

# Asymptotically good homological error correcting codes

Research Article

Jason McCullough\*, Heather Newman

**Abstract:** Let  $\Delta$  be an abstract simplicial complex. We study classical homological error correcting codes associated to  $\Delta$ , which generalize the cycle codes of simple graphs. It is well-known that cycle codes of graphs do not yield asymptotically good families of codes. We show that asymptotically good families of codes do exist for homological codes associated to simplicial complexes of dimension at least 2. We also prove general bounds and formulas for (co-)cycle and (co-)boundary codes for arbitrary simplicial complexes over arbitrary fields.

**2010 MSC:** 94B25, 05E45

**Keywords:** Error correcting codes, Simplicial complexes, Simplicial homology

## 1. Introduction

Fix a field (usually finite)  $\mathbb{K}$ . Given an abstract simplicial complex  $\Delta$ , one can define the reduced chain complex of  $\Delta$  with coefficients in  $\mathbb{K}$ , which naturally defines the cycles, boundaries and homology of  $\Delta$ . These are the classical homological codes, seemingly first studied by Salzer [10]. In the case when  $\Delta$  is a one-dimensional simplicial complex,  $\Delta$  may be viewed as a simple graph. Since the Moore bound gives a logarithmic bound on the girth of a graph in terms of the number of edges and vertices, it follows that one cannot define asymptotically good families of codes from cycle codes of graphs. By applying recent work in [4] and [2], we show in Theorem 4.2 that one can define asymptotically good families of codes from cycle codes of simplicial complexes of dimension two or higher.

The paper is organized as follows: Section 2 gathers preliminary results, definitions and related work. Section 3 contains our results about arbitrary homological codes over arbitrary fields. In it we recover the parameters for the  $i$ th cycle and boundary codes (and their duals) of a simplex of arbitrary dimension. This was previously done in [10] and [13] over  $\mathbb{F}_2$ . Section 4 contains our main results on the existence of asymptotically good homological codes.

\* This author was supported by a grant from the Simons Foundation (576107, JGM).

Jason McCullough (Corresponding Author); Department of Mathematics, Iowa State University, Ames, Iowa 50011, United States (email: [jmccullo@iastate.edu](mailto:jmccullo@iastate.edu)).

Heather Newman; Department of Mathematics, Drexel University, Philadelphia, PA 19104, United States (email: [hn385@drexel.edu](mailto:hn385@drexel.edu)).

## 2. Preliminaries

### 2.1. Simplicial complexes

Let  $\Delta$  be an abstract simplicial complex on vertex set  $[m] := \{1, \dots, m\}$ ; that is,  $\Delta$  is a collection of subsets of  $[m]$ , closed under taking subsets. Elements  $\sigma \in \Delta$  are called faces or simplices. Maximal faces are called facets. We note that  $\Delta$  is determined by its facets. For  $\sigma \in \Delta$ , we define the dimension of  $\sigma$  by setting  $\dim(\sigma) = |\sigma| - 1$ . Thus  $\dim(\emptyset) = -1$ . Then the dimension of  $\Delta$  is  $\dim(\Delta) = \max\{\dim(\sigma) \mid \sigma \in \Delta\}$ . We denote by  $\Delta^i$  the set of  $i$ -dimensional faces of  $\Delta$ , called the  $i$ -skeleton of  $\Delta$ , and call an element of  $\Delta^i$  an  $i$ -face. An  $m$ -simplex is then a simplicial complex on  $[m + 1]$  consisting of all possible facets with one  $m$ -dimensional facet. Set  $f_i(\Delta)$  to be the number of  $i$ -faces, that is,  $f_i(\Delta) = |\Delta^i|$ . The vector  $(f_{-1}(\Delta), f_0(\Delta), f_1(\Delta), \dots, f_{\dim(\Delta)}(\Delta))$  is called the  $f$ -vector of  $\Delta$ . We sometimes refer to 0-faces as vertices and 1-faces as edges. We say that an  $i$ -face  $\sigma \in \Delta^i$  has degree  $r$  if  $\sigma$  is contained in exactly  $r$  many  $(i + 1)$ -faces of  $\Delta$ ; that is,  $\deg(\sigma) = |\{\tau \in \Delta^{i+1} : \sigma \subset \tau\}|$ . Finally we define the  $i$ -degree of  $\Delta$  to be  $\deg(i, \Delta) = \min\{\deg(\sigma) \mid \sigma \in \Delta^i\}$ . Thus  $\deg(1, \Delta)$  is just the minimal degree of a vertex of the 1-skeleton of  $\Delta$  viewed as a simple graph.

Over any field  $\mathbb{K}$ , we may define a chain complex  $(C_\bullet(\Delta, \mathbb{K}), \partial_\bullet)$  by setting  $C_i(\Delta, \mathbb{K})$  to be the  $\mathbb{K}$ -vector space whose basis is identified with  $\Delta^i$  and such that given  $\sigma \in \Delta^i$ , write  $\sigma = \{j_0, j_1, \dots, j_i\} \subseteq [m]$ , where  $j_0 < \dots < j_i$ , then

$$\partial_i(\sigma) = \sum_{k=0}^i (-1)^k \{j_0, \dots, \widehat{j_k}, \dots, j_i\},$$

where  $\widehat{j_k}$  denotes that  $j_k$  has been removed. Thus  $C_{-1}(\Delta, \mathbb{K}) \cong \mathbb{K}$  as  $\emptyset$  is the only  $-1$ -dimensional face. We extend this linearly to define a map  $\partial_i : C_i(\Delta, \mathbb{K}) \rightarrow C_{i-1}(\Delta, \mathbb{K})$ . Since  $\partial_{i-1} \circ \partial_i = 0 \ \forall i$ , the resulting sequence is a chain complex of  $\mathbb{K}$ -vector spaces. Let  $Z_i(\Delta, \mathbb{K}) = \text{Ker}(\partial_i)$  and  $B_i(\Delta, \mathbb{K}) = \text{Im}(\partial_{i+1})$ . Note that  $B_i(\Delta, \mathbb{K}) \subseteq Z_i(\Delta, \mathbb{K}) \subseteq C_i(\Delta, \mathbb{K}) \ \forall i$ . Then we define the  $i$ th reduced simplicial homology of  $\Delta$  with coefficients in  $\mathbb{K}$  as

$$H_i(C_\bullet(\Delta, \mathbb{K})) = \widetilde{H}_i(\Delta, \mathbb{K}) = \frac{Z_i(\Delta, \mathbb{K})}{B_i(\Delta, \mathbb{K})}.$$

We define the co-chain complex of  $\Delta$  over  $\mathbb{K}$  to be the vector space dual  $C^\bullet(\Delta, \mathbb{K}) = (C_\bullet(\Delta, \mathbb{K}))^*$  of the chain complex, with coboundary maps  $\partial^i = \partial_i^*$ . We set  $Z^i(\Delta, \mathbb{K}) = \text{Ker}(\partial^{i+1})$  and  $B^i(\Delta, \mathbb{K}) = \text{Im}(\partial^i)$ . Then the  $i$ th reduced simplicial cohomology of  $\Delta$  over  $\mathbb{K}$  is

$$H^i(C^\bullet(\Delta, \mathbb{K})) = \widetilde{H}^i(\Delta, \mathbb{K}) = \frac{Z^i(\Delta, \mathbb{K})}{B^i(\Delta, \mathbb{K})}.$$

Since  $\text{Hom}_{\mathbb{K}}(\_, \mathbb{K})$  is exact, there is a canonical isomorphism  $\widetilde{H}_i(\Delta, \mathbb{K}) \cong \widetilde{H}^i(\Delta, \mathbb{K})$ . When the simplicial complex  $\Delta$  and the field  $\mathbb{K}$  are clear from context, we omit them from the notation, e.g. writing  $\widetilde{H}_i$  for  $\widetilde{H}_i(\Delta, \mathbb{K})$ .

Let  $S(\Delta)$  denote the suspension of the simplicial complex  $\Delta$ ; that is,  $S(\Delta)$  is the simplicial complex on vertex set  $\Delta^0 \cup \{a, b\}$ , where  $a, b$  are two new vertices. For any facet  $\sigma \in \Delta$ , we assert that  $\sigma \cup \{a\}$  and  $\sigma \cup \{b\}$  are facets of  $S(\Delta)$ . This uniquely determines  $S(\Delta)$ . The suspension can be viewed as the union of two cones of  $\Delta$  glued along their bases.

### 2.2. Linear codes

A linear code  $C$  of length  $n$  over a (usually finite) field  $\mathbb{K}$  is a  $\mathbb{K}$ -vector subspace of  $\mathbb{K}^n$ . Elements  $\mathbf{x} = (x_1, \dots, x_n) \in C$  are called codewords. We say that  $C$  is an  $[n, k, d]$ -linear code if  $C$  has length  $n$ , the dimension  $\dim_{\mathbb{K}}(C)$  is  $k$ , and the minimum Hamming distance of  $C$  is  $d$ . The minimum Hamming

distance is the minimum over all pairs  $\mathbf{x}, \mathbf{y} \in C$  of distinct codewords of the number of positions  $i$  such that  $x_i \neq y_i$ . The weight  $w(\mathbf{x})$  of a codeword  $\mathbf{x} \in C$  is the number of positions  $i$  such that  $x_i \neq 0$ . The minimum distance of a linear code  $C$  is equal to the minimum weight of a nonzero codeword of  $C$ .

A generator matrix  $G$  of a linear code  $C$  over  $\mathbb{K}$  is a matrix whose row space is  $C$ . A parity check matrix  $H$  of  $C$  is a matrix whose null space is  $C$ . In other words  $\mathbf{x} \in C$  if and only if  $H\mathbf{x}^T = \mathbf{0}$ . The dual code  $C^\perp$  of  $C$  is the code whose generator matrix is  $H$ . Thus  $G$  is a parity check matrix for  $C^\perp$ . We do not require the rows of  $G$  or  $H$  to be linearly independent.

We make use of the following well-known fact relating minimum distance of a code to its parity check matrix.

**Proposition 2.1** ([8, Lemma 1.2.3]). *Let  $H$  be a parity-check matrix for a linear code  $C$  over a field  $\mathbb{K}$ . Then every set of  $d - 1$  columns of  $H$  are linearly independent if and only if  $C$  has minimum distance at least  $d$ .*

One of the main challenges in classical error correcting code theory is to find codes with large minimum distance and dimension relative to their length. More precisely, we seek a family of asymptotically good codes. A family of codes  $C_i$  with parameters  $[n_i, k_i, d_i]$  with  $\lim_{i \rightarrow \infty} n_i \rightarrow \infty$  is *asymptotically good* if there exists  $\alpha, \beta > 0$  such that

$$\lim_{i \rightarrow \infty} \frac{k_i}{n_i} > \alpha \quad \text{and} \quad \lim_{i \rightarrow \infty} \frac{d_i}{n_i} > \beta,$$

where  $\frac{k_i}{n_i}$  and  $\frac{d_i}{n_i}$  define the information rate and relative minimum distance of  $C_i$ , respectively.

Of particular interest are Low Density Parity Check (LDPC) codes. A linear code  $C$  (usually defined over  $\mathbb{F}_2$ ) is an LDPC code if  $C$  has a parity check matrix  $H$  with relatively few nonzero entries. If the Hamming weight (number of nonzero entries) in each row and column of  $H$  is constant,  $C$  is called a regular LDPC code. If we relax one of these conditions,  $C$  is called an irregular LDPC code. The cycle codes in this paper have parity check matrices with fixed column weight (since the image of the boundary of any  $i$ -face has weight  $i + 1$ ), but only certain ones will have fixed row weight as well.

### 2.3. Homological error correcting codes

Let  $\Delta$  be an  $m$ -dimensional simplicial complex and let  $\mathbb{K}$  be a field. For each  $0 \leq i \leq m$ , we define four potentially distinct linear codes. These codes will be the main object of study in this paper. The definitions parallel those in [10] and [13]. We define the  $i$ th cycle code (respectively boundary code) as  $Z_i(\Delta, \mathbb{K})$  (resp.  $B_i(\Delta, \mathbb{K})$ ). Similarly, we define the  $i$ th cocycle code (respectively coboundary code) as  $Z^i(\Delta, \mathbb{K})$  (resp.  $B^i(\Delta, \mathbb{K})$ ). Note that all four codes have length  $f_i$  as they are subvector spaces of  $C_i$  or  $C_i^*$ . Moreover, by their definitions it is clear that  $Z^i(\Delta, \mathbb{K}) = B_i(\Delta, \mathbb{K})^\perp$  and  $B^i(\Delta, \mathbb{K}) = Z_i(\Delta, \mathbb{K})^\perp$ . In particular for all  $i$  we have,

$$\dim_{\mathbb{K}} Z^i(\Delta, \mathbb{K}) + \dim_{\mathbb{K}} B_i(\Delta, \mathbb{K}) = \dim_{\mathbb{K}} Z_i(\Delta, \mathbb{K}) + \dim_{\mathbb{K}} B^i(\Delta, \mathbb{K}) = f_i(\Delta).$$

### 2.4. Relation to previous work

The definition of LDPC codes originates in the thesis of Gallager [6], [5]. Such codes have recently attracted more attention because of the good iterative decoding algorithms associated to them which approach the Shannon limit. Asymptotically good codes can be shown to exist by pseudorandom methods and constructed algorithmically. (See e.g. [7]). Our Theorem 4.2 shows that they can be constructed as cycle codes of two-dimensional simplicial complexes.

The earliest references we can find to the definition of homological error-correcting codes of a simplicial complex are the papers by Salzer [10] and Thomeier [13], where the authors use the terms “topological codes” and “polyhedral codes,” respectively. Both describe the general construction of cycle and boundary

codes of a simplicial complex. Both authors then compute the parameters of the  $i$ th cycle codes for an  $m$ -dimensional simplex. We recover this as a special case of our more general formulas for the dimension and minimum distance. See Theorem 3.7.

In [9], Rytřř showed that any linear code over  $\mathbb{Q}$  or  $\mathbb{F}_p$  can be realized as a truncated homological code of some two dimensional simplicial complex. Without truncation, there are linear codes which cannot be realized. Our results give both restrictions on the parameters of homological codes and existence results for asymptotically good homological codes.

When  $\Delta$  is a one-dimensional simplicial complex, we may view  $\Delta$  as a simple graph  $G = (V, E)$  with vertex set  $V = \Delta^0$  and edge set  $E = \Delta^1$ . In this case  $Z_1(\Delta, \mathbb{F}_2)$  is called the cycle code of the graph  $G$  and has been well-studied. See [14] for a concise survey. It is easy to see that if  $G = (V, E)$  is a graph, then its cycle code has parameters  $[n, k, d]$ , where  $n = |E|$ ,  $k = |E| - |V| + 1$  and  $d$  is the girth of  $G$ , that is, the minimum length of a cycle in  $G$ . The well-known Moore bound states that for an  $r$ -regular graph  $G$ , that is a graph which has  $r$  edges incident to each vertex, the girth of  $G$  is bounded above by a term logarithmic in  $r$  and  $|V|$ . A similar bound for non-regular graphs was given in [1]. Thus if we have a family of graphs  $G_i = (V_i, E_i)$  with linear growth for  $|V_i|$  and  $|E_i|$ , the girth of  $G_i$  is at most logarithmic in  $|V_i|$  and  $|E_i|$ . Therefore, it is impossible to construct a family of asymptotically good error correcting codes from graphs or one-dimensional simplicial complexes. Our Theorem 4.2 shows that it is possible to do so with two-dimensional simplicial complexes.

Finally we comment that there is great interest in finding quantum error correcting codes, introduced in [11], with similar properties. A well-studied family are the CSS codes [3], [12]. Some constructions of CSS codes, also called homological codes, are defined from graphs, simplicial or polyhedral complexes. In the quantum setting, the existence of asymptotically good codes is still an open question. We do not discuss quantum codes further in this paper and refer the interested reader to the survey [14] and the references therein.

### 3. Results on arbitrary homological codes

In this section we collect general results which hold for each of the four types of homological codes for an arbitrary simplicial complex  $\Delta$  and an arbitrary field  $\mathbb{K}$ . When clear from context, we suppress  $\Delta$  and  $\mathbb{K}$  from the notation.

**Theorem 3.1.** *Let  $\Delta$  be an  $m$ -dimensional simplicial complex,  $\mathbb{K}$  a field and  $0 \leq i \leq m$ . Then  $Z_i = Z_i(\Delta, \mathbb{K})$  is an  $[n, k, d]$ -code with*

$$\begin{aligned} n &= f_i, \\ k &= \sum_{j=-1}^i (-1)^{i+j} f_j + \sum_{j=0}^{i-1} (-1)^{i+j+1} \dim_{\mathbb{K}} \tilde{H}_j, \\ d &\geq i + 2. \end{aligned}$$

**Proof.**  $Z_i$  is a sub-vector space of  $C_i$  and  $\dim_{\mathbb{K}} C_i = f_i$ . So  $n = f_i$ . By definition, the matrix  $H$  associated to  $\partial_i$  is a parity check matrix for  $Z_i$ . Since two  $i$ -faces of  $\Delta$  can only share at most one  $(i - 1)$ -face, any two columns of  $H$  can share at most one index where that entry is nonzero. Thus each set of  $i + 1$  columns of  $\partial_i$  are linearly independent. By Proposition 2.1,  $d \geq i + 2$ . For the computation of  $k$ , we proceed by induction on  $i$ . For  $i = 0$ , we have the short exact sequence

$$0 \rightarrow Z_0 \rightarrow C_0 \rightarrow C_{-1} \rightarrow 0.$$

Thus,

$$\dim_{\mathbb{K}} Z_0 = \dim_{\mathbb{K}} C_0 - \dim_{\mathbb{K}} C_{-1} = f_0 - f_{-1}$$

as desired. Now suppose

$$\dim_{\mathbb{K}} Z_p = \sum_{j=-1}^p (-1)^{j+p} f_j + \sum_{j=0}^{p-1} (-1)^{j+p+1} \dim_{\mathbb{K}} \tilde{H}_j$$

for some  $p \in \mathbb{N}$ . For  $i \geq 0$ , we have the short exact sequences

$$0 \rightarrow B_i \rightarrow Z_i \rightarrow \tilde{H}_i \rightarrow 0$$

and

$$0 \rightarrow Z_i \rightarrow C_i \rightarrow B_{i-1} \rightarrow 0.$$

Therefore,

$$\dim_{\mathbb{K}} B_i = \dim_{\mathbb{K}} Z_i - \dim_{\mathbb{K}} \tilde{H}_i$$

and

$$\dim_{\mathbb{K}} Z_i = \dim_{\mathbb{K}} C_i - \dim_{\mathbb{K}} B_{i-1}.$$

Then by induction we have

$$\begin{aligned} & \dim_{\mathbb{K}} Z_{p+1} \\ &= \dim_{\mathbb{K}} C_{p+1} - \dim_{\mathbb{K}} B_p \\ &= f_{p+1} - (\dim_{\mathbb{K}} Z_p - \dim_{\mathbb{K}} \tilde{H}_p) \\ &= f_{p+1} - \left[ \left( \sum_{j=-1}^p (-1)^{j+p} f_j + \sum_{j=0}^{p-1} (-1)^{j+p+1} \dim_{\mathbb{K}} \tilde{H}_j \right) - \dim_{\mathbb{K}} \tilde{H}_p \right] \\ &= \sum_{j=-1}^{p+1} (-1)^{j+p+1} f_j + \sum_{j=0}^p (-1)^{j+p} \dim_{\mathbb{K}} \tilde{H}_j. \end{aligned}$$

□

**Corollary 3.2.** *Let  $\Delta$  be a contractible simplicial complex and let  $\mathbb{K}$  be a field. Then,*

$$\dim_{\mathbb{K}} Z_i(\Delta, \mathbb{K}) = \sum_{j=-1}^i (-1)^{i+j} f_j.$$

**Proof.** If  $\Delta$  is contractible then  $\tilde{H}_i(\Delta, \mathbb{K}) = 0 \forall i$ . □

The following corollary provides an explicit formula for computing the dimension of cycle codes of graphs discussed in Section 2.3.

**Corollary 3.3.** *Let  $\Delta$  be a connected simplicial complex. Then,*

$$\dim_{\mathbb{K}} Z_1(\Delta, \mathbb{K}) = f_1(\Delta) - f_0(\Delta) + 1.$$

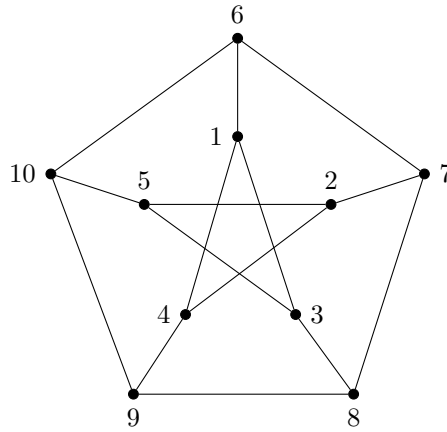
**Proof.** By Theorem 3.1, we have

$$\dim_{\mathbb{K}} Z_1 = f_1 - f_0 + f_{-1} + \dim_{\mathbb{K}} \tilde{H}_0.$$

Since  $\Delta$  is connected,  $\tilde{H}_0 = 0$ . Moreover,  $f_{-1} = 1$  for any simplicial complex. The result follows. □

In general, as the following example shows, the minimum distance of  $Z_i(\Delta, \mathbb{K})$  can be strictly bigger than  $i + 2$ . When  $i = 1$ ,  $d$  represents the girth of the 1-skeleton of  $\Delta$ .

**Example 3.4.** Consider the Petersen Graph  $P$  viewed as a simplicial complex of dimension 1. It has 10 vertices (0-faces) and 15 edges (1-faces). The graph  $P$  is 3-regular with girth 5. It follows from Theorem 3.1 and the discussion in Section 2.4 that  $Z_1(P, \mathbb{K})$  is a  $[15, 6, 5]$ -code for any field  $\mathbb{K}$ . Note that the minimum distance here is  $5 > 3 = 1 + 2$ .



If  $\mathbb{K} = \mathbb{F}_2$ , then the parity check matrix of  $Z_1(P, \mathbb{F}_2)$  is the incidence matrix of  $P$ , which is also the matrix associated to  $\partial_1(P, \mathbb{F}_2)$ . We index the rows by the vertices and the columns by the edges of  $P$ .

$$\begin{matrix}
 & 1,3 & 1,4 & 1,6 & 2,4 & 2,5 & 2,7 & 3,5 & 3,8 & 4,9 & 5,10 & 6,7 & 6,10 & 7,8 & 8,9 & 9,10 \\
 \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{matrix} & \left( \begin{array}{cccccccccccccc}
 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1
 \end{array} \right)
 \end{matrix}$$

Since the dimensions of the homologies of a simplicial complex can depend on the field  $\mathbb{K}$ , the parameters for a given cycle code can also depend on  $\mathbb{K}$ .

**Example 3.5.** Let  $\Delta$  be the standard triangulation of the real projective plane on vertex set  $\{1, 2, 3, 4, 5, 6\}$  with the following 2-faces:

$$\Delta^2 = \{123, 124, 126, 134, 135, 156, 235, 236, 346, 456\}.$$

If  $\mathbb{K} = \mathbb{F}_2$ , then  $\tilde{H}_1(\Delta, \mathbb{K}) = \mathbb{K}$ , whereas if  $\mathbb{K} = \mathbb{F}_p$  with  $p \neq 2$ , then  $\tilde{H}_1(\Delta, \mathbb{K}) = 0$ . So by Theorem 3.1, when  $\mathbb{K} = \mathbb{F}_2$ ,  $Z_2(\Delta, \mathbb{K})$  has dimension

$$\begin{aligned}
 k &= f_2 - f_1 + f_0 - f_{-1} + \tilde{H}_1(\Delta, \mathbb{F}_2) - \tilde{H}_0(\Delta, \mathbb{F}_2) \\
 &= 10 - 15 + 6 - 1 + 1 - 0 = 1,
 \end{aligned}$$

whereas when  $\mathbb{K} = \mathbb{F}_p$  for  $p \neq 2$ ,  $Z_2(\Delta, \mathbb{K})$  has dimension  $k = 0$ .

**Theorem 3.6.** *Let  $\Delta$  be an  $m$ -dimensional simplicial complex,  $\mathbb{K}$  a field and  $0 \leq i < m$ . Then  $B_i(\Delta, \mathbb{K})$  is an  $[n, k, d]$ -code with*

$$\begin{aligned} n &= f_i, \\ k &= \sum_{j=-1}^i (-1)^{i+j} f_j + \sum_{j=0}^i (-1)^{i+j+1} \dim_{\mathbb{K}} \tilde{H}_j, \\ d &= i + 2. \end{aligned}$$

**Proof.**  $B_i$  is a sub-vector space of  $C_i$  and  $\dim_{\mathbb{K}} C_i = f_i$ . So  $n = f_i$ . Since  $\dim_{\mathbb{K}} B_i = \dim_{\mathbb{K}} Z_i - \dim_{\mathbb{K}} \tilde{H}_i$ , the formula for  $k$  follows from Theorem 3.1. Since  $\dim(\Delta) \geq i + 1$ , the image of  $\partial_{i+1}$  is nonempty. Each  $(i + 1)$ -face has boundary with  $i + 2$  distinct faces. Hence there is at least one vector in  $B_i$  of weight exactly  $i + 2$ . So  $d \leq i + 2$ . But  $B_i \subseteq Z_i$  and the minimum distance of  $Z_i$  was at least  $i + 2$ . Thus  $d = i + 2$ .  $\square$

**Corollary 3.7.** *Let  $\Delta$  be an  $m$ -dimensional simplex,  $\mathbb{K}$  a field and  $0 \leq i < m$ . Then  $Z_i(\Delta, \mathbb{K}) = B_i(\Delta, \mathbb{K})$  is an  $[n, k, d]$ -code with*

$$\begin{aligned} n &= \binom{m+1}{i+1}, \\ k &= \binom{m}{i+1}, \\ d &= i + 2. \end{aligned}$$

**Proof.** By Theorem 3.1,  $n = f_i = \binom{m+1}{i+1}$ . Since  $\Delta$  is contractible,  $\tilde{H}_i(\Delta) = 0 \forall i$ . Hence,  $Z_i = B_i \forall i$  and  $d = i + 2$  by Theorem 3.6. Finally, we have

$$k = \sum_{j=-1}^i (-1)^{i+j} f_j = \sum_{j=-1}^i (-1)^{i+j} \binom{m+1}{j+1} = \binom{m}{i+1},$$

by Pascal’s Identity.  $\square$

Note that the parameters do not depend on the field  $\mathbb{K}$ .

**Remark 3.8.** *The family of  $m$ -simplex codes described here are not asymptotically good according to the definition given in Section 2.2. If  $i = m - 1$ , then  $k = \binom{m}{m} = 1$ ; that is, a one-dimensional code. In the most interesting scenario, we set  $i = m - 2$  and obtain the following:*

$$\begin{aligned} n &= \frac{(m^2 + m)}{2}, \\ k &= m, \\ d &= m. \end{aligned}$$

Note that there is quadratic growth in  $n$ , but linear growth in  $k$  and  $d$ . Thus,

$$\lim_{i \rightarrow \infty} \frac{k_i}{n_i} = 0 \quad \text{and} \quad \lim_{i \rightarrow \infty} \frac{d_i}{n_i} = 0.$$

Other choices of  $i < m - 2$  produce codes with similar asymptotics.

A similar analysis can be carried out on the dual codes  $Z^i(\Delta, \mathbb{K})$  and  $B^i(\Delta, \mathbb{K})$ . In general, the minimum distance depends on the minimum degree of an  $i$ -face in  $\Delta$ .

**Theorem 3.9.** *Let  $\Delta$  be an  $m$ -dimensional simplicial complex,  $\mathbb{K}$  a field and  $0 \leq i < m$ . Then  $Z^i = Z^i(\Delta, \mathbb{K})$  is an  $[n, k, d]$ -code with*

$$\begin{aligned} n &= f_i, \\ k &= \sum_{j=-1}^{i-1} (-1)^{i+j+1} f_j + \sum_{j=0}^i (-1)^{i+j} \dim_{\mathbb{K}} \tilde{H}_j, \\ d &\geq \deg(i, \Delta) + 1. \end{aligned}$$

**Proof.** Since  $Z^i \subseteq C^i$ ,  $n = f_i$ . Since  $Z^i = B_i^\perp$ ,  $\dim_{\mathbb{K}} Z^i + \dim_{\mathbb{K}} B_i = f_i$  and the value of  $k$  follows from Theorem 3.6. Each column of the matrix associated to  $\partial_{i+1}^*$  has at least  $\deg(i, \Delta)$  nonzero entries. Since any two distinct  $i$ -faces are contained in at most one  $(i + 1)$ -face, every set of  $\deg(i, \Delta)$  columns of  $\partial_{i+1}^*$  is linearly independent. Indeed suppose  $r$  columns of  $\partial_{i+1}^*$  are linearly dependent. The first such column has at least  $\deg(i, \Delta)$  nonzero entries and each other column can cancel at most one of these entries, meaning  $r > \deg(i, \Delta)$ . Applying Proposition 2.1 yields that  $d \geq \deg(i, \Delta) + 1$ .  $\square$

A similar argument gives the parameters for the coboundary codes.

**Theorem 3.10.** *Let  $\Delta$  be an  $m$ -dimensional simplicial complex,  $\mathbb{K}$  a field and  $0 \leq i \leq m$ . Then  $B^i = B^i(\Delta, \mathbb{K})$  is an  $[n, k, d]$ -code with*

$$\begin{aligned} n &= f_i, \\ k &= \sum_{j=-1}^{i-1} (-1)^{i+j+1} f_j + \sum_{j=0}^{i-1} (-1)^{i+j} \dim_{\mathbb{K}} \tilde{H}_j, \\ \deg(i, \Delta) + 1 &\leq d \leq \min\{\deg(\sigma) \mid \sigma \in \Delta^{i-1} \text{ and } \deg(\sigma) > 0\}. \end{aligned}$$

**Proof.** The formulas for  $n$  and  $k$  follow from Theorem 3.1. That  $\deg(i, \Delta) + 1 \leq d$  follows from Theorem 3.9 since  $B^i \subseteq Z^i$ . On the other hand, by definition there is a nonzero element in  $B^i$  of weight  $\min\{\deg(\sigma) \mid \sigma \in \Delta^{i-1} \text{ and } \deg(\sigma) > 0\}$ .  $\square$

When  $\Delta$  is the  $m$ -dimensional simplex, each  $i$ -face is contained in exactly  $m - i$  many  $(i + 1)$ -dimensional faces. Applying the previous two theorems yields the following calculation, which together with Theorem 3.7, recovers calculations in [10] and [13].

**Corollary 3.11.** *Let  $\Delta$  be an  $m$ -simplex,  $\mathbb{K}$  a field and  $0 \leq i \leq m$ . Then  $Z^i(\Delta, \mathbb{K}) = B^i(\Delta, \mathbb{K})$  is an  $[n, k, d]$ -code with*

$$\begin{aligned} n &= \binom{m+1}{i+1}, \\ k &= \binom{m}{i}, \\ d &= m - i + 1. \end{aligned}$$

It is easy to check that none of these codes form a regular LDPC family of codes that are asymptotically good.

## 4. Asymptotically good homological codes

The goal of this section is to prove the existence of asymptotically good cycle codes for specially chosen simplicial complexes  $\Delta$  with  $\dim(\Delta) \geq 2$ . We rely on the following result of Dotterer, Guth, and Kahle derived from work of Aronshtam, Linial, Łuczak and Meshulam [2]. Since a proof is omitted from [4] we sketch a proof here. For notation and details, we refer the reader to [2].



**Theorem 4.1** ([4, Theorem 6.2(1)]). *For any  $\alpha > 0$ , and for all  $n$  sufficiently large, there exist 2-dimensional simplicial complexes  $\Delta$  with  $n$  vertices and with at least  $\alpha n^2$  faces, such that every cycle in  $H_2(\Delta, \mathbb{F}_2)$  is supported on at least  $C_\alpha n^2$  faces (where  $C_\alpha > 0$  is a constant depending only on  $\alpha$ ).*

**Proof.** Let  $Y_2(n, p)$  denote the probability space of simplicial complexes of that contain every  $(d - 1)$ -dimensional face and also each  $d$ -dimensional face with independent probability  $p$ . Fix  $\alpha > 0$  and let  $c > 6\alpha$ . By [2, Theorem 4.1], there exists a constant  $\delta$  (depending only on  $c$ ) such that asymptotically almost surely, every every minimal core subcomplex (and thus every minimal cycle)  $K$  of  $Y \in Y_2(n, \frac{c}{m})$  with  $f_2(K) \leq \delta n^2$  is the boundary of a 3-dimensional simplex. Let  $X_m$  denote the number of 2-cycles in  $Y$  with exactly  $m$  faces so that  $X_m = 0$  for  $m \leq 3$  and  $X_4$  counts the number of boundaries of 3-dimensional simplices in  $Y$ . The expected number of such subcomplexes is

$$\mathbb{E}X_4 = \binom{n}{4} p^4 = \frac{n(n-1)(n-2)(n-3)}{3!} \frac{c^4}{n^4} \leq \frac{c^4}{6}.$$

Thus for sufficiently large  $n$ , we can find a  $d$ -dimensional simplicial complex  $Y$  with  $n$  vertices and no cycles of size  $\leq \delta n^2$  except for the at most  $\frac{c^4}{6}$  boundaries of 3-simplices. By removing one face from each cycle, we get a simplicial complex  $Y'$  with no 2-cycles of size  $\leq \delta n^2$ . The expected number of 2-faces in  $Y'$  is

$$\mathbb{E}f_2(Y') \geq \binom{n}{3} p - \frac{c^4}{6} = \frac{n(n-1)(n-2)}{6} \frac{c}{n} - \frac{c^4}{6} = \frac{c(n-1)(n-2) - c^4}{6},$$

and for  $n$  sufficiently large  $\frac{c(n-1)(n-2) - c^4}{6} > \alpha n^2$ . Thus we may take  $C_\alpha = \delta$ . □

With this result and Theorem 3.1 in hand, we can now prove our main result about the existence of asymptotically good cycle codes.

**Theorem 4.2.** *There is a positive constant  $c \in \mathbb{R}$  such that, for all  $m \in \mathbb{N}$  sufficiently large, there exists a 2-dimensional simplicial complex  $\Delta_m$  on  $m$  vertices such that  $Z_2(\Delta_m, \mathbb{F}_2)$  is an  $[n, k, d]$ -code with*

$$\begin{aligned} n &= m^2, \\ k &\geq \frac{m^2}{2}, \\ d &\geq cm^2. \end{aligned}$$

*In particular,  $Z_2(\Delta_m, \mathbb{F}_2)$  give a family of asymptotically good codes over  $\mathbb{F}_2$ .*

**Proof.** Let  $\alpha = 1$  and apply Theorem 4.1. We get an infinite family of 2-dimensional simplicial complexes  $\Delta_m$  for  $m \gg 0$  such that  $\Delta_m$  has  $m$  vertices and at least  $m^2$  2-faces and such that every cycle of  $H_2(\Delta_m, \mathbb{F}_2) = Z_2(\Delta_m, \mathbb{F}_2)$  is supported on at least  $cm^2$  faces where  $c = C_1$  above. First note that we may assume that  $f_2(\Delta_m) = m^2$  since deleting faces can only increase the size of the support of a cycle in  $H_2(\Delta_m, \mathbb{F}_2)$ . Moreover, we may assume that all  $\binom{m}{2}$  edges are present since adding edges does not affect  $Z_2(\Delta_m, \mathbb{F}_2)$ . Thus by Theorem 3.1,  $Z_2(\Delta_m, \mathbb{F}_2)$  is an  $[n, k, d]$ -code with  $n = f_2(\Delta_m) = m^2$ . Since every cycle is supported on at least  $cm^2$  faces,  $d \geq cm^2$ . Finally we have

$$\begin{aligned} k &= f_2 - f_1 + f_0 - f_{-1} + \tilde{H}_1 - \tilde{H}_0 \\ &= m^2 - \binom{m}{2} + m - 1 + \tilde{H}_1 - \tilde{H}_0 \\ &\geq \frac{m^2}{2} \end{aligned}$$

where the last inequality uses that  $\tilde{H}_0 = 0$ , since  $\Delta_m$  is connected, and that  $m \geq 2$ . □

The results in [2] and [4] rely on probabilistic methods to prove the existence of such simplicial complexes. It would be interesting to have explicit constructions of such simplicial complexes.

Finally we remark that the previous Theorem can be extended to higher dimension by taking suspensions.

**Proposition 4.3.** *Let  $\Delta$  be an  $m$ -dimensional simplicial complex and let  $\mathbb{K}$  be a field. If  $Z_m(\Delta, \mathbb{K})$  had parameters  $[n, k, d]$ , then  $Z_{m+1}(S(\Delta), \mathbb{K})$  has parameters  $[2n, k, 2d]$ .*

**Proof.** For all  $i \geq -1$ , there is a canonical isomorphism  $\phi : \tilde{H}_i(\Delta, \mathbb{K}) \cong \tilde{H}_{i+1}(S(\Delta), \mathbb{K})$ , coming from the Mayer-Vietoris sequence induced by sending an  $i$ -face  $\sigma$  to the sum  $\sigma \cup \{a\} + \sigma \cup \{b\}$ , where  $a, b$  are the two additional vertices. In particular,  $Z_m(\Delta, \mathbb{K}) = \tilde{H}_m(\Delta, \mathbb{K}) \cong \tilde{H}_{m+1}(S(\Delta), \mathbb{K}) = Z_{m+1}(S(\Delta), \mathbb{K})$  and the dimensions of the 2 codes agree. If  $n$  is the length of  $Z_m(\Delta, \mathbb{K})$ , then by construction the length of  $S(\Delta)$  is  $f_{m+1}(S(\Delta)) = 2f_m(\Delta) = 2n$ . Finally, it is clear that  $\phi$  doubles the support of any cycle in  $Z_m(\Delta, \mathbb{K})$  and hence doubles the minimum weight of the associated code.  $\square$

Note that since the dimension is preserved, we do not get asymptotically good families by taking repeated suspensions of a given simplicial complex. However, combining with Theorem 4.2 yields the following consequence.

**Corollary 4.4.** *For any integer  $r \geq 2$ , there is a family of  $r$ -dimensional simplicial complexes  $\{\Delta_m\}$  whose cycle codes  $Z_r(\Delta_m, \mathbb{F}_2)$  are asymptotically good.*

**Proof.** Fix  $r \geq 2$ . Let  $S^r(\Delta)$  denote the  $r$ -fold suspension of  $\Delta$ . Let  $\{\Delta_m\}$  be the family of 2-dimensional simplicial complexes from Theorem 4.2. Then by Proposition 4.3,  $Z_r(S^{r-2}(\Delta_m), \mathbb{F}_2)$  has parameters  $[2^{r-2}m^2, m^2/2, c2^{r-2}m^2]$ , and hence forms an asymptotically good family.  $\square$

Finally we remark that since the parity check matrix of  $Z_r(S^{r-2}(\Delta_m), \mathbb{F}_2)$  is the matrix associated to  $\partial_r(S^{r-2}(\Delta_m), \mathbb{F}_2)$ , it has exactly  $r + 1$  ones per column. So the corresponding family of codes is a (non-regular) LDPC family of codes.

**Acknowledgment:** We are very grateful to Tony Bahri and Lance Miller for feedback on an earlier version of this paper and Matthew Kahle for useful conversations. We also thank the referee for several helpful suggestions.

## References

- [1] N. Alon, S. Hoory, N. Linial, The Moore bound for irregular graphs, *Graphs Combin.* 18(1) (2002) 53–57.
- [2] L. Aronshtam, N. Linial, T. Łuczak, R. Meshulam, Collapsibility and vanishing of top homology in random simplicial complexes, *Discrete Comput. Geom.* 49(2) (2013) 317–334.
- [3] A. R. Calderbank, P. W. Shor, Good quantum error-correcting codes exist, *Physical Review A* 54(2) (1996) 1098–1105.
- [4] D. Dotterer, L. Guth, M. Kahle 2-complexes with large 2-girth, *Discrete Computational Geometry* 59(2) (2018) 383–412.
- [5] R. G. Gallager, Low-density parity-check codes, *IRE Trans.* 8(1) (1962) 21–28.
- [6] R. G. Gallager, *Low-Density Parity-Check Code*, MIT Press, 1963.
- [7] X.-Y. Hu, E. Eleftheriou, D. M. Arnold, Regular and irregular progressive edge-growth Tanner graphs, *IEEE Trans. Inform. Theory* 51(1) (2005) 386–398.
- [8] Steven Roman, *Coding and Information Theory*, Graduate Texts in Mathematics, vol. 134, Springer-Verlag, New York, 1992.

- [9] Pavel Rytíř, Geometric representations of linear codes, *Adv. Math.* 282(10) (2015) 1–22.
- [10] Charles Saltzer, Topological codes, *Error Correcting Codes (Proc. Sympos. Math. Res. Center, Madison, Wis., 1968)*, Wiley, New York (1968) 111–129.
- [11] P. W. Shor, Scheme for reducing decoherence in quantum memory, *Phys. Rev. A.* 52(4) (1995) R2493–R2496.
- [12] A. M. Steane, Multiple-particle interference and quantum error correction, *Proc. Roy. Soc. A.* 452(1954) (1996) 2551–2577.
- [13] S. Thomeier, Error-correcting polyhedral codes, *J. Comput. Inform.* 1(2) (1990) 91–101.
- [14] G. Zémor, On Cayley graphs, surface codes, and the limits of homological coding for quantum error correction, *Coding and cryptology, Lecture Notes in Comput. Sci.*, vol. 5557, Springer, Berlin (2009) 259–273.