
\alpha' \prod_{i' \in N(c_j) \setminus i} \text{sign}(V_{ij}) & \frac{\min_{i' \in N(c_j)} |V_{ij}|}{|V_{ij}|} \leq \tau, \\
\prod_{i' \in N(c_j) \setminus i} \text{sign}(V_{ij}) & \frac{\min_{i' \in N(c_j)} |V_{ij}|}{|V_{ij}|} \leq \tau, \\
\end{align*}

(8)

and $\beta' > 0$ is an offset parameter that reduces the check node LLR magnitudes. We denote this modified check node update algorithm as the threshold offset MSA (TOMSA).

Both the TAMSA and TOMSA selectively, or locally, reduce the magnitudes of the check node LLRs that are likely to belong to a problematic object without requiring knowledge of its location or structure. The TAMSA and TOMSA add a simple threshold test compared to the AMSA and OMSA, while the attenuation (offset) parameter only needs to be applied to a few check nodes at high SNR.
Fig. 3. Simulated performance of an (8000,4000) LDPC code decoded with the MSA, AMSA, OMSA, TAMSA, and TOMSA. Solid curves represent BER, dashed curves represent FER.

IV. RESULTS AND DISCUSSION

In this section, we present the simulated performance of the (8000, 4000) code of [21] (used in [3] and [4]), the progressive edge growth (PEG) (1008, 504) LDPC code of [21] (used in [3]), and the (155, 64) Tanner code of [22] with various decoding algorithms, including the proposed TAMSA and TOMSA with different parameters, each using a 5-bit uniform quantizer with $\Delta = 0.15$ and $\ell_{\text{max}} = 2.25$.

Fig. 3 shows the BER and FER performance of the (8000,4000) code for the MSA, the AMSA with $\alpha = 0.8$, the OMSA with $\beta = 0.15$, the TAMSA and TOMSA with parameters $[\alpha' = 0.8, \tau = 2]$, and the TOMSA with parameters $[\beta' = 0.15, \tau = 2]$.

We see that, for the chosen parameters, the TAMSA and TOMSA each have better error floor performance than the MSA, AMSA, and OMSA and maintain the same waterfall performance. (The parameters used here were chosen after running decoder simulations for various values of $\alpha'$, $\beta'$, and $\tau$.)

Fig. 4 shows the BER and FER performance of the (1008,504) PEG-LDPC code for the quantized AMSA, OMSA, and TOMSA with three sets of parameters $[\alpha' = 0.8, \tau = 2]$, $[\alpha' = 0.8, \tau = 1.75]$, and $[\alpha' = 0.75, \tau = 1.75]$. We see that best error floors are achieved with the TAMSA and that the TAMSA with $[\alpha' = 0.75, \tau = 1.75]$ exhibits more than one order of magnitude gain compared to the AMSA and OMSA at SNR = 4dB. However, the waterfall performance is about 0.1dB worse than the other algorithms with $\alpha = \alpha' = 0.8$. Again, depending on the application, the values of $\alpha'$ and $\tau$ should be tuned to obtain the most desirable performance.

Fig. 5 shows the BER performance of the (155, 64) Tanner code with the AMSA for various values of $\alpha$, the TAMSA for $[\alpha = 0.7, \tau = 1.5]$, the MSA, and the unquantized SPA. We see that there is a notable difference between the unquantized SPA performance and the quantized MSA variants. However, at high SNRs, the TAMSA significantly outperforms both the AMSA and the MSA and maintains a roughly constant 0.5dB lag compared to the unquantized SPA performance.

Based on these results, the optimal values of $\alpha$ and $\beta$ introduced in [3], [4], obtained using quantized DE, are good candidates for $\alpha'$ and $\beta'$ in the TAMSA and TOMSA as well. We found that, when $\tau$ is close to $\ell_{\text{max}}$, values of $\alpha'$ or $\beta'$ with less reduction of LLR magnitudes suffer some waterfall performance loss compared to the optimal $\alpha$ or $\beta$, as also found in [3], [4]. While values of $\alpha'$ or $\beta'$ with more reduction of LLR magnitudes also hurt the waterfall performance, this can in fact be beneficial for the error floor performance, since they slow down the convergence of the LLRs from check nodes that don’t satisfy the condition $\min_{c_j \in N(c_i)} |V_{c_j}| > \tau$ in (7) or (8), thus allowing LLRs from check nodes that do satisfy (7) or (8) to help “correct” errors. Hence, values of $\alpha'$ and $\beta'$ with more reduction of LLR magnitudes may be desirable, depending on the application, e.g., if error floor performance is critical.

Our results also indicate that the TAMSA and TOMSA have comparable waterfall performance to the AMSA and OMSA with parameters $\alpha' = \alpha$ and $\beta' = \beta$, respectively, even though the TAMSA and TOMSA only selectively, rather than uniformly, offset or attenuate. In the low SNR region, it is important to reduce the check node LLR magnitudes of the MSA to obtain good performance. In this region, many of the channel LLR magnitudes are small, and the corresponding variable node and check node LLR magnitudes are also small for the first few iterations of decoding. As a result, it is likely that the variable node LLR magnitudes received by the check nodes at low SNR are below the threshold $\tau$. Therefore, in the early iterations, the threshold condition $\min_{c_j \in N(c_i)} |V_{c_j}| > \tau$ in (7) or (8) is not satisfied for most check nodes $c_j$ and attenuation/offset is applied, which improves the performance compared to the MSA. However, in the high SNR region, the threshold condition is usually satisfied outside a problematic object $G(A)$ (and attenuation/offset is not applied), while it is often not satisfied inside $G(A)$ (and attenuation/offset is...
applied), which improves the error floor performance. Together, this results in the TAMSA (TOMSA) combining the best features of the MSA and the AMSA (OMSA).

In Figs. 2-4, we see that the OMSA slightly outperforms the AMSA at high SNRs. This follows from the fact that, for given values of $\ell_{\text{max}}$, $\alpha$, and $\beta$, the LLR magnitudes for the OMSA grow to larger values compared to the AMSA (quantized value of $\ell_{\text{max}} - \beta$ vs. quantized value of $\alpha \times \ell_{\text{max}}$), which helps the reliable check node LLRs of the OMSA “correct” errors inside a problematic object $G(A)$. However, in Fig. 3, we see that the TAMSA has better error floor performance than the TOMSA. While the check node LLRs that satisfy (7) or (8) for both the TAMSA and the TOMSA can grow to $\ell_{\text{max}}$, the check node LLRs that don’t satisfy (7) or (8) are limited to a value smaller than $\tau$ (quantized value of $\alpha' \times \tau$ vs. quantized value of $\tau - \beta'$). Consequently, for the parameter values chosen in our examples, the TAMSA makes the check node LLRs that are below $\tau$ smaller compared to the TOMSA, which helps “correct” errors by slowing down the check node LLR convergence.

V. CONCLUSION

In this paper, a modified version of the MSA was proposed to lower the error floor of quantized LDPC decoders. Based on the assumption that a problematic object is the dominant cause of the error floor, the proposed TAMSA (TOMSA) selectively attenuates (offsets) a check node LLR if the check node receives any variable node LLR with magnitude below some threshold $\tau$, while allowing a check node LLR magnitude to reach the maximum quantizer level if all the variable node LLRs received by the check node have magnitude greater than $\tau$. It was shown that this new approach can decode some received vectors $r$ that become stuck, and thus cannot be decoded correctly using MSA or AMSA (OMSA). Simulation results presented for several codes demonstrated that the TAMSA (TOMSA) combines the advantages of both the MSA and the AMSA (OMSA) to offer better error floor performance without sacrificing waterfall performance. Future work involves applying the TAMSA to techniques designed to reduce complexity, such as layered decoding [23] and single-minimum MSA implementations [13].

REFERENCES


