

CODING FOR THE UNSOURCED A-CHANNEL WITH ERASURES: THE LINKED LOOP CODE

William W. Zheng, Jamison R. Ebert, Stefano Rini, and Jean-Francois Chamberland

ABSTRACT

The A-channel is a noiseless multiple access channel in which users simultaneously transmit Q -ary symbols and the receiver observes the set of transmitted symbols, but not their multiplicities. An A-channel is said to be unsourced if, additionally, users' transmissions are encoded across time using a common codebook and decoding of the transmitted messages is done without regard to the identities of the active users [1]. An interesting variant of the unsourced A-channel is the unsourced A-channel with erasures (UACE), in which transmitted symbols are erased with a given independent and identically distributed probability. In this paper, we focus on designing a code that enables a list of transmitted codewords to be recovered despite the erasures of some of the transmitted symbols. To this end, we propose the linked-loop code (LLC), which uses parity bits to link each symbol to the previous M symbols in a tail-biting manner, i.e., the first symbols of the transmission are linked to the last ones. The decoding process occurs in two phases: the first phase decodes the codewords that do not suffer from any erasures, and the second phase attempts to recover the erased symbols using the available parities. We compare the performance of the LLC over the UACE with other codes in the literature and argue for the effectiveness of the construction. Our motivation for studying the UACE comes from its relevance in machine-type communication and coded compressed sensing.

1. INTRODUCTION

The A-channel with erasures is defined by the input/output relationship

$$Y = \bigcup_{k \in [K]} \delta_{E_k}(x_k), \quad (1)$$

This material is based upon work supported, in part, by the National Science Foundation (NSF) under Grants CCF-2131106 and CNS-2148354, by Qualcomm Technologies, Inc., through their University Relations Program, and by the Ministry of Science and Technology (MOST) under Grant 110-2221-E-A49-052. William Weijia Zheng is with the Department of Information Engineering, the Chinese University of Hong Kong, HKSAR (email: wjzheng@link.cuhk.edu.hk). J. R. Ebert and J.-F. Chamberland are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA (emails: {jrebert, chmbrlnd}@tamu.edu). S. Rini is with the Department of Electrical and Computer Engineering, National Yang Ming Chiao Tung University (NYCU), Hsinchu, 300, TW (email: stefano.rini@nycu.edu.tw).

where the k th users's symbol $x_k \in \mathcal{Q}$, $|\mathcal{Q}| = Q$, the erasure indicator $E_k \stackrel{\text{iid}}{\sim} \mathcal{B}(p_e)$, and the erasure function is defined as

$$\delta_E(x) := \begin{cases} x & E = 0 \\ \emptyset & E = 1. \end{cases} \quad (2)$$

Above, $\mathcal{B}(p)$ indicates the Bernoulli distribution with parameter $p \in [0, 1]$. In the A-channel with erasures of (1), each symbol is received by the receiver with probability $1 - p_e$ and erased with probability p_e . The unsourced A-channel with erasures (UACE) is similar to the model described in (1); however, the receiver is only interested in identifying the transmitted codewords, rather than the traditional source-message pairs. In this article, we introduce the UACE paradigm, and we propose a coding scheme for the challenge which we term linked-loop code (LLC). The LLC uses parity bits to link coded symbols in a tail-biting manner, thus allowing for codewords to be recovered even when some of their symbols have been erased. We evaluate the performance of the LLC and compare it to that of other coding schemes available in the literature.

Literature Review: The A-channel is described in [2], where it models a multiple-access channel that employs frequency division for transmission without the availability of signal intensity. The authors of [2] determine the mutual information of the A-channel under uniform inputs. The limit of the mutual information of the A-channel when K and Q tend to infinity, with the ratio $\lambda = K/Q$ kept fixed, is studied in [3]. Recent interest in the A-channel is due to the emergence of short-packet machine-type communication and IoT devices. In particular, the idea of unsourced random access (URA) is introduced in [4] as an alternative to the customary identify-then-transmit procedure. A common strategy in URA is to use concatenated coding techniques; when this strategy is adopted, in certain settings, the unsourced A-channel can serve as an abstraction for the channel between the outer encoders and the outer decoder. For example, in [5], a linear inner code is used to further encode outer-encoded messages, where the outer code allows for recovery over the unsourced A-channel. During inner decoding, each section is decoded independently, and a list of outer codeword symbols is produced for each section. The parity bits from the outer code designed for the unsourced A-channel are then used to link the outer codeword symbols corresponding to the same user across all the transmission sections. The ensuing unsourced

A-channel code is termed the tree code. The original tree code has been enhanced in [5, 6]. More recently, following the renewed interest in these applications of the A-channel, Lancho et al. [1] offered finite-blocklength achievability bounds for the unsourced A-channel.

Contributions: The major contributions of this article can be summarized as follows.

- **Sec. 2 :** We introduce the UACE model, we define the information transmission problem for this model, and we identify relevant performance metrics.
- **Sec. 3 :** We propose the linked-loop code (LLC) as the first code specifically designed for the UACE. Encoding and decoding procedures are described. The code parameters are also discussed.
- **Sec. 4 :** We review codes in the literature that are compatible with the UACE setting, although not specifically designed for this channel: (i) the tree code of [5], and the (ii) outer LDPC code (triadic design) of [6].
- **Sec. 5 :** Numerical simulations offer insight into the performance of the LLC in Sec. 3, and we provide a performance comparison with other systems in Sec. 4.

Notation: In the following, we adopt the short-hand notation $[N] \triangleq (0, \dots, N-1)$. Calligraphic scripts are used to denote sets (e.g., \mathcal{A}), and $|\cdot|$ is used to denote the cardinality of a set. $GF(2)$ indicates the Galois field of order 2. Given two row vectors \mathbf{v} and \mathbf{w} , we denote their concatenation as $\mathbf{v.w}$. Random variables are indicated with uppercase letters (e.g., N). With a certain abuse of notation, we sometimes interchange the notation for (random) vectors and sets when this is clear from the context.

The code for the numerical evaluations of this paper is provided online at <https://github.com/williamzhen90711/URAC23>.

2. CHANNEL MODEL

The UACE is a variation of the unsourced A-channel from [1] that also encompasses symbol erasures. In the (perfect) unsourced A-channel of [1], the channel output corresponding to the l^{th} channel use is defined as

$$Y(l) = \bigcup_{k \in [K]} x_k(l), \quad (3)$$

where $x_k(l) \in \mathcal{Q}$. When transmission takes place over L channel uses, we indicate the L channel outputs as $\mathbf{Y} = [Y(0), \dots, Y(L-1)]$. The L channel inputs corresponding to user k are labeled as $\mathbf{x}_k = [x_k(0), \dots, x_k(L-1)]$.

The UACE is the variation of the model in (3) defined in (1). Note that $|Y(l)|$ is a random variable which may be less than K due to both collisions and erasures. Here, a collision refers to the event when two distinct users $i \neq j$ share a same symbol $x_i(l) = x_j(l) \in \mathcal{Q}$. Note that the UACE in (1) is parametrized by the number of users K , the cardinality of the input set \mathcal{Q} , and the erasure probability p_e .

2.1. Binary Vector Representation

When describing coding schemes, it is often convenient to operate over a field; this is also the case also for the UACE. For this reason, we introduce a specific notation for the case in which $Q = 2^J$ for some $J \in \mathbb{N}$. Let $x_k(l)$ be denoted by a J -bit binary vector $\mathbf{v}_k(l)$. Then, the channel output may be expressed as

$$\mathbf{y}(l) = \bigcup_{k \in [K]} \delta_{E_{lk}} (\mathbf{v}_k(l)). \quad (4)$$

After L uses of the channel, what we obtain may be expressed as the tuple

$$\mathbf{Y} = [\mathbf{y}(0), \mathbf{y}(1), \dots, \mathbf{y}(L-1)]. \quad (5)$$

Moving forward, it shall be convenient to represent the concatenation of the L input binary vectors, each comprised of J bits, as a single row vector; that is

$$\mathbf{v}_k = \mathbf{v}_k(0) \cdot \mathbf{v}_k(1) \dots \mathbf{v}_k(L-1). \quad (6)$$

In this case, we have that the binary representation of the input of user k is a binary vector \mathbf{v}_k of dimension $N = JL$.

2.2. Coding for the UACE

Let us consider the problem in which each user $k \in [K]$ wishes to transmit a payload of B bits within L channel uses. Assume payload $\mathbf{W} = \{\mathbf{w}_k\}_{k \in [K]}$ is drawn uniformly at random from $[2]^B$. Let us refer to $\mathbf{w} \in \mathbf{W}$ as an *active* payload, whereas $[2]^B \setminus \mathbf{W}$ contains *inactive* payloads. Finally, if two users randomly select the same payload, let us refer to this event as a payload *collision*.

Definition 2.1 ((L, B) Code). A code of block-length L and payload length B for the UACE with parameters (K, Q, p_e) is defined as a (possibly) random *encoding* mapping to produce L channel inputs at each user $k \in [K]$ to encode their message \mathbf{w}_k , defined as

$$\mathbf{x}_k = f_{\text{enc}}(\mathbf{w}_k), \quad (7)$$

where $\mathbf{w}_k \in [2]^B$ and $\mathbf{x}_k \in \mathcal{Q}^L$, together with a (possibly) random *decoding* mapping which produces a list of transmitted payloads from the channel output, that is

$$\widehat{\mathbf{W}} = f_{\text{dec}}(\mathbf{Y}), \quad (8)$$

with $|\widehat{\mathbf{W}}| = \widehat{K}$.

In other words, the encoding function maps the B -bit message \mathbf{w}_k into the L -tuple of channel inputs \mathbf{x}_k . Upon observing the L channel outputs, \mathbf{Y} , the decoding function produces a list of payloads that are likely to be active. The list of possibly active payloads produced by the decoding function has length \widehat{K} . Note that the rate of the code in Def. 2.1 is obtained as $R = B/L$. Also note that in a UACE each encoder $k \in [K]$ utilizes the same encoding function, so that the identity of the transmitter remains immaterial.

Definition 2.2 (Payload Probability of Error). Given code with parameters (N, B) for a UACE with parameters (K, Q, p_e) , define

- the Payload Dropping Probability (PDP) as

$$P_{\text{PDP}} = \frac{1}{K} \sum_{k \in [K]} \mathbb{P} \left[\mathbf{w}_k \notin \widehat{\mathbf{W}}(\mathbf{Y}) | \mathbf{w}_k \in \mathbf{W} \right], \quad (9)$$

- the Payload Hallucination Probability (PHP) as

$$P_{\text{PHP}} = \frac{1}{K} \sum_{k' \in [K]} \mathbb{P} \left[\mathbf{w}_{k'} \notin \mathbf{W} | \mathbf{w}_{k'} \in \widehat{\mathbf{W}}(\mathbf{Y}) \right], \quad (10)$$

where the probability in (9) and (10) is taken over (i) the choice of the payload at each user which is selected uniformly at random over $[2]^B$, (ii) the realization of the erasure sequence, and (iii) the possible randomness in the encoding and decoding functions.

More informally, the PDP indicates the probability that a payload is declared inactive when it was actually active, while the PHP indicates the probability that a payload is declared active when it was indeed inactive. The PDP and PHP are somewhat analogous to the type I and type II errors, respectively, in the hypothesis testing literature.

2.3. Linear codes

Binary linear codes are a special case of the class of codes in Def. 2.1 in which the encoding function of (7) is a matrix multiplication over $GF(2)$. More precisely, a binary linear code is defined as a set of L generator matrices $\{\mathbf{G}_l\}_{l \in [L]}$ which specify how the parity bits for section $l \in [L]$ are generated from the payload \mathbf{w} . Let $\{m(l)\}_{l \in [L]}$ with $\sum_{l \in [L]} m(l) = B$ indicate the number of information bits allocated to section $l \in [L]$, and recall that $\mathbf{m} = [m(0), m(1), \dots, m(L-1)]$. Next, we partition the payload \mathbf{w} into L vectors,

$$\mathbf{w} = \mathbf{w}(0) \cdot \mathbf{w}(1) \dots \mathbf{w}(L-1), \quad (11)$$

with $\mathbf{w}(l) \in [2]^{m(l)}$ for all $l \in [L]$. Let the parity bits in section l be denoted by $\mathbf{p}(l) \in \mathbb{F}_2^{p(l)}$, where length $p(l) = J - m(l)$. The parity bits are obtained from the payload as

$$\mathbf{p}(l) = \mathbf{w} \overline{\mathbf{G}}_l \quad (12a)$$

$$= \sum_{l' \in [L]} \mathbf{w}(l') \mathbf{G}_{l'l}, \quad (12b)$$

where $\overline{\mathbf{G}}_l \in \mathbb{F}_2^{B \times (J - m(l))}$. When considering the partitioning of the payload as in (11), one naturally obtains a partitioning of the $\overline{\mathbf{G}}_l$ in sub-matrices as in (12b) where $\mathbf{G}_{l'l} \in \mathbb{F}_2^{m_{l'} \times (J - m(l))}$. Finally, the channel input is obtained as

$$\mathbf{v}(l) = \mathbf{w}(l) \cdot \mathbf{p}(l). \quad (13)$$

The parameters found in this section are summarized in Tab. 1.

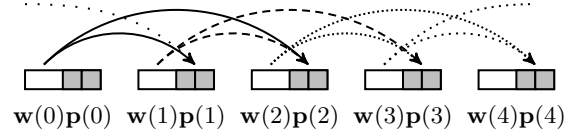


Fig. 1: This notional diagram depicts the formation of a linked-loop codeword.

Table 1: Relevant Symbols and Parameters

Notation	Parameter Description
B	Message length in bits
L	Number of coded sub-blocks per round
J	Length of coded sub-blocks
N	Length of coded B -bit message, $N = LJ$
M	Number of previous sub-blocks used to generate parity bits
$m(l)$	Number of information bits in l -th sub-block
$p(l)$	Number of parity bits in l -th sub-block, $p(l) = J - m(l)$

Remark. Throughout the section, we have omitted the subscript k found in (6). This simplifies notation and it stresses the fact that the encoding is the same at all users $k \in [K]$.

3. LINKED-LOOP CODE (LLC)

We now discuss the task of correctly identifying active payloads before introducing the LLC approach. For simplicity, let us consider the regime in which each user experiences no erasures. Also, we will assume no payload collision occurs, as the probability of this event is orders of magnitude smaller than the target PDP and PHP.

The decoder begins by selecting a single entry of $\mathbf{y}(0)$, and sets this entry as the root of a tree. Then, the decoder checks all entries in $\mathbf{y}(1)$ and identifies those entries that are parity consistent with the root entry. For each such entry, the decoder creates a path from section 0 to section 1. Proceeding, the decoder checks all entries in $\mathbf{y}(2)$ and identifies those entries that are parity consistent with the previously identified partial paths through $\mathbf{y}(0)$ and $\mathbf{y}(1)$. Retaining only parity-consistent paths, the decoder proceeds until all sections have been considered. In the tail-biting scenario, the first sections may need to be re-considered to ensure that they are parity-consistent with the last sections. The hope is that, by the end, a single parity-consistent path remaining, which corresponds to a transmitted codeword. As in [5], we refer to this operation as *stitching* and illustrate the algorithm in Fig. 2.

Next, let us discuss a decoding approach when only one erasure is encountered. When an erasure occurs for a certain user, the decoder is not able to insert the correct element in one of the lists, let's say \mathbf{y}_l . Consequently, the number of surviving paths becomes zero for some section l' with $l \leq l' \leq L$. This means that, after all initial paths have been linked, some of the elements of the root do not have a com-

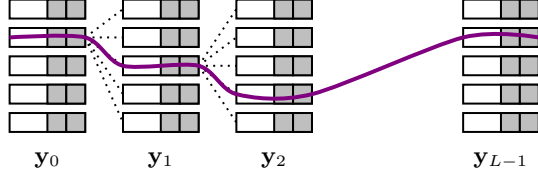


Fig. 2: The diagram illustrates the stitching of elements across sections. When successful, this process returns candidate codewords.

plete path of length L . For these remaining codewords, a *phase 2* of decoding is needed in which all non-terminating paths are considered. For these paths, we introduce a special symbol, $n \setminus a$, that enables each path to have a one-section loss tolerance. After all other sections not in outage have been stitched to some path, we calculate the lost section's information bits by reverse-engineering (12).

3.1. Encoding

The LLC is a binary linear code for the UACE defined by: (i) a memory depth M , and (ii) a set of M parity matrices $\{\mathbf{G}_r\}_{r=1}^M$. The LLC has a fixed number of information and parity bits in each section, i.e., $m(l) = m$ for some $m > 0$. Additionally, the $J - m$ parity bits depend on the information bits of the M previous sections, in a circular manner, that is

$$\mathbf{p}(l) = \sum_{l' \in [L]} \mathbf{w}(l') \mathbf{G}_{l'l} \quad (14a)$$

$$= \sum_{r=1}^M \mathbf{w}(l - r \bmod L) \mathbf{G}_r, \quad (14b)$$

where in (14b) we have set $\mathbf{G}_{l'l} = \mathbf{G}_r$ for $r = l - l' \bmod L$ and $1 \leq r \leq M$.

For example, if $L = 16$ and $M = 2$, the first three lists of parity bits are obtained as:

$$\mathbf{p}(0) = \mathbf{w}(14) \mathbf{G}_2 + \mathbf{w}(15) \mathbf{G}_1 \quad (15a)$$

$$\mathbf{p}(1) = \mathbf{w}(15) \mathbf{G}_2 + \mathbf{w}(0) \mathbf{G}_1 \quad (15b)$$

$$\mathbf{p}(2) = \mathbf{w}(0) \mathbf{G}_2 + \mathbf{w}(1) \mathbf{G}_1. \quad (15c)$$

3.2. Decoding

For the sake of compactness, we present the decoding algorithm through pseudo-code in Alg. 1. Therein, **checkParity**(\mathbf{u}) simply checks whether the parity bits of all sections of the input \mathbf{u} are consistent with (14b); thus, it returns **False** if and only if (14b) is violated with all involved sections that are present as explicit bits in \mathbf{u} . Hence when a path is of length $\leq M$, it will always be parity-consistent because no section's parity can be checked at all. The subroutine **checkParityWithRcv**(\mathbf{u}) serves a similar purpose, except for when \mathbf{u} contains an $n \setminus a$, it tries to uniquely determine the lost information bits according to (14b) by reverse-engineering. This can be done when we choose matrices in $\{\mathbf{G}_r\}_{r=1}^M$ that admit a right-inverse.

Algorithm 1 Pseudo code of the LLC decoding algorithm.

Require: UACE output $\mathbf{Y} = [\mathbf{y}(0), \dots, \mathbf{y}(L-1)]$

Ensure: $\forall l \in [L], |\mathbf{y}(l)| \leq K$;

▷ Decoding phase 1

```

1:  $L^1 \leftarrow \emptyset$ ;  $\widehat{\mathbf{W}}^1 \leftarrow \emptyset$ 
2: for each  $l \in [L]$  do
3:   for each  $\mathbf{u} \in (L^1 \text{ if } L^1 \neq \emptyset \text{ else } \{\emptyset\})$  do
4:      $L^1 \leftarrow L^1 - \{\mathbf{u}\}$ 
5:     for each  $x(l) \in \mathbf{y}(l)$  do
6:       if checkParity( $\mathbf{u}.x(l)$ ) then
7:          $L^1 \leftarrow L^1 \cup \{\mathbf{u}.x(l)\}$ 
8:  $\widehat{\mathbf{W}}^1 \leftarrow \text{extractInfoBits}(L^1)$ 

```

▷ Decoding phase 2

```

9:  $L^2 \leftarrow \emptyset$ ;  $\widehat{\mathbf{W}}^2 \leftarrow \emptyset$ 
10: for each  $x(0)^\times \in \mathbf{y}(0)$  do
11:    $L^2(x(0)^\times) \leftarrow \{x(0)^\times\}$ 
12:   for each  $l \in \{1, 2, \dots, L-1\}$  do
13:     for each  $\mathbf{u} \in L^2(x(0)^\times)$  do
14:        $L^2(x(0)^\times) \leftarrow L^2(x(0)^\times) - \{\mathbf{u}\}$ 
15:       if  $\mathbf{u}$  contains no  $n \setminus a$  then
16:          $L^2(x(0)^\times) \leftarrow L^2(x(0)^\times) \cup \{\mathbf{u}.n \setminus a\}$ 
17:       for each  $x(l) \in \mathbf{y}(l)$  do
18:         if checkParityWithRcv( $\mathbf{u}.x(l)$ ) then
19:            $L^2(x(0)^\times) \leftarrow L^2(x(0)^\times) \cup \{\mathbf{u}.x(l)\}$ 
20:        $L^2 \leftarrow L^2 \cup L^2(x(0)^\times)$ 
21:  $\widehat{\mathbf{W}}^2 \leftarrow \text{extractInfoBits}(L^2)$ 
22:  $\widehat{\mathbf{W}} \leftarrow \widehat{\mathbf{W}}^1 \cup \widehat{\mathbf{W}}^2$ 
23: return  $\widehat{\mathbf{W}}$ 

```

4. RELATED RESULTS

In our numerical simulations, we shall compare the performance of the LLC with the tree code from [5] and the LDPC code from [6]. We stress that both of these codes were not originally designed for the UACE model and, as such this comparison offers only limited insight. These schemes are selected as benchmarks because they are readily applicable to the problem at hand.

Tree code (TC): The tree code of [5] is a binary linear code in which the length of parity bits, $p(l)$, increases by each section. In decoding, it is assumed that Y in (1) does not contain any erasures but possibly contains additional, untransmitted symbols. For this reason, the decoding algorithm is substantially equivalent to phase 1 in the LLC decoding, with the difference that the number of parity bits increases with the section number, thus allowing for a more efficient pruning of surviving paths.

LDPC: Another candidate coding scheme for the UACE is the triadic LDPC code with belief propagation (BP) decoding of [6]. In this approach, sections are viewed as symbols over a Galois field and they are encoded using a non-binary LDPC code; therein, sections contain either all information bits or

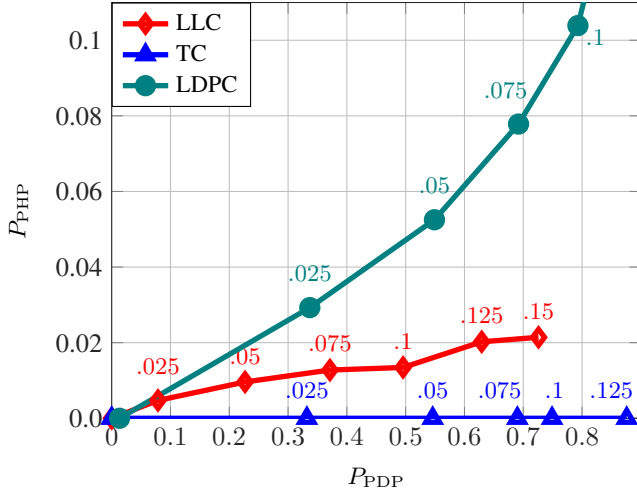


Fig. 3: Empirical P_{PHP} vs P_{PDP} for the LLC (with $M = 2$), tree code, and LDPC code. Here, $K = 100$, and the p_e associated with each point is labelled on the graph.

all parity bits. Parity sections are obtained from information sections as in a classic LDPC code. The FFT method for decoding non-binary LDPC codes is leveraged to efficiently implement message passing between parity sections and information sections.

5. SIMULATION RESULTS

In this section, we study the performance of the LLC on the UACE with hyperparameters tuned to practical applications. We provide a comparison with the tree code of [5] and the LDPC code of [6]. We consider a system with K active users, each having $B = 128$ bits of information to transmit. The corresponding codewords are divided into $L = 16$ sub-blocks, each of length $J = 16$. For the tree code, the specific value for $m(l)$ leads to different discriminating power and was chosen carefully after extensive numerical experimentation. Comparison of the three codes using fixed $K = 100$ and varying $p_e \in [0, 0.15]$ is shown in Fig. 3. Compared to TC and LDPC code, LLC offers a noticeable improvements in terms of greatly decaying the PDP due to phase 2 of decoding, while keeping PHP in a low level. Fig. 4 shows how LLC itself performs under different (K, p_e) with $K \in [50, 150]$.

6. CONCLUSION

This paper presents a novel coding scheme designed for UACE called linked-loop code (LLC). The LLC is managed to identify all payloads suffer no erasure, also to recover payloads suffer one erasure from the UACE. The latter greatly decreases PDP, while keeping PHP to be empirically low comparing with some other methods in the recent literature.

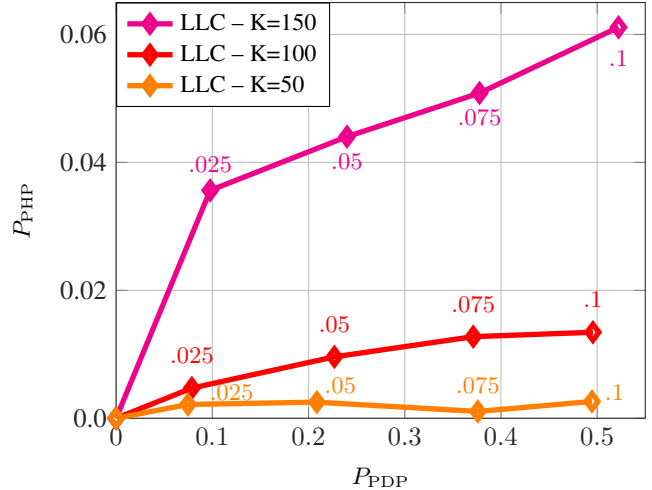


Fig. 4: Empirical P_{PHP} vs P_{PDP} of the LLC (with $M = 2$) for various values of K and p_e . The label associated with each data point indicates the p_e .

7. REFERENCES

- [1] Alejandro Lancho, Alexander Fengler, and Yury Polyanskiy, “Finite-blocklength results for the A-channel: Applications to unsourced random access and group testing,” in *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2022, pp. 1–8.
- [2] Shih-Chun Chang and Jack Wolf, “On the T -user M -frequency noiseless multiple-access channel with and without intensity information,” *IEEE Trans. Info. Theo.*, vol. 27, no. 1, pp. 41–48, 1981.
- [3] Leonid Bassalygo and Mark Pinsker, “Calculation of the asymptotically optimal capacity of a T -user M -frequency noiseless multiple-access channel,” *Numbers, Information and Complexity*, pp. 177–180, 2000.
- [4] Yury Polyanskiy, “A perspective on massive random-access,” in *Proc. Int. Symp. on Info. Theo.* IEEE, 2017, pp. 2523–2527.
- [5] Vamsi K. Amalladinne, Jean-Francois Chamberland, and Krishna R. Narayanan, “A coded compressed sensing scheme for unsourced multiple access,” *IEEE Trans. Info. Theo.*, vol. 66, no. 10, pp. 6509–6533, 2020.
- [6] Vamsi K. Amalladinne, Asit Kumar Pradhan, Cynthia Rush, Jean-Francois Chamberland, and Krishna R. Narayanan, “Unsourced random access with coded compressed sensing: Integrating AMP and belief propagation,” *IEEE Trans. Info. Theo.*, vol. 68, no. 4, pp. 2384–2409, 2021.