A LOW COMPLEXITY FEC SCHEME BASED ON THE INTERSECTION OF INTERLEAVED BLOCK CODES

Olivier Pothier†, Loïc Brunel‡, Joseph Boutros‡

† Laboratoires d'Electronique Philips,
22, av. Descartes, BP15, 94453 Limeil-Brévannes cedex, France
‡Communications & Electronics Department
Ecole Nationale Supérieure des Télécommunications
46, Rue Barrault, 75634 Paris cedex 13, France
Fax : +33 1 45 89 00 20 - Email : {pothier,brunel,boutros}@com.enst.fr

Abstract - We describe a class of asymptotically good codes built from the intersection of randomly permuted binary BCH codes. This family of pseudo-random error correcting codes, called *Generalized Low Density* (GLD) codes, is a direct generalization of Gallager's Low Density Parity Check (LDPC) codes. GLD codes belong to the larger family of Tanner codes based on a random bipartite graph. We study the GLD ensemble performance and prove the asymptotically good property. We also compare GLD codes minimum distance and performance to the Varshamov-Gilbert bound and BSC capacity respectively. The results show that Maximum-Likelihood decoding of GLD codes achieves near capacity efficiency. The sub-optimal iterative decoding of GLD codes is briefly presented. Experimental results of small and large blocklength codes are finally illustrated on both AWGN and Rayleigh fading channels.

I. INTRODUCTION

In his Ph.D. thesis [1], Robert Gallager studied error-correcting codes based on low density parity-check matrices. Two revolutionary ideas have been exploited by Gallager :

- The use of random permutations linking simple parity codes to build an efficient low complexity code that imitates random coding.
- An iterative decoding technique where a priori information and channel observation are both used to compute a posteriori and new a priori information.

Unfortunately, Gallager's work has been forgotten by the majority of the scientific community until the recent invention of Turbo codes [2]. The unbeatable parallel concatenated recursive convolutional codes presented by Berrou, Glavieux and Thitimajshima exploit both ideas cited above. Another important work on low complexity concatenated codes has been done by Tanner [3] who made a generalization of Gallager's codes. Tanner gave three different versions of an iterative decoding algorithm but he restricted his construction to deterministic permutations derived from structured graphs.

We describe a class of *Generalized Low Density* codes by replacing simple parity codes in Gallager's LDPC construction with binary BCH codes. Each parity check equation of an LDPC (N, K) code is replaced by the parity-check matrix of a small binary BCH (n, k) code called the *constituent code*. LDPC codes are usually defined by their sparse parity-check matrix, but they can also be described with a bipartite graph which is a Tanner representation of the whole code. Thus, a GLD code can be seen as a Tanner code built on a random graph where the constituent code C_0 is associated to each subcode node.

In the next section, the GLD code is defined by a graphical representation showing that the whole code is equal to the intersection of interleaved constituent codes. The GLD code parity-check matrix representation is given in [4]. GLD codes exhibit an excellent performance on both AWGN and Rayleigh channels and present a high BER slope at high SNR due to their large minimum distance. The ensemble performance is studied in Section III where GLD codes are proved to have a minimum Hamming distance proportional to

0-7803-5565-2/99/$10.00 © 1999 IEEE

their blocklength, i.e. GLD codes are asymptotically good. Since the capacity study on AWGN channels is intractable, we compared the performance limit of GLD codes to the BSC capacity limit. GLD codes efficiency is very close to capacity limit when maximum-likelihood decoding is applied.

The practical decoding of GLD codes is performed iteratively as briefly explained in section IV. The iterative decoding algorithm is based on a SISO (Soft Input, Soft Output) decoding of the small constituent codes and has a very low complexity. The decoding time can be dramatically reduced when all forward-backward procedures in the SISO decoding of the constituents are executed in parallel.

II. STRUCTURE OF THE GLD CODE

Figure 1 shows an LDPC code of length $N = 12$ and dimension $K = 3$. The code is defined by a regular bipartite graph. The left part contains 12 bit nodes (i.e. the codeword) and the right part holds 9 subcode nodes (i.e. the parity-check constraints). The 9 constituents are simple binary SPC(4,3,2) codes. Thus, a 12-bit word included in the left part is a codeword of the resulting LDPC code, if and only if the 9 right clumps of 4 incoming bits belong to the SPC code. The 36 branches of the graph are chosen randomly. This binary (12,3) LDPC code is equal to the intersection of 3 interleaved block codes, called *supercodes*, where each supercode is the direct sum of 3 SPC codes. If we take the bit ordering of the first supercode as an identity reference, only two random permutations have to be applied before the intersection. Consequently, the regular random graph of Figure 1 is equivalent to two random interleavers linking direct sums of SPC codes.

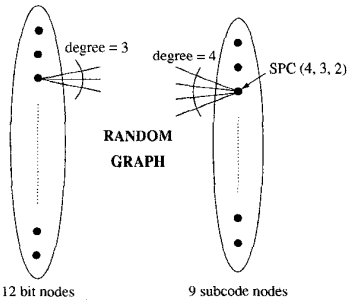

Figure 1: An LDPC code example.

Let us now describe the generalization of an LDPC code by defining a GLD code from its graphical representation given in Figure 2. The (n,n-1,2) SPC code is replaced by a $C_0(n,k)$ BCH code with a correction capacity equal to t, i.e. $d_{Hmin}(C_0) \geq 2t+1$. Note that the degree of the subcode nodes is equal to n, the length of the constituent code. The GLD code length is N bits, which is the number of nodes in the graph left part. If J denotes the degree of the bit nodes, we conclude that the right part holds JN/n identical BCH codes. Finally, an N-bit word included in the left part is a codeword of the resulting GLD code, if and only if the JN/n right clumps of n incoming bits belong to the BCH $C_0(n,k)$ code. The GLD code C graphically defined by Figure 2 can be written as follows [4]: $C = \bigcap_{j=1}^{J} C_j$, where $C_1 = C_0 \oplus \ldots \oplus C_0$ is the direct sum of N/n constituent codes, and $C_{j+1} = \pi_j(C_1)$ for $j = 1 \ldots J-1$, where π_j is a random permutation.

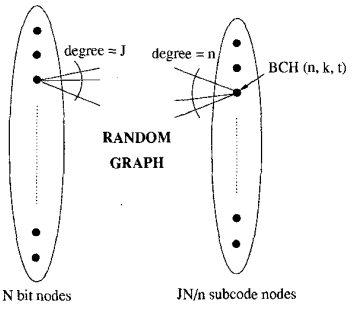

Figure 2: The GLD code graph.

It can be easily shown [3] that the coding rate of the GLD code $C(N,K)$, built from a regular bipartite graph with bit nodes of degree J and a constituent code $C_0(n,k)$, satisfies $R = K/N \geq 1 - J(1-k/n)$. The equality holds in most cases where the graph branches are selected by a random matching.

Tanner gave a weak but interesting lower bound for both the length N and the minimum distance $d_{Hmin}(C)$ function of the graph girth g. Roughly speaking, we can write for the GLD code length $N \geq [n]^{g/4}$ and for the GLD code minimum distance $d_{Hmin}(C) \geq [d_{Hmin}(C_0)]^{g/4}$. Note that a product code is defined by a complete graph ($g = 8$, $J = 2$) and satisfies $N = n^2$ and $d_{Hmin}(C) = [d_{Hmin}(C_0)]^2$. Note also that the girth is always a multiple of 4 when the subcode nodes are grouped into J sets each one having N/n nodes. So, to guarantee a minimum distance larger than or equal to $[d_{Hmin}(C_0)]^2$, the $J-1$ random interleavers π_j are chosen with no

cycles of length 4. The ensemble performance shows that in most cases the minimum distance of C is strictly greater than $[d_{Hmin}(C_0)]^2$. It approaches $[d_{Hmin}(C_0)]^3$ when the length N approaches n^3.

III. ENSEMBLE PERFORMANCE

In this section, we show how to compute the average weight distribution of the generalized low density code, and the latter is proved to be asymptotically good. We also consider a BSC with a transition probability $0 < p < 1/2$ and find the maximal value of p for which the word error probability P_e goes to zero when N goes to infinity.

Average Minimum Distance

Let us determine the average weight distribution of the ensemble of GLD codes built with a BCH code C_0 and random permutations $\pi_1, \ldots, \pi_{J-1}$. The weight coefficients are obtained by averaging over all possible interleavers π_j and depend on the moment-generating function $g(s)$ [1] (i.e. the normalized weight-enumerator polynomial) of the constituent code C_0. For example, the moment-generating function $g(s)$ of $C_0 = (7, 4, 3)$ is $g(s) = (1 + 7e^{3s} + 7e^{4s} + e^{7s})/16$.

The first supercode C_1 of length N is the direct sum of N/n independent codes C_0. Hence, its moment-generating function $G(s)$ is simply a power of $g(s)$, $G(s) = g(s)^{N/n} = \sum_\ell Q(\ell)e^{\ell s}$, where $Q(\ell)$ is the probability that a word of weight ℓ belongs to C_1. Since the total number of codewords in C_1 is $(2^k)^{N/n}$, then the average number in C_1 of codewords of weight ℓ is $N_1(\ell) = 2^{(kN/n)}Q(\ell)$. Exploiting the fact that C_1, $C_2 = \pi_1(C_1)$, ..., $C_J = \pi_{J-1}(C_1)$ are totally independent, the probability that an ℓ-weight word belongs to $C = C_1 \bigcap C_2 \bigcap \ldots \bigcap C_J$ can be written as :

$$P(\ell) = \left(\frac{N_1(l)}{\binom{N}{\ell}} \right)^J$$

Finally, the average number of codewords in C having weight ℓ is :

$$\overline{N(\ell)} = \binom{N}{\ell} \times P(\ell) = \frac{2^{(JkN/n)}Q(\ell)^J}{\binom{N}{\ell}^{J-1}}$$

By using exactly the same bounding technique as in [1], *i.e.* upper bounding each of the coefficients $Q(\ell)$ with $G(s)e^{-\ell s}$, and after applying the extended Stirling approximation (valid for large N), we get an upper bound on the average number of codewords of weight ℓ in the GLD code (details omitted) :

$$\overline{N(\ell)} \leq C(\lambda, N) \times e^{-NB(\lambda)}$$

where $\lambda = \ell/N$ is the normalized weight.
The two functions $C(\lambda, N)$ and $B(\lambda)$ are expressed as follows :

$$C(\lambda, N) = \sqrt{2\pi N\lambda(1-\lambda)}^{J-1} \times e^{(J-1)/(12N\lambda(1-\lambda))}$$

$$B(\lambda) = (J-1)H(\lambda) - \frac{J}{n}[\mu(s) + k\ log2] + Js\lambda$$

where $H(\lambda)$ is the natural entropy function and $\mu(s) = \log(g(s))$. The upper bound has been optimized and the optimal value of s is related to the weight by $\lambda = \mu'(s)/n$, where $\mu'(s)$ is the derivative of $\mu(s)$ relative to s.

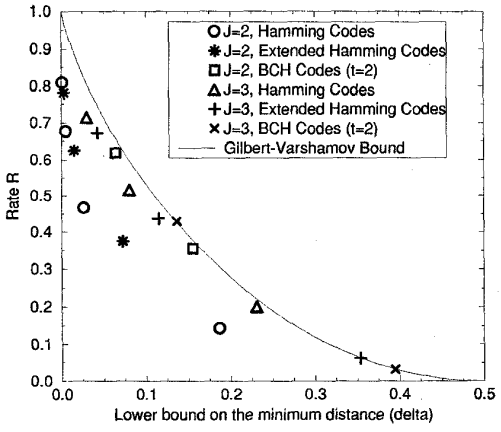

Figure 3: GLD rate vs minimum distance.

Asymptotically, when $N \to \infty$, the average number $\overline{N(\ell)}$ of codewords of weight ℓ goes to zero if $B(\lambda) > 0$. The smallest value δ of $\lambda \in]0 \ldots 1/2[$ satisfying $B(\delta) = 0$ and $B(\lambda) > 0$ when $\lambda \in]0 \ldots \delta[$, gives us a *lower bound* on the minimum distance $\delta(C) = d_{Hmin}(C)/N$ of the GLD code. Thus, C is **asymptotically good** when δ exists. This is the case when C_0 is a BCH code for any $J \geq 2$. Figure 3 shows the rate versus the lower bound of $\delta(C)$ for different constituent codes. These points are not far from the Varshamov-Gilbert bound and get close enough when $J = 3$ (intersection of 3 supercodes) or $t = 2$ (double error-correcting BCH codes).

Near capacity performance

Now, let us compute the maximal value of p for which the word error probability P_e of an ML decoder goes to zero when N is arbitrarily large. An upper bound on P_e is obtained by assuming that a decoding error occurs when at least half of the codeword non zero bits are covered. If j denotes the channel error weight, ℓ the weight of a codeword and i the number of covered non zero bits, we have the following upper bound :

$$P_e \leq \sum_{\substack{j=1 \\ \ell=d_{Hmin}}}^{N} p^j (1-p)^{N-j} \overline{N(\ell)} \sum_{i=\ell/2}^{\ell} \binom{\ell}{i} \binom{N-\ell}{j-i}$$

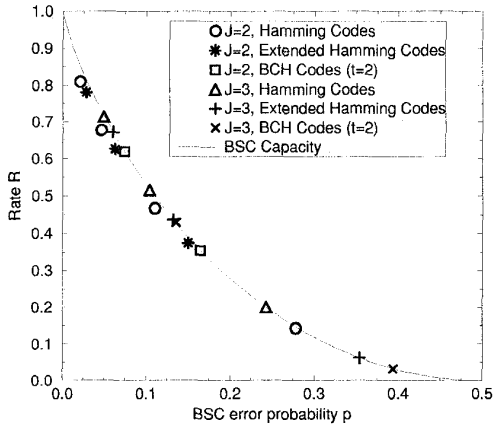

Figure 4: GLD rate vs BSC error probability.

When N is large enough, an expression similar to the upper bound on $\overline{N(\ell)}$ can be found for P_e. After some algebraic manipulations, we get :

$$P_e \leq D(N,p) \times e^{-NE(p)}$$

where the exponent function $E(p)$ is given by :

$$E(p) = Min_\lambda \left[B(\lambda) + H(p) - \lambda log2 - (1-\lambda) H(\frac{p-\lambda/2}{1-\lambda}) \right]$$

We conclude that an ML decoder for this GLD code achieves $P_e \to 0$ if $p < p_{th}$ where the threshold p_{th} is the first value corresponding to a sign transition of $E(p)$. Figure 4 shows the rate versus p_{th} for different GLD codes. We conclude that all GLD codes achieve near capacity performance when their length is arbitrarily large.

IV. ITERATIVE DECODING

The iterative decoding of a GLD code with one interleaver π linking $J = 2$ supercodes is briefly described below. The first supercode C_1 is called *upper code* and the second supercode $C_2 = \pi(C_1)$ is called *lower code.* If $J > 2$, the information propagation strategy from one supercode to another has a great influence on the iterative decoding performance.

Let $r = (r_1, \ldots, r_N)$ denote the channel output. The decoder starts by converting r into a conditional probability vector $y = (p(r_1|0), \ldots, p(r_N|0))$. The decoder structure (one iteration step) is illustrated in Figure 5. Two identical SISO decodings are performed in one iteration (similar to Turbo decoding [2]), the SISO decoding of the upper code followed by the SISO of the lower one. At iteration i, each SISO reads the observation y and the a priori information generated by the preceding decoder. $P_\ell(i-1)$ is the a priori (extrinsic) information generated by the lower SISO at time $i-1$ and $P_u(i)$ is the a priori generated by the upper decoder at time i. The a posteriori probabilities APP of the N coded bits are also computed by the SISO.

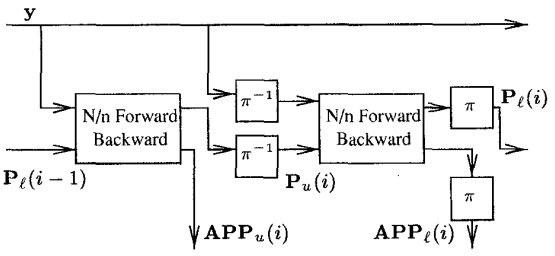

Figure 5: One decoding iteration.

Notice that a turbo decoder has two forward-backward algorithms [5], one for each supercode. In the GLD code case, since $C_1 = C_0 \oplus \ldots \oplus C_0$, the SISO decoding of the upper code is performed by N/n parallel low complexity forward-backward algorithms applied to the trellis of the constituent code C_0. The same procedure is valid for the lower code $\pi(C_1)$.

V. SIMULATION RESULTS

Two GLD codes have been tested for an additive white Gaussian noise channel (AWGN) and a Rayleigh channel with BPSK modulation.

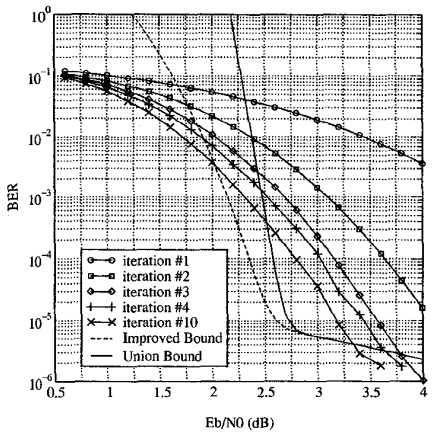

Figure 6: GLD code on AWGN channel : length $N = 420$, rate $R = 0.466$, $C_0 = (15, 11)$

The first code, suitable for mobile radio transmissions or small frame systems, has length $N = 420$ and dimension $K = 196$. Its constituent code is the (15,11,3) BCH code and the graph degree is $J = 2$. The performance on AWGN channel is shown in Figure 6 for different iteration steps. The bit error rate is compared to the ML Union Bound and to the improved Gallager bound [6]. Figure 7 shows its performance on the Rayleigh fading channel. The loss caused by the fading is about 2.5dB.

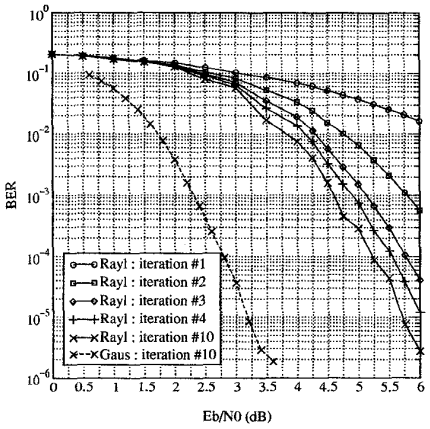

Figure 7: GLD code on Rayleigh channel : length $N = 420$, rate $R = 0.466$, $C_0 = (15, 11)$

The second code, suitable for deep space communications or image transmissions, has length $N = 2 \times 65536$. Figure 8 shows its BER versus the signal-to-noise ratio per bit. The performance of a rate 1/2 Turbo code (punctured PCCC with octal generators 23,35) having the same length is plotted on the same figure. The GLD code is 0.23dB away from the Turbo code.

ACKNOWLEDGMENT

We would like to thank Gilles Zémor for all what he taught us on Tanner codes.

References

[1] R.G. Gallager: Low-density parity-check codes, MIT Press, 1963.

[2] C. Berrou, A. Glavieux, P. Thitimajshima : "Near Shannon limit error-correcting coding and decoding : turbo-codes," *Proceedings of ICC'93*, Genève, pp. 1064-1070, May 1993.

[3] R.M. Tanner: "A recursive approach to low complexity codes", *IEEE Trans. on Inf. Theory*, Vol. IT-27, Sept 1981.

[4] J. Boutros, O. Pothier, G. Zémor : "Generalized Low Density (Tanner) Codes," *Proceedings of ICC'99*, Vancouver, June 1999.

[5] L.R. Bahl, J. Cocke, F. Jelinek and J. Raviv : "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Trans. on Inf. Theory*, Vol. IT-20, pp. 284-287, March 1974.

[6] A.J. Viterbi, J.K. Omura : Principles of digital communications and coding, McGraw-Hill, 1979.

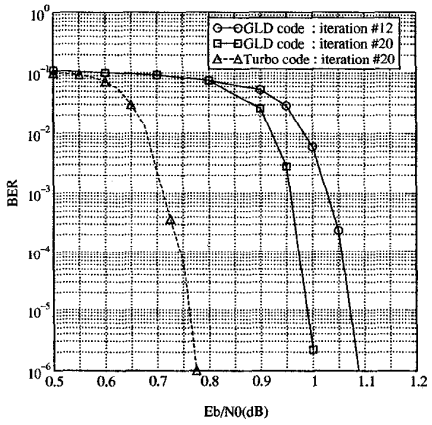

Figure 8: GLD code on AWGN channel : length $N = 2 \times 65536$, rate $R = 0.466$, $C_0 = (15, 11)$