Correctable Erasure Patterns in Product Topologies

Lukas Holzbaur∗, Sven Puchinger†, Eitan Yaakobi‡, and Antonia Wachter-Zeh∗

∗Technical University of Munich, {lukas.holzbaur, antonia.wachter-zeh}@tum.de
†Technical University of Denmark, svepu@dtu.dk
‡Technion — Israel Institute of Technology, yaakobi@cs.technion.ac.il

Abstract—Locality enables storage systems to recover failed nodes from small subsets of surviving nodes. The setting where nodes are partitioned into subsets, each allowing for local recovery, is well understood. In this work we consider a generalization introduced by Gopalan et al., where, viewing the codewords as arrays, constraints are imposed on the columns and rows in addition to some global constraints. Specifically, we present a generic method of adding such global parity-checks and derive new results on the set of correctable erasure patterns. Finally, we relate the set of correctable erasure patterns in the considered topology to those correctable in tensor-product codes.

I. INTRODUCTION

Distributed data storage systems consist of an ever growing number of nodes/servers. To protect against data loss, the data is encoded with error/erasure correcting codes, most commonly, maximum distance separable (MDS) codes. While these offer the optimal trade-off between probability of data loss and storage overhead, the recovery of a failed node must involve a large number of nodes. This undesirable overhead lead to the introduction of locally recoverable codes (LRCs) [1], which allow for the recovery of a small number of failed nodes by contacting only a limited number of surviving nodes. In coding theoretic terms, LRCs can recover any small number of erasures from small subsets of surviving positions. For a fixed storage overhead, this improvement comes at the cost of a reduced minimum distance [1], [2], i.e., number of failures that is guaranteed to be correctable. However, codes of optimal minimum distance are not necessarily also optimal in terms of mean time to data loss, as larger numbers of failed nodes might also be correctable. Codes that are able to recover all such combinations of failed nodes are coined maximally recoverable (MR) [3]. Such MR LRCs (also referred to as PMDS codes) have received considerable attention in recent years. Constructions for different parameters regimes are given in [4], [5], [6], [7], [8], [9] and the set of correctable erasure patterns has been characterized [5], [6].

While LRCs represent an important notion of locality, some real world systems require more complex locality constraints. For example, Facebook’s s4 storage system [10] employs MDS codes of length 14 tolerating 4 failed nodes within each data center, plus an additional parity check across 3 data centers. A theoretical framework for such restraints, referred to as the topology of the system, was established in [6]. A topology is a restriction on the support of the parity-check matrix of a code, i.e., it specifies a subset of positions in this matrix that must be zero. One case of particular interest is given by grid-like/product topologies, which, viewed as an array, impose constraints on each row and column, plus an additional number of global constraints. Even in this simplification of the general setting, it is a highly non-trivial problem to determine the erasure patterns that are correctable, i.e., the patterns that an MR code must be able to correct. The seminal works of [6], [8] initiated studies of the classification of these patterns for some restricted cases (see Table I) [11] and bounds on the required field size [12], [13]. In particular, [8] established a necessary condition for an erasure pattern to be correctable, which is then shown to also be sufficient for the case of one column constraint and no global constraints. Further, the sufficiency of this condition is conjectured to hold for the case of more column constraints.

The main results of this work are twofold. First, we establish a generic method of adding global constraints to MR codes, which allows us to establish the set of correctable erasure patterns as a function of the patterns which are correctable without the global constraints. Second, we show that the necessary condition for correctability given in [8] is not sufficient in general. Finally, we connect MR codes for product topologies to tensor-product (TP) codes [14], which are another class of codes with application to storage.

II. PRELIMINARIES

A. Notation

We denote the set of integers {a, b} = {i | a ≤ i ≤ b} and write [1, b] = [b]. For an a × b matrix A and a set I ⊆ [b] we denote by A I the restriction of A to the columns indexed by I. We write 0 a×b to denote the all-zero vector of length b. Let F q denote the finite field of size q. We denote a code C of length n and dimension dim(C) = k over F q by [n, k] q. If the field size q is not of interest we write F and [n, k]. A set I ⊆ [n] with |I| = k is said to be an information set of C if dim(CI) = k.

For sets U ⊆ [n] and V ⊆ [m] we denote the Cartesian product U × V = \{(u, v) | u ∈ U, v ∈ V \}. For two codes C 1 ⊆ F q n and C 2 ⊆ F q m we write C 1 ⊗ C 2 = (G 1 ⊗ G 2) , where G 1, G 2 denote some generator matrix of C 1, C 2, respectively. ⊗ is the Kronecker product, and (G) is the row span of the matrix G. Note that the resulting code is independent of the specific choice of the generator matrices. We have C 1 ⊗ C 2 ⊆ F q n×m but we frequently use the equivalent interpretation as a subset of F q n×m.

B. Codes for Grid-Like Topologies

In this work we are concerned with codes for grid-like topologies Tn×m(a, b, h), i.e., non-rigorously, product codes with h additional parity constraints that are allowed to depend on all codeword positions. We adapt the notation of [8].

Definition 1 (Code for grid-like topology [8, Definition 2.11]. Let Ccol be an [m, m − a] code and Crow be an [n, n − b] code. We define a code for the topology Tn×m(a, b, h) to be any code with parity-check matrix

\[ H = \begin{pmatrix} H_{\text{local}} \\ H_{\text{global}} \end{pmatrix}, \]

where Hlocal is a parity-check matrix of the code Ccol ⊗ Crow and Hglobal is an arbitrary h × mn matrix.
We denote the set of all such codes by $\mathbb{C}_{m \times n}(a, b, h)$.

**Definition 2** (Correctable erasure pattern [8, Definition 2.2]).

Let $E \subseteq [m] \times [n]$ denote a set of erased positions. We say the erasure pattern $E$ is correctable in the topology $T_{m \times n}(a, b, h)$ if and only if there exists a code in $\mathbb{C}_{m \times n}(a, b, h)$ that can correct this erasure pattern.

We denote the set of erasure patterns which are correctable in $T_{m \times n}(a, b, h)$ by $\mathbb{E}_{m \times n}(a, b, h)$ and by $\mathbb{E}_{m \times n}^\text{max}(a, b, h)$ those that are not a proper subset of any correctable pattern.

**Definition 3** (Maximally recoverable code [8, Definition 2.3]).

We say a code $C \in \mathbb{C}_{m \times n}(a, b, h)$ is maximally recoverable (MR) if it corrects every erasure pattern in $\mathbb{E}_{m \times n}(a, b, h)$.

We denote the set of codes that are MR for a topology $T_{m \times n}(a, b, h)$ by $\mathbb{C}_{m \times n}^\text{MR}(a, b, h)$.

Observe that, any code that can correct an erasure pattern $E$ can also correct any erasure pattern $E' \subseteq E$. A maximally recoverable code is equivalently defined as being able to correct any pattern in $\mathbb{E}_{m \times n}^\text{max}(a, b, h)$ instead of $\mathbb{E}_{m \times n}(a, b, h)$.

By the same argument, the set of correctable erasure patterns $\mathbb{E}_{m \times n}(a, b, h)$ is uniquely defined by $\mathbb{E}_{m \times n}^\text{max}(a, b, h)$.

We summarize several important properties of codes for grid-like topologies in terms of our notation. Throughout this work we let $m, n, a, b, h$ be non-negative integers which satisfy $m > a$ and $n > b$.

**Proposition 1** (Properties of codes for grid-like topologies [8, Proposition 2.1]). For any $C \in \mathbb{C}_{m \times n}^\text{MR}(a, b, h)$ it holds that

1. the dimensions of $C, C_{\text{col}},$ and $C_{\text{row}}$ are $\dim(C) = (m - a)(n - b) - h$

2. $\dim(C_{\text{col}}) = m - a$ and $\dim(C_{\text{row}}) = n - b$

3. the codes $C_{\text{col}}$ and $C_{\text{row}}$ are MDS.

**Definition 4** ([8, Definition 3.1]). Consider the topology $T_{m \times n}(a, b, h)$ and an erasure pattern $E \subseteq [m] \times [n]$. We say that $E$ is regular\(^1\) if for all $U \subseteq [m], |U| = a \geq a$, and $V \subseteq [n], |V| = b \geq b$, we have

$$|E \cap (U \times V)| \leq va + ub - ab$$

Remark 1. Intuitively, the restriction of the erasure pattern to $U \times V$ can be interpreted as the shortening of the respective code in the positions outside of this grid. As we will show in Section IV-B, this shortened code needs to be able to decode the remaining erasures for the pattern to be correctable. Definition 4 is a necessary (in general insufficient, see Section IV-A) condition for this to be possible, namely, that the number of erasures remaining in the shortened code does not exceed its redundancy.

\(^1\)Note that the original definition does not include the restriction $a \geq a$ and $v \geq b$. However, it is easy to check that this restriction is indeed necessary.

It was shown in [8] that all erasure patterns that are not regular are not correctable in the topology $T_{m \times n}(a, b, h)$. On the other hand, for some cases (see Table I) it is known that all regular patterns are correctable, which led to the following conjecture.

**Conjecture 1** ([8, Conjecture 3.1]). An erasure pattern $E$ is correctable for the topology $T_{m \times n}(a, b, 0)$ if and only if it is regular.

In Section IV-A we disprove this conjecture.

**III. Global Redundancy $h > 0$**

For some topologies with $h = 0$ the set of correctable erasure patterns is fully characterized (as the regular patterns) and, for some cases, explicit constructions are known. For $h > 0$ only the special cases of $a = 0, b, h \geq 0$ and $a = b = h = 1$ are known (see Table I). In this section, we characterize the set of correctable erasure patterns $\mathbb{E}_{m \times n}^\text{max}(a, b, h)$ as a function of $\mathbb{E}_{m \times n}(a, b, 0)$, for any $m, n, a, b, h$. A similar result for the extension of codes defined by a binary parity-check matrix, i.e., where $H_{\text{local}}$ is in $F_2$, has been derived in [6, Section V.A].

**A. Construction**

We generalize the Gabidulin-based code construction [2] for $a = 0$. These codes have been shown to be MR for the topology $T_{m \times n}(0, b, h)$ in [15]. We show that by the same two-stage encoding procedure we can “add global redundancy” to codes for any grid-like topology. In other words, we give a general construction that, given a code $C_{\text{out}} \in \mathbb{C}_{m \times n}^\text{MR}(a, b, 0)$, returns a code $C \in \mathbb{C}_{m \times n}(a, b, h)$.

**Definition 5** (Gabidulin codes). Let $g = (g_1, \ldots, g_n) \in \mathbb{F}_q^n$ be such that the $g_i$ are linearly independent over $\mathbb{F}_q$. The $[n, k, g]$ Gabidulin code $\text{Gab}(n, k, g)$ is defined as

$$\text{Gab}(n, k, g) = \langle G \rangle$$

with

$$G = \begin{pmatrix} g_1 & g_2 & \cdots & g_n \\ g_1^q & g_2^q & \cdots & g_n^q \\ \vdots & \vdots & \ddots & \vdots \\ g_1^{q^{k-1}} & g_2^{q^{k-1}} & \cdots & g_n^{q^{k-1}} \end{pmatrix}$$

Observe that Gabidulin codes exist for any $n \leq s$. Gabidulin codes are designed for the rank metric, but it is also well-known that they are MDS in the Hamming metric.

**Lemma 1** ([16, Lemma 3]). Let $G \in \mathbb{F}_q^{k \times n}$ be a generator matrix of an $[n, k, d]$ Gabidulin code $\text{Gab}(n, k, g)$. Then, for any full-rank matrix $A \in \mathbb{F}_q^{n \times n}$, the code $C' = \langle G \cdot A \rangle$ is an $[n, k, d]$ Gabidulin code $\text{Gab}(n, k, g')$ with $g' = g \cdot A$.

**Lemma 2.** Let $\mathcal{I}$ be the set of information sets of $C_{\text{out}}[k_{\text{out}}, k_{\text{out}}, g]$ and $C_{\text{in}}$ be a $\text{Gab}(k_{\text{out}}, k_{\text{in}}, g')$. Gabidulin code. Then the code $\langle G_{\text{in}} \cdot C_{\text{out}} \rangle_{\mathcal{I}}$ is a $\text{Gab}(k_{\text{out}}, k_{\text{in}}, g')$. Gabidulin code for any $\mathcal{I} \in \mathcal{L}_{\text{out}}$.

---

**TABLE I**

Properties of different topologies. Lines 1-4 collect the known results. Lines 5/6 summarize the results on global redundancy of Section III, where the $\sim$ symbol implies that a code for the respective topology can be constructed given a code for the same topology with $h = 0$. Lines 7/8 give the results on the characterization of correctable erasure patterns presented in Section IV.
Proof: By definition of an information set, the matrix $G_{\text{out}}|z$ is of full rank. The statement follows from observing that $(G_{\text{in}} \cdot G_{\text{out}})|z = (G_{\text{in}} \cdot (G_{\text{out}}|z))$ and applying Lemma 1.

Theorem 1. Let $C_{\text{out}} \in \mathbb{C}^{MN}_{m \times n}(a,b,0)$ be an $[n_{\text{out}} = mn, k_{\text{out}} = (m-n)(n-b)]$ code and

$$C_{\text{in}} = \text{Gab}(m-n)(n-b), (m-a)(n-b) - h, g)$$

with $s \geq (m-a)(n-b)$. Then the code $(G_{\text{in}}, \cdot G_{\text{out}})$ corrects all erasure patterns in

$$\{E' \cup I \mid E' \in \mathbb{E}^{\text{max}}_{m \times n}(a,b,0), I \subseteq [(m \times [n] \setminus E'), |I| = h\}.$$

Proof: Let $E$ be an erasure pattern of the set above. Then, there is an erasure pattern $E'$ of the set $\mathbb{E}^{\text{max}}_{m \times n}(a,b,0)$ with $E = E' \cup I$. The complement of $E'$ is, by definition, an information set of the outer code $C_{\text{out}}$. By restricting the overall code to this information set, we thus obtain a Gabidulin code with parameters $[(m-a)(n-b), (m-a)(n-b) - h]$ by Lemma 2. Note that there are exactly $h$ remaining erasures (given by $I$) in the remaining positions. Since any Gabidulin code is MDS, the restricted code can correct exactly $h$ erasures, which concludes the proof.

Example 1. Let $C_{\text{out}}[mn, m(n-b)]$ be the code spanned by

$$\text{diag}(G, G, \ldots, G),$$

where $G$ is the generator matrix of an arbitrary $[n, n-b]_q$ MDS code. Observe that $C_{\text{out}} \in \mathbb{C}^{MN}_{m \times n}(0, b, 0)$ and the set of its information sets is given by

$$I_{\text{out}} = \{\text{From each block pick arbitrary } n-b \text{ positions}\}.$$

Choose $C_{\text{in}}$ to be an $[m(n-b), m(n-b) - h_{\text{in}}(m-n)]$ Gabidulin code to obtain the PMDS code construction of [2]. Observe that $I_{\text{out}}$ are exactly the subsets of positions that must give an MDS code in a PMDS code.

B. Correctable Erasure Patterns

Let us characterize the patterns correctable in $T_{m \times n}(a,b,h)$ given the set of patterns correctable in $T_{m \times n}(a,b,0)$

Theorem 2. We have

$$\mathbb{E}^{\text{max}}_{m \times n}(a,b,h) = \{E' \cup I \mid E' \in \mathbb{E}^{\text{max}}_{m \times n}(a,b,0),$$

$$I \subseteq [(m \times [n] \setminus E'), |I| = h\},$$

i.e., any $E \in \mathbb{E}^{\text{max}}_{m \times n}(a,b,h)$ can be obtained by adding $h$ erasures to some $E' \in \mathbb{E}^{\text{max}}_{m \times n}(a,b,0)$.

Proof: "\(\Rightarrow\)" follows by the construction of Theorem 1.

The other direction is implied by the following argument. Let $E \in \mathbb{E}^{\text{max}}_{m \times n}(a,b,h)$. Denote by $G$ the generator matrix of a code $C \subseteq \mathbb{C}_{m \times n}(a,b,h)$ that corrects $E$ and let $H_{\text{local}}, H_{\text{global}}$ be as in Definition 1. Denote by $C' \subseteq \mathbb{C}_{m \times n}(a,b,0)$ the code obtained by setting $H_{\text{global}} = 0$. Then there exists a generator matrix of $C'$ of the form

$$G' = \begin{pmatrix} G \\ G_{\text{global}} \end{pmatrix},$$

for some $G_{\text{global}} \in \mathbb{F}^{n \times mn}$. Trivially, we have

$$\text{rank}(G'_|[mn]|E) \geq \text{rank}(G|[mn]|E) = \dim(C),$$

where the last equality holds because $E$ is correctable in $C$ by definition. By basic linear algebra arguments there exists a subset $I \subseteq E$ with $|I| = \dim(C) - \dim(C') \leq h$ such that

$$\text{rank}(G'|([mn]|E)|I) = \text{rank}(G') = \dim(C').$$

It follows that $E \setminus I$ is correctable in $T_{m \times n}(a,b,0)$. This concludes the proof.

Corollary 1. Denote $n = mn$. Then, for any $E' \in \mathbb{E}^{\text{max}}(a,b,0)$ and any $[N,k]$ code $C \subseteq \mathbb{C}^{MN}_{m \times n}(a,b,h)$ the code $C|_{[mn]|E'}$ must be an $[N - |E'|, k, h + 1]$ MDS code.

Proof: It follows trivially from the dimension of $C$ that the code $C|_{[mn]|E'}$ can never correct more than $h$ erasures. The existence of a code for which this restriction is an MDS code follows from Corollary 2.

Corollary 2. Let $C_{\text{out}}[N,k] \in \mathbb{C}^{MN}_{m \times n}(a,b,0)$. Then there exists an $[N, k - h|I|, a]$ code $C \subseteq \mathbb{C}^{MN}_{m \times n}(a,b,h)$.

IV. NEGATIVE RESULTS ON CORRECTABLE ERASURE PATTERNS

In this section we turn to providing negative results on the set of correctable erasure patterns for given topology. First, we prove that a specific erasure pattern is never correctable in the topology $T_{5 \times 5}(2,2,0)$, thereby disproving Conjecture 1 as given in [8]. Then, using this result, we provide generic methods of constructing uncorrectable, regular erasure patterns for larger topologies.

A. Disproving a Conjecture on the Correctability of Regular Erasure Patterns

We begin with some general observations on the relation between low-weight codewords.

Lemma 3. Consider a linear $[5,3]$ code $C_{\text{row}}$ with generator matrix

$$G_{\text{row}} = \begin{pmatrix} 1 & 0 & 0 & p_{1,1}^{\text{row}} & p_{1,2}^{\text{row}} \\ 0 & 1 & 0 & p_{2,1}^{\text{row}} & p_{2,2}^{\text{row}} \\ 0 & 0 & 1 & p_{3,1}^{\text{row}} & p_{3,2}^{\text{row}} \\ \end{pmatrix}$$

and a codeword

$$(1 \ 0 \ 0 \ 0 \ 0) \in C_{\text{row}}. \ (1)$$

Then

$$p_{1,1}^{\text{row}} = - (a_2 p_{2,1}^{\text{row}} + a_3 p_{3,1}^{\text{row}}) \ (2)$$

and

$$p_{1,2}^{\text{row}} = - (a_2 p_{2,2}^{\text{row}} + a_3 p_{3,2}^{\text{row}}). \ (3)$$

Proof: The statement follows from

$$(1 \ a_2 \ a_3 \ 0 \ 0) = (1 \ a_2 \ a_3) \cdot \begin{pmatrix} 1 & 0 & 0 & p_{1,1}^{\text{row}} & p_{1,2}^{\text{row}} \\ 0 & 1 & 0 & p_{2,1}^{\text{row}} & p_{2,2}^{\text{row}} \\ 0 & 0 & 1 & p_{3,1}^{\text{row}} & p_{3,2}^{\text{row}} \end{pmatrix} = (0 \ 0). \ (4)$$

Observe that each row of $G_{\text{row}}$ is a codeword of $C_{\text{row}}$, in particular,

$$(1 \ 0 \ 0 \ p_{1,1}^{\text{row}} \ p_{1,2}^{\text{row}}) \in C_{\text{row}}. \ (5)$$

We define the same notions in similar notation for a $[5,3]$ code $C_{\text{col}}$, i.e.,

$$\begin{pmatrix} 1 & \gamma_2 & \gamma_3 & 0 & 0 \end{pmatrix} \in C_{\text{col}} \ (5)$$

and

$$\begin{pmatrix} 1 & 0 \ p_{1,1}^{\text{col}} \ p_{1,2}^{\text{col}} \end{pmatrix} \in C_{\text{col}} \ (6)$$

and

$$\begin{pmatrix} 1 & \gamma_2 \ p_{2,1}^{\text{col}} & \gamma_3 \ p_{3,1}^{\text{col}} \end{pmatrix} \in C_{\text{col}} \ (7)$$

and

$$\begin{pmatrix} 1 & \gamma_2 \ p_{2,2}^{\text{col}} & \gamma_3 \ p_{3,2}^{\text{col}} \end{pmatrix} \in C_{\text{col}} \ (8)$$

Note that if $C_{\text{row}}$ (or $C_{\text{col}}$) is MDS we have $p_{1,1}^{\text{row}} \neq 0, p_{1,2}^{\text{row}} \neq 0, p_{1,1}^{\text{col}} \neq 0$ and the codewords given in (1) and (4) ((5) and (6)) are unique. This follows from the well-known facts that all symbols in the parity part of a systematic generator matrix of an MDS code must be non-zero and that every codeword of weight $d$ of a linear code is unique up to scalar multiplication with an element of $\mathbb{F}^*$. With these general relations established, we are now ready to prove that there exists a regular erasure pattern that is never
correctable in the topology $T_{5 \times 5}(2, 2, 0)$, by constructing a non-zero codeword that is zero in all non-erased positions.

**Lemma 4.** The regular pattern $E$ given by

$$E = \begin{pmatrix}
0 & \cdots & 0 \\
\star & \cdots & \star \\
0 & \cdots & 0 \\
\cdots & \cdots & \cdots \\
0 & \cdots & \star
\end{pmatrix}$$

is not correctable in $T_{5 \times 5}(2, 2, 0)$.

**Proof:** Let $C \in C_{m \times n}(a, b, 0)$. We prove the statement by constructing a non-zero codeword of $C$ that is zero in all non-erased positions. This implies that the code $C$ cannot correct the pattern uniquely since there are at least two codewords that coincide on all non-erased positions—the constructed non-zero codeword and the all-zero codeword.

To construct such a codeword, we replace the $\star$-symbols in $E$ by elements of $F$ (not all zero) in the following manner:

(a) Choose the element in position $(2, 1)$ to be $\gamma_2 \in F^*$. Note that, since the code is linear, we can always normalize one single non-zero position to be an arbitrary element of $F^*$ and w.l.o.g. we assume position $(2, 1)$ to be non-zero.

(b) Choose the second row to be the $\gamma_2$-multiple of $(1)$.

(c) Choose the second and third column to be the corresponding multiples of $(5)$.

(d) To obtain a multiple of $(1)$ as the third row, set its first position to be $\gamma_3$.

(e) Encode the first row and column with $G_{\text{row}}$ and $G_{\text{col}}$, respectively.

(f) Replace the entries according to $(2),(3),(7)$, and $(8)$.

(g) Fill in the fourth and fifth rows with the corresponding multiples of $(4)$.

The individual steps are given by

$$\Rightarrow \begin{pmatrix}
0 & \cdots & 0 \\
\cdots & \cdots & \cdots \\
0 & \cdots & 0 \\
\cdots & \cdots & \cdots \\
0 & \cdots & \star
\end{pmatrix} \Rightarrow \begin{pmatrix}
0 & \cdots & 0 \\
\cdots & \cdots & \cdots \\
0 & \cdots & 0 \\
\cdots & \cdots & \cdots \\
0 & \cdots & \star
\end{pmatrix} \Rightarrow \begin{pmatrix}
0 & \cdots & 0 \\
\cdots & \cdots & \cdots \\
0 & \cdots & 0 \\
\cdots & \cdots & \cdots \\
0 & \cdots & \star
\end{pmatrix} \Rightarrow \begin{pmatrix}
0 & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & \cdots & \cdots
\end{pmatrix} \Rightarrow \begin{pmatrix}
0 & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & \cdots & \cdots
\end{pmatrix}
$$

It is easy to see that the fourth and fifth columns are also multiples of $(6)$. Hence, this array contains a codeword of the row code in every row and a codeword of the column code in every column. It is therefore a valid codeword of any code for $T_{5 \times 5}(2, 2, 0)$. As the used properties hold for any linear code we conclude that this pattern is never correctable.

We note that if both $G_{\text{row}}$ and $G_{\text{col}}$ are MDS, the first step, i.e., choosing the element in position $(2, 1)$, determines the whole matrix, as each of the subsequent steps is unique in this case.

**Remark 2.** The proof of Lemma 4 can be carried out for any column/row permutation of $E$, i.e., all these permutations are not correctable in $T_{5 \times 5}(2, 2, 0)$. Further, computer search shows that these 450 permutations are exactly the regular patterns that are not correctable in this topology.

**Theorem 3.** Conjecture 1 (cf. [8, Conjecture 3.1]) is false.

**Proof:** Follows immediately from Lemma 4.

**B. Implications of Uncorrectable Regular Patterns for Larger Topologies**

In this section, we show that the incorrectable erasure pattern for the topology $T_{5 \times 5}(2, 2, 0)$ given in Lemma 4 has implications on the set of correctable erasure for almost all topologies. To this end, we employ arguments based on shortening and puncturing.

**Definition 6.** Let $C$ be an $[N, k]$ code and $I \subseteq [N]$ be a set of integers. We define the shortening operator as

$$\text{short}_I(C) = \{c|_I|_I | c \in C, c|_I = 0|_I\}$$

and the puncturing operator as

$$\text{punct}_I(C) = \{c|_I|_I | c \in C\}.$$

We collect some well-known/basic properties in the following proposition.

**Proposition 2.** Let $G$ and $H$ denote a generator and parity-check matrix of an $[N, k]$ code $C$, respectively. Then,

1. $H|_{[N]|_I}$ is a parity-check matrix of $\text{short}_I(C)$ and $G|_{[N]|_I}$ is a generator matrix of $\text{punct}_I(C)$.
2. An erasure pattern $E \subset [N]$ is correctable if and only if it fulfills the equivalent conditions

$$\dim(\text{short}_{[N]|_I}(C)) = 0$$

$$\dim(\text{punct}_{[N]|_I}(C)) = k.$$

3. An erasure pattern $E \subset [N]$ is correctable only if the pattern $E \setminus I$ is correctable in the code $\text{short}_I(C)$ for any $I \subseteq [N]$.

4. An erasure pattern $E \subset [N]$ is correctable only if the pattern $E \setminus I$ is correctable in the code $\text{punct}_I(C)$ for any $I \subseteq [E]$.

**Proof:** Properties 1 and 2 are well-known.

**Proof of 3:** By property 2 we know that $E \setminus I$ is correctable in $\text{short}_I(C)$ if and only if $\dim(\text{short}_{[N]|_I \setminus E}(C)) = 0$. Now observe that

$$\text{short}_{[N]|_I \setminus E}(C) = \text{short}_{[N]|_I}(C) \oplus \text{short}_{E}(C) = \text{short}_{[N]|_I \setminus E}(C) \oplus \text{short}_{E}(C) = \text{short}_{[N]|_I \setminus E}(C) \oplus \text{short}_{E}(C).$$

As shortening does not increase the code dimension we have

$$\dim(\text{short}_{[N]|_I \setminus E}(C)) \leq \dim(\text{short}_{[N]|_I}(C)).$$

The statement follows from observing that $C$ is correctable in $C$ if and only if $\dim(\text{short}_{[N]|_I}(C)) = 0$.

**Proof of 4** By property 2 an erasure pattern is correctable if and only if $\dim(\text{punct}_{E}(C)) = 0$. The statement follows from observing that $\text{punct}_{[N]|_I}(\text{punct}_{E}(C)) = \text{punct}_{[N]|_I}(C)$.

**Lemma 5.** Let $C \in C_{m \times n}(a, b, 0)$. Denote by $\mathbb{I}_\text{col}$ and $\mathbb{I}_\text{row}$ the sets of information sets of the respective column and row code. Then for any $U \subseteq [m]$ such that $[m]|_U \cup \mathbb{I}_\text{col} \subseteq \mathbb{I}_\text{col}$ and $V \subseteq [n]$ such that $[n]|_V \cup \mathbb{I}_\text{row} \subseteq \mathbb{I}_\text{row}$ we have

$$\text{short}_{[m]|_U \cup [n]|_V}(C) \in C_{\mathbb{I}_\text{col} \times \mathbb{I}_\text{row}}(a-b, b-(n-v), 0).$$

Further, for any $U \subseteq \mathbb{I}_\text{col}$ and $V \subseteq \mathbb{I}_\text{row}$ we have

$$\text{punct}_{[m]|_U \cup [n]|_V}(C) \in C_{\mathbb{I}_\text{col} \times \mathbb{I}_\text{row}}(a-m-U, b-(n-v), 0).$$
Proof: As $h = 0$, by Definition 1, the generator matrix of $C$ is given by $G_{\text{col}} \otimes G_{\text{row}}$. Let $G_{\text{col}}^{\text{short}, \mathbf{a}, \mathbf{b}}$ and $G_{\text{row}}^{\text{short}, \mathbf{a}, \mathbf{b}}$ denote the generator matrices of the column/row code short/punctured in the positions indexed by $J$. It follows directly from the definition of the Kronecker product that $$\text{short}(\{m\} \times \{n\}) \otimes (\mathbf{u} \times \mathbf{v}) \equiv \{ \mathbf{c} \} = \left\{ G_{\text{col}}^{\text{short}, \mathbf{a}, \mathbf{b}} \otimes G_{\text{row}}^{\text{short}, \mathbf{a}, \mathbf{b}} \right\}.$$ The statement follows from observing that shortening a position of an information set of an $\{m, m-b\}$ code gives an $\{m-1, m-b-1\}$ code. Similarly, puncturing a position in the complement of an information set gives an $\{m-1, m-b\}$ code.

Lemma 6. If there exists a regular erasure pattern $E$ that is not correctable in $T_{m,n}(a,b,0)$ then there exists a regular erasure pattern $E'$ that is not correctable in $T_{m+\delta,n+\gamma}(a,b,0)$ for any $\delta, \gamma$.

Proof: We show that $E' = E$ is not correctable. As $E'$ does not contain additional erasures compared to $E$ and the restriction of Definition 4 does not depend on $m$ or $n$, we conclude that the pattern is regular, i.e., the regularity of a pattern in $T_{m,n}(a,b,0)$ directly implies the regularity of the same pattern in $T_{m+\delta,n+\gamma}(a,b,0)$ for any $\delta, \gamma$.

By Lemma 5 we have $\text{short}(\{m+\delta\} \times \{n+\gamma\}) \otimes (\mathbf{u} \times \mathbf{v}) \equiv \{ \mathbf{c} \} = \left\{ G_{\text{col}}^{\text{short}, \mathbf{a}, \mathbf{b}} \otimes G_{\text{row}}^{\text{short}, \mathbf{a}, \mathbf{b}} \right\}$. By assumption, the restricted pattern $E' \equiv \{ \mathbf{u} \times \mathbf{v} \} = E$ is not correctable in $T_{m,n}(a,b,0)$ and the statement follows from 3) in Proposition 2.

Lemma 7. If there exists a regular erasure pattern $E$ that is not correctable in $T_{m,n}(a,b,0)$ then there exists a regular erasure pattern $E'$ that is not correctable in $T_{m+\delta,n+\gamma}(a+b,0)$ for any $\delta, \gamma$.

Proof: Denote by $E_{\text{col}}'$ and $E_{\text{row}}'$ the column and row code of a code $C' \in C_{m+\delta,n+\gamma}(a+b,0)$. Without loss of generality assume that the first $m - a$ and $n - b$ positions are a (subset of an) information set of the code $C_{\text{col}}'$ and $C_{\text{row}}'$. Let $E'$ be the erasure pattern obtained by adding $\delta$ rows and $\gamma$ columns of erasures to $E$, i.e., $E' \equiv E \cup \left\{ (\mathbf{m}+\mathbf{d}) \times \mathbf{n}+\mathbf{g} \right\}$, where $\mathbf{d}$ is the size of the erasures in the columns and $\mathbf{g}$ is the size of the erasures in the rows. We show that this pattern is regular for the topology $T_{m+n}(\mathbf{d}, \mathbf{g})$, i.e., fulfills Definition 4 for any $U \subseteq \mathbf{d}$, $V \subseteq \mathbf{g}$, $U \cap V = \emptyset$. As the erasures in the columns and rows consist only of erasures it suffices to show that the subsets of rows/columns of $[m+\delta, n+\gamma] \subseteq T_{m+n}(\mathbf{d}, \mathbf{g})$, $\mathbf{d} \times \mathbf{g} \subseteq T_{m+n}(\mathbf{d}, \mathbf{g})$ fulfill the condition. Define $U = U' \cap [m]$ and $V = V' \cap [n]$ and observe that $E'$ can be partitioned into two disjoint subsets with $$|E' \cap (U \times V)| \leq |E| - (v-\gamma) + (a-\delta) - ab.$$ By Lemma 5 we have $\text{punct}(\{m+\delta\} \times \{n+\gamma\}) \otimes (\mathbf{u} \times \mathbf{v}) \equiv \{ \mathbf{c} \} = \left\{ G_{\text{col}}^{\text{punct}, \mathbf{a}, \mathbf{b}} \otimes G_{\text{row}}^{\text{punct}, \mathbf{a}, \mathbf{b}} \right\}$. By Definition 1, the generator matrix of $C$ is given by $G_{\text{col}} \otimes G_{\text{row}}$. Let $G_{\text{col}}^{\text{short}, \mathbf{a}, \mathbf{b}}$ and $G_{\text{row}}^{\text{short}, \mathbf{a}, \mathbf{b}}$ denote the generator matrices of the column/row code short/punctured in the positions indexed by $J$. It follows directly from the definition of the Kronecker product that $$\text{short}(\{m\} \times \{n\}) \otimes (\mathbf{u} \times \mathbf{v}) \equiv \{ \mathbf{c} \} = \left\{ G_{\text{col}}^{\text{short}, \mathbf{a}, \mathbf{b}} \otimes G_{\text{row}}^{\text{short}, \mathbf{a}, \mathbf{b}} \right\}.$$ The statement follows from observing that shortening a position of an information set of an $\{m, m-b\}$ code gives an $\{m-1, m-b-1\}$ code. Similarly, puncturing a position in the complement of an information set gives an $\{m-1, m-b\}$ code.

Lemma 6. If there exists a regular erasure pattern $E$ that is not correctable in $T_{m,n}(a,b,0)$ then there exists a regular erasure pattern $E'$ that is not correctable in $T_{m+\delta,n+\gamma}(a+b,0)$ for any $\delta, \gamma$.

Proof: We show that $E' = E$ is not correctable. As $E'$ does not contain additional erasures compared to $E$ and the restriction of Definition 4 does not depend on $m$ or $n$, we conclude that the pattern is regular, i.e., the regularity of a pattern in $T_{m,n}(a,b,0)$ directly implies the regularity of the same pattern in $T_{m+\delta,n+\gamma}(a,b,0)$ for any $\delta, \gamma$.

By Lemma 5 we have $\text{short}(\{m+\delta\} \times \{n+\gamma\}) \otimes (\mathbf{u} \times \mathbf{v}) \equiv \{ \mathbf{c} \} = \left\{ G_{\text{col}}^{\text{short}, \mathbf{a}, \mathbf{b}} \otimes G_{\text{row}}^{\text{short}, \mathbf{a}, \mathbf{b}} \right\}$. By assumption, the restricted pattern $E' \equiv \{ \mathbf{u} \times \mathbf{v} \} = E$ is not correctable in $T_{m,n}(a,b,0)$ and the statement follows from 3) in Proposition 2.

Lemma 7. If there exists a regular erasure pattern $E$ that is not correctable in $T_{m,n}(a,b,0)$ then there exists a regular erasure pattern $E'$ that is not correctable in $T_{m+\delta,n+\gamma}(a+b,0)$ for any $\delta, \gamma$.

Proof: Denote by $E_{\text{col}}'$ and $E_{\text{row}}'$ the column and row code of a code $C' \in C_{m+\delta,n+\gamma}(a+b,0)$. Without loss of generality assume that the first $m - a$ and $n - b$ positions are a (subset of an) information set of the code $C_{\text{col}}'$ and $C_{\text{row}}'$. Let $E'$ be the erasure pattern obtained by adding $\delta$ rows and $\gamma$ columns of erasures to $E$, i.e., $E' \equiv E \cup \left\{ (\mathbf{m}+\mathbf{d}) \times \mathbf{n}+\mathbf{g} \right\}$, where $\mathbf{d}$ is the size of the erasures in the columns and $\mathbf{g}$ is the size of the erasures in the rows. We show that this pattern is regular for the topology $T_{m+n}(\mathbf{d}, \mathbf{g})$, i.e., fulfills Definition 4 for any $U \subseteq \mathbf{d}$, $V \subseteq \mathbf{g}$, $U \cap V = \emptyset$. As the erasures in the columns and rows consist only of erasures it suffices to show that the subsets of rows/columns of $[m+\delta, n+\gamma] \subseteq T_{m+n}(\mathbf{d}, \mathbf{g})$, $\mathbf{d} \times \mathbf{g} \subseteq T_{m+n}(\mathbf{d}, \mathbf{g})$ fulfill the condition. Define $U = U' \cap [m]$ and $V = V' \cap [n]$ and observe that $E'$ can be partitioned into two disjoint subsets with $$|E' \cap (U \times V)| \leq (v-\gamma) + (a-\delta) - ab.$$

$\square$
REFERENCES