On the Lower Bound of the Free Distance of Partial Unit Memory Codes Based on LDPC Codes

Konstantin Kondrashov
Institute for Information Transmission Problems
Russian Academy of Science
Moscow, Russia
Email: k_kondrashov@iitp.ru

Viktor Zyablov
Institute for Information Transmission Problems
Russian Academy of Science
Moscow, Russia
Email: zyablov@iitp.ru

***Abstract*—In this paper we describe ensemble of binary partial unit memory (PUM) codes based on Low–Density Parity–Check (LDPC) block codes. We study the lower bound on the free distance of the proposed codes and show that the increase α of these codes has positive value.**

I. Introduction

Unit Memory (UM) codes were introduced by Lee in 1976 [1]. These are convolutional codes with rate $R = k/n$, memory $m = 1$ and overall constraint length $\nu \leq k$. In the case when $\nu < k$ the latest codes are called *Partial Unit Memory* (PUM) codes. (P)UM codes are constructed based on block codes, e.g. *Reed-Solomon* (RS) [2], [3] or BCH codes [4], [5]. The use of block codes makes an algebraic description of these convolutional codes possible and simplifies their study.

There are two important characteristics of a convolutional code having strong impact on its error correcting capabilities: the *free distance* d_{free} and the $increase$ (slope) of the extended row distance α. The *extended row distance* d_l^r is defined [6] to be the minimum Hamming weight of all paths in the minimal code trellis that diverge from zero state and then return for the first time back to the zero state only after l branches. The free distance is defined as $d_{free} = \min\limits_{l=1,2,\ldots} \{d_l^r\}$. The α gives average linear increase of d_l^r: $\alpha = \lim\limits_{l\to\infty} d_l^r/l$.

In this contribution we consider PUM codes based on LDPC block codes and derive lower bounds for the free distance and the slope.

The paper is organized as follows. In section II we describe ensemble of (P)UM codes based on LDPC block codes. We derive the lower bound on the free distance of the proposed codes and the increase α in section III. In section IV we give numerical results. In section V we give a conclusion.

II. Ensemble of (P)UM Codes Based on LDPC Codes

It is possible to use any linear block code to build linear convolutional (P)UM code. In this contribution we consider using LDPC block codes [7] for this purpose.

Any linear code may be defined by either generator or parity-check matrix and LDPC codes are defined by the last option. Therefore, we define a (P)UM code by its semi-infinite transposed parity-check matrix $\mathbf{H}^{\mathrm{T}}$:

$$\mathbf{H}^{\mathrm{T}} = \begin{pmatrix} \mathbf{H}_0^{\mathrm{T}} & \mathbf{H}_1^{\mathrm{T}} & & \\ & \mathbf{H}_0^{\mathrm{T}} & \mathbf{H}_1^{\mathrm{T}} & \\ & & \ddots & \ddots \end{pmatrix}, \tag{1}$$

where $\mathbf{H}_0$, $\mathbf{H}_1$ are $r \times n$ matrices, $r = n - k$. For either UM or PUM codes, block matrix $\mathbf{H}_0$ must have full rank and $\mathbf{H}_1$ may have less rank if the code is PUM: $\mathrm{rank}\,(\mathbf{H}_0) = r$, $\mathrm{rank}\,(\mathbf{H}_1) = r_1 \leq r$.

We build an ensemble $\mathcal{C}(n, k, k_1)$ of (P)UM codes by choosing randomly and independently LDPC codes from an ensemble of regular Gallager LDPC codes [7]. The parity-check matrix of such a Gallager code consists of a number of so-called $layers$. The parity-check of the first layer $\mathbf{H}^*$ is obtained by combining n_0 identity matrices

$$\mathbf{H}^* = \left(\underbrace{\mathbf{I}_b \ \mathbf{I}_b \ \ldots \ \mathbf{I}_b}_{n_0 \ times} \right), \tag{2}$$

where identity matrix $\mathbf{I}_b$ has size $b \times b$. Having l layers in the LDPC code, its parity-check matrix will be defined as

$$\mathbf{H}_{LDPC} = \left(\pi_1 \left(\mathbf{H}^*\right) \ \pi_2 \left(\mathbf{H}^*\right) \ \ldots \ \pi_l \left(\mathbf{H}^*\right) \right), \tag{3}$$

where π_i is random column permutation. Resulting parity-check matrix dimensions are $r \times n$, where $r = lb$ and $n = bn_0$. By construction, such matrix has l ones and each column and n_0 ones in each row.

To get a (P)UM code from ensemble $\mathcal{C}(n, k, k_1)$, we pick two random LDPC codes with check matrices $\mathbf{H}'$ and $\mathbf{H}''$, $\mathrm{rank}\,(\mathbf{H}') = r$, $\mathrm{rank}\,(\mathbf{H}'') = r_1 \leq r$. These two define a check matrix (1) of a (P)UM code. Now we build generator matrix $\mathbf{G}$ from parity-check matrix $\mathbf{H}$. Having both $\mathbf{H}$ and $\mathbf{G}$ in minimal basic encoding form, their overall constraint lengths ν must coinside [8]. Thus,

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_0 & \mathbf{G}_1 & & \\ & \mathbf{G}_0 & \mathbf{G}_1 & \\ & & \ddots & \ddots \end{pmatrix}, \tag{4}$$

where $\mathbf{G}_0$ and $\mathbf{G}_1$ are $k \times n$ matrices, $\mathrm{rank}(\mathbf{G}_0) = k$ and $\mathrm{rank}(\mathbf{G}_1) = k_1 \leq k$, $k_1 = r_1$. For PUM codes with overall

constraint length $\nu = k_1 < k$, $\mathrm{rank}(\mathbf{G}_1) = k_1 < k$ and sub-matrices $\mathbf{G}_0$, $\mathbf{G}_1$ may be represented as follows:

$$\mathbf{G}_0 = \begin{pmatrix} \mathbf{G}_{00} \\ \mathbf{G}_{01} \end{pmatrix}, \ \mathbf{G}_1 = \begin{pmatrix} \mathbf{0} \\ \mathbf{G}_{11} \end{pmatrix}, \tag{5}$$

where $\mathbf{G}_{00}$ is $k - k_1 \times n$ matrix and $\mathbf{G}_{01}$, $\mathbf{G}_{11}$ are $k_1 \times n$ matrices. Without loss of generality, we will consider $\mathbf{G}_0$ and $\mathbf{G}_1$ in the form (5) further.

III. Free Distance Lower Bound

In this section we derive the lower bound on the free distance d_{free} of PUM codes from an ensemble $\mathcal{C}(n, k, k_1)$, where $k + k_1 < n$. To do that, consider extended row distances d_l^r step by step at different branch lengths $l = 1, 2, \ldots$

Let $\mathbf{u} = [\mathbf{u}_0, \ \mathbf{u}_1, \ \mathbf{u}_2, \ \ldots, \ \mathbf{u}_i, \ \ldots]$ with blocks $\mathbf{u}_j$ of length k denote information sequence. We split each information block $\mathbf{u}_j$ by two parts consisting of $k - k_1$ and k_1 bits correspondingly: $\mathbf{u}_i = (\mathbf{u}_{i,0} \ \mathbf{u}_{i,1})$. For a PUM code described by generator matrix (4) with block sub-matrices (5), corresponding code sequence $\mathbf{v}$ is defined by product

$$\mathbf{v} = \begin{bmatrix} \vdots \\ \mathbf{u}_{i,0}^{\mathrm{T}} \\ \mathbf{u}_{i,1}^{\mathrm{T}} \\ \mathbf{u}_{i+1,0}^{\mathrm{T}} \\ \mathbf{u}_{i+1,1}^{\mathrm{T}} \\ \mathbf{u}_{i+2,0}^{\mathrm{T}} \\ \mathbf{u}_{i+2,1}^{\mathrm{T}} \\ \vdots \end{bmatrix}^{\mathrm{T}} \cdot \begin{pmatrix} \ddots & & & & & \\ \mathbf{G}_{00} & \mathbf{0} & & & & \\ \mathbf{G}_{01} & \mathbf{G}_{11} & & & & \\ & \mathbf{G}_{00} & \mathbf{0} & & & \\ & \mathbf{G}_{01} & \mathbf{G}_{11} & & & \\ & & & \mathbf{G}_{00} & \mathbf{0} & \\ & & & \mathbf{G}_{01} & \mathbf{G}_{11} & \\ & & & & & \ddots \end{pmatrix}. \tag{6}$$

Let us examine all possible outputs at different lengths of information sequence and determine corresponding Hamming weights.

A. d_1^r

Consider information sequence $\mathbf{u} = [\ldots, \ \mathbf{0}, \ \mathbf{u}_i, \ \mathbf{0}, \ \ldots]$, where $\mathbf{u}_i = (\mathbf{u}_{i,0} \ \mathbf{u}_{i,1})$ is the only non-zero block. Let us examine all code sequences generated by information sequence of such kind. From equation (6) we have:

$$\begin{aligned} \mathbf{v} &= [\ldots, \ \mathbf{0}, \ \mathbf{v}_i, \ \mathbf{v}_{i+1}, \ \mathbf{0}, \ \ldots], \\ \mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00} + \mathbf{u}_{i,1}\mathbf{G}_{01}, \\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11}. \end{aligned}$$

For random non-zero $\mathbf{u}_i = (\mathbf{u}_{i,0} \ \mathbf{u}_{i,1})$, there are 3 cases possible.

1) $\mathbf{u}_{i,0} \neq \mathbf{0}$, $\mathbf{u}_{i,1} = \mathbf{0}$.
   Output code blocks:

$$\begin{aligned} \mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00}, \\ \mathbf{v}_{i+1} &= \mathbf{0}. \end{aligned}$$

   Let C_{00} denote code defined by generator matrix $\mathbf{G}_{00}$. Then, $\mathbf{v}_i \in C_{00}$ and summarily Hamming weight $\mathrm{wt}(\mathbf{v}_i \ \mathbf{v}_{i+1}) \geq d(C_{00})$.

2) $\mathbf{u}_{i,0} \neq \mathbf{0}$, $\mathbf{u}_{i,1} \neq \mathbf{0}$.
   Output code blocks:

$$\begin{aligned} \mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00} + \mathbf{u}_{i,1}\mathbf{G}_{01}, \\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11}. \end{aligned}$$

   Let C_0 denote code defined by generator matrix $\mathbf{G}_0$, C_{11} denote code defined by generator matrix $\mathbf{G}_{11}$. Then, $\mathbf{v}_i \in C_0$, $\mathbf{v}_{i+1} \in C_{11}$ and $\mathrm{wt}(\mathbf{v}_i \ \mathbf{v}_{i+1}) \geq d(C_0) + d(C_{11})$.

3) $\mathbf{u}_{i,0} = \mathbf{0}$, $\mathbf{u}_{i,1} \neq \mathbf{0}$.
   Output code blocks:

$$\begin{aligned} \mathbf{v}_i &= \mathbf{u}_{i,1}\mathbf{G}_{01}, \\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11}. \end{aligned}$$

   Let C_{01} denote code defined by generator matrix $\mathbf{G}_{01}$. Then, $\mathbf{v}_i \in C_{01}$, $\mathbf{v}_{i+1} \in C_{11}$ and $\mathrm{wt}(\mathbf{v}_i \ \mathbf{v}_{i+1}) \geq d(C_{01}) + d(C_{11})$.

Extended row distance d_1^r is defined by the output with the minimum Hamming weight among these cases. Hereinafter we assume that among two codes of the same length the one with highest rate has lowest distance. All C_{ij} belong to LDPC codes which satisfy our assumption. Thus, $d(C_0) < d(C_{01})$ and minimum Hamming weight in case 2 is less than in case 3. This yields $d_1^r = \min(d(C_{00}), d(C_0) + d(C_{11}))$. Compare now cases 1 and 2. Minimum distance of code C_{00} is greater than minimum distance of code C_0. However, $d(C_{00})$ may be less than sum $d(C_0) + d(C_{11})$. The later is defined by relation between k and k_1, *e.g.* if $k - k_1 > k_1$, then $d(C_{00}) < d(C_{11})$ and $d_1^r = d(C_{00})$.

B. d_2^r

Consider information sequence $\mathbf{u}$ having 2 subsequent non-zero information blocks: $\mathbf{u} = [\ldots, \ \mathbf{0}, \ \mathbf{u}_i, \ \mathbf{u}_{i+1}, \ \mathbf{0}, \ \ldots]$, where $\mathbf{u}_j = (\mathbf{u}_{j,0} \ \mathbf{u}_{j,1})$, $j = i, i+1$. Let us examine all possible code sequences. Equation (6) yields:

$$\begin{aligned} \mathbf{v} &= [\ldots, \ \mathbf{0}, \ \mathbf{v}_i, \ \mathbf{v}_{i+1}, \ \mathbf{v}_{i+2}, \ \mathbf{0}, \ \ldots], \\ \mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00} + \mathbf{u}_{i,1}\mathbf{G}_{01}, \\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11} + \mathbf{u}_{i+1,0}\mathbf{G}_{00} + \mathbf{u}_{i+1,1}\mathbf{G}_{01}, \\ \mathbf{v}_{i+2} &= \mathbf{u}_{i+1,1}\mathbf{G}_{11}. \end{aligned}$$

Consider essential cases.

1) $\mathbf{u}_{i,0} \neq \mathbf{0}$, $\mathbf{u}_{i,1} = \mathbf{0}$, random $\mathbf{u}_{i+1}$.
   Output code blocks:

$$\begin{aligned} \mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00}, \\ \mathbf{v}_{i+1} &= \mathbf{u}_{i+1,0}\mathbf{G}_{00} + \mathbf{u}_{i+1,1}\mathbf{G}_{01}, \\ \mathbf{v}_{i+2} &= \mathbf{u}_{i+1,1}\mathbf{G}_{11}. \end{aligned}$$

   Code block $\mathbf{v}_i \in C_{00}$ and code blocks $\mathbf{v}_{i+1}$, $\mathbf{v}_{i+2}$ belong to codes covered in d_1^r study. Thus, $\mathrm{wt}(\mathbf{v}_i \ \mathbf{v}_{i+1} \ \mathbf{v}_{i+2}) \geq d(C_{00}) + d_1^r$.

2) $\mathbf{u}_{i,0} \neq \mathbf{0}$, $\mathbf{u}_{i,1} \neq \mathbf{0}$, $\mathbf{u}_{i+1,0} \neq \mathbf{0}$, $\mathbf{u}_{i+1,1} = \mathbf{0}$.
Output code blocks:

$$\begin{aligned}\mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00} + \mathbf{u}_{i,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11} + \mathbf{u}_{i+1,0}\mathbf{G}_{00},\\ \mathbf{v}_{i+2} &= \mathbf{0}.\end{aligned}$$

If matrix

$$\mathbf{G}_{m0} = \begin{pmatrix}\mathbf{G}_{11}\\ \mathbf{G}_{00}\end{pmatrix}$$

defines code C_{m0}, then $\mathbf{v}_{i+1} \neq \mathbf{0}$, since its information vector is non-zero: $\mathbf{u}_{i,1} \neq \mathbf{0}$, $\mathbf{u}_{i+1,0} \neq \mathbf{0}$. That yields $\mathbf{v}_i \in C_0$, $\mathbf{v}_{i+1} \in C_{m0}$ and $\mathrm{wt}\left(\mathbf{v}_i\ \mathbf{v}_{i+1}\ \mathbf{0}\right) \geq d\left(C_0\right) + d\left(C_{m0}\right)$.

3) $\mathbf{u}_{i,0} \neq \mathbf{0}$, $\mathbf{u}_{i,1} \neq \mathbf{0}$, $\mathbf{u}_{i+1,0} = \mathbf{0}$, $\mathbf{u}_{i+1,1} \neq \mathbf{0}$.
Output code blocks:

$$\begin{aligned}\mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00} + \mathbf{u}_{i,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11} + \mathbf{u}_{i+1,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+2} &= \mathbf{u}_{i+1,1}\mathbf{G}_{11}.\end{aligned}$$

If matrix

$$\mathbf{G}_{m1} = \begin{pmatrix}\mathbf{G}_{11}\\ \mathbf{G}_{01}\end{pmatrix}$$

defines code C_{m1}, then $\mathbf{v}_{i+1} \neq \mathbf{0}$. That yields $\mathbf{v}_i \in C_0$, $\mathbf{v}_{i+1} \in C_{m1}$, $\mathbf{v}_{i+2} \in C_{11}$ and $\mathrm{wt}\left(\mathbf{v}_i\ \mathbf{v}_{i+1}\ \mathbf{v}_{i+2}\right) \geq d\left(C_0\right) + d\left(C_{m1}\right) + d\left(C_{11}\right)$.

4) $\mathbf{u}_{i,0} \neq \mathbf{0}$, $\mathbf{u}_{i,1} \neq \mathbf{0}$, $\mathbf{u}_{i+1,0} \neq \mathbf{0}$, $\mathbf{u}_{i+1,1} \neq \mathbf{0}$.
Output code blocks:

$$\begin{aligned}\mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00} + \mathbf{u}_{i,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11} + \mathbf{u}_{i+1,0}\mathbf{G}_{00} + \mathbf{u}_{i+1,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+2} &= \mathbf{u}_{i+1,1}\mathbf{G}_{11}.\end{aligned}$$

If matrix

$$\mathbf{G}_{\alpha} = \begin{pmatrix}\mathbf{G}_{11}\\ \mathbf{G}_{00}\\ \mathbf{G}_{01}\end{pmatrix}$$

defines code C_{α}, then $\mathbf{v}_{i+1} \neq \mathbf{0}$. We have $\mathbf{v}_i \in C_0$, $\mathbf{v}_{i+1} \in C_{\alpha}$, $\mathbf{v}_{i+2} \in C_{11}$ and $\mathrm{wt}\left(\mathbf{v}_i\ \mathbf{v}_{i+1}\ \mathbf{v}_{i+2}\right) \geq d\left(C_0\right) + d\left(C_{\alpha}\right) + d\left(C_{11}\right)$.

5) $\mathbf{u}_{i,0} = \mathbf{0}$, $\mathbf{u}_{i,1} \neq \mathbf{0}$, random $\mathbf{u}_{i+1}$.
Depending on $\mathbf{u}_{i+1}$, this case will be equal to cases from 2 to 4 with only exception that $\mathbf{v}_i \notin C_0$, rather $\mathbf{v}_i \in C_{01}$. $d\left(C_{01}\right) > d\left(C_0\right)$, therefore this case could not give minimum weight compared to cases 2 – 4.

Now we should determine case corresponding to code sequence with minimum Hamming weight.
Compare cases 3 and 4:

$$d\left(C_0\right) + d\left(C_{m1}\right) + d\left(C_{11}\right) > d\left(C_0\right) + d\left(C_{\alpha}\right) + d\left(C_{11}\right),$$

since $d\left(C_{m1}\right) > d\left(C_{\alpha}\right)$: $\dim\left(G_{m1}\right) = 2k_1$, $\dim\left(G_{\alpha}\right) = k + k_1$ and $2k_1 < k + k_1$.
Compare cases 4 and 2:

$$d\left(C_0\right) + d\left(C_{\alpha}\right) + d\left(C_{11}\right) > d\left(C_0\right) + d\left(C_{m0}\right),$$

since $d\left(C_{11}\right) > d\left(C_{m0}\right)$: $\dim\left(G_{11}\right) = k_1$, $\dim\left(G_{m0}\right) = k$ and $k_1 < k$.
Compare cases 1 and 2:

$$d\left(C_{00}\right) + \min\left(d\left(C_{00}\right),\ d\left(C_0\right) + d\left(C_{11}\right)\right) > d\left(C_0\right) + d\left(C_{m0}\right),$$

since $d\left(C_{00}\right) > d\left(C_{m0}\right)$, $d\left(C_{00}\right) > d\left(C_0\right)$ and second summable is greater then $d\left(C_0\right)$ in any cases. Thus, $d_2^r = d\left(C_0\right) + d\left(C_{m0}\right)$.

C. d_3^r

Consider $\mathbf{u} = [\ldots,\ \mathbf{0},\ \mathbf{u}_i,\ \mathbf{u}_{i+1},\ \mathbf{u}_{i+2},\ \mathbf{0},\ \ldots]$, where $\mathbf{u}_j = \left(\mathbf{u}_{j,0}\ \mathbf{u}_{j,1}\right)$, $j = i, i+1, i+2$. Let us examine all possible code sequences. Equation (6) yields:

$$\begin{aligned}\mathbf{v} &= [\ldots,\ \mathbf{0},\ \mathbf{v}_i,\ \mathbf{v}_{i+1},\ \mathbf{v}_{i+2},\ \mathbf{v}_{i+3},\ \mathbf{0},\ \ldots],\\ \mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00} + \mathbf{u}_{i,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11} + \mathbf{u}_{i+1,0}\mathbf{G}_{00} + \mathbf{u}_{i+1,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+2} &= \mathbf{u}_{i+1,1}\mathbf{G}_{11} + \mathbf{u}_{i+2,0}\mathbf{G}_{00} + \mathbf{u}_{i+2,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+3} &= \mathbf{u}_{i+2,1}\mathbf{G}_{11}.\end{aligned}$$

Consider essential cases.

1) $\mathbf{u}_{i,0} \neq \mathbf{0}$, $\mathbf{u}_{i,1} = \mathbf{0}$, random $\mathbf{u}_{i+1}$, random $\mathbf{u}_{i+2}$.
Output code blocks:

$$\begin{aligned}\mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00},\\ \mathbf{v}_{i+1} &= \mathbf{u}_{i+1,0}\mathbf{G}_{00} + \mathbf{u}_{i+1,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+2} &= \mathbf{u}_{i+1,1}\mathbf{G}_{11} + \mathbf{u}_{i+2,0}\mathbf{G}_{00} + \mathbf{u}_{i+2,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+3} &= \mathbf{u}_{i+2,1}\mathbf{G}_{11}.\end{aligned}$$

$\mathbf{v}_i \in C_{00}$ and distribution of non-zero parts in $\mathbf{u}_{i+1}$ and $\mathbf{u}_{i+2}$ gives results for $\mathbf{v}_{i+1}$, $\mathbf{v}_{i+2}$, $\mathbf{v}_{i+3}$ covered in d_2^r analysis. Thus, $\mathrm{wt}\left(\mathbf{v}_i\ \mathbf{v}_{i+1}\ \mathbf{v}_{i+2}\ \mathbf{v}_{i+3}\right) \geq d\left(C_{00}\right) + d_2^r$.

2) $\mathbf{u}_{i,0} \neq \mathbf{0}$, $\mathbf{u}_{i,1} \neq \mathbf{0}$, $\mathbf{u}_{i+1,0} \neq \mathbf{0}$, $\mathbf{u}_{i+1,1} = \mathbf{0}$, random $\mathbf{u}_{i+2}$.
Output code blocks:

$$\begin{aligned}\mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00} + \mathbf{u}_{i,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11} + \mathbf{u}_{i+1,0}\mathbf{G}_{00},\\ \mathbf{v}_{i+2} &= \mathbf{u}_{i+2,0}\mathbf{G}_{00} + \mathbf{u}_{i+2,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+3} &= \mathbf{u}_{i+2,1}\mathbf{G}_{11}.\end{aligned}$$

$\mathbf{v}_i \in C_0$, $\mathbf{v}_{i+1} \in C_{m0}$ and distribution of non-zero parts in $\mathbf{u}_{i+1}$ gives results for $\mathbf{v}_{i+2}$, $\mathbf{v}_{i+3}$ covered in d_1^r analysis. Thus, $\mathrm{wt}\left(\mathbf{v}_i\ \mathbf{v}_{i+1}\ \mathbf{v}_{i+2}\ \mathbf{v}_{i+3}\right) \geq d\left(C_0\right) + d\left(C_{m0}\right) + d_1^r$.

3) $\mathbf{u}_{i,0} \neq \mathbf{0}$, $\mathbf{u}_{i,1} \neq \mathbf{0}$, $\mathbf{u}_{i+1,0} \neq \mathbf{0}$, $\mathbf{u}_{i+1,1} \neq \mathbf{0}$, random $\mathbf{u}_{i+2}$.
Output code blocks:

$$\begin{aligned}\mathbf{v}_i &= \mathbf{u}_{i,0}\mathbf{G}_{00} + \mathbf{u}_{i,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+1} &= \mathbf{u}_{i,1}\mathbf{G}_{11} + \mathbf{u}_{i+1,0}\mathbf{G}_{00} + \mathbf{u}_{i+1,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+2} &= \mathbf{u}_{i+1,1}\mathbf{G}_{11} + \mathbf{u}_{i+2,0}\mathbf{G}_{00} + \mathbf{u}_{i+2,1}\mathbf{G}_{01},\\ \mathbf{v}_{i+3} &= \mathbf{u}_{i+2,1}\mathbf{G}_{11}.\end{aligned}$$

$\mathbf{v}_i \in C_0$, $\mathbf{v}_{i+1} \in C_{\alpha}$ and distribution of non-zero parts in $\mathbf{u}_{i+2}$ gives results for $\mathbf{v}_{i+2}$, $\mathbf{v}_{i+3}$

covered by cases 2 to 4 in d_2^r analysis. Thus, $\mathrm{wt}\left(\mathbf{v}_i\ \mathbf{v}_{i+1}\ \mathbf{v}_{i+2}\ \mathbf{v}_{i+3}\right) \geq d\left(C_0\right) + d\left(C_\alpha\right) + d\left(C_{m0}\right)$. Note, that $d\left(C_0\right) + d\left(C_{m0}\right) = d_2^r$.

We omit cases when $\mathbf{u}_{i,0} = \mathbf{0}$, $\mathbf{u}_{i,1} \neq \mathbf{0}$ or $\mathbf{u}_{i+1,0} = \mathbf{0}$, $\mathbf{u}_{i+1,1} \neq \mathbf{0}$ since they give results similar to cases 2 or 3 with the only difference that one of summable code vectors will have grater distance.
Compare cases 1 and 3:

$$d\left(C_{00}\right) + d_2^r > d\left(C_\alpha\right) + d_2^r.$$

Compare cases 2 and 3:

$$d\left(C_0\right) + d\left(C_{m0}\right) + d_1^r > d\left(C_0\right) + d\left(C_{m0}\right) + d\left(C_\alpha\right).$$

Thus, $d_3^r = d\left(C_0\right) + d\left(C_{m0}\right) + d\left(C_\alpha\right) = d_2^r + d\left(C_\alpha\right)$.

Let us show now that there exists PUM codes such that matrices $\mathbf{G}_{m0}$, $\mathbf{G}_{m1}$ and $\mathbf{G}_\alpha$ define codes C_{m0}, C_{m1} and C_α correspondingly.

Lemma 1: For $k + k_1 < n$ there exist PUM codes from ensemble $\mathcal{C}(n, k, k_1)$, such that matrix $\mathbf{G}_\alpha = \left(\mathbf{G}_{11}^T\ \mathbf{G}_{00}^T\ \mathbf{G}_{01}^T\right)^T$ has rank $k + k_1$.

Proof: Consider product of matrices $\mathbf{G}$ defined by (4), (5) and $\mathbf{H}^{\mathrm{T}}$ defined by (1):

$$\begin{pmatrix} \ddots & & & & \\ \mathbf{G}_{00}\mathbf{H}_0^{\mathrm{T}} & \mathbf{G}_{00}\mathbf{H}_1^{\mathrm{T}} & & & \\ \mathbf{G}_{01}\mathbf{H}_0^{\mathrm{T}} & \mathbf{G}_{01}\mathbf{H}_1^{\mathrm{T}} + \mathbf{G}_{11}\mathbf{H}_0^{\mathrm{T}} & \mathbf{G}_{11}\mathbf{H}_1^{\mathrm{T}} & & \\ & \mathbf{G}_{00}\mathbf{H}_0^{\mathrm{T}} & \mathbf{G}_{00}\mathbf{H}_1^{\mathrm{T}} & & \\ & \mathbf{G}_{01}\mathbf{H}_0^{\mathrm{T}} & \mathbf{G}_{01}\mathbf{H}_1^{\mathrm{T}} + \mathbf{G}_{11}\mathbf{H}_0^{\mathrm{T}} & \mathbf{G}_{11}\mathbf{H}_1^{\mathrm{T}} & \\ & & & & \ddots \end{pmatrix}.$$

Since $\mathbf{G}\mathbf{H}^{\mathrm{T}} = \mathbf{0}$, it implies among others

$$\begin{cases} \mathbf{G}_{00}\mathbf{H}_1^{\mathrm{T}} = \mathbf{0} \\ \mathbf{G}_{11}\mathbf{H}_1^{\mathrm{T}} = \mathbf{0} \end{cases}. \tag{7}$$

Let us show that $\mathbf{G}_{00}$ and $\mathbf{G}_{11}$ may have linear independent rows. Recall that any linear code with generator matrix $\mathbf{G}'$ and check matrix $\mathbf{H}'$ is equivalent to some linear operator and $\dim\left(\mathbf{G}'\right) + \dim\left(\mathbf{H}'\right) = n$. That means that we can find not more than $n - r_1$ independent rows which will give zero when multiplied by $\mathbf{H}_1^{\mathrm{T}}$. But $\mathbf{G}_{00}$ and $\mathbf{G}_{11}$ has k rows summarily and $n - r_1 > k$. Thus, there exist $\mathbf{G}_{00}$ and $\mathbf{G}_{11}$ with linear independent rows, $\mathbf{G}_{00}$ and $\mathbf{G}_{01}$ have independent rows by definition and we can always find $\mathbf{G}_\alpha = \left(\mathbf{G}_{11}^{\mathrm{T}}\ \mathbf{G}_{00}^{\mathrm{T}}\ \mathbf{G}_{01}^{\mathrm{T}}\right)^{\mathrm{T}}$ with rank $k + k_1$, unless $k + k_1 < n$. ■

Now we can determine information sequences that generate code sequences with minimum extended row distance at corresponding lengths.

Lemma 2: Extended row distance d_l^r, $l > 1$ of a PUM code from ensemble $\mathcal{C}\left(n, k, k_1\right)$, $k + k_1 < n$ is defined by the case when information sequence

$$\mathbf{u} = [\ldots,\ \mathbf{0},\ \mathbf{u}_{i,0}\ \mathbf{u}_{i,0},\ \ldots,\ \mathbf{u}_{i+l-1,0}\ \mathbf{u}_{i+l-1,1},\ \mathbf{0},\ \ldots]$$

has all information sub-blocks except the last one being non-zero: $\mathbf{u}_{s,j} \neq \mathbf{0},\ s = i, \ldots\ i + l - 1,\ j = 0, 1,\ \mathbf{u}_{i+l-1,0} \neq \mathbf{0},\ \mathbf{u}_{i+l-1,1} = \mathbf{0}$.

Proof: Let us examine each code block $\mathbf{v}_i$ of code sequence $\mathbf{v}$ generated by $\mathbf{u}$ in details. Code block $\mathbf{v}_i$ may belong to 3 different LDPC codes: $C_0,\ C_{00},\ C_{01}$. First of these codes is defined by generator matrix

$$\mathbf{G_0} = \begin{pmatrix} \mathbf{G}_{00} \\ \mathbf{G}_{01} \end{pmatrix}$$

and two others are obtained by crossing out upper or lower sub-matrix. Code block $\mathbf{v}_{i+1}$ and further may belong to 4 different codes which are defined by generator matrix

$$\mathbf{G}_\alpha = \begin{pmatrix} \mathbf{G}_{11} \\ \mathbf{G}_{00} \\ \mathbf{G}_{01} \end{pmatrix}$$

and its possible crossings out. The last code block $\mathbf{v}_{i+l}$ is exception. It either belongs to code C_{11} or equals zero. Recall now that having two LDPC codes of the same length, code with higher rate will have lower distance. Thus, d_l^r is defined by the case when $\mathbf{v}_i \in C_0$, code blocks inside code sequence $\mathbf{v}_{i+1} \in C_\alpha$, last but one $\mathbf{v}_{i+l-1} \in C_m$ and last $\mathbf{v}_{i+l} = \mathbf{0}$. Such code sequence is generated by input sequence

$$\mathbf{u} = [\ldots,\ \mathbf{0},\ \mathbf{u}_{i,0}\ \mathbf{u}_{i,0},\ \ldots,\ \mathbf{u}_{i+l-1,0}\ \mathbf{0},\ \mathbf{0},\ \ldots],$$

$\mathbf{u}_{s,j} \neq \mathbf{0},\ s = i, \ldots\ i + l - 1,\ j = 0, 1,\ \mathbf{u}_{i+l-1,0} \neq \mathbf{0},\ \mathbf{u}_{i+l-1,1} = \mathbf{0}$. ■

Lemma 2 gives following bounds:

$$\begin{aligned} d_l^r &\geq d\left(C_0\right) + (l-2)d\left(C_\alpha\right) + d\left(C_{m0}\right),\ l \geq 2, \\ \alpha &\geq d\left(C_\alpha\right), \\ d_{free} &= \min\left(d_1^r, d_2^r\right). \end{aligned} \tag{8}$$

Note the inequality sign. It appears in (8) because summing code distances assumes that code words from different code blocks are independent of each other. However this is not true. Two neighboring code blocks are always generated by information words having common parts. This restricts code blocks to be from some sub-codes and thus, their distance may increase. Compare $d_1^r \geq \min\left(d\left(C_{00}\right), d\left(C_0\right) + d\left(C_{11}\right)\right)$ and $d_2^r \geq d\left(C_0\right) + d\left(C_{m0}\right)$. It always holds that $d\left(C_0\right) + d\left(C_{11}\right) > d\left(C_0\right) + d\left(C_{m0}\right)$. However, $d\left(C_{00}\right)$ may be less than $d\left(C_0\right) + d\left(C_{m0}\right)$ depending on relation between k and k_1. Thus, we have following bound for d_{free}:

$$d_{free} \geq \min\left(d\left(C_{00}\right), d\left(C_0\right) + d\left(C_{m0}\right)\right). \tag{9}$$

IV. Numerical Results

It is possible to obtain numerical values for relative distances of LDPC codes [7]. Thus, if we associate LDPC codes defined by generator matrices from our analysis of distances with their parity-check matrices, we could obtain numerical estimations for bounds (8), (9).

Let us consider code sequence $\mathbf{v}$ checked by its transposed semi-infinite parity-check matrix $\mathbf{H}^{\mathrm{T}}$:

$$\begin{bmatrix} \vdots \\ \mathbf{v}_{i-1}^{\mathrm{T}} \\ \mathbf{v}_{i}^{\mathrm{T}} \\ \mathbf{v}_{i+1}^{\mathrm{T}} \\ \vdots \end{bmatrix}^{\mathrm{T}} \cdot \begin{pmatrix} \ddots & & & & \\ \mathbf{H}_0^{\mathrm{T}} & \mathbf{H}_1^{\mathrm{T}} & & & \\ & \mathbf{H}_0^{\mathrm{T}} & \mathbf{H}_1^{\mathrm{T}} & & \\ & & \mathbf{H}_0^{\mathrm{T}} & \mathbf{H}_1^{\mathrm{T}} & \\ & & & \ddots & \end{pmatrix}.$$

Since $\mathbf{v}\mathbf{H}^{\mathrm{T}} = \mathbf{0}$, code blocks must satisfy equation:

$$\mathbf{v}_{i-1}\mathbf{H}_1^{\mathrm{T}} + \mathbf{v}_i\mathbf{H}_0^{\mathrm{T}} = \mathbf{0}. \tag{10}$$

We may determine check matrices corresponding to codes C_{00}, C_{m0} and C_α from this recurrent equation. Recall that code block $\mathbf{v}_i$ corresponding to C_{00} appears only in sequence $\mathbf{v} = [\ldots,\ \mathbf{0}, \mathbf{v}_i,\ \mathbf{0}, \ldots]$. Thus, $\mathbf{v}_i$ must satisfy system

$$\begin{cases} \mathbf{v}_i\mathbf{H}_0^{\mathrm{T}} = \mathbf{0} \\ \mathbf{v}_i\mathbf{H}_1^{\mathrm{T}} = \mathbf{0} \end{cases}. \tag{11}$$

Therefore, we conclude that C_{00} has parity-check matrix $\mathbf{H}_{00}$:

$$\mathbf{H}_{00} = \begin{pmatrix} \mathbf{H}_1 \\ \mathbf{H}_0 \end{pmatrix}.$$

Code block $\mathbf{v}_i$ corresponding to code C_{m0} occurs in sequences $\mathbf{v} = [\ldots, \mathbf{v}_{i-1},\ \mathbf{v}_i,\ \mathbf{0}, \ldots]$, where $\mathbf{v}_{i-1} \neq \mathbf{0}$ and must satisfy

$$\begin{cases} \mathbf{v}_{i-1}\mathbf{H}_1^{\mathrm{T}} + \mathbf{v}_i\mathbf{H}_0^{\mathrm{T}} = \mathbf{0} \\ \mathbf{v}_i\mathbf{H}_1^{\mathrm{T}} = \mathbf{0} \end{cases}. \tag{12}$$

We may not obtain explicit check matrix for C_{m0} from (12), however we may estimate $d(C_{m0})$. Code block $\mathbf{v}_{i+1} \in C_{m0}$ and $d(C_{m0})$ should be not less than distance of LDPC code defined by parity-check matrix $\mathbf{H}_1$ to satisfy equation $\mathbf{v}_{i+1}\mathbf{H}_1^{\mathrm{T}} = \mathbf{0}$.

Code block $\mathbf{v}_i$ corresponding to code C_α occurs in sequences $\mathbf{v} = [\ldots, \mathbf{v}_{i-1},\ \mathbf{v}_i,\ \mathbf{v}_{i+1}, \ldots]$, where $\mathbf{v}_{i-1} \neq \mathbf{0}$ and $\mathbf{v}_{i+1} \neq \mathbf{0}$. Parity-check matrix of C_α may be obtained by solving system of recurrent equations (10). We may not rewrite it in explicit form or give any estimations: this will be the parity-check of some irregular LDPC code for which no existing methods of distance estimation could be applied. We assume only that decreasing k_1 will increase its distance, since $\dim(G_\alpha) = k + k_1$ will decrease.

Numerical results for d_{free} of a PUM code based on (n, l, b) LDPC codes along with Gilbert-Varshamov relative bound δ_{gv} are provided in Table I, where $\delta = d/n$ denotes relative code distance of LDPC code defined by $\mathbf{H}_0$ and $\delta_{free} = d_{free}/n$. For estimation we used parity-check matrices $\mathbf{H}_0$ and $\mathbf{H}_1$ such that $\mathrm{rank}(\mathbf{H}_0)/\mathrm{rank}(\mathbf{H}_1) = r/r_1 = 0.5$. Variables l, n_0 define LDPC and PUM code rate $R = 1 - l/n_0$ and are chosen to maximize δ.

TABLE I
RELATIVE BOUNDS FOR CODE DISTANCES AT DIFFERENT RATES

Rate $R = 1 - l/n_0$	l	n_0	δ_{gv}	δ	δ_{free}
0.65	11	32	0.066	0.062	0.103
0.70	12	40	0.053	0.052	0.094
0.75	12	50	0.042	0.038	0.068
0.80	12	60	0.031	0.026	0.053
0.85	15	100	0.022	0.021	0.035
0.90	10	100	0.013	0.010	0.021
0.95	10	200	0.006	0.004	0.007

V. CONCLUSION

We have considered binary PUM codes based on LDPC block codes and studied their characteristics. These codes may be decoded iteratively with two iteration loops where at inner iteration each block is decoded as LDPC code and at outer iteration they share mutual information. In such scheme their decoding complexity is defined by underlying LDPC codes complexity. Thus, these codes inherit encoding complexity of LDPC codes and their low decoding complexity and outperform them in the sense of distance. By combining LPDC codes lying near GV-bound it is possible to obtain PUM codes lying above. Using PUM codes also offers a kind of trade-off between d_{free} and α – decreasing k_1 will increase α but may decrease d_{free}.

REFERENCES

[1] L. nan Lee, "Short unit-memory byte-oriented binary convolutional codes having maximal free distance (corresp.)," *Information Theory, IEEE Transactions on*, vol. 22, no. 3, pp. 349 – 352, May 1976.

[2] V. V. Zyablov and V. Sidorenko, "On periodic (partial) unit memory codes with maximum free distance," in *Selected papers from the Workshop on Information Protection, Error Control, Cryptology, and Speech Compression*. London, UK: Springer-Verlag, 1994, pp. 74–79.

[3] J. Justesen, "Bounded distance decoding of unit memory codes," *Information Theory, IEEE Transactions on*, vol. 39, no. 5, pp. 1616 –1627, Sep. 1993.

[4] U. Dettmar and U. Sorger, "New optimal partial unit memory codes based on extended bch codes," *Electronics Letters*, vol. 29, no. 23, pp. 2024 – 2025, 1993.

[5] U. Dettmar and S. Shavgulidze, "New optimal partial unit memory codes," *Electronics Letters*, vol. 28, no. 18, pp. 1748 –1749, 1992.

[6] C. Thommesen and J. Justesen, "Bounds on distances and error exponents of unit memory codes," *Information Theory, IEEE Transactions on*, vol. 29, no. 5, pp. 637 – 649, Sep. 1983.

[7] R. Gallager, "Low-density parity-check codes," *Information Theory, IRE Transactions on*, vol. 8, no. 1, pp. 21 –28, 1962.

[8] J. Forney, G.D., R. Johannesson, and Z.-X. Wan, "Minimal and canonical rational generator matrices for convolutional codes," *Information Theory, IEEE Transactions on*, vol. 42, no. 6, pp. 1865 –1880, Nov. 1996.