Local Cohomology and Stratification

Vidit Nanda

Abstract. We outline an algorithm to recover the canonical (or, coarsest) stratification of a given regular CW complex into cohomology manifolds, each of which is a union of cells. The construction proceeds by iteratively localizing the poset of cells about a family of subposets; these subposets are in turn determined by a collection of cosheaves which capture variations in cohomology of cellular neighborhoods across the underlying complex. The result is a finite sequence of categories whose colimit recovers the canonical strata via (isomorphism classes of) its objects. The entire process is amenable to efficient distributed computation.

1. Introduction

Setting aside technicalities for the moment, an $n$-dimensional stratification of a given topological space $X$ is a nested sequence of closed subspaces

$$\varnothing = X_{-1} \subset X_0 \subset \cdots \subset X_n = X$$

where each successive difference $X_d - X_{d-1}$ resembles a (possibly empty) $d$-dimensional manifold whose connected components are the $d$-dimensional strata. The suppressed technicalities guarantee uniformity of sufficiently small neighborhoods in $X$ around (points in) every stratum. Stratified spaces are of fundamental importance in any branch of mathematics which seriously confronts singularities. Every topological manifold admits a straightforward stratification into its connected components; less trivially, the following spaces all admit, and are often analyzed using, stratifications: finite CW complexes, quotients of properly discontinuous Lie group actions on smooth manifolds, algebraic varieties, semialgebraic and subanalytic sets.

The 2-dimensional singular space below (let’s call it $Y$) is built by pinching a torus along a meridian and attaching a disk across an equator:

Any regular CW structure is a stratification of $Y$ by skeleta, where $d$-dimensional strata are precisely the $d$-cells; passing to a subdivision further refines this skeletal stratification in the sense that every new cell is entirely contained in the interior of an old
cell. On the other hand, the coarsest stratification of $Y$ into manifolds requires only one 0-stratum, one 1-stratum and two 2-strata:

- \[ \_ + \_ + \_ \]

Our main result here involves algorithmically recovering the coarsest stratification of a finite regular CW complex where all strata are unions of cells. This is called the **canonical stratification** of the complex, and its existence and uniqueness play a central role in Goresky and MacPherson’s proof of the topological invariance of intersection homology for stratified pseudomanifolds [14, Sec 4]. Our argument, like theirs, has an intuitive geometric core but invokes algebraic and categorical machinery. For the purposes of this introductory section, we focus on geometry and ask: given a finite cellulation of $Y$, how might one identify the canonical strata and determine which cells lie in each canonical stratum? Attempts to formulate answers highlight the hazards of investigating local topology of small neighborhoods around cells in the 0, 1 and 2-dimensional canonical strata of $Y$:

- \[ \_ \]
- \[ \_ \]
- \[ \_ \]

Each neighborhood above deformation-retracts to its central cell and is therefore contractible; however, compactifying these three classes of neighborhoods by collapsing their boundaries to points creates new stratified spaces with potentially interesting topology. The compactified neighborhood around the vertex in the canonical 0-stratum of $Y$ is (homeomorphic to) two 2-spheres joined at their north and south poles with a spanning disk across the middle. The compactified neighborhood of any cell in the canonical 1-stratum resembles a 2-sphere whose interior has been partitioned in two by a disk attached along a great circle passing through the north and south poles. And finally, the compactified neighborhood around any cell in either of the two canonical 2-strata is homeomorphic to a 2-sphere:

- \[ \_ \]
- \[ \_ \]
- \[ \_ \]

One hopes to recover intrinsic strata by clustering cells whenever they exhibit similar compactified neighborhoods. But already in this example, we encounter two significant difficulties: first, the compactified neighborhoods do not distinguish cells in the two 2-strata from each other. The second difficulty is somewhat subtler — although the compactified neighborhoods of cells in the 0-stratum and 1-stratum are not homeomorphic, both are homotopy-equivalent to a wedge of two 2-spheres. Therefore, they can not be distinguished by weaker, more computable topological invariants such as cohomology. We tackle the first problem by constructing a **complex of cosheaves** which encodes
how local topology varies across cells, and we bypass the second problem by working exclusively in the category of cohomologically stratified spaces.

Consider the task of determining the canonical stratification of a finite regular CW complex $X$ into $R$-cohomology manifolds, where $R$ is a fixed nontrivial commutative ring with unity. In this case, our complex of cosheaves assumes a particularly appealing form: it is a functor

$$L^\bullet : Fc(X) \rightarrow \mathbf{Ch}(R)$$

from the poset of cells in $X$ (where $x > y$ denotes that $y$ is a face of $x$, or equivalently, that $x$ is a co-face of $y$) to the category of lower-bounded $R$-module cochain complexes. To each cell $x$ of $X$, it assigns a cochain complex $L^\bullet(x)$:

$$L^0(x) \xrightarrow{\beta^0_x} L^1(x) \xrightarrow{\beta^1_x} L^2(x) \xrightarrow{\beta^2_x} \cdots,$$

where $L^d(x)$ is the free $R$-module generated by all $d$-dimensional co-faces of $x$ — in particular, the module $L^d(x)$ is trivial for $d < \dim x$ and $d > \dim X$. The co-differentials $\beta^\bullet_x$ are inherited verbatim from incidence degrees among cells of $X$ taking values in $R$. Thus, $L^\bullet(x)$ computes the reduced cohomology of $x$’s compactified neighborhood in $X$. Given another cell $y$ satisfying $x \geq y$, there is an inclusion map

$$L^\bullet(x \geq y) : L^\bullet(x) \hookrightarrow L^\bullet(y)$$

arising from the fact that co-faces of $x$ are also co-faces of $y$. Here is a simple version of our main result.

**Theorem 1.1.** There is a category $S$ obtained from $Fc(X)$ by formally inverting a particular subset of those face relations $x \geq y$ for which $L^\bullet(x \geq y)$ induces isomorphisms on cohomology; two cells lie in the same canonical stratum of $X$ if and only if they are isomorphic in $S$.

Before proceeding to the $L^\bullet$-induced definition of $S$ and a proof of this theorem, we highlight three salient features. First, the construction of $S$ is algorithmic and isomorphism classes of its objects are explicitly (and efficiently) computable. Second, the non-isomorphisms in $S$ reveal when a canonical stratum contains another in its boundary. And third, our construction actually yields a nested sequence of intermediate categories

$$S_0 \hookrightarrow S_1 \hookrightarrow \cdots \hookrightarrow S_{\dim X} = S,$$

each of which receives a functor from $Fc(X)$ compatible with the previous ones; and all the canonical strata of dimension exceeding $(\dim X - d)$ may be recovered from certain isomorphism classes in $S_d$.

This paper is organized as follows. Section 2 tersely collects relevant background material involving stratifications and cosheaves, while Section 3 describes the local cohomology complex of cosheaves $L^\bullet$ associated to a regular CW complex $X$. In Section 4 we use $L^\bullet$ to define an initial functor $Fc(X) \rightarrow S_0$ via categorical localization, and prove that the top-dimensional canonical strata of $X$ correspond bijectively with isomorphism classes of (images of) top-cells in $S_0$. Section 5 contains the heart of our argument: it concludes the proof of Theorem 1.1 by describing the inductive construction of $Fc(X) \rightarrow S_d$ from the preceding $Fc(X) \rightarrow S_{d-1}$. The final Section 6 may be read independently of the preceding ones — it is devoted to computational matters, and includes an efficient distributed algorithm for recovering canonical strata in practice.
Context. Our goal is to provide a principled topological pre-processor for various manifold learning algorithms. The fundamental aim of manifold learning is to automatically infer the intrinsic dimension of a compact submanifold \( M \) (often with boundary) of Euclidean space from a finite point sample \( P \subset \mathbb{R}^n \) lying on, or at least near, \( M \). An enormous amount of work has been done in this field, and we will make no attempt to even summarize it here — the curious reader is referred to [20] and its myriad citations. In most cases, one constructs a graph around the points in \( P \) by inserting all edges of Euclidean length bounded above by a parameter threshold \( \epsilon > 0 \). Shortest paths along this graph then serve as proxies for intrinsic geodesics of \( M \) and are used to find distance-preserving projections to lower-dimensional subspaces of \( \mathbb{R}^n \) via spectral methods.

When \( M \) is not a single manifold, but rather a union of different manifolds across several dimensions, our standard manifold learning methods necessarily result in over-fitting along regions where \( M \) has low intrinsic dimension. Theorem 1.1 provides a remedy by conferring the ability to partition \( P \) into distinct clusters, each of which (at the given scale \( \epsilon \)) is guaranteed to lie on a cohomology manifold of known dimension. These clusters may then be independently subjected to the usual manifold learning techniques. Perhaps more important from a practical perspective is the fact that the computations involved in partitioning \( P \) are easily distributed across several processors. And as a straightforward byproduct, one also gains the capacity to identify (those points of \( P \) which lie in) every anomalous, singular region of \( M \).

There now exists a substantial body of work where filtered (co)homology groups of cell complexes built around data points play a central role [12, 24]. With the knowledge of canonical strata of such cell complexes comes the prospect of efficiently computing far more refined topological invariants such as intersection cohomology [13] groups. Also accessible, thanks to suitable flavors of Morse theory [23, 25], is the ability to cluster canonical \( d \)-strata by \((d + 1)\)-dimensional cobordisms internal to the ambient dataset.

Acknowledgments. I am grateful to Justin Curry (who pointed me to Definition 3.1), Rob Ghrist and Ulrike Tillmann for insightful discussions. This work was supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1. It is dedicated to Mark Goresky and Bob MacPherson, whose efforts over four decades have built some of my favorite playgrounds.

2. Preliminaries

Our primary references for stratification theory, categorical localization and cellular (co)sheaves are [15, 19, 32], [8, 10] and [5, 28] respectively. The interplay between cell structures and general stratifications on the same underlying space has been explored thoroughly in [29]. Here we will be concerned entirely with cohomologically stratified spaces, which the reader may have encountered before in [14, Sec 4.1], [11, Sec 3.3], or [27, Sec 5].

2.1. Cohomological stratifications. Fix a non-trivial commutative ring \( R \) with identity and recall that an \( R \)-cohomology manifold of dimension \( d \geq 0 \) is a locally compact Hausdorff topological space \( Z \) where each point \( z \) has an open neighborhood \( U_z \subset Z \) whose compactly-supported \( R \)-valued cohomology agrees with that of \( d \)-dimensional
Euclidean space $\mathbb{R}^d$, i.e.,

$$H^\bullet_c(U_z; R) \simeq \begin{cases} R & \text{if } \bullet = d, \\ 0 & \text{otherwise}. \end{cases} \quad (1)$$

Readers may safely replace the compactly supported cohomology of $U_z$ (and any open set encountered henceforth) by the standard cohomology of its one-point compactification relative to the additional point.

**Definition 2.1.** An $n$-dimensional (cohomological) stratification of a locally compact Hausdorff topological space $X$ is an ascending sequence of closed subspaces

$$\emptyset = X_{-1} \subset X_0 \subset \cdots \subset X_{n-1} \subset X_n = X,$$

where the connected components of $X_d - X_{d-1}$, called $d$-dimensional strata, are required to obey the following axioms.

- **Frontier:** if a stratum $\sigma$ intersects the closure of another stratum $\tau$, then in fact $\sigma$ is completely contained in the closure of $\tau$ and $\dim \tau \geq \dim \sigma$ (with equality occurring if and only if $\tau = \sigma$). This relation, denoted by $\tau \succeq \sigma$ henceforth, forms a locally finite partial order on the set of all strata.

- **Link:** for each $d$-stratum $\sigma$, there exists an $(n - d - 1)$-dimensional stratified space $L = L(\sigma)$, called the link of $\sigma$:

$$\emptyset = L_{-1} \subset L_0 \subset \cdots \subset L_{n-d-1} = L,$$

so that every open neighborhood around a point $p$ in $\sigma$ admits a basic open subneighborhood $U_p \subset X$ with the following structure. The intersections $U_p \cap X_i$ are empty for $i < d$, while for $d \leq i \leq n$ there are stratified quasi-isomorphisms of compactly-supported singular cochain complexes (of $R$-modules):

$$C^\bullet_c((\mathcal{C}L_{i-d-1} \times \mathbb{R}^d) \simeq C^\bullet_c(U_p \cap X_i), \quad (2)$$

where $\mathcal{C}L_\bullet$ denotes the open cone\(^1\) on $L_\bullet$.

By quasi-isomorphism here (and elsewhere) we mean a cochain map which induces isomorphisms on all cohomology groups. The additional requirement that the maps from (2) be stratified guarantees that they behave naturally in three separate ways: they should be filtered, stratum-preserving and refinable as explained below.

1. They should respect (contravariant maps induced by) the inclusions $L_\bullet \subset L_{\bullet+1}$ in their domains and $X_\bullet \subset X_{\bullet+1}$ their codomains so that the following squares commute:

$$\begin{array}{ccc}
C^\bullet_c((\mathcal{C}L_{i-d-1} \times \mathbb{R}^d) & \simeq & C^\bullet_c(U_p \cap X_i) \\
\uparrow & & \uparrow \\
C^\bullet_c((\mathcal{C}L_{i-d-1} \times \mathbb{R}^d) & \simeq & C^\bullet_c(U_p \cap X_{i+1})
\end{array}$$

\(^1\)The open cone on $Z$ is the quotient of $Z \times [0,1)$ by the relation which identifies $(z, 0)$ with $(z', 0)$ for any $z$ and $z'$ in $Z$; by convention, the open cone on the empty set is the one-point space.
They should preserve strata in the sense that there exist bijections 
\[ \Phi_i : \{i - \text{strata } \tau \succeq \sigma \text{ of } X\} \xrightarrow{\cong} \{(i - d - 1) - \text{strata of } L\} \]
so that the maps from (2) restrict to quasi-isomorphisms 
\[ C^\bullet_c(\mathcal{C} \Phi_i(\tau) \times \mathbb{R}^d) \xrightarrow{\cong} C^\bullet_c(U_p \cap \tau). \]

(3) They should refine to smaller basic neighborhoods \( V_p \) so that the following triangles of quasi-isomorphisms commute:
\[
\begin{align*}
C^\bullet_c(\mathcal{C} L_{i-d-1} \times \mathbb{R}^d) &\xrightarrow{\cong} C^\bullet_c(U_p \cap X_i) \\
\downarrow \cong & \downarrow \cong \\
C^\bullet_c(V_p \cap X_i)
\end{align*}
\]
(here the vertical map is induced by the inclusion \( V_p \subset U_p \)).

Remark 2.2. The following observations concern Definition 2.1.

(1) By a theorem of Eilenberg and Zilber [9], links are well-defined (only) up to filtered quasi-isomorphism type: we may replace \( L \) by a filtered cochain complex by substituting a tensor product for the left side of (2).

(2) The link axiom (when \( i = d \)) guarantees that each \( d \)-stratum is indeed a \( d \)-dimensional \( \mathbb{R} \)-cohomology manifold.

(3) This definition, unlike similar notions which typically appear in the intersection homology literature, does not require the union of top strata to be dense in \( X \). In other words, \( X_{n-1} \) need not equal \( X_{n-2} \).

We call one stratification a **coarsening** of another whenever each stratum of the former is a union of strata of the latter. All stratifications encountered henceforth will be coarsenings of the **skeletal** stratification of a fixed finite regular\(^2\) CW complex \( X \). The skeletal stratification is defined as follows: writing the face partial order among cells as \( \geq \), we recall that each cell \( y \) of \( X \) has

- an open star \( \text{st}(y) \) containing all cells \( x \) which satisfy \( x \succeq y \), and
- a link \( \text{lk}(y) \), containing all cells \( x \) that share a co-face but no face with \( y \).

The \( d \)-dimensional skeletal strata are the \( d \)-cells of \( X \), and the frontier partial order coincides with \( \geq \). Since \( \text{st}(y) \) of a \( d \)-cell \( y \) is simultaneously homeomorphic to all sufficiently small neighborhoods around points in \( y \) and to \( \mathcal{C} \text{lk}(y) \times \mathbb{R}^d \), the link \( L(y) \) in the sense of Definition 2.1 is precisely \( \text{lk}(y) \).

Definition 2.3. The **canonical** stratification of a regular CW complex \( X \) is the smallest coarsening of the cellular stratification. In other words, it is the coarsest stratification of \( X \) whose strata are all unions of cells.

The existence and uniqueness of canonical stratifications was established (in greater generality than ours) by Goresky and MacPherson [14, Sec 4]. Here we only identify a simple consequence of the frontier axiom for cells lying in top strata.

Proposition 2.4. Let \( X \) be a finite regular CW complex of dimension \( n \) equipped with any stratification coarser than its skeletal stratification. If a cell \( y \) of \( X \) lies in an \( n \)-dimensional stratum \( \sigma \), then so must every cell \( x \) which satisfies \( x \succeq y \).

\(^2\)Regularity in this context means that the characteristic map of every cell is an embedding.
Proof. Let \( \tau \) be the unique coarse stratum containing \( x \). Since \( y \) lies in the boundary of \( x \) by assumption, the closure of \( \tau \) intersects \( \sigma \) non-trivially at \( y \). The desired conclusion now follows from the axiom of the frontier in Definition 2.1 and the fact that there are no strata of dimension exceeding \( n \), since \( \dim \tau \geq \dim \sigma = n \) forces \( \tau = \sigma \). \( \square \)

Thus, membership of cells in top-dimensional canonical strata is upward-closed with respect to the face partial order.

2.2. Localizations of posets. The localization of a poset \( P \) about a subcollection \( \Sigma \) of its order relations is the minimal category containing \( P \) in which the relations of \( \Sigma \) have been rendered invertible \cite[Ch I.1]{10}.

Definition 2.5. Let \( (P, \geq) \) be a poset and \( \Sigma = \{(x \geq y)\} \) a subset of its relations which is closed\(^3\) under composition. The localization of \( P \) about \( \Sigma \) is a category \( P[\Sigma^{-1}] \) whose objects are precisely the elements of \( P \), while morphisms from \( p \) to \( q \) are given by equivalence classes of finite (but arbitrarily long) \( \Sigma \)-zigzags of order relations in \( P \)

\[
p \geq y_0 \leq x_0 \geq \cdots \geq y_k \leq x_k \geq q,
\]

where:

1. only relations in \( \Sigma \) and equalities can point backwards (i.e., \( \leq \)),
2. composition is given by concatenation, and
3. the trivial zigzag \( p = p \) represents the identity of each element \( p \).

The equivalence between \( \Sigma \)-zigzags is generated by the transitive closure of the following basic relations. Two such zigzags are related

- horizontally if one is obtained from the other by excision of equalities:

\[
(\cdots \leq x \geq y = y \geq x' \leq \cdots) \sim (\cdots \leq x \geq x' \leq \cdots),
\]

\[
(\cdots \geq y \leq x = x \leq y' \geq \cdots) \sim (\cdots \geq y \leq y' \geq \cdots),
\]

- or vertically, if they form the rows in a grid:

\[
\begin{array}{cccccc}
p & \geq & y_0 & \leq & x_0 & \geq \cdots \geq x_k & \geq q \\
\| & \| & \| & \| & \| & \| & \|
\end{array}
\]

\[
\begin{array}{cccccc}
p & \geq & y_0' & \leq & x_0' & \geq \cdots \geq x_k' & \geq q \\
\| & \| & \| & \| & \| & \|
\end{array}
\]

where all vertical face relations (also) lie in \( \Sigma \).

The category \( P[\Sigma^{-1}] \) enjoys the following universal property \cite[Lem I.1.3]{10} — there is a canonical localization functor \( P \to P[\Sigma^{-1}] \) which acts by inclusion. In particular, it sends each element of \( P \) to itself and each relation \( x \geq y \) to its own equivalence class of \( \Sigma \)-zigzags. Moreover, any functor \( P \to C \) which maps order relations in \( \Sigma \) to isomorphisms in the target category \( C \) factorizes uniquely across the localization functor.

\(^3\)That is, if \((p \geq q)\) and \((q \geq r)\) both lie in \( \Sigma \), then so does \((p \geq r)\).
2.3. Cellular cosheaves. Sheaves (and their dual cosheaves) assign algebraic objects to open sets of topological spaces [2]. Our interest here is in a particularly tame and computable sub-class of cosheaves, where the base space is always a finite regular regular CW complex $X$ and the assigned algebraic objects are cell-wise constant. We write $Fc(X)$ to indicate the poset of cells with order relation $x > y$ denoting that $y$ is a face of $x$.

Definition 2.6. A cellular cosheaf over $X$ taking values in an Abelian category $A$ is a functor $F : Fc(X) \rightarrow A$. Thus, it assigns to each cell $x$ an $A$-object $F(x)$ and to each face relation $x \geq y$ an $A$-morphism $F(x \geq y) : F(x) \rightarrow F(y)$ so that
- $F(x = x)$ is the identity on $F(x)$ for each cell $x$, and
- $F(y \geq z) \circ F(x \geq y)$ equals $F(x \geq z)$ across any triple of cells $x \geq y \geq z$.

We call $F(x)$ the stalk of $F$ at $x$, and call $F(x \geq y)$ the extension map of $F$ at $x \geq y$.

A morphism of cellular cosheaves $\eta : F \rightarrow G$ over $X$ is a natural transformation; it assigns to each cell $x$ an $A$-morphism $\eta_x : F(x) \rightarrow G(x)$ so that for each $x \geq y$ the relevant square in $A$ commutes:

\[
\begin{array}{ccc}
  F(x) & \xrightarrow{F(x \geq y)} & F(y) \\
  \downarrow{\eta_x} & & \downarrow{\eta_y} \\
  G(x) & \xrightarrow{G(x \geq y)} & G(y)
\end{array}
\]

Cosheaf morphisms may be composed stalkwise, and there is always a zero morphism $0 : F \rightarrow G$ which (as one might expect) assigns the zero map $0_x : F(x) \rightarrow G(x)$ in $A$ to each cell $x$ of $X$.

Definition 2.7. A (lower-bounded) complex of cellular cosheaves $F^\bullet$ (over $X$, taking values in $A$) is a sequence of $A$-valued cellular cosheaves over $X$ connected by cosheaf morphisms:

\[
F^0 \xrightarrow{\eta^0} F^1 \xrightarrow{\eta^1} F^2 \xrightarrow{\eta^2} \cdots
\]

so that every successive composition $\eta^{\bullet+1} \circ \eta^\bullet$ yields the zero morphism.

Every complex of cosheaves $F^\bullet$ on $X$ may be reinterpreted as a single cosheaf which takes values in the category $\text{Ch}(A)$ of lower-bounded cochain complexes in $A$ — consider the collection of commuting squares in $A$ that are assigned to face relations $x \geq y \geq z \geq \cdots$ of $X$:

\[
\begin{array}{ccccccc}
  F^0(x) & \xrightarrow{\cdots} & F^1(x) & \xrightarrow{\cdots} & F^2(x) & \xrightarrow{\cdots} \\
  \downarrow & & \downarrow & & \downarrow & & \cdots \\
  F^0(y) & \xrightarrow{\cdots} & F^1(y) & \xrightarrow{\cdots} & F^2(y) & \xrightarrow{\cdots} \\
  \downarrow & & \downarrow & & \downarrow & & \cdots \\
  F^0(z) & \xrightarrow{\cdots} & F^1(z) & \xrightarrow{\cdots} & F^2(z) & \xrightarrow{\cdots} \\
  \downarrow & & \downarrow & & \downarrow & & \\
  \vdots & & \vdots & & \vdots & & \ddots
\end{array}
\]
The horizontal slices of this diagram confirm that a cochain complex in $A$ is allocated to each cell, while vertical arrows between these slices prescribe the desired extension maps. Taking cohomology horizontally, one therefore obtains an ordinary cellular cosheaf of (graded) $A$-objects over $X$. We call this the cohomology of $F^*$ and denote it by $H^*F$.

3. Local cohomology of CW complexes

Let $X$ be a finite regular CW complex with face poset $Fc(X)$ and $R$ a fixed nontrivial commutative ring with identity $1_R$. We write $\text{Mod}(R)$ to denote the category of $R$-modules, and $\text{Ch}(R)$ – rather than the cumbersome $\text{Ch}(\text{Mod}(R))$ – to indicate the category of cochain complexes of $R$-modules indexed by the non-negative natural numbers.

**Definition 3.1.** [6, Def 7.1] The local cohomology $L^*$ of $X$ is a complex of cosheaves

$$L^0 \xrightarrow{\beta^0} L^1 \xrightarrow{\beta^1} L^2 \xrightarrow{\beta^2} \cdots$$

over $X$ taking values in $\text{Mod}(R)$, prescribed by the following data.

1. For each dimension $d \geq 0$ and cell $x$ of $X$, the cosheaf $L^d$ has as its stalk $L^d(x)$ the free $R$-module with basis

$$\{z \in Fc(X) \mid z \geq x \text{ and } \dim z = d\}.$$

When $x \geq y$, the extension $L^d(x \geq y) : L^d(x) \rightarrow L^d(y)$ is determined by the obvious inclusion of basis cells.

2. The cosheaf morphism $\beta^d$ assigns to each cell $x$ the map $\beta^d_x : L^d(x) \rightarrow L^{d+1}(x)$ defined by (linearly extending) the following action on basis cells. For each $d$-cell $z \geq x$, we have

$$\beta^d_x(z) = \sum_w \langle \partial w, z \rangle_R \cdot w,$$

where the sum is taken over all $(d + 1)$-cells $w \geq x$, and $\langle w, z \rangle_R \in \{0, \pm 1\}$ is the $R$-valued degree of the attaching map in $X$ from the boundary of $w$ onto $z$.

Routine calculations confirm that $\beta^{d+1} \circ \beta^d$ is always zero, and that the $\beta_x^*$’s commute with all relevant extension maps. In light of the discussion following Definition 2.7, we will shift perspective and regard $L^*$ as a single cosheaf valued in $\text{Ch}(R)$. In this setting, the stalk $L^*(x)$ lying over each cell $x$ is the entire cochain complex

$$L^0(x) \xrightarrow{\beta^0_x} L^1(x) \xrightarrow{\beta^1_x} L^2(x) \xrightarrow{\beta^2_x} \cdots,$$

and $H^*L(x)$ is the compactly-supported cohomology of $x$’s open star in $X$. Equivalently, one may describe $H^*L(x)$ as the ordinary relative cohomology of the pair $(\text{st}(x), \partial \text{st}(x))$ where $\text{st}(x)$ is the closure of $x$’s open star in $X$ and $\partial$ denotes the topological boundary:

$$\partial \text{st}(x) = \text{st}(x) - \text{st}(x).$$

When $y$ is a zero-dimensional cell, we have $\partial \text{st}(y) = \text{lk}(y)$.

For each $d \geq 0$, we write $R[d]^*$ to indicate the special cochain complex in $\text{Ch}(R)$ which is trivial except for a single copy of $R$ in the $d$-th position:

$$0 \rightarrow \cdots \rightarrow 0 \rightarrow R \rightarrow 0 \rightarrow \cdots.$$
Thus, $H^\bullet R[d]$ is the cohomology of small neighborhoods in $d$-dimensional $R$-cohomology manifolds, as in (1). By Definition 2.1, if a cell $x$ lies in some top-dimensional stratum of $X$, then we must have an isomorphism $H^\bullet L(x) \simeq H^\bullet R[\dim X]$. For cells of high dimension, this requirement has strong consequences.

**Proposition 3.2.** If $n = \dim X$, then every $n$-cell $x$ has $L^\bullet (x) = R[n]^\bullet$. And moreover, every $(n - 1)$-cell $y$ with $H^\bullet L(y) \simeq H^\bullet R[n]$ lies in the boundary of exactly two $n$-cells.

**Proof.** The first assertion follows from Definition 3.1 and the fact that there are no cells of dimension $\geq (n + 1)$ in $X$. Turning to the second assumption, note that the local cohomology of an $(n - 1)$-cell $y$ is only supported in dimensions $(n - 1)$ and $n$:

$$L^\bullet (y) = 0 \to \cdots \to 0 \to R^\beta_{y - 1} \to R^k \to 0 \to \cdots$$

where $k$ counts the number of $n$-cells $x$ satisfying $x > y$. If $L^\bullet (y)$ is quasi-isomorphic to $R[n]^\bullet$, then $\beta_{y - 1}$ must be injective (since $H^\bullet L(y)$ is trivial), so $k > 0$. By definition, $\beta_{y - 1}$ acts by sending a basis element of $R$ to a linear combination of basis elements in $R^k$, and by the regularity of $X$, all coefficients in this linear combination lie in $\{ \pm 1_R \}$. The $n$-th cohomology of $L^\bullet (y)$ is the quotient

$$H^n L(y) = \frac{R^k}{\text{im}(\beta_{y - 1})},$$

and it must be isomorphic to $R = H^n R[n]$ by assumption. Since $R$ is a commutative ring, it satisfies the invariant basis property and we have $k = 2$. □

In light of the preceding result and (1), one might hope to identify top-dimensional canonical strata by simply clustering together adjacent cells $x \geq y$ in $X$ whenever there exists a top-dimensional cell $w$ satisfying $w \geq x$ so that all extension maps below are quasi-isomorphisms:

$$R[\dim X]^\bullet = L^\bullet (w) \to L^\bullet (x) \to L^\bullet (y).$$

Our next example shows that this (necessary) condition for $x$ and $y$ to belong to the same top stratum is not sufficient when $\dim w - \dim y > 1$.

**Example 3.3.** Consider the 2-dimensional simplicial complex depicted below: it is obtained by subdividing a parallelogram into four 2-simplices along diagonals, and attaching an extra 2-simplex along one of the four resulting half-diagonals.

Let $w, x$ and $y$ respectively denote the extra 2-simplex, the shared half-diagonal, and the central vertex. It is easy to confirm that $L^\bullet (y)$ is quasi-isomorphic to $R[2]^\bullet$ since its link has the cohomology of a circle. And moreover, given any sequence $w' > x' > y$ of cells where $w'$ and $x'$ differ from $w$ and $x$ respectively, the inclusion maps

$$R[2]^\bullet = L^\bullet (w') \to L^\bullet (x') \to L^\bullet (y)$$
are all quasi-isomorphisms. Note, however, that $L^\bullet(w)$ is not quasi-isomorphic to $R[2]^\bullet$ by Proposition 3.2 since there are three 2-simplices containing $x$ as a face. Thus, $x$ does not belong to any canonical 2-stratum, and hence (by Proposition 2.4), neither does $y$.

4. Extracting top strata

Throughout this section, $X$ denotes a finite $n$-dimensional regular CW complex, $R$ is a fixed commutative ring and $L^\bullet : Fc(X) \to \text{Ch}(R)$ is the local cohomology complex of $X$ from Definition 3.1. Consider the collection of face relations $(x \geq y)$ in $Fc(X)$ sent by $L^\bullet$ to quasi-isomorphisms, i.e.,

$$E_0 = \{(x \geq y) \in Fc(X) \mid H^\bullet L(x \geq y) \text{ is an isomorphism}\} \quad (3)$$

and define the subcollection

$$W_0 = \{(x \geq y) \in E_0 \mid H^\bullet L(y) \simeq H^\bullet R[n], \quad (x' \geq y) \in E_0 \text{ for all } x' \geq y \in Fc(X)\}. \quad (4)$$

In other words, $W_0$ is what remains of $E_0$ when we impose

- **dimensionality**: retain only those face relations in which both cells have the local cohomology (isomorphic to that) of a top-dimensional cell, and
- **upward-closure**: remove $(x \geq y)$ if there exists some $x' \geq y$ for which $L^\bullet(x' \geq y)$ is not a quasi-isomorphism.

It is clear that if $W_0$ contains both $(x \geq y)$ and $(y \geq z)$, then it also contains $(x \geq z)$.

**Definition 4.1.** The category $S_0$ is the localization of the face poset $Fc(X)$ about $W_0$.

Recall that $S_0$ has the cells of $X$ as objects, while its morphisms are equivalence classes of $W_0$-zigzags (as described in Definition 2.5). And, there is a canonical functor $Fc(X) \to S_0$ which is universal with respect to rendering all the face relations from $W_0$ invertible.

**Proposition 4.2.** Two cells $w$ and $z$ are isomorphic in $S_0$ if and only if there is a $W_0$-zigzag

$$w \geq y_0 \leq x_0 \geq \cdots \geq y_k \leq x_k \geq z$$

whose last relation $(x_k \geq z)$ lies in $W_0$.

**Proof.** Given any $W_0$-zigzag from $w$ to $z$ such as the one depicted above, recall that every $(x_\bullet \geq y_\bullet)$ is in $W_0$. And by the upward-closure requirement, each preceding $(x_{\bullet-1} \geq y_\bullet)$ is automatically in $W_0$ (with the understanding that $x_{-1} = w$). Since the last relation $(x_k \geq z)$ is assumed to lie in $W_0$ as well, one can directly construct an inverse zigzag in $S_0$:

$$z = z \leq x_k \geq y_k \leq \cdots \leq x_0 \geq y_0 \leq w = w.$$  

Conversely, if $w$ and $z$ are isomorphic in $S_0$, then there exist $W_0$-zigzags from $w$ to $z$ and vice-versa whose composites are (equivalent to) identity morphisms. We label some relevant cells in these zigzags:

$$w \geq \cdots \leq x \geq z \quad \text{and} \quad z \geq y \leq \cdots \geq w.$$  

A brief examination of the horizontal and vertical relations from Definition 2.5 confirms that any zigzag in the equivalence class of the identity must have all its forward-pointing $\geq$’s lying in $W_0$. As a consequence, when we compose the zigzags labeled above, we
obtain \((x \geq y) \in W_0\). An appeal to the upward-closure requirement now forces \((x \geq z) \in W_0\), as desired.

Here is the main result of this section.

**Proposition 4.3.** Two \(n\)-dimensional cells lie in the same canonical \(n\)-stratum of \(X\) if and only if they are isomorphic in \(S_0\).

**Proof.** To show that the first assertion implies the second, let \(\sigma\) be the \(n\)-stratum containing two cells \(w\) and \(z\). Since \(\sigma\) is connected by Definition 2.1, there exists a zigzag of cells lying entirely in \(\sigma\) from \(w\) to \(z\), say
\[
w \geq v_0 \leq v_1 \geq \cdots \geq v_k \leq z.
\]
By the same definition and Proposition 2.4, it follows that every face relation appearing above lies in \(W_0\). By adding an identity to the end of this zigzag, we may therefore construct a morphism in \(S_0\) from \(w\) to \(z\),
\[
w \geq v_0 \leq v_1 \geq \cdots \geq v_k \leq z = z.
\]
It is now evident from Proposition 4.2 that \(w\) and \(z\) are isomorphic in \(S_0\). In order to show the reverse implication, assume the existence of an isomorphism from \(w\) to \(z\) in \(S_0\), say
\[
w \geq y_0 \leq x_0 \geq \cdots \geq y_k \leq x_k = z,
\]
where the last equality follows from the fact that \(X\) contains no cells of dimension \(> n\).

Define the union
\[
Z = \bigcup_{i=0}^{k} \text{st}(y_i)
\]
of open stars, and note that \(Z\) is connected — every constituent \(\text{st}(y_i)\) is connected and intersects the subsequent \(\text{st}(y_{i+1})\) non-trivially (since both stars must contain \(x_i\)). By the dimensionality requirement on \(W_0\), every point in \(Z\) has an open neighborhood with compactly supported cohomology isomorphic to \(H^\bullet (\mathbb{R}^n) = H_c^\bullet (\mathbb{R}^n)\), and by the upward-closure requirement no cell lying in \(Z\) may lie in the boundary of a cell lying outside \(Z\). Thus, the frontier and link requirements on strata imposed by Definition 2.1 are satisfied by the connected set \(Z \subset X\), whence all cells lying in \(Z\) must belong to the same canonical \(n\)-stratum \(\sigma\) of \(X\). Thus, \(\sigma\) contains both \(w\) and \(z\) as desired.

Finally, note that if an arbitrary cell \(y\) lies in a (not necessarily canonical) \(n\)-stratum \(\sigma\) of \(X\), then there must exist an \(n\)-cell \(w \geq y\) also lying in \(\sigma\) by Proposition 2.4, otherwise we arrive at the contradiction \(H^n L(y) = 0\). Thus, we have the following consequence of Proposition 4.3.

**Corollary 4.4.** The canonical \(n\)-strata of \(X\) correspond bijectively with isomorphism classes of its \(n\)-cells in \(S_0\).

5. Uncovering lower strata

As before, let \(X\) be the finite \(n\)-dimensional CW complex whose canonical strata we wish to discover. In this section, we will inductively construct the desired stratification
\[
X = Y_0 \supset Y_1 \supset \cdots \supset Y_n \supset Y_{n+1} = \varnothing,
\]
where each \( Y_\bullet \subset X \) is a regular CW subcomplex, so that the connected components of each difference \( Y_d - Y_{d+1} \) are precisely the canonical \((n-d)\)-strata of \( X \). We have numbered the \( Y_\bullet \) in reverse (compare Definition 2.1) merely to indicate the order in which they are constructed. In addition to the initial complex \( Y_0 = X \), the first step of our construction also requires using the sets \( E_0 \) and \( W_0 \) which were defined in (3) and (4) respectively.

**Definition 5.1.** Given the triple \((Y_{d-1}, E_{d-1}, W_{d-1})\), the \((n-d)\)-dimensional regular CW subcomplex \( Y_d \) consists of all cells \( y \) in \( Y_{d-1} \) for which the trivial face relation \((y = y)\) does not lie in \( W_{d-1} \). Letting \( L_d^\bullet : \text{Fc}(Y_d) \to \text{Ch}(R) \) be the local cohomology of \( Y_d \), define the set of face relations

\[
E_d = E_{d-1} \cap \{ (x \geq y) \text{ in } \text{Fc}(Y_d) \mid H^\bullet L_d(x \geq y) \text{ is an isomorphism} \},
\]

and also

\[
W_d = W_{d-1} \cup \{ (x \geq y) \text{ in } E_d \mid H^\bullet L_d(y) \simeq H^\bullet R[n-d], \text{ and} \\
(x' \geq y) \in E_d \text{ for all } x' \geq y \text{ in } \text{Fc}(Y_d) \}.
\]

This produces the subsequent triple \((Y_d, E_d, W_d)\).

Note that \( W_\bullet \) is an increasing sequence of sets while both \( Y_\bullet \) and \( E_\bullet \) are decreasing. As a motivating example for the preceding constructions, consider the 2-dimensional stratified space \( Z_0 \) obtained as the closed cone over a wedge of two circles:

If \( Z_0 \) is equipped with a regular CW decomposition, then all cells lying in its boundary (i.e., the figure-eight on the left side) will have trivial local cohomology. The local cohomology of \( Z_0 \) thus fails to distinguish generic cells lying on either of the two boundary circles from the special vertex which forms their intersection. But once we identify all cells lying on canonical 2-strata (as described in Section 4), removing those cells produces a 1-dimensional CW subcomplex \( Z_1 \subset Z_0 \) whose local cohomology correctly distinguishes the special vertex:

Similar top-down constructions have been employed implicitly by Whitney [33] and explicitly by Goresky-MacPherson [14] in non-cellular contexts.

**Remark 5.2.** Two properties of the sequence \( Y_\bullet \) mentioned in Definition 5.1 may be inductively justified as follows, using the fact that \( Y_0 \) is an \( n \)-dimensional regular CW complex as a base case.
(1) To see that $Y_d \subset Y_{d-1}$ is a regular CW subcomplex, note that if a cell $y$ lies in the difference $Y_{d-1} - Y_d$, then we have $(y = y) \in W_{d-1}$. And given any cell $x \geq y$ in $Y_{d-1}$, we have $(x \geq y) \in W_{d-1}$ by upward closure, whence $(x = x) \in W_{d-1}$. Thus, $x$ also lies in the difference $Y_{d-1} - Y_d$. Since the collection of cells removed from $Y_{d-1}$ to obtain $Y_d$ is upward-closed with respect to the face partial order, $Y_d$ is a regular CW subcomplex of $Y_{d-1}$.

(2) To see that $\dim Y_d \leq (n - d)$, assume $\dim Y_{d-1} \leq (n - d + 1)$. But now, for each cell $w$ in $Y_{d-1}$ of dimension $(n - d + 1)$, we have $(w = w) \in W_{d-1}$ by Proposition 3.2, so $w$ is not in $Y_d$. Since $Y_d$ contains no cells of dimension $(n - d + 1)$, its dimension can not exceed $(n - d)$.

**Definition 5.3.** For each $d > 0$ in $[n]$, the category $S_d$ is the localization of $Fc(X)$ about the set $W_d$ from (7).

Recall from Section 2.2 that there is a universal functor $Fc(X) \rightarrow S_d$ for each $d$; and since $W_{d-1} \subset W_d$ holds by (7), the universal property of localization guarantees a unique functor $S_{d-1} \rightarrow S_d$ which makes the following diagram commute:

$$\begin{array}{ccc}
Fc(X) & \longrightarrow & S_d \\
\downarrow & & \downarrow \\
S_{d-1} & \longrightarrow & S_d
\end{array}$$

In fact, $S_{d-1} \rightarrow S_d$ admits a convenient explicit description: note that every $W_{d-1}$-zigzag is automatically a $W_d$-zigzag, so $S_{d-1}$ is a subcategory of $S_d$ and the inclusion $S_{d-1} \hookrightarrow S_d$ satisfies this universal property. Our next result establishes a monotonicity property of morphisms across this entire nested sequence $S_*$; in its statement and beyond, $[p]$ denotes the set $\{0, \ldots, p\}$ for each integer $p \geq 0$.

**Lemma 5.4.** Fix $d$ in $[n]$ along with cells $w$ and $z$ in $X$. For each $W_d$-zigzag from $w$ to $z$,

$$w \geq y_0 \leq x_0 \geq \cdots \geq y_k \leq x_k \geq z,$$

define the function $\varphi : [k] \rightarrow [d]$ by setting

$$\varphi(p) = \min \{q \geq 0 \mid (x_p \geq y_p) \in W_q\}. \quad (8)$$

Then, $\varphi$ is monotone increasing in the sense that $\varphi(p - 1) \leq \varphi(p)$ for all $p \geq 1$ in $[k]$.

**Proof.** Set $\varphi(p) = q$ for a fixed $p$, so that $(x_p \geq y_p) \in W_q - W_{q-1}$, whence both $x_p$ and $y_p$ are cells in $Y_q - Y_{q+1}$. Since $x_{p-1} \geq y_p$ holds in $X$, the cell $x_{p-1}$ must lie in $Y_r - Y_{r+1}$ for some $r \leq q$ because $Y_*$ is a decreasing sequence of CW subcomplexes of $X$. Thus, $(x_{p-1} \geq y_{p-1}) \in W_r$, so we have

$$\varphi(p - 1) \leq r \leq q = \varphi(p),$$

as desired. \hfill $\square$

The monotonicity from the preceding lemma guarantees the non-existence of certain morphisms in the entire sequence $S_*$.

**Corollary 5.5.** Given $i \leq j$ in $[n]$ along with cells $w$ in $Y_i$ and $z$ in $Y_i - Y_j$, there are no morphisms in $S_d$ from $w$ to $z$ for any $d$ in $[n]$. 
Proof. Note that a direct face relation \( w \geq z \) can not hold in \( X \) since \( Y_j \) is a CW subcomplex. Thus, any morphism in \( S_d \) from \( w \) to \( z \) admits a genuine zigzag representative. Proceeding by contradiction, assume the existence of such a zigzag:

\[
w \geq y_0 \leq x_0 \geq \cdots \geq y_k \leq x_k \geq z.
\]

The face relations \( w \geq y_0 \) and \( x_k \geq z \) in \( X \) guarantee \( j \leq \varphi(0) \) and \( \varphi(k) \leq i \), where \( \varphi \) is the function from (8). The monotonicity of \( \varphi \) from Lemma 5.4 now forces \( j \leq i \), and hence, \( i = j \). The cell \( z \) is thus constrained to lie in \( Y_i - Y_j = \emptyset \), which yields the desired contradiction.

As a direct consequence of the preceding result, isomorphism classes in \( S_i \) of \( (n - i) \)-dimensional cells from \( Y_i \) remain unchanged across all the inclusions \( S_i \hookrightarrow S_j \) for \( i \leq j \). The next result gives convenient alternate descriptions of such isomorphism classes across the entire sequence \( S_\ast \).

**Proposition 5.6.** For each \( d \in [n] \), the following are equivalent. Two cells \( w \) and \( z \) in \( X \)
1. lie in the same connected component of \( Y_d - Y_{d+1} \);
2. are connected by a \( W_d \)-zigzag, whose relations \( (\leq \) and \( \geq \) all lie in \( W_d - W_{d-1} \);
3. are isomorphic in \( S_d \) to a common \( (n-d) \)-dimensional cell from \( Y_d \).

Proof. We will show (1) \( \Rightarrow \) (2) \( \Rightarrow \) (3) \( \Rightarrow \) (1). If \( w \) and \( z \) lie in the same connected component of \( Y_d - Y_{d+1} \), then there is a zigzag of cells

\[
w \geq v_0 \leq v_1 \geq \cdots \geq v_k \leq z,
\]

lying entirely in \( Y_d - Y_{d+1} \). By Definition 5.1, each face relation \( (\geq \) or \( \leq \) appearing in this zigzag lies in \( W_d \), and by Lemma 5.4, no such relation can lie in \( W_d - W_{d-1} \). So by (7), the cell \( w \) (and indeed, each cell in this zigzag) has its local cohomology \( H^\bullet R^\ast [n-d] \) isomorphic to \( H^\bullet L_d^\ast (v) \) isomorphic to \( H^\bullet L_d^\ast (w) \). By the argument used while proving Corollary 4.4, and recalling that \( \dim Y_d \leq (n - d) \), there exists an \( (n - d) \)-cell \( v \geq w \) in \( Y_d \) which is isomorphic in \( S_d \) to \( w \), and hence to every other cell in our zigzag (including \( z \)). Finally, if there is an \( (n - d) \)-cell \( v \) from \( Y_d \) isomorphic to both \( w \) and \( z \) in \( S_d \), then there is a \( W_d \)-zigzag from \( w \) to \( z \) obtained by composing zigzags \( w \to v \to z \). By Lemma 5.4, all cells in this composite lie in \( W_d - W_{d-1} \), so \( w \) and \( z \) lie in the same connected component of \( Y_d - Y_{d+1} \).

Our next result describes the canonical strata of \( X \) as isomorphism classes in \( S_\ast \).

**Proposition 5.7.** For each \( d \in [n] \) and canonical \( (n-d) \)-stratum \( \sigma \) of \( X \), every cell lying in \( \sigma \) is isomorphic in \( S_d \) to some \( (n-d) \)-cell from \( Y_d \).

Proof. Let \( \emptyset = L_{-1} \subset \cdots \subset L_{d-1} = L \) be the link of \( \sigma \) in the sense of Definition 2.1. Assume, proceeding via induction, that for each \( j < d \) the canonical \( (n-j) \)-strata of \( X \) correspond to isomorphism classes in \( S_j \) of \( (n-j) \)-cells from \( Y_j \). By Proposition 5.6, these strata are precisely the connected components of \( Y_j - Y_{j+1} \). Thus, for each cell \( x \) lying in \( \sigma \) and \( j < d \), we have quasi-isomorphisms

\[
C_\ast^c (\emptyset L_{d-j-1} \times \mathbb{R}^{n-d}) \xrightarrow{\sim} C_\ast^c (\text{st}(x) \cap Y_{j+1}),
\]

where \( \text{st}(x) \) the the open star of \( x \) in \( X = Y_0 \). The cochain complexes on the right side compute the compactly-supported cohomology of \( x \)'s open star in \( Y_{j+1} \), and (by the equivalence of singular and cellular cohomology), they are filtered chain homotopy equivalent to the local cohomology complexes \( L_j^\ast (x) \). By refinability of the quasi-isomorphisms
above, whenever \( x \geq y \) holds in \( X \) for another cell \( y \) lying in \( \sigma \), we obtain commuting triangles of filtered quasi-isomorphisms

\[
C^*_c(\mathcal{C}L_{d-j-1} \times \mathbb{R}^{n-d}) \xrightarrow{\sim} L^*_j(x) \xrightarrow{\sim} \bigcup L^*_j(x \geq y) \xrightarrow{\sim} L^*_j(y)
\]

Thus, \( L^*_j(x \geq y) \) is a quasi-isomorphism whenever \( j < d \), and so by \( (6) \) we have \( (x \geq y) \in E_{d-1} \). Being an \((n - d)\)-stratum, \( \sigma \) is an \( R \)-cohomology manifold, and there are quasi-isomorphisms

\[
R[n - d]^* \xrightarrow{\sim} L^*_d(x) \xrightarrow{\sim} L^*_d(y).
\]

Since \( \sigma \) is a top-dimensional stratum of the subcomplex \( Y_d \), we have \( (x' \geq y) \in E_d \) whenever \( x' \geq y \) in \( Y_d \) by Proposition \( 2.4 \). Thus, \( (x \geq y) \in W_d - W_{d-1} \) by \( (7) \). Since \( \sigma \) is connected, any two cells of \( X \) that it contains may be connected by a zigzag of cells lying in \( \sigma \); and by the preceding argument, each face relation appearing in such a zigzag lies in \( W_d - W_{d-1} \). The desired conclusion now follows from Proposition \( 5.6 \). \( \square \)

The following result establishes the converse of Proposition \( 5.7 \).

**Proposition 5.8.** For each \( d \) in \([n]\), the isomorphism class in \( S_d \) of any \((n - d)\)-dimensional cell from \( Y_d \) is a canonical \((n - d)\)-stratum of \( X \).

**Proof.** Proceeding once again by induction, we assume that the statement holds for all \( j \) in \([d - 1]\). Fix an \((n - d)\)-dimensional cell \( w \) in \( Y_d \subset X \), and let \( \sigma \) be the collection of all cells which are isomorphic to \( w \) in \( S_d \). We first show that \( \sigma \) satisfies the link axiom from Definition \( 2.1 \). Writing \( \text{st}(w) \) and \( \text{lk}(w) \) for \( w \)'s star and link in \( X \), for each \( i \leq d \) we have the intersections

\[
\text{st}_i(w) = \text{st}(w) \cap Y_i \quad \text{and} \quad \text{lk}_i(w) \cap Y_i,
\]

which form the star and link of \( w \) in \( Y_i \subset X \). Since \( \text{st}_i(w) \) admits a stratified homeomorphism to the corresponding \( \mathcal{C} \text{lk}_i(w) \times \mathbb{R}^{n-d} \), we get stratified quasi-isomorphisms of singular cochain complexes

\[
C^*_c(\mathcal{C} \text{lk}_i(w) \times \mathbb{R}^{n-d}) \xrightarrow{\sim} C^*_c(\text{st}_i(w)),
\]

whose codomains are, in turn, quasi-isomorphic to the local cohomology stalks \( L^*_i(w) \). If \( z \) is any other cell which becomes isomorphic to \( w \) in \( S_d \), then by Proposition \( 5.6 \) there is a zigzag from \( w \) to \( z \) whose face relations all lie in \( W_d - W_{d-1} \):

\[
w \geq v_0 \leq v_1 \geq \cdots \geq v_{k-1} \leq v_k \geq z,
\]

and hence zigzags of quasi-isomorphisms in \( \text{Ch}(\mathbb{R}) \) for all \( i \in [d] \)

\[
L^*_i(w) \rightarrow L^*_i(v_0) \leftarrow \cdots \leftarrow L^*_i(v_k) \rightarrow L^*_i(z).
\]

Each complex in sight is bounded and free (hence projective), so the backwards-pointing maps admit quasi-inverses (see [31, Cor 10.4.7]). Thus, one may turn them around to obtain maps \( L^*_i(w) \rightarrow L^*_i(z) \) which fit into a string of quasi-isomorphisms

\[
C^*_c(\mathcal{C} \text{lk}_i(w) \times \mathbb{R}^{n-d}) \rightarrow C^*_c(\text{st}_i(w)) \rightarrow L^*_i(w) \rightarrow L^*_i(z) \rightarrow C^*_c(\text{st}_i(z)).
\]
Setting \( L_j = \text{lk}_{d-j-1}(w) \) for all \( j \in [d-1] \), one obtains a candidate \((d-1)\)-dimensional stratified space \( L \) to serve as the link of \( \sigma \).

Turning now to the frontier axiom, for each \( i < d \) let \( K_i^\bullet \) be the quotient complex \( L_i^\bullet / L_{i+1}^\bullet \) suitably restricted so that it assigns zero stalks to all cells lying outside \( \sigma \). Over each cell \( z \) in \( \sigma \), the cohomology \( H^*K_i(z) \) coincides with the (compactly supported) cohomology of the difference \( \text{st}_i(z) - \text{st}_{i+1}(z) \). By construction, the closure of a canonical \((n-i)\)-stratum \( \tau \) in \( X \) intersects \( \sigma \) at the cell \( z \) if and only if the cohomology \( H^{n-i}K_i(z) \) is non-trivial. By definition, \( K_i^\bullet \) fits into a short exact sequence (of complexes of cosheaves over \( \sigma \)):

\[
0 \rightarrow L_{i+1}^\bullet \hookrightarrow L_i^\bullet \rightarrow K_i^\bullet \rightarrow 0.
\]

Applying the five-lemma [31, Ex 1.3.3] to the resulting long exact sequence, we note that all the extension maps of \( K_i^\bullet \) induce isomorphisms on cohomology. Thus, if \( H^{n-i}K_i \) is non-zero at \( z \), then it is also non-zero at every other cell in \( \sigma \), whence every cell in \( \sigma \) admits a co-face from \( \tau \). Thus, \( \sigma \subset \tau \), as desired.

Finally, to see that \( \sigma \) is canonical, we use Proposition 5.7 to pick an \((n-d)\)-cell \( w \) from \( Y_d \) so that all cells in \( \sigma \) are isomorphic in \( S_d \) to \( w \). Proposition 5.6 now guarantees that no additional cells may be added to \( \sigma \) while preserving its connectedness.

The three preceding propositions establish that the sequence \( Y_\bullet \subset X \) from Definition 5.1 constitutes the canonical stratification of \( X \). The localization functors \( \text{Fc}(X) \rightarrow S_\bullet \) and the universal functors \( S_\bullet \rightarrow S_{\bullet+1} \) arising from the inclusions \( W_\bullet \subset W_{\bullet+1} \) fit into a commutative diagram

\[
\begin{array}{ccccccccc}
& & \text{Fc}(X) & & \\
& \downarrow & & \downarrow & & \\
S_0 & \longrightarrow & S_1 & \longrightarrow & \cdots & \longrightarrow & S_n = S,
\end{array}
\]

which one may view as analogous to an injective resolution of the functor \( \text{Fc}(X) \rightarrow S_0 \) from Section 4. Proceeding from left to right in the sequence above, we obtain canonical strata of decreasing dimension by examining isomorphism classes in \( S_\bullet \). We collect these results into the following expanded version of Theorem 1.1.

**Theorem 5.9.** Given a finite regular \( n \)-dimensional CW complex \( X \), let \( Y_\bullet \) and \( W_\bullet \) be as in Definition 5.1, and let \( S_\bullet \) denote the localization of the face poset \( \text{Fc}(X) \) about \( W_\bullet \). For each \( d \in [n] \), the canonical \((n-d)\)-strata of \( X \) correspond bijectively with isomorphism classes of \((n-d)\)-dimensional cells from \( Y_d \) in \( S_k \) for any \( k \geq d \).

The morphisms in \( S = S_n = \text{colim}(S_\bullet) \) also recover the frontier partial order among canonical strata.

**Proposition 5.10.** The frontier relation \( \tau \succeq \sigma \) holds among two canonical strata of \( X \) if and only if there is a morphism in \( S \) from a cell lying in \( \tau \) to a cell lying in \( \sigma \).

**Proof.** If \( \tau \succeq \sigma \) holds, then there are cells \( w \) and \( z \) in \( \tau \) and \( \sigma \) respectively so that \( w \geq z \) in \( \text{Fc}(X) \). This face relation includes into \( S \) as a morphism from \( w \) to \( z \). Conversely, given cells \( w \) and \( z \) in \( \tau \) and \( \sigma \) respectively, assume the existence of a \( W \)-zigzag

\[
w \geq y_0 \leq x_0 \geq \cdots \geq y_k \leq x_k \geq z
\]
and let $\varphi: [k] \to [n]$ be the associated monotone function from Lemma 5.4. By construction, this zigzag contains cells in canonical strata $\tau_i$ of dimension $n - \varphi(i)$ for all $i$ in $[k]$. If $\varphi(i) = \varphi(i + 1)$ then $\tau_i = \tau_{i+1}$, otherwise $\tau_i \succ \tau_{i+1}$ (because $x_i$ is a co-face of $y_{i+1}$). Thus, there is a descending sequence of canonical strata

$$\tau \succeq \tau_0 \succeq \tau_1 \succeq \cdots \succeq \tau_k \succeq \sigma,$$

so we obtain $\tau \succeq \sigma$ as desired. \hfill \Box

The argument above shows that the zigzag paths which form morphisms in $S$ may remain indefinitely in a single stratum or descend to lower-dimensional adjacent strata, but they can never ascend to higher strata. In other words, $S$ is a (cellular, 1-categorical) version of the \textit{entrance path category} associated to the canonical stratification of $X$ — see [23, Def 3.1], [30, Sec 7] or [34, Sec 2].

6. Algorithms

Progress in applied algebraic topology [12, 24] has inspired, and been sustained by, a combination of theoretical advances [1, 3, 4] on one hand and increased algorithmic sophistication on the other [7, 16, 21]. As a pleasant consequence, one has recourse to a profusion of advanced software [26] dedicated to the very specific task of efficiently computing (co)homology groups of vast cell complexes [17, 22]. From a computational perspective, the local cohomology $L^\bullet$ of a finite regular CW complex enjoys the obvious, but enormous, advantage of being \textit{local}. One only ever needs to construct cochain complexes corresponding to open stars of cells (rather than holding the entire complex in system memory). Since each such star-complex may be processed independently of the others, the construction of $L^\bullet$ and the computation of $H^\bullet L$ are inherently distributable operations.

6.1. Subroutines. There are several avenues for efficiently storing a CW complex for the purpose of computing its cohomology: one may employ a \textit{Hasse graph}, whose vertices are cells and whose edges, usually adorned with an orientation, connect neighboring cells across codimension one. Alternately, one may directly store (sparse avatars of) \textit{coboundary matrices} over $R$, whose nonzero entries encode adjacencies between cells of codimension one. In the presence of extra structure – for instance, if the complex involved is simplicial or cubical – the additional rigidity affords access to highly optimized data structures such as simplex trees or bitmaps. Rather than tethering the prospective implementer to a fixed and possibly inconvenient representation, we will present all our algorithms at a high level and work directly with the poset of cells.

Our first subroutine \texttt{UpSet} employs a minor modification of breadth-first search on a finite poset $P$ to extract the subposet of all elements which lie above a fixed element $v$ but \textit{not} above a fixed element $u$. In all instances of interest here, our input $P$ will be the face poset $\text{Fc}(X)$ of a finite regular CW complex $X$, while $u$ and $v$ will be adjacent cells across codimension one:

$$u >_1 v \text{ if } u > v \text{ and } \dim u - \dim v = 1.$$

By definition, the output $P^u_v$ in such cases is precisely the subposet of $\text{Fc}(X)$ containing cells which lie in the difference of open stars $\text{st}(v) - \text{st}(u)$. We also reserve the right to set $u = \emptyset$, so that line 07 does not check $x$ against $u$ and the output is just the poset of cells in $\text{st}(v)$.
Algorithm: \( \text{UpSet}(P, u, v) \)

**In:** A finite poset \( P \) and elements \( u, v \in P \)

**Out:** The subposet \( P^u_v \) of elements \( \geq v \) but not \( \geq u \)

01 initialize an empty queue \( Q \) of elements
02 add \( v \) to \( Q \)
03 while \( Q \) is nonempty
04 remove \( w \) from \( Q \)
05 if \( w \) is not in \( P^u_v \)
06 add \( w \) to \( P^u_v \)
07 for each element \( x > w \) in \( P \) with \( x \not\geq u \)
08 add \( x \) to \( Q \) (uniquely)
09 return \( P^u_v \)

The next required subroutine is \( \text{Cohom} \), which accepts a poset of cells as input and returns the sequence of \( R \)-modules corresponding to the cohomology of the associated cochain complex. When \( R \) is a field, the cohomology modules are all vector spaces and in this case the output can just be a sequence of integers storing their dimensions. If \( R \) is the ring of integers, then one has to encode the torsion subgroups (if any) in addition to the ranks of the free parts in every dimension. In either case, since the efficient computation of (co)homology has been extensively discussed elsewhere [7, 16, 18, 21], we will treat this subroutine as a black box and not explicitly write it down here.

### 6.2. Description

Our main algorithm \( \text{StratCast} \) accepts as input the poset of cells in a finite \( n \)-dimensional regular CW complex \( X \) and uses the constructions of Definition 5.1 to assign each cell \( x \) a number \( \text{codim}(x) \) in \( [n] = \{0, 1, \ldots, n\} \) so that \( x \) lies in a canonical stratum of dimension \( n - \text{codim}(x) \). Once codimensions have been determined for all cells, finding the actual canonical strata reduces to the simple task of locating connected components of a fixed codimension.

Algorithm: \( \text{StratCast}(\text{Fc}(X)) \)

**In:** The face poset of an \( n \)-dim regular CW complex

**Out:** Assigns canonical codimensions to cells

01 for each \( d \) in \( (0, 1, \ldots, n - 1, n) \)
02 for each cell \( w \) in \( \text{Fc}(X) \) // dist
03 assign \( h^\star(x) = \text{Cohom}(\text{UpSet}(\text{Fc}(X), \emptyset, w)) \)
04 for each \( (x >_1 y) \) in \( \text{Fc}(X) \) with \( \ell(x >_1 y) = d - 1 \) // dist
05 assign \( c^\star(x, y) = \text{Cohom}(\text{UpSet}(\text{Fc}(X), y, x)) \)
06 if \( c^\star(x, y) \) is trivial
07 assign \( \ell(x >_1 y) = d \)
08 assign \( \text{codim}(z) = d \) to each \( (n - d) \)-cell \( z \) in \( \text{Fc}(X) \)
09 for each \( i \)-cell \( u \) in \( \text{Fc}(X) \) with \( i \) in \( (n - d - 1, \ldots, 1, 0) \) // dist
10 if \( h^{n-d}(u) \simeq R \) and all other \( h^\star(u) = 0 \)
11 if \( \text{codim}(v) = d \) and \( \ell(v >_1 u) = d \) for all \( v >_1 u \) in \( \text{Fc}(X) \)
12 assign \( \text{codim}(u) = d \)
13 remove all cells from \( \text{Fc}(X) \) with \( \text{codim} = d \)
The main loop (in line 01) increments the current codimension \( d \) from 0 up to \( n \); as it executes each iteration, the cells lying on \((n - d)\)-dimensional canonical strata are identified and removed from \( X \), so that in the \( d \)-th iteration we pass from \( Y_d \) to \( Y_{d+1} \) in the language of Definition 5.1. There are three secondary loops (at lines 02, 04 and 09); all three are trivially distributable across several processors, as indicated by the \textit{dist} comments in those lines. The first loop in line 04 is easiest to explain: it uses our subroutines \texttt{UpSet} and \texttt{Cohom} to compute the compactly-supported cohomology of each open star \( \text{st}(w) \cap Y_d \) where \( w \) is a cell in \( Y_d \subset X \).

The second loop from line 04 computes whether or not inclusions of open stars \( \text{st}(x) \subset \text{st}(y) \) for \( x >_1 y \) induce isomorphisms on compactly supported cohomology by examining their cokernels. In other words, we use the fact that the map induced by \( L_d^*(x >_1 y) \) (namely, the inclusion \( \text{st}(x) \cap Y_d \hookrightarrow \text{st}(y) \cap Y_d \)) on cohomology is an isomorphism if and only if the difference \( (\text{st}(y) - \text{st}(x)) \cap Y_d \) has trivial cohomology. We begin by assigning to each face relation \( (x >_1 y) \) the number \( \ell(x >_1 y) = -1 \). As the algorithm executes, \( \ell(x >_1 y) \) gets incremented to the largest integer \( k \) so that the extension maps \( L_i^*(x >_1 y) \) are quasi-isomorphisms for all \( i \in [k] \). Thus, we are guaranteed to have \((x >_1 y) \in E_k \) where \( E_k \) denotes the sets from (6). Since this loop is completely independent of the loop in line 02, it may be executed simultaneously with that loop.

The third loop, which spans lines 09 through 12, must be executed after the first two intermediate loops have terminated. It uses the cohomology groups \( h^* \) computed by the loop in line 02 and the cohomology groups \( \ell \)-values computed by the loop in line 04 to select cells lying on canonical strata of dimension \((n - d)\) via (7). Note that line 10 ensures the correct dimensionality and line 11 enforces upward closure. Line 12 removes all cells lying on canonical \((n - d)\)-strata from \( Y_d \), so we are left with \( Y_{d+1} \) and the outer loop of line 01 enters its next iteration.

### 6.3. Complexity.

The intermediate loops in line 02 and line 04 of \texttt{StratCast} require breadth-first search (for constructing open stars via \texttt{UpSet}) and matrix diagonalization via row and column operations (for computing cohomology via \texttt{Cohom}). Define the \textit{star size} complexity parameter

\[
p = \max \{ \text{number of cells in } \text{st}(x) \mid x \text{ is a cell in } X \} \tag{9}
\]

(it suffices to take the maximum over zero-dimensional cells, since those are guaranteed to have the largest open stars). The worst-case complexity of computing cohomology of any open star, or a difference of open stars, via row and column operations is thus \( O(p^3) \) provided we incur a constant cost performing ring operations in \( R \). The cost of running \texttt{Cohom} clearly dominates the \( O(p^2) \) cost of executing breadth-first search in \texttt{UpSet}. Since the loops of line 02 and 04 may be executed in a completely distributed fashion, their combined cost within a single iteration of the outer loop of line 01 is \( O(p^3) \).

The third intermediate loop (on line 09) may be reinterpreted as two nested loops: an outer \( i \)-loop which decrements dimension from \((n - d - 1)\) to 0, and an inner \( u \)-loop which examines all remaining cells of dimension \( i \). Since we require knowledge of \textit{codim} values of all cells \( v >_1 u \) while processing each \( u \), only the \( u \)-loop is actually distributable. During each iteration of the \( u \)-loop, one incurs an \( O(p) \) cost while checking all remaining cells \( v >_1 u \) in line 11, and an \( R \)-dependent cost of testing \( R \)-module isomorphisms in line 10. At least when \( R \) is a finite field, the rationals or the integers, checking whether a given module has rank one or zero incurs \( O(1) \) cost. Thus, in typical
cases we expect line 11 to dominate the burden of executing the \( u \)-loop. Noting that the \( i \)-loop runs at most \( n \) times regardless of \( d \), each iteration of the loop on line 09 incurs a computational cost of \( O(np) \).

Since all cells with \( \text{codim} = d \) are identified on lines 08 and 12 in a given iteration of the outer \( d \)-loop of StratCast, we assume that removing them from \( X \) incurs no additional cost. Putting all our estimates together, each iteration of the \( d \)-loop incurs a worst-case cost of \( O(p^3 + np) \). Since this loop executes exactly \( (n+1) \) times, the total complexity of running StratCast (assuming that the number of available processors exceeds the number of cells and codimension one face relations in \( X \)) is \( O((n+1)(p^3 + np)) \). And once StratCast has terminated, one may find all desired canonical strata of \( X \) by computing connected components of cells with the same \( \text{codim} \) value. This is linear in the number of cells in \( X \), so we obtain the following result.

**Proposition 6.1.** Let \( X \) be an \( n \)-dimensional regular CW complex with \( m \) cells and star size \( p \) as in (9). The time complexity of computing its canonical stratification with coefficients in a (commutative, unital) ring \( R \) via StratCast is

\[
O((n+1)(p^3 + np) + m),
\]

provided that

1. ring operations in \( R \) incur \( O(1) \) cost,
2. isomorphism-testing against 0 and \( R \) in \( \text{Mod}(R) \) is \( O(1) \), and
3. the number of available processors exceeds \( mp \).

The first two conditions above are satisfied by typical choices of \( R \) (such as the integers or finite fields). In practice, one expects to have \( m \gg (n+1)(p^3 + np) \), so the observed cost of computing canonical strata is essentially linear in the size of \( X \).

**References**

[1] A. Blumberg and M. Lesnick. Universality of the homotopy interleaving distance. *arXiv:1705.01690 [math.AT]*, 2017.
[2] G. Bredon. *Sheaf theory*. Springer, 1997.
[3] P. Bubenik, V. de Silva, and V. Nanda. Higher interpolation and extension for persistence modules. *SIAM Journal on Applied Algebra and Geometry*, 1(1):272–284, 2017.
[4] F. Chazal, V. de Silva, M. Glisse, and S. Oudot. *The structure and stability of persistence modules*. Springer-Briefs, 2016.
[5] J. Curry. Sheaves, cosheaves and applications. *arXiv:1303.3255v2 [math.AT]*, 2013.
[6] J. Curry. Dualities between cellular sheaves and cosheaves. *arXiv:1607.06942 [math.AT]*, 2016.
[7] J. Curry, R. Ghrist, and V. Nanda. Discrete Morse theory for computing cellular sheaf cohomology. *Foundations of Computational Mathematics*, 16(4):875–897, 2016.
[8] W. Dwyer and D. Kan. Calculating simplicial localizations. *Journal of Pure and Applied Algebra*, 18:17–35, 1980.
[9] S. Eilenberg and J. Zilber. On products of complexes. *American Journal of Mathematics*, 75:200–204, 1953.
[10] P. Gabriel and M. Zisman. *Calculus of fractions and homotopy theory*. Number 35 in Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag, 1967.
[11] J. Getz and M. Goresky. *Hilbert Modular Forms with Coefficients in Intersection Homology and Quadratic Base Change*. Birkhäuser, 2012.
[12] R. Ghrist. *Elementary Applied Topology*. Createspace, 2016.
[13] M. Goresky and R. MacPherson. Intersection homology theory. *Topology*, 2(19):135 – 162, 1980.
[14] M. Goresky and R. MacPherson. Intersection homology II. *Inventiones Mathematicae*, 71:77–129, 1983.
[15] M. Goresky and R. MacPherson. *Stratified Morse theory*, volume 14 of *Ergebnisse Der Mathematik Und ihrer Grenzgebiete*. Springer-Verlag, 1988.

[16] S. Harker, K. Mischaikow, M. Mrozek, and V. Nanda. Discrete Morse theoretic algorithms for computing homology of complexes and maps. *Foundations of Computational Mathematics*, 14(1):151 – 184, 2014.

[17] G. Henselman. Eirene, computational homology software, gregoryhenselman.org/eirene.

[18] G. Henselman and R. Ghrist. Matroid filtrations and computational persistent homology. *arXiv:1606.00199 [math.AT]*, 2016.

[19] F. Kirwan. *An Introduction to Intersection Homology Theory*, volume 187 of *Pitman Research Notices in Mathematics*. Longman Scientific and Technical, Harlow, 1988.

[20] J. A. Lee and M. Verleysen. *Nonlinear dimensionality reduction*. Springer-Verlag, 2008.

[21] K. Mischaikow and V. Nanda. Morse theory for filtrations and efficient computation of persistent homology. *Discrete and Computational Geometry*, 50(2):330–353, 2013.

[22] V. Nanda. Perseus: the persistent homology software, people.maths.ox.ac.uk/nanda/perseus.

[23] V. Nanda. Discrete Morse theory and localization. *arXiv:1510.01907v2 [math.AT]*, 2015.

[24] V. Nanda and R. Sazdanović. Simplicial models and topological inference for biological systems. In N. Jonoska and M. Saito, editors, *Discrete and Topological Models in Molecular Biology*, chapter 6, pages 109–141. Springer, 2014.

[25] V. Nanda, D. Tamaki, and K. Tanaka. Discrete Morse theory and classifying spaces. *arXiv:1612.08429 [math.AT]*, 2016.

[26] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington. A roadmap for the computation of persistent homology. *arXiv:1506.08903v6 [math.AT]*, 2017.

[27] C. Rourke and B. Sanderson. Homology stratifications and intersection homology. *Geometry and Topology Monographs*, 2:455 – 472, 1999.

[28] A. Shepard. *A Cellular Description of the Derived Category of a Stratified Space*. Brown University PhD Thesis, 1985.

[29] D. Tamaki. Cellular stratified spaces. *arXiv:1609.04500 [math.AT]*, 2016.

[30] D. Treumann. Exit paths and constructible stacks. *Compositio Mathematica*, 145:1504–1532, 2009.

[31] C. Weibel. *Homological Algebra*. Cambridge University Press, 1994.

[32] S. Weinberger. *The Topological Classification of Stratified Spaces*. University of Chicago Press, 1994.

[33] H. Whitney. Tangents to an analytic variety. *Annals of Mathematics*, 81:469 – 549, 1965.

[34] J. Woolf. The fundamental category of a stratified space. *Journal of Homotopy and Related Structures*, 4(1):359–387, 2009.