Combinatorics of Singly-Repairable Families

Eugene M. Luks
Computer Science Department,
University of Oregon,
Eugene, OR 97403.
luks@cs.uoregon.edu

Amitabha Roy
Computer Science Department,
Boston College,
Chestnut Hill, MA 02167.
aroy@cs.bc.edu

Submitted: Jan 17, 2005; Accepted: Oct 18, 2005; Published: Nov 15, 2005
Mathematics Subject Classifications: 05D05, 68R05, 94B65

Abstract

A non-empty set $\mathcal{F}$ of n-bit vectors over alphabet $\{0,1\}$ is called singly repairable, if every vector $u \in \mathcal{F}$ satisfies the following conditions:

(i) if any bit of u is changed (from 0 to 1 or vice versa), the new vector does not belong to $\mathcal{F}$

(ii) there is a unique choice of a different bit that can then be changed to give another vector $\neq u$ in $\mathcal{F}$.

Such families $\mathcal{F}$ exist only for even n and we show that $2^{n/2} \leq |\mathcal{F}| \leq \frac{2^{n+1}}{(n+2)}$. The lower bound is tight for all even n and we show that the families of this size are unique under a natural notion of isomorphism (namely, translations and permutation of coordinates). We also construct families that achieve the upper bound when n is of the form $2^m - 2$. For general even n, we construct families of size at least $2^n/n$. Of particular interest are *minimal* singly-repairable families. We show that such families have size at most $2^n/n$ and we construct families achieving this upper bound when n is a power of 2. For general even n, we construct minimal families of size $\Omega(2^n/n^2)$. The study of these families was inspired by a computational scheduling problem.

1 Introduction

In this paper, we study the extremal combinatorics of a family $\mathcal{F}$ of n-bit vectors from $\{0,1\}^n$ such that every vector $u \in \mathcal{F}$ satisfies the following properties:

(a) negating any bit of u in $\mathcal{F}$ produces a vector v not in $\mathcal{F}$ (we call such a bit flip a "break").

(b) there is a unique choice of some other bit (we call this the "repair" bit) of v which when negated produces a vector in $\mathcal{F}$.

By "negating a bit" we mean flipping the value from 0 to 1 (or from 1 to 0). These families, which we call singly repairable, arise in the context of *fault-tolerant solutions* (formulated by [2]) for scheduling problems. These are special solutions to optimization problems (e.g., resource allocation) that are tolerant to unforeseen events, e.g., a resource suddenly becoming unavailable. In the event of such a "break", there is some other resource which could be brought into play as a "repair" and maintain optimality. In this paper, we are concerned with the combinatorics of families of vectors which admit the break-repair property.

We prove the following:

Theorem 1.1. *Let $n > 0$ be even and let $\mathcal{F}$ be a collection of vectors from $\{0,1\}^n$. If $\mathcal{F}$ is singly repairable, then*

$$2^{n/2} \leq |\mathcal{F}| \leq \frac{2^{n+1}}{n+2}.$$

The lower bound is achieved for all even n. Moreover, the families achieving the lower bound are unique up to permutation of coordinates and translations. The upper bound is achieved when n is of the form $2^m - 2$. For arbitrary even n, there exists a singly-repairable family of size at least $2^n/n$.

Of particular interest are *minimal* singly-repairable families. In terms of our applications, minimal singly-repairable families connect any two fault tolerant solutions via some sequence of breaks and repairs.

Theorem 1.2. *Let $n > 0$ be even and let $\mathcal{F}$ be a family of vectors from $\{0,1\}^n$. If $\mathcal{F}$ is minimal singly repairable, then*

$$2^{n/2} \leq |\mathcal{F}| \leq 2^n/n.$$

The lower bound is achieved for all even n. The upper bound is achieved when n is a power of 2. For general even n, there exists a minimal singly-repairable family of size $\Omega(2^{n+1-r}/n)$ where r is the number of 1's in the binary representation of n.

More generally, one may consider repairable families where we place no restriction on the number of repairs. We intend to study the combinatorics of these families in a future paper. The computational complexity of finding robust solutions, inspired by research in [2], appears in [4].

Organization of the paper: In Section 2, we introduce definitions and notation used in the rest of the paper. In Section 3, we prove upper and lower bounds on the sizes of singly-repairable families and construct families achieving these bounds. Then in Section 4, we consider minimal singly-repairable families and give constructions for families achieving the largest possible size.

2 Definitions and Notation

The intended objects of study are n-bit vectors over $\{0,1\}$. The collection of all such vectors is denoted as $\mathbb{Z}_2^n$ or $\{0,1\}^n$. Frequently, we shall consider $\mathbb{Z}_2^n$ as a vector space (and not just as a collection of vectors) over $\{0,1\}$. We assume that vectors are indexed by $i \in \{0,1,\ldots,n-1\}$, where v_i refers to the i-th bit of v.

A translation of $\mathcal{F} \subseteq \mathbb{Z}_2^n$ by a vector $v \in \mathbb{Z}_2^n$ is the set $\mathcal{F} + v = \{u+v | \ u \in \mathcal{F}\}$. Two families of vectors are said to be isomorphic if they are related by an element of the group generated by permutations of coordinates and translations.

The (Hamming) weight of a vector is the number of coordinates with a 1. A vector has even (resp. odd) parity if its weight is even (resp. odd). Let $\mathcal{E}(n)$ (resp. $\mathcal{O}(n)$) denote all the even weight (resp. odd weight) vectors of length n. Given two vectors $u, v \in \mathbb{Z}_2^n$, the (Hamming) distance between u and v, denoted by $d(u,v)$, is the weight of $u+v$ (equivalently, it is the number of positions where u and v differ).

Let $X_i \subseteq \mathbb{Z}_2^{n_i}$ for $1 \le i \le r$ be non-empty families of vectors. Define $(X_1 \mid X_2 \cdots \mid X_r) \subseteq \mathbb{Z}_2^{\sum_i n_i}$ to be the collection of $\prod_{i=1}^r |X_i|$ vectors, each denoted by $(x_1 \mid x_2 \ldots \mid x_r)$ formed by concatenating, in order, vectors $x_1 \in X_1$, $x_2 \in X_2, \ldots, x_r \in X_r$.

The basic operations on vectors are bit flips (negations): changing a specified bit from a one to a zero (or from zero to one). Given an n-bit vector u, let $\partial_i(u)$ denote the vector u with the i-th bit flipped. We can extend this definition to a set of bit flips: $\partial_S(u)$ represents the vector with bits in the set $S \subseteq \{0,1,\ldots,n-1\}$ flipped. When $|S| = 2$ (say $S = \{i,j\}$), we write $\partial_{ij}(u)$ for simplicity (and when we write $\partial_{ij}(u)$ it will be implicitly understood that $i \neq j$).

Definition 2.1. *Let $\mathcal{F} \subseteq \mathbb{Z}_2^n$ be a family of vectors. We say that $\mathcal{F}$ is <u>singly repairable</u> if every vector $u \in \mathcal{F}$ satisfies the following conditions:*

(i) for all i, $0 \le i \le n-1$, $\partial_i(u) \notin \mathcal{F}$.

(ii) for all i, $0 \le i \le n-1$, there exists a unique j where $0 \le j \le n-1$ and $j \neq i$ such that $\partial_{ij}(u) \in \mathcal{F}$.

***Remark*.** (i) We interpret bit flips as breaks and repairs. Let $u \in \mathcal{F}$ and suppose the i-th bit of u is flipped, we refer to ∂_i as a *break* since $\partial_i(u) \notin \mathcal{F}$. The *repair* to that break is flipping the j-th bit for some unique coordinate j, $j \neq i$ such that $\partial_{ij}(u) \in \mathcal{F}$. In other words, singly-repairable families are such that every member of the family has a unique repair for every break.

(ii) The set of all singly-repairable subfamilies of $\mathbb{Z}_2^n$ is closed under isomorphisms. For simplifying proofs, we often translate a given $\mathcal{F} \subseteq \mathbb{Z}_2^n$ by a suitable vector to obtain an isomorphic copy which contains 0^n.

We disallow vectors at Hamming distance 1 from each other in any singly-repairable family. To emphasize this crucial property, we call any family $\mathcal{F}$ (not necessarily singly-repairable) in $\mathbb{Z}_2^n$ *diffuse* if no pair of distinct vectors in $\mathcal{F}$ are at distance 1 from each other.

Example 2.1. Let $n \geq 2$ be even. The set of n-bit vectors v that satisfy the formula

$$(v_0 = v_1) \wedge (v_2 = v_3) \wedge \cdots \wedge (v_{n-2} = v_{n-1})$$

is a singly-repairable family of size $2^{n/2}$. This easy example achieves a lower bound (see Theorem 3.7) on the size of singly-repairable subfamilies of $\mathbb{Z}_2^n$. Even more strikingly, it is the unique family, up to isomorphisms, that achieves this lower bound (see Theorem 3.10).

3 General Bounds

In the following discussion, let $\mathcal{F} \subseteq \mathbb{Z}_2^n$ be singly repairable. A vector $u \in \mathcal{F}$ induces a relation on $\{0, 1, \ldots, n-1\}$ as follows: $i \sim_u j$ if $\partial_{ij}(u) \in \mathcal{F}$. This implies that $\sim_u$ is a symmetric relation on $\{0, 1, \ldots, n-1\}$ which partitions it into break-repair sets, each of size 2. This also implies that n has to be even, a fact which we will henceforth assume throughout the paper unless explicitly mentioned otherwise.

We will, on occasion, treat a singly-repairable family $\mathcal{F}$ as an undirected graph: the vertices are the vectors in $\mathcal{F}$ and the edges are $\{u, v\}$ where $v = \partial_{ij}(u), u, v \in \mathcal{F}$ for some $i \neq j, 0 \leq i, j \leq n-1$. We shall refer to this graph as the *break-repair* graph of $\mathcal{F}$. Without risk of confusion, we sometimes call $\mathcal{F}$ a graph (when we really mean the break-repair graph of $\mathcal{F}$) and refer to vertices, paths, cycles etc., in $\mathcal{F}$. This graph theoretic view of $\mathcal{F}$ enables us to study the lattice of singly-repairable subfamilies of $\mathcal{F}$. In particular, the connected components of this graph correspond to *minimal* singly-repairable subfamilies. Lemma 3.1 below records these easily provable facts.

Lemma 3.1. *Let $\mathcal{F} \subseteq \mathbb{Z}_2^n$ be a non-empty singly-repairable family. Then the following hold:*

a) *n is even.*

b) *If $v = \partial_{ij}(u)$ and $w = \partial_{kl}(u)$ where u, v, w are three distinct vectors in $\mathcal{F}$, then $\{i, j\} \cap \{k, l\} = \varnothing$.*

c) *$\mathcal{F}$ is minimal iff it is connected (as a graph).*

Remark. Note that it is important to include u in the definition of the relation $\sim_u$, different u's might give rise to different relations. In Section 3.2, we show the special role of $\sim_u$ by showing that the smallest minimal families are essentially unique.

3.1 Upper Bounds

We now prove upper bounds on the size of singly-repairable subfamilies of $\mathbb{Z}_2^n$.

Proposition 3.1. *If $\mathcal{F} \subseteq \mathbb{Z}_2^n$ is a singly-repairable family consisting of vectors of even weight, then $|\mathcal{F}| \leq \frac{2^n}{n}$.*

Proof. Each vector in $\mathcal{F}$ has n neighbors at distance 1 and since $\mathcal{F}$ is singly repairable, each such neighbor is counted exactly twice (otherwise Lemma 3.1 (b) would be violated). Thus $\mathcal{F}$ has $|\mathcal{F}|n/2$ neighbors of odd weight. So we have $|\mathcal{F}|n/2 \leq 2^{n-1}$, from which the result follows. $\square$

Corollary 3.2. *If $\mathcal{F} \subseteq \mathbb{Z}_2^n$ is a minimal singly-repairable family, then $|\mathcal{F}| \leq \frac{2^n}{n}$.*

Proof. We can translate $\mathcal{F}$ by a suitable vector in $\mathbb{Z}_2^n$ to ensure that $0^n \in \mathcal{F}$. Since $\mathcal{F}$ is connected (Lemma 3.1 (c)), every vector in $\mathcal{F}$ has even weight. The result now follows from Proposition 3.1. $\square$

More generally, we have the following bound.

Corollary 3.3. *If $\mathcal{F} \subseteq \mathbb{Z}_2^n$ is singly repairable, then $|\mathcal{F}| \leq \frac{2^{n+1}}{n+2}$.*

Proof. Let $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_E \subseteq \mathbb{Z}_2^n$ be singly repairable, where $\mathcal{F}_O$ (resp. $\mathcal{F}_E$) consist of the odd weight (resp. even weight) vectors in $\mathcal{F}$. Then consider $\mathcal{F}' \subseteq \mathbb{Z}_2^{n+2}$ where

$$\mathcal{F}' = (\mathcal{F}_O \,|\, \{01, 10\}) \cup (\mathcal{F}_E \,|\, \{00, 11\})$$

Observe that there is no vector in $\mathcal{F}_E$ at distance 1 from a vector in $\mathcal{F}_0$. Thus $\mathcal{F}'$ is singly repairable and consists of vectors of even weight. Since $|\mathcal{F}'| = 2|\mathcal{F}|$, the result follows from Proposition 3.1. $\square$

Remark. Single repairability implies the absence of equilateral triangles of side length 2 but the latter is a weaker condition. In fact, Problem B-6 of the 61st William Lowell Putnam Examination (2000) essentially established a bound of $2^{n+1}/n$ for families that exclude equilateral triangles of side length 2. [1]

We now describe a class of examples of singly-repairable subfamilies of $\mathbb{Z}_2^n$ of size $2^{n/2}$. These families are then used to construct singly-repairable families that achieve the maximum size of $2^{n+1}/(n+2)$ (from Corollary 3.3) for infinitely many values of n.

Example 3.1. Let $s \in \mathbb{Z}_2^{n/2}$. Let $\mathcal{B}_s \subseteq \mathbb{Z}_2^n$ denote the set of vectors v such that

$$s_i = 0 \Rightarrow v_{2i+1} \neq v_{2i}$$
$$s_i = 1 \Rightarrow v_{2i+1} = v_{2i}$$

where $0 \leq i \leq n/2 - 1$. Each $\mathcal{B}_s$ is singly repairable and has size $2^{n/2}$. Moreover, any pair of families $\mathcal{B}_s, \mathcal{B}_t$ are isomorphic (they are related by a translation).

Recall that an $[n, d]$ code [3] is a subset of Z_2^n such that the minimum Hamming distance between any two distinct vectors is d. An $[n, k, d]$ linear code is an $[n, d]$ code that is a k-dimensional subspace of $\mathbb{Z}_2^n$.

[1]The first author served on the 2000 Putnam Questions Committee.

Lemma 3.4. *Let $\mathcal{F}$ be an $[n/2, 3]$ code. Then $\bigcup_{s\in\mathcal{F}} \mathcal{B}_s$ is a singly-repairable subfamily of $\mathbb{Z}_2^n$.*

Proof. Let $s, t \in \mathcal{F}$ be two distinct vectors in $\mathcal{F}$. Any vector $v \in \mathbb{Z}_2^n$ that satisfies $v_{2i+1} = v_{2i}$ is at least distance 1 away from any vector $w \in \mathbb{Z}_2^n$ that satisfies $w_{2i+1} \neq w_{2i}$, where $0 \leq i \leq n/2 - 1$. Since $d(s,t) \geq 3$, $v \in \mathcal{B}_s$ is at least distance 3 away from $w \in \mathcal{B}_t$. Hence, $\mathcal{B}_s \cup \mathcal{B}_t$ is singly repairable and more generally, $\bigcup_{s\in\mathcal{F}} \mathcal{B}_s$ is singly repairable. ☐

Theorem 3.5.

(i) There exist singly-repairable subfamilies of $\mathbb{Z}_2^n$ of size $\Theta(2^n/n)$ for all even n.

(ii) There exist singly-repairable subfamilies of $\mathbb{Z}_2^n$ of size $\frac{2^{n+1}}{n+2}$ when n is of the form $2^m - 2$.

Proof. (i) There is a $[m, m - \lfloor \log_2 m \rfloor - 1, 3]$ linear code (also called the shortened Hamming code, see [1], section 2.6, page 47) which, for $m = n/2$, is a linear code in $\mathbb{Z}_2^{n/2}$ of size at least $\frac{2^{n/2}}{n}$. Using this code as the family $\mathcal{F}$ in Lemma 3.4, we construct a singly-repairable subfamily of $\mathbb{Z}_2^n$ of size at least $2^n/n$.

(ii) It is well-known via the Gilbert Varshamov bound [[3], page 33, Theorem 12], that a linear code with parameters $[n, k, d]$ exists if

$$\sum_{i=0}^{d-2} \binom{n-1}{i} \leq 2^{n-k}.$$

Hence there is a $[n/2, k, 3]$ linear code when $n = 2^m - 2$ and $k = 2^{m-1} - m$ for some integer m. Using this code as $\mathcal{F}$ in Lemma 3.4, our construction produces a singly-repairable family of size $(2^{n/2})\, 2^k = 2^{n+1}/(n+2)$.

☐

While we have achieved the theoretical upper bound for singly-repairable families, we were particularly interested in what can happen for *minimal* singly-repairable families. In Section 4, we show that the upper bound for minimal families is achievable for every value of n which is a power of 2.

3.2 Lower Bounds

In this section, we prove that any singly-repairable subfamily of $\mathbb{Z}_2^n$ has size at least $2^{n/2}$. We first introduce a notion of *partial repairable* subfamilies of $\mathbb{Z}_2^n$. This concept makes sense even when n is odd and we temporarily suspend the restriction that n is even in our discussions involving partial repairability.

Recall that a *diffuse* family is a family $\mathcal{F} \subseteq \mathbb{Z}_2^n$ such that no two vectors in $\mathcal{F}$ are at distance 1 from each other.

Definition 3.1. *Let $\mathcal{F} \subseteq \mathbb{Z}_2^n$ be a diffuse family. Then a break-repair pair for $u \in \mathcal{F}$ is a pair $\{i, j\} \subseteq \{0, 1, \ldots, n-1\}$, with $i \neq j$, such that*

(i) $\partial_{ij}(u) \in \mathcal{F}$ *and*

(ii) for all k, $0 \leq k \leq n-1$ where $k \neq i, j$, $\partial_{ik}(u) \notin \mathcal{F}$ and $\partial_{jk}(u) \notin \mathcal{F}$.

Notation: If $\mathcal{F}$ is a diffuse subfamily of $\mathbb{Z}_2^n$, we denote:

$$\mathcal{E}_{\mathcal{F}}(u) = \{\{i, j\} | \{i, j\} \text{ is a break-repair pair for } u \text{ in } \mathcal{F}\}.$$

Definition 3.2. *Let $\mathcal{F} \subseteq \mathbb{Z}_2^n$, $r \leq n/2$. Then $\mathcal{F}$ is called r-singly repairable if $\mathcal{F}$ is diffuse and $|\mathcal{E}_{\mathcal{F}}(u)| \geq r$ for all $u \in \mathcal{F}$ (i.e., every vector in $\mathcal{F}$ has at least r break-repair pairs).*

Note that when n is even, an $n/2$-singly repairable family is our usual singly-repairable family (Definition 2.1).

Remark. *Every* diffuse subfamily in $\mathbb{Z}_2^n$ is r-singly repairable for some r, where $0 \leq r \leq n/2$.

Lemma 3.6. *If $\mathcal{F} \subseteq \mathbb{Z}_2^n$ is a non-empty r-singly repairable family ($r \geq 1$), then $|\mathcal{F}| \geq 2^r$.*

Proof. By induction on r. The result is clear for $r = 1$: a 1-singly repairable family has to have a vector u which has at least one break-repair pair, thereby forcing another vector $v = \partial_{ij}(u)$ (for some $0 \leq i, j \leq n-1$) to also be a member of the family. Then assume that the result is true for $r = s-1$ where $s \geq 2$. We prove it true for $r = s$.

Let $\mathcal{F}$ be s-singly repairable. Choose a coordinate i, where $0 \leq i \leq n-1$ for which there is some vector $u \in \mathcal{F}$ such that $u_i = 1$ and some $v \in \mathcal{F}$ such that $v_i = 0$. Such an i must exist since $s \geq 2$. So $\mathcal{F}_{i,1} = \{w \in \mathcal{F} \mid w_i = 1\}$ and $F_{i,0} = \{w \in \mathcal{F} \mid w_i = 0\}$ are both non-empty. They are clearly diffuse.

Observe that both $\mathcal{F}_{i,0}$ and $\mathcal{F}_{i,1}$ are $(s-1)$-singly repairable (since i can be a member of at most one break-repair pair for each vector w in $\mathcal{F}$). By the induction hypothesis, this means that $|\mathcal{F}_{i,0}| \geq 2^{s-1}$ and $\mathcal{F}_{i,1} \geq 2^{s-1}$. Since $\mathcal{F}_{i,0} \cap \mathcal{F}_{i,1} = \varnothing$, $|\mathcal{F}| \geq 2^s$. $\square$

Theorem 3.7. *If $\mathcal{F} \subseteq \mathbb{Z}_2^n$ is singly repairable, then $|\mathcal{F}| \geq 2^{n/2}$.*

Proof. When $\mathcal{F} \subseteq \mathbb{Z}_2^n$ is singly repairable, it is $n/2$-repairable and hence the desired bound follows from Lemma 3.6. $\square$

Remark. It is worth noting that a more direct inductive approach, whereby we take a singly-repairable family $\mathcal{F} \subseteq \mathbb{Z}_2^n$ and consider

$$\mathcal{F}_1 = \{u \in \mathcal{F} | \ \{0, 1\} \text{ is a break-repair pair for } u \in \mathcal{F}\}$$

and $\mathcal{F}_2 = \mathcal{F} \setminus \mathcal{F}_1$ and inducting on $\mathcal{F}_1$ or $\mathcal{F}_2$ fails, as neither may be singly repairable (the family in Figure 1 is the smallest counterexample).

3.3 Uniqueness of Family Achieving Lower Bound

In this section, we prove that a singly-repairable family in $\mathbb{Z}_2^n$ of size $2^{n/2}$ is isomorphic (under permutations of coordinates or affine translations) to any $\mathcal{B}_s$, where $s \in \mathbb{Z}_2^{n/2}$.

This will imply that there is one canonical smallest singly-repairable family up to isomorphisms, for example, $\mathcal{B}_{1^n}$.

Definition 3.3. *A family $\mathcal{F} \subseteq \mathbb{Z}_2^n$ is called pure if it is diffuse and if for every $u, v \in \mathcal{F}$, $\mathcal{E}_{\mathcal{F}}(u) = \mathcal{E}_{\mathcal{F}}(v)$. A diffuse family that is not pure is called impure.*

Lemma 3.8. *Let $\mathcal{F} \subseteq \mathbb{Z}_2^n$ be pure, singly repairable and of size $2^{n/2}$. Then $\mathcal{F}$ is isomorphic to $\mathcal{B}_{1^n}$,*

Proof. Translating and permuting coordinates in $\mathcal{F}$, if necessary, we obtain an isomorphic family $\mathcal{F}'$ such that $0^n \in \mathcal{F}'$ and $\{2i+1, 2i\} \in \mathcal{E}_{\mathcal{F}'}(u)$ for every vector $u \in \mathcal{F}'$ where $0 \leq i \leq n/2 - 1$. Since $\mathcal{F}'$ has size $2^{n/2}$, it is connected and hence $\mathcal{F}'$ is isomorphic to $\mathcal{B}_{1^n}$. $\square$

Let $\mathcal{F} \subseteq \mathbb{Z}_2^n$ be impure r-singly repairable. Similar to the proof of Lemma 3.6, we define for $0 \leq i \leq n-1$, the family $\mathcal{F}_{i,1} = \{w \in \mathcal{F} |\ w_i = 1\}$ and $\mathcal{F}_{i,0} = \{w \in \mathcal{F} |\ w_i = 0\}$. Both $\mathcal{F}_{i,1}$ and $\mathcal{F}_{i,0}$ are diffuse and the argument used in the proof of Lemma 3.6 shows that if $u \in \mathcal{F}$, $\mathcal{E}_{\mathcal{F}}(u) \supset \mathcal{E}_{\mathcal{F}_{i,j}}(u)$ for $j = 0, 1$ and if both $\mathcal{F}_{i,0}$ and $\mathcal{F}_{i,1}$ are non-empty, then they are both $(r-1)$-singly repairable (however, they may be pure).

Lemma 3.9. *Let $\mathcal{F} \subseteq \mathbb{Z}_2^n$ be a non-empty r-singly repairable impure family where $r \geq 1$. Then $|\mathcal{F}| \geq 2^r + 1$.*

Proof. (By induction) If $r = 1$: without loss of generality, an impure 1-singly repairable family includes vectors $u, v, w \in \mathbb{Z}_2^n$, where $v = \partial_{12}(u)$ and $w = \partial_{34}(z)$ (where either $z = u$ or v or some fourth vector in $\mathcal{F}$). Since there are at least 3 vectors in $\mathcal{F}$, the result holds for $r = 1$.

Assume that $r \geq 2$. We show that there must be a coordinate $i, 0 \leq i \leq n-1$, such that $\mathcal{F}_{i,0}$ and $\mathcal{F}_{i,1}$ are both non-empty, $(r-1)$-singly repairable and at least one of them is impure. Then $|\mathcal{F}| = |\mathcal{F}_{i,0}| + |\mathcal{F}_{i,1}| \geq (2^{r-1} + 1) + 2^{r-1} = 2^r + 1$ (by induction), thereby establishing the bound for impure r-singly repairable families.

Suppose first that for some distinct indices $i, j, k \in \{0, 1, \ldots, n-1\}$, there exist vectors $u, v \in \mathcal{F}$ such that $\{i, j\} \in \mathcal{E}_{\mathcal{F}}(u)$ and $\{j, k\} \in \mathcal{E}_{\mathcal{F}}(v)$. Then u and $\partial_{ij}(u)$ are split between $\mathcal{F}_{i,0}$ and $\mathcal{F}_{i,1}$ and whichever of these contains v is impure since $\{j, k\}$ is not in $\mathcal{E}_{\mathcal{F}}(u)$ or $\mathcal{E}_{\mathcal{F}}(\partial_{ij}(u))$. So we may assume that for all $X, Y \in \bigcup_{u \in \mathcal{F}} \mathcal{E}_{\mathcal{F}}(u)$, either $X = Y$ or $X \cap Y = \emptyset$.

Since $\mathcal{F}$ is impure, there is some pair $\{i, j\} \subseteq \{0, 1, \ldots, n-1\}$ that belongs to at least one $\mathcal{E}_{\mathcal{F}}(u)$ for some $u \in \mathcal{F}$ but does not belong to all $\mathcal{E}_{\mathcal{F}}(w)$ for all $w \in \mathcal{F}$. Without loss of generality, assume that such a vector u is $(0, 0, \ldots, 0) \in \mathcal{F}$. Since $r \geq 2$, there is some other break-repair pair in $\mathcal{E}_{\mathcal{F}}(u)$. Again, without loss of generality, assume that $\{0,1\}$ is such a break-repair pair (so that $\{0,1\} \neq \{i,j\}$ and both $\{0,1\}$ and $\{i,j\}$ are in $\mathcal{E}_{\mathcal{F}}(u)$). Let $v = (1,1,0,\ldots,0)$ so that $v \in \mathcal{F}$ and $\{0,1\} \in \mathcal{E}_{\mathcal{F}}(v)$. Then $v \in \mathcal{F}_{0,1}$ and $u \in \mathcal{F}_{0,0}$ so that both $\mathcal{F}_{0,0}$ and $\mathcal{F}_{0,1}$ are non-empty. If either $\mathcal{F}_{0,0}$ or $\mathcal{F}_{0,1}$ is impure, we are done.

So assume that $\mathcal{F}_{0,0}$ and $\mathcal{F}_{0,1}$ are both pure. Since $\{i,j\}$ is a break-repair pair for $u \in \mathcal{F}_{0,0}$, this means that $\{i,j\}$ is a break-repair pair for every vector in $\mathcal{F}_{0,0}$. Moreover, $\{i,j\}$ cannot be a break-repair pair for any vector in $\mathcal{F}_{0,1}$: if so, then it would be a break-repair pair for *every* vector in the pure family $\mathcal{F}_{0,1}$. This in turn would make $\{i,j\}$ a break repair pair for every vector in $\mathcal{F}$ and that contradicts the choice of $\{i,j\}$. Our assumption that break-repair pairs are disjoint in $\bigcup_{u\in\mathcal{F}} \mathcal{E}_{\mathcal{F}}(u)$ implies that i does not participate in any break-repair pair for v in $\mathcal{F}_{0,1}$. In this situation, choose any coordinate $l \in \{0,1,\ldots,n-1\} \setminus \{0,1,i,j\}$, which takes part in a break-repair pair for v in $\mathcal{F}$, such an l (in another break-repair pair) exists since $r \geq 2$. Furthermore, if $\{l,l'\} \in \mathcal{E}_{\mathcal{F}}(v)$ then $l' \notin \{0,1,i,j,l\}$. Then $\mathcal{F}_{l,0}$ and $\mathcal{F}_{l,1}$ are both non-empty (since coordinate l is broken in some vector in $\mathcal{F}$) and u and v belong to $\mathcal{F}_{l,0}$. Since u and v have different break-repair pairings in $\mathcal{F}_{l,0}$, it follows that $\mathcal{F}_{l,0}$ is an impure family. $\square$

Theorem 3.10. *A singly-repairable family $\mathcal{F} \subseteq \mathbb{Z}_2^n$ of size $2^{n/2}$ is isomorphic to $\mathcal{B}_{1^n}$.*

Proof. Lemma 3.9 implies that any impure singly-repairable family $\mathcal{F} \subseteq \mathbb{Z}_2^n$ has size $> 2^{n/2}$. Hence, $\sim_u$ must be constant for all $u \in \mathcal{F}$ and Lemma 3.8 implies that $\mathcal{F} \cong \mathcal{B}_{1^n}$. $\square$

4 Minimal Singly-Repairable Families

The singly-repairable families constructed in Theorem 3.5 have a large number of connected components, each component being a minimal singly-repairable family. We now consider the problem of finding whether we can attain these bounds with just one connected component. Corollary 3.2 tells us the best we can hope to do and we show that we can indeed achieve this upper bound for infinitely many n.

More specifically, we prove that when n is a power of 2, there are minimal singly-repairable subfamilies of $\mathbb{Z}_2^n$ of size $2^n/n$ (Section 4.1). For general even n, we construct a minimal singly-repairable family of size $\Omega(2^n/n^2)$ (Section 4.2).

Notation. For an integer $s \in \{0,1,\ldots 2^r-1\}$, we let s_i $(0 \leq i \leq r-1)$ denote the i-th least-significant bit in the binary representation of s. For $0 \leq i \leq r-1$, $e_r(i) \in \mathbb{Z}_2^r$ is a vector such that $e_r(i)_j = 1$ iff $i = j$.

An m-bit binary Gray code [3] is an ordering $(u_0, u_1, \ldots, u_{2^m-1})$ of vectors in $\mathbb{Z}_2^m$ such that any two successive vectors differ in exactly one bit. A Gray code is cyclic if u_{2^m-1} also differs from u_0 in one bit. There are many constructions known for non-isomorphic cyclic Gray codes; two such examples are the binary reflected Gray code and the balanced Gray code [5].

Definition 4.1. *A bipartite singly-repairable system (BSR) is a pair $(\mathcal{U}, \mathcal{V})$ of disjoint subsets of $\mathbb{Z}_2^n$ such that $\mathcal{U} \cup \mathcal{V}$ is singly repairable with breaks in $\mathcal{U}$ repaired in $\mathcal{V}$, and vice-versa. That is, we require that $\mathcal{U} \cup \mathcal{V}$ is singly repairable and for all $u \in \mathcal{U}$,*

$$\partial_{ij}(u) \in \mathcal{U} \cup \mathcal{V} \to \partial_{ij}(u) \in \mathcal{V}$$

and for all $v \in \mathcal{V}$,

$$\partial_{ij}(v) \in \mathcal{U} \cup \mathcal{V} \to \partial_{ij}(v) \in \mathcal{U}$$

for all distinct $i, j \in \{0, 1, \ldots, n-1\}$.

A *minimal BSR* is a BSR $(\mathcal{U}, \mathcal{V})$ such that $\mathcal{U} \cup \mathcal{V}$ is minimal singly repairable. The size of a BSR $(\mathcal{U}, \mathcal{V})$ is $|\mathcal{U} \cup \mathcal{V}|$.

4.1 Construction when n is a power of 2

We will now provide an explicit construction for minimal singly-repairable subfamilies of $\mathbb{Z}_2^n$ of size $2^n/n$ (the largest possible size, via Proposition 3.1).

Fix $m > 0$, $n = 2^m$. Let $G(0), G(1), \ldots, G(n-1)$ be some fixed m-bit cyclic Gray code.

Notation. Define the function $f : \{0, 1, \ldots, n-1\} \to \{0, 1, \ldots, m-1\}$ such that $G(j+1) = G(j) + e_m(f(j))$ (we assume that the index j is taken modulo n so that $G(n) = G(0)$). Note that $f(j)$ specifies which bit in $G(j)$ has to be flipped to get $G(j+1)$, the next term in the Gray code.

Recall that $\mathcal{E}(n)$ refers to the set of all even weight $\{0,1\}$ vectors of length n.

Define the $m \times n$ matrix $\mathcal{A}$ with columns indexed $\{0, 1, \ldots, n-1\}$ and rows indexed by $\{0, 1, \ldots, m-1\}$ as follows. The j-th column of $\mathcal{A}$, where $0 \leq j \leq n-1$ is the m-bit binary representation of the integer j, with the least significant bit appearing in row 0.

For $0 \leq i \leq n-1$, we let

$$\mathcal{G}_i^{(n)} = \{v \in \mathcal{E}(n) \mid \mathcal{A}\, v^T = G(i)^T \bmod 2\}. \tag{1}$$

When n is obvious from the context, we simply write $\mathcal{G}_i$.

Remark. Interpret each $\mathcal{G}_i$ as follows. Equation (1) is equivalent to saying that $v \in \mathcal{G}_i$ iff $v \in \mathcal{E}(n)$ and

$$\sum_{0 \leq k \leq n-1} v_k k_j \equiv G(i)_j \pmod 2, \text{ for } 0 \leq j \leq m-1$$

For any $0 \leq j \leq m-1$, half of the k's in $0, 1, \ldots, n-1$ have $k_j = 1$. The sum determines the number of 1's in the corresponding positions of v, and we want the parity of this sum to be coordinated with the j-th bit of $G(i)$. Since the total number of 1's is even, one has the same parity for the number of 1's in the rest of the positions of v.

Lemma 4.1. *Any vector w of odd weight in $\mathbb{Z}_2^n$ is at Hamming distance 1 from a unique element of each $\mathcal{G}_r$, where $0 \le r \le n-1$.*

Proof. Let $l,\ 0 \le l \le n-1$ be such that

$$l_j = 1 \text{ iff } \sum_{0 \le k \le n-1} w_k k_j \not\equiv G(r)_j \pmod 2$$

for each $j,\ 0 \le j \le m-1$. Then $\partial_l(w) \in \mathcal{G}_r$ is the unique element at distance 1 from w. $\square$

We have $\mathcal{E}(n) = \bigcup_{i=0}^{n-1} \mathcal{G}_i$, a disjoint union and for each i, $|\mathcal{G}_i| = \frac{|\mathcal{E}(n)|}{n} = \frac{2^{n-1}}{n}$. Define

$$\mathcal{H}_j = \bigcup_{i=0}^{n-1} (\mathcal{G}_{i+j} \,|\, \mathcal{G}_i)$$

where we assume that the indices $j,\ 0 \le j \le n-1$, are taken modulo n (so that $\mathcal{H}_n = \mathcal{H}_0$). Note that $|\mathcal{H}_j| = \frac{2^{2n-2}}{n}$. We will now show that $\mathcal{H}_i \cup \mathcal{H}_{i+1}$ is a minimal singly-repairable family for all $i,\ 0 \le i \le n-1$.

Lemma 4.2. *For $1 \le i, j \le n-1$ with $i \ne j$, $(\mathcal{H}_i, \mathcal{H}_j)$ is a BSR.*

(Observe that $\mathcal{H}_i \cup \mathcal{H}_j$ will then be a singly-repairable subfamily of $\mathbb{Z}_2^{2n}$ of size $\frac{2^{2n}}{2n}$.)

Proof. Given a odd-weighted string $u \in \mathbb{Z}_2^{2n}$, we claim that u is at distance 1 from precisely two elements of $\mathcal{H}_i \cup \mathcal{H}_j$ (this implies a unique repair for every break). To see this, say $u = (v \mid w)$ with $v, w \in \mathbb{Z}_2^n$. Then exactly one of v, w is in $\mathcal{E}(n)$. Say, for example, $v \in \mathcal{E}(n)$. Then $v \in \mathcal{G}_r$ for some unique r. If u is at distance 1 from $u' \in \mathcal{H}_i$, then $u' = (v \mid w')$ with w' the unique element of $\mathcal{G}_{r-i}$ of distance 1 from w (Lemma 4.1). The analysis for $w \in \mathcal{E}(n)$ is similar. Similarly u is at distance 1 from a unique element of $\mathcal{H}_j$.

It follows then that $\mathcal{H}_i \cup \mathcal{H}_j$ is singly repairable and since breaks in $\mathcal{H}_j$ are repaired in $\mathcal{H}_i$ (and vice versa), $(\mathcal{H}_i, \mathcal{H}_j)$ is a BSR. $\square$

Remark. More generally, consider families

$$\mathcal{A} = \bigcup_{j=0}^{n-1} (\mathcal{G}_{j^\sigma} \mid \mathcal{G}_j), \qquad \mathcal{B} = \bigcup_{j=0}^{n-1} (\mathcal{G}_{j^\tau} \mid \mathcal{G}_j)$$

where $\sigma, \tau \in \mathrm{Sym}(n)$ with $j^\sigma \ne j^\tau$ for any j (where $\mathrm{Sym}(n)$ refers to the symmetric group acting on the elements of $\{0, 1, \ldots, n-1\}$). Then the same proof shows that $(\mathcal{A}, \mathcal{B})$ is a BSR.

Remark. We can also explicitly state the break-repairs pairs for each vector in $\mathcal{F} = \mathcal{H}_i \cup \mathcal{H}_{i+1}$. Say $(v \mid w) \in \mathcal{H}_i$ and then specifically $(v \mid w) \in (\mathcal{G}_{i+j} \mid \mathcal{G}_j)$ for some j. Given a break in the k-th bit of v, repair in the k'-th bit where k' is obtained from k by changing the bit $k_{f(i+j)}$; the repair is then in $(\mathcal{G}_{i+j+1} \mid \mathcal{G}_j) \subseteq \mathcal{H}_{i+1}$. Given a break in the k-th bit of w, repair in the k'-th bit where k' is obtained from k by changing the bit $k_{f(j-1)}$; the repair is then in $(\mathcal{G}_{i+j} \mid \mathcal{G}_{j-1}) \subseteq \mathcal{H}_{i+1}$. The repairs for breaks in $\mathcal{H}_{i+1}$ are viewed similarly.

We want to establish that $(\mathcal{H}_i, \mathcal{H}_{i+1})$ is a *minimal* BSR, for $0 \le i \le n-1$. We prove that $\mathcal{F}_{01} = \mathcal{H}_0 \cup \mathcal{H}_1$ is minimal (the general case can be proved using similar arguments). We first prove a limited form of connectivity in $\mathcal{F}_{01}$ in Lemmas 4.3 and 4.4.

Lemma 4.3. *Let $u_0, v_0 \in \mathcal{G}_0$ and let p, q, r, s $(0 \le p, q, r, s \le n-1)$ be such that p and q differ only in the k-th bit of their m-bit binary representation, and r and s also differ only in the k-th bit of their m-bit binary representation where $0 \le k \le m-1$. Then there is a path in $\mathcal{F}_{01}$ between $(u_0 \,|\, v_0) \in (\mathcal{G}_0 \,|\, \mathcal{G}_0)$ and $(u_0 \,|\, v_0') \in (\mathcal{G}_0 \,|\, \mathcal{G}_0)$ where v_0' is obtained from v_0 by flipping the bits in positions p, q, r, s (i.e., $v_0' = \partial_{\{p,q,r,s\}}(v_0)$).*

Proof. We may assume that $p \ne r$ (otherwise, $v_0 = v_0'$ and the start and end vertex are the same, so we have nothing to prove). We exhibit a path in $\mathcal{F}_{01}$ between $(u_0 \,|\, v_0)$ and $(u_0 \,|\, v_0')$ below:

$$
\begin{array}{cc}
 & (u_0 \,|\, v_0) \in (\mathcal{G}_0 \,|\, \mathcal{G}_0) \\
\to & (u_1 \,|\, v_0) \in (\mathcal{G}_1 \,|\, \mathcal{G}_0) \\
\to & (u_1 \,|\, v_1) \in (\mathcal{G}_1 \,|\, \mathcal{G}_1) \\
\to & (u_2 \,|\, v_1) \in (\mathcal{G}_2 \,|\, \mathcal{G}_1) \\
\vdots & \vdots \\
\to & (u_{n-1} \,|\, v_{n-1}) \in (\mathcal{G}_{n-1} \,|\, \mathcal{G}_{n-1}) \\
\to & (u_0 \,|\, v_{n-1}) \in (\mathcal{G}_0 \,|\, \mathcal{G}_{n-1}) \\
\to & (u_0 \,|\, v_0') \in (\mathcal{G}_0 \,|\, \mathcal{G}_0)
\end{array}
$$

We now define the intermediate vertices u_i, v_i. Fix a coordinate t, where $0 \le t \le n-1$. We set $u_i = \partial_{tt'}(u_{i-1})$ where t' is obtained from t by flipping the $f(i-1)$-th bit of t, so that $u_i \in \mathcal{G}_i$. That is, when we break on the left half, we always choose to do so at coordinate t. The repair, dictated by the next sequence in the Gray code, is made by flipping coordinate t'. For any i, $0 \le i \le m-1$, consider the changes of bit i. Since i gets changed an even number of times in the cycle $(\mathcal{G}_0 \mid \mathcal{G}_0) \to \cdots \to (\mathcal{G}_0 \mid \mathcal{G}_0)$, every coordinate is flipped an even number of times. Thus the final vector will be of the form $(u_0 \mid \cdot)$.

On the right side, we will almost do the same thing. We always break at the p-th position, *except* only during one transition. We pick one instance when successive terms in the Gray code are related by the k-th bit, where instead of breaking in the p-th bit (and being forced to repair in q-th bit), we break in the r-th bit (and hence are forced to repair by flipping the s-th bit). The choice of which transition to choose is arbitrary; for example, we choose it to be the last time that the k-th bit is involved. So let j be the largest index where $0 \leq j \leq n-1$ such that $f(j) = k$. Then we let

$$v_{j+1} = \partial_{rs}(v_j)$$

and for all i such that $i \neq j,\ 0 \leq i \leq n-1$, we let

$$v_{i+1} = \partial_{pp'}(v_i)$$

where p' is obtained from p by flipping the $f(i)$-th bit in the binary representation of p (so that $p' = q$ exactly when $f(i) = k$).

Then p, q, r, s will be flipped an odd number of times, every other coordinate being flipped an even number of times. Thus the final vector in the path is of the form $(u_0 \,|\, v_0')$ where v_0' is obtained from v_0 by flipping coordinates p, q, r, s. □

Lemma 4.4. *Any two vectors in $(\mathcal{G}_0 \,|\, \mathcal{G}_0)$ are in the same connected component in $\mathcal{F}_{01}$.*

Proof. Let $(u_0 \,|\, v_0)$ and $(u_1 \,|\, v_1)$ be any two distinct vertices in $(\mathcal{G}_0 \,|\, \mathcal{G}_0)$. We show that there is a path in $\mathcal{F}_{01}$ joining the two vertices.

We first show that one can go from $(u_0 \,|\, v_0)$ to $(u_0 \,|\, v_1)$ in $\mathcal{F}_{01}$ using a sequence of cycles in $\mathcal{F}_{01}$ each cycle changing v_0 in exactly 4 places (p, q, r, s) (where the integers in each pair (p, q) and (r, s) differ by the same bit, as required by Lemma 4.3). Without loss of generality, we take $v_1 = 0^n$. It is therefore enough to show that there is a partial ordering $\prec$ on $\mathcal{G}_0$ such that 0^n is the unique minimal element and if $0^n \neq w \in \mathcal{G}_0$, there is a choice of (p, q, r, s) as in Lemma 4.3 that moves w to an element $\prec w$. For this, let $w \in \mathcal{G}_0$ and define $\alpha(w) \in \mathbb{Z}^m$ by

$$\alpha(w)_j = |\{k \in \{0, 1, \ldots, n-1\} \mid w_k \cdot k_j = 1\}|, \quad \text{for } 0 \leq j \leq m-1$$

(by the definition of $\mathcal{G}_0$, $\alpha(w)_j$ is even); we say $v \prec w$ iff $\alpha(v)$ strictly precedes $\alpha(w)$ lexicographically. Let k be minimal such that $\alpha(w)_k \neq 0$. There exist (at least) two coordinates p, r such that $p_k = r_k = 1$ and $w_p = w_r = 1$. Form q and s from p and r respectively by changing the k-th bit to 0. Say p, q, r, s moves w to $\hat{w}$; then $\hat{w} \prec w$.

Repeating the same argument for the first half, (we use the version of Lemma 4.3 that changes bits on the first half instead of the second) we can then similarly show that there is a path from $(u_0 \,|\, v_1)$ to $(u_1 \,|\, v_1)$. □

Lemma 4.5. *For $0 \leq i \leq n-1$, $\mathcal{H}_i \cup \mathcal{H}_{i+1}$ is a minimal singly-repairable subfamily of $\mathbb{Z}_2^{2n}$.*

Proof. As before, for ease of notation in our proof, we consider the case $\mathcal{F}_{01} = \mathcal{H}_0 \cup \mathcal{H}_1$. The argument easily generalizes.

We need to show that $\mathcal{F}_{01}$ is connected. Since there is a path between any vertex in $\mathcal{F}_{01}$ to some vertex in $(\mathcal{G}_0 \,|\, \mathcal{G}_0)$, it clearly suffices to show that there is a break-repair path between any two elements of $(\mathcal{G}_0 \mid \mathcal{G}_0)$ (the intermediate vertices are not all within $(\mathcal{G}_0 \mid \mathcal{G}_0)$). Lemma 4.4 guarantees the existence of such a path. $\square$

We thus have an explicit construction for the following theorem.

Theorem 4.6. *When n is a power of 2, there exist minimal singly-repairable subfamilies of $\mathbb{Z}_2^n$ of size $2^n/n$ (the largest possible, via Proposition 3.1).*

Remark. Choosing non-isomorphic Gray codes (e.g., the binary reflected Gray code or the balanced Gray code[5]) for $G(i)$'s in this construction yields non-isomorphic minimal singly-repairable families of size $2^n/n$ when n is a power of 2. For small values of n, we have other examples which indicate that the Gray code construction is not unique. Since these largest minimal singly-repairable families are not unique up to translations and permutations, the question is whether one can somehow characterize these families.

4.2 Construction for general n

In this section, we show that for general even n, there exist minimal singly-repairable subfamilies of $\mathbb{Z}_2^n$ of size $\Omega(2^n/n^2)$.

Definition 4.2. *A BSR sequence for n is a sequence $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_{r-1})$ such that $\mathcal{A}_i \subseteq \mathbb{Z}_2^n$ and*

(i) $(\mathcal{A}_i, \mathcal{A}_j)$ is a BSR for all $i \neq j$, $0 \leq i, j \leq r-1$.

(ii) $(\mathcal{A}_i, \mathcal{A}_{i+1})$ is a minimal BSR for $0 \leq i \leq r-2$.

Remark: Observe that we do not require that $(\mathcal{A}_{r-1}, \mathcal{A}_0)$ is a minimal BSR and so trivially, we have that $(\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_{s-1})$ is also a BSR sequence for n for all $2 \leq s \leq r$.

Our goal is to construct a BSR sequence $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_{r-1})$ for n, when n is not necessarily a power of 2. Then $\mathcal{A}_0 \cup \mathcal{A}_1$ (or more generally, $\mathcal{A}_i \cup \mathcal{A}_{i+1}$) is the desired family. Note first that we already have a construction of BSR sequences when n is a power of 2.

Lemma 4.7. *Let n be a power of 2. There exists a BSR sequence*

$$(\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_{(n/2)-1})$$

for n with each $|\mathcal{H}_i| = 2^{n-1}/n$.

Proof. Immediate from Lemmas 4.2 and 4.5. $\square$

We show how to construct a BSR sequence for $n+m$ given a BSR sequence for n when m is a power of 2.

Lemma 4.8. *Let $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_{m-1})$ be a BSR sequence for n where m is a power of 2 and $m \geq 2$. Then $\mathcal{B} = (\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_{m-1})$ is a BSR sequence for $n+m$ where $\mathcal{B}_j = \bigcup_{i=0}^{m-1} (\mathcal{A}_{i+j} \,|\, \mathcal{G}_i)$ where $0 \leq j \leq m-1$ and $\mathcal{A}_m = \mathcal{A}_0$.*

Proof. To see that $(\mathcal{B}_i, \mathcal{B}_j)$ is a BSR, consider $x = (a \mid g) \in (\mathcal{A}_{i+k} \mid \mathcal{G}_k) \subseteq \mathcal{B}_i$. Since $(\mathcal{A}_{i+k}, \mathcal{A}_{j+k})$ is a BSR, a break in the a part of x is uniquely repaired in $(\mathcal{A}_{j+k} \mid \mathcal{G}_k) \subseteq \mathcal{B}_j$; since a break in $\mathcal{G}_k$ is uniquely repaired in $\mathcal{G}_{i+k-j}$ (cf. proof of Lemma 4.2), a break in the g part of x is uniquely repaired in $(\mathcal{A}_{i+k} \mid \mathcal{G}_{i+k-j}) \subseteq \mathcal{B}_j$.

We prove that $(\mathcal{B}_i, \mathcal{B}_{i+1})$ is minimal. For simplicity of notation, we consider the case when $i = 0$. Since there is an edge from any element of $\mathcal{A}_i$ to an element of $\mathcal{A}_{i\pm1}$ and from any element of $\mathcal{G}_i$ to an element of $\mathcal{G}_{i\pm1}$, there is a break/repair path from any element of the $\mathcal{B}_i \cup \mathcal{B}_{i+1}$ to some element of $(\mathcal{A}_0 \mid \mathcal{G}_0)$. So it suffices to show that there is a break/repair path from any $(a \mid g) \in (\mathcal{A}_0 \mid \mathcal{G}_0)$ to any $(a' \mid g') \in (\mathcal{A}_0 \mid \mathcal{G}_0)$ (the intermediate points are not confined to $(\mathcal{A}_0 \mid \mathcal{G}_0)$). Using the idea in the connectivity proof in Lemma 4.3, we know there is a path from $(a \mid g)$ to $(a'' \mid g')$ for *some* $a'' \in \mathcal{A}_0$ (it wanders around the $\mathcal{G}_i$ cycle to get there, using arbitrary break/repairs on the $\mathcal{A}_i$ side). But since $\mathcal{A}_0 \cup \mathcal{A}_1$ is connected, there is a break/repair path within $\mathcal{A}_0 \cup \mathcal{A}_1$ from a'' to a'. This path induces a break/repair path from $(a'' \mid g')$ to $(a' \mid g')$. $\square$

Remark: Observe that

$$|\mathcal{B}_j| = \frac{2^{m-1}}{m} \sum_{i=0}^{m-1} |\mathcal{A}_i| \text{ for } 0 \leq j \leq m-1 \tag{2}$$

Corollary 4.9. *There exists a minimal singly-repairable subfamily of $\mathbb{Z}_2^n$ of size at least $2^{n+1-r}/n$ where r is the number of 1's in the binary representation of n.*

Proof. Let $n = 2^{k_1} + 2^{k_2} + \ldots + 2^{k_r}$ and let $n_i = 2^{k_i}$ for $1 \leq i \leq r$ where $k_1 > k_2 \cdots > k_r$. Starting with a BSR sequence $\mathcal{B}^1 = (\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_{n_1/2-1})$ for n_1 from Lemma 4.7, we construct $r-1$ sequences $\mathcal{B}^j$ for $2 \leq j \leq r$, where

$$\mathcal{B}^j = (\mathcal{B}_0^j, \mathcal{B}_1^j, \mathcal{B}_2^j, \ldots, \mathcal{B}_{n_j-1}^j)$$

and $\mathcal{B}^j$ is a BSR sequence for $n_1 + n_2 + n_3 + \cdots + n_j$, . Each $\mathcal{B}^j$ is constructed by choosing an appropriate prefix of $\mathcal{B}^{j-1}$ of length n_j (such a prefix exists since $n_j \leq n_{j-1}$ for $j \geq 3$ and $n_1/2 \geq n_2$), which is itself a BSR sequence and then applying Lemma 4.8 (with $m = n_j$).

Equation (2) implies that

$$|\mathcal{B}_i^r| = \frac{2^{n_1+n_2+\ldots+n_r-r}}{n_1} \text{ where } 0 \leq i \leq n_r - 1$$

This provides a minimal BSR $(\mathcal{B}_0^r, \mathcal{B}_1^r)$ of size $|\mathcal{B}_0^r \cup \mathcal{B}_1^r| = 2^n/2^{k_1+r-1}$. $\square$

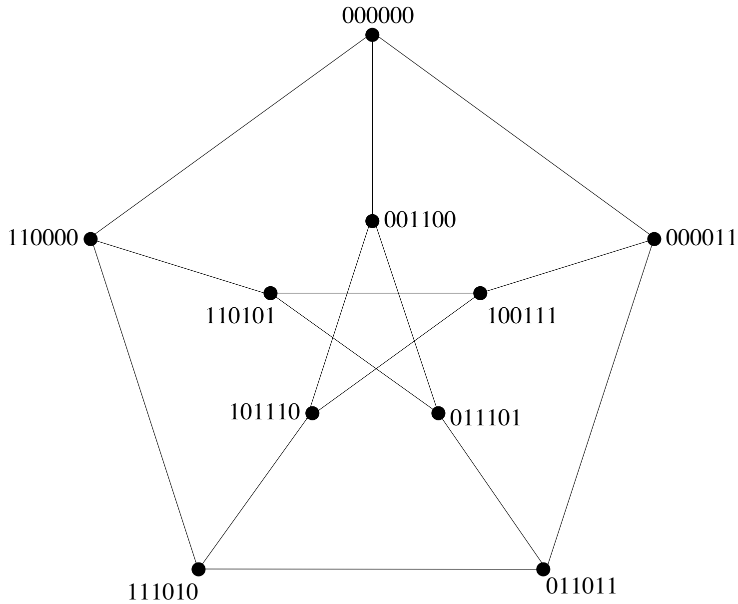

Figure 1: Graph for Singly-Repairable Subfamily of $\mathbb{Z}_2^6$ of size 10

Since $r = O(\log n)$, Corollary 4.9 guarantees families of size $\Omega(2^n/n^2)$.

Remark: An explicit formula for $\mathcal{B}_i^r$ is

$$\mathcal{B}_i^r = \bigcup_{i_1 \in \mathbb{Z}_{n_1/2}} \bigcup_{i_j \in \mathbb{Z}_{n_j}} (\mathcal{G}_{i_1}^{n_1/2} \mid \mathcal{G}_{i_1-i_2}^{n_1/2} \mid \mathcal{G}_{i_2-i_3}^{n_2} \mid \ldots \mid \mathcal{G}_{i_r-i}^{n_r})$$

where $i \in \mathbb{Z}_{n_r}$ (where $\mathbb{Z}_m$ represents the set of integers mod m).

Remark. The construction of a minimal singly-repairable subfamily for general n, in Corollary 4.9, already falls short of the largest possible size for $n = 6$. While our construction gives a family of size 8, there is a minimal singly-repairable family of size 10. Moreover, one can prove that, up to permutation of coordinates and translations, this family of size 10 is unique. Furthermore, the break-repair graph (Figure 1) of this family corresponds to the well-known Petersen graph.

5 Remaining Gaps

While it was possible to attain the upper bound of $2^n/n$ (Proposition 3.1) for minimal singly-repairable subfamilies of $\mathbb{Z}_2^n$ via an explicit construction (Corollary 4.6) when n is a power of 2, the best construction for arbitrary n, of size $2^n/n^2$ (Corollary 4.9) falls short of this upper bound when $n = 6$ (the size 10 singly-repairable family ($\lfloor 2^6/6 \rfloor = 10$) from Figure 1 already achieves the largest possible size). So this indicates the possibility that one can do better for arbitrary even n. A similar gap exists for non-minimal families.

While we could construct non-minimal singly-repairable subfamilies of $\mathbb{Z}_2^n$ of the largest possible size $2^{n+1}/(n+2)$ when $n+2$ is a power of 2, our method also falls short of the maximum possible size for general even n.

References

[1] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein. *Covering Codes*. Elsevier Science, 1987.

[2] M. Ginsberg, A. Parkes, and A. Roy. Supermodels and robustness. In *Proceedings of the Fifteenth National Conference of the American Association for Artificial Intelligence and the tenth conference on Innovative Applications of Artificial Intelligence, 1998, Madison, WI*, pages 334–339, 1998.

[3] F.J. MacWilliams and N.J.A. Sloane. *The Theory of Error-Correcting Codes*. North Holland, 1977.

[4] A. Roy. Fault tolerant boolean satisfiability. submitted, 2004.

[5] C. Savage. A survey of combinatorial Gray codes. *SIAM Review*, 39(4):605–629, 1997.