

Neural Normalized Min-Sum Message-Passing vs. Viterbi Decoding for the CCSDS Line Product Code

Jonathan Nguyen*, Linfang Wang*, Chester Hulse*, Sahil Dani*, Amaael Antonini*, Todd Chauvin‡, Dariush Divsalar†, Richard Wesel*

*Department of Electrical and Computer Engineering, University of California, Los Angeles, Los Angeles, California 90095

†Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109

‡SA Photonics, Los Gatos, California 95032

Email: {nguyen.j, lfwang, chulse, sahildani, amaael, wesel}@ucla.edu, t.chauvin@saphotonics.com, Dariush.Divsalar@jpl.nasa.gov

Abstract—The Consultative Committee for Space Data Systems (CCSDS) 141.11-O-1 Line Product Code (LPC) provides a rare opportunity to compare maximum-likelihood decoding and message passing. The LPC considered in this paper is intended to serve as the inner code in conjunction with a (255,239) Reed Solomon (RS) code whose symbols are bytes of data. This paper represents the 141.11-O-1 LPC as a bipartite graph and uses that graph to formulate both maximum likelihood (ML) and message passing algorithms. ML decoding must, of course, have the best frame error rate (FER) performance. However, a fixed point implementation of a Neural-Normalized MinSum (N-NMS) message passing decoder closely approaches ML performance with a significantly lower complexity.

Index terms— line product code, LDPC decoders, neural network, maximum likelihood, FPGA

I. INTRODUCTION

Line codes describe a set of encoding maps used to transmit digital data. The primary purpose of a line code is to manage the disparity of a transmission, which is defined as the difference between the number of transmitted 1s and 0s. Managing bit disparity has the benefit of minimizing DC components in transmissions which cannot be reliably transmitted over most long-distance communication channels. In this paper, we reference the Consultative Committee for Space Data Systems' (CCSDS) 141.11-O-1 proposed line code, known as the Line Product Code (LPC) [1].

While often impractical since their complexity scales at a rate of $O(2^n)$, maximum likelihood (ML) decoders represent the best possible decoding performance. Previous work including [2] and [3] have proposed methods of reducing the complexity of ML decoding for linear block codes by representing them as a trellis and performing Viterbi decoding. While still on the order of $O(2^n)$, these methods drastically reduce number of required operations, enough so that for a short blocklength code such as the LPC, ML decoding is considered.

This research is supported by SA Photonics and the Air Force Research Lab (AFRL). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect views of SA or AFRL. Research was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

Message passing algorithms, such as belief propagation or MinSum, are low-complexity iterative decoders for linear block codes. However, message passing algorithms are sub-optimal because they assume that the Tanner graph defined by the parity check matrix has no cycles. As a result, for short block length codes with short cycles, message passing decoders do not provide satisfying performance.

Recently, numerous works have focused on improving the performance of message passing decoders with the aid of neural networks [4]–[12]. Nachmani *et al.* and Lugosch *et al.* in [4], [5], [7] proposed Neural Normalized MinSum (N-NMS) and Neural Offset MinSum (N-OMS) decoders to improve the performance of the NMS and OMS decoders. Unlike NMS and OMS, which use a constant multiplicative or offset weight, N-NMS and N-OMS assign distinct trainable weights to each edge in each iteration. Simulations in [4], [5], [7] show that N-NMS and N-OMS have the capability to drastically improve the decoding performance of NMS and OMS for short-blocklength codes. Therefore, this paper aims to reformulate the LPC as a linear block code to leverage the recent advancements in neural network-based decoders.

The rest of the paper is organized as follows: Section II introduces the LPC, how it manages bit disparity, and how it is encoded. Section III describes how to represent the LPC as a bipartite graph and how to perform both maximum likelihood and message passing decoding on it. Section IV details the process of implementing the considered decoders on a Field Programmable Logic Array (FPGA). Simulation results and corresponding discussion are shown in Section V and Section VI concludes this paper.

II. LINE PRODUCT CODE ENCODING

The LPC encoder operates on blocks of 25 bits denoted by $LPCEncIn[24:0]$. The most significant bit $LPCEncIn[24]$ is channel system data denoted as S . The LPC encoder discards $LPCEncIn[23:16]$ (legacy implementation of the laser communication terminal encoding process), and maps $LPCEncIn[15:0]$ to the following 4×4 matrix:

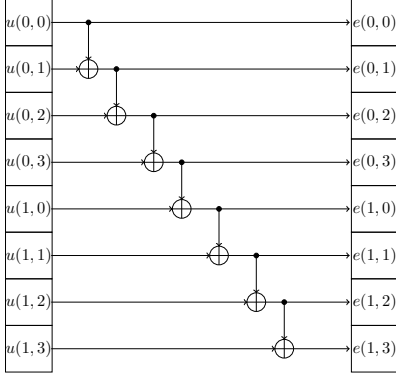


Fig. 1. Differential encoding to calculate sub-block 1. where \oplus = logical XOR

$u(0,0)$	$u(0,1)$	$u(0,2)$	$u(0,3)$
$u(1,0)$	$u(1,1)$	$u(1,2)$	$u(1,3)$
$u(2,0)$	$u(2,1)$	$u(2,2)$	$u(2,3)$
$u(3,0)$	$u(3,1)$	$u(3,2)$	$u(3,3)$

The LPC encoder generates the codewords of 24 bits. The codewords along with S forms a 5×5 matrix:

$e^*(0,0)$	$e^*(0,1)$	$e^*(0,2)$	$e^*(0,3)$	$ph(0)$
$e^*(1,0)$	$e^*(1,1)$	$e^*(1,2)$	$e^*(1,3)$	$ph(1)$
$e^*(2,0)$	$e^*(2,1)$	$e^*(2,2)$	$e^*(2,3)$	$ph(2)$
$e^*(3,0)$	$e^*(3,1)$	$e^*(3,2)$	$e^*(3,3)$	$ph(3)$
$pv(0)$	$pv(1)$	$pv(2)$	$pv(3)$	S

The encoding of LPC consists of the following steps:

- 1) Calculate $e(i, j)$ ($i, j = 0, \dots, 3$) using *LPCEncIn*[15:0] via differential encoding. In particular, we refer to $\{e(i, j) | i = 0, 1, j = 0, \dots, 3\}$ as sub-block 1 and $\{e(i, j) | i = 2, 3, j = 0, \dots, 3\}$ as sub-block 2.
- 2) Calculate the horizontal parity bits $ph(i)$ ($i = 0, \dots, 3$) and vertical parity bits $pv(i)$ ($i = 0, \dots, 3$).
- 3) Apply bit-wise inversion of sub-block 1 and/or 2 in order to minimize difference between the number of transmitted ones and zeros, which is also referred as disparity. We denote $e^*(i, j)$ ($i, j = 0, \dots, 3$) as the sub-block bits after inversion process.

The following section describes these three steps in detail.

A. Differential Encoding

Differential encoding is performed on the two sub-blocks separately. For sub-block 1, initialize $e(0, 0) = u(0, 0)$ and $e(1, 0) = u(1, 0) \oplus e(0, 3)$:

$$e(i, j) = u(i, j) \oplus e(i, j - 1), \quad j > 0 \quad (1)$$

For sub-block 2, initialize $e(2, 0) = u(2, 0)$ $e(2, 0) = u(2, 0) \oplus e(0, 3)$ as shown in Fig. 1. The remaining bits can be derived using equation (1). Fig. 1 shows the differential encoding on sub-block 1.

B. Horizontal Parity bits

After calculating the $\{e(0, 0) \dots e(3, 3)\}$ bits, the horizontal and vertical parity bits must be determined. The horizontal parity bit $ph(0)$ is always calculated for odd parity of the first row, meaning

$$\left[ph(0) + \sum_{k=0}^3 e(0, k) \right] \bmod 2 = 1 \quad (2)$$

The other three parity bits $ph(1)$, $ph(2)$, and $ph(3)$ are determined not only by their corresponding rows, but also by $e(0, 0)$. Specifically, if $e(0, 0) = 0$, then odd parity is used for $ph(1)$, $ph(2)$, and $ph(3)$ and its corresponding rows. Otherwise, the even parity must be satisfied. Therefore,

$$\left[ph(i) + \sum_{k=0}^3 e(i, k) \right] \bmod 2 = 1 \oplus e(0, 0), \quad i = 1, 2, 3. \quad (3)$$

C. Vertical Parity Bits

The vertical parity bit $pv(0)$ is always calculated for even parity of the first row, meaning that

$$\left[pv(0) + \sum_{k=0}^3 e(k, 0) \right] \bmod 2 = 0 \quad (4)$$

The other vertical parity bits, $pv(1)$, $pv(2)$, and $pv(3)$ are calculated using their corresponding rows and $e(2, 0)$. If $e(2, 0) = 0$, then even parity is used for $pv(1)$, $pv(2)$, and $pv(3)$ and its corresponding rows. Otherwise, the odd parity must be satisfied. Therefore:

$$\left[pv(i) + \sum_{k=0}^3 e(k, i) \right] \bmod 2 = e(2, 0), \quad i = 1, 2, 3 \quad (5)$$

D. Minimization of Disparity

The disparity of each sub-block is defined as the difference between the number of transmitted ones and zeros. The goal of the LPC encoder is to minimize the disparity of each sub-block so that a relatively equal number of 0s and 1s are transmitted. Consequently, sub-block 1, sub-block 2, both, or neither are inverted at the encoder's end depending on the value of the disparity bits of each sub-block. Define $DispSum[i]$ ($i = 0, \dots, 3$) as the disparity in the 5×5 matrix, after the inversion of none, one or both sub-blocks. The following table lists inversion rules corresponding to each $DispSum[i]$ bit where $i = 0, \dots, 3$:

	Inversion of Sub-block 1	Inversion of Sub-block 2
$DispSum[0]$	No	No
$DispSum[1]$	Yes	No
$DispSum[2]$	No	Yes
$DispSum[3]$	Yes	Yes

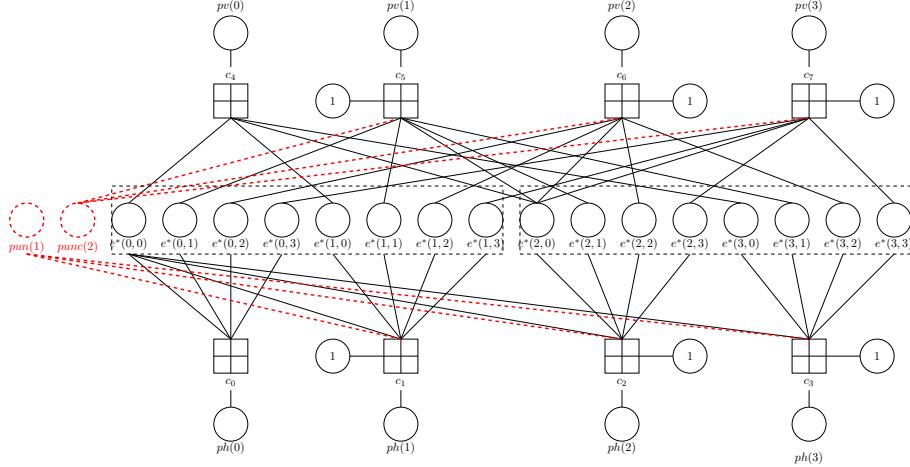


Fig. 2. Bipartite graph of LPC. The red dashed circles are punctured variable node which indicate the sub-block inversion.

The LPC encoder performs sub-block inversion based on the rules shown in the previous table that provide minimum $DispSum[i]$. Based on these inversion rules, the $e^*(i, j)$'s are calculated as follows:

$$e^*(i, j) = \begin{cases} 1 - e(i, j) & \text{Inversion} \\ e(i, j) & \text{No Inversion} \end{cases}, \quad (6)$$

where $i = 0, 1$ for sub-block 1, $i = 2, 3$ for sub-block 2, and $j = 0, \dots, 3$ for both sub-blocks.

III. LINE PRODUCT CODE DECODING

As a linear code, LPC can be represented by a parity check matrix H and corresponding bipartite graph \mathcal{G} . Let \mathbf{v} be a codeword of LPC, and define $\mathbf{s}(\mathbf{v})$ by:

$$\mathbf{s}(\mathbf{v}) = H\mathbf{v}^T. \quad (7)$$

Note that for a conventional linear block code, $\mathbf{s}(\mathbf{v})$ is a vector that is independent with \mathbf{v} . However, in this case, there are four distinct \mathbf{s} with each one corresponding to a single inversion rule. One possible solution is to perform the decoding process using four different \mathbf{s} separately. This, however, will inevitably increase the hardware usage and decoding latency.

The following section shows that the four matrices can be combined into one by introducing two punctured variable nodes which indicate the inversion rule. As a result, decoding can be performed using only one matrix. The next sections describes the application of decoding methods including maximum likelihood and message passing to the LPC.

A. Parity Check Matrix Representation

Equations (2) through (5) put eight parity check constraints on $e(i, j)$, horizontal parity bits and vertical parity bits. The black, solid line portions in Fig. 2 represent the bipartite graph defined by these eight parity checks. The "box-plus" symbols and circles represent check nodes and variable nodes, respectively. Circles with a 1 represent a special variable node

whose value is a constant 1. The eight check nodes are denoted as c_0, \dots, c_7 . The bipartite graph is drawn such that the modulo-2 sum of all variable nodes connected to each check node must equal zero. These are known as the parity checks.

Given a valid codeword, incorrectly inverting one sub-block will cause the new codeword to fail some parity checks. More specifically, the check nodes that connect to an odd number of variable nodes in one sub-block will no longer satisfy all parity checks if that sub-block gets inverted.

Therefore $\{c_1, c_2, c_3\}$ do not satisfy the parity check condition when sub-block 1 gets inverted and $\{c_5, c_6, c_7\}$ do not satisfy the parity check condition when sub-block 2 gets inverted. Two extra variable nodes $punc(1)$ and $punc(2)$ are introduced in order to make sure that the check nodes still satisfy the parity check condition after the sub-block inversion. $punc(1)$ connects the check nodes that have an odd number of variable node neighbors belonging to sub-block 1. When sub-block 1 gets inverted, $punc(1)$ equals 1 such that for each check node connected to $punc(1)$, all variable nodes connected to that check node sum to zero. Similarly, $punc(2)$ connects the check nodes that have an odd number of variable node neighbors belonging to sub-block 2. Fig. 2 shows the complete bipartite graph of LPC.

B. Maximum Likelihood Decoding via the Parity Check Matrix

In this section, we describe how to utilize Wolf's work in [2] to represent the Line Product Code (LPC) as a trellis for performing maximum likelihood (ML) decoding. Since ML decoding represents the theoretical limit of decoding performance, practical decoders which achieve frame error rates closer to it are more desirable. Furthermore, because the LPC is a short blocklength code, reduction in complexity via a trellis representation such as [2] may be feasible to implement in hardware.

Naive ML decoders simply compare the received codeword against all valid codewords. As a result, its complexity scales on the order of 2^k where k is the number of information

bits in the codeword. According to the bipartite graph for the LPC shown in Figure 2, there are 262,144 unique codewords making it infeasible in hardware.

As such, we leverage Wolf's framework in [2] to create a trellis representation of the LPC. While sacrificing some parallelism, the final trellis contains only 2,764 edges which represents a 94-fold reduction in complexity compared to brute force ML decoding. To construct the trellis, we follow the procedure in [2] with one important caveat: we terminate the trellis at the syndrome of $(1, 1, 1, 1, 0, 0, 0, 0)$ instead of the all-zeros syndrome. This is because the Line Product Code is NOT strictly linear due to the use of odd parity (XOR with a constant 1). This means that the trellis must terminate at a syndrome with 1s in the indices whose corresponding check node contains this constant 1, and 0 elsewhere.

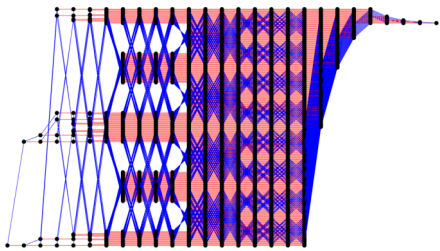


Fig. 3. LPC Trellis Derived from Wolf's Method Following Pruning, with 2942 states and 5372 branches

C. Maximum Likelihood Decoding via the Generator Matrix

The trellis representation of a code can also be constructed using a trellis oriented generator matrix [3]. A code's generator matrix G can be transformed to become "trellis oriented" via row operations. A permutation of the columns of G yields a code G' equivalent to G on memoryless channels [3]. This permutation may yield a simpler trellis. However, finding this permutation is known to be an NP-hard problem. Using heuristics, a permutation that simplifies the trellis may be found. For example, a graph of a trellis obtained with one such permutation of the LPC yielded a trellis with 1098 states and 1908 branches, down from 2942 states and 5372 branches of the original code.

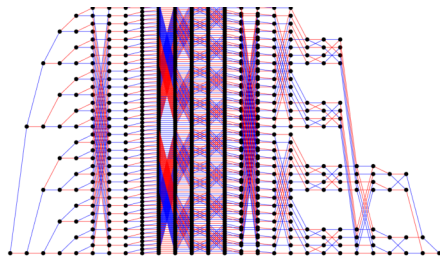


Fig. 4. LPC Trellis Derived from a Generator Matrix Allowing Permutations, with 1098 states and 1908 branches

D. Message Passing Decoding

Message passing decoding algorithms, such as belief propagation and MinSum, provide an excellent decoding perfor-

mance for linear block codes with large girths, defined as length of the shortest cycle in its Tanner graph. For the LPC, message passing decoders do not perform well, because its girth is only 4.

Recently, the neural-network-aided message passing decoders [4], [5], [7] have shown substantial improvements compared to conventional message passing decoders. Neural-network-aided message passing decoders assign distinct weights to each message in each iteration, such that the decoder can overcome trapping sets with short cycle lengths. This paper considers a neural normalized MinSum (N-NMS) decoder with a flooding schedule. In the t^{th} decoding iteration, N-NMS updates the check-to-variable node message, $u_{c_j \rightarrow v_i}^{(t)}$, the variable-to-check node message, $l_{v_i \rightarrow c_j}^{(t)}$, and posterior of each variable node, $l_{v_i}^{(t)}$, by:

$$u_{c_i \rightarrow v_j}^{(t)} = \beta_{(c_i, v_j)}^{(t)} \times \prod_{v_{j'} \in \mathcal{N}(c_i) / \{v_j\}} \text{sgn}(l_{v_{j'} \rightarrow c_i}^{(t-1)}) \quad (8)$$

$$\times \min_{v_{j'} \in \mathcal{N}(c_i) / \{v_j\}} |l_{v_{j'} \rightarrow c_i}^{(t-1)}| \quad (9)$$

$$l_{v_j \rightarrow c_i}^{(t)} = l_{v_i}^{ch} + \sum_{c_{i'} \in \mathcal{N}(v_j) / \{c_i\}} u_{c_{i'} \rightarrow v_j}^{(t)} \quad (9)$$

$$l_{v_j}^{(t)} = l_{v_i}^{ch} + \sum_{c_{i'} \in \mathcal{N}(v_j)} u_{c_{i'} \rightarrow v_j}^{(t)} \quad (10)$$

$\mathcal{N}(c_i)$ represents the set of the variable nodes connected to c_i and $\mathcal{N}(v_j)$ represents the set of the check nodes that are connected to v_j . $l_{v_j}^{ch}$ is the LLR of the channel observation of v_j . $\beta_{(c_i, v_j)}^{(t)}$ are multiplicative weights to be trained. The decoding process terminates when all parity checks are satisfied or when the maximum iteration count, I_T , is reached. In this paper, we follow the steps of [12] to train the neural network.

IV. HARDWARE IMPLEMENTATION OF MESSAGE PASSING DECODERS

Despite ML decoding being the most optimal, its computational complexity for both the parity check and generator matrix derived trellises is too high to meet timing constraints for practical hardware implementation. Table I shows the worst case number of operations for each decoding method considered here. An operation is defined as an addition or multiplication, in the case of belief propagation (BP), we consider arctan, exp, and log as a single operation since they are typically implemented via a Lookup Table (LUT). Table I indicates that message passing algorithms, except belief propagation, utilize significantly fewer operations than ML decoders, indicating that they are the most feasible to implement on hardware. It should also be noted that the Table I assumes that the message passing decoders always run for 8 iterations. In reality, for higher E_b/N_0 (around 7.5 dB for the LPC), the number of required iterations approaches 1, making message passing even more attractive. As such, our focus for this section will be on the N-NMS decoder, with MS and NMS decoders included for the purpose of comparison. The field-programmable gate array (FPGA) device used for hardware implementation was the Zynq ZCU106 MPSoC.

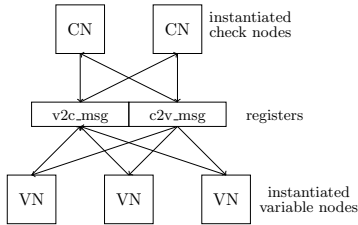


Fig. 5. Block Diagram of FPGA architecture.

The overall architecture consists of a bank of registers storing messages between check and variable nodes, and small modules to perform check node (CN) and variable node (VN) operations, as seen in Figure 5. The overall decoder controls the timing and coordinates the messages passed between the check and variable node modules for each decoding iteration. It also controls the terminating point by checking if the codeword estimate is valid or if the maximum number of iterations has been reached.

The initial FPGA implementation is a simple MinSum decoder, where check nodes search for the two minimums among its messages, and variable nodes compute simple summations. MinSum will be used as baseline to compare against the Normalized MinSum (NMS) and Neural-Normalized MinSum (N-NMS) implementations. However, in order to implement the N-NMS decoder, we must dynamically assign edge weights depending on the iteration of the decoder. This task is divided between 2 modules: the main decoder and the check node module. The multiplicative edge weights are first quantized to a 6-2 scheme, meaning the first 6 bits represent the integer part of the number and the last 2 bits represent the fractional part. Testing and simulations on the LPC showed that the 6-2 quantization achieved a satisfactory middle ground between accuracy and bit width.

Once the edge weights are quantized, they are stored in Block RAM (BRAM). The structure of the BRAM can be represented as a 2-D matrix where each element represents a register that stores an 8-bit quantized edge weight. The index of a certain edge weight in the matrix also contains information regarding the iteration count and edge number in the bipartite graph. The main decoder module uses this information to assign weights to the proper edge depending on the iteration count. After the check node module calculates the check-to-variable node message, it then multiplies that by the incoming edge weight. The fixed point FER curves in the following section are generated using C++ simulations. To match the FPGA implementation, all edge weights and calculated messages are quantized according to the 6-2 scheme discussed earlier.

V. SIMULATION RESULTS

A. Frame Error Rates for Various Decoders

In this section, we showcase our floating point simulation results for the Frame Error Rate (FER) of various decoding methods. The maximum likelihood FER was simulated using

TABLE I
COMPARISON OF DECODING COMPLEXITY VIA NUMBER OF OPERATIONS.

Decoder	Worst case number of operations
BP*	10,192
standard MS*	3,200
NMS* / N-NMS*	3,616
ML (Brute force)	524,288
ML (Generator matrix trellis)	6,130

* Message passing algorithms are assumed to always run for 8 iterations

the trellis method described in sections III-B and III-C. Additionally, in line with practical limitations on actual decoding hardware, we limit the number of decoding iterations to two and eight. At high E_b/N_0 , most received codewords are low-noise, making the average number of iterations approach 1. However, errors involving trapping sets require more iterations to correct which explains the gap between the 2 and 8 iteration decoders.

The N-NMS decoder performed the closest to Maximum Likelihood out of the three decoders considered, even beating out Belief Propagation. These results line up with the findings of [12]. To summarize, via its training process, the N-NMS decoder was able to adapt its weights to the particular structure of the LPC unlike belief propagation or normalized MinSum. In particular, the N-NMS weights are specifically trained to mitigate the decoding loss caused by trapping sets. With the LPC being such a short block-length code, its cycles have particularly small girths making N-NMS the ideal decoding method for it.

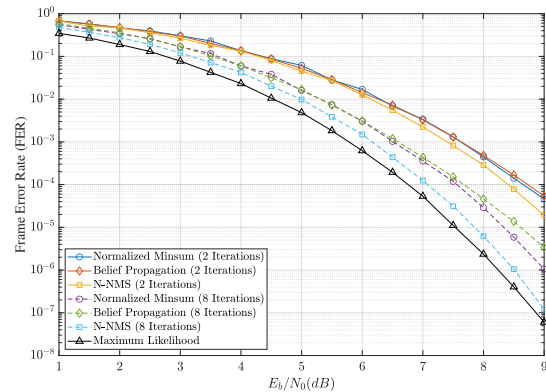


Fig. 6. FER for various floating point decoding methods capped at 2 and 8 decoding iterations

B. Quantization Loss for Fixed Point Decoders

In Table II, we observe a noticeable increase in the Look Up Tables (LUTs) used by the N-NMS implementations as compared to the baseline MS and NMS ones. However, it is important to note that the N-NMS decoders also perform better than their counterparts. So, essentially, we are trading extra hardware utilization for better performance.

The 6-2 quantization used on our FPGA is inherently different than typical software simulations which utilize 64-

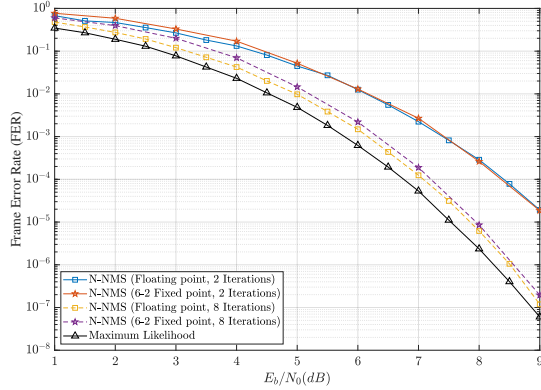


Fig. 7. Floating vs fixed point FER for the N-NMS decoder capped at 2 and 8 decoding iterations

TABLE II
DECODER FER PERFORMANCE AND RESOURCE USAGE

decoder	E_b/N_0 (dB) ^a	LUT	Reg.
baseline MS (8)	9.93	4953 (100%)	2201 (100%)
NMS	9.63	5205 (105%)	2201 (100%)
N-NMS(1 ^b)	13.36	5793 (117%)	2201 (100%)
N-NMS(2)	10.54	5784 (117%)	2206 (100%)
N-NMS(4)	9.28	5795 (117%)	2201 (100%)
N-NMS(8)	9.17	5796 (117%)	2206 (100%)

^aEstimated $\left(\frac{E_b}{N_0}\right)$ to achieve FER of 10^{-7} .

^b(n) number of iterations spent decoding.

bit floating point numbers. Since we utilize fewer bits in our fixed point implementation, its precision is comparatively diminished to floating point and we expect some deterioration in FER. The purpose of the simulations shown in Figure 7 is to demonstrate that with the N-NMS decoder, utilizing a fixed point quantization as opposed to floating point presents an almost negligible loss in frame error rate.

C. Reed Solomon Frame Error Rate

As noted in the CCSDS specification, the LPC serves as the inner code in conjunction to a (255,239) Reed-Solomon (RS) operating on GF(256). Since the RS code has 16 parity bytes, it can correct for up to 8 byte errors. Considering that the LPC encodes 2 bytes of data at a time, the RS code can correct for up to 4 LPC frame errors. Given the FER of the LPC, the corresponding FER of the RS code can be modeled via a binomial expression: $P_{RS}(e) = 1 - P(X < 5)$, where $X \sim B(128, P_{LPC}(e))$. Figure 8 shows the Reed Solomon FER for the N-NMS fixed and floating point implementations at various E_b/N_0 with the LPC as the inner code.

VI. CONCLUSION

This paper compares both the feasibility and decoding performance of maximum likelihood (ML), belief propagation, standard MinSum, normalized MinSum, and Neural Normalized MinSum (NNMS) decoders on the Line Product Code (LPC). An initial exploration of ML decoding was considered due to the LPC's short blocklength. However, attempted

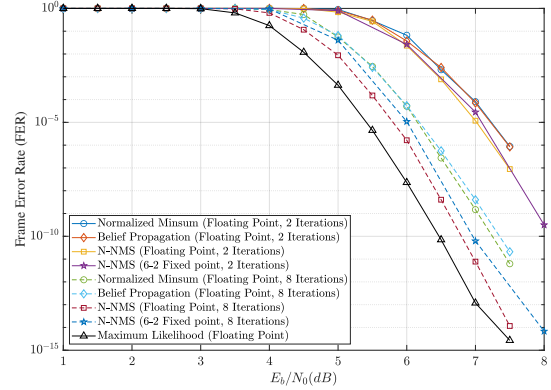


Fig. 8. Reed Solomon Code FER for Floating vs Fixed point N-NMS and Maximum Likelihood Decoding on the LPC

simulation on an FPGA board showed that, even with complexity reduction via a trellis, ML decoding failed to meet timing requirements. In lieu of this, we considered message passing algorithms, while less optimal than ML decoding, can be performed iteratively and with fewer operations than it. Simulation results on the LPC show that, with sufficient iterations, these message passing algorithms approach the frame error rate achieved by ML decoding. In particular, the NNMS decoder, with only 8 iterations, show only a 0.5 dB loss compared to ML making it the most promising decoder considered here.

REFERENCES

- [1] "Optical high data rate (HDR) communication — 1064 nm," Consultative Committee for Space Data Systems (CCSDS), Washington, DC, USA, 2008.
- [2] J. Wolf, "Efficient maximum likelihood decoding of linear block codes using a trellis," *IEEE Transactions on Information Theory*, vol. 24, no. 1, pp. 76–80, 1978.
- [3] V. Pless, W. C. Huffman, and R. A. Brualdi, *Handbook of coding theory*. Amsterdam: Elsevier, 1998.
- [4] E. Nachmani, Y. Be'ery, and D. Burshtein, "Learning to decode linear codes using deep learning," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing*, Sep. 2016, pp. 341–346.
- [5] L. Lugosch and W. J. Gross, "Neural offset min-sum decoding," in *2017 IEEE Inter. Symp. on Info. Theory (ISIT)*, Jun. 2017, pp. 1361–1365.
- [6] E. Nachmani, E. Marciano, D. Burshtein, and Y. Be'ery, "RNN decoding of linear block codes," *CoRR*, vol. abs/1702.07560, 2017. [Online]. Available: <http://arxiv.org/abs/1702.07560>
- [7] E. Nachmani, E. Marciano, L. Lugosch, W. J. Gross, D. Burshtein, and Y. Be'ery, "Deep learning methods for improved decoding of linear codes," *IEEE J. Sel. Top. Sig. Pro.*, vol. 12, no. 1, pp. 119–131, Feb. 2018.
- [8] F. Liang, C. Shen, and F. Wu, "An iterative BP-CNN architecture for channel decoding," *IEEE J. Sel. Top. Signal Process.*, vol. 12, no. 1, pp. 144–159, Feb. 2018.
- [9] X. Wu, M. Jiang, and C. Zhao, "Decoding optimization for 5G LDPC codes by machine learning," *IEEE Access*, vol. 6, pp. 50179–50186, 2018.
- [10] L. Lugosch and W. J. Gross, "Learning from the syndrome," in *2018 52nd Asilomar Conf. on Signals, Systems, and Computers*, Oct. 2018, pp. 594–598.
- [11] W. Lyu, Z. Zhang, C. Jiao, K. Qin, and H. Zhang, "Performance evaluation of channel decoding with deep neural networks," in *2018 IEEE Inter. Conf. on Comm. (ICC)*, May 2018, pp. 1–6.
- [12] L. Wang, S. Chen, J. Nguyen, D. Dariush, and R. Wesel, "Neural-network-optimized degree-specific weights for ldpc minsum decoding," *arXiv preprint arXiv:2107.04221*, 2021.