Generalized Sphere Packing Bound

Arman Fazeli,* Alexander Vardy,* and Eitan Yaakobi†

*University of California San Diego, La Jolla, CA 92093, USA
†California Institute of Technology, Pasadena, CA 91125, USA

{afazelic, avardy}@ucsd.edu, yaakobi@caltech.edu

Abstract—Kulkarni and Kiyavash recently introduced a new method to establish upper bounds on the size of deletion-correcting codes. This method is based upon tools from hypergraph theory. The deletion channel is represented by a hypergraph whose edges are the deletion balls (or spheres), so that a deletion-correcting code becomes a matching in this hypergraph. Consequently, a bound on the size of such a code can be obtained from bounds on the matching number of a hypergraph. Classical results in hypergraph theory are then invoked to compute an upper bound on the matching number as a solution to a linear-programming problem: the problem of finding fractional transversals.

The method by Kulkarni and Kiyavash can be applied not only for the deletion channel but also for other error channels. This paper studies this method in its most general setup. First, it is shown that if the error channel is regular and symmetric then the upper bound by this method coincides with the well-known sphere packing bound and thus is called here the generalized sphere packing bound. Even though this bound is explicitly given by a linear programming problem, finding its exact value may still be a challenging task. The art of finding the exact upper bound (or slightly weaker ones) is the assignment of weights to the hypergraph’s vertices in a way that they satisfy the constraints in the linear programming problem. In order to simplify the complexity of the linear programming, we present a technique based upon graph automorphisms that in many cases significantly reduces the number of variables and constraints in the problem. We then apply this method on specific examples of error channels. We start with the $Z$ channel and show how to exactly find the generalized sphere packing bound for this setup. Next studied is the non-binary limited magnitude channel both for symmetric and asymmetric errors, where we focus on the single-error case. We follow up on the deletion channel, which was the original motivation of the work by Kulkarni and Kiyavash, and show how to improve upon their upper bounds for single-deletion-correcting codes. Since the deletion and grain-error channels resemble a very similar structure for a single error, we also improve upon the existing upper bounds on single-grain error-correcting codes. Finally, we apply this method for projective spaces and find its generalized sphere packing bound for the single-error case.

I. INTRODUCTION

One of the basic and fundamental results in coding theory asserts that an upper bound on a length-$n$ binary code $C$ with minimum Hamming distance $2r+1$ is

$$|C| \leq \frac{2^n}{B(r)},$$

where $B(r) = \sum_{i=0}^{r} \binom{n}{i}$. This is known as the classical sphere packing bound. This bound can be applied for other cases as well. Let $X$ be a finite set with some distance function $d : X \times X \to \mathbb{N}$. Assume that the volume of every ball is the same, that is, if $B_r(x) = \{ y \in X \mid d(x, y) \leq r \}$ then for all $x \in X$, $|B_r(x)| = \Delta_r$ for some fixed value $\Delta_r$. Then, the resulting sphere packing bound on a code $C \subseteq X$ with minimum distance $2r+1$ becomes $|X|/\Delta_r$. However, what happens if the size of all balls is not the same? Clearly, a naive solution is to use $\Delta_r$ as the minimum size of all balls and then to apply the same bound, but this approach can give a very weak upper bound. The goal of this paper is to study a generalization of the sphere packing bound for setups where the size of all balls is not necessarily the same.

The lower counter bound for the sphere packing one is the well-known Gilbert-Varshamov bound [10], [22]. This bound states that if the size of all balls of radius $r$ is the same, $\Delta_r$, then a lower bound on a code $C \subseteq X$ with minimum distance $r+1$ becomes $|X|/\Delta_r$. In [21], a similar study was carried for the Gilbert-Varshamov bound in case that the size of all balls is not necessarily the same. Using Turán’s theorem, it was shown that the size derivation on a lower bound of a code still holds, with the modification of using the average size of the balls. That is, if $\Delta_r \triangleq (\sum_{x \in X} |B_r(x)|)/|X|$, then a generalized Gilbert-Varshamov bound asserts that there exists a code with minimum distance $r+1$ and of size at least $|X|/\Delta_r$. Thus, an immediate question to ask is whether the same analogy holds for the sphere packing bound: Is $|X|/\Delta_r$ an upper bound on a code $C \subseteq X$ with minimum distance $2r+1$? Even though in most of the cases we study in this work this derivation does hold, the answer in general to this question is negative. However, it is interesting to find some conditions under which this bound will always be satisfied.

The deletion channel [18] is one of the examples where the balls can have different sizes. Recently, in [15], Kulkarni and Kiyavash showed a technique, based upon tools from hypergraph theory [2], in order to derive explicit non-asymptotic upper bounds on the cardinalities of deletion-correcting codes. These upper bounds were given both for binary and non-binary codes as well as for deletion-correcting codes for constrained sources. Since the method in [15] can be applied for other similar setups, more results were presented shortly after for different channel models. Upper bounds on the cardinalities of grain-error-correcting codes were given in [8] and [12] and similar bounds for multipermutations codes with the Kendall’s $\tau$ distance were derived in [4].

This paper has two main goals. First, we extend the method studied for the deletion channel by Kulkarni and Kiyavash [15] and analyze it in its most general setting. We assume that the error channel is characterized by a directed graph, which depicts for a given transmitted word, its set of possible received words. Then, an upper bound will be given on codes which can correct $r$ errors, for some fixed $r$. This bound is established by the solution of a linear programming given from a hypergraph that is derived from the error channel graph. In particular, it is shown that the sphere packing bound is a special case of this bound. We also study properties of this bound and show a scheme, based upon graph automorphisms, that in
many cases can significantly reduce the complexity of the linear programming problem. In the second part of this work, we provide specific examples on the application of this method to setups where the balls have different sizes. These examples include the Z channel, non-binary channels with limited magnitude errors (symmetric and asymmetric), deletion channel, grain-error channel, and finally, projective spaces. In some of these examples we improve upon the existing results which use this method to calculate the upper bound on the code cardinalities. When possible in these examples, we compare the bounds we receive with the state-of-the-art ones.

In order to describe our results, we need to introduce some notation. Let \( H = (X, E) \) be a hypergraph, where \( X = \{x_1, \ldots, x_n\} \) is its vertices set and \( E = \{E_1, \ldots, E_m\} \) is its hyperedges set. Let \( A \) be the \( n \times m \) incidence matrix of \( H \), so \( A(i, j) = 1 \) if \( x_i \in E_j \). A transversal in \( H \) is a subset \( T \subseteq X \) that intersects every hyperedge in \( E \). The transversal number of \( H \), denoted by \( \tau(H) \), is the size of the smallest transversal. Every transversal can be represented by a binary vector \( w \in \{0, 1\}^n \) which needs to satisfy \( A^T \cdot w \geq 1 \). However, if the vector \( w \) can have values over \( \mathbb{R}_+ \) and still satisfies the last inequality, then it is called a fractional transversal. Under this setup, it is known that \( \tau^*(H) \leq \tau(H) \), where \( \tau^*(H) \) is the linear programming relaxation of \( \tau(H) \), defined as

\[
\tau^*(H) = \min \left\{ \sum_{i=1}^n w_i : A^T \cdot w \geq 1, w \in \mathbb{R}_+^n \right\}.
\]

Let \( G = (X, E) \) be a directed graph which describes an error channel. The vertices set \( X \) is the set of all possible transmitted words, and the edges set \( E \) consists of all pairs of vertices of distance one. The distance between \( x, y \in X \) is the path metric in \( G \) and is denoted by \( d(x, y) \). Note that since the graph is directed, it is possible to have \( d(x, y) \neq d(y, x) \). For every \( x \in X \), its radius-\( r \) ball is the set \( B_r(x) \) which was defined above and its degree is \( \deg_r(x) = |B_r(x)| \). The largest cardinality of a length-\( n \) code in \( G \) with minimum distance \( d \) is denoted by \( A_G(n, d) \). Given some positive integer \( r \), the graph \( G \) is associated with a hypergraph \( \mathcal{H}(G, r) = (X_r, E_r) \) where \( X_r = X \) and \( E_r = \{B_r(x) \mid x \in X\} \). Observing that every code \( C \subseteq X \) of minimum distance \( 2r+1 \) is a matching in \( \mathcal{H}(G, r) \) (which is a collection of pairwise disjoint edges), the following upper bound on \( A_G(n, 2r+1) \) was verified in \[15\]:

\[
A_G(n, 2r+1) \leq \tau^*(\mathcal{H}(G, r)). \tag{2}
\]

One of the first properties we present asserts that if the graph \( G \) is regular such that \( \deg_r(x) = \Delta_r \) for all \( x \in X \), and the distance function \( d \) is symmetric, then the bound \( \tau^*(\mathcal{H}(G, r)) \) coincides with the sphere packing bound, that is, \( \tau^*(\mathcal{H}(G, r)) = \frac{n}{\Delta_r} \). Therefore, in this work the bound \( \tau^*(\mathcal{H}(G, r)) \) is called the generalized sphere packing bound.

The expression \( \tau^*(\mathcal{H}(G, r)) \) provides an explicit upper bound on \( A_G(n, 2r+1) \). However, it may still be a hard problem to calculate this value since it requires the solution of a linear programming problem that can have an exponential number of variables and constraints. Clearly, one would inspire to find this exact value, but if this is not possible to accomplish, it is still valuable to give an upper bound on \( \tau^*(\mathcal{H}(G, r)) \), which, in essence, is an upper bound on \( A_G(n, 2r+1) \) as well. Such an upper bound will be given by finding any fractional transversal and the goal will be to find one with small weight. In fact, all the upper bound results presented in \[4, 8, 12, 15\] follow this approach and an upper bound on the value \( \tau^*(\mathcal{H}(G, r)) \) in each case is given.

The rest of the paper is organized as follows. Section \[II\] establishes the rest of the definitions and tools required in this paper and demonstrates them on the \( Z \) channel. This channel will be used throughout the paper as a running example and a case study we rigorously investigate. In Section \[III\] we start with basic properties on the generalized sphere packing bound. In particular, we show upper and lower bounds on its value and prove that if the graph \( G \) is regular and symmetric then the sphere packing bound coincides with the generalized sphere packing bound. We also show several examples which establish a dissenting answer to the question brought earlier about the upper bound validity of an average sphere packing value. We then proceed to define a special monotonicity property on the graph \( G \) which states that a graph is monotone if for all \( r \) and two vertices \( x \) and \( y \), if \( y \in B_r(x) \) then \( \deg_r(y) \leq \deg_r(x) \). This property is useful in order to give a general formula for a fractional transversal and a corresponding upper bound. In fact, this property and fractional transversal were used in the previous works \[8, 12, 15\]. Lastly in this section, we use tools from automorphisms on graphs in order to simplify the complexity of the linear programming problem in \[1\]. Noticing that in many channels there are groups of vertices with similar behavior motivates us to treat them as the same vertex and thus significantly reduce the number of variables and constraints in the linear programming \[1\]. In Section \[IV\] we study the \( Z \) channel. Our main contribution here is finding a method to calculate the generalized sphere packing bound for all radii. In Section \[V\] we carry a similar task for the limited-magnitude channel with symmetric and asymmetric errors. We focus only the single error case of radius one in both cases and find fractional transversals and corresponding upper bounds. Section \[VI\] follows upon the original work of \[15\], improving the bounds derived therein for the deletion channel (for the case of a single deletion). Since the structure of the deletion and grain-error channel is very similar, especially for a single error, we continue with the same approach to improve upon the existing upper bounds from \[8, 12\] on the cardinalities of single-grain error-correcting codes. Section \[VII\] studies bounds on projective spaces and in particular we give an optimal solution for the radius-one case under this channel. Finally, Section \[VIII\] concludes the paper and proposes some problems which remained open.

II. Definitions and Preliminaries

In this section we formally define the tools and definitions used throughout the paper. We mainly follow the same definitions and properties from \[15\].

Let \( H = (X, E) \) be a hypergraph where \( X = \{x_1, \ldots, x_n\} \), \( E = \{E_1, \ldots, E_m\} \) and \( A \) its \( n \times m \) incidence matrix. A matching in \( H \) is a collection of pairwise disjoint hyperedges and the matching number of \( H \), denoted by \( \nu(H) \),
is the size of the largest matching. The matching number of $\mathcal{H}$, $\nu(\mathcal{H})$, is the solution of the integer linear programming problem

$$
\nu(\mathcal{H}) = \max \left\{ \sum_{i=1}^{m} z_i : A \cdot z \leq 1, z \in \{0,1\}^m \right\}.
$$

Note that the transversal number $\tau(\mathcal{H})$, defined in the previous section, is the solution of the integer linear programming problem

$$
\tau(\mathcal{H}) = \min \left\{ \sum_{i=1}^{n} w_i : A^T \cdot w \geq 1, w \in \{0,1\}^n \right\}.
$$

These two problems satisfy weak duality and thus $\nu(\mathcal{H}) \leq \tau(\mathcal{H})$. Furthermore, they can be slightly modified such that the vectors in the minimization and maximization problems can have values in $\mathbb{Z}_+$, and still they give the values of $\nu(\mathcal{H})$ and $\tau(\mathcal{H})$, that is,

$$
\nu(\mathcal{H}) = \max \left\{ \sum_{i=1}^{m} z_i : A \cdot z \leq 1, z \in \mathbb{Z}_+^m \right\},
$$

$$
\tau(\mathcal{H}) = \min \left\{ \sum_{i=1}^{n} w_i : A^T \cdot w \geq 1, w \in \mathbb{Z}_+^n \right\}.
$$

The relaxation of these integer linear programs allows the variables $z$ and $w$ to take values in $\mathbb{R}_+$, which are not necessarily integers. The value of this linear program relaxation for the matching number is denoted by

$$
\nu^*(\mathcal{H}) = \max \left\{ \sum_{i=1}^{m} z_i : A \cdot z \leq 1, z \in \mathbb{R}_+^m \right\},
$$

and the corresponding one for the transversal number is the value $\tau^*(\mathcal{H})$, stated in [1]. Note that the real solutions can be significantly different than the integer solutions and since $\nu^*(\mathcal{H})$ and $\tau^*(\mathcal{H})$ satisfy strong duality, the following property holds [15]

$$
\nu(\mathcal{H}) \leq \nu^*(\mathcal{H}) = \tau^*(\mathcal{H}) \leq \tau(\mathcal{H}),
$$

and in particular, for any fractional transversal $w$,

$$
\nu(\mathcal{H}) \leq \tau^*(\mathcal{H}) \leq \sum_{i=1}^{n} w_i.
$$

Lastly, we mention here that we will usually denote the fractional transversal by $w = (w_1, \ldots, w_n)$, such that $w_i$ corresponds to the value that is assigned to the vertex $x_i$. However, when it will be clear from the context, the notation $w_x$ will be used to refer to the value of $w_i$, where $x = x_i$.

Every error channel studied in this work will be depicted by some directed graph $\mathcal{G} = (X, E)$, where the set $E$ defines the set of all pairs of vertices of distance one from each other. The distance between every two vertices $x, y \in X$, denoted by $d(x, y)$, is the length of the shortest path from $x$ to $y$ in the graph $\mathcal{G}$, and $d(x, y) = \infty$ if such a path does not exist. Note that this definition of distance is not necessarily symmetric and thus it may happen that $d(x, y) \neq d(y, x)$. However if for all $x, y \in X$, $d(x, y) = d(y, x)$, then we say that $\mathcal{G}$ is symmetric, and otherwise it is not symmetric. For any $x \in X$, we let $B^\text{out}_r(x), B^\text{in}_r(x)$ be the sets $B^\text{out}_r(x) = \{ y \in X \mid d(x, y) \leq r \}$ and $B^\text{in}_r(x) = \{ y \in X \mid d(y, x) \leq r \}$. The out-degree of $x$ is $\deg^\text{out}_r(x) = |B^\text{out}_r(x)|$ and the in-degree is $\deg^\text{in}_r(x) = |B^\text{in}_r(x)|$. The definition of $B^\text{out}_r(x)$ and $B^\text{in}_r(x)$ coincide with the ones in the Introduction for $B_r(x)$ and $\deg_r(x)$, respectively. To ease the notation in the paper we will follow the ones from the Introduction for the “out” case and use the ones defined above for the “in” case.

If a word $x \in X$ is transmitted and at most $r$ errors occurred then any word in $B_r(x)$ can be received. A code $C \subseteq X$ in this graph is said to have minimum distance $d$ if for all $x, y \in C$, $d(x, y) \geq d$. We let $A_G(n, d)$ be the largest cardinality of a code in $\mathcal{G}$ of length $n$ and minimum distance $d$. If for every $r \geq 0$, there exists some fixed $\Delta_r$ such that for every $x \in X$, $\deg_r(x) = \Delta_r$, then we say that the graph $\mathcal{G}$ is regular and otherwise it is called non-regular.

For any positive integer $r$, $\mathcal{H}(\mathcal{G}, r) = (X_r, E_r)$ is a hypergraph associated with $\mathcal{G}$ such that $X_r = X$ and $E_r = \{ B_r(x) : x \in X \}$. As was stated in (2), the value $\tau^*(\mathcal{H}(\mathcal{G}, r))$ is an upper bound on $A_G(n, 2r + 1)$ and is called in this work the generalized sphere packing bound.

The average size of a ball of radius $r$ in $\mathcal{G}$ is defined to be

$$
\Delta_r = \frac{1}{|X|} \sum_{x \in X} \deg_r(x).
$$

In [21], using Turán’s theorem a generalized Gilbert-Varshamov bound was shown to hold also for the cases where the size of all balls is not the same. This bound asserts that a lower bound on $A_G(n, d)$ is given by

$$
\frac{|X|}{\Delta_{d-1}} \leq A_G(n, d).
$$

Let us remind the question we brought in the Introduction about the analogy of the last bound to the sphere packing bound. Namely, does the following inequality hold

$$
A_G(n, 2r + 1) \leq \frac{|X|}{\Delta_r}?
$$

We call the value $\frac{|X|}{\Delta_r}$ the average sphere packing value and denote it by $\text{ASPV}(\mathcal{G}, r)$. We do not call this value a bound since, as we shall see later, it is not necessarily a valid upper bound.

The following example demonstrates the definitions and concepts introduced in this section for the $Z$ channel.

**Example 1.** The $Z$ channel is a channel with binary inputs and outputs where the errors are asymmetric. Here, we assume that errors can only change a 1 to 0 with some probability $0 < p < 1$, but not vice versa; see Fig [1]. The corresponding graph is $\mathcal{G}_Z = (X_Z, E_Z)$, where $X_Z = \{0,1\}^n$ and $E_Z = \{(x, y) : x, y \in \{0,1\}^n, x \geq y, w_H(x) = w_H(y) + 1\}$, and $w_H(x)$ denotes the Hamming weight of $x$. Let $r$ be some fixed positive integer. For every $x \in \{0,1\}^n$,

$$
B_{r,Z}(x) = \{ y \in \{0,1\}^n : x \geq y, w_H(x) - w_H(y) \leq r \},
$$

and $\deg_{Z,r}(x) = \sum_{i=0}^{(a_H(x))}$.
The corresponding hypergraph is \( \mathcal{H}(G_Z, r) = (X_Z, r, E_Z) \), such that \( X_Z = \{0, 1\}^n \) and \( E_Z = \{B_Z(x) : x \in \{0, 1\}^n\} \). The generalized sphere packing bound becomes

\[
\tau^*(\mathcal{H}(G_Z, r)) = \min \left\{ \sum_{x \in \{0, 1\}^n} w_x : \forall x \in \{0, 1\}^n, \sum_{y \in B_Z(x)} w_y \geq 1, w_x \geq 0 \right\}.
\]

The average size of a ball with radius \( r \) is

\[
\bar{\Delta}_{Z, r} = \frac{1}{2^n} \sum_{x \in \{0, 1\}^n} \sum_{i=0}^{r} \binom{n}{i} w_H(x)_i \geq \frac{1}{2^n} \sum_{w=0}^{n} \binom{n}{w} \sum_{i=0}^{r} \binom{n}{i} w_i.
\]

For \( 0 \leq i \leq r \), \( \sum_{w=0}^{n} \binom{n}{i} w_i = \binom{n}{i} 2^{n-i} \) and thus we get

\[
\bar{\Delta}_{Z, r} = \frac{1}{2^n} \sum_{i=0}^{r} \binom{n}{i} 2^{n-i} = \frac{1}{2^n} \sum_{i=0}^{r} \binom{n}{i} 2^i.
\]

Therefore, the average sphere packing value in this case becomes

\[
ASPV(G_Z, r) = \frac{2^n}{\bar{\Delta}_{Z, r}} = \frac{2^n}{\sum_{i=0}^{r} \binom{n}{i} 2^i}.
\]

In particular, for \( r = 1 \) we get

\[
ASPV(G_Z, 1) = \frac{2^n}{\bar{\Delta}_{Z, 1}} = \frac{2^n}{1 + n/2} = \frac{2^{n+1}}{n + 2}.
\]

In the sequel it will be verified that the average sphere packing value for \( r = 1 \) is a valid upper bound for the Z channel. \( \square \)

Even though the generalized sphere packing bound \( \tau^*(\mathcal{H}(G, r)) \) gives an explicit upper bound on the cardinality of error-correcting codes, it is not necessarily immediate to calculate it. To accomplish this task, one needs to solve a linear programming which, in general, does not necessarily have an efficient solution. Furthermore, note that in many of the communication channels the number of variables and constraints can be very large and in particular exponential with the length of the words. Our main discussion in this paper will be dedicated towards approaches for deriving the value \( \tau^*(\mathcal{H}(G, r)) \) for different graphs \( G \). However, in cases where it will not be possible to derive this explicit value, we note that every fractional transversal provides a valid upper bound and thus we inspire to give the best fractional transversal we can find.

### III. General Results and Observations

In this section we start by proving basic properties on the value of the generalized sphere packing bound \( \tau^*(\mathcal{H}(G, r)) \) as specified in (1). We then show some approaches for finding fractional transversals. Finally, we present a scheme, based upon automorphisms on graphs, that in many cases can significantly reduce the complexity of the linear programming problem for calculating the value \( \tau^*(\mathcal{H}(G, r)) \). As specified in Section [II], we assume throughout this section that the error channel is depicted by some directed graph \( G = (X, E) \) and for a fixed integer \( r \geq 1 \), \( \mathcal{H}(G, r) = (X_r, E_r) \) is its associated hypergraph.

#### A. Basic Properties of the Generalized Sphere Packing Bound

We start here by proving some basic properties and giving insights on the value of \( \tau^*(\mathcal{H}(G, r)) \). The next lemma proves a lower bound on the generalized sphere packing bound in case that its in-degree is upper bounded.

**Lemma 1.** If for all \( x \in X \), \( \deg_r(x) \leq \Delta \), then

\[
\tau^*(\mathcal{H}(G, r)) \geq \frac{|X|}{\Delta}.
\]

**Proof:** Since \( \deg_r(x) \leq \Delta \), for all \( x \in X \), the weight of every column of the incidence matrix \( A \) of \( \mathcal{H}(G, r) \) is at most \( \Delta \), that is, \( \sum_{i=1}^{n} a_{ij} \leq \Delta \) for all \( 1 \leq j \leq n \). Let \( \vec{w} \) be a fractional transversal in \( \mathcal{H}(G, r) \). Then, for every \( 1 \leq i \leq n \), \( \sum_{j=1}^{n} a_{ij} \vec{w}_j \geq 1 \), and thus

\[
n \leq \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \vec{w}_j.
\]

However, note that

\[
n \leq \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \vec{w}_j = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} \vec{w}_j = \sum_{j=1}^{n} \vec{w}_j \sum_{i=1}^{n} a_{ij} \leq \Delta \sum_{j=1}^{n} \vec{w}_j,
\]

and therefore

\[
\sum_{j=1}^{n} \vec{w}_j \geq \frac{n}{\Delta}.
\]

Hence, we conclude that \( \tau^*(\mathcal{H}(G, r)) \geq \frac{|X|}{\Delta} \). \( \square \)

Next, we show an upper bound on the generalized sphere packing bound in case that its out-degree is lower bounded.

**Lemma 2.** If for all \( x \in X \), \( \deg_r(x) \geq \Delta \), then

\[
\tau^*(\mathcal{H}(G, r)) \leq \frac{|X|}{\Delta}.
\]

**Proof:** If \( \deg_r(x) \geq \Delta \) for all \( x \in X \) then the vector \( \vec{w} = 1/\Delta \) is a fractional transversal and thus \( \tau^*(\mathcal{H}(G, r)) \leq |X|/\Delta \). \( \square \)

According to the last two lemmas we can show that if the graph \( G \) is regular and symmetric then the generalized sphere packing bound coincides with the sphere packing bound.

**Corollary 3.** If the graph \( G \) is symmetric and regular then the generalized sphere packing bound and the sphere packing
bound coincide. Furthermore, \( \tau^*(\mathcal{H}(G, r)) = \frac{|X|}{\Delta_r} \), where for all \( x \in X \), \( \deg_r(x) = \deg^m_r(x) = \Delta_r \).

**Proof:** Since \( G \) is regular then for all \( x \in X \), \( \deg_r(x) = \Delta_r \) and according to Lemma 2 we have
\[
\tau^*(\mathcal{H}(G, r)) \leq \frac{|X|}{\Delta_r}.
\]
Since \( G \) is also symmetric we have that for all \( x \in X \), \( \deg^m_r(x) = \Delta_r \) and according to Lemma 1 we get
\[
\tau^*(\mathcal{H}(G, r)) \geq \frac{|X|}{\Delta_r}.
\]
Therefore, \( \tau^*(\mathcal{H}(G, r)) = \frac{|X|}{\Delta_r}. \)

The next example proves that the requirement on the graph \( G \) to be symmetric is necessary in order to have equality between the sphere packing and the generalized sphere packing bound.

**Example 2.** In this example the graph \( G_2 = (X_2, E_2) \) has six vertices, so \( X_2 = \{x_1, x_2, x_3, x_4, x_5, x_6\} \). For \( 2 \leq i \leq 6 \), there is an edge from \( x_i \) to \( x_1 \) and finally there is an edge from \( x_1 \) to \( x_2 \); see Fig. 2. Therefore, \( b_1(x_i) = 2 \) for all \( 1 \leq i \leq 6 \), so the graph \( G_2 \) is regular and the sphere packing bound becomes \( \frac{|X_2|}{2} = 3 \). However, the vector \( \mathbf{w} = (1, 0, 0, 0, 0, 0) \) is a fractional transversal, which is optimal, and thus the generalized sphere packing bound of \( G_2 \) equals 1.

In the next example, we show a graph that does not obey to the average sphere packing value. This provides a negative answer to the earlier question we asked in the Introduction regarding the validity of the average sphere packing value as a valid bound.

**Example 3.** The graph \( G_3 = (X_3, E_3) \) in this example has five vertices, so \( X_3 = \{x_1, x_2, x_3, x_4, x_5\} \). There is an edge from the first vertex to all other four vertices; see Fig. 3. The average size of a ball is \( \frac{15+4+1}{5} = 9/5 \) and thus the average sphere packing value becomes \( \frac{5}{9/5} = 25/9 \). However, the minimum distance of the code \( C = \{x_2, x_3, x_4, x_5\} \) in \( G_3 \) is \( \infty \), and in particular, it can be a code with minimum distance 3, which contradicts the average sphere packing value.

**Example 4.** Assume there are \( n = k^2 \) vertices partitioned into two groups: the first one consists of \( k \) vertices and the other group of the remaining \( n - k \) vertices. Every vertex from the first group is connected (symmetrically) to a set of exactly \( \frac{n-k}{k} = k - 1 \) vertices from the second group such that there is no overlap between these \( k \) sets. The \( n - k \) vertices in the second group are all connected to each other. Thus, the average radius-one ball size is
\[
\bar{\Delta}_1 = \frac{k \cdot k + (n - k)(n - k + 1)}{n} = n - 2\sqrt{n} + 3 - \frac{1}{\sqrt{n}} > n/2.
\]

Therefore, the average sphere packing value is less than 2. However, it is possible to construct a single-error correcting code with the \( k \) vertices of the first group.

**B. Monotonicity and Fractional Transversals**

Remember that a vector \( \mathbf{w} \) is a fractional transversal if \( \mathbf{w} \geq 0 \) and for \( 1 \leq i \leq n \),
\[
\sum_{y \in B_r(x_i)} w_y \geq 1.
\]

A first example for choosing a fractional transversal is stated in the next lemma.

**Lemma 4.** The vector \( \mathbf{w} \) given by
\[
w_i = \frac{1}{\min_{x \in B_r^p(x_i)} \{\deg_r(x)\}},
\]
for \( 1 \leq i \leq n \), is a fractional transversal.

**Proof:** It is easy to verify that \( \mathbf{w} \geq 0 \). For every \( 1 \leq i \leq n \), if \( y \in B_r(x_i) \), then \( x_i \in B_r^p(y) \) and thus
\[
w_y = \frac{1}{\min_{x \in B_r^p(y)} \{\deg_r(x)\}} \geq \frac{1}{\deg_r(x_i)}.
\]
Therefore, we get
\[
\sum_{y \in B_r(x_i)} w_y \geq \sum_{y \in B_r(x_i)} \frac{1}{\deg_r(x_i)} = 1.
\]
A graph $G$ is said to satisfy the monotonicity property, or $G$ is monotone, if for every $r \geq 1$, $x \in X$ and $y \in B_r(x)$, \[
\deg_G(y) \leq \deg_G(x).
\]
In this case, the fractional transversal from Lemma 4 can be stated more explicitly.

**Lemma 5.** If $G$ is monotone then the vector $w$ given by \[
\forall i \leq n, w_i = \frac{1}{\deg_G(x_i)},
\]
for $1 \leq i \leq n$, is a fractional transversal.

**Proof:** If $G$ is monotone then for every $x \in B_r^n(x_1), \deg_G(x) \geq \deg_G(x_1)$. Therefore, the fractional transversal $w$ from Lemma 4 simply becomes \[
w_i = \frac{1}{\deg_G(x_i)}.
\]

As a result of Lemma 5 if $G$ is monotone, then the following expression is an upper bound on $A_G(n, 2r + 1)$, \[
A_G(n, 2r + 1) \leq \sum_{i=1}^{n} w_i = \sum_{i=1}^{n} \frac{1}{\deg_G(x_i)}. \tag{4}
\]
We call this bound the monotonicity upper bound, which holds in case that $G$ is monotone, and denote it by $MB(G, r)$. We will build upon Example 1 to exemplify the monotonicity upper bound for the $Z$ channel.

**Example 5.** It is straightforward to verify that the graph $G_Z$ from Example 1 satisfies the monotonicity property since for every $x, y \in \{0, 1\}^n$, if $y \in B_Z(x)$ then $w_H(y) \leq w_H(x)$. Thus, according to Lemma 5 the vector $w = (w_x)_{x \in \{0, 1\}^n}$ given by \[
w_x = \frac{1}{\deg_G(x)} = \frac{1}{\sum_{i=0}^{n} \binom{w_H(x)}{i}},
\]
is a fractional transversal. Therefore, the monotonicity upper bound $MB(G_Z, r)$ derived in 4 is calculated to be \[
MB(G_Z, r) = \sum_{x \in \{0, 1\}^n} w_x = \sum_{x \in \{0, 1\}^n} \frac{1}{\deg_G(x)} = \sum_{i=0}^{n} \binom{n}{i} \frac{1}{\sum_{i=0}^{n} \binom{w_H(x)}{i}}.
\]
For example, for $r = 1$, we get \[
MB(G_Z, 1) = \sum_{i=0}^{n} \binom{n}{w} \frac{1}{\sum_{i=0}^{n} \binom{w}{i}} = \sum_{i=0}^{n} \frac{n}{i+1} \frac{1}{\sum_{i=0}^{n} \binom{w_H(x)}{i}} = \frac{2^{n+1}}{n+1}.
\]
Note that the average sphere packing value, calculated in Example 1 for $r = 1$ is $\frac{2^{n+1}}{n+2}$, is stronger than the monotonicity upper bound. In fact, this hints that in some cases, which will be studied in the sequel, it is possible to improve upon the monotonicity upper bound. Indeed, it is possible to verify that in this case the fractional transversal according to Lemma 5 is not optimal by showing that the vector \[
w' = (w'_{x})_{x \in \{0, 1\}^n}, \quad w'_{x} = \frac{1}{w_H(x) + 1} w_H(x) + \frac{2}{w_H(x) + 3},
\]
for $x \neq 0$ and $w'_{0} = 1$, is a fractional transversal. The corresponding bound for this fractional transversal becomes \[
\frac{2^{n+1}}{n+3} \leq \frac{2^{n+1}}{n+2},
\]
which verifies the validity of the average sphere packing value. However, this choice of fractional transversal is still suboptimal and hence we seek to find a further improvement. Finding the exact value $\tau^*(H(G_Z, 2))$ will be the topic and problem we solve in Section IV.

The deletion channel which was studied in [15], the overlapping grain-error model studied in [8] and the non-overlapping grain error-error model for $r = 1, 2, 3$ studied in [12] all satisfy the monotonicity property. Indeed, all these works applied the monotonicity upper bound in order to derive upper bounds on the cardinalities of error-correcting codes in every channel. However, as will be shown in this work, the choice of the fractional transversal according to Lemma 5 is not necessarily optimal. This will be verified by providing different fractional transversals which yield stronger upper bounds than the ones achieved by the monotonicity upper bound.

**C. Automorphisms on Graphs**

One of the main obstacles in calculating the value of $\tau^*(H(G, r))$ is the large number of variables and constraints in the linear programming in 1. However, most of the graphs studied in this work contain symmetries between their vertices. For example, the linear programming in Example 1 for the $Z$ channel has $2^n$ variables and $2^n$ constraints in order to find the value of $\tau^*(H(G_Z, r))$, but it is not hard to notice that vectors of the same weight have identical behavior, and thus, one would expect to assign the same weight to these vertices. This will reduce the number of variables and constraints from $2^n$ to $n + 1$, which significantly simplifies the linear programming problem in 3. This subsection presents a scheme, based upon graph automorphisms, that in many cases can be used in order to significantly reduce the number of variables and constraints to calculate the bound $\tau^*(H(G, r))$. We will show the general scheme along with a demonstration how it is applied on our continued example of the $Z$ channel.

Let us first remind some tools derived from properties on automorphisms of graphs. Let $G = (X, E)$ be a directed graph with $n$ vertices. An automorphism of $G$ is a permutation of its vertices that preserves adjacency. That is, an automorphism of $G$ is a permutation $\pi: X \rightarrow X$ such that for all $(x, y) \in X \times X$, $(\pi(x), \pi(y)) \in E$ if and only if $(x, y) \in E$. Assume $|X| = n$, let $S_n$ be the set of all permutations of $n$ elements. The set of all automorphisms of $G$ is \[
Aut(G) = \{ \pi \in S_n | \pi \text{ is an automorphism of } G \}.
\]
It is known that $Aut(G)$ is a subgroup of the symmetric group $S_n$ under the operation of functions composition. The group $Aut(G)$ induces a relation $R$ on $X$ such that $(x, y) \in R$ if and only if there exists $\pi \in Aut(G)$ where $\pi(x) = y$. It is possible to verify that $R$ is an equivalence
order and hence $X$ is partitioned into $1 \leq n(G) \leq n$ equivalence classes, denoted by $X_1, \ldots, X_{n(G)}$. Furthermore, we denote $\mathcal{X}(G) = \{X_1, \ldots, X_{n(G)}\}$.

For any $c > 0$, let us define the set $W_c = \{w : w$ is a fractional transversal and $\sum_{i=1}^{n} w_i = c\}$.

Given a partition $X = \{X_1, \ldots, X_k\}$ of $X$, we say that a fractional transversal $w$ is $X$-regular if for all $1 \leq j \leq k$ and every $x, y \in X_j$, $w_x = w_y$.

Given a fractional transversal $w$ and an automorphism $\pi \in \text{Aut}(G)$, the vector $w^\pi$ is defined by $w^\pi = w_{\pi(i)}$. The next lemma proves that the vector $w^\pi$ is a fractional transversal as well.

**Lemma 6.** Let $w$ be a fractional transversal and $\pi$ an automorphism. Then, the vector $w^\pi$ is a fractional transversal as well.

*Proof:* It is clear to verify that $w^\pi \geq 0$. We need to show that for all $1 \leq i \leq n$ the following inequality holds

$$\sum_{y \in B_r(x_i)} w^\pi(y) \geq 1.$$ 

Since $\pi$ is an automorphism, $y \in B_r(x_i)$ if and only if $\pi(y) \in B_r(\pi(x_i))$ and therefore

$$\sum_{y \in B_r(x_i)} w^\pi(y) = \sum_{y \in B_r(\pi(x_i))} w_{\pi(y)} = \sum_{y \in B_r(\pi(x_i))} w_y \geq 1,$$

where the last inequality holds since $w$ is a fractional transversal.

Our main result in this part is stated in the next theorem and corollary.

**Theorem 7.** For every $c > 0$, if $W_c \neq \emptyset$ then $W_c$ contains an $\mathcal{X}(G)$-regular fractional transversal.

*Proof:* Let $w \in W_c$ be a fractional transversal. If $w$ is $\mathcal{X}(G)$-regular then the property holds. Otherwise, let $\pi \in \text{Aut}(G)$ and $w^\pi$ as defined above. Note that

$$\sum_{i=1}^{n} w_i^\pi = \sum_{i=1}^{n} w_{\pi(i)} = \sum_{i=1}^{n} w_i = c,$$

and together with Lemma[3] we get that $w^\pi \in W_c$. Similarly, we can show that $w^\pi \in W_c$. Let $\pi_1, \pi_2, \ldots, \pi_N$ be some order of the automorphisms in $\text{Aut}(G)$. We can similarly derive that the vector

$$w^* = \frac{\sum_{i=1}^{N} w_{\pi_i}}{N}$$

belongs to $W_c$ as well.

We finally show that $w^*$ is $\mathcal{X}(G)$-regular. For all $1 \leq j \leq n(G)$ and $x_{n_1}, x_{n_2} \in X_j$

$$w_{n_2}^* = \frac{\sum_{i=1}^{N} w_{\pi_i(n_1)}}{N} = \frac{\sum_{i=1}^{N} w_{\pi_i(n_2)}}{N}.$$ 

Now, let $\pi^* \in \text{Aut}(G)$ be such that $\pi^*(n_2) = n_1$ and note that

$$\{\pi_1, \ldots, \pi_N\} = \{\pi^* \circ \pi_1, \ldots, \pi^* \circ \pi_N\}.$$

Thus, we get

$$w_{n_2}^* = \frac{\sum_{i=1}^{N} w_{\pi_i(n_1)}}{N} = \frac{\sum_{i=1}^{N} w_{(\pi^* \circ \pi_i)(n_2)}}{N}.$$ 

Lastly, we note that Theorem[7] holds not only for the automorphism group $\text{Aut}(G)$ but also for every subgroup $H$ of $\text{Aut}(G)$. Given a subgroup $H$ of $\text{Aut}(G)$, assume it partitions the vertices set $X$ into $n_H$ equivalence classes $X_H = \{X_1, \ldots, X_{n_H}\}$. Let $A_H$ be an $n_H \times n_H$ adjacency matrix corresponding to the subgroup $H$, such that for $1 \leq i, j \leq n_H$,

$$A_H(i, j) = \frac{|\{(x, y) : x \in X_i, y \in B_r(x) \cap X_j\}|}{|X_i|}.$$

The next Corollary summarizes this discussion.

**Corollary 8.** Let $H$ be a subgroup of $\text{Aut}(G)$ and $X_H = \{X_1, \ldots, X_{n_H}\}$ is its partition of $X$ into $n_H$ equivalence classes. Then, the generalized sphere packing bound $\tau^*(\mathcal{H}(G, r))$ from[1] becomes

$$\tau^*(\mathcal{H}(G, r)) = \min \left\{ \sum_{i=1}^{n_H} |X_i| w_i : A_H \cdot w \geq 1, w \in \mathbb{R}_{+}^{n_H} \right\}. (6)$$

*Proof:* According to Theorem[7] it is enough to consider only fractional transversals which are $X_H(G)$-regular. Such a fractional transversal can be represented by a vector $w \in \mathbb{R}_{+}^{n_H}$ such that for $1 \leq i \leq n_H$, $w_i$ is the weight given to all the vectors in the set $X_i$.

The condition $A^T \cdot w$ from[1] can be stated as for all $x \in X$, $\sum_{y \in B_r(x)} w_y \geq 1$. However, for all $x \in X_i$ the number of vertices $y \in B_r(x)$ which belong to some set $X_j$ is fixed and is given by the value $A_H(i, j)$. Therefore, for every $x \in X_i$, this condition can be written as $\sum_{j=1}^{n_H} A_H(i, j) w_j \geq 1$. Finally, since there are $|X_i|$ vectors which are assigned with weight $w_i$ we get that the weight of this $X_H(G)$-regular fractional transversal is $\sum_{i=1}^{n_H} |X_i| w_i$ and thus the corollary holds.

The next example shows how to apply the automorphisms scheme presented in this subsection for the $Z$ channel.

**Example 6.** In Example[1] we saw that in order to find the value $\tau^*(\mathcal{H}(G, Z, r))$ according to[3], it is required to solve a linear programming with $2^n$ variables and $2^n$ constraints. Let us demonstrate how the automorphism scheme studied in this subsection can reduce both the number of variables and constraints to be $n+1$.

First, we define the following set of automorphisms on $G_Z$. For every $\sigma \in S_n$, a permutation $\pi_\sigma : \{0, 1\}^n \rightarrow \{0, 1\}^n$ is defined such that for all $x \in \{0, 1\}^n$, $(\pi_\sigma(x))_i = x_{\sigma(i)}$. It is possible to verify that the set $H = \{\pi_\sigma \mid \sigma \in S_n\}$ is a subgroup of $\text{Aut}(G_Z)$. Furthermore, the set $\{0, 1\}^n$ is partitioned under $H$ into $n+1$ equivalence classes $X_H(G_Z) = \{X_0, X_1, \ldots, X_n\}$, where $X_i = \{x \in \{0, 1\}^n : w_H(x) = i\}$, for $0 \leq i \leq n$. Therefore, according to equation (6) in Corollary[8] it is enough to limit our search and find only fractional
transversals \( w \) which are \( X_H(G_Z) \)-regular. Hence, the problem in (3) is simplified to be

\[
\tau^*(\mathcal{H}(G_Z, r)) = \min \left\{ \sum_{i=0}^{n} n^i \omega_i : \sum_{i=0}^{\min\{m, n\}} \left( \begin{array}{c} n \\ i \end{array} \right) w_{-i} \geq 1, 0 \leq i \leq n \right\}.
\]

(7)

In the next section we will continue Example 3 and show exactly how to solve the problem in (7).

IV. THE Z CHANNEL

The Z channel was already discussed before in Examples 1, 2, and 3. We derived the linear programming problem to find the value \( \tau^*(\mathcal{H}(G_Z, r)) \) in (3) and calculated its average sphere packing value. Then, we saw that \( G_Z \) is monotone and thus we calculated its monotonicity upper bound. Finally, we showed how to use the graph automorphism approach in order to derive a more compact linear programming problem to calculate \( \tau^*(\mathcal{H}(G_Z, r)) \) in (7).

The goal of this section is to solve the linear programming problem in (7) by finding the appropriate transversal property for all radius \( r \) and look at \( w_k^*(n) \) in this subsection.

A. Equivalence of the two formulas

In order to see the equivalence of two definitions, we fix \( r \) and look at \( w_k^* \) as a function of both \( k \) and \( n \) denoted by \( w_k^*(n) \) in this subsection. Let the sequence \( \Delta_k(n) \) as \( \Delta_k(n) = w_k^*(n) - w_k^*(n-1) \) for all \( n \). So,

\[
\Delta_k(n) = \begin{cases} \frac{1}{r} & \text{if } k = n - r \\ 0 & \forall k \geq n - r \\ -\sum_{i=1}^{k} \Delta_{k+i}(n) \frac{(k+i)}{(k+r)} & \forall k < n - r \end{cases}
\]

Now, we define another sequence \( D_i(n) \) as \( D_i(n) = \Delta_{n-i-1}(n) \cdot \frac{n!}{r!(n-r-1)!} \) to normalize and reverse the direction of the recursion:

\[
D_0(n) = D_1(n) = \cdots = D_{r-2}(n) = 0,
D_{r-1}(n) = 1,
\frac{D_i(n)}{r!} + \frac{D_{i-1}(n)}{(r-1)!} + \cdots + \frac{D_{i-r}(n)}{0!} = 0 \quad \forall i \geq r.
\]

Note that \( D_i(n) \) is independent of \( n \). So, we drop \( n \) and write \( w_k^*(n) \) as

\[
w_k^*(n) = \Delta_k(n) + \Delta_k(n-1) + \cdots + \Delta_k(k+r)
= \frac{r!k!}{n!} D_{n-k-1} + \frac{r!k!}{(n-1)!} D_{n-k-2} + \cdots + \frac{r!k!}{(k+r)!} D_{r-1}(n)
= r!k! D_{m=k-1}
\]

from (3) and calculated its average sphere packing value. Then, we saw that \( G_Z \) is monotone and thus we calculated its monotonicity upper bound. Finally, we showed how to use the graph automorphism approach in order to derive a more compact linear programming problem to calculate \( \tau^*(\mathcal{H}(G_Z, r)) \) in (7).

B. Transversal property for \( w^* \)

The definition of \( w^* \) in (8) ensures that the inequality constraints in (7) are satisfied. So, the non-negativity of \( w^* \) is enough to show \( w^* \) is a valid transversal.

First, we study the case \( r = 1 \). A simple induction on \( i \), shows that \( D_i = (-1)^i \). Therefore,

\[
w_k^* = \sum_{m=k+1}^{n} \frac{(-1)^{m-k-1}}{m!}
= \left( \frac{1}{k+1} \right) \left( \frac{1}{k+1} \right) \left( k+1 \right) \left( k+2 \right) \left( k+3 \right) \cdots > 0.
\]

In general, it is not easy derive an explicit formula for \( w^* \) for \( r \geq 2 \). However, we show that \( D_m \) is bounded by an exponential function of \( 2r \) and hence, the first few terms in (3)
are dominant comparing to the rest and \( w_k^* \geq 0 \) is mostly the case. Let us first verify the statement \( w_k^* > 0 \) for \( k \geq 3r - 1 \):

\[
w_k^* = r!k! \sum_{m=r+k}^{n} \frac{D_{m-k-1}}{m!} = r!k! \left( \frac{1}{(r+k)!} + \sum_{m=r+k+1}^{n} \frac{D_{m-k-1}}{m!} \right) \\
\geq r!k! \left( \frac{1}{(r+k)!} + \sum_{m=r+k+1}^{n} \frac{|D_{m-k-1}|(r+k)!}{m!} \right) \\
\geq r!k! \left( \frac{1}{(r+k)!} - \sum_{m=r+k+1}^{n} \frac{|D_{m-k-1}|}{(4r)^{m-r-k}} \right) \\
\geq r!k! \left( 1 - \sum_{m=r+k+1}^{n} \frac{(2r)^{m-r-k}}{m!} \right) \\
= r!k! \left( 1 - \sum_{m=1}^{n-r-k} 2^{-m} \right) = r!k! 2^{-(n-k-r)} > 0,
\]

where the last inequality comes from Lemma 24 in appendix A. In other words, for \( k \geq 3r - 1 \), the first term in (8) is larger than the sum of the absolute values of the remaining terms and they cannot cancel it out. The proof of the case \( k < 3r - 1 \) is incomplete for arbitrary radius \( r \). However, we introduce a method to verify the feasibility (transversal property) of \( w^* \) for any fixed \( r \) in the following fashion:

Given \( k < 3r - 1 \), we look for a number \( n_k \) such that

\[
\sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} \geq \frac{1}{(2r)^{r+k}} \left( e^{2r} - \sum_{m=0}^{n_k} \frac{(2r)^m}{m!} \right), \tag{12}
\]

which means for all \( n > n_k \) we have

\[
w_k^* = r!k! \sum_{m=r+k}^{n} \frac{D_{m-k-1}}{m!} \\
= r!k! \left( \sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} + \sum_{m=n_k+1}^{n} \frac{D_{m-k-1}}{m!} \right) \\
\geq r!k! \left( \sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} - \sum_{m=n_k+1}^{n} \frac{(2r)^{m-r-k}}{m!} \right) \\
> r!k! \left( \sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} - \sum_{m=n_k+1}^{n} \frac{(2r)^{m-r-k}}{m!} \right) \\
= r!k! \left( \sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} - e^{2r} \sum_{m=0}^{n_k} \frac{(2r)^m}{m!} \right) \geq 0;
\]

And then we check the values of \( w_k^* \) for the finite set of \( k < 3r - 1 \) and \( n \leq n_k \). Note that,

\[
\lim_{n_k \to \infty} e^{2r} - \sum_{m=0}^{n_k} \frac{(2r)^m}{m!} = 0.
\]

Also, \( D_i \) is bounded by an exponential function (see Lemma 24) and hence the following limit exists

\[
\ell_k := \lim_{n_k \to \infty} \sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!}.
\]

Finally, if \( w_k^* > \epsilon_k > 0 \) for all \( n > k + r \), then \( \ell_k \geq \frac{\epsilon_k}{(2r)^{r+k}} > 0 \). So, the number \( n_k \) should exists. As an example, when \( r = 2 \) we have \( n_1 = n_2 = 6 \), and \( n_3 = n_4 = 7 \). Using the above approach, we have verified the feasibility for all \( r \leq 20 \).

Our calculations also show that \( n_k \leq 4r - 1 \) for all \( n \leq 20 \). In appendix A we prove that \( w^* \) defined in (8), is also the optimal transversal assignment and gives us the best bound using these approach.

In order to evaluate the results, we compared between the different upper bounds for the Z channel. The first bound is the monotonicity bound (MB in short), which was calculated in Example 5; the second one is the average sphere packing value (ASPV in short), which was calculated in Example 1; and the third bound is the generalized sphere packing bound (GSPB in short). The best known (to us) upper bound for the Z channel, due to Weber, De Vroedt, and Boekee [24], appears in the last column of Table I. We see from Table I that this bound is better than the GSPB even under optimal weight assignment. However, the bound of [24] involves solving an integer programming problem, and the authors of [24] have computed this bound only for \( n \leq 23 \). In contrast, our bound in Theorem 3 is easy to compute for all \( n \), and we give its values for \( r = 1, 2, 3, 4 \) up to \( n \leq 32 \) in Tables I, II, III, and IV.

\begin{table}[h]
\centering
\caption{Z channel: upper bounds comparison for \( r = 1 \)}
\begin{tabular}{|c|c|c|c|}
\hline
\( n \) & MB & ASPV & GSPB \\
\hline
5 & 10 & 9 & 3 \\
6 & 18 & 16 & 14 \\
7 & 32 & 28 & 26 \\
8 & 56 & 51 & 47 \\
9 & 102 & 93 & 86 \\
10 & 186 & 170 & 159 \\
11 & 341 & 315 & 295 \\
12 & 630 & 585 & 551 \\
13 & 1170 & 1092 & 1032 \\
14 & 2184 & 2048 & 1940 \\
15 & 4095 & 3855 & 3662 \\
16 & 7710 & 7281 & 6935 \\
17 & 14563 & 13797 & 13170 \\
18 & 2794 & 26214 & 25075 \\
19 & 5248 & 49932 & 47853 \\
20 & 99864 & 95324 & 91514 \\
21 & 190650 & 182361 & 175351 \\
22 & 364722 & 349525 & 336586 \\
23 & 699050 & 671088 & 647131 \\
24 & 1342177 & 1290555 & 1264606 \\
25 & 2581110 & 2485513 & 2402690 \\
26 & 4971026 & 4793400 & 4638907 \\
27 & 9586980 & 9256395 & 8967211 \\
28 & 18512790 & 17856957 & 17353537 \\
29 & 35791394 & 34636833 & 33618332 \\
30 & 69273666 & 67108864 & 65191862 \\
31 & 13421728 & 130150524 & 12653913 \\
32 & 260301048 & 252645135 & 245818070 \\
\hline
\end{tabular}
\end{table}

In the next section, we will extend the study of the Z channel for non-binary symbols.

V. LIMITED MAGNITUDE CHANNELS

We turn in this section to generalize the Z channel for the non-binary case. In this setup, every symbol can have \( q \) values, \( 0, 1, \ldots, q - 1 \) and we denote \([q] = \{0, 1, \ldots, q - 1\}\). We study the limited magnitude model and focus solely on the single error setup which is carried for two cases. Namely, the
error can be asymmetric (Fig. 4(a)) or symmetric (Fig. 4(b)). This error-channel is motivated by the feature of the errors in non-binary flash memories. The cells in flash memories are charged with electrons and due to the inaccuracy in cell-programming and electrons leakage, the charge level of a cell can either increase or decrease by limited magnitude. For more details see for example [5], [6], [13], [17], [28].

### A. Asymmetric Errors

In the asymmetric non-binary channel, the value of every symbol can only decrease, and in this study we only consider the case where the value of each symbol can decrease by one. The corresponding graph is $G_{A,q} = (X_{A,q}, E_{A,q})$, where
Thus, the average sphere packing value in this case is given by

$$X_{A,q} = [q]^n$$

and

$$E_{A,q} = \left\{ (x, y) : x, y \in [q]^n, x \geq y, \sum_{i=1}^n x_i = \sum_{i=1}^n y_i + 1 \right\}.$$ 

Given some $x \in [q]^n$, its ball of radius one is described by the set $B_{A,q,1}(x) = \{ y \in [q]^n : x \geq y, \sum_{i=1}^n x_i - \sum_{i=1}^n y_i \leq 1 \}$, and $d_{E_{A,q,1}}(x) = w_H(x) + 1$. The hypergraph in this case is $H(G_{A,q,1}) = \langle X_{A,q,1}, E_{A,q,1} \rangle$, where $X_{A,q,1} = [q]^n$ and $E_{A,q,1} = \{ B_{A,q,1}(x) : x \in [q]^n \}$.

According to the above definitions it is immediate to verify that for all $y \in B_{A,q,1}(x)$, $w_H(y) \leq w_H(x)$ and thus the graph $G_{A,q}$ is monotone. In the next two lemmas we calculate the monotonicity upper bound and the average sphere packing value under this setup.

**Lemma 10.** The monotonicity upper bound of the graph $G_{A,q}$ for $r = 1$ is

$$MB(G_{A,q,1}) = \frac{q^{n+1}}{(q-1)(n+1)}.$$ 

**Proof:** Since the graph $G_{A,q}$ is monotone, according to Lemma 5 the following vector $w = (w_x)_{x \in [q]^n}$ is a fractional transversal,

$$w_x = \frac{1}{\deg_{E_{A,q,1}}(x)} = \frac{1}{w_H(x) + 1}.$$ 

Thus, the monotonicity upper bound from Equation (2) becomes

$$MB(G_{A,q,1}) = \sum_{x \in [q]^n} w_x = \sum_{x \in [q]^n} \frac{1}{w_H(x) + 1} = \sum_{0 \leq i_1 + \cdots + i_{q-1} \leq n} \frac{n}{i_1 + \cdots + i_{q-1} + 1} = \frac{q^{n+1}}{(q-1)(n+1)}.$$ 

**Lemma 11.** The average sphere packing value of the graph $G_{A,q}$ for $r = 1$ is

$$ASPV(G_{A,q,1}) = \frac{q^{n+1}}{(q-1)(n+1) + 1}.$$ 

**Proof:** The value of the average ball size is

$$\frac{1}{q^n} \sum_{x \in [q]^n} (w_H(x) + 1) = \frac{1}{q^n} \sum_{0 \leq i_1 + \cdots + i_{q-1} \leq n} \frac{n}{i_1 + \cdots + i_{q-1} + 1} \left( i_1 + \cdots + i_{q-1} + 1 \right) = \frac{1}{q^n} \cdot \left( nq^{n-1}(q - 1) + q^n \right) = n + 1 - n/q.$$ 

Thus, the average sphere packing value in this case is given by

$$\frac{q^n}{n + 1 - n/q} = \frac{q^{n+1}}{(q-1)(n+1) + 1}.$$ 

The linear programming problem from Equation (3) becomes

$$\tau^*(\mathcal{H}(G_{A,q,1})) = \min \left\{ \sum_{x \in [q]^n} w_x : \sum_{y \in B_{A,q,1}(x)} w_y \geq 1 \right\}.$$ 

However, it can be significantly simplified according to the tools developed in Section 3.6.2. Similarly to the set of automorphisms from Example 6 for every permutation $\sigma \in S_n$ we define a permutation $\pi_\sigma = [q]^n \rightarrow [q]^n$ such that for all $x \in [q]^n$, $(\pi_\sigma(x))_i = x_{\sigma(i)}$. Hence, also here the set $H_A = \{ \pi_\sigma : \sigma \in S_n \}$ is a subgroup of $\text{Aut}(G_{A,q})$. However, now the subgroup $H_A$ partitions the set $[q]^n$ into the following $n_A = (n+q-1)$ equivalence classes

$$\mathcal{X} = \left\{ X_i : i = (i_0, \ldots, i_{q-1}) \geq 0, \sum_{j=0}^{q-1} i_j = n \right\},$$ 

where $X_i$ is characterized as follows

$$X_i = \{ x \in [q]^n : x^{-1}(j) = i_j, 0 \leq j \leq q-1 \},$$ 

and $x^{-1}(j) = \{ 1 \leq k \leq n : x_k = j \}$. We denote the set $I_A$ to be $I_A = \{ i : i = (i_0, \ldots, i_{q-1}) \geq 0, \sum_{j=0}^{q-1} i_j = n \}$ and define an $n_A \times n_A$ matrix $A_I$ such that its entries are the vectors $(i, j) \in I_A \times I_A$. We assign the values $A_{H,I}(i, i) = 1$ and $A_{H,I}(i, j) = \frac{1}{(q-1)(n+1)}$ if there exists $1 \leq k \leq n$ such that $j_k = i_k - 1$ and $j_{k+1} = i_{k+1} + 1$ and for all $1 \leq k \leq n$ \{ $k, k+1$ \}, $j_k = i_k$. All other values in the matrix $A_I$ are assigned with the value $0$. Finally, according to Corollary 3, we proved the following theorem.

**Theorem 12.** The generalized sphere packing bound for $H(G_{A,q,1})$ is given by

$$\tau^*(\mathcal{H}(G_{A,q,1})) = \min \left\{ \sum_{i \in I_A} |X_i| w_i : A_{I,I} \cdot w \geq 1, w = (w_i)_{i \in I_A} \in \mathbb{R}^{n_A}_+ \right\}.$$ 

We finish this section by showing an improvement upon the suboptimal monotonicity upper bound from Lemma 10. In the fractional transversal notation of Theorem 12 if one applied the monotonicity upper bound, then the fractional transversal assignment would be $w_i = 1/(n - i_0 + 1)$ for $i \in I_A$. However, under this assignment almost all of the constraints hold with strict inequality. We show that it is possible to reduce the weights in this assignment without violating the constraints and thus receive a stronger upper bound.

**Theorem 13.** The vector $w = (w_i)_{i \in I_A}$ given by

$$w_i = \frac{1}{n - i_0 + 1 + \frac{i_0 - 1}{2(n - i_0)}}$$ 

if $i_0 \neq n$ and otherwise $w_i = 1$ is a fractional transversal for $\tau^*(\mathcal{H}(G_{A,q,1}))$ as stated in Theorem 12.

**Proof:** It is straightforward to verify that $w_i \geq 0$ for all $i \in I_A$. According to the conditions for $\tau^*(\mathcal{H}(G_{A,q,1}))$.
from Theorem 12 we need to show that for all \(i = (i_0, i_1, \ldots, i_{q-1}) \in \mathcal{I}_A\) the following inequality holds

\[
w_{(i_0i_1, \ldots, i_{q-1})} + \frac{1}{2(n-i_0)} + i_1 w_{(i_1)} + i_2 w_{(i_1, i_2)} + \cdots + i_{q-1} w_{(i_1, \ldots, i_{q-1}, 1)} \geq 1.
\]

If \(i_0 = n\) then this inequality holds with equality and it is possible to verify that it holds for \(i_0 = n - 1\) as well. Thus, we can assume that \(i_0 < n - 1\). After placing the values of \(w_i\) stated in the theorem, we need to show the following

\[
\frac{1}{n - i_0 + 1 + \frac{i_0 - 1}{2(n-i_0)}} + \frac{i_1}{n - i_0 + \frac{i_1 - 2}{2(n-i_0-1)}} + \frac{i_2}{n - i_0 + \frac{i_2 - 1}{2(n-i_0-2)}} + \cdots + \frac{i_{q-1}}{n - i_0 + 1 + \frac{i_{q-1} - 1}{2(n-i_0+1-q)}} \geq 1.
\]

Note that

\[
\frac{1}{n - i_0 + 1 + \frac{i_0 - 1}{2(n-i_0)}} \geq \frac{1}{n - i_0 + 1 + \frac{i_1}{2(n-i_0)}},
\]

and thus it is enough to show that

\[
\frac{i_2 + i_3 + \cdots + i_{q-1} + 1}{n - i_0 + 1 + \frac{i_1}{2(n-i_0)}} \geq \frac{i_1}{n - i_0 + \frac{i_1 - 2}{2(n-i_0-1)}}
\]

or, since \(i_0 + i_1 + \cdots + i_{q-1} = n\),

\[
\frac{n - i_0 - i_1 + 1}{n - i_0 + 1 + \frac{i_1 - 2}{2(n-i_0-1)}} + \frac{i_1}{n - i_0 + \frac{i_1 - 2}{2(n-i_0-1)}} \geq 1
\]

and

\[
\frac{i_1}{n - i_0 + \frac{i_1 - 2}{2(n-i_0-1)}} \geq \frac{i_1 + i_{q-1}}{n - i_0 + 1 + \frac{i_{q-1}}{2(n-i_0+1)}}.
\]

Let us denote \(n - i_0 = M\) and we need to show that

\[
\frac{1}{M + \frac{i_1}{2(M-1)}} \geq \frac{1}{M + 1 + \frac{i_1}{2M}},
\]

and equivalently

\[
\frac{2(M-1)}{2M^2 - 2M + i_1 - 2} \geq \frac{2M + 1}{2M^2 + 2M + i_1},
\]

or

\[
2M^2 + 2M + 2 \geq 3i_1,
\]

which holds since \(M = n - i_0 \geq i_1\).

Table V summarizes the upper bounds results we derived in this section for \(q = 3\). The first column is the monotonicity upper bound we found in Lemma 10. The second column is the average sphere packing value from Lemma 11. The third column is the improvement in Theorem 13 over the monotonicity upper bound. Lastly, the last column is the value of the generalized sphere packing bound from Theorem 12 which we solved numerically. Note that there is no upper bound we know of in the literature for this error channel.

### B. Symmetric Errors

Since this model and graph are very similar to the asymmetric case, they are briefly presented. The graph is given by \(G_{S,q} = (X_{S,q}, E_{S,q})\), where \(X_{S,q} = \{[q]^n\}\) and

\[E_{S,q} = \{(x, y) : (x, y) \in E_{A,q} \text{ or } (y, x) \in E_{A,q}\}.
\]

Similarly, for every \(x \in [q]^n\), its corresponding ball of radius one is the set \(B_{S,q,1}(x) = \{y \in [q]^n : y \in B_{A,q,1}(x) \text{ or } x \in B_{A,q,1}(y)\}\). The hypergraph is \(\mathcal{H}(G_{S,q} = 1) = (X_{S,q,1}, E_{S,q,1})\), where \(X_{S,q,1} = [q]^n\) and \(E_{S,q,1} = \{B_{S,q,1}(x) : x \in [q]^n\}\).

This setup is different than all other error channels studied so far in the sense that it does not satisfy the monotonicity property. Thus, we cannot conclude the corresponding fractional transversal of the monotonicity upper bound. However, we can still calculate the average sphere packing value.

**Lemma 14.** The average sphere packing value of the graph \(G_{S,q}\) for \(r = 1\) is

\[\text{ASPV}(G_{S,q}, 1) = \frac{q^n}{2n + 1 - 2n/q}.
\]

**Proof:** First we calculate the value of the expected ball size, which is given by

\[
\frac{1}{q^n} \sum_{x \in [q]^n} (2n + 1 - x^{-1}(i_0) - x^{-1}(i_{q-1}))
\]

\[
= \frac{1}{q^n} \cdot \sum_{i_0 + \cdots + i_{q-1} = n} \left(\sum_{i_1, \ldots, i_{q-1}} (i_1 + \cdots + i_{q-2})\right)
\]

\[
= \frac{1}{q^n} \cdot ((n + 1)q^n + n(q - 2)q^{n-1}) = 2n + 1 - 2n/q.
\]

Thus, the average sphere packing value becomes

\[\frac{q^n}{2n + 1 - 2n/q}.
\]
Next, we define the set of automorphisms to be used here. One can verify that every permutation in $H_A$ is an automorphism in $G_{S,q}$. However, in this case we can expand and use more automorphisms. For every binary vector $b \in \{0,1\}^n$, we define the permutation
\[
\pi_b : [q]^n \rightarrow [q]^n,
\]
as follows. For every $x \in [q]^n$, $\pi_b(x)$ is the vector defined as
\[
\pi_b(x)_i = \begin{cases} 
q - 1 - x_i & \text{if } b_i = 1 \\
 x_i & \text{if } b_i = 0
\end{cases}
\]
Then, the set $H_S = H_A \cup \{\pi_b : b \in \{0,1\}^n\}$ is a subgroup of $Aut(G_{S,q})$. The subgroup $H_S$ partitions the set $[q]^n$ into $n_S = (\frac{n+1}{2}) - 1$ equivalence classes
\[
\mathcal{X} = \left\{ x_i : i = (i_0,\ldots,i_{\lfloor q/2 \rfloor - 1}) \in 0, \sum_{j=0}^{\lfloor q/2 \rfloor - 1} i_j = n \right\},
\]
and $X_i$ is the set
\[
X_i = \{ x \in [q]^n : x^{-1}(j) + x^{-1}(q-1-j) = i_j, 0 \leq j \leq \lfloor q/2 \rfloor - 1 \}.
\]
We define
\[
I_S = \{ i : i = (i_0,\ldots,i_{\lfloor q/2 \rfloor - 1}) \in 0, \sum_{j=0}^{\lfloor q/2 \rfloor - 1} i_j = n \}
\]
and the $n_S \times n_S$ matrix $A_S$ with the following entries $(i,j) \in I_S \times I_S$:
\[
1) \text{ For all } i \in I_S, A_S(i,i) = 1,
2) \text{ } A_S(i,j) = k \text{ if there exists } 1 \leq k \leq \lfloor q/2 \rfloor - 1 \text{ such that } j = i_k - 1 \text{ and } j_{k-1} = i_{k-1} + 1 \text{ and for all } \ell \in \lfloor [q/2] \rfloor \setminus \{k,k-1\}, j_{\ell} = i_{\ell}.
\]
To conclude, according to Corollary 8, the generalized sphere packing for $H(G_{S,q+1})$ becomes
\[
\tau^*(H(G_{S,q+1})) = \min \left\{ \sum_{i \in I_S} X_i | w_i \in A_S^2 \wedge (w_i) \in \mathbb{R}^n_{+} \right\}.
\]
Even though this graph does not satisfy the monotonicity property we can still derive a similar bound, which will be stated in the following theorem.

**Theorem 15.** The vector $w = (w_x)_{x \in [q]^n}$ given by
\[
w_x = \frac{1}{\deg_{S,q+1}(x) - 1}
\]
is a fractional transversal.

**Proof:** Let $x \in [q]^n$ and let $i_j = x^{-1}(j)$ for $j \in [q]$. Then, $\deg_{S,q+1}(x) = 2n - i_0 - i_{q-1} + 1$. We need to show that
\[
1 + i_1 + 2i_2 + \cdots + 2i_{q-3} + i_{q-2} + \frac{i_0 + i_{q-1}}{2n - i_0 - i_{q-1} + 1} \geq 1.
\]

Since $1/(2n - i_0 - i_{q-1} + 1) \geq 1/(2n - i_0 - i_{q-1} + 1)$ and $1/(2n - i_0 - i_{q-1} + 1) \geq 1/(2n - i_0 - i_{q-1} + 1)$, it is enough to show that
\[
1 + i_1 + 2i_2 + \cdots + 2i_{q-3} + i_{q-2} + \frac{i_0 + i_{q-1}}{2n - i_0 - i_{q-1} + 1} \geq 1.
\]
which holds with equality.

Comparison results for $q = 3$ and $q = 4$ are summarized in Tables VI and VII. The first column is the average sphere packing value which was calculated in Lemma 14. The second column is the upper bound we found in Theorem 15. The last column is the value of the generalized sphere packing bound that we solved numerically. Note that in this example the value of the average sphere packing value is less than the one of the generalized sphere packing value, however, that doesn’t mean that it is a valid upper bound.

**TABLE VI**

| $n$ | ASPV | Theorem 15 | GSPB |
|-----|-------|------------|------|
| 5   | 31    | 37         | 32   |
| 6   | 81    | 93         | 82   |
| 7   | 211   | 238        | 216  |
| 8   | 562   | 624        | 572  |
| 9   | 1514  | 1663       | 1538 |
| 10  | 4119  | 4484       | 4177 |
| 11  | 11307 | 12217      | 11449|
| 12  | 31261 | 33564      | 31618|
| 13  | 86963 | 92872      | 87872|
| 14  | 243201 | 258535    | 245544|
| 15  | 683281 | 723466    | 689388|
| 16  | 1927465 | 2033685  | 1943532|
| 17  | 5456626 | 5739520  | 5499244|
| 18  | 15496819 | 16255303 | 15610684|

**TABLE VII**

| $n$ | ASPV | Theorem 15 | GSPB |
|-----|-------|------------|------|
| 5   | 120   | 139        | 125  |
| 6   | 409   | 463        | 417  |
| 7   | 1424  | 1586       | 1449 |
| 8   | 5041  | 5540       | 5115 |
| 9   | 18078 | 19666      | 18313|
| 10  | 65536 | 70707      | 66297|
| 11  | 239674 | 256844    | 242193|
| 12  | 853011 | 940934    | 891482|

VI. DELETION AND GRAND-ERROR CHANNELS

In this section we shift our attention to the deletion channel, which was the original usage of the generalized sphere packing
bound in [15]. We will only focus on the single-deletion case. First, we revisit the fractional transversal given in [15] to verify that the graph in the deletion channel satisfies a similar property to the monotonicity property from Section [II-B]. However, it will be noticed that this choice of fractional transversal is suboptimal and thereof an improvement will be presented. This will be our main result in this section, namely, an explicit expression of a fractional transversal which improves upon the one from [15]. Since the structure of the deletion and grain-error-correcting codes.

A. Deletions

As studied in the previous examples and sections, we first introduce the graph for the deletion channel. However, note that the graph in this setup is different than the previous ones studied so far. Specifically, a length $n$ vector which suffers a single deletion will result with a vector of length $n-1$. To accommodate this structure, the vertices in the graph are defined to be both vectors of length $n$ and $n-1$, so the graph is $G_D = (X_D, E_D)$, where $X_D = \{0, 1\}^n \cup \{0, 1\}^{n-1}$ and $E_D = \{ (x, y) \in \{0, 1\}^n \times \{0, 1\}^{n-1} : y = (x_1, \ldots, x_i, x_{i+2}, \ldots, x_n) \text{ for some } 1 \leq i \leq n \}$. For any $x \in \{0, 1\}^n$, its radius one ball is the set $B_{D,1}(x) = \{ y \in \{0, 1\}^{n-1} : (x, y) \in E_D \}$, and for $x \in \{0, 1\}^{n-1}$, $B_{D,1}(x) = \emptyset$. Therefore, $1 \leq \deg_{D,1}(x) \leq n$ for $x \in \{0, 1\}^n$, and $\deg_{D,1}(x) = 0$ for $x \in \{0, 1\}^{n-1}$.

At this point, we could basically construct the hypergraph for the deletion channel as was done in the previous examples such that its set of vertices is $X_D = \{0, 1\}^n \cup \{0, 1\}^{n-1}$. However, since the length-$n$ vectors do not participate in the balls we can eliminate them in the hypergraph construction, which coincides with the hypergraph construction in [15]. Thus the hypergraph for the single deletion channel is $H(G_D, 1) = (X_{D,1}, E_{D,1})$, where $X_{D,1} = \{0, 1\}^{n-1}$ and $E_{D,1} = \{ B_{D,1}(x) : x \in \{0, 1\}^{n-1} \}$. This definition does not change the analysis of the upper bounds studied in this paper. Thus, the generalized sphere packing bound in this setup becomes

$$\tau^*(H(G_D, 1)) = \min \left\{ \sum_{x \in \{0, 1\}^{n-1}} w_x : \sum_{y \in B_{D,1}(x)} w_y \geq 1, \forall x \in \{0, 1\}^n \right\}$$

For a vector $x \in \{0, 1\}^n$, we denote by $\rho(x)$ the number of runs in $x$. For example, if $x = 001010010$, then $\rho(x) = 7$. It is easily verified that for $x \in \{0, 1\}^n$, $\deg_{D,1}(x) = \rho(x)$. [15]. It is also known that the number of length-$n$ vectors with $1 \leq \rho \leq n$ runs is $2^{\rho-1}$. Let us first calculate the average sphere packing value for the hypergraph $H(G_D, 1)$. This will be done in the next lemma.

Lemma 16. The average sphere packing value of the graph $G_D$ for $r = 1$ is

$$\text{ASPV}(G_D, 1) = \frac{2^n}{n+1}.$$
Lemma 17. For $2 \leq \rho \leq n$ and $0 \leq \mu \leq \rho - 2$, 
\[ N_n(\rho, \mu) = 2 \left( \frac{\rho - 2}{\mu} \right) \left( \frac{n - \rho + 1}{\rho - \mu - 1} \right). \]

Proof: For every $x = (x_1, \ldots, x_n) \in \{0, 1\}^n$, let $x' = (x'_1, \ldots, x'_{n-1}) \in \{0, 1\}^{n-1}$ be a vector of length $n - 1$ such that for $1 \leq i \leq n - 1$, $x'_i = x_i + x_{i+1}$. Note that $w_H(x') = \rho(x) - 1$. Let $c(x')$ denote the number of times that two consecutive ones appear in $x'$, so we have $c(x') = \mu(x)$.

Let $x$ be a vector of length $n$ such that $\rho(x) = \rho$ and $\mu(x) = \mu$, where $1 \leq \rho \leq n$ and $0 \leq \mu \leq \rho - 2$. Assume the vector $x'$ has $\rho$ runs of ones of length $h_1, \ldots, h_{\rho}$. Then, first we have that 
\[ \sum_{i=1}^{\rho} h_i = w_H(x') = \rho - 1. \] (14)

Every run of ones of length $h_i$ in $x'$ contributes $h_i - 1$ pairs of two consecutive ones. Therefore, 
\[ \sum_{i=1}^{\rho} (h_i - 1) = \mu. \] (15)

Together, from (14) and (15), we conclude that $\rho = \rho - \mu - 1$. Furthermore, the number of solutions to (14) (or (15)) is 
\[ \binom{n - \rho + 1}{\rho - \mu - 1} = \binom{n - \rho + 1}{\rho - 2}. \] For every solution $h_1, \ldots, h_{\rho - \mu - 1}$, let $k_0, k_1, \ldots, k_{\rho - \mu - 1}$ be the number of zeros between the blocks of ones in $x'$, where $k_0 \geq 0, k_1, \ldots, k_{\rho - \mu - 2} \geq 1$, and $k_{\rho - \mu - 1} \geq 0$. Note that their sum is $n - 1 - (\rho - 1) = n - \rho$.

Finally, the number of options to choose the vector $x'$ is the number of solutions to choose the runs of ones $h_1, \ldots, h_{\rho - \mu - 1}$ and runs of zeros $k_0, \ldots, k_{\rho - \mu - 1}$. Every choice of the vector $x'$ determines the vector $x$ up to choosing whether it starts with zero or one. Therefore, we get 
\[ N_n(\rho, \mu) = 2 \left( \frac{\rho - 2}{\mu} \right) \left( \frac{n - \rho + 1}{\rho - \mu - 1} \right). \]

Next, the main result in this section is proved.

Theorem 18. The vector $w = (w_x)_{x \in \{0, 1\}^{n-1}}$ defined by 
\[ w_x = \begin{cases} \frac{1}{\rho(x)} & \text{if } \mu(x) \leq 1 \\ \frac{1}{\rho(x)} \left( 1 - \frac{\mu(x)}{\rho(x) \tau} \right) & \text{otherwise} \end{cases} \]

is a fractional transversal.

Proof: Let $x$ be a length-$n$ binary vector with $\rho$ runs and $\mu$ middle-$1$-runs. We need to show that $\sum_{y \in B_{D,1}(x)} w_y \geq 1$. It can be verified that this claim holds for $\rho = 1, 2, 3$ or $\mu = 0, 1$ and thus we assume for the rest of the proof that $\rho \geq 4$ and $\mu \geq 2$. Note that for a fixed $\rho$, $w_x$ is decreases when $\mu$ increases.

If a vector $y \in B_{D,1}(x)$ is received by deleting a middle-$1$-run bit then $\rho(y) = \rho - 2$ and $\mu(y) \leq \mu - 1$. Otherwise, $\rho(y) = \rho$ and $\mu(y) \leq \mu + 1$ or, if it is the first or last bit which is a single-bit run, $\rho(y) = \rho - 1$ and $\mu(y) \leq \mu$, however, the worst case in terms of the value of $w_y$ is achieved for $\rho(y) = \rho$ and $\mu(y) = \mu + 1$. Therefore,
\[ \sum_{y \in B_{D,1}(x)} w_y \geq \frac{\rho - 1}{\rho - 2} \left( 1 - \frac{\mu - 1}{(\rho - 2)^2} \right) + \frac{\mu}{\rho} \left( 1 - \frac{\mu + 1}{\rho^2} \right) \]
\[ = 1 + \frac{2\mu}{\rho(\rho - 2)} - \frac{\mu(\mu - 1)}{(\rho - 2)^2} + \frac{\mu(\mu + 1)}{\rho^2} \]
and thus it is enough to show that for $\rho \geq 4, 2 \leq \mu \leq \rho - 2$,
\[ \frac{2\mu}{\rho(\rho - 2)} - \frac{\mu(\mu - 1)}{(\rho - 2)^2} + \frac{1}{\mu^{\rho}} \geq 0 \]

The function
\[ f(\mu) = \frac{\mu - 1}{(\rho - 2)^3} + \frac{1}{\mu^3} \]
in the range $2 \leq \mu \leq \rho - 2$ is maximized either when $\mu = 2$ or $\mu = \rho - 2$ and thus we need to show that
\[ \frac{2}{\rho(\rho - 2)} - \frac{1}{\rho^3} \geq 0 \]
and
\[ \frac{2}{\rho(\rho - 2)} - \frac{3}{\rho^3} \leq 0 \]
which holds for all $\rho \geq 4$.

For a vector $x$ with $\rho$ runs and $\mu$ middle-$1$-runs, we denote its weight by $w(\rho, \mu)$, as specified in Theorem 18. From Lemma 17 and Theorem 18, we conclude with the following upper bound on $\tau^*(H(G_D, 1))$.

Theorem 19. The value $\tau^*(H(G_D, 1))$ satisfies
\[ \tau^*(H(G_D, 1)) \leq 2 + \sum_{\rho = 2, \mu = 0}^{n-1} N_{n-1}(\rho, \mu) w(\rho, \mu). \]

Proof: We calculate the upper bound on $\tau^*(H(G_D, 1))$ according to the fractional transversal from Theorem 18. 
\[ w = (w_x)_{x \in \{0, 1\}^{n-1}}. \] Every vector $x$ is assigned with a weight $w_x = w(\rho, \mu)$ according to its number of runs $\rho$ and number of middle-$1$-runs $\mu$. Thus, we get this upper bound to be
\[ \sum_{x \in \{0, 1\}^{n-1}} w_x \geq 2 + \sum_{\rho = 2}^{n-1} \sum_{\mu = 0}^{\rho - 2} N_{n-1}(\rho, \mu) w(\rho, \mu). \]

Table VIII summarizes the results of the different bounds discussed in this subsection. MB corresponds to the equivalent of the monotonicity upper bound, which is the value $2^{n-2}$ from [15]. ASPV corresponds to the average sphere packing value $2^{n-1}$ from Lemma 16. The third column is our upper bound results from Theorem 19. The column titled GPSB [15] is the exact value of $\tau^*(H(G_D, 1))$ from [13], which this linear programming problem was numerically solved in [15] for.
Theorem 19

26216

\[ \text{MB} \left( \mathcal{G}_G, 1 \right) = \frac{2^{n+1} - 2}{n}. \]

This bound is slightly improved in [9] by noticing that if there is a code with odd number of codewords, then there exists a code with one more codeword, and thus this upper bound becomes \( 2 \left[ \frac{2^{n+1} - 2}{n} \right] \).

The average sphere packing value in this case is calculated in the next lemma.

**Lemma 20.** The average sphere packing value of the graph \( \mathcal{G}_G \) for \( r = 1 \) is

\[
\text{ASPV} \left( \mathcal{G}_G, 1 \right) = \frac{2^{n+1}}{n+1}.
\]

**Proof:** The size of the radius one ball centered in \( x \in \{0, 1\}^n \) is \( \deg_{G, 1}(x) = \rho(x) \). Thus, the average size of a ball is

\[
\frac{1}{2^n} \sum_{x \in \{0, 1\}^n} \deg_{G, 1}(x) = \frac{1}{2^n} \sum_{\rho = 1}^{n} 2 \left( \frac{n-1}{\rho-1} \right) \rho = \frac{1}{2^n} \sum_{\rho = 0}^{n-1} \left( \frac{n-1}{\rho+1} \right) (\rho+1) = \frac{1}{2^n} (2^{n-1} + (n-1)2^{n-2}) = \frac{n+1}{2}.
\]

Thus, the average sphere packing value becomes

\[
\frac{2^n}{(n+1)/2} = \frac{2^{n+1}}{n+1}.
\]

Note that very similarly to the deletion channel, the fractional transversal given by the monotonicity property is suboptimal. We carry similar steps as in the previous subsection in order to give a better fractional transversal, stated in the next theorem.

**Theorem 21.** The vector \( w = (w_x)_{x \in \{0, 1\}^n} \) defined by

\[
w_x = \begin{cases} \frac{1}{\rho(x)} & \text{if } \mu(x) \leq 1 \\ \frac{1}{\rho(x)} \left( 1 - \frac{\mu(x)}{\rho(x)^2} \right) & \text{otherwise} \end{cases}
\]

is a fractional transversal.

**Proof:** Let \( x \) be a binary vector of length \( n \) with \( \rho \) runs and \( \mu \) middle-1-runs. We will show that \( \sum_{y \in B_{G, 1}(x)} w_y \geq 1 \). As in the proof of Theorem 18, it is possible to verify that this property holds for \( \rho = 1, 2, 3 \) and \( \mu = 0, 1 \), so we assume that \( \rho \geq 4 \) and \( \mu \geq 2 \).

If a vector \( y \in B_{G, 1}(x) \) is received by a single grain-error of a middle-1-run bit then \( \rho(y) = \rho - 2 \) and \( \mu(y) \leq \mu - 1 \).
Otherwise, \( \rho(y) = \rho \) and \( \mu(y) \leq \mu + 1 \) or, in case the last bit err, \( \rho(y) = \rho - 1 \) and \( \mu(y) \leq \mu \), however, the worst case is achieved for \( \rho(y) = \rho \) and \( \mu(y) = \mu + 1 \). Hence, we get
\[
\sum_{y \in B_{\rho,\mu}(x)} w_y \geq \frac{1}{\rho} \left( 1 - \frac{\mu}{\rho^2} \right) + \frac{\mu}{\rho - 2} \left( 1 - \frac{\mu - 1}{(\rho - 2)^2} \right) + \frac{\mu - 1}{\rho} \left( 1 - \frac{\mu + 1}{(\rho - 2)^2} \right)
\geq 1 + \frac{2\mu}{\rho(\rho - 2)} - \frac{\mu(\mu - 1)}{\rho^2} - \frac{\mu + 1}{\rho^2} + \frac{\mu(\mu + 1)}{\rho^3}.
\]

The rest of the proof is identical to the proof of Theorem 19.

Finally, we conclude with the following theorem.

**Theorem 22.** The value \( \tau^*(\mathcal{H}(G_C, 1)) \) satisfies
\[
\tau^*(\mathcal{H}(G_C, 1)) \leq 2 + \sum_{\rho = 2}^{\mu = 0} N_n(\rho, \mu) w(\rho, \mu).
\]

Table IX summarizes the improvements and results discussed in this section on the cardinalities of single-grain error-correcting codes. In the last column we gave the cardinalities of the best known to us codes taken from [9], [19], and [20]. For \( 5 \leq n \leq 8 \) the lower bound coincides with the best known upper bound [19]. For \( n = 9, 10, 11 \) the best known upper bound is 88, 176, 352, respectively [20], while for \( n \geq 12 \) the best known upper bound is the monotonic upper bound from [9], [12].

| \( n \) | MB [9, 12] | ASPV | Theorem 22 | LB |
|---|---|---|---|---|
| 5 | 12 | 10 | 12 | 8 [19] |
| 6 | 20 | 18 | 20 | 16 [19] |
| 7 | 36 | 32 | 36 | 20 [19] |
| 8 | 62 | 56 | 60 | 44 [19] |
| 9 | 112 | 102 | 108 | 72 [19] |
| 10 | 204 | 186 | 196 | 112 [20] |
| 11 | 372 | 341 | 358 | 210 [19] |
| 12 | 682 | 630 | 656 | 372 [20] |
| 13 | 1260 | 1170 | 1212 | 702 [19] |
| 14 | 2340 | 2184 | 2250 | 1272 [20] |
| 15 | 4368 | 4096 | 4202 | 2400 [19] |
| 16 | 8190 | 7710 | 7882 | 4522 [20] |
| 17 | 15420 | 14563 | 14844 | 8428 [20] |
| 18 | 29126 | 27594 | 28058 | 15348 [19] |
| 19 | 55188 | 52428 | 53202 | 27596 [20] |
| 20 | 104856 | 99964 | 101612 | 52432 [20] |
| 21 | 199728 | 190650 | 192850 | 99880 [20] |
| 22 | 381300 | 364722 | 368478 | 190652 [19] |
| 23 | 729444 | 690050 | 705510 | 364724 [19] |

### Table IX

**Grain-Error Channel Comparison**

**VII. Projective Spaces**

In this section, we explain an example where there is no monotonicity property, yet we benefit from the graph automorphisms and we simplify the linear programming again.

Koetter and Kschischang [14] modeled codes as subsets of projective space \( \mathbb{P}^n_q \), the set of linear subspaces of \( \mathbb{F}_q^n \), or of Grassmann space \( \mathcal{G}(n, k) \), the subset of linear subspaces of \( \mathbb{F}_q^n \) having dimension \( k \). Subsets of \( \mathbb{P}^n_q \) are called **projective codes** and similar to previous sections, it is desired to select elements with large distance from each other.

Let us first introduce the graph \( \mathcal{G}_P = (X_P, E_P) \) for projective codes, where \( X_P \) is the set of all linear subspaces in \( \mathbb{F}_q^n \) and \( E_P = \{ \{x, y\} : x \subset y \text{ or } y \subset x \, \text{and} \ |\dim(x) - \dim(y)| = 1 \} \), and using the path distance \( d_P(x, y) \) defined on graph \( \mathcal{G}_P \) we define
\[
B_{\mathcal{G}_P}(x) = \{ y \in X_P : d_P(x, y) \leq r \}.
\]

The corresponding hypergraph is \( \mathcal{H}(\mathcal{G}_P, r) = (X_{\mathcal{G}_P}, E_{\mathcal{G}_P}) \), such that \( X_{\mathcal{G}_P} = X_P \) and \( E_{\mathcal{G}_P} = \{ B_{\mathcal{G}_P}(x) : x \in X_P \} \). The generalized sphere packing bound becomes
\[
\tau^*(\mathcal{H}(\mathcal{G}_P, r)) = \min \left\{ \sum_{y \in B_{\mathcal{G}_P}(x)} w(x) : x \in X_P, \sum_{y \in B_{\mathcal{G}_P}(x)} w(x) \geq 1, w(x) \geq 0 \right\}.
\]

Assume \( x_1 \) and \( x_2 \) are elements in \( X_P \) with same dimension \( k \). There exist an injective linear transform \( T : \mathbb{P}^n_q \to \mathbb{P}^n_q \) mapping the basis of \( x_1 \) into a basis for \( x_2 \). Note that \( x \subset y \) if and only if \( T(x) \subset T(y) \). Hence, all such linear transforms are automorphisms on \( \mathcal{G}_P \), which means for any \( x_1, x_2 \in X_P \) of the same dimensions, there exist an automorphism mapping between them. Therefore, they lie in a same equivalence class. So we assign a same transversal weight to all the subspaces with the same dimension. We also need to find the size and the distribution of elements in \( B_{\mathcal{G}_P}(x) \). The general formula is given in [7] but we only study the case \( r = 1 \). Given \( x \) with dimension \( k \) in \( X_P \), there are \( \binom{k}{1} = 2^k - 1 \) subspaces of dimension \( k - 1 \) in \( B_{\mathcal{G}_P}(x) \), where
\[
\left[ \begin{array}{c} n \\ m \end{array} \right] = \frac{(2^n - 1)(2^n - 2 - 1) \cdots (2^n - m - 1)}{m!} = \frac{(2^n - 1)(2^n - 2 - 1) \cdots (2^n - m - 1)}{m!}
\]

is the number of subspaces of dimension \( m \) in a space of dimension \( n \). There are also \( 2^n - 2^k = 2^{n-k} - 1 \) subspaces of dimension \( k + 1 \) in \( B_{\mathcal{G}_P}(x) \) that include \( x \). Therefore, there are \( (2^n - 1) + (2^n - k - 1) \) elements in \( B_{\mathcal{G}_P}(x) \).

So,
\[
\tau^*(\mathcal{H}(\mathcal{G}_P, r)) = \min \left\{ \sum_{k=0}^{n} w_k \left[ \begin{array}{c} n \\ k \end{array} \right] : 0 \leq k \leq n \right\}
\]

\[
\tau^*(\mathcal{H}(\mathcal{G}_P, r)) = \left\{ \sum_{k=0}^{n} w_k \left[ \begin{array}{c} n \\ k \end{array} \right] : 0 \leq k \leq n \right\}
\]

is not surprising since \( \left[ \begin{array}{c} n \\ k \end{array} \right] \) is the largest coefficient in the cost function. This leads us to a greedy approach of starting from the middle, which has the highest impact on cost function; minimizing it, i.e. \( \left[ \begin{array}{c} n \\ k \end{array} \right] = 0 \); and then moving toward the tails where we pick the least possible value to satisfy the constraints. We call it as the **greedy** weight assignment, which is expressed as

\[
\]
It is clear that the greedy output has the transversal property well using the scheme presented in the paper.

Each error channel, still remains an interesting open problem.

In order to find the generalized sphere packing bound for each error channel, we didn’t completely solve the other setups studied here, namely the limited magnitude, deletion, and grain-error channels, we didn’t completely solve many cases is not easy to find. We found this solution for the Z channel and projective spaces in case of radius one. In the other setups studied here, namely the limited magnitude, deletion, and grain-error channels, we didn’t completely solve the linear programming problem but found a corresponding upper bound, which is a valid upper bound on the codes cardinalities in each case. Thus, solving the linear programming, in order to find the generalized sphere packing bound for each error channel, still remains an interesting open problem. We also mention that other error channels can be studied as well using the scheme presented in the paper.

Lastly, we follow up on the question we asked in the Introduction about the validity of the average sphere packing value. Even though in general it is not a valid upper bound, we believe that there are some conditions under which this value will hold as an upper bound, and finding these conditions remains as open problem.

**REFERENCES**

[1] C. Bachoc, F. Vallentin, and A. Passuello, “Bounds for projective codes from semidefinite programming,” *arXiv:1205.6406*, Apr. 2013.

[2] C. Berge, “Packing problems and hypergraph theory: a Survey,” *Annals of Discrete Mathematics*, vol. 4, pp. 3–37, 1979.

[3] M. Braun, T. Etzion, and A. Vardy, “Linearity and complements in projective space,” *Linear Algebra and its Applications*, vol. 438, pp. 57–70, Jan. 2013.

[4] S. Buzzaglo, E. Yaakobi, T. Etzion, and J. Bruck, “Error-correcting for multipermutations,” *IEEE Int. Symp. on Information Theory*, pp. 724–728, Istanbul, Turkey, July 2013.

[5] Y. Cassuto, M. Schwartz, V. Bohossian, and J. Bruck, “Codes for asymmetric limited-magnitude errors with application to multi-level flash memories,” *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1582–1595, April 2010.

[6] N. Elarief and B. Bose, “Optimal, systematic, q-ary codes correcting all asymmetric and symmetric errors of limited magnitude,” *IEEE Trans. Inform. Theory*, vol. 56, no. 3, pp. 979–983, March 2010.

[7] T. Etzion and A. Vardy, “Error-correcting codes in projective space,” *IEEE Trans. on Information Theory*, vol. 57, no. 2, pp. 1165–1173, Feb. 2011.

[8] R. Gabrys, E. Yaakobi, and L. Doleck, “Correcting grain-errors in magnetic media,” *IEEE Int. Symp. on Information Theory*, pp. 689–693, Istanbul, Turkey, July 2013.

[9] R. Gabrys, E. Yaakobi, and L. Doleck, “Correcting grain-errors in magnetic media,” submitted to *IEEE Trans. on Information Theory*, 2013.

[10] E.N. Gilbert, “A comparison of signalling alphabets,” *Bell System Technical Journal*, vol. 31, pp. 504–522, 1952.

[11] A.R. Iyengar, P.H. Siegel, and J.K. Wolf, “Write channel model for bit-patterned media recording,” *IEEE Transactions on Magnetics*, vol. 47, no. 1, pp. 35–45, Jan. 2011.

[12] N. Kasyap and G. Zémor, “Upper bounds on the size of grain-correcting codes,” *arXiv:1302.6154*, Jun. 2013.

[13] T. Kløve, B. Bose, and N. Elarief, “Systematic single limited magnitude asymmetric error correcting codes,” *Proc. IEEE Inform. Theory Workshop*, Cairo, Egypt, January 2010.

[14] R. Koetter and F.R. Kschischang, “Coding for errors and erasures in random network coding,” *IEEE Trans. on Information Theory*, vol. 54, no. 8, pp. 3579–3591, Aug. 2008.

[15] A. Kulkarni and N. Kiyavash, “Non-asymptotic upper bounds for deletion correcting codes,” *arXiv:1211.3289*, November 2012.

[16] A. Mazumdar, A. Barg, and N. Kasyap, “Coding for high-density recording on a 1-d granular magnetic medium,” *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7403–7417, Nov. 2011.

[17] M. Schwartz, “On the non-existence of lattice tilings by quasi-circles,” *European J. of Combinatorics*, vol. 36, pp. 130–142, Feb. 2014.

[18] F.F. Sellers, Jr, “Bit loss and gain correcting code,” *IRE Trans. on Information Theory*, vol. 8, no. 1, pp. 35–38, January 1962.

[19] A. Sharov and R.M. Roth, “Bounds and constructions for granular media coding,” *IEEE Int. Symp. on Information Theory*, pp. 2343–2347, St. Petersburg, Russia, August 2011.

[20] A. Sharov and R.M. Roth, “Improved bounds and constructions for granular media coding,” *Proc. Allerton Conf. Commun., Control and Comput.*, Allerton Retreat Center, Monticello, IL, 2013.

[21] J.M.G.M. Tolhuizen, “The generalized Gilbert-Varshamov bound is implied by Turán’s theorem,” *IEEE Trans. on Information Theory*, vol. 43, pp. 1605–1606, September 1997.

[22] R.R. Varshamov, “Estimate of the number of signals in error correcting codes,” *Dokl. Akad. Nauk SSSR*, vol. 117, pp. 739–741, 1957.

[23] R.R. Varshamov and G.M. Tenengolts, “Codes which correct single asymmetric errors (in Russian),” *Automatica i Telemekhanika*, vol. 26, no. 2, pp. 288–292, 1965. English translation in *Automation and Remote Control*, vol. 26, no. 2, pp. 286–290, 1965.

[24] J. Weber, C. De Vroedt, and D. Boekee, “Bounds and constructions for binary codes of length less than 24 and asymmetric distance less than 6,” *IEEE Trans. on Information Theory*, vol. 34, no. 5, pp. 1321–1331, September 1988.
Appendix A
Optimal Transversal Weight for Z Channel

In this section, we go over lemmas and proofs that has been used in section [IV].

Lemma 24. For all $r \in \mathbb{N}$, $|D_m| \leq (2r)^{m-r+1}$.

Proof: The proof is based on induction on $m$. Let us assume $|D_i| \leq (2r)^{i-r+1}$ for all $i \leq m - 1$. Therefore,

$$D_m = -r! \frac{D_{m-1}}{(r-1)!} + \frac{r!}{(r-2)!} \frac{D_{m-2}}{(r-2)!} + \cdots + \frac{r!}{0!} \frac{D_{m-r}}{(r-1)!}$$

$$\leq r! \frac{(2r)^{m-1}}{(r-1)!} + \frac{r!}{(r-2)!} (2r)^{m-2} + \cdots + \frac{r!}{0!} (2r)^{m-r+1}$$

$$= (2r)^{m-r+1} \left( \frac{r!}{(r-1)!} + \frac{r!}{(r-2)!} + \cdots + \frac{r!}{0!} \right)$$

$$\leq (2r)^{m-r+1} (2^{-1} + 2^{-2} + \cdots + 2^{-r})$$

$$\leq (2r)^{m-r+1}.$$

The idea behind the optimality proof is to write the function $f(w)$ as a non-negative linear combination of some other cost functions denoted by $f_i(w)$ and show $w^*$ is the feasible point which minimizes them all, and hence $w^*$ also minimizes the cost function and is the desired optimal solution.

Let us define cost functions $f_i(w)$ for all $0 \leq i \leq n$ as

$$f_k(w) = w_0 \quad k = 0,$n

$$f_k(w) = \sum_{i=0}^{r} w_{k+i} \binom{k+r}{r-i} \quad 1 \leq k \leq n - r,$n

$$f_k(w) = w_k \quad \forall n - r < k \leq n.$n

From [7], any feasible $w$ should satisfy

$$f_k(w) \geq \frac{1}{(k+r)!} \sum_{m=1}^{n} \frac{n!}{(n-m)!} D_{k-m+r-1} \quad \forall 1 \leq k \leq n - r,$n

$$f_k(w) \geq 0 \quad \forall n - r < k \leq n;$$

And, $w = w^*$ gives us equalities in all of them. We are required to show $w^*$ also minimizes $f(w)$, where

$$f(w) = \sum_{i=0}^{n} w_i \binom{n}{i}.$$

In order to prove the optimality, we show $f(w)$ can be written as

$$f(w) = y_0 f_0(w) + y_1 f_1(w) + \cdots + y_n f_n(w),$$

where $y_i$’s are some non-negative constants. Hence, for any transversal weight $w$ we have

$$f(w) \geq y_0 f_0(w^*) + y_1 f_1(w^*) + \cdots + y_n f_n(w^*) = y_0 + y_1 + \cdots + y_{n-r} = f(w^*).$$

We first show the choice of $y$ is unique. Then the problem reduces to show the non-negativity of $y$. Note that the cost functions $f_i(w)$ are inner products of $w$ with some non-negative vectors $m_i$, i.e. $f_i(w) = \langle m_i, w \rangle$, where

$$m_{ij} = \begin{cases} 1, & \text{if } i = j = 0 \\ 1, & \text{if } i = j, \text{and } n - r < i \leq n \\ 0, & \text{otherwise.} \end{cases}$$

Now, if we form the $(n+1) \times (n+1)$ matrix $M$ with elements $\{m_{ij}\}$, the problem of finding $y$ will be equivalent to solving $M^T y = c$, where $c = \left( \binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n} \right)$. $M$ is an upper triangular matrix with non-zero elements on the main diagonal, and hence $M$ is non-singular and the solution is unique. Since $M$ is upper triangular, $M^T y = c$ gives us the following recursions on $y_i$’s:

$$y_0 = 0,$n

$$y_i = \left( \frac{1}{r+1} \right) \binom{n}{i} - \sum_{j=1}^{i-1} \binom{r+j}{i} y_j \quad (19)$$

$$\forall i : 1 \leq i \leq n - r,$n

and

$$y_i = \binom{n}{i} - \sum_{j=1}^{n-r} \binom{r+j}{i} y_j \quad \forall i : n - r < i \leq n.$n

Lemma 25 gives an explicit formula for $y_k$ when $0 \leq k \leq n - r$, which again uses the sequence $D_i$ defined in [10]. We put the proof at the end.

Lemma 25. If $y$ is the solution to $M^T y = c$, then

$$y_0 = 1,$n

$$y_k = \frac{r!}{(k+r)!} \sum_{m=1}^{n} \frac{n!}{(n-m)!} D_{k-m+r-1} \quad \forall 1 \leq k \leq n - r.$n

Let us benefit from a simple change of variables to get

$$z_k := y_{n-k} \frac{(n-k+r)!}{r!} = \sum_{m=0}^{n-k-1} \frac{D_{n-k-m+r-1}}{(n-m)!}$$

$$= \sum_{m=0}^{n-k-1} \frac{D_{m+r-1}}{(m+k)!} \frac{1}{k!} + \sum_{m=0}^{n-k-1} \frac{D_{m+r-1}}{(m+k)!}.$$n

which is proven to be positive if $k \geq 4r - 1$ by using the same argument as $w_k^* > 0$ when $k \geq 3r - 1$ in [11]. Also, we can
verify $z_k \geq 0$ for the values of $r \leq k \leq 4r - 2$ by finding some numbers $n'_k$ such that
\[
\sum_{m=k}^{n'_k} \frac{D_{m+r-k-1}}{m!} \geq \frac{1}{(2r)^k} \left( e^{2r} - \sum_{m=0}^{n'_k} \frac{(2r)^m}{m!} \right),
\]
and we get $z_k > 0$ for all $n > n_k$ because
\[
z_k = \sum_{m=k}^{n-1} \frac{D_{m+r-k-1}}{m!} = \sum_{m=k}^{n_k} \frac{D_{m+r-k-1}}{m!} + \sum_{m=n_k+1}^{n-1} \frac{D_{m+r-k-1}}{m!} \\
\geq \sum_{m=k}^{n_k} \frac{D_{m+r-k-1}}{m!} - \frac{1}{(2r)^k} \sum_{m=n_k+1}^{\infty} \frac{(2r)^m}{m!} \\
= \sum_{m=k}^{n_k} \frac{D_{m+r-k-1}}{m!} - \frac{e^{2r} - \sum_{m=0}^{n_k} \frac{(2r)^m}{m!}}{(2r)^k} \geq 0.
\]

Finally, we check the values of $z_k$ for the finite set of $k < 4r - 1$ and $n \leq n_k$.

In order to show $y_k \geq 0$ for $k > n - r$, we rewrite the expression for $y_k$ when $k > n - r$ as
\[
y_k = \binom{n}{k} - \sum_{j=\max\{k-r,1\}}^{n-r} \binom{r+j}{k} y_j = \binom{n}{k} - y_{n-r} \binom{n}{k} - y_{n-r-1} \binom{n-1}{k} - \cdots - y_{n-r-p} \binom{n-p}{k},
\]
where $n - r - p = \max\{k-r,1\}$. So, $y_k \geq 0$ is equivalent to
\[
y_{n-r} \leq 1 - \frac{n-1}{n} y_{n-r-1} - \frac{n-2}{n} y_{n-r-2} - \cdots - \frac{n-p}{n} y_{n-r-p}.
\]
By assuming $y_i \geq 0$ for all $i \leq n - r$ and the recursive formula for $y_k$'s, we have
\[
y_{n-r} = 1 - \frac{n-1}{n} y_{n-r-1} - \frac{n-2}{n} y_{n-r-2} - \cdots - \frac{n-p}{n} y_{n-r-p} \\
\leq 1 - \frac{n-1}{n} y_{n-r-1} - \frac{n-2}{n} y_{n-r-2} - \cdots - \frac{n-p}{n} y_{n-r-p},
\]
where $n - r - q = \max\{n-2r,1\} \leq \max\{k-r,1\} = n - r - p$. Now, it suffices to show
\[
\frac{n-k}{n} \leq \frac{n-r}{n} \implies (n-k)! \leq (n-r)! (r-k)! \\
\iff (n-k)(n-k-1) \cdots (n-k-r) \leq (r-k)(r-k-1) \cdots (r-k),
\]
which always holds since $n-k < r$. In short, for any fixed radius, one can verify the optimality in a very same fashion as feasibility by just proving $z_k \geq 0$ for all $r \leq k \leq 4r - 2$ and the non-negativity of the remaining $y_k$'s follows immediately. Doing so, we proved the optimality of our transversal weight assignment for all $r \leq 20$.

**Proof of Lemma**\[25\] We define sequences $\{y'_i(n)\}$ for $1 \leq j \leq n - r - 1$ as
\[
y'_i(n) = \begin{cases} 0 & 1 \leq i < j, \\ n!r! & i = j, \\ \frac{(n-i)!}{(n-r)!} \sum_{\ell=\max\{i-r,1\}}^{j} \binom{r}{i} y'_j(n) & j < i \leq n - r, \end{cases}
\]
And, define $y'_i(n)$ as $y'_i(n) = \sum_{j=1}^{n-r} y'_i(n)$. It is easy to see
\[
y_i(n) = \frac{1}{r!} \binom{n}{i} \left( \frac{n-i}{r} \sum_{\ell=\max\{i-r,1\}}^{j} \binom{r}{i} y'_j(n) \right)
\]
forall $i : 1 \leq i \leq n - r$.

So, $y'_i$ satisfies the same recursive relation as \[19\] and hence it is nothing but $y_i$. On the other hand, using the change of variables $\delta'_i(n) = y'_i(n)(r+i)\binom{n-i}{nr}$ gives us
\[
\delta'_i(n) = 0, \quad 1 \leq i < j, \\
\delta'_i(n) = 1, \quad i = j, \\
\delta'_i(n) + \frac{\delta'_{i-1}(n)}{r} + \cdots + \frac{\delta'_i(n)}{(n-r)!} = 0, \quad j < i \leq n - r.
\]
So, $\delta'_i(n)$ is nothing but $D_{i+j+r-1}$ defined in \[10\]. Substituting this value in the formula for $y'_i$ gives
\[
y_i = \sum_{j=1}^{n-r} y'_i(n) = \sum_{j=1}^{n-r} \frac{n!}{(r+i)! (n-j)!} \left\{ \frac{D_{i+j+r-1}}{(r+i)!} \right\}
\leq \frac{r!}{(r+i)!} \sum_{j=1}^{n-r} \frac{n!}{(n-j)!} D_{i+j+r-1}
\leq \frac{r!}{(r+i)!} \sum_{j=1}^{n-r} \frac{n!}{(n-j)!} D_{i+j+r-1}.
\]

**APPENDIX B**

**OPTIMAL TRANSVERSAL WEIGHT FOR PROJECTIVE CODES**

The idea behind the proof is to again write the cost function $\sum_{k=0}^{n} w_k n^k/2^k$ as a non-negative linear combination of some other cost functions $f_k(w)$, where $w^*$ minimizes them all and so minimizes the cost function. We strongly benefit from the symmetry in the optimal transversal i.e. $w^*_k = w^*_{n-k}$ and we discuss the proof for all indices $k \leq \lfloor \frac{n}{2} \rfloor$ without loss of generality. Let us define the partial cost functions $f_k(w)$ as
\[
f_k(w) = \begin{cases} w_{k+1} + (2^{k+1} - 1)w_k + (2^n - k - 1)w_{k+2} & \text{if } k \equiv \lfloor n/2 \rfloor \mod 2 \\ w_k & \text{otherwise}, \end{cases}
\]
with the only exception of
\[
w_{k+1} + (2^{k+1} - 1)w_k + (2^{n-k-1} - 1)w_{k+2} & \text{if } k \equiv \lfloor n/2 \rfloor \mod 2 \\ w_0 + w_1(2^{n-1}) & \text{if } k = 0 \text{ and } \lfloor \frac{n}{2} \rfloor \neq 1 \text{ or } 2 \mod 4
\]

The idea is to write \( f(w) = \sum_{k=0}^{n_2} w_{k[k]} \) as \( \sum_{k=0}^{n_2} y_k f_k(w) \), where \( y_k \)'s are some fixed non-negative real numbers. Note that \( w^* \) is the minimizer of all these cost functions in the feasible set and non-negativity of \( y_k \)'s automatically proves the optimality of \( w^* \) for \( f(w) \).

Given arbitrary \( k \) such that \( k \equiv \left\lfloor \frac{n}{2} \right\rfloor + 1 \mod 4 \) if none of the indices \( \{k, k-1, k-2, k-3\} \) fall into the two exceptional categories in (20), we have,

\[
\begin{align*}
  f_k(w) &= w_k, \\
  f_{k-1}(w) &= w_{k-1}, \\
  f_{k-2}(w) &= w_{k-2}(2^{k-1} - 1) + w_{k-1} + w_k(2^{n-k+1} - 1), \\
  f_{k-3}(w) &= w_{k-3}(2^{k-2} - 1) + w_{k-2} + w_{k-1}(2^{n-k+2} - 1).
\end{align*}
\]

Furthermore, \( \{w_k, w_{k-1}, w_{k-2}, w_{k-3}\} \) show up only in these \( f_i(w) \)'s. So, by comparing it to the corresponding coefficients in \( f_i(w) \), we must have

\[
\begin{align*}
  \left[ \begin{array}{c}
  n \\
  k \\
  \end{array} \right]_2 &= y_k + y_{k-2}(2^{n-k+1} - 1) \\
  \left[ \begin{array}{c}
  n \\
  k-1 \\
  \end{array} \right]_2 &= y_{k-1} + y_{k-2} + y_{k-3}(2^{n-k+2} - 1) \\
  \left[ \begin{array}{c}
  n \\
  k-2 \\
  \end{array} \right]_2 &= y_{k-2}(2^{k-1} - 1) + y_{k-3} \\
  \left[ \begin{array}{c}
  n \\
  k-3 \\
  \end{array} \right]_2 &= y_{k-3}(2^{k-2} - 1).
\end{align*}
\]

By solving the system of equations above, we get

\[
\begin{align*}
  y_{k-3} &= \left[ \begin{array}{c}
  n \\
  k-3 \\
  \end{array} \right]_2 / (2^{k-2} - 1) \geq 0 \\
  y_{k-2} &= \left( \left[ \begin{array}{c}
  n \\
  k-2 \\
  \end{array} \right]_2 - y_{k-3} \right) / (2^{k-1} - 2) \geq 0 \\
  y_{k-1} &= \left[ \begin{array}{c}
  n \\
  k-1 \\
  \end{array} \right]_2 - y_{k-2} - y_{k-3}(2^{n-k+2} - 1) \geq 0 \\
  y_k &= \left[ \begin{array}{c}
  n \\
  k \\
  \end{array} \right]_2 - y_{k-2}(2^{n-k+1} - 1) \geq 0.
\end{align*}
\]

Story is not much different on the edges i.e. \( y_0 \) and \( y_{\left\lfloor \frac{n}{2} \right\rfloor} \) and we can prove the non-negativity of the first half of the \( y_k \)'s, which is followed by the non-negativity of the other half due to the symmetry. So \( y_k \)'s are non-negative and \( w^* \) is the optimal transversal weight.