Construction of Protograph LDPC Codes with Linear Minimum Distance

Dariush Divsalar
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109-8099
Email: Dariush.Divsalar@jpl.nasa.gov

Sam Dolinar
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109-8099
Email: sam@shannon.jpl.nasa.gov

Christopher Jones
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109-8099
Email: christop@jpl.nasa.gov

***Abstract*— A construction method for protograph-based LDPC codes that simultaneously achieve low iterative decoding threshold and linear minimum distance is proposed. We start with a high-rate protograph LDPC code with variable node degrees of at least 3. Lower rate codes are obtained by splitting check nodes and connecting them by degree-2 nodes. This guarantees the linear minimum distance property for the lower-rate codes. Excluding checks connected to degree-1 nodes, we show that the number of degree-2 nodes should be at most one less than the number of checks for the protograph LDPC code to have linear minimum distance. Iterative decoding thresholds are obtained by using the reciprocal channel approximation. Thresholds are lowered by using either precoding or at least one very high-degree node in the base protograph. A family of high- to low-rate codes with minimum distance linearly increasing in block size and with capacity-approaching performance thresholds is presented. FPGA simulation results for a few example codes show that the proposed codes perform as predicted.**

I. Introduction

Low-density parity-check (LDPC) codes were proposed by Gallager [1] in 1962. After introduction of turbo codes by Berrou et al [2] in 1993, researchers revisited LDPC codes, and extended the work of Gallager using the code graphs introduced by Tanner [3] in 1981. After 1993 there have been many contributions to the design and analysis of LDPC codes; see for example [10], [12], [4], [13], [14], [15], and references there. 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 [9] and [11].

II. Protograph LDPC codes

For high-speed decoding, it is advantageous for an LDPC code to be constructed from a protograph [7] or projected graph [8]. A protograph is a Tanner graph with a relatively small number of nodes. A "copy-and-permute" operation [7] 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 π is chosen to be decomposable into permutations along each edge of the protograph. The iterative decoding 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.

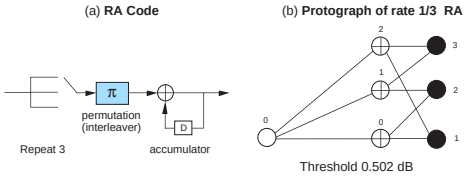


Fig. 1. (a) A rate-1/3 RA code with repetition 3, and (b) its corresponding protograph.

Repeat-Accumulate (RA) [5], Irregular Repeat-Accumulate (IRA) [6], and recently proposed Accumulate-Repeat-Accumulate (ARA) [16] 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 in this paper follows a fast and accurate approximation to density evolution originally proposed in [22]. Less than 0.005 dB deviations from true density evolution thresholds have been observed by the application of this approximation

to protograph codes over binary-input additive white Gaussian noise (BI-AWGN) channels.

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 $\vec{s}_e$ are passed along edges leaving variable nodes ($\vec{s}_e = s_{chan}$ from transmitted nodes and $\vec{s}_e = 0$ from punctured nodes). The transformation $R(\vec{s}_e)$ is applied and an extrinsic return message, $\overleftarrow{r}_e$, is determined by computing the sum of all incoming messages save the one along edge e. Transformation $R(\cdot)$ is then reapplied to produce $\overleftarrow{s}_e$. The process continues and a threshold is determined by the smallest value of s_{chan} for which unbounded growth of all messages $\vec{s}_e$ can be achieved.

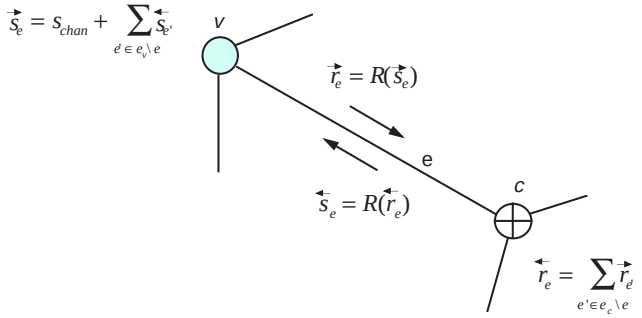


Fig. 2. The reciprocal channel approximation in use on a protograph.

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 [22] when applied to the binary erasure channel (BEC). In the case of a BEC with erasure probability ϵ 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 $C(s)+C(r)=1$.

IV. Precoded Protograph LDPC codes

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 1.11 dB while their ensemble asymptotic minimum distance grows like $0.023n$ as n goes to infinity. For comparison the asymptotic minimum distance of random codes grows as $0.11n$. We express the normalized logarithmic asymptotic weight distribution of a code as $r(\delta)=\frac{\ln(A_d)}{n}$ where d is Hamming weight, $\delta=\frac{d}{n}$, and A_d is the ensemble weight distribution. If $r(\delta)$ starts out negative near $\delta=0$ and has a first zero crossing at $\delta=\delta_{min}>0$, then the typical minimum distance of the code ensemble is $d_{min}=n\delta_{min}$, which grows linearly with n at the rate δ_{min}. This growth rate δ_{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 [18] and [19].

Precoding places a constraint node between a degree-1 variable node and a higher degree variable, which is then optionally erased. Precoding often lowers the iterative decoding threshold of a given protograph without altering its rate [16]. Precoding is generally most useful at lower rates, because iterative decoding thresholds for very high-rate regular LDPC codes are already satisfactory.

Fig. 3 compares the asymptotic weight distribution of (3,6) regular LDPC codes with and without precoding to that of rate-1/2 random codes. Precoding improves both the iterative decoding threshold and the asymptotic growth rate of d_{min}.

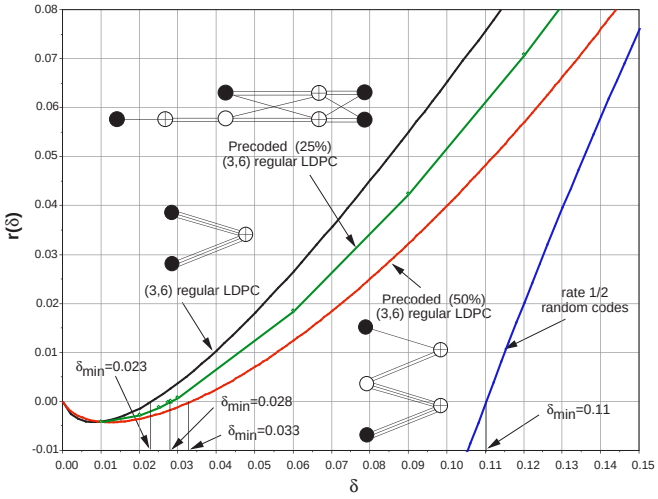


Fig. 3. Asymptotic weight distributions and zero crossings for (3,6) regular LDPC with: no precoding, 25 percent precoded nodes, 50 percent precoded nodes, and rate 1/2 random codes.

V. A method to construct protographs with $\delta_{min}>0$ despite having degree-2 variable nodes

Computation of ensemble weight enumerators for protograph LDPC codes [19] requires knowledge of the partial weight enumerator $A_{w_1,w_2,\ldots,w_m}$ for every check with degree m in the protograph. Any degree-m check node can be split into an equivalent subgraph with two check nodes of degree m_1+1 and m_2+1 connected by a degree-2 variable node, such that $m_1+m_2=m$. Figure 4 shows a degree-m check and its equivalent representation for partial weight enumeration. The partial weight enumerator for the check with degree m, expanded to represent m binary sequences each of length N, is obtained from the partial weight enumerators of the two

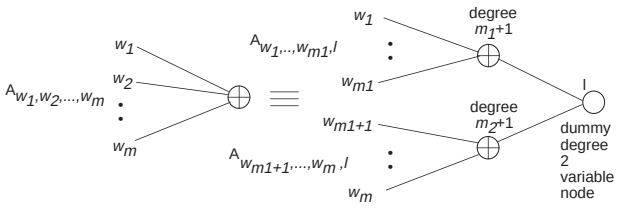


Fig. 4. Degree-m check and its equivalent representation.

checks with degrees $m_1 + 1$ and $m_2 + 1$ as

$$A^c_{w_1,w_2,\ldots,w_m} = \sum_{l=1}^{N} \frac{A_{w_1,\ldots,w_{m_1},l} A_{w_{m_1+1},\ldots,w_m,l}}{\binom{N}{l}} \quad (1)$$

We use this idea to construct protograph LDPC codes that include degree-2 variable nodes to achieve good iterative decoding thresholds, yet also have minimum distance growing linearly with block size. Start with a high-rate protograph LDPC code where the degrees of all variable nodes are at least 3. We know that such a code ensemble has minimum distance that grows linearly with block size. Next we split a check node in the protograph into two checks and distribute the total number of edges into the original check between the two new checks, and then connect these two checks with a non-transmitted degree-2 variable node. The resulting protograph has one additional check node and one new non-transmitted degree-2 variable node. The corresponding protograph LDPC code ensemble will have the same average weight enumerator, and so its ensemble minimum distance will grow linearly with block size with the same linearity coefficient. The overall code rate remains the same. Finally, if we change the new degree-2 variable node from an untransmitted node to a transmitted node, we obtain a lower-rate code, but the property that the ensemble minimum distance grows linearly with block size will be preserved. We can continue spliting additional check nodes to generate lower-rate protograph LDPC codes.

Figure 5 shows an example of such a construction to obtain a rate-1/2 AR4JA code [17], starting with a rate-2/3 code. In the last step of this construction we also attached an accumulator as a precoder to lower the iterative decoding threshold. The iterative decoding threshold for this rate-1/2 code is 0.64 dB, and the asymptotic growth rate of the ensemble minimum distance is $\delta_{min} = 0.015$.

After using our new check node splitting technique to design a particular low-rate code such as the rate-1/2 AR4JA code in Fig. 5 with minimum distance guaranteed to grow linearly with block size, this property will be preserved if we attach additional variable nodes of degree-3 and higher to this low-rate protograph. Thus, we can conclude that the entire AR4JA family described in [17] for rates $r = (n+1)/(n+2), n = 0, 1, 2, \ldots$, has ensemble minimum distance growing linearly with block size. Protographs for this AR4JA family are shown in Fig. 6. The thresholds achieved by this family compared to the corresponding capacity limits are also shown in Fig. 6.

The next example illustrates that precoding is not essential for constructing a code having both a low iterative decoding

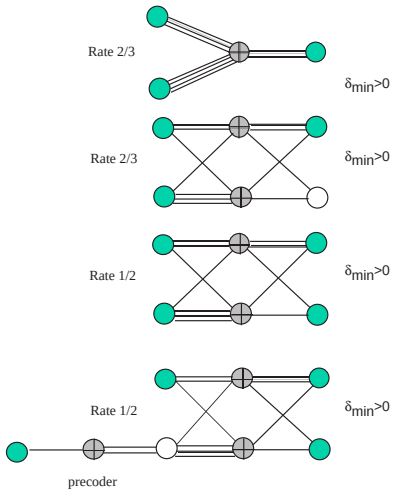
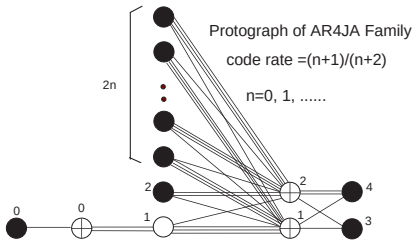


Fig. 5. Constructing rate 1/2 AR4JA LDPC code from a rate 2/3 protograph LDPC code.



Code Rate	Protograph Threshold	Capacity	Difference
1/2	0.628	0.187	0.441
2/3	1.450	1.059	0.391
3/4	2.005	1.626	0.379
4/5	2.413	2.040	0.373
5/6	2.733	2.362	0.371
6/7	2.993	2.625	0.368
7/8	3.209	2.845	0.364

Fig. 6. Protograph of AR4JA family with rates 1/2 and higher.

threshold and linearly growing minimum distance. Instead a high-degree variable node is used to lower the iterative decoding threshold. We start with a rate-4/5 code and apply check node splitting to obtain rate-2/3 and rate-1/2 codes as shown in Fig. 7. Reversing the construction process, we note that the higher-rates protographs in Fig. 7 can be obtained by simply puncturing some of the degree-2 nodes of the rate-1/2 protograph.

Note that the rate-1/2 code in Fig. 7 has $\lambda^{'}(0)\rho^{'}(1) = 1.37$. Thus the code does not satisfy the relation $\lambda^{'}(0)\rho^{'}(1) < 1$ [20], where $\lambda(x), \rho(x)$ are the degree distributions for variable and constraint nodes. For a protograph code this condition is a sufficient, but not a necessary, condition for minimum distance growing with n. Specifically, the ensemble asymptotic minimum distance over block size for this protograph is a small but positive number, $\delta_{min} = 0.005$.

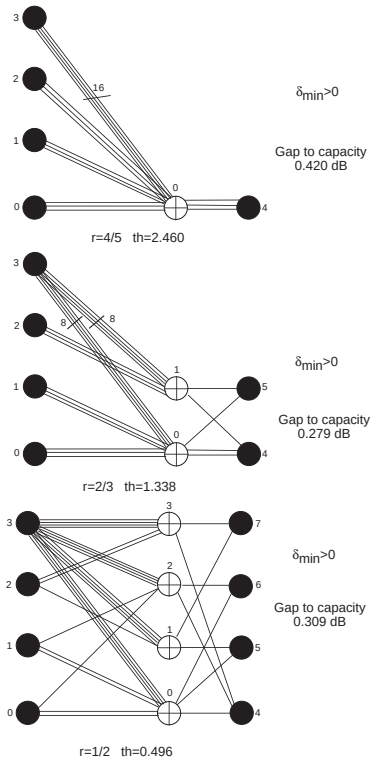


Fig. 7. Constructing rate-2/3 and rate-1/2 codes from a rate-4/5 protograph LDPC code.

VI. Generality of the Check Node Splitting Technique for Constructing Protographs with Linearly Growing Ensemble Minimum Distance

Up to now we have used our check node splitting technique to generate examples of protograph codes having degree-2 variable nodes that nonetheless possess the desirable property that their ensemble minimum distance grows linearly with block size. In this section, we tackle the converse problem and identify a condition which, if satisfied by a given protograph having degree-2 variables, will insure that the corresponding ensemble minimum distance grows linearly with block size.

Definition: Let $\mathcal{P}$ denote any protograph containing only transmitted variable nodes of degree 2 and higher. Let $\mathcal{P}'$ denote the subgraph of $\mathcal{P}$ that contains only its degree-2 variables and their attached edges and checks. If this subgraph is not connected, decompose it into its disjoint connected pieces, $\mathcal{P}' = \cup_i \mathcal{P}'_i(m_i)$, where m_i is the number of degree-2 variable nodes in the connected subgraph $\mathcal{P}'_i(m_i)$. We say that the original protograph $\mathcal{P}$ satisfies the *check node splitting condition* if each connected subgraph $\mathcal{P}'_i(m_i)$ has exactly $m_i + 1$ checks.

Lemma: If and only if a protograph $\mathcal{P}$ satisfies the check node splitting condition, then it can be constructed by our node splitting technique starting with a higher-rate ancestral protograph $\mathcal{P}^+$ having only transmitted variable nodes of degree 3 and higher.

Theorem: If a protograph $\mathcal{P}$ satisfies the check node splitting condition, then $\mathcal{P}$ will inherit the property from its ancestral protograph $\mathcal{P}^+$ that its ensemble minimum distance grows linearly with block size.

Proof: The theorem follows from the lemma by the arguments given in the previous section. The proof of the lemma is trivial in one direction. Every check node i in the ancestral protograph $\mathcal{P}^+$ that is split m_i times in succession will generate a corresponding subgraph $\mathcal{P}'_i(m_i)$ of m_i degree-2 variable nodes and m_i+1 attached checks, because one new check node and one new degree-2 variable node are created with each split. Furthermore, this subgraph $\mathcal{P}'_i(m_i)$ will be disconnected from the subgraph $\mathcal{P}'_j(m_j)$ created by splitting any other check node $j \neq i$ in the ancestral protograph $\mathcal{P}^+$.

To prove the converse, it is sufficient to show that each connected subgraph $\mathcal{P}'_i(m_i)$ of m_i degree-2 variable nodes and $m_i + 1$ attached check nodes can be derived by applying the node splitting technique to a single check node in an ancestral protograph $\mathcal{P}^+$. First select any check node of degree 1 within the subgraph $\mathcal{P}'_i(m_i)$. This is always possible, because there are more checks than variables in $\mathcal{P}'_i(m_i)$, and all variables in this subgraph are degree-2. The single degree-2 variable node attached to the selected check node is also attached to one other uniquely determined check node in $\mathcal{P}'_i(m_i)$. This second check node must have degree ≥ 2 in $\mathcal{P}'_i(m_i)$, because otherwise the subgraph consisting of these two checks and their connecting variable would be disconnected from the remainder of $\mathcal{P}'_i(m_i)$. Merge these two checks and delete the connecting variable. The result is a subgraph $\mathcal{P}'_i(m_i - 1)$ consisting of $m_i - 1$ degree-2 variables and m_i checks. This subgraph $\mathcal{P}'_i(m_i-1)$ is also connected, since the second check included in the merger from $\mathcal{P}'_i(m_i)$ must have been connected to the remainder of the graph. By the same procedure applied to $\mathcal{P}'_i(m_i)$, merge another pair of checks from $\mathcal{P}'_i(m_i-1)$ and delete the connecting variable to obtain a smaller subgraph $\mathcal{P}'_i(m_i - 2)$ with the same properties. Continue this process until obtaining $\mathcal{P}'_i(0)$ consisting of one check node and zero degree-2 variable nodes. Finally, reverse this process starting with $\mathcal{P}'_i(0)$ to obtain the original connected subgraph $\mathcal{P}'_i(m_i)$. Each check node merger and variable node deletion is reversed by a corresponding check node split and new variable node creation and attachment. ■

Note that while the lemma provides an if-and-only-if condition for determining whether our node splitting method can be applied to prove asymptotically growing ensemble minimum distance, the theorem does not rule out the possibility of protographs failing this condition for which linearly growing minimum distance could be proved by other means.

VII. Simulation results

Fig. 8 shows bit and frame error rate simulation results for LDPC codes with dimension $k = 4096$ expanded from the rate-1/2 through rate-4/5 AR4JA protographs in Fig. 6 for which precoding was used to lower the decoding threshold. Fig. 9 shows simulation results for LDPC codes with $k = 3680$ built from the rate-1/2 protograph in Fig. 7, for which a high-degree node was used in the base protograph. The rate-2/3 and rate-4/5 protographs in Fig. 9 are obtained by

puncturing the rate-1/2 protograph. Protographs were lifted using the ACE algorithm [21] to find circulants for each edge of the protograph. The simulation results are also compared with Gallager's bound [1] for random codes. All simulations were performed on a field-programmable gate array (FPGA) implementation of an LDPC decoder developed at JPL.

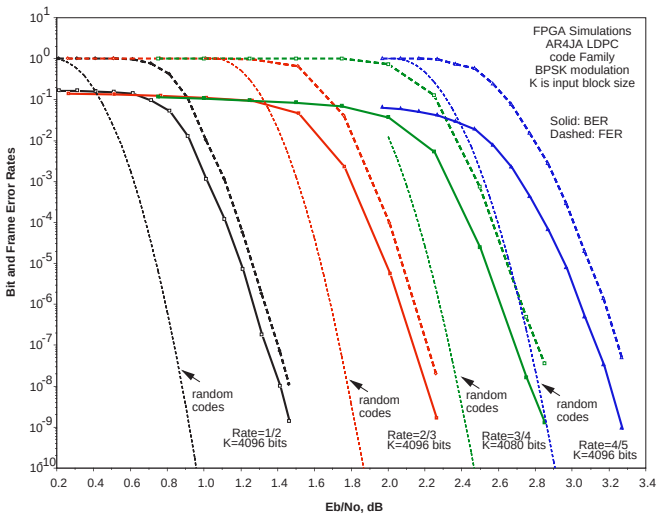


Fig. 8. Performance of AR4JA code family with input block size $k = 4096$.

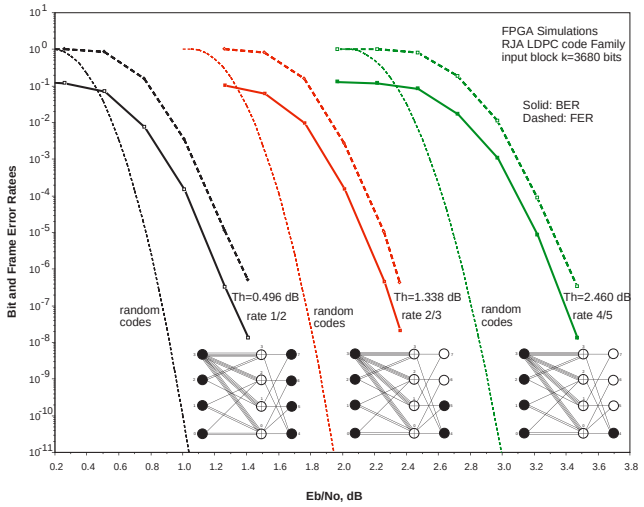


Fig. 9. Performance of LDPC codes with input block size $k = 3680$ lifted from the rate-1/2 protograph of Fig. 7, and from rate-2/3 and rate-4/5 protographs obtained by puncturing the rate-1/2 protograph.

VIII. Conclusion

In this paper we introduced a new construction technique for designing ensembles of structured codes that exhibit both good threshold performance and minimum distance that for a typical instance from the ensemble increases linearly with blocklength.

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

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

[2] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, Vol. 44, pp. 1261–1271, October 1996.

[3] M. R. Tanner, "A recursive approach to low complexity codes," IEEE Trans. Inform. Theory, vol. 27, pp. 533-547, 1981.

[4] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low- density parity-check codes," IEEE Trans. Inform. Theory, vol. 47, pp. 619-637, 2001.

[5] D. Divsalar, H. Jin, and R. McEliece, "Coding theorems for Turbo-like codes," in Proceedings of the 1998 Allerton Conference,1998.

[6] H. Jin, A. Khandekar, and R. McEliece, "Irregular repeat-accumulate codes," in Proc. 2nd International Symposium on Turbo Codes, 2000.

[7] Jeremy Thorpe, "Low Density Parity Check (LDPC) Codes Constructed from Protographs," JPL INP Progress Report 42-154, August 15, 2003.

[8] Richardson, et al., "Methods and apparatus for decoding LDPC codes," United States Patent 6,633,856, October 14, 2003.

[9] T. Richardson, "Multi-Edge Type LDPC Codes," presented at the Workshop honoring Prof. Bob McEliece on his 60th birthday, California Institute of Technology, Pasadena, California, May 24-25, 2002.

[10] D.J.C. MacKay, R.M. Neal, "Near Shannon limit performance of low density parity check codes," Electronics Letters ,Vol. 32, Issue 18, 29 Aug. 1996, Page(s) 1645.

[11] T. Richardson and R. Urbanke,"The Renaissance of Gallager's Low-Density Parity-Check Codes," IEEE Communications Magazine, pages 126-131, August 2003

[12] M. Luby, M. Mitzenmacher, A. Shokrollahi, and D. Spielman, "Analysis of low density codes and improved designs using irregular graphs," IEEE Trans. Inform. Theory, vol. 47, pp. 585-598, 2001.

[13] T. Richardson and R. Urbanke, "The capacity of low-density parity check codes under message- passing decoding," IEEE Trans. Inform. Theory, vol. 47, pp. 599-618, 2001.

[14] Y. Kou, S. Lin, and M.P.C. Fossorier, "Low-density parity-check codes based on finite geometries: a rediscovery and new results," IEEE Transactions on Information Theory, vol. 47, Nov. 2001, pp. 2711 -2736.

[15] F.R. Kschischang, "Codes defined on graphs," IEEE Communications Magazine, Vol. 41, Issue 8 , Aug. 2003, Pages 118 -125.

[16] A. Abbasfar, D. Divsalar, and K. Yao, "Accumulate Repeat Accumulate Codes," (abstract) IEEE ISIT 2004, Chicago, IL, June 27-July 2, and IEEE Globecom 2004, Dallas, Texas, 29 November - 3 December, 2004.

[17] D. Divsalar, C. Jones, S. Dolinar, J. Thorpe; Protograph Based LDPC Codes with Minimum Distance Linearly Growing with Block Size , IEEE Globecom 2005.

[18] S.L. Fogal, Robert McEliece, Jeremy thorpe "Enumerators for Protograph Ensembles of LDPC Codes," IEEE ISIT 2005, Adelaide, Australia 4-9 September,2005.

[19] D. Divsalar, "Finite Length Weight Enumerators for Protograph Based LDPC Code Ensembles," IEEE Communication Theory Workshop, Park City, Utah, June 12-15, 2005.

[20] Changyan Di; Urbanke, R.; Richardson, T.; "Weight distributions: how deviant can you be? " Proceedings, 2001 IEEE International Symposium on Information Theory, 2001. 24-29 June 2001 Page(s):50

[21] Tian, T.; Jones, C.; Villasenor, J.; and Wesel, R. D.; "Characterization and selective avoidance of cycles in irregular LDPC code construction," IEEE Transactions on Communications, Aug. 2004, pp. 1242–1247.

[22] Chung S. Y.,"On the Construction of Some Capacity-Approaching Coding Schemes," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge,Massachusetts", Sept., 2000.