Abstract—In this paper we design protograph-based LDPC codes with short block sizes. Mainly we consider rate 1/2 codes with input block sizes 64, 128, and 256 bits. To simplify the encoder and decoder implementations for high data rate transmission, the structure of the codes is based on protographs and circulants. These codes are designed for short block sizes based on maximizing the minimum distance and stopping set size subject to a constraint on the maximum variable node degree. In particular, we consider codes with variable node degrees between 3 and 5. Increasing the node degree leads to larger minimum distances, at the expense of smaller girth. Therefore, there is a trade-off between undetected error rate performance (improved by increasing minimum distance) and the degree of sub-optimality of the iterative decoders typically used (which are adversely affected by graph loops). Various LDPC codes are compared and simulation results are provided.

I. INTRODUCTION

Low-density parity-check (LDPC) codes were proposed by Gallager [1] in 1962. Recently, there have been many contributions to the design and analysis of LDPC codes; see for example [2], [8], [9], and references therein. For ensemble weight enumerators for unstructured irregular LDPC codes see for example [18], [19]. Recently a flurry of work has been conducted on the design of LDPC codes with imposed sub-structures, starting with the introduction of multi-edge-type codes in [7].

Short block length codes are desirable in scenarios where a frame length constraint is imposed on the physical layer. This occurs, for instance, when command and control data is transmitted from ground to a spacecraft. In this paper we use degree-3 and higher degree nodes. By avoiding degree-2 nodes we obtain a code structure whose minimum distance grows linearly with blocksize $n$. Limiting code design to the use of degree-3 and higher variable nodes is a sufficient, but not necessary condition for minimum distance to grow linearly with $n$ [11]. In this paper we demonstrate that small protograph based codes can in fact achieve competitively low required bit signal-to-noise ratio without the use of degree-2 variable nodes.

These LDPC codes can improve data link and network layer protocols in support of communication networks. In addition, these codes are suitable for transmission of uplink command and control data since they have very low undetected error rates. For an input block size of 64, the maximum undetected error rate for all signal-to-noise ratios is less than $3 \times 10^{-5}$. This maximum occurs at a bit signal-to-noise ratio between 1 and 2 dB. The undetected error rate is lower than this maximum value both below and above this SNR. Short codes having low undetected error rates at all SNRs are useful for a spacecraft if it tumbles and loses its antenna pointing.

II. PROTOGRAPH LDPC CODES

To aid in implementation of high-speed decoding, it is advantageous for an LDPC code to be constructed from a protograph [5] or projected graph [6]. A protograph is a Tanner graph with a relatively small number of nodes. A “copy-and-permute” operation [5] can be applied to the protograph to obtain larger derived graphs of various sizes. This operation consists of first making $N$ copies of the protograph, and then permuting the endpoints of each edge among the $N$ variable and $N$ check nodes connected to the set of $N$ edges copied from the same edge in the protograph. The derived graph is the graph of a code $N$ times as large as the code corresponding to the protograph, with the same rate and the same distribution of variable and check node degrees. LDPC codes with protograph structure are a subclass of multi-edge-type LDPC codes.
As an example for protograph based LDPC codes we consider the rate-1/3 Repeat-Accumulate (RA) code depicted in Fig. 1(a). For this code the minimum $E_b/N_0$ threshold with iterative decoding is 0.502 dB. This code has a protograph representation shown in Fig. 1(b), as long as the interleaver $\pi$ is chosen to be decomposable into permutations along each edge of the protograph. The iterative decoding $E_b/N_0$ threshold is unchanged despite the additional constraint imposed by the protograph. The protograph consists of 4 variable nodes and 3 check nodes, connected by 9 edges. Three variable nodes are connected to the channel (transmitted nodes) and are shown as dark filled circles. One variable node is not connected to the channel (i.e., it is punctured) and is depicted by a blank circle. The three check nodes are depicted by circles with a plus sign inside.

Repeat-Accumulate (RA) [3], Irregular Repeat-Accumulate (IRA) [4], and recently proposed Accumulate-Repeat-Accumulate (ARA) [10] codes, with suitable definitions of their interleavers, all have simple protograph representations. These codes provide fairly low iterative decoding thresholds but have sublinear minimum distance. However for certain applications linear minimum distance is required for low error floor performance.

III. RECIPROCAL CHANNEL APPROXIMATION IN PROTOGRAPHS

Computation of iterative decoding thresholds for the protographs [15] used in this paper follows a fast and accurate approximation to density evolution called reciprocal channel approximation (RCA) originally proposed in [16] for regular LDPC codes.

The reciprocal channel approximation (RCA) makes use of a single real-valued parameter, in this case signal-to-noise ratio (SNR) $s$, as a stand-in for full density evolution. For every value of $s$, a reciprocal of SNR, $r$, is defined such that $C(s) + C(r) = 1$, where $C(x)$ denotes the capacity of the binary-input AWGN channel with SNR $x$. In the reciprocal channel approximation, the parameter $s$ is additive at variable nodes, and the reciprocal parameter $r$ is additive at check nodes. Chung’s self-inverting reciprocal energy function, $R(x) = C^{-1}(1 - C(x))$, transforms between the parameters $s$ and $r$, namely $r = R(s)$ and $s = R(r)$.

To apply the RCA technique to a protograph we first identify all transmitted variable nodes and select a target channel SNR $s_{chan}$. As shown in Fig. 2 messages $\tilde{s}_e$ are passed along edges leaving variable nodes ($\tilde{s}_e = s_{chan}$ from transmitted nodes and $\tilde{s}_e = 0$ from punctured nodes). The transformation $R(\tilde{s}_e)$ is applied and an extrinsic return message, $\tilde{r}_e$, is determined by computing the sum of all incoming messages save the one along edge $e$. Transformation $R(\cdot)$ is then reapplied to produce $\tilde{\tilde{r}}_e$. The process continues and a threshold is determined by the smallest value of $s_{chan}$ for which unbounded growth of all messages $\tilde{s}_e$ can be achieved. Less than 0.005 dB deviations from true density evolution thresholds have been observed by the application of this approximation to protographs in BI-AWGN channels.

**Fig. 2.** The reciprocal channel approximation in use on a protograph.

Motivation for applying RCA to the BI-AWGN channel most likely derived from the fact that a similar reciprocal channel definition yields exact density evolution results [16] when applied to the binary erasure channel (BEC). In the case of a BEC with erasure probability $\epsilon$ and capacity $C = 1 - \epsilon$, a parameter $s = -\log \epsilon$ is additive at variable nodes, a reciprocal parameter $r = -\log(1 - \epsilon)$ is additive at check nodes, and $s$ and $r$ are related by $C(s) + C(r) = 1$.  

---

**Fig. 1.** (a) A rate-1/3 RA code with repetition 3, and (b) its corresponding protograph.
EXIT-based analysis [20] or analysis based on other one-dimensional functionals acting on intermediate densities provides additional insight into the design of such codes.

IV. PROTOGRAPH OF REGULAR LDPC CODES

We express the normalized logarithmic asymptotic weight distribution of a code as 
\[ r(\delta) = \lim_{n \to \infty} \frac{\text{sup} r_n(\delta)}{n} \]
where 
\[ r_n(\delta) = \frac{\ln(A_n)}{n} \]
d is Hamming distance, and \( A_d \) is the ensemble weight distribution.

If the first zero crossing of \( r(\delta) \) (i.e. \( r(\delta_{\text{min}}) = 0 \) for \( \delta_{\text{min}} > 0 \)) exists, and \( r(\delta) < 0 \) for all \( 0 < \delta < \delta_{\text{min}} \) then \( \delta_{\text{min}} \) is called the typical minimum distance ratio. Consider \( P(d \leq \delta_{\text{min}} n) \leq \sum_{d=1}^{\delta_{\text{min}} n} A_d \). If the contribution of codewords with weights less than \( \delta_{\text{min}} n \) to this sum becomes negligible as \( n \) becomes large, then we can say that the normalized minimum distance is almost surely greater than \( \delta_{\text{min}} - \epsilon \) for any \( \epsilon > 0 \). In other words, with high probability the minimum distance of most codes in the ensemble increases linearly with \( n \).

Classic regular LDPC codes, in addition to simplicity, have low error floors. However, their iterative decoding thresholds are high. For example the (3,6) regular LDPC codes have an iterative decoding threshold of 1.102 dB while their typical minimum distance ratio is 0.023. For comparison the typical minimum distance ratio of random codes is 0.11. This typical minimum distance ratio or growth rate \( \delta_{\text{min}} \) is a characteristic of the specific protograph from which the LDPC code ensemble is constructed. Methods to compute the asymptotic weight enumerators for LDPC codes with protograph structure are presented in [12] and [13].

Fig. 3 compares the asymptotic weight distribution of (3,6) regular LDPC codes to that of rate 1/2 random codes.

we chose a design with degrees at least 3 to maximize the growth rate \( \delta_{\text{min}} \) and to allow the simplest possible base protograph with a single check node.

Fig. 4. Protograph of an LDPC code with variable node degrees 3 and 4.

The protograph associated with the rate 1/2 code was lifted by a factor of 4 to remove all multiple parallel edges. The lifted protograph is shown in Fig. 5. The resulting graph was then lifted by factor of 16 using progressive edge growth [17] and the ACE algorithm [14] to find phases associated with circulants of size 16. Performance of the rate 1/2 protograph code with \( k = 64 \)

V. PROTOGRAPH CODES FOR SHORT BLOCKS

We wish to construct protograph-based LDPC codes for short block lengths that simultaneously achieve low iterative decoding thresholds and linear minimum distance (\( \delta_{\text{min}} > 0 \)) such that error floors may be effectively suppressed. We start with a rate 1/2, 2-node protograph with one variable node having degree 3 and the other having degree 4 as shown in Fig. 4. Protographs with variable node degrees at least 3 are guaranteed to have the linear minimum distance property. Protographs with some degree-2 variable nodes can also have the linear minimum distance property if the number of degree-2 nodes is less than the number of checks [13]. However,
and \( n = 128 \) for various number of iterations is shown in Fig. 6. The performance of this code is compared with the \( k = 128 \) and \( k = 256 \) codes are shown in Fig. 8 for 50 iterations.

![Fig. 6](attachment:fig6.png)

**Fig. 6.** Performance of rate 1/2 protograph code in Fig. 5 for short block size, \( k = 64 \).

![Fig. 8](attachment:fig8.png)

**Fig. 8.** Performance of rate 1/2 protograph codes with input block sizes 64, 128, and 256.

The minimum distance of the \( k = 64, n = 128 \) code is \( d_{min} = 12 \). The maximum undetected error rate over all SNRs for this code is \( 2 \times 10^{-4} \). To increase the minimum distance, and further to decrease the maximum undetected error rate, next we consider a protograph with degrees 3 and 5. The smallest protograph with 2 nodes is shown in Fig. 9.

![Fig. 9](attachment:fig9.png)

**Fig. 9.** Protograph of an LDPC code with variable node degrees 3 and 5.

Again this small protograph is first lifted by a factor of 4. The lifted protograph is shown in Fig. 10. The resulting graph was then lifted by a factor of 16 using progressive edge growth [17] and the ACE algorithm [14] to find phases associated with circulants of size 16. The minimum distance of this rate 1/2 (128,64) code is 14.

![Fig. 10](attachment:fig10.png)

Performance of our (3,5) rate 1/2 protograph code with \( k = 64, 128, 256 \) is shown in Fig. 11. Also shown for comparison is the performance of the (512,256) AR4JA code and a (512,256) regular LDPC code proposed in [22]. These two codes are 0.25 dB better than our (3,5) rate 1/2 (512,256) protograph code in the waterfall region, but our proposed code does better in the error floor region.

![Fig. 11](attachment:fig11.png)

**Fig. 11.** Performance comparison of protograph codes with input block sizes 64, 128, 256.

Fig. 12 shows undetected word error rate (UER) of the short block AR4JA code if the block size \( k \) is less than 256. Larger block sizes are constructed by first lifting the \( k = 64 \) code by a factor of 2 to get \( k = 128 \) and next lifting the \( k = 128 \) code by a factor of 2 to obtain the \( k = 256 \) code. The performances of the degree (3, 4) code is better than that of the short block AR4JA code.
short protograph codes with block size $k = 64$ and variable node degrees $(3,4)$ and $(3,5)$. The behavior of the undetected error rate can be explained as follows. Scale the received vector by the standard deviation, so that the decoding regions are not a function of the SNR. Then the mean of the received vector is proportional to the SNR. Suppose the decoding region of each codeword is represented as a cone. A noiseless codeword lies on the axis of the cone at high SNR far from the origin, at low SNR close to the origin. The noisy received vectors form a unit sphere around the noiseless codeword. The undetected error rate is the intersection of this unit sphere with the decoding cones of the other codewords. Intuitively, the UER goes to a very small value as the SNR goes to zero or is large. The intersection with nearest-neighbor cones is largest at some moderate SNR [21] (see Fig. 13).

At asymptotically low SNR, the channel symbols are nearly useless, and the UER will be the probability that a totally random channel symbol vector happens to be decodable to some (any) codeword. This is a low value, since the vast majority of the space of received channel symbols is undecodable. At medium SNR: the channel symbols are now nearer to the correct codeword, but they are also quite near the nearest neighbors, so the UER is higher than in the asymptotically-low SNR region. At high SNR: the UER goes to zero as the SNR grows unbounded.

Finally, we show the performance of various codes compared with the sphere packing bound in Fig. 14.
VI. CONCLUSION

In this paper we introduced a new construction technique for designing LDPC codes for short block lengths. These codes exhibit good threshold performance, good minimum distance and low undetected error rates.

ACKNOWLEDGMENT

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES