CS 6170: Computational Topology, Spring 2019
Lecture 28

Topological Data Analysis for Data Scientists

Dr. Bei Wang

School of Computing
Scientific Computing and Imaging Institute (SCI)
University of Utah
www.sci.utah.edu/~beiwang
beiwang@sci.utah.edu

April 18, 2019

Cohomology

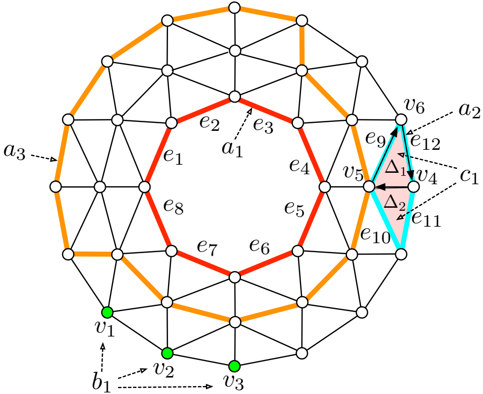
Homology: elementary chain

- Simplicial complex K, e.g. a triangulation of an annulus.
- Vertices, edges and triangles: $K^0 = \{v_i\}$, $K^1 = \{e_i\}$ and $K^2 = \{\Delta_i\}$.
- By abuse of notation, each 0-, 1- and 2-simplex in K corresponds to an *elementary chain* of the same dimension, denoted as v_i, e_i and Δ_i.
- 0-, 1- and 2-chains: formal sums of 0-, 1- and 2-simplexes (elementary chains) with integer coefficients $\mathbb{Z}_2 = \{0, 1\}$:

$$\begin{aligned} C_0 &= C_0(K; \mathbb{Z}_2) = \{b = \sum g_i v_i \mid g_i \in \mathbb{Z}_2\}, \\ C_1 &= C_1(K; \mathbb{Z}_2) = \{a = \sum g_i e_i \mid g_i \in \mathbb{Z}_2\}, \\ C_2 &= C_2(K; \mathbb{Z}_2) = \{c = \sum g_i \Delta_i \mid g_i \in \mathbb{Z}_2\}. \end{aligned}$$

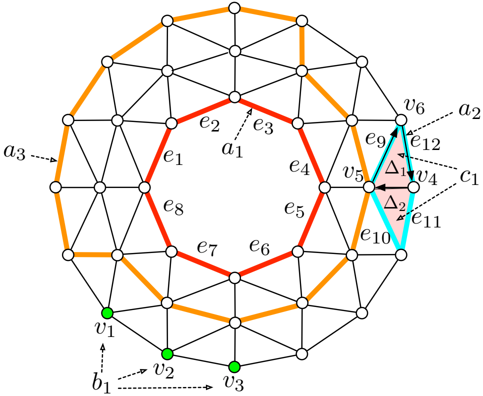
- Two elements $a, a' \in C_1$ are homologous iff $a - a' = \partial(c)$, for some 2-chain c, denoted as $a \sim a'$. In this case $[a] = [a']$.

Homology



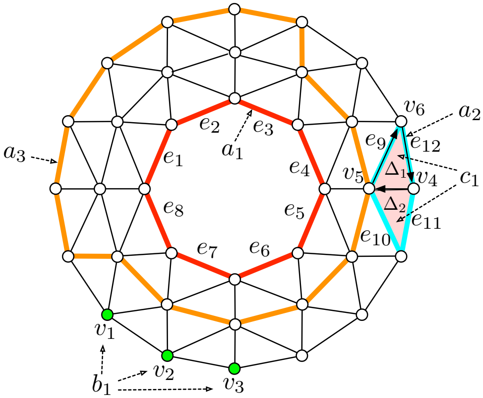
- 0-chain: $b_1 = v_1 + v_2 + v_3$ (solid green),
- 1-chain: $a_1 = e_1 + e_2 + e_3 + e_4 + e_5 + e_6 + e_7 + e_8$ (bold red),
- 2-chain: $c_1 = \Delta_1 + \Delta_2$ (bold pink).
- Here a_1 (bold red) $\sim a_3$ (bold orange): they bound the "same" tunnel.

Homology



- *Boundary maps*: $\partial_2 : C_2 \to C_1$ and $\partial_1 : C_1 \to C_0$
- *1-dimensional* homology groups: $\mathsf{H}_1 = \mathsf{Z}_1/\mathsf{B}_1 = \ker\partial_1/\mathrm{im}\partial_2$

Homology



- a_1 (bold red) is a 1-cycle since $\partial(a_1) = 0$.
- $a_2 = e_9 + e_{10} + e_{11} + e_{12}$ (bold cyan) is a 1-boundary since it is the boundary of the 2-chain c_1.
- a_1 is a 1-cycle, but not a 1-boundary, which makes $[a_1]$ a non-trivial element of H_1.

Cohomology: elementary cochain

- Simplicial complex K, e.g. a triangulation of an annulus.
- Homology groups are "dual" to homology groups
- Vertices, edges and triangles: $K^0 = \{v_i\}$, $K^1 = \{e_i\}$ and $K^2 = \{\Delta_i\}$.
- By abuse of notation, each 0-, 1- and 2-simplex in K corresponds to an *elementary cochain* of the same dimension, denoted as v_i^*, e_i^* and Δ_i^*.
- 1-simplex e has a corresponding elementary 1-cochain e^*, which is a function on 1-chain whose value is 1 on e and 0 on all other edges.
- $e^* : C_1 \to \mathbb{Z}_2$, where $e^*(e) = 1$ and $e^*(e') = 0$ for all $e' \in K^1, e' \neq e$.
- Similarly, we have elementary 0-cochains, v^* associated with the 0-simplices v; and elementary 2-cochains Δ^* associated with the 2-simplices Δ.
- 0-, 1- and 2-cochains can be considered as sums of elementary cochains.

Cohomology: elementary cochain

- 0-, 1- and 2-cochains: functions on $0-$, $1-$ and 2-chain groups.

$$\begin{aligned}
C^0 &= C^0(K;\mathbb{Z}_2) = \{\beta : C_0 \to \mathbb{Z}_2, \beta = \sum g_i v_i^* \mid g_i \in \mathbb{Z}_2\}, \\
C^1 &= C^1(K;\mathbb{Z}_2) = \{\alpha : C_1 \to \mathbb{Z}_2, \alpha = \sum g_i e_i^* \mid g_i \in \mathbb{Z}_2\}, \\
C^2 &= C^2(K;\mathbb{Z}_2) = \{\gamma : C_2 \to \mathbb{Z}_2, \gamma = \sum g_i \Delta_i^* \mid g_i \in \mathbb{Z}_2\}.
\end{aligned}$$

Cohomology: coboundary map

- *Coboundary maps* are dual to the boundary maps, $\delta_0 : C^0 \to C^1$, $\delta_1 : C^1 \to C^2$
- Let $\beta \in C^0$, $\alpha \in C^1$, we have:

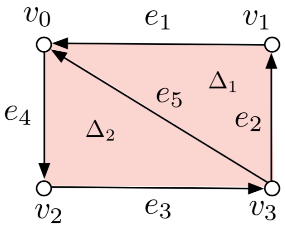
$$\begin{aligned}(\delta_0\beta)([v_0,v_1]) &= \beta(\partial_1([v_0,v_1]) = \beta(v_1) + \beta(v_0),\\ (\delta_1\alpha)([v_0,v_1,v_2]) &= \alpha(\partial_2([v_0,v_1,v_2]))\\ &= \alpha([v_1,v_2]) + \alpha([v_0,v_2]) + \alpha([v_0,v_1]).\end{aligned}$$

- If $\alpha = \sum g_i e_i^*$, then $\delta(\alpha) = \sum g_i(\delta e_i^*)$.
- To compute δe^* for each oriented simplex e, we have $\delta e^* = \sum \Delta_j^*$, where the summation extends over all Δ_j having e as a face.

Cohomology: cocycle, cobounday

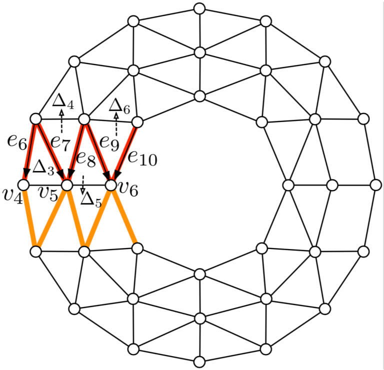
- For a cochain $\alpha \in C^1$, we call α a 1-*cocycle* if $\delta_1(\alpha) = 0$.
- We call α a 1-*coboundary* if there exists a cochain $\beta \in C^0$ such that $\delta_0(\beta) = 0$.
- It is easy to verify that $\delta \circ \delta = 0$.
- 1-coboundaries are always 1-cocycles, we have $\mathrm{im}(\delta_0) \subseteq \ker(\delta_1)$.
- We define the 1-*cohomology* of K to be the quotient group, $\mathsf{H}^1 = \mathsf{Z}^1/\mathsf{B}^1 = \ker(\delta_1)/\mathrm{im}(\delta_0)$.
- Two 1-cocycles α and α' are *cohomologous* if $\alpha + \alpha'$ is a coboundary.

Cohomology



- $e_5^* : C_1 \to Z$ has value 1 on e_5 and 0 on other edges. Ignore orientations (for now).
- δe_5^* has values 1 on Δ_1 and Δ_2, because e_5 appears in $\partial\Delta_2$ and $\partial\Delta_1$.
- $\delta e_5^* = \Delta_2^* + \Delta_1^*$.
- $\delta v_1^* = e_2^* + e_1^*$, $\delta v_3^* = e_3^* + e_2^* + e_5^*$.
- 1-cochain $\alpha = e_1^* + e_5^* + e_3^*$ is a 1-cocycle since $\delta(\alpha) = \delta(e_1^*) + \delta(e_5^*) + \delta(e_3^*) = (\Delta_1^*) + (\Delta_2^* + \Delta_1^*) + (\Delta_2^*) = 0$.
- α is also a 1-coboundary since $\alpha = \delta(v_1^* + v_3^*)$.

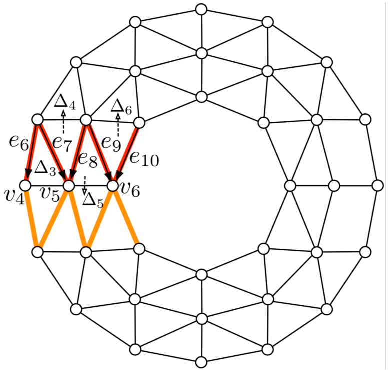
Cohomology



- 1-chain $\alpha_1 = e_6^* + e_7^* + e_8^* + e_9^* + e_{10}^*$ is a 1-cocycle

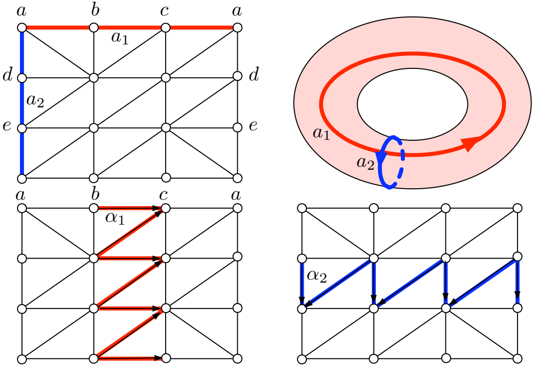
$$\begin{aligned}\delta(\alpha_1) &= \delta(e_6^*) + .. + \delta(e_{10}^*) \\ &= \Delta_3^* + (\Delta_4^* + \Delta_3^*) + (\Delta_5^* + \Delta_4^*) + (\Delta_6^* + \Delta_5^*) + \Delta_6^* = 0.\end{aligned}$$

Cohomology



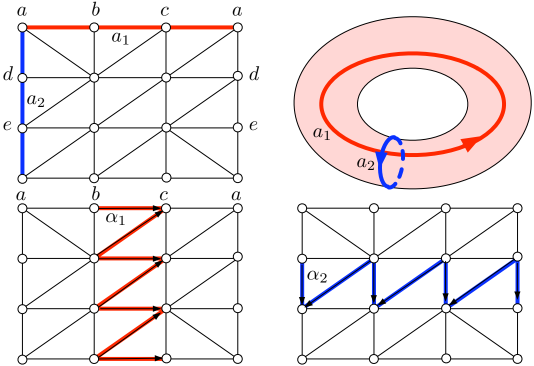
- 1-chain $\alpha_1 = e_6^* + e_7^* + e_8^* + e_9^* + e_{10}^*$ is not a 1-coboundary.
- $[\alpha_1] \in \mathsf{H}^1$, and α_1 can be used as the representative of the 1-dimensional cohomology class.
- α_1 (bold red) is cohomologous to α_2 (bold orange), as we can check $\alpha_1 + \alpha_2 = \delta(v_4^* + v_5^* + v_6^*)$.

Homology groups of a torus



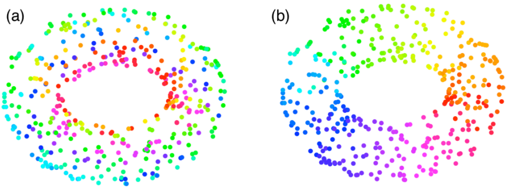
- H_1 is generated by the 1-chains a_1 (red) and a_2 (blue).
- $a_1 = [a, b] + [b, c] + [c, a]$ and $a_2 = [a, d] + [d, e] + [e, a]$.
- a_1 and a_2 are 1-cycles, as $\partial(a_1) = \partial(a_2) = 0$.
- a_1 and a_2 are not 1-boundaries.
- In addition, a_1 and a_2 are not homologous.

Cohomology groups of a torus



- H^1 is generated by the 1-cochains α_1 (red) and α_2 (blue). α_1 and α_2 are 1-cocycles, not 1-coboundaries, and are not cohomologous.
- The duality between cohomology and homology generators is counter-intuitive.
- Here, $\alpha_1 \in \mathsf{H}^1$ (bold red) is dual to $a_1 \in \mathsf{H}_1$ (bold red), while $\alpha_2 \in \mathsf{H}^1$ (bold blue) is dual to $a_2 \in \mathsf{H}_1$ (bold blue).

Cohomology based parametrization



de Silva et al. (2009)

- Using a principle from homotopy theory: relates circular coordinates with cohomology.
- Let $[X, \mathbb{S}^1]$ be the set of equivalence classes of continuous maps from space X to $\mathbb{S}^1$ under the homotopy relation.
- For topological spaces with the homotopy type of a cell complex, there is an isomorphism $\mathsf{H}^1(X; \mathsf{Z}) \cong [X, \mathbb{S}^1]$
- This implies that if X has a non-trivial 1-dimensional cohomology class $[\alpha] \in \mathsf{H}^1(X; \mathsf{Z})$, we can construct a continuous function $\theta : X \to \mathbb{S}^1$ from a representative α.

Future Directions and Discussions

What is the most valuable tool
you have leant in this class?

What is on your wish list?

What do you think are the future directions for topological data analysis?

Advanced Topics

- Topology in visualization: vector field topology, tensor field topology
- TDA for biomedicine: high-dimensional data analysis, mapper
- TDA for materials science, astronomy, music, signal processing
- Multi-parameter persistent homology
- Discrete Morse theory, discrete Stratified Morse theory

Future directions

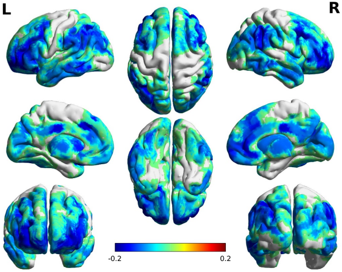
- Better understanding of multi-parameter persistent homology
- Scalable computation
- Machine learning
- New visualization tools, uncertainty visualization
- New theory leads to new algorithms and applications!

Topological Data Analysis of Functional MRI Connectivity in Time and Space Domains

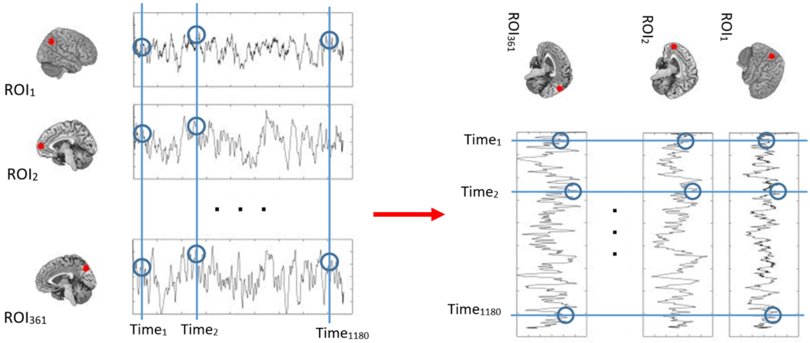
Anderson et al. (2018), Best Paper at MICCAI CNI 2018.

Key findings

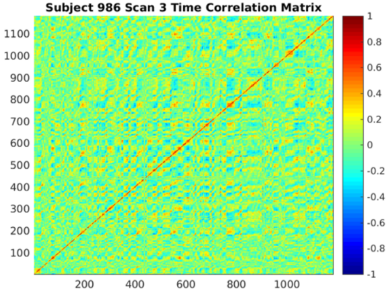
- Functional connectivity in time and space domains produced complementary information about brain function.
- Results from topological data analysis is significantly correlated with cognitive performance, especially fluid intelligence.
- Time and Space topology are both correlated with fluid intelligence, but the spatial domain picks up additional behavior and allows us to localize brain regions.

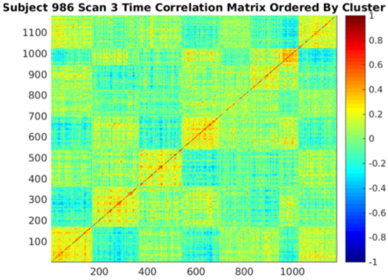


Flipping the time series

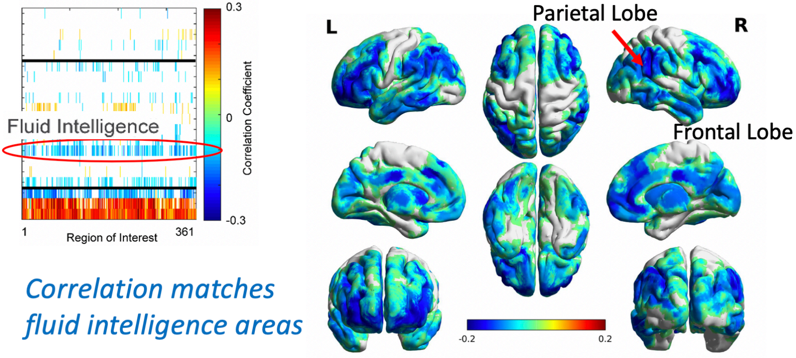


Time domain connectivity matrix



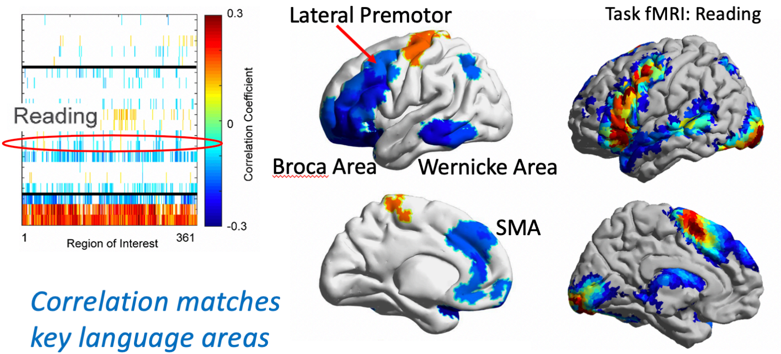


TDA: Significant correlation with fluid intelligence



Correlation matches fluid intelligence areas

TDA: Significant correlation with reading comprehension

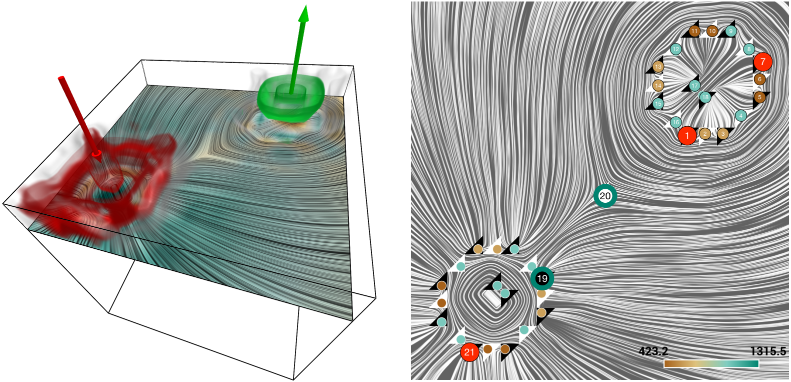


Correlation matches key language areas

Extract Robust Features From Stress Tensor Fields

Wang and Hotz (2017); Jankowai et al. (2018)

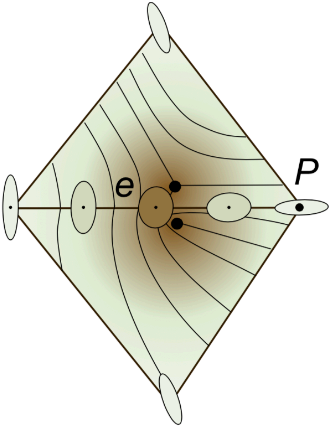
Stress tensor field: two point load



(a) (b)

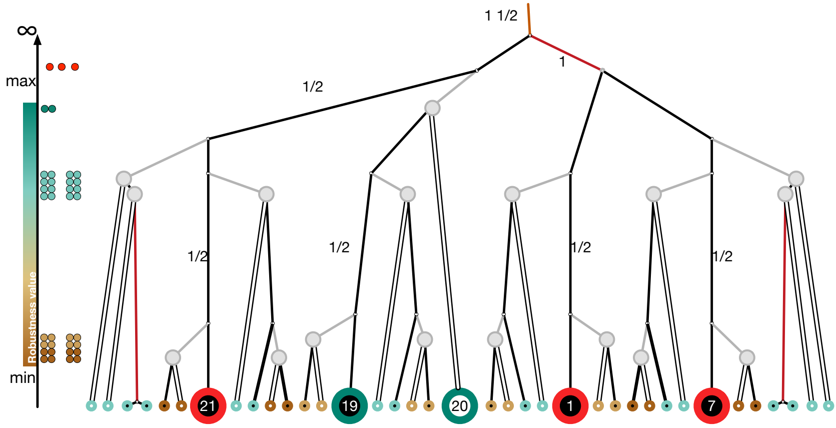
Figure: Visualization of the robustness of a slice of a stress tensor field. (a) A single slice of the data embedded in a 3D context visualized using volume rendering. (b) A textured slice with degenerate cells. White triangles represent trisectors and black ones represent wedges. Degenerate points in each cell are visualized with a brown-to-turquoise colormap. Degenerate points with infinite robustness are in red.

Robustness of degenerate points



(a) (b)

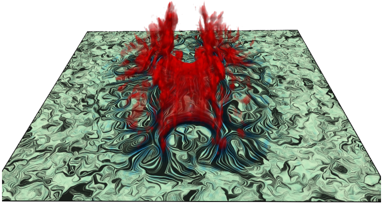
Stress tensor field: two point load



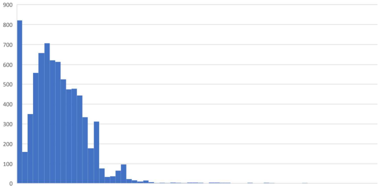
(c)

Figure: Visualization of the robustness of a slice of a stress tensor field. (c) The full merge tree.

Diffusion tensor imaging



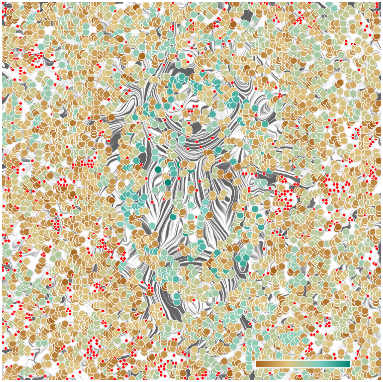
(a)



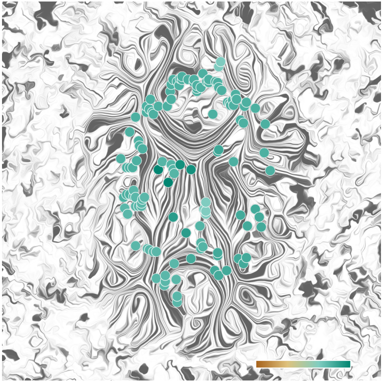
(b)

Figure: A slice of a diffusion tensor imaging data set. (a) 3D context visualization using volume rendering of the anisotropy. (b) The histogram of the robustness values of degenerate points.

Diffusion tensor imaging



(c)



(d)

Figure: (c) All degenerate cells are color-coded according to their robustness values. (d) Most robust degenerate cells are highlighted in turquoise.

Thank you for a wonderful and energetic semester!

Anderson, K. L., Anderson, J. S., Palande, S., and Wang, B. (2018). Topological data analysis of functional mri connectivity in time and space domains. In *Connectomics Neuroimaging (Lecture Notes in Computer Science, Proceedings of International Workshop on Connectomics in NeuroImaging), accepted*, volume 11083. Springer.

de Silva, V., Morozov, D., and Vejdemo-Johansson, M. (2009). Persistent cohomology and circular coordinates. *Proceedings 25th Annual Symposium on Computational Geometry*, pages 227–236.

Jankowai, J., Wang, B., and Hotz, I. (2018). Robust extraction and simplification of 2D tensor field topology. Under review.

Wang, B. and Hotz, I. (2017). Robustness for 2D symmetric tensor field topology. In Schultz, T., Özarslan, E., and Hotz, I., editors, *Modeling, Analysis, and Visualization of Anisotropy*, pages 3–27. Springer International Publishing.