## HYPERGRAPHS AND THEIR HOMOLOGY

By

Christopher Potvin

## A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Mathematics—Doctor of Philosophy

2023

#### ABSTRACT

It is commonplace today for scientists to study networks where interactions among the participants happen at higher orders, where any number of participants can be related at once. Graphs, the typical mathematical model for networks, can only display pairwise relationships, and are sometimes inadequate in these situations. The combinatorial structures that have been used to model higher order networks and generalize graphs, and which will play a major role in this thesis, are hypergraphs and simplicial complexes. Hypergraphs have grown in popularity over the past decades, and are becoming standard combinatorial objects used in network representations. They differ from graphs only in that an edge in a hypergraph can have any number of vertices, not just two.

Simplicial complexes, another generalization of graphs, have become ubiquitous in algebraic topology. In part, this is because the homology of simplicial complexes is standard. Homology is a tool for studying properties of the shape of topological spaces, and has proven to be useful in topology for classifying spaces as it is a topological invariant. In data science, researchers study the homology of simplicial complexes that change with the data. Keeping track of the simplicial complexes over time allows them to ascertain when changes in the homology are caused by meaningful changes in the data.

Hypergraphs are also generalizations of simplicial complexes, however, the homology of hypergraphs is currently a problematic gap in the theory. No universal theory of homology for hypergraphs has been established, and the definition of homology for simplicial complexes does not extend obviously. Nonetheless, there has been research done into the homology of hyper-graphs. The two homology theories for hypergraphs studied in this thesis are called the restricted barycentric homology and the relative barycentric homology. We present novel combinatorial definitions, topological and classification results, and methods for computation. This thesis aids in the development of theory, interpretability, and data science of topological hypergraph analytics.

Copyright by CHRISTOPHER POTVIN 2023

#### ACKNOWLEDGEMENTS

I would like to begin by thanking my wife Cassie, my parents, sister, in-laws, and extended family for their patience and support while I was a graduate student. They always gave me the time and space I needed when I was studying, reading, or writing. I am also grateful for my friends, both among fellow graduate students at MSU and outside of school, who have helped to keep me sane-ish.

My advisor, Dr. Liz Munch, was incredibly helpful for the past three years. She was willing to take on a new mentee as the pandemic started and allowed me to choose a research area not directly related to her own. She always reminded me of the big picture, and had time for me when I needed it. I appreciate everyone whom I have met through MunchLab's weekly meetings, which have always served to motivate me.

I was fortunate to do an internship at Pacific Northwest National Laboratory (PNNL), which was a great opportunity and experience for me. I would like to thank everyone I met there for including me as one of their own; the authors of [22] form a nonexhaustive list. Dr. Emilie Purvine, in particular, served on my graduate committee, encouraged me to apply for the internship, and mentored me while I was interning. Throughout this thesis, the symbol † will be used to denote a result that I worked on during my time at PNNL.

# TABLE OF CONTENTS

LIST OF FIGURES
CHAPTER 1: INTRODUCTION
CHAPTER 2:  BACKGROUND  5    2.1:  Hypergraphs  5    2.2:  Simplicial Complexes and their Homology  14    2.3:  Hypergraph Homology Theories  26
CHAPTER 3:  RESULTS ON THE RESTRICTED BARYCENTRIC HOMOLOGY
CHAPTER 4:  RESULTS ON THE RELATIVE BARYCENTRIC HOMOLOGY  43    4.1:  Two Versions of a Mayer-Vietoris Sequence  44    4.2:  Results on Maximum Edge Hypergraphs  52    4.3:  Results on General Hypergraphs  63
CHAPTER 5:  BRIDGING THE THEORIES AND COMPUTATION
CHAPTER 6:  DYNAMIC HYPERGRAPHS  87    6.1:  Adding an Edge  88    6.2:  Adding a Vertex  96
CHAPTER 7: CONCLUSIONS AND FUTURE WORK
BIBLIOGRAPHY

# LIST OF FIGURES

Figure 2.1	A hypergraph $\mathcal{H}$ and a reduced hypergraph $\mathcal{H}'$ .	7
Figure 2.2	A hypergraph $\mathcal{H}$ and its line graph $L(\mathcal{H})$	8
Figure 2.3	A hypergraph $\mathcal{H}$ and its complement $\text{Comp}(\mathcal{H})$	9
Figure 2.4	A hypergraph $\mathcal{H}$ , its complement $\text{Comp}(\mathcal{H})$ , and its supplement $\text{Supp}(\mathcal{H})$	10
Figure 2.5	Of these two hypergraphs, $\mathcal{H}$ and $K$ , only $K$ is a simplicial complex	15
Figure 2.6	A simplicial complex <i>K</i> , its barycentric subdivision <i>T</i> , and a generated sub- complex	17
Figure 2.7	Comparing a hypergraph and simplicial model of a biological system	27
Figure 2.8	A hypergraph $\mathcal{H}$ and its barycentric subdivision $T$ with the restricted barycentric subdivision highlighted.	31
Figure 2.9	A hypergraph $\mathcal{H}$ and its barycentric subdivision $T$ with the missing subcomplex highlighted.	34
Figure 3.1	Hypergraphs with different fence components than connected components	41
Figure 3.2	Hypergraphs used as examples of Theorem 3.2.6	42
Figure 4.1	A hypergraph $\mathcal{H}$ with computation of $H^{rel}(\mathcal{H})$	55
Figure 4.2	A hypergraph $\mathcal{H}$ used as an example for Theorem 4.2.7	58
Figure 4.3	Hypergraphs used as examples for Theorem 4.2.9	60
Figure 4.4	A hypergraph $\mathcal{H}$ used as an example for Theorem 4.2.10	63
Figure 4.5	A hypergraph $\mathcal{H}$ with multiple connected components	64
Figure 4.6	A non maximum edge hypergraph $\mathcal{H}$ with contractible $K$	65
Figure 4.7	A non maximum edge hypergraph $\mathcal{H}$ with contractible S	68
Figure 4.8	A hypergraph for which the map $H_1(S) \rightarrow H_1(K)$ is not zero	69
Figure 4.9	A hypergraph for which $H_n^{rel}(\mathcal{H}) = H_n(K) \bigoplus H_{n-1}(S)$	70
Figure 4.10	Three different hypergraphs with the same toplexes	75
Figure 5.1	A hypergraph $\mathcal{H}$ used as an example for Theorem 5.1.2	79

Figure 5.2	A hypergraph $\mathcal{H}$ and its missing subcomplex and restricted barycentric sub- division
Figure 6.1	Three hypergraphs used as examples for Proposition 6.1.1
Figure 6.2	A hypergraph where adding the edge $CD$ merged two simplicial components 91
Figure 6.3	A hypergraph where adding the edge A reduced $\beta_0^{res}(\mathcal{H})$
Figure 6.4	A hypergraph used as an example for Theorem 6.2.1
Figure 6.5	A hypergraph used as an example for Theorem 6.2.2

#### **CHAPTER 1**

#### INTRODUCTION

Studying mathematics and analyzing data with network representations is a concept as old as particular trip of Euler through Königsberg, and a generic traveling salesperson looking for the shortest path to travel to all of the cities on their itinerary. Traditionally, those networks have been modeled using graphs. However, over the course of the 20th century, graphs have occasionally been found lacking because, by default, they only encode pairwise interactions. It is commonplace today for scientists to study networks where interactions among the participants happen at higher orders. For example, researchers in biology have found evidence of a three way symbiotic relationship where no two of the organisms would be able to survive without the third [23]. They discovered this network while studying a plant that was surviving in warmer soil than it would typically be able to live. Originally, they found a fungus on the plant and thought the fungus and the plant were in a symbiotic relationship, but later, they found a virus using the fungus as a host. Researchers isolated and removed the virus, and the fungus and the plant died. The fungus and the virus could not live without the plant, and the virus and the plant died without the fungus. If data scientists wanted to use a combinatorial object to model this data, a graph would be insufficient because there is a three way relationship without any of the two way subsets that would need to be displayed in a graph. The combinatorial structures that have been used to generalize graphs, and which will play a major role in this thesis, are hypergraphs and simplicial complexes. Surveys of historical and current methods of studying higher order networks can be found in the following: [3, 4, 6, 31].

Hypergraphs have grown in popularity over the past decades, and are becoming standard combinatorial objects used in network representations [5, 8]. Similar to graphs, hypergraphs can be studied via notions of their path components [17, 1, 21], or via ideas of acyclicity [15, 10]. Some fields where researchers have studied hypergraphs include biology for gene identification [16], urban traffic networks [17], and the domain name system [20].

Simplicial complexes, another generalization of graphs to higher order networks, have become ubiquitous in algebraic topology. In part, this is because the homology of simplicial complexes is

standard [18, 24]. Homology is a tool for studying properties of the shape of topological spaces. Given a topological space, the homology of that space is represented by a group in every integer dimension greater than or equal to zero. The ranks of these groups are called the Betti numbers. In dimension zero, the Betti number is the number of connected components of the space. In dimension one, it is the number of loops, or one dimensional holes. This increases with the dimension, so in dimension two, the Betti number represents the number of hollow voids in the space, and so on. Homology has proven to be useful in topology of simplicial complexes that change with the data. Keeping track of the simplicial complexes over time allows them to ascertain when changes in the data cause changes in the homology. This process, called persistent homology, is one of the major successes of the field of topological methods to model and analyze data, an umbrella under which this thesis also falls.

Hypergraphs can also be viewed as generalizations of simplicial complexes, however, the lack of a notion of homology of hypergraphs is currently a problematic gap in the theory. No universal theory of homology for hypergraphs has been established, and the definition of homology for simplicial complexes does not extend obviously. Nonetheless, there has been research done into defining a homology of hypergraphs. First, researchers embedded hypergraphs into a particular simplicial complex and found the homology of that complex [26, 25]. However, many different hypergraphs will have the same simplicial complex. One group of researchers studying hypergraph homology defined the embedded homology [7], which fits a hypergraph to the closest chain complex (the algebraic structure needed for homology). They have also expanded that theory with typical homology results like relative homology, a Mayer-Vietoris sequence, and a Künneth formula [7, 33, 29, 30]. One drawback of the embedded homology is that it abstractifies the original hypergraph from this homology theory. Other theories on the homology of hypergraphs have crept up as well, includin algebraic theories on rings and ideals associated to hypergraphs [14] and 1-dimensional

homology on oriented hypergraphs [12].

There are two additional homology theories for hypergraphs that rely on the barycentric subdivision of the associated simplicial complex [28], which are the focus of this thesis. The first of these constructions, the restricted barycentric homology, only considers the subset of simplices of the barycentric subdivision which are included as edges in the hypergraph. The second is the relative barycentric homology for hypergraphs, where given the same barycentric subdivision, the homology is computed relative to the subcomplex of the simplices that are not edges in the hypergraph. This latter definition of homology does a good job of accounting for the missing subedges, thus discriminating between different hypergraphs with the same associated simplicial complex.

This thesis contributes to the study of hypergraph homology by developing the theory of the restricted and relative barycentric homology. We accomplish this simultaneously in three directions: combinatorics, topology, and data science. We start by manipulating standard results from algebraic topology like the Mayer-Vietoris sequence and the long exact sequence of a pair to make it easier to calculate the homology of complicated hypergraphs. These tools allow us to completely classify the relative barycentric homology in several dimensions of a specific type of hypergraph called a maximum edge hypergraph. In order to interpret these results, we also needed novel combinatorial definitions not found in the hypergraph literature. These new concepts include the ideas of the supplement of a hypergraph, and its fence components, an alternative to the traditional idea of connected components. Lastly, we made advancements towards using these methods in data science by finding techniques to simplify the computation and developing results that account for changes in the hypergraph. The computations discussed in this thesis were implemented in HyperNetX, an open source python package for hypergraph analytics [27]. We have studied how the changes in a hypergraph over a filtration will affect the relative and restricted barycentric homology.

We will conclude the introduction with an outline of the contents of the paper. The next chapter contains background information: Section 2.1 on hypergraphs and Section 2.2 on simplicial complexes and their homology. Section 2.3 gives the definitions of the restricted and relative

barycentric homology for hypergraphs. Chapter 3 contains development of the restricted barycentric homology. The chapter on the relative barycentric homology, Chapter 4, is split into three sections: Section 4.1 on Mayer-Vietoris theorems for hypergraphs, Section 4.2 on results for maximum edge hypergraphs, and Section 4.3, which contains results on the relative barycentric homology that apply to general hypergraphs, and not just maximum edge hypergraphs. Chapter 5 discusses some results that bridge the relative and restricted versions of the theory, and also contains Section 5.2, which is about computation. Dynamic hypergraphs are studied in Chapter 6. This thesis finishes with some concluding remarks and discussions of potential directions for future research in Chapter

7.

#### **CHAPTER 2**

#### BACKGROUND

In this chapter, we will introduce the main objects of study, and the primary techniques used to study them. First up will be an introduction to hypergraphs, including several novel definitions. Then, we will introduce simplicial complexes, objects that are both combinatorial and topological in nature. One major advantage of simplicial complexes is that they have a well studied homology theory [24, 18], which is not true for hypergraphs. We will define simplicial homology and give a few traditional results that we will analogize to hypergraphs later in this thesis. The last thing in this chapter will be the definitions of the restricted and relative barycentric homology of hypergraphs. These definitions were originally given by Emilie Purvine and collaborators [28]. The bulk of this thesis will be developing the homology theory further for these two definitions.

## 2.1 Hypergraphs

A hypergraph is a combinatorial object that generalizes vertex-edge graphs. Informally, hypergraphs are graphs where an edge can have any number of vertices, not just two. Hypergraphs are becoming increasingly utilized because of their ability to represent multirelational network data. Two standard references for hypergraphs are [5, 8].

This section will be split into two parts. The first contains standard hypergraph definitions. Versions of these definitions can be found in [8] unless otherwise noted. The second part consists of new definitions and set-theoretic results on hypergraphs.

## 2.1.1 Standard Hypergraph Definitions

We will begin with the definition of a hypergraph. Similarly to a graph, a hypergraph consists of vertices and edges, with the difference being that an edge can have any number of vertices.

**Definition 2.1.1** (Hypergraph). A hypergraph  $\mathcal{H}$  consists of two sets: a set of vertices

$$V = \{v_1, v_2, \dots, v_n\}$$

and a set of edges, each of which is a nonempty subset of the vertex set,

$$E = \{e_1, e_2, \ldots, e_m \mid e_i \subset V\}.$$

We will assume in perpetuity that neither V nor E is empty and that every vertex is in some edge, i.e. that there are no *isolated vertices* [8]. Because of this property, a hypergraph is well-defined given only the edge set. Henceforth, we will use the notation  $e \in \mathcal{H}$  in place of  $e \in E$  to mean that e is an edge in the hypergraph  $\mathcal{H}$ , and will not consider the hypergraph any different from its edge set. We will occasionally specify the edge set E or the vertex set V when it is convenient. The next definition comes from [27], where it is used to simplify hypergraphs that have repeated edges.

**Definition 2.1.2** (Edge-Collapsed Hypergraph). A hypergraph is said to be edge-collapsed if for all  $i \neq j$ ,  $e_i \neq e_j$ .

All hypergraphs are henceforth assumed to be edge-collapsed, so there will not be any hypergraphs that have more than one copy of the same edge. This is analogous to avoiding multigraphs when studying graph theory. A hypergraph that is a subset of another hypergraph will be called a *subhypergraph*. Throughout this thesis, it will often be important whether or not an edge is contained in another edge. We will use the terminology subedge [11] to describe an edge that is a subset of another edge, and toplex [28] to describe an edge that is not a subset of another edge.

**Definition 2.1.3** (Subedge). Let  $\mathcal{H}$  be a hypergraph and  $e_i, e_j \in \mathcal{H}$  such that  $e_i \subset e_j$ . Then  $e_i$  is called a subedge in  $\mathcal{H}$  and in particular it is a subedge of  $e_j$ .

**Definition 2.1.4** (Toplex). Let  $\mathcal{H}$  be a hypergraph and  $e_i \in E$  such that  $\forall e \in E, e_i \notin e$ . Then  $e_i$  is called a toplex of  $\mathcal{H}$ .

Since an edge is either a subset of some other edge or a subset of no other edges, every edge in a hypergraph is a subedge or a toplex. Hypergraphs that have no subedges will also play an important role, particularly in Theorems 3.2.2 and 4.3.9.

**Definition 2.1.5** (Reduced Hypergraph). Let  $\mathcal{H}$  be a hypergraph. It will be called a reduced hypergraph if it does not have any subedges, or equivalently, if every edge is a toplex.

A reduced hypergraph is also called a simple hypergraph in [8] and other places. A reduced hypergraph  $\mathcal{H}'$  that has the same toplexes as a hypergraph  $\mathcal{H}$  will be called its reduction. Figure



Figure 2.1 A hypergraph  $\mathcal{H}$  and a reduced hypergraph  $\mathcal{H}'$ .

2.1 gives our first two examples of a hypergraph,  $\mathcal{H}$  and  $\mathcal{H}'$ . Notice that both hypergraphs have the same vertex set,  $\{A, B, C, D\}$ , but there are different edge sets,  $\mathcal{H} = \{ABC, BCD, B, C\}$  and  $\mathcal{H}' = \{ABC, BCD\}$ . Here,  $\mathcal{H}'$  does not have any edges that are subedges, and so it is a reduced hypergraph, and since  $\mathcal{H}$  has the same toplexes,  $\mathcal{H}'$  is the reduction of  $\mathcal{H}$ .

There are multiple ways of turning a hypergraph back into a graph. One can turn a hypergraph into a bipartite graph by using *V* and *E* as the vertex sets of the bipartite graph, with graph edges from  $v \in V$  to  $e \in E$  if  $v \in e$  in the hypergraph [21]. Another method is the *two-section* [8] of a hypergraph, which is a graph on the same vertex set *V*, with graph edges between two vertices that share an edge in the hypergraph. Most useful for us will be the line graph, defined next. In [8], the line graph is also called the intersection graph or representative graph.

**Definition 2.1.6** (Line Graph). Let  $\mathcal{H}$  be a hypergraph with edge set E. The line graph of  $\mathcal{H}$ , denoted  $L(\mathcal{H})$ , is a graph built as follows:

- 1. The vertex set of the line graph is the edge set of the hypergraph,  $V_L := E$
- 2. There is an edge between  $e_i$  and  $e_j$  in the line graph if  $e_i \cap e_j \neq \emptyset$  in  $\mathcal{H}$ .

For an example of a line graph, see Figure 2.2. The hypergraph  $\mathcal{H}$  in the figure has edge set  $E = \{ABC, BCD, B, C\}$ , which are the vertices of the line graph L(H). Notice that there is not a graph edge between B and C in  $L(\mathcal{H})$  because  $B \cap C = \emptyset$ , but all other pairs of edges in  $\mathcal{H}$  intersect nontrivially, and that is represented by a graph edge in  $L(\mathcal{H})$ .



Figure 2.2 A hypergraph  $\mathcal{H}$  and its line graph  $L(\mathcal{H})$ .

The next definition is that of a complement hypergraph. Like with graphs, the complement hypergraph consists of all of the edges that were not in the original hypergraph, retaining the same vertex set.

**Definition 2.1.7** (Complement of a Hypergraph). Let  $\mathcal{H} = \{V, E\}$  be a hypergraph. The complement of  $\mathcal{H}$  is all subsets of the vertex set that are not edges in  $\mathcal{H}$ , and is denoted Comp $(\mathcal{H})$ . Hence, if  $\mathcal{P}(V)$  denotes the power set of V, Comp $(\mathcal{H}) = \mathcal{P}(V) \setminus E$ .

This set theoretic standard notion of the complement serves as the building block for several results herein, as well as the inspiration for Definition 2.1.11 in the novel definitions section below. See Figure 2.3 for an example. Another definition brought over from graph theory is the definition of a walk along a hypergraph. Below is the definition of 1-walk along a hypergraph from [21], herein just referred to as a walk.

**Definition 2.1.8** (Hypergraph Edge Walk). Let  $\mathcal{H}$  be a hypergraph with edge set E. A walk is a sequence of edges  $e_{i_1}, e_{i_2}, \ldots, e_{i_k}$ , where  $e_{i_j} \cap e_{i_{j+1}} \neq \emptyset$  and  $e_{i_j} \neq e_{i_{j+1}}$ .

Note that there is not a standard definition of walk in the hypergraph literature. In [8], a walk lists both edges and vertices that are in their intersection, and is called a path. In [21], a path is a specific type of walk. The idea of an *s*-walk, as defined in [21], leverages the more general nature of hypergraphs by stipulating that each pair of consecutive edges in the walk must intersect in at least *s* different vertices. There would also be a natural notion of walk where only the vertices are



Figure 2.3 A hypergraph  $\mathcal{H}$  and its complement Comp $(\mathcal{H})$ .

listed and adjacent vertices in the walk must share an edge. We will use the given definition of walks to define connected components in hypergraphs.

**Definition 2.1.9** (Connected Components of a Hypergraph). Let  $\mathcal{H}$  be a hypergraph with edge set *E*. Define an equivalence relation on *E*:  $e_i \sim e_j$  if there is a walk from  $e_i$  to  $e_j$ . The equivalence classes of this relation are the connected components of the hypergraph.

If walks were defined here in terms of vertices then the definition above could also be reworked to give connected components in terms of vertices, as done in [8].

## 2.1.2 New Definitions and Results

Next we will include some novel definitions and some set-theoretic results on the complement and supplement hypergraphs. The first definition is that of a maximum edge hypergraph. This definition was necessitated and motivated by the results of Section 4.2.

**Definition 2.1.10** (Maximum Edge Hypergraph). Let  $\mathcal{H} = \{V, E\}$  be a hypergraph. It will be called a maximum edge hypergraph if there is some edge *e* such that e = V. Oftentimes, *e* will be referred to as the maximum edge of  $\mathcal{H}$ .

An important property of maximum edge hypergraphs is that if e is the maximum edge,  $e_i \subset e \forall e_i \in E$ . In words, every other edge of  $\mathcal{H}$  is a subedge of the maximum edge e.

It will become apparent later (as soon as Section 2.3) that a special subset of the complement hypergraph is particularly important. This subset consists of the edges that are subsets of edges that



Figure 2.4 A hypergraph  $\mathcal{H}$ , its complement Comp $(\mathcal{H})$ , and its supplement Supp $(\mathcal{H})$ .

are in the hypergraph, and will be called the supplement hypergraph. (The † is used throughout this thesis to indicate work partially supported by PNNL.)

**Definition 2.1.11** (Supplement Hypergraph<sup>†</sup>). Let  $\mathcal{H}$  be a hypergraph with edge set E and vertex set V. Consider the set  $\mathcal{E}$  of all subsets of edges in E. Define  $G := \mathcal{E} \setminus E$ . Here, G is the set of all subedges of edges in  $\mathcal{H}$  that are not themselves edges, i.e.

$$G = \{ e' \subset e \text{ for some } e \in E \mid e' \neq e \text{ for any } e \in E \}.$$

The hypergraph with edge set G is called the supplement hypergraph and denoted  $\text{Supp}(\mathcal{H})$ .

Figure 2.4 gives an example of a hypergraph with both its complement and supplement. Here, the hypergraph  $\mathcal{H}$  has the vertices  $\{A, B, C\}$ , but *ABC* is not an edge, and so *ABC* is an edge in Comp( $\mathcal{H}$ ). However, *ABC* is not a subset of any existing edge in  $\mathcal{H}$ , and is thus not an edge in Supp( $\mathcal{H}$ ). The complement Comp( $\mathcal{H}$ ) is a maximum edge hypergraph, and the other two hypergraphs are not.

Note that the vertex set of  $\text{Supp}(\mathcal{H})$  may not be the same as the vertex set of  $\mathcal{H}$ , as in Figure 2.4. Supp $(\mathcal{H})$  inherits only the vertices in edges in G. Since the complement includes all subsets of V that are not edges in  $\mathcal{H}$ , Supp $(\mathcal{H}) \subset \text{Comp}(\mathcal{H})$ . Furthermore, for maximum edge hypergraphs, the supplement and the complement agree, as shown next.

**Proposition 2.1.1.** <sup>†</sup> Let  $\mathcal{H}$  be a maximum edge hypergraph. Then  $Comp(\mathcal{H}) = Supp(\mathcal{H})$ .

*Proof.* Let *e* be the maximum edge of  $\mathcal{H}$ , and *V* the vertex set of  $\mathcal{H}$ . Since  $\mathcal{H}$  is a maximum edge hypergraph,  $e = \{V\}$ . By definition,  $\operatorname{Supp}(\mathcal{H}) \subset \operatorname{Comp}(\mathcal{H})$ . Let  $f \in \operatorname{Comp}(\mathcal{H})$ . By the definition of complement, *f* is not an edge in  $\mathcal{H}$ , and  $f \subset V$ . However,  $e = \{V\}$ , so  $f \subset e$ . Therefore, *f* is not an edge in  $\mathcal{H}$ , but is a subset of an edge in  $\mathcal{H}$ . Thus,  $f \in \operatorname{Supp}(\mathcal{H})$ . So,  $\operatorname{Comp}(\mathcal{H}) \subset \operatorname{Supp}(\mathcal{H})$ , and by double inclusion, the two are equal, proving the proposition.

Maximum edge hypergraphs will turn out to be quite important, so we will want ways to generate them even when starting with a hypergraph that is not maximum edge. One such way is to take the complement. The next result says that for  $\mathcal{H}$  that is not maximum edge, its complement is maximum edge.

**Lemma 2.1.2.** Suppose  $\mathcal{H}$  is not a maximum edge hypergraph. Then its complement,  $Comp(\mathcal{H})$ , is a maximum edge hypergraph.

*Proof.* Since  $\mathcal{H}$  is not a maximum edge hypergraph, there is no edge containing all of its vertices V. Therefore, by definition of the complement,  $\overline{e} = V$  is an edge in the complement hypergraph. Since all other edge in the complement will be subsets of  $\overline{e} = V$ , the complement hypergraph is a maximum edge hypergraph.

The following corollary combines the previous results with the law of double complements to write any hypergraph that is not maximum edge as a supplement of a maximum edge hypergraph. Hence, every hypergraph is either a maximum edge hypergraph or the supplement of one.

**Corollary 2.1.3.** Suppose *H* is not a maximum edge hypergraph. Then *H* is the supplement of its complement,

$$\mathcal{H} = Supp(Comp(\mathcal{H})).$$

*Proof.* This corollary follows from the previous results. Recall the law of double complements  $\mathcal{H} = \text{Comp}(\text{Comp}(\mathcal{H}))$ . Since  $\text{Comp}(\mathcal{H})$  is a maximum edge hypergraph by Lemma 2.1.2, its complement is equal to its supplement by Proposition 2.1.1. Thus we get the following string of equalities that prove the corollary:  $\mathcal{H} = \text{Comp}(\text{Comp}(\mathcal{H})) = \text{Supp}(\text{Comp}(\mathcal{H}))$ .

The next few definitions of this section are moving towards an alternative to the definition of connected components in hypergraphs (Definition 2.1.9), which are called fence components. Instead of being based off of intersections, fence components will be based off of inclusions of subedges. These will come up often, particularly in Chapter 3. The preliminary definition will be that of a hypergraph fence, which is a specific type of walk (Definition 2.1.8).

**Definition 2.1.12** (Hypergraph Fence). Let  $\mathcal{H}$  be a hypergraph. A fence in  $\mathcal{H}$  is a hypergraph walk  $e_{i_1}, e_{i_2}, \ldots, e_{i_k}$  such that for all  $e_{i_j}$  either:

- 1.  $e_{i_{j-1}} \subset e_{i_j} \supset e_{i_{j+1}}$ , or
- 2.  $e_{i_{i-1}} \supset e_{i_i} \subset e_{i_{i+1}}$ .

A fence is more restrictive than a walk because for a fence, each consecutive pair of edges needs to have an inclusion relation instead of only a nontrivial intersection. The definition for connected components in Definition 2.1.9 uses the notion of walks to form an equivalence relation and partition the edge set. The next definition does an analogous thing with fences instead of walks.

**Definition 2.1.13** (Fence Components). Let  $\mathcal{H}$  be a hypergraph with edge set E. Define a relation  $\sim$  on E with  $e_i \sim e_j$  if there is a fence from  $e_i$  to  $e_j$ . The equivalence classes of this relation are called the fence components of  $\mathcal{H}$ , and the number of fence components of a hypergraph will be denoted  $\Gamma(\mathcal{H})$ .

A good example of the difference between fence components and connected components can be found in Figure 2.1. Both hypergraphs  $\mathcal{H}$  and  $\mathcal{H}'$  have one connected component, as the two toplexes intersect in both hypergraphs. In  $\mathcal{H}$ ,  $ABC \supset B \subset BCD$  is a fence between ABC and BCD, and  $C \subset ABC$ , so all edges are in the same fence component. In  $\mathcal{H}'$ , though, there is no fence between ABC and BCD, since they do not have a mutual subedge. Therefore  $\mathcal{H}'$  has two fence components whilst having one connected component.

Inclusion of edges in a hypergraph gives a partial order on  $\mathcal{H}$ . The below definition comes from [22], where it is called the edge containment partial order. This poset will be used later

on, especially in Section 5.2. The name face poset was inspired by the face poset of a simplicial complex from [32].

**Definition 2.1.14** (Face Poset of a Hypergraph). Let  $\mathcal{H}$  be a hypergraph. Define a relation on the edges of  $\mathcal{H}$  by  $e_i < e_j$  if  $e_i \subset e_j$ . This is a partial order, and it forms what is called the face poset of  $\mathcal{H}$ ,  $FP(\mathcal{H})$ .

Recalling Definition 2.1.6, the line graph is a graph on the edge set of  $\mathcal{H}$  that displays information about the intersection of edges. The next definition is another graph called the edge inclusion graph and it is analogous to the line graph, but displays information about the inclusion relations among edges in  $\mathcal{H}$ . That difference is comparable to the difference between the connected components of a hypergraph and its fence components, as discussed in the previous example.

**Definition 2.1.15** (Edge Inclusion Graph). Let  $\mathcal{H}$  be a hypergraph. The edge inclusion graph is a graph  $EIG(\mathcal{H})$  constructed as follows:

- 1. The vertex set of  $EIG(\mathcal{H})$  is the edge set of the hypergraph,  $V_{EIG} = \{v_e \mid e \in \mathcal{H}\}$ .
- 2. There is an edge in  $EIG(\mathcal{H})$  between  $v_{e_i}$  and  $v_{e_j}$  if  $e_i \subset e_j$  or  $e_i \supset e_j$  in  $\mathcal{H}$ .

The final definition on hypergraphs is reserved for a special type of connected component (Definition 2.1.9), called a simplicial component, which will prepare our transition to talking about simplicial complexes.

**Definition 2.1.16** (Simplicial Components). Let  $\mathcal{H}$  be a hypergraph. Suppose there is a connected component  $C \subset \mathcal{H}$  that is closed under taking subsets, i.e.

$$\forall e \in C, e' \subset e \in C \implies e' \in C.$$

Then *C* is called a simplicial component.

Another way of phrasing the subset closure property is that there is a connected component that contains all possible subedges of its toplexes. A hypergraph for which every component is a simplicial complex, as we will see in the next section.

## 2.2 Simplicial Complexes and their Homology

Like hypergraphs, simplicial complexes are combinatorial objects that generalize graphs. In fact, all graphs are 1-dimensional simplicial complexes. Moreover, simplicial complexes could be considered as a special class of hypergraphs. They are hypergraphs for which every component is a simplicial component (Definition 2.1.16), i.e. a hypergraph for which all possible subedges are present in the hypergraph. They are commonly used in algebraic topology. One nice property that simplicial complexes have that general hypergraphs do not have is a well-defined, standard homology theory. Abstract simplicial complexes were originally defined in [34] in the late 1930s and many of the definitions below appear in some form there. Unless another source is cited, versions of the definitions in this section can be found in [18] and/or [24].

## 2.2.1 Simplicial Complexes

We will begin with the definition of an abstract simplex (plural - simplices). Informally, an abstract simplex is a set of distinct vertices, just like an edge in a hypergraph.

**Definition 2.2.1** (Simplex). An abstract simplex  $\sigma$  is a set of vertices  $\sigma = \{v_0, v_1, ..., v_n\}$  such that  $v_i \neq v_j$  for any  $i \neq j$ , i.e. a set without any repeated elements. The dimension of the simplex  $\sigma$  is n, one less than the number of vertices it has.

In this thesis we will consider abstract simplices, but simplices have geometric realizations as well. Suppose the n + 1 vertices of an n-simplex  $\sigma$  lie in  $\mathbb{R}^n$  such that  $\{v_1 - v_0, \ldots, v_n - v_0\}$  is a linearly independent set. Then the geometric realization of  $\sigma$  is the convex hull of its vertices in  $\mathbb{R}^n$  [24]. Henceforth, all simplices will be abstract simplices (the word abstract will be omitted), and figures will show a geometric realization.

Simplices are the building blocks for simplicial complexes. An abstract simplicial complex is a set of distinct simplices that is closed under taking subsets of its elements. Simplicial complexes also have geometric realizations, as the union of the geometric realization of their simplices, and similarly, outside of figures, all simplicial complexes will be abstract simplicial complexes.

**Definition 2.2.2** (Simplicial Complex). A (finite) simplicial complex K is a collection of simplices



Figure 2.5 Of these two hypergraphs,  $\mathcal{H}$  and K, only K is a simplicial complex.  $K = \{\sigma_1, \sigma_2, ..., \sigma_m\}$  such that:

- 1.  $\sigma_i \neq \sigma_j$  for any  $i \neq j$  and
- 2. for all  $\sigma_i \in K$ , if  $\tau \subset \sigma_i$ , then  $\tau \in K$ .

The dimension of a simplicial complex is the maximum dimension over all of its simplices. All simplicial complexes considered will be assumed to be finite in both the number of simplices and their dimensions.

Hypergraphs were defined after simplicial complexes, historically, and are generalizations of simplicial complexes. Every simplicial complex is a hypergraph, but the converse is not true. The following lemma relates to Definition 2.1.16 of simplicial components in a hypergraph. It says that a hypergraph with all of its possible subedges is a simplicial complex.

**Lemma 2.2.1** (Hypergraphs that are Simplicial Complexes). Let  $\mathcal{H}$  be a hypergraph. If for all  $e \in \mathcal{H}, e' \subset e \implies e' \in \mathcal{H}$ , then  $\mathcal{H}$  is a simplicial complex.

*Proof.* Since all hypergraphs are assumed to be edge-collapsed (Definition 2.1.2), any hypergraph satisfies the first condition of Definition 2.2.2. The second condition of Definition 2.2.2 is exactly the assumption made in the lemma, proving that a hypergraph with subedge closure is a simplicial complex.

A first example of viewing a hypergraph as a simplicial complex is in Figure 2.5. The hypergraph  $\mathcal{H}$  is missing some subedges, like AC, and so is not a simplicial complex. However, the hypergraph

*K* is closed under taking subedges, and is thus a simplicial complex. The rightmost image in the figure shows the geometric realization of *K* as a shaded triangle. Both the second hypergraph and the triangle represent the exact same set of sets, *K*, which is the minimal simplicial complex containing  $\mathcal{H}$ . It might be important to consider only a part of a simplicial complex and the first definition about simplicial complexes is that of a special subset of a simplicial complex, called a subcomplex.

**Definition 2.2.3** (Subcomplex of a Simplicial Complex). A subset of a simplicial complex that is itself a simplicial complex is called a subcomplex.

Here is one important type of subcomplex:

**Definition 2.2.4** (Generated Subcomplex). Given a simplicial complex K, let  $V' \subset V$  be a subset of the vertex set of K. Then, the subcomplex generated by V' is the set of simplices in K whose vertices are contained in V', i.e.

$$K_{V'} = \{ \sigma \in K \mid \sigma \subseteq V' \}.$$

The next special type of subcomplex is the *n*-skeleton associated with a simplicial complex.

**Definition 2.2.5** (Skeleta of a Simplicial Complex and Underlying Graph). The *n*-skeleton of a simplicial complex is the subcomplex consisting of all of the simplices that are at most dimension *n*. The 1-skeleton of a simplicial complex is a graph, called the underlying graph in [2].

Many important results in algebraic topology rely on the idea of refining a simplicial complex (see, for example, the Simplicial Approximation Theorem in [24]). One standard way of doing so is called the barycentric subdivision, defined below. As the names restricted and relative barycentric homology suggest, both of the main homology theories developed herein utilize the barycentric subdivision.

**Definition 2.2.6** (Barycentric Subdvision). Given a simplicial complex K, its barycentric subdivision T is a simplicial complex such that the vertex set of T is the set of simplices of K, i.e.

$$V_T = \{ v_{\sigma_i} \mid \sigma_i \in K \}$$



Figure 2.6 A simplicial complex K, its barycentric subdivision T, and a generated subcomplex. and the simplices of T are given by the face relations on K. That is,

 $T := \{\Omega = \{v_{\sigma_1}, \dots, v_{\sigma_m}\} \subset K \mid \text{upto reordering of vertices in } \Omega, \sigma_1 \subset \sigma_2 \subset \dots \subset \sigma_m \in K\}.$ 

Figure 2.6 gives an example of both a barycentric subdivision and a generated subcomplex (Definition 2.2.4). It starts with the simplicial complex *K*, given by the triangle *ABC* and all of its subsets,  $K = \{ABC, AB, AC, BC, A, B, C\}$ . Each of these becomes a vertex in *T*, the barycentric subdivision. Notice, for example, that there is an edge between the vertices *A* and *AB* in *T*. This is because in *K*,  $A \subset AB$ . Similarly, there is a triangle on the vertices *C*, *AC*, *ABC* in *T* because in *K*,  $C \subset AC \subset ABC$ . The rightmost picture highlights the subcomplex in *T* generated by the vertices  $\{A, B, AC, BC\}$ , which includes all simplices in *T* only consisting of those vertices.

Recall that in graph theory, clique is the term for a complete subgraph. If a set of *m* vertices of a graph is a clique, then all smaller subsets of those same vertices will also be a clique. Thus, it might be advantageous to turn the cliques of a graph into simplices of a simplicial complex. This construction is called a clique complex.

**Definition 2.2.7** (Clique Complex). Let *G* be a graph. Cliques in *G* are complete subgraphs, i.e. sets of vertices in which every two vertices are joined by an edge. Form a simplicial complex X(G) by letting every clique in *G* be a simplex in X(G). Then X(G) is called the clique complex of *G*. If *K* is a simplicial complex, it is called a clique complex (or flag complex) if it is the clique complex of its 1-skeleton.

The last sentence of that definition is equivalent to the condition that every set of vertices that are pairwise connected by 1-simplices are mutually contained in a single simplex [2]. In hypergraphs, this condition is called conformality [2, 8]. As an example, the barycentric subdivision is always a clique complex. This is because, for a set of simplices in a simplicial complex, if they are pairwise related to each other, then there is a total order on that set. This can be seen in Figure 2.6. The vertices {C, AC, ABC} in T are each pairwise connected by an edge since  $C \subset AC, C \subset ABC$ , and  $AC \subset ABC$  in K. These can be rearranged into a single chain,  $C \subset AC \subset ABC$  in K, and so the corresponding triangle in T is shaded.

Recall the face poset FP from Definition 2.1.14. Since simplicial complexes are hypergraphs, it also makes sense to use this definition for the face poset of a simplicial complex. There is also a natural way to generate a simplicial complex given a poset. This is called the order complex [32]. It uses chains of relations in the poset as simplices for the simplicial complex.

**Definition 2.2.8** (Order Complex). Let (P, <) be a poset. The order complex of P, denoted  $\Delta(P)$ , is the simplicial complex whose vertices are the elements of P, and whose simplices are chains of the relation on P. In other words,  $\{v_0, v_1, \ldots, v_k\}$  is a simplex in  $\Delta(P)$  exactly when  $v_0 < v_1 < \ldots < v_k$  is a linearly ordered set in P.

One result that uses the order complex is that the order complex of the face poset of a simplicial complex is the same as the barycentric subdivision of that simplicial complex [32]. This gives an alternative definition of the barycentric subdivision.

**Proposition 2.2.2** (Alternative Definition of Barycentric Subdivision). Let *K* be a simplicial complex, and let FP(K) be its face poset. The order complex  $\Delta(FP(K))$  is the same simplicial complex as the barycentric subdivision, i.e.  $\Delta(FP(K)) = T_K$ .

*Proof.* By Definition 2.2.6 of the barycentric subdivision, the vertex set of  $T_K$  is

$$V_{T_K} = \{ v_\sigma \mid \sigma \in K \}$$

and the simplices in  $T_K$  are given by chains of face relations in K,

$$\Omega = \{v_{\sigma_1}, v_{\sigma_2} \dots, v_{\sigma_m}\} \in T_K$$

if and only if, for some reordering of the  $\sigma_i$ ,

$$\sigma_1 \subset \sigma_2 \subset \ldots \subset \sigma_m \in K.$$

The elements of the face poset of *K* are the simplices in *K*, with  $\sigma_i < \sigma_j$  in *FP*(*K*) if and only if  $\sigma_i \subset \sigma_j$  in *K*. The order complex  $\Delta(FP(K))$  has as vertices the elements of *FP*(*K*), which are the simplices in *K*. Therefore,  $\Delta(FP(K))$  and  $T_K$  have the same vertices. The simplices in  $\Delta(FP(K))$  are given by chains of relations in *FP*(*K*). Since in *FP*(*K*),

$$\sigma_1 < \sigma_2 < \ldots < \sigma_m$$

if and only if

$$\sigma_1 \subset \sigma_2 \subset \ldots \subset \sigma_m \in K,$$

the simplices of  $\Delta(FP(K))$  are the same as the simplices of  $T_K$ , and thus they are the same simplicial complex.

The last definitions on simplicial complexes before we get into their homology are that of the star and link. The star and link are important for many results on simplicial complexes that are of a topological nature, as they essentially analogize the concepts of an open neighborhood and its boundary to the combinatorial objects of simplicial complexes.

**Definition 2.2.9** (Stars and Links of Simplices). Let *K* be a simplicial complex, and  $\sigma \in K$  be a simplex. The (open) star of  $\sigma$ , denoted St( $\sigma$ ), is defined as the set of simplices having  $\sigma$  as a face:

$$\operatorname{St}(\sigma) = \{ \tau \in K \mid \sigma \subseteq \tau \}.$$

Let  $L = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$  be a subset of simplices in *K*, not necessarily a subcomplex. The star of *L* is the union of the stars of its simplices:

$$\operatorname{St}(L) = \bigcup_{k=1}^{n} \operatorname{St}(\sigma_k).$$

The closed star of L,  $\overline{\operatorname{St}(L)}$ , is the smallest subcomplex of K containing  $\operatorname{St}(L)$ .

The link of L, Lk(L), is the set of simplices in the closed star that are not in the star. This is the boundary of the star, given by

$$Lk(L) = St(L) \setminus St(L).$$

## 2.2.2 Simplicial Homology

Next we will move into defining the homology of simplicial complexes. Earlier, we mentioned that the presence of a standard homology theory was a feature that set simplicial complexes apart from hypergraphs. Recall that the difference between general hypergraphs and simplicial complexes is that simplicial complexes contain all of their subsets. This property makes a critical difference when it comes to defining homology. In short, simplices have natural boundaries while general hypergraph edges do not. Consider a 2-simplex with three vertices; its geometric realization is a triangle. Assuming it is in a simplicial complex, the three sides of the triangle (one dimensional simplices) are also guaranteed to be geometric realizations of simplices, and they make up the boundary of the simplex. On the other hand, in a general hypergraph, there might be an edge with three vertices, but not all (or possibly even none) of the two vertex subsets of that edge might be edges themselves.

This property plays an important role in the boundary map, a requirement for homology. In particular we will define a group for each dimension of simplices, and then the boundary map descends in dimension by one, as in the triangle to edges example above. This sequence of groups and maps forms the underlying structure required for a homology theory.

Let *K* be a simplicial complex. Define the *n*th chain group of *K* as the group of finite formal sums of *n*-dimensional simplices:

$$C_n(K) \coloneqq \left\{ \sum a_i \sigma_i \mid \sigma_i \text{ is an } n \text{-dimensional simplex } \in K \right\}.$$

The  $a_i$  are coefficients from a specified group; in this thesis the coefficient group is always assumed to be  $\frac{\mathbb{Z}}{2\mathbb{Z}}$ . Henceforth, when there is no ambiguity the *K* will be dropped, and the chain groups will

be denoted by  $C_n$ . There is a map  $\partial_n : C_n \to C_{n-1}$  called the boundary map defined as follows. If  $\sigma = \{v_0, ..., v_n\}$  is an *n*-simplex in *K*, then

$$\partial_p(\sigma) \coloneqq \sum_{i=0}^p (-1)^i [v_0, v_1, \cdots, \hat{v_i}, \cdots, v_n]$$

where the hat indicates removal of that vertex. The  $(-1)^i$  is not necessary when working with  $\frac{\mathbb{Z}}{2\mathbb{Z}}$  coefficients, but is included in the general definition of the boundary map for any coefficient group. The map  $\partial_n$  is then extended linearly to the chains of  $C_n$ . Importantly,  $\partial_{n-1} \circ \partial_n = 0$ . So there is a sequence of groups and maps:

$$\dots \xrightarrow{\partial} C_n \xrightarrow{\partial} C_{n-1} \xrightarrow{\partial} \dots \xrightarrow{\partial} C_0 \longrightarrow 0.$$

A sequence of groups and maps of this form where  $\partial^2 = 0$  is called a chain complex. If  $Ker(\partial_n) \cong Im(\partial_{n+1})$ , then the chain complex is said to be exact at *n*, and if the sequence is exact at every step, it is called an exact sequence. While  $C_n$  is not necessarily exact at *n*, we can measure its failure to be exact using homology.

**Definition 2.2.10** (Simplicial Homology). Given a simplicial complex K, form the chain complex C(K) of its chain groups and boundary maps as above. Then, the *n*th homology group of K,  $H_n(K)$ , is defined as

$$H_n(K) \coloneqq \frac{Ker(\partial_n)}{Im(\partial_{n+1})}.$$

The rank of this group is called the *n*-th Betti number of *K*, denoted  $\beta_n(K)$ .

Viewing a single vertex v as a 0-dimensional simplicial complex, we can compute that  $\beta_0(v) = 1$ , while all other Betti numbers are 0.

In some sense, homology generalizes the concept of finding, counting, and listing cycles in a graph. Since every graph is a 1-dimensional simplicial complex, it makes sense to talk about the simplicial homoloy of a graph. Let G be a graph. Then, it turns out that  $\beta_0(G)$  is equal to the number of connected components of the graph, and  $\beta_1(G)$  is equal to the number of linearly independent cycles in the graph. In higher dimensional simplicial complexes, however, some of those graph cycles might be filled in with two dimensional simplices. Perhaps there is even a two dimensional analogue of a cycle, this could look like an enclosed void in the geometric realization. Such a void would show up as a class in  $H_2(K)$ , the second dimensional homology group. This generalizes up to the dimension of the simplicial complex, above which the homology groups are guaranteed to be zero, since there are no simplices in those dimensions.

Instead of  $\beta_0 = 1$  for a single point, it would perhaps be more intuitive for a point to have no homology whatsoever. This is the case in the theory of reduced homology (see, e.g. [18]), which we will define next. Reduced homology is occasionally used when a result relies on it, for example in the proof of Theorem 4.2.9. This could occur when it will be convenient to say that, for a vertex  $v, \beta_i(v) = 0$  for all *i*.

**Definition 2.2.11** (Reduced Homology). Given a simplicial complex *K* with chain groups  $C_n$ , the reduced chain complex is the usual chain complex with  $\frac{\mathbb{Z}}{2\mathbb{Z}}$  augmented after  $C_0$ , i.e.

$$\ldots \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} C_{n-1} \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_1} C_0 \xrightarrow{\epsilon} \frac{\mathbb{Z}}{2\mathbb{Z}} \longrightarrow 0$$

where  $\epsilon$  maps a 0-chain to the sum of its coefficients. Then, the reduced homology of *K* is the homology of the reduced chain complex, denoted  $\widetilde{H}(K)$ , with reduced Betti numbers  $\widetilde{\beta}(K)$ .

The main property we need from reduced homology is that  $\tilde{\beta}_0(K) = \beta_0(K) - 1$ , and for  $i \neq 0$ ,  $\tilde{H}_i(K) \cong H_i(K)$ . Reduced homology in dimension zero does not represent the number of connected components. It instead represents gaps between connected components;  $\tilde{\beta}_0(K)$  is the number of 1-dimensional simplices that would need to be added to *K* to have a connected simplicial complex.

Another useful property of simplicial complexes is that the barycentric subdivision operator preserves homology, creating a refinement of a simplicial complex with the same topological properties, as shown in [18, 24].

**Lemma 2.2.3** (Homology of Barycentric Subdivision). Let K be a simplicial complex, with T its barycentric subdivision. Then,

$$H_n(T) \cong H_n(K)$$
 for all  $n$ .

Simplicial complexes that have the same topological properties as a point play an important role, and are given the name contractible.

**Definition 2.2.12** (Contractible Simplicial Complex). A simplicial complex K is said to be contractible if it is homotopy equivalent to a single point.

We will not define homotopy equivalence here (see [18], but this definition implies (among other things) that for reduced homology,  $\tilde{H}_n(K) = 0$  for all *n*. For the regular simplicial homology of a contractible simplicial complex,

$$H_n(K) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{when } n = 0\\ 0 & \text{else.} \end{cases}$$

In this thesis, the simplicity of the homology groups of a contractible simplicial complex will be often used to ease calculation of homology groups of more complicated complexes.

The last homological definition we will require is the relative homology of a pair  $S \subseteq K$ , where K is a simplicial complex, and S is a subcomplex. Geometrically, this gives the homology of a simplicial complex, after collapsing the subcomplex S to a point. Unfortunately, the quotient space K/S is not generally a simplicial complex, so it does not make sense to talk about the simplicial homology of K/S.

Let C(K) and C(S) be the chain groups of K and S, as defined above for the definition of simplicial homology. Then  $C_n(S) \subset C_n(K)$ , and in particular,  $C_n(K)/C_n(S)$  is a well defined quotient group in all dimensions. This sequence of groups, with the inherited boundary maps, still forms a chain complex [18]. This chain complex is denoted C(K, S) and is called the relative chain complex. Since it is a chain complex, the homology can be computed, and that is the definition of relative homology. **Definition 2.2.13** (Relative Homology). Let *K* be a simplicial complex, with *S* a subcomplex. Then, the homology of *K* relative to *S*, denoted H(K, S), is defined as the homology of the chain complex C(K, S) defined above.

Relative homology can be difficult to calculate. Some of the results in the next section will give additional ways to calculate it.

#### 2.2.3 Simplicial Homology Theorems

There are some important theorems of simplicial homology that we will analogize to hypergraphs in this thesis. The first of these is the Mayer-Vietoris Theorem, which relates the homology of a space to the homology of two of its subsets that cover it, as well as their intersection. This theorem can be found in [18, 24].

**Theorem 2.2.4** (Mayer-Vietoris Theorem). *Let K be a simplicial complex, and*  $L_1$ ,  $L_2$  *be subcomplexes such that*  $L_1 \cup L_2 = K$ . *Then the sequence of homology groups* 

$$\dots \to H_n(L_1 \cap L_2) \to H_n(L_1) \oplus H_n(L_2) \to H_n(K) \to H_{n-1}(L_1 \cap L_2) \to \dots$$

where the first two maps are induced by inclusions, and the third map is derived through techniques of homological algebra, is exact.

This long exact sequence is often called the Mayer-Vietoris sequence. It is commonly used in algebraic topology to calculate the homology of complicated spaces when the homology of a cover of them is known. It can be used with regular homology or reduced homology.

A first application of the Mayer-Vietoris sequence is to find the homology of the spheres,  $S^n$ , via induction, assuming that the homology of  $S^1$  has been computed as the base case. Recall that

$$H_i(S^1) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & i = 0, 1\\ 0 & \text{else} \end{cases}$$

For the inductive step, assume that  $H_i(S^i) = \frac{\mathbb{Z}}{2\mathbb{Z}}$  and consider  $S^{i+1}$ . Note that  $S^{i+1}$  can be split into slightly overlapping northern and southern hemisphere, which are both i + 1 dimensional disks,

 $D^{i+1}$ , so

$$S^{i+1} = D^{i+1} \cup D^{i+1}.$$

Each disk is homologically trivial, and they intersect in an equator of the sphere,

$$D^{i+1} \cap D^{i+1} = S^i.$$

Each of these pieces can now be plugged into the Mayer Vietoris sequence, giving

$$\dots \to H_{i+1}(S^i) \to H_{i+1}(D^{i+1}) \oplus H_{i+1}(D^{i+1}) \to H_{i+1}(S^{i+1}) \to H_i(S^i) \to H_i(D^{i+1}) \oplus H_i(D^{i+1}) \to \dots$$

Since disks are homologically trivial, it follows from the exactness of the sequence that

$$H_{i+1}(S^{i+1}) \to H_i(S^i)$$

is an isomorphism, and thus,

$$H_{i+1}(S^{i+1}) \cong \frac{\mathbb{Z}}{2\mathbb{Z}}.$$

The rest of the Mayer-Vietoris sequence yields the expected result, and can be applied in the same way. We will discuss another application of the Mayer-Vietoris sequence in Theorem 6.2.1. However, there is also a version of the Mayer-Vietoris sequence that uses relative homology, and that is the version we will be analogizing to hypergraphs in Theorem 4.1.1.

**Theorem 2.2.5** (Relative Mayer-Vietoris Theorem). Let K be a simplicial complex with subcomplexes  $L_1, L_2, S \subset K$  such that  $L_1 \cup L_2 = K$ . Further assume we have subcomplexes  $R_1 \subset L_1 \cap S$ , and  $R_2 \subset L_2 \cap S$  such that  $R_1 \cup R_2 = S$ . Then the sequence

$$\dots \to H_n(L_1 \cap L_2, R_1 \cap R_2) \to H_n(L_1, R_1) \oplus H_n(L_2, R_2) \to H_n(K, S) \to H_{n-1}(L_1 \cap L_2, R_1 \cap R_2) \to \dots$$

is exact.

Next up is the main result of relative homology. This is another long exact sequence, one that relates the relative homology H(K, S) to H(K) and H(S). This sequence will play a major role in Chapters 4 and 5.

**Theorem 2.2.6** (Long Exact Sequence of a Pair). *Given a simplicial complex K and a subcomplex S, the following is a long exact sequence of homology groups:* 

$$\dots \to H_n(S) \to H_n(K) \to H_n(K, S) \to H_{n-1}(S) \to H_{n-1}(K) \to \dots$$

The exactness of this sequence implies that in order to compute the homology H(K, S), we only need H(K), H(S), and knowledge of the linear maps between them. Computationally, leveraging this sequence will be vital as an alternative to taking the relative homology directly, since K/S is in general not a simplicial complex.

The final result of this section is a way to remove simplices from simplicial complexes without changing their homology, called a collapse. This will be analogized to hypergraphs in Theorem 3.2.6 and also used in Proposition 4.3.10.

**Theorem 2.2.7** (Simplicial Collapse). *Assume*  $\sigma < \tau$  *are simplices in a simplicial complex K such that:* 

- 1.  $\tau$  is a maximal face of K, and
- 2.  $\tau$  is the only maximal coface of  $\sigma$ .

*Then, if*  $\overline{K} := K \setminus \{ \omega \mid \sigma \leq \omega \leq \tau \}, K \sim \overline{K}$  are homotopy equivalent.

Simplicial complexes that collapse to a single vertex are called collapsible. Collapsible simplicial complexes are always contractible, but the converse is not true.

### 2.3 Hypergraph Homology Theories

We are now ready to merge the topics of the prior two sections - hypergraphs, simplicial complexes, and simplicial homology. Hypergraphs are more general than simplicial complexes. The lack of the subset closure requirement means that more data can be modeled and that some data can be modeled more accurately with hypergraphs than simplicial complexes or graphs [1, 3, 4, 6, 16, 17, 19, 21, 20, 22, 28, 31].



Figure 2.7 Comparing a hypergraph and simplicial model of a biological system.

Here is an example from biology [23]. Researchers found a species of panic grass that was surviving in hostile conditions that it would not typically have been able to survive. Originally, they located a fungus on the plant that was in a symbiotic relationship with the plant. After more research, though, they also found a virus that was on the fungus that was a part of this symbiosis. Researchers isolated and removed the virus, and both the fungus and the plant died. Researchers isolated and removed the fungus, and both the virus and the plant died. Researchers killed the plant, and both the fungus and virus died. No pair of them could survive without the third, in what appears to be a three-way symbiotic relationship. This is much easier to represent with a hypergraph than a graph or simplicial complex. If attempting to represent this with a graph, all three (the virus, fungus, and plant) would need to be connected somehow, but each edge in the graph can only connect two of them, which might mislead the reader into thinking that two of them connected by an edge could survive. As a simplicial complex, there would be a 2-dimensional simplex that represents the three way relationship, but this simplex would also necessarily contain all of its subsets, including all of the two-way edges. As a hypergraph, there only needs to be a three vertex edge and no possibly misleading subedges. This can be seen in Figure 2.7.

Homology has a history of being a useful tool for mathematicians. In topology, it is an important invariant for classifying spaces. In topological data analysis, researchers use persistent homology to measure how their data is changing over a parameter. This is done by taking the homology of simplicial complexes built at critical points in the parameter's range. Being able to define a homology theory for hypergraphs will allow researchers to use these results to classify and study hypergraphs, and analyze data that is better modeled using a hypergraph.

It is when we reach the point of defining a homology theory on hypergraphs that their generality becomes a double-edged sword. Because of the lack of the downward closure property that simplicial complexes have, hypergraphs can be used to model data in ways that sometimes better reflects its internal structure. However, the lack of that property also means there is not a natural boundary map on hypergraph edges that immediately analogizes the definition of simplicial homology. From the example in Figure 2.7, in order to more accurately represent the data the hypergraph does not contain the pairwise edges between the vertices. The lack of those edges means that there is not a natural boundary map for the edge that is present, and so it does not fit into our understood framework for a homology theory.

The main complication, therefore, in defining a hypergraph homology theory is finding a chain complex and boundary map that adequately and accurately represent the hypergraph. The bulk of my thesis consists of development of the theory for two related definitions of hypergraph homology, as originally defined by Emilie Purvine and collaborators [28]. These are the restricted barycentric homology and the relative barycentric homology. There also exist other homology and cohomology theories for hypergraphs, like the embedded homology [7, 29, 30, 33], and others [9, 12, 14, 25, 26].

Since we want to use simplicial homology, we will first start by taking a hypergraph, and mapping it into a simplicial complex. One natural way of turning a hypergraph into a simplicial complex is called the associated simplicial complex of the hypergraph [21, 25].

**Definition 2.3.1** (Associated Simplicial Complex). Let  $\mathcal{H}$  be a hypergraph. The associated simplicial complex of  $\mathcal{H}$ , denoted K, is the simplicial complex constructed as follows:

$$K = \{ \sigma \subset V \mid \sigma \subset e \text{ for some } e \in E \}.$$

The associated simplicial complex is the smallest simplicial complex that contains the hypergraph  $\mathcal{H}$ . The toplexes (Definition 2.1.4) of  $\mathcal{H}$  are the maximal faces of K. An initial approach to hypergraph homology, then, may be to define the homology of a hypergraph as the homology of its associated simplicial complex, as done in [26, 25]. However, many distinct hypergraphs can have the same associated simplicial complex, as we will explain. Because of the downward closure property, a simplicial complex is completely characterized by its maximal faces. Thus, the toplexes of a hypergraph determine its associated simplicial complex. Take, for example, a hypergraph whose only toplex is the edge *ABC*. There are six possible subedges: *A*, *B*, *C*, *AB*, *AC*, *BC*. This means that there are  $2^6 = 64$  different hypergraphs that have that identical associated simplicial complex, even with such a small example. The set of hypergraphs with the same toplexes, and hence, associated simplicial complex, is called a hyperblock [21]. In order to distinguish between different hypergraphs with the same associated simplicial complex, a more nuanced homology definition is needed.

The associated simplicial complex is the simplicial complex attained by forcing the hypergraph to be closed under taking subedges. Essentially, we are adding all of the missing subedges to the hypergraph to get the associated simplicial complex. Each edge  $e \in \mathcal{H}$  has a corresponding simplex  $\sigma_e \in K$ . Typically,  $e \in K$  will be used to indicate a simplex in K that represents an edge e in the hypergraph. However, if that e was also used in close proximity to talk about the hypergraph edge, we will use  $\sigma_e \in K$  when talking about the simplex in K. Another notation will come up when talking about the barycentric subdivision T of K. Edges in the hypergraph are simplices in the associated simplicial complex, and hence, vertices in its barycentric subdivision. When there is no risk of confusion,  $e \in T$  will be used to denote a vertex in the barycentric subdivision that represents an edge in the hypergraph. In a situation where that might seem ambiguous,  $v_e$  or  $v_{\sigma_e} \in T$  will denote that a vertex in the barycentric subdivision represents the simplex  $\sigma$  in K.

Both homology theories we will develop herein utilize the associated simplicial complex of the hypergraph and its barycentric subdivision. We will start by defining the restricted barycentric homology. These homology theories were originally defined by Emilie Purvine and collaborators [28]. The notation and formal statements of the definitions are original to this thesis.
#### 2.3.1 Restricted Barycentric Homology

The idea behind the restricted barycentric homology is that we should restrict our study to only the simplices that represent edges in the hypergraph. If one were to attempt to restrict the associated simplicial complex K to the simplices e representing edges in  $\mathcal{H}$ , the resulting set is not necessarily a simplicial complex, as it is the hypergraph  $\mathcal{H}$  exactly. Thus it is ill-defined to discuss its simplicial homology. The definition of restricted barycentric homology works around this pitfall by first passing to the barycentric subdivision T. Each edge  $e \in \mathcal{H}$  is represented by a vertex  $v_e \in T$ . Build the subcomplex generated by those vertices as in Definition 2.2.4. This subcomplex contains only simplices whose vertices all represent edges in the hypergraph and so, better reflects the structure of the hypergraph. During the results section in Chapter 3, it will become apparent that the restricted barycentric homology does well quantifying information about inclusion relations among edges that are present in the hypergraph.

**Definition 2.3.2** (Restricted Barycentric Subdivision). Let  $\mathcal{H}$  be a hypergraph. Let K be its associated simplicial complex, and let T be the barycentric subdivision of K. The restricted barycentric subdivision R of  $\mathcal{H}$  is a subcomplex  $R \subset T$  constructed as follows:

$$R = \{ \Omega \in T \mid \forall v \in \Omega, v = v_e \text{ for some } e \in \mathcal{H} \}$$

Each simplex in *R* consists of vertices that represent edges in  $\mathcal{H}$ . The next result says that if a set of vertices  $v_{e_i} \in R$  forms a simplex in *R*, those edges  $e_i \in \mathcal{H}$  are totally ordered by inclusion in the hypergraph. This is useful to help get some intuition about what the restricted barycentric subdivision can tell us about the hypergraph. It will also be utilized later when discussing alternate ways of constructing *R*, for example in 3.1.1.

**Proposition 2.3.1.** Let  $\mathcal{H}$  be a hypergraph, and let  $\Omega = \{v_{e_1}, v_{e_2}, \dots, v_{e_k}\}$  be a simplex in R, its restricted barycentric subdivision. Then for some ordering of the  $e_i, e_1 \subset e_2 \subset \ldots \subset e_k$  as edges in  $\mathcal{H}$ .

*Proof.* By Definition 2.3.2 of the restricted barycentric subdivision, each vertex  $v_{e_i}$  in  $\Omega$  is an edge  $e_i$  in  $\mathcal{H}$ , and a simplex  $\sigma_{e_i} \in K$ . By Definition 2.2.6, simplices in the barycentric subdivision are



Figure 2.8 A hypergraph  $\mathcal{H}$  and its barycentric subdivision T with the restricted barycentric subdivision highlighted.

given by chains of face relations in *K*. Thus for some ordering,  $\sigma_{e_1} \subset \sigma_{e_2} \subset \ldots \subset \sigma_{e_k}$  as simplices in *K*. However, the simplices  $\sigma_{e_i}$  are set theoretically the same in *K* and as edges  $e_i$  in  $\mathcal{H}$ , so  $e_1 \subset e_2 \subset \ldots \subset e_k$  as edges in  $\mathcal{H}$  as well, proving the proposition.

We will now define the restricted barycentric homology for hypergraphs. By first defining the restricted barycentric subdivision R, we have done most of the leg work. The restricted barycentric homology is defined as the simplicial homology of R.

**Definition 2.3.3** (Restricted Barycentric Homology). Let  $\mathcal{H}$  be a hypergraph, and R its restricted barycentric subdivision as constructed above. The restricted barycentric homology of  $\mathcal{H}$ , denoted  $H^{res}(\mathcal{H})$  is defined to be the simplicial homology of R, i.e.

$$H_n^{res}(\mathcal{H}) \coloneqq H_n(R).$$

The rank of  $H_n^{res}(\mathcal{H})$  will be denoted  $\beta_n^{res}(\mathcal{H})$  and called the *n*th restricted Betti number of the hypergraph.

We will go through an example of the restricted barycentric homology with the hypergraph in Figure 2.8. This hypergraph has two toplexes, ABC and BCD, so its associated simplicial complex (not shown) is the triangles ABC and BCD, glued along their shared edge BC. Shown in the picture is the barycentric subdivision T of that associated simplicial complex. The restricted barycentric subdivision R is highlighted. Here, R is generated by the four vertices in T that represent the

four edges in  $\mathcal{H}$ : *ABC*, *BCD*, *B*, *C*. Since  $B \subset ABC$ , there is an edge in *R* between *B* and *ABC*, and similarly for the other highlighted edges. This creates a hollow diamond shape, which has simplicial homology

$$H_n(R) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{when} n = 0, 1\\ 0 & \text{else.} \end{cases}$$

Therefore, the restricted barycentric homology of  $\mathcal H$  is

$$H_n^{res}(\mathcal{H}) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{when } n = 0, 1\\ 0 & \text{else.} \end{cases}$$

### 2.3.2 Relative Barycentric Homology

Whereas the restricted barycentric homology theory accounted for the subedges that were missing in the hypergraph by simply omitting them, the relative barycentric homology takes a different approach. The process begins the same way, by taking the associated simplicial complex K of a hypergraph  $\mathcal{H}$ , and then moving to the barycentric subdivision T. Once again, we take a particular subcomplex of T, which we call the missing subcomplex, defined below. The missing subcomplex is defined in almost the same way as the restricted barycentric subdivision, as a generated subcomplex on a vertex set. The difference is that the vertex set of the missing subcomplex consists of the vertices  $v \in T$  that do not represent edges in  $\mathcal{H}$ . It is the subcomplex in T built by vertices representing all of the missing subedges of edges in  $\mathcal{H}$ .

**Definition 2.3.4** (Missing Subcomplex). Let  $\mathcal{H}$  be a hypergraph. Let K be its associated simplicial complex, and let T be the barycentric subdivision of K. The missing subcomplex S of  $\mathcal{H}$  is a subcomplex  $S \subset T$  constructed as follows:

$$S = \{ \Omega \in T \mid \forall v \in \Omega, v \neq v_e \text{ for any } e \in \mathcal{H} \}.$$

Similarly to the restricted case, defining the missing subcomplex was most of the work towards defining the relative barycentric simplicial homology. The next definition is purely for notational convenience while discussing examples.

**Definition 2.3.5.** Let  $\mathcal{H}$  be a hypergraph, with T the barycentric subdivision of its associated simplicial complex K. Let S be the missing subcomplex as defined above. Then the relative barycentric subdivision (T, S) of  $\mathcal{H}$  will be defined as the geometric realization of the barycentric subdivision T with the missing subcomplex S labeled.

Now we are ready to define the relative barycentric homology of a hypergraph  $\mathcal{H}$ . It is defined as the homology of the barycentric subdivision T relative to the missing subcomplex S, using relative homology from Definition 2.2.13. In Chapter 4, it will become apparent that one of the strengths of the relative barycentric homology is to quantify information about the subedges of  $\mathcal{H}$ , and in particular the relationship between the subedges that are present and the possible subedges that are missing from  $\mathcal{H}$ .

**Definition 2.3.6** (Relative Barycentric Homology). Let  $\mathcal{H}$  be a hypergraph with barycentric subdivision *T*, and missing subcomplex *S* as above. Then the relative barycentric homology of  $\mathcal{H}$ , denoted  $H^{rel}(\mathcal{H})$ , is defined as the homology of *T* relative to *S*:

$$H_n^{rel}(\mathcal{H}) \coloneqq H_n(T,S)$$

The rank of  $H_n^{rel}(\mathcal{H})$  will be denoted  $\beta_n^{rel}(\mathcal{H})$  and called the *n*-th relative Betti number of the hypergraph.

For a first example of the relative barycentric homology, see Figure 2.9. The hypergraph  $\mathcal{H}$  has a single edge *ABC*. Its associated simplicial complex is the two-simplex *ABC* and its subsets. Notice that this is, topologically, a 2-dimensional disk. The barycentric subdivision *T* is shown. The missing subcomplex is generated by all of the missing subsets of *ABC* in  $\mathcal{H}$ , which is all of its proper subsets {*AB*, *AC*, *BC*, *A*, *B*, *C*}. These are connected to form the missing subcomplex, highlighted in the figure. The missing subcomplex *S* in this case forms the entire boundary of *T*. Now the relative barycentric homology  $H_n^{rel}(\mathcal{H}) = H_n(T, S)$ . A disk relative to its boundary is the sphere of the same dimension [18], so here  $H_n(T, S) = \widetilde{H_n}(\mathbb{S}^2)$ , where  $\mathbb{S}^2$  is the 2-dimensional



Figure 2.9 A hypergraph  $\mathcal{H}$  and its barycentric subdivision T with the missing subcomplex highlighted.

sphere. Recalling the homology groups of spheres from algebraic topology texts such as [18, 24] yields that

$$H_n^{rel}(\mathcal{H}) = H_n(T, S) = \widetilde{H_n}(\mathbb{S}^2) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{when } n = 2\\ 0 & \text{else.} \end{cases}$$

The rest of this thesis contains new results, almost all of which utilize either the restricted barycentric homology or the relative barycentric homology of hypergraphs. We start with the restricted barycentric homology in the next chapter.

### **CHAPTER 3**

#### **RESULTS ON THE RESTRICTED BARYCENTRIC HOMOLOGY**

We will begin the results portion of this thesis with the restricted barycentric homology. Throughout this chapter, R will be used to denote the restricted barycentric subdivision (Definition 2.3.2) of a hypergraph  $\mathcal{H}$ .

### 3.1 Restricted Barycentric Subdivision

We begin this section with results on the combinatorial structure of the restricted barycentric subdivisiOn R itself before moving into results on the homology. A main component of this thesis, discussed in Section 5.2, is the computation of these homology theories. In general, computing the barycentric subdivision of a simplicial complex is expensive, as the number of simplices grows very quickly. Sepcifically, the barycentric subdivision of an n-dimensional simplex has (n + 1)! n-dimensional simplices. Therefore, where possible, we will use alternate constructions that do not require a full barycentric subdivision. To that end, the first two results of this section are alternative constructions of the restricted barycentric subdivision R.

The following proposition is used as the definition of the restricted barycentric subdivision in [22]. It says that R is the order complex (Definition 2.2.8) of the face poset (Definition 2.1.14) of the hypergraph.

**Proposition 3.1.1** (Poset Construction of *R*). *The restricted barycentric subdivisIon R of a hyper*graph  $\mathcal{H}$  is the same simplicial complex as the order complex of the face poset of  $\mathcal{H}$ ,

$$R = \Delta(FP(\mathcal{H})).$$

*Proof.* Let  $\mathcal{H}$  be a hypergraph. Denote its restricted barycentric subdivision by R. From  $\mathcal{H}$ , form the poset FP and its order complex  $\Delta(FP)$ . Let  $\Omega$  be a simplex in R. Then, after a relabelling if necessary,  $\Omega = \{v_{e_1}, v_{e_2}, \dots, v_{e_k}\}$  where  $e_i \subset e_j$  in  $\mathcal{H}$  for all i < j. Note

$$e_1 \subset e_2 \subset e_3 \ldots \subset e_k$$
 in  $\mathcal{H}$ 

by Proposition 2.3.1. Thus,

in 
$$FP(\mathcal{H}), e_1 < e_2 < \dots < e_k$$
,

which implies  $\sigma$  is a simplex in  $\Delta(FP)$  as well. Therefore  $R \subset \Delta(FP)$ .

Let  $\sigma = \{e_1, e_2, \dots, e_k\}$  be a simplex in  $\Delta(FP(\mathcal{H}))$ . By definition, a simplex in the order complex of a poset represents a chain of inclusions in the poset,

$$e_1 < e_2 < \ldots < e_k \text{ in } FP(\mathcal{H}).$$

The relation on the poset  $FP(\mathcal{H})$  was given to be edge inclusion in  $\mathcal{H}$ , so this means

$$e_1 \subset e_2 \subset \ldots \subset e_k textin \mathcal{H}$$

. Therefore,  $\sigma = \{v_{e_1}, v_{e_2}, \dots, v_{e_k}\}$  is also a simplex in *R* by Proposition 2.3.1, and  $\Delta(FP) \subset R$ , so  $\Delta(FP) = R$ .

Recall the edge inclusion graph EIG (Definition 2.1.15) of a hypergraph. The second result is that its clique complex (Definition 2.2.7) is also equal to R. This is the construction that will be used later for computation, since it does not require the entire barycentric subdivision.

**Theorem 3.1.2** (Clique Construction of *R*). *The restricted barycentric subdivision of a hypergraph*  $\mathcal{H}$  is the same simplicial complex as the clique complex (Definition 2.2.7) of the edge inclusion graph (Definition 2.1.15) of  $\mathcal{H}$  that is R = X(EIG).

*Proof.* Let  $\mathcal{H}$  be a hypergraph. Denote its restricted barycentric subdivision with R. Build the edge inclusion graph, *EIG*. The vertices of *EIG* are edges in  $\mathcal{H}$  so let k vertices,  $v_{e_1}, ..., v_{e_k}$ , be a clique in *EIG*. This yields a (k - 1)-simplex in X(EIG). Inclusion is a partial order, and these k vertices represent edges  $e_1, ..., e_k$  that are each pairwise comparable since it is a clique. Thus there is a linear order (potentially after reordering),

$$e_1 \subset e_2 \subset \ldots \subset e_{k-1} \subset e_k \text{ in } \mathcal{H}.$$

This chain of edge inclusions gives rise to a matching (k - 1)-simplex in R, so  $X(EIG) \subset R$ .

Let  $\sigma$  be a k simplex in R. This corresponds to a chain of k + 1 edges,

$$e_1 \subset e_2 \subset \ldots \subset e_k \subset e_{k+1}$$
 in  $\mathcal{H}$ .

Each of these edges is a vertex in the edge inclusion graph, and since subset is a transitive relation, each of those edges has a pairwise inclusion relation with every other, forming a (k + 1)-clique in the graph, which corresponds to a *k* simplex in is clique complex X(EIG), so  $R \subset X(EIG)$ . Therefore, R = X(EIG).

When starting with the associated simplicial complex and barycentric subdivision, all of the possible subedges are needed as vertices, not just the subedges that are present in the hypergraph. Even worse, information about all possible chains of inclusion among all the possible edges is needed. In this construction, only the edges of  $\mathcal{H}$  are stored as vertices, and only their possible inclusion relations need to be checked.

### **3.2 Restricted Barycentric Homology**

Now we will develop the theory of the restricted barycentric homology. The first result is a plausibility check. If  $\mathcal{H}$  is a hypergraph that also happens to be a simplicial complex, its restricted barycentric homology is the same as the simplicial homology. We first note a minor technicality: as a result of the downward closure property, a simplicial complex will always contain the empty set, and by Definition 2.1.1, a hypergraph never will. So, we say that a hypergraph is the same as its associated simplicial complex when their set of nonempty elements are the same.

**Proposition 3.2.1.** Let *H* be a hypergraph that is the same as its associated simplicial complex, *K*. Then for all *n*,

$$H_n^{res}(\mathcal{H}) \cong H_n(K).$$

*Proof.* Since  $\mathcal{H}$  is already a simplicial complex, there are no missing subedges to add to obtain the associated simplicial complex *K* (Definition 2.3.1), and thus  $\mathcal{H} = K \setminus \{\emptyset\}$ . Let *T* be the barycentric subdivision of *K* as in Definition 2.2.6.

Let  $\Omega$  be a simplex in *T*. Then

$$\Omega = \{v_{\sigma_1}, v_{\sigma_2}, \ldots, v_{\sigma_k}\},\$$

where each  $\sigma_i$  is a simplex in *K*. However,  $K \setminus \{\emptyset\} = \mathcal{H}$ , and so each  $v_{\sigma_i} = v_{e_i}$  represents an edge  $e_i$  in  $\mathcal{H}$ . Thus by the definition of the restricted barycentric subdivision (Definition 2.3.2),  $\Omega \in R$ , and hence  $T \subset R$ . Since by definition  $R \subset T$ , R = T. We now have the following equalities and isomorphism:

$$H^{res}(\mathcal{H}) = H(R) = H(T) \cong H(K).$$

The first equality comes from Definition 2.3.3 of the restricted barycentric homology. The second equality comes from the set theoretic result R = T, proved above, and the isomorphism  $H(T) \cong H(K)$  was given as Lemma 2.2.3. This proves the proposition.

In words, when  $\mathcal{H}$  is a simplicial complex, the restricted barycentric subdivision is equal to the barycentric subdivision because all of the possible edges are actually present in the hypergraph. In the next result, let  $H^{res}(e_i)$  denote the restricted barycentric homology of the hypergraph with only one edge,  $e_i$ . Recall the definition of a reduced hypergraph from Definition 2.1.5 as a hypergraph with no subedges. Since the restricted barycentric homology is concerned primarily with inclusion relations, and a reduced hypergraph has no inclusion relations, a reduced hypergraph has a fairly simple, characterizable restricted barycentric homology. This characterization is given in the next theorem.

**Theorem 3.2.2** (Restricted Homology of Reduced Hypergraphs<sup>†</sup>). Let  $\mathcal{H}$  be a reduced hypergraph. Then  $\beta_0^{res}(\mathcal{H}) = |\mathcal{H}|$  and  $\beta_i^{res}(\mathcal{H}) = 0$  for i > 0.

*Proof.* Recall the restricted barycentric subdivision (Definition 2.3.2) is a simplicial complex with a vertex for each edge in the hypergraph, and simplices based on chains of inclusions of edges. In a reduced hypergraph, no edges are subedges, so each vertex is isolated and the restricted barycentric subdivision is a disjoint union of vertices, each representing one of the  $e_i$ . Since the homology of

disjoint topological spaces is equal to the direct sum of the homology of each of their components [18], this gives that  $H^{res}(\mathcal{H}) = \bigoplus_{i=1}^{m} H^{res}(e_i)$ , where *m* is the number of edges in the hypergraph.

Since the hypergraph with only the edge  $e_i$  is a maximum edge hypergraph, we can apply Theorem 3.2.3 to find its restricted Betti numbers,  $\beta_0^{res}(e_i) = 1$  and  $\beta_n^{res}(e_i) = 0$  for n > 0. Thus each vertex of R contributes a rank one subgroup to the zeroth dimensional homology of the simplicial complex. This means that  $\beta_0^{res}(\mathcal{H}) = |\mathcal{H}|$  and  $\beta_n^{res}(\mathcal{H}) = 0$  for n > 0, as was to be shown.

In a reduced hypergraph, every edge is a toplex and there are no subedges. The next result is about maximum edge hypergraphs (Definition 2.1.10), in which there is only one toplex. The restricted barycentric homology has useful implications for the inclusion relationships between toplexes and their subedges. In a maximum edge hypergraph, all of the edges are subedges of the single toplex. This leads to a classification of the restricted barycentric homology of maximum edge hypergraphs with nontrivial homology in only dimension zero.

**Theorem 3.2.3** (Restricted Homology of Maximum Edge Hypergraphs). Let  $\mathcal{H}$  be a maximum edge hypergraph. Then the restricted barycentric homology  $H_n^{res}(\mathcal{H})$  is as follows:

$$H_n^{res}(\mathcal{H}) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{if } n = 0\\ 0 & \text{else.} \end{cases}$$

*Proof.* Let  $\mathcal{H}$  be a maximum edge hypergraph, with maximum edge e. Construct R, the restricted barycentric subdivision (Definition 2.3.2) of  $\mathcal{H}$ . Note that e is a vertex in R, since it is an edge in the hypergraph and a simplex in K. Because all other edges are included in e,  $v_e$  is connected via an edge in R to all other vertices in R. If there are any higher dimensional simplices  $\Omega$  with vertices in  $R \setminus v_e$ , representing inclusions among subedges of e in the hypergraph, all of those edges are still contained in e, so  $\Omega \cup v_e$  is a simplex in R. Thus, R is a cone with apex  $v_e$ , and hence contractible ([18]), proving the lemma.

A maximum edge hypergraph has a single fence component (Definition 2.1.12). The next theorem says that this knowledge is enough to know the zeroth dimensional restricted barycentric

homology of the hypergraph. It turns out that the number of fence components of the hypergraph is equal to the rank of the zeroth dimensional homology group, as shown next.

**Theorem 3.2.4** (Zeroth Dimensional Restricted Homology<sup>†</sup>). The rank of the zeroth dimensional restricted barycentric homology group,  $\beta_0^{res}(\mathcal{H})$ , is equal to the number of fence components,  $\Gamma(\mathcal{H})$ , of the hypergraph.

*Proof.* Recall that  $\beta_0^{res}(\mathcal{H}) = \beta_0(R)$  is the number of path components of R where R is the restricted barycentric subdivision of  $\mathcal{H}$  (Definition 2.3.3). Suppose two vertices of R,  $v_e$  and  $v_f$ , are in the same path component of R. This means that there is a sequence of vertices  $v_e, v_{e_1}, v_{e_2}, ..., v_f$  such that there is an inclusion relationship between the hyperedges representing each pair of adjacent vertices in the sequence. In other words, in the hypergraph there is a sequence of edges  $e, e_1, e_2, ..., f$  where each pair of adjacent edges has an inclusion relationship in either direction. This is the definition of e and f being in the same fence component of  $\mathcal{H}$ .

Reversing that line of thought yields the inverse statement, and so two vertices in *R* are in the same path component of *R* if and only if their corresponding edges are in the same fence component of  $\mathcal{H}$ . Since there is a bijection between vertices of *R* and edges of  $\mathcal{H}$ , there will be the same number of path components of *R* and fence components of  $\mathcal{H}$ . Thus  $\beta_0^{res}(\mathcal{H}) = \Gamma(\mathcal{H})$ , proving the theorem.

In the above proof, a bijection is found between fence components of the hypergraph and path components of R. Each fence component of the hypergraph gives rise to a path component of R. Since the homology of a simplicial complex is the direct sum of the homologies of its path components [18], we get the following corollary.

**Corollary 3.2.5.** Let  $\mathcal{H}$  be a hypergraph, with fence components  $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_k$ . Then

$$H_n^{res}(\mathcal{H}) = \bigoplus_{i=1}^k H_n^{res}(\mathcal{H}_i)$$



Figure 3.1 Hypergraphs with different fence components than connected components.

As an application of Theorem 3.2.4, see the hypergraphs in Figure 3.1. The difference between  $\mathcal{H}$  and  $\mathcal{H}'$  is the edge B, which is only in  $\mathcal{H}'$ . This edge does not change the number of connected components as per Definition 2.1.9, but does change the number of fence components. In  $\mathcal{H}$ , there is no fence between the edge AB and BC, because they do not share a mutual subedge or toplex. However, in  $\mathcal{H}$ , AB, B, BC is a fence. Thus  $\beta_0^{res}(\mathcal{H}) = 2$  and  $\beta_0^{res}(\mathcal{H}') = 1$ .

The last theorem in this section uses the idea of a simplicial collapse (Theorem 2.2.7). It gives conditions on which edges can be removed from a hypergraph without changing the restricted barycentric homology.

**Theorem 3.2.6.** <sup>†</sup> Let  $\mathcal{H}$  be a hypergraph with an edge e such that there is exactly one edge f with  $e \subset f$  or  $f \subset e$ . Denote by  $\mathcal{H}'$  the hypergraph  $\mathcal{H} \setminus \{e\}$ . Then  $\forall n, H_n^{res}(\mathcal{H}) \cong H_n^{res}(\mathcal{H}')$ .

*Proof.* Let R, R' be the restricted barycentric subdivisions of  $\mathcal{H}, \mathcal{H}'$  respectively. Note that  $R = R' \cup \{ef, e\}$ . In R, ef is a maximal simplex. If ef was not maximal, then it would be a face of efg for some edge  $g \in E$ . This would imply that eg is a simplex in R, which would imply that either  $e \subset g$  or  $g \subset e$ , which cannot be true as e was assumed to only be related by inclusion to f. Since ef is the only maximal simplex that e is a face of, R can be simplicially collapsed along e and ef by Theorem 2.2.7. Removing ef and e from R leaves the same simplicial complex as R'. Therefore, R simplicially collapses to, and is thus homotopy equivalent to, R', proving the theorem.



Figure 3.2 Hypergraphs used as examples of Theorem 3.2.6.

The two hypergraphs in Figure 3.2 have the same restricted barycentric homology because the edge CD in  $\mathcal{H}$  only has C as a subedge, and thus can be removed by the theorem. We will further discuss the computation of the restricted barycentric homology in Section 5.2. As part of the conclusion, we will share some ideas for future research on the restricted barycentric homology of hypergraphs. For now, we are ready to move on to the results on the second definition of hypergraph homology studied in this thesis, the relative barycentric homology.

### **CHAPTER 4**

#### **RESULTS ON THE RELATIVE BARYCENTRIC HOMOLOGY**

In this chapter, we will discuss some results in the development of the theory of the relative barycentric homology for hypergraphs. Throughout, *S* will denote the missing subcomplex (Definition 2.3.4), and *T* will denote the barycentric subdivision (Definition 2.2.6) of the associated simplicial complex *K* (Definition 2.3.1) of the hypergraph.

This chapter has three sections. The first section contains two analogues of the Mayer-Vietoris sequence from Theorem 2.2.4. These allow for breaking large and complex hypergraphs down and taking the relative barycentric homology of subhypergraphs, before gluing them back together with the Mayer-Vietoris sequences. The first version is at the level of the hypergraph edges. It allows for computing the relative barycentric homology of a hypergraph by knowing the homology of two subhypergraphs and of the hypergraph whose edges are their intersection. The second method applies at the simplicial complex level, using the relative Mayer-Vietoris Theorem to compute the homology after taking the barycentric subdivision and finding the missing subcomplex. Therefore, the second Mayer-Vietoris Theorem of this section will be called the simplicial version. The drawback of the second method is that there is not a nice way to frame this theorem using edges of the original hypergraph, but it has the advantage of requiring one fewer condition. We will give an example of a hypergraph where the first version cannot be used to compute its relative barycentric homology, but the simplicial version can be used to show why both results are useful. The last topic of this section will be some implications of the hypergraph versions of the Mayer-Vietoris Theorem.

The second section, then, is useful for studying the subhypergraphs. Maximum edge hypergraphs (Definition 2.1.10) have a relative barycentric homology that is simpler to study than general hypergraphs. This is due to their associated simplicial complex *K* being contractible. We will utilize that fact and the long exact sequence of a pair of spaces from Theorem 2.2.6 as shortcuts to computing the relative barycentric homology of maximum edge hypergraphs. We will see that the relative barycentric homology  $H^{rel}(\mathcal{H})$  for a maximum edge hypergraph  $\mathcal{H}$  depends heavily on the homology of its missing subcomplex H(S). After showing that in generality, we move on to some results that help interpret what information the relative barycentric homology relates about the maximum edge hypergraph. This is done in dimensions 0, 1, m - 2, and m - 1, where m is the number of vertices in the hypergraph. Lastly, a condition is presented that would cause a maximum edge hypergraph to have no relative barycentric homology at all.

The third section gives results on the relative barycentric homology for general hypergraphs. These include an overall classification of the relative barycentric homology of reduced hypergraphs in all dimensions and of any hypergraph in dimension zero, comparable to Theorems 3.2.4 and 3.2.2 for the restricted barycentric homology. In dimension zero, the relative barycentric homology turns out to be entirely dependent on the simplicial components (Definition 2.1.16) of the hypergraph. There are more results about manipulating the long exact sequence of the pair (T, S), which maintains usefulness even when  $\mathcal{H}$  is not a maximum edge hypergraph. The last result of this chapter gives a condition on which simplicial collapses can be applied during relative barycentric homology computations.

#### 4.1 Two Versions of a Mayer-Vietoris Sequence

We will begin the results for the relative barycentric homology with two versions of the Mayer-Vietoris Theorem. Since this is a relative homology theory, both versions are built off of Theorem 2.2.5. In algebraic topology, the Mayer-Vietoris Theorem is very useful for finding the homology of a space when it can be decomposed into the union of two subspaces whose homology is already known, and the intersection of those two subspaces also has a known homology. Recall the *n*-dimensional sphere from the earlier example when we defined the Mayer-Vietoris sequence. It is the union of two hemispheres, which are *n*-dimensional disks. The intersection of the hemispheres is the equator, which is an (n - 1)-dimensional sphere. The Mayer-Vietoris theorem can thus inductively be used to prove the homology groups of the spheres.

The first version allows for the hypergraph  $\mathcal{H}$  to be written as  $\mathcal{H} = \mathcal{H}_1 \bigcup \mathcal{H}_2$ , and these  $\mathcal{H}_1$ and  $\mathcal{H}_2$  form the analogous constructions to the subspaces in the regular Mayer-Vietoris sequence (Theorem 2.2.4). This is helpful because it can be used to break down hypergraphs into smaller components for relative barycentric homology computations. Alternatively, it can be used to glue together hypergraphs along some intersecting edges and compute the homology of the new hypergraph.

**Theorem 4.1.1** (Mayer-Vietoris Theorem for Hypergraphs). *Fix a hypergraph*  $\mathcal{H} = \{V, E\}$ . *Let*  $\mathcal{H}_1 = \{V_1, E_1\}$  and  $\mathcal{H}_2 = \{V_2, E_2\}$  be subhypergraphs, chosen such that

- *1.*  $E_1, E_2 \subset E$  with  $E_1 \cup E_2 = E$ .
- 2. Let  $k \in \{1, 2\}$ . For any  $e_i \in \mathcal{H}$  and  $e_j \in E_k$  with  $e_i \subset e_j$ , we have that  $e_i \in E_k$ .
- 3. Let  $E' := E_1 \cap E_2$  and call the hypergraph it induces  $\mathcal{H}'$ . For all  $e_1 \in E_1$  and  $e_2 \in E_2$ ,  $\exists e' \in E'$  s.t.  $e_1 \cap e_2 \subset e'$ .

Then the following is a long exact sequence on the relative barycentric homology of hypergraphs:

$$\dots \to H_n^{rel}(\mathcal{H}') \to H_n^{rel}(\mathcal{H}_1) \oplus H_n^{rel}(\mathcal{H}_2) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}^{rel}(\mathcal{H}') \to \dots$$

*Proof.* The main objective of the proof is to show that the barycentric subdivisions and missing subcomplexes of  $\mathcal{H}, \mathcal{H}_1, \mathcal{H}_2$  and  $\mathcal{H}'$  satisfy the requirements of Theorem 2.2.5.

Note that because we have inclusions on hypergraphs,

$$\mathcal{H}_1$$
 $\mathcal{H}' \qquad \mathcal{H}$ 
 $\mathcal{H}' \qquad \mathcal{H}_2$ 
 $\mathcal{H}_2$ 

we have inclusions on the associated simplicial complexes,

$$\begin{array}{cccc}
 & K_1 \\
 & \Im & \Im \\
 & K' & K_2 \\
 & K_2
\end{array}$$

Define pairs (T, S), (T', S'),  $(T_1, S_1)$ , and  $(T_2, S_2)$  as in Definition 2.3.5 for  $\mathcal{H}$ ,  $\mathcal{H}'$ ,  $\mathcal{H}_1$ , and  $\mathcal{H}_2$  respectively. In each case, T is the barycentric subdivision of the relevant associated simplicial

complex *K*, and *S* is the missing subcomplex of *T* induced by the relevant hypergraph. For example, S' and T' are defined as

$$T' = \{ \Omega = \{ v_{\sigma_1}, \cdots, v_{\sigma_n} \} \mid \sigma_1 \subset \sigma_2 \subset \ldots \subset \sigma_n \subset K' \}, \text{ and}$$
$$S' = \{ \Omega \in T \mid \forall v_{\sigma_i} \in \Omega, \sigma_i \text{ does not represent an edge in } \mathcal{H}' \}$$

With this notation in place, the long exact sequence becomes

$$\dots \to H_n(T', S') \to H_n(T_1, S_1) \oplus H_n(T_2, S_2) \to H_n(T, S) \to H_{n-1}(T', S') \to \dots$$

We will show that our setting fits the setup of the relative Mayer-Vietoris sequence from Definition 2.2.5, which will give the theorem. In particular, we need to show four things: (i)  $T = T_1 \cup T_2$ ; (ii)  $S = S_1 \cup S_2$ ; (iii)  $T' = T_1 \cap T_2$ ; and (iv)  $S' = S_1 \cap S_2$ .

This will be done in order. First, we will show  $T = T_1 \cup T_2$ . If  $\Omega = \{v_{\sigma_1}, v_{\sigma_2}, \dots, v_{\sigma_k}\}$  is a simplex in *T*, then for some ordering of those vertices,  $\sigma_1 \subset \sigma_2 \subset \dots \subset \sigma_k \in K$ . Further,  $\sigma_k$  represents a subedge of an edge *e* in  $\mathcal{H}$ , in which all of the other  $\sigma_i$  represented by vertices in  $\Omega$  must also be contained. For the rest of this proof, the  $v_{\sigma_k}$  that represents the simplex in *K* containing all other vertices of a simplex  $\Omega$  in *T* will be called the maximal vertex of  $\Omega$ . This edge *e* must then be in  $E_1$  or  $E_2$ , hence  $\sigma_i$  for all  $i \in \{1, 2, \dots, k\}$  is in  $K_1$  or  $K_2$ , which implies then that  $\Omega$  is a simplex in  $T_1$  or  $T_2$ , i.e.  $\Omega \in T_1 \cup T_2$ .

Similarly, a simplex  $\Phi$  in  $T_1$  or  $T_2$  has a maximal vertex  $v_{\tau}$  corresponding to a simplex  $\tau$  in  $K_1$ or  $K_2$ , such that  $\sigma \subset \tau$  for all  $v_{\sigma} \in \Phi$ . This simplex  $\tau$  represents a subedge of an edge e' in  $E_1$  or  $E_2$ . Both  $E_1$  and  $E_2$  are subsets of E, so  $e' \in E$ . Thus  $\tau \in K$ , and, hence,  $\Phi \in T$ . Therefore, any simplex in  $T_1$  or  $T_2$  must also be in T. So,  $T_1 \cup T_2 = T$ .

Next, it is necessary to show  $S = S_1 \cup S_2$ . Recall that *S* is the subcomplex of *T* generated by the vertices  $v_{\sigma}$  that correspond to the simplices  $\sigma$  in *K* that are missing as subedges in  $\mathcal{H}$ . Suppose  $\Omega$  is a simplex in *S*. Then every vertex comprising  $\Omega$  corresponds to a simplex in *K* that is a missing subedge in  $\mathcal{H}$ . In particular, the maximal vertex of  $\Phi$  corresponds to a subedge of an edge *e* in *E*. Now, *e* must be in  $E_1$  or  $E_2$ . Thus,  $\Phi$  is in  $S_1$  or  $S_2$ , so  $S \subseteq S_1 \cup S_2$ .

Now let  $\Phi$  be in  $S_1 \cup S_2$ , say WLOG  $\Phi \in S_1$ . Then every vertex in  $\Phi$  represents a missing subedge of an edge in  $\mathcal{H}_1$ . By assumption two of the theorem for the construction of  $E_1$  and  $E_2$ , all of those vertices represent edges that are missing in E as well, thus  $\Phi$  is also in S. Therefore,  $S = S_1 \cup S_2$ .

The third step is to show that  $T' = T_1 \cap T_2$ . Let  $\Omega \in T'$ . Then all of the vertices in  $\Omega$  represent simplices in K' contained in the simplex representing the maximal vertex of  $\Omega$ , meaning they are all subedges of the edge e' in E' corresponding to that simplex. Since  $E' = E_1 \cap E_2$ , e' is in  $E_1$ and  $E_2$ , so all of its subedges are represented in  $K_1$  and  $K_2$ . Therefore  $\Omega \in T_1 \cap T_2$ , and hence  $T' \subseteq T_1 \cap T_2$ . Similarly, a simplex  $\Phi$  in both  $T_1$  and  $T_2$  has the same maximum edge in  $E_1$  and  $E_2$ , so that edge is in their intersection E', and thus  $\Phi$  is in T'. Therefore,  $T' = T_1 \cap T_2$ .

Lastly, it is necessary to verify that  $S' = S_1 \cap S_2$ . Let  $\Phi$  be a simplex in S', then every vertex in  $\Phi$  represents a missing subedge of an edge in E', hence a missing subedge of an edge in both  $E_1$  and  $E_2$ . Again by the second condition in the theorem on  $E_1$  and  $E_2$ ,  $\Phi$  is necessarily a simplex in both  $S_1$  and  $S_2$ .

If  $\Omega$  is a simplex in  $S_1 \cap S_2$ , then every vertex in  $\Omega$  represents a missing subedge of a hyperedge in  $E_1 \cap E_2 = E'$ , and so  $\Omega \in S'$ . Thus,  $S' = S_1 \cap S_2$ . Therefore the complexes satisfy the requirements of Definition 2.2.5, and so we have that the long exact sequence

$$\ldots \to H_n^{rel}(\mathcal{H}') \to H_n^{rel}(\mathcal{H}_1) \oplus H_n^{rel}(\mathcal{H}_2) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}^{rel}(\mathcal{H}') \to \ldots$$

is exact as required.

Condition 3 seems, and is, restrictive, but it is a necessary condition. Recall that the condition states that for any  $e_1$  in  $E_1$  and  $e_2$  in  $E_2$ , there is an edge e' in  $E_1 \cap E_2$  such that  $e_1 \cap e_2 \subset e'$ . In words, this means that every pairwise intersection of edges in  $E_1$  and  $E_2$  is contained in an edge in both. Consider the following hypergraph, which can be seen in Figure 2.1:

$$\mathcal{H} = \{ABC, BCD, B, C\}.$$

A natural choice for a partition of this hypergraph is to separate the 3-vertex edges, and then in order to satisfy condition 2 of the theorem,  $E_1 = \{ABC, B, C\}$  and  $E_2 = \{BCD, B, C\}$ . This partition, however, does not satisfy condition 3. The edges  $ABC \in E_1$  and  $BCD \in E_2$  intersect in BC, which is not an edge of E' or a subset therein. The vertices generating  $T' = T_1 \cap T_2$  are  $\{BC, B, C\}$ , but since  $B, C \in E$ , the only vertex generating  $S' = S_1 \cap S_2$  is  $\{BC\}$ . If there was an "intersection hypergraph" that would allow us to use the Mayer-Vietoris theorem for hypergraphs here, it would have to contain BC as both a missing subedge and a maximal edge. Of course, an edge can't be both a maximal edge and a missing subedge, so this is impossible. Thus, for this  $\mathcal{H}$  and the most natural partition, there is not a hypergraph whose relative barycentric division is (T', S'), and it will therefore be impossible to write the Mayer-Vietoris sequence where the intersection term is the relative barycentric homology of any hypergraph. The following lemma gives this result in generality.

**Lemma 4.1.2.** Suppose  $E_1$ ,  $E_2$  form a partition of a hypergraph and satisfy condition two of the Mayer-Vietoris Theorem for hypergraphs, but do not satisfy condition three. Then, (T', S') is not the relative barycentric subdivision of any hypergraph.

*Proof.* Let  $e_1 \in E_1$ ,  $e_2 \in E_2$  be such that  $e_1 \cap e_2$  is not a subset of any edge in  $E_1 \cap E_2$ ; in other words  $e_1$  and  $e_2$  are edges that contradict condition three. It suffices to assume that  $e_1 \cap e_2$  is maximal among intersections of edges in  $E_1$  and  $E_2$ . This is because if an intersection is not maximal, then a maximal intersection that it is contained in would also contradict the third condition. Since  $e_1 \cap e_2$  is maximal among intersections  $\{e_i \cap e_j \mid e_i \in E_1, e_j \in E_2\}$ ,  $\sigma_{e_1 \cap e_2}$  is a maximal simplex in  $K' = K_1 \cap K_2$ . There is thus also a corresponding vertex  $v_{\sigma_{e_1} \cap e_2}$  of  $T' = T_1 \cap T_2$ . Since  $e_1 \cap e_2$  is not a subedge of any edge in  $E_1 \cap E_2$ , by condition two  $v_{e_1} \cap v_{e_2}$  is a vertex in T' that is part of the generating set for S'. It will form part of the missing subcomplex. Suppose there was a hypergraph such that K' was its associated simplicial complex, with (T', S') being its relative barycentric subdivision. By the preceding sentences,  $e_1 \cap e_2$  is both a maximal edge and a missing subedge of that hypergraph, which is a contradiction. Therefore, no hypergraph with relative barycentric subdivision (T', S') can exist, proving the lemma.

All hope is not lost for a useful Mayer-Vietoris sequence on hypergraphs, though. If we pass through to the level of the relative barycentric subdivision, the relative version of the Mayer-Vietoris sequence (Theorem 2.2.5) will still hold and be useful to compute the homology of a hypergraph in terms of a direct sum of two of its subgraphs. In this case, simply defining T' as  $T_1 \cap T_2$  and S'as  $S_1 \cap S_2$  still ensures

$$\dots \to H_n(T', S') \to H_n(T_1, S_1) \oplus H_n(T_2, S_2) \to H_n(T, S) \to H_{n-1}(T', S') \to \dots$$

is an exact sequence. The difference here is that  $H_*(T', S')$  is not the relative barycentric homology of the hypergraph  $\mathcal{H}_{E_1 \cap E_2}$ , which is the hypergraph whose edgeset is  $E_1 \cap E_2$ . The simplicial complex T' still can be given a physical interpretation in terms of the hypergraph, though. A simplex in T' is a simplex in  $T_1 \cap T_2$  and the simplices in  $T_i$  are the simplices in T that only contain the representations of subsets of edges in  $E_i$  as their vertices. Thus T' is generated by (as in Definition 2.2.4) the subsets  $V_j = \{v_{\sigma_0}, ..., v_{\sigma_k} \mid \forall i, \sigma_i \subset e_1 \in E_1, \text{ and } \sigma_i \subset e_2 \in E_2\}$ . Physically, then, T' is still the simplicial complex whose vertices are subedges of edges in  $E_1$  and  $E_2$  and whose simplices are all the simplicies in T containing just some subset of those vertices. The caveat is just that if the toplexes of T' are not hyperedges, then there is no hypergraph interpretation of T', as the toplexes that are not hyperedges cannot be phrased as "missing subedges". In our example above from Figure 2.1, BC was a toplex in T', but was not an edge of  $\mathcal{H}$ . The missing subcomplex S'is still generated by any vertex of T' that isn't a hyperedge, but there may be a vertex in S' whose corresponding simplex in K is not a subset of any edge in  $E' = E_1 \cap E_2$ .

We get the following theorem:

**Theorem 4.1.3** (Mayer-Vietoris Theorem - Simplicial Version). Let  $\mathcal{H} = \{V, E\}$  be a hypergraph. If  $\mathcal{H}_1 = \{V_1, E_1\}$  and  $\mathcal{H}_2 = \{V_2, E_2\}$  are subhypergraphs chosen such that

- *1.*  $E_1, E_2 \subset E$  with  $E_1 \cup E_2 = E$ , and
- 2. *if*  $e_i \subset e_j$ , and  $e_j \in E_1$  (wlog), then  $e_i \in E_1$ .

If  $T' = T_1 \cap T_2$  and  $S' = S_1 \cap S_2$ , where  $T_k$ ,  $S_k$  are the barycentric subdvision and missing subcomplex of  $\mathcal{H}_k$ , then the following is a long exact sequence using both the relative barycentric homology of hypergraphs and the relative homology of simplicial complexes:

$$\dots \to H_n(T', S') \to H_n^{rel}(\mathcal{H}_1) \oplus H_n^{rel}(\mathcal{H}_2) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(T', S') \to \dots$$

Using this simplicial version of the theorem, the homology of the hypergraph in the example above from Figure 2.1 can be computed. Recall that  $E_1 = \{ABC, B, C\}$  and  $E_2 = \{BCD, B, C\}$ . We will see in the next section that

$$H_n^{rel}(\mathcal{H}_1) \cong H_n^{rel}(\mathcal{H}_2) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & n = 1\\ 0 & \text{else.} \end{cases}$$

The associated simplicial complex of  $\mathcal{H}$  is two triangles joined together at an edge, so  $\mathcal{H}$  is not going to have any relative barycentric homology in dimensions higher than 2. Note that  $T_1$  and  $T_2$  are each barycentric subdivisions of triangles, with intersection  $T' = \{\{BC, B\}, \{BC, C\}, BC, B, C\}$ . Of these, only  $BC \in S'$ , as B and C are both edges in the hypergraph. Therefore, (T', S') is homotopic to a line segment, which is contractible and so  $H_n(T', S') = \widetilde{H_n}(T') = 0$  for all n.

Since every third term in the long exact sequence of Theorem 4.1.3 is 0, there is an isomorphism,  $H_n^{rel}(\mathcal{H}_1) \oplus H_n^{rel}(\mathcal{H}_2) \cong H_n^{rel}(\mathcal{H}).$  Therefore,

$$H_n^{rel}(\mathcal{H}) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}} & n = 1\\ 0 & \text{else.} \end{cases}$$

Applications of both Theorem 4.1.1 for hypergraphs and the simplicial version in Theorem 4.1.3 follow. If  $E_1$  and  $E_2$  can be chosen as disjoint sets (including if, for example, the hypergraph has multiple connected components), then the Mayer-Vietoris sequence from Theorem 4.1.1 reduces to  $H_n^{rel}(\mathcal{H}_1) \oplus H_n^{rel}(\mathcal{H}_2) \cong H_n^{rel}(\mathcal{H}') \forall n$ . This is displayed in the second hypergraph in Figure 2.1. Here,  $E_1 = \{ABC\}$  and  $E_2 = \{BCD\}$ . Their intersection is empty, so the homology of the hypergraph is the direct sum of their homologies.

In Figure 2.9, we computed that

$$H_n^{rel}(\mathcal{H}_1) = H_n^{rel}(\mathcal{H}_2) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{when } n = 2\\ 0 & \text{else,} \end{cases}$$

and so

$$H_n^{rel}(\mathcal{H}') = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{when } n = 2\\ 0 & \text{else.} \end{cases}$$

The second condition of Theorem 4.1.3 which says that the sets of edges chosen to partition  $\mathcal{H}$  must be closed under taking subedges in the hypergraph, means choosing  $E_1, E_2$  to be disjoint is not always possible. For example, if  $\mathcal{H}$  is a maximum edge hypergraph, then the Mayer-Vietoris sequence will not be useful, as whichever subhypergraph contains the maximum edge must contain all of the edges in  $\mathcal{H}$ .

If  $E_1$  and  $E_2$  are not disjoint, but the space (T', S') is contractible (whether or not condition 3 holds), then very nearly the same thing can be said. This was the case in our previous example from Figure 2.1, where condition 3 does not hold. After passing through to the relative barycentric subdivision level of the Mayer-Vietoris sequence, (T', S') being contractible implies that  $H_n(T', S') = 0$  when n > 0, so the isomorphism  $H_n^{rel}(\mathcal{H}_1) \oplus H_n^{rel}(\mathcal{H}_2) \cong H_n^{rel}(\mathcal{H})$  holds for all n > 1, with a still exact sequence

$$0 \to H_1^{rel}(\mathcal{H}_1) \oplus H_1^{rel}(\mathcal{H}_2) \to H_1^{rel}(\mathcal{H}) \to H_0(T', S') \to H_0^{rel}(\mathcal{H}_1) \oplus H_0^{rel}(\mathcal{H}_2) \to H_0^{rel}(\mathcal{H}) \to 0.$$

The Mayer-Vietoris sequence can reduce any hypegraph to a collection of maximum edge hypergraphs and their intersections, by letting the subhypergraphs be induced by the toplexes. It is easier to find the homology of those subhypergraphs and their intersections, and then the Mayer-Vietoris sequence will help splice them together. In fact, the next section will focus on computing the homology of hypergraphs with maximum edges.

#### 4.2 **Results on Maximum Edge Hypergraphs**

Maximum edge hypergraphs have many nice properties in the relative barycentric homology, running contrary to Theorem 3.2.3 which says the restricted barycentric homology does not distinguish any maximum edge hypergraph from any other. For the duration of this section, let  $\mathcal{H}$  be a maximum edge hypergraph as per Definition 2.1.10. Let e be the maximum edge, and |e| = m be the number of vertices in the hypergraph. In this section, we will begin by leveraging the long exact sequence of a pair, Theorem 2.2.6, as a computational tool for the relative barycentric homology. This will lead into Theorem 4.2.3 and Corollary 4.2.4 which give computational aides in high and low dimensions, respectively. Using those results, this section will conclude with several interpretability results that glean information about the hypergraph edge structure based on its relative barycentric homology, with examples along the way.

## 4.2.1 Leveraging the Long Exact Sequence of a Pair

As the relative barycentric homology of a hypergraph  $\mathcal{H}$  is a relative homology, it is a good idea to try to fit it into the long exact sequence of a pair of spaces from Theorem 2.2.6. The missing subcomplex *S* is a subcomplex of the barycentric subdivision *T*. Since the definition of the relative barycentric homology in Definition 2.3.5 is  $H_n^{rel}(\mathcal{H}) = H_n(T, S)$ , there is always a long exact sequence:

$$\ldots \to H_n(S) \to H_n(T) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S) \to \ldots$$

In the case of a maximum edge hypergraph, this sequence becomes a lot simpler, which the results at the beginning of this section show. Specifically, the first lemma of this section fully details the associated simplicial complex K of a maximum edge hypergraph  $\mathcal{H}$ .

**Lemma 4.2.1.** The associated simplicial complex K of a maximum edge hypergraph  $\mathcal{H}$  is the standard (m-1)-simplex.

*Proof.* Recall from Definition 2.3.1 that *K* is obtained from  $\mathcal{H}$  by adding all of the subedges of edges in  $\mathcal{H}$ . However, all of the subedges of a maximum edge hypergraph are the subedges of *e*, its maximum edge. The simplex associated to *e* is of dimension m - 1 since *e* has *m* vertices. The

only other simplices in K are the faces of e, so K is an (m - 1)-simplex, consisting of  $\sigma_e$  and its subsets.

Since the homology of the standard (m - 1)-simplex is well documented [18], we have the following corollary based on Lemma 2.2.3, which says that the barycentric subdivision *T* has the same homology as the associated simplicial complex *K*.

**Corollary 4.2.2.** Let  $\mathcal{H}$  be a maximum edge hypergraph, with associated simplicial complex K and barycentric subdivision T. The homology of K, and hence of T is

$$H_n(K) = H_n(T) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{when } n = 0\\ 0 & \text{else.} \end{cases}$$
(4.1)

We will substitute the preceding corollary into the long exact sequence to get the following theorem that relates the relative barycentric homology of the hypergraph  $\mathcal{H}$  to the homology of its missing subcomplex *S*. In high dimensions, this is an isomorphism.

**Theorem 4.2.3.** In high dimensions  $(n \ge 2)$ , the relative barycentric homology of a maximum edge hypergraph  $\mathcal{H}$  depends only on the (n - 1)-dimensional homology of the missing subcomplex S:  $\forall n \ge 2$ ,

$$H_n^{rel}(\mathcal{H}) \cong H_{n-1}(S).$$

*Proof.* This theorem follows from the long exact sequence:

According to Corollary 4.2.2, this part of the long exact sequence can be written with zeros for  $H_n(T)$  and  $H_{n-1}(T)$ , since  $n \ge 2$ . This makes the sequence  $0 \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S) \to 0$ , which implies, by exactness, that  $H_n^{rel}(\mathcal{H}) \cong H_{n-1}(S)$ .

Although there is not as nice of a theorem in low dimensions, substituting Corollary 4.2.2 into the long exact sequence can still be fruitful. This next corollary yields an exact sequence that will aid in studying the relative barycentric homology in dimensions 0 and 1.

**Corollary 4.2.4.** If *H* is a maximum edge hypergraph, then the following is an exact sequence:

$$\begin{array}{ccc} H_1(T) \longrightarrow H_1^{rel}(\mathcal{H}) \longrightarrow H_0(S) \longrightarrow H_0(T) \longrightarrow H_0^{rel}(\mathcal{H}) \longrightarrow 0. \\ \| & & \| \\ 0 & & \frac{\mathbb{Z}}{2\mathbb{Z}} \end{array}$$

*Proof.* The first 0 and the  $\frac{\mathbb{Z}}{2\mathbb{Z}}$  come from substituting Corollary 4.2.2 into the long exact sequence, which already terminates in a 0 after  $H_0^{rel}(\mathcal{H})$ .

Next, we will go through a few examples utilizing Theorem 4.2.3 and Corollary 4.2.4. We have already done the first example, which is of the hypergraph  $\mathcal{H} = \{ABC\}$ . A picture of  $\mathcal{H}, T$ , and S can be found in Figure 2.9. Here, S is homotopic to a 1-sphere, and so

$$H_n(S) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & n = 0, 1\\ 0 & \text{else.} \end{cases}$$

Theorem 4.2.3 then gives that  $H_2^{rel}(\mathcal{H}) \cong H_1(S) = \frac{\mathbb{Z}}{2\mathbb{Z}}$ , and Corollary 4.2.4 is now

$$\begin{array}{cccc} H_1(T) & \longrightarrow & H_1^{rel}(\mathcal{H}) & \longrightarrow & H_0(S) & \longrightarrow & H_0(T) & \longrightarrow & H_0^{rel}(\mathcal{H}) & \longrightarrow & 0. \\ \\ \| & & \| & & \| \\ 0 & & \frac{\mathbb{Z}}{2\mathbb{Z}} & & \frac{\mathbb{Z}}{2\mathbb{Z}} \end{array}$$

Since the map  $\frac{\mathbb{Z}}{2\mathbb{Z}} \to \frac{\mathbb{Z}}{2\mathbb{Z}}$  is induced by  $S \subset T$ , it is the identity map, leaving (by exactness)  $H_1^{rel}(\mathcal{H}) \cong H_0^{rel}(\mathcal{H}) = 0$ . Therefore,

$$H_n^{rel}(\mathcal{H}) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & n = 2\\ 0 & \text{else.} \end{cases}$$

Another example is given in Figure 4.1. This hypergraph is  $\mathcal{H} = \{ABC, AB, C\}$ . Its barycentric subdivision *T* is shown on the right and the missing subcomplex *S* is highlighted. Notice that *S* 



Figure 4.1 A hypergraph  $\mathcal{H}$  with computation of  $H^{rel}(\mathcal{H})$ .

is two line segments and so  $H_0(S) = \frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}}$ , and it has no homology in other dimensions. By Theorem 4.2.3,  $H_n^{rel}(\mathcal{H}) = 0$  for all  $n \ge 2$ . In this case, the exact sequence from Corollary 4.2.4 reads as follows:

$$\begin{array}{cccc} H_1(T) & \longrightarrow & H_1^{rel}(\mathcal{H}) & \longrightarrow & H_0(S) & \longrightarrow & H_0(T) & \longrightarrow & H_0^{rel}(\mathcal{H}) & \longrightarrow & 0. \\ \\ \| & & \| & & \| \\ 0 & & \frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}} & & \frac{\mathbb{Z}}{2\mathbb{Z}} \end{array}$$

Both path components of *S* (each represented by a  $\frac{\mathbb{Z}}{2\mathbb{Z}}$ ) lie in the same path component of *T*, and so the map  $\frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}} \to \frac{\mathbb{Z}}{2\mathbb{Z}}$  is surjective with a rank one kernel. By exactness then,  $H_1^{rel}(\mathcal{H}) = \frac{\mathbb{Z}}{2\mathbb{Z}}$ and  $H_0^{rel}(\mathcal{H}) = 0$ . Thus,

$$H_n^{rel}(\mathcal{H}) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & n = 1\\ 0 & \text{else.} \end{cases}$$

### 4.2.2 Interpretations

For the remainder of this section, we will give interpretative results on the relative barycentric homology. These results will allow us to utilize relative barycentric homology to glean information about the structure of the hypergraph itself, thus making these tools potentially useful for data analysis. There are interpretative results in dimensions 0, 1, m - 2 and m - 1, where m is the cardinality of the maximum edge (and hence also the vertex set). A condition is also given that

causes a hypergraph to have no relative barycentric homology in any dimension. The first result, however, classifies the relative barycentric homology of a hypergraph that only has one edge.

**Theorem 4.2.5.** Let  $\mathcal{H}$  be a hypergraph with one edge, e. Let |e| = m. Then

$$H_n^{rel}(\mathcal{H}) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{when } n = m-1 \\ 0 & \text{else.} \end{cases}$$

*Proof.* Since a one edge hypergraph is a maximum edge hypergraph, it follows from Lemma 4.2.1 that *T* is homotopic to an (m - 1)-disk. Let us construct *S*. It is the subcomplex of *T* given by all simplices whose vertices do not represent edges in the hypergraph  $\mathcal{H}$ . Since the only edge in  $\mathcal{H}$  is  $e, S = T \setminus St(v_{\sigma_e})$ , where  $St(v_{\sigma_e})$  is the star of the vertex  $v_{\sigma_e}$  in *T* (Definition 2.2.9). Since  $\sigma_e$  is the only (m - 1)-simplex in *K*, any of the (m - 1)-simplices of *T* must contain the vertex  $v_{\sigma_e}$ , and so *S* is constrained to simplices of dimension at most m - 2 that do not contain  $v_{\sigma_e}$ . Thus, as can be seen for m = 3 in Figure 2.9, *S* is the boundary of *T*. The boundary of an (m - 1)-disk is an (m - 2)-sphere, and, therefore, (T, S) is homotopic to an (m - 1)-sphere.

Recall  $H_n^{rel}(\mathcal{H}) = H_n(T, S)$  by definition, and from the reduced homology of the (m-1)-sphere [18].

$$H_n^{rel}(\mathcal{H}) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{when } n = m - 1\\ 0 & \text{else.} \end{cases}$$

For an example of this theorem, see Figure 2.9, where we computed the relative barycentric homology of  $\mathcal{H} = \{ABC\}$ , a one edge hypergraph. The next theorem says that the only way for a maximum edge hypergraph to have  $H_0^{rel}(\mathcal{H}) \neq 0$  is if that hypergraph is a simplicial complex, combining with Corollary 4.2.2 for a sufficient and necessary condition to get a nontrivial  $H_0^{rel}(\mathcal{H})$ .

**Theorem 4.2.6.** A maximum edge hypergraph  $\mathcal{H}$  has nontrivial homology in dimension zero if and only if it is a simplicial complex, i.e. there are no missing subedges and S is empty. In this case,  $H_0^{rel}(\mathcal{H}) = \frac{\mathbb{Z}}{2\mathbb{Z}}$ .

*Proof.* Suppose that there are no missing subedges in the hypergraph  $\mathcal{H}$ . This means S is empty and thus,  $H_0(S) = 0$ . Substitute that into Corollary 4.2.4 to get the following exact sequence:

$$0 \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to H_0^{rel}(\mathcal{H}) \to 0.$$

This implies that  $H_0^{rel}(\mathcal{H}) = \frac{\mathbb{Z}}{2\mathbb{Z}}$  as required.

Next suppose the hypergraph  $\mathcal{H}$  is not a simplicial complex. Then there are some missing subedges and so *S* is not empty. The map from  $H_0(S) \to \frac{\mathbb{Z}}{2\mathbb{Z}}$  from Corollary 4.2.4 is surjective then, since any representative in  $H_0(S)$  of a path component in *S* will also represent the path component of *T*, and thus be in  $H_0(T)$ . Since the generator of  $\frac{\mathbb{Z}}{2\mathbb{Z}}$  is in the image of the map  $H_0(S) \to H_0(T)$ , it is in the kernel of the map  $\frac{\mathbb{Z}}{2\mathbb{Z}} \to H_0^{rel}(\mathcal{H})$ . That map then factors through 0, leaving an exact sequence  $0 \to H_0^{rel}(\mathcal{H}) \to 0$ , proving that  $\mathcal{H}$  not simplicial implies  $H_0^{rel}(\mathcal{H}) = 0$ .

Because a maximum edge hypergraph that has a nontrivial  $H_0^{rel}$  must be a simplicial complex, Corollary 4.2.2 says that it will have no other nontrivial homology groups. Therefore, a maximum edge hypergraph cannot have both a 0-dimensional homology cycle and a homology cycle in any higher dimension. The next result heavily utilizes Corollary 4.2.4 and relates  $\beta_1^{rel}(\mathcal{H})$ , the rank of the first dimensional relative barycentric homology group, to the number of path components of the missing subcomplex *S*.

**Theorem 4.2.7.** Let  $\mathcal{H}$  be a maximum edge hypergraph with missing subcomplex S. If  $\mathcal{H}$  is a simplicial complex,  $H_1^{rel}(\mathcal{H}) = 0$ . If  $\mathcal{H}$  is not a simplicial complex, then  $\beta_1^{rel}(\mathcal{H}) = \beta_0(S) - 1 = \widetilde{\beta_0}(S)$ 

*Proof.* Suppose that the hypergraph  $\mathcal{H}$  is a simplicial complex. This means S is empty, and so  $H_0(S) = 0$ . Substitute that into Corollary 4.2.4 to get the exact sequence  $0 \to H_1^{rel}(\mathcal{H}) \to 0$ , implying that  $H_1^{rel}(\mathcal{H}) = 0$ .

Next, assume *S* is nontrivial. Then the map  $H_0(S) \to \frac{\mathbb{Z}}{2\mathbb{Z}}$  from Corollary 4.2.4 relates the path components of *S* to the single path component of *T* and is surjective. This means that the kernel has rank  $\beta_0(S) - 1$ , by rank-nullity. Since  $0 \to H_1^{rel}(\mathcal{H}) \to H_0(S)$  is exact,  $H_1^{rel}(\mathcal{H}) \to H_0(S)$ 



Figure 4.2 A hypergraph  $\mathcal{H}$  used as an example for Theorem 4.2.7.

is injective. So the rank of  $H_1^{rel}(\mathcal{H})$  is equal to the rank of the kernel of the map  $H_0(S) \to \frac{\mathbb{Z}}{2\mathbb{Z}}$ . Therefore,  $\beta_1^{rel}(\mathcal{H}) = \beta_0(S) - 1$ , proving the theorem.

We have already computed two examples of the relative barycentric homology in dimension 1. In the one edge hypergraph of Figure 2.9, *S* only has one path component, and hence,  $\beta_1^{rel}(\mathcal{H}) = 0$ . The hypergraph in Figure 4.1 has a missing subcomplex *S* with two path components, making  $\beta_1^{rel}(\mathcal{H}) = 1$ . In Figure 4.2, *S* has three path components, which leads to  $\beta_1^{rel}(\mathcal{H}) = 2$ .

Moving towards higher dimensions now, recall that the relative barycentric homology of the single edge hypergraph is concentrated in dimension m - 1 (Theorem 4.2.5). This is actually an if and only if statement, of which the next theorem proves the other half. The only way for a maximum edge hypergraph with m vertices in the maximum edge to have homology in dimension m - 1 is if the maximum edge is the only edge in the hypergraph.

**Theorem 4.2.8.** Given a maximum edge hypergraph H,

$$\beta_{m-1}^{rel}(\mathcal{H}) = \begin{cases} 1 & |\mathcal{H}| = 1 \\ 0 & else. \end{cases}$$

*Proof.* The  $|\mathcal{H}| = 1$  part of the statement follows from Theorem 4.2.5. Now let  $|\mathcal{H}| \ge 2$ , and let f be an edge that is not the maximum edge. Recall from Theorem 4.2.3 that  $\beta_{m-1}^{rel}(\mathcal{H}) = \beta_{m-2}(S)$ .

Recall from the proof of Theorem 4.2.5 that *S* is a subspace of the boundary of *T*. This boundary is an (m - 2)-sphere. However, since *f* is an edge in the hypergraph, any simplices containing the vertex representing *f* are not part of *S*. Since *S* is not the entire (m - 2)-sphere, there cannot be any (m - 2)-homology. Thus  $\beta_{m-1}^{rel}(\mathcal{H}) = \beta_{m-2}(S) = 0$ , as was to be shown.

A hypergraph can have relative barycentric homology in dimension zero if and only if it is a simplicial complex, and in dimension m-1 if and only if it is only a single edge. This way of relating the 0th dimensional homology and the (m - 1)-dimensional homology of an (m - 1)-dimensional simplicial complex is reminiscent of some of the duality theorems from algebraic topology like Poincare, Lefschetz, or Alexander Duality (see [24, 18]). The next theorem is another indicator that there may be a duality at work, because it gives a way of relating the (m - 2)-dimensional homology to the homology in dimension 1. The pairing of dimensions that add up to the dimension of the simplicial complex is a staple of the aforementioned duality theories. Recall from Theorem 4.2.7 that the relative barycentric homology in dimension 1 of a maximum edge hypergraph was related to the number of path components in the missing subcomplex. The next result relates the relative barycentric homology in dimension m - 2 to the number of path components in the subcomplex of T generated by the subedges that are present in the hypergraph, which is a complementary notion to the missing subcomplex.

**Theorem 4.2.9.** Given a maximal edge hypergraph,  $\mathcal{H}$ , with m > 3, let  $\mathcal{H}' = \mathcal{H} \setminus \{e\}$  be the hypergraph after removing the maximum edge. Recall from Definition 2.1.12 that  $\Gamma(\mathcal{H}')$  is the number of fence components of  $\mathcal{H}'$ . Then

$$\beta_{m-2}^{rel}(\mathcal{H}) = \begin{cases} 0 & \mathcal{H}' \text{ is empty,} \\ \\ \Gamma(\mathcal{H}') - 1 & else. \end{cases}$$

*Proof.* From Theorem 4.2.8,  $H_{m-2}(S) = 0$  as long as there are any present subedges in the hypergraph. If there are not, then  $\beta_{m-2}^{rel}(\mathcal{H}) = \Gamma(\mathcal{H}') = 0$  by Corollary 4.2.5, taking care of the first case. Now, we can assume that  $\mathcal{H}'$  is not empty for the remainder of the proof.



Figure 4.3 Hypergraphs used as examples for Theorem 4.2.9.

Since  $\mathcal{H}$  is a maximum edge hypergraph, its associated simplicial complex K is the standard (m-1)-simplex (Lemma 4.2.1). Recall from Theorem 4.2.3 that  $H_{m-2}^{rel}(\mathcal{H}) \cong H_{m-3}(S)$ , where S is the missing subcomplex (Definition 2.3.4). Note that, geometrically, S is entirely contained in the boundary of the barycentric subdivision T (since the only interior point represents the maximal edge), and this boundary is homeomorphic to  $\mathbb{S}^{m-2}$ , the (m-2) dimensional sphere.

In particular, an application of the long exact sequence of the pair  $(\partial T, S)$  of topological spaces leads to the following exact sequence:

Because  $\mathcal{H}'$  is not empty,  $H_{m-2}(S) = 0$ . Thus,  $\beta_{m-2}(\partial T, S) - 1 = \beta_{m-3}(S)$  by the exactness of the above sequence. Furthermore,  $\beta_{m-3}(S) = \beta_{m-2}^{rel}(\mathcal{H})$  by Theorem 4.2.3. Thus we get the following string of equalities:

$$\beta_{m-2}(\partial T, S) - 1 = \beta_{m-3}(S) = \beta_{m-2}^{rel}(\mathcal{H}).$$

So, to prove the theorem, we need to show that that  $\beta_{m-2}(\partial T, S)$  is the number of fence components of  $\mathcal{H}'$ ,  $\Gamma(\mathcal{H}')$  which was already seen (Theorem 3.2.4) to be the number of path components in the restricted barycentric subdivision R'. Recall from Definition 2.3.2 that the vertices in R' are  $v_{\sigma_{e'}}$ , where e' is a proper subedge in  $\mathcal{H}$  (and hence an edge in  $\mathcal{H}'$ ). This means R' will be entirely contained in  $\partial T$ . We prove  $\beta_{m-2}(\partial T, S) = \Gamma(\mathcal{H}') = \beta_0(R')$  by showing that each path component of R' gives rise to an (m-2)-dimensional homology class in  $H_{m-2}(\partial T, S)$ .

We start by showing the (m - 2)-faces of the star (Definition 2.2.9) of a path component of R' comprise an (m - 2)-cycle in  $C_{m-2}(\partial T, S)$ , the chain group. To see this, it needs to be shown that its boundary is contained in S. However, the boundary of the star is the link, and the link must be entirely contained in S, as shown next. If there was a vertex in the link that was not in S, then that vertex would have to represent an existing subedge from the hypergraph, but then it would be part of the path component in question, as it would be connected to the path component via the star. So, the link, and hence boundary, of a path component in R' is in S. Thus, any fence component of the hypergraph  $\mathcal{H}'$  gives rise to an (m - 2)-cycle in  $C_{m-2}(\partial T, S)$ , i.e. is in the kernel of  $C_{m-2}(\partial T, S)$ . Since there are no (m - 1)-simplices in  $\partial T$ , the image of  $C_{m-1}(\partial T, S)$  via  $\partial_{m-1}$  is empty. Thus, by the definition of homology, any fence component of the hypergraph  $\mathcal{H}'$  gives rise to an (m - 2)-cycle in  $H_{m-2}(\partial T, S)$ .

Therefore, the number of fence components of  $\mathcal{H}$  is equal to  $\beta_{m-2}(\partial T, S)$ , which is equal to  $\beta_{m-3}(S) + 1$ , which is equal to  $\beta_{m-2}^{rel}(\mathcal{H}) + 1$ , proving the theorem.

For two examples of the m = 4 case, see Figure 4.3. Each of these hypergraphs has  $\beta_2^{rel} = 1$  since they have two fence components after removing the maximum edge. Theorem 4.2.9 is useful because it is giving information about the edges that are actually present in the hypergraph, while the result in dimension 1 is giving information about the missing subedges. This complementary relationship between the results in dimension 1 and m-2 suggests a duality in the relative barycentric homology. For a maximum edge hypergraph with m > 3, combining Theorem 4.2.3 and Theorem 4.2.9 creates an easy way to study the subedges of the hypergraph by building only *S*, and taking its (m - 3)-dimensional homology, like in the proof of the latter theorem.

Table 4.1 displays  $\beta^{rel}(\mathcal{H})$  for the different dimensions and types of maximum edge hypergraphs that have been discussed thus far. In this table, *m* is the number of vertices in the hypergraph,  $\Gamma$  is the number of fence components, Supp( $\mathcal{H}$ ) is the supplement of  $\mathcal{H}$  (Definition 2.1.11), and  $\mathcal{H}'$  is as in the previous theorem. The final theorem in this section gives a condition on maximum edge

Dimension	H Simplicial	${\cal H}$ One Edge	${\cal H}$ Other
$eta_0^{rel}(\mathcal{H})$	1	0	0
$\beta_1^{rel}(\mathcal{H})$	0	0	$\Gamma(\operatorname{Supp}(\mathcal{H})) - 1$
$eta_{m-2}^{rel}(\mathcal{H})$	0	0	$\Gamma(\mathcal{H}') - 1$
$\beta^{rel}_{m-1}(\mathcal{H})$	0	1	0

Table 4.1 with  $\beta^{rel}(\mathcal{H})$  for Maximum Edge Hypergraphs, *m* is the number of vertices in the hypergraph,  $\Gamma$  is its number of fence components.

hypergraphs that would cause  $\mathcal{H}$  to have no nontrivial relative barycentric homology groups. This will happen when the missing subcomplex *S* turns out to be contractible (Definition 2.2.12).

**Theorem 4.2.10.** Let  $\mathcal{H}$  be a maximum edge hypergraph such that S is nonempty and contractible. Then  $\beta_n^{rel}(\mathcal{H}) = 0$  for all n.

*Proof.* Since S is contractible,

$$H_n(S) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} \text{ when } n = 0\\ 0 \text{ else.} \end{cases}$$

Theorem 4.2.3 gives that  $\beta_n^{rel}(\mathcal{H}) = 0$  for all  $n \ge 2$ . Now, plug in  $H_0(S) = \frac{\mathbb{Z}}{2\mathbb{Z}}$  to the exact sequence from Corollary 4.2.4 to get

$$\begin{array}{cccc} H_1(T) & \longrightarrow & H_1^{rel}(\mathcal{H}) & \longrightarrow & H_0(S) & \longrightarrow & H_0(T) & \longrightarrow & H_0^{rel}(\mathcal{H}) & \longrightarrow & 0. \\ \\ \| & & \| & & \| \\ 0 & & \frac{\mathbb{Z}}{2\mathbb{Z}} & & \frac{\mathbb{Z}}{2\mathbb{Z}} \end{array}$$

The map  $\frac{\mathbb{Z}}{2\mathbb{Z}} \to \frac{\mathbb{Z}}{2\mathbb{Z}}$  is the identity isomorphism since the path component of *S* is included into the path component of *T*. Therefore,  $H_1^{rel}(\mathcal{H}) = 0 = H_0^{rel}(\mathcal{H})$ .

Thus, for all 
$$n$$
,  $\beta_n^{rel}(\mathcal{H}) = 0$ .

Figure 4.4 gives an example of a hypergraph that has a contractible S. This hypergraph is  $\mathcal{H} = \{ABC, AB, A\}$ . In the figure, we can see that S is one connected line segment, and so it is contractible. Therefore, by Theorem 4.2.10,  $H_n^{rel}(\mathcal{H}) = 0$  for all n. Even if  $\mathcal{H}$  is not a maximum edge hypergraph, the condition of contractibility of S eases the computation of the



Figure 4.4 A hypergraph  $\mathcal{H}$  used as an example for Theorem 4.2.10.

relative barycentric homology, which is the result of Theorem 4.3.4 in the next section. That theorem, and the rest of the results therein, generalizes the results on the relative barycentric homology of maximum edge hypergraphs of this section to hypergraphs that are not necessarily maximum edge.

# 4.3 **Results on General Hypergraphs**

One of the first results we did for the restricted barycentric homology was Proposition 3.2.1, saying that the restricted barycentric homology of a hypergraph that happens to be a simplicial complex agrees with its simplicial homology. This is true for the relative baryencetric homology as well, as stated in the first result of this section. This is largely due to the fact that the missing subcomplex, *S*, will be empty in this case.

**Proposition 4.3.1.** Let  $\mathcal{H}$  be a hypergraph that is the same as K, its associated simplicial complex. Then for all n,  $H_n^{rel}(\mathcal{H}) \cong H_n(K)$ .

*Proof.* Since  $\mathcal{H} = K$  as sets, there are no missing subedges of  $\mathcal{H}$ , and S, the missing subcomplex, is empty. Recall that  $H_n^{rel}(\mathcal{H})$  was defined as  $H_n(T, S)$ . Since S is empty,  $H_n(T, S) \cong H_n(T)$ for all n. By definition of barycentric subdivision,  $H_n(T) \cong H_n(K)$  for all n. Thus for all n,  $H_n^{rel}(\mathcal{H}) \cong H_n(K)$ .

A standard result of simplicial homology is that different path components do not interact during



Figure 4.5 A hypergraph  $\mathcal{H}$  with multiple connected components.

the homology computation, and so the simplicial homology of a simplicial complex with multiple path components is the direct sum of the simplicial homology of its path components. This turns out to be true for the relative barycentric homology as well. Compare this to Corollary 3.2.5, which wrote the restricted barycentric homology as the direct sum of the homology of its fence components. For the relative barycentric homology, we will go back to the standard definition of hypergraph connected components from Definition 2.1.9.

**Proposition 4.3.2.** Let  $\mathcal{H}$  be a hypergraph, with connected components  $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_k$ . Then

$$H_n^{rel}(\mathcal{H}) = \bigoplus_{i=1}^k H_n^{rel}(\mathcal{H}_i)$$

*Proof.* Let  $T_i$  and  $S_i$  be the barycentric subdivision and missing subcomplex corresponding to each component  $\mathcal{H}_i$ . Let  $\sqcup$  be used to denote the disjoint union of two spaces. The proof is the following string of isomorphisms:

$$\bigoplus_{i=1}^{k} H_n^{rel}(\mathcal{H}_i) \cong \bigoplus_{i=1}^{k} H_n(T_i, S_i) \cong H_n(\bigsqcup_{i=1}^{k} T_i, \bigsqcup_{i=1}^{k} S_i) \cong H_n(T, S) \cong H_n^{rel}(\mathcal{H}).$$

Here the first and last isomorphisms are the definition of the relative barycentric homology from Definition 2.2.13, and the middle two isomorphisms are drawn from the fact that the result holds for simplicial homology, which can be found in [18].  $\Box$ 

Although the proof above uses properties of simplicial homology, one could also have approached that proof with the Mayer-Vietoris sequence from Theorem 4.1.1 and inducted on the number of path components, since distinct path components have empty intersections.



Figure 4.6 A non maximum edge hypergraph  $\mathcal{H}$  with contractible *K*.

The hypergraph in Figure 4.5 has two connected components. By the theorem, it would suffice to compute the relative barycentric homology of each and then the relative barycentric homology of  $\mathcal{H}$  would be the direct sum. Each component is a maximum edge hypergraph, so we could use the results of the last section to do this computation more easily than if we started with the barycentric subdivision of all of  $\mathcal{H}$  at once.

### **4.3.1** Manipulating the Long Exact Sequence of a Pair

In the last section, we manipulated the long exact sequence of homology groups of a pair of topological spaces as in Definition 2.2.6. As we noted there, in the case of the relative barycentric homology of hypergraphs, it is as follows:

$$\ldots \to H_n(S) \to H_n(T) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S) \to \ldots$$

This sequence was used in the last section in the case where  $\mathcal{H}$  was a maximum edge hypergraph, which meant T was homotopic to a disk. Since the homology of a disk is well known, this allowed us to easily study the relative barycentric homology of a maximum edge hypergraph, as in Theorem 4.2.3 and Corollary 4.2.4. Theorem 4.2.10 gave a condition where, if S is contractible and  $\mathcal{H}$  is a maximum edge hypergraph,  $H^{rel}(\mathcal{H}) = 0$  for all n. However, the long exact sequence above is useful in more cases besides just when  $\mathcal{H}$  is a maximum edge hypergraph. The next few results give more situations where the long exact sequence is useful for studying the relative barycentric homology.

Corollary 4.2.2 gave us the homology for the barycentric subdivision of a maximum edge
hypergraph, and that allowed us to utilize the long exact sequence, as in Theorem 4.2.3 and other results. Any hypergraph whose barycentric subdivision has the same homology groups as a disk, i.e. is contractible, will allow us to do the same thing. See, for example, Figure 4.6 where the hypergraph  $\mathcal{H}$  has a contractible associated simplicial complex *K* even though it is not a maximum edge hypergraph. Since the homology groups have not changed, we also have not fundamentally changed the result (see Theorem 4.2.3 and Corollary 4.2.4), just extended it to a wider class of hypergraphs. The following theorem now holds for any hypergraph whose associated simplicial complex is contractible.

**Theorem 4.3.3.** Let  $\mathcal{H}$  be a hypergraph such that its associated simplicial complex K is contractible. Then  $\forall n \geq 2$ ,

$$H_n^{rel}(\mathcal{H}) \cong H_{n-1}(S),$$

and the following is an exact sequence:

$$\begin{array}{cccc} H_1(T) & \longrightarrow & H_1^{rel}(\mathcal{H}) & \longrightarrow & H_0(S) & \longrightarrow & H_0(T) & \longrightarrow & H_0^{rel}(\mathcal{H}) & \longrightarrow & 0 \\ \\ \| & & & \| & \\ 0 & & & \frac{\mathbb{Z}}{2\mathbb{Z}} \end{array}$$

*Proof.* This result (and proof) is identical to the one in Theorem 4.2.3 and Corollary 4.2.4. Since *K* is contractible, *T* is also contractible, and so (by Definition 2.2.12) they have homology groups

$$H_n(K) \cong H_n(T) \cong \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} \text{ when } n = 0\\ 0 \text{ else.} \end{cases}$$
(4.2)

Plugging these groups into the long exact sequence of the pair (T, S) gives the result.  $\Box$ 

Even if  $\mathcal{H}$  is not a maximum edge hypergraph as in Theorem 4.2.10, using the long exact sequence when *S* is contractible can simplify the calculation of the relative barycentric homology. The case where *S* is contractible is even more enlightening than when *K* is contractible, as it was above.

**Theorem 4.3.4.** Let  $\mathcal{H}$  be a hypergraph such that its missing subcomplex, S, is contractible. Then  $\forall n \geq 1$ ,

$$H_n^{rel}(\mathcal{H}) \cong H_n(T) \cong H_n(K),$$

and

$$\beta_0^{rel}(\mathcal{H}) = \beta_0(T) - 1 = \beta_0(K) - 1.$$

*Proof.* Since *S* is contractible, its homology groups are as follows:

$$H_n(S) \cong \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} \text{ when } n = 0\\ 0 \text{ else.} \end{cases}$$
(4.3)

Plug these into the long exact sequence of a pair

$$\dots \to H_n(S) \to H_n(T) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S) \to H_{n-1}(T) \to \dots$$

to get an exact sequence

$$\begin{array}{ccc} H_n(S) & \longrightarrow & H_n(T) & \longrightarrow & H_n^{rel}(\mathcal{H}) & \longrightarrow & H_{n-1}(S) \\ \| & & \| \\ 0 & & 0 \end{array}$$

for dimensions  $n \ge 2$ . This shows that for  $n \ge 2$ ,  $H_n^{rel}(\mathcal{H}) \cong H_n(T)$ . We must consider low dimensions separately since  $H_0(S) = \frac{\mathbb{Z}}{2\mathbb{Z}}$ . The end of the exact sequence then reads

$$0 \to H_1(T) \to H_1^{rel}(\mathcal{H}) \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to H_0(T) \to H_0^{rel}(\mathcal{H}) \to 0$$

The map  $\frac{\mathbb{Z}}{2\mathbb{Z}} \to H_0(T)$  is rank one since the representative of the path component in *S* is also a representative of the path component in *T*. Thus by exactness, the map  $H_1^{rel}(\mathcal{H}) \to \frac{\mathbb{Z}}{2\mathbb{Z}}$  is the zero map. So, by exactness,  $H_1(T) \cong H_1^{rel}(\mathcal{H})$  and  $\beta_0(T) - 1 = \beta_0^{rel}(\mathcal{H})$ , as was to be shown. The last equality in each part of the theorem are given by the equivalence of the simplicial homology of *T* and *K*, from Lemma 2.2.3.

Figure 4.7 shows a hypergraph  $\mathcal{H}$  where the last theorem can be applied to compute the relative barycentric homology. The missing subcomplex *S* is highlighted; note that it is contractible. Therefore by the theorem,  $\beta_1^{rel}(\mathcal{H}) = 1$  since  $\beta_1(T) = 1$ , and  $\beta_0^{rel}(\mathcal{H}) = 0$  because  $\beta_0(T) = 1$ .



Figure 4.7 A non maximum edge hypergraph  $\mathcal{H}$  with contractible S.

It is also possible to factor long exact sequences via maps, and not just groups. One example of this is when the map  $H_n(S) \to H_n(T)$  is zero for all n > 0. Since, as topological spaces,  $S \subset T$ , the map  $H_n(S) \to H_n(T)$  could only be zero in dimension 0 if S is empty, but there is a rather large class of hypergraphs that satisfy this condition for n > 0. We will describe what this condition says about the hypergraph, its missing subedges and its associated simplicial complex. What the expression  $H_n(S) \to H_n(T) = 0$  means is that none of the (higher than 0-dimensional) homology in S survives the inclusion into T, i.e. that any homology cycle in S is a boundary in T. Figure 4.8 has an example of a hypergraph that does not satisfy this condition. The chain of 1-simplices  $\{E, CE, C, BC, B, BE\}$  is a homology generator in both S and T, so the map  $H_1(S) \to H_1(T)$  is not the zero map. Hypergraphs that do satisfy this condition can decompose their relative barycentric homology as a direct sum of the homology of T and the homology of S, in the following way.

**Theorem 4.3.5.** Suppose that  $\mathcal{H}$  is a hypergraph such that the map on homology,  $H_n(S) \to H_n(T)$ , induced by the inclusion  $S \to T$ , is the zero map for all n > 0.

Then, for all  $n \ge 2$ ,

$$H_n^{rel}(\mathcal{H}) \cong H_n(T) \bigoplus H_{n-1}(S).$$

*Proof.* Since the map  $H_n(S) \to H_n(T)$  is the zero map for all n > 0, the long exact sequence factors into short exact sequences that can be split up between the *S* and *T* step:

$$0 \to H_n(T) \to H^{rel}(\mathcal{H}) \to H_{n-1}(S) \to 0$$



Figure 4.8 A hypergraph for which the map  $H_1(S) \rightarrow H_1(K)$  is not zero.

is exact when  $n \ge 2$ . Because the homology groups in question are vector spaces over the field  $\frac{\mathbb{Z}}{2\mathbb{Z}}$ , the sequence splits, and for all  $n \ge 2$ ,

$$H_n^{rel}(\mathcal{H}) \cong H_n(T) \bigoplus H_{n-1}(S),$$

completing the proof.

Often, the sequence will split even if the coefficient group is not a field. This will be true, if, for example, the group  $H_{n-1}(S)$  is a free module. A slight modification of the hypergraph in Figure 4.8 is shown in Figure 4.9. Notice the difference is that now, E is an edge in the hypergraph, and so the map  $H_n(S) \rightarrow H_n(K)$  is zero for all  $n \neq 0$ . Even though there still is a nontrivial homology class in  $H_1(S)$ , it is homologically trivial when included into T, because it is the boundary of the 2-simplices it surrounds. Thus the thereom applies, and  $\beta_2^{rel}(\mathcal{H}) = \beta_1(S) + \beta_2(K) = 1$ .

This theorem is useful because instead of having a relative homology, we now have the direct sum of two nonrelative homologies that are easier to compute than the relative homology. Moreover, there is a large class of hypergraphs for which this decomposition holds. One type of hypergraph to which Theorem 4.3.5 always applies, and for which we would not even need to construct S to check the condition, is classified in the next theorem.



Figure 4.9 A hypergraph for which  $H_n^{rel}(\mathcal{H}) = H_n(K) \bigoplus H_{n-1}(S)$ .

**Theorem 4.3.6.** Let  $\mathcal{H}$  be a hypergraph such that its associated simplicial complex K satisfies the following condition:

For all n, for any 
$$\sum_{i=1}^{k} \sigma_i$$
 generating a class in  $H_n(K)$ ,  
 $\exists j \in \{1, 2, ..., k\}$  such that  $\sigma_j$  is a toplex in K.

Then for  $\mathcal{H}$ ,  $H_n(S) \to H_n(T)$  is 0 for all n > 0, and the previous theorem applies to  $\mathcal{H}$ .

*Proof.* Suppose  $\mathcal{H}$  is a hypergraph with *K* satisfying the condition given in the theorem statement. We will approach the proof by contradiction: suppose that for some n > 0,  $H_n(S) \to H_n(T)$  is not the zero map. Let  $\sum_{i=1}^k \sigma_i$  be in the image of the map from  $C_n(S) \to C_n(T)$  such that  $\left[\sum_{i=1}^k \sigma_i\right]$  is not homologous to 0 in  $H_n(T)$ . Since the map  $H_n(S) \to H_n(T)$  is generated by the inclusion map  $S \to T$ , the preimage of  $\sum_{i=1}^k \sigma_i$  in *T* is  $\sum_{i=1}^k \sigma_i$  in *S*. Therefore, each of the  $\sigma_i$  represents a missing subedge in  $\mathcal{H}$ . However, according to the condition given in the theorem statement, at least one of the  $\sigma_i$  is a toplex in *K*. Toplexes in *K* are exactly the same as toplexes in  $\mathcal{H}$ , by definition of the associated simplicial complex (Definition 2.3.1). Thus, for some *i*,  $\sigma_i$  is both a missing subedge of  $\mathcal{H}$  and a present toplex in  $\mathcal{H}$ . This is a contradiction. Thus,  $H_n(S) \to H_n(T) = 0$  for all n > 0, and Theorem 4.3.5 can be applied to  $\mathcal{H}$ .

At least for this section, we are finished manipulating the long exact sequence of a pair. This idea, and especially the map  $H_n(S) \rightarrow H_n(T)$  will play a major role in computing the relative barycentric homology in Section 5.2. We will use the long exact sequence to avoid computing a relative homology, instead using the linear algebra of the maps of the exact sequence for computation.

## 4.3.2 Classification of Relative Barycentric Homology for General Hypergraphs

We will continue developing the relative barycentric homology with some classification theorems. The next theorem is the analog of Theorem 3.2.4 for relative barycentric homology. It relates the zeroth dimensional relative barycentric homology to the simplicial components (Definition 2.1.16).

**Theorem 4.3.7** (Zeroth Dimensional Relative Homology<sup>†</sup>). Let  $\mathcal{H}$  be a hypergraph. Then  $\beta_0^{rel}$  equals the number of simplicial components of  $\mathcal{H}$ .

The proof will be given by induction on the number of connected components (Definition 2.1.9). The base case ( $\mathcal{H}$  is connected) is written as a separate lemma, and was already proven for maximum edge hypergraphs in Theorem 4.2.6. Note that not every connected hypergraph is maximum edge, but the proofs are similar.

**Lemma 4.3.8.** <sup>†</sup> Let  $\mathcal{H}$  be a connected hypergraph. Then  $\beta_0^{rel} \leq 1$ , and  $\beta_0^{rel} = 1$  if and only if  $\mathcal{H}$  is a simplicial complex.

*Proof.* First, note that since  $\mathcal{H}$  is connected, its associated simplicial complex K and barycentric subdivision T are also connected. Since  $H_*(\mathcal{H}) := H_*(T, S)$ , recall the tail end of the long exact exact sequence of a pair of topological spaces:

$$\dots \to H_1(T, S) \to H_0(S) \to H_0(T) \to H_0(T, S) \to 0.$$

Because *T* is connected,  $H_0(T) = \frac{\mathbb{Z}}{2\mathbb{Z}}$ . Noting that  $H_0(T, S)$  maps to 0 in the exact sequence, everything (if anything) in  $H_0(T, S)$  must be in its kernel. Thus by exactness,  $\operatorname{rk}(H_0(T, S)) \leq 1$ .

Now let's consider the exactness at the  $H_0(T)$  step. Recall that  $rk(H_0(T) = 1$ , and so if the kernel of the map  $H_0(T) \rightarrow H_0(T, S)$  is rank 1, then  $H_0(T, S) = 0$  by exactness. That will happen exactly when the map  $H_0(S) \rightarrow H_0(T)$  has rank 1.

That map is the map on homology induced by the inclusion  $S \to T$ . If *S* is not empty, then *S* includes nontrivially into *T*, and so the map on homology cannot be trivial. Thus as long as *S* is nonempty,  $H_0(S) \to H_0(T)$  has rank 1, and therefore  $H_0(T, S) = 0$ .

Recall that with respect to the hypergraph, *S* is the missing subcomplex, i.e. the subcomplex of *T* generated by the simplices in *K* that are not edges of  $\mathcal{H}$ . So *S* will be empty only when there are no missing subedges in  $\mathcal{H}$ . A hypergraph with no missing subedges is a simplicial complex, completing the proof.

*Proof of Theorem 4.3.7.* The lemma will serve as the base case for the proof of the theorem. Next, suppose that the theorem is true for hypergraphs with k - 1 connected components.

Let  $\mathcal{H}$  be a hypergraph with k connected components. WLOG, choose a connected component of  $\mathcal{H}$  and call it  $\mathcal{H}_1$  and its edge set  $E_1$ . Let the rest of the hypergraph be  $\mathcal{H}_2$  with edge set  $E_2 = E \setminus E_1$ . Since  $E_1 \cap E_2 = \emptyset$ , the Mayer-Vietoris theorem for hypergraphs (Theorem 4.1.1) applies in this situation, and every term corresponding to the intersection is 0. This yields an isomorphism

$$H_0(\mathcal{H}_1) \oplus H_0(\mathcal{H}_2) \cong H_0(\mathcal{H}).$$

By the induction hypothesis and the base case, each of the terms on the left has rank equal to its number of simplicial components. Since the map giving the isomorphism is induced by the inclusions of  $\mathcal{H}_1$  and  $\mathcal{H}_2$  into  $\mathcal{H}$ , and will be the identity on homology generators (which are representatives of those simplicial components),  $\beta_0(\mathcal{H}) = \beta_0(\mathcal{H}_1) + \beta_0(\mathcal{H}_2)$  will be the number of simplicial components, finishing the proof.

As an example, consider again the hypergraph in Figure 4.5. This hypergraph has two connected components. The component with the edge *ABC* is not simplicial, but the component with the edge *DEF* is simplicial. Therefore  $\beta_0^{rel}(\mathcal{H}) = 1$ . Throughout the next theorem, we will use  $H^{rel}(e_i)$  to denote the relative barycentric homology of the hypergraph with only one edge,  $e_i$ , which was given in Theorem 4.2.5. The next result classifies the relative barycentric homology for reduced

hypergraphs (Definition 2.1.5). This result compares to Theorem 3.2.2 for the restricted barycentric homology.

**Theorem 4.3.9** (Relative Homology of Reduced Hypergraphs<sup>†</sup>). Let  $\mathcal{H}$  be a reduced hypergraph, with edge set  $E = \{e_1, ..., e_m\}$ . Let  $E^k$  denote the set of edges with exactly k vertices, i.e.  $E^k := \{e \in E \mid |e| = k\}.$ Then  $H^{rel}(\mathcal{H}) = \bigoplus_{i=1}^m H^{rel}(e_i)$ , and  $\beta_{k-1}^{rel}(\mathcal{H}) = |E^k|.$ 

*Proof.* The proof is by induction on the number of edges in the hypergraph.

Base Case: Suppose m = 1, then since  $e_1$  is the only edge in the hypergraph,  $H^{rel}(\mathcal{H}) = H^{rel}(e_1)$ . Let  $|e_1| = k$ . Recall the homology of the single edge hypergraph (Theorem 4.2.5):

$$H_n^{rel}(\mathcal{H}) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} \text{ when } n = k - 1\\ 0 \text{ else.} \end{cases}$$

Since  $|E^k| = 1$ , and  $|E^j| = 0$  for  $j \neq k$ , the base case is proved.

Induction Hypothesis: Assume that the theorem holds for a reduced hypergraph with m - 1 edges.

Induction Step: Let  $\mathcal{H}$  be a reduced hypergraph with *m* edges. Let *K* be the associated simplicial complex of  $\mathcal{H}$  and *T* its barycentric subdivision. Let *S* denote the missing subcomplex.

Partition E into  $E_1$  and  $E_2$  with  $E_1 = \{e_1\}$  and  $E_2 = E \setminus \{e_1\}$ . Let  $\mathcal{H}_1$  and  $\mathcal{H}_2$  be the hypergraphs with edge sets  $E_1$  and  $E_2$ , respectively. Let  $K_1, K_2, T_1$ , and  $T_2$  be the corresponding associated simplicial complexes and barycentric subdivisions. Let  $S_1$  and  $S_2$  be the corresponding missing subcomplexes.

Then recall the simplicial version of the Mayer-Vietoris sequence (Theorem 4.1.3). Let  $T' = T_1 \cap T_2$  and  $S' = S_1 \cap S_2$ . Then there is an exact sequence:

$$\dots \to H_n(T', S') \to H_n(T_1, S_1) \oplus H_n(T_2, S_2) \to H_n(T, S) \to H_{n-1}(T', S') \to \dots$$

Apply the definition of the relative barycentric homology to see that  $H_n(T, S) = H_n^{rel}(\mathcal{H})$ ,  $H_n(T_1, S_1) = H_n^{rel}(\mathcal{H}_1) = H_n^{rel}(e_1)$  by the base case, and  $H_n(T_2, S_2) = H_n^{rel}(\mathcal{H}_2) = \bigoplus_{i=2}^m H_n^{rel}(e_i)$  by the induction hypothesis.

Next, we will show  $H_n(T', S') = 0$  for all n. This is done by showing that S' = T'. By definition,  $S' \subset T'$ , since  $S_1 \subset T_1$ , and  $S_2 \subset T_2$ .

Suppose  $\Omega \in T'$ . Then,  $\Omega \in T_1$  and  $T_2$ . Recall that simplices in the barycentric subdivision represent chains of inclusions of simplices in the original simplicial complex, so  $\Omega = \{v_{\sigma_0} \subset v_{\sigma_1} \subset \dots \subset v_{\sigma_r}\}$  for  $r = \dim \Omega + 1$  and every  $\sigma_i$  is a subset of an edge in  $E_1$  and and edge in  $E_2$ .

Remember that  $E_1 = \{e_1\}$  and  $e_1 \notin E_2$ . Since the hypergraph is reduced, any subset of  $e_1$  that is also a subset of an edge in  $E_2$  cannot be present as an edge in the hypergraph, as it would be a proper subedge. Therefore,  $\Omega$  is built on the vertices  $\sigma$  that are all missing subedges of the hypergraph, and thus  $\Omega \in S'$ .

Since  $\Omega$  was an arbitrary simplex in  $T', T' \subset S'$ , and thus, T' = S'. Therefore  $H_n(T', S') = 0$  for all *n*. We get the following exact sequence:

which yields the following isomorphism of homology:

$$H_n^{rel}(e_1) \oplus (\bigoplus_{i=2}^m H_n^{rel}(e_i)) \cong H_n^{rel}(\mathcal{H}).$$

This proves the first part of the theorem. The second part of the theorem about  $\beta^{rel}$  follows from the base case and induction hypothesis, with  $\beta_n^{rel}(\mathcal{H}) = \beta_n^{rel}(\mathcal{H}_2)$  for all  $n \neq |e_1|$ , and  $\beta_{|e_1|}^{rel}(\mathcal{H}) = \beta_{|e_1|}^{rel}(\mathcal{H}_2) + 1.$ 

For this theorem, it does not matter whether the reduced hypergraph has a high degree of intersection, or is made up of many connected components. It only considers that there are no subedges present in the hypergraph. This seems to imply that the relative barycentric homology is giving information primarily about the subedge structure of the hypergraph, and not the structure



Figure 4.10 Three different hypergraphs with the same toplexes.

and intersectionality of its toplexes. For example, the three hypergraphs in Figure 4.10 all have the same relative barycentric homology.

Recall the definition of a simplicial collapse from Theorem 2.2.7. Simplicial collapses make computing the homology of a simplicial complex easier by shrinking the complex without changing its homology. Unfortunately, not every allowable collapse can be applied to the barycentric subdivision T of a hypergraph  $\mathcal{H}$  without changing its relative barycentric homology. In the barycentric subdivision T of Figure 4.4, the edge {AC, C} can be simplicially collapsed with the triangle {ABC, AC, C}. However, this would split the missing subcomplex S, highlighted in the image, into two path components, thereby changing  $H_1^{rel}(\mathcal{H})$ . This theorem gives a condition on the simplices of the barycentric subdivision T allowing collapses depending on whether or not the collapsed simplices lie in the missing subcomplex S.

**Proposition 4.3.10.** Let  $\mathcal{H}$  be a hypergraph, with barycentric subdivision T and missing subcomplex S. Let  $\Phi < \Omega$  be two simplices in T that meet the requirements for a simplicial collapse. Suppose either

- 1. All of the vertices of  $\Phi$  and  $\Omega$  are in S, or
- 2. All of the vertices of  $\Phi$  and  $\Omega$  are not in S.

Then the collapse can be made at the simplicial level without changing the relative barycentric homology of the hypergraph.

*Proof.* Recall the definition of the relative barycentric homology as  $H_n^{rel}(\mathcal{H}) = H_n(T, S)$ . Let T' denote  $T \setminus \{\Phi, \Omega\}$  and S' denote the analogous construction for S. In the second case, S' = S.

Since simplicial collapses preserve homology,  $H_n(T) \cong H_n(T')$  for all n, and  $H_n(S) \cong H_n(S')$ . Using the long exact sequences of the pairs (T, S) and (T', S') and maps on homology generated by inclusions, we get the following commutative diagram:

The four outside vertical maps are all isomorphisms, thus the middle map is an isomorphism as well. Therefore, for all n,  $H_n^{rel}(\mathcal{H}) \cong H_n(T, S) \cong H_n(T', S')$ , proving the theorem.  $\Box$ 

Initial collapses can only use case 2 of not in *S*, but subsequent collapses might use the in *S* case. An example of a hypergraph  $\mathcal{H}$  and corresponding barycentric subdivision where Proposition 4.3.10 does apply can be found in Figure 4.1. The edge {*AC*, *C*} can be collapsed with the triangle {*ABC*, *AC*, *C*}. After that collapse, {*ABC*, *AC*, *A*} is the only maximal simplex that {*ABC*, *AC*} is a face of, and so they can be collapsed. Then, finally, the only maximal simplex that *AC* is a face of is {*AC*, *A*}, and they are both simplices in *S*. The pair (*T'*, *S'*) after these collapses has the same relative homology.

In this section, we gave some results helping to compute and interpret the relative barycentric homology for hypergraphs. Several of these results, like Theorem 4.3.3 and Theorem 4.3.5, leveraged the ideas of the long exact sequence of a pair. Because of the exactness of the sequence

$$\dots \to H_n(S) \to H_n(T) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S) \to H_{n-1}(T) \to \dots$$

every homology class in  $H_n^{rel}(\mathcal{H})$  either comes from  $H_n(T)$  or maps into  $H_{n-1}(S)$ . This concept will form the backbone for computing the relative barycentric homology at the end of the next chapter in Section 5.2. First, however, we will state some results that help to form a bridge between the restricted and relative barycentric homology theories, allowing us to connect the two.

#### **CHAPTER 5**

#### **BRIDGING THE THEORIES AND COMPUTATION**

This chapter has two main objectives, which both serve the purpose of uniting the relative and restricted barycentric homology theories. The second goal will be to discuss the computation of the relative and restricted barycentric homology. During my internship with PNNL, I wrote code in Python for computing the Betti numbers for both the restricted barycentric homology and the relative barycentric homology. That code, and an accompanying tutorial, will soon be avilable on HyperNetX, a PNNL developed Python package for studying hypergraphs [27]. In this thesis, we will not spend much time discussing the code and its efficiency. Instead, in Section 5.2, we will look to show the theory behind why the algorithm returns the correct answer. This is necessary because the computations are not done using the definitions given in Section 2.3. In particular, a main goal when writing the code was to avoid building the entire barycentric subdivision since this is computationally expensive, and we will discuss how this goal was accomplished. Before we discuss computation however, we will give some theoretical results merging the two homology theories. These are mainly set-theoretic results on the supplement and complement of hypergraphs and how the definition of supplement from Definition 2.1.11 interacts with the definition of the missing subcomplex from Definition 2.3.4. Recall Lemma 2.1.2 which says that for every hypergraph, either it is a maximum edge hypergraph or its complement is maximum edge (sometimes both). Since the relative barycentric homology of maximum edge hypergraphs is easier to study, benefits of the results in this section include ways of relating the homology of a complement or supplement hypergraph to the homology of the original hypergraph.

## 5.1 Results Bridging the Relative and Restricted Barycentric Homologies

For this section, recall the definition of the complement (Definition 2.1.7) and supplement (Definition 2.1.11) of a hypergraph. Briefly, given a hypergraph, its complement is all of the subsets of the vertex set that are not edges in the hypergraph, while the supplement is all subsets of existing edges that are not themselves edges. The first lemma says that given a hypergraph, its missing subcomplex (Definition 2.3.4) is the restricted barycentric subdivision (Definition 2.3.2)

of its supplement.

**Lemma 5.1.1.** <sup>†</sup> Let  $\mathcal{H}$  be a hypergraph and Supp( $\mathcal{H}$ ) be its supplement. The restricted barycentric subdivison R of Supp( $\mathcal{H}$ ) is the same as the missing subcomplex S of  $\mathcal{H}$ .

*Proof.* Let *R* be the restricted barycentric subdivision of  $\text{Supp}(\mathcal{H})$ . Then the vertices in *R* are the edges of  $\text{Supp}(\mathcal{H})$  and the simplices of *R* are given by their inclusion relations in  $\text{Supp}(\mathcal{H})$ . By definition of the supplement, the edges in  $\text{Supp}(\mathcal{H})$  are the missing subedges of edges in  $\mathcal{H}$ . Let *S* be the missing subcomplex of the relative barycentric subdivision of  $\mathcal{H}$ . The vertices of *S* are the missing subedges of edges in  $\mathcal{H}$ , with simplices given by the chains of inclusions in the associated simplicial complex  $\mathcal{H}$ . Thus, these simplicial complexes *R* and *S* have the same vertex set with simplices constructed in the same way, and are therefore the same simplicial complex.

This means that given a hypergraph  $\mathcal{H}$ , we can alter the long exact sequence for the relative barycentric homology

$$\ldots \to H_n(S) \to H_n(T) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S) \to \ldots$$

by replacing  $H_n(S)$  at all steps with  $H^{res}$ Supp $(\mathcal{H})$ , as in the following theorem, which is based off of Theorem 4.2.3 and Corollary 4.2.4 and holds when  $\mathcal{H}$  is a maximum edge hypergraph.

**Theorem 5.1.2.** <sup> $\dagger$ </sup> Let  $\mathcal{H}$  be a maximum edge hypergraph, and  $\mathcal{G}$  be its supplement. Then

$$H_{n+1}^{rel}(\mathcal{H}) \cong H_n^{res}(\mathcal{G})$$

when n > 1, and there is a long exact sequence

$$0 \to H_1^{rel}(\mathcal{H}) \to H_0^{res}(\mathcal{G}) \to \frac{\mathbb{Z}}{2\mathbb{Z}} \to H_0^{rel}(\mathcal{H}) \to 0.$$

*Proof.* Since  $\mathcal{H}$  is a maximum edge hypergraph, we have the following results for its relative homology from Theorem 4.2.3 and Corollary 4.2.4.

$$H_{n+1}^{rel}(\mathcal{H}) \cong H_n(S)$$



Figure 5.1 A hypergraph  $\mathcal{H}$  used as an example for Theorem 5.1.2.

when n > 1 and an exact sequence

$$0 \to H_1^{rel}(\mathcal{H}) \to H_0(S) \to \mathbb{Z} \to H_0^{rel}(\mathcal{H}) \to 0.$$

However, by the lemma, we can substitute *R* in for *S*. Moreover, since *R* is the restricted barycentric subdivision of  $\text{Supp}(\mathcal{G})$ ,  $H_n(R) \cong H_n^{res}(\text{Supp}(\mathcal{H})) \forall n$ , so in the above expressions,  $H_n(S)$  can be replaced by  $H_n^{res}(\text{Supp}(\mathcal{G}))$ . This proves the theorem.

The prior theorem relates the relative barycentric homology of a maximum edge hypergraph to the restricted barycentric homology of its supplement. Figure 5.1 shows a hypergraph  $\mathcal{H}$  and its supplement Supp( $\mathcal{H}$ ). Notice that the missing subcomplex of  $\mathcal{H}$ , highlighted in T, is the same as the restricted barycentric subdivision of Supp( $\mathcal{H}$ ). The following corollary, which uses Corollary 2.1.3 to rephrase Theorem 5.1.2, relates the restricted barycentric homology of a hypergraph that is not maximum edge to the relative barycentric homology of its complement.

**Corollary 5.1.3.** <sup>†</sup> Let  $\mathcal{H}$  be a hypergraph that is not maximum edge, and let  $Comp(\mathcal{H})$  be its complement. Then

$$H_{n+1}^{rel}(Comp(\mathcal{H})) \cong H_n^{res}(\mathcal{H})$$

when n > 1, and there is a long exact sequence

$$0 \to H_1^{rel}(Comp(\mathcal{H})) \to H_0^{res}(\mathcal{H}) \to \mathbb{Z} \to H_0^{rel}(Comp(\mathcal{H})) \to 0.$$

*Proof.* This corollary is a result of Corollary 2.1.3, which says that  $\mathcal{H}$  is the supplement of Comp( $\mathcal{H}$ ), and so the previous theorem can be applied to yield the corollary.



Figure 5.2 A hypergraph  $\mathcal{H}$  and its missing subcomplex and restricted barycentric subdivision.

This sheds some light on why it makes sense to talk about the restricted and relative barycentric homologies as complementary theories. The restricted barycentric subdivision (Definition 2.3.2) is built from the barycentric subdivision using the vertices that represent edges in the hypergraph. The missing subcomplex of Definition 2.3.4 is built from the barycentric subdivision using the vertices that do not represent edges in the hypergraph. A vertex in the barycentric subdivision either represents an edge in the hypergraph or it does not. Therefore, the restricted barycentric subdivision. However, the same cannot be said for the higher dimensional simplices. Simplices that have some vertices representing edges in the hypergraph and others that do not will not be in either the restricted barycentric subdivision or the missing subcomplex. Figure 5.2 illustrates this interesting concept. The missing subcomplex *S* is highlighted in red and the restricted barycentric subdivision *R* is highlighted in blue. These partition the vertex set of the barycentric subdivision *R* is highlighted in blue.

In the next section, which discusses computation, the connection between the restricted and relative barycentric homologies will become even more apparent. During computation, there are many parallels between the theories, and some of the shortcuts that can be taken to reduce computation complexity work for both versions of the homology, as we will see next.

#### 5.2 Computation of the Restricted and Relative Barycentric Homologies

This section<sup>†</sup> contains a discussion of the computation of the restricted and relative barycentric homology theories. There are two main success stories of this section. First, for both the restricted

and relative barycentric homology, the computation is made without taking the barycentric subdivision. Avoiding this computationally unwieldy construction saves both space and time. The second contribution, exclusively pertaining to the relative barycentric homology, is that the computation happens without needing to take the quotient space or a relative homology at all. This is shown in Theorem 5.2.3. The code, and an accompanying tutorial, which was written during my internship with PNNL, will soon be included in the python package HyperNetX [27]. Finding the associated simplicial complex of a hypergraph was already implemented in HyperNetX, and taking the homology of a simplicial complex is done using the standard Smith Normal Form [13], so my contribution besides the two successes outlined above is building the specific simplicial complexes needed for the computations. We will begin with the computation of the restricted barycentric homology.

#### 5.2.1 Computation of Restricted Barycentric Homology

The restricted barycentric subdivision is built in the code as follows: the edges of the hypergraph are used as vertices of a graph. Each pair of edges is checked for an inclusion, and an edge is added to the graph if there is an inclusion relation. The restricted barycentric subdivision is the clique complex of this graph, as demonstrated in Theorem 3.1.2, and the chain complex of that is what is used in the computation. The main advantage of this approach is that it does not require taking all subsets of the edges, as is necessary for finding the full barycentric subdivision. There is still a bottleneck in having to take all maximal cliques of a graph.

This construction follows closely the understanding of the restricted barycentric subdivision as the order complex of the edge containment poset, as in Proposition 3.1.1. Each element of the poset is an edge in the hypergraph, and the relations in the poset are inclusions of edges. After building the restricted barycentric subdivision, one builds the chain complex, boundary matrices, and takes the homology as implemented in HyperNetX.

Before we move on to the relative barycentric homology, we will follow along with the algorithm to compute the homology of the hypergraph in Figure 2.8. This hypergraph has four edges:  $\mathcal{H} = \{ABC, BCD, B, C\}$ . First, the algorithm will build a graph with those as its vertices. Then, *B* will be connected to *ABC* and *BCD* because it is a subset of them, but *B* will not be connected to *C*, because there is not an inclusion relation in either direction between *B* and *C*. Next, *C* will be connected via an edge to *ABC* and *BCD* since it is a subset of both of them. Then *ABC* and *BCD* will not be connected. This gives the restricted barycentric subdivision *R*, as highlighted in the figure, but without having to construct *T*. HyperNetX will build the chain complex to *R* and take its homology using Smith Normal Form, returning [1, 1] to indicate that  $\beta_0^{res}(\mathcal{H}) = 1$  and  $\beta_1^{res}(\mathcal{H}) = 1$ . Since there are no simplices in higher dimension, it does not return any higher Betti numbers, which are all zero.

## 5.2.2 Computation of Relative Barycentric Homology

The first step in taking the relative barycentric homology is to use the native HyperNetX functions to find the chain complex, boundary matrices, and homology of the associated simplicial complex K. In order to build the missing subcomplex, it is necessary to go through almost the same process as it takes to build the barycentric subdivision. It still starts by taking all subsets of the edges of the hypergraph. Instead of using every subset like in the barycentric subdivision, it only stores those subsets that are not edges of  $\mathcal{H}$ . These form the vertices of the missing subcomplex. Again the full missing subcomplex S is the clique complex of a graph built on these vertices based on inclusions in subsets of the hypergraph, as shown in the proposition below.

**Proposition 5.2.1** (Poset Construction of S). The missing subcomplex S of a hypergraph  $\mathcal{H}$  is the same simplicial complex as the order complex (Definition 2.2.8) of the face poset (Definition 2.1.14) of Supp( $\mathcal{H}$ ),  $S_{\mathcal{H}} = \Delta(FP_{Supp(\mathcal{H})})$ .

*Proof.* This result is a combination of Lemma 5.1.1 and Proposition 3.1.1. Lemma 5.1.1 says that *S* is the same simplicial complex as the restricted barycentric subdivision *R* of Supp( $\mathcal{H}$ ), and Proposition 3.1.1 says that *R* is the same simplicial complex as the mentioned poset construction.  $\Box$ 

This is analogous to Proposition 3.1.1 for the restricted barycentric subdivision. These two constructions being so closely connected lends credence to the idea that the restricted and relative

barycentric homology are complementary theories. The next theorem, which is how the missing subcomplex is built in the code, is a clique complex construction of *S* similar to Theorem 3.1.2.

**Theorem 5.2.2** (Clique Construction of S). The missing subcomplex S of a hypergraph  $\mathcal{H}$  is the same simplicial complex as the clique complex (Definition 2.2.7) of the edge inclusion graph (Definition 2.1.15) of its supplement Supp( $\mathcal{H}$ ),  $S_{\mathcal{H}} = X(EIG_{Supp(\mathcal{H})})$ .

*Proof.* Similarly to the above Proposition, utilizing Lemma 5.1.1 gives that *S* is the same simplicial complex as  $R_{\text{Supp}\mathcal{H}}$ , and then applying Theorem 3.1.2 constructs  $R_{\text{Supp}(\mathcal{H})}$  as stated.

Once the missing subcomplex *S* is constructed, we use the built in HyperNetX functions for generating its chain complex for use in the computation. Recall our long exact sequence for relative barycentric homology:

$$\dots \to H_n(S) \to H_n(T) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S) \to H_{n-1}(T) \dots$$

In this sequence, the map  $H_n(S) \to H_n(T)$  is the map induced on homology by the inclusion  $S \to T$ . The map  $H_n(T) \to H_n^{rel}(\mathcal{H})$  is the map induced on homology by the quotient map  $T \to (T, S)$ , and the map  $H_n^{rel}(\mathcal{H}) \to H_{n-1}(S)$  is derived using the snake lemma. Because of this exactness, the terms and maps listed above are enough to determine  $H_n^{rel}(\mathcal{H})$  without actually computing the relative homology  $H_n(T, S)$ .

We will next leverage another important fact from algebraic topology, that of Lemma 2.2.3:  $H_n(T) \cong H_n(K)$  for all *n*, and they are furthermore homotopy equivalent. So anywhere in the above discussion, it is possible to replace H(T) with H(K). (Computationally, this will not affect Betti numbers, but will affect generators.) We will use  $i_n$  to denote the composition  $H_n(S) \to H_n(T) \to H_n(K)$ , where the second map is a homotopy equivalence. This leads to an exact sequence:

$$\ldots \to H_n(S) \xrightarrow{i_n} H_n(K) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S) \xrightarrow{i_{n-1}} H_{n-1}(K) \ldots$$

There are large classes of hypergraphs for which  $i_n$  is guaranteed to be the zero map (Theorem 4.3.6) and it will always be possible to write the above sequence as a direct sum decomposition.

In order for a homology in *S* to also be in *K*, *K* must contain a homology class that is entirely generated by simplices that are missing subedges in  $\mathcal{H}$ , as seen in Figure 4.8. This means that *K* needs a homology class that is generated entirely by simplices that are not toplexes. Therefore, any hypergraph whose associated simplicial complex does not have such a homology can have its relative barycentric homology written as  $H^{rel}(\mathcal{H}) \cong H_n(K) \oplus H_{n-1}(S)$  by Theorem 4.3.5.

This leads to the best, so far, physical interpretation of the relative barycentric homology. Even if the hypergraph does not lie in that special case where  $H^{rel}(\mathcal{H}) \cong H_n(K) \oplus H_{n-1}(S)$ , all relative barycentric homologies still come from either K or S (because of the exactness of the sequence). So the relative homology can be separated as either homology native to the associated simplicial complex of that hypergraph or homology coming from the poset structure of the missing subedges of the hypergraph. The latter piece (those homologies coming from S) can also be viewed as the restricted barycentric homology of the supplement hypergraph. In fact, even if  $i_n : H_n(S) \to H_n(K)$ is not 0 for all n, that map is all of the information we need to find  $H^{rel}(\mathcal{H})$ . This is summarized in the next theorem.

**Theorem 5.2.3** (Computing Relative Barycentric Homology). Let *H* be a hypergraph. Recall the long exact sequence of homology groups:

$$\dots \to H_n(S) \xrightarrow{i_n} H_n(K) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S) \xrightarrow{i_{n-1}} H_{n-1}(K) \dots$$

Then  $H_n^{rel}(\mathcal{H})$  is determined entirely by the maps  $i_n : H_n(S) \to H_n(K)$  and  $i_{n-1} : H_{n-1}(S) \to H_{n-1}(K)$ . In fact,

$$H_n^{rel}(\mathcal{H}) = H_n(T, S) \cong Cok(i_n) \oplus Ker(i_{n-1}), and \beta_n^{rel}(\mathcal{H}) = \beta_n(K) + \beta_{n-1}(S) - rk(i_n) - rk(i_{n-1}).$$

*Proof.* This is again because of the exactness of the sequence

$$\dots \to H_n(S)\overrightarrow{i_n}H_n(K) \to H_n^{rel}(\mathcal{H}) \to H_{n-1}(S)\overrightarrow{I_{n-1}}H_{n-1}(K) \to \dots$$

The composition of the middle two maps is zero since the sequence is exact. This means that all homology generators of  $H_n^{rel}(\mathcal{H})$  either come from  $H_n(K)$  or map into  $H_{n-1}(S)$ , but not both.

First, we consider the generators that come from  $H_n(K)$ . Since they are in the image of the map  $H_n(K) \to H_n^{rel}(\mathcal{H})$ , they were not in the image of  $i_n : H_n(S) \to H_n(K)$  by exactness. Thus, they must be exactly the generators of the cokernel of  $i_n$ , denoted  $\operatorname{Cok}(i_n)$ . Recall that  $\operatorname{rk}(\operatorname{Cok}(i_n)) = \beta_n(K) - \operatorname{rk}(i_n)$ .

Then there are the generators of  $H_n^{rel}(\mathcal{H})$  that map into  $H_{n-1}(S)$ . Since these are the image of that map, they become the kernel of  $i_{n-1} : H_{n-1}(S) \to H_{n-1}(K)$ , denoted Ker $(i_{n-1})$ . Recall that  $\operatorname{rk}(\operatorname{Ker}(i_{n-1})) = \beta_n(S) - \operatorname{rk}(i_{n-1})$ .

As stated earlier, because of the exactness of the sequence, all homology generators of  $H_n^{rel}(\mathcal{H})$ either come from  $H_n(K)$  or map into  $H_{n-1}(S)$ , but not both. The generators that come from  $H_n(K)$ are exactly  $\operatorname{Cok}(i_n)$ , and the generators that map into  $H_{n-1}(S)$  are exactly  $\operatorname{Ker}(i_{n-1})$ , so

$$H_n^{rel}(\mathcal{H}) \cong \operatorname{Cok}(i_n) \oplus \operatorname{Ker}(i_{n-1}).$$

This means that

$$\beta_n^{rel}(\mathcal{H}) = \operatorname{rk}(\operatorname{Cok}(i_n)) + \operatorname{rk}(\operatorname{Ker}(i_{n-1})) = \beta_n(K) - \operatorname{rk}(i_n) + \beta_n(S) - \operatorname{rk}(i_{n-1}),$$

as was to be shown.

Note that if  $i_n$  is the zero map, then its kernel is everything in  $H_n(S)$  and its cokernel is everything in  $H_n(K)$ , yielding the direct sum decomposition  $H^{rel}(\mathcal{H}) \cong H_n(K) \oplus H_{n-1}(S)$  presented in Theorem 4.3.5 and recalled above.

Once the missing subcomplex *S* is built, we can find the homology of *S* using HyperNetX. Then, it is necessary to construct the matrix for the map  $S \to K$ . The inclusion from  $S \to T$  is the identity on each simplex in *S*, but we need to describe the map  $T \to K$ . This map is induced on simplices of *T* by a vertex map. Each vertex in *T* represents a simplex in *K*, so if  $v_{\sigma}$  is a vertex in *S* representing  $\{v_0, ..., v_n\} \in K$ , the map  $S \to K$  takes  $v_{\sigma}$  to  $v_0$ . It always chooses the lexicographically first vertex contained in that simplex in *K*.

The algorithm computes the matrix (in each dimension) for this map by looking at every simplex in *S*, writing down all of the first entries in each of the vertices comprising that simplex, and then

checking the simplices in K to look for a match. If there is a match, the corresponding matrix element is made a 1.

After computing that matrix, each of the homology generators of *S* are multiplied by it to put them in terms of the chain groups for *K*. Cycles in *S* are automatically cycles in *K*. So, using the same techniques as the homology algorithm with Smith Normal Form uses, they are projected into the cokernel to see if they are also homology generators in *K*. Homologies in *S* that are also homologies in *K* result in obstructions to the direct sum decomposition, and the number of these obstructions in each dimension is tracked in a vector called Bettiob. Lastly, the relative Betti numbers are computed, using the following formula:  $\beta_n^{rel}(\mathcal{H}) = \beta_n(K) + \beta_{n-1}(S) - Bettiob_n - Bettiob_{n-1}$ .

To close this section, we will follow along with the algorithm to compute the relative barycentric homology of the hypergraph in Figure 2.9. This is the hypergraph  $\mathcal{H} = \{ABC\}$ . First, HyperNetX takes the chain complex of the associated simplicial complex K and, using Smith Normal Form, its homology. It records this as [1,0,0], denoting that  $\beta_0(K) = 1$  and  $\beta_i(K) = 0$  for all  $i \neq i$ 0. Next, we have to build S. The algorithm notes that, of the subsets of ABC, the collection {AB, AC, BC, A, B, C} is missing from the hypergraph, and so it forms the vertex set of S. Edges are made between A and AB, and A and AC, as well as the rest of the pairs that have a subset relation, giving the graph highlighted in the figure, which is S. HyperNetX takes its chain complex and homology, returning [1, 1] as the sequence of Betti numbers for S. The map  $S \to K$  sends A, AB, AC to A, B and BC to B, and C to C. The algorithm builds the corresponding matrix and then the homology generators of S are multiplied by that map to see if they still generate nontrivial homology in K. In this case, the class in dimension zero maps nontrivially (since the only path component of S is also the path component of K), but the class in dimension one does not as there is no homology in dimension one in K. So, the code returns [1,0] to denote the rank of the map induced on homology by the map  $S \rightarrow K$ . These vectors are added according to the formula to get a final answer for this hypergraph of [0, 0, 1], as expected.

#### **CHAPTER 6**

#### **DYNAMIC HYPERGRAPHS**

If the relative and restricted barycentric homology of hypergraphs are going to be useful for studying real world data, it will be necessary to quantify how changing a hypergraph changes the homology, as real world data is constantly being modified. Our methods need to be able to account for and analyze changes in the hypergraph. This chapter takes the first steps towards doing just that. There are three ways in which a hypergraph  $\mathcal{H}$  can change: the vertex set V could change, the edge set E could change, or the incidence of the vertices and edges can change. However, the third option of changing vertex-edge relationships is covered by the second case of changing the edge set. Changing vertex-edge incidence is done by removing all of the edges that were not a part of the original edge set. Therefore, all dynamic hypergraphs can be traced by keeping track of the vertices and edges being added or removed.

Moreover, because of the invertability of the operation of adding or removing vertices and edges, it is enough to only consider the consequences of adding. The upcoming results in this chapter are written for adding vertices and edges, but they can also be read to have the opposite effect for removing a vertex or edge with the same conditions. For example, in Theorem 6.2.1, conditions are given for which adding a vertex to a hypergraph increases the dimension of all of the homology classes. It should be understood that this also implies that removing a vertex under the same conditions decreases the dimension of all the homology cycles.

This chapter consists of two sections. The first gives results on the effects on the restricted and relative barycentric homology when adding an edge to a hypergraph. It is not assumed that the vertex set stays the same when adding edges, so new edges can consist partially or entirely of new vertices. These results are concentrated in dimension zero. The second section contains results on the effects of adding a vertex to a hypergraph, while keeping the edge set the same. The main result of that section, Theorem 6.2.1, is for the relative barycentric homology and applies for maximum edge hypergraphs in all dimensions.

#### 6.1 Adding an Edge

Throughout this section, let  $\mathcal{H} := \{V, E\}$  be a hypergraph, with  $V = \{v_1, ..., v_n\}$  and  $E = \{e_1, ..., e_m \mid e_i \subset V\}$ . The edge added to the hypergraph will be denoted  $e' \notin E$ . Let  $V' = V \cup e'$  and  $E' = E \cup \{e'\}$ , then  $\mathcal{H}' = \{V', E'\}$  denotes the new hypergraph with added edge. We will start by outlining the effects of adding an edge on the relative barycentric homology in dimension zero.

## 6.1.1 The Effects of Adding an Edge on Zeroth Dimensional Relative Homology

Let  $\mathcal{H}$  be a hypergraph. Recall that  $\beta_0^{rel}(\mathcal{H})$ , the 0th relative Betti number, is the rank of  $H_0^{rel}(\mathcal{H})$ . This is the number of path components of  $\mathcal{H}$  that are simplicial complexes, called simplicial components from Definition 2.1.16.

We will consider when adding an edge can change  $H_0^{rel}$ . First, how can adding an edge to the hypergraph increase the number of connected components that are simplicial complexes? One option is to add a new connected component that was not there before. This would have to be a single new edge around one new vertex. If it is around more vertices, then there will be missing subedges, since we have only added one edge. An isolated vertex/edge pair like this that does not interact with the rest of the hypergraph in any way could be interesting and is worth tracking. This would increase  $\beta_0^{rel}$  by 1. This case is the topic of the first proposition.

**Proposition 6.1.1.** <sup>†</sup> Suppose  $e' \cap V = \emptyset$ . Then,

$$\beta_0^{rel}(\mathcal{H}') = \begin{cases} \beta_0^{rel}(\mathcal{H}) + 1 & if |e'| = 1\\ \\ \beta_0^{rel}(\mathcal{H}) & else. \end{cases}$$

*Proof.* Since  $e' \cap V = \emptyset$ , the new edge e' is its own connected component in the hypergraph. This means that it cannot reduce  $\beta_0^{rel}$ , by Theorem 4.3.7. It will increase  $\beta_0^{rel}$  when it is a simplicial component. If |e'| = 1, then it is simplicial. If |e'| > 1, then it is not simplicial as it will have some missing subedges, since e' is the only edge in its connected component.

Figure 6.1 has three hypergraphs:  $\mathcal{H}$ ,  $\mathcal{H}_1$ , and  $\mathcal{H}_2$ . Each of  $\mathcal{H}_1$  and  $\mathcal{H}_2$  adds one edge that is a new connected component to  $\mathcal{H}$ . By the above proposition,  $\beta_0^{rel}(\mathcal{H}_1) = \beta_0^{rel}(\mathcal{H}) + 1$  and  $\beta_0^{rel}(\mathcal{H}_2) = \beta_0^{rel}(\mathcal{H})$  based on the number of vertices in the added edge.



Figure 6.1 Three hypergraphs used as examples for Proposition 6.1.1.

The second option for increasing the number of simplicial components is to turn an existing connected component into a simplicial complex by adding an edge. This would mean that there was a connected component in  $\mathcal{H}$  that was only missing one edge, thereby finishing a simplicial component. This would increase  $\beta_0^{rel}$  by 1. It represents the densest and second most dense types of connected components possible in a hypergraph. This is proved in the next proposition. Recall that a subedge is an edge that is contained in another edge, while a toplex is an edge that is not contained in any other edges. Every edge in a hypergraph is either a toplex or subedge.

**Proposition 6.1.2.** <sup>†</sup> Suppose  $e' \subset e$  for some  $e \in E$ . Denote by C' the connected component of  $\mathcal{H}'$  containing e'. Then,

$$\beta_0^{rel}(\mathcal{H}') = \begin{cases} \beta_0^{rel}(\mathcal{H}) + 1 & C' \text{ is simplicial} \\ \beta_0^{rel}(\mathcal{H}) & else. \end{cases}$$

*Proof.* Since e' is a subedge, it only intersects with one connected component of  $\mathcal{H}$ , and that component was missing an edge, so it was not simplicial to begin with. Therefore, adding e' cannot reduce  $\beta_0^{rel}$ , by Theorem 4.3.7. Adding e' will only increase  $\beta_0^{rel}$  if it finishes filling in a connected component, turning it into a simplicial component.

Something else that may be interesting here is that, if a connected component is only missing one edge, it will have trivial relative homology (Theorem 4.3.4). So if a connected component has nontrivial relative homology in dimension greater than 0, adding a subedge to it cannot add a 0 dimensional homology, proving the following corollary.

**Corollary 6.1.3.** Suppose  $\mathcal{H}$  is a connected hypergraph with  $\beta_i^{rel} > 0$  for some i > 0. Then adding an edge to  $\mathcal{H}$  cannot increase  $\beta_0^{rel}$ .

The previous results showed how adding an edge to a hypergraph can increase the number of simplicial components. Next, how can adding an edge to a hypergraph *reduce* the number of connected components that are simplicial complexes? This is done by adding an edge that intersects with a connected component that was a simplicial complex. If this is done in a way that the new edge is missing subedges, then any previous simplicial components that were connected to this new edge are no longer simplicial. What follows is the case when e' intersects exactly one simplicial component.

**Proposition 6.1.4.** <sup>†</sup> Suppose e' is an that intersects one simplicial component, C. Let |e'| = d. *Then*,

$$\beta_0^{rel}(\mathcal{H}') = \begin{cases} \beta_0^{rel}(\mathcal{H}) & \text{all size } d-1 \text{ subsets of } e' \text{ are in } \mathcal{H} \\ \beta_0^{rel}(\mathcal{H}) - 1 & \text{else} \end{cases}$$

*Proof.* Since e' intersects exactly one simplicial component C, adding e' will reduce  $\beta_0^{rel}$  by 1 unless C remains simplicial after appending e', by Theorem 4.3.7. This can only occur if all d - 1 subsets of e' are edges in  $\mathcal{H}$ . All of the other subsets will also be there because it was a simplicial component to begin with.

However, a newly added edge can intersect any number of components at once, by, for example, including multiple components in a single edge. If the new edge intersects with any number other than two connected components that were simplicial complexes, it will make them all nonsimplicial. This will reduce  $\beta_0^{rel}$  by the number of connected components that were connected via the new edge.

If the new hyperedge intersects with exactly two connected components that were simplicial complexes, it is possible that instead of killing both of them, it merges them into one simplicial complex, only reducing  $\beta_0^{rel}$  by 1. This will happen exactly when the new hyperedge is between two vertices, one each in a connected component that was a simplicial complex. If there were any more



Figure 6.2 A hypergraph where adding the edge CD merged two simplicial components. than two vertices in it, (as would have to happen if it were intersecting more than two simplicial components), then there would be missing subedges inside of the new edge. This result is proved next.

**Proposition 6.1.5.** <sup>†</sup> Suppose e' is an edge that intersects k > 1 simplicial components, and |e'| = d. Then,

$$\beta_0^{rel}(\mathcal{H}') = \begin{cases} \beta_0^{rel}(\mathcal{H}) - 1 & k = d = 2\\ \beta_0^{rel}(\mathcal{H}) - k & else \end{cases}$$

*Proof.* Note that, in order to intersect k different components,  $d \ge k$ . We will consider when d > 2 and d = 2 separately. If d > 2, let  $\{v_i, v_j\}$  be vertices in e' where  $v_i$  is in a different simplicial component then  $v_j$ . So  $\{v_i, v_j\}$  is a missing subedge of e', and so the new connected component containing e' is not a simplicial component. That means none of the formerly simplicial components that intersect e' will be simplicial components in  $\mathcal{H}'$  and  $\beta_0^{rel}$  will decrease by the number of simplicial components that intersect e', byt Theorem 4.3.7.

In the case k = d = 2, e' intersects 2 simplicial components and is an edge with 2 vertices. Notice both vertices (and hence, all of the proper subsets) in e' are already edges in the hypergraph. Instead of making both components nonsimplicial, this merges two simplicial components into one, therefore only reducing  $\beta_0^{rel}$  by 1.

For an example of the k = d = 2 case, see Figure 6.2. The hypergraph  $\mathcal{H}'$  in this figure is the hypergraph after the edge *CD* was added. Notice how, if *CD* was removed from the hypergraph,

there would be two simplicial components, with toplexes *ABC* and *DEF*, so  $\beta_0^{rel}(\mathcal{H}) = 2$ . The edge *CD* connects two simplicial components and has two vertices, so  $\beta_0^{rel}(\mathcal{H}) = 1$ .

Lastly, there are ways of adding an edge without changing  $\beta_0^{rel}$ . This will probably be the most common result of adding an edge. One way to do this is by adding a new connected component which has more than one vertex, as in  $\mathcal{H}_2$  of Figure 6.1. This does not affect any of the other existing connected components, and adds a connected component that is not a simplicial complex, so it will not change  $\beta_0^{rel}$ .

You could also add a new toplex intersecting connected components that were not simplicial complexes to begin with. This could be one or more connected components, but as long as none of them were simplicial complexes, adding toplexes will not affect  $\beta_0^{rel}$ . Adding subedges to connected components that are missing at least two hyperedges will also not change  $\beta_0^{rel}$ .

So far we have enumerated what can happen to  $\beta_0^{rel}$  when an isolated edge or a subedge is added to the hypergraph. Now let e' be a toplex intersecting at least one edge in E. The change in  $\beta_0^{rel}$ from adding e' depends primarily on how many simplicial components it intersects. Note that the addition of a toplex cannot turn a non-simplicial component into a simplicial one. If e' is a toplex that does not intersect any simplicial components,  $\beta_0^{rel}$  will not change.

**Proposition 6.1.6.** <sup>†</sup> Suppose e' is a toplex that intersects no simplicial components, but does intersect  $\mathcal{H}$ . Then  $\beta_0^{rel}(\mathcal{H}') = \beta_0^{rel}(\mathcal{H})$ .

*Proof.* Let *C* be a simplicial component in  $\mathcal{H}$ . Because *e'* does not intersect *C*, *C* is still a simplicial components in  $\mathcal{H}'$ . Therefore,  $\beta_0^{rel}(\mathcal{H}') > \beta_0^{rel}(\mathcal{H})$ . Propositions 6.1.1 and 6.2 detailed all ways that adding an edge can increase  $\beta_0^{rel}$ , neither of which apply here. Therefore,  $\beta_0^{rel}(\mathcal{H}') = \beta_0^{rel}(\mathcal{H})$ .

Next, we will move on to the effect of adding an edge on the restricted barycentric homology in dimension zero.

#### 6.1.2 Effects of Adding an Edge on Zeroth Dimensional Restricted Barycentric Homology

Recall that given a hypergraph, the rank of the zeroth dimensional restricted barycentric homology group, denoted by  $\beta_0^{res}$ , is the number of fence components of the hypergraph. Next we will detail what happens to  $\beta_0^{res}$  when edges are added to the hypergraph. Again, the effects of removing an edge are the inverse of the effects of adding the same edge.

How can adding an edge to the hypergraph add a new fence component to the hypergraph? The fence components of the hypergraph are given by the inclusion relations of the hypergraph, so adding a subedge will never add a new fence component to the hypergraph. In fact, even for adding toplexes, in order to create a new fence component, that new edge must not include any other edges. So in order to add an edge to increase  $\beta_0^{res}$ , it must be a toplex that does not have any subedges, although it can intersect other edges.

Sometimes adding an edge can reduce  $\beta_0^{res}$  by merging fence components. As with some of the above, a single edge wrapped around what were formerly separate connected components will reduce  $\beta_0^{res}$  by an arbitrary amount. Moreover, it is possible to add subedges and merge fence components. Imagine a star type hypergraph where multiple edges that do not otherwise intersect all share a single vertex, as in Figure 6.3. If this vertex is not an edge in the hypergraph, then each edge is its own fence component, but adding in the "star vertex" will create a hypergraph with only one fence component. Unlike the case with relative barycentric homology, there will always be a fence component in whatever is left, so the merger will reduce  $\beta_0^{res}$  by 1 less than the number of components that were involved.

Note that it is not possible to add an edge that is simultaneously a subedge in one fence component and a toplex in another fence component, as the larger fence component must already have contained the smaller one. This means that in order for this kind of merge to take place, the new edge must be exclusively a subedge or toplex of the fence components it is merging.

Oftentimes, adding an edge to the hypergraph will not change  $\beta_0^{res}$ . This will happen when the edge added is part of exactly one fence component that was already present. It can be either a toplex or a subedge. If an edge is added that only has an inclusion relation with one other edge, then it

will not change the restricted homology at all (as in Theorem 3.2.6). If the edge added is a toplex, then it must have at least one subedge and all of its subedges must be totally ordered. Otherwise it would have merged more than one fence component. If the added edge is a subedge, then all of the edges that it is contained in must be able to be totally ordered. (Any subedges of the new edge would be contained in its toplexes as well, so those will not affect  $\beta_0^{res}$ .) The first result is what can happen to  $\beta_0^{res}$  when e' is a subedge.

**Proposition 6.1.7.** <sup>†</sup> Suppose e' is a subedge of edges representing r different fence components. *Then* 

$$\beta_0^{res}(\mathcal{H}') = \beta_0^{res}(\mathcal{H}) - (r-1)$$

*Proof.* Note that it is impossible for r to be 0, since e' is a subedge of some edge  $e \in E$ , and that edge must be in a fence component. If r = 1, then all of the edges e with  $e' \subset e$  are already in the same fence component and so e' is also only in that fence component and the number does not change. So let e > 1 and choose  $e_1, ..., e_r$  edges such that each is in a different fence component of  $\mathcal{H}$  and  $e' \subset e_i \forall 1 \leq i \leq r$ . Then for i > 1,  $\{e_1, e', e_i\}$  is a valid fence. Thus all of the  $e_i$  are now in the same fence component, and the number of fence components has been reduced by r - 1. Hence,

$$\beta_0^{res}(\mathcal{H}') = \beta_0^{res}(\mathcal{H}) - (r-1)$$

in all cases.

One example of this theorem is a hypergraph where many edges share a central vertex, as in Figure 6.3. In the hypergraph  $\mathcal{H}$ , each of those edges is a separate fence component, but after the edge A is added into the hypergraph  $\mathcal{H}'$ , all of the edges are in the same fence component. Therefore, adding the edge A reduced  $\beta_0^{res}(\mathcal{H})$  by four. The second result in this section is what can happen to  $\beta_0^{res}$  when e' is a toplex.

**Proposition 6.1.8.** <sup>†</sup> Suppose e' is a toplex with subedges representing r different fence components. Then

$$\beta_0^{res}(\mathcal{H}') = \beta_0^{res}(\mathcal{H}) - (r-1).$$



Figure 6.3 A hypergraph where adding the edge A reduced  $\beta_0^{res}(\mathcal{H})$ 

*Proof.* If r = 0, then e' doesn't have any subedges, and therefore cannot form a fence with any other edges, so it is a fence component all by itself, thereby increasing  $\beta_0^{res}$  by 1. So let  $r \ge 1$  and choose  $e_1, ..., e_r$  edges such that each is in a different fence component of  $\mathcal{H}$  and  $e' \supset e_i \forall 1 \le i \le r$ . Then for i > 1,  $\{e_1, e', e_i\}$  is a valid fence. Thus all of the  $e_i$  are now in the same fence component, and the number of fence components has been reduced by r - 1. Thus

$$\beta_0^{res}(\mathcal{H}') = \beta_0^{res}(\mathcal{H}) - (r-1),$$

proving the proposition.

In the first lemma of this section, it is impossible for r = 0, but in the second lemma this is possible, and that is the only case where adding an edge can increase  $\beta_0^{res}$ . Figure 6.1 serves as a good example for this result. Here both  $\mathcal{H}_1$  and  $\mathcal{H}_2$  add a new fence component, increasing  $\beta_0^{res}$ from  $\mathcal{H}$ . This is different from the relative case, in which only  $\mathcal{H}_1$  increased  $\beta_0^{rel}$ .

In both the restricted and relative barycentric homology theories, increasing  $\beta_0$  by adding an edge seems to have the fewest number of possible causes. It will thus usually be easy to track why  $\beta_0^{rel}$  or  $\beta_0^{res}$  went up, and which new edge caused the increase. If  $\beta_0^{rel}$  and  $\beta_0^{res}$  both increased, this can only be the result of a new singleton edge around a vertex that was not present before, which is certainly worth tracking. In the next section, we will dive into results about the effects of adding a vertex into some existing edges.

## 6.2 Adding a Vertex

It is natural to consider the effects of adding a vertex to a hypergraph. There might be a new data point that enters into existing relationships, like a new species in an ecosystem. There are two main results in this section. The first applies to the relative barycentric homology of maximum edge hypergraphs, and says that adding a vertex inside of just the maximum edge increases the dimension of all homology cycles by one. We will also give some remarks about under what conditions this can apply to hypergraphs that are not necessarily maximum edge. The second result will give a condition on which adding a vertex to a hypergraph can affect the zeroth dimensional restricted barycentric homology.

# 6.2.1 The Effects of Adding a Vertex to a Maximum Edge Hypergraph on the Relative Barycentric Homology

Let  $\mathcal{H}$  be a maximum edge hypergraph, as in Definition 2.1.10, with maximum edge e. Let |e| = m. Suppose a new vertex  $v_{m+1}$  is added to  $\mathcal{H}$  such that  $v_{m+1} \in e$  is only in the maximal edge. Denote the new hypergraph by  $\mathcal{H}'$ . Therefore,

$$V' = \{v_1, v_2, ..., v_m, v_{m+1}\},\$$

with new maximum edge e' = V'. The edge set is

$$E' = \{\{e'\} \cup E \setminus \{e\}\}\},\$$

and, thus the hypergraph  $\mathcal{H}'$  is

$$\mathcal{H}' = \{V', E'\}.$$

**Theorem 6.2.1** (Add a Vertex Theorem). Suppose  $\mathcal{H}'$  is built from a maximum edge hypergraph  $\mathcal{H}$  by adding a single vertex into only the maximum edge of  $\mathcal{H}$ . Then  $H_n^{rel}(\mathcal{H}') \cong H_{n-1}^{rel}(\mathcal{H})$  for all  $n \ge 0$  (assuming that  $H_{-1}^{rel}(\mathcal{H}) = 0$ ).

*Proof.* This will be proved by showing that the missing subcomplex S' of  $\mathcal{H}'$  is homotopy equivalent to a suspension of S, the missing subcomplex of  $\mathcal{H}$ . Recall from [18] that the suspension  $\Sigma X$  of a topological space X is the join of X and two points. Another way of thinking about the suspension

is as a union of two cones of X, with the copy of X on the bases of the cones identified with itself. Recalling that cones are contractible, and an application of the Mayer-Vietoris sequence from Theorem 2.2.4, gives the result that suspending a space increases the dimension of its homology groups by one, i.e.  $\tilde{H}_{n+1}(\Sigma X) \cong \tilde{H}_n(X)$  for all  $n \ge 0$ . Notice that this result uses the reduced homology of Definition 2.2.11, and is known as the suspension theorem.

To ease notation, we will add ' to constructions pertaining to  $\mathcal{H}'$ . Hence, K' will be the associated simplicial complex of  $\mathcal{H}'$ , and in this case,  $K' = \{\sigma \mid \sigma \subset e'\}$ . Further, T' will be the barycentric subdivision of K', and S' will be the subcomplex generated by vertices corresponding to missing subedges of  $\mathcal{H}'$ , so, by definition,  $H_n^{rel}(\mathcal{H}') = H_n(T', S')$ . To avoid confusion with the vertex set of the hypergraph, we will use  $u_{\sigma}$  to denote vertices in T and T' during this proof.

The proof that S' is homotopy equivalent to a suspension of S will proceed as follows. First, we will describe the set of simplices in S', and split it up into three groups. Second, we will show that two of those groups contract to the vertices  $u_{\sigma_e}$  and  $u_{\sigma_{v_{m+1}}}$ , respectively. Lastly, we will form two cones with base S within S', which will finish showing that S' is homotopy equivalent to S.

We begin by describing which simplices of T' are in S'. Recall that a simplex in S',  $\Omega$ , consists of vertices representing simplices in K' that are missing subedges of  $\mathcal{H}'$ . By assumption, e' is the only edge of  $\mathcal{H}'$  containing the new vertex  $v_{m+1}$ . Therefore, every simplex in K' containing the vertex  $v_{m+1}$  except  $\sigma_{e'}$  represents a missing subedge in  $\mathcal{H}'$ . Therefore, if  $u_{\tau}$  is a vertex in T'representing a simplex  $\tau$  in K' that contains  $v_{m+1}$ ,  $u_{\tau}$  is part of the generating set for S'. We will denote the set of simplices in T' containing at least one such vertex  $u_{\tau}$  by  $S[v_{m+1}]$  to indicate their reliance on the new vertex  $v_{m+1}$ .

Let  $\Omega$  be a simplex in S' that is not in  $S[v_{m+1}]$ . Since none of the vertices  $u_{\sigma}$  in  $\Omega$  represent simplices  $\sigma$  in K' that contain  $v_{m+1}$ ,  $\Omega$  is in T. Now there are two cases,  $\Omega \in S$  or not. Note  $S \subset S'$ because if a subedge was missing in  $\mathcal{H}$ , it is still a missing subedge in  $\mathcal{H}'$ , as the only edge added to  $\mathcal{H}'$  was e', which is not in S. Suppose though that  $\Omega$  was not in S. This would mean that at least one simplex represented by a vertex in  $\Omega$  was not in S, but is missing in  $\mathcal{H}'$ . The only edge that is missing in  $\mathcal{H}'$  but was an edge in  $\mathcal{H}$  is e. We will refer to simplices  $\Omega$  of S' where at least one of the vertices represents the simplex e in K' as S[e] to denote its containment of the vertex  $u_{\sigma_e}$  in S'.

Note that *S* does not intersect *S*[*e*] or *S*[ $v_{m+1}$ ] since neither  $u_{\sigma_e}$  nor  $u_{\sigma_{v_{m+1}}}$  is in *S*. Also, by the definition of how simplices are constructed in the barycentric subdivision, *S*[*e*] does not intersect *S*[ $v_{m+1}$ ], since in *K'*, *e* and  $v_{m+1}$  do not have an inclusion relation. Therefore  $S' = S \cup S[e] \cup S[v_{m+1}]$  and those sets are disjoint.

The next step of the proof is to show that S[e] and  $S[v_{m+1}]$  are contractible. By definition, S[e] is contractible, since any simplex in S[e] has  $u_{\sigma_e}$  as a vertex. Next we will show that  $S[v_{m+1}]$  is contractible. Let  $\Omega$  be a simplex in  $S[v_{m+1}]$ . This means that one of the vertices  $u_{\tau}$  of  $\Omega$  represents a simplex  $\tau$  in K' that contains  $v_{m+1}$  as a vertex. Since  $v_{m+1} \subset \tau$ , that subset relationship is represented by an edge in the barycentric subdivision. Therefore,  $\Omega \cup u_{\sigma_{v_{m+1}}}$  is a simplex in S'. The set of simplices of the form  $\Omega \cup u_{\sigma_{v_{m+1}}}$  is contractible to the vertex  $u_{\sigma_{v_{m+1}}}$ , showing that  $S[v_{m+1}]$  is contractible.

The last step in the proof that S' is homotopy equivalent to a suspension of S is to show that S connects to both S[e] and  $S[v_{m+1}]$ , which were just shown to be contractible. Recalling that S is disjoint from both of those other sets, let  $\Omega$  be a simplex in S. Because every vertex  $u_{\sigma}$  in  $\Omega$  represents a simplex  $\sigma$  in K that is a subset of e,  $\Omega \cup u_{\sigma_e}$  is a simplex in S' and, in particular, in S[e]. Therefore, S connects to S[e], and since S[e] is contractible, it forms a cone of S. Similarly, if  $\Omega$  is a simplex in S,  $\Omega \cup u_{\sigma_{v_{m+1}}}$  is a simplex in S'. This is now a part of  $S[v_{m+1}]$ , and forms the other cone of S. We have shown that S' is homotopy equivalent to two distinct cones of S. Thus, S' is homotopy equivalent to a suspension of S.

So, from the suspension theorem, for all n,  $\tilde{H}_n(S) \cong \tilde{H}_{n+1}(S')$ . From the results on the homology of maximal edge hypergraphs (Theorem 4.2.3), this can be extended to the string of isomorphisms, for  $n \ge 1$ ,  $H_{n+1}^{rel}(\mathcal{H}) \cong H_n(S) \cong H_{n+1}(S') \cong H_{n+2}^{rel}(\mathcal{H}')$ , proving the theorem in dimensions higher than 2.

It remains to consider the homology groups  $H_m^{rel}(\mathcal{H}')$  when m = 0, 1, 2.  $H_0^{rel}(\mathcal{H}') = 0$  by Theorem 4.2.6 as S' is not empty since  $u_{\sigma_e} \in S'$ .

 $H_1^{rel}(\mathcal{H}')$  depends on  $H_0(S')$  as noted in Corollary 4.2.4. First, suppose that  $\mathcal{H}$  was a simplicial



Figure 6.4 A hypergraph used as an example for Theorem 6.2.1.

complex to begin with, so that *S* is empty. Then  $H_0^{rel}(\mathcal{H}) = \frac{\mathbb{Z}}{2\mathbb{Z}}$ , by Theorem 4.2.6. The suspension of the empty set is two disjoint points, so  $H_0(S') = \frac{\mathbb{Z}}{2\mathbb{Z}} \oplus \frac{\mathbb{Z}}{2\mathbb{Z}}$ . Therefore, by Theorem 4.2.7,  $H_1^{rel}(\mathcal{H}') = \frac{\mathbb{Z}}{2\mathbb{Z}} = H_0^{rel}(\mathcal{H}).$ 

Next, suppose that *S* was not empty. Then  $H_0^{rel}(\mathcal{H}) = 0$  by Theorem 4.2.6. In this case, *S'* must be path connected as a suspension of *S*, so by Theorem 4.2.7,  $H_1^{rel}(\mathcal{H}') = 0$  In both cases,  $H_1^{rel}(\mathcal{H}') \cong H_0(\mathcal{H})$ .

Recall that  $H_2^{rel}(\mathcal{H}') \cong H_1(S')$  by Theorem 4.2.3, and since S' is the suspension of S,  $H_1(S') \cong \tilde{H}_0(S)$ . Again, when S is empty,  $H_1^{rel}(\mathcal{H}) \cong \tilde{H}_0(S) = 0$ , so  $H_2^{rel}(\mathcal{H}') = 0$  as well. If S is not empty, then the following string of isomorphisms holds from the fact that S' is a suspension of S and Theorems 4.2.7 and 4.2.3:

$$H_1^{rel}(\mathcal{H}) \cong \tilde{H}_0(S) \cong H_1(S') \cong H_2^{rel}(\mathcal{H}').$$

In both cases,  $H_2^{rel}(\mathcal{H}') \cong H_1^{rel}(\mathcal{H})$ , and this concludes the proof of the theorem in all dimensions.

The images in Figure 6.4 serve to illustrate this theorem. The hypergraph  $\mathcal{H}$  has  $\beta_1^{rel}(\mathcal{H}) = 1$ , with all other relative Betti numbers equal to zero. The hypergraph  $\mathcal{H}_1$  is an example of a single vertex being added to  $\mathcal{H}$ . As long as the new vertex is added inside only the maximum edge, the theorem holds. Thus  $\beta_2^{rel}(\mathcal{H}_1) = 1$ , and all other Betti numbers are zero. The other hypergraph in the figure,  $\mathcal{H}_2$ , is different in that the vertex D was added into the edge C to make a new subedge CD as well. This hypergraph has the same relative Betti numbers, and we conjecture that this theorem holds no matter which subedges the new vertex is added into, as long as it is added

into the maximum edge. Even if the hypergraph is not a maximum edge hypergraph, it will still sometimes be possible to apply the above theorem. As the theorem shows, if the vertex is added to a maximum edge hypergraph, it suspends the missing subcomplex. Similarly, if a vertex is added into an existing toplex, the subcomplex of the missing subcomplex corresponding to missing subedges of that toplex will likewise be suspended. Therefore, a homology cycle that is entirely contained in the missing subedges of that toplex will see its dimension increase by one. This could also be seen by using the Mayer-Vietoris sequence, if either version applies for that hypergraph. It will not apply in the situation where the new vertex takes an edge that was a subedge and is not included in the edges which that edge was contained in, therefore creating a toplex. That situation will be addressed next for its effects on the restricted barycentric homology.

## 6.2.2 The Effects of Adding a Vertex on the Zeroth Restricted Barycentric Homology

Recall from Theorem 3.2.3 that the restricted barycentric homology of a maximum edge hypergraph  $\mathcal{H}$  is as follows:

$$H^{res}(\mathcal{H}) = \begin{cases} \frac{\mathbb{Z}}{2\mathbb{Z}} & \text{if } n = 0\\ 0 & \text{else.} \end{cases}$$

Adding a vertex inside of the maximum edge of a maximum edge hypergraph will not change the fact that is it a maximum edge hypergraph, and so it cannot change the restricted barycentric homology. One case where the number of fence components, and hence the zeroth dimensional restricted barycentric homology, can change is detailed below as the final result of this section.

**Theorem 6.2.2.** Let  $\mathcal{H}$  be a hypergraph with an edge e that is a subedge. Suppose a vertex  $v_{n+1}$  is added to the hypergraph such that  $v_{n+1}$  is in exactly e and its subedges. Then,  $\beta_0^{res}$  will increase, by an amount depending on the number of toplexes containing e.

*Proof.* Let  $\mathcal{H}'$  denote the hypergraph with  $v_{n+1}$ . Recall from Theorem 3.2.4 that  $\beta_0^{res}$  is the number of fence components. Let f be an edge of  $\mathcal{H}$  such that  $e \subset f$ . Therefore, by Definition 2.1.13, in  $\mathcal{H}$ , e and f are part of the same fence component. In order to prove the theorem, it will suffice to show that in  $\mathcal{H}'$ , e and f are not part of the same fence component.



Figure 6.5 A hypergraph used as an example for Theorem 6.2.2.

Suppose, to the contrary, that there is a fence  $e, e_{i_1}, e_{i_2}, \ldots, e_{i_k}, f$  between e and f. By assumption, e is the largest edge containing  $v_{n+1}$ , and is thus a toplex. Therefore,  $e_{i_1}$  must be a subedge of e, and, by assumption, must contain  $v_{n+1}$  as well. By definition of fence (Definition 2.1.12),  $e_{i_1} \subset e_{i_2}$ , so  $e_{i_2}$  also contains  $v_{n+1}$ , and is thus a subset of e. Therefore,  $e_{i_3} \subset e_{i_2} \subset e$  and all of the edges in the fence will likewise be subsets of e. Since f is not a subset of e, there cannot be a fence from e to f, and they must be in different fence components.

As an example of this theorem, see the hypergraphs in Figure 6.5. The vertex D has been added to hypergraph  $\mathcal{H}$  to get  $\mathcal{H}'$ . That vertex was added inside the edge BC to create a new edge BCD. In  $\mathcal{H}$ , BC is a subedge, but in  $\mathcal{H}'$ , BCD is a toplex. Note that  $\mathcal{H}$  has one fence component, but  $\mathcal{H}'$ has two. Therefore,  $\beta_0^{res}(\mathcal{H}) = 1$ ,  $\beta_0^{res}(\mathcal{H}') = 2$ , and, as the theorem states,  $\beta_0^{res}$  has increased.
## **CHAPTER 7**

## **CONCLUSIONS AND FUTURE WORK**

We will close this thesis by offering some concluding remarks about the implication of its contents as well as some interesting ideas for future research along these lines. Data is rapidly gaining importance in most fields of study, including many that would not have been considered as data science fields even twenty years ago. As technology improves, we have the ability to create larger and more complex data sets. To handle complex data sets, researchers and analysts need more sophisticated methods to model data. One such method is that of hypergraphs. Hypergraphs are useful tools because they serve as generalizations for two common combinatorial objects: graphs and simplicial complexes. Graphs have been used as data models for over 150 years, whereas the science of using simplicial complexes to model data is much more recent. Hypergraphs combine useful features of both graphs and simplicial complexes. Like graphs, they are simply a list of vertices and a list of edges, and like simplicial complexes, they are naturally able to model higher order relationships. Because of the generality of hypergraph structure, they are natural models for many different types of data sets, and sometimes serve as better representations than graphs or simplicial complexes.

However, with greater generality comes less rigidity and structure. When researchers study graphs, one question they often ask is about the cycles in those graphs. There is not a canonical notion of what constitutes a cycle in a hypergraph, and, in fact, there are at least five different definitions of a cycle in a hypergraph that all reduce to the definition of a cycle if the hypergraph happens to be a graph. Moreover, since hypergraphs can display higher order relationships, a natural question to ask is if there is any notion of what it would mean for a hypergraph to have a higher order cycle.

This is where the idea of homology is useful. A graph can be viewed as a one dimensional simplicial complex, and when it is, the zeroth dimensional homology gives information about the connected components of that graph, and the first dimensional homology gives information about the structure of cycles in the graph. In general, the homology of simplicial complexes gives an

interpretation of higher order cycles. Unfortunately, the way in which hypergraphs generalize simplicial complexes means that the definition of the homology of a simplicial complex does not immediately translate to hypergraphs. The question of defining a homology theory for hypergraphs which accurately reflects the way in which hypergraphs generalize simplicial complexes is an important one as data scientists increase their usage of hypergraphs to model data. This thesis contributes towards an answer to that question.

Two particular homology theories for hypergraphs, the restricted and relative barycentric homology, were put forth by Emilie Purvine and collaborators [28]. This thesis takes those two theories and significantly develops them in three ways: combinatorially, topologically, and computationally. Combinatorially, we discussed new definitions on hypergraphs that allow for these homology theories to be more readily interpreted, such as the supplement of a hypergraph or its fence components. We took concepts standard in algebraic topology, such as the Mayer-Vietoris Sequence or the Long Exact Sequence of a Pair, and adapted them for use with these hypergraph homology theories. These adaptations are useful to simplify calculations of the homology of complicated hypergraphs. Lastly, for data science purposes, everything needs to be computable. We developed methods of building the necessary structures for computation without using the barycentric subdivision, which is computationally unwieldy, and discussed the novel algorithm that is used by HyperNetX for the computation of the restricted and relative barycentric homology of hypergraphs. This thesis helps to make the theories computable and the results interpretable, both necessary ingredients for successful data analysis.

There is still more work to be done. We will close by briefly discussing some ideas for future research. To align with the three main areas of contribution, we will discuss potential questions in the combinatorial, topological, and computational directions. As mentioned above, there is a hierarchy of hypergraph acyclicities. One potential question to ask is how those different acyclicity theories fit with the relative and restricted barycentric homology of a hypergraph. Is a hypergraph with certain homology groups guaranteed a specific type of acyclicity? In graph theory, researchers often look for specific subgraphs and the effect that they might have on the properties of the entire

graph. Can we do something similar with hypergraph minors, and in particular, is there a certain subedge structure that always generates a homology class?

In topology, there is also a notion of singular homology which applies to continuous spaces. For geometric realizations of simplicial complexes, these two notions coincide [18, 24]. One advantage of singular homology is that you can consider the homology of a subspace of a simplicial complex without it needing to be a subcomplex. In particular, we would like to prove that the restricted barycentric homology of a hypergraph  $\mathcal{H}$  is isomorphic to the singular homology of the associated simplicial complex K restricted to the simplices that represent edges in the hypergraph. Similarly, we conjecture that the relative barycentric homology of  $\mathcal{H}$  is isomorphic to the singular homology of K relative to the simplices that represent edges missing from the hypergraph. Recall that some of the theorems on the relative barycentric homology, in particular Theorem 4.2.7, give more information about the edges that are missing from the hypergraph than the edges that are present. To switch this result to a result that gives information about the edges present in the hypergraph, we could define a slightly different version of the relative barycentric homology, where instead of collapsing the missing subcomplex, the restricted barycentric subdivision is collapsed instead. Then, the analogous version of Theorem 4.2.7 would be giving information about the edges that are present in the hypergraph. This would equate Theorem 4.2.7 to Theorem 4.2.9. In general, these two theorems seem to be complementary, and when there is a maximum edge hypergraph on three vertices, they are exactly the same. Just as with many duality theorems in topology, the dimensions of homology discussed in the reference theorems add up to the dimension of the simplicial complex. One interesting question for future research is if this concept can be generalized to a duality theorem for the relative barycentric homology of hypergraphs.

For areas relating to computation and data science, there are also some natural next steps. First, the code currently only returns the Betti numbers, and it would be useful to have it also give the generators for the homology classes, to add to the interpretability of the results. Also, real world data is constantly changing and it is important to continue to develop tools to handle dynamic hypergraphs and hypergraph filtrations, along the lines of [25]. Lastly, we would like to apply these

methods to real data sets from other disciplines, such as biology or social science, to glean insights useful to researchers in those fields, as well as motivate needed advances in the theory.

Because of how quickly the field of data science is growing, researchers are continually trying to develop new methods that can be used to analyze increasingly large and complex data sets while simulataneously maintaining computational feasibility and accurate, interpretable results. This thesis not only gives computational tools towards topological methods for hypergraph analytics in the form of hypergraph homology theories; it also further develops the topology and combinatorics of these methods to make the results of the computational tools meaningfully interpretable.

## BIBLIOGRAPHY

- [1] Sinan G. Aksoy et al. "Hypernetwork science via high-order hypergraph walks". In: *EPJ Data Science* 9.1 (2020). DOI: 10.1140/epjds/s13688-020-00231-0.
- [2] Hans-Jürgen Bandelt and Victor Chepoi. "Metric graph theory and geometry: a survey". In: 2006.
- [3] Federico Battiston et al. "Networks beyond pairwise interactions: Structure and dynamics". In: *Physics Reports* 874 (2020). Networks beyond pairwise interactions: Structure and dynamics, pp. 1–92. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2020.05.004.
   URL: https://www.sciencedirect.com/science/article/pii/S0370157320302489.
- [4] Austin R Benson, David F Gleich, and Desmond J Highan. *Higher-order network analysis takes off, fueled by old ideas and new data*. 2021. URL: https://sinews.siam.org/Details-Page/higher-order-network-analysis-takes-off-fueled-by-old-ideas-and-new-data.
- [5] C. Berge. Hypergraphs: Combinatorics of Finite Sets. North-Holland Mathematical Library. Elsevier Science, 1984. ISBN: 9780080880235. URL: https://books.google.com/books?id= jEyfse-EKf8C.
- [6] Christian Bick et al. "What are higher-order networks?" In: *CoRR* abs/2104.11329 (2021). arXiv: 2104.11329. URL: https://arXiv.org/abs/2104.11329.
- [7] Stephane Bressan et al. *The Embedded Homology of Hypergraphs and Applications*. 2018. arXiv: 1610.00890.
- [8] A. Bretto. Hypergraph Theory: An Introduction. Mathematical Engineering. Springer International Publishing, 2013. ISBN: 9783319000800. URL: https://books.google.com/ books?id=lb5DAAAAQBAJ.
- [9] F. R. K. Chung and R. L. Graham. "Cohomological Aspects of Hypergraphs". In: *Transactions of the American Mathematical Society* 334.1 (1992), pp. 365–388. ISSN: 00029947. URL: http://www.jstor.org/stable/2153987.
- [10] Alessandro D'Atri and Marina Moscarini. "On Hypergraph Acyclicity and Graph Chordality". In: *Inf. Process. Lett.* 29 (1988), pp. 271–274.
- France Dacar. "The cyclicity of a hypergraph". In: *Discrete Mathematics* 182.1 (1998), pp. 53–67. ISSN: 0012-365X. DOI: https://doi.org/10.1016/S0012-365X(97)00133-7. URL: https://www.sciencedirect.com/science/article/pii/S0012365X97001337.
- [12] Reinhard Diestel. Homological aspects of oriented hypergraphs. 2021. arXiv: 2007.09125.
- [13] Jean-Guillaume Dumas et al. "Computing Simplicial Homology Based on Efficient Smith

Normal Form Algorithms". In: *Algebra, Geometry and Software Systems*. Ed. by Michael Joswig and Nobuki Takayama. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 177–206. ISBN: 978-3-662-05148-1.

- [14] Eric Emtander. "Betti Numbers of Hypergraphs". In: Communications in Algebra 37.5 (2009), pp. 1545–1571. DOI: 10.1080/00927870802098158. eprint: https://doi.org/10.1080/00927870802098158. URL: https://doi.org/10.1080/00927870802098158.
- [15] Ronald Fagin. "Degrees of Acyclicity for Hypergraphs and Relational Database Schemes".
  In: J. ACM 30.3 (July 1983), pp. 514–550. ISSN: 0004-5411. DOI: 10.1145/2402.322390.
  URL: https://doi.org/10.1145/2402.322390.
- [16] Song Feng et al. Hypergraph Models of Biological Networks to Identify Genes Critical to Pathogenic Viral Response. 2020. arXiv: 2010.03068.
- [17] Giorgio Gallo et al. "Directed hypergraphs and applications". In: Discrete Applied Mathematics 42.2 (1993), pp. 177–201. ISSN: 0166-218X. DOI: https://doi.org/10. 1016/0166-218X(93)90045-P. URL: https://www.sciencedirect.com/science/article/pii/ 0166218X9390045P.
- [18] A. Hatcher, Cambridge University Press, and Cornell University. Department of Mathematics. *Algebraic Topology*. Algebraic Topology. Cambridge University Press, 2002. ISBN: 9780521795401. URL: https://pi.math.cornell.edu/~hatcher/AT/AT.pdf.
- [19] Jeffrey Johnson. Hypernetworks in the Science of Complex Systems. IMPERIAL COLLEGE PRESS, 2014. DOI: 10.1142/p533. eprint: https://www.worldscientific.com/doi/pdf/10. 1142/p533. URL: https://www.worldscientific.com/doi/abs/10.1142/p533.
- [20] Cliff A. Joslyn et al. "Hypergraph Analytics of Domain Name System Relationships". In: *Algorithms and Models for the Web Graph*. Ed. by Bogumił Kamiński, Paweł Prałat, and Przemysław Szufel. Cham: Springer International Publishing, 2020, pp. 1–15. ISBN: 978-3-030-48478-1.
- [21] Cliff A. Joslyn et al. *Hypernetwork Science: From Multidimensional Networks to Computational Topology.* 2020. arXiv: 2003.11782.
- [22] Bill Kay et al. "Hypergraph Topological Features for Autoencoder-Based Intrusion Detection for Cybersecurity Data". In: International Conference in Machine Learning. ICML-ML4Cyber, July 2022.
- [23] Luis M. Márquez et al. "A Virus in a Fungus in a Plant: Three-Way Symbiosis Required for Thermal Tolerance". In: *Science* 315.5811 (2007), pp. 513–515. DOI: 10.1126/science. 1136237.
- [24] J.R. Munkres. *Elements Of Algebraic Topology*. CRC Press, 2018. ISBN: 9780429962462.

URL: https://books.google.com/books?id=-mdQDwAAQBAJ.

- [25] Audun Myers et al. *Topological Analysis of Temporal Hypergraphs*. 2023. DOI: 10.48550/ ARXIV.2302.02857. URL: https://arxiv.org/abs/2302.02857.
- [26] A.D. Parks, S.L. Lipscomb, and NAVAL SURFACE WARFARE CENTER DAHLGREN VA. Homology and Hypergraph Acyclicity: A Combinatorial Invariant For Hypergraphs. Defense Technical Information Center, 1991. URL: https://books.google.com/books?id= kUIJkAEACAAJ.
- [27] Brenda Praggastis. *PNNL/HyperNetX: Python package for hypergraph analysis and visualization.* 2021. URL: https://github.com/pnnl/HyperNetX.
- [28] Emilie Purvine. *Homology of Graphs and Hypergraphs*. 2021. URL: https://www.youtube. com/watch?v=XeNBysFcwOw.
- [29] Shiquan Ren and Jie Wu. *Stability of persistent homology for hypergraphs*. 2020. arXiv: 2002.02237.
- [30] Shiquan Ren, Jie Wu, and Mengmeng Zhang. *Relative Embedded Homology of Hypergraphs*. 2021.
- [31] Leo Torres et al. "The why, how, and when of representations for complex systems". In: *SIAM Rev.* 63 (2021), pp. 435–485.
- [32] Michelle L. Wachs. "Poset topology: Tools and applications". In: *Geometric Combina*torics, IAS/Park City Mathematics Series. 2004.
- [33] Chong Wang, Shiquan Ren, and Jian Liu. *A Künneth Formula of Hypergraphs*. 2021. arXiv: 2003.05142.
- [34] J. H. C. Whitehead. "Simplicial Spaces, Nuclei and m-Groups". In: *Proceedings of the London Mathematical Society* s2-45.1 (1939), pp. 243–327. DOI: https://doi.org/10.1112/plms/s2-45.1.243. eprint: https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/plms/s2-45.1.243. URL: https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/plms/s2-45.1.243.