On the Loop Homology of a Certain Complex of RNA Structures

Thomas J. X. Li 1,† and Christian M. Reidys 1,2,*,†

Abstract: In this paper, we establish a topological framework of τ-structures to quantify the evolutionary transitions between two RNA sequence–structure pairs. τ-structures developed here consist of a pair of RNA secondary structures together with a non-crossing partial matching between the two backbones. The loop complex of a τ-structure captures the intersections of loops in both secondary structures. We compute the loop homology of τ-structures. We show that only the zeroth, first and second homology groups are free. In particular, we prove that the rank of the second homology group equals the number γ of certain arc-components in a τ-structure and that the rank of the first homology is given by γ − χ + 1, where χ is the Euler characteristic of the loop complex.

Keywords: topology; simplicial complex; homology; Mayer–Vietoris sequence; RNA; secondary structure

1. Introduction

Ribonucleic acid (RNA) is a biomolecule that folds into a helical configuration of its primary sequence by forming hydrogen bonds between pairs of nucleotides. The most prominent class of coarse-grained structures are the RNA secondary structures [1,2]. These are contact structures that can be represented as diagrams where vertices are nucleotides and arcs are base pairs drawn in the upper half-plane. As a secondary structure can be uniquely decomposed into loops, the free energy of a structure is calculated as the sum of the energy of its individual loops [3]. The loop-based energy models for RNA secondary structure prediction have played an important role in unveiling the various regulatory functions of RNA [4].

RNA molecules evolve under selection pressures on both sequence and structure. One prominent example are RNA riboswitches [5], which is a class of RNA molecules that can express two alternative secondary structures, each of which appears in a specific biophysical context. In order to analyze riboswitch sequences, Huang et al. developed a Boltzmann sampler of sequences that are simultaneously compatible with a pair of structures, i.e., a bi-secondary structure, based on the loop energy model [6]. The time complexity of computing such bi-compatible sequences turns out to be intimately related to the number of so called exposed nucleotides, i.e., the number of vertices shared by multiple loops appearing in the algorithm. In some sense, finding bi-compatible sequences is easy if the intersection of such loops is small and becomes difficult when the intersection is large. In order to address the complexity problem, Bura et al. studied the simplicial homology of the loop complex (derived from the intersections of loops) of a bi-secondary structure [7,8]. Subsequently, they introduced the weighted homology of bi-structures, in which simplices are endowed with specific weights allowing the expression of the size of the loop intersections [9].
The aforementioned work focuses on the structural transition without changing the underlying sequence. However, selection pressures act on both sequence and structure, resulting in the simultaneous change of genotypes and phenotypes. This motivates the study of the evolutionary transition between two distinct sequence–structure pairs. In what follows, we take the first step by, analogous to [7,8], computing the simplicial homology of their loop complex.

In this paper we study triples, \((S, T, \phi)\), consisting of an ordered pair of RNA secondary structures, \((S, T)\), [1,2,10–13] together with a non-crossing partial matching, \(\phi\), between the two backbones. We shall denote such a triple a transition structure or \(\tau\)-structure and note that the mapping \(\phi\) relates homologous bases between two underlying RNA sequences.

The topological framework of \(\tau\)-structures developed here is designed to quantify the probability of evolutionary transitions between RNA sequence-structure pairs subject to specific sequence constraints. In an algorithmic guise, \(\tau\)-structures play a central role in solving the following computational problem: Given two arbitrary RNA secondary structures, determine a pair of specifically related sequences that minimizes the total free energies of the two sequence–structure pairs. Our main objective is to compute the simplicial homology of the loop complex of a \(\tau\)-structure. In order to state our main theorem, we next introduce the notions of \(\phi\)-crossing and arc-component.

Two arcs \((i_1, i_2)\) and \((j_1, j_2)\) are \(\phi\)-crossing, i.e., if either one of the following is the case:

1. Both \(i_1\) and \(i_2\) are incident to \(\phi\)-arcs such that \(\phi(i_1) < j_1 < \phi(i_2) < j_2\) or \(j_1 < \phi(i_1) < j_2 < \phi(i_2)\), or;
2. Both \(j_1\) and \(j_2\) are incident to \(\phi\)-arcs such that \(i_1 < \phi^{-1}(j_1) < i_2 < \phi^{-1}(j_2)\) or \(\phi^{-1}(j_1) < i_1 < \phi^{-1}(j_2) < i_2\).

\(\phi\)-crossing induces an equivalence relation for which nontrivial equivalence classes are called arc-components. An arc-component is of type 1 if both endpoints of any arc are incident to \(\phi\)-arcs.

We shall prove the following main result.

**Main Theorem.** Let \(I = (S, T, \phi)\) be a \(\tau\)-structure and \(H_n(I)\) the \(n\)-th homology group of its loop complex, \(K(I)\). Then, the following is the case:

\[
\begin{align*}
H_3(I) &= 0, \\
H_2(I) &\cong \bigoplus_{k=1}^{\gamma} \mathbb{Z}, \\
H_1(I) &\cong \bigoplus_{k=1}^{\gamma - \chi + 1} \mathbb{Z},
\end{align*}
\]

where \(\gamma\) denotes the number of arc-components of type 1 and \(\chi\) the Euler characteristic of \(K(I)\).

Our framework generalizes the work of [7,8] on bi-structures. Bi-structures can be viewed as particular \(\tau\)-structures, namely those over two identical sequences. The loop complex of a bi-structure exhibits only a nontrivial zeroth and second homology, the latter being freely generated by crossing components [7,8]. On the algorithmic side, \(\tau\)-structures play a similar role as bi-structures in [6] where a Boltzmann sampler of sequences that are simultaneously compatible with two structures is presented.

The loop complex of \(\tau\)-structures exhibits a variety of new features. First it has in general a nontrivial first homology group. Secondly, it features two types of arc-components: Those of type 1, which correspond to spheres, and arc-components of type 2, which correspond to surfaces with a nontrivial boundary (see Section 7).

In order to prove the main theorem, we first establish basic properties of the loop complex \(X = K(I)\). Secondly, we manipulate \(X\) by means of simplicial collapses and then dissect \(K_1\), which is a certain sub-complex of dimension one. As a result, we derive \(\tilde{X}_2\), which is a specific \(X\)-sub-complex, and we organize it into its crossing components. We
note that $X_2$ is not unique, as its construction depends on certain choices. We then proceed with a combinatorial analysis of $\tilde{X}_2$-crossing components, which lays the foundation for proving that their geometric realizations are either spheres or surfaces with boundary.

The combinatorics of crossing components, specifically Lemma 6, further controls the manner crossing components are glued and this observation facilitates to prove the main result in two stages. First, by computing the homology of components in isolation and, secondly, by computing the homology of the loop complex by Mayer–Vietoris sequences. Several of our arguments are reminiscent of ideas of classical facts, for instance the Jordan curve theorem in the proof of Lemma 7.

The paper is organized as follows: In Section 2, we recall the definitions of RNA secondary structures, bi-structures and their loop complexes. In Section 3, we introduce $\tau$-structures and present some basic facts of their loop complexes. In Section 4, we simplify the loop complex in multiple rounds via simplicial collapses and topological stratification. In Section 5, we investigate the combinatorics of crossing components in the simplified complex. We then compute the homology of components in Section 6 and integrate this information in order to obtain the homology of $\tau$-structures in Section 7. Finally we discuss our results in Section 8.

2. Secondary Structures, Bi-Structures and the Loop Complex

An RNA secondary structure encapsulates the nucleotide interactions within a single RNA sequence. It can be represented as a diagram, which consists of a labeled graph over the vertex set $[n] := \{1, \ldots, n\}$ for which the vertices are arranged in a horizontal line and arcs are drawn in the upper half-plane. Each vertex corresponds to a nucleotide in the vertex set $\mathbb{X}$; any two arcs do not have a common vertex; (2) any two arcs do not have a common vertex; (3) any two arcs are non-crossing.

Given a secondary structure $S$, the set $[1, n] := \{1, \ldots, n\}$ is called the backbone and an interval $[i, j] \subseteq [1, n]$ denotes the set of consecutive vertices $\{i, i + 1, \ldots, j - 1, j\}$. A vertex $k$ is covered by an arc $(i, j)$ if $i \leq k \leq j$ and there exists no other arc $(p, q)$ such that $i < p < k < q < j$. The set of vertices covered by an arc $(i, j)$ is called a loop, $s$, and we refer to $(i, j)$ as the covering arc of $s$. We shall equip each diagram with two “normal” vertices at positions 0 and $n + 1$ together with the arc $(0, n + 1)$, which we call the rainbow. Its associated distinguished loop is called the exterior loop. Each loop, $s$, can be represented as a disjoint union of intervals on the backbone of $S$, $s = \bigcup_{k=1}^{l} [s_k]$, such that $(s_1, s_2)$ and $(s_k, s_{k+1})$ for $1 \leq k \leq l - 1$ form arcs and any other vertices in $s$ are unpaired. By construction, a diagram $S$ is in one-to-one correspondence with its set of arcs as well as its set of loops $\{s_i\}$.

Let $X$ be a collection of finite sets. A subset $\{X_0, X_1, \ldots, X_d\} \subseteq X$ is defined to be a $d$-simplex of $X$ if the set intersection $\cap_{k=0}^{d} X_k \neq \emptyset$. Let $K_d(X)$ be the set of all $d$-simplices of $X$. The nerve of $X$ is a simplicial complex given by $K(X) = \bigcup_{d=0}^{\infty} K_d(X) \subseteq 2^X$.

The nerve formed by the collection $\{s_i\}$ of $S$-loops is referred to as the loop complex $K(S)$ of a secondary structure $S$. The fact that secondary structures correspond to non-crossing diagrams immediately implies that the loop complex of a secondary structure is a tree.

Let $S$ and $T$ be two secondary structures. The pair of secondary structures $S$ and $T$ over $[n]$, $R = (S, T)$, is called a bi-structure. We represent a bi-structure as a diagram on a horizontal backbone with the $S$-arcs drawn in the upper and the $T$-arcs drawn in the lower half plane. The nerve formed by $S$-loops and $T$-loops of a bi-structure, $R$, is called the loop complex, $K(R)$.

In bi-structures [7], two arcs $a, a'$ are equivalent if there exists a sequence of arcs $a_1 = a, a_2, \ldots, a_l = a'$ such that two consecutive arcs $a_k$ and $a_{k+1}$ are crossing for $1 \leq k \leq l - 1$. This equivalence relation is used to simplify the combinatorial analysis ofbi-crossing components.
A crossing component of $R = (S, T)$ is an equivalence class that contains at least two crossing arcs.

The loop complex of a bi-structure has a nontrivial zeroth and second homology group. In particular, the first homology $H_1(R)$ is zero and the second homology $H_2(R)$ is free and its rank equals the number of crossing components of the bi-structure $R$ [8].

### 3. Some Basic Facts

Let $S$ and $T$ be two secondary structures over $[n]$ and $[m]$, respectively. We shall refer to their backbones as $[n]_S = \{1_S, 2_S, \ldots, n_S\}$ and $[m]_T = \{1_T, 2_T, \ldots, m_T\}$. We refer to the pair $(i, \phi(i))$ as a $\phi$-arc between $i \in [n]_S$ and $\phi(i) \in [m]_T$. Two $\phi$-arcs $(i_1, \phi(i_1))$ and $(i_2, \phi(i_2))$ are crossing if $i_1 < i_2$ and $\phi(i_1) > \phi(i_2)$.

**Definition 1.** A $\tau$-structure $I$ is a triple $(S, T, \phi)$ consisting of an ordered pair of secondary structures, $(S, T)$, together with a partial matching $\phi$ between their vertex sets $[n]_S$ and $[m]_T$ such that any two $\phi$-arcs are non-crossing.

In analogy to the case of bi-structures, we shall represent the diagram of a $\tau$-structure as a labeled graph over two horizontal backbones in which $S$-arcs are drawn in the upper halfplane and $T$-arcs are drawn in the lower halfplane with the edges drawn in between the two backbones representing $\phi$ (see Figure 1).

A bi-structure corresponds to a particular $\tau$-structure having two backbones of the same length and $\phi$ being the identity map.

![Figure 1. A $\tau$-structure.](image)

The partial matching $\phi$ induces an equivalence relation $\sim$ on the set of vertices of the two backbones by identifying for each pair $(i, \phi(i))$ the vertices $i$ and $\phi(i)$. The $\phi$-arcs then correspond one-to-one to nontrivial equivalence classes of size two. This equivalence relation gives rise to the definition of a loop in a $\tau$-structure.

**Definition 2.** Given a $S$-vertex or $T$-vertex $v \in [n]_S \cup [m]_T$, an $I$-vertex, $\bar{v}$, is an equivalence class induced by $v$, i.e., $\bar{v} \in [n]_S \cup [m]_T / \sim$. Given an $S$-loop $s$, the $\bar{S}$-loop $\bar{s}$ is the set of equivalence classes induced by the set of $s$-vertices. Similarly, we define the $T$-loop. An $I$-loop is defined to be either an $S$-loop or a $T$-loop.

Now we are in position to introduce the loop complex of a $\tau$-structure.

**Definition 3.** The nerve formed by $I$-loops is called the loop complex $K(I) = \bigcup_{d=0}^\infty K_d(I)$, where $K_d(I)$ denotes the set of $d$-simplices, i.e., the non-empty intersections of $d$ distinct loops. $H_n(I)$ shall denote the $n$-th homology group of the loop complex of $I$.

Let us proceed by collecting some basic properties of the loop complex.
Proposition 1. Let $I = (S, T, \phi)$ be a $\tau$-structure and $v_1, v_2$ be two distinct $S$-vertices. Then $\sigma_1 \neq \sigma_2$. In particular there exists a bijection between vertices in an $S$-loop $s$ and vertices in the corresponding $S$-loop $\bar{s}$.

Proof. By definition, any two $\phi$-arcs in the partial matching $\phi$ do not share common endpoints. Thus, $\phi$ maps $v_1$ and $v_2$ to different $T$-vertices. Therefore, the equivalence classes $\bar{v}_1$ and $\bar{v}_2$ are different.

Proposition 2. Let $I = (S, T, \phi)$ be a $\tau$-structure and $s_1, s_2, s_3$ be three different $S$-loops. Then, it follows that:

1. either $s_1 \cap s_2 = \emptyset$ or $|s_1 \cap s_2| = 2$;
2. $s_1 \cap s_2 \cap s_3 = \emptyset$.

Proof. For the secondary structure $S$, any two $S$-loops $s_1$ and $s_2$ either have trivial intersection or intersect at two endpoints of one arc, i.e., either $s_1 \cap s_2 = \emptyset$ or $|s_1 \cap s_2| = 2$. Proposition 1 guarantees that $|s_1 \cap s_2| = |s_1 \cap s_2|$, whence assertion (1). Any $S$-vertex is contained in at most two distinct $S$-loops. Thus $s_1 \cap s_2 \cap s_3 = \emptyset$. By Proposition 1, $s_1 \cap s_2 \cap s_3 = \emptyset$. By abuse of notation, we refer to $\bar{s}$-loops and $T$-loops $\bar{s}, \bar{t}$ simply as $S$-loops and $T$-loops $s, t$, respectively.

Proposition 3. Let $I$ be a $\tau$-structure with loop complex $K(I)$. Then, any 2-simplex $\sigma \in K_2(I)$ has the form $\{s_1, s_2, t_1\}$ or $\{t_1, t_2, s_1\}$, where $s_1, s_2$ are $S$-loops and $t_1, t_2$ are $T$-loops.

Proof. By construction, the three loops in a 2-simplex $\sigma$ have a non-empty intersection. In view of Proposition 2, not all three loops are from the same secondary structure. Therefore, $\sigma$ is of the form $\{s_1, s_2, t_1\}$ or $\{t_1, t_2, s_1\}$. A 1-simplex $\sigma = \{r_1, r_2\} \in K(I)$ is called pure if both $r_1$ and $r_2$ belong to the same secondary structure and mixed. Proposition 3 shows that any 2-simplex contains exactly one pure edge and two mixed edges.

Proposition 4. Let $I$ be a $\tau$-structure with loop complex $K(I)$. Then, any 3-simplex $\sigma \in K_3(I)$ has the form $\{s_1, s_2, t_1, t_2\}$, where $s_1, s_2$ are $S$-loops and $t_1, t_2$ are $T$-loops.

Proof. Suppose that $\sigma = \{r_1, r_2, r_3, r_4\} \in K_2(I)$. By Proposition 3, we can w.l.o.g. set that $r_1 = s_1, r_2 = s_2$ and $r_3 = t_1$. Similarly, the face $\{r_1, r_2, r_4\} \in K_2(I)$ implies that $r_4$ is necessarily a $T$-loop, i.e., $r_4 = t_2$.

Definition 4. Given a simplex $\sigma \in K(I)$, a face $\omega \subset \sigma$ is called $\sigma$-free if, for any simplex $\sigma_1 \in K(I)$ containing $\omega$, $\sigma_1$ is a face of $\sigma$. Clearly, $\sigma$ is maximal in $K(I)$. Moreover, $\sigma$ is the unique maximal simplex that contains $\omega$.

Proposition 5. Let $I$ be a $\tau$-structure with loop complex $K(I)$. Then, any 3-simplex $\sigma \in K_3(I)$ contains a $\sigma$-free, mixed 1-face $\omega$.

Proof. Let $\sigma = \{s_1, s_2, t_1, t_2\}$. Suppose that the covering arcs of the loops $s_1, s_2, t_1, t_2$ are $(i_0, i_3), (i_1, i_2), (j_0, j_3)$ and $(j_1, j_2)$, respectively, and that we have the distinguished $\phi$-arc $(i_1, j_1)$ with $\phi(i_1) = j_1$. Then, the loops $s_1, s_2, t_1, t_2$ intersect at the equivalence class of $i_1 = j_1$ and by w.l.o.g. we may assume $i_0 < i_1 < i_2 < i_3$ and $j_0 < j_1 < j_2 < j_3$ (see Figure 2).
We shall prove the proposition by contradiction. Assume that none of the mixed 1-faces of \( \sigma \) are free, we consider the two particular mixed edges \( \omega_1 = \{ s_1, t_2 \} \) and \( \omega_2 = \{ s_2, t_1 \} \). By assumption, there exist 2-simplices \( \Delta_1 \) and \( \Delta_2 \), such that \( \omega_1 \subset \Delta_1 \) and \( \omega_2 \subset \Delta_2 \), neither of which are faces of \( \sigma \).

In view of \( \{ s_1, t_2 \} \subset \Delta_1 \), it follows from Proposition 3 that \( \Delta_1 \) is of the form \( \{ s_1, t_2, s' \} \) or \( \{ s_1, t_2, t' \} \). By w.l.o.g. we set \( \Delta_1 = \{ s_1, t_2, t' \} \). Since \( \Delta_1 \not\subset \sigma \), we have \( t' \neq t_1 \). By construction, \( s_1 \cap t_2 \cap t' \neq \emptyset \). It implies that there exists an \( I \)-vertex \( \bar{p} = \tilde{p} \in s_1 \cap t_2 \cap t' \), which corresponds to a \( \phi \)-arc \( (i', j') \). As \( i' \in s_1 \cap s_2 \cap t_1 \cap t_2 \) and \( \bar{p} = s_1 \cap t_2 \cap t' \), the \( \phi \)-arc \( (i', j') \) is distinct from \( (i_1, j_1) \). Since \( \omega_1 = \{ s_1, t_2 \} \subset \Delta_1 \), the \( \phi \)-arc \( (i', j') \) connects two intervals \( [i_0, i_1]_S \cup [i_2, i_3]_S \) and \( [j_1, j_2]_T \), i.e., \( i' \in [i_0, i_1]_S \cup [i_2, i_3]_S \) and \( j' \in (j_1, j_2)_T \). Note that if \( i' \in [i_0, i_1]_S \), then \( i' < i_1 \) and \( j' > j_1 \), i.e., the \( \phi \)-arc \( (i', j') \) and \( (i_1, j_1) \) are crossing, which is a contradiction. Hence, we derive that the \( \phi \)-arc \( (i', j') \) connects \( [i_2, i_3]_S \) and \( (j_1, j_2)_T \) (see Figure 2).

Similarly, \( \Delta_2 \) corresponds to a \( \phi \)-arc \( (i'', j'') \) distinct from \( (i_1, j_1) \). Since \( \omega_2 \subset \Delta_2 \), it follows that the \( \phi \)-arc \( (i'', j'') \) connects two intervals \( [i_1, i_2]_S \) and \( [j_2, j_3]_T \) (see Figure 2). In view of \( i'' < i_1 \) and \( j'' > j_3 \), the \( \phi \)-arcs \( (i', j') \) and \( (i'', j'') \) are crossing, which contradicts the fact that \( \phi \)-arcs are non-crossing.

Accordingly, the non-crossing property of \( \phi \)-arcs implies that either one, \( \{ s_1, t_2 \} \) or \( \{ s_2, t_1 \} \), is \( \sigma \)-free.

**Proposition 6.** Given a loop complex \( K(I) \), then any 3-simplex \( \sigma \in K_3(I) \) contains at least two \( \sigma \)-free 2-faces.

**Proof.** By Proposition 5, \( \sigma \) contains at least one \( \sigma \)-free mixed edge \( \omega \). Let \( \Delta_1 \) and \( \Delta_2 \) denote the 2-faces of \( \sigma \) containing \( \omega \). Suppose that there exists a simplex \( \sigma_1 \) satisfying \( \Delta_1 \subset \sigma_1 \) and \( \sigma_1 \not\subset \sigma \). Then \( \sigma_1 \) also contains \( \omega \), which contradicts \( \omega \) being \( \sigma \)-free. Therefore, \( \Delta_1 \) is \( \sigma \)-free and, analogously, we derive that \( \Delta_2 \) is \( \sigma \)-free.

**Proposition 7.** Let \( \tau \) be a \( \tau \)-structure with loop complex \( K(I) \). Then \( H_0(I) \cong \mathbb{Z} \) and \( H_d(I) = 0 \) for \( d \geq 4 \).

**Proof.** By assumption, \( \phi \) induces at least one \( \phi \)-arc, which in turn gives rise to at least a 1-simplex connecting an S-loop and a T-loop. Therefore, \( K(I) \) is path-connected and \( H_0(I) \cong \mathbb{Z} \). By the pigeonhole principle, for any set of \( d + 1 \) \( I \)-loops with \( d \geq 4 \), at least three belong to the same secondary structure. By Proposition 2, these three loops intersect trivially. Thus, for \( d \geq 4 \), the intersection of any \( d + 1 \) loops is trivial. As a result, \( K(I) \) does not contain simplices of dimension four or higher. Therefore, \( K_d(I) = \emptyset \) and \( H_d(I) = 0 \) for \( d \geq 4 \).
4. From $X$ to $\tilde{X}_2$

We now begin computing the homology of the loop complex of a $\tau$-structure. Directly constructing its discrete Morse function [14,15] or gradient vector field seems rather involved. We shall take a slightly different approach along the lines of discrete Morse theory. We shall simplify the loop complex in multiple rounds and render it amendable to its homological analysis. One of these simplifications is obtained via elementary collapses, which is a reduction introduced by J.H.C. Whitehead [16]. These collapses can be understood in terms of non-critical points of the Morse function.

4.1. Simplicial Collapses

The notion of elementary collapse is closely related to the freeness of a simplex. Given a complex $K$, let $\sigma$ be a maximal simplex and $\Delta$ be a $\sigma$-free face with codimension 1. The removal of $\sigma$ and $\Delta$ constitutes elementary collapse [16,17] and gives rise to a new complex $K \setminus \{\sigma, \Delta\}$. We say that $K$ simplicial collapses to a sub-complex $K'$, which is denoted by $K \searrow K'$ if there is a sequence of complexes $K_0 = K, K_1, \ldots, K_q = K'$ such that $K_i$ is obtained from $K_{i-1}$ via an elementary collapse. Clearly, $K$ and $K'$ are homotopy equivalent.

In the following, we shall employ simplicial collapses in order to reduce the loop complex in two steps: first, eliminating all tetrahedra and then removing triangles possessing a free edge.

**Step 1**: Given a 3-simplex $s_1 = \{s_1, s_2, t_1, t_2\}$, Proposition 5 guarantees that $s_1$ always has a free mixed edge $\omega$ and two free triangles $\Delta_1, \Delta'_1$ containing $\omega$. The triple $(\omega, \Delta_1, \Delta'_1)$ is called a $s_1$-butterfly. In analogy to the analysis in case of bi-structures [7], we perform the butterfly removal on the loop complex. Specifically, the simplicial collapse consists of two elementary collapses with respect to $(s_1, \Delta_1)$ and $(\Delta'_1, \omega)$ and provides us a sub-complex $K_1 = K(\omega) \setminus \{s_1, \Delta_1, \Delta'_1, \omega\}$, which is homotopy equivalent to $K(\omega)$.

We observe that $K_1$ is again the loop complex of a $\tau$-structure $I_1 = (S_1, T_1, \phi_1)$ obtained by splicing vertices in $I$ as follows: For any $I$-vertex $q \in s_1 \cap s_2 \cap t_1 \cap t_2$, suppose that $q$ is the equivalence class of an $S$-vertex $i$ and a $T$-vertex $j$. Then, there is a $\phi$-arc connecting $i$ and $j$, i.e., $\phi(i) = j$. We now splice $i$ into two consecutive vertices $i_1, i_2$ and splice $j$ into two consecutive vertices $j_1, j_2$. The $\phi$-arc $\phi(i) = j$ is replaced with two new $\phi$-arcs $\phi_1(i_1) = j_1$ and $\phi_1(i_2) = j_2$. The arcs are connected in a manner such that the mixed edge $\omega$ is not formed in $K_1$. In the case of $\omega = \{s_1, t_2\}$, we depict the resulting $\tau$-structure $I_1$ in Figure 3 and note that all other cases are treated similarly. Furthermore, we remove any other $\phi$-arcs that give rise to $\Delta_1, \Delta'_1$ or $\omega$. It is straightforward to verify that the loop complex of $I_1$ is exactly $K_1$.

![Figure 3. The $s_1$-butterfly removal. The simplicial collapse of the butterfly $(\omega, \Delta_1, \Delta'_1)$ corresponds to splicing vertices in the $\tau$-structure. Here, $s_1 = \{s_1, s_2, t_1, t_2\}, \Delta_1 = \{s_1, t_1, t_2\}, \Delta'_1 = \{s_1, s_2, t_2\}$ and $\omega = \{s_1, t_2\}$. The vertex $i$ is split into two consecutive vertices $i_1, i_2$ and $j$ is split into two consecutive vertices $j_1, j_2$. The $\phi$-arc $\phi(i) = j$ is replaced with two new $\phi$-arcs $\phi_1(i_1) = j_1$ and $\phi_1(i_2) = j_2$. Depending on the situation, the arcs are connected in a manner such that the loop $s_1$ and the loop $t_2$ do not intersect.]

By iteratively eliminating all 3-simplices, $\sigma$ and the associated $\sigma$-butterflies, we derive a sequence of sub-complexes $K_0 = K(I), K_1, \ldots, K_q = X$ such that $X$ does not contain any
3-simplices. Accordingly, we obtain a sequence of \( \tau \)-structures \( I_0 = I, I_1, \ldots, I_q = 1' \) by splicing the corresponding vertices of \( \sigma \) such that the loop complex of \( I_j \) is given by \( K_j \). It is clear that the sub-complex \( X \) satisfies the following property.

**Lemma 1.** Let \( K(I) \) be the loop complex of a \( \tau \)-structure \( I \) and \( X \) be the complex obtained by removing from \( K(I) \) all 3-simplices \( \sigma \) and the associated \( \sigma \)-butterflies. Then the following is the case:

1. \( X \) is the complex of a \( \tau \)-structure \( I' \) such that each \( \phi \)-arc is incident to at most one arc;
2. \( X \) is a sub-complex of \( K(I) \) such that the 0-th skeleton \( X^{(0)} = K(I)^{(0)} \) and \( X \) does not contain any 3-simplices;
3. \( H_n(X) \cong H_n(I), \) for \( n \geq 0 \).

A complex possessing the properties of Lemma 1 is called **lean**. As an immediate consequence we obtain the following.

**Corollary 1.** Let \( I \) be a \( \tau \)-structure, then \( H_3(I) = 0 \).

**Proof.** In view of Proposition 1, as a sub-complex of \( K(I) \), \( X \) does not contain any 3-simplices. For \( X \) holds \( \text{ker} \, \partial_3 = 0 \), whence \( H_3(I) \cong H_3(X) = 0 \).

**Step 2:** Suppose that \( \Delta_1 \) is a 2-simplex in \( X \) having a free 1-face \( \omega_1 \). By collapsing \( (\Delta_1, \omega_1) \), we obtain a sub-complex \( X_1 = X \setminus \{\Delta_1, \omega_1\} \). By successively deleting 2-simplices \( \omega \) together with \( \tau \)-free edges, we derive a sequence of sub-complexes \( X_0 = X, X_1, \ldots, X_m = \hat{X} \) such that each 2-simplex of \( \hat{X} \) does not contain any free 1-face.

**Proposition 8.** Let \( \hat{X} \) be a complex obtained from \( X \) by iteratively removing all 2-simplices \( \omega \) together with \( \tau \)-free 1-face. Then, the following is the case:

1. \( \hat{X} \) is a sub-complex of \( X \) such that the 0-th skeleton \( \hat{X}^{(0)} = X^{(0)} \) and each 2-simplex of \( \hat{X} \) does not have a free 1-face;
2. For \( n \geq 0 \), we have \( H_n(\hat{X}) \cong H_n(X) \).

Figure 4 displays the simplicial collapse from \( X \) to \( \hat{X} \).

**Remark 1.** The sub-complex \( \hat{X} \) is not unique. By construction, \( \hat{X} \) depends on the particular sequence of the simplicial collapses, as well as which free face of a simplex is being removed (see Figure 5).
Figure 5. $\tilde{X}$ is not unique. Performing the following sequence of elementary collapses on the complex $X$ and is displayed in Figure 4; $\{2, 4, 9\}$, $\{2, 7, 9\}$, $\{2, 9\}$, $\{1, 7, 9\}$, $\{7, 9\}$, $\{1, 5, 9\}$, $\{1, 9\}$, $\{1, 5, 11\}$, $\{1, 11\}$ and $\{5, 9, 11\}$, $\{5, 9\}$ produces a different $\tilde{X}$.

Lemma 1 guarantees that $\tilde{X}$ is derived from a lean complex.

**Lemma 2.** Let $\tilde{X}$ be the sub-complex derived from a lean complex, then the following is the case:

1. For any 2-simplex $\Delta = \{s_1, s_2, t_1\} \in \tilde{X}$, holds $s_1 \cap s_2 \cap t_1 = \emptyset$, i.e., any $2$-simplex $\Delta$ corresponds to the $\phi$-arc $(i, \phi(i))$.

2. Any pure 1-simplex $\omega$ of $\tilde{X}$ is either maximal or contained in exactly two 2-simplices $\Delta, \Delta'$. In the latter case, $\omega$ is associated with an arc $(i, i')$ in $I$, such that vertices $i, i'$ are incident to $\phi$-arcs $(i, j)$ and $(i', j')$, respectively.

**Proof.** For any $\Delta = \{s_1, s_2, t_1\} \in \tilde{X}$, we have $|s_1 \cap s_2 \cap t_1| \leq |s_1 \cap s_2| = 2$. By construction, in $\tilde{X}$, $\omega$ is not $\Delta$-free, whence there exists another 2-simplex $\Delta' = \{s_1, s_2, t'_1\} \neq \Delta$ containing $\omega$. Lemma 1 guarantees that $\tilde{X}$ does not contain any 3-simplices, i.e., we have the following:

$$(s_1 \cap s_2 \cap t_1) \cap (s_1 \cap s_2 \cap t'_1) = s_1 \cap s_2 \cap t_1 \cap t'_1 = \emptyset.$$ 

Therefore, sets $s_1 \cap s_2 \cap t_1$ and $s_1 \cap s_2 \cap t'_1$ are disjoint and both have cardinality 1, whence assertion (1) follows. Let $s_1 \cap s_2 \cap t_1 = \emptyset$, then $\Delta$ is in one-to-one correspondence with the vertex $i$ as well as its induced $\phi$-arc $(i, \phi(i))$ in $I$.

Suppose the pure 1-simplex $\omega$ is not maximal, then there exists at least one 2-simplex $\Delta$ having $\omega$ as a face. Since we work in the complex $\tilde{X}$, $\omega$ cannot be $\Delta$-free, whence there exists another 2-simplex $\Delta'$ containing $\omega$. Assertion (1) then implies that $\Delta$ and $\Delta'$ are the only such 2-simplices, which completes the proof. $\square$

4.2. Topological Stratification

We shall next show that the complex $\tilde{X}$ is topologically stratified and decomposes into two sub-complexes $\tilde{X}_2$ and $K_1$, where $\tilde{X}_2$ is induced by all $\tilde{X}$-2-simplices and $K_1$ by all 1-simplices of $\tilde{X} \setminus \tilde{X}_2$ (see Figure 6)

**Figure 6.** Topological stratification. $\tilde{X}$ stratifies into a sub-complex $\tilde{X}_2$ induced by all 2-simplices and $K_1$ induced by the remaining 1-simplices.

**Proposition 9.** We have $H_2(\tilde{X}) \cong H_2(\tilde{X}_2)$ and $H_1(\tilde{X}) \cong H_1(\tilde{X}_2) \oplus \oplus_{i=1}^{k} \mathbb{Z}$, where the non-negative integer $k$ depends on $K_1$ and $\tilde{X}_2 \cap K_1$. 
Mathematics 2021, 9, 1749

6. Some Combinatorics of Crossing Components

In order to compute the homology of $\tilde{X}$, we introduce an equivalence relation partitioning the set of 2-simplices. To this end, we observe that 2-simplices of $\tilde{X}$ appear in pairs: Given a 2-simplex $\Delta$ with pure edge $\omega$, Lemma 2 guarantees that there exists a unique 2-simplex $\Delta'$ such that $\Delta \cap \Delta' = \omega$. Each pair $(\Delta, \Delta')$ is associated with a unique arc $(i_1, i_2)$ and $i_1$ and $i_2$ are incident to the $\phi$-arcs, $(i_1, \phi(i_1))$ and $(i_2, \phi(i_2))$. We refer to $(\Delta, \Delta')$ together with their pure and $\phi$-arcs as a couple.

Two couples $(\Delta_1, \Delta'_1)$ and $(\Delta_2, \Delta'_2)$ are $\phi$-crossing if their corresponding arcs $(i_1, i_2)$ and $(j_1, j_2)$ are $\phi$-crossing, i.e., if either the following is the case:

1. Both $i_1$ and $i_2$ are incident to $\phi$-arcs such that $\phi(i_1) < j_1 < \phi(i_2)$ or $j_1 < \phi(i_1) < j_2 < \phi(i_2)$, or;
2. Both $j_1$ and $j_2$ are incident to $\phi$-arcs such that $i_1 < \phi^{-1}(j_1) < i_2 < \phi^{-1}(j_2)$ or $\phi^{-1}(j_1) < i_1 < \phi^{-1}(j_2) < i_2$.

This notion induces an equivalence relation $\sim_\phi$ on all $\tilde{X}$-couples to be the transitive closure with respect to $\phi$-crossing. Two couples $(\Delta_1, \Delta'_1)$ and $(\Delta_2, \Delta'_2)$ are $\phi$-equivalent, denoted by $(\Delta_1, \Delta'_1) \sim_\phi (\Delta_2, \Delta'_2)$, if there exists a sequence of pairs $(\Delta_1, \Delta'_1), (\Delta_2, \Delta'_2), \ldots, (\Delta_k, \Delta'_k)$, such that $(\Delta_1, \Delta'_1)$ and $(\Delta_{i+1}, \Delta'_{i+1})$ are $\phi$-crossing for $1 \leq i \leq k - 1$. The relation $\sim_\phi$ partitions the set of $\tilde{X}$-pairs into equivalence classes, $[(\Delta, \Delta')]$. A $\phi$-crossing component or component, $C$, of $\tilde{X}$ is the set of 2-simplices contained in an equivalence class $[(\Delta, \Delta')]$. 

Proof. By construction, $\tilde{X}_2$ is a sub-complex of $\tilde{X}$. We have the short exact sequence of chain complexes $0 \to C_*(\tilde{X}_2) \to C_*(\tilde{X}) \to C_*(\tilde{X}, \tilde{X}_2) \to 0$, which induces the long exact sequence.

$$
\cdots \to H_{n+1}(\tilde{X}, \tilde{X}_2) \xrightarrow{\partial} H_n(\tilde{X}_2) \xrightarrow{i_*} H_n(\tilde{X}) \xrightarrow{j_*} H_n(\tilde{X}, \tilde{X}_2) \to \cdots
$$

By construction, $\tilde{X}$ does not have any 3-simplices, which implies that $C_3(\tilde{X}, \tilde{X}_2) = 0$. Since $\tilde{X}_2$ contains all 2-simplices in $\tilde{X}$, we have $C_2(\tilde{X}, \tilde{X}_2) = 0$. Then we derive $H_2(\tilde{X}, \tilde{X}_2) = H_2(\tilde{X}, \tilde{X}_2) = 0$. In view of the long exact sequence, we have the following:

$$
0 = H_3(\tilde{X}, \tilde{X}_2) \xrightarrow{\partial} H_2(\tilde{X}_2) \xrightarrow{i_*} H_2(\tilde{X}) \xrightarrow{j_*} H_2(\tilde{X}, \tilde{X}_2) = 0,
$$

in other words, $H_2(\tilde{X}) \cong H_2(\tilde{X}_2)$. For $n = 1$, the long exact sequence reads as follows:

$$
0 = H_2(\tilde{X}, \tilde{X}_2) \xrightarrow{\partial} H_1(\tilde{X}_2) \xrightarrow{i_*} H_1(\tilde{X}) \xrightarrow{j_*} H_1(\tilde{X}, \tilde{X}_2),
$$

which induces the following exact sequence.

$$
0 \xrightarrow{\partial} H_1(\tilde{X}_2) \xrightarrow{i_*} H_1(\tilde{X}) \xrightarrow{j_*} \text{Im} j_* \xrightarrow{} 0.
$$

The relative homology $C_1(\tilde{X}, \tilde{X}_2)$ is freely generated by the 1-simplices in $\tilde{X} \setminus \tilde{X}_2$. By definition, we have $H_1(\tilde{X}, \tilde{X}_2) = \ker(\partial_2) / \text{Im}(\partial_2)$, where $C_2(\tilde{X}, \tilde{X}_2) \xrightarrow{\partial_2} C_1(\tilde{X}, \tilde{X}_2) \xrightarrow{\partial_1} C_0(\tilde{X}, \tilde{X}_2)$. In view of $\text{Im}(\partial_2) = 0$, $H_1(\tilde{X}, \tilde{X}_2) \cong \ker(\partial_1)$ and $H_1(\tilde{X}, \tilde{X}_2)$ is free. As a subgroup of the free group $H_1(\tilde{X}, \tilde{X}_2)$, $\text{Im} j_*$ is free and thus projective. Then, the short exact sequence is split exact and the following is the case:

$$
H_1(\tilde{X}) \cong H_1(\tilde{X}_2) \oplus \text{Im} j_* \cong H_1(\tilde{X}_2) \bigoplus_{i=1}^k \mathbb{Z},
$$

where the non-negative integer $k = \text{rnk}(\text{Im} j_*)$ depends on $K_1$ and $\tilde{X}_2 \cap K$. 

Proposition 9 shows that the sub-complex $K_1$ gives rise to only free generators in the first homology of $\tilde{X}$. 
By construction, the set of all 2-simplices in $\tilde{X}$ is partitioned into components, $C$, and each component $C$ induces a $\tilde{X}$-sub-complex. By the abuse of notation, we shall denote a component as well as its induced $\tilde{X}$-sub-complex by $C$. We refer to the $\varphi$- arcs of a component $C$ as $C$-$\varphi$-arcs.

Suppose that $(i_1,j_1),(i_2,j_2),\ldots,(i_l,j_l)$, where $i_1 < i_2 < \cdots < i_l$ and $\varphi(i_k) = j_k$ are $C$-$\varphi$-arcs. These partition the two backbones into blocks, the outer block $\beta_1 = [1,i_1]_S \cup (i_1,n)_S \cup [1,j_1]_T \cup (j_1,m)_T$ and the inner blocks $\beta_r = (i_{r-1},i_r)_S \cup (i_r,j_r)_T, 2 \leq r \leq l$. A block, $\beta$, contains a couple $(\Delta, \Delta')$ if all endpoints of $\varphi$-arcs associated with $(\Delta, \Delta')$ are contained in $\beta$ and the component $C_1$ if $\beta$ contains all $C_1$-couples.

Figure 7 demonstrates a decomposition of $\tilde{X}$ into components. In this example, the complex $K(I')$ does not a priori contain any free 1-face, whence $K(I') = \tilde{X}$. It is straightforward to verify that $\tilde{X}$ features the five components:

$C_1 = \{\{2,3,8\}, \{2,3,7\}, \{2,7,8\}, \{3,7,8\}\}$,
$C_2 = \{\{2,4,9\}, \{2,4,10\}, \{2,9,10\}, \{4,9,10\}\}$,
$C_3 = \{\{1,5,9\}, \{1,5,11\}, \{1,9,11\}, \{5,9,11\}\}$,
$C_4 = \{\{1,6,7\}, \{1,6,12\}, \{1,7,12\}, \{6,7,12\}\}$,
$C_5 = \{\{2,7,9\}, \{1,7,9\}\}$.

The $\varphi$-arcs of $C_5$ induce the outer block $[1,5]_S \cup (19,22]_S \cup [1,5]_T \cup (18,21]_T$ and the inner block $[7,17]_S \cup [7,16]_T$. The components $C_1, C_4$ are contained in the outer and $C_2, C_3$ in the inner block, respectively.

**Figure 7.** Components. LHS: A $\tau$-structure $I'$ with labeled loops. RHS: The complex $K(I')$ of $I'$ embedded in the Euclidean space $\mathbb{R}^3$. Note that $K(I')$ is identical to $\tilde{X}$. The 2-simplices in $K(I')$ are partitioned into five components. The insert (RHS) depicts the component in green, which is an empty tetrahedron.

**Lemma 3.** Let $C$ be a $\tilde{X}$-component with blocks $\beta_1, \beta_2, \ldots, \beta_l$. Then, for any other component $C_1 \neq C$, there exists a unique block, $\beta_k$, that contains $C_1$.

**Proof.** Suppose that $C$-$\varphi$-arcs are given by $(i_1,j_1),(i_2,j_2),\ldots,(i_l,j_l)$, where $i_1 < i_2 < \cdots < i_l$ and $\varphi(i_k) = j_k$.

**Claim:** Given a $C_1$-couple $(\Delta, \Delta')$ with $\varphi$-arcs $(i,j)$ and $(i',j')$, then there exists a block $\beta_k$ that contains both $(i,j)$ and $(i',j')$.

Firstly, both endpoints $i, j$ are contained in the same $C$-block otherwise $(i,j)$ would cross a $C$-$\varphi$-arc. Suppose $(i,j)$ and $(i',j')$ are contained in different $C$-blocks. Since $C$ and $C_1$ are distinct components, $C$-couples and $C_1$-couples are mutually non-crossing. As a result, $C_1$ partitions the $C$-couples into two non-empty subsets, one contained in $(i,i')_S \cup (j,j')_T$ and the other in $[1,i]_S \cup (i',n)_S \cup [1,j]_T \cup (j',m)_T$. By construction, any couple contained in the former subset and any pair in the latter are non-crossing, which is impossible since $C$ is a component.

It follows that there exists a $C$-block $\beta_k$ that contains both $(i,j)$ and $(i',j')$.

Any $C_1$-couple $(\Delta_1, \Delta_1')$ that crosses $(\Delta, \Delta')$ has at least one $\varphi$-arc contained in $\beta_k$. By the above argument, both of its $\varphi$-arcs are then contained in $\beta_k$, i.e., $\beta_k$ contains $(\Delta_1, \Delta_1')$. Since $C_1$ is a component this implies that any $C_1$-couple is contained in $\beta_k$. $\square$
Lemma 3 gives rise to a relation over the set of $\tilde{X}$-components as follows: $C_1 \prec C_2$ if all $C_1$-$\varphi$-arcs are contained in an inner block of $C_2$. In view of Lemma 3, we can verify that the relation $\prec$ is a well-defined partial order. In Figure 7, $C_2 \prec C_3$, $C_3 \prec C_5$ and $C_1, C_4, C_5$ are maximal.

We next identify how two 2-simplices that share a mixed edge affect the location of couples along the two backbones. This reflects the planarity of $\tau$-structures and has important consequences for how components are glued in the loop complex.

Lemma 4. Let $\Delta$ and $\Delta_1$ be $\tilde{X}$-2-simplices that share the mixed edge, $\omega$, and that correspond to the $\varphi$-arcs $(i, j)$ and $(i_1, j_1)$, where $i < i_1$. Then any couple $(\Delta_2, \Delta'_2)$ where $\{\Delta_2, \Delta'_2\} \cap \{\Delta, \Delta_1\} = \emptyset$ is contained in either

$$(i, i_1)_S \cup (j, j_1)_T \quad \text{or} \quad [1, i)_S \cup (i_1, n)_{S} \cup [1, j)_T \cup (j_1, m)_T.$$

Proof. Without loss of generality we may assume that $(\Delta_2, \Delta'_2)$ features the $S$-arc $(i_2, i'_2)$.

Let $\omega = \{s, t\}$, since $\Delta \cap \Delta_1 = \omega$, by construction $i, i_1 \in s$ and $j, j_1 \in t$. In case one of the vertices, $i_2, i'_2$ is contained in $(i, i_1)_S$ and the other is contained in $[1, i)_S \cup (i_1, n)_S$, the arc $(i_2, i'_2)$ organizes $i, i_1$ into distinct $S$-loops (see Figure 8), which is a contradiction. Thus, $i_2, i'_2$ are contained in $(i, i_1)_S$ or $[1, i)_S \cup (i_1, n)_S$, i.e., the $\varphi$-arcs associated with $(\Delta_2, \Delta'_2)$ are contained in $(i, i_1)_S \cup (j, j_1)_T$ or $[1, i)_S \cup (i_1, n)_S \cup [1, j)_T \cup (j_1, m)_T$. \quad \Box

![Figure 8. The planarity of a $\tau$-structure. Triangles $\Delta$ and $\Delta_1$ share the mixed edge, $\omega = \{s, t\}$ and correspond to the $\varphi$-arcs $(i, j)$ and $(i_1, j_1)$, respectively. Without loss of generality, we may assume that $i < i_1$. Any $\varphi$-arcs are non-crossing, we have $j < j_1 < j_2$.](image)

Lemma 5. Let $C$ be a $\tilde{X}$-component. Then for any 2-simplex $\Delta \in C$ and any 1-face $\omega$, there exists at most one 2-simplex $\Delta_1 \in C$ such that $\Delta \cap \Delta_1 = \omega$.

Proof. Clearly, the statement holds when $\omega$ is pure. For mixed $\omega$, by w.l.o.g. let $\Delta = \{s_1, s_2, t\}$ $\omega = \{s_2, t\}$ and assume the $S$-loop $s_2$ corresponds to the arc $(i, i')$ where $i < i'$.

Suppose there exist distinct 2-simplices $\Delta_1, \Delta_2 \in C$ having $\omega$ as a 1-face. We shall denote the $\Delta, \Delta_1$ and $\Delta_2$-$\varphi$-arcs by $(i, j)$, $(i_1, j_1)$, $(i_2, j_2)$, respectively. Without loss of generality, we may assume that $i < i_1 < i_2$. Since any $\varphi$-arcs are non-crossing, we have $j < j_1 < j_2$.

We now apply Lemma 4: Since $\Delta \cap \Delta_2 = \omega$, the couple $(\Delta_1, \Delta'_1)$ is contained in $(i, i_2)_S \cup (j, j_2)_T$ or $[1, i)_S \cup (i_2, n)_S \cup [1, j)_T \cup (j_2, m)_T$. Since $i < i_1 < i_2$, $(\Delta_1, \Delta'_1)$ is contained in $(i, i_2)_S \cup (j, j_2)_T$.

In view of $(\Delta, \Delta')$ and $(\Delta_1, \Delta'_1)$ belonging to $C$, there exists a sequence of pairs $(a_1, a'_1) = (\Delta_1, \Delta'_1), (a_2, a'_2), \ldots, (a_k, a'_k) = (\Delta, \Delta')$ such that $(a_i, a'_i)$ and $(a_{i+1}, a'_{i+1})$ are crossing.

Lemma 4 applies to any couple $(a_i, a'_i)$, locating it in $(i, i_2)_S \cup (j, j_2)_T$ or $[1, i)_S \cup (i_2, n)_S \cup [1, j)_T \cup (j_2, m)_T$. In addition, since $(a_2, a'_2)$ and $(\Delta_1, \Delta'_1)$ are crossing, $(a_2, a'_2)$ is contained in $(i, i_2)_S \cup (j, j_2)_T$, as is $(a_1, a'_1) = (\Delta_1, \Delta'_1)$. Analogously, it follows that $(a_k, a'_k) = (\Delta, \Delta')$ is contained in $(i, i_2)_S \cup (j, j_2)_T$, which is, by construction, impossible.

Accordingly, there exists at most one 2-simplex $\Delta_1 \in C$ such that $\Delta \cap \Delta_1 = \omega$. \quad \Box
Let \( C \) be a component and \( \omega \) a \( C \)-1-face. Lemma 5 implies that \( \omega \) is contained in either one or two \( C \)-2-simplices. We call a 1-face \( \omega \) that is contained in a unique \( C \)-2-simplex \( C \)-free. Furthermore we call a component \( C \) complete if, for any of its 1-faces \( \omega \), there exist exactly two 2-simplices \( \Delta, \Delta' \in C \) such that \( \Delta \cap \Delta' = \omega \) and incomplete, otherwise.

Let \( C \) be a component with \( \phi \)-arcs \( (i_1,j_1), (i_2,j_2), \ldots, (i_l,j_l) \), where \( i_1 < i_2 < \cdots < i_l \) and \( \phi(i_l) = j_k \). Let \( \Delta_1, \Delta_2, \ldots, \Delta_l \) denote the 2-simplices associated with the \( \phi \)-arcs \((i_1,j_1), (i_2,j_2), \ldots, (i_l,j_l)\), respectively. Suppose that \( C \)-\( \phi \)-arcs partition the two backbones into the sequence of blocks \( \beta_1, \beta_2, \ldots, \beta_l \) ordered from left to right and where \( \beta_1 = [1,i_1)s \cup (i_1,n]s \cup [1,j_1)_T \cup (j_1,m)_T \) denotes the outer-block. Let \( \omega_k \) and \( \omega'_k \) denote the mixed 1-faces of \( \Delta_k \) that are associated with \( \beta_k \) and \( \beta_{k+1} \), respectively. We call the mixed edges \( \omega_k \) and \( \omega'_k \) associated with the outer block the \( C \)-boundary. Note that \( \omega_1 \) and \( \omega'_1 \) coincide if and only if \( \Delta_1 \) and \( \Delta_1 \) share a common mixed edge.

The next lemma constitutes a key observation which facilitates the computation of the homology of \( \tau \)-structures via the Mayer–Vietoris sequence, see Theorem 2.

**Lemma 6.** Let \( C \) be an \( \bar{X} \)-component having boundary \( \omega_1 \) and \( \omega'_1 \), and let \( C' \) be an \( \bar{X} \)-component with the property \( C_1 \neq C' \). Suppose that \( \omega \) is a 1-simplex shared by \( C \) and \( C_1 \), then \( \omega = \omega_1 \) or \( \omega = \omega'_1 \).

**Proof.** Let \( \Delta_1, \Delta_2, \ldots, \Delta_l \) denote the 2-simplices of \( C \) associated with the \( \phi \)-arcs \((i_1,j_1), (i_2,j_2), \ldots, (i_l,j_l)\), ordered from left to right. Suppose that \( \Delta_0 \in C_1 \) and \( \Delta_0 \in C \) share a common mixed edge \( \omega \) and have the \( \phi \)-arcs \((i_0,j_0) \) and \((i_k,j_k) \), respectively. In view of \( C_1 \neq C, C_1 \) is contained in the outer \( C \)-block and we may by w.l.o.g. assume that \( i_0 < i_1 \).

The fact \( \Delta_0 \cap \Delta_0 = \omega \), now organizes \( C \) relative to \( \Delta_0 \) in a particular fashion, see Figure 9. Employing Lemma 4 we obtain that, except of \( (\Delta_k, \Delta'_k) \), all \( \Delta \)-couples partition into the following:

\[
A_1 \subset (i_0,i_k)s \cup (j_0,j_k)_T \quad \text{and} \quad A_2 \subset [1,i_0)s \cup (i_k,n)s \cup [1,j_0)_T \cup (j_k,m)_T.
\]

Now we use the fact that that \( C \) is a component: Firstly, any \( A_1 \)-couple and any \( A_2 \)-couple are by construction non-crossing. Secondly, by construction, \((\Delta_k, \Delta'_k)\) is crossing into either only one \( A_1 \) or \( A_2 \), which implies that either one, \( A_1 \) or \( A_2 \), is trivial.

Suppose \( A_1 = \emptyset \). Then all \( C \)-\( \phi \)-arcs are contained in \([i_k,n]s \) which implies \( \Delta_0 = \Delta_1 \). In case \( A_2 = \emptyset \), all \( C \)-\( \phi \)-arcs are contained in \((i_0,i_k) \) which implies \( \Delta_k = \Delta_1 \). Therefore, for the shared 1-edge, \( \omega \), holds \( \omega = \omega_1 \) or \( \omega = \omega'_1 \). \( \square \)

Figure 9. The fact \( \Delta_0 \cap \Delta_0 = \omega \) organizes \( C \)-couples into two subsets \( A_1 \subset (i_0,i_k)s \cup (j_0,j_k)_T \) and \( A_2 \subset [1,i_0)s \cup (i_k,n)s \cup [1,j_0)_T \cup (j_k,m)_T \), marked by dashed parallelograms.

### 6. The Homology of Components

In order to compute the homology of \( \tau \)-structures, we shall first compute the homology of components separately and then integrate this information via the Mayer–Vietoris sequence [18,19]. To this end we adopt a topological perspective of crossing components, taking a closer look at the manifold formed by 2-simplices of a component obtained via gluing along their edges.

Given a component \( C \), two couples \((\Delta_1, \Delta'_1) \sim_1 (\Delta_2, \Delta'_2) \) if there exists a sequence of pairs \((\Delta_1, \Delta'_1), (\Delta_2, \Delta'_2), \ldots, (\Delta_k, \Delta'_k) \) such that \((\Delta_i, \Delta'_i) \) and \((\Delta_{i+1}, \Delta'_{i+1}) \) share at least one
1-face for \( 1 \leq i \leq k - 1 \). Clearly, \( \sim_1 \) is an equivalence relation and partitions all pairs in \( C \) into equivalence classes \( D_1, D_2, \ldots, D_k \), which we refer to as \( C \)-ribbons.

**Lemma 7.** The geometric realization of a complete \( \tilde{X} \)-component, \( C \) is a sphere.

**Proof.** Suppose that \( \sim_1 \) partitions all \( C \)-pairs into the ribbons \( D_1, D_2, \ldots, D_k \).

**Claim 1:** The geometric realization of a ribbon, \( D_i \), is a surface without boundary.

Since \( C \) is complete, for each 2-simplex \( \Delta \in D_i \) and its 1-face \( \omega \), there exists a unique 2-simplex \( \Delta_1 \in C \) such that \( \Delta \cap \Delta_1 = \omega \). By construction, \( \Delta_1 \) is further contained in \( D_i \). This implies that the geometric realization of \( D_i \) is a compact and connected 2-manifold, i.e., a surface without boundary.

Let \( T_{D_i} \) denote the \( D_i \)-sub-complex consisting of all pure edges for which its loops are \( T \)-loops. We shall employ \( T_{D_i} \) in order to relate the crossing relation with \( \sim_1 \).

**Claim 2:** \( T_{D_i} \) is connected, i.e., \( T_{D_i} \) is a subtree of \( T \).

Suppose that \( T_{D_1} \) contains at least two connected components \( T_1 \) and \( T_2 \). Let \( D' \) denote the subset of 2-simplices of \( D_i \) containing at least one \( T_1 \)-vertex. Since all 2-simplices of \( D' \) are incident to the connected component \( T_1 \), the complex formed by \( D' \) is connected. \( D' \) does not contain any \( T_2 \)-vertices, since otherwise there would, by construction of \( D' \), exist a pure edge connecting a \( T_1 \)-vertex and a \( T_2 \)-vertex, which contradicts the assumption that \( T_{D_i} \) is disconnected. Thus, \( D' \) is a connected, which is a proper subset of \( D_i \), and its geometric realization is a surface having boundary \( B \), see Figure 10.

![Figure 10](image-url)

Figure 10. The boundary of \( D' \). The trees \( T_1 \) and \( T_2 \) (blue) represent the two connected components of \( T_{D_i} \). The sub-complex \( D' \) is formed by 2-simplices of \( D_i \) containing at least one \( T_1 \)-vertex. Then the boundary \( B \) of \( D' \) (red) consists of only \( S \)-loops.

The set of 2-simplices containing a fixed \( T_1 \)-vertex, \( t \), forms a neighborhood of \( t \) in the \( D_i \)-induced surface, whence any \( T_1 \)-vertex is contained in the interior of \( D' \). Consequently, the boundary does not contain any \( T_1 \)-vertex. From this, we conclude that the boundary, being a cycle, consists of only \( S \)-loops. This, however, is impossible since the complex of a secondary structure \( S \) is a tree.

**Claim 3:** \( C \) consists of a unique ribbon.

Suppose that \( C \) contains at least two ribbons \( D_1 \) and \( D_2 \). Since \( C \) is a component, there exists a couple \( (\Delta, \Delta') \) in \( D_1 \) and \( (\Delta_0, \Delta_0') \) not contained in \( D_1 \), such that \( (\Delta, \Delta') \) and \( (\Delta_0, \Delta_0') \) are crossing.

Without loss of generality, we may consider \( \Delta = \{s_1,s_2,t_1\}, \Delta' = \{s_1,s_2,t_2\}, \Delta_0 = \{t_1,t_2,s_1'\} \) and \( \Delta_0' = \{t_1,t_2,s_2'\} \), see Figure 11. Since \( (\Delta, \Delta') \) and \( (\Delta_0, \Delta_0') \) are crossing, the \( T \)-loops \( t_1',t_2' \) are on the path connecting \( T \)-loops \( t_1 \) and \( t_2 \) in the complex of the secondary structure \( T \).
By Claim 2, $T_{D_1}$ is connected and thus $t_1, t_2 \in T_{D_1}$ implies that $\{t'_1, t'_2\} \in T_{D_1}$. Claim 1 guarantees that each 1-face of $D_1$ is not $D_1$-free, whence there exist two $D_1$-2-simplices having the face $\{t'_1, t'_2\}$. However, $(\Delta_0, \Delta'_0)$ is the unique couple, for which the 2-simplices $\Delta_0, \Delta'_0$ both contain the pure 1-simplex $\{t'_1, t'_2\}$. As a result the fact that $D_1$ is a surface without boundary guarantees that $(\Delta_0, \Delta'_0)$ is contained in $D_1$, which is a contradiction.

Therefore $C$ is organized as a single ribbon, for which its geometric realization is a surface without boundary. We shall proceed by computing its Euler characteristic $\chi$.

Suppose that $C$ contains $n$ pairs $(\Delta, \Delta')$. Then the complex $C$ features $2n$ 2-simplices and $3n$ 1-simplices. Claim 2 stipulates the connectivity of $S_C$ and $T_C$, whence both $S_C$ and $T_C$ are connected subsets of trees and as such trees themselves. We shall use this in order to count 0-simplices as follows: each pair $(\Delta, \Delta')$ corresponds to one pure 1-simplex in $S_C$ or $T_C$. Thus $S_C$ and $T_C$ contain $n$ 1-simplices, which implies that $C$ contains $n + (1 + 1)$ 0-simplices. From this, it follows $\chi = 2$ and $C$ is homeomorphic to a sphere. □

**Remark 2.** The key in the proof is to show that the projection of $D_1$ onto each secondary structure, $T_{D_1}$, is connected. If $T_{D_1}$ is disconnected, then we can construct a cycle separating different connected components and consisting of only $S$-loops, resulting in a contradiction. The proof is reminiscent of the Jordan curve theorem in the plane [20,21].

While any complete component is organized as a distinguished ribbon, an incomplete component can consist of multiple ribbons. In Figure 12, the two couples $(\{2, 5, 6\}, \{4, 5, 6\})$ and $(\{1, 3, 5\}, \{1, 3, 6\})$ form an incomplete component, which contains two ribbons and each of which are a couple.

**Figure 11.** A couple $(\Delta, \Delta') \in D_1$ and $(\Delta_0, \Delta'_0) \notin D_1$ that are crossing, where $\Delta = \{s_1, s_2, t_1\}$, $\Delta' = \{s_1, s_2, t_2\}$, $\Delta_0 = \{t'_1, t'_2, s'_1\}$ and $\Delta'_0 = \{t'_1, t'_2, s'_2\}$.

**Figure 12.** LHS: A $\tau$-structure having an incomplete component (colored in blue and with its loops labeled by numbers and its 2-simplices labeled by triangles). RHS: The complex induced by the incomplete component indicates that it contains two ribbons.

**Lemma 8.** Given an incomplete $\bar{X}$-component $C$, having the ribbons $D_1, D_2, \ldots, D_k$. Then, for each $D_i$, there exists some 1-face $\omega \in D_i$ that is $D_i$-free.

**Proof.** Suppose there exists a ribbon $D_i$, in which each 1-face $\omega \in D_i$ is not $D_i$-free. Lemma 5 guarantees that there exist exactly two 2-simplices $\Delta, \Delta' \in D_i$ such that $\Delta \cap \Delta' = \omega$. Using the argument of Claim 2 of Lemma 7, we can conclude that $T_{D_i}$ is connected.

$C$ consists of at least two distinct ribbons since otherwise $C = D_i$ contains no $C$-free 1-face, which contradicts the assumption that $C$ is incomplete. Thus there exist a couple
\((\Delta, \Delta') \in D_i\) and \((\Delta_0, \Delta_0') \not\in D_i\), such that \((\Delta, \Delta')\) and \((\Delta_0, \Delta_0')\) are crossing. In analogy to Claim 3 of Lemma 7, we can conclude that both \(\Delta_0\) and \(\Delta_0'\) are contained in \(D_i\), which is impossible.

Therefore there exists a 1-face \(\omega \in D_i\) that is \(D_i\)-free. \(\square\)

**Theorem 1.** Let \(C\) be a component of \(\tilde{X}\). Then the following is the case:

\[
H_2(C) \cong \begin{cases} 
\mathbb{Z} & \text{if } C \text{ is complete} \\
0 & \text{if } C \text{ is incomplete} 
\end{cases} 
\tag{1}
\]

and the following is also the case:

\[
H_1(C) \cong \begin{cases} 
0 & \text{if } C \text{ is complete} \\
\oplus_{k=1}^r \mathbb{Z} & \text{if } C \text{ is incomplete,} 
\end{cases} 
\tag{2}
\]

where the non-negative integer \(r\) depends on \(C\). Furthermore, a complete component, \(H_2(C)\) is freely generated by the sum of all \(C\)-2-simplices.

**Proof.** In case \(C\) is complete, by Lemma 7, \(C\) is homeomorphic to a sphere, whence \(H_2(C) \cong \mathbb{Z}\) and \(H_1(C) \cong 0\). Clearly, \(H_2(C)\) is freely generated by the sum of 2-simplices of \(C\).

In case \(C\) is incomplete, suppose that \(D_1, D_2, \ldots, D_k\) are the \(C\)-ribbons. In the following, we prove Equations (1) and (2) by induction on the number \(k\) of ribbons.

For \(k = 1\), \(C\) is a ribbon and, as such, is connected. Since \(C\) is incomplete, it contains by Lemma 8 at least one \(C\)-free 1-face. This implies that the geometric realization of \(C\) is a surface with boundary. Therefore \(H_2(C) \cong 0\) and \(H_1(C)\) is free.

For the induction step, we shall combine ribbons in order to compute the homology of an incomplete component. In view of the fact that \(C\) is the union of the sub-complexes \(D_k\) and \(D' = \bigcup_{j \leq k} D_j\), we have the following inclusion maps:

\[
\begin{array}{ccc}
D_k & \xrightarrow{i} & D_k \\
\downarrow & & \downarrow h \\
D_k \cap D' & \xrightarrow{j} & C = D_k \cup D' \\
\downarrow & & \downarrow i \\
D' & \xrightarrow{j} & \end{array}
\]

Each inclusion map induces a chain map on the corresponding simplicial chain groups and a homomorphism between the corresponding homology groups.

\[
\begin{array}{ccc}
H_n(D_k) & \xrightarrow{i_*} & H_n(D_k) \\
\downarrow & & \downarrow h_* \\
H_n(D_k \cap D') & \xrightarrow{j_*} & H_n(C) \\
\downarrow & & \downarrow l_* \\
H_n(D') & \xrightarrow{j_*} & \end{array}
\]

Accordingly, we obtain the Mayer–Vietoris sequence

\[
\cdots \rightarrow H_n(D_k \cap D') \xrightarrow{\Phi} H_n(D_k) \oplus H_n(D') \xrightarrow{\Psi} H_n(C) \xrightarrow{\partial} H_{n-1}(D_k \cap D') \rightarrow \cdots
\]

where \(\Phi = i_* \oplus j_*\), \(\Psi = h_* - l_*\) and \(\partial\) is the connecting homomorphism given by the zig-zag lemma (see [22,23] for more details).
By Lemma 8, \( D_k \) contains at least one \( D_k \)-free 1-face, whence its geometric realization of \( D_k \) is a surface with boundary. As a result, \( H_2(D_k) \cong 0 \) and \( H_1(D_k) \) is free.

From the definition of ribbon follows, that the intersection \( D_k \cap D_j \) of any two ribbons cannot contain any 2-simplices. Furthermore, \( D_k \cap D_j \) does not contain any 1-simplices, whence \( D_k \cap D_j \) contains only 0-simplices.

In view of \( D_k \cap (\bigcup_{j<k} D_j) = \bigcup_{j<k} (D_k \cap D_j) \), we conclude that \( D_k \cap D' \) consists of only 0-simplices. Therefore, \( H_2(D_k \cap D') \cong H_1(D_k \cap D') \cong 0 \) and \( H_0(D_k \cap D') \) are free.

In case of \( n = 2 \), the Mayer–Vietoris sequence reads as follows:

\[
0 = H_2(D_k \cap D') \xrightarrow{\Phi} H_2(D_k) \oplus H_2(D') \xrightarrow{\Psi} H_2(C) \xrightarrow{\partial} H_1(D_k \cap D') = 0.
\]

Thus, \( \Psi \) is an isomorphism.

\[
H_2(C) \cong H_2(D_k) \oplus H_2(D').
\]

By induction hypothesis, we have \( H_2(D') \cong 0 \) and due to the fact that \( D_k \) contains a free \( D_k \)-1-face, \( H_2(D_k) \cong 0 \). Accordingly, we obtain \( H_2(C) \cong 0 \).

In case of \( n = 1 \), we observe the following:

\[
0 = H_1(D_k \cap D') \xrightarrow{\Phi} H_1(D_k) \oplus H_1(D') \xrightarrow{\Psi} H_1(C) \xrightarrow{\partial} H_0(D_k \cap D')
\]

which gives rise to the following exact sequence.

\[
0 \xrightarrow{\Phi} H_1(D_k) \oplus H_1(D') \xrightarrow{\Psi} H_1(C) \xrightarrow{\partial} \text{Im} \partial \xrightarrow{} 0.
\]

As a subgroup of the free group \( H_0(D_k \cap D') \), \( \text{Im} \partial \) is free and thus projective. As a result, the short exact sequence is split exact and the following:

\[
H_1(C) \cong H_1(D_k) \oplus H_1(D') \oplus \text{Im} \partial
\]

holds. Since both \( H_1(D_k) \) and \( H_1(D') \) are free, we conclude that \( H_1(C) \) is free, which completes the proof. \( \square \)

Theorem 1 shows that, while the complete components contribute only to the second homology, the incomplete components provide generators of the first homology. In terms of discrete Morse theory, each complete component contains a critical point of dimension 2 and each incomplete component can feature multiple critical points of dimension 1.

7. The Main Theorem

In this section, we compute the homology of a \( \tau \)-structure. The key tool here is the Mayer–Vietoris sequence, which allows us to connect and compose the homology data of the sub-complexes.

A certain ordering by which the components are glued in combination with Lemma 6 are critical for the application of the Mayer–Vietoris sequence, since they constitute the determinants of how components intersect.

**Theorem 2.** Let \( \tilde{X} \) be the complex obtained from the loop complex of a \( \tau \)-structure \( X \). Then, \( H_1(\tilde{X}) \) is free and the following is the case:

\[
H_2(\tilde{X}) \cong \bigoplus_{i=1}^M \mathbb{Z},
\]

where \( M \) denotes the number of complete \( \tilde{X} \)-components. Furthermore, \( H_2(\tilde{X}) \) is freely generated by \( \{ c_i \}_i^M \), where \( c_i \) denotes the sum of 2-simplices contained in a complete component, \( C_i \).
Proof. Let $\tilde{X}_2$ be the $\tilde{X}$-sub-complex induced by the $\tilde{X}$-2-simplices. Suppose that $\tilde{X}_2$ is partitioned into $j$ components $C_1, \ldots, C_j$. The set $\{C_i\}$ and any of its subsets are partially ordered and, by recursively removing maximal components, we can obtain a “descending” sequence $(C_{i_1}, \ldots, C_{i_k})$ such that the following obtains.

$$C_{i_h} \neq C_{i_k} \quad \text{for} \quad h < k.$$ 

This sequence gives rise to a sequence of sub-complexes, $(Y_0, Y_1, \ldots, Y_j)$ obtained by recursively adding a component as follows: $Y_0 = \emptyset$ and $Y_k = Y_{k-1} \cup C_{i_k}$. By construction, we have $Y_j = \tilde{X}_2 = \bigcup_{i=1}^j C_i$.

We next prove by induction on the number of components, $k$, that $H_1(Y_k)$ is free and $H_2(Y_k) \cong \bigoplus_{i=1}^{M_k} \mathbb{Z}$, where $M_k$ denotes the number of complete components in $C_{i_1}, \ldots, C_{i_k}$.

For the induction basis $k = 1$, $Y_1 = C_{i_1}$ itself is a component and its homology has been computed in Theorem 1. Clearly, when $C_{i_1}$ is complete, $H_2(Y_1)$ is freely generated by the sum of 2-simplices of $C_{i_1}$.

For the induction step we consider $Y_k$ as the union of two sub-complexes $C_{i_k}$ and $Y_{k-1}$ and shall combine $C_{i_k}$ and $Y_{k-1}$ by means of the Mayer–Vietoris sequence as follows.

$$\cdots \rightarrow H_n(C_{i_k} \cap Y_{k-1}) \rightarrow H_n(C_{i_k}) \oplus H_n(Y_{k-1}) \rightarrow H_n(Y_k) \rightarrow H_{n-1}(C_{i_k} \cap Y_{k-1}) \rightarrow \cdots$$

By construction, $C_{i_h} \neq C_{i_k}$ for any $h < k$, which enables the application of Lemma 6.

We accordingly conclude that the 1-skeleton of $C_{i_k} \cap C_{i_h}$ is contained in the $\{\omega_{i_1}, \omega_{i_2}\}$.

In view of $C_{i_j} \cap (\cup_{j < j} C_i) = \cup_{j < j} (C_{i_j} \cap C_i)$, for any $i$ and $j$, we derive that the 1-skeleton of $C_{i_k} \cap Y_{k-1}$ is contained in $\{\omega_{i_1}, \omega_{i_2}\}$. This severely constrains $C_{i_k} \cap Y_{k-1}$, which consists of 0-simplices and at most two 1-simplices.

As a result, $H_2(C_{i_k} \cap Y_{k-1}) \cong H_1(C_{i_k} \cap Y_{k-1}) \cong 0$ and $H_0(C_{i_k} \cap Y_{k-1})$ are free.

In case $n = 2$, we have the following.

$$0 = H_2(C_{i_k} \cap Y_{k-1}) \xrightarrow{\Phi} H_2(C_{i_k}) \oplus H_2(Y_{k-1}) \xrightarrow{\Psi} H_2(Y_k) \xrightarrow{\partial} H_1(C_{i_k} \cap Y_{k-1}) = 0.$$ 

Thus, $\Psi$ is an isomorphism and $H_2(Y_k) \cong H_2(C_{i_k}) \oplus H_2(Y_{k-1})$. By induction hypothesis, $H_2(Y_{k-1}) \cong \bigoplus_{i=1}^{M_{k-1}} \mathbb{Z}$, where $M_{k-1}$ denotes the number of complete components in $C_{i_1}, \ldots, C_{i_{k-1}}$. Due to Theorem 1, $H_2(C_{i_k}) \cong \mathbb{Z}$ in case of $C_{i_k}$ being complete and $H_2(C_{i_k}) \cong 0$ in case of $C_{i_k}$ being incomplete. Thus, we can conclude by induction

$$H_2(Y_k) \cong \bigoplus_{i=1}^{M_k} \mathbb{Z}.$$ 

Clearly, $H_2(Y_k)$ is freely generated by $\{c_{i_k}^{1} M_l\}$, where $c_{i_k}$ is given by the sum of 2-simplices of the $i_k$-th complete component.

In case $n = 1$, we have the following.

$$0 = H_1(C_{i_k} \cap Y_{k-1}) \xrightarrow{\Phi} H_1(C_{i_k}) \oplus H_1(Y_{k-1}) \xrightarrow{\Psi} H_1(Y_k) \xrightarrow{\partial} H_0(C_{i_k} \cap Y_{k-1})$$

and the following short exact sequence follows.

$$0 \xrightarrow{\Phi} H_1(C_{i_k}) \oplus H_1(Y_{k-1}) \xrightarrow{\Psi} H_1(Y_k) \xrightarrow{\partial} \text{Im} \partial \xrightarrow{} 0.$$ 

As a subgroup of the free group $H_0(C_{i_k} \cap Y_{k-1})$, $\text{Im} \partial$ is free and thus projective. Then the short exact sequence is split exact and the following is the case.

$$H_1(Y_k) \cong H_1(C_{i_k}) \oplus H_1(Y_{k-1}) \oplus \text{Im} \partial.$$ 

Since both $H_1(C_{i_k})$ and $H_1(Y_{k-1})$ are free, we can conclude that $H_1(Y_k)$ is free.
Therefore it follows by induction that \( H_1(\tilde{X}_2) = H_1(Y) \) is free and \( H_2(\tilde{X}_2) = H_2(Y) \cong \bigoplus_{i=1}^M \mathbb{Z}_i \), where \( M \) denotes the number of complete components in \( C_1, \ldots, C_l \).

In view of Proposition 9, we have the following:

\[
H_2(\tilde{X}) \cong H_2(\tilde{X}_2) \cong \bigoplus_{i=1}^M \mathbb{Z}_i
\]

and \( H_1(\tilde{X}) \cong H_1(\tilde{X}_2) \oplus \bigoplus_{i=1}^N \mathbb{Z}_i \) is free, which completes the proof. \( \square \)

The transitive closure with respect to \( \phi \)-crossing produces an equivalence relation, i.e., \( (i_1, i_2) \sim_\phi (j_1, j_2) \), if there exists a sequence of arcs such as the following:

\[
\alpha_1 = (i_1, i_2), \alpha_2, \ldots, \alpha_k = (j_1, j_2)
\]

such that two consecutive arcs \( \alpha_i \) and \( \alpha_{i+1} \) are \( \phi \)-crossing for \( 1 \leq k \leq l - 1 \).

An arc-component, \( A \), is an \( \sim_\phi \)-equivalence class of arcs such that \( A \) contains at least two arcs. \( A \) is of type 1 if both endpoints of any of its arcs are incident to \( \phi \)-arcs and of type 2, otherwise.

Suppose that \( (i_1, j_1), (i_2, j_2), \ldots, (i_l, j_l) \), where \( i_1 < j_1 < \cdots < j_l \) and \( \phi(i_k) = j_k \) are \( \phi \)-arcs associated with \( A \). These partition the two backbones into blocks: the outer block \( \beta_1 = [1, i_1] \cup (i_1, n] \cup [1, j_1] \cup (j_1, m] \) and the inner blocks \( \beta_r = (i_r, i_{r-1}] \cup (j_r, j_{r-1}] \), \( 2 \leq r \leq l \).

**Lemma 9.** In \( \Gamma' \), let \( A \) be an arc-component of type 1 with blocks \( \beta_2, \beta_3, \ldots, \beta_l \). Then, for any arc \( (i, i') \notin A \), there exists a unique block, \( \beta_{i'} \), that contains \( i \) and \( i' \).

**Proof.** Without loss of generality, we may assume that \( (i, i') \) is an \( S \)-arc. Since each \( \phi \)-arc is incident to at most one arc due to Lemma 1, \( (i, i') \) is not incident to any \( \phi \)-arc associated with \( A \) and thus its endpoints are contained in \( A \)-blocks.

By construction, any \( T \)-arc of \( A \), \( (j_p, j_q) \) and \( (i, i') \) are non-crossing, since otherwise \( (i, i') \) would belong to the arc-component \( A \). Accordingly, \( \phi^{-1}(j_p) \) and \( \phi^{-1}(j_q) \) are either contained in the interval \( (i, i') \) or \( [1, i] \cup (i', n] \) and we can conclude that \( \phi \)-arcs contained in either \( Y_0 \subset (i, i') \cup [1, i] \cup (i', n] \).

Suppose that \( i \) and \( i' \) are contained in different blocks, then \( Y_0 \) and \( Y_1 \) are nontrivial. By construction, any \( \phi \)-arc and any \( Y_1 \)-arc are non-crossing, which is impossible since \( A \) is an arc-component. As a result \( i \) and \( i' \) are contained in a single \( A \)-block. \( \square \)

**Remark 3.** Lemma 9 is the “arc”-analogue of Lemma 3 for arc-components of type 1. Note that the statement does not hold for arc-components of type 2.

Given an arc-component \( A \) of type 1 in the interaction structure \( \Gamma' \), each \( \phi \)-arc is associated with exactly one 2-simplex, as Lemma 1 guarantees that each \( \phi \)-arc is incident to at most one arc. Let \( \Delta_1, \Delta_2, \ldots, \Delta_l \) denote the 2-simplices associated with the \( \phi \)-arcs \( (i_1, j_1), (i_2, j_2), \ldots, (i_l, j_l) \), respectively. Let \( \omega_k \) and \( \omega'_k \) denote the mixed 1-faces of \( \Delta_k \) that are associated with \( \beta_k \) and \( \beta_{k+1} \), respectively.

**Lemma 10.** Let \( A \) be an \( \Gamma' \)-arc-component of type 1 with blocks \( \beta_1, \beta_2, \ldots, \beta_l \). Let \( \Delta_1, \Delta_2, \ldots, \Delta_l \) denote the 2-simplices associated with \( A \) and \( \omega_k \) and \( \omega'_k \) be the mixed 1-faces of \( \Delta_k \) that are associated with \( \beta_k \) and \( \beta_{k+1} \), respectively. Then \( \omega'_k = \omega_{k+1} \) for \( 1 \leq k \leq l - 1 \) and \( \omega'_l = \omega_1 \).

**Proof.** Let \( (i_1, j_1), (i_2, j_2), \ldots, (i_l, j_l) \) denote the \( \phi \)-arcs associated with \( A \) ordered from left to right. It suffices to prove \( \omega'_s = \omega_s \). Set \( \omega'_t = \{s, t\} \), where \( s \) and \( t \) are \( S \)-loops and \( T \)-loops associated with \( \beta_2 \). Clearly, \( i_s \in s \) and \( j_l \in t \).

By Lemma 9, any \( \Gamma' \)-arc is either contained in \( \beta_2 \) or its complement. This guarantees that \( i_1 \) and \( i_2 \) belong to the same \( S \)-loop \( s \) and, furthermore, that \( j_1 \) and \( j_2 \) belong to the same \( T \)-loop \( t \).
As a result the 2-simplex $\Delta_2$ associated with $(i_2, j_2)$ contains $\{s, l\}$ as a mixed face, i.e., $\omega'_1 = \omega_2$. \qed

Now we are in position to prove the Main Theorem.

**Proof of the Main Theorem.** The triviality of the third homology group, $H_3(I)$, follows from Corollary 1. In view of Lemma 1, Proposition 8 and Theorem 2, we have the following:

$$H_2(I) \cong H_2(\tilde{X}) \cong \oplus_{k=1}^\gamma \mathbb{Z},$$

where $M$ denotes the number of complete $\tilde{X}$-components.

First we observe that any $\sigma$-butterfly removal does not change the crossing status of any two arcs. Thus, there is a natural bijection between the set of $I$-arc-components and that of $I'$-arc-components. Moreover, each $I$-arc-component of type 1 corresponds to an $I'$-arc-component of type 1.

It suffices to show that when passing from the complex $X = K(I')$ to $\tilde{X}$, the number of $I'$-arc-components of type 1 equals the number $M$ of complete $X$-components.

**Claim:** there exists a bijection between the set of $I'$-arc-components of type 1 and that of complete components of $\tilde{X}$.

Given an $I'$-arc-component of type 1, $A$, let $C$ denote the set of couples $(\Delta, \Delta')$ associated with arcs of $A$. In view of Lemma 10, any 1-simplex that appears as a face in $C$-couples is shared by at least two 2-simplices. As a result, $C$ does not contain any free 1-simplex and passing from $X$ to $\tilde{X}$, simplicial collapses do not affect $C$-2-simplices.

Consequently, all $C$-couples in $I'$ simply persist when passing to $\tilde{X}$, where $C$ itself becomes a $\tilde{X}$-component. Lemma 10 further guarantees that $C$ does not contain any free 1-face, whence Lemma 8 implies that $C$ is complete.

Accordingly, the Ansatz $\Phi(A) = C$ produces a well-defined mapping between $I'$ arc-components of type 1 and complete $\tilde{X}$-components.

$\Phi$ is, by construction, injective since the mapping constitutes a mere reinterpretation of an arc-component of type 1.

In order to establish surjectivity, let $C_1$ be a complete $\tilde{X}$-component. $C_1$ induces a distinguished, unique set of $I'$-arcs, $A_1$. By construction, any two $A_1$-arcs $a_1, a_2 \in A_1$ satisfy $a_1 \sim_\phi a_2$, whence $A_1$ is contained in a nontrivial and distinguished arc-component $A'$.

We proceed by proving that $A_1 = A'$ is an arc-component of type 1. Suppose $A_1 \subsetneq A'$, then there exist an arc $(i, i') \in A' \setminus A_1$ and $(j, j') \in A_1$ such that $(i, i')$ and $(j, j')$ are $\phi$-crossing. Note that $(j, j')$ is associated to a $C_1$-couple $(\Delta_j, \Delta_j')$. Let $\{s, s'\}$ be the pure edge associated with $(i, i')$. We showed in the proof of Lemma 7, that $S_{C_1}$ is connected.

Since $(i, i')$ and $(j, j')$ are $\phi$-crossing, the $S$-loops $s_1, s_2$ are vertices of $\Delta_j, \Delta_j'$ from which follows that both $s$ and $s'$ are on a path of two $S$-loops in $C_1$. Accordingly $\{s, s'\}$ is contained in $S_{C_1}$. Lemma 7 guarantees that $C_1$ is a sphere, whence there exists a $C_1$-couple $(\Delta_0, \Delta_0')$ for which $\{s, s'\}$ is a pure face. As a result we obtain $(i, i') \in A_1$, which is a contradiction from which $A' = A_1$ follows.

By Lemma 2, both endpoints of an $A_1$-arc are incident to $\phi$-arcs, whence $A_1$ is an arc-component of type 1 and $\Phi(A_1) = C_1$, whence $\Phi$ is surjective and, thus, bijective.

Therefore $\chi = M$ and $H_2(I) \cong \oplus_{k=1}^\gamma \mathbb{Z}.

The Euler characteristic $\chi$ of $K(I)$ can be expressed as the alternating sum of ranks of its homology groups, i.e., $\chi = \text{rk}(H_0(I)) - \text{rk}(H_1(I)) + \text{rk}(H_2(I)) - \text{rk}(H_3(I))$, whence $\text{rk}(H_1(I)) = \gamma - \chi + 1$. Theorem 2 shows that $H_1(I) \cong H_1(\tilde{X})$ is free, from which we can conclude $H_1(I) \cong \oplus_{k=1}^\gamma \mathbb{Z}$. \qed

8. Discussion

In this paper we computed the simplicial homology of $\tau$-structures. $\tau$-structures represent a meaningful generalization of bi-structures \cite{7, 8} as they allow us to study transitions between sequence-structure pairs where the underlying sequences differ in
specific manners. Bi-structures would only facilitate the analysis of such pairs, where the underlying sequences are equal. Intuitively, the free energy of such sequence–structure pairs, which is a quantity that is straightforward to compute, characterizes the probability of finding such a pair of sequences in the course of evolution.

We are currently extending the results of this paper by computing the weighted homology of \( \tau \)-structures over a discrete valuation ring, \( R \). This entails the study of weighted complexes in which simplices are endowed with weights \([9,24,25]\). These weighted complexes feature a new boundary map, \( \partial^n_v : C_n(R)(X) \to C_{n-1}(R)(X) \), where \( C_n(R)(X) \) denotes the free \( R \)-module generated by all \( n \)-simplices contained in \( X \). \( \partial^n_v \) is given by the following.

\[
\partial^n_v(\sigma) = \sum_{i=0}^{n} \frac{v(\hat{\sigma}_i)}{v(\sigma)} \cdot (-1)^i \hat{\sigma}_i.
\]

As it is the case for bi-structures \([9]\), the weighted homology plays a crucial role in the Boltzmann sampling of sequence–structure pairs that minimize the free energy of the \( \tau \)-structure.

Our approach differs from the purely algebraic proofs for the simplicial loop homology of bi-structures \([8]\). In the case of bi-structures, the computation for the second homology employs the fact that their first homology group is trivial. This allows us to understand the second homology group via a long exact sequence of relative homology groups. However, for \( \tau \)-structures, the first homology is, in general, nontrivial and therefore requires a different approach.

We first reduce the loop complex \( K(I) \) to the sub-complex \( \tilde{X} \) via simplicial collapses, retaining the homology of the original space. We then dissect a certain 1-dimensional sub-complex in Proposition 9 and then decompose \( \tilde{X} \) into components based on the \( \phi \)-crossing of couples. This decomposition allows us to identify different generators of the homology, with incomplete and complete components contributing to the first and second homologies, respectively. Finally, we compute the homology of \( \tilde{X} \) by gluing components in a particular manner. This makes use of the planarity of \( \tau \)-structures in Lemma 6 and assures that we encounter particularly simple intersections, when applying the Mayer–Vietoris sequence. It is worth pointing out that in the proof of Theorem 2 the particular ordering in which the components are glued is crucial.

In view of Proposition 9, Theorems 1 and 2, the generators of the first homology of the loop complex originate from the combining of the sub-complexes and the gluing of different components. The detailed descriptions of all these generators are a work in progress and here we restrict ourselves to using the Euler characteristic to express the rank of the first homology group. Only complete crossing components contribute to the second homology and their geometric realizations are spheres. We provide a combinatorial characterization of the generators in \( I \) in terms of arc-components of type 1.

As for applications of this framework, we currently employ \( \tau \)-structures to investigate the evolutionary trajectories of viruses, such as flus and the Coronavirus. Specifically, we compute the loop homology of evolutionary transitions to gain deeper insight into sequence-structure-function relationships of the virus.

**Author Contributions:** Conceptualization, T.J.X.L. and C.M.R.; Investigation, T.J.X.L. and C.M.R.; Writing—original draft, T.J.X.L and C.M.R. Both authors have contributed equally to this work. Both authors have read and agreed to the possible publication of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** main text.

**Informed Consent Statement:** main text.

**Data Availability Statement:** main text.

**Acknowledgments:** We gratefully acknowledge the comments and discussions from Andrei Bura, Qijun He and Fenix Huang.
Conflicts of Interest: The authors declare no conflicts of interest.

References
1. Waterman, M. Secondary Structure of Single-Stranded Nucleic Acids. In Studies on Foundations and Combinatorics, Advances in Mathematics Supplementary Studies; Rota, G.C., Ed.; Academic Press: New York, NY, USA, 1978; Volume 1, pp. 167–212.
2. Smith, T.F.; Waterman, M.S. RNA secondary structure. *Math. Biol.* **1978**, *42*, 31–49.
3. Zuker, M.; Stiegler, P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* **1981**, *9*, 133–148.
4. Eddy, S.R. Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* **2001**, *2*, 919–929.
5. Breaker, R.R. Riboswitches and the RNA World. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, a003566.
6. Huang, F.W.D.; Barrett, C.; Reidys, C.M. The energy-spectrum of bicompatible sequences. *arXiv: 2021*, arXiv:1910.00190v1.
7. Bura, A.C.; He, Q.; Reidys, C.M. Loop homology of bi-secondary structures. *Discret. Math.* **2021**, *344*, 112371.
8. Bura, A.C.; He, Q.; Reidys, C.M. Loop Homology of Bi-secondary Structures II. *arXiv 2019*, arXiv:1909.01222.
9. Bura, A.; He, Q.; Reidys, C. Weighted Homology of Bi-Structures over Certain Discrete Valuation Rings. *Mathematics 2021*, *9*, 744.
10. Waterman, M. Combinatorics of RNA Hairpins and Cloverleaves. *Stud. Appl. Math.* **1979**, *60*, 317–323.
11. Penner, R.; Waterman, M. Linear trees and RNA secondary structure. *Adv. Math.* **1993**, *217*, 31–49.