Fourteenth International Workshop on Algebraic and Combinatorial Coding Theory
September 7–13, 2014, Svetlogorsk (Kaliningrad region), Russia *pp. 157–162*

On the Hamming-Like Upper Bound on the Minimum Distance of LDPC Codes [1]

Alexey Frolov alexey.frolov@iitp.ru
Inst. for Information Transmission Problems
Russian Academy of Sciences
Moscow, Russia

Abstract. In [1] a Hamming-like upper bound on the minimum distance of regular binary LDPC codes is given. In this paper we extend the bound to the case of irregular and generalized LDPC codes over $\mathbb{F}_q$. The bound is shown to lie under the Varshamov–Gilbert bound at high rates.

1 Introduction

In this paper we investigate the minimum code distance of LDPC codes [2, 3] over $\mathbb{F}_q$. Such codes have good error-correcting capabilities, efficient encoding and decoding algorithms. All of these makes the codes very popular in practical applications.

In [1] a Hamming-like upper bound on the minimum distance of regular binary LDPC codes is given. In this paper we extend the bound to the case of irregular and generalized LDPC codes over $\mathbb{F}_q$.

Our contribution is as follows. First we derive the upper bound for generalized LDPC codes (we assume the Tanner graph [3] to be regular) over $\mathbb{F}_q$. The bound depends on the weight enumerator of the constituent code. Second we derive the upper bound for irregular LDPC codes (we assume the Tanner graph to be irregular) over $\mathbb{F}_q$. The constituent code in this case is a single parity-check (SPC) code over $\mathbb{F}_q$. We note that the bound also improves the result from [1] for the binary case. At last we show the obtained bounds to lie under the Varshamov–Gilbert bound at high rates.

2 Generalized LDPC codes

In this section we obtain the upper bound on the minimum distance of generalized LDPC codes. We use so-called syndrome counting method [1].

[1]This research has been supported by RFBR, research projects No. 13-01-12458 and No. 14-07-31197.

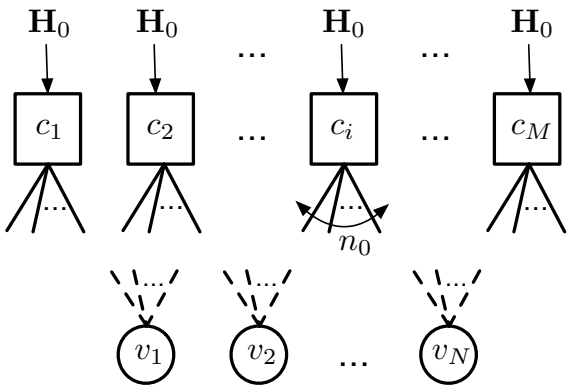

Figure 1: Tanner graph

Let us briefly consider the construction of generalized LDPC code $\mathcal{C}$. To construct such a code we use a bipartite graph, which is called the Tanner graph [3] (see Fig. 1). The graph consists of N variable nodes and M check nodes. In this section we assume all the check nodes to have the same degree n_0 (such Tanner graphs are called right regular). We associate constituent codes to each of the check nodes. In this section all the constituent codes are the same (we denote the constituent code by $\mathcal{C}_0$). We assume $\mathcal{C}_0$ to be an $[n_0, R_0, d_0]$ code over $\mathbb{F}_q$. Let Let us denote the parity-check matrix of the constituent codes by $\mathbf{H}_0$. The matrix has size $m_0 \times n_0$, where $m_0 = (1 - R_0)n_0$.

Let $G(s, n_0, d_0)$ be the weight enumerator of the code $\mathcal{C}_0$, i.e.

$$G(s, n_0, d_0) = 1 + \sum_{i=d_0}^{n_0} A(i)s^i,$$

where $A(i)$ is the number of codewords of weight i in a code $\mathcal{C}_0$.

To check if $\mathbf{v} = (v_1, v_2, \ldots, v_N) \in \mathbb{F}_q^N$ is a codeword of $\mathcal{C}$ we associate the symbols of $\mathbf{v}$ to the variable nodes (see Fig. 1). The word $\mathbf{v}$ is called a codeword of $\mathcal{C}$ if all the constituent codes are satisfied (the symbols which come to the codes via the edges of the Tanner graph form codewords of the constituent codes).

Consider all the possible vectors of length N, weight $W = \omega N$ over $\mathbb{F}_q$. We introduce an equiprobable distribution on such vectors. Let us consider the i-th check, let $\mathbf{S}_i$ denote the syndrome of the i-th constituent code. Note, that $\mathbf{S}_i$ is a random variable and it is easy to see that

$$p_0 = \Pr(\mathbf{S}_i = \mathbf{0}) \quad = \quad \frac{1}{\binom{N}{W}(q-1)^W} \left[\sum_{i=0}^{n_0} \left\{ A(i) \binom{N - n_0}{W - i} (q-1)^{W-i} \right\} \right].$$

In what follows we are interesting in asymptotic estimate when $N \to \infty$. In this case we have

$$\frac{\binom{N-n_0}{W-i}}{\binom{N}{W}} \to \omega^i (1-\omega)^{n_0-i}$$

and

$$\begin{aligned} p_0 &= \left[\sum_{i=0}^{n_0} \left\{A(i)\omega^i(1-\omega)^{n_0-i}(q-1)^{-i}\right\}\right] + o(1) \\ &= (1-\omega)^{n_0} G\left(\frac{\omega}{(1-\omega)(q-1)}, n_0, d_0\right) + o(1). \end{aligned}$$

Let $H(X)$ be the *binary* entropy of the random variable X, let us formulate the following lemma

Lemma 1. *For the random variable $\mathbf{S}_i$ we have*

$$\begin{aligned} H(\mathbf{S}_i) &= -\sum_{j=0}^{q^{m_0}-1} \Pr(\mathbf{S}_i = j) \log_2 \Pr(\mathbf{S}_i = j) \\ &\le -p_0 \log_2 p_0 - (1-p_0)\log_2 \frac{1-p_0}{q^{m_0}-1}. \end{aligned}$$

Proof. The last transition is done in accordance to the log-sum inequality. □

Let us introduce some additional notation. By $\mathbf{S} = (\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_M)$ we denote the resulting syndrome of generalized LDPC code. Let

$$h_Q(x) = -x\log_Q x - (1-x)\log_Q(1-x) + x\log_Q(Q-1).$$

be Q-ary entropy function.

We are ready to prove a theorem

Theorem 1. *Let $\mathcal{C}$ be a generalized LDPC code of length N, rate R, minimum distance δN, with constituent $[n_0, R_0, d_0]$ code $\mathcal{C}_0$ over $\mathbb{F}_q$. Let $G(s, n_0, d_0)$ be the weight enumerator of $\mathcal{C}_0$. Then for sufficiently large N the following inequality holds*

$$R \le 1 - \frac{h_q(\delta/2)}{h_{q^{m_0}}\left[1-(1-\delta/2)^{n_0} G\left(\frac{\delta/2}{(1-\delta/2)(q-1)}\right)\right]} + o(1).$$

Proof. First note, that for $\omega \le \delta/2$

$$\frac{1}{N} H(\mathbf{S}) = h_q(\omega)\log_2 q + o(1),$$

as all the syndromes corresponding to such vectors are different.

After applying such an inequality

$$H(X,Y) \le H(X) + H(Y)$$

we obtain

$$H(\mathbf{S}) \le \sum_{i=1}^{M} H(\mathbf{S}_i) = M h_{q^{m_0}}(1-p_0)\log_2 q^{m_0}. \quad (1)$$

Finally we have

$$R \le 1 - \frac{h_q(\omega)}{h_{q^{m_0}}(1-p_0)} + o(1). \quad (2)$$

After substituting of p_0 and $\omega = \delta/2$ into (2) we obtain the needed result. □

3 Irregular LDPC codes

In this section we derive the upper bound for irregular LDPC codes over $\mathbb{F}_q$. We assume the Tanner graph to be irregular. The constituent code in this case is a single parity-check (SPC) code over $\mathbb{F}_q$.

First we note that an SPC code over $\mathbb{F}_q$ is an MDS code. For the MDS code the number of codewords of weight W can be calculated as follows

$$A(W) = [s^W]G(s, d_0, n_0) = \binom{n_0}{W}(q-1)\sum_{j=0}^{W-d_0}\left\{(-1)^j\binom{W-1}{j}q^{W-d_0-j}\right\}.$$

Thus the enumerator of an SPC code over $\mathbb{F}_q$ is as follows

$$G(s, d_0 = 2, n_0) = \frac{1}{q}\left(1 + (q-1)s\right)^{n_0} + \frac{q-1}{q}(1-s)^{n_0}.$$

To formulate a theorem we need a notion of row degree polynomial

$$\rho(x) = \sum_{i=r_{\min}}^{r_{\max}} \rho_i x^i,$$

where ρ_i is a fraction of rows of the parity check matrix of weight i, $r_{\min}$ and $r_{\max}$ are the minimal and maximal row weights accordingly.

Theorem 2. *Let $\mathcal{C}$ be an LDPC code of length N, rate R, minimum distance δN, with row degree polynomial $\rho(x)$. Then for sufficiently large N the following inequality holds*

$$R \le \overline{R}(q, \rho(x)) = 1 - \frac{h_q(\delta/2)}{h_q\left[\frac{q-1}{q}\left(1 - \rho\left(1 - \frac{q}{q-1}\delta/2\right)\right)\right]} + o(1).$$

Proof. Consider the right part of (1), we have

$$\begin{aligned}\frac{1}{\log_2 q}\sum_{i=1}^{M} H(\mathbf{S}_i) &= (1-R)\sum_{i=r_{\min}}^{r_{\max}} \rho_i h_q\left[1-(1-\omega)^{n_0} G\left(\frac{\omega}{(1-\omega)(q-1)}\right)\right] \\ &= (1-R)\sum_{i=r_{\min}}^{r_{\max}} \rho_i h_q\left[\frac{q-1}{q}-\frac{q-1}{q}\left(1-\frac{q}{q-1}\omega\right)^i\right] \\ &\le (1-R)h_q\left[\frac{q-1}{q}-\frac{q-1}{q}\rho\left(1-\frac{q}{q-1}\omega\right)\right].\end{aligned}$$

These completes the proof. □

Remark 1. *We note that the bound improves the result from* [1] *for the binary case. Recall that in* [1] *in case of irregular LDPC code it is suggested to just substitute* $r_{\max}$ *to the bound for regular code.*

At last we prove that the upper bound is better for regular codes (with the same average row degree as irregular codes).

Proposition 1. *Let* $\ell > 0$ *be an integer, let* $\rho(x)$ *be the row degree distribution of irregular code, such that* $\sum_{i=r_{\min}}^{r_{\max}} i\rho_i = \ell$ *and let* $\rho_{reg} = x^\ell$*, then*

$$\overline{R}(q,\rho(x)) \le \overline{R}(q,\rho_{reg}(x)).$$

Proof. Let $\alpha > 0$. By the concavity of the function α^x we have

$$\rho(\alpha) \ge \alpha^{\sum_{i=r_{\min}}^{r_{\max}} i\rho_i} = \rho_{\mathrm{reg}}(\alpha).$$

These completes the proof. □

4 Numerical results

In Table 1 the results for $q = 8$ are shown. As an example we choose regular $(\ell = 3, n_0)$ LDPC codes. We see that the new bound improves existing upper bounds for linear codes such as the Plotkin bound [4], the Bassalygo–Elias bound [5] and the first McEliece–Rodemich–Rumsey–Welch bound [6]. We also see that at very high rates ($R > 0.994$) the bound lies below the Varshamov–Gilbert bound. We note that the interval of rates in which we observe this behavior is decreasing when q grows. For $q = 2$ the interval is $R > 0.985$, for $q = 16$ the interval is $R > 0.997$.

Table 1: Results for $q = 8$

$(\ell, n_0); R$	(3,10); 0.7	(3,50); 0.94	(3,100); 0.97	(3,200); 0.985	(3,500); 0.994	(3,600); 0.995
VG	0.1260	0.0179	0.0080	0.0036	**0.0013**	**0.0011**
New	0.2282	0.0263	0.0106	0.0043	**0.0013**	**0.0010**
PL	0.2625	0.0525	0.0262	0.0131	0.0052	0.0044
BE	0.2338	0.0355	0.0160	0.0073	0.0026	0.0021
MRRW	0.2494	0.0545	0.0281	0.0144	0.0059	0.0050

References

[1] Y. Ben-Haim and S. Litsyn. Upper bounds on the rate of ldpc codes as a function of minimum distance. *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2092–2100, May 2006.

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

[3] R. Tanner. A recursive approach to low complexity codes. *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 533–547, Sep. 1981.

[4] W. W. Peterson and E. J. Weldon, Jr. *Error-Correcting Codes*. Cambridge: MIT Press, 1972.

[5] L. A. Bassalygo. New Upper Bounds for Error Correcting Codes. *Problems Inf. Transm.*, vol. 1, no. 4, pp. 32–35, 1965.

[6] R. J. McEliece, E. R. Rodemich, H. Rumsey, Jr., and L. R. Welch. New Upper Bounds on the Rate of a Code via the Delsarte–MacWilliams Inequalities. *IEEE Trans. Inf. Theory* vol. 23, no. 2, pp. 157–166, 1977.