Generalized Sphere Packing Bound

Arman Fazeli, *Student Member, IEEE*, Alexander Vardy, *Fellow, IEEE*, and Eitan Yaakobi, *Member, IEEE*

***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 nonbinary 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 have a 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.**

***Index Terms*— Sphere packing bound, linear programming, asymmetric errors, *Z* channel, deletion channel, subspace codes, grain errors, fractional transversals.**

I. Introduction

ONE of the basic and fundamental results in coding theory asserts that an upper bound on a length-n binary code $\mathcal{C}$ that can correct r errors is

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

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 $B_r(x) \triangleq \{y \in X \mid d(x, y) \leqslant r\}$ then for all $x \in X$, $|B_r(x)| = \Delta_r$ for some fixed value Δ_r. Then, the resulting sphere packing bound on an r-error-correcting code $\mathcal{C} \subseteq X$ becomes $|X|/\Delta_r$. However, ***what happens if the size of all balls is not the same?*** Clearly, a naive solution is to use Δ_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 [11], [24]. This bound states that if the size of all balls of radius r is the same, Δ_r, then a lower bound on a code $\mathcal{C} \subseteq X$ with minimum distance $r+1$ becomes $|X|/\Delta_r$. In [23], a similar study was carried out 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 same expression on a lower bound of a code still holds, with the modification of using the average size of the balls. That is, if $\overline{\Delta}_r \triangleq (\sum_{x \in X} |B_r(x)|)/|X|$, then a generalized Gilbert-Varshamov bound asserts that there exists a code with minimum distance $r + 1$ and of size at least $|X|/\overline{\Delta}_r$. Thus, an immediate question to ask is whether the same analogy holds for the sphere packing bound: ***Is*** $|X|/\overline{\Delta}_r$ ***an upper bound on the cardinality of an*** r***-error-correcting code*** $\mathcal{C} \subseteq X$***?*** 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.

The deletion channel [19] is one of the examples where the balls can have different sizes. Recently, in [16], 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 [16] 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 [10] and [13] and similar bounds for multipermutations codes with the Kendall's τ distance were derived in [4].

This paper has two main goals. First, we study the method studied for the deletion channel by Kulkarni and Kiyavash [16]

Manuscript received February 13, 2014; revised March 1, 2015; accepted March 6, 2015. Date of publication March 16, 2015; date of current version April 17, 2015. This paper was presented at the 2014 IEEE International Symposium on Information Theory [8] and [9].

A. Fazeli is with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: afazeli@ucsd.edu).

A. Vardy is with the Department of Electrical and Computer Engineering, the Department of Computer Science Engineering, and the Department of Mathematics, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: avardy@ucsd.edu).

E. Yaakobi is with the Department of Computer Science, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: yaakobi@cs.technion.ac.il).

Communicated by N. Kashyap, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2015.2413418

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), the deletion channel, the 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 obtain with the state-of-the-art ones.

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

$$\tau^*(\mathcal{H}) = \min \left\{ \sum_{i=1}^{n} w_i : A^T \cdot \boldsymbol{w} \geqslant \mathbf{1}, \boldsymbol{w} \in \mathbb{R}_+^n \right\}. \quad (1)$$

Let $\mathcal{G} = (X, E)$ be a directed graph which describes an error channel. The vertex set X is the set of all possible transmitted words, and the edges set E consists of all pairs of vertices at distance one, which correspond to a single error in the channel. The distance between $x, y \in X$, is the path metric in $\mathcal{G}$ 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 an r-error-correcting code in $\mathcal{G}$ of length n is denoted by $A_{\mathcal{G}}(n, r)$. Given some positive integer r, the graph $\mathcal{G}$ is associated with a hypergraph $\mathcal{H}(\mathcal{G}, r) = (X_r, \mathcal{E}_r)$ where $X_r = X$ and $\mathcal{E}_r = \{B_r(x) \mid x \in X\}$. Observing that every r-error-correcting code $\mathcal{C} \subseteq X$ is a matching in $\mathcal{H}(\mathcal{G}, r)$ (which is a collection of pairwise disjoint edges), the following upper bound on $A_{\mathcal{G}}(n, r)$ was verified in [16],

$$A_{\mathcal{G}}(n, r) \leqslant \tau^*(\mathcal{H}(\mathcal{G}, r)). \quad (2)$$

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

The expression $\tau^*(\mathcal{H}(\mathcal{G}, r))$ provides an explicit upper bound on $A_{\mathcal{G}}(n, r)$. 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 aspire to find this exact value, but if this is not possible to accomplish, it is still valuable to give an upper bound on $\tau^*(\mathcal{H}(\mathcal{G}, r))$, which is an upper bound on $A_{\mathcal{G}}(n, r)$ 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, where the weight of a fractional transversal is the sum of its entries. In fact, all the upper bound results presented in [4], [10], [13], and [16] follow this approach and an upper bound on the value $\tau^*(\mathcal{H}(\mathcal{G}, r))$ 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 of the generalized sphere packing bound. In particular, we show upper and lower bounds on its value and prove that if the graph $\mathcal{G}$ 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 $\mathcal{G}$ which states that a graph is monotone if for all r and two vertices x and y, if $y \in B_r(x)$ then $\deg_r(y) \leqslant \deg_r(x)$. 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 [10], [13], [16]. 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 out 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 [16], 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 [10] and [13] 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 [16].

Let $\mathcal{H} = (X, \mathcal{E})$ be a hypergraph where $X = \{x_1, \dots, x_n\}$, $\mathcal{E} = \{E_1, \dots, E_m\}$ and A its $n \times m$ incidence matrix. A ***matching*** in $\mathcal{H}$ is a collection of pairwise disjoint hyperedges and the ***matching number*** of $\mathcal{H}$, denoted by $\nu(\mathcal{H})$, 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

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

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

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

These two problems satisfy weak duality and thus $\nu(\mathcal{H}) \leqslant \tau(\mathcal{H})$. The relaxation of these integer linear programmings allows the variables $\boldsymbol{z}$ and $\boldsymbol{w}$ to take values in $\mathbb{R}_+$, which are not necessarily integers. The value of this linear programming relaxation for the matching number is denoted by

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

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 [16]

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

and in particular, for any fractional transversal $\boldsymbol{w}$,

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

Lastly, we mention here that we will usually denote the fractional transversal by $\boldsymbol{w} = (w_1, \dots, w_n)$, 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 at 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 $B_r^{\text{out}}(x), B_r^{\text{in}}(x)$ be the sets $B_r^{\text{out}}(x) = \{y \in X \mid d(x, y) \leqslant r\}$ and $B_r^{\text{in}}(x) = \{y \in X \mid d(y, x) \leqslant r\}$. The out-degree of x is $\deg_r^{\text{out}}(x) = |B_r^{\text{out}}(x)|$ and the in-degree is $\deg_r^{\text{in}}(x) = |B_r^{\text{in}}(x)|$. The definitions of $B_r^{\text{out}}(x)$ and $\deg_r^{\text{out}}(x)$ 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 $\mathcal{C} \subseteq X$ in this graph is said to be an r-error-correcting code if for all for all $x, y \in \mathcal{C}$, $B_r(x) \cap B_r(y) = \emptyset$. We let $A_{\mathcal{G}}(n, r)$ be the largest cardinality of an r-error-correcting code in $\mathcal{G}$ of length n. If for every $r \geqslant 0$, there exists some fixed Δ_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, $\mathcal{H}(\mathcal{G}, r) = (X_r, \mathcal{E}_r)$ is a hypergraph associated with $\mathcal{G}$ such that $X_r = X$ and $\mathcal{E}_r = \{B_r(x) : x \in X\}$. As was stated in (2), the value $\tau^*(\mathcal{H}(\mathcal{G}, r))$ is an upper bound on $A_{\mathcal{G}}(n, r)$ 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

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

In [23], 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 $A_{\mathcal{G}}(n, r)$ is given by

$$\frac{|X|}{\overline{\Delta}_{2r}} \leqslant A_{\mathcal{G}}(n, r).$$

In the Introduction we asked about the analogy of the last bound to the sphere packing bound. Namely, does the following inequality hold

$$A_{\mathcal{G}}(n, r) \leqslant \frac{|X|}{\overline{\Delta}_r}?$$

We call the value $\frac{|X|}{\overline{\Delta}_r}$ the ***average sphere packing value*** and denote it by $ASPV(\mathcal{G}, r)$. 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 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

$$E_Z = \{(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n, \boldsymbol{x} \geqslant \boldsymbol{y}, w_H(\boldsymbol{x}) = w_H(\boldsymbol{y}) + 1\},$$

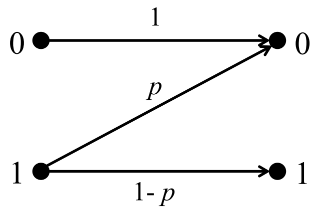


Fig. 1. The Z-channel.

and $w_H(\boldsymbol{x})$ denotes the Hamming weight of $\boldsymbol{x}$. Let r be some fixed positive integer. For every $\boldsymbol{x} \in \{0,1\}^n$,

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

and $\deg_{Z,r}(\boldsymbol{x}) = \sum_{i=0}^{r} \binom{w_H(\boldsymbol{x})}{i}$.

The corresponding hypergraph is $\mathcal{H}(\mathcal{G}_Z, r) = (X_{Z,r}, \mathcal{E}_{Z,r})$, such that $X_{Z,r} = \{0,1\}^n$ and $\mathcal{E}_{Z,r} = \{B_{Z,r}(\boldsymbol{x}) \ : \ \boldsymbol{x} \in \{0,1\}^n\}$. The generalized sphere packing bound becomes

$$\tau^*(\mathcal{H}(\mathcal{G}_Z, r)) = \min\left\{ \sum_{\boldsymbol{x}\in\{0,1\}^n} w_{\boldsymbol{x}} : \forall \boldsymbol{x} \in \{0,1\}^n, \sum_{\boldsymbol{y}\in B_{Z,r}(\boldsymbol{x})} w_{\boldsymbol{y}} \geqslant 1, w_{\boldsymbol{x}} \geqslant 0 \right\}. \quad (3)$$

The average size of a ball with radius r is

$$\overline{\Delta}_{Z,r} = \frac{1}{2^n} \sum_{\boldsymbol{x}\in\{0,1\}^n} \sum_{i=0}^{r} \binom{w_H(\boldsymbol{x})}{i} = \frac{1}{2^n} \sum_{w=0}^{n} \binom{n}{w} \sum_{i=0}^{r} \binom{w}{i} = \frac{1}{2^n} \sum_{i=0}^{r} \sum_{w=0}^{n} \binom{n}{w}\binom{w}{i}.$$

For $0 \leqslant i \leqslant r$, $\sum_{w=0}^{n} \binom{n}{w}\binom{w}{i} = \binom{n}{i} 2^{n-i}$ and thus we get

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

Therefore, the average sphere packing value in this case becomes

$$ASPV(\mathcal{G}_Z, r) = \frac{2^n}{\overline{\Delta}_{Z,r}} = \frac{2^n}{\sum_{i=0}^{r} \frac{\binom{n}{i}}{2^i}}.$$

In particular, for $r = 1$ we get

$$ASPV(\mathcal{G}_Z, 1) = \frac{2^n}{\overline{\Delta}_{Z,1}} = \frac{2^n}{1 + n/2} = \frac{2^{n+1}}{n+2}.$$

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. □

Even though the generalized sphere packing bound $\tau^*(\mathcal{H}(\mathcal{G}, r))$ 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 $\tau^*(\mathcal{H}(\mathcal{G}, r))$ for different graphs $\mathcal{G}$. 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 aspire 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 $\tau^*(\mathcal{H}(\mathcal{G}, r))$ 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 $\tau^*(\mathcal{H}(\mathcal{G}, r))$. As specified in Section II, we assume throughout this section that the error channel is depicted by some directed graph $\mathcal{G} = (X, E)$ and for a fixed integer $r \geqslant 1$, $\mathcal{H}(\mathcal{G}, r) = (X_r, \mathcal{E}_r)$ 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 $\tau^*(\mathcal{H}(\mathcal{G}, r))$. 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 $x \in X$, $\deg_r^{in}(x) \leqslant \Delta$, then

$$\tau^*(\mathcal{H}(\mathcal{G}, r)) \geqslant \frac{|X|}{\Delta}.$$

Proof: Since $\deg_r^{\text{in}}(x) \leqslant \Delta$, for all $x \in X$, the weight of every row of the incidence matrix A of $\mathcal{H}(\mathcal{G}, r)$ is at most Δ, that is, $\sum_{j=1}^{n} a_{i,j} \leqslant \Delta$ for all $1 \leqslant i \leqslant n$, where $n = |X|$. Let $\boldsymbol{w}$ be a fractional transversal in $\mathcal{H}(\mathcal{G}, r)$. Then, for every $1 \leqslant j \leqslant n$, $\sum_{i=1}^{n} a_{i,j} w_i \geqslant 1$, and thus

$$n \leqslant \sum_{j=1}^{n} \sum_{i=1}^{n} a_{i,j} w_i.$$

However, note that

$$n \leqslant \sum_{j=1}^{n} \sum_{i=1}^{n} a_{i,j} w_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} w_i = \sum_{i=1}^{n} w_i \sum_{j=1}^{n} a_{i,j} \leqslant \Delta \sum_{i=1}^{n} w_i,$$

and therefore

$$\sum_{i=1}^{n} w_i \geqslant \frac{n}{\Delta}.$$

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

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 $x \in X$, $\deg_r(x) \geqslant \Delta$, then

$$\tau^*(\mathcal{H}(\mathcal{G}, r)) \leqslant \frac{|X|}{\Delta}.$$

Proof: If $\deg_r(x) \geqslant \Delta$ for all $x \in X$ then the vector $\boldsymbol{w} = \mathbf{1}/\Delta$ is a fractional transversal and thus $\tau^*(\mathcal{H}(\mathcal{G}, r)) \leqslant |X|/\Delta$. ■

The last two lemmas readily imply the following.

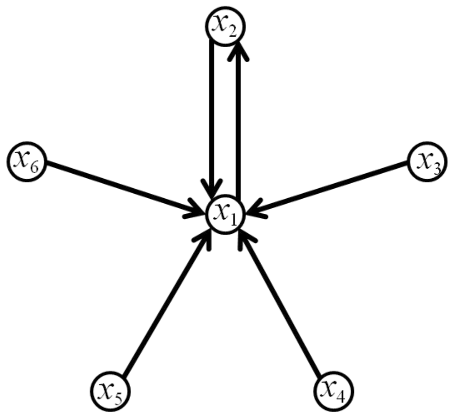


Fig. 2. The graph $\mathcal{G}_2$.

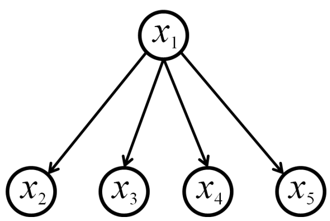


Fig. 3. The graph $\mathcal{G}_3$.

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

The next example shows that for a non-symmetric graph $\mathcal{G}$, the sphere packing bound and the generalized sphere packing bound need not be equal.

Example 2: In this example the graph $\mathcal{G}_2 = (X_2, E_2)$ has six vertices, so $X_2 = \{x_1, x_2, x_3, x_4, x_5, x_6\}$. For $2 \leqslant i \leqslant 6$, 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, $\deg_1(x_i) = 2$ for all $1 \leqslant i \leqslant 6$, so the graph $\mathcal{G}_2$ is regular and the sphere packing bound becomes $\frac{|X_2|}{2} = 3$. However, the vector $\boldsymbol{w} = (1, 0, 0, 0, 0, 0)$ is a fractional transversal, which is optimal, and thus the generalized sphere packing bound of $\mathcal{G}_2$ equals 1. □

In the next example, we show a graph for which the average sphere packing value does not hold as a valid bound. This provides a negative answer to the question we asked in the Introduction regarding the validity of the average sphere packing value as an upper bound.

Example 3: The graph $\mathcal{G}_3 = (X_3, E_3)$ 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 radius-one ball is $\frac{1 \cdot 5 + 4 \cdot 1}{5} = 9/5$ and thus the average sphere packing value becomes $\frac{5}{9/5} = 25/9$. However, the minimum distance of the code $\mathcal{C} = \{x_2, x_3, x_4, x_5\}$ in $\mathcal{G}_3$ is ∞, and in particular, it can be a single-error-correcting code, which contradicts the average sphere packing value. □

Example 3 depicts a directed, i.e. not symmetric, graph where the sphere packing value is not an upper bound to the code size. Next we show an example of a symmetric graph that does not satisfy the average sphere packing value either.

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

$$\begin{aligned}\overline{\Delta_1} &= \frac{k \cdot k + (n-k)(n-k+1)}{n} \\ &= n - 2\sqrt{n} + 3 - \frac{1}{\sqrt{n}} > n/2.\end{aligned}$$

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. □

Examples 3 and 4 prove that the average sphere packing value does not hold in all cases. In fact, from Example 4, we do not only conclude that it does not hold in general, but also that the ratio between this value and a size of a code can be arbitrarily small. However, it is still very interesting to find some minimal conditions such that this bound holds.

B. Monotonicity and Fractional Transversals

Remember that a vector $\boldsymbol{w}$ is a fractional transversal if $\boldsymbol{w} \geqslant \boldsymbol{0}$ and for $1 \leqslant i \leqslant n$,

$$\sum_{y \in B_r(x_i)} w_y \geqslant 1.$$

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

Lemma 4: The vector $\boldsymbol{w}$ given by

$$w_i = \frac{1}{\min_{x \in B_r^{in}(x_i)}\{\deg_r(x)\}},$$

for $1 \leqslant i \leqslant n$, is a fractional transversal.

Proof: It is easy to verify that $\boldsymbol{w} \geqslant \boldsymbol{0}$. For every $1 \leqslant i \leqslant n$, if $y \in B_r(x_i)$, then $x_i \in B_r^{\text{in}}(y)$ and thus

$$w_y = \frac{1}{\min_{x \in B_r^{\text{in}}(y)}\{\deg_r(x)\}} \geqslant \frac{1}{\deg_r(x_i)}.$$

Therefore, we get

$$\sum_{y \in B_r(x_i)} w_y \geqslant \sum_{y \in B_r(x_i)} \frac{1}{\deg_r(x_i)} = 1.$$

■

A graph $\mathcal{G}$ is said to satisfy the ***monotonicity property***, or $\mathcal{G}$ is ***monotone***, if for every $r \geqslant 1$, $x \in X$ and $y \in B_r(x)$,

$$\deg_r(y) \leqslant \deg_r(x).$$

In this case, the fractional transversal from Lemma 4 can be stated more explicitly.

Lemma 5: If $\mathcal{G}$ is monotone then the vector $\boldsymbol{w}$ given by

$$w_i = \frac{1}{\deg_r(x_i)},$$

for $1 \leqslant i \leqslant n$, is a fractional transversal.

Proof: If $\mathcal{G}$ is monotone then for every $x \in B_r^{\text{in}}(x_i)$, $\deg_r(x) \geqslant \deg_r(x_i)$. Therefore, the fractional transversal $\boldsymbol{w}$ from Lemma 4 simply becomes

$$w_i = \frac{1}{\deg_r(x_i)}.$$

■

As a result of Lemma 5, if $\mathcal{G}$ is monotone, then the following expression is an upper bound on $A_{\mathcal{G}}(n, r)$,

$$A_{\mathcal{G}}(n, r) \leqslant \sum_{i=1}^{n} w_i = \sum_{i=1}^{n} \frac{1}{\deg_r(x_i)}. \quad (4)$$

We call this bound the ***monotonicity upper bound***, which holds in case that $\mathcal{G}$ is monotone, and denote it by $MB(\mathcal{G}, r)$. 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 $\mathcal{G}_Z$ from Example 1 satisfies the monotonicity property since for every $\boldsymbol{x}, \boldsymbol{y} \in \{0, 1\}^n$, if $\boldsymbol{y} \in B_{Z,r}$ then $w_H(\boldsymbol{y}) \leqslant w_H(\boldsymbol{x})$. Thus, according to Lemma 5, the vector $\boldsymbol{w} = (w_{\boldsymbol{x}})_{\boldsymbol{x} \in \{0,1\}^n}$ given by

$$w_{\boldsymbol{x}} = \frac{1}{\deg_r(\boldsymbol{x})} = \frac{1}{\sum_{i=0}^{r} \binom{w_H(\boldsymbol{x})}{i}},$$

is a fractional transversal. Therefore, the monotonicity upper bound $MB(\mathcal{G}_Z, r)$ derived in (4) is calculated to be

$$\begin{aligned} MB(\mathcal{G}_Z, r) &= \sum_{\boldsymbol{x} \in \{0,1\}^n} w_{\boldsymbol{x}} = \sum_{\boldsymbol{x} \in \{0,1\}^n} \frac{1}{\sum_{i=0}^{r} \binom{w_H(\boldsymbol{x})}{i}} \\ &= \sum_{w=0}^{n} \binom{n}{w} \frac{1}{\sum_{i=0}^{r} \binom{w}{i}}. \end{aligned}$$

For example, for $r = 1$, we get

$$\begin{aligned} MB(\mathcal{G}_Z, 1) &= \sum_{w=0}^{n} \binom{n}{w} \frac{1}{\sum_{i=0}^{1} \binom{w}{i}} \\ &= \sum_{w=0}^{n} \binom{n}{w} \frac{1}{w+1} = \frac{2^{n+1} - 1}{n+1}. \end{aligned}$$

Note that the average sphere packing value, calculated in Example 1, for $r = 1$ is $\frac{2^{n+1}}{n+2}$, is smaller 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 $\boldsymbol{w}' = (w'_{\boldsymbol{x}})_{\boldsymbol{x} \in \{0,1\}^n}$, where

$$w'_{\boldsymbol{x}} = \frac{1}{w_H(\boldsymbol{x}) + 1} \cdot \frac{w_H(\boldsymbol{x}) + 2}{w_H(\boldsymbol{x}) + 3},$$

for $\boldsymbol{x} \neq \boldsymbol{0}$ and $w'_{\boldsymbol{0}} = 1$, is a fractional transversal. The corresponding bound for this fractional transversal becomes

$$2^{n+1} \cdot \frac{1}{n + 3 - \frac{2n+6}{n^2+3n+4}} \leqslant \frac{2^{n+1}}{n+2},$$

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 $\tau^*(\mathcal{H}(\mathcal{G}_Z, r))$ will be the topic and problem we solve in Section IV. □

The exact value $\tau^*(\mathcal{H}(\mathcal{G}_Z, r))$ will be obtained in Section IV.

C. Automorphisms on Graphs

One of the main obstacles in calculating the value of $\tau^*(\mathcal{H}(\mathcal{G}, r))$ 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 $\tau^*(\mathcal{H}(\mathcal{G}_Z, r))$, 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 $\tau^*(\mathcal{H}(\mathcal{G}, r))$. 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 $\pi : X \to X$ such that for all $(x, y) \in X \times X$, $(x, y) \in E$ if and only if $(\pi(x), \pi(y)) \in E$. Assume $|X| = n$, we let $\mathbb{S}_n$ be the set of all permutations of n elements. The set of all automorphisms of G is

$$Aut(G) = \{\pi \in \mathbb{S}_n \mid \pi \text{ is an automorphism of } G\}.$$

It is known that $Aut(G)$ is a subgroup of the symmetric group $\mathbb{S}_n$ 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 $1 \leqslant n(G) \leqslant n$ 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

$$\mathcal{W}_c = \left\{ \boldsymbol{w} \ : \boldsymbol{w} \text{ is a fractional transversal and } \sum_{i=1}^{n} w_i = c \right\}.$$

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

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

Lemma 6: Let $\boldsymbol{w}$ be a fractional transversal and π an automorphism. Then, the vector $\boldsymbol{w}^\pi$ is a fractional transversal as well.

Proof: It is clear to verify that $\boldsymbol{w}^\pi \geqslant \boldsymbol{0}$. We need to show that for all $1 \leqslant i \leqslant n$ the following inequality holds

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

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

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

where the last inequality holds since $\boldsymbol{w}$ 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 $\mathcal{W}_c \neq \emptyset$ then $\mathcal{W}_c$ contains an $\mathcal{X}(\mathcal{G})$-regular fractional transversal.

Proof: Let $\boldsymbol{w} \in \mathcal{W}_c$ be a fractional transversal. If $\boldsymbol{w}$ is $\mathcal{X}(\mathcal{G})$-regular then the property holds. Otherwise, let $\pi \in Aut(G)$ and $\boldsymbol{w}^\pi$ 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 6 we get that $\boldsymbol{w}^\pi \in \mathcal{W}_c$. Similarly, we can show that $\frac{\boldsymbol{w}+\boldsymbol{w}^\pi}{2} \in \mathcal{W}_c$. Let $\pi_1, \pi_2, \ldots, \pi_N$ be some order of the automorphisms in $Aut(\mathcal{G})$. We can similarly derive that the vector

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

belongs to $\mathcal{W}_c$ as well.

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

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

Now, let $\pi^* \in Aut(\mathcal{G})$ 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

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

■

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

$$A_H(i,j) = \frac{|\{(x,y) : x \in X_i, y \in B_r(x) \cap X_j\}|}{|X_i|}. \quad (5)$$

The next Corollary summarizes this discussion.

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

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

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

The condition $A^T \cdot \boldsymbol{w} \geqslant 1$ from (1) can be stated as for all $x \in X$, $\sum_{y \in B_r(x)} w_y \geqslant 1$. 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 $\sum_{j=1}^{n_H} A_H(i,j) w_j \geqslant 1$. Finally, since there are $|X_i|$ vectors which are assigned with weight w_i we get that the weight of this $\mathcal{X}_H(\mathcal{G})$-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 $\tau^*(\mathcal{H}(\mathcal{G}_Z, r))$ 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 $\mathcal{G}_Z$. For every $\sigma \in \mathbb{S}_n$, a permutation $\pi_\sigma : \{0,1\}^n \to \{0,1\}^n$ is defined such that for all $\boldsymbol{x} \in \{0,1\}^n$, $(\pi_\sigma(\boldsymbol{x}))_i = x_{\sigma(i)}$. It is possible to verify that the set $H = \{\pi_\sigma : \sigma \in \mathbb{S}_n\}$ is a subgroup of $Aut(\mathcal{G}_Z)$. Furthermore, the set $\{0,1\}^n$ is partitioned under H into $n+1$ equivalence classes $\mathcal{X}_H(\mathcal{G}_Z) = \{X_0, X_1, \ldots, X_n\}$, where $X_i = \{\boldsymbol{x} \in \{0,1\}^n : w_H(\boldsymbol{x}) = i\}$, for $0 \leqslant i \leqslant n$. Therefore, according to equation (6) in Corollary 8, it is enough to limit our search and find only fractional transversals $\boldsymbol{w}$ which are $\mathcal{X}_H(\mathcal{G}_Z)$-regular. Hence, the problem in (3) is simplified to be

$$\tau^*(\mathcal{H}(\mathcal{G}_Z, r)) = \min\left\{ \sum_{\ell=0}^{n} \binom{n}{\ell} w_\ell : \sum_{i=0}^{\min\{\ell, r\}} \binom{\ell}{i} w_{\ell-i} \geqslant 1, \right. \\ \left. w_\ell \geqslant 0, 0 \leqslant \ell \leqslant n \right\}. \quad (7)$$

□

In the next section we will continue with Example 6 and show exactly how to solve the problem stated in (7).

IV. The Z Channel

The Z channel was already discussed before in Examples 1, 5, and 6. We derived the linear programming problem to find the value $\tau^*(\mathcal{H}(\mathcal{G}_Z, r))$ in (3) and calculated its average sphere packing value. Then, we saw that $\mathcal{G}_Z$ 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 $\tau^*(\mathcal{H}(\mathcal{G}_Z, r))$ in (7).

The goal of this section is to solve the linear programming problem in (7) by finding the appropriate fractional transversal and prove that it gives the value of $\tau^*(\mathcal{H}(\mathcal{G}_Z, r))$. This result is proved in the next theorem.

Theorem 9: For all $r \leqslant 20$, the optimal fractional transversal which solves the linear programming in (7) is given by the following recursive formula

$$w_n^* = w_{n-1}^* = \cdots = w_{n-r+1}^* = 0, \tag{8}$$
$$w_0^* = 1,$$
and for all k: $1 \leqslant k \leqslant n-r$ we have
$$w_k^* = \left(1 - \sum_{i=1}^{r} w_{k+i}^* \binom{k+r}{r-i}\right) / \binom{k+r}{r}.$$

Soon, we will show the equivalent formula

$$w_0^* = 1, \text{ and for all } k \geqslant 1$$
$$w_k^* = r!k! \sum_{m=r+k}^{n} \frac{D_{m-k-1}}{m!}. \tag{9}$$

where D_i is given by another recursive relation independent from n:

$$D_0 = D_1 = \cdots = D_{r-2} = 0,$$
$$D_{r-1} = 1, \text{ and for all } i \geqslant r$$
$$\frac{D_i}{r!} + \frac{D_{i-1}}{(r-1)!} + \cdots + \frac{D_{i-r}}{0!} = 0. \tag{10}$$

We divide the proof into three parts. First, we show the equivalence of the two formulas above. Then, we show that $\mathbf{w}^*$ is in fact a transversal or in other words, it is in the feasibility region of the linear programming. Next, we discuss its optimality. Our method shows both feasibility and optimality for all $r \leqslant 20$ and we conjecture that $\mathbf{w}^*$ is the optimal transversal weight for all radii $r \in \mathbb{N}$. We note that the weight assignment in Theorem 9 could also be the optimal weight assignment for $r > 20$, but one needs to verify that for each r using the method presented in the following.

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. Lets define the sequence $\Delta_k(n)$ as $\Delta_k(n) = w_k^*(n) - w_k^*(n-1)$ for all n. So,

$$\Delta_k(n) = \frac{1}{\binom{n}{r}} \qquad \text{if } k = n-r,$$
$$\Delta_k(n) = 0 \qquad \text{if } k > n-r,$$
$$\Delta_k(n) = -\sum_{i=1}^{r} \Delta_{k+i}(n) \cdot \frac{\binom{k+r}{r-i}}{\binom{k+r}{r}} \quad \text{if } k < n-r.$$

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

$$D_0(n) = D_1(n) = \cdots = D_{r-2}(n) = 0,$$
$$D_{r-1}(n) = 1, \text{ and for all } i \geqslant r$$
$$\frac{D_i(n)}{r!} + \frac{D_{i-1}(n)}{(r-1)!} + \cdots + \frac{D_{i-r}(n)}{0!} = 0.$$

Note that the initialization and the recursion for sequence $D_i(n)$ are independent of n. So the sequence $D_i(n)$ is also independent of n and we drop n to write $w_k^*(n)$ as

$$\begin{aligned} w_k^*(n) &= \Delta_k(n) + \Delta_k(n-1) + \cdots + \Delta_k(k+r) \\ &= \frac{r!k!}{n!} D_{n-k-1} + \frac{r!k!}{(n-1)!} D_{n-k-2} + \cdots + \frac{r!k!}{(k+r)!} D_{r-1} \\ &= r!k! \sum_{m=r+k}^{n} \frac{D_{m-k-1}}{m!}. \end{aligned}$$

B. Transversal Property for $\mathbf{w}^*$

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

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

$$\begin{aligned} w_k^* &= \sum_{m=k+1}^{n} \frac{(-1)^{m-k-1} k!}{m!} \\ &= \left(\frac{1}{k+1} - \frac{1}{(k+1)(k+2)}\right) \\ &\quad + \left(\frac{1}{(k+1)(k+2)(k+3)} - \frac{1}{(k+1)(k+2)(k+3)(k+4)}\right) + \cdots > 0. \end{aligned}$$

In general, it is not easy to derive an explicit formula for $\mathbf{w}^*$ for $r \geqslant 2$. However, we show that D_m is bounded by an exponential function of $2r$ (see Appendix A) and hence, the first few terms in (9) are dominant compared to the rest and $w_k^* \geqslant 0$ is always the case. We divide the proof into two parts. First, we claim the positivity of w_k^* for $k \geqslant 3r-1$.

Theorem 10: Let $\mathbf{w}^*$ be the transversal weight assignment defined in (8). For all $k \geqslant 3r-1$ we have

$$w_k^* \geqslant \frac{r!k!}{(r+k)!} 2^{-(n-k-r)} > 0.$$

We refer to the Appendix A for the complete proof of the Theorem 10. 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 $\mathbf{w}^*$ for any fixed r in the following fashion:

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

$$\sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} \geqslant \frac{1}{(2r)^{r+k}} \left(e^{2r} - \sum_{m=0}^{n_k} \frac{(2r)^m}{m!}\right), \tag{11}$$

which implies the following for all $n > n_k$:

$$\begin{aligned} w_k^* &= r!k! \sum_{m=r+k}^{n} \frac{D_{m-k-1}}{m!} \\ &= r!k! \left(\sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} + \sum_{m=n_k+1}^{n} \frac{D_{m-k-1}}{m!} \right) \\ &\overset{(a)}{\geqslant} r!k! \left(\sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} - \sum_{m=n_k+1}^{n} \frac{(2r)^{m-k-r}}{m!} \right) \\ &> r!k! \left(\sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} - \frac{1}{(2r)^{k+r}} \sum_{m=n_k+1}^{\infty} \frac{(2r)^m}{m!} \right) \\ &= r!k! \left(\sum_{m=r+k}^{n_k} \frac{D_{m-k-1}}{m!} - \frac{e^{2r} - \sum_{m=0}^{n_k} \frac{(2r)^m}{m!}}{(2r)^{r+k}} \right) \geqslant 0, \end{aligned}$$

where (a) is verified by Lemma 25 in Appendix A.

Now, we are left with checking only the values of w_k^* for the finite set of $k < 3r - 1$ and $n \leqslant n_k$. Note that,

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

Also, D_i is bounded by an exponential function (see Lemma 25) 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 $\ell_k \geqslant \frac{\epsilon_k}{r!k!} > 0$. 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 $r \leqslant 20$. Numerical calculations also show that $n_k \leqslant 4r - 1$ for all $r \leqslant 20$. In Appendix A, we prove that $\mathbf{w}^*$ defined in (8), is also the optimal transversal assignment and gives us the best bound using this 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 *et al.* [26], 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 [26] involves solving an integer programming problem, and the authors of [26] have computed this bound only for $n \leqslant 23$. In contrast, our bound in Theorem 9 is easy to compute for all n, and we give its values for $r = 1, 2, 3, 4$ up to $n \leqslant 32$ in Tables I, II, III, and IV.

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

TABLE I

Z CHANNEL: UPPER BOUNDS COMPARISON FOR $r = 1$

n	MB	ASPV	GSPB	[26]
5	10	9	8	6
6	18	16	14	12
7	31	28	26	18
8	56	51	47	36
9	102	93	86	62
10	186	170	159	117
11	341	315	295	210
12	630	585	551	410
13	1170	1092	1032	786
14	2184	2048	1940	1500
15	4095	3855	3662	2828
16	7710	7281	6935	5430
17	14563	13797	13170	10374
18	27594	26214	25075	19898
19	52428	49932	47853	38008
20	99864	95325	91514	73174
21	190650	182361	175351	140798
22	364722	349525	336586	271953
23	699050	671088	647131	523586
24	1342177	1290555	1246069	?
25	2581110	2485513	2402690	?
26	4971026	4793490	4638907	?
27	9586980	9256395	8967211	?
28	18512790	17895697	17353537	?
29	35791394	34636833	33618332	?
30	69273666	67108864	65191862	?
31	134217727	130150524	126535913	?
32	260301048	252645135	245818070	?

TABLE II

Z CHANNEL: UPPER BOUNDS COMPARISON FOR $r = 2$

n	MB	ASPV	GSPB	[26]
5	7	5	4	2
6	12	8	6	4
7	19	13	9	4
8	31	21	16	7
9	51	35	27	12
10	84	59	46	18
11	140	101	79	32
12	238	174	138	63
13	407	303	243	114
14	703	532	432	218
15	1224	942	772	398
16	2151	1680	1388	739
17	3806	3013	2510	1279
18	6780	5433	4562	2380
19	12153	9845	8327	4242
20	21902	17924	15260	8069
21	39672	32768	28068	14374
22	72190	60133	51802	26679
23	131914	110740	95904	50200
24	241977	204600	178065	?
25	445447	379146	331499	?
26	822696	704555	618679	?
27	1524039	1312642	1157328	?
28	2831211	2451465	2169652	?
29	5273303	4588640	4075740	?
30	9845788	8607148	7670997	?
31	18424950	16176901	14463616	?
32	34553129	30460760	27317244	?

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

TABLE III

Z CHANNEL: UPPER BOUNDS COMPARISON FOR $r = 3$

n	MB	ASPV	GSPB	[26]
5	7	4	2	2
6	11	6	3	2
7	17	9	5	2
8	26	13	7	4
9	40	20	11	4
10	63	31	18	6
11	99	50	29	8
12	156	80	48	12
13	248	130	81	18
14	400	214	136	34
15	650	357	231	50
16	1066	601	395	90
17	1764	1020	682	168
18	2946	1744	1186	320
19	4960	3006	2076	616
20	8418	5216	3653	1144
21	14395	9108	6462	2134
22	24786	15993	11486	4116
23	42956	28232	20507	7346
24	74902	50081	36768	?
25	131345	89240	66176	?
26	231537	159687	119534	?
27	410164	286866	216639	?
28	729924	517216	393863	?
29	1304514	935722	718180	?
30	2340710	1698286	1313176	?
31	4215629	3091572	2407381	?
32	7618868	5643846	4424196	?

TABLE IV

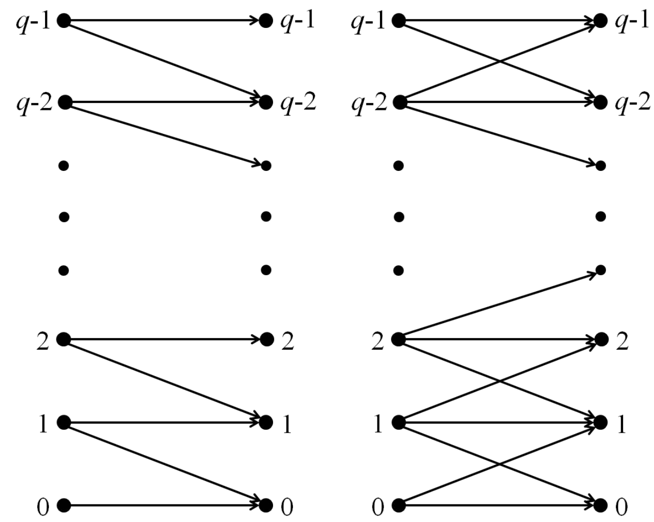
Z CHANNEL: UPPER BOUNDS COMPARISON FOR $r = 4$

n	MB	ASPV	GSPB	[26]
5	7	4	2	2
6	11	5	2	2
7	17	7	3	2
8	25	10	4	2
9	38	15	6	2
10	58	22	9	4
11	89	33	14	4
12	135	49	21	4
13	207	76	34	6
14	320	118	54	8
15	496	185	87	12
16	774	294	143	16
17	1217	472	236	26
18	1927	767	393	44
19	3073	1258	660	76
20	4939	2081	1118	134
21	7998	3470	1905	229
22	13050	5829	3266	423
23	21450	9862	5632	745
24	35509	16791	9763	?
25	59192	28761	17010	?
26	99330	49540	29772	?
27	167749	85775	52333	?
28	285019	149239	92366	?
29	487070	260846	163640	?
30	836918	457873	290949	?
31	1445509	806964	519048	?
32	2508896	1427610	928919	?

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], [14], [18], [30].

A. Asymmetric Errors

In the asymmetric non-binary channel, the value of every symbol can only decrease, and in this study we only consider



(a) (b)

Fig. 4. Two cases of the non-binary channel: (a) asymmetric errors, (b) symmetric errors.

the case where the value of each symbol can decrease by one. The corresponding graph is $\mathcal{G}_{A,q} = (X_{A,q}, E_{A,q})$, where $X_{A,q} = [q]^n$ and

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

Given some $\boldsymbol{x} \in [q]^n$, its ball of radius one is described by the set $B_{A,q,1}(\boldsymbol{x}) = \{\boldsymbol{y} \in [q]^n : \boldsymbol{x} \geqslant \boldsymbol{y}, \sum_{i=1}^{n} x_i \leqslant \sum_{i=1}^{n} y_i + 1\}$, and $\deg_{A,q,1}(\boldsymbol{x}) = w_H(\boldsymbol{x}) + 1$. The hypergraph in this case is $\mathcal{H}(\mathcal{G}_{A,q}, 1) = (X_{A,q,1}, \mathcal{E}_{A,q,1})$, where $X_{A,q,1} = [q]^n$ and $\mathcal{E}_{A,q,1} = \{B_{A,q,1}(\boldsymbol{x}) : \boldsymbol{x} \in [q]^n\}$.

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

Lemma 11: The monotonicity upper bound of the graph $\mathcal{G}_{A,q}$ for $r = 1$ is

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

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

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

Thus, the monotonicity upper bound from Equation (4) becomes

$$\begin{aligned} MB(\mathcal{G}_{A,q}, 1) &= \sum_{\boldsymbol{x} \in [q]^n} w_{\boldsymbol{x}} = \sum_{\boldsymbol{x} \in [q]^n} \frac{1}{w_H(\boldsymbol{x}) + 1} \\ &= \sum_{i_0 + i_1 + \cdots + i_{q-1} = n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} \frac{1}{i_1 + \cdots + i_{q-1} + 1} \\ &= \frac{q^{n+1} - 1}{(q-1)(n+1)}. \end{aligned} \quad (12)$$

The last equality follows from the identity

$$(1+(q-1)x)^n = \sum_{i_0+\cdots+i_{q-1}=n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} x^{i_1+\cdots+i_{q-1}},$$

and thus

$$\begin{aligned}\frac{q^{n+1}-1}{(q-1)(n+1)} &= \int_0^1 (1+(q-1)x)^n dx \\ &= \int_0^1 \sum_{i_0+\cdots+i_{q-1}=n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} x^{i_1+\cdots+i_{q-1}} dx \\ &= \sum_{i_0+\cdots+i_{q-1}=n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} \frac{1}{i_1+\cdots+i_{q-1}+1}.\end{aligned}$$

■

Lemma 12: The average sphere packing value of the graph $\mathcal{G}_{A,q}$ for $r = 1$ is

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

Proof: The value of the average ball size is

$$\begin{aligned}&\frac{1}{q^n} \cdot \sum_{\boldsymbol{x}\in[q]^n} (w_H(\boldsymbol{x})+1) \\ &= \frac{1}{q^n} \cdot \sum_{i_0+i_1+\cdots+i_{q-1}=n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} (i_1+\cdots+i_{q-1}+1) \\ &= \frac{1}{q^n} \cdot (nq^{n-1}(q-1)+q^n) = n+1-n/q.\end{aligned}$$

Here, the second equality is a result of the identity

$$x(1+(q-1)x)^n = \sum_{i_0+\cdots+i_{q-1}=n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} x^{i_1+\cdots+i_{q-1}+1},$$

and hence

$$\begin{aligned}&(1+(q-1)x)^n + xn(q-1)(1+(q-1)x)^{n-1} \\ &\quad = \sum_{i_0+\cdots+i_{q-1}=n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} \\ &\qquad\qquad \times (i_1+\cdots+i_{q-1}+1)x^{i_1+\cdots+i_{q-1}},\end{aligned}$$

and thus by assigning $x = 1$ we get the required result. ■

The linear programming problem from (1) becomes

$$\tau^*(\mathcal{H}(\mathcal{G}_{A,q,1})) = \min\left\{ \sum_{\boldsymbol{x}\in[q]^n} w_{\boldsymbol{x}} \;:\; \sum_{\boldsymbol{y}\in B_{A,q,1}(\boldsymbol{x})} w_{\boldsymbol{y}} \geqslant 1 \right\}.$$

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

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

where $X_{\boldsymbol{i}}$ is characterized as follows

$$X_{\boldsymbol{i}} = \{\boldsymbol{x} \in [q]^n : \boldsymbol{x}^{-1}(j) = i_j, 0 \leqslant j \leqslant q-1\},$$

and $\boldsymbol{x}^{-1}(j) = |\{1 \leqslant k \leqslant n : x_k = j\}|$. We denote the set $\mathbf{I}_A$ to be $\mathbf{I}_A = \{\boldsymbol{i} \;:\; \boldsymbol{i} = (i_0, \ldots, i_{q-1}) \geqslant \boldsymbol{0}, \sum_{j=0}^{q-1} i_j = n\}$ and define an $n_A \times n_A$ matrix A_H such that its entries are indexed by the vectors $(\boldsymbol{i}, \boldsymbol{j}) \in \mathbf{I}_A \times \mathbf{I}_A$. We assign the values $A_H(\boldsymbol{i}, \boldsymbol{i}) = 1$ and $A_H(\boldsymbol{i}, \boldsymbol{j}) = i_k$ if there exists $1 \leqslant k \leqslant q-1$ such that $j_k = i_k - 1$ and $j_{k-1} = i_{k-1} + 1$ and for all $\ell \in [q] \setminus \{k, k-1\}$, $j_\ell = i_\ell$. All other values in the matrix A_H are assigned with the value 0. Finally, according to Corollary 8 we proved the following theorem.

Theorem 13: The generalized sphere packing bound for $\mathcal{H}(\mathcal{G}_{A,q,1})$ is given by

$$\begin{aligned}&\tau^*(\mathcal{H}(\mathcal{G}_{A,q,1})) \\ &\quad = \min\left\{ \sum_{\boldsymbol{i}\in\mathbf{I}_A} |X_{\boldsymbol{i}}| w_{\boldsymbol{i}} \;:\; A_H^T \cdot \boldsymbol{w} \geqslant 1, \boldsymbol{w} = (w_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbf{I}_A} \in \mathbb{R}_+^{n_A} \right\}.\end{aligned}$$

We finish this section by showing an improvement upon the monotonicity upper bound from Lemma 11. In the fractional transversal notation of Theorem 13, if one applied the monotonicity upper bound, then the fractional transversal assignment would be $w_{\boldsymbol{i}} = 1/(n-i_0+1)$ for $\boldsymbol{i} \in \mathbf{I}_A$. 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 obtain a stronger upper bound.

Theorem 14: The vector $\boldsymbol{w} = (w_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbf{I}_A}$ given by

$$w_{\boldsymbol{i}} = \frac{1}{n - i_0 + 1 + \frac{i_1-1}{2(n-i_0)}},$$

if $i_0 \neq n$ and otherwise $w_{\boldsymbol{i}} = 1$ is a fractional transversal for $\tau^*(\mathcal{H}(\mathcal{G}_{A,q,1}))$ as stated in Theorem 13.

Proof: It is straightforward to verify that $w_{\boldsymbol{i}} \geqslant 0$ for all $\boldsymbol{i} \in \mathbf{I}_A$. According to the conditions for $\tau^*(\mathcal{H}(\mathcal{G}_{A,q,1}))$ from Theorem 13, we need to show that for all $\boldsymbol{i} = (i_0, i_1, \ldots, i_{q-1}) \in \mathbf{I}_A$ the following inequality holds

$$\begin{aligned}&w_{(i_0,i_1,\ldots,i_{q-1})} + i_1 w_{(i_0+1,i_1-1,\ldots,i_{q-1})} + i_2 w_{(i_0,i_1+1,i_2-1\ldots,i_{q-1})} \\ &\qquad + \cdots + i_{q-1} w_{(i_0,i_1,\ldots,i_{q-2}+1,i_{q-1}-1)} \geqslant 1.\end{aligned}$$

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 $w_{\boldsymbol{i}}$ stated in the theorem, we need to show the following

$$\begin{aligned}&\frac{1}{n-i_0+1+\frac{i_1-1}{2(n-i_0)}} + \frac{i_1}{n-i_0+\frac{i_1-2}{2(n-i_0-1)}} \\ &\qquad + \frac{i_2}{n-i_0+1+\frac{i_1}{2(n-i_0)}} + \frac{i_3+\cdots+i_{q-1}}{n-i_0+1+\frac{i_1-1}{2(n-i_0)}} \geqslant 1.\end{aligned}$$

Note that

$$\frac{1}{n-i_0+1+\frac{i_1-1}{2(n-i_0)}} \geqslant \frac{1}{n-i_0+1+\frac{i_1}{2(n-i_0)}},$$

and thus it is enough to show that

$$\frac{i_2+i_3+\cdots+i_{q-1}+1}{n-i_0+1+\frac{i_1}{2(n-i_0)}} + \frac{i_1}{n-i_0+\frac{i_1-2}{2(n-i_0-1)}} \geqslant 1,$$

TABLE V
NON-BINARY CHANNEL, ASYMMETRIC ERRORS: UPPER BOUNDS COMPARISON FOR $q = 3$

n	MB	ASPV	Theorem 14	GSPB
5	60	56	60	55
6	156	145	154	144
7	410	385	402	381
8	1093	1035	1071	1021
9	2952	2811	2888	2770
10	8052	7702	7877	7591
11	22143	21252	21673	20955
12	61320	59049	60056	58235
13	170820	164929	167424	162744
14	478296	462867	469156	456987
15	1345210	1304446	1320524	1288583
16	3798240	3689718	3731321	3646657
17	10761680	10470824	10579575	10353898
18	30585828	29801576	30088394	28464819
19	87169610	85043521	85805885	84168158
20	249056028	243264027	245304388	230986164
21	713205900	697356880	702851238	690706260
22	2046590844	2003046358	2017923470	1984633746
23	5883948676	5763868091	5804351676	5712720517

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

$$\frac{n - i_0 - i_1 + 1}{n - i_0 + 1 + \frac{i_1}{2(n-i_0)}} + \frac{i_1}{n - i_0 + \frac{i_1-2}{2(n-i_0-1)}} \geqslant 1$$

and

$$\frac{i_1}{n - i_0 + \frac{i_1-2}{2(n-i_0-1)}} \geqslant 1 - \frac{n - i_0 - i_1 + 1}{n - i_0 + 1 + \frac{i_1}{2(n-i_0)}}$$

which is

$$\frac{i_1}{n - i_0 + \frac{i_1-2}{2(n-i_0-1)}} \geqslant \frac{i_1 + \frac{i_1}{2(n-i_0)}}{n - i_0 + 1 + \frac{i_1}{2(n-i_0)}}.$$

Let us denote $n - i_0 = M$ and we need to show that

$$\frac{1}{M + \frac{i_1-2}{2(M-1)}} \geqslant \frac{1 + \frac{1}{2M}}{M + 1 + \frac{i_1}{2M}},$$

and equivalently

$$\frac{2(M-1)}{2M^2 - 2M + i_1 - 2} \geqslant \frac{2M+1}{2M^2 + 2M + i_1},$$

or

$$2M^2 + 2M + 2 \geqslant 3i_1,$$

which holds since $M = n - i_0 \geqslant i_1$. ∎

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 11. The second column is the average sphere packing value from Lemma 12. The third column is the improvement in Theorem 14 over the monotonicity upper bound. Lastly, the last column is the value of the generalized sphere packing bound from Theorem 13, 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 $\mathcal{G}_{S,q} = (X_{S,q}, E_{S,q})$, where $X_{S,q} = [q]^n$ and

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

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

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 15: The average sphere packing value of the graph $\mathcal{G}_{S,q}$ for $r = 1$ is

$$ASPV(\mathcal{G}_{S,q}, 1) = \frac{q^n}{2n + 1 - 2n/q}.$$

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

$$\frac{1}{q^n} \cdot \sum_{\boldsymbol{x} \in [q]^n} (2n + 1 - \boldsymbol{x}^{-1}(0) - \boldsymbol{x}^{-1}(q-1))$$
$$= \frac{1}{q^n} \cdot \sum_{i_0 + \cdots + i_{q-1} = n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} (n+1+(i_1 + \cdots + i_{q-2}))$$
$$= \frac{1}{q^n} \cdot ((n+1)q^n + n(q-2)q^{n-1}) = 2n + 1 - 2n/q.$$

The second equality is derived in a similar manner to the one in Lemma 12 using the identity

$$x^{n+1}(2 + (q-2)x)^n = \sum_{i_0 + \cdots + i_{q-1} = n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} x^{i_1 + \cdots + i_{q-2} + n + 1},$$

so

$$(n+1)x^n(2 + (q-2)x)^n + x^{n+1}n(q-2)(2 + (q-2)x)^{n-1} = \sum_{i_0 + \cdots + i_{q-1} = n} \binom{n}{i_0, i_1, \ldots, i_{q-1}} \cdot (i_1 + \cdots + i_{q-2} + n + 1)x^{i_1 + \cdots + i_{q-2}},$$

and we get the required equality by assigning $x = 1$. ∎

Next, we define the set of automorphisms to be used here. One can verify that every permutation in H_A is an automorphism in $\mathcal{G}_{S,q}$. However, in this case we can expand and use more automorphisms. For every binary vector $\boldsymbol{b} \in \{0, 1\}^n$, we define the permutation

$$\pi_{\boldsymbol{b}} : [q]^n \to [q]^n,$$

as follows. For every $\boldsymbol{x} \in [q]^n$, $\pi_{\boldsymbol{b}}(\boldsymbol{x})$ is the vector defined as

$$\pi_{\boldsymbol{b}}(\boldsymbol{x})_i = \begin{cases} q - 1 - x_i & \text{if } b_i = 1, \\ x_i & \text{if } b_i = 0. \end{cases}$$

Then, the set $H_S = H_A \cup \{\pi_{\boldsymbol{b}} \ : \ \boldsymbol{b} \in \{0,1\}^n\}$ is a subgroup of $Aut(\mathcal{G}_{S,q})$. The subgroup H_S partitions the set $[q]^n$ into $n_S = \binom{n+\lceil q/2\rceil-1}{\lceil q/2\rceil-1}$ equivalence classes

$$\mathcal{X} = \left\{X_{\boldsymbol{i}} \ : \ \boldsymbol{i} = (i_0, \ldots, i_{\lceil q/2\rceil-1}) \geqslant \boldsymbol{0}, \ \sum_{j=0}^{\lceil q/2\rceil-1} i_j = n\right\},$$

and $X_{\boldsymbol{i}}$ is the set

$$X_{\boldsymbol{i}} = \{\boldsymbol{x} \in [q]^n : \boldsymbol{x}^{-1}(j) + \boldsymbol{x}^{-1}(q-1-j) = i_j, 0 \leqslant j \leqslant \lceil q/2\rceil - 1\}.$$

We define

$$\mathbf{I}_S = \left\{\boldsymbol{i} \ : \ \boldsymbol{i} = (i_0, \ldots, i_{\lceil q/2\rceil-1}) \geqslant \boldsymbol{0}, \ \sum_{j=0}^{\lceil q/2\rceil-1} i_j = n\right\}$$

and the $n_S \times n_S$ matrix A_S with the following entries $(\boldsymbol{i},\boldsymbol{j}) \in \mathbf{I}_S \times \mathbf{I}_S$.

1) For all $\boldsymbol{i} \in \mathbf{I}_S$, $A_S(\boldsymbol{i},\boldsymbol{i}) = 1$,
2) $A_S(\boldsymbol{i},\boldsymbol{j}) = k$ if there exists $1 \leqslant k \leqslant \lceil q/2\rceil - 1$ such that $j_k = i_k - 1$ and $j_{k-1} = i_{k-1} + 1$ and for all $\ell \in [\lceil q/2\rceil] \setminus \{k, k-1\}$, $j_\ell = i_\ell$.

To conclude, according to Corollary 8, the generalized sphere packing for $\mathcal{H}(\mathcal{G}_{S,q,1})$ becomes

$$\tau^*(\mathcal{H}(\mathcal{G}_{S,q,1})) = \min\left\{\sum_{\boldsymbol{i}\in\mathbf{I}_S} |X_{\boldsymbol{i}}| w_{\boldsymbol{i}} : A_S^T \cdot \boldsymbol{w} \geqslant 1, \boldsymbol{w} = (w_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbf{I}_S} \in \mathbb{R}_+^{n_S}\right\}.$$

We can derive a bound similar to Theorem 14.

Theorem 16: The vector $\boldsymbol{w} = (w_{\boldsymbol{x}})_{\boldsymbol{x}\in[q]^n}$ given by

$$w_{\boldsymbol{x}} = \frac{1}{\deg_{S,q,1}(\boldsymbol{x}) - 1}$$

is a fractional transversal.

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

$$\frac{1}{2n - i_0 - i_{q-1}} + \frac{i_0}{2n - (i_0 - 1) - i_{q-1}} + \frac{i_1}{2n - (i_0 + 1) - i_{q-1}} + \frac{i_1 + 2i_2 + \cdots + 2i_{q-3} + i_{q-2}}{2n - i_0 - i_{q-1}} + \frac{i_{q-2}}{2n - i_0 - (i_{q-1} + 1)} + \frac{i_{q-1}}{2n - i_0 - (i_{q-1} - 1)} \geqslant 1,$$

or

$$\frac{1 + i_1 + 2i_2 + \cdots + 2i_{q-3} + i_{q-2}}{2n - i_0 - i_{q-1}} + \frac{i_0 + i_{q-1}}{2n - i_0 - i_{q-1} + 1} + \frac{i_1 + i_{q-2}}{2n - i_0 - i_{q-1} - 1} \geqslant 1.$$

Since $1/(2n - i_0 - i_{q-1}) \geqslant 1/(2n - i_0 - i_{q-1} + 1)$ and $1/(2n - i_0 - i_{q-1} - 1) \geqslant 1/(2n - i_0 - i_{q-1} + 1)$, it is enough to show that

$$\frac{1 + i_1 + 2i_2 + \cdots + 2i_{q-3} + i_{q-2}}{2n - i_0 - i_{q-1} + 1} + \frac{i_0 + i_{q-1}}{2n - i_0 - i_{q-1} + 1} + \frac{i_1 + i_{q-2}}{2n - i_0 - i_{q-1} + 1} \geqslant 1,$$

which holds with equality. ■

TABLE VI
NON-BINARY CHANNEL, SYMMETRIC ERRORS: UPPER BOUNDS COMPARISON FOR $q = 3$

n	ASPV	Theorem 16	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
NON-BINARY CHANNEL, SYMMETRIC ERRORS: UPPER BOUNDS COMPARISON FOR $q = 4$

n	ASPV	Theorem 16	GSPB
5	120	139	123
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	883011	940934	891482

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 15. The second column is the upper bound we found in Theorem 16. 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 not a valid upper bound.

VI. DELETION AND GRAIN-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 [16]. We will only focus on the single-deletion case. First, we revisit the fractional transversal given in [16] to verify that the graph in the deletion channel satisfies a similar property to the monotonicity property from Section III-B. Then we present our main result in this section, namely, an explicit expression of a fractional transversal which improves upon the one from [16]. Since the structure of the deletion and grain-error channels is very similar, especially for a single error, in the second part of this section we show also how to improve upon the upper bound from [10] and [13] on the cardinality of single grain-error-correcting codes.

A. Deletions

As in the previous examples, we first introduce the graph for the deletion channel. However, note that the graph in this setup is different than the previous ones. Specifically, a length n

vector which suffers a single deletion will result in 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 $\mathcal{G}_D = (X_D, E_D)$, where $X_D = \{0,1\}^n \cup \{0,1\}^{n-1}$ and

$$E_D = \{(\boldsymbol{x}, \boldsymbol{y}) \in \{0,1\}^n \times \{0,1\}^{n-1} : \boldsymbol{y} = (x_1, \ldots, x_i, x_{i+2}, \ldots, x_n) \text{ for some } 1 \leqslant i \leqslant n\}.$$

For any $\boldsymbol{x} \in \{0,1\}^n$, its radius one ball is the set $B_{D,1}(\boldsymbol{x}) = \{\boldsymbol{y} \in \{0,1\}^{n-1} : (\boldsymbol{x}, \boldsymbol{y}) \in E_D\}$, and for $\boldsymbol{x} \in \{0,1\}^{n-1}$, $B_{D,1}(\boldsymbol{x}) = \emptyset$. Therefore, $1 \leqslant \deg_{D,1}(\boldsymbol{x}) \leqslant n$ for $\boldsymbol{x} \in \{0,1\}^n$, and $\deg_{D,1}(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in \{0,1\}^{n-1}$.

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 [16]. Thus the hypergraph for the single deletion channel is $\mathcal{H}(\mathcal{G}_D, 1) = (X_{D,1}, \mathcal{E}_{D,1})$, where $X_{D,1} = \{0,1\}^{n-1}$ and $\mathcal{E}_{D,1} = \{B_{D,1}(\boldsymbol{x}) : \boldsymbol{x} \in \{0,1\}^n\}$. 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

$$\tau^*(\mathcal{H}(\mathcal{G}_D, 1)) = \min\left\{ \sum_{z \in \{0,1\}^{n-1}} w_z : \sum_{y \in B_{D,1}(\boldsymbol{x})} w_{\boldsymbol{y}} \geqslant 1, \forall \boldsymbol{x} \in \{0,1\}^n \right\}. \tag{13}$$

For a vector $\boldsymbol{x} \in \{0,1\}^n$, we denote by $\rho(\boldsymbol{x})$ the number of runs in $\boldsymbol{x}$. For example, if $\boldsymbol{x} = 001010010$, then $\rho(\boldsymbol{x}) = 7$. It is easily verified that for $\boldsymbol{x} \in \{0,1\}^n$, $\deg_{D,1}(\boldsymbol{x}) = \rho(\boldsymbol{x})$, [16]. It is also known that the number of length-n vectors with $1 \leqslant \rho \leqslant n$ runs is $2\binom{n-1}{\rho-1}$. Let us first calculate the average sphere packing value for the hypergraph $\mathcal{H}(\mathcal{G}_D, 1)$. This will be done in the next lemma.

Lemma 17: The average sphere packing value of the graph $\mathcal{G}_D$ for $r = 1$ is

$$ASPV(\mathcal{G}_D, 1) = \frac{2^n}{n+1}.$$

Proof: Every vector $\boldsymbol{x} \in \{0,1\}^n$ generates a ball, i.e. a hyperedge, in $\mathcal{H}(\mathcal{G}_D, 1)$. Thus, the average size of a ball is given by

$$\frac{1}{2^n} \sum_{\boldsymbol{x} \in \{0,1\}^n} \deg_{D,1}(\boldsymbol{x}) = \frac{1}{2^n} \sum_{\rho=1}^{n} 2\binom{n-1}{\rho-1}\rho$$
$$= \frac{1}{2^{n-1}} \sum_{\rho=0}^{n-1} \binom{n-1}{\rho}(\rho+1) = \frac{1}{2^{n-1}}(2^{n-1} + (n-1)2^{n-2})$$
$$= \frac{n+1}{2}.$$

Thus, the average sphere packing value becomes

$$\frac{2^{n-1}}{(n+1)/2} = \frac{2^n}{n+1}.$$

■

Note that if one chose the hypergraph to contain all binary vectors of length $n-1$ and n, the resulting average sphere packing value would have been weaker. We specifically chose the hypergraph this way as it is the smallest one where any single-deletion code can be studied and analyzed.

In the setup and structure of the graph $\mathcal{G}_D$, it is not possible to indicate whether the graph $\mathcal{G}_D$ satisfies the monotonicity property. The vectors in the ball centered at some $\boldsymbol{x} \in \{0,1\}^n$ are of length $n-1$ and thus do not have corresponding balls. However, there is still a very similar property to the monotonicity one. Namely, for every $\boldsymbol{y} \in B_{D,1}(\boldsymbol{x})$, where $\boldsymbol{x} \in \{0,1\}^n$,

$$\rho(\boldsymbol{y}) \leqslant \rho(\boldsymbol{x}) = \deg_{D,1}(\boldsymbol{x}).$$

This property was established in [16] and thus a choice of a fractional transversal $(w_{\boldsymbol{x}})_{\boldsymbol{x} \in \{0,1\}^{n-1}}$, was given by

$$w_{\boldsymbol{x}} = \frac{1}{\rho(\boldsymbol{x})}.$$

The corresponding upper bound, which we call here the monotonicity upper bound, was calculated in [16] to be

$$\sum_{\boldsymbol{x} \in \{0,1\}^{n-1}} \frac{1}{\rho(\boldsymbol{x})} = \sum_{\rho=1}^{n-1} 2\binom{n-2}{\rho-1} \cdot \frac{1}{\rho} = \frac{2^n - 2}{n-1}.$$

However, it is possible to verify that for this fractional transversal many of the constraints in the linear programming in (13) hold with strict inequality, which implies that a better one could be found. This will be the focus in the rest of this subsection, that is, an improvement upon the last upper bound by the equivalent of the monotonicity property.

For a vector $\boldsymbol{x}$, let $\mu(\boldsymbol{x})$ be the number of middle runs (i.e., not on the edges) of length 1 in $\boldsymbol{x}$. We call these runs *middle-1-runs*. For example, for $\boldsymbol{x} = 001010010$, $\mu(\boldsymbol{x}) = 4$. First notice that if $\rho(\boldsymbol{x}) \geqslant 2$ then $0 \leqslant \mu(\boldsymbol{x}) \leqslant \rho(\boldsymbol{x}) - 2$. Let $N_n(\rho, \mu)$ denote the number of vectors of length n with ρ runs and μ middle-1-runs. For $\rho = 1$ and $\mu = 0$, we have $N_n(1, 0) = 2$. For $2 \leqslant \rho \leqslant n$ and $0 \leqslant \mu \leqslant \rho - 2$, the value of $N_n(\rho, \mu)$ is calculated in the next lemma. For all other values of ρ and μ the value of $N_n(\rho, \mu)$ is zero.

Lemma 18: For $2 \leqslant \rho \leqslant n$ and $0 \leqslant \mu \leqslant \rho - 2$,

$$N_n(\rho, \mu) = 2\binom{\rho-2}{\mu}\binom{n-\rho+1}{\rho-\mu-1}.$$

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

Let $\boldsymbol{x}$ be a vector of length n such that $\rho(\boldsymbol{x}) = \rho$ and $\mu(\boldsymbol{x}) = \mu$, where $1 \leqslant \rho \leqslant n$ and $0 \leqslant \mu \leqslant \rho - 2$. Assume the vector $\boldsymbol{x}'$ has p runs of ones of length $h_1, \ldots, h_p$. Then, first we have that

$$\sum_{i=1}^{p} h_i = w_H(\boldsymbol{x}') = \rho - 1. \tag{14}$$

Every run of ones of length h_i in $\boldsymbol{x}'$ contributes $h_i - 1$ pairs of two consecutive ones. Therefore,

$$\sum_{i=1}^{p}(h_i - 1) = \mu. \tag{15}$$

Together, from (14) and (15), we conclude that $p = \rho - \mu - 1$. Furthermore, the number of solutions to (14) (or (15)) is $\binom{\mu+p-1}{\mu} = \binom{\rho-2}{\mu}$. 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 $\boldsymbol{x}'$, where $k_0 \geqslant 0, k_1, \ldots, k_{\rho-\mu-2} \geqslant 1$, and $k_{\rho-\mu-1} \geqslant 0$. Note that their sum is $n-1-(\rho-1) = n-\rho$, and thus, under the above constraints, the number of solutions to

$$\sum_{j=0}^{\rho-\mu-1} k_j = n - \rho$$

is $\binom{n-\rho+1}{\rho-\mu-1}$.

Finally, the number of options to choose the vector $\boldsymbol{x}'$ 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 $\boldsymbol{x}'$ determines the vector $\boldsymbol{x}$ up to choosing whether it starts with zero or one. Therefore, we get

$$N_n(\rho, \mu) = 2\binom{\rho-2}{\mu}\binom{n-\rho+1}{\rho-\mu-1}.$$

■

Next, the main result in this section is proved.

Theorem 19: The vector $\boldsymbol{w} = (w_{\boldsymbol{x}})_{\boldsymbol{x}\in\{0,1\}^{n-1}}$ defined by

$$w_{\boldsymbol{x}} = \begin{cases} \frac{1}{\rho(\boldsymbol{x})} & \text{if } \mu(\boldsymbol{x}) \leqslant 1 \\ \frac{1}{\rho(\boldsymbol{x})}\left(1 - \frac{\mu(\boldsymbol{x})}{\rho(\boldsymbol{x})^2}\right) & \text{otherwise} \end{cases}$$

is a fractional transversal.

Proof: Let $\boldsymbol{x}$ be a length-n binary vector with ρ runs and μ middle-1-runs. We need to show that $\sum_{\boldsymbol{y}\in B_{D,1}(\boldsymbol{x})} w_{\boldsymbol{y}} \geqslant 1$. 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 $\rho \geqslant 4$ and $\mu \geqslant 2$. Note that for a fixed ρ, $w_{\boldsymbol{x}}$ decreases when μ increases.

If a vector $\boldsymbol{y} \in B_{D,1}(\boldsymbol{x})$ is received by deleting a middle-1-run bit then $\rho(\boldsymbol{y}) = \rho - 2$ and $\mu - 3 \leqslant \mu(\boldsymbol{y}) \leqslant \mu - 1$. Otherwise, $\rho(\boldsymbol{y}) = \rho$ and $\mu(\boldsymbol{y}) \leqslant \mu + 1$ or, if it is the first or last bit which is a single-bit run, $\rho(\boldsymbol{y}) = \rho - 1$ and $\mu(\boldsymbol{y}) \leqslant \mu$, however, the worst case in terms of the value of $w_{\boldsymbol{y}}$ is achieved for $\rho(\boldsymbol{y}) = \rho$ and $\mu(\boldsymbol{y}) = \mu + 1$. Therefore,

$$\sum_{\boldsymbol{y}\in B_{D,1}(\boldsymbol{x})} w_{\boldsymbol{y}} \geqslant \frac{\mu}{\rho-2}\left(1 - \frac{\mu-1}{(\rho-2)^2}\right) + \frac{(\rho-\mu)}{\rho}\left(1 - \frac{\mu+1}{\rho^2}\right)$$
$$= 1 + \frac{2\mu}{\rho(\rho-2)} - \frac{\mu(\mu-1)}{(\rho-2)^3} - \frac{\mu+1}{\rho^2} + \frac{\mu(\mu+1)}{\rho^3},$$

and thus it is enough to show that for $\rho \geqslant 4, 2 \leqslant \mu \leqslant \rho - 2$,

$$\frac{2\mu}{\rho(\rho-2)} - \frac{\mu(\mu-1)}{(\rho-2)^3} - \frac{\mu+1}{\rho^2} + \frac{\mu(\mu+1)}{\rho^3} \geqslant 0,$$

or

$$\frac{2}{\rho(\rho-2)} - \frac{1}{\rho^2} \geqslant \frac{\mu-1}{(\rho-2)^3} + \frac{1}{\mu\rho^2} - \frac{\mu+1}{\rho^3}.$$

TABLE VIII

DELETION CHANNEL COMPARISON

n	MB [16]	ASPV	Theorem 20	GSPB [16]	LB [25]
5	7	5	7	6	6
6	12	9	12	10	10
7	21	16	20	17	16
8	36	28	35	30	30
9	63	51	61	53	52
10	113	93	109	96	94
11	204	170	197	175	172
12	372	315	358	321	316
13	682	585	657	593	586
14	1260	1092	1212	1104	1096
15	2340	2048	2251	?	2048
16	4368	3855	4202	?	3856
17	8191	7281	7882	?	7286
18	15420	13797	14845	?	13798
19	29127	26214	28059	?	26216
20	55188	49932	53202	?	49940
21	104857	95325	101163	?	95326
22	199728	182361	192850	?	182362
23	381300	349525	368478	?	349536

The function

$$f(\mu) = \frac{\mu-1}{(\rho-2)^3} + \frac{1}{\mu\rho^2} - \frac{\mu+1}{\rho^3}$$

in the range $2 \leqslant \mu \leqslant \rho - 2$ is maximized either when $\mu = 2$ or $\mu = \rho - 2$ and thus we need to show that

$$\frac{2}{\rho(\rho-2)} - \frac{1}{\rho^2} \geqslant \frac{1}{(\rho-2)^3} + \frac{1}{2\rho^2} - \frac{3}{\rho^3},$$

and

$$\frac{2}{\rho(\rho-2)} - \frac{1}{\rho^2} \geqslant \frac{\rho-3}{(\rho-2)^3} + \frac{1}{(\rho-2)\rho^2} - \frac{\rho-1}{\rho^3},$$

which holds for all $\rho \geqslant 4$. ■

For a vector $\boldsymbol{x}$ with ρ runs and μ middle-1-runs, we denote its weight by $w(\rho, \mu)$, as specified in Theorem 19. From Lemma 18 and Theorem 19 we conclude with the following upper bound on $\tau^*(\mathcal{H}(\mathcal{G}_D, 1))$.

Theorem 20: The value $\tau^*(\mathcal{H}(\mathcal{G}_D, 1))$ satisfies

$$\tau^*(\mathcal{H}(\mathcal{G}_D, 1)) \leqslant 2 + \sum_{\rho=2}^{n-1}\sum_{\mu=0}^{\rho-2} N_{n-1}(\rho, \mu) w(\rho, \mu).$$

Proof: We calculate the upper bound on $\tau^*(\mathcal{H}(\mathcal{G}_D, 1))$ according to the fractional transversal from Theorem 19, $\boldsymbol{w} = (w_{\boldsymbol{x}})_{\boldsymbol{x}\in\{0,1\}^{n-1}}$. Every vector $\boldsymbol{x}$ is assigned with a weight $w_{\boldsymbol{x}} = w(\rho, \mu)$ according to its number of runs ρ and number of middle-1-runs μ. Thus, we get this upper bound to be

$$\sum_{\boldsymbol{x}\in\{0,1\}^{n-1}} w_{\boldsymbol{x}} = 2 + \sum_{\rho=2}^{n-1}\sum_{\mu=0}^{\rho-2} N_{n-1}(\rho, \mu) w(\rho, \mu).$$

■

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 $\frac{2^n-2}{n-1}$ from [16]. ASPV corresponds to the average sphere packing value $\frac{2^n}{n+1}$ from Lemma 17. The third column is our upper bound results from Theorem 20. The column titled GSPB [16] is the exact value of $\tau^*(\mathcal{H}(\mathcal{G}_D, 1))$ from (13), which this linear programming problem was numerically solved in [16] for $n \leqslant 14$. Since this linear programming has a large number of constraints and variables it is numerically hard to solve it for larger values of n. The last column LB corresponds to the lower bound, which is the best known construction of single-deletion codes from [25]. We notice here that the single-deletion codes from [25] show that the ASPV is not a valid upper bound in case n is not of the form $2^m - 1$. However, we also note that it depends on the choice of the hypergraph. As mentioned earlier, if the hypergraph consists of all binary vectors of length $n-1$ and n then the ASPV would have been a valid upper bound.

B. Grain Errors

The grain-error channel is a recent model which was studied mainly for granular media with applications to magnetic recording technologies [28], [29]. In this medium, the information is stored in individual grains and every grain can store a single information bit according to the polarity of the grain. However, it may happen that a single grain holds more than a single information bit (we assume here two), in which case the polarity of the cell is determined by the last bit that was written to it. Thus, if the value of the two information bits sharing the same grain is not the same then one of them will be in error, called ***grain-error***. We will follow the model studied by previous works which assume that the first bit smears its adjacent one to the right. There are several recent studies of this model which analyzed its information theory behavior [12], proposed code constructions, and upper bounds [10], [13], [17], [20], [21].

The grain-error channel for a single grain-error is very similar to the single deletion setup in the sense that in both case the size of the radius one ball depends only on the number of runs in the center of the ball. However in this case the length of the received words remains the same. The graph describing this channel model is $\mathcal{G}_G = (X_G, E_G)$, where $X_G = \{0,1\}^n$ and

$$E_G = \{(\boldsymbol{x},\boldsymbol{y}) : \ \boldsymbol{x},\boldsymbol{y} \in \{0,1\}^n, \text{ and there exists } 2 \leqslant i \leqslant n \text{ such that } \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{e}_i \text{ and } x_i \neq x_{i-1}\},$$

where '+' denotes modulo 2 addition. The radius one ball for some $\boldsymbol{x} \in \{0,1\}^n$ is the set $B_{G,1}(\boldsymbol{x}) = \{\boldsymbol{y} \in \{0,1\}^n : (\boldsymbol{x},\boldsymbol{y}) \in E_G\}$. The hypergraph for the single grain-error channel becomes $\mathcal{H}(\mathcal{G}_G, 1) = (X_{G,1}, \mathcal{E}_{G,1})$, where $X_{G,1} = \{0,1\}^n$ and $\mathcal{E}_{G,1} = \{B_{G,1}(\boldsymbol{x}) : \boldsymbol{x} \in \{0,1\}^n\}$. Finally, the generalized sphere packing bound for the single grain-error channel is

$$\tau^*(\mathcal{H}(\mathcal{G}_G,1)) = \min\left\{ \sum_{\boldsymbol{x}\in\{0,1\}^n} w_{\boldsymbol{x}} : \sum_{\boldsymbol{y}\in B_{G,1}(\boldsymbol{x})} w_{\boldsymbol{y}} \geqslant 1, \forall \boldsymbol{x} \in \{0,1\}^n \right\}.$$

The size of the ball $B_{G,1}(\boldsymbol{x})$ can be given by $\deg_{G,1}(\boldsymbol{x}) = \rho(\boldsymbol{x})$, where, as before, $\rho(\boldsymbol{x})$ is the number of runs in $\boldsymbol{x}$. It is also verified that if $\boldsymbol{y} \in B_{G,1}(\boldsymbol{x})$ then $\rho(\boldsymbol{y}) \leqslant \rho(\boldsymbol{x})$ and thus the graph $\mathcal{G}_G$ satisfies the monotonicity property. These results were verified both in [10] and [13] and showed that the vector $\boldsymbol{w} = (w_{\boldsymbol{x}})_{\boldsymbol{x}\in\{0,1\}^n}$ given by $w_{\boldsymbol{x}} = \frac{1}{\rho(\boldsymbol{x})}$, is a fractional transversal. Accordingly, the corresponding upper bound, called here the monotonicity upper bound, on $\tau^*(\mathcal{H}(\mathcal{G}_G,1))$ becomes

$$MB(\mathcal{G}_G,1) = \frac{2^{n+1}-2}{n}.$$

This bound is slightly improved in [10] 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 $2\lfloor\frac{2^{n+1}-2}{2n}\rfloor$.

The average sphere packing value in this case is calculated in the next lemma. The proof is omitted since it is identical to the one of Lemma 17.

Lemma 21: The average sphere packing value of the graph $\mathcal{G}_G$ for $r = 1$ is

$$ASPV(\mathcal{G}_G,1) = \frac{2^{n+1}}{n+1}.$$

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 22: The vector $\boldsymbol{w} = (w_{\boldsymbol{x}})_{\boldsymbol{x}\in\{0,1\}^n}$ defined by

$$w_{\boldsymbol{x}} = \begin{cases} \frac{1}{\rho(\boldsymbol{x})} & \text{if } \mu(\boldsymbol{x}) \leqslant 1, \\ \frac{1}{\rho(\boldsymbol{x})}\left(1 - \frac{\mu(\boldsymbol{x})}{\rho(\boldsymbol{x})^2}\right) & \text{otherwise,} \end{cases}$$

is a fractional transversal.

Proof: Let $\boldsymbol{x}$ be a binary vector of length n with ρ runs and μ middle-1-runs. We will show that $\sum_{\boldsymbol{y}\in B_{G,1}(\boldsymbol{x})} w_{\boldsymbol{y}} \geqslant 1$. As in the proof of Theorem 19, it is possible to verify that this property holds for $\rho = 1,2,3$ and $\mu = 0,1$, so we assume that $\rho \geqslant 4$ and $\mu \geqslant 2$.

If a vector $\boldsymbol{y} \in B_{G,1}(\boldsymbol{x})$ is received by a single grain-error of a middle-1-run bit then $\rho(\boldsymbol{y}) = \rho - 2$ and $\mu(\boldsymbol{y}) \leqslant \mu - 1$. Otherwise, $\rho(\boldsymbol{y}) = \rho$ and $\mu(\boldsymbol{y}) \leqslant \mu + 1$ or, in case the last bit errs, $\rho(\boldsymbol{y}) = \rho - 1$ and $\mu(\boldsymbol{y}) \leqslant \mu$, however, the worst case is achieved for $\rho(\boldsymbol{y}) = \rho$ and $\mu(\boldsymbol{y}) = \mu + 1$. Hence, we get

$$\begin{aligned}\sum_{\boldsymbol{y}\in B_{G,1}(\boldsymbol{x})} w_{\boldsymbol{y}} &\geqslant \frac{1}{\rho}\left(1-\frac{\mu}{\rho^2}\right) + \frac{\mu}{\rho-2}\left(1-\frac{\mu-1}{(\rho-2)^2}\right) \\ &\quad + \frac{(\rho-\mu-1)}{\rho}\left(1-\frac{\mu+1}{\rho^2}\right) \\ &\geqslant 1 + \frac{2\mu}{\rho(\rho-2)} - \frac{\mu(\mu-1)}{(\rho-2)^3} - \frac{\mu+1}{\rho^2} + \frac{\mu(\mu+1)}{\rho^3}.\end{aligned}$$

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

Finally, we conclude with the following theorem.

Theorem 23: The value $\tau^*(\mathcal{H}(\mathcal{G}_G,1))$ satisfies

$$\tau^*(\mathcal{H}(\mathcal{G}_G,1)) \leqslant 2 + \sum_{\rho=2}^{n}\sum_{\mu=0}^{\rho-2} N_n(\rho,\mu) w(\rho,\mu).$$

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 codes known to us taken from [10], [20], and [21]. The best known upper is given

TABLE IX

GRAIN-ERROR CHANNEL COMPARISON

n	MB [10], [13]	ASPV	Theorem 23	Best UP	LB
5	12	10	12	8 [20]	8 [20]
6	20	18	20	16 [20]	16 [20]
7	36	32	34	26 [20]	26 [20]
8	62	56	60	44 [20]	44 [20]
9	112	102	108	88 [21]	72 [21]
10	204	186	196	176 [21]	112 [21]
11	372	341	358	352 [21]	210 [10]
12	682	630	656	682	372 [21]
13	1260	1170	1212	1260	702 [10]
14	2340	2184	2250	2304 [22]	1272 [21]
15	4368	4096	4202	4368	2400 [10]
16	8190	7710	7882	8190	4522 [21]
17	15420	14563	14844	15420	8428 [21]
18	29126	27594	28058	29126	15348 [10]
19	55188	52428	53202	55188	27596 [10]
20	104856	99864	101162	104856	52432 [10]
21	199728	190650	192850	199728	99880 [10]
22	381300	364722	368478	381300	190652 [10]
23	729444	699050	705510	729444	364724 [10]

in the column before last. For $5 \leqslant n \leqslant 11$ it is taken from [20] and [21], and for $n = 14$ is a result from [22]. All other values in this column are the ones calculated by the monotonicity upper bound [10], [13]. For $12 \leqslant n \leqslant 23$, Theorem 23 improves on the best known so far, that is the monotonicity upper bound and the result for $n = 14$ from [22].

VII. PROJECTIVE SPACES

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

Koetter and Kschischang [15] modeled codes as subsets of projective space $\mathbb{F}_q^n$, 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 $\mathbb{F}_q^n$ 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 the 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

$$\mathcal{B}_{P,r}(x) = \{y \in \mathcal{X}_P : d_P(x, y) \leqslant r\}.$$

The corresponding hypergraph is $\mathcal{H}(\mathcal{G}_P, r) = (X_{P,r}, \mathcal{E}_{P,r})$, such that $X_{P,r} = X_P$ and $\mathcal{E}_{P,r} = \{\mathcal{B}_{P,r}(x) : x \in X_P\}$. The generalized sphere packing bound becomes

$$\tau^*(\mathcal{H}(\mathcal{G}_P, r)) = \min\left\{\sum_{x \in X_P} w(x) : \forall x \in \mathcal{X}_P, \sum_{y \in \mathcal{B}_{P,r}(x)} w_y \geqslant 1, w_x \geqslant 0\right\}.$$

Assume x_1 and x_2 are elements in X_P with same dimension k. There exist an injective linear transform $\mathcal{T} : \mathbb{F}_q^n \to \mathbb{F}_q^n$ mapping the basis of x_1 into a basis for x_2. Note that $x \subset y$ if and only if $\mathcal{T}(x) \subset \mathcal{T}(y)$. Therefore, 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 one to another and they are 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 $\mathcal{B}_{P,r}(x)$. 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 $\begin{bmatrix} k \\ k-1 \end{bmatrix}_2 = 2^k - 1$ subspaces of dimension $k - 1$ in $\mathcal{B}_{P,1}(x)$, where

$$\begin{bmatrix} n \\ m \end{bmatrix}_2 = \frac{(2^n - 1)(2^{n-1} - 1) \cdots (2^{n-m+1} - 1)}{(2^m - 1)(2^{m-1} - 1) \cdots (2^1 - 1)}$$

is the number of subspaces of dimension m in a space of dimension n. There are also $\frac{2^n - 2^k}{2^k} = 2^{n-k} - 1$ subspaces of dimension $k + 1$ in $\mathcal{B}_{P,1}(x)$ that include x. Therefore, there are $(2^k - 1) + (2^{n-k} - 1) + 1$ elements in $\mathcal{B}_{P,1}(x)$. And,

$$\tau^*(\mathcal{H}(\mathcal{G}_P, 1)) = \min\left\{\sum_{k=0}^{n} w_k \begin{bmatrix} n \\ k \end{bmatrix}_2 : \forall\, 0 \leqslant k \leqslant n, w_k + (2^k - 1)w_{k-1} + (2^{n-k} - 1)w_{k+1} \geqslant 1, w_k \geqslant 0\right\}. \quad (16)$$

It is also shown that there exist automorphisms mapping a fixed subspace of dimension k to a fixed subspace of dimension $n - k$ (see [3].) So, subspaces of dimension k and $n - k$ are also in same equivalence classes and we assign same weights to them. Note that $\begin{bmatrix} n \\ k \end{bmatrix}_2 = \begin{bmatrix} n \\ n-k \end{bmatrix}_2$. So, we can benefit from the symmetry and we set $w_k = w_{n-k}$ to halve both the number of constraints and the number of parameters in the linear programming.

TABLE X

PROJECTIVE CODES: UPPER BOUNDS AND WEIGHTS FOR $r = 1$

n	$w_0^*, w_1^*, \cdots, w_{\lfloor \frac{n}{2} \rfloor}^*$	ASPV	GSPB	[1]
2	1, 0	1	1	-
3	1, 0	3	2	-
4	0.83, 0.17, 0	8	6	6
5	0.67, 0.34, 0	30	22	20
6	0, 0.30, 0.07, 0	159	132	124
7	0, 0.29, 0.15, 0	1142	834	776
8	1, 0, 0.14, 0.03, 0	11364	9460	9268
9	1, 0, 0.13, 0.07, 0	157860	116656	107419
10	1, 0, 0, 0.066, 0.016, 0	3073031	2566390	-
11	1, 0, 0, 0.065, 0.032, 0	84047153	62462160	-

Optimal transversal weights for $n \leqslant 11$ are listed in Table X. To avoid repetition we only list the first half of the optimal transversal weight in the second column. The third column is the average sphere packing value (ASPV) and the fourth column is the generalized sphere packing bound, which together validate the ASPV as an upper bound on the size of the code for $n \leqslant 11$ and $r = 1$. The last column is the best known upper bound from [1], which uses the semidefinite programming to improve the best previously known upper bounds from [7] for projective codes. While the GSPB is not smaller than the already existing upper bound, it is easy to derive it for all n as the linear programming in this case is solved and the optimal solution is given, see (17) below.

It is interesting to see that $w_{\lfloor \frac{n}{2} \rfloor} = 0$ for all $n > 2$, which is not surprising since $\begin{bmatrix} n \\ \lfloor \frac{n}{2} \rfloor \end{bmatrix}_2$ is the largest multiplier 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. $w_{\lfloor \frac{n}{2} \rfloor} = 0$; and moving toward the tails by picking the least possible value that satisfies the constraints. We call it as the ***greedy*** weight assignment, which is expressed as

$$\begin{aligned} & w^*_{\lceil \frac{n}{2} \rceil} = 0, \\ & \text{and for all } k\text{:} \quad 0 \leqslant k < \lfloor \frac{n}{2} \rfloor, \\ & w^*_k = \max\left\{ \frac{1 - w^*_{k+1} - (2^{n-k-1} - 1) w^*_{k+2}}{2^{k+1} - 1}, 0 \right\}; \qquad (17) \\ & \text{But, if } w^*_0 = w^*_1 = 0, \text{ we reset } w^*_0 \text{ to } 1. \\ & \text{Finally, we use the symmetry and set } w^*_k = w^*_{n-k}. \end{aligned}$$

It is clear that the greedy output has the transversal property and lies in the feasible set. The following theorem also shows the optimality of the greedy assignment in (17) (See Appendix B for the proof.)

Theorem 24: Let $\mathcal{G}_P$ be the associated graph with projective code when $\mathbb{F}_q^n$ is the space and $\mathbf{w}^*$ be defined as (17), then

$$\tau^*(\mathcal{H}(\mathcal{G}_P, 1)) = \sum_{k=0}^{n} w^*_k \begin{bmatrix} n \\ k \end{bmatrix}_2.$$

VIII. Conclusions and Discussion

This paper follows up on a recent work by Kulkani and Kiyavash for deriving upper bounds on the cardinality of deletion-correcting codes. We study this method in order to give a generalization of the sphere packing bound. This scheme can provide an upper bound on the cardinality of codes according to any error channel. The main challenge in deriving this upper bound is the solution of a linear programming problem, which in 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 lastly also mention that other error channels can be studied as well using the method presented in the paper.

Appendix A
Optimal Transversal Weight for Z Channel

In this section, we first complete the proof of the transversal property for the proposed transversal weight assignment $\mathbf{w}^*$ in (8) by using Lemma 25 to prove Theorem 10. Next, we show that this choice of $\mathbf{w}^*$ is also the optimal transversal assignment for the Z Channel and hence the generalized sphere packing bounds in Tables I, II, III, and VI are the bests one can get with this method.

Lemma 25: For all $r \in \mathbb{N}$, $|D_m| \leqslant (2r)^{m-r+1}$.

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

$$\begin{aligned} D_m &= -\left(\frac{r!}{(r-1)!} D_{m-1} + \frac{r!}{(r-2)!} D_{m-2} + \cdots + \frac{r!}{0!} D_{m-r} \right) \\ &\leqslant \frac{r!|D_{m-1}|}{(r-1)!} + \frac{r!|D_{m-2}|}{(r-2)!} + \cdots + \frac{r!|D_{m-r}|}{0!} \\ &\leqslant \frac{r!(2r)^{m-r}}{(r-1)!} + \frac{r!(2r)^{m-r-1}}{(r-2)!} + \cdots + \frac{r!(2r)^{m-2r+1}}{0!} \\ &= (2r)^{m-r+1} \left(\frac{r!(2r)^{-1}}{(r-1)!} + \frac{r!(2r)^{-2}}{(r-2)!} + \cdots + \frac{r!(2r)^{-r}}{0!} \right) \\ &\leqslant (2r)^{m-r+1} (2^{-1} + 2^{-2} + \cdots + 2^{-r}) \\ &\leqslant (2r)^{m-r+1}. \end{aligned}$$

■

Now, we are ready to prove Theorem 10.

Theorem 10 (Recall): Let $\mathbf{w}^*$ be the transversal weight assignment defined in (8). For all $k \geqslant 3r-1$ we have

$$w^*_k \geqslant \tfrac{r!k!}{(r+k)!} 2^{-(n-k-r)} > 0.$$

Proof:

$$\begin{aligned} w^*_k &= r!k! \sum_{m=r+k}^{n} \frac{D_{m-k-1}}{m!} \\ &= r!k! \left(\frac{1}{(r+k)!} + \sum_{m=r+k+1}^{n} \frac{D_{m-k-1}}{m!} \right) \\ &\geqslant \frac{r!k!}{(r+k)!} \left(1 - \sum_{m=r+k+1}^{n} \frac{|D_{m-k-1}|(r+k)!}{m!} \right) \\ &\geqslant \frac{r!k!}{(r+k)!} \left(1 - \sum_{m=r+k+1}^{n} \frac{|D_{m-k-1}|}{(4r)^{m-r-k}} \right) \\ &\geqslant \frac{r!k!}{(r+k)!} \left(1 - \sum_{m=r+k+1}^{n} \frac{(2r)^{m-r-k}}{(4r)^{m-r-k}} \right) \\ &= \frac{r!k!}{(r+k)!} \left(1 - \sum_{m=1}^{n-k-r} 2^{-m} \right) = \frac{r!k!}{(r+k)!} 2^{-(n-k-r)} > 0, \end{aligned} \tag{18}$$

where (18) comes from Lemma 25. In other words, for $k \geqslant 3r-1$, the first term in (8) is greater than the sum of the absolute values of the remaining terms and they cannot cancel it out. ■

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

Let us define the cost functions $f_i(\mathbf{w})$ for all $0 \leqslant i \leqslant n$ as

$$\begin{aligned} f_k(\mathbf{w}) &= w_0 && k = 0, \\ f_k(\mathbf{w}) &= \sum_{i=0}^{r} w_{k+i} \binom{k+r}{r-i} && \forall 1 \leqslant k \leqslant n-r, \\ f_k(\mathbf{w}) &= w_k && \forall n-r < k \leqslant n. \end{aligned}$$

From (7), any feasible $\mathbf{w}$ should satisfy

$$\begin{aligned} f_k(\mathbf{w}) &\geqslant 1 && k=0,\\ f_k(\mathbf{w}) &\geqslant 1 && \forall 1 \leqslant k \leqslant n-r,\\ f_k(\mathbf{w}) &\geqslant 0 && \forall n-r < k \leqslant n; \end{aligned}$$

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

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

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

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

where y_i's are some non-negative constants. Hence, for any transversal weight $\mathbf{w}$ we have

$$\begin{aligned} f(\mathbf{w}) &\geqslant y_0 f_0(\mathbf{w}^*) + y_1 f_1(\mathbf{w}^*) + \cdots + y_n f_n(\mathbf{w}^*)\\ &= y_0 + y_1 + \cdots + y_{n-r} = f(\mathbf{w}^*). \end{aligned}$$

We first show that the choice of $\mathbf{y}$ is unique. Then the problem reduces to show the non-negativity of $\mathbf{y}$. Note that the cost functions $f_i(\mathbf{w})$ are inner products of $\mathbf{w}$ with some non-negative vectors $\mathbf{m}_i$, i.e. $f_i(\mathbf{w}) =< \mathbf{m}_i, \mathbf{w} >$, where

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

Now, if we form the $(n+1) \times (n+1)$ matrix M with elements $\{m_{i,j}\}$, the problem of finding $\mathbf{y}$ will be equivalent to solving $M^T\mathbf{y} = \mathbf{c}$, where $\mathbf{c} = (\binom{n}{0}, \binom{n}{1}, \cdots, \binom{n}{n})$. 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:

$$\begin{aligned} &y_0 = 1,\\ &\forall i:\ 1 \leqslant i \leqslant n-r,\\ &y_i = \frac{1}{\binom{r+i}{i}}\left(\binom{n}{i} - \sum_{j=\max\{i-r,1\}}^{i-1} \binom{r+j}{i} y_j\right); \end{aligned} \quad (19)$$

And,

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

Lemma 26 gives an explicit formula for y_k when $0 \leqslant k \leqslant n-r$, which again uses the sequence D_i defined in (10). We present the proof at the end of this subsection.

Lemma 26: If $\mathbf{y}$ is the solution to $M^T\mathbf{y} = \mathbf{c}$, then

$$y_0 = 1, \text{ and for all } 1 \leqslant k \leqslant n-r \text{ we have}$$
$$y_k = \frac{r!}{(k+r)!} \sum_{m=1}^{k} \frac{n!}{(n-m)!} D_{k-m+r-1}.$$

Let us benefit from a simple change of variables to get

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

which is proven to be positive if $k \geqslant 4r-1$ by using the same argument as $w_k^* > 0$ when $k \geqslant 3r-1$ in (18). Also, we can verify $z_k \geqslant 0$ for the values of $r \leqslant k \leqslant 4r-2$ by finding some numbers n'_k such that

$$\sum_{m=k}^{n'_k} \frac{D_{m+r-k-1}}{m!} \geqslant \frac{1}{(2r)^k}\left(e^{2r} - \sum_{m=0}^{n_k} \frac{(2r)^m}{m!}\right),$$

and we get $z_k > 0$ for all $n > n'_k$ because

$$\begin{aligned} z_k &= \sum_{m=k}^{n-1} \frac{D_{m+r-k-1}}{m!}\\ &= \sum_{m=k}^{n_k} \frac{D_{m+r-k-1}}{m!} + \sum_{m=n_k+1}^{n-1} \frac{D_{m+r-k-1}}{m!}\\ &\geqslant \sum_{m=k}^{n_k} \frac{D_{m+r-k-1}}{m!} - \sum_{m=n_k+1}^{n-1} \frac{(2r)^{m-k}}{m!}\\ &> \sum_{m=k}^{n_k} \frac{D_{m+r-k-1}}{m!} - \frac{1}{(2r)^k} \sum_{m=n_k+1}^{\infty} \frac{(2r)^m}{m!}\\ &= \sum_{m=k}^{n_k} \frac{D_{m+r-k-1}}{m!} - \frac{e^{2r} - \sum_{m=0}^{n_k} \frac{(2r)^m}{m!}}{(2r)^k} \geqslant 0. \end{aligned}$$

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

In order to show $y_k \geqslant 0$ for $k > n-r$, we rewrite the expression for y_k when $k > n-r$ as

$$\begin{aligned} y_k &= \binom{n}{k} - \sum_{j=\max\{k-r,1\}}^{n-r} \binom{r+j}{k} y_j\\ &= \binom{n}{k} - y_{n-r}\binom{n}{k}\\ &\quad - y_{n-r-1}\binom{n-1}{k} - \cdots - y_{n-r-p}\binom{n-p}{k}, \end{aligned}$$

where $n-r-p = \max\{k-r,1\}$. So, $y_k \geqslant 0$ is equivalent to

$$y_{n-r} \leqslant 1 - \frac{\binom{n-1}{k}}{\binom{n}{k}} y_{n-r-1} - \frac{\binom{n-2}{k}}{\binom{n}{k}} y_{n-r-2} - \cdots - \frac{\binom{n-p}{k}}{\binom{n}{k}} y_{n-r-p}.$$

By assuming $y_i \geqslant 0$ for all $i \leqslant n-r$ and the recursive formula for y_k's, we have

$$\begin{aligned} y_{n-r} &= 1 - \frac{\binom{n-1}{n-t}}{\binom{n}{n-r}} y_{n-r-1} - \frac{\binom{n-2}{n-r}}{\binom{n}{n-r}} y_{n-r-2} - \cdots - \frac{\binom{n-q}{n-r}}{\binom{n}{n-r}} y_{n-r-q}\\ &\leqslant 1 - \frac{\binom{n-1}{n-t}}{\binom{n}{n-r}} y_{n-r-1} - \frac{\binom{n-2}{n-r}}{\binom{n}{n-r}} y_{n-r-2} - \cdots - \frac{\binom{n-p}{n-r}}{\binom{n}{n-r}} y_{n-r-p}, \end{aligned}$$

where $n-r-q = \max\{n-2r, 1\} \leqslant \max\{k-r, 1\} = n-r-p$. Now, it suffices to show

$$\frac{\binom{n-\ell}{k}}{\binom{n}{k}} \leqslant \frac{\binom{n-\ell}{n-r}}{\binom{n}{n-r}} \iff \frac{(n-k)!}{(n-\ell-k)!} \leqslant \frac{r!}{(r-\ell)!}$$
$$\iff (n-k)(n-k-1)\cdots(n-k-\ell) \leqslant r(r-1)\cdots(r-\ell),$$

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 $z_k \geqslant 0$ for all $r \leqslant k \leqslant 4r-2$ 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 $r \leqslant 20$.

Proof of Lemma 26: We define sequences $\{y_i^j(n)\}$ for $1 \leqslant j \leqslant n-r$ as

$$y_i^j(n) = 0, \qquad 1 \leqslant i < j,$$
$$y_i^j(n) = \binom{n}{j} \Big/ \binom{r+j}{j} = \frac{n!r!}{(r+j)!(n-j)!}, \qquad i = j,$$
$$y_i^j(n) = -\sum_{\ell=\max\{i-r,1\}}^{i-1} \frac{\binom{r+\ell}{i}}{\binom{r+i}{i}} y_\ell^j(n) \qquad j < i \leqslant n-r;$$

And, define $y_i'(n)$ as $y_i'(n) = \sum_{j=1}^{n-r} y_i^j(n)$. It is easy to see that for all $1 \leqslant i \leqslant n-r$ we have

$$y_i'(n) = \frac{1}{\binom{r+i}{i}} \left(\binom{n}{i} - \sum_{\ell=\max\{i-r,1\}}^{i-1} \binom{r+\ell}{i} y_j'(n) \right).$$

So, y_i' is nothing but the exact sequence y_i defined in (19). Now we use the variable change

$$\delta_i^j(n) = y_i^j(n)(r+i)!\frac{(n-j)!}{n!r!}$$

to get the recursion for $\delta_i^j(n)$ as

$$\delta_i^j(n) = 0, \qquad 1 \leqslant i < j,$$
$$\delta_i^j(n) = 1, \qquad i = j,$$
$$\frac{\delta_i^j(n)}{r!} + \frac{\delta_{i-1}^j(n)}{(r-1)!} + \cdots + \frac{\delta_{i-r}^j(n)}{0!} = 0, \quad j < i \leqslant n-r.$$

We observe that $\delta_i^j(n)$ is also equal to $D_{i-j+r-1}$ defined in (10). Substituting this value in the formula for y' gives the desired:

$$y_i = \sum_{j=1}^{n-r} y_i^j(n) = \sum_{j=1}^{n-r} \delta_i^j(n) \frac{n!r!}{(r+i)!(n-j)!}$$
$$= \frac{r!}{(r+i)!} \sum_{j=1}^{n-r} \frac{n!}{(n-j)!} D_{i-j+r-1}$$
$$= \frac{r!}{(r+i)!} \sum_{j=1}^{i} \frac{n!}{(n-j)!} D_{i-j+r-1}.$$

Appendix B
Optimal Transversal Weight for Projective Codes

The idea behind the proof is to again write the cost function $f(\mathbf{w}) = \sum_{k=0}^n w_k \begin{bmatrix} n \\ k \end{bmatrix}_2$ as a non-negative linear combination of some other cost functions $f_k(\mathbf{w})$, where $\mathbf{w}^*$ minimizes them all and so minimizes the cost function $f(\mathbf{w})$. We benefit from the symmetry in the optimal transversal i.e. $w_k^* = w_{n-k}^*$ so we limit the discussion to $k \leqslant \lfloor \frac{n}{2} \rfloor$. Let us define the cost functions $f_k(\mathbf{w})$ as

$$f_k(\mathbf{w}) = w_{k+1} + (2^{k+1}-1)w_k + (2^{n-k-1}-1)w_{k+2},$$

if $k \equiv \lfloor n/2 \rfloor - 1 (\text{mod} 4)$ or $k \equiv \lfloor n/2 \rfloor - 2 (\text{mod} 4)$ and otherwise

$$f_k(\mathbf{w}) = w_k.$$

with the only exceptions of

$$f_k(\mathbf{w}) = w_{k-1} + (2^{n-k+1}-1)w_k + (2^{k-1}-1)w_{k-2}, \tag{20}$$

if $k = \frac{n}{2}$ for n even, and

$$f_k(\mathbf{w}) = w_0 + w_1(2^n - 1), \tag{21}$$

if $k = 0$ and $\lfloor \frac{n}{2} \rfloor \not\equiv 1$ or 2 (mod4).

The idea is to write $f(\mathbf{w}) = \sum_{k=0}^n w_k \begin{bmatrix} n \\ k \end{bmatrix}_2$ as $\sum_{k=0}^n y_k f_k(\mathbf{w})$, where y_k's are some fixed non-negative real numbers. Note that all cost functions $f_k(\mathbf{w})$ take their minimum values on $\mathbf{w}^*$. So, the non-negativity of y_k (for all k) is enough to have the optimality of $\mathbf{w}^*$ for $f(\mathbf{w})$.

Assume that k is given such that

$$k \equiv \lfloor \frac{n}{2} \rfloor + 1 \mod 4,$$

and non of the indices in $\{k, k-1, k-2, k-3\}$ fall into the two exceptional categories in (20) and (21). We have,

$$f_k(\mathbf{w}) = w_k,$$
$$f_{k-1}(\mathbf{w}) = w_{k-1},$$
$$f_{k-2}(\mathbf{w}) = w_{k-2}(2^{k-1}-1) + w_{k-1} + w_k(2^{n-k+1}-1),$$
$$f_{k-3}(\mathbf{w}) = w_{k-3}(2^{k-2}-1) + w_{k-2} + w_{k-1}(2^{n-k+2}-1).$$

It is easy to see that $\{w_k, w_{k-1}, w_{k-2}, w_{k-3}\}$ show up only in these $f_i(\mathbf{w})$'s. Comparing the corresponding coefficients in $f(\mathbf{w})$ results in the following system of equations for y_k, y_{k-1}, y_{k-2}, and y_{k-4}:

$$\begin{bmatrix} n \\ k \end{bmatrix}_2 = y_k + y_{k-2}(2^{n-k+1}-1)$$
$$\begin{bmatrix} n \\ k-1 \end{bmatrix}_2 = y_{k-1} + y_{k-2} + y_{k-3}(2^{n-k+2}-1)$$
$$\begin{bmatrix} n \\ k-2 \end{bmatrix}_2 = y_{k-2}(2^{k-1}-1) + y_{k-3}$$
$$\begin{bmatrix} n \\ k-3 \end{bmatrix}_2 = y_{k-3}(2^{k-2}-1).$$

And, the solution to the system of equations is given by

$$y_{k-3} = \begin{bmatrix} n \\ k-3 \end{bmatrix}_2 / (2^{k-2}-1) \geqslant 0$$
$$y_{k-2} = \left(\begin{bmatrix} n \\ k-2 \end{bmatrix}_2 - y_{k-3} \right) / (2^{k-1}-2) \geqslant 0$$
$$y_{k-1} = \begin{bmatrix} n \\ k-1 \end{bmatrix}_2 - y_{k-2} - y_{k-3}(2^{n-k+2}-1) \geqslant 0$$
$$y_k = \begin{bmatrix} n \\ k \end{bmatrix}_2 - y_{k-2}(2^{n-k+1}-1) \geqslant 0.$$

The story is not much different on the edges i.e. where y_0 or $y_{\lfloor \frac{n}{2} \rfloor}$ are involved. Due to the symmetry, the non-negativity of the first half of the $\mathbf{y}$ is followed by the non-negativity of the other half. So y_k $(\forall k)$ are all non-negative and $\mathbf{w}^*$ is the optimal transversal weight.

ACKNOWLEDGMENT

The authors thank the three anonymous reviewers for their valuable comments and suggestions, which have contributed for the clarity of the paper and its presentation.

REFERENCES

[1] C. Bachoc, A. Passuello, and F. Vallentin, "Bounds for projective codes from semidefinite programming," *Adv. Math. Commun.*, vol. 7, no. 2, pp. 127–145, May 2013.
[2] C. Berge, "Packing problems and hypergraph theory: A survey," *Ann. Discrete Math.*, vol. 4, pp. 3–37, 1979.
[3] M. Braun, T. Etzion, and A. Vardy, "Linearity and complements in projective space," *Linear Algebra Appl.*, vol. 438, no. 1, pp. 57–70, Jan. 2013.
[4] S. Buzaglo, E. Yaakobi, T. Etzion, and J. Bruck, "Error-correcting codes for multipermutations," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 724–728.
[5] Y. Cassuto, M. Schwartz, V. Bohossian, and J. Bruck, "Codes for asymmetric limited-magnitude errors with application to multilevel flash memories," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1582–1595, Apr. 2010.
[6] N. Elarief and B. Bose, "Optimal, systematic, q-ary codes correcting all asymmetric and symmetric errors of limited magnitude," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 979–983, Mar. 2010.
[7] T. Etzion and A. Vardy, "Error-correcting codes in projective space," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 1165–1173, Feb. 2011.
[8] A. Fazeli, A. Vardy, and E. Yaakobi, "Generalized sphere packing bound: Applications," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 1261–1265.
[9] A. Fazeli, A. Vardy, and E. Yaakobi, "Generalized sphere packing bound: Basic principles," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 1256–1260.
[10] R. Gabrys, E. Yaakobi, and L. Dolecek, "Correcting grain-errors in magnetic media," *IEEE Trans. Inf. Theory*, to be published.
[11] E. N. Gilbert, "A comparison of signalling alphabets," *Bell Syst. Tech. J.*, vol. 31, no. 3, pp. 504–522, 1952.
[12] A. R. Iyengar, P. H. Siegel, and J. K. Wolf, "Write channel model for bit-patterned media recording," *IEEE Trans. Magn.*, vol. 47, no. 1, pp. 35–45, Jan. 2011.
[13] N. Kashyap and G. Zémor, "Upper bounds on the size of grain-correcting codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4699–4709, Aug. 2014.
[14] T. Kløve, B. Bose, and N. Elarief, "Systematic single limited magnitude asymmetric error correcting codes," in *Proc. IEEE Inf. Theory Workshop*, Cairo, Egypt, Jan. 2010, p. 1.
[15] R. Koetter and F. R. Kschischang, "Coding for errors and erasures in random network coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3579–3591, Aug. 2008.
[16] A. A. Kulkarni and N. Kiyavash, "Nonasymptotic upper bounds for deletion correcting codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 8, pp. 5115–5130, Aug. 2013.
[17] A. Mazumdar, A. Barg, and N. Kashyap, "Coding for high-density recording on a 1-D granular magnetic medium," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7403–7417, Nov. 2011.
[18] M. Schwartz, "On the non-existence of lattice tilings by quasi-crosses," *Eur. J. Combinat.*, vol. 36, pp. 130–142, Feb. 2014.
[19] F. F. Sellers, Jr., "Bit loss and gain correction code," *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 35–38, Jan. 1962.
[20] A. Sharov and R. M. Roth, "Bounds and constructions for granular media coding," in *Proc. IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, Jul./Aug. 2011, pp. 2343–2347.
[21] A. Sharov and R. M. Roth, "Improved bounds and constructions for granular media coding," in *Proc. 51st Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Oct. 2013, pp. 637–644.
[22] A. Sharov and R. M. Roth, "New upper bounds for grain-correcting and grain-detecting codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 1121–1125.
[23] L. M. G. M. Tolhuizen, "The generalized Gilbert–Varshamov bound is implied by Turan's theorem [code construction]," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1605–1606, Sep. 1997.
[24] R. R. Varshamov, "Estimate of the number of signals in error correcting codes," *Doklady Akademii Nauk SSSR*, vol. 117, pp. 739–741, 1957.
[25] R. R. Varshamov and G. M. Tenengolts, "Codes which correct single asymmetric errors," (in Russian), *Autom. Telemekhanika*, vol. 26, no. 2, pp. 288–292, 1965.
[26] J. H. Weber, C. de Vroedt, and D. E. Boekee, "Bounds and constructions for binary codes of length less than 24 and asymmetric distance less than 6," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 1321–1331, Sep. 1988.
[27] J. H. Weber, C. de Vroedt, and D. E. Boekee, "New upper bounds on the size of codes correcting asymmetric errors (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 33, no. 3, pp. 434–437, May 1987.
[28] R. L. White, R. M. H. Newt, and R. F. W. Pease, "Patterned media: A viable route to 50 Gbit/in^2 and up for magnetic recording?" *IEEE Trans. Magn.*, vol. 33, no. 1, pp. 990–995, Jan. 1997.
[29] R. Wood, M. Williams, A. Kavcic, and J. Miles, "The feasibility of magnetic recording at 10 terabits per square inch on conventional media," *IEEE Trans. Magn.*, vol. 45, no. 2, pp. 917–923, Feb. 2009.
[30] E. Yaakobi, J. Ma, L. Grupp, P. H. Siegel, S. Swanson, and J. K. Wolf, "Error characterization and coding schemes for flash memories," in *Proc. Workshop Appl. Commun. Theory Emerg. Memory Technol.*, Miami, FL, USA, Dec. 2010, pp. 1856–1860.

Arman Fazeli (S'14) was born in Tehran, Iran, in 1989. He received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2012. Since then, he has been working towards his Ph.D. degree in the electrical and computer engineering at the University of California, San Diego, where he is supervised by Prof. Alexander Vardy. His current research interests include information and coding theory, with particular emphasis on coding for distributed storage systems. Arman received silver and bronze medals at the International Mathematical Olympiad (IMO) in 2006 and 2007, when he was in Iran national mathematics team.

Alexander Vardy (S'88–M'91–SM'94–F'98) was born in Moscow, U.S.S.R., in 1963. He earned his B.Sc. (summa cum laude) from the Technion, Israel, in 1985, and Ph.D. from the Tel-Aviv University, Israel, in 1991. During 1985-1990 he was with the Israeli Air Force, where he worked on electronic counter measures systems and algorithms. During the years 1992 and 1993 he was a Visiting Scientist at the IBM Almaden Research Center, in San Jose, CA. From 1993 to 1998, he was with the University of Illinois at Urbana-Champaign, first as an Assistant Professor then as an Associate Professor. Since 1998, he has been with the University of California San Diego (UCSD), where he is the Jack Keil Wolf Endowed Chair Professor in the Department of Electrical and Computer Engineering, with joint appointments in the Department of Computer Science and the Department of Mathematics. While on sabbatical from UCSD, he has held long-term visiting appointments with CNRS, France, the EPFL, Switzerland, and the Technion, Israel.

His research interests include error-correcting codes, algebraic and iterative decoding algorithms, lattices and sphere packings, coding for digital media, cryptography and computational complexity theory, and fun math problems.

He received an IBM Invention Achievement Award in 1993, and NSF Research Initiation and CAREER awards in 1994 and 1995. In 1996, he was appointed Fellow in the Center for Advanced Study at the University of Illinois, and received the Xerox Award for faculty research. In the same year, he became a Fellow of the Packard Foundation. He received the IEEE Information Theory Society Paper Award (jointly with Ralf Koetter) for the year 2004. In 2005, he received the Fulbright Senior Scholar Fellowship, and the Best Paper Award at the IEEE Symposium on Foundations of Computer Science (FOCS). During 1995-1998, he was an Associate Editor for *Coding Theory* and during 1998-2001, he was the Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY. From 2003 to 2009, he was an Editor for the *SIAM Journal on Discrete Mathematics*. He is currently serving on the Executive Editorial Board for the IEEE TRANSACTIONS ON INFORMATION THEORY. He has been a member of the Board of Governors of the IEEE Information Theory Society during 1998-2006, and again starting in 2011.

Eitan Yaakobi (S'07–M'12) is an Assistant Professor at the Computer Science Department at the Technion Israel Institute of Technology. He received the B.A. degrees in computer science and mathematics, and the M.Sc. degree in computer science from the Technion - Israel Institute of Technology, Haifa, Israel, in 2005 and 2007, respectively, and the Ph.D. degree in electrical engineering from the University of California, San Diego, in 2011. Between 2011-2013, he was a postdoctoral researcher in the department of Electrical Engineering at the California Institute of Technology. His research interests include information and coding theory with applications to non-volatile memories, associative memories, data storage and retrieval, and voting theory. He received the Marconi Society Young Scholar in 2009 and the Intel Ph.D. Fellowship in 2010-2011.