

Quantifying flow and mixing through porous media using Topological Data Analysis

Aakash Karlekar, Kevin Steiger, Connor Senor, Abril Montes

May 13, 2026

Abstract

Porous media are ubiquitous in nature and industry, playing a fundamental role in processes such as environmental cleanup, enhanced oil recovery, and geological carbon sequestration. These processes typically involve the mixing of miscible, multi-phase fluids within the intricate pore networks within porous media. Characterizing the complex internal pore structure of these materials and understanding how this micro-scale geometry mechanically drives fluid mixing can help optimize such processes.

Recent research highlights the profound impact of structural topology and phase saturation on these dynamics. For instance, a lower liquid phase saturation in porous media has been shown to increase the spatial convergence of flow streamlines and strengthen helical flow. This phenomenon increases both shear and vorticity dominated fluid deformation, which significantly enhances the overall mixing efficiency.

This research aims to build upon these insights by using techniques in Topological Data Analysis — specifically persistent homology — to analyze high-resolution 3D images capturing multiphase mixing within artificial porous media samples. Our primary objective is to comprehensively quantify the topological structure of the fluid phases and formulate topological measures, which we will ultimately use to correlate the structural complexity of the system with its mixing performance.

1 Background

Topological Data Analysis (TDA) [18], and particularly **persistent homology (PH)**, can be successfully applied across a wide range of scenarios to analyze and measure the topological properties of materials using two-dimensional (2D) or three-dimensional (3D) images. Consequently, PH has found diverse applications across numerous fields. For example, it is utilized in medical settings for analyzing brain networks [1], protein structures [24], bone morphology [19], and tumor shapes [20]. In physics, it is used to analyze composite materials [16], better understand phase transitions [5], and characterize the cosmic web of the universe [25]. Furthermore, it is widely applied in computational settings for image, network, and clustering analyses [17].

This project focuses on the specific application of TDA to 3D digitized images of porous media, which are highly relevant in natural and industrial processes such as environmental remediation, secondary and tertiary oil recovery, and CO₂ sequestration. These images, frequently obtained from 3D scans of porous material samples [2, 22, 21], are typically either binary or grayscale. In a binary representation, each voxel takes on one of two values, representing either void space or impermeable solid material. Conversely, a grayscale representation captures a continuous variation in density, permeability, or fluid concentration by assigning a specific intensity value to each voxel.

We are specifically analyzing high-resolution images captured by Velásquez-Parra et al. [23], who utilized 4D (three spatial dimensions + time) synchrotron X-ray micro-tomography to observe real-time, multi-phase fluid dynamics within artificial porous media. These experiments captured the precise geometric arrangement of the liquid phase under varying saturation levels and flow rates, mapping out complex internal flow phenomena such as streamline convergence and vorticity-driven helical flow.

Our analysis utilizes two distinct sets of imaging data: First, 3D binary morphology images that capture the system’s initial structural state; and second, 4D grayscale images that track the temporal evolution of an aqueous tracer spreading through the liquid phase. To analyze these images, we will employ a thresholding process whereby features of the images appear and/or disappear as some threshold parameter is varied. For the static morphology data, this involves the use of **alpha complexes** and the **Euclidean Distance Transform (EDT)**, while for the dynamic tracer data, we will use grayscale-based thresholding. The result of this thresholding process is a nested sequence of geometric layers called a topological filtration. Specifically, at a given threshold level, the subset of the image comprising the voxels that meet the threshold criteria constitutes the current geometric layer.

We can study the topological properties of each filtration layer by computing its relevant homology groups. These groups are most simply described in terms of Betti numbers. For a 3D space, there are three relevant Betti numbers ($\beta_0, \beta_1, \beta_2$), corresponding to the numbers of connected components, tunnels (handles), and enclosed cavities found within each layer, respectively [15, 6]. PH synthesizes this information by quantifying the topological features that persist across many layers of a filtration. As we progress through a filtration, the threshold value at which a feature appears is referred to as its birth value. Similarly, the threshold value at which a feature disappears is its death value. We can then visualize the persistence of these features as points (topological generators) in a 2D graph, where one axis represents the birth value and the other represents the death value. By plotting the

corresponding birth-death pairs for all features encountered in our filtration, we obtain a **Persistence Diagram (PD)** for a given homology group. A PD can then be used to generate Persistence Barcodes, Persistence Landscapes (PLs), and Persistence Images (PIs), all of which represent various aspects of the topological composition of the porous material.

From the outputs of our PH process, we will derive a robust set of topological measures designed to quantitatively characterize both the static matrix of the porous medium and the topological deformation of the tracer front. This builds on our previous, successful experience of using topological measures to quantify the effects of noising/denoising on porous media images. Ultimately, by correlating these static structural metrics with the dynamic mixing measures, we aim to uncover deeper mechanistic insights into how the pore geometry of the system directly governs its mixing performance.

2 Significance

This research holds profound implications for a broad range of industrial and environmental sectors. Beyond the domains previously discussed [1, 24, 19, 20, 16, 5, 25, 17], miscible multiphase mixing is fundamental to the operation of established technologies such as fuel cell energy storage [13], agricultural soil management [3], geothermal energy extraction [12], and pharmaceutical chromatography [14]. Furthermore, characterizing these micro-scale dynamics is essential for the progression of emerging fields, such as advanced water filtration systems [7] and biomedical tissue engineering [9]. The insights gleaned from our experimental analysis regarding the relationship between complex pore architecture and fluid mixing can help optimize and advance these critical technologies.

This research is also distinguished by its application of TDA to complex fluid systems. Traditional macroscopic metrics found in the literature frequently fail to capture the intricate spatial connectivity and structural bottlenecks that are found within a porous matrix. Such structural features are responsible for governing local flow pathways and vorticity. By using robust topological measures to quantitatively characterize these structures, and subsequently correlating them with dynamic 4D mixing behaviors, our work significantly advances the current theoretical framework of how the physical geometry of porous media impacts critical mixing properties such as streamline convergence and helical flow.

3 Methodology

We work with high-resolution images captured by Velásquez-Parra et al. [23], who utilized 4D (three spatial dimensions + time) synchrotron X-ray micro-tomography to observe real-time, multi-phase fluid dynamics within artificial porous media. Experiments were conducted at the beamline for Tomographic Microscopy and Coherent Radiology Experiments (TOMCAT) of the Swiss Light Source (SLS) at the Paul Scherrer Institute (PSI) in Villigen, Switzerland. The porous medium consisted of a sintered packing of irregularly shaped borosilicate glass grains forming a cylindrical sample 8 mm in height and 4 mm in diameter, with a bulk porosity of 0.28, an intrinsic permeability of 10^{-10} m², and an average pore size of 30 μ m — resembling a poorly graded sandy soil with a relatively homogeneous pore size distribution.

Transport experiments were conducted in two stages. In the first stage, gas (air) and liquid phases were co-injected simultaneously at a constant flow rate using an automatically controlled syringe pump until steady-state flow conditions were reached, after which injection was halted to yield a static, immobile air phase within the pore space. Saturation degrees ($S_w = V_w/V_T$, where V_w is the liquid phase volume and V_T is the total pore space volume) spanning from $S_w = 0.75$ to $S_w = 0.92$ were tested across five experiments. In the second stage, a potassium iodide (KI) aqueous solution was used as both the tracer (0.90 M) and the resident solution (0.06 M), exploiting KI’s chemical inertness with borosilicate glass and its high effective atomic number to achieve strong X-ray contrast. Tracer injection was performed at low capillary numbers (average $C_a = 2.67 \times 10^{-6}$) and at three flow rates ($q = 0.125, 0.25, \text{ and } 0.50 \text{ mm}^3 \text{ s}^{-1}$, corresponding to Péclet numbers $Pe = 1.69\text{--}7.28$) to ensure the air phase remained immobile throughout.

Across these experiments, Velásquez-Parra et al. found that lower liquid phase saturation consistently enhances solute mixing efficiency. Quantified through the temporal evolution of the variance of the concentration field $\sigma_c^2(t)$, lower S_w was shown to produce a faster decay of concentration segregation, reflecting more rapid homogenization of the tracer with the resident solution. This enhancement is driven by a greater deformation of the solute front: as saturation decreases, the larger volume of immobile air clusters forces flow streamlines to converge more strongly, increasing their tortuosity and promoting the formation of a preferential flow backbone. This streamline convergence in turn strengthens helical flow within the liquid phase, as reflected in a linearly increasing average absolute helicity density $\langle |h| \rangle$ with decreasing S_w . The resulting flow field exhibits a broader distribution of both shear- and vorticity-dominated deformation regions, quantified via the Q -criterion, both of which grow in magnitude as saturation falls. Together, these mechanisms — greater front deformation, enhanced helical motion, and intensified shear and rotational strain — produce the super-diffusive scaling of the solute front area observed experimentally, and establish liquid phase saturation as the primary control on pore-scale mixing dynamics in unsaturated porous media.

Resolving these dynamics required imaging at exceptional spatial and temporal fidelity. Each tomographic dataset consisted of $2016 \times 2016 \times 1100$ isotropic voxels with a lateral size of $2.75 \text{ }\mu\text{m}$, acquired over a total rotation time of 2.4 s using nearly monochromatic synchrotron radiation at 21 keV — representing approximately 15% finer spatial resolution and 60% faster acquisition than comparable prior studies [8]. Each experiment yielded up to 250 tomograms (time steps). Prior to analysis, noise inherent to the tomographic acquisition was addressed by applying a 2D non-local means denoising algorithm slice-by-slice across the XY -planes of each 32-bit tomographic stack, a strategy found to perform comparably to its 3D counterpart while remaining computationally tractable. The cylindrical sample boundary was then isolated using an Active Contours segmentation mask, excluding the surrounding imaging region from further processing, and phase segmentation of the liquid, gas, and solid constituents was performed via K-means clustering on the 16-bit tomograms at each time step.

Together, these steps reduce the effective image domain from the raw $2016 \times 2016 \times 1100$ acquisition grid to the sample-bounded volume of $1518 \times 1502 \times 1000$ voxels used in all subsequent analysis. A linear concentration calibration curve, established from fourteen static reference scans at known KI concentrations (0.06–1.81 M), was then applied to convert

voxel intensities within the liquid phase mask into molar concentration values. The resulting four-dimensional concentration fields $c(\mathbf{x}, t)$ were complemented by pore-scale numerical simulations of Stokes flow (via COMSOL Multiphysics finite element method), providing fully resolved velocity, vorticity, and shear rate tensor fields at every node. Together, these experiments captured the precise geometric arrangement of the liquid phase under varying saturation levels, mapping out complex internal flow phenomena such as streamline convergence and vorticity-driven helical flow.

Our proposed methodology leverages this dataset through two topological analyses, targeting the static pore geometry and the dynamic tracer evolution respectively.

i. Topological Analysis of the Solid Morphology

To characterize the structural topology of the porous medium, we analyze the 3D binary images of the solid matrix using two complementary persistent homology (PH) filtration approaches, cross-validating the resulting topological summaries to ensure the physical accuracy of our structural models.

- (a) **Alpha Complex Filtration.** In this approach, expanding spheres of radius α are centered at the coordinates of solid matrix voxels to construct a nested sequence of simplicial complexes. As the filtration parameter α increases, sphere intersections give rise to edges, faces, and higher-dimensional simplices. Tracking the birth and death of these complexes allows us to continuously capture the emergence and vanishing of structural features — connected components, loops, and enclosed cavities — across a range of spatial scales.
- (b) **EDT Filtration.** Complementing the alpha complex approach, the Euclidean distance transform (EDT) assigns each void voxel a scalar value corresponding to its L^2 distance to the nearest solid boundary. Using this distance field as the filtration function — by progressively thresholding across encountered distance values and tracking the resulting simplicial complexes — yields a geometrically grounded characterization of pore space structure that is directly comparable to the alpha complex results.

Since the full image domain measures $1518 \times 1502 \times 1000$ voxels, direct computation over the entire solid matrix presents significant computational challenges. We therefore address the subproblem of identifying a sufficient level of spatial subsampling that faithfully represents the global topology of the domain.

ii. Topological Characterization of the Tracer Front Evolution

To analyze the dynamic mixing process, we apply grayscale-based thresholding to the 4D time-evolving images using a sliding window approach. The grayscale intensity at each voxel corresponds to the local concentration of the aqueous tracer, and by using this scalar field as the filtration parameter, we track how the topological structure of the tracer front — its connected components, loops, and enclosed regions — evolves and deforms over time. This provides a quantitative, scale-resolved description of the mixing dynamics that complements the flow simulations of Velásquez-Parra et al. [23].

4 Mathematical Framework and Theoretical Foundation

This section establishes the mathematical and computational framework necessary to quantify the complex structures within our porous media sample. We begin with a detailed exposition of topological concepts (**Section 4.1**) that were briefly described earlier, followed by a comparative analysis that aligns our experimental system with the interaction complex framework proposed by Kramár et al. [11] (**Section 4.2**), and finally outline the validation strategies employed to ensure our findings are consistent (**Section 4.3**).

4.1 Topological Concepts

Traditional analysis of porous media often relies on bulk geometric parameters like pore volume, pore throat diameters, and pore size distributions. While these metrics provide detailed descriptions of the state of the system, they fail to capture the underlying complexity of realistic, multiscale samples. In nature, porous media rarely conform to the uniform shapes used in idealized simulations. Instead, they are characterized by significant structural heterogeneity, where even minor variations in pore connectivity can result in vastly different transport phenomena. These discrepancies arise because geometric properties are inherently tied to exact shapes and distances, making them highly sensitive to local perturbations. Consequently, a small deformation in a pore space can significantly alter geometric values and lead to inaccurate flow predictions. This highlights the need for a more robust framework that prioritizes the fundamental structure of the system: **topology**.

In our porous media sample, the specific shape of individual grains and the exact metric dimensions of the pore space are informative, but are less critical than the network they form. Instead, the primary focus lies in the system’s connectivity. We want to determine whether pores are isolated or interconnected, identify the existence of continuous or redundant pathways, and characterize the presence of trapped gas or dead-end pores. These properties are known as **topological features**, which are characteristics that remain invariant under continuous deformation (stretching, shrinking, bending), and are independent of minor geometric variations between samples. By prioritizing these features, we can establish a robust structural description that remains valid even as the local geometry fluctuates during the mixing process.

While our analysis focuses on the complex interconnectedness of pore networks, these topological principles can be illustrated through a classic example in **algebraic topology**. Let us consider a standard coffee mug with one handle and a doughnut. In this context, both of these objects are considered to be topologically equivalent because they consist of a single connected piece with exactly one hole. As shown in Figure 1, the coffee mug can be continuously deformed into the shape of a doughnut without tearing the material or joining pieces together, demonstrating that the fundamental connectivity of the space remains unchanged regardless of their outward appearance. The thickness of the doughnut or the depth of the mug is of minor importance, since what really matters is the underlying structure. By counting these features, such as the number of pieces and the number of holes, we can mathematically group objects into distinct classes regardless of their physical shape.

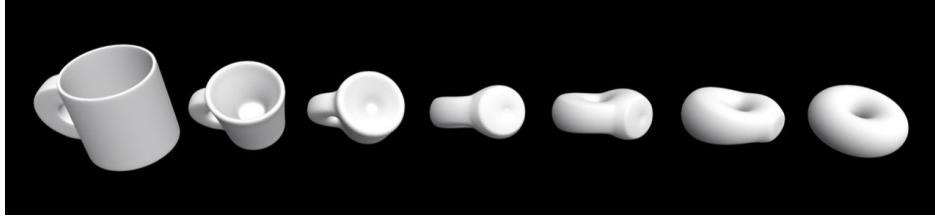


Figure 1: A continuous deformation of a coffee mug into a doughnut (torus)[4]

To translate this intuition into a rigorous mathematical framework, we utilize the concept of **homology**. This allows us to describe the global structure of a space by distinguishing between objects based on the number and type of ‘holes’ across different dimensions [18]. These counts are represented by **homology groups**, denoted H_k , where each group captures features of a specific dimension k . The rank of each homology group is defined as its **Betti number** β_k , which is a single value analogous to the dimension of a vector space that counts the occurrences of these k -dimensional features. For instance, the zeroth Betti number (β_0) represents the number of distinct connected components, the first Betti number (β_1) counts one-dimensional cycles or loops, and the second Betti number (β_2) identifies three-dimensional cavities or voids.

i. Interpreting Betti Numbers in Fluid Transport

While the mathematical definitions of Betti numbers remain constant, their utility in our study lies in their ability to quantify distinct physical phenomena across the solid and fluid phases:

- 1 **The Solid Phase (Static Framework):** The topology of the glass grains defines the geometric constraints of the experiment. Here, β_0 represents the number of individual, disconnected grains within the domain, while β_1 identifies the contact points between these grains. These one-dimensional loops are physically significant as they define the pore throats and tortuous pathways that the fluid is forced to navigate.
- 2 **The Fluid Phase (Dynamic Mixing):** For the tracer, the Betti numbers serve as a direct metric for the evolution of the mixing front. Here, β_0 describes fragmentation, which represents the number of disconnected tracer clusters present before they merge into a continuous stream. β_1 describes folding, the primary signature of mixing. It represents the fluid being stretched into filaments and ‘folding’ around obstacles (grains or air bubbles) to create loops. β_2 describes entrapment, occurring when the tracer front completely encloses a volume of air, thereby capturing a ‘dead zone’ within the fluid phase.

By establishing these interpretations, we can move from a purely visual assessment of flow to a quantified analysis of the mixing regimes present in our datasets.

Computing Betti numbers provides a snapshot of a system’s topology at a single scale. However, we must remember that porous media are inherently multiscale systems. A single Betti number cannot distinguish between a large, physically significant pore and a tiny,

transient artifact caused by noise. To resolve this, we must investigate how these features survive across a range of spatial scales. This is where **persistent homology** becomes essential, allowing us to track the lifespan of topological features.

In persistent homology, we monitor the ‘birth’ and ‘death’ of topological features as we vary a filtration parameter [18]. The difference between the scale at which a feature first appears and the scale at which it disappears (or merges with a larger feature) is defined as its **persistence**. This value represents the importance, or robustness, of a feature. Those with high persistence are considered to be significant structural elements of the pore network, and those with low persistence are typically dismissed as noise. To visualize this data, we utilize **persistence diagrams**, which are essentially scatter plots that map the birth and death coordinates of each feature. Points located far from the diagonal – a line representing features with zero persistence, where the birth value equals the death value – represent high persistence features that define the true structure of the sample, whereas points near the diagonal indicate short-lived artifacts. Additionally, by summing the persistence of all features within a specific dimension, we can calculate the **total persistence**. This single-value metric characterizes the overall topological complexity of the system, allowing for a direct comparison of connectivity between different samples.

The utility of persistent homology depends on the reliability of the features it identifies, a challenge amplified by the nature of our dataset. In the experiment conducted by Velásquez-Parra et al. [23], we must reconcile the immense complexity of high-resolution, four-dimensional volumes with the computational necessity of significant subsampling. Every stage of this data reduction, compounded by scanner noise, calibration uncertainties, and motion artifacts from sample rotation, introduces the risk of generating results that are mere processing artifacts rather than physical realities. This fundamental concern is addressed by the stability properties established by Kramár et al. [11], who provide the theoretical guarantee that our topological descriptions remain robust under such perturbations. Central to this robustness is the **Interaction Complex**, a framework that ensures our analysis remains structurally representative despite the inherent noise and multiscale challenges of the dataset.

4.2 The Interaction Complex and Methodological Stability

The significance of the **Interaction Complex** lies in its ability to resolve the ambiguity inherent in multiphase systems. We utilize this framework because it functions as a mathematical map describing how the solid, liquid, and gas phases co-exist within the porous medium. Rather than analyzing these phases in isolation, which can obscure the critical physical events like mixing and entrapment, the interaction complex focuses on the specific boundaries where they meet. In our 4D dataset, this allows us to track the tracer as it navigates fluid-solid interfaces (moving past glass grains) and fluid-gas interfaces (displacing or wrapping around air bubbles). By analyzing the entire network of interfaces as a unified system, we ensure that our topological characterization is physically grounded in the complex interfacial connectivity that defines the transport regime.

To translate this interface-focused framework into a functional analysis, we distinguish between two fundamental categories of implementation: **value-based (interaction)** and **position-based (geometric)**. By comparing these approaches, we can classify and deter-

mine which of our methods best captures the physical reality of our system while remaining robust against experimental error.

- i. **Interaction Complexes (EDT and Grayscale):** Our primary approaches function as interaction complexes because they define connectivity through continuous scalar fields. In the EDT approach, each voxel is assigned a distance value to the nearest solid boundary, while Grayscale thresholding utilizes the scalar intensity of the tracer. These methods do not rely on discrete coordinates. Instead, they ‘sweep’ through a landscape of physical values.
- ii. **Position Complexes (Alpha Complexes):** In contrast, the alpha complex remains a strictly geometric, position-based method. It determines connectivity by the spatial proximity of discrete points, specifically through a Delaunay triangulation controlled by a radial parameter (α). While this provides a complementary perspective on the solid network, it focuses on where points are located in space rather than how the phases physically ‘touch’.

The practical advantage of selecting value-based interaction complexes is most clearly demonstrated through the stability properties established by Kramár et al [11]. Their work provides a theoretical guarantee, formally defined in the Lipschitz continuity of topological features (Theorem 7.2), that our persistence diagrams will respond to data perturbations in a gradual, predictable manner. This stability is vital for a dataset of this complexity, where minor measurement errors or imaging artifacts could otherwise be amplified into significant topological discrepancies.

To quantify this, we perform a sensitivity analysis comparing the error amplification of position-based methods against the interaction complex framework. Consider a scenario where a marginal measurement error of $\epsilon = 0.001$ is introduced, a value consistent with the noise levels analyzed by Kramár and his colleagues. In a position-based alpha complex, this small positional shift can cause a ‘jump’ in the topology, such as an edge between two grains suddenly appearing or disappearing, which leads to significant discrepancies. However, because our EDT and Grayscale methods are anchored to continuous scalar interactions, they remain invariant to such noise.

As illustrated in Table 1, which compares the resulting bottleneck distances (a metric used in their study to analyze the stability of persistence diagrams) under identical noise conditions, we characterize the impact of this noise by defining error amplification as the ratio of the output topological discrepancy to the input error (d_b/ϵ). While the position complex shows a significant amplification ($2615\times$), effectively ‘jumping’ to an unstable topological state, the interaction complex acts as a filter, actually reducing the impact of the noise ($0.4\times$). Comparing these frameworks reveals that the interaction complex demonstrates up to a 6,500-fold reduction in error sensitivity compared to strictly geometric constructions. This dramatic difference in error resilience confirms that the topological signatures we identify are stable representations of the physical mixing regime rather than transient artifacts of experimental noise.

Ultimately, this framework provides a theoretical guarantee that our different coding approaches are measuring the same underlying physical structure. Despite using different

Complex	Stability	Error Amplification	Status
Position	$d_b = 2.6151$	2615x (amplifies)	Unstable
Interaction	$d_b = 0.0004$	0.4x (reduces)	Stable

Table 1: Stability comparison of complexes under a 0.001 measurement error

data types, geometric distance for EDT and tracer intensity for Grayscale, both are mathematically categorized as interaction complexes. This ensures that the resulting topological signatures (fragmentation, folding, and entrapment) are not artifacts of our subsampling or scanner noise, but are stable, convergent representations of the global mixing regime.

4.3 Validation Strategies

To ensure the reliability of the observed topological signatures, we employed two primary validation strategies: spatial homogeneity and scale independence. These tests confirm that the detected features are intrinsic properties of the medium rather than localized artifacts of the observation scale.

1. **Spatial Homogeneity:** The first validation strategy assesses whether the topology remains invariant across different regions of the sample. We systematically shifted the observation window through various X and Y displacements from the central axis to determine if the topological plots exhibited regional consistency. To provide a rigorous baseline for this comparison, we performed a parallel static geometry analysis, calculating the solid fraction and pore size distribution across the entire domain. The convergence of both the topological signatures and the geometric metrics confirms that the sample is spatially homogeneous, ensuring that our findings are representative of the global medium.
2. **Scale Independence:** The second strategy addresses the impact of window size on the stability of the results. In topological analysis, it is critical to identify the scale at which the observation window becomes large enough to capture the essential physics of the mixing regime. By systematically increasing the dimensions of the analysis window, we evaluated the scale independence of our Betti numbers. The stabilization of the persistence patterns across these varying scales validates that our 200-voxel sub-volumes are sufficiently sized to capture robust physical structures while remaining localized enough for computational efficiency. Together, these strategies confirm that our results are not transient artifacts of noise but are stable, scale-independent descriptors of the transport process.

5 Morphology Data Analysis

5.1 Alpha Complexes

While the initial aims for this project were to try and analyze the porous media with Alpha Complexes, we had to pivot a bit in terms of our research question. In particular, our Alpha Complexes library, GUDHI, is not particularly efficient for extremely large point clouds.

Even subvolumes of the larger domain were taking as much as 10 minutes for just a fraction of the total volume. As a result, we developed a sub-question that we’ve tried to answer:

How small of a subvolume still yields an accurate topological characterization of the porous media?

In order to compare topology of the different subvolumes, we considered a wide variety of different measures. The measure that we ended up selecting for our analysis was total persistence, given by the following equation:

$$\text{TP}_k = \sum_i (d_i^{(k)} - b_i^{(k)}) \quad (1)$$

Total persistence is the sum of all the lifespans from every generator. We sum up the lifespans of generators at dimension 0, dimension 1, and dimension 2. A lifespan for a generator is the difference between the alpha value at which a generator dies, and the alpha value at which a generator is born. Total persistence was a good measure due to the fact that it is so simple. We prioritized simplicity because we wanted to work with a measure that could be simply understood so that it would be easy to troubleshoot in this new context. However, as we will later explain there were some issues with this raw total persistence metric.

In order to answer our research question, we developed the following methodology. First we would do all of our analysis on a 200x200x200 subvolume within the center of the porous media. This size was chosen because the time to run Alpha Complexes was still reasonable at this particular size. We then tested various downsample ratios. Originally, we were placing an initial generator at the center of every single solid voxel. When we began downsampling at say a ratio of $\frac{1}{2}$, we would only place initial generators at the center of $\frac{1}{2}$ of all the solid voxels in the subvolume. Then we would calculate total persistence using the downsampled initial generators. Those generators were randomly selected, and in order to reduce variance, we would run 10 trials at each downsampling level, and calculate the average total persistence across all the trials. This gave us a mean total persistence for every downsampling level. We then plotted average total persistence as a function of initial generators.

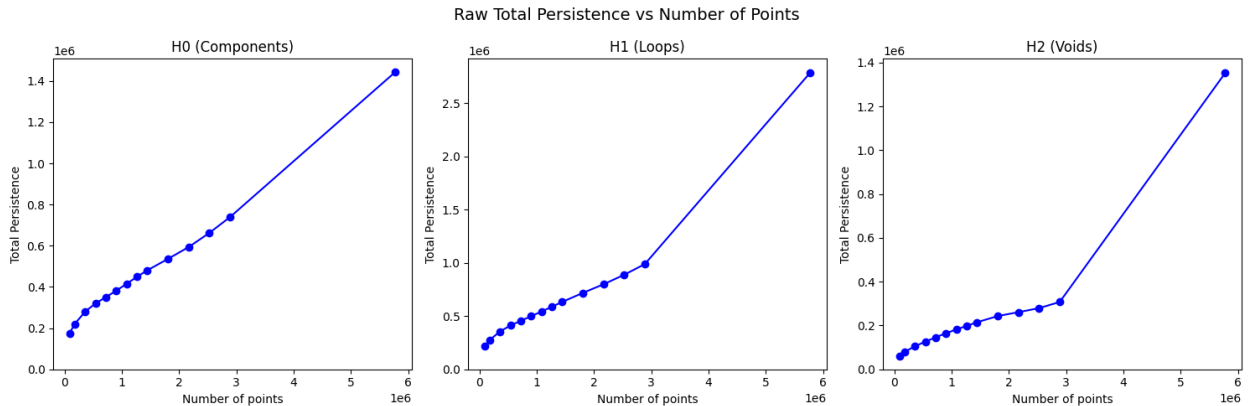


Figure 2: Raw total persistence vs. number of sampled solid voxels across H0 (Components), H1 (Loops), and H2 (Voids).

As shown in the plot above, total persistence seemed to grow with the initial number of generators. This intuitively makes sense as more initial generators cause a higher overall

total persistence. However this scaling made it difficult to compare how well the actual topology was being captured at various downsampling levels. Another issue that we faced was this idea of "grid introduced noise".

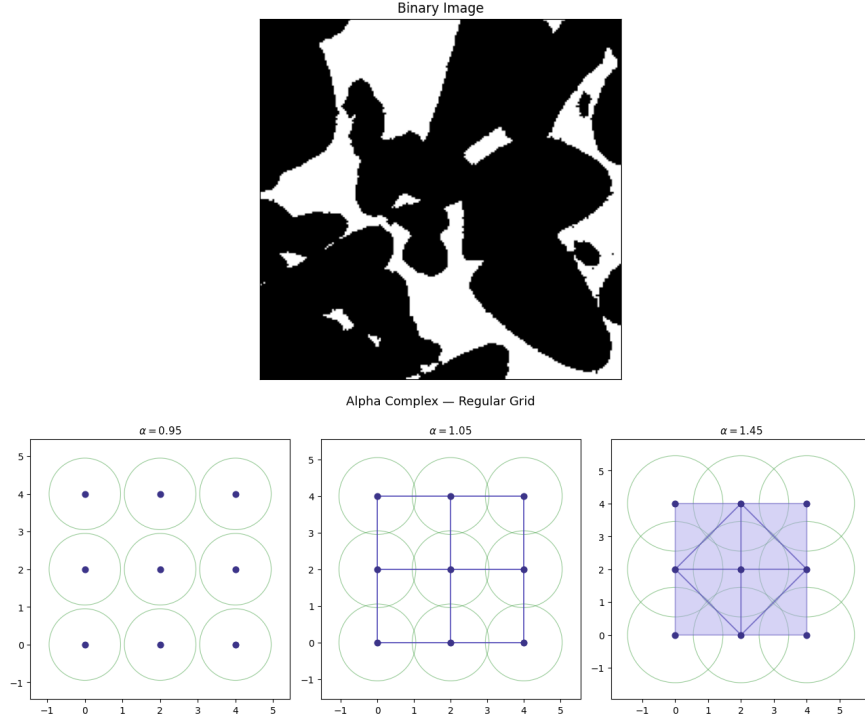


Figure 3: Top: 2D slice of the porous medium. Bottom: Alpha complex on a regular grid at three values of α , illustrating spurious features introduced by lattice geometry.

A significant source of noise in our total persistence metric comes from the regular grid structure of the voxel data as shown in figure 2. Due to how generators are placed at voxel centroids on a regular lattice, the alpha complex filtration introduces many short-lived features that reflect the grid structure rather than the true pore structure. These "fake" features contribute to total persistence but carry no meaningful topological information about the media that we care about. In short we had two main issues that we had encountered thus far. First, total persistence was scaling with the initial number of generators, and second we had a decent amount of noise due to the grid structure of our data contributing to total persistence.

In order to solve the scaling problem, we ended up finding a paper by Jaquette and Schweinhart [10] that gave us the scaling relationship between initial number of generators and total persistence. This is given in the following equation:

$$E_{\alpha}^i \sim n^{\frac{d-\alpha}{d}} \quad (2)$$

Where n is the number of generators, d is the number of dimensions, and α is the exponent applied to each feature's lifespan in the following weighted total persistence formula:

$$E_{\alpha}^i = \sum_{(b, d) \in PH_i} (d - b)^{\alpha} \quad (3)$$

Based on this relationship, if we set $\alpha = 3$, we can force the exponent of the scaling relationship to 0. This theoretically stabilizes total persistence, normalizing for the initial number of generators. This means that we can finally compare total persistence across the initial number of generators and observe just how accurately the true topology is being captured. Another nice side-effect of cubing of lifespans before adding them to total persistence, is that this helps reduce overall noise, as smaller lifespans end up contributing less to total persistence than longer lifespans.

Once we used this new metric, Norm-3 total persistence, we were able to once again plot total persistence as a function of the number of initial generators, creating the plots below in figure 3.

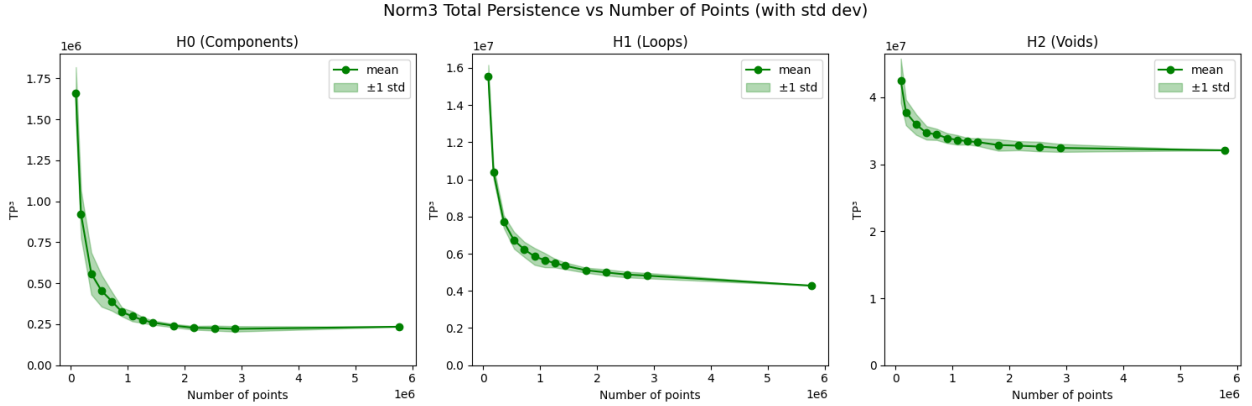


Figure 4: Norm-3 total persistence (TP_3) vs. number of sampled solid voxels across H0 (Components), H1 (Loops), and H2 (Voids), with ± 1 standard deviation shaded.

H0, and H2 seem to show a nice, converging relationship as we increase the number of generators, which is important to us. This apparent stabilization may seem to indicate that there is a threshold at which alpha complexes with downsampling can accurately capture the same topology as alpha complexes without any downsampling. Further work might be to research what that exact threshold may be, as well as try and understand why H1 does not seem to have a similar "converging" relationship. Both of these directions may help us understand how we can downsample with Alpha Complexes and still get an accurate understanding of the underlying topology.

5.2 Euclidean Distance Transform and HomCloud

A major goal of this part of the project was to compare the topological structure of the porous-medium sample across different computational representations. In the midterm version of this report, the focus was mainly on generating persistence diagrams from selected two-dimensional slices and three-dimensional subdomains. In the final stage of the project, this analysis was extended by computing normalized total persistence across multiple domain sizes and by connecting the topological convergence results to a physical pore-scale length using SNOW2.

The Euclidean Distance Transform (EDT) provides a way to convert the geometry of a binary porous medium into a scalar field. Starting with a binary image of the sample, each

voxel is assigned a value based on its Euclidean distance to the nearest phase boundary. In this setting, the phase boundary is the interface separating the solid phase from the liquid and gas phase. Voxels near the interface receive values close to zero, while voxels deeper inside a phase receive values with larger magnitude. This scalar field can then be used as the input for persistent homology calculations in HomCloud.

For this project, HomCloud was used to compute persistence diagrams from EDT fields using a superlevel filtration. Persistent homology tracks the appearance and disappearance of topological features throughout the filtration. The three homology dimensions considered were

$$H_0, \quad H_1, \quad H_2,$$

where H_0 represents connected components, H_1 represents loops or tunnels, and H_2 represents enclosed voids. Each persistence diagram records the birth and death values of these features. The lifetime of a feature is the difference between its death and birth values, and total persistence was computed by summing these lifetimes:

$$TP_k = \sum_i |d_i - b_i|,$$

where b_i and d_i are the birth and death values of the i th feature in homology dimension H_k .

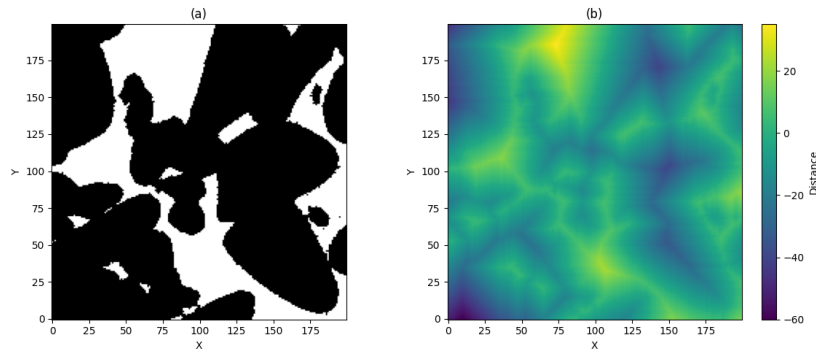


Figure 5: Example EDT field computed from a binary slice of the porous-medium sample. The EDT converts geometric structure into a scalar field that can be used to define a filtration for persistent homology.

An important part of the analysis was understanding the effect of using a signed Euclidean distance transform. Initially, some persistence diagrams were confusing because many birth and death values appeared negative. This does not mean that the Euclidean distances themselves were negative. Instead, it comes from the signed distance convention, where the magnitude of the value represents distance to the nearest interface, while the sign indicates which side of the interface the voxel belongs to. Values near zero lie close to the phase boundary, while positive and negative values correspond to voxels on opposite sides of that boundary. Since the porous-medium sample contains large regions belonging to different phases, negative values are expected in the signed EDT field.

After verifying the EDT workflow on two-dimensional slices, the same method was applied to three-dimensional subdomains. This made it possible to compare topological measurements across different spatial locations and across different domain sizes. To make the values

comparable across domain sizes, total persistence was normalized by the number of voxels in the subdomain:

$$\frac{TP_k}{\text{voxel}}.$$

This normalization allowed the analysis to focus on whether the measured topology stabilized as the size of the sampled domain increased.

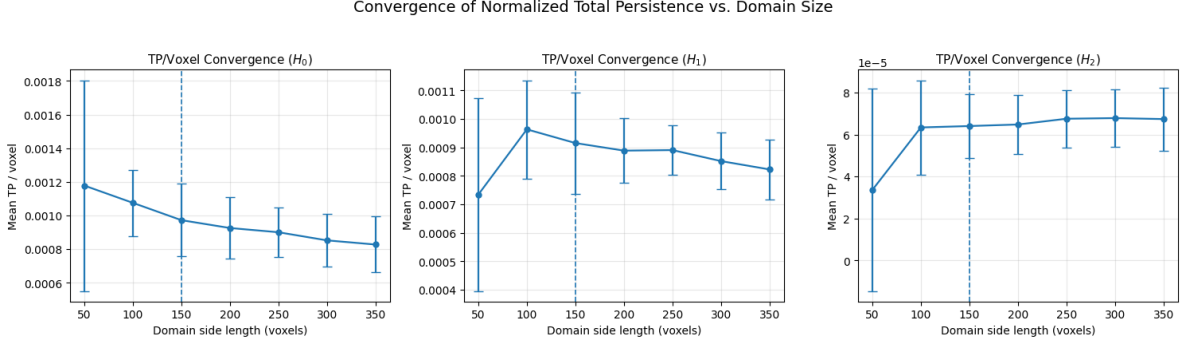


Figure 6: Mean normalized total persistence versus domain side length for H_0 , H_1 , and H_2 . Error bars show ± 1 standard deviation across sampled spatial locations.

Figure 6 shows the convergence behavior of normalized total persistence as the domain side length increases. For small domains, the normalized total persistence varies more strongly between sampled locations, meaning that the measured topology depends heavily on the particular region selected. As the domain size increases, the curves begin to stabilize. This suggests that the selected domain has become large enough to capture representative topological structure from the sample.

To examine this more directly, the variance of normalized total persistence was also plotted as a function of domain size for each homology dimension. These plots show how much the topological measurement changes across different spatial locations at the same domain size.

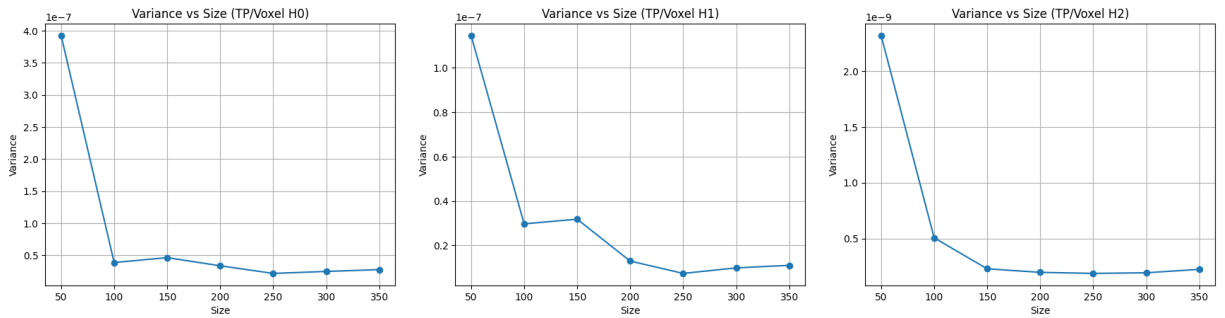


Figure 7: Variance of normalized total persistence across domain sizes for H_0 , H_1 , and H_2 . Lower variance indicates that the sampled regions are producing more consistent topological measurements.

The convergence and variance results suggest that normalized total persistence becomes

reasonably stable at a domain side length of approximately

$$L_c \approx 175 \pm 25 \text{ voxels.}$$

This value should not be interpreted as an exact universal threshold. Instead, it is an approximate representative domain size for this experimental sample. Below this range, the sampled domains may be too small to consistently capture the larger-scale topology of the porous structure. Near and above this range, the total persistence measurements become more stable across sampled locations.

5.3 Connecting Topological Convergence to Pore Size Using SNOW2

To give the topological convergence result a physical interpretation, SNOW2 was used to estimate the average pore size of the sample. SNOW2 is a pore-network extraction method implemented in PoreSpy. It uses a distance transform to identify pore centers and then applies watershed segmentation to separate the connected pore space into individual pore regions. This makes it possible to estimate quantities such as average pore size and the number of segmented pore regions.

The purpose of adding SNOW2 to the analysis was to connect the topological representative domain size to a physical pore-scale length. Instead of reporting only that the persistence measurements stabilize around a certain number of voxels, the result can be interpreted in terms of how many average pore diameters are needed for stable topological measurements.

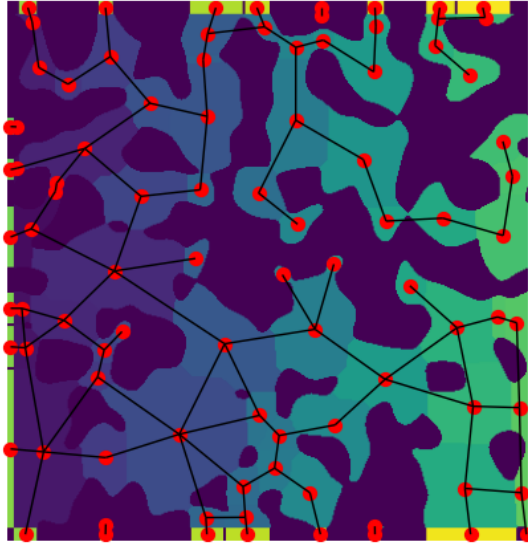


Figure 8: Example SNOW2 segmentation/network output. SNOW2 uses distance-transform information and watershed segmentation to separate connected pore space into distinct pore regions.

The SNOW2 segmentation output identified individual pore regions in the voxel data. From this segmentation, the average pore size was estimated as

$$\bar{d}_{pore} = 11.81 \pm 0.05 \text{ voxels.}$$

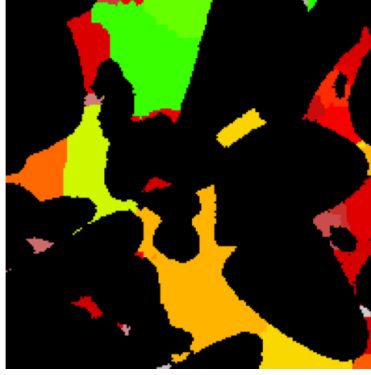


Figure 9: Pore regions identified using SNOW2 segmentation.

Using the convergence estimate from the normalized total persistence results,

$$L_c \approx 175 \pm 25 \text{ voxels},$$

and the SNOW2 average pore size,

$$\bar{d}_{pore} = 11.81 \pm 0.05 \text{ voxels},$$

the representative domain size can be expressed in terms of average pore diameters:

$$\frac{\text{domain length}}{\text{average pore size}} = \frac{175 \pm 25}{11.81 \pm 0.05} \approx 14.8 \pm 2.1.$$

Therefore, for this experimental sample, stable normalized total persistence measurements require a domain side length of roughly

$$15 \pm 2$$

average pore diameters.

This result provides a useful bridge between the computational topology and the physical morphology of the sample. The EDT and HomCloud workflow gives a quantitative topological measure of the porous structure, while SNOW2 gives a pore-scale physical length. Together, these results suggest that future topological measurements on similar porous-medium samples should use domains at least about fifteen average pore diameters across in order to obtain stable and representative measurements. This should be interpreted as an approximate representative length scale for this sample, rather than as a universal threshold for all porous materials.

5.4 Updated Interpretation

The midterm analysis showed that HomCloud could successfully generate persistence diagrams from signed EDT fields for both two-dimensional slices and three-dimensional subdomains. The final analysis extends this by using normalized total persistence to study convergence with respect to domain size. The key result is that the topology becomes more

stable as the sampled domain size increases, with stabilization occurring near 175 ± 25 voxels for this sample.

By combining this with the SNOW2 pore-size estimate, the domain-size result can be interpreted physically: representative topological measurements require a domain length of approximately 15 ± 2 average pore diameters. This gives a more meaningful criterion for selecting future subdomain sizes. Rather than choosing a domain size only in terms of voxel count, the selected domain can now be related to the underlying pore morphology of the material.

6 Tracer Front Analysis

The 4D grayscale volumetric images provided by Velasquez-Parra et al. [23] encompass a spatial domain of $1518 \times 1502 \times 1000$ voxels. These datasets are comprised of several experiments where a 0.9M KI tracer was injected into the artificial porous media system for various values of S_w and q . We first begin by conducting exploratory data analysis of a specific experiment corresponding to a steady-state liquid saturation regime of $S_W = 0.92$ and a constant tracer injection rate of $q = 0.25 \frac{\text{mm}^3}{\text{s}}$.

Fig. 10 illustrates the temporal evolution of the grayscale intensity distributions across the sampled time instances. An analysis of these distributions reveals that the population of voxels completely devoid of the tracer remains relatively stable throughout the observational period. Conversely, the active grayscale intensities within the fluid pathways exhibit a progressive, temporal shift, ultimately converging toward an asymptotic maximum value. This observed trend aligns with underlying physical principles, as the localized grayscale intensity of a given voxel serves as a direct proxy for the aqueous tracer concentration within that specific micro-region of the porous medium. Because the tracer is continuously introduced into the system at a steady rate, the localized concentration is physically constrained to monotonically increase, eventually approaching a limiting value. Consequently, the corresponding voxel intensities reliably track this physical accumulation, asymptotically approaching a fixed maximum in accordance with the established concentration-intensity calibration curve detailed by Velásquez-Parra et al. [23].

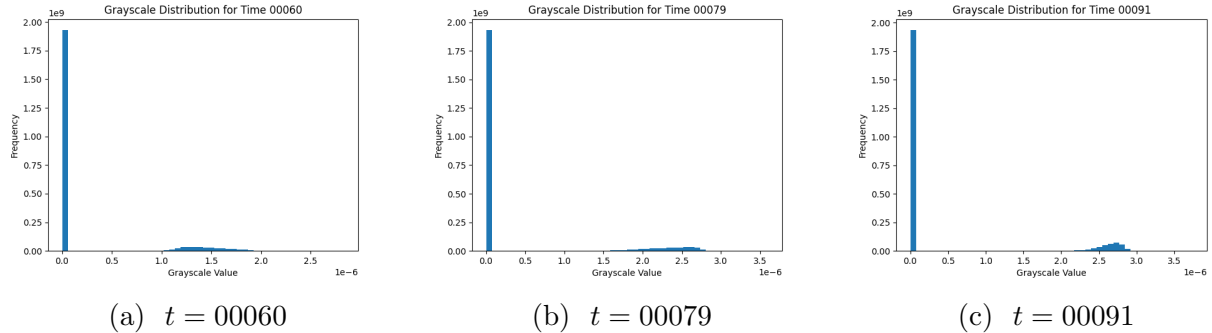


Figure 10: Distribution of grayscale values in images across time steps.

To analyze the localized topological evolution of the system, Persistence Diagrams (PDs) were computed for a $200 \times 200 \times 200$ voxel sub-volume extracted from the center of the

spatial domain across all time instances. Fig. 11 presents the PDs corresponding to time step $t = 00060$. Across all homology dimensions, a high density of topological generators is concentrated near the birth-death diagonal, indicative of small-scale, transient concentration fluctuations (topological noise). Conversely, a distinct subset of generators exhibits high persistence, deviating significantly from the diagonal. These highly persistent features are hypothesized to represent robust, macro-scale flow pathways within the tracer plume that critically govern the fluid mixing process.

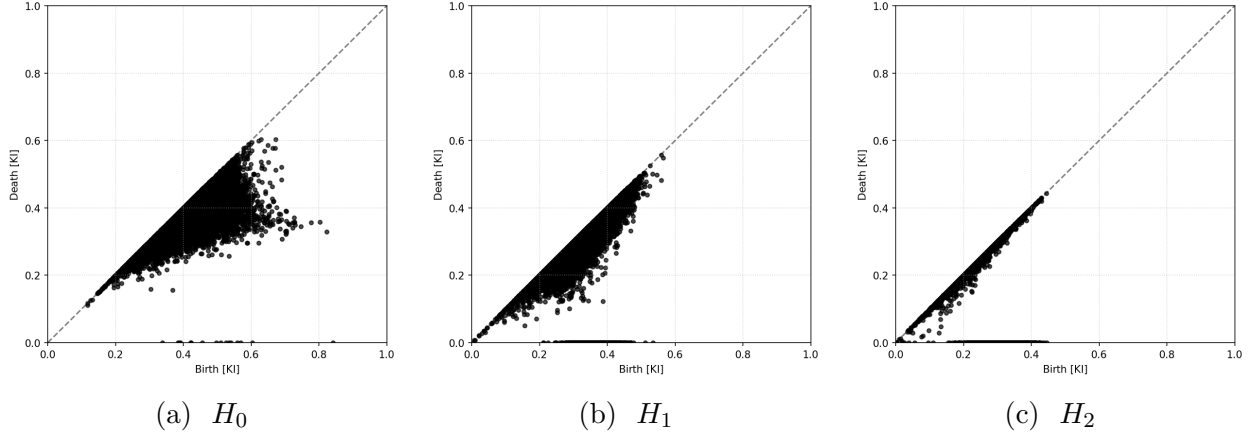


Figure 11: Persistence diagrams for $t = 00060$.

Given this exploratory analysis, we devise a sliding window approach in which a 200×200 voxel window is centered at the midpoint of the xy -plane on the z -slice whose average concentration is closest to 0.5 M, which we take as the location of the tracer front. To ensure our analysis primarily captures significant topological features, we filter out generators with a lifespan less than 0.1. Fig. 12 illustrates the temporal evolution of the Persistence Diagrams (PDs) corresponding to H_0 . An analysis of these diagrams indicates that the topological generators exhibit a systematic translation toward higher threshold values (a rightward shift in the birth-death plane) as time progresses. This behavior is a direct mathematical consequence of the overall increase in localized tracer concentration. Because the non-zero grayscale intensity distributions shift toward higher values over time, the intensity thresholds at which features emerge and merge are elevated, inherently increasing both the birth and death coordinates of the generators during the superlevel thresholding process.

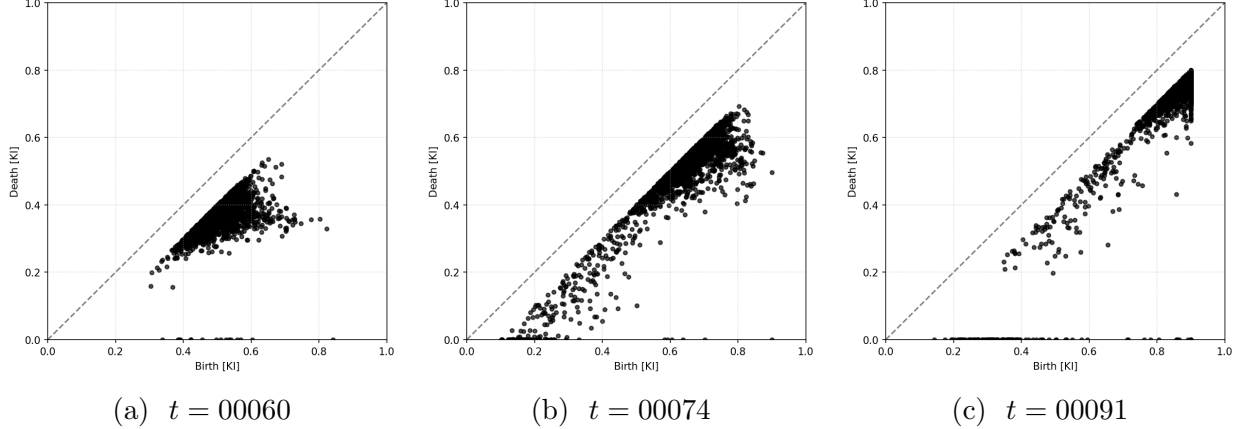


Figure 12: Persistence diagrams for H_0 across time steps $\left(S_w = 0.92, q = 0.025 \frac{\text{mm}^3}{s}\right)$.

This is also verified through the positive trend seen in the average birth statistic for the same experiment shown in Fig. 13, alongside the number of generators (NG_i), average lifespan (AL_i), and total persistence (TP_i) statistics. Recalling that $TP_i = AL_i \cdot NG_i$ by definition, the interplay between these two quantities offers a more nuanced picture of the evolving tracer front topology than either statistic alone. For $i = 0$, NG_0 is large throughout the experiment and exhibits a pronounced peak near $t \approx 75s$, yet AL_0 remains consistently low (near 0.15), indicating that while many connected components are born, they are short-lived and quickly merge. As a result, TP_0 remains comparatively stable, reflecting a topological landscape dominated by a large number of transient components rather than a smaller number of persistent ones. For $i = 1$, NG_1 peaks sharply near $t \approx 72\text{--}75s$ — coinciding with the period of maximal solute front deformation — while AL_1 dips slightly at the same time before recovering. The early-time peak in TP_1 , occurring near $t \approx 65s$, is thus driven primarily by a period of elevated AL_1 rather than a surge in generator count, suggesting that the loops forming in this phase are geometrically significant and long-lived. The most striking behavior is observed for $i = 2$: while AL_2 increases monotonically over the experiment, NG_2 undergoes a sharp collapse near $t \approx 75s$, driving a corresponding collapse in TP_2 despite the rising average lifespan. This indicates that the enclosed cavities (voids) present in the early tracer front are progressively eliminated as the solute homogenizes within the pore space, and that the few cavities persisting at later times are individually long-lived but too few in number to sustain a large total persistence. Taken together, these trends illustrate how the product structure $TP_i = AL_i \cdot NG_i$ can mask competing dynamics: a growing AL_i and a collapsing NG_i can produce a declining TP_i , as seen for $i = 2$, while a stable TP_i , as seen for $i = 0$, can arise from a consistently large generator count offset by uniformly short feature lifespans.

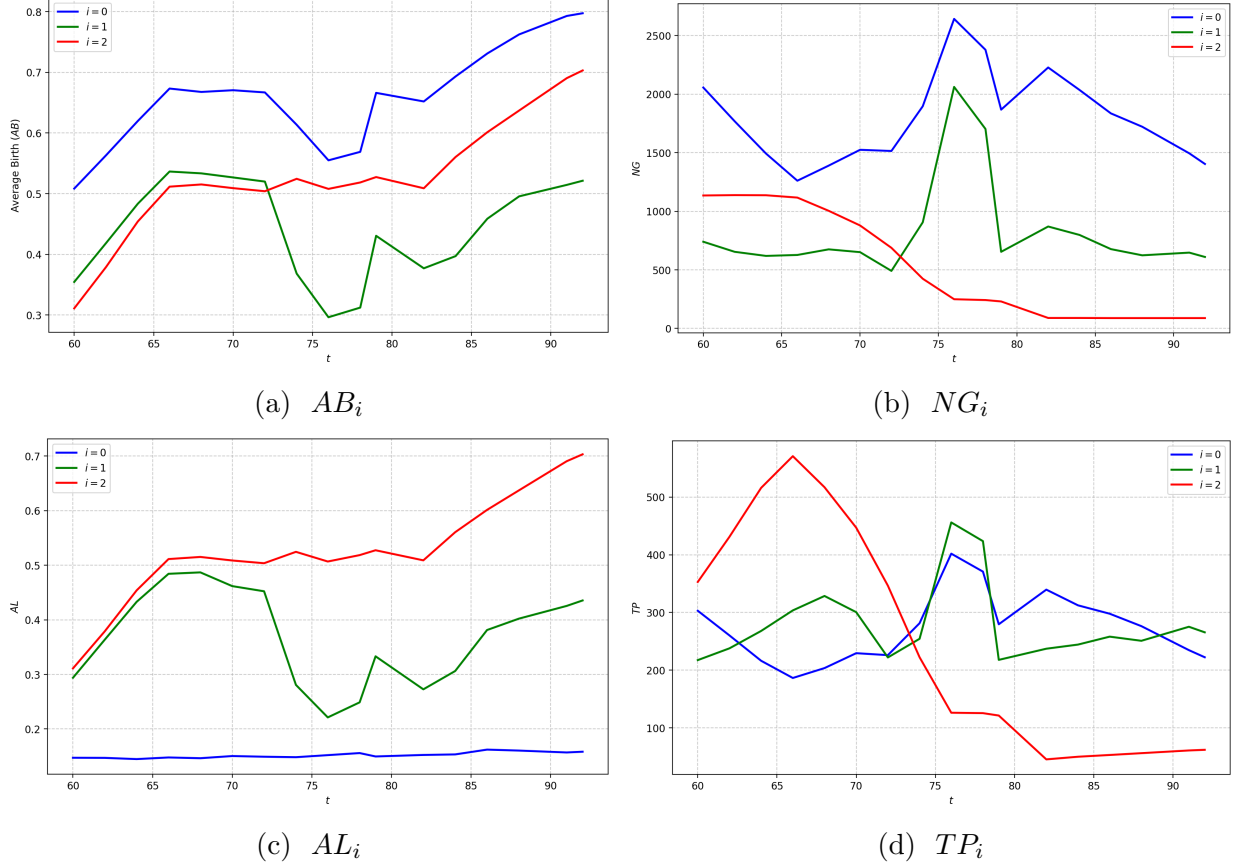


Figure 13: Persistence statistics $\left(S_w = 0.92, q = 0.025 \frac{\text{mm}^3}{s}\right)$.

Since we are interested in understanding the overall significance of the H_0, H_1, H_2 tracer features across various configurations of S_w and q , we examine the total persistence measure across several experiments in Fig. 14. At high saturation ($S_w = 0.92$), TP_0 dominates throughout the experiment, indicating that the mixing front advances primarily via stable, isolated tracer fingers. This is consistent with the near-fully saturated liquid phase forming a wide, contiguous 3D pore network, in which, as Velásquez-Parra et al. report, the solute front undergoes relatively modest deformation and the concentration variance σ_c^2 decays slowly [23]. A pronounced transient peak in TP_2 is also observed at early times for this experiment: because the liquid phase is spatially continuous, advancing tracer fingers retain the three-dimensional freedom to wrap completely around pockets of resident fluid, momentarily forming enclosed cavities before molecular diffusion rapidly homogenizes them. This rapid collapse of H_2 features is consistent with the paper’s description of the scalar dissipation rate being proportional to the square of the local concentration gradient — the highly localized gradients bounding a closed cavity trigger an intense, short-lived burst of diffusive flux that destroys the encapsulation almost immediately after it forms.

As saturation decreases toward $S_w = 0.75$, a qualitative shift in the dominant topological signature is observed: TP_1 grows relative to TP_0 , and the transient H_2 peak is substantially suppressed. This inversion directly reflects the physical fragmentation of the liquid phase by the trapped air. Velásquez-Parra et al. show that at lower S_w , immobile air clusters

occupy the centers of pore bodies, confining the wetting liquid to thin grain-coating films, pendular rings at grain contacts, and narrow crevices [23]. In this geometrically constrained environment, the tracer cannot achieve the three-dimensional spatial continuity required to fully encapsulate resident fluid, suppressing H_2 generation. Instead, the tortuous pathways force strong convergence of flow streamlines, which the paper quantifies through a linearly decreasing Euler characteristic χ with S_w and a corresponding linear increase in the average absolute helicity density $\langle|h|\rangle$. The resulting vorticity-dominated flow mechanically tears advancing tracer fingers apart and folds them back on themselves, producing the loops and tunnels captured by H_1 . The elevated TP_1 signal is therefore a direct topological signature of the chaotic, helical mixing regime that Velásquez-Parra et al. identify as the primary driver of mixing enhancement at lower saturation. Together, the relative amplitudes of TP_0 , TP_1 , and TP_2 across saturation levels constitute a compact topological fingerprint of the transition from advection-dominated fingering to vorticity-driven folding as S_w decreases.

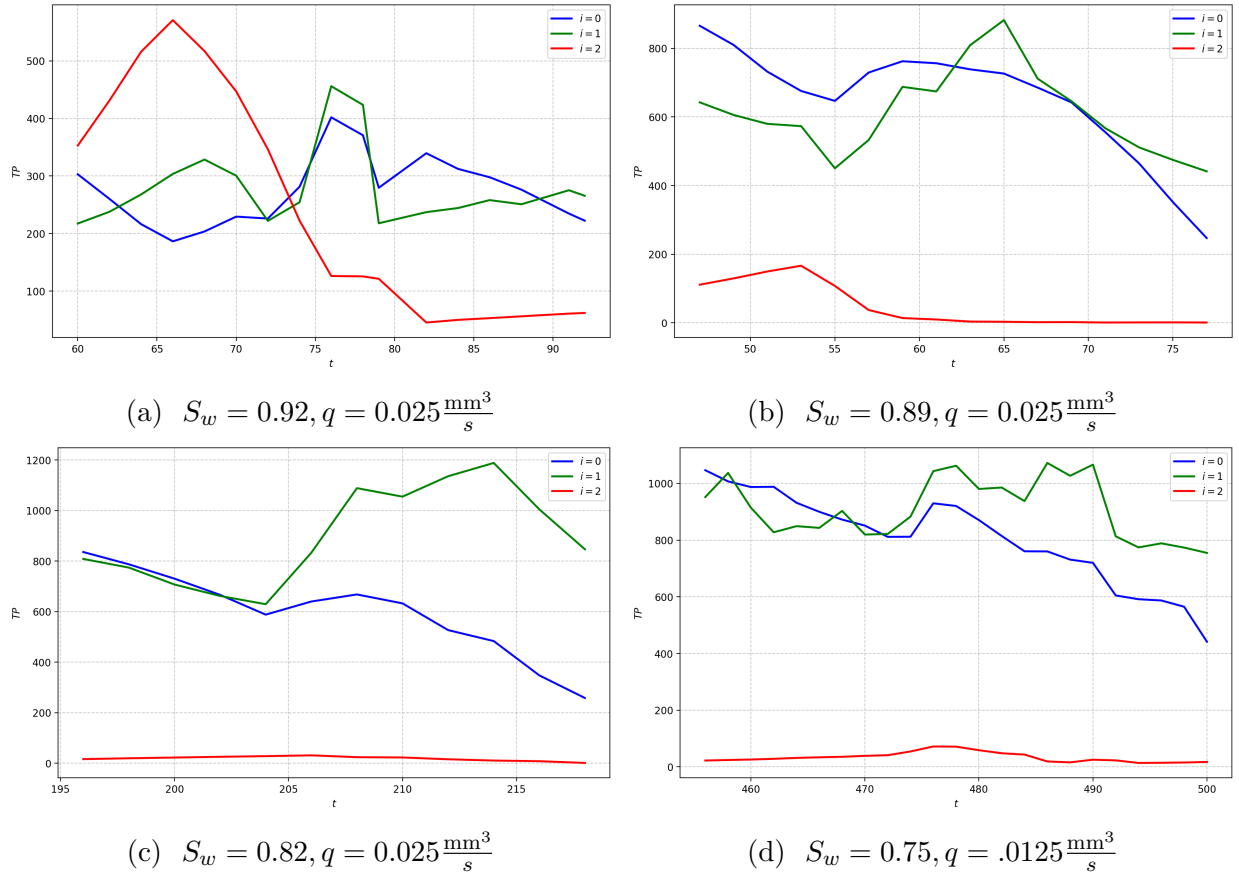


Figure 14: TP_i across several experiments.

To ensure that our observed trends are robust, we perform a sensitivity analysis by varying the sliding window along two axes: (i) altering its location in the xy -plane as it tracks the z -position of the tracer front, and (ii) adjusting the spatial dimensions of the window itself.

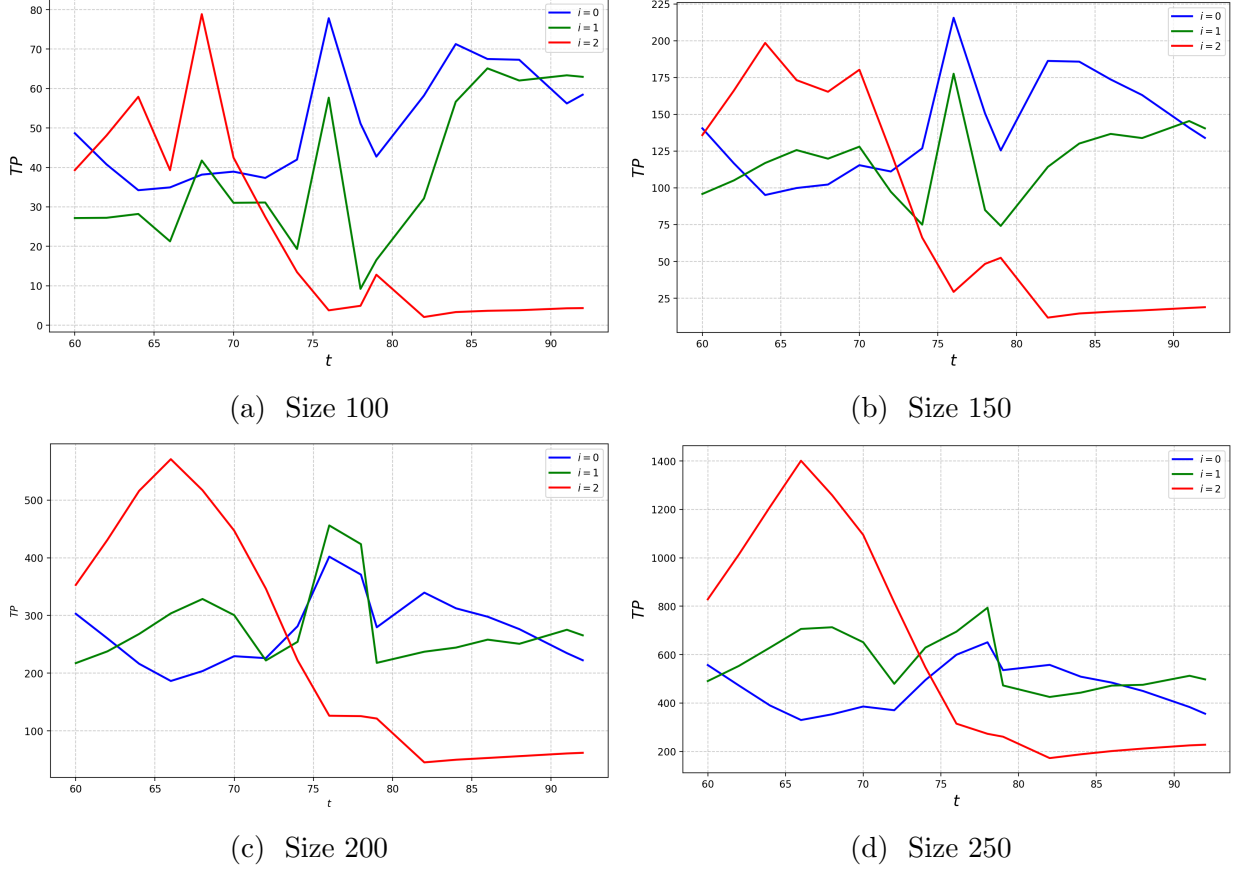


Figure 15: TP_i for various window sizes ($S_w = 0.92, q = 0.025 \frac{\text{mm}^3}{s}$).

Fig. 15 shows the effect of varying window dimensions for the experiment corresponding to $S_w = 0.92, q = 0.025 \frac{\text{mm}^3}{s}$. At smaller window dimensions, the TP curves for all homology groups exhibit noticeably higher volatility, with sharp fluctuations that obscure the underlying trends. This is expected: a small window captures only a limited portion of the tracer front, making the topological summary highly sensitive to local, transient features that may not be representative of the global mixing dynamics. As the window size increases, the TP values grow in a roughly cubic fashion — a natural consequence of the fact that total persistence accumulates contributions from a volume of generators, so scaling the window linearly scales TP as the cube of that factor. Crucially, however, the qualitative trends established in our earlier analysis are preserved across all window sizes: the prominent transient H_2 peak remains visible, and the relative ordering $TP_0 > TP_1$ at high saturation persists regardless of window dimension. Beyond a window size of $200 \times 200 \times 200$ voxels, further increases yield diminishing changes in the observed trends, with the curves stabilizing in both shape and relative magnitude. We therefore adopt $200 \times 200 \times 200$ as our window dimension for all subsequent analyses, as it provides a favorable balance between capturing a representative cross-section of the tracer front and avoiding the inclusion of regions far from the mixing interface where concentration gradients are negligible.

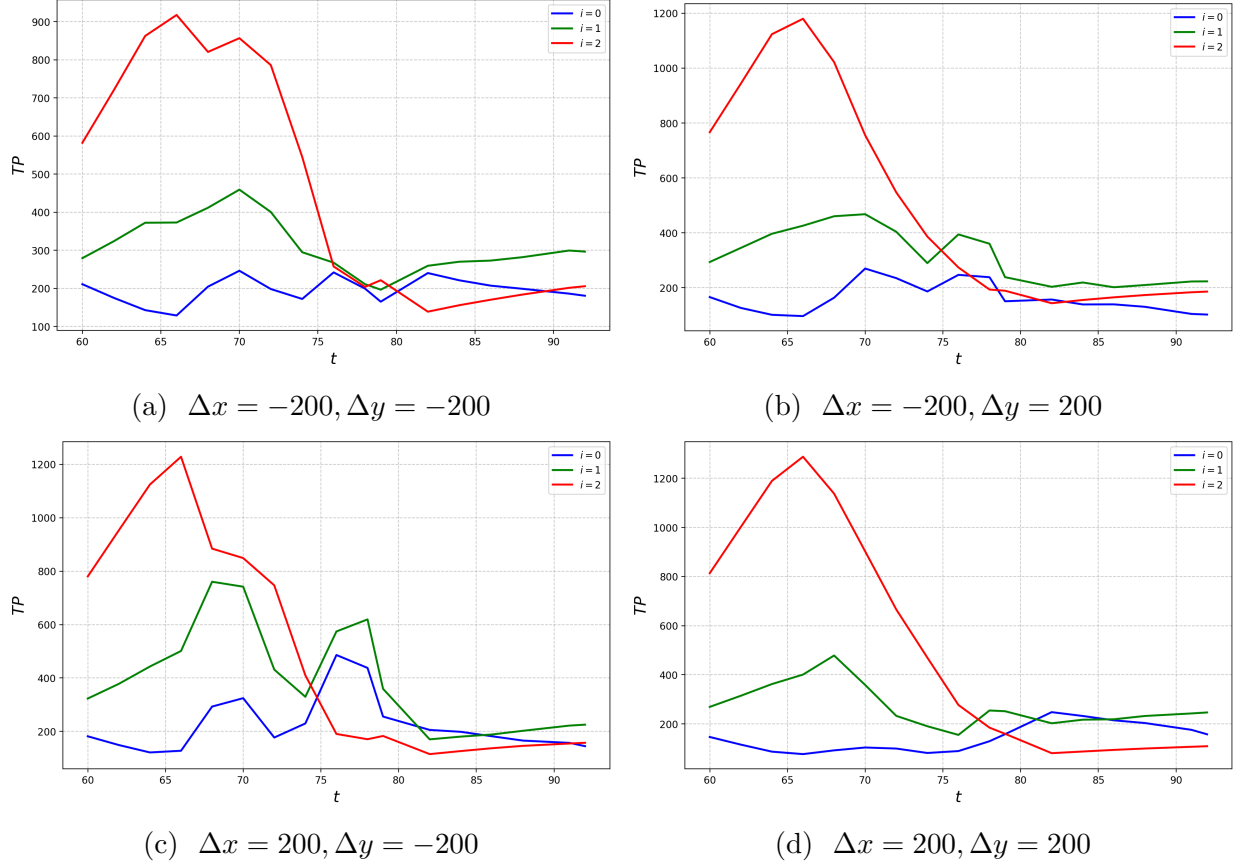


Figure 16: TP_i for various window XY displacements $\left(S_w = 0.92, q = 0.025 \frac{\text{mm}^3}{s}\right)$.

The spatial robustness of our results is further confirmed in Fig. 16, which shows the total persistence curves obtained by repositioning the window at several distinct locations within the xy -plane while continuing to track the same z -position of the tracer front. Across all tested locations, the transient H_2 peak remains consistently present, and the dominance of H_0 over H_1 features at high saturation is preserved. The close agreement between curves at different window positions indicates that the topological signature of the tracer front is a global property of the mixing dynamics rather than an artifact of any particular spatial sampling choice, lending further confidence to the generalizability of our persistent homology framework.

7 Discussion

The Euclidean Distance Transform (EDT) filtration applied to the 3D binary images of the solid matrix reveals that the porous medium is topologically homogeneous at sufficiently large sampling scales. By systematically subsampling the domain and computing persistent homology summaries across subvolumes of increasing size, we find that the topological statistics stabilize once the sampled region contains at least 15 ± 2 pore sizes across each spatial dimension. Below this threshold, the persistence diagrams exhibit high variance

between subvolumes, reflecting the stochastic heterogeneity inherent to any finite porous sample. Above it, the diagrams converge to a consistent signature, confirming that the sintered borosilicate glass grain packing used by Velásquez-Parra et al. behaves as a statistically homogeneous medium at the scales relevant to our analysis [23]. This finding is physically consistent with the relatively narrow pore size distribution reported for the sample, and provides a principled, topology-informed criterion for selecting representative subvolume sizes in future pore-scale studies. Importantly, it also validates our choice of the $200 \times 200 \times 200$ voxel sliding window for the tracer front analysis: the window is large enough to contain a statistically representative cross-section of the pore geometry while remaining small enough to be localized to the mixing interface.

The persistent homology analysis of the tracer front dynamics yields several physically meaningful results that are consistent with and complementary to the findings of Velásquez-Parra et al. Across all experiments, the total persistence curves exhibit a clear transient peak followed by a decay toward a dynamic steady state, reflecting the competition between advective front deformation and diffusive homogenization. At high saturation ($S_w = 0.92$), H_0 features dominate H_1 throughout the experiment, indicating that the mixing front advances primarily through stable, isolated tracer fingers in the wide, contiguous pore channels characteristic of near-fully saturated conditions. The prominent transient H_2 peak observed at this saturation confirms that the three-dimensional continuity of the liquid phase permits complete encapsulation of resident fluid, consistent with the paper’s description of a relatively homogeneous, piston-like solute front at high S_w . As saturation decreases toward $S_w = 0.75$, the H_1 signal grows relative to H_0 , and the H_2 peak is substantially suppressed. This topological inversion constitutes a direct mathematical signature of the transition from advection-dominated fingering to vorticity-driven folding that Velásquez-Parra et al. identify as the primary mechanism of mixing enhancement at lower saturation [23]. Together, the relative amplitudes and temporal evolution of TP_0 , TP_1 , and TP_2 provide a compact, quantitative characterization of the mixing regime that is both physically interpretable and robust to the choice of window size and location, as demonstrated by our sensitivity analysis.

Several directions remain open for future work. On the solid matrix side, the EDT-based homogeneity analysis should be complemented and cross-validated using Alpha Complex filtrations, which probe the geometry of the solid phase directly through the coordinates of grain boundary voxels rather than through the distance field of the void space. Since the two filtrations encode structural information differently — EDT emphasizing proximity to solid boundaries and alpha complexes capturing grain-scale connectivity — agreement between their homogeneity thresholds would provide stronger evidence that the identified pore size criterion reflects an intrinsic geometric property of the medium rather than an artifact of the filtration choice. Additionally, a more exhaustive subsampling analysis, exploring a wider range of subvolume shapes, orientations, and spatial positions within the domain, would further quantify the uncertainty bounds on the homogeneity threshold and establish whether any weak spatial gradients in the pore structure — for instance, near the sample boundary — affect the topological statistics.

For the tracer front analysis, the most significant methodological extension is to move beyond the sliding window approach toward a full isosurface-based analysis. The current window approach, while robust and well-validated, necessarily captures only a local cross-section of the tracer front at any given time step. A natural generalization is to apply the

persistent homology filtration directly to the full $c = 0.5$ M concentration isosurface reconstructed across the entire sample volume, as Velásquez-Parra et al. use to quantify solute front deformation through the growth of its surface area [23]. This would allow our topological characterization to operate on the same global geometric object as the paper’s own mixing analysis, enabling a direct quantitative comparison between the super-diffusive scaling of the isosurface area and the temporal evolution of TP_0 , TP_1 , and TP_2 . Such a reconciliation would establish whether total persistence provides an independent or redundant characterization of front deformation relative to surface area, and could reveal topological aspects of the mixing dynamics — such as the formation and destruction of loops and encapsulations — that are invisible to scalar area measures alone. Achieving this will require addressing the computational challenges associated with running persistent homology on full-domain, high-resolution concentration fields, motivating further work on efficient filtration algorithms and domain decomposition strategies suited to the image dimensions encountered in synchrotron micro-tomography datasets.

References

- [1] Mehmet E. Aktas, Esra Akbas, and Ahmed El Fatmaoui. “Persistence homology of networks: Methods and applications”. In: *Applied Network Science* 4.1 (2019), p. 61.
- [2] M. J. Blunt et al. “Pore-scale imaging and modelling”. In: *Adv. Water Resour.* 51 (2013), pp. 197–216.
- [3] S. Bottero et al. “4D microvelocimetry reveals multiphase flow field perturbations in porous media”. In: *Proceedings of the National Academy of Sciences* 121 (2024).
- [4] *Coffee mug deforming into a donut - 3D model by Henry Segerman - Sketchfab* — *sketchfab.com*. <https://sketchfab.com/models/ed470a3541ff4aabb5972b02f74e4ac3/embed>. [Accessed 13-05-2026].
- [5] Irene Donato et al. “Persistent homology analysis of phase transitions”. In: *Phys. Rev. E* 93.5 (May 2016), p. 052138.
- [6] Herbert Edelsbrunner, John Harer, et al. “Persistent homology—a survey”. In: *Contemporary Mathematics*. Vol. 453. 2008, pp. 257–282.
- [7] I. M. Griffiths, A. Kumar, and P. S. Stewart. “Flow and fouling in membrane filters: effects of membrane morphology”. In: *Journal of Fluid Mechanics* 824 (2017).
- [8] Sharul Hasan et al. “Direct characterization of solute transport in unsaturated porous media using fast X-ray synchrotron microtomography”. In: *Proceedings of the National Academy of Sciences* 117.38 (2020), pp. 23443–23449. DOI: [10.1073/pnas.2011716117](https://doi.org/10.1073/pnas.2011716117). eprint: <https://www.pnas.org/doi/pdf/10.1073/pnas.2011716117>. URL: <https://www.pnas.org/doi/abs/10.1073/pnas.2011716117>.
- [9] A. Hernandez et al. “Optimizing scaffold pore size for tissue engineering: insights across various tissue types”. In: *Frontiers in Bioengineering and Biotechnology* (2024).
- [10] Jonathan Jaquette and Benjamin Schweinhart. “Fractal dimension estimation with persistent homology: A comparative study”. In: *Communications in Nonlinear Science and Numerical Simulation* 84 (May 2020), p. 105163. ISSN: 1007-5704. DOI: [10.1016/j.cnsns.2019.105163](https://doi.org/10.1016/j.cnsns.2019.105163). URL: <http://dx.doi.org/10.1016/j.cnsns.2019.105163>.
- [11] Miroslav Kramár et al. “Quantifying force networks in particulate systems”. In: *Physica D* 283 (2014), pp. 37–55.
- [12] V. Martin, J. Jaffré, and J. E. Roberts. “Modeling Fractures and Barriers as Interfaces for Flow in Porous Media”. In: *SIAM Journal on Scientific Computing* 26.5 (2005), pp. 1667–1691.
- [13] G. Milan et al. “Pore-Scale Transport and Two-Phase Fluid Structures in Fibrous Porous Layers: Application to Fuel Cells and Beyond”. In: *Transport in Porous Media* 136.1 (2021), pp. 245–270.
- [14] M. R. Morad et al. “Localization in Flow of Non-Newtonian Fluids Through Disordered Porous Media”. In: *Frontiers in Physics* (2021).
- [15] James R. Munkres. *Elements of algebraic topology*. Perseus Publ, 1997.

- [16] Ryo Muso et al. “Persistent homology-based microstructural feature extraction and its correlation with mechanical properties in Mo–TiC alloys”. In: *International Journal of Refractory Metals and Hard Materials* (2026), p. 107721. ISSN: 0263-4368.
- [17] Nina Otter et al. “A roadmap for the computation of persistent homology”. In: *EPJ Data Science* 6.1 (2017), p. 17.
- [18] M. A. Porter, M. Feng, and E. Katifori. “The topology of data”. In: *Physics Today* 76.1 (Jan. 2023), pp. 36–42. ISSN: 0031-9228.
- [19] Ysanne Pritchard et al. “Persistent homology analysis distinguishes pathological bone microstructure in non-linear microscopy images”. In: *Frontiers in Endocrinology* 14 (2023), p. 1065113.
- [20] Eashwar Somasundaram et al. “Persistent homology of tumor CT scans is associated with survival in lung cancer”. In: *Medical Physics* 48.11 (2021), pp. 7043–7051.
- [21] G. Sundaramoorthi et al. “3D Membrane Imaging and Porosity Visualization”. In: *Ind. Eng. Chem. Res.* 55 (2016), pp. 3689–3695.
- [22] K. Taghizadeh et al. “X-ray 3D imaging based microunderstanding of granular mixtures: Stiffness enhancement by adding small fractions of soft particles”. In: *Proc. Nat. Acad. Sciences* 120 (2023), e2219999120.
- [23] Andrés Velásquez-Parra et al. “Phase Saturation Control on Vorticity Enhances Mixing in Porous Media”. In: *Water Resources Research* 60 (2024), e2023WR036628.
- [24] Xiaoqi Wei and Guo-Wei Wei. “Persistent Topological Laplacians – a Survey”. In: *arXiv* (2023).
- [25] Georg Wilding et al. “Persistent homology of the cosmic web – I. Hierarchical topology in Λ CDM cosmologies”. In: *Monthly Notices of the Royal Astronomical Society* 507.2 (Aug. 2021), pp. 2968–2990.