

Quantifying flow and mixing through porous media using Topological Data Analysis

Aakash Karlekar Kevin Steiger Abril Montes Connor Senor ¹

¹Department of Mathematical Sciences,
New Jersey Institute of Technology, Newark, NJ 07102, USA

May 1, 2026

Table of Contents

- 1 Introduction
- 2 Mathematical Framework and Theoretical Foundations
- 3 Analyzing Solid Morphology Data
 - Alpha Complexes
 - Euclidean Distance Transform (EDT)
- 4 Tracer Front Deformation
- 5 Discussion

Background

- Porous media play a fundamental role in critical natural and industrial processes, such as enhanced oil recovery and CO₂ sequestration^{1,2}.
- The mixing of multi-phase fluids is mechanically driven by the complex, micro-scale geometry of internal pore networks.
- Topological Data Analysis (TDA), specifically persistent homology (PH), offers a robust mathematical framework to quantify these topological properties³.
- **Objective:** Apply PH to high-resolution 3D images to characterize porous media structure and understand mixing performance.

¹M. J. Blunt et al. "Pore-scale imaging and modelling". In: *Adv. Water Resour.* 51 (2013), pp. 197–216.

²G. Sundaramoorthi et al. "3D Membrane Imaging and Porosity Visualization". In: *Ind. Eng. Chem. Res.* 55 (2016), pp. 3689–3695.

³M. A. Porter, M. Feng, and E. Katifori. "The topology of data". In: *Physics Today* 76.1 (Jan. 2023), pp. 36–42. ISSN: 0031-9228.

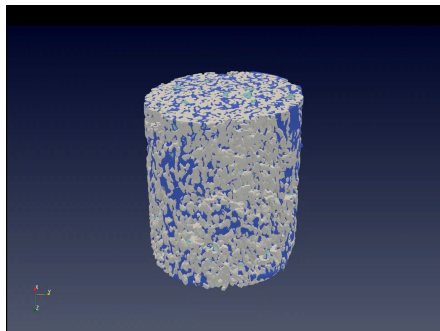
Experiment

- **Imaging:** High-resolution 4D synchrotron X-ray micro-tomography ($2.75\mu m$, 2.4 s) sourced from Velásquez-Parra et al.⁴.
- **Medium:** Sintered glass grains mimicking poorly graded sandy soil.
- **Fluids:** Immobile liquid and gas phase along with Potassium iodide (KI) aqueous tracer.
- **Procedure:** Tracer injected at constant flow rates (q) under varying steady-state liquid saturations (S_w).
- **Observation:** Real-time tracking of spatial concentration and multi-phase mixing dynamics.
- **Conclusions:** Lower $S_w \Rightarrow$ better liquid-liquid mixing

⁴Andrés Velásquez-Parra et al. "Phase Saturation Control on Vorticity Enhances Mixing in Porous Media". In: *Water Resources Research* 60 (2024), e2023WR036628.

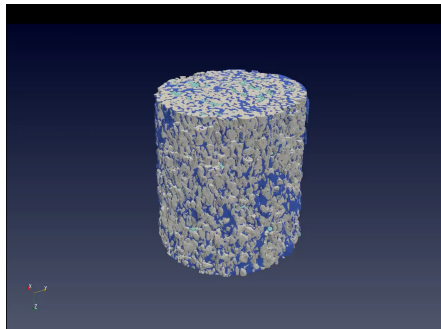
Morphology Data

- 3D binary images that capture the initial structural state of the multi-phase system (Solid–Liquid–Gas).
- Focus on solid phase
- Look to employ two complementary PH filtration methods to characterize matrix:
 - **Alpha Complex Filtration**
 - **Euclidean Distance Transform (EDT) Filtration**
- Cross-validate outputs to ensure results are robust



Tracer Data

- 4D grayscale images tracking the temporal evolution of an aqueous tracer spreading through the liquid phase.
- Grayscale values correspond to concentration of tracer fluid
- Track the topological evolution and deformation of the tracer front over time
 - Sliding Window
 - **Grayscale-Based PH**



Significance

- Advances beyond traditional macroscopic metrics, which often fail to capture intricate spatial connectivity and structural bottlenecks.
- Mathematically substantiates how lower liquid phase saturation alters active flow pathways and increases spatial convergence of streamlines⁵.
- Broad implications for optimizing technologies in fuel cells⁶, soil management⁷, water filtration⁸, and tissue engineering⁹.

⁵Andrés Velásquez-Parra et al. "Phase Saturation Control on Vorticity Enhances Mixing in Porous Media". In: *Water Resources Research* 60 (2024), e2023WR036628.

⁶G. 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.

⁷S. Bottero et al. "4D microvelocimetry reveals multiphase flow field perturbations in porous media". In: *Proceedings of the National Academy of Sciences* 121 (2024).

⁸I. 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).

⁹A. Hernandez et al. "Optimizing scaffold pore size for tissue engineering: insights across various tissue types". In: *Frontiers in Bioengineering and Biotechnology* (2024).

Table of Contents

- 1 Introduction
- 2 Mathematical Framework and Theoretical Foundations**
- 3 Analyzing Solid Morphology Data
 - Alpha Complexes
 - Euclidean Distance Transform (EDT)
- 4 Tracer Front Deformation
- 5 Discussion

Topological Characterization

Topology allows us to extract the underlying structure of the flow paths

- **Physical Mapping of Betti Numbers (β_n):**

- β_0 (connected components): separate blobs

- β_1 (loops/tunnels): paths that wrap around an obstacle

- β_2 (voids/cavities): fully enclosed empty spaces

- **Note:** Physical interpretation depends on the phase we are analyzing (fluid vs. solid)

- **Persistence as a Filter:** use persistence diagrams to separate real structural features (long-lived) from imaging artifacts and surface roughness (short-lived)

The Challenge

Our Data Challenges

- **Scale and Subsampling Constraints**

Full (raw) volume: $2016 \times 2016 \times 1100 = 4$ billion voxels

Current analysis: $1518 \times 1502 \times 1000$ (post-processed)

- **Scanner Noise and Calibration Uncertainties:** voxelization artifacts, background correction errors, conversion of pixel values to concentration
- **Mechanical and Motion Artifacts:** restricted sample movement during rotation

Big Question: Are our topological features real structures, or artifacts of subsampling and noise?

Interaction Complex and Stability Theorem

Methodological Comparison

- Position Complex (position-based): features are defined by geometric proximity
- Interaction Complex (value-based): features are defined by physical values

Stability Theorem (*Kramár et al.*¹⁰)

"The persistence diagram is Lipschitz continuous with respect to the filtration function"

- **translation**: a small change in the input results in small changes in the output

The Significance: The **interaction complex** provides a universal stability guarantee that traditional methods cannot maintain

¹⁰Miroslav Kramár et al. "Quantifying force networks in particulate systems". In: *Physica D* 283 (2014), pp. 37–55.

Application to Our Project

Our methods follow this framework:

Approach	Filtration	Framework	Stability
Alpha Complexes	Ball radius	Position	Less stable
EDT	Distance	Interaction	Stable
Grayscale	Concentration	Interaction	Stable

Table: Categorization of topological approaches by stability and framework

- **Primary analyses:** EDT and Grayscale-based thresholding
- **Validation check:** Alpha complexes

All three approaches analyze the same structures from different perspectives. Agreement across these frameworks confirms that our topological features are real, not artifacts.

Validation Test 1: Spatial Homogeneity

- **Question:** Do different spatial locations within the sample yield consistent topology?
- **Idea:** Topology does not change by location
- **Method:** Systematically shift the observation window to various X and Y displacements from the spatial center of the sample
Validation: If plots across these spatial locations show similar topology, it suggests that the observed mixing regimes are representative of the global medium

Additional Verification: Static Geometry

- **Methods:** Solid fraction and pore size distribution analysis

Validation Test 2: Scale Independence

- **Question:** At what scale does the topology stabilize and become representative of the global mixing regime?
- **Idea:** Topology is independent of scale (once a sufficiently large window is established)
- **Method:** Systematically increase dimensions of tracer front observation window
Validation: If plots across varying scales show stable topology, it suggests that we have found the ideal observation scale

Why this matters

- **Theoretical Foundation:** interaction complex framework proven stable (approximately 6500x better than alternative approaches)
- **Smart Method Choices:** EDT and grayscale-based thresholding use a stable framework, alpha complex provides geometric validation
- **Rigorous Validation:** spatial homogeneity and scale independence tests

Table of Contents

- 1 Introduction
- 2 Mathematical Framework and Theoretical Foundations
- 3 Analyzing Solid Morphology Data
 - Alpha Complexes
 - Euclidean Distance Transform (EDT)
- 4 Tracer Front Deformation
- 5 Discussion

Table of Contents

- 1 Introduction
- 2 Mathematical Framework and Theoretical Foundations
- 3 Analyzing Solid Morphology Data**
 - Alpha Complexes
 - Euclidean Distance Transform (EDT)
- 4 Tracer Front Deformation
- 5 Discussion

Characterizing Porous Microstructure: Motivation & Approach

- Full dataset contains over 1 billion voxels
 - Computing TDA on the full domain is computationally difficult
- **Key question:** How small a subvolume still yields an accurate topological characterization?
- Two complementary analysis techniques explored:
 - **Alpha Complexes** — generators in solid voxel centroids
 - **EDT** — distance transform filtration on pore space

Total Persistence

Total Persistence summarizes the overall topological complexity of a filtration:

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

where $b_i^{(k)}$ and $d_i^{(k)}$ are the birth and death times of the i -th feature in dimension k .

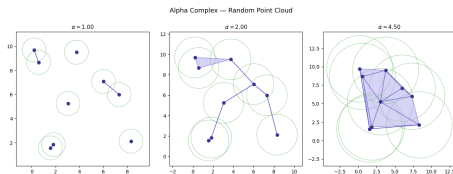


Figure: Persistence diagram — random point cloud

Raw Total Persistence

- Alpha complex computed on random subsamples of solid voxels within a fixed $200 \times 200 \times 200$ subvolume
- For each down-sample, take 10 trials of total persistence, calculate mean standard deviation
- Raw TP grows with the number of sampled points n
- No convergence observed — cannot reliably compare topological features across sampling fractions

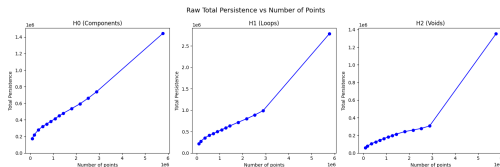


Figure: Raw total persistence vs. number of sampled solid voxels

Grid Introduced Noise

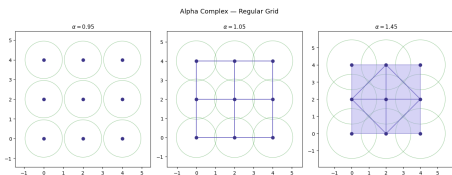


Figure: Alpha complex — regular grid



Figure: 2D slice of porous medium

- Generators at regular grid positions introduce spurious short-lived features — noise from lattice geometry, not true pore structure

Norm 3 Approach

- Raw TP scales as $n^{\frac{d-\alpha}{d}}$ with sample size — a normalization is needed to enable comparison across sampling fractions
- The α -weighted lifetime sum is defined as:¹¹

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

- Setting $\alpha = d = 3$ forces the scaling exponent to zero — E_3^i is theoretically stable with respect to n :

$$n^{\frac{d-\alpha}{d}} = n^{\frac{3-3}{3}} = n^0 = 1$$

- Cubing lifetimes also reduces the impact of short-lived spurious features relative to true long-lived pore structure

¹¹Jonathan 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. URL: <http://dx.doi.org/10.1016/j.cnsns.2019.105163>.

Results: Norm 3 Total Persistence

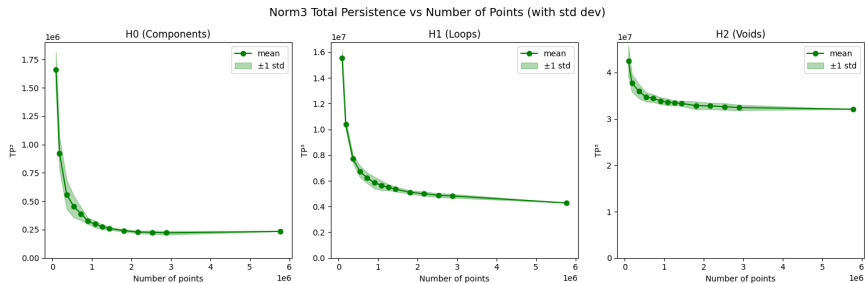


Figure: TP_3 vs. number of sampled solid voxels

- H2 (Voids) shows the strongest stabilization — consistent with convergence to a stable topological signature
- H0 and H1 show partial convergence at large n ; a precise downsampling threshold requires further analysis

Table of Contents

- 1 Introduction
- 2 Mathematical Framework and Theoretical Foundations
- 3 Analyzing Solid Morphology Data
 - Alpha Complexes
 - Euclidean Distance Transform (EDT)
- 4 Tracer Front Deformation
- 5 Discussion

EDT: From Geometry to Scalar Field

- Binary porous medium $\rightarrow$ Euclidean Distance Transform
- Each voxel is assigned its distance to the nearest interface
- Produces a scalar field encoding pore geometry
- Large values $\rightarrow$ deep inside pores
- Small values $\rightarrow$ near boundaries
- Serves as input for topological analysis in HomCloud

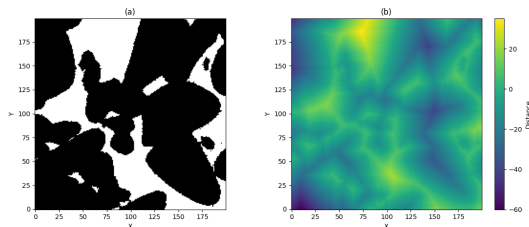


Figure: Example slice of the EDT field

Persistent Homology via HomCloud

- Filtration applied to the EDT field using a superlevel filtration
- Tracks topological features:
 - H_0 : connected components
 - H_1 : loops
 - H_2 : voids
- Persistence diagrams summarize when features appear and disappear
- Total persistence is computed as the sum of feature lifetimes
- Total persistence was computed across multiple domain sizes

¹²I. Obayashi, T. Nakamura, and Y. Hiraoka, *Persistent Homology Analysis for Materials Research and Persistent Homology Software: HomCloud*, J. Phys. Soc. Jpn. 91 (2022), 091013. Available at: <https://homcloud.dev/python-api/index.html>

Variance of Total Persistence vs Domain Size

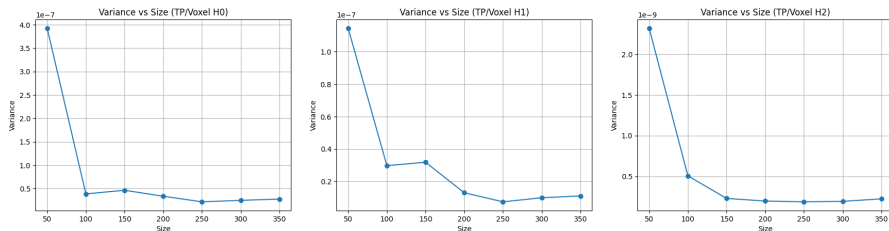


Figure: Variance of normalized total persistence across domain sizes for H_0 , H_1 , and H_2

SNOW2: Extracting Pore Structure

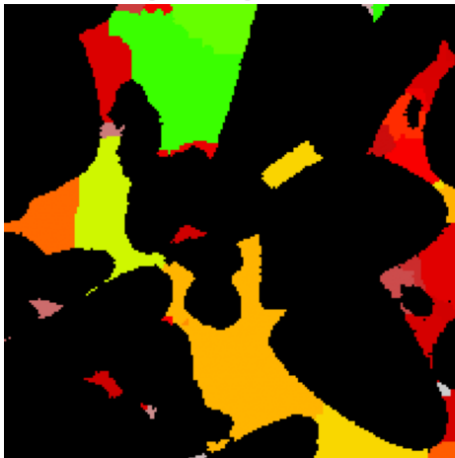
- SNOW2 converts voxel data into a pore network representation
- Step 1: Compute distance transform
 - Identifies distance from pore space to solid boundary
- Step 2: Detect local maxima
 - Points farthest from boundaries represent pore centers
- Step 3: Remove redundant peaks
 - Helps ensure one node per physical pore
- Step 4: Watershed segmentation
 - Divides the domain into distinct pore regions

12

¹²E. Dagdelen et al., *Machine Learning and Topological Approaches for Predicting Permeability of Porous Media*, 2026.

SNOW2 Output

SNOW2 segmented regions - middle slice



From Topology to Physical Scale

- SNOW2 provides an average pore size in voxels
- Topological results show that normalized total persistence stabilizes near domain size $\sim 150^3$
- This connects the topological result to a physical length scale:

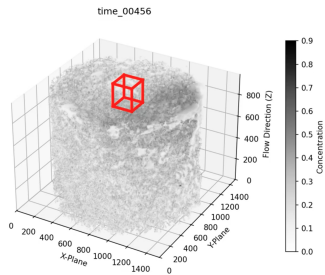
$$\frac{\text{domain length}}{\text{average pore size}} = \frac{150}{11.8} \approx 12.7$$

- Interpretation:
 - The domain must contain enough pores to give stable topological measurements
 - A domain length of about 13 average pore diameters is sufficient for this sample

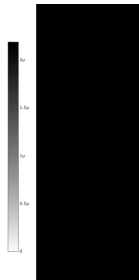
Table of Contents

- 1 Introduction
- 2 Mathematical Framework and Theoretical Foundations
- 3 Analyzing Solid Morphology Data
 - Alpha Complexes
 - Euclidean Distance Transform (EDT)
- 4 Tracer Front Deformation
- 5 Discussion

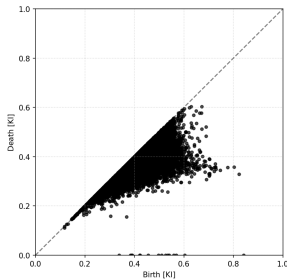
Background



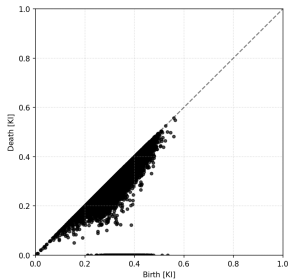
Superlevel Threshold: 2.19e-06



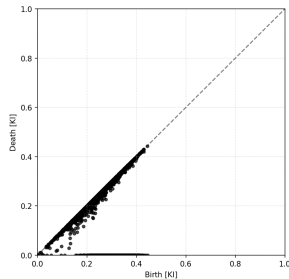
Results



(a) H_0



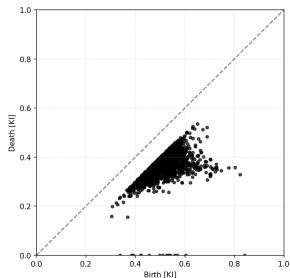
(b) H_1



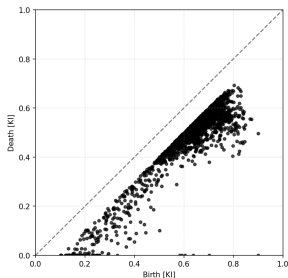
(c) H_2

Figure: Persistence diagrams for Time $t = 00060$ ($S_w = 0.92, q = 0.025 \frac{\text{mm}^3}{s}$).

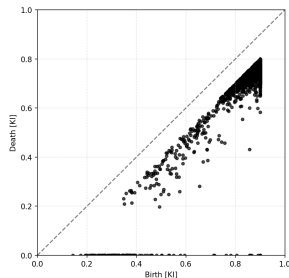
Results (cont.)



(a) $t = 00060$



(b) $t = 00074$



(c) $t = 00091$

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

Results (cont.)

Number of Generators

We define the NG_i to be the number of topological generators (points) present in the persistence diagrams corresponding to H_i .

Average Birth

A generator within a persistence diagram can be represented as (b, d) where b, d refer to the birth and death number of the generator respectively. We define the Average Birth of the generators corresponding to H_i as

$$AB_i := \frac{1}{NG_i} \sum_{n=1}^{NG_i} b_n$$

a birth coordinate b and a death coordinate

Results (cont.)

Average Lifespan

We define the Average Lifespan of the generators corresponding to H_i as

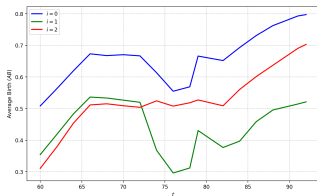
$$AL_i := \frac{1}{NG_i} \sum_{n=1}^{NG_i} |b_n - d_n|$$

Total Persistence

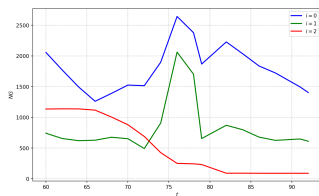
We define the Total Persistence of the generators corresponding to H_i as

$$TP_i := \sum_{n=1}^{NG_i} |b_n - d_n| = AL_i \times NG_i$$

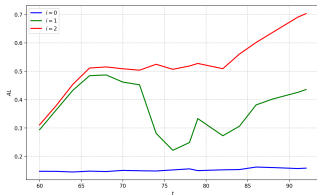
Results (cont.)



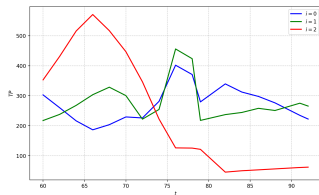
(a) AB_i



(b) NG_i



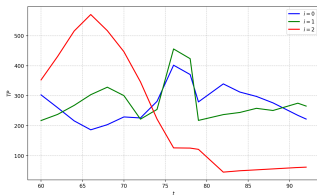
(c) AL_i



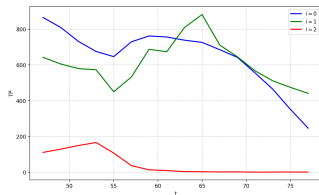
(d) TP_i

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

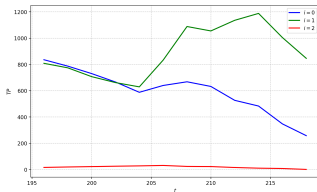
Results (cont.)



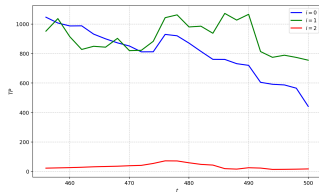
(a) $S_w = 0.92, q = 0.025 \frac{\text{mm}^3}{\text{s}}$



(b) $S_w = 0.89, q = 0.025 \frac{\text{mm}^3}{\text{s}}$



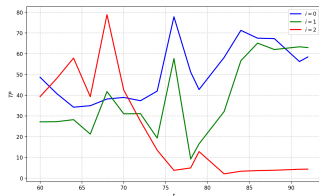
(c) $S_w = 0.82, q = 0.025 \frac{\text{mm}^3}{\text{s}}$



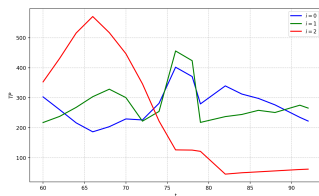
(d) $S_w = 0.75, q = .0125 \frac{\text{mm}^3}{\text{s}}$

Figure: TP_i across several experiments.

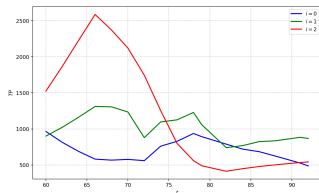
Robustness (Window Size)



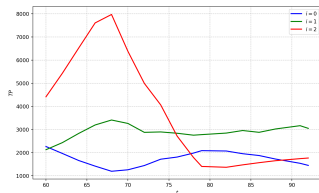
(a) Size 100



(b) Size 200



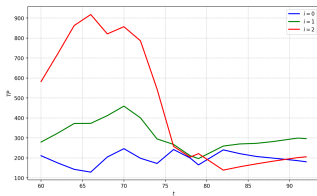
(c) Size 300



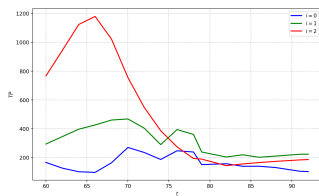
(d) Size 400

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

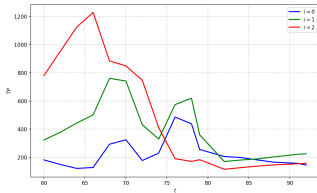
Robustness (Window Location)



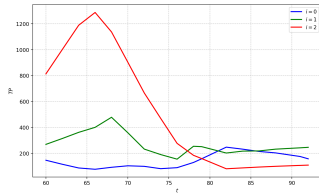
(a) $\Delta x = -200, \Delta y = -200$



(b) $\Delta x = -200, \Delta y = 200$



(c) $\Delta x = 200, \Delta y = -200$



(d) $\Delta x = 200, \Delta y = 200$

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

Table of Contents

- 1 Introduction
- 2 Mathematical Framework and Theoretical Foundations
- 3 Analyzing Solid Morphology Data
 - Alpha Complexes
 - Euclidean Distance Transform (EDT)
- 4 Tracer Front Deformation
- 5 Discussion

- Analyzing Solid Morphology Data
 - Solid domain is homogeneous
 - Ratio of domain length to number of sampling points is roughly 12.7
- Tracer Front Deformation
 - Lower $S_w \Rightarrow$ reduced H_2 peak and H_1 features $\gg H_2$ features
 - Findings consistent for various sliding window locations and sizes

Future Work

- Analyzing Solid Morphology Data
 - Continue developing Alpha Complexes approach
 - Connect work to ongoing work about subsampling domain
- Tracer Front Deformation
 - Correlate results with existing measures from paper¹³
 - Expand sliding window approach to consider iso-surface of tracer front

¹³ [Andrés Velásquez-Parra et al.](#) "Phase Saturation Control on Vorticity Enhances Mixing in Porous Media". In: *Water Resources Research* 60 (2024), e2023WR036628.

Thanks!
Questions?