



# LUND UNIVERSITY

## Chained Gallager codes

Zyablov, Victor; Hug, Florian; Johannesson, Rolf

2009

[Link to publication](#)

*Citation for published version (APA):*

Zyablov, V., Hug, F., & Johannesson, R. (2009). *Chained Gallager codes*. Paper presented at International Symposium on Problems of Redundancy in Information and Control Systems.

*Total number of authors:*

3

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# Chained Gallager Codes

Victor Zyablov<sup>1</sup>, Florian Hug<sup>2</sup>, and Rolf Johannesson<sup>2</sup>

<sup>1</sup>Inst. for Information Transmission Problems  
Russian Academy of Sciences  
Moscow 101447, Russia  
Email: zyablov@iitp.ru

<sup>2</sup>Dept. of Electrical and Information Technology  
Lund University  
P. O. Box 118, SE-22100 Lund, Sweden  
Email: {florian, rolf}@eit.lth.se

## Abstract

The ensemble of regular Low-Density Parity-Check (LDPC) codes introduced by Gallager is considered. Using probabilistic arguments a lower bound on the normalized minimum distance is derived. Chained Gallager codes are introduced as a combination of two Gallager codes and their error correcting capabilities are studied.

## I. TWO TRANSMISSION SCHEMES

Consider two different binary transmission schemes for communication over two parallel, independent channels. In the first transmission scheme, denoted  $\mathcal{T}_1$  and illustrated in Fig. 1(a), the message sequence  $\mathbf{u}_1$  is split into two parts  $\mathbf{u}_1^{(1)}$  and  $\mathbf{u}_1^{(2)}$ , that is,  $\mathbf{u}_1 = \mathbf{u}_1^{(1)}\mathbf{u}_1^{(2)}$ , where  $\mathbf{u}_1^{(1)}$  and  $\mathbf{u}_1^{(2)}$  have equal length. The messages  $\mathbf{u}_1^{(i)}$ ,  $i = 1, 2$ , are encoded separately by the codes  $\mathcal{C}_1^{(i)}$ ,  $i = 1, 2$ , both of rate  $R$ . Then, the two codewords  $\mathbf{v}_1^{(1)}$  and  $\mathbf{v}_1^{(2)}$  are transmitted over independent channels before the received sequences  $\mathbf{r}_1^{(1)}$  and  $\mathbf{r}_1^{(2)}$  are decoded. Finally, the two partial message sequences  $\hat{\mathbf{u}}_1^{(1)}$  and  $\hat{\mathbf{u}}_1^{(2)}$ , decided by the two decoders, are combined into the decided message sequence  $\hat{\mathbf{u}}_1 = \hat{\mathbf{u}}_1^{(1)}\hat{\mathbf{u}}_1^{(2)}$ .

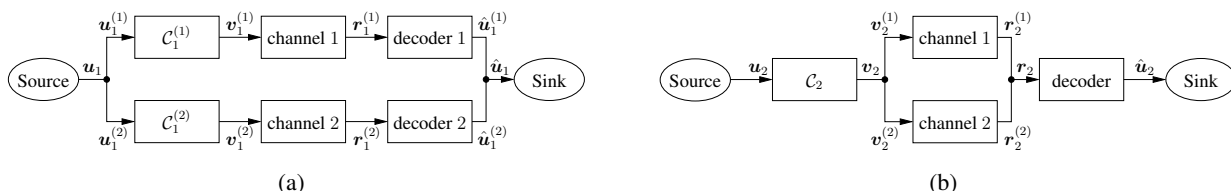


Fig. 1. Transmission scheme  $\mathcal{T}_1$  (a) and  $\mathcal{T}_2$  (b) with two parallel, independent channels.

In the second transmission scheme denoted  $\mathcal{T}_2$  and shown in Fig. 1(b), the message sequence  $\mathbf{u}_2$ , where  $\mathbf{u}_2$  is twice as long as  $\mathbf{u}_1^i$ ,  $i = 1, 2$ , is encoded by a single encoder  $\mathcal{C}_2$  of rate  $R$ . Clearly, the encoding matrix for  $\mathcal{T}_2$  has twice as many rows and columns as each of the two encoding matrices for  $\mathcal{T}_1$ . The obtained codeword  $\mathbf{v}_2$  is split into two partial codewords  $\mathbf{v}_2^{(1)}$  and  $\mathbf{v}_2^{(2)}$ , that is,  $\mathbf{v}_2 = \mathbf{v}_2^{(1)}\mathbf{v}_2^{(2)}$ , where  $\mathbf{v}_2^{(1)}$  and  $\mathbf{v}_2^{(2)}$  have equal length, and are transmitted over parallel, independent channels. Finally, the two received sequences  $\mathbf{r}_2^{(1)}$  and  $\mathbf{r}_2^{(2)}$  are concatenated into  $\mathbf{r}_2 = \mathbf{r}_2^{(1)}\mathbf{r}_2^{(2)}$  and decoded by a single decoder to obtain the decided message sequence  $\hat{\mathbf{u}}_2$ .

## II. GALLAGER CODES

Considering large encoding matrices, it is not far-fetched to use Low-Density Parity-Check (LDPC) codes due to their low decoding complexity. We will consider the Gallager ensemble of binary LDPC  $(j, k)$ -regular codes of block length  $n$  [1]. The parity-check matrix of such a Gallager code has  $j$  ones in each column and  $k$  ones in each row. An integer  $m$  is chosen such

that  $n = km$  and  $l = jm$  denote the total number of columns and rows of its parity-check matrix  $H$ , respectively.

Combining  $k$  identity matrices, the so-called *first layer* parity-check matrix  $H^*$  is obtained, that is,

$$H^* = \underbrace{\left( I_m \quad I_m \quad \dots \quad I_m \right)}_{k \text{ times}} \quad (1)$$

where  $I_m$  is the  $m \times m$  identity matrix. The parity-check matrix  $H^*$  has a single one in each column and  $k$  ones in each row, dimensions  $m$  and  $km$ , and corresponds to a  $(1, k)$ -regular LDPC code with rate  $R = 1 - \frac{1}{k}$ .

Combining  $j$  of the  $\binom{km}{k}$  column permutations of the first layer parity-check matrix  $H^*$  row-wise, we obtain the parity-check matrix  $H$  as

$$H = \left( H^* \Pi_1 \quad H^* \Pi_2 \quad \dots \quad H^* \Pi_j \right)^T \quad (2)$$

with different permutation matrices  $\Pi_i$ ,  $i = 1, 2, \dots, j$ . Such a parity-check matrix has  $j$  ones in each column and  $k$  ones in each row. The dimensions are given by  $l$  and  $n$  with  $l = jm$  and  $n = km$  and a (design) rate<sup>1</sup> of  $R = 1 - \frac{j}{k}$ .

Hereinafter we will consider the ensemble  $\mathcal{C}(n, j, k)$  of LDPC Gallager codes where the first layer parity-check matrix is given by (1) and the permutation matrices  $\Pi_i$ ,  $i = 1, 2, \dots, j$ , are chosen randomly and independently of each other [2].

### III. A LOWER BOUND ON THE MINIMUM DISTANCE

Following Gallager's approach [1], it can be shown, that the probability of randomly chosen sequence of length  $n$  and weight  $w$  being a valid codeword of  $\mathcal{C}(n, j, k)$  is given by

$$\left( \frac{N(w)}{\binom{n}{w}} \right)^j \binom{n}{w} \quad (3)$$

where  $N(w)$  is the number of codewords of weight  $w$  of the first layer parity check matrix  $H^*$  and  $1/\binom{n}{w}$  is the probability of randomly chosen sequence of length  $n$  and weight  $w$ . We conclude that as long as the sum over (3) from  $w = 2$  to  $w_0$  is smaller than 1, there exists a Gallager code among  $\mathcal{C}(n, j, k)$ , whose minimum distance is at least  $w_0 + 1$ . Upper-bounding  $N(w)$  by  $s^{-w} E(s)$ , where  $E(s)$  is the generating function of the number of codewords, and exploiting the regular structure of the first layer parity-check matrix we obtain,

$$N(w) < \min_{s>0} \left\{ s^{-w} \left( \frac{(1+s)^k + (1-s)^k}{2} \right)^m \right\}. \quad (4)$$

By combining (3) and (4), we conclude that, as long as

$$f(\omega) \simeq j \min_{s>0} \left\{ -\omega \log(s) + \frac{1}{k} \left( \log \left( (1+s)^k + (1-s)^k \right) - 1 \right) \right\} - (j-1)h(\omega) \quad (5)$$

with  $\omega = \frac{w}{n}$ , is negative, there exists a code among  $\mathcal{C}(n, j, k)$  with a normalized minimum distance greater than or equal to  $\omega$ . We will summarize this in the following theorem:

*Theorem 1:* Given the ensemble of Gallager codes  $\mathcal{C}(n, j, k)$  with its parity-check matrices having  $j$  ones per column,  $k$  ones per row, and (design) rate of  $R = 1 - \frac{j}{k}$ . The dimensions of

<sup>1</sup>The actual code rate may be slightly greater than the design rate since there may exist linear dependent rows in  $H$ .

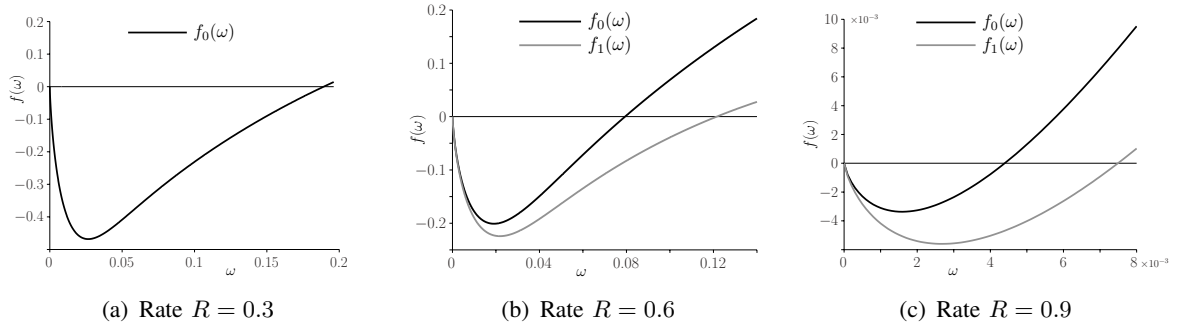


Fig. 2. Obtained lower bound on the normalized minimum distance for Chained Gallager codes ( $f_0(\omega)$ ) of different rates  $R$  compared to one of its underlying Gallager code ( $f_1(\omega)$ ) of lower rate, normalized by the same length  $n$ .

these parity-check matrices are given by  $l$  and  $n$ , where  $l = jm$  and  $n = km$ . Denoting the largest zero of  $f(\omega)$  by  $\omega_0$ , the normalized minimum distance of the ensemble of Gallager codes  $\mathcal{C}(n, j, k)$  is lower bounded by  $\omega_0$ . That is, among  $\mathcal{C}(n, j, k)$ , there exists a Gallager code whose normalized minimum distance is greater than or equal to  $\omega_0$ .  $\square$

Applying this lower bound to two previously introduced transmission schemes we have the following theorem:

*Theorem 2:* Consider the previously introduced transmission schemes  $\mathcal{T}_1$  and  $\mathcal{T}_2$ . Let  $\mathcal{C}_1^{(1)}, \mathcal{C}_1^{(2)} \in \mathcal{C}(n, j, k)$ , and  $\mathcal{C}_2 \in \mathcal{C}(2n, j, k)$ , all with the same rate  $R = 1 - \frac{j}{k}$ . More error patterns can be corrected by the transmission scheme  $\mathcal{T}_2$  than by scheme  $\mathcal{T}_1$ .  $\square$

#### IV. CHAINED GALLAGER CODES

As a variant of Gallager codes, we will introduce what we call *Chained Gallager codes* of rate  $R = 1 - \frac{j}{2k}$  as a combination of two Gallager codes of the lower rate  $R = 1 - \frac{j}{k}$ .

Given two parity-check matrices  $H_1, H_2 \in \mathcal{C}(n, j, k)$ , each of rate  $R = 1 - \frac{j}{k}$ , the parity-check matrix of a Chained Gallager code  $H_{\text{cg}}$  can be written as

$$H_{\text{cg}} = \begin{pmatrix} H_1 & H_2 \end{pmatrix}. \quad (6)$$

With  $2k$  ones per row and  $j$  ones per column,  $H_{\text{cg}}$  corresponds to a  $(j, 2k)$ -regular LDPC code with rate  $R = 1 - \frac{j}{2k}$ . However,  $H_{\text{cg}}$  belongs only to a subclass of  $\mathcal{C}(2n, j, 2k)$ , as each half is an independently Gallager code chosen from  $\mathcal{C}(n, j, k)$ .

Following the definition of  $\mathcal{C}(n, j, k)$ , the ensemble of randomly chosen Chained Gallager codes with  $j$  ones per column and  $k$  ones per rows is denoted by  $\mathcal{C}_{\text{cg}}(n, j, k)$ . The corresponding parity-check matrix  $H_{\text{cg}}$  has dimensions  $l$  and  $n$  such that  $l = jm$  and  $n = km$ , with its two halves belonging to  $\mathcal{C}(n/2, j, k/2)$ .

Hereinafter we denote the number of codewords that have 1s only in either the left or the right half by  $N_1(w)$ , that is,

$$N_1(w) = \left\{ \mathbf{v}_2^{(1)} \mathbf{v}_2^{(2)} \mid (w_{\text{H}}(\mathbf{v}_2^{(1)}) > 0 \wedge w_{\text{H}}(\mathbf{v}_2^{(2)}) = 0) \vee (w_{\text{H}}(\mathbf{v}_2^{(1)}) = 0 \wedge w_{\text{H}}(\mathbf{v}_2^{(2)}) > 0) \right\}. \quad (7)$$

Similarly, denote by  $N_2(w)$  the number of codewords that have 1s in both halves, that is,

$$N_2(w) = \left\{ \mathbf{v}_2^{(1)} \mathbf{v}_2^{(2)} \mid w_{\text{H}}(\mathbf{v}_2^{(1)}) > 0 \wedge w_{\text{H}}(\mathbf{v}_2^{(2)}) > 0 \right\}. \quad (8)$$

Removing the restrictions to codewords in (7) and (8), the corresponding number of sequences is given by  $M_1(w)$  and  $M_2(w)$ , respectively. The probability that a fixed sequence among the

Rate $R = 1 - j/k$	$j$	$k$	$\omega_0$	$\delta_{gv}$	$\Delta$
$R = 0.3$	35	50	0.1893	0.1893	$1.4412 \times 10^{-7}$
	70	100	0.1893		$1.4445 \times 10^{-7}$
$R = 0.6$	20	50	0.0794	0.0794	$3.6326 \times 10^{-4}$
	70	175	0.0794		$1.4654 \times 10^{-6}$
$R = 0.9$	5	50	0.0044	0.0130	0.6618
	70	700	0.0130		$8.7523 \times 10^{-5}$

TABLE I  
NUMERICAL RESULTS

set of all possible sequences fulfills the conditions in (7) and (8) is denoted by  $P(\varepsilon_1)$  and  $P(\varepsilon_2)$ , respectively. Clearly,  $P(\varepsilon_1) + P(\varepsilon_2) = 1$ .

Having introduced these notations, the probability  $P(\varepsilon_{cw})$  that a randomly chosen sequence coincides with a codeword of  $\mathcal{C}_{cg}(n, j, k)$  is

$$P(\varepsilon_{cw}) = 2 \frac{N_1(w)}{M_1(w)} P(\varepsilon_1) + \frac{N_2(w)}{M_2(w)} P(\varepsilon_2). \quad (9)$$

With in total  $\binom{n}{w}$  sequences of length  $n$  and weight  $w$ , we have

$$P(\varepsilon_1) = \frac{M_1(w)}{\binom{n}{w}} \quad \text{and} \quad P(\varepsilon_2) = \frac{M_2(w)}{\binom{n}{w}}. \quad (10)$$

Substituting (10) into (9), we finally obtain

$$P(\varepsilon_{cw}) = 2 \frac{N_1(w)}{M_1(w)} \frac{M_1(w)}{\binom{n}{w}} + \frac{N_2(w)}{M_2(w)} \frac{M_2(w)}{\binom{n}{w}} = \frac{2N_1(w) + N_2(w)}{\binom{n}{w}} = \frac{N_{cg}(w)}{\binom{n}{w}} \quad (11)$$

with  $N_{cg}(w)$  denoting the total number of codewords of weight  $w$  in  $\mathcal{C}_{cg}(n, j, k)$ . Thereby we conclude, that for both  $\mathcal{C}(n, j, k)$  and  $\mathcal{C}_{cg}(n, j, k)$ , the same lower bound on the normalized minimum distance  $\omega_0$ , namely (5), holds. Moreover we note that the additional restriction on  $\mathcal{C}_{cg}(n, j, k)$  that each half belongs to  $\mathcal{C}(n/2, j, k/2)$  has no influence on the derivation of the lower bound on the normalized minimum distance  $\omega_0$ . Thus we have the following theorem:

*Theorem 3:* The lower bound on the normalized minimum distance  $w_0$  of  $\mathcal{C}(n, j, k)$  coincides with the lower bound on the normalized minimum distance of  $\mathcal{C}_{cg}(n, j, k)$ .  $\square$

## V. NUMERICAL RESULTS

The lower bound on the normalized minimum distance  $\omega_0$  in (5) is calculated for (Chained) Gallager codes of rates  $R = 0.3$ ,  $R = 0.6$ , and  $R = 0.9$ . As the rate  $R = 1 - \frac{j}{k}$  depends only on  $j$  and  $k$ , we keep one of these parameters constant while we vary the other, obtaining different rates. The numerical results are given in Table I together with the corresponding absolute and relative Gilbert-Varshamov bound  $\delta_{gv}$  and  $\Delta$  defined by

$$\delta_{gv} = h^{-1}(1 - R) \quad \text{and} \quad \Delta = \frac{(\delta_{gv} - \omega_0)}{\delta_{gv}} \quad (12)$$

where  $h^{-1}(x)$  is the inverse binary entropy function. Although the lower bound in (5) is restricted to the class of (Chained) Gallager codes, we obtain almost the Gilbert-Varshamov lower bound. Note that the parameters  $j$  and  $k$  have to be chosen sufficiently large.

## VI. CHANNEL STATE INFORMATION

Now, we restrict the transmission scheme  $\mathcal{T}_2$  further by assuming that during any transmission interval errors can occur only in one of the two parallel, independent channels. Moreover, this information is available as Channel State Information (CSI) at the receiver side, but not necessarily known to the sender.

The parity-check matrix of a Chained Gallager code from  $\mathcal{C}_{\text{cg}}(n, j, k)$  with rate  $R = 1 - \frac{j}{k}$  consists of two parity-check matrices from  $\mathcal{C}(n/2, j, k/2)$  with rate  $R' = 1 - \frac{2j}{k} < R$  (cf. (6)).

Using a Chained Gallager code of rate  $R > 0.5$  and decoding only its underlying Gallager code of the channel being in the error free state, more error patterns can be corrected. The lower bounds on the normalized minimum distance for Chained Gallager codes and Gallager codes,  $f_0(\omega)$  and  $f_1(\omega)$ , obtained from (5) and normalized by the same block length  $n$  are illustrated in Figs. 2(b) and 2(c), respectively. Moreover, if the Chained Gallager code has rate  $R < 0.5$ , each of its two underlying parity-check matrices from  $\mathcal{C}(n/2, j, k/2)$  has full rank, and by using the CSI all possible error patterns can be corrected.

We will summarize these observations in the following two theorems:

*Theorem 4:* Consider transmission scheme  $\mathcal{T}_2$  over two parallel, independent channels, with at least one channel being in the error free state during any transmission interval. Let this information be available as CSI at the receiver side. By using a code from  $\mathcal{C}_{\text{cg}}(n, j, k)$  instead of  $\mathcal{C}(n, j, k)$  the number of correctable error patterns is increased as long as  $\frac{j}{k} < 0.5$ , that is, the rate  $R > 0.5$ . □

*Theorem 5:* Consider transmission scheme  $\mathcal{T}_2$  over two parallel, independent channels, with at least one channel being in the error free state during any transmission interval. Let this information be available at the receiver side as CSI. By using a code from  $\mathcal{C}_{\text{cg}}(n, j, k)$  with rate  $R = 1 - \frac{j}{k}$  all possible error patterns can be corrected as long as  $\frac{j}{k} > 0.5$ , that is, the rate  $R < 0.5$ . □

## VII. CONCLUSIONS

We have introduced two different transmission schemes  $\mathcal{T}_1$  and  $\mathcal{T}_2$ . Using probabilistic arguments, we obtained a tight lower bound on the normalized minimum distance of the ensemble of  $(j, k)$ -regular LDPC codes with block length  $n$ , showing that more errors patterns can be corrected using scheme  $\mathcal{T}_2$ .

Chained Gallager codes, a combination of two Gallager codes of lower rate, have been introduced for which the same lower bound on the normalized minimum distance holds.

Moreover, we compared Gallager codes and Chained Gallager code of same rate  $R$  using scheme  $\mathcal{T}_2$  with at least one channel being in the error free state during any transmission interval. Assuming CSI being available at the receiver side and a rate  $R > 0.5$ , more error patterns can be corrected by using Chained Gallager codes. Furthermore, for rate  $R < 0.5$  Chained Gallager codes can correct any error pattern since each of the underlying Gallager codes has full rank.

## ACKNOWLEDGEMENTS

This research was supported in part by the Swedish Research Council under Grant 621-2007-6281.

## REFERENCES

- [1] R. G. Gallager, *Low Density Parity Check Codes*. Cambridge: MIT Press, 1963.
- [2] K. S. Zigangirov, A. E. Pusane, D. K. Zigangirov, and D. J. Costello Jr., "On the error-correcting capability of LDPC codes," *Problems on Information Transmission*, vol. 44, no. 3, pp. 214–225, Sep. 2008.