Sparse Graph Codes for the 2-User Unsourced MAC

Alexander Fengler*, Gianluigi Liva†, Yury Polyanskiy*

2022 56th Asilomar Conference on Signals, Systems, and Computers | 978-1-6654-5906-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IEEECONF56349.2022.10051894

***Abstract*—We study the design of low-density parity check (LDPC) codes for the 2-user binary adder channel (BAC) in the unsourced setting. That is, both users employ the same code. We show that classic design criteria for LDPC codes do not capture the special requirements needed to deal with the interference in an unsourced BAC (UBAC). Instead, we give a graph theoretic analysis that allows to calculate the expected fraction of good pivots for any LDPC ensemble with a given degree distribution pair. A good pivot is a variable node associated with a received "erasure" symbol for which revealing its value allows to recover both transmitted messages up to a vanishing small fraction of residual erasures by simple BP decoding. The analysis of the expected fraction of good pivots reveals a surprising connection between graph theory and density evolution which can be used to compute optimized degree distributions. Results over the 2-user UBAC with additive white Gaussian noise are also provided, and the implications of the presence of noise are discussed. Finally, provide a simple multi-edge type LDPC code construction that can provably achieve the 2-user UBAC limit for linear codes.**

***Index Terms*—Unsourced Multiple Access, Internet of Things, Low-Density-Parity Check Codes**

I. Introduction

The number of compact battery-powered, wireless devices is steadily increasing and their wide deployment enables novel applications in, e.g., smart homes, infrastructure monitoring, factory automation, or healthcare systems. The traffic generated in such applications differs from classical human-generated data traffic in that it is characterized by small duty cycles and small payloads. Establishing coordination between a large number of such devices is time and energy consuming and makes classical scheduling-based multiple-access solutions inefficient. In this context, random access (RA) protocols become an appealing solution [1]. Recently, an information-theoretic perspective on RA was introduced, which relies on the so-called unsourced multiple access (UMAC) framework. UMAC exploits the idea of same codebook communication [2]. This approach allows separating the different messages based purely on the structure of the codebook, i.e., the set of allowed messages. It was shown that good unsourced code designs can approach the capacity of the additive white Gaussian noise (AWGN) adder channel without the need for coordination [2], [3]. While many unsourced code constructions have been proposed [3]–[9], most of them lack analytic understanding. For this reason, we study the design of unsourced codebooks on a simplified channel, the 2-user unsourced binary adder channel (UBAC). In particular, we focus on the design of low-density parity check (LDPC) codes for the 2-user UBAC. In the 2-user UBAC, the channel output Y_i is the sum of two channel inputs $X_{1,i}, X_{2,i}$, i.e.,

$$Y_i = X_{1,i} + X_{2,i} \qquad i = 1, ..., n$$

where n is the blocklength and $X_{1,i}, X_{2,i} \in \{0, 1\}$. Note, that the addition is over the reals. Note that a received symbol $Y_i = 1$ leaves ambiguity about the values of $X_{1,i}$ and $X_{2,i}$. From a decoding perspective, channel outputs equal to 1 can be treated as erasures, and the decoding problem turns into the problem of decoding over a binary erasure channel (BEC) [10]. When a binary linear block code is used, if the code does not possess idle coordinates, the average number of erasures at the channel output is half of the blocklength. One may expect that designing an LDPC code from an ensemble with BP decoding threshold $> 1/2$ would be sufficient to achieve reliable communication. However, the code design problem turns to be more complex, since in the 2-user UBAC the erasure patterns at the channel output match the support set of codewords. This fact renders a density evolution (DE) analysis of the BP decoder hard, hindering the use of BEC threshold-based arguments for the code design. Moreover, the BP decoder requires an initial bias (by guessing one erased bit) to bootstrap iterative erasure decoding. Based on this observation, a low-complexity extension of BP decoding, called pivot decoding, was proposed in [10]. In this type of decoding, a variable node (VN) associated with an erasure (*pivot*) is chosen at random and its value is fixed to 0 or 1. A graph theoretic analysis was given in [10] that allows calculating the expected fraction of "good pivots" for any regular LDPC ensemble. A good pivot is a VN associated with an erasure, for which revealing its value allows resolving the erasure pattern via BP decoding (up to a vanishing-small fraction of the erased bits). The work in [10] has shown that the per-user rates for *regular* LDPC codes with a non-vanishing fraction of good pivots are limited by $R = 1/4$, where R denotes the per-user rate.

In the first part of this work, we extend the analysis of [10] to unstructured irregular LDPC code ensembles. Furthermore, we show that the fraction of erased messages can be tracked by density evolution on the induced graph. We give a criterion that guarantees the existence of good pivots resulting in a tool to optimize the degree distributions. Our results show that the use of irregular distributions allows for rates up to $R = 1/2$, which for linear codes it the best achievable per-user-rate on the 2-user UBAC [10]. In the second part of this work, we investigate how the designed codes perform over a 2-user UBAC with AWGN. We find that the noise actually helps the decoding process. This leads to the counterintuitive conclusion that for

A. Fengler and Y.Polyanskiy are with the Massachusetts Institute of Technology (MIT), MA, USA (email:{fengler,yp}@mit.edu).

G. Liva is with German Aerospace Center (DLR), Germany (email: gianluigi.liva@dlr.de)

AF was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Grant 471512611. AF and YP were also supported by the National Science Foundation under Grant No CCF-2131115.

UBAC-type interference with a low natural noise level it is beneficial to artificially add a controlled amount of noise to the received signal before decoding. In the final part of the paper, we propose a multi-edge type LDPC code construction that provably achieves the $R = 1/2$ limit over the 2-user UBAC.

II. Pivot Decoding

Maximum-likelihood decoding over the 2-user UBAC can be formulated as an erasure decoding problem. Given the observation $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, we may construct a vector $\tilde{\mathbf{y}}$ by setting $\tilde{y}_i = 0$ if $y_i = 0$, $\tilde{y}_i = 1$ if $y_i = 2$, and $\tilde{y}_i =?$ if $y_i = 1$. Here, the symbol "?" represents an erasure. The set of coordinates in which $\tilde{\mathbf{y}}$ is set to "?" is $\mathcal{E} = \{i|\tilde{y}_i =?\}$ and it is referred to as the *erasure set*. Decoding proceeds by creating a list $\mathcal{L}$ of all codewords that are compatible with the modified observation $\tilde{Y}^n$, i.e., $\mathcal{L} = \{\mathbf{c}|c_i = \tilde{y}_i, \forall i \notin \mathcal{E}\}$. The final step consists of searching all unordered codeword pairs $\{\mathbf{v}, \mathbf{w}\} \in \mathcal{L} \times \mathcal{L}$ whose real sum yields $\mathbf{y}$. We refer to a pair $\{\mathbf{v}, \mathbf{w}\}$ satisfying $\mathbf{y} = \mathbf{v}+\mathbf{w}$ as a *valid (codeword) pair*. If the solution is unique, then the final list is $\{\mathbf{v}, \mathbf{w}\}$ and decoding succeeds. In case of multiple valid pairs, a pair is picked at random. We refer to the event where decoding yields multiple valid pairs as a decoding failure.

LDPC codes admit a simpler (yet, suboptimum) decoding algorithm that relies on a simple modification of the BP algorithm. Observe that for any binary linear block code the erasure pattern at the output of a 2-user UBAC matches the support set of a codeword. It follows that the subset of VNs with indexed in $\mathcal{E}$ form a stopping set, and a plain application BP decoding must result in a decoding failure [10]. To solve the issue, one needs to pick a VN with index in $\mathcal{E}$ and fix its value to either zero or one. We refer to the selected VN as the *pivot*. Then the BP decoder is run. We name the overall decoding algorithm as *pivot decoding*. A *good pivot* is a VN with the property that fixing its value will allow the BP decoder the recover all remaining VNs with exception of a set that scales at most sublinearly with n. I.e., good pivots allow for recovery with a vanishing induced bit erasure probability. [1] Note that a residual, small fraction of erasures can be recovered by concatenating an outer high-rate code with the inner LDPC code. In particular, by properly choosing the outer code, it is possible to construct ensemble sequences for which a vanishing bit erasure probability after BP decoding of the inner code results in a vanishing block error probability after outer code decoding, as n grows large [11], with an outer code rate that converges to 1.

III. LDPC Code Design for Pivot Decoding

Let n and m denote the number of VNs and CNs respectively. We analyze irregular LDPC ensembles defined by a pair (L, R) of degree distributions, which can be written in polynomial form in a dummy variable x as

$$L(x) = \sum_{i=0}^{L_{\max}} L_i x^i \qquad \text{and} \qquad R(x) = \sum_{i=0}^{R_{\max}} R_i x^i.$$

Here L_i and R_i denote the fractions of VNs and CNs with degree i.The corresponding edge oriented degree distributions are

$$\lambda(x) = \sum_{i=1}^{L_{\max}} \lambda_i x^{i-1} \qquad \text{and} \qquad \rho(x) = \sum_{i=1}^{R_{\max}} \rho_i x^{i-1}$$

where λ_i and ρ_i are the fractions of edges that connect to degree-i VNs and CNs. They satisfy $\lambda(x) = L'(x)/L'(1)$ and $\rho(x) = R'(x)/R'(1)$. The design rate can be expressed as $\mathrm{R} = 1 - L'(1)/R'(1) = 1 - \int_0^1 \rho(x)\mathrm{d}x / \int_0^1 \lambda(x)\mathrm{d}x$. The average number of edges per VN and CN is given by $L'(1)$ and $R'(1)$. For two given codewords $\mathbf{c}_1, \mathbf{c}_2$ and the corresponding channel output $\mathbf{y} = \mathbf{c}_1 + \mathbf{c}_2$ we use the notation $\mathcal{S}_i = \{j : y_j = i\}$ to partition the channel output into distinct sets. For $j \in \mathcal{S}_0$ $\mathbf{c}_{1,j} = \mathbf{c}_{1,j} = 0$ while for $j \in \mathcal{S}_2$ $\mathbf{c}_{1,j} = \mathbf{c}_{1,j} = 1$. Only the bits in $\mathcal{S}_1 = \mathcal{E}$ are ambiguous. Let the codebook $\mathcal{C}$ be a random LDPC code from an (L, R) ensemble. (See [12] for details on the ensemble.) For the analysis we assume, w.l.o.g., that $\mathbf{c}_1 = 0$. With this assumption $\mathcal{S}_1 = \mathrm{supp}(\mathbf{c}_2)$, and $\mathcal{S}_2 = \emptyset$. In the following we compute the *induced degree distribution* induced by $\mathcal{S}_1$, i.e., the left and right degree distributions of the induced graph obtained after removing all VNs in $\mathcal{S}_0 \cup \mathcal{S}_2$ together with their connecting edges. Since $\mathcal{S}_1$ constitutes the support of a codeword, all connected CNs must have even degree. We have the following result:

Lemma 1. As $n \to \infty$ the degree distribution of the induced graph concentrates around $(L_{\text{ind}}, R_{\text{ind}})$ with $L_{\text{ind}}(x) = L(x)$ and

$$R_{\text{ind}}(x) = R\left(\frac{1+x}{2}\right) + R\left(\frac{1-x}{2}\right) \tag{1}$$

as $n \to \infty$.

We omit the proof due to space constraints and give only a sketch: The VN degrees in the induced graph depend only on the distribution of the channel output and will concentrate around L. Given the VN degree distribution L, we can then enumerate all graph configurations such that the outgoing edges from the VNs in the induced graph are connected to even degree check nodes. This reveals that the number of graphs that give rise to (1) is exponentially larger than any other configuration. We call a graph from the ensemble (L, R_{ind}) a *dominant* induced graph. In [10] a method was developed to compute the fraction of good pivots for regular LDPC codes. The result of [10] can be extended to irregular ensembles. The key to the analysis is to consider the subgraph of the induced graph that contains only degree 2 CNs and the connected VNs. The probability that a VN is connected to a degree 2 CN in the dominant induced graph is exactly $p := \rho_{\text{ind},2}$, where ρ_{ind} is the edge oriented version of the dominant induced degree distribution. Therefore, the number of edges that connect a VN of degree l to a CN of degree 2 is binomially-distributed with parameters (l, p). Averaging over L gives the VN degree

[1] This definition deviates from [10] where a good pivot was defined in the stricter sense that it allowed to recover all codeword bits. We refer to those as good pivots *in the strong sense*

distribution $\tilde{L}$ of the subgraph with only degree 2 CNs:

$$\tilde{L}_i = \sum_l L_l \binom{l}{i} p^i (1-p)^{l-i} \tag{2}$$

We can also view this subgraph as a (unipartite) graph with nodes given by the VNs and edges given by the degree 2 CNs. Such a graph will have a degree distribution $\tilde{L}$. According to the Molloy-Reed criterion [13], a random graph with given degree distribution $\tilde{L}$ has a *giant component*, i.e., a connected component of size $\epsilon_D n$ almost surely as $n \to \infty$ if

$$\sum i(i-2)\tilde{L}_i > 0 \tag{3}$$

Moreover, $\epsilon_D > 0$ can be calculated as $\epsilon_D = 1 - \tilde{L}(u)$ where u is the smallest positive solution of $u = \tilde{\lambda}(u)$ where $\tilde{\lambda}$ is the edge oriented version of $\tilde{L}$ [13], [14].

Theorem 1. As $n \to \infty$ a code from an (L, R) ensemble has at least a fraction ϵ_D of good pivots *in the strong sense* with probability approaching 1 if

$$\epsilon_D + \delta^* > 1$$

where $\delta^* = \inf\{\delta > 0 : \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}[A_{\delta n}] > 0\}$ with $\mathbb{E}[A_{\delta n}]$ being the average number of stopping sets of size δn in the dominant induced graph.

Note that Theorem 1 is in general not tight since it does not account for the reduced CN degrees in the induced graph, after removing the giant component. Nonetheless, we can obtain a stronger achievability result from the general density evolution framework [12]. For a parameter $\epsilon \in [0, 1]$ the density evolution (DE) recursion is defined as

$$x_l = \epsilon\lambda(1 - \rho(1 - x_{l-1})) \tag{4}$$

with $x_{-1} = 1$. x_l describes the average fraction of erased messages emitted from VNs at iteration l. An equivalent recursion that describes the average fraction y_l of erased messages emitted from CNs at iteration l is [12]

$$1 - y_l = \rho(1 - \epsilon\lambda(y_{l-1})) \tag{5}$$

We are interested in the case where all VNs are erased, i.e. $\epsilon = 1$. The condition that the recursion (4) has no fixed points except for $x = 0$ and $x = 1$ if and only if

$$x - \lambda(1 - \rho(1 - x)) < 0 \quad \forall x \in (0, 1). \tag{6}$$

Note, that $x = 0$ and $x = 1$ are always fixed points of (4). We first show that (6) implies $\epsilon_D > 0$:

Theorem 2. If

$$f_\lambda(x) := x - \lambda(1 - \rho_{\text{ind}}(1 - x)) < 0 \quad \forall x \in (0, 1) \tag{7}$$

then $\epsilon_D > 0$.

Proof: A necessary criterion for (7) is that $f'_\lambda(1) < 0$, which can be shown to be equivalent to (3). We have that $f'_\lambda(1) = 1 - \lambda'(1)\rho'_{\text{ind}}(0)$. Furthermore, $\rho'_{\text{ind}}(0) = \rho_{\text{ind},2} = p$ and $\lambda'(1) = L''(1)/L'(1) = \sum_l l(l-1)L_l/(\sum_l lL_l)$. Therefore,

$$\begin{aligned} f'(1) < 0 &\Leftrightarrow \lambda'(1) > \frac{1}{p} \\ &\Leftrightarrow \sum_l l[l - (1 + 1/p)]L_l > 0 \end{aligned} \tag{8}$$

By using (2) we get

$$\begin{aligned} \sum_i i(i-2)\tilde{L}_i &= \sum_l L_l \sum_i i(i-2)\binom{l}{i} p^i (1-p)^{l-i} \\ &= \sum_l L_l \sum_i (i^2 - 2i)\binom{l}{i} p^i (1-p)^{l-i} \\ &\overset{(a)}{=} \sum_l L_l(lp(1-p) + l^2p^2 - 2lp) \\ &= p^2 \sum_l L_l(l^2 - l - l/p) \\ &= p^2 \sum_l L_l[l(l - (1 + 1/p)] \\ &\overset{(b)}{>} 0 \end{aligned}$$

where (a) follows by computing the first and second moments of the binomial distributions with parameters (l, p) and (b) follows by (8). This shows that $f'_\lambda(1) < 0$ is equivalent to the Molloy-Reed criterion in the subgraph of only degree 2 CNs which implies $\epsilon_D > 0$. ■

Theorem 2 shows that there are $\epsilon_D n$ VNs that form a cycle, i.e., fixing one value in the giant component will determine the values of all other VNs in the giant component. In other words, there exist a subset of relative size ϵ_D that, if chosen as pivots, will reveal a fraction ϵ_D of VNs. After this, the fraction of erased messages emitted from VNs x_1 will satisfy $x_1 < 1$. We make the following Ansatz: The density evolution recursion describes the evolution of the fraction of erased messages in the erased graph. This statement does not follow immediately from standard arguments like [12] because the set of $\epsilon_D n$ VNs that is revealed depends on the graph structure. With this Ansatz and condition (7) the recursion (4) with $\epsilon = 1$ will converge to $x_\infty = 0$ with enough iterations. The criterion in Theorem 2 then can be used to optimize left and right degree distributions such that $x_\infty = 0$. For example, for a fixed induced CN degree distribution, good VN degree distributions that satisfy (7) can be found by a linear program. Good degree distribution pairs can be obtained by alternating optimization [12], as presented in the next Section.

IV. Optimization

For fixed ρ_{ind}, (7) is linear in the coefficients of λ. Recall that the design rate is $\mathrm{R} = 1 - \frac{\sum_i \rho_i/i}{\sum_i \lambda_i/i}$. Therefore, the search for λ that maximizes the design rate can be formulated as the linear program [12]

$$\max_\lambda \left\{ \sum_{i\geq 2} \frac{\lambda_i}{i} \middle| \lambda_i \geq 0; \sum_{i\geq 2} \lambda_i = 1; f_\lambda(x) < 0 \ \forall x \in (0, 1) \right\}$$

A similar condition can be formulated by means of (5) as

$$g_\rho(y) := 1 - y - \rho(1 - \lambda(y)) < 0 \quad \forall y \in (0, 1).$$

which leads to the linear program

$$\max_\rho \left\{ -\sum_{i\geq 1} \frac{\rho_i}{i} \middle| \rho_i \geq 0; \sum_{i\geq 1} \rho_i = 1; g_\rho(y) < 0 \ \forall y \in (0, 1) \right\}$$

TABLE I
DEGREE DISTRIBUTIONS WITH RESTRICTED MAXIMAL VN AND CN DEGREE FOUND WITH THE PROCEDURE IN SECTION IV.

$R_{\max} = L_{\max}$	L_2	L_3	L_7	R_4	R_5	R_7	R
4	0.43	0.57	–	1	–	–	0.36
5	0.36	0.64	–	0.44	0.56	–	0.42
7	0.33	0.57	0.1	–	0.73	0.27	0.44

We can find good degree distributions by alternating between these two linear programs. Recall that the per-user rates achievable by linear codes on the 2-user UBAC are limited by $R < 0.5$ [10]. The presented approach allows to construct LDPC code ensembles with design rates that closely approach this limit. Examples are provided in Table I.

V. AWGN PERFORMANCE

On the AWGN adder MAC with binary phase-shift keying (BPSK), the channel input-output relation is

$$Y_i = \sqrt{\mathsf{SNR}}(2X_{1,i}-1)+\sqrt{\mathsf{SNR}}(2X_{2,i}-1)+Z_i, i = 1, ..., n$$

where SNR denotes the channel signal-to-noise ratio (SNR), and where $Z_i \sim \mathcal{N}(0, 1)$. Note that the symmetry of the BAC is broken in the sense that the random Gaussian noise will make one of the codewords being closer to the channel output than the other. The result is that a BP decoder is often able to recover one codeword correctly even without pivoting. The found codeword can be subtracted from the channel output (i.e., interference cancellation is performed) possibly enabling the decoding of the second codeword. This leads to the decoder architecture in Figure 1, where $\mathcal{L}$ denotes the final list (initialized to $\mathcal{L} = \emptyset$). The initial log-likelihood ratios at the BP decoder input $\text{LLR}_i = \log \frac{p(y_i|x_{1,i}=0)}{p(y_i|x_{1,i}=1)}$ are computed by accounting the interference term, resulting in

$$\begin{aligned}\text{LLR}_i = &\log(1 + \exp[-2\sqrt{\mathsf{SNR}}(\sqrt{\mathsf{SNR}} - y_i)]) \\ &- \log(1 + \exp[-2\sqrt{\mathsf{SNR}}(\sqrt{\mathsf{SNR}} + y_i)])\end{aligned}$$

As already mentioned, in principle, BP decoding does not require pivoting. Yet, we found that adding a pivoting stage can significantly improve the performance at the cost of a moderate additional complexity. The pivoting decoder architecture is depicted in Figure 2. Here, before BP decoding, an index i (the pivot) with LLR_i close to zero is chosen and set to either $+\infty$ or $-\infty$. Then two versions of the BP decoder with interference cancellation in Figure 1 are run to produce two output lists. The list with the higher overall likelihood is chosen as final output. In principle, this strategy can be extended to run multiple times with different choices for the pivot. The analysis of the previous section gives an estimate of how many guesses are necessary on average to find a good pivot: Since ϵ_D is the asymptotic expected fraction of good pivots, the probability of not finding a good pivot in K tries is $p_e = (1 - \epsilon_D)^K$.The performance of the decoders described in Figure 1 and and in Figure 2 are compared in Figure 3 in terms of the per-user block error probability over $E_b/N_0 = \mathsf{SNR}/(2R)$. In additional, we plot the performance of a single-user with random interference and noise, i.e., over

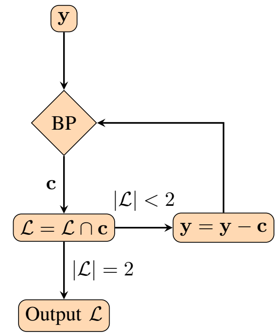

Fig. 1. BP decoder architecture with interference cancellation without pivoting (AWGN adder MAC).

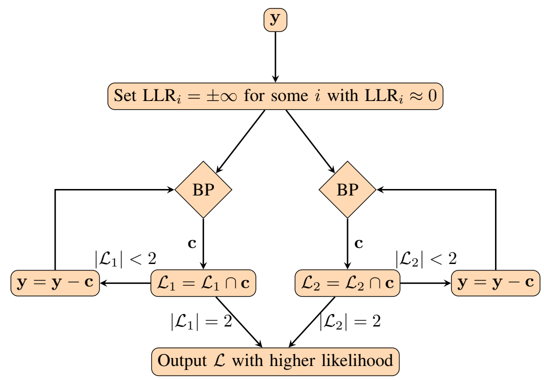

Fig. 2. BP decoder architecture with interference cancellation and with pivoting (AWGN adder MAC).

the channel $Y_i = \sqrt{\mathsf{SNR}}(2X_i - 1) + \sqrt{\mathsf{SNR}} I_i + Z_i$ where $P(I_i = \pm 1) = 1/2$. As expected, for high SNR the error probability converges to the error probability on the BAC. For E_b/N_0 is the range $10 - 15$ dB we can observe a lower error probability than in the high SNR regime, with a minimum achieved at around 12 dB. This can be explained qualitatively by the two contradicting effects of the additive noise. On the one hand, the AWGN helps with breaking the symmetry, while on the other hand it deteriorates the channel output.

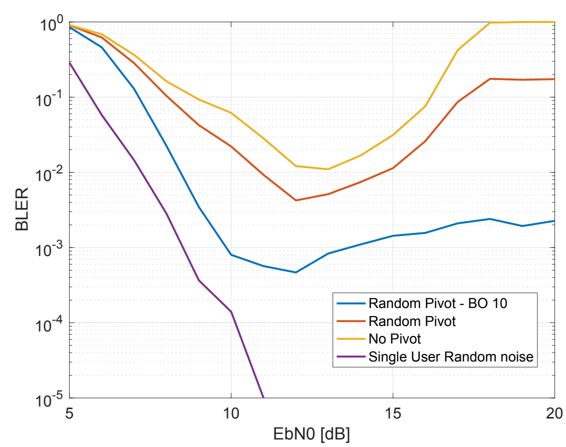

Fig. 3. Block error rate for the 2-user AWGN adder MAC with BPSK modulation. The block length is $n = 800$ and the rate is $R = 0.35$. The code is obtained from an LDPC code ensemble defined by the degree distributions $L(x) = 0.4x^2 + 0.6x^3$ and $R(x) = x^5$.

VI. MULTI-EDGE TYPE LDPC CODE CONSTRUCTION

In this section we describe a multi-edge type LDPC code [12, Chapter 7] construction for the 2-user UBAC which relies on two simple building blocks, i.e., an (n_1, k_1) LDPC code $\mathcal{C}_1$ designed for 2-user UBAC, and an (n_2, k_2) LDPC code $\mathcal{C}_2$ constructed from a capacity-achieving ensemble for the BEC with erasure probability $\epsilon = 0.5$. We denote by $R_1 = k_1/n_1$ the rate of the UBAC code $\mathcal{C}_1$, while $R_2 = k_2/n_2$ is the rate of the code $\mathcal{C}_2$. We add the further constraint $n_2 - k_2 = k_1$. The reason for this will become clear in the following. We construct the parity-check matrix $\mathbf{H}$ of a multi-edge type LDPC code $\mathcal{C}$ with blocklength $n = n_1 + Sn_2$ and dimension $k = k_1 + Sk_2$ from the parity-check matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ of $\mathcal{C}_1$ and $\mathcal{C}_2$ as

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{1,u} & \mathbf{H}_{1,p} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{I} & \mathbf{0} & \mathbf{H}_2 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{H}_2 & \mathbf{0} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (9)$$

where $\mathbf{H}_1 = [\mathbf{H}_{1,u}|\mathbf{H}_{1,p}]$ with $\mathbf{H}_{1,u}$ being an $(n_1 - k_1) \times k_1$ matrix whose columns correspond w.l.o.g. to an information set of $\mathcal{C}_1$, and where the number of row blocks is $S + 1$. It follows that the codewords of $\mathcal{C}$ can be partitioned in a first block of n_1 bits, followed by S blocks of n_2 bits each. By observing the structure of the parity-check matrix in (9), it is clear that $n_2 - k_2$ must equal k_1. The rate of the code $\mathcal{C}$ is

$$R = \frac{k}{n} = \frac{k_1 + Sk_2}{n_1 + Sn_2}$$

and that R approaches R_2 as the parameter S grows. We study the performance of this scheme by resorting to a particular BP decoding schedule. In particular, in a first phase (Phase 1), the code $\mathcal{C}_1$ is used to decode the first n_1 bits of the two codewords transmitted by the users. If decoding succeeds, the first k_1 bits of both vectors are extracted. Let us denote them by $\mathbf{u}_1^{(1)}$ and $\mathbf{u}_1^{(2)}$. The two vectors are then used, in a second phase (Phase 2) to decode the remaining S blocks composing each codeword. To do so, consider the block of n_2 bits following the first block of n_1 bits and denote such block as $\mathbf{c}_2^{(1)}$ and $\mathbf{c}_2^{(2)}$ for users 1 and 2, respectively. Following from (9), we have that $\mathbf{c}_2^{(1)}\mathbf{H}_2^{\mathsf{T}} = \mathbf{u}_1^{(1)}$ and $\mathbf{c}_2^{(2)}\mathbf{H}_2^{\mathsf{T}} = \mathbf{u}_1^{(2)}$, i.e., the vectors $\mathbf{u}_1^{(1)}$ and $\mathbf{u}_1^{(2)}$ define two cosets of the code $\mathcal{C}_2$. More specifically, since $\mathbf{u}_1^{(1)}$ and $\mathbf{u}_1^{(2)}$ can take all possible values in $\mathbb{F}_2^{k_1}$ with uniform probability, the interference generated by a user on the other one over the second block results in erasure patterns that are statistically identical to those generated by a BEC with erasure probability $\epsilon = 0.5$. The same observation holds for all the following $S - 1$ blocks. Decoding of each of the S blocks in Phase 2 proceeds by performing BP decoding over the local graph defined by $\mathbf{H}_2$. Let us analyze the block error probability under this decoding schedule. We define the events $E_1, E_2, \ldots, E_{S+1}$ where E_i is the probability that BP decoding fails for the ith block, while we decode the complementary event as $\bar{E}_i$. The event where we are not able to decode the n-bit codeword of a user is E. We have that

$$\begin{aligned} P(E) &= P\left(\bigcup_{i=1}^{S+1} E_i\right) \\ &= P\left(\bigcup_{i=2}^{S+1} E_i \middle| \bar{E}_1\right) P\left(\bar{E}_1\right) + P\left(E_1\right) \\ &\leq P\left(\bigcup_{i=2}^{S+1} E_i \middle| \bar{E}_1\right) + P\left(E_1\right) \\ &\leq \sum_{i=2}^{S+1} P\left(E_i|\bar{E}_1\right) + P\left(E_1\right) \\ &= SP\left(E_2|\bar{E}_1\right) + P\left(E_1\right). \end{aligned}$$

It follows that $P(E)$ can be made arbitrarily small by choosing an LDPC code $\mathcal{C}_1$ designed for the 2-user BAC with an arbitrary rate (this allows to make $P\left(E_1\right)$ arbitrarily small), and by choosing an LDPC code $\mathcal{C}_2$ with a BEC BP decoding threshold larger than or equal to $1/2$ (this allows to make $P\left(E_2|\bar{E}_1\right)$ arbitrarily small). If $\mathcal{C}_2$ has rate close to $1/2$, i.e., if $\mathcal{C}_2$ is capacity-approaching over a BEC with erasure probability $\epsilon = 0.5$ [15], by choosing S large enough it is possible to approach the linear code limit of the 2-user UBAC.

REFERENCES

[1] N. Abramson, "The ALOHA system - another alternative for computer communications," in *Proc. 1970 Fall Joint Computer Conference*. AFIPS Press, 1970.

[2] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. IEEE Int. Symp. Inf. Theory ISIT*, Jun. 2017, pp. 2523–2527.

[3] A. Fengler, P. Jung, and G. Caire, "SPARCs for Unsourced Random Access," *IEEE Trans. Inf. Theory*, vol. 67, no. 10, pp. 6894–6915, Oct. 2021.

[4] A. K. Pradhan, V. K. Amalladinne, K. R. Narayanan, and J. Chamberland, "Polar Coding and Random Spreading for Unsourced Multiple Access," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[5] E. Marshakov, G. Balitskiy, K. Andreev, and A. Frolov, "A Polar Code Based Unsourced Random Access for the Gaussian MAC," in *Proc IEEE Vehicular Technology Conference*, Sep. 2019, pp. 1–5.

[6] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, "Energy efficient coded random access for the wireless uplink," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4694–4708, Aug. 2020.

[7] V. K. Amalladinne, J.-F. Chamberland, and K. R. Narayanan, "A Coded Compressed Sensing Scheme for Unsourced Multiple Access," *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6509–6533, Oct. 2020.

[8] D. Truhachev, M. Bashir, A. Karami, and E. Nassaji, "Low-Complexity Coding and Spreading for Unsourced Random Access," *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 774–778, Mar. 2021.

[9] A. Fengler, O. Musa, P. Jung, and G. Caire, "Pilot-Based Unsourced Random Access with a Massive MIMO Receiver, Interference Cancellation, and Power Control," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 5, pp. 1522–1534, May 2022.

[10] G. Liva and Y. Polyanskiy, "On Coding Techniques for Unsourced Multiple-Access," in *Proc. 55th Asilomar Conf. Signals Syst. Comput.*, Oct. 2021, pp. 1507–1514.

[11] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, Jun. 2006.

[12] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge ; New York: Cambridge University Press, Mar. 2008.

[13] M. Molloy and B. Reed, "The Size of the Giant Component of a Random Graph with a Given Degree Sequence," *Combinator. Probab. Comp.*, vol. 7, no. 3, pp. 295–305, Sep. 1998.

[14] R. Albert and A.-L. Barabasi, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47–97, Jan. 2002.

[15] P. Oswald and A. Shokrollahi, "Capacity-achieving sequences for the erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 364–373, Dec. 2002.