

Topological Signatures of Immune Response: Characterizing T-Cell Dynamics using Persistent Homology

Nathan Yong
ny79@njit.edu

Ho Joon Han
hh346@njit.edu

Amogh Dalal
asd33@njit.edu

May 10, 2026

Abstract

T-cells are important parts of the immune and lymphatic system responsible for triggering immune responses to combat the entrance of foreign entities in the human body, and their effectiveness is largely determined by their cooperative motility - how quickly they can help each other move through crowded environments. Prior work has studied their movement under a simplified 2-dimensional model, which simulated cells with changing morphology as circles with oscillating radius, and showed that for intermediate spatial densities of cells, immune response was maximized for certain frequencies of radius oscillation. In this paper, we consider the same model and use Topological Data Analysis (TDA) to explain and characterize the physical behavior of this model. We find that through using Persistent Homology, we can correlate dimension 0 and 1 topological features with immune response and further explain the dynamics of cells by relating it to the topology and intercellular forces of the system. We also show that in our simulated model, the topological behavior within the global cell domain is also present within local clusters of cells, and that the overall correlation of features gets more stable as cells operate in denser environments.

1 Introduction

1.1 Background

T-cells are crucial parts of the lymphatic and immune system that protect the body against foreign pathogens and infected cells. When pathogens enter the lymphoid tissues, they can be detected by immature dendritic cells [1]. The dendritic cells can then absorb the pathogen and break it down to produce antigens, which are used by cells to uniquely identify the type of pathogen present. During this process, the cells mature into antigen-presenting cells (hereby referred to as APC cells) which can provide their antigens to T-cells through physical contact. If the antigens fit the receptors of the contacted T-cells, then the T-cells are activated and alerted of the presence of pathogens, thereby inducing an immune response.

In order to provide an effective immune response, T-cells must efficiently scan the surface of APC cells to determine whether a threat is present [2]. Because each T-cell has different receptors that match different antigens, it is optimal for many cells to concurrently scan APC cells to produce a timely immune response. Notably, while searching for APC cells, T-cells travel in extremely crowded environments. However, unlike vehicular or pedestrian traffic, contacts between T-Cells in this crowded environment are non-obstructing. That is, T-Cells exhibit a phenomenon known as cooperative motility that allows them to continue movement instead of inhibiting one another in crowded environments [3]. It is of biological interest to quantify the immune response of these cells and study their dynamics to explain the mechanisms behind their cooperative motility.

In order to quantify immune response in such crowded environments, the scanning rate, or the average number of first time contacts of T-Cells to an APC cell, is utilized. Past work by Wiebke and Parisi has shown that in a simplified 2-dimensional system, the deformation in T-Cell morphology, - modeled as circles with oscillating radius under a fixed frequency - can greatly affect the scanning rate against an APC cell [4]. Notably, there was an optimal frequency which maximized scanning rate, as shown in Figure 1. In this paper, we study the same 2-D system and its properties by applying persistent-homology-based topological data analysis to the system. The analysis provides insights into how the effects of deforming cell morphology and scanning rate are reflected within the topological features of

the spatial distribution of T-cells, and how the features vary in environments with varying crowding of cells.

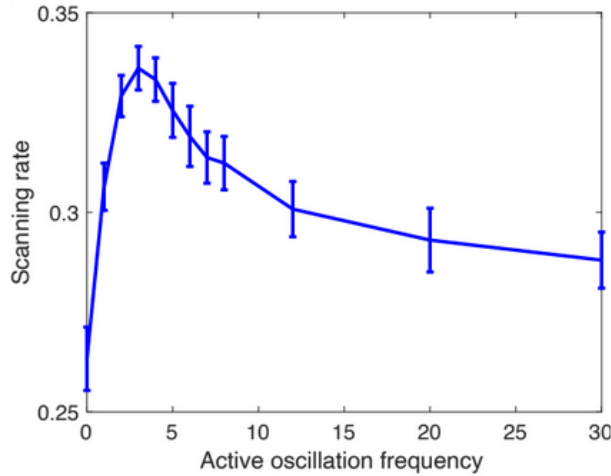


Figure 1: Figure from Wiebke and Parisi [4]. The plot shows the scanning rate trend over frequency. The error bars represent the standard error of the mean scanning rate across realizations. The precise method of computing the scanning rate is explained in the Experimental Setup section.

1.2 Goals

The goal of this study is to better understand the topology of the spacial arrangement of T-Cells in the 2-dimensional system, and relate the cell dynamics arising from the force laws to topological measures. In doing so, we hope to explain and characterize the process of T-cell-dominated immune response from a topological point of view and provide a deeper understanding of the dynamics of T-Cells and the biological phenomena they exhibit.

1.3 Related Work

Much work has been done on studying the dynamics of T-cells in the immune system. [4] Wiebke and Parisi studied T-cells and scanning rate in various 1-D and 2-D scenarios, which forms the basis of much of our work. [5] Vroomans et al. demonstrated through computational simulations that chemotactic attraction of T-cells toward dendritic cells, particularly when combined with post-contact desensitization, enhances scanning efficiency and increases the likelihood that rare antigen-specific T-cells locate dendritic cells carrying cognate antigen. [6] Celli et al. combined flow cytometry, two-photon imaging, and computational modeling to show that a minimum threshold of dendritic cells is required to ensure a high probability of T-cell encounters and initiation of an immune response. [7] Wortel et al. used microfluidic experiments and computational modeling to show that T-cells in crowded, confined environments synchronize their speeds and form cooperative moving trains, a collective behavior driven by cell-to-cell contact maintenance and pushing, which is absent in neutrophils and may help T-cells navigate dense tissue environments without jamming.

Topological data analysis has applications in many other biological settings. [8] Porter et al. provide an accessible overview of TDA and persistent homology, and how it can be used to characterize the shape of data within applications in granular matter, vascular networks, and spatial systems.

2 Model

We utilize the same discrete-time model for 2-D cell interactions as given by Wiebke and Parisi [4]. T-cells are assumed to be circles, and the deformation in their morphology is modeled by the change in the radius of the circle. While this assumption produces a model which is much simpler than the real T-cell system, it is sufficient to test whether changes in cell morphology affects the cooperative motility of cells.

The position of each cell, $x_i(t)$, is governed by

$$x_i(t + \Delta t) = x_i(t) + v_i(t)\Delta t, \quad (1)$$

where the velocities $v_i(t)$ are given by

$$v_i(t) = v_{0i} + \frac{F_i(t)}{\gamma}, \quad (2)$$

In this setup, v_{0i} is the self-propulsion velocity of cell i , which is an innate velocity that the cell prefers. γ is the damping coefficient (which could be affected by the fluid medium), and $F_i(t)$ is the net force on cell i , which is given by the sum over all repulsive and adhesive forces to the cell from all the other cells j , F_{ij}^{rep} and F_{ij}^{adh} . Rather than identifying the exact physical origin of these forces, it is assumed that any intercellular force is either a part of F_{ij}^{rep} or F_{ij}^{adh} , and that the parameters in their definitions can be tuned to make the model accurate.

This simplifying assumption is made since there is a myriad of different factors that contribute to the velocity change, including uncontrolled physical forces (e.g., electrostatic forces) and forces induced by cells (e.g., chemotaxis signals sent by APC cells which can attract activated T-cells to certain positions) [9]. Incorporating these phenomena individually into the model would over-complicate it for the purposes of this study. To reiterate, the primary focus of this study is to analyze how morphology deformation influences the cooperative motility of cells in the presence of intercellular forces; tracking the exact time evolution of cell positions is secondary.

Note that the mass of the cells is not included in the velocity equation since all cells are assumed to have a fixed average radius, therefore fixed average mass, and the F_{rep} and F_{adh} parameters can be varied if the average cell mass changes.

F_{ij}^{rep} occurs when two cells are in contact with each other, i.e., when two cells overlap each other, as shown in Figure 2. The force is defined as

$$F_{ij}^{\text{rep}} = \begin{cases} F_{\text{rep}} \frac{d_{ij} - r_{eq}}{r_{eq}} \mathbf{e}_{ij}, & d_{ij} < r_{eq}, \\ 0, & d_{ij} \geq r_{eq} \end{cases} \quad (3)$$

where F_{rep} represents the maximum magnitude of the repulsive force, d_{ij} is the distance between the centers of two cells, r_{eq} is the sum of the radius of the two cells r_i and r_j (i.e., if $d_{ij} = r_{eq}$, then the two circles are in contact at a single point), and $\mathbf{e}_{ij}$ is the unit vector along the line joining the cell centers. Thus, the repulsive force increases as the centers of the cells are closer to each other.

F_{ij}^{adh} occurs when two cells are not in contact. The force is defined as

$$F_{ij}^{\text{adh}} = \begin{cases} 0, & d_{ij} \leq r_{eq}, \\ F_{\text{adh}} \frac{d_{ij} - r_{eq}}{d_{\text{cut}} - r_{eq}} \mathbf{e}_{ij}, & r_{eq} < d_{ij} < d_{\text{cut}}, \\ 0, & d_{ij} \geq d_{\text{cut}} \end{cases} \quad (4)$$

where F_{adh} represents the maximum magnitude of the adhesive force and d_{cut} represents the cutoff distance. Namely, the adhesive force becomes stronger the farther the centers of the cells are away from each other, as long as their distance does not exceed the cutoff distance.

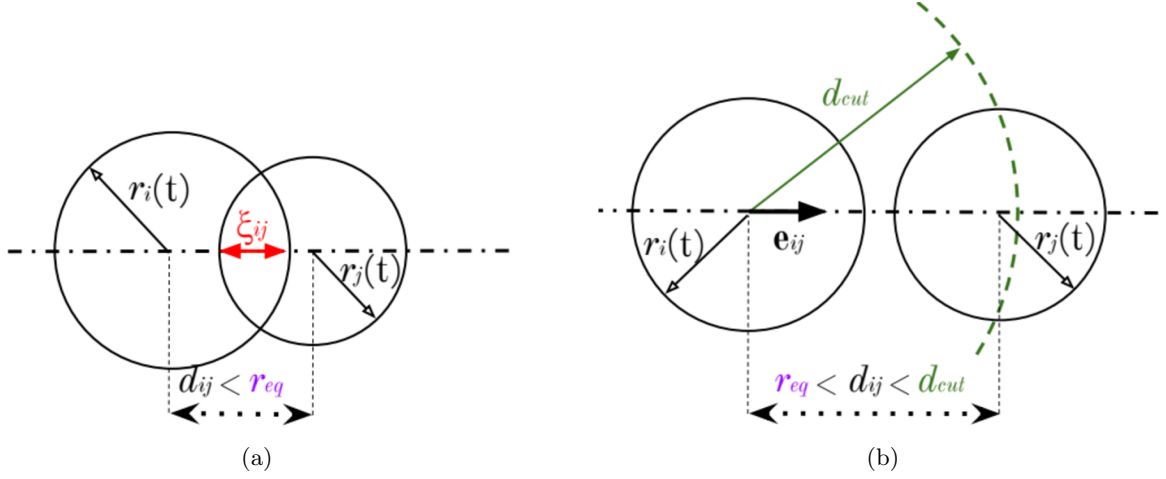


Figure 2: Figures from Wiebke and Parisi [4]. In (a), two cells are in contact with each other due to their overlap, where $\xi_{ij} = r_i + r_j - d_{ij}$ represents the overlap distance. Since $\xi_{ij} > 0$, the two cells exert nonzero F_{ij}^{rep} onto each other, and $F_{ij}^{\text{adh}} = 0$. In (b), the two cells are not in contact and within d_{cut} from each other, so the two cells exert nonzero F_{ij}^{adh} onto each other, and $F_{ij}^{\text{rep}} = 0$.

To model deformation, the cells are assumed to be circles with oscillating radius. Namely, the radius of each cell $r_i(t)$ follows

$$r_i(t) = r_0 + A \sin(\omega t + \phi_i), \quad (5)$$

where r_0 is the initial fixed radius of all cells, ω is the frequency of oscillation, A is the amplitude, and the phase ϕ_i for the i th cell is drawn uniformly from $[0, 2\pi)$.

Note that all equations are non-dimensionalized before simulating them; length is scaled by r_0 and time by $\frac{r_0}{v_0}$, where every cell is assumed to have identical self-propulsion velocity v_0 .

3 Experimental Setup

To simulate the model, 300 cells are generated in a 43 by 43 square grid with a pseudo-periodic boundary. This means that when a cell exists the domain, it re-enters the domain from the opposite side. However, unlike a true periodic boundary, the newly entering cell is considered distinct from the cell which left the domain since it receives a new cell ID upon re-entering. Consequently, if a cell contacts the APC cell, leaves and returns to the domain, and contacts the APC cell again, the simulation considers that two first-contacts have occurred. In other words, the simulation is zoning into a neighborhood of an APC cell within a larger region of T-cells, rather than assume there are repeating square tiles in every direction with an APC cell centered in each square; the pseudo-periodic boundaries are implemented to make the simulation computationally feasible. Consequently, when performing topological data analysis, the boundaries are assumed to be non-periodic.

The simulation is run from $T_i = 0$ to $T_f = 1500$, with step size $dt = 0.005$, along with maximum repulsive force $F_{\text{rep}} = 20$ and maximum adhesive force $F_{\text{adh}} = 0.75$. The cells are all set to radius $r_0 = 1$, so that the average radius of the cells is approximately 1 for all time steps. Cells are allowed to randomly reorient the direction of their movement at some rate λ , and we set $\lambda = 0$ as in analysis done by [4]. The APC cell is fixed at the center of the square grid since its movement is nearly stationary relative to that of T-cells in lymph nodes [10]. An example of this setup is shown in Figure 3, and a summary of parameter values is shown in Table 1. While most of the study used the listed values, whenever noted, certain parts of the study used different values.

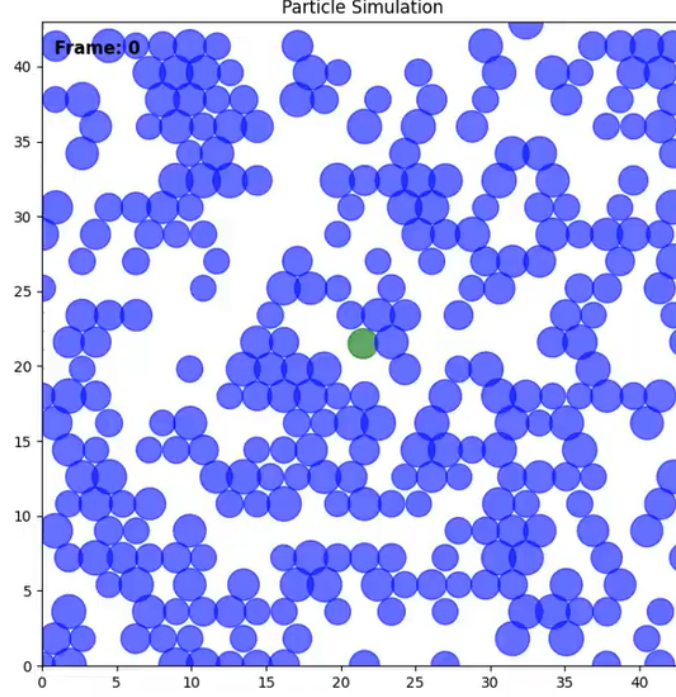


Figure 3: T-cells (blue) initialized in a 43 by 43 square grid. The green cell at the center is the APC cell.

Table 1: Parameter values used in the study.

Parameter	Symbol	Value
Initial time	T_i	0
Final time	T_f	1500
Cutoff distance for the adhesive force	d_{cut}	0.3333
Time step size	dt	0.005
Number of particles	N	300
Average cell radius	r_0	1
Average rate of reorientation	λ	0
Maximum repulsive force	F_{rep}	20
Maximum adhesive force	F_{adh}	0.75

For each realization, the starting position and phase of the cells are randomized. Every time the nondimensionalized time becomes a multiple of 50 ($t = 0, 50, 100, \dots$), a frame of data is recorded, which includes the position, velocity, and radius of the t-cells and the APC cell at that time step. An example of recorded data is shown in Table 2.

Table 2: An example of data output from the T-cell simulation.

Frame at $t = 0.000$					
Cell ID	x-Coordinate	y-Coordinate	x-Velocity	y-Velocity	Radius
1	21.500	21.500	-3.121	0.493	1.000
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

In addition, the first contact time of cells - the time at which each T-cell contacts the APC cell for the first time - is recorded, along with their ID. This is used to compute the scanning rate, the average rate over time in which the first contacts occur (e.g., scanning rate of 5 means on average, 5 T-cells contact the APC cell for the first time per unit time). Namely, the total number of first-contacts is plotted against time, and a linear fitting is performed to the data, as shown in Figure 4. The slope of the line is the scanning rate.

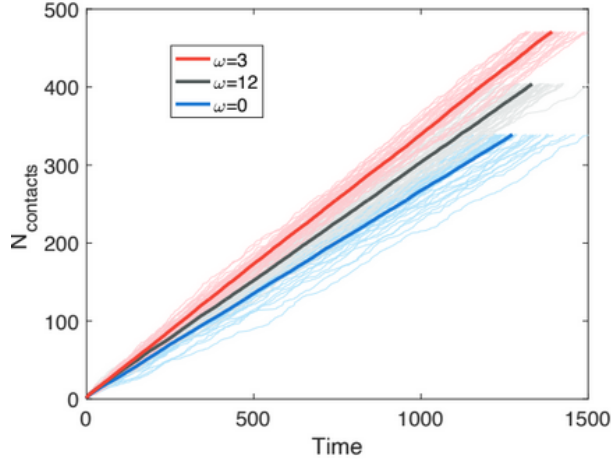


Figure 4: Figure from Wiebke and Parisi [4]. The graph plots the total number of first-contacts over time. The darker lines are the average linear fitting over realizations for each frequency value.

For each ω value, we run 30 realizations. Values of ω ranging from 0 to 30 are simulated to identify the optimal frequency that maximizes the scanning rate. For every frame in a realization, persistence diagrams are generated, and corresponding topological data analysis measures are computed. The details of TDA measures are explained in the following section. The mean of the measures across the frames generates a measure which corresponds to a realization. The mean of the measures across all realizations for a given ω value generates a measure which corresponds to the ω value.

4 Topological Data Analysis (TDA)

4.1 Persistent Homology

Given a well-defined structure, such as a graph of vertices and edges or a continuous manifold, it is relatively easy to identify the topological features present in a data set. However, when given a more irregular data set, such as the one shown in Figure 3, it is not immediately obvious which topological features are present in the data. One approach to this problem is to apply persistent homology to the data set.

To do so, as noted by Adler et al., a threshold value (whose exact nature depends on the specific TDA method) is defined, such that for a given threshold value, a specific manifold is assigned to the data set [11]. Within the manifold, topological features can be identified. For 2-D data sets, there are two types of topological features - dimension 0 features which represent connected components (e.g., vertices and edges) and dimension 1 features which represent loops (e.g., vertices connected into a square). As the threshold value is varied, the manifolds associated with the data set will change, so different topological features will be present for each threshold value.

In this study, threshold values begin at the smallest possible values and are increased continuously. Throughout the process, two distinct topological features can merge into one, thereby resulting in the disappearance of a feature [12]. The creation and disappearance of topological features can be tracked, and birth and death coordinate values (i.e., the threshold values in which a topological feature is created and destroyed, respectively) are assigned to each feature. These birth and death coordinates are graphed on a Cartesian plane, and these graphs are referred to as persistence diagrams. For a 2-D data set, two persistence diagrams are generated - one for dimension 0 features and another for dimension 1 features.

In this study, two methods of persistent homology, alpha complex and Euclidean distance transform, are applied to the T-cell data for single-species analysis (where every cell is considered the same species of data). Each method uses different thresholding techniques, so topological features which can be detected by both methods are more likely to be significant features of the data rather than noise. Specific examples of persistent homology are shown for each method.

4.2 Alpha Complex

Alpha complex is a graph-complex-based persistent homology method that takes a point cloud as an input and builds a simplicial complex depending on the threshold value α [13] [14]. In other words, it constructs a graph-like object (specifically a type of hypergraph) to build topological features with increasing α value. Specifically, alpha complex begins with vertices at $\alpha = 0$ (where the number and position of vertices can be chosen by the user) and has circles with radius α centered at each vertex; as the threshold value α increases, the circles grow in size, as shown in Figure 5. Once two circles intersect, an edge is created to connect their centers.

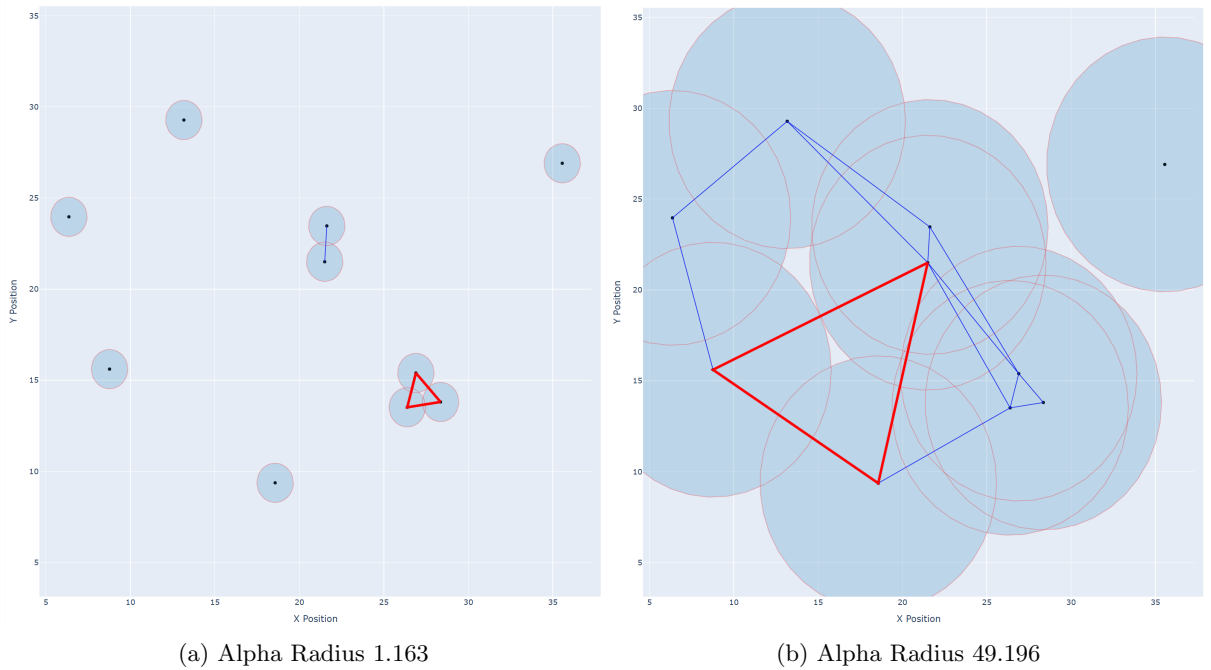
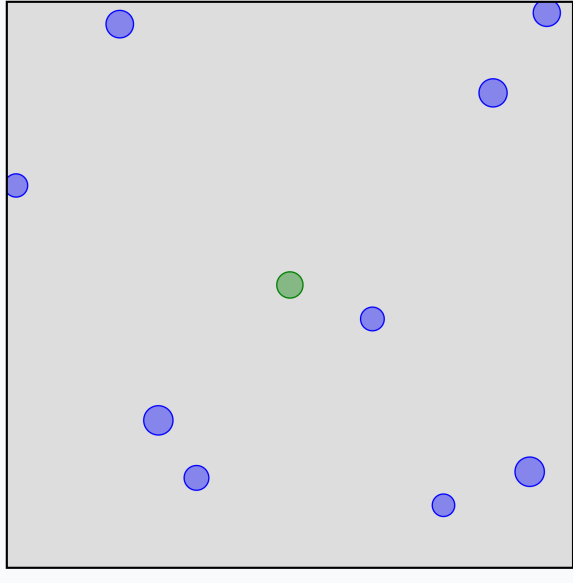


Figure 5: Alpha complex operating on a point cloud at two different α values.

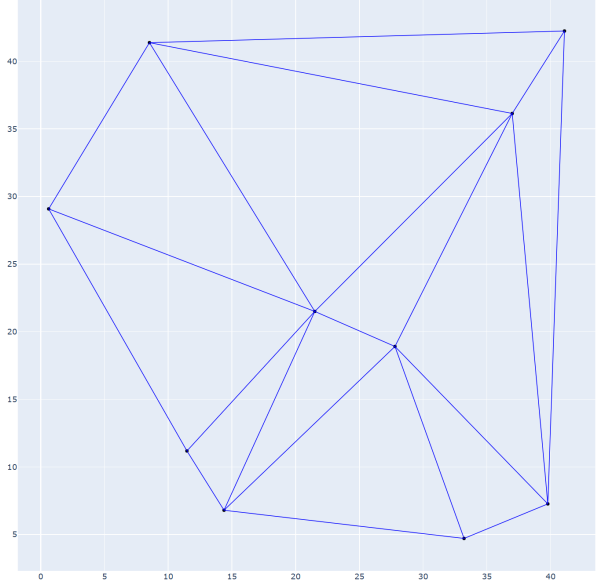
Dimension 0 features are always born from the start at $\alpha = 0$, when the vertices appear. When an edge forms between two vertices, as in Figure 5a, one of the features dies as there is now only one dimension 0 feature (the edge). This feature is then plotted on a persistence diagram, based on the α value the feature was created and destroyed. It should be noted that there is always one dimension 0 feature that never dies (as everything becomes connected to one graph at a sufficiently high α value), which is not plotted on the persistence diagram. Intuitively, the lifetime (death α value - birth α value scaled by a constant) value of dimension 0 features can be interpreted as how far each vertex is away from its nearest neighbor.

Dimension 1 features are born when a loop is formed at some α value, as shown in Figure 5b (the red outline). They then die when the radius of the circle grows sufficiently large to the point where the expanding circles fill the interior of the triangle.

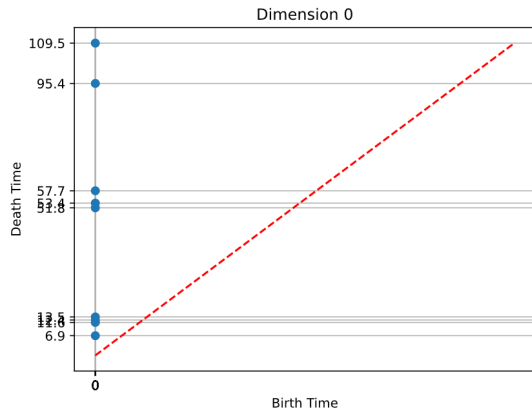
In this study, we consider the cell centers (i.e., circle centers) as points of a point cloud to input into alpha complexes, as shown in Figure 6a and Figure 6b. For both dimension 0 and dimension 1 features, persistence diagrams are generated as α increases, as seen in Figure 6c and Figure 6d. To confirm the accuracy of the persistence diagrams, consider 6e and 6f, where a dimension 1 feature is born at $\alpha = 89.749$ and dead at $\alpha = 103.243$. In Figure 6d, this precise coordinate, $(89.749, 103.243)$, is present, thereby confirming the validity of the TDA.



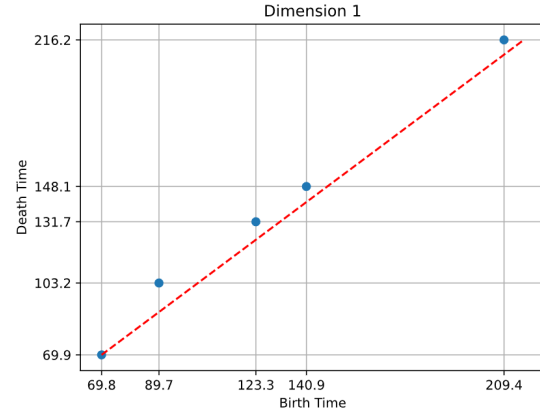
(a) Cell frame example



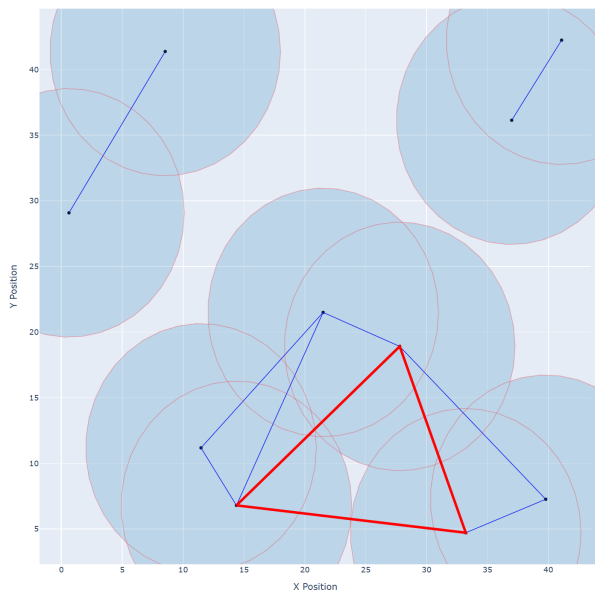
(b) Simplicial complex constructed from cell centers



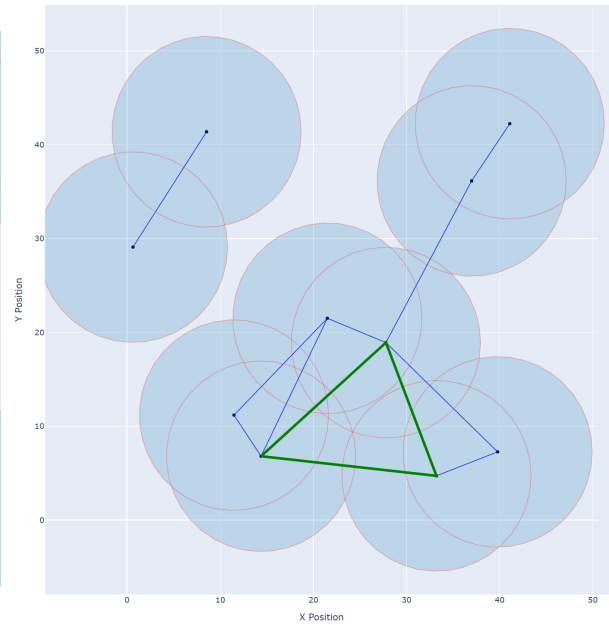
(c) Persistence Diagram for Dimension 0 features



(d) Persistence Diagram for Dimension 1 features



(e) Dimension 1 feature birth at 89.749 (red)



(f) Dimension 1 feature death at 103.243 (green)

Figure 6: Alpha complexes performed on 10 cells

The GUDHI library was used to perform alpha complex analysis in this study [15].

4.3 Distance Transform (DT)

Distance transform takes a binary image as its input, and it performs a distance transform on every pixel [16]. In this study, the transformation and subsequent TDA were performed using the HomCloud library [17]. For every frame, the information regarding the position and radius of cells are used to produce a corresponding binary image, where the colors distinguish the interior and exterior of cells. Every pixel is then assigned a scalar value which is equal to the distance from the pixel to its nearest boundary between the two colors [18]. There are multiple different choices for the distance metric, including Euclidean, taxicab, and max norms. For this study, a signed Euclidean distance transform (EDT) was performed (thus using the Euclidean distance metric), where pixels at the interior of the cells were assigned negative distance values, to differentiate the region inside and outside the cells. Once the distance transform is complete, the original binary image is transformed into a distance height map.

An example of distance transform is shown in Figure 7. Note that while the Euclidean distance was used as the metric in this study, the example below uses taxicab (L_1) norm since it is easier to visualize on a low-resolution grid.

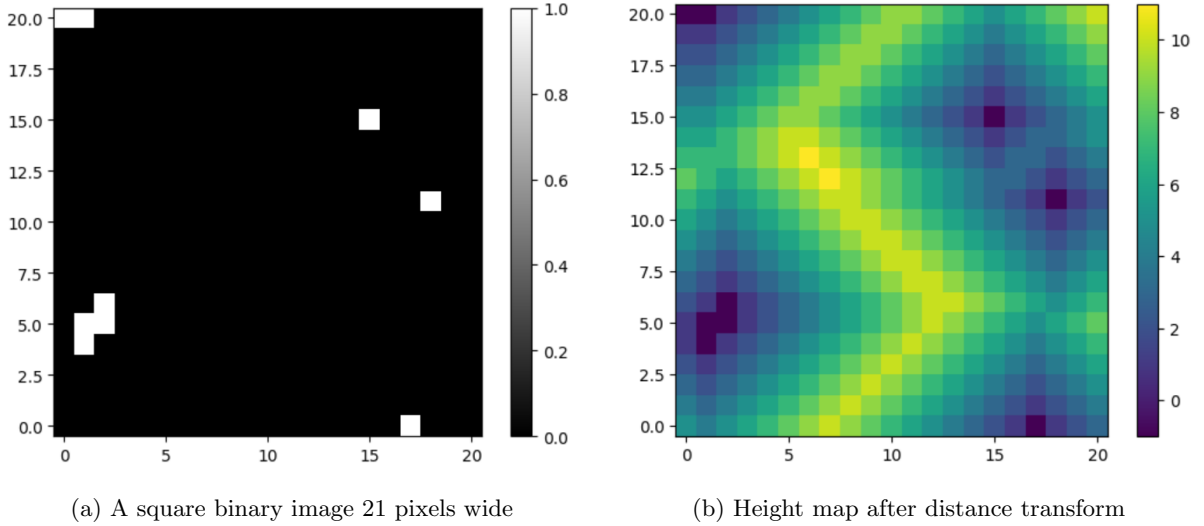


Figure 7: (a) A binary image of five T-cells, where the interior of the cells is white. For this study, square binary images 1000 pixels wide were used. (b) A height map, where each pixel contains the signed DT value.

The height map is then scanned by thresholding the height (i.e., the DT value). In this study, sublevel filtration was applied by scanning the height values from the smallest value to the largest value. Before scanning, there is an empty grid of pixels, referred to as the mapped grid, where each pixel corresponds to the pixel in the height map with the same coordinates. Once the scanning begins, pixels whose height values are less than the threshold value appear on the mapped grid. The scanning process is visualized in Figure 8.

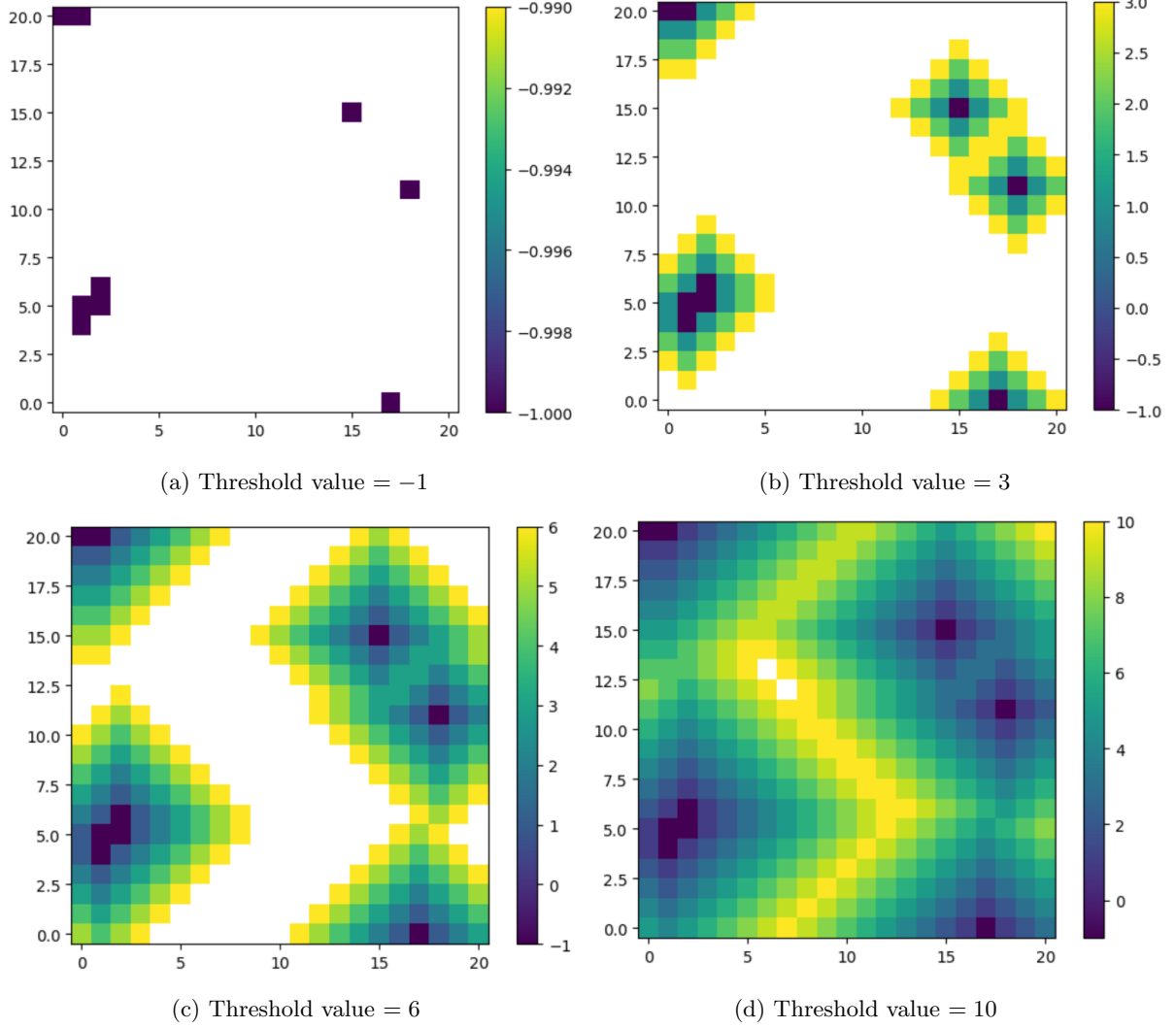


Figure 8: The mapped grid at different threshold values.

First, the pixel with the smallest height value appears on the mapped grid. This pixel starts the dimension 0 topological feature which will persist throughout the entire scanning process (as other features are connected to it, they die out and get absorbed to it). Consequently, this dimension 0 topological feature is not plotted on the persistence diagram, for it does not have a proper death coordinate. If multiple pixels have the minimum height value, then only the one which starts the all-absorbing dimension 0 feature is absent in the persistence diagram; every other feature which has a proper death coordinate is recorded.

In this example, the local minima that first appear on the mapped grid are the centers of cells, which start dimension 0 features (Figure 8a). As the filtration threshold increases, more pixels appear on the mapped grid, which enlarges the dimension 0 features as balls (in the chosen metric sense). Eventually, some of these features become connected to each other, in which case some of them die out. In Figure 8b, two dimension 0 features connect when the threshold value is 3. When the threshold value is 4, as shown in Figure 8c, another dimension 0 feature becomes connected to the joined feature. In each instance, a dimension 0 feature is eliminated.

It should be noted that in addition to the centers of cells, points which lie at the intersecting regions of cells are local minima in the distance transform, as shown in Figure 9. Consequently, for DT, the number of dimension 0 features/ generators will in general not be equal to the number of cells. This is contrary to alpha complex which is guaranteed to have the same number of dimension 0 features/ generators as the number of cells by design (every generator is chosen to be a center of a cell).

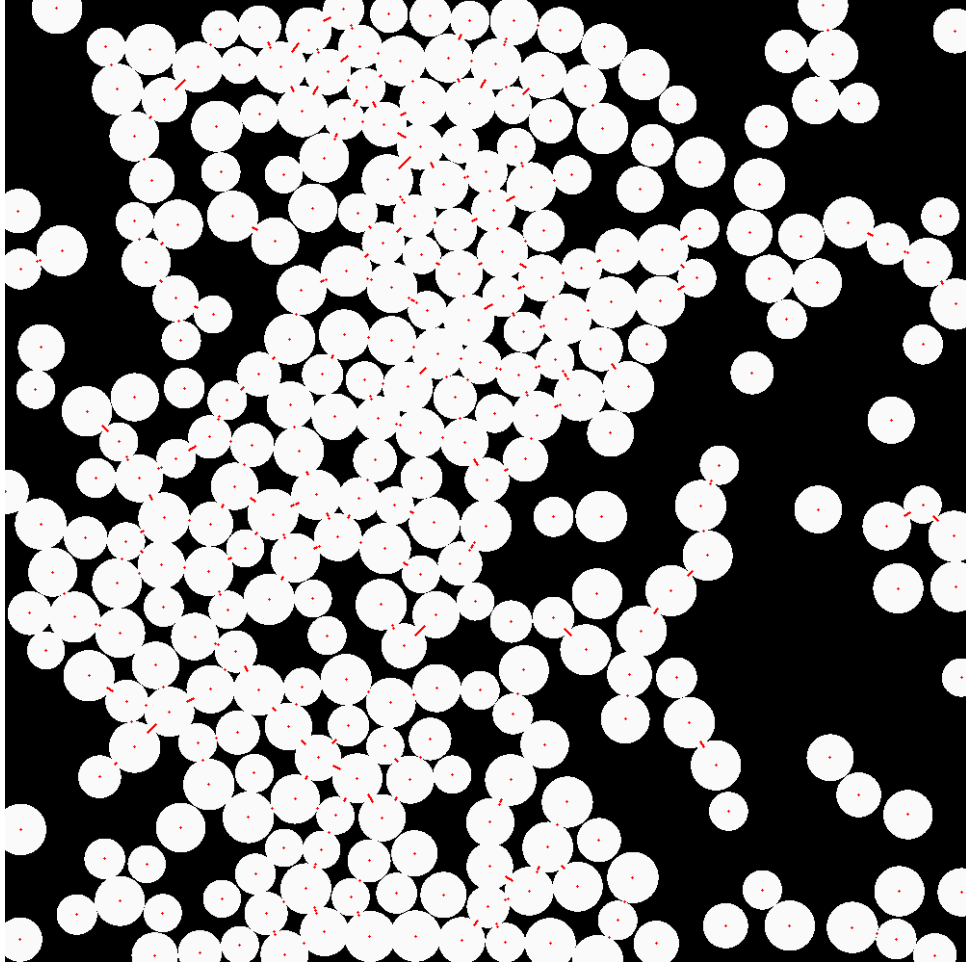


Figure 9: The red dots represent the birth positions of dimension 0 features using EDT, which are also local minima in the distance transform image. The birth positions are overlotted on top of the binary image of T-cells.

Returning to the thresholding process, dimension 1 features can emerge as some of the manifolds are connected in a loop, and they can die out as the mapped grid fills up and severs the loop. In Figure 8d, a loop is created when the threshold value becomes 10. The loop is eliminated when the value becomes 11, thereby filling the two empty pixels inside the loop. Once the filtration threshold reaches the global maximum distance, the scanning stops.

Persistence diagrams are then generated to plot the birth and death coordinates. For the example from Figure 8, the persistence diagrams are shown in Figure 10, which is consistent with what happens during the scanning process.

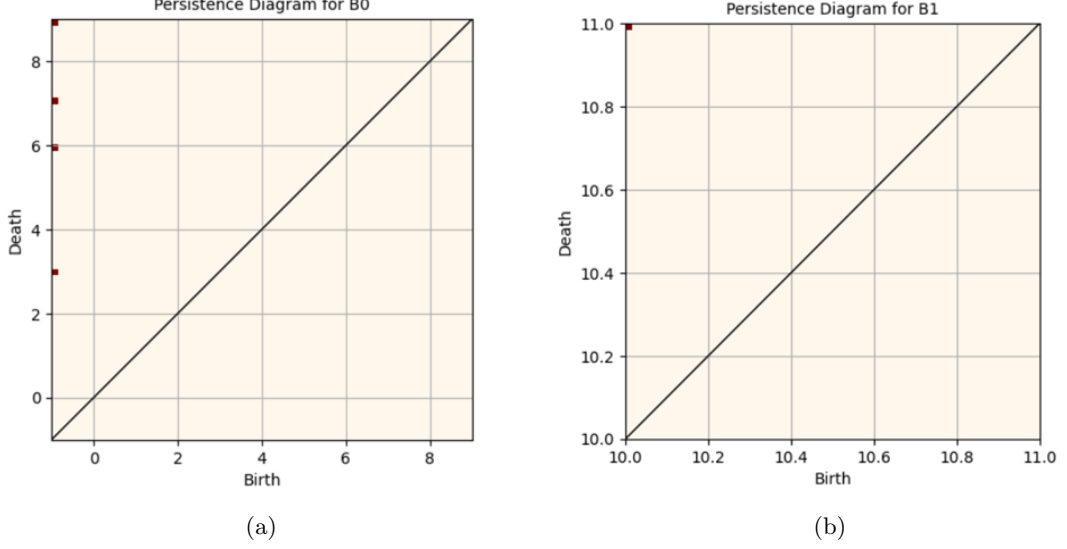


Figure 10: Persistence diagrams for the example from Figure 8.

In summary, when applying DT to the T-cell model, dimension 0 features are born at the centers and intersecting regions of cells, and dimension 1 features connect the dimension 0 birth positions into a loop.

4.4 Lifetime/ Lifespan

Persistence diagrams resulting from persistent homology methods result in a set of birth and death coordinate values. For a birth-death coordinate pair, a lifetime/ lifespan measure can be calculated by subtracting the birth time from the death time. To give this value a more visual meaning on the persistence diagram, this value is divided by $\sqrt{2}$ to get the distance of each birth-death point from the $y = x$ line.

4.5 Choosing the Optimal Image Resolution

For DT, the resolution of the binary image determines its computational cost. Consequently, to choose the optimal image resolution, the following test was performed.

EDT was applied to a sample T-cell data for $\omega = 3$. For a given image resolution (n by n), there were 30 realizations. For each realization, the mean lifespan of dimension 0 and 1 features for each frame was recorded, thereby forming a vector of mean lifespan values per dimension. This vector was averaged across the realizations, so that each image resolution had a lifespan vector associated with it for each dimension.

To quantify the change in TDA measures, the Euclidean distance between the mean lifespan vectors of two consecutive image resolution values was computed. The distance was then divided by the Euclidean norm of the mean lifespan vector corresponding to the lower resolution value to obtain the relative change in the lifespan measure. In summary, the relative change in lifespan between consecutive resolutions at the resolution n_i , $\Delta L_{(rel)}(n_i)$, followed:

$$\Delta L_{(rel)}(N_i) = \frac{\|L(n_{i+1}) - L(n_i)\|_2}{\|L(n_i)\|_2} \quad (6)$$

where $L(n_i)$ represents the mean lifespan vector corresponding to an image resolution of n_i by n_i .

Figure 11 shows the test result, where the relative change in lifespan is plotted across image resolutions:

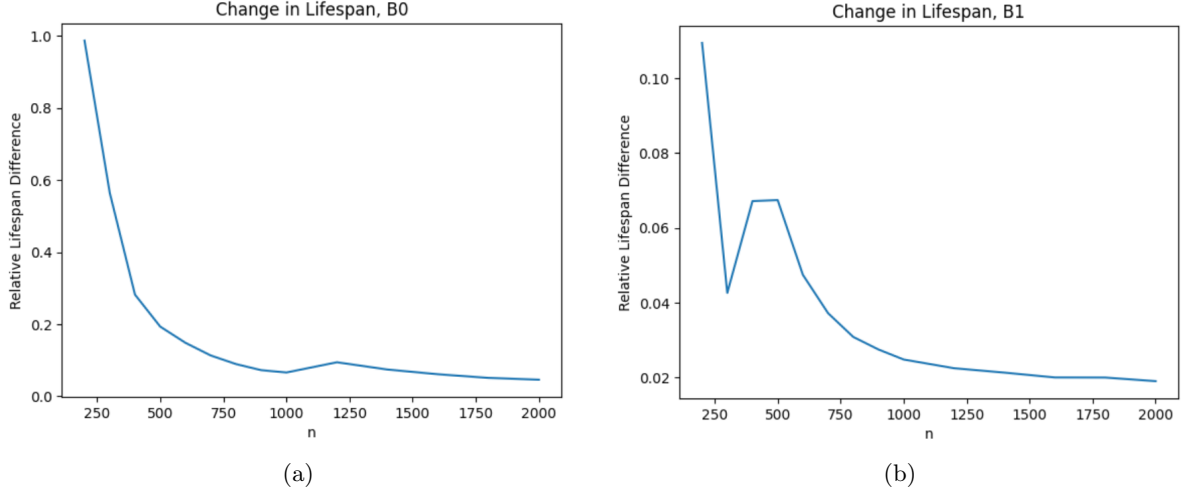


Figure 11: Convergence of the EDT lifespan measure as the image resolution, n , increases for dimension 0 (a) and dimension 1 (b) features.

The mean lifespan measures converge at around $n = 1000$, so 1000 by 1000 grids were used for the EDT analysis in this study.

5 Results and Discussion

5.1 Force Dynamics and Topological Consequences

For both alpha complex and Euclidean distance transform analysis, the mean lifespan of dimension 0 features is positively correlated with frequency, as seen in Figure 12a and Figure 12c. This is likely due to the force dynamics that depends on frequency. In particular, when cells are in contact but not overlapping, they experience zero adhesive and repulsive force. If cells do not oscillate their radius, they are more likely to slide past each other due to the lack of forces. However, if their radius does oscillate, then the force equilibrium can be broken more easily. If the radius shrinks away, then the adhesive force can pull the cells together. If the radius grows and an overlapping region forms, then the cells can push each other away.

Without loss of generality, assume the radius initially grows and an overlapping area forms. This creates a repulsive force which pushes the cells away from each other and terminates their overlap. The termination of the overlap creates an adhesive force between the cells, as long as the cells are within the cutoff distance from each other. The adhesive force then pulls the cells towards each other until they overlap again.

Overall, cells that oscillate undergo a cycle of pushing and pulling each other which effectively locks them in place; the cells are always in motion, but their range of motion is restricted to a certain distance from each other. In this study, this phenomenon is referred to as "position locking". It should be noted that position locking is not guaranteed in a crowded environment since intercellular forces from other cells can break the position locking between two cells. Rather, cells which oscillate more frequently are more likely to break equilibrium and stay position locked with each other for longer periods of time, thereby increasing the average distance between them. This then increases the mean lifespan of dimension 0 features at higher frequencies since it will take a larger threshold value to connect the cell centers.

For dimension 1 features of EDT and alpha complex, beyond a certain frequency value, the mean lifespan decreases as the frequency increases (Figures 12b and 12d). This is also due to position locking. At higher frequencies, position locking between T-cells becomes stronger (it becomes more common among cells and lasts longer), which induces a regular grid structure between the centers of the cells. This occurs since a regular grid structure maximizes the number of position lockings that can occur within the simulation domain. An example of this is shown in Figure 13. In contrast, at lower frequencies, there

is less position locking, and cells can slide past each other more easily, thereby forming a more irregular structure of cell centers.

Cell centers which lie in a regular grid structure tend to be equidistant from each other (since there is no obvious directional preference from the force laws), and the distance to non-adjacent cell centers tends to be minimized. In the case of Figure 13, the square minimizes the diagonal length compared to other quadrilateral grids. Since dimension 1 features are terminated once the diagonals become connected, minimal diagonal distance minimizes their lifespan. Consequently, as the frequency increases, cell centers tend to form regular grid patterns, and the mean lifespan of dimension 1 features decreases.

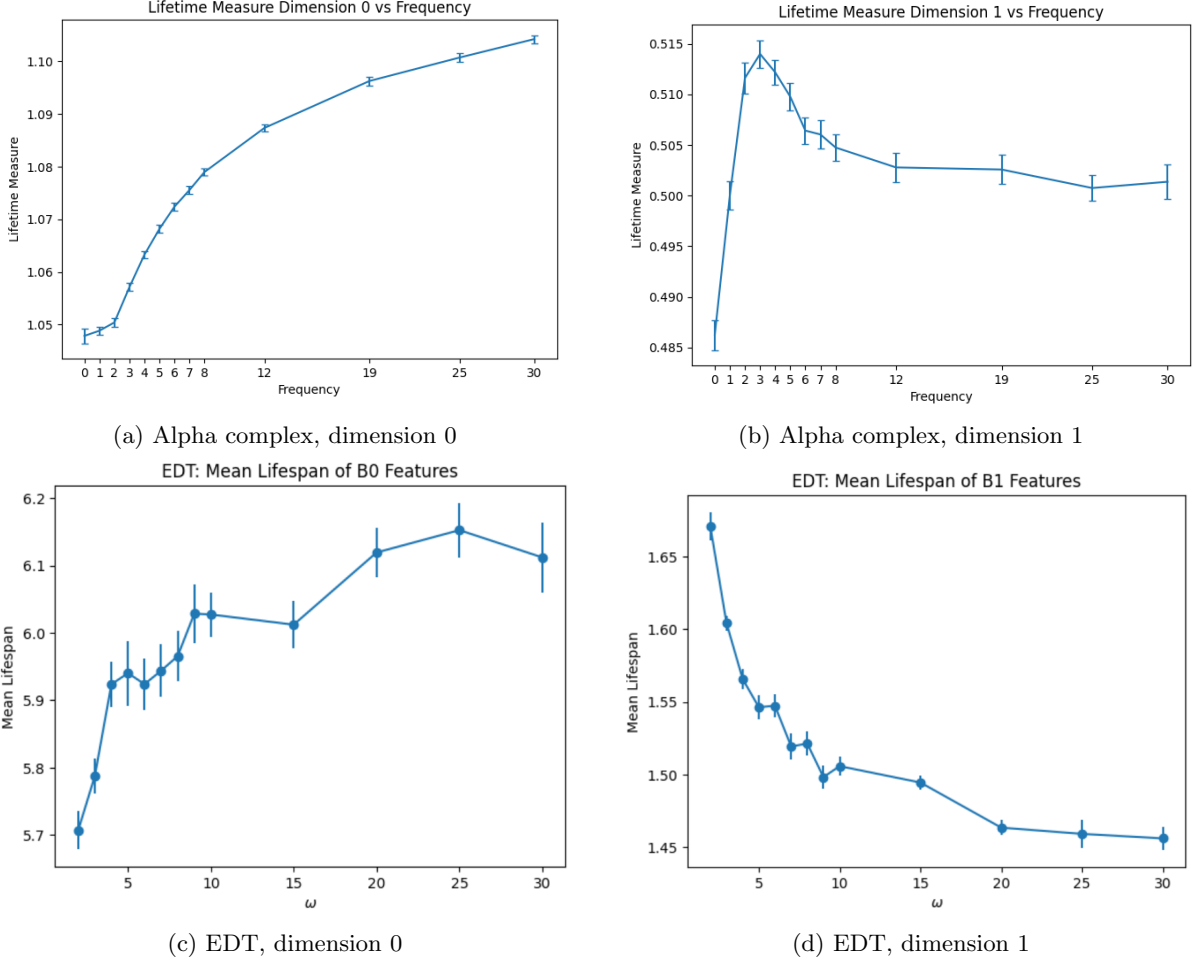


Figure 12: Mean lifespan of dimension 0 and 1 features across frequency using alpha complex (top) and Euclidean distance transform (bottom). Error bars indicate sampling variance across realizations.

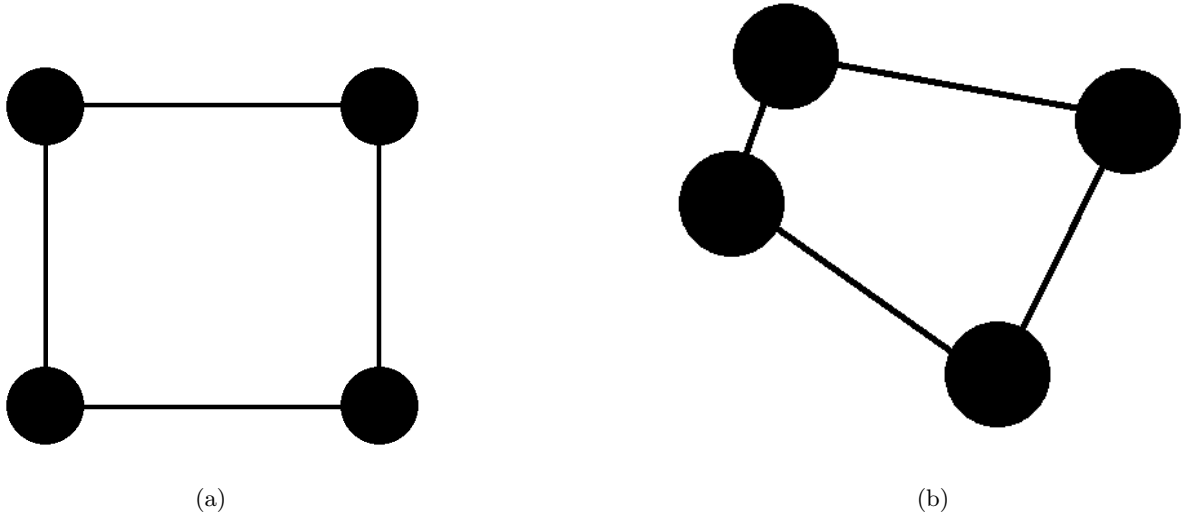


Figure 13: (a) Cells in a regular grid structure. (b) Cells in an irregular grid structure.

While both EDT and alpha complex measures display the same limiting behavior at high frequency values, they have a crucial difference. In contrast to EDT, the alpha complex method shows that the mean lifespan of dimension 1 features reaches a local maximum at $\omega = 3$ (Figure 12b). Notably, this is identical to the optimal frequency for the scanning rate of T-cells [4]. This implies that the topological lifetime measure resulting from alpha complex could be connected to the scanning rate of T-cells. In the context of the TDA measure, at the optimal frequency $\omega = 3$, dimension 1 loops formed via the simplex construction last longer on average. Topologically, this implies that cells form more loop-like structures at the optimal frequency. This also implies that cells tend to cluster more often at non-optimal frequency values, which decreases the presence of dimension 1 loops within the system.

To explain this, consider Figure 14. According to TDA, the larger mean lifetime value at the frequency of 3 implies that on average, the system at the optimal frequency contains dimension 1 holes which are topologically most apparent. From Figure 14b, at $\omega = 3$, there are more dimension 1 holes which are visually more apparent (i.e., empty space surrounded by cells lying in a loop), which corresponds to the large mean lifespan of dimension 1 features at the optimal frequency.

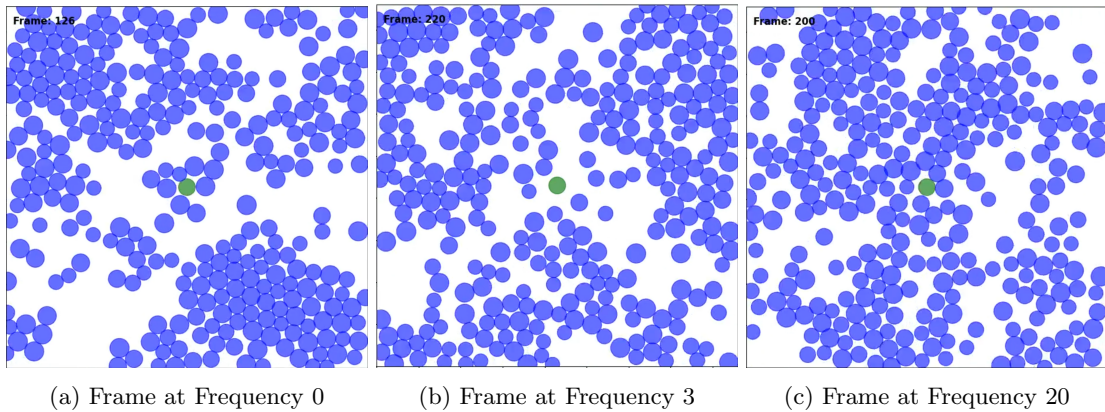


Figure 14: Frames of cells at different frequencies

Scanning-rate wise, the peak at frequency 3 comes about due to the balancing of extreme dynamics at extremely low and high frequency values. At lower frequency values, cells tend to clump together and stick, as their oscillation is not fast enough to break the force equilibrium, thereby causing them to slide past each other with near zero net intercellular force. The sliding motion is inefficient at making cells move past obstacles (e.g., other cells), which lowers their cooperative motility. In contrast, at higher frequency values, cells oscillate so fast that they get locked in position with other cells, which causes them to be effectively stuck in place. Since the cells cannot move around effectively at extreme frequency values, the scanning rate is relatively low for both cases.

At the optimal frequency 3, the two extreme behaviors balance out. The cells' oscillations induce position locking which creates sufficient space between cells and prevent sliding. Simultaneously, the oscillations are slow enough to allow cells to move past each other, rather than make them become completely locked in place. This allows the cells to move past each other with minimal resistance, which increases the overall scanning rate since more cells can contact the APC cell within the time limit.

Moreover, the balancing of the two extreme behavior creates ample space between cells. Notably, the space is spread throughout the domain, rather than gathered in one place; as can be seen in Figure 14, there is noticeable space between the cells. In contrast, the empty space is not as evenly distributed at extreme frequency values. Consequently, at the optimal frequency, it is more common for cells to enclose the evenly spread-out empty space, thereby forming loops or dimension 1 features with the longest average lifespan. Therefore, dimension 1 holes are topologically significant evidence of the cooperative motility of T-cells.

Overall, the force dynamics between T-cells explains the connection between scanning rate and topological measures of T-cells.

5.2 Verifying the Force Dynamics

The explanation of results from the previous subsection assumes that the trend in the scanning rate and topological features of T-cell positions over frequency arise from force laws governing the system which induce position locking. To verify this, the force parameters, F_{adh} and F_{rep} which represent the maximum strength of the adhesive and repulsive forces respectively, were modified and simulated; if the force dynamics explained the connection between scanning rate and TDA measures, then both measures should significantly change with the new simulation. Namely, the repulsive force was set to zero while the adhesive force maintained the same value from the previous simulation. TDA was performed on the resulting data using alpha complex. Results from the new simulation is shown in Figure 15.

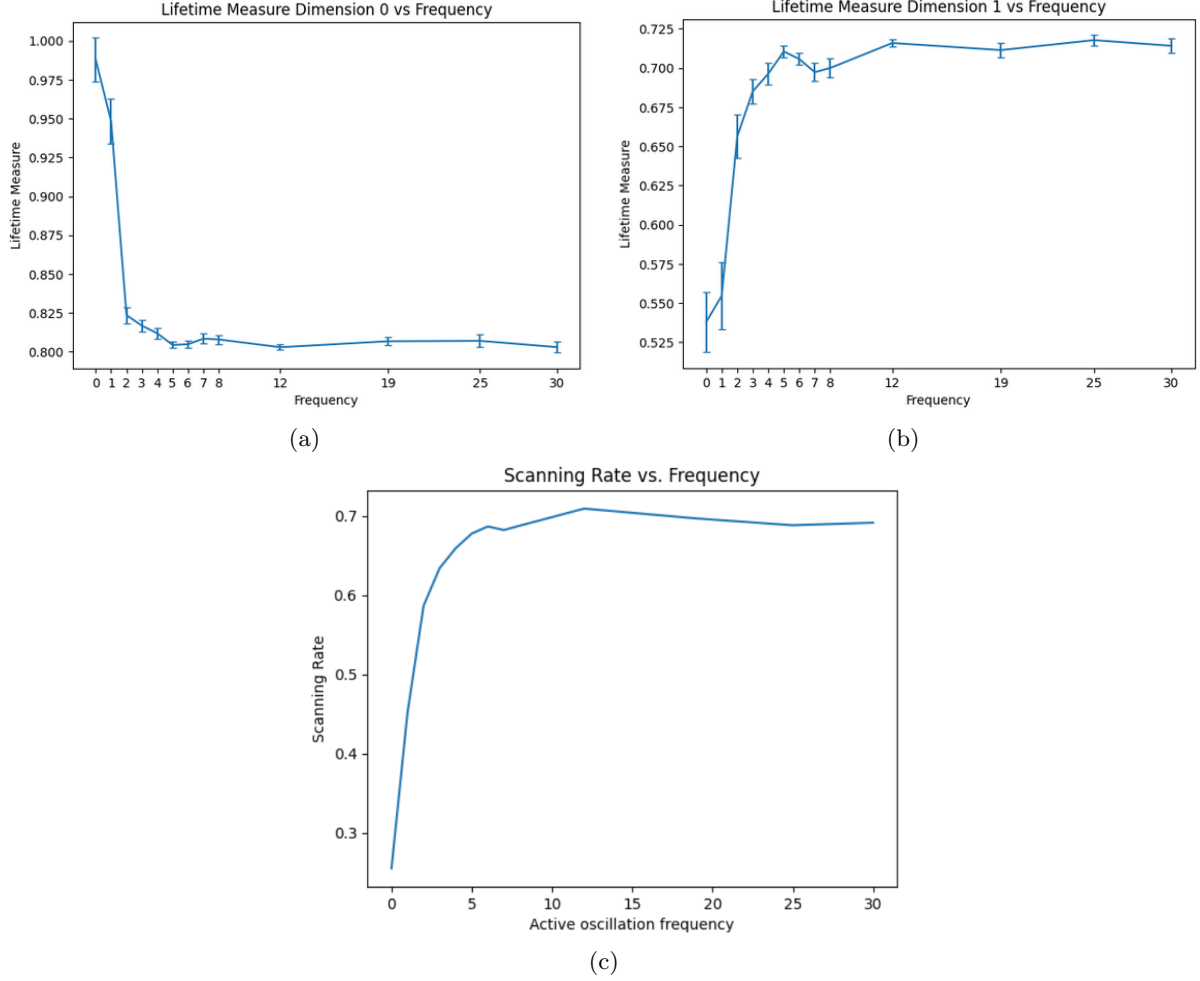


Figure 15: (a) and (b) Mean lifespan of dimension 0 and 1 features using alpha complex, (c) scanning rate across frequencies ("active oscillation frequency" is identical to "frequency", i.e., the frequency at which cells oscillate their radius). $F_{adh} = 0.75$ and $F_{rep} = 0$.

The new results display two notable outcomes. First, the trend in both TDA measures and scanning rate over frequency is significantly different from the original simulation (Figure 12); there is no statistically significant local extrema. This is because position locking is impossible without the repulsive force - position locking occurs due to a cycle of pushing and pulling which only preserves the radial component of velocity (i.e., the component parallel to the line connecting the centers of the cells), so it is impossible without the "pushing" part of the cycle. Since the optimal frequency from the original simulation originates from the balancing of two opposite phenomena (sliding at low frequencies and position locking at high frequencies), lacking one of the phenomena prohibits the existence of an optimal frequency for scanning rate. Moreover, this implies that there is no optimal frequency for cooperative motility, so there is no optimal frequency for creating holed space between cells. Consequently, there is also no optimal frequency for the TDA measures.

Second, there is a similarity between the trend over frequency in dimension 1 mean lifespan and scanning rate. This is also true for the original simulation, as shown in Figure 16. Notably, the similarity is most apparent at lower frequency values (approximately 12 or less); for the new simulation, the two trends even agree with the relatively insignificant (statistically insignificant at the current sample size) local peaks at $\omega = 5$ and $\omega = 12$. Consequently, there is a potential connection between scanning rate and mean lifespan of dimension 1 features.

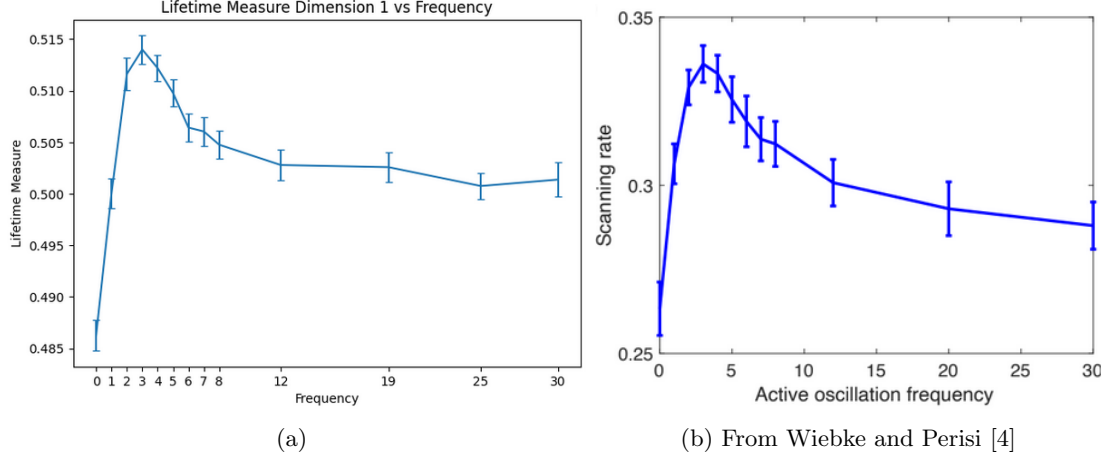


Figure 16: (a) mean lifespan of dimension 1 features using alpha complex, (b) scanning rate for the original simulation.

Overall, results from the new simulation are consistent with the hypothesis that the force dynamics between T-cells determine their topological features and scanning rate. Moreover, it indicates that the mean lifespan measure of dimension 1 features using alpha complex could be an alternative measure for the cooperative motility of T-cells. This is likely because apparent dimension 1 holes/ loops (i.e., large mean dimension 1 lifespan) are evidence of effective cooperative motility (i.e., large scanning rate); cells which move around efficiently must have sufficient space between them, and they consequently form loops. Future studies on this topic should perform a similar analysis for different combinations of force values to further test these ideas and compute which ranges of force values result in the existence of an optimal frequency.

Moreover, to verify whether position locking is truly occurring within the T-cell system, future studies should track the velocity vector of the T-cells. Position locking between two cells only preserves the radial component of their velocity, so if position locking is a real phenomenon, then the velocity vector of a T-cell should be nearly parallel to the radial vector to the nearest T-cell. For instance, one can compute the average angle (ranging from 0 to $\frac{\pi}{2}$) between the velocity vector of each cell and the radial vector to the nearest neighboring cell, and observe how the measure changes for different frequency values. If position locking is real, then the measure should decrease to zero at large frequencies, for position locking is expected to be more common and longer-lasting at larger frequency values.

5.3 The Difference Between Alpha Complex and EDT

As noted above, while both alpha complex and EDT TDA measures show identical limiting behavior at high frequencies, there is a significant difference between the methods. Namely, alpha complex is able to detect dimension 1 holes in the cell positions at $\omega = 3$ while EDT is unable to do so. This is because the methods create their dimension 0 generators differently.

For the T-cell data, alpha complex selects the centers of cells as generators for dimension 0 features. Consequently, it is by design efficient at capturing the topological patterns in the arrangement of cells, the biggest contributing factor to scanning rate. In contrast, EDT selects the local minima of the distance transform of the binary image of the cell positions as generators for dimension 0 features. While this includes the centers of T-cells, it also includes points from regions of intersection between cells (as shown in Figure 8d). These additional generators create significant noise for the detection of dimension 1 features since they generate many loops which last very shortly. So, in the case of this study, EDT's ability to capture relatively small-scale features of images prevents it from detecting larger-scale topological features.

5.4 Revised EDT and Alpha Complex

To improve EDT, the process of computing the mean lifespan of topological features was modified. In each frame, rather than sampling the lifespan of every feature, the mean lifespan value was obtained by accounting for only the top p percent of the longest-lasting features in the frame to eliminate the noise features originating from unwanted generators. An example of this filtration is shown in Figure 17. For this study, p was set to 98 since the trend of the lifespan plot changed significantly at that value, as shown in Figure 18. The filtration process was also applied to alpha complex with $p = 50$, as shown in Figure 19.

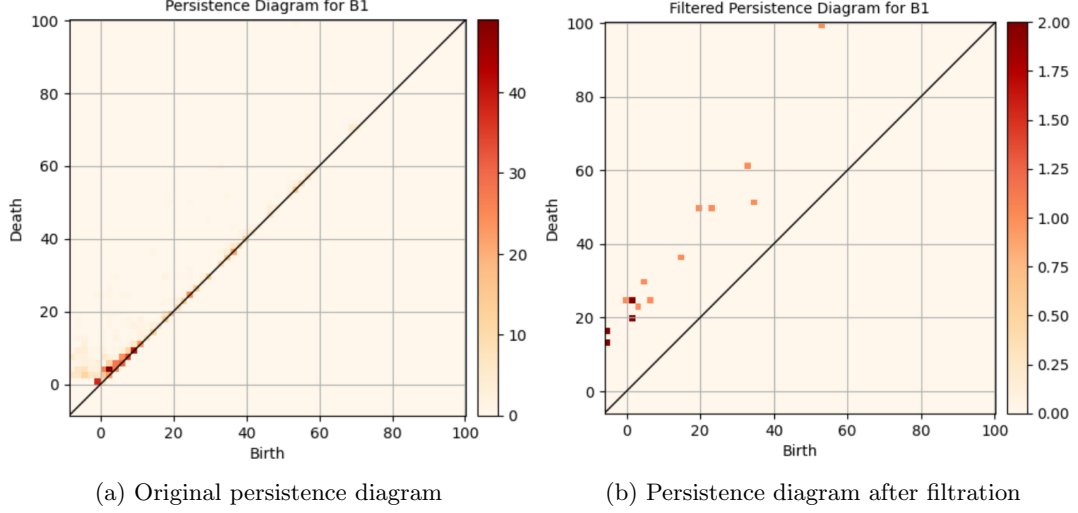


Figure 17: Filtration (keeping only the top 98th percentile of the longest-lasting features) removes the noise features, i.e., the features near the diagonal $y = x$, in the persistence diagram. The color bar represents how many features/ generators share the same birth-death coordinate. The longest-lasting features are nearly invisible in the original persistence diagram since they occur significantly less often than noise features.

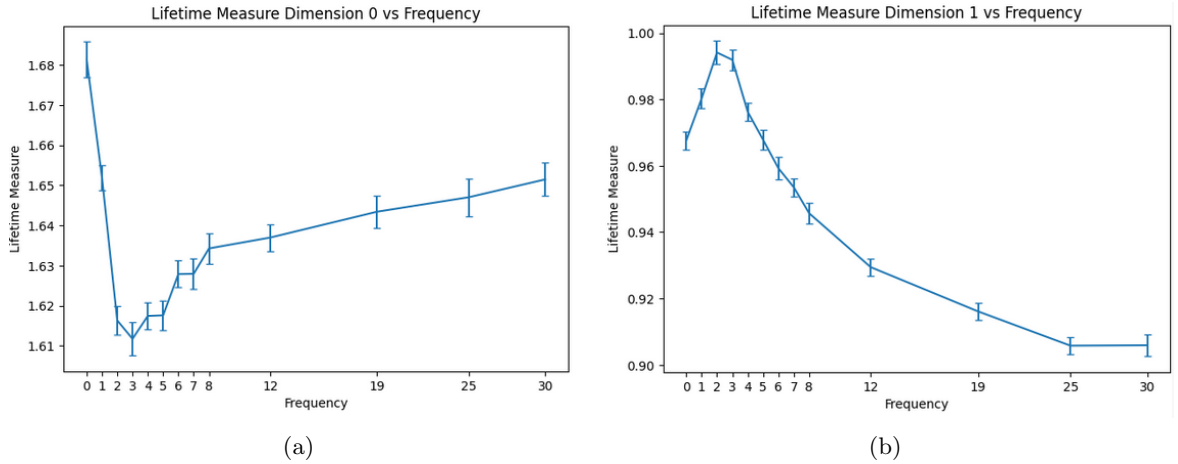


Figure 18: Mean lifetime of dimension 0 and 1 features for EDT across frequency after filtering out 98 percent of features.

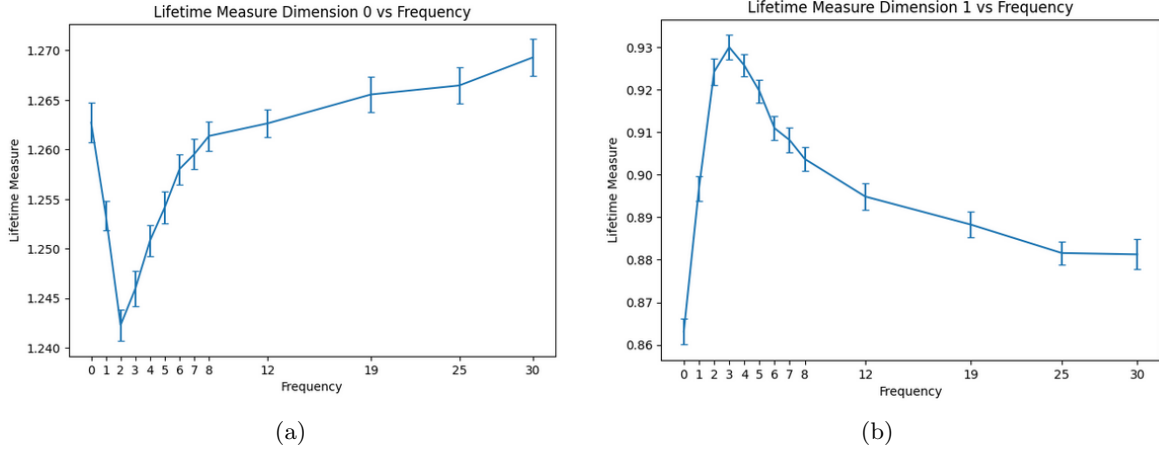


Figure 19: Mean lifetime of dimension 0 and 1 features for alpha complexes across frequency after filtering out 50 percent of features.

As expected, upon removing the noise from irrelevant generators, EDT is able to detect the significant presence of dimension 1 features at $\omega = 3$, just like alpha complexes. This is because the filtration by lifespan only preserves the most topologically significant features, which originate from the dimension 0 generators located at the center of T-cells. Consequently, the filtered TDA accounts for topological features created from cell positions alone, just like alpha complexes.

Moreover, a noticeable local minimum appears at $\omega = 3$ for dimension 0 features of both EDT and alpha complex (Figures 18a and 19a). By only considering the longest-lasting dimension 0 features, the filtered TDA and alpha complex account for only the cells with a large minimum distance to the nearest cell. Consequently, the local minimum at $\omega = 3$ indicates that at that frequency value, the lifespan of such a subset of dimension 0 features is minimized. It is crucial to note that because the filtered TDA considers the position of a only small subset of T-cells, the local force dynamics cannot directly explain the new result.

Instead, this result can be explained by the peak at $\omega = 3$ in Figure 18b. Namely, the creation of holes between cells reduces the overall amount of space that the cells can inhabit within the square domain. With less space available, cells are pushed close to each other, which reduces the average lifespan of dimension 0 features.

The reason why this phenomenon is not reflected within the unfiltered TDA is because cells which are close to each other, especially those which overlap with each other, are unaffected by the creation of dimension 1 holes. This is due to the repulsive force; the lack of space caused by the dimension 1 holes cannot push overlapping cells much closer due to the force, especially since the repulsive force grows as the overlap region increases in area. In contrast, non-overlapping cells with sufficient space around them can be pushed towards other cells much more easily due to the lack of the repulsive force. And since cells with the most empty space around them produce dimension 0 features with the longest lifespan, their experience becomes much more pronounced once only they are considered after the filtration by lifespan.

5.5 Generator Count

When alpha complex is executed on any 2D set of points, the number of dimension 0 features that are generated is known before the execution of the algorithm. This is because the number of dimension 0 features directly corresponds to the number of points in the point cloud input. In contrast, the number of dimension 1 features can vary between data even if the number of dimension 0 generators remains fixed. The average number of generators per realization per frequency value was calculated for each frequency value, and the results are shown in Figure 20.

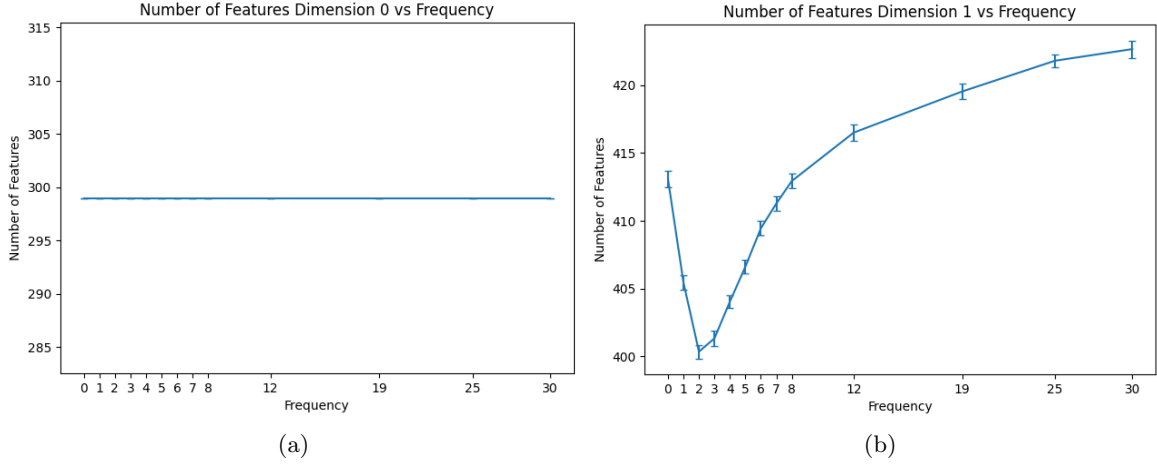


Figure 20: Mean number of dimension 0 and 1 features for alpha complex across frequency.

As can be seen in Figure 20a, the number of dimension 0 generators is always 299 because there are 300 cells, and the center of the APC cell is excluded from the analysis. Notably, however, as can be seen in Figure 20b, a local minimum emerges at $\omega = 3$, which is the optimal frequency for both scanning rate and mean lifetime of dimension 1 features. Future studies should further explore this trend since the number of generators could be another topological evidence of cooperative motility.

5.6 Clustering Analysis

5.6.1 K Means Clustering

TDA analysis was performed on local clusters of cells to test whether the global topological trends were also present in local groups.

To cluster the cells into local groups, the following algorithm was used. Given a set D containing n cell positions of fixed dimension $d = 2$, and a fixed parameter k , the algorithm searches for centroids

μ_k such that the quantity $\sum_{j=1}^k \sum_{x_i \in C_j} \|x_i - \mu_j\|^2$ is minimized. Informally, each k th cluster has a centroid

that is the mean of all the elements contained in it, and the algorithm minimizes the sum of squared distances from all elements in the clusters to their mean.

The algorithm starts by randomly initializing k points, $\mu_k \in \mathbb{R}^2$. Then, all n cells in D are assigned to any one of the k clusters by finding the nearest μ_k and assigning the corresponding label. After this step, the value μ_k is updated for all k clusters to be the mean of each corresponding cluster. All cells in D are reassigned according to the new values μ_k , and the process repeats until the values μ_k converge to a local minimum.

5.6.2 Topological Cluster Analysis

T-Cells exhibiting both oscillatory and static behavior exhibit tendencies to cluster during simulation. While previous metrics describe topological behavior across the entire grid of cells, it is also of interest to analyze how clusters form and what topological behavior occurs within these groups. K-Means is performed on the simulated positions of each T-Cell in the grid for various frequencies, across time, and report the results.

We find that across frequencies and throughout time, cells tend to cluster into $k = 4$ groups across the grid, centered in each quadrant. This relationship is shown in Figure 21, where the silhouette score is highest for $k = 4$, and adding more clusters past this point leads to diminishing returns in minimizing the intra-cluster sum of squares, as shown in the elbow graph. In addition, on average, clusters remain in their respective quadrants across time, as shown by Figure 22. Finally, the topological behavior within clusters appears to be the same as global behavior, as indicated by Figure 23.

From this analysis, we pose that T-Cells, on average, tend to cluster around the 4 quadrants in the grid. This behavior is supported by the inherent geometry of the grid, in that the single APC cell lies statically in the center of the of simulation and cells tend to either stay in their respective quadrants or move around if they perform a contact with the APC. We see that this behavior is consistent across time, and that the local dynamics within each quadrant are very similar to the global dynamics of the entire system.

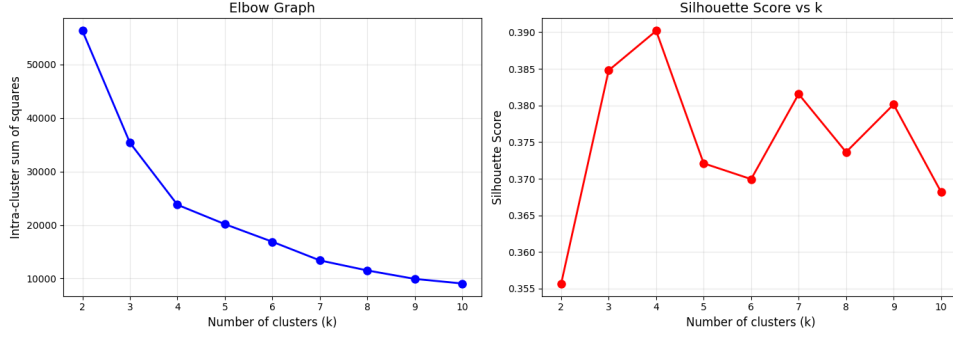


Figure 21: Elbow Graph (1) and Silhouette Score (2) across simulated frequencies and time frames. The elbow in (1) appears to be at $k = 4$, meaning adding more clusters results in marginal minimization of loss. The Silhouette score, a measure from $[-1, 1]$, that measures the distinctiveness of clusters, is also maximized at $k = 4$.

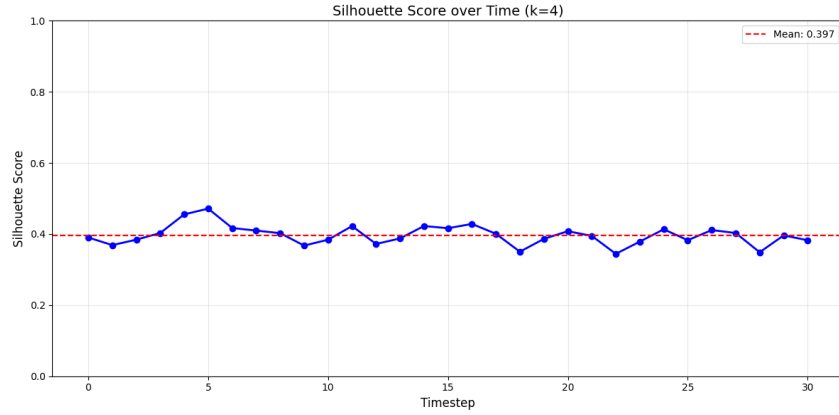


Figure 22: Silhouette Score throughout time for optimal frequency $w = 3$. The quality of cluster distinction does not change throughout time.

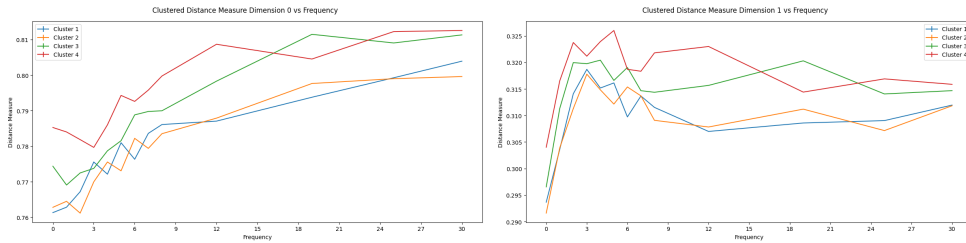


Figure 23: Dimension 0 and 1 measures vs frequency for each $k = 4$ cluster using alpha complex. The local measures appear to be consistent with the global relationship in figure 12a.

5.7 Effect of Cell Density on Topological Features

As T-Cells operate in extremely crowded environments, it is important to understand their behavior in such scenarios. Varying the number of particles, the domain size, and the size of cells all contributes to the changing the density of cells within the region. As in [4], we introduce the area fraction (packing fraction), as a measure from 0 to 1, estimated by taking the approximate area of cells over the area of the grid, or $\frac{N\pi}{L^2}$, that serves as a metric for the density of the system. It has been shown that for intermediate area fractions, the scanning rate of the system is optimized, as cell flow is optimal for intermediate densities, reminiscent of other crowded systems under which flow is studied such as vehicular transport. Since scanning rate has been shown to be correlated with topological measures, it is of interest to see the relationship between area fraction topological features.

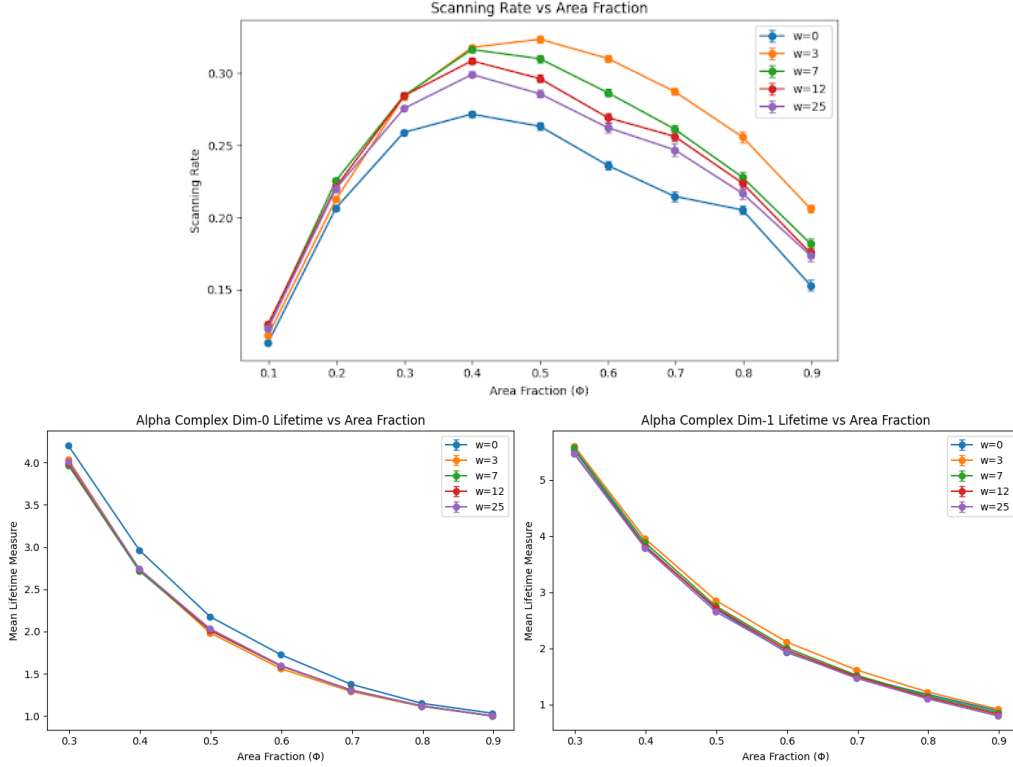


Figure 24: Intermediate area fraction optimizes the scanning rate for all ω values. However, the mean lifetime of dimension 0 and 1 features decreases monotonically as area fraction increases even though for fixed area fraction values, the trend of lifetime measures across frequency can detect an optimal frequency value which maximizes mean lifetime.

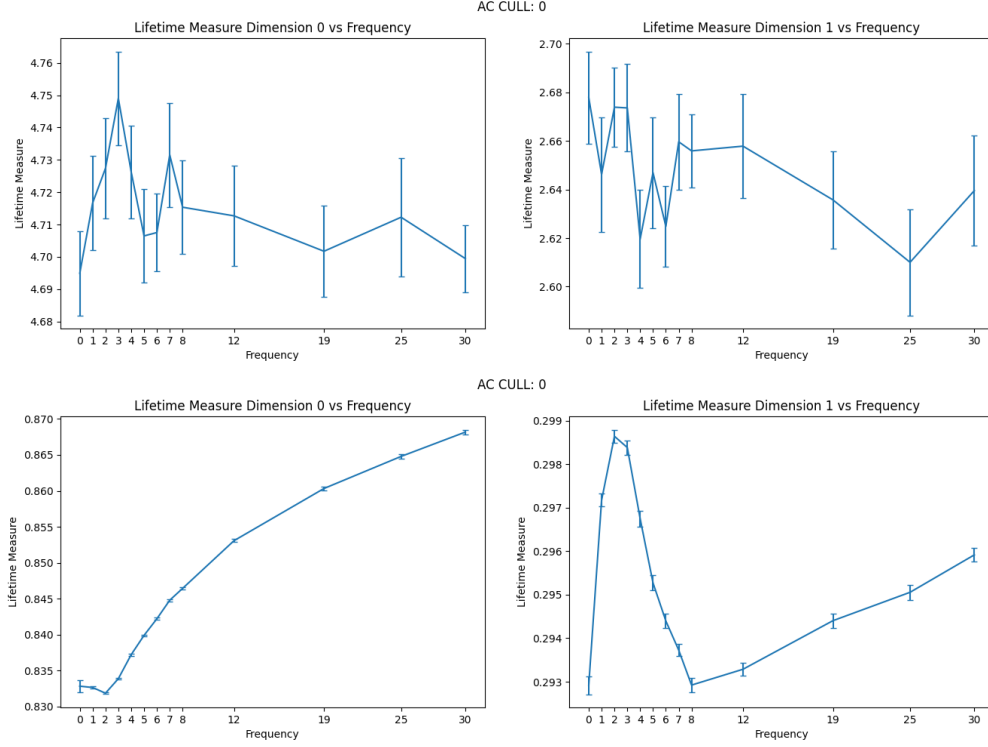


Figure 25: The average lifetime of dimension 0 and 1 features using alpha complex (without filtering by lifetime) for Area fraction $\Phi = 0.1$ (top) and $\Phi = 0.9$ (bottom). The peak in the mean dimension 1 lifetime measure emerges with less standard error for the higher area fraction value.

As shown in Figure 24, the scale of topological features for both Dimension 0 and Dimension 1 trends downwards, even when applying culling/ filtering the generators by lifetime. For a fixed area fraction, the peak of $\omega = 3$ holds, but the general lifetime of the features in denser environments is shorter. In Figure 25, the overall correlation becomes more stable as area fraction increases (without filtering), and the peak is more refined for extremely high area fraction. This shows that the correlation in topological features and frequency is more prominent in crowded environments, and that the biological phenomena explained by such features holds more robustly in crowded environments. This aligns with the intuition for cooperative motility, where T-Cells in a crowded environment help to increase overall flow instead of inhibiting each others' movements.

6 Conclusions

In this study, topological data analysis was utilized to study the dynamics of 2-D T-cells with oscillating radius. Quantitative topological measures were developed to describe and reinforce the phenomenon of cooperative motility of T-Cells via deformation. To do this, a model of oscillating self-propelled particles was simulated, and persistent homology techniques were applied to the resulting data to identify a trend in topological measures over frequency that was significantly similar to the scanning rate. Notably, the mean lifetime of dimension 0 features had a local minimum and that of dimension 1 had a local maximum at frequency $\omega = 3$, which was the optimal frequency for cooperative motility determined by the scanning rate. These results were accomplished by using two different persistent homology methods, alpha complex and Euclidean distance transform, to verify that the topological trends were significant.

In addition, a theory explaining the correlation between topological measures and scanning rate was established. Namely, the force laws of the system caused sliding between cells at low frequency values and position locking at high frequency values, which gave rise to the optimal frequency value as the extreme phenomena were balanced at that value. This theory was tested by changing the force parameters of the cell simulation, which confirmed that the force laws influenced scanning rate and TDA measures, and that the correlation held under different parameter regimes.

Furthermore, it was verified that the trend in the global topological measures also held within local clusters of cells. Finally, it was demonstrated that the scale of the topological measures decreased while overall correlation in measures became more statistically significant for large area fraction, corresponding to the dense environment real T-cells inhabit.

6.1 Future Work

Many directions for future work can be considered regarding the application of TDA to the T-cell system and studying its cooperative motility. Further experiments should be conducted to verify the current theory based on the force dynamics of the system since the theory explains why an optimal frequency exists for scanning rate, and why dimension 1 lifespan measure matches trends in the scanning rate over frequency. In addition to the methods explained prior (varying force parameters and measuring the angle between cell velocity and the radial vector), network-based media methods such as in [19] can also be used to further explain the trends from a non-topological point of view. Also, the simulation should be extended to 3 dimensions to identify if the same trends in scanning rate and TDA measures hold or new behaviors emerge.

In addition to the generator count analysis, additional topological measures should be studied further. In particular, consecutive bottleneck distance - which measures how similar persistence diagrams are between consecutive time frames [20] - and Dowker persistence - which compares topological similarity between two species within the data (e.g., T-cells which have contacted the APC cell and those which have not), [21] - are promising avenues for deeper exploration of the topological properties of the T-cell system. Another promising avenue is studying the simulated system under stochastic cell reorientation, where $\lambda > 0$ unlike in our given simulations ($\lambda = 0$), which gives a more realistic simulation of T-Cells, as they exhibit similar behavior in real life [4]. Finally, further experiments in clustering can be conducted to verify whether global trends in the new measures also hold within local subdomains.

References

- [1] C. J. Janeway, P. Travers, M. Walport, and et al., *Immunobiology: The Immune System in Health and Disease*. New York: Garland Science, 5 ed., 2001.
- [2] P. Bousso, “T-cell activation by dendritic cells in the lymph node: lessons from the movies,” *Nature Reviews Immunology*, vol. 8, no. 9, pp. 675–684, 2008.
- [3] M. J. Miller, S. H. Wei, I. Parker, and M. D. Cahalan, “Two-photon imaging of lymphocyte motility and antigen response in intact lymph node,” *Science*, vol. 296, no. 5574, pp. 1869–1873, 2002.
- [4] L. E. Wiebke and D. R. Parisi, “Active deformation enhances target scanning in crowded two-dimensional self-propelled particle simulations,” *Rev. Mod. Phys.*, October 2025. Dated: October 24, 2025.
- [5] R. M. A. Vroomans, A. F. M. Marée, R. J. de Boer, and J. B. Beltman, “Chemotactic migration of T cells towards dendritic cells promotes the detection of rare antigens,” *PLOS Computational Biology*, 2012.
- [6] S. Celli, M. Day, A. J. Müller, C. Molina-Paris, G. Lythe, and P. Bousso, “How many dendritic cells are required to initiate a T-cell response?,” *Blood*, vol. 120, no. 19, pp. 3945–3948, 2012.
- [7] I. M. N. Wortel, J. Postat, M. Mihaylova, M. Merino, A. Bhagath, M. Harris, L. Wouters, L. Wiebke, D. R. Parisi, J. N. Mandl, and J. Textor, “Cooperative motility emerges in crowds of T cells but not neutrophils,” *bioRxiv*, 2024.
- [8] M. A. Porter, M. Feng, and E. Katifori, “The topology of data,” *Physics Today*, vol. 76, no. 1, pp. 36–41, 2023.
- [9] M. Krummel, F. Bartumeus, and A. Gérard, “T cell migration, search strategies and mechanisms,” *Nat Rev Immunol.*, vol. 16, pp. 193–201, 2016.
- [10] J. Anaya, Y. Shoenfeld, and A. Rojas-Villarraga, *Autoimmunity: From Bench to Bedside*. Bogota, Colombia: El Rosario University Press, 2013.

- [11] R. J. Adler, O. Bobrowski, M. S. Borman, E. Subag, and S. Weinberger, “Persistent homology for random fields and complexes,” *Festschrift for Lawrence D. Brown*, vol. 6, pp. 1–21, 2010.
- [12] H. Edelsbrunner, D. Letscher, and A. Zomorodian, “Topological persistence and simplification,” *Discrete Computational Geometry*, no. 28, pp. 511–533, 2002.
- [13] Z. Su, X. Liu, L. B. Hamdan, V. Maroulas, J. Wu, G. Carlsson, and G.-W. Wei, “Topological data analysis and topological deep learning beyond persistent homology: a review,” *Artificial Intelligence Review*, vol. 59, p. 58, Dec 2025.
- [14] Edelsbrunner, “Alpha shapes - a survey,” *Tessellations in the Sciences: Virtues, Techniques and Applications of Geometric Tilings*, 2011.
- [15] The GUDHI Project, *GUDHI User and Reference Manual*. GUDHI Editorial Board, 3.12.0 ed., 2026.
- [16] T. Saito and J.-I. Toriwaki, “New algorithms for euclidean distance transformation of an n-dimensional digitized picture with applications,” *Pattern Recognition*, vol. 27, no. 11, pp. 1551–1565, 1994.
- [17] I. Obayashi, T. Nakamura, and Y. Hiraoka, “Persistent homology analysis for materials research and persistent homology software: HomCloud,” *Journal of the Physical Society of Japan*, vol. 91, no. 9, p. 091013, 2022. Software version 5.3.0, released March 25, 2026.
- [18] C.-S. Hu, A. Lawson, Y.-M. Chung, and K. Keegan, “Two-parameter persistence for images via distance transform,” Oct 2021.
- [19] J. T. Gostick, Z. A. Khan, T. G. Tranter, M. D. Kok, M. Agnaou, M. Sadeghi, and R. Jervis, “Porespy: A python toolkit for quantitative analysis of porous media images,” *Journal of Open Source Software*, vol. 4, no. 37, p. 1296, 2019.
- [20] E. Munch, “A user’s guide to topological data analysis,” *Journal of Learning Analytics*, vol. 4, pp. 47–61, 2017.
- [21] B. J. Stolz, J. Dhesi, J. A. Bull, H. A. Harrington, H. M. Byrne, and I. H. R. Yoon, “Relational persistent homology for multispecies data with application to the tumor microenvironment,” *Bulletin of Mathematical Biology*, vol. 86, p. 128, Sep 2024.